

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 371 853**

51 Int. Cl.:

C12N 15/54 (2006.01)

C12N 9/12 (2006.01)

C12N 1/19 (2006.01)

C12N 1/21 (2006.01)

C12N 5/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **99909610 .0**

96 Fecha de presentación: **25.02.1999**

97 Número de publicación de la solicitud: **1056868**

97 Fecha de publicación de la solicitud: **06.12.2000**

54

Título: **HOMÓLOGO HUMANO DE LA PROTEÍNA "FUSED" DE DROSOPHILA.**

30

Prioridad:
26.02.1998 US 31563

73

Titular/es:
GENENTECH, INC.
1 DNA WAY
SOUTH SAN FRANCISCO, CA 94080-4990, US

45

Fecha de publicación de la mención BOPI:
10.01.2012

72

Inventor/es:
DE SAUVAGE, Frederic, J. y
ROSENTHAL, Arnon

45

Fecha de la publicación del folleto de la patente:
10.01.2012

74

Agente: **Ponti Sales, Adelaida**

ES 2 371 853 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

Homólogo humano de la proteína “*fused*” de *Drosophila*

CAMPO DE LA INVENCION

[0001] La presente invención se refiere de manera general a moléculas de señalización, específicamente a moléculas de señalización y mediadoras en la cascada hedgehog (Hh) que se encuentran implicadas en la proliferación y diferenciación celulares.

ANTECEDENTES DE LA INVENCION

[0002] El desarrollo de los organismos multicelulares depende, por lo menos en parte, de mecanismos que especifican, dirigen o mantienen la información posicional para la formación de células, tejidos u órganos. Se han asociado diversas moléculas de señalización secretadas, tales como elementos de las familias del factor beta de crecimiento transformante (TGF- β), de Wnt, de factores de crecimiento fibroblástico y hedgehog, con la actividad de formación de diferentes células y estructuras en *Drosophila*, así como en vertebrados (Perrimon, Cell 80:517-520, 1995).

[0003] Se identificó hedgehog (*Hh*) por primera vez como un gen de polaridad de segmento mediante un cribado genético en *Drosophila melanogaster*, Nusslein-Volhard et al., Roux, Arch. Dev. Biol. 193:267-282, 1984, que desempeña una amplia diversidad de funciones en el desarrollo (Perrimon, supra). Aunque únicamente se ha identificado un gen de *Drosophila*, se han aislado tres homólogos *Hh* de mamífero: *Hh Sonic* (*SHh*), *Hh Desert* (*DHh*) y *Hh Indian* (*Ihh*) (Echelard et al., Cell 75:1417-30, 1993; Riddle et al., Cell 75:1401-16, 1993). *SHh* se expresa a nivel elevado en el notocordio y en la placa del piso de embriones de vertebrados en desarrollo. Los ensayos *in vitro* de explantes, así como la expresión ectópica de *SHh* en animales transgénicos demuestran que *SHh* desempeña un papel clave en la formación del tubo neuronal, Echelard et al., supra, 1993; Krauss et al., Cell 74:1431-44 (1993), Riddle et al. Cell 75: 1401-16 (1993), Roelink et al. Cell 81: 445-55 (1995). Los ensayos *in vitro* de explantes, así como la expresión ectópica de *SHh* en animales transgénicos demuestran que *SHh* desempeña un papel clave en la formación del tubo neuronal, Echelard et al., supra, 1993; Ericson et al., Cell 81:745-56, 1995; Marti et al., Nature 375:322-5, 1995; Roelink et al., supra, 1995; Hynes et al., Neuron 19:15-26, 1997). *Hh* también desempeña un papel en el desarrollo de las extremidades (Krauss et al., Cell 75:1431-44, 1993; Laufer et al., Cell 79:993-1003, 1994), de las somitas (Fan y Tessier-Lavigne, Cell 79:1175-86, 1994; Johnson et al., Cell 79:1165-73, 1994), de los pulmones (Bellusci et al., Develop. 124:53-63, 1997), y de la piel (Oro et al., Science 276:817-21, 1997). De manera similar, *Ihh* y *DHh* se encuentran implicados en el desarrollo de hueso, intestino y células germinales (Apelqvist et al., Curr. Biol. 7:801-4, 1997; Bellusci et al., Development 124:53-63, 1997; Bitgood et al., Curr. Biol. 6:298-304, 1996; Roberts et al., Development 121:3163-74, 1995). Los ratones sin *SHh* fortalecieron además la noción de que el *SHh* es crítico para muchos aspectos del desarrollo de vertebrados, Chiang et al., Nature 383: 407-13 (1996). Estos ratones muestran defectos en las estructuras en la línea media, tales como el notocordio y la placa del piso, ausencia de tipos de células ventrales en el tubo neural, ausencia de estructuras de las extremidades distales, ciclopía, y ausencia de la columna espinal y la mayoría de costillas.

[0004] En la superficie celular, se cree que las señales de Hh son transmitidas por la proteína de 12 dominios transmembrana *Patched* (*Ptch*) [Hooper and Scott, Cell 59: 751-65 (1989); Nakano et al., Nature 341: 508-13 (1989)] y el receptor del tipo acoplado a proteína G Smoothed (*Smo*) [Alcedo et al., Cell 86: 221-232 (1996); van den Heuvel and Ingham, Nature 382: 547-551 (1996)]. La evidencia tanto genética como bioquímica apoyan un modelo de receptor en el que *Ptch* y *Smo* son parte de un complejo de receptores multicomponentes, Chen y Struhl, Cell 87: 553-63 (1996); Marigo et al., Nature 384: 176-9 (1996); Stone et al., Nature 384: 129-34 (1996). Tras la unión de *Hh* a *Ptch*, el efecto inhibitor normal de *Ptch* sobre *Smo* disminuye, permitiendo que *Smo* transduzca la señal de *Hh* a través de la membrana plasmática. La pérdida de mutaciones de función en el gen *Ptch* se ha identificado en pacientes con el síndrome del nevo de células basales (BCNS), una enfermedad hereditaria caracterizada por múltiples carcinomas de células basales (BCCs). Las mutaciones del gen *Ptch* disfuncionales también se han asociado con un gran porcentaje de tumores esporádicos de carcinoma de célula basal, Chidambaram et al., Cancer Research 56: 4599-601 (1996); Gailani et al. Nature Genet. 14: 78-81 (1996); Hahn et al., Cell 85: 841-51 (1996); Johnson et al., Science 272: 1668-71(1996); Unden et al., Cancer Res. 56: 4562-5 (1996), Wicking et al., Am. J. Hum. Genet. 60: 21-6 (1997). Se cree que la pérdida de función de *Ptch* causa una señalización de *Smo* incontrolada en carcinoma de células basales. De manera similar, se han identificado mutaciones de *Smo* activantes en tumores esporádicos de BCC (Xie et al., Nature 391: 90-2 (1998)), enfatizando el papel de *Smo* como subunidad de señalización en el complejo de receptores para *SHh*. Sin embargo, el mecanismo exacto por el cual *Ptch* controla la actividad de *Smo* aún no se ha aclarado y los mecanismos de señalización por los cuales la señal de *Hh* se transmite desde el receptor a dianas corriente abajo en la cascada también están por elucidar. El análisis epistático

genético en *Drosophila* ha identificado varios genes de polaridad de segmentos que parecen actuar como componentes del mecanismo de transducción de señales de *Hh*, Ingham, Curr. Opin. Genet. Dev. 5: 492-8 (1995); Perrimon, supra. Entre estos se incluyen una molécula de tipo quinesina, Costal-2 (*Cos-2*) [Robbins et al., Cell 90: 225-34 (1997); Sisson et al., Cell 90: 235-45 (1997)], una proteína designada como *fused* [Preat et al. Genetics 135: 1047-62 (1993); Therond et al., Proc. Natl Acad Sci. USA 93: 4224-8 (1996)], una molécula nueva con función desconocida designada como Supresor de *fused* [Pham et al., Genetics 140: 587-98 (1995); Preat. Genetics 132: 725-36 (1992)] y una proteína de dedos de zinc Ci. [Alexandre et al., Genes Dev. 10: 2003-13 (1996); Dominguez et al., Science 272: 1621-5 (1996); Orenic et al, Genes Dev. 4: 1053-67 (1990)]. Entre los elementos adicionales implicados en la señalización de *Hh* se incluyen el factor de transcripción CBP [Akimaru et al., Nature 386: 735-738 (1997)], el regulador negativo *slimb* [Jiang and Struhl, Nature 391: 493-496 (1998)] y el elemento de respuesta a *SHh* COUP-TFII [Krishnan et al., Science 278: 1947-1950 (1997)].

[0005] Los mutantes en *Cos-2* son embrionariamente letales y muestran un fenotipo similar a la sobreexpresión de *Hh*, incluyendo duplicaciones del componente central de cada segmento y dominio de expansión de genes sensibles a *Hh*. En cambio, los embriones mutantes para *fused* y Ci muestran un fenotipo similar a la pérdida de función de *Hh* incluyendo la delección de la parte posterior de cada segmento y la sustitución de la duplicación de la imagen de tipo especular de la parte anterior o cada segmento y sustitución de una duplicación de tipo especular de la parte anterior, Busson et al. Roux. Arch. Dev. Biol. 197: 221-230 (1988). Las caracterizaciones moleculares de Ci sugirieron que es un factor de transcripción que activa directamente los genes sensibles a *Hh*, tales como Wingless and Dpp, Alexandre et al., (1996) supra, Dominguez et al., (1996) supra. Así mismo, el análisis molecular de *fused* revela que está estructuralmente relacionado con serina-treonina quinasas y que tanto un dominio N-terminal quinasa intacto como una región reguladora C-terminal son necesarios para su función correcta. Preat et al., Nature 347: 87-9 (1990); Robbins et al., (1997), supra. Therond et al., Proc. Natl. Acad Sci. USA 93: 4224-8 (1996). En concordancia con las supuestas funciones opuestas de *Cos-2* y *fused*, las mutaciones de *fused* son suprimidas por mutantes de *Cos-2* y también por mutantes del Supresor de *fused*, Preat et al., Genetics 135: 1047-62 (1993). Sin embargo, mientras que las mutaciones nulas de *fused* y las mutaciones del dominio N-terminal quinasa se pueden suprimir totalmente por las mutaciones del Supresor de *fused*, las mutaciones en C-terminal de *fused* muestran un fuerte fenotipo de *Cos-2* en una base del Supresor de *fused*. Esto sugiere que el dominio quinasa de *fused* puede actuar como un activador constitutivo de la señalización de *SHh* cuando no está presente el Supresor de *fused*. Los estudios recientes han mostrado que el *fused*, *Cos-2* y Ci de *Drosophila* de 92 kDa están presentes en un complejo de multiproteínas asociado con microtúbulos y que la señalización de *Hh* conduce a la disociación de este complejo de los microtúbulos, Robbins et al, Cell 90: 225-34 (1997); Sisson et al. Cell 90: 235-45 (1997). Tanto *fused* como *Cos-2* se fosforilan en respuesta al tratamiento con *Hh*, Robbins et al., supra; Therond et al., Genetics 142: 1181-98 (1996), pero la quinasa o quinasas responsables de esta actividad o actividades aún deben caracterizarse. Hasta ahora, los únicos homólogos de vertebrados conocidos para estos componentes son miembros de la familia de proteínas Gli (por ejemplo, Gli-1, Gli-2 y Gli-3). Éstas son supuestos factores de transcripción de dedos de zinc que están estructuralmente relacionados con Ci. Entre éstas Gli-1 mostró ser un mediador candidato de la señal de *SHh* [Hynes et al., Neuron 15: 35-44 (1995), Lee et al., Development 124: 2537-52 (1997); Alexandre et al., Genes Dev. 10: 2003-13 (1996)], sugiriendo que el mecanismo de activación de genes en respuesta a *Hh* se puede conservar entre mosca y vertebrados. Para determinar si otros componentes de señalización en la cascada de *Hh* se conservan de forma evolutiva y examinar la función de *fused* en la cascada de señalización de *Hh* a nivel bioquímico, los solicitantes han aislado y caracterizado el ADNc de *fused* humano. La distribución tisular en el ratón indica que *fused* se expresa en tejidos sensibles a *SHh*. Los estudios bioquímicos demuestran que *fused* es una quinasa funcional. Los estudios funcionales proporcionan evidencias de que *fused* es un activador de Gli y que una forma negativa dominante es capaz de bloquear la señalización de *SHh* en embriones de *Xenopus*. Juntos, estos datos demostraron que *fused* está implicado directamente en la señalización de *Hh*.

[0006] Los solicitantes han identificado ADNc que codifica un polipéptido *fused* humano (*hfused*) y, de este modo, han proporcionado por primera vez una molécula *fused* de vertebrado.

CARACTERÍSTICAS DE LA INVENCION

[0007] En una realización, la presente invención proporciona un ácido nucleico aislado que comprende ADN que codifica: (i) un polipéptido *fused* humano que tiene por lo menos un 80% de identidad en la secuencia de aminoácidos con la secuencia de aminoácidos 1 a 1315 de la figura 1 (SEQ ID NO. 2); o (ii) un fragmento de polipéptido *fused* de vertebrado que tiene residuos de aminoácidos 1 a 260 de la Figura 1 (SEQ ID NO. 24). El ácido nucleico aislado de la presente invención codifica un polipéptido que tiene actividad biológica de *fused*, donde dicha actividad biológica de *fused* comprende una o más entre: la unión a y la influencia en la señalización de hedgehog; regulación de la patogénesis de carcinoma de células basales; fosforilación o modulación de la fosforilación de Gli; fosforilación de proteína básica de mielina, sustratos de proteína de levadura, sustratos de péptidos sintéticos o sustratos de polímeros.

[0008] En un aspecto, el ácido nucleico aislado comprende DNA que codifica un polipéptido *fused* que tiene por lo menos aproximadamente un 80%, preferiblemente por lo menos aproximadamente un 85%, más preferiblemente por lo menos aproximadamente un 90%, y lo más preferiblemente por lo menos aproximadamente un 95% de identidad en la secuencia de aminoácidos con un polipéptido que tiene los residuos de aminoácidos 1 a aproximadamente 1315 de la figura 1 (SEQ ID NO. 2). Preferiblemente, el grado más elevado de identidad en la secuencia aparece en el dominio quinasa (aminoácidos 1 a aproximadamente 260 (SEQ ID NO:24 tal como se muestra en la figura 1). Especialmente preferidos son las moléculas de ácido nucleico que contienen una secuencia codificante para una lisina en la posición de aminoácido 33. En otro aspecto, la presente invención proporciona un ácido nucleico aislado que comprende ADN que codifica el mismo polipéptido maduro codificado por el ADNc en el Depósito ATCC No. 209637 (designación: pRK5tkneo.h*Fused*-1272). En un aspecto adicional, la presente invención proporciona un ácido nucleico que comprende ADN que codifica el mismo polipéptido maduro codificado por el ADNc en el Depósito ATCC No. 209637 (designación: pRK5tkneo.h*Fused*-1272) o una secuencia que se hibrida al mismo bajo condiciones rigurosas y que codifica un polipéptido que tiene una actividad biológica de *fused*.

[0009] En otra realización, la presente invención proporciona un vector que comprende ADN que codifica un polipéptido *fused* de vertebrado. También se proporciona una célula huésped que comprende dicho vector. A modo de ejemplo, las células huésped pueden ser células de mamífero, (por ejemplo células CHO), células procariotas (por ejemplo, *E. coli*) o células de levadura (por ejemplo, *Saccharomyces cerevisias*). También se proporciona un proceso para producir un polipéptido *fused* de vertebrados y comprende cultivar células huésped bajo condiciones adecuadas para la expresión de *fused* de vertebrado y recuperar el mismo del cultivo celular.

[0010] En otra realización, la presente invención proporciona un polipéptido *fused* de vertebrado aislado tal como se establece en las reivindicaciones. Por ejemplo, la presente invención proporciona un polipéptido *fused* de vertebrado de secuencia nativa aislado, que en una realización es un *fused* humano que incluye una secuencia de aminoácidos que comprende los residuos 1 a aproximadamente 1315 de la Figura 1 (SEQ ID NO. 2). Se incluyen específicamente un polipéptido *fused* de vertebrado humano y otros nativos con o sin la metionina de iniciación. Alternativamente, la presente invención proporciona a polipéptido *fused* de vertebrado codificado por el ácido nucleico depositado bajo el número de acceso ATCC 209637.

[0011] Los polipéptidos *fused* descritos en la presente invención tienen actividad biológica *fused*, que comprende una o más entre:

- unión a e influencia en la señalización de hedgehog;
- regulación de la patogénesis de carcinoma de células basales;
- fosforilación o modulación de la fosforilación de Gli;
- fosforilación de proteína básica de mielina, sustratos de proteínas de levadura, sustratos de péptidos sintéticos o sustratos de polímeros.

[0012] En otra realización, la presente invención proporciona moléculas quiméricas que comprenden un polipéptido *fused* de vertebrado fusionado a un polipéptido o secuencia de aminoácidos heteróloga. Un ejemplo de dicha molécula quimérica comprende un polipéptido *fused* de vertebrado fusionado a una secuencia epítipo etiqueta o una región constante de una inmunoglobulina.

[0013] En la presente invención se describe un marcador de secuencia expresada (EST) que comprende las secuencias de nucleótidos identificadas en la figura 2 como 2515662 (SEQ ID NO. 3).

[0014] También en la presente invención se describen compuestos y métodos para desarrollar antagonistas contra y agonistas que inducen la modulación de *fused* de la señalización de Hedgehog. En particular, un antagonista de *fused* de vertebrado que bloquea, previene, inhibe y/o neutraliza el funcionamiento normal de *fused* en el mecanismo de señalización de SH, incluyendo tanto moléculas bioorgánicas pequeñas como nucleótidos antisentido.

[0015] Además, en la presente invención se describen variantes de corte y empalme alternativo de *fused* humano. En una realización adicional, la presente invención proporciona un método de cribado o análisis para identificar moléculas que modulan la activación de *fused* de la señalización de hedgehog. Preferiblemente, las moléculas previenen la interacción de *fused* con sus proteínas complejantes asociativas o evitan o inhiben la disociación de complejos. El análisis comprende la incubación de una mezcla que comprende *fused* y un sustrato (por ejemplo, *Gli*, *COUP-TFII*, *slimb*, *CBP*, *MBP*) con una molécula candidata y la detección de la capacidad de la molécula candidata de modular la fosforilación de *fused* a su sustrato. Las moléculas cribadas son preferiblemente candidatos de fármaco de molécula pequeña. En particular, el método se refiere a una técnica para cribar antagonistas de actividad biológica de *fused* que comprende:

- (a) exponer las células diana que expresan *fused* en un cultivo a un compuesto candidato; y
- (b) analizar los lisados celulares para valorar el nivel y/o identidad de fosforilación; o
- (c) valorar los cambios fenotípicos o funcionales en las células tratadas;

5 y comparar los resultados con las células de control que no se expusieron al compuesto candidato.

[0016] En otra realización, el método se refiere a una técnica de diagnóstico para determinar si un trastorno particular está modulado por la señalización de hedgehog, que comprende:

- 10 (a) cultivar células o tejidos de análisis;
- (b) administrar un compuesto que puede inhibir la señalización de hedgehog modulada por *fused*; y
- (c) medir el grado de atenuación de quinasa en el sustrato de *fused* en lisados celulares o efectos fenotípicos mediados por hedgehog en las células de análisis.

15 BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0017]

20 Las Figuras 1A-1F muestran la secuencia de nucleótidos (SEQ ID NO. 1) y aminoácidos derivada (SEQ ID NO. 2) de una secuencia nativa de polipéptido *fused* humano. Se incluyen el dominio quinasa (residuos 1 a aproximadamente 260) (SEQ ID No. 24) y el sitio de unión a ATP aproximadamente en la posición de aminoácido 33.

La figura 2 muestra la EST 2515662 (SEQ ID NO. 3) que se utilizó en la clonación de la secuencia de longitud completa humana.

25 Las Figuras 3A-3E muestran una comparación entre el *fused* humano y de *Drosophila* (SEQ ID NOS. 2 y 23, respectivamente). Los espacios introducidos para la alineación óptima se indican con guiones. Los aminoácidos idénticos están encuadrados. El residuo de lisina mutado en *fused*-DN (negativo dominante, lisina en la posición de aminoácido 33) se muestra con una estrella.

30 Las Figuras 4A-4F muestran la secuencia de DNA28494 (SEQ ID NO. 6) que era una variante por corte y empalme incorrecta de *fused* humano aislado de una biblioteca de pulmones fetales. Este clon contiene una potencial metionina de iniciación en la posición 116 seguido de un marco de lectura abierto de 1944 pb. Está presente un segundo marco de lectura desde aproximadamente la posición 2295 hasta 4349. Existe una diferencia de un nucleótido entre el clon DNA28495 (SEQ ID NO. 4) y el clon DNA28494 (SEQ ID NO. 6) localizado en el primer ORF en la posición 1863 del clon 28495 (SEQ ID NO. 4) (A vs. G) que cambia la secuencia codificante de una Gln a una Arg en la posición 583.

40 **[0018]** El primer marco de lectura abierto de DNA28494 (SEQ ID NO. 6) empieza en el residuo 115 y está seguido por un marco de lectura abierto largo de 630 aminoácidos.

Las Figuras 5A-5F muestran la secuencia de DNA28494 (SEQ ID NO. 6) que era otra variante de corte y empalme incorrecta de *fused* humano aislado de una biblioteca de pulmones fetales.

45 La Figura 6 es una transferencia western del producto de PCR de un epítipo etiqueta de DNA28495 (SEQ ID NOS. 5 & 21) y DNA28494 (SEQ ID NOS. 7 & 22). Se detectó una banda específica de 150 kDa en el residuo celular de células transfectadas con la construcción correspondiente al clon DNA28494 (SEQ ID NO. 6) y se pudo detectar una banda específica de aproximadamente 100 kDa para el clon DNA28495 (SEQ ID NO. 4) (Fig. 6). Estas bandas no estaban presentes en el control transfectado de simulación. La presencia de la banda de 100 kDa sugiere que los dos marcos de lectura abiertos de DNA21494 (SEQ ID NO. 6) se pueden empalmar juntos para dirigir la síntesis de una proteína grande de 150 kDa. La ausencia de esta banda para DNA28495 (SEQ ID NO. 4) sugirió que este clon aparentemente no se empalma correctamente.

55 La Figura 7 es un análisis de transferencia northern de *fused* humano (SEQ ID NO 1). Múltiples transferencias Northern de tejido fetal y adulto humano eran sondas con una sonda de ADNc de *fused* humano.

60 Las Figuras 8A-8F muestran una fotografía que muestra la hibridación in situ de tejidos embrionarios y adultos con *fused* (SEQ ID NO 1). Secciones sagitales de embriones de ratón E11.5 (Figura 8A) y E13.5 (Figura 8B). Sección coronal hasta la médula espinal de embriones de ratón E11.5 (Figura 8C) y E13.5 (Figura 8D). Sección sagital hasta ratón P1 (Figura 8E) y adulto (Figura 8F). Cp, plexo coroideo; hb, rombencéfalo; hip, formación hipocámpal; ht, coraón; hy, hipotálamo, kd, riñón; lg, pulmón; mb, mesencéfalo; md, derivado del intestino medio; mnd, componente

mandibular del primer arco branquial: sc, médula espinal; st, estómago; tec, tectum de mesencéfalo; vh, asta ventral de la médula espinal; vm, mesencéfalo ventral. Barras de escala: Figura 8A, 1,0 mm; Figura 8B, 1,62 mm; Figura 8C, 0,14 mm; Figura 8D, 0,17 mm; Figura 8E, 2,0 mm; y Figura 8F, 3,1 mm.

5 Las Figuras 9A-9C son una fotografía que muestra la hibridación *in situ* que muestra la presencia de ARNm de *fused* en niveles elevados en los testículos de ratones adultos (figura 9A). Una amplificación elevada revela las diferencias en los niveles de expresión de túbulos seminíferos (figura 9C). La hibridación de los testículos con una sonda de control de cadena de sentido a *fused* no produjo hibridación (figura 8B).

10 Las Figuras 10A-10B son un gráfico de barras que representa la activación de Gli por *fused*. (Figura 10A): se transfectaron células C3H10T1/2 con una p9XGliLus, luciferasa en ptkRenilla y *fused* o varios mutantes de *fused*. Las células se recogieron 48 horas después de la transfección y se analizó la actividad de luciferasa tal como se ha descrito en el ejemplo 7. (Figura 10B): Transactivación con *fused* de una construcción de informador Gli. Se cotransfectaron células C3H10T1/2 con una construcción informadora p9XGliLuc, luciferasa en ptkRenilla y un vector de expresión impulsado por CMV para *fused* o varios mutantes de *fused*. Las células se recogieron 48 horas después de la transfección y se analizó la actividad de luciferasa tal como se ha descrito en los ejemplos. Los datos representan el promedio de las determinaciones por duplicado.

20 Las Figuras 11A-11E son una fotografía que muestra que *fused*-DN (SEQ ID NO 25) inhibe la señalización de *SHh* en el desarrollo temprano de *Xenopus*. Se representan: (Figura 11A) Visión dorsal de embriones en la fase renacuajo. El embrión superior es la inyección con *fused*-DN (SEQ ID NO 25) y el embrión de abajo es el control; (Figura 11B) Visión lateral del embrión en fase renacuajo. El embrión superior es la inyección con *fused*-DN (SEQ ID NO 25) y el embrión de abajo es el control; (Figuras 11C & 11D) Tinción con Pax-6 de embriones de neurula de fase 16 inyectados con ADN de control y *fused*-DN (SEQ ID NO 25), respectivamente; (Figura 11E) Expresión de *SHh* en la placa del piso del embrión de control en fase de neurula (izquierda) o embrión inyectado con *fused*-DN (SEQ ID NO 25) (derecha).

30 La Figura 12 es una fotografía que confirma la actividad quinasa de *fused* (SEQ ID NO 2) y su activación de Gli. Se representan células 293 transfectadas con construcciones de *fused* etiquetadas con HA tal como se indica en el ejemplo 10 e inmunoprecipitadas con anticuerpos anti-HA y proteína A sefarosa. Las partículas de proteína A se sometieron a un ensayo de quinasa *in vitro* tal como se describe en el ejemplo 10 en presencia de MBP.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES PREFERIDAS

35 I. Definiciones

[0019] Los términos "*fused* de vertebrado" y "polipéptido *fused* de vertebrado", cuando se utilizan en la presente invención, comprenden *fused* de vertebrado de secuencia nativa y variantes de *fused* de vertebrado (que se desfinen posteriormente en la presente invención) que tienen actividad biológica de *fused*. *Fused* se puede aislar de variedad de fuentes, tales como tipos de tejidos humanos o de otras fuentes, o se pueden preparar mediante métodos recombinantes o sintéticos.

45 [0020] Un "*fused* de vertebrado de secuencia nativa" comprende un polipéptido que presenta la misma secuencia de aminoácidos que *fused* de vertebrado de origen natural. Dicho *fused* de vertebrado de secuencia nativa puede aislarse de la naturaleza o puede producirse por medios recombinantes y/o sintéticos. El término "*fused* de vertebrado de secuencia nativa" comprende específicamente formas truncadas naturales de *fused* vertebrado, formas variantes naturales (por ejemplo, formas de corte y empalme alternativo) y variantes alélicas naturales de *fused* de vertebrado. *Fused* de vertebrado nativo incluye, por ejemplo, *fused* en mamíferos, tales como humano, murino, bovino, porcino, equino, felino, canino, etc... y preferiblemente se refiere a humano. De esta manera, en una realización de la presente invención, el *fused* de vertebrado humano de secuencia nativa es un *fused* de vertebrado humano maduro o de longitud completa que comprende los aminoácidos 1 a 1315 tal como se muestra en la figura 1 con o sin la metionina de inicio en la posición 1.

55 [0021] El término "variante de *fused* de vertebrado" significa *fused* de vertebrado activo tal como se define a continuación que tiene por lo menos aproximadamente un 80% de identidad en la secuencia de aminoácidos con (a) una molécula de ADN que codifica un polipéptido *fused* de vertebrado, o (b) el complemento de la molécula de ADN de (a). En una realización particular, la variante de *fused* de vertebrado tiene por lo menos aproximadamente un 80% de homología en la secuencia de aminoácidos con el *fused* de vertebrado que tiene la secuencia de aminoácidos deducida (SEQ ID NO:2) mostrada en la figura 1 para un *fused* de vertebrado de secuencia nativa de longitud completa. Dichas variantes de *fused* de vertebrado incluyen, sin limitación, polipéptidos *fused* de vertebrado en los que se añaden, o eliminan uno o más residuos de aminoácidos en el extremo N o C-terminal de la secuencia

de la figura 1 (SEQ ID NO 2). Preferiblemente, el ácido nucleico o identidad en la secuencia de aminoácidos es por lo menos aproximadamente del 85%, más preferiblemente por lo menos aproximadamente del 90%, e incluso más preferiblemente por lo menos aproximadamente del 95%.

[0022] La expresión “porcentaje (%) de identidad en la secuencia de aminoácidos” referida a las secuencias de *fused* de vertebrado identificadas en la presente invención se define como el porcentaje de residuos aminoácidos en una secuencia candidata que son idénticos a los residuos aminoácidos en la secuencia de *fused* de vertebrado tras la alineación de las secuencias y la introducción de espacios, en caso necesario, para conseguir el máximo porcentaje de identidad en las secuencias, y sin considerar ninguna sustitución conservativa como parte de la identidad de secuencias. La alineación con fines de determinación del porcentaje de identidad en la secuencia de aminoácidos puede conseguirse de diversas maneras que se encuentran comprendidas dentro de los conocimientos del experto en la materia, por ejemplo utilizando software informático disponible para el público, tal como el programa BLAST, BLAST-2, ALIGN o Megalign (DNASTAR). Los expertos en la materia podrán determinar los parámetros apropiados para medir la alineación, incluyendo cualquier algoritmo necesario para conseguir la alineación máxima a lo largo de la longitud completa de las secuencias que se comparan. La expresión “porcentaje (%) de identidad en la secuencia de ácidos nucleicos” con respecto a las secuencias de *fused* de vertebrado identificadas en la presente invención se define como el porcentaje de nucleótidos en una secuencia candidata que son idénticos a los nucleótidos en la secuencia de *fused* de vertebrado tras la alineación de la secuencias y la introducción de espacios, en caso necesario, para conseguir el máximo porcentaje de identidad en las secuencias. La alineación con el fin de determinar el porcentaje de identidad en las secuencias de ácidos nucleicos puede conseguirse de diversas maneras que se encuentran comprendidas dentro de los conocimientos del experto en la materia, por ejemplo utilizando software informático disponible para el público, tal como el programa BLAST, BLAST-2, ALIGN o Megalign (DNASTAR). Los expertos en la materia pueden determinar los parámetros apropiados para medir la alineación, incluyendo cualquier algoritmo necesario para conseguir la máxima alineación a lo largo de la longitud completa de las secuencias que se comparan.

[0023] El término “etiquetado con epítipo”, cuando se utiliza en la presente invención, se refiere a un polipéptido quimérico que comprende el polipéptido *fused* de vertebrado o una parte del mismo, unido a un “polipéptido etiqueta”. El polipéptido etiqueta presenta suficientes residuos para proporcionar un epítipo contra el que puede prepararse un anticuerpo, aunque es suficientemente corto para no interferir con la actividad del polipéptido *fused* de vertebrado. El polipéptido etiqueta también es bastante único, de manera que el anticuerpo no reacciona de forma cruzada en grado sustancial con otros epítipos. Los polipéptidos etiqueta adecuados generalmente presentan por lo menos seis residuos de aminoácidos y habitualmente entre aproximadamente 8 y aproximadamente 50 residuos de aminoácidos (preferiblemente entre aproximadamente 10 y aproximadamente 20 residuos).

[0024] Tal como se utiliza en la presente invención, el término “inmunoadhesina” se refiere a moléculas de tipo anticuerpo que combinan la especificidad de unión de una proteína heteróloga (una “adhesina”) con las funciones efectoras de los dominios constantes de inmunoglobulina. Estructuralmente la inmunoadhesina comprende la fusión de una secuencia de aminoácidos con la especificidad de unión deseada, que es diferente que el sitio de reconocimiento y unión a antígeno de un anticuerpo (es decir, es “heterólogo”), y una secuencia de dominio constante de inmunoglobulina. La parte adhesina de una molécula de inmunoadhesina habitualmente es una secuencia de aminoácidos contigua que comprende por lo menos el sitio de unión de un receptor o de un ligando. La secuencia de dominio constante de inmunoglobulina en las inmunoadhesinas puede obtenerse a partir de cualquier inmunoglobulina, tal como los subtipos IgG-1, IgG-2, IgG-3 o IgG-4, IgA (incluyendo IgA-1 e IgA-2, IgE, IgD o IgM). La inmunoadhesión descrita en la literatura incluye fusiones del receptor de células T' [Gascoigne et al., Proc. Natl. Acad. Sci. USA 84: 2936-2940 (1987)]; CD4' [Capron et al., Nature 337: 525-53 (1989); Traunecker et al., Nature 339: 68-70 (1989); Zettmeissl et al., DNA Cell Biol. USA 9: 347-353 (1990); Bym et al., Nature 344, 667-670 (1990)]; L-selectina (receptor mensajero) [Watson et al. J. Cell. Biol. 110, 2221-2229 (1990); Watson et al., Nature 349, 164-167 (1991 1)]; CD44' [Aruffo et al., Cell 61. 1303-1313 (1990)]; CD28' y B7' [Linsley et al. J. Exp. Med 173, 721-730 (1991)]; CTLA-4* [Lisley et al., J. Exp. Med 174, 561-569 (1991)]; CD22* [Stamenkovic et al., Cell 66. 1133-1144 (1991)]; receptor de TNF [Ashkenazi et al., Proc. Natl. Acad. Sci. USA 88, 10535-10539 (1991); Lesslauer et al., Eur. J. Immunol. 27, 2883-2886 (1991); Peppel et al., J. Exp. Med. 174, 1483-1489 (1991)]; receptores de NP [Bennett et al. J. Biol. Chem. 266, 23060-23067 (1991)]; cadena α de receptor de IgE* [Ridgway y Gorman, J. Cell. Biol. 115, resumen 1448 (1991)]; receptor de HGF [Mark, M.R. et al., 1992, proporcionado], donde el asterisco (*) indica que el receptor es miembro de la superfamilia de inmunoglobulinas.

[0025] La “rigurosidad” de las reacciones de hibridación resulta fácilmente determinable por un experto ordinario en la materia, y generalmente es un cálculo empírico dependiente de la longitud de la sonda, de la temperatura de lavado y de la concentración de sales. En general, las sondas más largas requieren temperaturas más altas para la correcta hibridación, mientras que sondas más cortas requieren temperaturas inferiores. La hibridación generalmente depende de la capacidad del ADN desnaturalizado de hibridarse nuevamente cuando se encuentran presentes

cadenas complementarias en un medio cercano aunque inferior a su T_f (temperatura de fusión). Cuanto más elevado sea el grado de homología deseado entre la sonda y la secuencia hibridable, más elevada es la temperatura relativa que puede utilizarse. En consecuencia, se deduce que a temperaturas relativas más elevadas se tiende a que las condiciones de reacción resulten más rigurosas, mientras que a temperaturas inferiores este efecto es menor. Además, la rigurosidad también es inversamente proporcional a las concentraciones de sales. Para detalles adicionales y una explicación de la rigurosidad de las reacciones de hibridación, ver Ausubel et al., *Current Protocols in Muscular Biology*, 1995.

[0026] Las "condiciones rigurosas", tal como se define en la presente invención, pueden identificarse como aquéllas que: (1) utilizan una fuerza iónica baja y una temperatura elevada para el lavado, por ejemplo cloruro sódico 0,015 M/citrato sódico 0,0015 M/dodecil sulfato sódico al 0,1% a 50°C, (2) utilizan durante la hibridación un agente desnaturante, tal como formamida, por ejemplo formamida al 50% (v/v) con albúmina de suero bovino al 0,1%/Ficoll al 0,1%/polivinilpirrolidona al 0,1%/tampón de fosfato sódico 50 mM a pH 6,5 con cloruro sódico 750 mM, citrato sódico 75 mM a 42°C, (3) utilizan formamida al 50%, 5 x SSC (NaCl 0,75 M, citrato sódico 0,075 M), fosfato sódico 50 mM (pH 6,8), pirofosfato sódico al 0,1%, 5 x solución de Denhardt, ADN de esperma de salmón sonificado (50 µg/ml), SDS al 0,1% y sulfato de dextrano al 10% a 42°C, con lavados a 42°C en 0,2 x SSC (cloruro sódico/citrato sódico) y formamida al 50% a 55°C, seguido de un lavado de elevada rigurosidad consistente en 0,1 x SSC que contenía EDTA a 55°C.

[0027] Las "condiciones moderadamente rigurosas" pueden identificarse como describen Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Nueva York: Cold Spring Harbor Press, 1989, e incluyen el uso de una solución de lavado y condiciones de hibridación (por ejemplo, temperatura, concentración iónica y % de SDS) menos rigurosas que las descritas anteriormente. Un ejemplo de condiciones moderadamente rigurosas es una condición como la incubación durante toda la noche a 37°C en una solución que comprenda: formamida al 20%, 5X SSC (NaCl 0,75 M, citrato de trisodio 15 mM), fosfato de sodio 50 mM (pH 7,6), 5X solución de Denhardt, sulfato de dextrano al 10%, y 20 mg/ml de ADN de esperma de salmón fragmentado y desnaturado, seguida con un lavado de los filtros en 1X SSC a 37 - 50°C. Los expertos en la materia sabrán cómo ajustar la temperatura, concentración iónica, etc., según sea necesario para acomodar factores tales como longitud de la sonda y similares.

[0028] "Aislado," cuando se usa para describir los diversos polipéptidos descritos en la presente invención, significa un polipéptido que se ha identificado y separado y/o recuperado a partir de un componente de su entorno natural. Los componentes contaminantes de su entorno natural son materiales que típicamente interferirían con los usos diagnósticos o terapéuticos del polipéptido, y pueden incluir enzimas, hormonas, y otros solutos proteicos o no proteicos. En las realizaciones preferidas, el polipéptido se purificará (1) hasta un grado suficiente para obtener al menos 15 residuos de la secuencia de aminoácidos N-terminal o interna mediante el uso de un secuenciador de copa rotatoria, o (2) hasta su homogeneidad según SDS-PAGE, en condiciones reductoras o no reductoras, usando azul de Coomassie o, preferiblemente, tinción con plata. El polipéptido aislado incluye el polipéptido *in situ* dentro de células recombinantes, puesto que al menos un componente de su entorno natural del *fused* de vertebrado no estará presente. Sin embargo, ordinariamente, el polipéptido aislado se preparará mediante al menos una etapa de purificación.

[0029] Una molécula de ácido nucleico de *fused* vertebrado "aislads" es una molécula de ácido nucleico que se identifica y separa de por lo menos una molécula de ácido nucleico contaminante con la que se asocia normalmente en la fuente natural del ácido nucleico de *fused* vertebrado. Una molécula de ácido nucleico de *fused* de vertebrado aislada es aquella que es diferente a la forma o disposición en que se halla en la naturaleza. Por tanto, las moléculas de ácido nucleico de *fused* de vertebrado aisladas se diferencian de las correspondientes moléculas de ácido nucleico de *fused* de vertebrado nativa ya que éstas existen en células naturales.

[0030] El término "secuencias de control" se refiere a secuencias de ADN necesarias para la expresión de una secuencia codificante unida operativamente a un organismo huésped particular. Las secuencias de control que son adecuadas para procariotas, por ejemplo, incluyen un promotor, opcionalmente una secuencia operativa, y un sitio de unión a ribosomas. Se sabe que las células eucariotas utilizan promotores, señales de poliadenilación y potenciadores.

[0031] El ácido nucleico está "unido operativamente" cuando está colocado en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN para una presecuencia o un líder secretorio está unido operativamente a un ADN para un polipéptido, si éste se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o un potenciador están operativamente unidos a una secuencia codificante si éstos afectan la transcripción de la secuencia; o un sitio de unión de ribosomas está unido operativamente a una secuencia codificante si está ubicado de forma que facilita la traducción. Generalmente, "operativamente unido" significa que las secuencias de ADN que están unidas son contiguas, y, en el caso de un líder secretor, contiguas y

en pauta de lectura. Sin embargo, los potenciadores no tienen porque ser contiguos. La unión se consigue mediante la ligación en los sitios de restricción convenientes. Si tales sitios no existen, se usan adaptadores o enlazadores de oligonucleótidos sintéticos de acuerdo con la práctica convencional.

- 5 **[0032]** El término "anticuerpo" se usa en el sentido más amplio y abarca específicamente anticuerpos monoclonales individuales (incluyendo anticuerpos agonistas y antagonistas), composiciones de anticuerpos con especificidad poliepitópica, así como fragmentos de anticuerpo (por ejemplo, Fab, F(ab')₂ y Fv), siempre que muestren la actividad biológica deseada.
- 10 **[0033]** El término "anticuerpo monoclonal", tal como se usa en la presente invención, se refiere a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprenden la población son idénticos salvo por posibles mutaciones naturales que puedan estar en cantidades minoritarias. Los anticuerpos monoclonales son muy específicos, dirigiéndose contra un único sitio antigénico. Además, a diferencia de las preparaciones de anticuerpos convencionales (policlonales), que habitualmente incluyen distintos anticuerpos dirigidos contra distintos determinantes (epítopos), cada anticuerpo monoclonal se dirige contra un único determinante en el antígeno. Además de su especificidad, los anticuerpos monoclonales son ventajosas en que se sintetizan mediante el cultivo de hibridomas, no contaminados por otras inmunoglobulinas. El modificador "monoclonal" indica el carácter del anticuerpo por obtenerse de una población sustancialmente homogénea de anticuerpos y no debe interpretarse que se requiere la producción de anticuerpos mediante ningún procedimiento en particular. Por ejemplo, los anticuerpos monoclonales a utilizar según la presente invención pueden fabricarse mediante el procedimiento del hibridoma descrito por primera vez por Kohler y Milstein, Nature, 256: 495 (1975), o pueden fabricarse mediante procedimientos de ADN recombinante [véase, por ejemplo, la patente de EE.UU. 4.816.567 (Cabilly et al.)].
- 25 **[0034]** Los anticuerpos monoclonales de la presente invención incluyen específicamente anticuerpos "quiméricos" (inmunoglobulinas) en los que una fracción de la cadena pesada y/o ligera es idéntica u homóloga a las secuencias correspondientes en anticuerpos derivados de una especie en particular o pertenecientes a una clase o subclase de anticuerpo en particular, mientras el resto de la cadena o cadenas es idéntico u homólogo a las secuencias correspondientes en anticuerpos derivados de otras especies o pertenecientes a otra subclase o clase de anticuerpo, así como fragmentos de dichos anticuerpos, siempre y cuando muestren la actividad biológica deseada (patente U.S. 4.816.567 (Cabilly et al.); Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984)).
- 35 **[0035]** Las formas "humanizadas de anticuerpos no humanos (por ejemplo, murinos) son inmunoglobulinas quiméricas, cadenas de inmunoglobulinas o fragmentos de las mismas (tales como Fv, Fab, Fab', F(ab')₂ u otras subsecuencias de anticuerpos de unión a antígeno) que contienen una secuencia mínima derivada de una inmunoglobulina no humana. En su mayor parte, los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpo receptor) en las que los residuos de la región determinante de complementariedad (CDR) del receptor son sustituidos por residuos de una CDR de especies no humanas (anticuerpo donante) tales como ratón, rata, o conejo que presentan la especificidad, afinidad y capacidad deseadas. En algunos casos, se reemplazan residuos de la región de armazón (FR) de la inmunoglobulina humana por los residuos no humanos correspondientes. Además los anticuerpos humanizados pueden comprender residuos que no se encuentran en el anticuerpo receptor ni en las secuencias de CDR o armazón importadas. Estas modificaciones se realizan para refinar y optimizar la acción del anticuerpo. En general, el anticuerpo humanizado comprenderá sustancialmente todos de por lo menos uno, y habitualmente dos, dominios variables, en el cual todas o sustancialmente todas las regiones de CDR corresponden a las de una inmunoglobulina no humana y todas o sustancialmente todas las regiones FR son las de una secuencia consenso de inmunoglobulina humana. El anticuerpo humanizado también comprenderá de manera óptima por lo menos una parte de una región constante (Fc) de inmunoglobulina, habitualmente la de una inmunoglobulina humana. Para más detalles, véase Jones et al., Nature 321:522-525 (1986); Reichmann et al., Nature 332:323-329 (1988); Presta, Curr. Op. Strict. Biol.2:593-596 (1992) y la patente de Estados Unidos 325.539 (Winter) concedida el 6 de Julio de 1993.
- 45 **[0036]** "Activo" o "actividad" para los objetivos de la presente invención se refiere a una forma o formas de *fused* vertebrado que mantiene las actividades biológicas y/o inmunogénicas de *fused* de vertebrado nativo o natural. Una actividad preferida es la capacidad de unirse e influir, por ejemplo, bloquear o en cualquier caso modular, la señalización de hedgehog. La actividad implica preferiblemente la regulación de la patogénesis de carcinoma de células basales. Otra actividad biológica preferida es la capacidad de fosforilar o modular la fosforilación de Gli.
- 55 **[0037]** El término "antagonista" se utiliza en la presente invención en el sentido más amplio para incluir cualquier molécula que bloquea, evita, inhibe, neutraliza el funcionamiento normal de *fused* en el mecanismo de señalización de Hh. Una forma particular de antagonista incluye una molécula que interfiere con la interacción entre *fused* y sus proteínas de unión o complejantes. De una manera similar, el término "agonista" se utiliza en la presente invención

para incluir cualquier molécula que induce, aumenta o estimula el funcionamiento de *fused* en el mecanismo de señalización de Hh. Las moléculas adecuadas que afectan a la interacción de proteína-proteína de *fused* y sus proteínas de unión incluyen fragmentos del último o moléculas bioorgánicas pequeñas, por ejemplo, peptidomiméticos, que evitarán o aumentarán, según sea el caso, la interacción de la formación del complejo adecuado. Ejemplos no limitantes incluyen proteínas, péptidos, glicoproteínas, glicopéptidos, glicolípidos, polisacáridos, oligosacáridos, ácidos nucleicos, moléculas bioorgánicas, peptidomiméticos, agentes farmacológicos, y sus metabolitos, secuencias de control de la transcripción y traducción, y similares. Otra forma preferida de antagonista incluye nucleótidos antisentido que inhiben la transcripción correcta del *fused* de tipo salvaje. Las formas preferidas de antagonistas son moléculas pequeñas que se unen específicamente o bloquean la unión del sitio de unión a ATP de *fused*.

[0038] El término "modulación" o "modular" significa la sobrerregulación o subregulación de un mecanismo de señalización. Los procesos celulares bajo el control de la transducción de señales puede incluir, pero no limitación, la transcripción de genes específicos; funciones celulares normales, tales como metabolismo, proliferación, diferenciación, adhesión, apoptosis y supervivencia, así como procesos anormales, tales como la transformación, bloqueo de la diferenciación y metástasis.

[0039] Las técnicas de "reacción en cadena de la polimerasa" o "PCR", tal como se utiliza en la presente invención, se refiere en general a un procedimiento en el que cantidades muy pequeñas de una porción específica de ácido nucleico, ARN y/o ADN se amplifica tal como se ha descrito en la patente de Estados Unidos No. 4,683,195 concedida el 28 de julio de 1987. En general, necesita que esté disponible la información de la secuencia desde los extremos de la región de interés o más allá, de manera que se puedan diseñar cebadores de oligonucleótidos: estos cebadores serán idénticos o similares en la secuencia a las cadenas opuestas de la plantilla a amplificar. Los nucleótidos 5' terminales de los dos cebadores pueden coincidir con los extremos del material amplificado. Las secuencias de PCR forman secuencias de ADN genómico total, y ADNc transcrito a partir del ARN celular, bacteriófago, o plásmido, etc. Véase, en general, Mullis et al., Cold Spring Harbor Symp. Quant. Biol. 51: 263 (1987); Erlich, Ed., PCR Technology, (Stockton Press, NY, 1989). Tal como se utiliza en la presente invención. La PCR se considera uno, pero no el único, ejemplode una muestra de prueba de ácido nucleico que comprende la utilización de un ácido nucleico conocido como cebador y una ácido nucleico polimerasa para amplificar o generar una porción específica de ácido nucleico.

II. Composiciones y métodos de la presente invención

A. *Fused* de vertebrado de longitud completa

[0040] La presente invención proporciona secuencias de nucleótidos recién identificadas y aisladas que codifican polipéptidos referidos en la presente solicitud como *fused* de humano y vertebrado. En particular, los solicitantes han identificado y aislado ADNc que codifica un polipéptido *fused* de vertebrado, tal como se describe en detalle en los ejemplos siguientes. Utilizando los programas informáticos de alineación BLAST, BLAST-2 y FastA, los solicitantes hallaron un *fused* humano de secuencia nativa y longitud completa (mostrado en la figura 3 (SEC ID NO: 2)) que tiene una identidad del 28% en la secuencia de aminoácidos con *fused* de *Drosophila* (SEC ID NO 23). Por consiguiente, actualmente se cree que el *fused* humano descrito en la presente solicitud es un miembro recién identificado de la cascada de señalización de hedgehog.

[0041] La secuencia nativa de longitud completa del gen de *fused* de vertebrado humano, o partes de la misma, puede utilizarse como sondas de hibridación para una biblioteca de ADNc para aislar el gen de longitud completa o para aislar otros genes homólogos de vertebrado (por ejemplo, los que codifican variantes naturales de *fused* de vertebrado o *fused* de vertebrado de otras especies) que tienen una identidad de secuencia deseada con la secuencia de *fused* de vertebrado descrita en la figura 1 (SEC ID No: 1). Opcionalmente, la longitud de las sondas será de aproximadamente 20 hasta aproximadamente 50 bases. Las sondas de hibridación pueden derivarse de la secuencia de nucleótidos de la figura 1 (SEC ID No. 1) o de secuencias genómicas incluyendo promotores, elementos potenciadores e intrones de *fused* de vertebrado de secuencia nativa. A modo de ejemplo, un procedimiento de cribado comprenderá el aislamiento de la región codificante del gen de *fused* de vertebrado utilizando la secuencia de ADN conocida para sintetizar una sonda seleccionada de aproximadamente 40 bases. Las sondas de hibridación pueden marcarse mediante una serie de marcadores, incluyendo radionucleótidos como el ³²P o ¹⁵S, o marcadores enzimáticos tales como la fosfatasa alcalina acoplada a la sonda mediante sistemas de acoplamiento de avidina/biotina. Las sondas marcadas que tienen una secuencia complementaria con aquellas del gen de *fused* de vertebrado de la presente invención se pueden utilizar para cribar bibliotecas de ADNc humano, ADN genómico o ARNm para determinar a qué miembros de dichas bibliotecas se hibrida la sonda.

B. Variantes de *fused* de vertebrado

[0042] Además del *fused* de vertebrado de secuencia nativa y longitud completa descrita aquí, se contempla que se pueden preparar variantes de *fused* de vertebrado. Las variantes de *fused* de vertebrado se pueden preparar mediante la introducción de cambios de nucleótidos apropiados en un ADN de *fused* de vertebrado, o mediante la síntesis del polipéptido *fused* de vertebrado deseado. Los expertos en la materia entenderán que los cambios en los aminoácidos pueden alterar los procesos post-traduccionales del *fused* de vertebrado.

[0043] Las variaciones en el *fused* de vertebrado de secuencia de longitud completa o en varios dominios del *fused* de vertebrado descritos en la presente invención, se pueden realizar, por ejemplo, utilizando cualquiera de las técnicas y directrices para mutaciones conservativas y no conservativas establecidas, por ejemplo, en la Patente de Estados Unidos No. 5.364.934. Las variaciones pueden ser una sustitución, una delección o una inserción de uno o más codones que codifican el *fused* de vertebrado que dan lugar a un cambio en la secuencia de aminoácidos del *fused* de vertebrado en comparación con el *fused* de vertebrado de secuencia nativa. Opcionalmente, la variación es por sustitución de por lo menos un aminoácido por cualquier otro aminoácido en uno o más de los dominios del *fused* de vertebrado. Al determinar qué residuo de aminoácido se puede insertar, sustituir o eliminar sin afectar de forma adversa la actividad deseada, se pueden encontrar directrices mediante la comparación de la secuencia del *fused* de vertebrado con la de las moléculas de proteínas conocidas homólogas y minimizando el número de cambios en la secuencia de aminoácidos realizados en regiones con elevada homología. Las sustituciones de aminoácidos pueden ser el resultado de la sustitución de un aminoácido por otro aminoácido que tiene una estructura y/o propiedades químicas similares, tales como la sustitución de una leucina por una serina, es decir, sustituciones conservativas de aminoácidos. Las inserciones o eliminaciones pueden estar opcionalmente en el intervalo de aproximadamente 1 a 5 aminoácidos. La variación permitida se puede determinar realizando sistemáticamente inserciones, eliminaciones o sustituciones de aminoácidos en la secuencia y analizando en las variantes resultantes la actividad en el ensayo *in vitro* descrito en los siguientes ejemplos.

[0044] Las variaciones se pueden realizar utilizando procedimientos conocidos en la técnica tales como la mutagénesis mediada por oligonucleótidos (dirigida de sitio), el rastreo de alanina, y mutagénesis por PCR. Para fabricar el ADN variante del *fused* de variante se puede llevar a cabo sobre el ADN clonado una mutagénesis dirigida de sitio [Carter et al., Nucl. Acids. Res., 13: 4331 (1986); Zoller et al., Nucl. Acids. Res., 10: 6487 (1987)], mutagénesis de cassette [Wells et al., Gene, 34 315 (1985)], mutagénesis de selección de restricción [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] u otras técnicas conocidas.

[0045] El análisis de aminoácidos por rastreo también se puede utilizar para identificar uno o más aminoácidos a lo largo de una secuencia contigua. Entre los aminoácidos de rastreo preferidos están los aminoácidos relativamente pequeños y neutros. Entre dichos aminoácidos se incluyen alanina, glicina, serina y cisteína. La alanina es habitualmente un aminoácido de rastreo preferido de este grupo ya que elimina la cadena lateral más allá del carbono beta y es menos probable que altere la conformación de la cadena principal de la variante. La alanina es también habitualmente preferida ya que es el aminoácido más habitual. Además, se encuentra frecuentemente tanto en posiciones escondidas como expuestas [Creighton, The Proteins, (W.H. Freeman and Co., N.Y.), Chothia, J. Mol. Biol., 150:1 (1976)]. Si la sustitución de alanina no produce cantidades adecuadas de variante, se puede utilizar un aminoácido isotérico.

[0046] En la secuencia de *fused* humano descrito en la figura 1, el dominio quinasa está representado por los residuos de aminoácidos 1-260 (SEC ID NO: 24) de la cual la posición de lisina 33 parece ser necesaria para la unión a ATP y, de este modo, la actividad enzimática.

C. Modificaciones de *fused* de vertebrado

[0047] Las modificaciones covalentes de *fused* de vertebrado están incluidas en el alcance de la presente invención. Un tipo de modificación covalente incluye la reacción de residuos de aminoácidos marcados de *fused* de vertebrado con un agente derivatizante orgánico que es capaz de reaccionar con cadenas laterales seleccionadas o los residuos N- o C-terminales del *fused* de vertebrado. La derivatización con agentes bifuncionales es útil, por ejemplo, para reticular el *fused* de vertebrado con una matriz o superficie de soporte insoluble en agua para su utilización en el procedimiento para purificar anticuerpos anti-*fused* de vertebrado, y viceversa. Entre los agentes entrecruzadores utilizados habitualmente se incluyen, por ejemplo, 1,1-bis(diazoacetil)-2-feniletano, glutaraldehído, ésteres de N-hidroxisuccinimida, por ejemplo, ésteres con ácido 4-azido salicílico, imidoésteres homobifuncionales, incluyendo ésteres disuccinimidílicos, tales como 3,3'-ditiobis(succinimidilpropionato), maleimidas bifuncionales, tales como bis-N-maleimido-1,8-octano y agentes, tales como metil-3-[(p-azidofenil)ditio]propioimidato.

[0048] Otras modificaciones incluyen la desamidación de residuos glutamilo y asparaginilo a los correspondientes residuos glutamilo y aspartilo, respectivamente, la hidroxilación de prolina y lisina, la fosforilación de grupos hidroxilo

de residuos de serilo o treonilo, la metilación de los grupos α -amino de las cadenas laterales de lisina, arginina e histidina [T.E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, páginas 79-86 (1983)], acetilación de la amina N-terminal, y la amidación de cualquier grupo carboxilo C-terminal.

[0049] Otro tipo de modificación covalente de *fused* de vertebrado comprende la unión del polipéptido a uno del conjunto de polímeros no proteináceos, por ejemplo, polietilenglicol (PEG), polipropilenglicol o polioxialquilenos, de la forma establecida en las Patentes de Estados Unidos Nos: 4.640.835; 4.496.689; 4.301.144; 4.670.417; 4.791.192 ó 4.179.337. Se esperaría que dichas modificaciones incrementaren la vida media de las moléculas en circulación en un sistema mamífero. La mayor vida media de las moléculas *fused* podría ser útil bajo ciertas circunstancias, tales como cuando la variante *fused* se administra como agente terapéutico.

[0050] El *fused* de vertebrado de la presente invención también se puede modificar de manera que formen una molécula quimérica que comprende *fused* de vertebrado unido a otra secuencia de aminoácidos o polipéptido heterólogo. En una realización, dicha molécula quimérica comprende una fusión del *fused* de vertebrado con un polipéptido etiqueta, lo cual proporciona un epítipo al cual se puede unir selectivamente un anticuerpo anti-etiqueta. El epítipo etiqueta se sitúa generalmente en el extremo amino o carboxilo terminal del *fused* de vertebrado. La presencia de dichas formas etiquetadas con epítipo del *fused* de vertebrado se puede detectar utilizando un anticuerpo contra el polipéptido etiqueta. Además, la disposición del epítipo etiqueta permite que el *fused* de vertebrado se purifique fácilmente mediante purificación de afinidad utilizando un anticuerpo anti-etiqueta u otro tipo de matriz de afinidad que se une al epítipo etiqueta. En una realización alternativa, la molécula quimérica puede comprender una fusión del *fused* de vertebrado con una inmunoglobulina o una región particular de una inmunoglobulina. Para una forma bivalente de la molécula quimérica, dicha fusión podría ser con la región Fc o una molécula de IgG.

[0051] Normalmente, el extremo C-terminal de una secuencia de aminoácidos contigua de un dominio de unión a ligando-(IFN- γ) de un receptor IFN- γ se fusiona al extremo N-terminal de una secuencia de aminoácidos contigua de una región constante de inmunoglobulina, en lugar de la región o regiones variables, aunque también son posibles fusiones N-terminales.

[0052] Habitualmente, dichas fusiones mantienen por los menos los dominios bisagra, CH2 y CH3 funcionalmente activos de la región constante de una cadena pesada de inmunoglobulina. Las fusiones se realizan también al extremo C-terminal de la parte Fc de un dominio constante o inmediatamente N-terminal al CH1 de la cadena pesada o la correspondiente región de la cadena ligera. Esto se realiza normalmente mediante la construcción de la secuencia de ADN apropiada y su expresión en un cultivo de células recombinante. Alternativamente, se pueden sintetizar inmunoadhesinas según métodos conocidos.

[0053] El sitio preciso en el que se realiza la fusión no es crítico. Se conocen sitios particulares y se pueden seleccionar a efectos de optimizar la actividad biológica, la secreción o las características de unión de las inmunoadhesinas.

[0054] En una realización preferida, el extremo c-terminal de una secuencia de aminoácidos contigua que comprende el sitio o sitios de unión para IFN- γ se fusiona, en extremo N-terminal, a la parte C-terminal de un anticuerpo (en particular el dominio Fc), que contiene las funciones efectoras de una inmunoglobulina, por ejemplo, inmunoglobulina G1 (IgG-1). Tal como se ha mencionado aquí anteriormente, es posible fusionar la región constante de cadena pesada completa a la secuencia que contiene el sitio o sitios de unión. Sin embargo, más preferiblemente, se utiliza en la fusión una secuencia que empiece en la región bisagra justo en dirección 5' del sitio de división para papaína (que define la Fc de IgG químicamente; residuo 216, tomando el primer residuo de la región constante de la cadena pesada como el 114 [Kobet et al., supra], o sitios análogos de otras inmunoglobulinas). Aunque antes se pensó que en las inmunoadhesinas sería necesaria la cadena ligera de inmunoglobulina para una secreción eficiente de las proteínas de fusión de proteína heteróloga-cadena pesada, se ha observado que incluso las inmunoadhesinas que contienen la cadena pesada de IgG1 completa se secretan de manera eficaz en ausencia de cadena ligera. Dado que la cadena ligera es innecesaria, la secuencia del dominio constante de cadena pesada de inmunoglobulina utilizada en la construcción de las inmunoadhesinas de la presente invención puede estar carente de un sitio de unión a cadena ligera. Esto se puede conseguir eliminando o alterando suficientemente elementos de la secuencia de cadena pesada de inmunoglobulina a los que la cadena ligera se une normalmente, de manera que dicha unión ya no es posible. De este modo, el dominio CH I se puede eliminar completamente en ciertas realizaciones de las quimeras de receptor de IFN- γ -inmunoglobulina.

[0055] En una realización particularmente preferida, la secuencia de aminoácidos que contiene el dominio extracelular de un receptor de IFN- γ se fusiona a la región bisagra y CH2 y CH3: o los dominios CH1, bisagra, CH2 y CH3 de una cadena pesada de IgG-1, IgG-2, IgG-3 o IgG-4. La construcción de una estructura típica se describe en

el ejemplo 1.

[0056] En algunas realizaciones, las moléculas de receptor de IFN- γ -inmunoglobulina (inmunoadhesinas) se ensamblan como monómeros, dímeros o multímeros y particularmente como dímeros o tetrámeros. En general, estas inmunoadhesinas ensambladas tendrán estructuras unitarias conocidas similares a las de las inmunoglobulinas correspondientes. Una unidad estructural básica de cuatro cadenas (un dímero de dos parejas de cadena pesada-cadena ligera de inmunoglobulina) es la forma en la que existen IgG, IgA e IgE. En las inmunoglobulinas de peso molecular elevado se repite una unidad de cuatro cadenas; IgM existe en general como un pentámero de unidades de cuatro cadenas básicas mantenidas juntas mediante enlaces disulfuro. La globulina IgA, y ocasionalmente la globulina IgG, también pueden existir en una forma multimérica en suero. En el caso de multímeros, cada unidad de cuatro cadenas puede ser la misma o diferente.

[0057] No es necesario que la parte de inmunoglobulina completa de las quimeras receptor de IFN- γ -inmunoglobulina sean de la misma inmunoglobulina. Se pueden combinar varias partes de diferentes inmunoglobulinas y se pueden fabricar variantes y derivados de inmunoglobulinas nativas tal como se ha descrito aquí anteriormente con respecto a IFN- γ , a efectos de optimizar las propiedades de las moléculas de inmunoadhesina. Por ejemplo, se observó que las construcciones de inmunoadhesina en las que la bisagra de IgG1 se substituyó por las de IgG-3 eran funcionales y mostraban una farmacocinética comparable con la de inmunoadhesinas que comprenden la cadena pesada de IgG-1 completa.

[0058] En la técnica se conocen varios polipéptidos etiqueta y sus respectivos anticuerpos. Entre los ejemplos se incluyen etiquetas de poli-histidina (poli-his) o poli-histidina-glicina (poli-his-gly); el polipéptido etiqueta de gripe HA y su anticuerpo 12C45 [Field et al., Mol. Cell. Biol., 8: 2159-2165 (1988)]; la etiqueta c-myc y los anticuerpos 8F9, 3C7, 6E10, G4, B7 y 9E10 de la misma [Evan et al., Molecular and Cellular Biology, 5: 3610-3616 (1985)]; y la etiqueta de glicoproteína D (gD) del virus del Herpes Simplex y su anticuerpo [Paborsky et al., Protein Engineering, 3 (6): 547-553 (1990)]. Entre otros polipéptidos etiqueta se incluyen el péptido Flag [Hopp et al., BioTechnology, 6: 1204-1210 (1988)]; el péptido epítipo KT3 [Martin et al., Science, 255: 192-194 (1992)]; un péptido epítipo de α -tubulina [Skinner et al., J. Biol. Chem., 266: 15163-15166 (1991)]; y el péptido etiqueta de la proteína T7 del gen 10 [Lutz-Freyemuth et al., Proc. Natl. Acad. Sci. USA, 87: 6393-6397 (1990)]. Una etiqueta preferida es la etiqueta de la gripe HA.

D. Preparación de *fused* de vertebrado

[0059] La siguiente descripción se refiere principalmente a la producción de un *fused* de vertebrado particular mediante el cultivo de células transformadas o transfectadas con un vector que contiene ácido nucleico que codifica el *fused* de vertebrado. Naturalmente, se prevé que se puedan utilizar procedimientos alternativos, que se conocen bien en la técnica, para preparar el *fused* de vertebrado. Por ejemplo, la secuencia de *fused* de vertebrado, o partes de la misma, se pueden producir mediante síntesis directa de péptidos utilizando técnicas de fase sólida [véase, por ejemplo, Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85: 2149-2154 (1963)]. La síntesis de proteínas *in vitro* se puede realizar utilizando técnicas manuales o mediante automatización. La síntesis automatizada se puede realizar, por ejemplo, utilizando un Applied Biosystems Peptide Synthesizer (Foster City, CA) utilizando las instrucciones del fabricante. Se pueden sintetizar químicamente por separado varias partes del *fused* de vertebrado y combinarse utilizando procedimientos químicos o enzimáticos para producir el *fused* de vertebrado de longitud completa.

1. Aislamiento de ADN que codifica *fused* de vertebrado

[0060] El ADN que codifica *fused* de vertebrado se puede obtener a partir de una biblioteca de ADNc preparada a partir de tejido que se cree que posee el ARNm del *fused* de vertebrado y lo expresa a un nivel detectable. Por consiguiente, el ADN de *fused* de vertebrado humano se puede obtener convenientemente a partir de una biblioteca de ADNc preparada a partir de tejido humano, tal como se describe en los ejemplos. El gen que codifica *fused* de vertebrado también se puede obtener a partir de una biblioteca genómica o mediante métodos de síntesis de oligonucleótidos.

[0061] Las bibliotecas se pueden cribar con sondas (tales como anticuerpos para el *fused* de vertebrado u oligonucleótidos de por lo menos aproximadamente 20-80 bases) diseñados para identificar el gen de interés o la proteína codificada por el mismo. El cribado del ADNc o biblioteca genómica con la sonda seleccionada se puede realizar utilizando procedimientos estándar, tal como se describe en Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). Un medio alternativo para aislar el gen que codifica *fused* de vertebrado es utilizar la metodología de PCR [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

[0062] Los ejemplos siguientes describen técnicas para cribar una biblioteca de ADNc. Las secuencias de oligonucleótidos seleccionadas como sondas deberían ser de longitud suficiente y suficientemente inequívoca que se minimizan los falsos positivos. El oligonucleótido está preferiblemente marcado de manera que se puede detectar tras la hibridación a ADN en la biblioteca que se criba. Los procedimientos de marcado son bien conocidos en la técnica, e incluyen la utilización de marcador radioactivos como ATP marcado con ³²P, biotilación o marcaje enzimático. Las condiciones de hibridación, incluyendo la rigurosidad moderada y la rigurosidad elevada, se proporcionan en Sambrook et al., supra.

[0063] Las secuencias identificadas en dichos procedimientos de cribado de bibliotecas se pueden comparar y alinear con otras secuencias conocidas depositadas y disponibles en bases de datos públicos, tales como el Banco de Genes u otras bases de datos privadas de secuencias. La identidad de secuencia (a nivel de aminoácido o nucleótido) en las regiones definidas de la molécula o a lo largo de la secuencia de longitud completa se puede determinar a través de la alineación de secuencias utilizando programas informáticos, tales como BLAST, BLAST-2, ALIGN, DNASTar. e INHERIT que utilizan diversos algoritmos para medir la homología.

[0064] El ácido nucleico que tiene la secuencia de codificación de la proteína se puede obtener mediante el cribado del ADNc seleccionado o las bibliotecas genómicas utilizando la secuencia de aminoácidos deducida descrita en la presente invención por primera vez, y, si es necesario, utilizando procedimientos convencionales de extensión con cebadores tal y como se describe en Sambrook et al., supra, para detectar precursores y procesando intermedios de ARNm que no se han transcrito de forma inversa en ADNc.

2. Selección y transformación de células huésped

[0065] Las células huéspedes se transforman o transfectan con los vectores de clonación y expresión descritos aquí para la producción de *fused* de vertebrado y se cultivan en medios convencionales con nutrientes modificados según sea apropiado para inducir promotores, seleccionar transformantes, y amplificar los genes que codifican las secuencias deseadas. Las condiciones de cultivo, tales como el medio, temperatura, pH y similares, se pueden seleccionar por expertos en la materia sin excesiva experimentación. En general, principios, protocolos, y técnicas prácticas para maximizar la productividad de los cultivos celulares se pueden encontrar en Mammalian Cell Biotechnology: A Practical Approach. Butlder, ed. (IRL Press, 1991) y Sambrook y col., supra.

[0066] Los procedimientos para las transfecciones son conocidos por los expertos en la materia, por ejemplo CaPO y electroporación. Dependiendo de la célula huésped utilizada, la transformación se realiza utilizando técnicas estándar adecuadas a dichas células. El tratamiento con calcio que utiliza cloruro cálcico, tal y como se describe en Sambrook et al., supra, o la electroporación se utilizan generalmente para procariontes u otras células que contienen barreras sustanciales de pared celular. La infección con *Agrobacterium tumefaciens* se utiliza para la transformación de ciertas células vegetales, tal como describe Shaw et al., Gene, 23:315 (1983) y WO 89/05859 publicada el 29 de junio de 1989. Para las células de mamíferos sin dichas paredes celulares, se puede utilizar el procedimiento de precipitación con fosfato cálcico de Graham y van der Eb, Virology, 52: 456-457 (1978). En la Patente de Estados Unidos No. 4.399.216 se han descrito aspectos generales de transfecciones de sistemas de células huésped de mamíferos. Las transformaciones en levadura se llevan a cabo habitualmente según el procedimiento de Van Solingen et al., J. Bact., 130: 946 (1977) y Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76: 3829 (1979). Sin embargo, también se pueden utilizar otros procedimientos para introducir ADN en células, tales como mediante microinyección nuclear, electroporación, fusión de protoplasto bacteriano con células intactas, o policondones, por ejemplo, polibreno, poliornitina. Para varias técnicas para transformar células de mamífero, ver Keown et al., Methods in enzymology, 185:527-537 (1990) y Manssur et al., Nature, 336. 348-352 (1988).

[0067] Entre las células huésped adecuadas para la clonación o expresión del ADN en los vectores de la presente invención se incluyen células procariontes, levadura, o eucariotas superiores. Entre las procariontes adecuadas se incluyen, pero no se limitan a, eubacterias, tales como organismos Gram-negativo o Gram-positivo, por ejemplo, Enterobacteriaceae, tal como *E. coli*. Varias cepas de *E. coli* están disponibles públicamente, tales como la cepa de *E. coli* K12 MM294 (ATCC 31.446); *E. coli* X1776 (ATCC 31.537); cepa de *E. coli* W3110 (ATCC 27.325) y cepa de *E. coli* K5772 (ATCC 53.635).

[0068] Además de los procariontes, los microbios eucariotas, tales como hongos filamentosos o levaduras, son huéspedes de clonación o expresión adecuados para vectores que codifican el *fused* de vertebrado. El *Saccharomyces cerevisiae* es un microorganismo huésped eucariótico inferior utilizado habitualmente.

[0069] Las células huésped adecuadas para la expresión de *fused* de vertebrado derivan de organismos multicelulares. Entre los ejemplos de células de invertebrados se incluyen células de insectos, tales como *Drosophila*

S2 y Spodoptera Sf9, así como células vegetales. Entre los ejemplos de líneas celulares de huéspedes mamíferos útiles se incluyen células de ovario de hámster chino (CHO) y COS. Ejemplos más específicos incluyen la línea CV1 de riñón de mono transformada por SV40 (COS-7, ATCC CRL 1651); línea de riñón de embrión humano (células 293 ó 293 subclonadas para el crecimiento en cultivo de suspensión, Graham et al., *J. Gen Virol.*, 36:59 (1977)); células de ovario de hámster chino/DHFR (CHO, Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); células de sertoli de ratón (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); células de pulmón humano (W138, ATCC CCL 75); células de hígado humano (Hep G2, HB 8065) y tumor mamario de ratón (MMT 060562, ATCC CCL51). La selección de la célula huésped apropiado se estima que está dentro del conocimiento de la técnica.

3. Selección y utilización de un vector replicable

[0070] El ácido nucleico (por ejemplo, ADNc o ADN genómico), que codifica el *fused* de vertebrado se puede insertar en un vector replicable para la clonación (amplificación del ADN) o para la expresión. Existen varios vectores disponibles públicamente. El vector puede estar, por ejemplo, en forma de plásmido, cósmido, partícula viral, o fago. La secuencia de ácidos nucleicos apropiada se puede insertar en el vector mediante una serie de procedimientos. En general, el ADN se inserta en un sitio o sitios de endonucleasa de restricción apropiados utilizando técnicas conocidas en el sector. Los componentes de los vectores incluyen generalmente, pero no se limitan a, una o más secuencias señal, un origen de replicación, uno o más genes marcadores, un elemento potenciador, un promotor y una secuencia de terminación de la transcripción. La construcción de vectores adecuados que contienen uno o más de estos componentes utiliza técnicas de unión estándar que son conocidas por un experto en la materia.

[0071] Tanto los vectores de expresión como de clonación contienen una secuencia de ácidos nucleicos que permite la replicación del vector en una o más células huésped seleccionadas. Dichas secuencias son conocidas para un conjunto de bacterias, levadura, y virus. El origen de la replicación del plásmido pBR322 es adecuado para la mayoría de bacterias gram-negativas, el origen del plásmido 2 μ es adecuado para la levadura, y orígenes virales varios (SV40, polioma, adenovirus, VSV o BPV) son útiles para vectores de clonación en células de mamíferos. Un vector de expresión replicable preferido es el plásmido pRK5. Holmes et al., *Science*, 253:1278-1280 (1991).

[0072] Los vectores de clonación y expresión contendrán habitualmente un gen de selección, también denominado como marcador seleccionable. Los genes de selección habituales codifican proteínas que (a) confieren resistencia a antibióticos u otras toxinas, por ejemplo, ampicilina, neomicina, metotrexato o tetraciclina, (b) complementan las deficiencias auxotróficas, o (c) suministran nutrientes críticos no disponibles del medio complejo, por ejemplo, el gen que codifica la D-alanina racemasa para *Bacilli*.

[0073] Un ejemplo de marcadores seleccionables adecuados para células de mamíferos son aquellos que permiten la identificación de células competentes para captar el ácido nucleico de *fused* de vertebrado, tal como DHFR o timidina quinasa. Una célula huésped apropiada cuando se utiliza DHFR de tipo salvaje es la línea de células CHO deficiente en actividad de DHFR, preparada y propagada tal y como se describe por Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980). Un gen de selección adecuado para su utilización en la levadura es el gen *trp1* presente en el plásmido de la levadura YRp7 [Stinchcomb et al., *Nature*, 282:39 (1979); Kingsman et al., *Gene*, 7:141 (1979); Tschemper et al., *Gene*, 10:157 (1980)]. El gen *trp1* proporciona un marcador de selección para una cepa mutante de levadura que carece de la capacidad de crecer en triptófano, por ejemplo, ATCC No. 44076 o PEP4-1 [Jones, *Genetics*, 85:12 (1977)].

[0074] Los vectores de clonación y expresión contienen habitualmente un promotor unido operativamente a la secuencia de ácido nucleico que codifica de *fused* de vertebrado para dirigir la síntesis de ARNm. Son conocidos promotores reconocidos por un conjunto de células huésped potenciales. Entre los promotores adecuados para utilizar con huéspedes procariotas se incluyen los sistemas de promotores de β -lactamasa y lactosa [Chang et al., *Nature*, 275:615 (1978); Goeddel et al., *Nature*, 281:544 (1979)], fosfatasa alcalina, un sistema de promotor de triptófano (*trp*) [Goeddel, *Nucleic Acids Res.*, 8:4057 (1980); EP 36.776], y promotores híbridos, tales como el promotor tac [deBoer et al., *Proc. Natl. Acad. Sci. USA*, 80:21-25 (1983)]. Los promotores para utilizar en sistemas bacterianos también contendrán una secuencia de Shine-Dalgarno (S.D.) unida operativamente al ADN que codifica *fused* de vertebrado.

[0075] Entre los ejemplos de secuencias promotoras adecuadas para su utilización con huéspedes de levadura se incluyen promotores para 3-fosfoglicerato quinasa [Hitzeman et al., *J. Biol. Chem.*, 255:2073 (1980)] u otros enzimas glucolíticos [Hess et al., *J. Adv. Enzyme Reg.*, 7:149 (1968); Holland, *Biochemistry*, 17:4900 (1978)], tal como enolasa, gliceraldehído-3-fosfato deshidrogenasa, hexoquinasa, piruvato descarboxilasa, fosfofructoquinasa, glucosa-6-fosfato isomerasa, 3-fosfoglicerato mutasa, piruvato quinasa, triosafosfato isomerasa, fosfoglucoisomerasa y glucoquinasa.

[0076] Otros promotores de levaduras, que son promotores inducibles que tienen la ventaja adicional de una transcripción controlada por las condiciones de crecimiento, son las regiones promotoras para alcohol deshidrogenasa 2, isocitocromo C, fosfatasa ácida, enzimas degradativas asociadas con el metabolismo del nitrógeno, metalotioneína, gliceraldehído-3-fosfato deshidrogenasa, y enzimas responsables de la utilización de maltosa y galactosa. Los vectores y promotores adecuados para la utilización en la expresión en levaduras se describen en detalle en EP 73.657.

[0077] La transcripción del *fused* de vertebrado a partir de vectores en células huésped de mamíferos está controlada, por ejemplo, mediante promotores obtenidos de los genomas de virus, tales como virus del poliovirus, virus de la viruela aviar (Patente UK 2.211.504 publicada el 5 de julio de 1989), adenovirus (tal como el Adenovirus 2), el virus de papiloma bovino, virus del sarcoma aviar, citomegalovirus, un retrovirus, virus de la hepatitis-B y virus de simio 40 (SV40), de promotores de mamíferos heterólogos, por ejemplo, el promotor de actina o un promotor de inmunoglobulina, y de promotores de choque térmico, siempre y cuando dichos promotores sean compatibles con los sistemas de células huésped.

[0078] Se puede incrementar la transcripción de un ADN que codifica el *fused* de vertebrado por eucariotas superiores mediante la inserción de una secuencia de potenciador en el vector. Los potenciadores son elementos que actúan en cis de ADN, habitualmente aproximadamente de 10 a 300 pb, que actúan en un promotor para aumentar su transcripción. Actualmente se conocen muchas secuencias de potenciadores de genes de mamíferos (globina, elastasa, albúmina, α -fetoproteína e insulina). Habitualmente, sin embargo, se utilizará un potenciador de un virus de célula eucariota. Entre los ejemplos se incluyen el potenciador de SV40 en la cara tardía del origen de replicación (pb 100-270), el potenciador de promotor temprano de citomegalovirus, el potenciador de poliovirus en la cara tardía del origen de replicación, y los potenciadores de adenovirus. El potenciador se puede cortar y empalmar ("splice") en el vector en una posición 5' ó 3' con respecto a la secuencia codificante del *fused* de vertebrado, pero se sitúa preferiblemente en un sitio 5' con respecto al promotor.

[0079] Los vectores de expresión utilizados en células huéspedes eucariotas (levadura, hongos, insectos, plantas, animales, humanos, o células nucleadas de otros organismos multicelulares) también contendrán secuencias necesarias para la terminación de la transcripción y para la estabilización del ARNm. Dichas secuencias están disponibles habitualmente de las regiones no traducidas 5' y, ocasionalmente 3', de ADNs o ADNcs eucariotas o virales. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la parte no traducida del ARNm que codifica el anticuerpo el *fused* de vertebrado.

[0080] En Gething et al., *Nature*, 293:620-625 (1981); Mantei et al., *Nature*, 281:40-46 (1979); EP 117.060 y EP 117.058 se describen adicionalmente otros procedimientos, vectores, y células huésped adecuados para la adaptación a la síntesis del *fused* de vertebrado en cultivos de células de vertebrados recombinantes.

4. Detección de la amplificación/expresión de los genes

[0081] La amplificación y/o expresión de los genes se puede medir en una muestra directamente, por ejemplo, mediante transferencia Southern convencional, transferencia Northern convencional para cuantificar la transcripción de ARNm [Thomas, *Proc. Natl. Acad. Sci. USA*, 77:5201-5205 (1980)], transferencia de puntos (análisis de ADN), o hibridación *in situ*, utilizando una sonda marcada apropiadamente, basada en las secuencias proporcionadas en la presente invención. Alternativamente, se pueden utilizar anticuerpos que pueden reconocer dobles cadenas específicas, incluyendo dobles cadenas de ADN, dobles cadenas de ARN, dobles cadenas híbridas de ADN-ARN o dobles cadenas de ADN-proteína. Los anticuerpos a su vez se pueden marcar y el ensayo se puede llevar a cabo cuando la doble cadena está unida a una superficie, de manera que tras la formación de la doble cadena en la superficie, se puede detectar la presencia de anticuerpos unidos a la doble cadena.

[0082] La expresión génica, alternativamente, se puede medir mediante procedimientos inmunológicos, tales como tinción inmunohistoquímica de células o secciones de tejido y el ensayo de cultivo de células o fluidos corporales, para cuantificar directamente la expresión del producto génico. Los anticuerpos útiles para la tinción inmunohistoquímica y/o el ensayo de fluidos de muestra pueden ser monoclonales o policlonales, y se pueden preparar en cualquier mamífero. Convenientemente, los anticuerpos se pueden preparar contra un polipéptido *fused* de vertebrado de secuencia nativa o contra un péptido sintético basado en las secuencias de ADN proporcionadas en la presente invención o contra una secuencia exógena fusionada a ADN de *fused* de vertebrado y que codifican un epítipo de anticuerpo específico.

5. Purificación de polipéptido

[0083] Las formas de *fused* de vertebrado se pueden recuperar de los lisados de células huésped. Si están unidas a

membrana, se pueden liberar de la membrana utilizando una solución de detergente adecuada (por ejemplo, Triton X-100) o mediante división enzimática. Las células utilizadas en la expresión del *fused* de vertebrado se pueden romper mediante diversos medios físicos o químicos, tales como ciclos de congelación-descongelación, sonicación, destrucción mecánica, o agentes para lisar células.

[0084] Se puede desear purificar el *fused* de vertebrado a partir de proteínas o polipéptidos de células recombinantes. Los siguientes procedimientos son ejemplos de procedimientos de purificación adecuados: mediante fraccionamiento en una columna de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía sobre sílice o una resina de intercambio catiónico, tal como DEAE; "cromatofocusing"; SDS-PAGE; precipitación con sulfato amónico; filtración en gel utilizando, por ejemplo, Sephadex G-75; columnas de proteína A sefrosa para eliminar contaminantes, tales como IgG, y columnas quelantes de metales para unir formas etiquetadas con epítipo del anticuerpo *fused* de vertebrado. Se pueden utilizar varios métodos de purificación de proteínas y dichos métodos son conocidos en la técnica y se describen, por ejemplo, en Deutscher, *Methods in Enzymology*, 182 (1990); Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, Nueva York (1982). La etapa o etapas de purificación seleccionadas dependerán, por ejemplo, de la naturaleza del proceso de producción utilizado y el *fused* de vertebrado concreto producido.

E. Uso para *fused* de vertebrado

(1) *Fused* es un mediador universal de la señalización de Hh

[0085] La molécula de longitud completa de *fused* humano de la figura 1 (SEC ID No. 1) codifica una proteína con un peso molecular predicho de 150 kDa que es significativamente mayor que el *fused* de *Drosophila* (100 kDa, *dfused* (SEQ ID NO 23)). El *fused* humano (*hfused*) muestra una notable homología con el homólogo de *Drosophila* en el dominio quinasa, pero poca homología con el *dfused* o cualquier otra proteína conocida sobre los restantes ≈1000 aminoácidos. El dominio quinasa se extiende desde el residuo 1 hasta aproximadamente el residuo 260, tal como se representa en la figura 1 (SEQ ID NOS. 24 & 2). Esta divergencia en el extremo C-terminal de las moléculas es inesperada dado que el extremo C-terminal de la molécula de *Drosophila* es necesaria para su actividad, Preat et al., *Nature* 347: 87-9 (1990). Un sitio de unión a ATP se encuentra aproximadamente en la posición de aminoácido 33 y es necesario para la actividad de quinasa.

[0086] Estudios anteriores en *Drosophila* indican que *dfused* es necesario para que tenga lugar la señal de Hh, pero no han dirigido la cuestión a si el *fused* es suficiente para activar este sistema de señalización. Tal como se describe en los ejemplos, los solicitantes han utilizado en la presente invención un elemento de unión a ADN de Gli presente en el promotor HNF3β, delante de un mediador de luciferasa de la cascada de Hh, que demuestra claramente que el *fused* solo es capaz de activar la transcripción mediada por Gli en este sistema. También es evidente que tanto un dominio de quinasa intacto como un dominio no catalítico C-terminal intacto son necesarios para esta activación, lo cual apoya la noción de que *fused* actúa como quinasa y que el extremo C-terminal puede jugar un papel en el reconocimiento del sustrato o en la regulación de la actividad quinasa.

[0087] Los solicitantes han demostrado en esta solicitud que *hfused* es una quinasa que es capaz de fosforilar sustratos artificiales, tales como MBP. Sin embargo, la identidad del sustrato fisiológico para *hfused* aún está por determinar. Un candidato obvio es el propio Gli-I, ya que la fosforilación de Gli-I por *hfused* se puede detectar *in vitro*.

[0088] Para determinar si *fused* humano es esencial para la señalización de Hh en vertebrados, se construyó un mutante mediante la alteración de una lisina conservada en el sitio de unión a ATP (aproximadamente el residuo de aminoácido 33). Habitualmente, dichos mutantes actúan como inhibidor de la correspondiente quinasa de tipo salvaje mediante el bloqueo del acceso al sustrato y/o factores reguladores, He et al., *Nature* 374: 617-22 (1995). Cuando se sobreexpresó en embriones de *Xenopus* en la etapa celular 2, el fenotipo más destacable fue la presencia de ojos con *fused* en aproximadamente el 30% de los embriones inyectados. Varias líneas de evidencia indican que este fenotipo es probable que resulte de la inhibición de la señalización de Hh. En primer lugar, los "knockout" de *SHh* muestran un fenotipo de ciclopía atribuido recientemente a mutaciones en el gen de *SHh*, Chiang et al., *Nature* 383: 407-13 (1996). En segundo lugar, los embriones del pez zebra (cíclopes) con una expresión reducida de *SHh* o inyectados con una forma constitutivamente activa de PKA, un regulador negativo del mecanismo de Hh, son cíclopes. En tercer lugar, se ha observado que *SHh*, que emana de la placa precordial, inhibe la expresión de Pax-6, un factor de transcripción clave requerido para el desarrollo del ojo en el centro de un campo ocular continuo, Ekker et al., *Curr. Biol.* 5: 944-55 (1995); Li et al., *Development* 124: 603-15 (1997); Macdonald et al., *Development* 121: 3267-78 (1995). Finalmente, la tinción para embriones con Pax-6 inyectados con *fused*-DN reveló un campo de expresión único sugiriendo la incapacidad de *SHh* de emenar de la placa precordial para subregular la expresión de Pax-6 en el centro del campo ocular.

[0089] Para confirmar la posición de *fused* en el mecanismo de señalización de Hh, la expresión de *SHh* en la placa del piso de embriones de *Xenopus* inyectados con *hfused*-DN se podía rescatar mediante la coinfección de Gli-I. Esto sugiere que *fused* actúa en asociación con Gli en el mecanismo de señalización de *SHh*.

[0090] La distribución en tejido de *fused* muestra que se expresa en todas las células sensibles a *SHh*. En particular, su patrón de expresión se solapa bien con Ptch, el componente de la unión del receptor Hh que es en sí un gen diana del mecanismo de señalización de *SHh*. Estos datos sugieren que *fused* está implicado en la mediación de una amplia variedad del efecto que *SHh* tiene en diferentes tejidos. Funcionalmente, esto se observó de nuevo en embriones de rana donde *fused*-DN inhibió el desarrollo del ojo, asó como la expresión de *SHh* en la placa del piso.

[0091] *hfused*-DN también parece afectar al desarrollo normal de tejidos, tales como los intestinos de la rana (que está regulada por la Hh india). Esto, combinado con el hecho de que *fused* se expresa en los intestinos y testículos, los sitios de la acción de *lHh* y *DHh* respectivamente, sugieren que *fused* puede ser un mediador universal de señalización para todos los miembros de la familia de proteínas de Hh.

[0092] Se observaron niveles muy elevados de ARNm de *fused* en células germinales, el desarrollo del cual parece estar regulado por *DHh*. Los ratones mutantes homocigóticos para *DHh* no son capaces de desarrollar células germinales y son viables, pero estériles (Bitgood et al., Curr. Biol. 6: 298-304 (1996)). Sin embargo, Patched, un receptor de Hedgehog se expresa en células Leydig intestinales y no en células germinales donde se expresa *fused*, Bitgood et al. supra. Esta discrepancia sugiere que puede haber receptores de hedgehog adicionales.

[0093] Los solicitantes han mostrado en los ejemplos que *hfused* de tipo salvaje es capaz de activar Gli en un ensayo de informador. Además, la expresión de *SHh* en la placa del piso de los embriones de ranas inyectados con *hfused*-DN se podía rescatar mediante la coinfección de Gli-I. Conjuntamente, estas observaciones son consistentes con la afirmación de que *fused* actúa cascada debajo de Smo y cascada arriba de Gli en este mecanismo de señalización, que es consistente con la evidencia genética en *Drosophila* hasta la fecha.

(2) Usos generales para *fused* de vertebrado

[0094] Las secuencias de nucleótidos (o su complemento) que codifican *fused* de vertebrado tienen varias aplicaciones en el sector de la biología molecular, incluyendo usos como sondas de hibridación, en mapeo cromosómico y génico y en la generación de ARN y ADN anti-sentido. El ácido nucleico de *fused* de vertebrado también será útil para la preparación de polipéptidos *fused* de vertebrado mediante las técnicas recombinantes descritas en la presente invención.

[0095] El gen de *fused* de vertebrado de secuencia nativa y longitud completa, o partes del mismo, puede utilizarse como sondas de hibridación para una biblioteca de ADNc para aislar el gen de longitud completa o para aislar otros genes (por ejemplo, los que codifican variantes naturales de *fused* de vertebrado o *fused* de vertebrado de otras especies) que tienen una identidad de secuencia deseada con la secuencia de *fused* de vertebrado descrita en la figura 1 (SEC ID No. 1). Opcionalmente, la longitud de las sondas será de aproximadamente 20 hasta aproximadamente 50 bases. Las sondas de hibridación pueden derivarse de la secuencia de nucleótidos de la figura 1 (SEC ID No. 1) o a partir de secuencias genómicas incluyendo promotores, elementos potenciadores e intrones de *fused* de vertebrado de secuencia nativa. A modo de ejemplo, un procedimiento de cribado comprenderá el aislamiento de la región codificante del gen de *fused* de vertebrado utilizando la secuencia de ADN conocida para sintetizar una sonda seleccionada de aproximadamente 40 bases. Las sondas de hibridación pueden marcarse mediante una serie de marcadores, incluyendo radionucleótidos como el ³²P o ¹⁵S, o marcadores enzimáticos tales como la fosfatasa alcalina acoplada a la sonda mediante sistemas de acoplamiento de avidina/biotina. Las sondas que tienen una secuencia complementaria con aquellas del gen de *fused* de vertebrado de la presente invención se pueden utilizar para cribar bibliotecas de ADNc humano, ADN genómico o ARNm para determinar a qué miembros de dichas bibliotecas se hibrida la sonda. En los Ejemplos de más adelante se describen técnicas de hibridación con mayor detalle.

[0096] Las sondas también pueden emplearse en técnicas de PCR para generar un grupo de secuencias para la identificación de secuencias de *fused* de vertebrado estrechamente relacionadas.

[0097] También pueden utilizarse secuencias de nucleótidos que codifican un *fused* de vertebrado para construir sondas de hibridación para localizar el gen que codifica el *fused* de vertebrado y para el análisis genético de individuos con trastornos genéticos. Las secuencias de nucleótidos proporcionadas en la presente invención pueden localizarse en un cromosoma y en regiones específicas de un cromosoma utilizando técnicas conocidas, como por ejemplo la hibridación *in situ*, análisis de unión contra marcadores cromosómicos conocidos, y el cribado por

hibridación con bibliotecas.

[0098] El polipéptido *fused* de vertebrados se puede utilizar en ensayos para identificar las otras proteínas o moléculas implicadas en la complejación con *fused* que en último término da lugar a la modulación de la señalización de hedgehog. Alternativamente, estas moléculas pueden modular la fosforilación con quinasa de *fused* de su sustrato. Mediante dichos procedimientos, pueden identificarse inhibidores de la interacción de la unión. También pueden utilizarse proteínas implicadas en dichas interacciones de unión para cribar inhibidores de péptidos o moléculas pequeñas o agonistas de la interacción de unión. Además, puede utilizarse el sustrato de *fused* de vertebrado para aislar proteínas complejantes correlativas. Pueden diseñarse ensayos de cribado para hallar compuestos de partida que mimetizan la actividad biológica de un *fused* de vertebrado nativo o hallar aquellos que actúan como sustrato para el *fused* de vertebrado. Dichos ensayos de cribado incluirán ensayos susceptibles de cribado de alto rendimiento de bibliotecas químicas, haciendo que sean especialmente adecuados para identificar candidatos de fármacos de molécula pequeña. Dichos inhibidores de molécula pequeña podrían bloquear la acción enzimática de *fused* e inhibir, de este modo, la señalización de hedgehog. Las moléculas pequeñas contempladas incluyen compuestos sintéticos orgánicos o inorgánicos. Pueden realizarse los ensayos en una serie de formatos, incluyendo ensayos de unión proteína-proteína, ensayos de cribado bioquímico, inmunoensayos y ensayos basados en células, que están bien caracterizados en el sector.

[0099] También pueden utilizarse ácidos nucleicos que codifican *fused* de vertebrado o sus formas modificadas para generar animales transgénicos o animales "knock out" que, a su vez, son útiles en el desarrollo y el cribado de reactivos terapéuticamente útiles. Un animal transgénico (por ejemplo, un ratón o una rata) es un animal que tiene células que contienen un transgén, el cual se introdujo en el animal o en un progenitor del animal en una fase prenatal, por ejemplo, embrionaria. Un transgén es un ADN el cual está integrado dentro del genoma de una célula a partir de la cual se desarrolla un animal transgénico. En una realización, puede utilizarse el ADNc que codifica *fused* de vertebrado para clonar el ADN genómico que codifica *fused* de vertebrado de acuerdo con las técnicas establecidas y las secuencias genómicas utilizadas para generar animales transgénicos que contienen células que expresan ADN que codifica *fused* de vertebrado. Los procedimientos para la generación de animales transgénicos, en concreto animales tales como ratones o ratas, se han convertido en convencionales en el sector y están descritos, por ejemplo, en las Patentes de U.S. Nºs 4,736,866 y 4,870,009. Habitualmente, las células concretas serían marcadas para la incorporación de transgén de *fused* de vertebrado con potenciadores específicos de tejido. Los animales transgénicos que incluyen una copia de un transgén que codifica *fused* de vertebrado introducida en la línea germinal del animal en una fase embrionaria pueden utilizarse para examinar el efecto de la expresión incrementada de ADN que codifica *fused* de vertebrado. Pueden utilizarse dichos animales como animales de muestra para reactivos pensados para proporcionar protección frente a, por ejemplo, estados patológicos asociados con su sobreexpresión. Por ejemplo, para el carcinoma de células basales, el *fused* se sobreexpresa en la capa de células basales de la piel utilizando un promotor de Queratina 5 ó 14. Según este aspecto de la invención, se trata a un animal con el reactivo y una incidencia reducida del estado patológico, en comparación con los animales no tratados que llevan el transgén, indicaría una intervención terapéutica potencial para el estado patológico.

[0100] Pueden utilizarse homólogos de *fused* de vertebrado no humano para construir un animal "knock out" con *fused* de vertebrado que tiene un gen defectuoso o alterado que codifica *fused* de vertebrado como resultado de la recombinación homóloga entre el gen endógeno que codifica *fused* de vertebrado y el ADN genómico alterado que codifica *fused* de vertebrado introducido en una célula embrionaria del animal. Por ejemplo, puede utilizarse el ADNc que codifica *fused* de vertebrado para clonar ADN genómico que codifica *fused* de vertebrado de acuerdo con las técnicas establecidas. Puede suprimirse una parte del ADN genómico que codifica *fused* de vertebrado o sustituirse por otro gen, como por ejemplo un gen que codifica un marcador seleccionable que puede utilizarse para monitorizar la integración. Habitualmente, se incluyen diversas kilobases del ADN flanqueante no alterado (en los extremos 5' y 3') en el vector [ver por ejemplo, Thomas y Capecchi, Cell, 51:503 (1987) para una descripción de vectores de recombinación homóloga]. Se introduce el vector en una línea de células madre embrionarias (por ejemplo, mediante electroporación) y se seleccionan células en las que el ADN introducido se ha recombinado homológamente con ADN endógeno [ver por ejemplo, Li et al., Cell, 69:915 (1992)]. A continuación se inyectan las células seleccionadas en un blastocisto de un animal (por ejemplo, un ratón o una rata) para formar quimeras de agregación [ver por ejemplo, Bradley, en Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson, ed. (IRL, Oxford, 1987), págs. 113-152]. A continuación, puede implantarse un embrión quimérico en un animal de acogida hembra pseudopreñada y se hace nacer el embrión para crear un animal "knock out". Puede identificarse la progenie que alberga el ADN homológamente recombinado en sus células germinales mediante técnicas estándar y utilizarse para criar animales en los que todas las células del animal contienen el ADN homológamente recombinado. Pueden caracterizarse animales knockout por ejemplo, por su capacidad de defenderse contra ciertos estados patológicos y por su desarrollo en estados patológicos debido a la ausencia del polipéptido *fused* de vertebrado.

[0101] Como *fused* está implicado como mediador universal para todos los miembros de la familia de Hh (*SHh*, *IHh*,

DHh), los estados patológicos o trastornos que están asociados con la señalización de Hh general también serían ratables con *fused* y antagonistas o agonistas del mismo. Por ejemplo, la activación de *SHh* (por ejemplo, agonistas de *fused*) se ha promovido recientemente como tratamiento para varios trastornos degenerativos del sistema nervioso, por ejemplo, enfermedad de Parkinson, déficits de memoria, enfermedad de Alzheimer, enfermedad de Lou Gehrig, la enfermedad de Huntington, esquizofrenia, apoplejía y adicción a drogas. Estudios recientes sugieren que los machos mutantes en *DHh* son infértiles debido a la incapacidad de los espermatoцитos de completar su diferenciación en espermatozoides maduros. Bitgood et al., Curr. Biol. 6: 298-304 (1996); Bitgood et al., Dev. Biol. 172: 126-138 (1995). Adicionalmente, se podría utilizar los agonistas de *fused* para tratar enfermedades intestinales, enfermedades óseas, enfermedades de la piel, enfermedades de los testículos, úlceras, enfermedades pulmonares, enfermedades del páncreas, diabetes, osteoporosis.

[0102] La presencia del dominio de proteína quinasa sugiere que *fused* puede actuar de forma similar como miembros de la familia de proteínas quinasa en la modulación de la señalización de Hh. Las proteínas quinasa son elementos esenciales de circuitos reguladores en células diferenciadas, así como células en crecimiento; Preat et al., Nature 347: 87-89 (1990). Muchas de estas enzimas están implicadas en la transducción de señales extracelulares y operan a través de una cascada de sucesos de fosforilación que amplifican y diseminan los efectos de una señal primaria. Tal como se ha descrito anteriormente, el *fused* de *Drosophila* soporta una homología significativa con otras serina/treonina quinasa intracelulares. Muchas serina/treonina quinasa está implicadas en el control del ciclo celular en levaduras y en mamíferos, Hunter, Cell 50: 823-829 (1987); Dunphy & Newport, Cell 55: 925-928 (1988); Lee & Nurse, Trend Genet. 4: 287-290 (1988).

[0103] La supresión o inhibición de la señalización de Hh también es un objetivo de las estrategias terapéuticas. Dado que se ha observado que el *fused* inactivo inhibe la señalización de Hh, se deduce que también se esperaría que un antagonista de *fused* fuera antagonista a la señalización de Hh. La limitación de la señalización de Hh sería útil en estados patológicos o trastornos caracterizados por la señalización de Hh. Por ejemplo, se sabe que *SHh* es activo en el carcinoma de células basales; se sabe que *DHh* es activo en la espermatogénesis. Los inhibidores o antagonistas de la señalización de Hh serían productos terapéuticos efectivos en el tratamiento del carcinoma de células masales o la contracepción masculina, respectivamente.

[0104] La estimulación de la señalización de Hh es también un objetivo de estrategias terapéuticas. La activación de señalización de Hh sería útil en estados patológicos o trastornos caracterizados por una señalización de Hh inactiva o insuficiente, por ejemplo, trastornos degenerativos del sistema nervioso, por ejemplo, enfermedad de Parkinson, déficits de memoria, enfermedad de Alzheimer, enfermedad de Lou Gehrig, la enfermedad de Huntington, esquizofrenia, apoplejía y adicción a drogas. Adicionalmente, los agonistas de *fused* se podrían utilizar para tratar enfermedades intestinales, enfermedades óseas, enfermedades de la piel, enfermedades de los testículos (incluyendo fertilidad), úlceras, enfermedades pulmonares, enfermedades del páncreas, diabetes, osteoporosis.

F. Anticuerpos anti-*fused* de vertebrados

[0105] La presente invención proporciona además anticuerpos anti-*fused* de vertebrado. Ejemplos de anticuerpos incluyen anticuerpos policlonales, monoclonales, humanizados, biespecíficos, y heteroconjugados.

I. Anticuerpos policlonales

[0106] Los anticuerpos anti-*fused* de vertebrado pueden comprender anticuerpos policlonales. Los procedimientos para la preparación de anticuerpos policlonales son conocidos por los expertos. Los anticuerpos policlonales pueden generarse en un mamífero, por ejemplo, mediante una o más inyecciones de un agente inmunizante y, si se desea, un adyuvante. Habitualmente, el agente inmunizante y/o el adyuvante se inyectarán en el mamífero mediante múltiples inyecciones subcutáneas o intraperitoneales. El agente inmunizante puede incluir el polipéptido *fused* de vertebrado o una proteína de fusión del mismo. Puede ser útil la conjugación del agente inmunizante a una proteína de la que se conoce su poder inmunogénico en el mamífero que va a ser inmunizado. Entre los ejemplos de dichas proteínas inmunogénicas se incluyen, pero sin limitarse a, la hemocianina de lapa californiana, la albúmina sérica, la tiroglobulina bovina, y el inhibidor de la tripsina de soja. Entre los ejemplos de los adyuvantes que pueden ser utilizados se incluyen el adyuvante completo de Freund's y el adyuvante MPL-TDM (monofosforil Lípido A, dicorinomicolato sintético de trehalosa). El protocolo de inmunización puede seleccionarse por los expertos en la materia sin excesiva experimentación.

2. Anticuerpos monoclonales

[0107] Los anticuerpos anti-*fused* de vertebrado pueden ser, alternativamente, anticuerpos monoclonales. Los anticuerpos monoclonales pueden prepararse usando procedimientos del hibridoma, como los descritos por Kohler y

Milstein, Nature, 256:495 (1975). En un procedimiento del hibridoma, un ratón, un hámster, u otro animal huésped apropiado, es habitualmente inmunizado con un agente inmunizante que obtiene linfocitos que producen o son capaces de producir anticuerpos que se unirán específicamente con el agente inmunizante. Alternativamente, los linfocitos pueden ser inmunizados *in vitro*.

[0108] El agente inmunizante habitualmente incluirá el polipéptido *fused* de vertebrado o una proteína de fusión del mismo. Generalmente, se usan tanto linfocitos de sangre periférica ("PBLs") si se desean células de origen humano, como células de bazo o de nódulos linfáticos si se desean fuentes de mamífero no humano. Los linfocitos se fusionan a continuación con una línea celular immortalizada usando un agente de fusión adecuado, como polietilenglicol, para formar una célula de hibridoma [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Las líneas celulares immortalizadas son habitualmente células de mamífero transformadas, particularmente células de mieloma de origen murino, bovino y humano. Habitualmente, se utilizan líneas celulares de mieloma de rata o ratón. Las células de hibridoma pueden cultivarse en un medio de cultivo adecuado que contenga preferentemente una o más sustancias que inhiban el crecimiento o la supervivencia de las células immortalizadas no fusionadas. Por ejemplo, si las células parentales carecen de la enzima hipoxantina guanina fosforibosil transferasa (HGPRT o HPRT), el medio de cultivo de los hibridomas incluirá habitualmente hipoxantina, aminopterina, y timidina (medio "HAT"), cuyas sustancias evitan el crecimiento de las células deficientes en HGPRT.

[0109] Las líneas celulares immortalizadas preferidas son aquellas que se fusionan de forma eficiente, soportan un nivel de expresión elevado y estable del anticuerpo por las células productoras de anticuerpo seleccionadas, y son sensibles a un medio como el medio HAT. Las líneas celulares immortalizadas más preferibles son líneas de mieloma murino, que pueden obtenerse, por ejemplo, a partir del Salk Institute Cell Distribution Center, San Diego, California y a partir del American Type Culture Collection (ATCC), Rockville, Maryland. Las líneas celulares de mieloma humano de heteromieloma ratón-humano también se han descrito para la producción de anticuerpos monoclonales humanos [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

[0110] El medio de cultivo en el que se cultivan las células de hibridoma puede analizarse a continuación para determinar la presencia de anticuerpos monoclonales dirigidos contra el *fused* de vertebrado. Preferentemente, la especificidad de unión de los anticuerpos monoclonales producidos por las células de hibridoma se determina por inmunoprecipitación o mediante ensayos de unión *in vitro* como el radioinmunoensayo (RIA) o el ensayo de inmunoabsorción ligado a enzimas (ELISA). Dichas técnicas y ensayos son conocidos en el sector. La afinidad de unión del anticuerpo monoclonal puede, por ejemplo, determinarse a través del análisis de Scatchard de Munson y Pollard, Anal. Biochem., 107:220 (1980).

[0111] Después de identificar las células de hibridoma deseadas, es posible subclonar los clones mediante procedimientos de selección por dilución limitante y crecerlos mediante procedimientos estándares [Goding, *supra*]. Entre los medios de cultivo adecuados para este propósito se incluyen, por ejemplo, el Medio de Eagle modificado por Dulbecco y el medio RPMI-1640. Alternativamente, las células de hibridoma pueden crecerse *in vivo* como ascitos en un mamífero.

[0112] Los anticuerpos monoclonales secretados por los subclones pueden aislarse o purificarse a partir del medio de cultivo o del fluido ascítico mediante procedimientos convencionales de purificación de inmunoglobulinas tales como, por ejemplo, proteína A-Sepharose, cromatografía de hidroxilapatita, electroforesis en gel, diálisis, o cromatografía de afinidad.

[0113] Los anticuerpos monoclonales pueden prepararse también mediante procedimientos de ADN recombinante, tales como los descritos en la patente americana U.S. 4.816.567. El ADN que codifica los anticuerpos monoclonales de la invención puede aislarse fácilmente y secuenciarse usando procedimientos convencionales (por ejemplo, usando sondas de oligonucleótidos que son capaces de unirse específicamente a genes que codifican la cadena pesada y la cadena ligera de anticuerpos murinos). Las células de hibridoma de la invención sirven como una fuente preferente de dicho ADN. Una vez aislado, el ADN puede introducirse en vectores de expresión, que son transfectados a continuación en células huésped, tales como las células COS de simio, las células de ovario de hámster chino (CHO), o las células de mieloma que de lo contrario no producirían proteína inmunoglobulina, para obtener la síntesis de anticuerpos monoclonales en las células huésped recombinantes. El ADN también puede modificarse, por ejemplo, sustituyendo la secuencia codificante de los dominios constantes de las cadenas pesada y ligera humanas en lugar de las secuencias murinas homólogas [Patente de Estados Unidos U.S. 4.816.567; Morrison et al., *supra*] o mediante unión covalente a la secuencia codificante de la inmunoglobulina de toda o parte de la secuencia codificante de un polipéptido que no es inmunoglobulina. Dicho polipéptido no inmunoglobulínico puede sustituirse por los dominios constantes de un anticuerpo de la invención o puede sustituirse por los dominios

variables de una región de un sitio de combinación con el antígeno de un anticuerpo de la invención para crear un anticuerpo bivalente quimérico.

[0114] Los anticuerpos pueden ser anticuerpos monovalentes. Los procedimientos para la preparación de anticuerpos monovalentes son bien conocidos en la materia. Por ejemplo, un procedimiento implica la expresión recombinante de la cadena ligera de una inmunoglobulina y la cadena pesada modificada. La cadena pesada generalmente se trunca en cualquier punto de la región Fc para evitar el entrecruzamiento de la cadena pesada. Alternativamente, los residuos cisteína relevantes son sustituidos por otro residuo aminoácido o se eliminan para evitar el entrecruzamiento.

[0115] Para la preparación de anticuerpos monovalentes también son apropiados procedimientos *in vitro*. Puede realizarse la digestión de anticuerpos para producir fragmentos de los mismos, particularmente, fragmentos Fab, usando técnicas rutinarias conocidas la materia.

3. Anticuerpos humanizados

[0116] Los anticuerpos anti-*fused* de vertebrado pueden además comprender anticuerpos humanizados o anticuerpos humanos. Las formas humanizadas de anticuerpos no humanos (por ejemplo, murinos) son las inmunoglobulinas quiméricas, cadenas de inmunoglobulinas o fragmentos de los mismos (como Fv, Fab, Fab', F(ab')₂ u otras subsecuencias de unión a antígeno de los anticuerpos) que contienen una secuencia mínima derivada de una inmunoglobulina no humana. Entre los anticuerpos humanizados se incluyen inmunoglobulinas humanas (anticuerpo receptor) en las que se sustituyen residuos de una región determinante de complementariedad (CDR) del receptor por los correspondientes residuos de un CDR de una especie no humana (anticuerpo donante) como ratón, rata o conejo que tengan la especificidad, afinidad y capacidad deseadas. En algunos casos, se sustituyen residuos Fv de la región de armazón de la inmunoglobulina humana por los correspondientes residuos no humanos. Los anticuerpos humanizados también pueden comprender residuos que no se encuentran ni en el anticuerpo receptor ni en las secuencias de CDR o armazón importadas. En general, el anticuerpo humanizado comprenderá substancialmente todos de al menos uno, y habitualmente dos, dominios variables, en los que todas o substancialmente todas las regiones CDR corresponden a las de una inmunoglobulina no humana y todas o substancialmente todas las regiones FR son las de una secuencia consenso de una inmunoglobulina humana. Óptimamente, el anticuerpo humanizado comprenderá al menos una parte de una región constante de una inmunoglobulina (Fc), habitualmente la de una inmunoglobulina humana [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); y Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

[0117] Los procedimientos para humanizar anticuerpos no humanos son bien conocidos en la técnica. Generalmente, un anticuerpo humanizado tiene uno o más de un residuo aminoácido introducido en él a partir de una fuente que no es humana. Estos residuos aminoácidos no humanos se denominan frecuentemente como residuos "importados", que típicamente se toman a partir de un dominio variable "importado". La humanización puede realizarse esencialmente, siguiendo el procedimiento de Winter y colaboradores [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], mediante la sustitución de secuencias CDR o CDRs de roedores por las secuencias correspondientes de un anticuerpo humano. Por consiguiente, tales anticuerpos "humanizados" son anticuerpos quiméricos (patente de Estados Unidos 4.816.567), en donde substancialmente menos de un dominio variable humano intacto se ha sustituido por la secuencia correspondiente de una especie no humana. En la práctica, los anticuerpos humanizados son habitualmente anticuerpos humanos en los que algunos residuos CDR y posiblemente algunos residuos FR se sustituyen por residuos de sitios análogos en anticuerpos de roedores.

[0118] Los anticuerpos humanos pueden producirse también usando diversas técnicas conocidas en la materia, incluyendo bibliotecas de expresión en fagos [Hoogenboom y Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. Las técnicas de Cole et al., y Boerner et al., también están disponibles para la preparación de anticuerpos monoclonales humanos (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p.77 (1985) y Boerner et al., J. Immunol., 147(1): 86-95 (1991)].

4. Anticuerpos biespecíficos

[0119] Los anticuepos biespecíficos son anticuerpos monoclonales, preferentemente humanos o humanizados, que tienen especificidades de unión para al menos dos antígenos diferentes. En el caso presente, una de las especificidades de unión es para el *fused* de vertebrado y la otra para cualquier otro antígeno, y preferentemente por una proteína de superficie celular o por un receptor o subunidad de un receptor.

[0120] Son conocidos en la materia los procedimientos para la preparación de anticuerpos biespecíficos. Tradicionalmente, la producción recombinante de anticuerpos biespecíficos se basa en la coexpresión de dos pares de cadena pesada/cadena ligera de inmunoglobulina, en donde las dos cadenas pesadas tienen distintas especificidades (Milstein y Cuello, Nature, 305:537-539 [1983]). Debido a la reordenación aleatoria de las cadenas pesadas y ligeras de inmunoglobulina, estos hibridomas (cuadromas) producen una mezcla potencial de diez moléculas de anticuerpo distintas, de los cuales sólo una presenta la estructura biespecífica correcta. La purificación de la molécula correcta se consigue habitualmente mediante procedimientos de cromatografía de afinidad. Se describen procedimientos similares en la patente internacional WO 93/08.829, publicada el 13 de Mayo de 1993, y en Traunecker et al., EMBO J., 10:3655-3659 (1991).

[0121] Los dominios variables de un anticuerpo con las especificidades de unión deseadas (sitios de combinación anticuerpo-antígeno) pueden fusionarse a secuencias de dominios constantes de una inmunoglobulina. La fusión preferentemente es con un dominio constante de una cadena pesada de inmunoglobulina, que comprenda al menos parte de las regiones bisagra, CH2 y CH3. Se prefiere tener presente la primera región constante de la cadena pesada (CH1) que contenga el sitio necesario para la unión de la cadena ligera en al menos una de las fusiones. Los ADNs que codifican las fusiones de la cadena pesada de la inmunoglobulina y, si se prefiere, la cadena ligera de la inmunoglobulina, se insertan en vectores de expresión separados, y se co-transfectan en un organismo huésped apropiado. Para más detalles sobre la generación de anticuerpos biespecíficos véase, por ejemplo, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Anticuerpos heteroconjugados

[0122] Los anticuerpos heteroconjugados también se encuentran en el alcance de la presente invención. Los anticuerpos heteroconjugados se componen de dos anticuerpos unidos covalentemente. Se ha propuesto que dichos anticuerpos, por ejemplo, dirigen las células del sistema inmune contra las células no deseadas [patente americana U.S. 4.676.980], y sirven para el tratamiento de la infección por VIH [WO 91/00360; WO 92/200.373; patente europea EP 03.089]. Se considera que los anticuerpos se pueden preparar *in vitro* usando procedimientos conocidos en la química sintética de proteínas, incluyendo aquellos que implican agentes de entrecruzamiento. Por ejemplo, se pueden generar inmunotoxinas usando una reacción de intercambio de disulfuro o formando un enlace tioéter. Entre los ejemplos de agentes apropiados para este objetivo se incluye el iminotiolato y el metil-4-mercaptobutirimidato y aquellos descritos, por ejemplo, en la patente de Estados Unidos 4.676.980.

G. Usos para anticuerpos anti-*fused* de vertebrado

[0123] Los anticuerpos anti-*fused* de vertebrado presentan varias utilidades. Por ejemplo, los anticuerpos anti-*fused* de vertebrado se pueden utilizar en ensayos de diagnóstico para *fused* de vertebrado, por ejemplo, detectando su expresión en células, tejidos o suero específicos. Se pueden utilizar varias técnicas de ensayo de diagnóstico conocidas en el sector, tales como ensayos de unión competitiva, ensayos sándwich directo o indirecto y ensayos de inmunoprecipitación realizados en fases heterogénea u homogénea [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pág. 147-158]. Los anticuerpos utilizados en ensayos de diagnóstico se pueden marcar con un grupo detectable. El grupo detectable debe ser capaz de producir, directa o indirectamente, una señal detectable. Por ejemplo, el grupo detectable puede ser un isótopo radioactivo, tal como ^3H , ^{14}C , ^{32}P , ^{35}S , o ^{125}I , un compuesto flúorescente o quimioluminiscente, tal como isotiocianato de fluoresceína, rodamina o luciferina, o una enzima, tal como fosfatasa alcalina, beta-galactosidasa o peroxidasa de rábano picante. Se puede utilizar cualquier método conocido en la técnica para conjugar el anticuerpo al grupo detectable, incluyendo los métodos descritos por Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); y Nygren, J. Histochem. and Cytochem., 30:407 (1982).

[0124] Los anticuerpos anti-*fused* de vertebrado también son útiles para la purificación por afinidad de *fused* de vertebrado a partir de un cultivo de células recombinantes o fuentes naturales. En este proceso, los anticuerpos contra *fused* de vertebrado se inmovilizan en un soporte sólido, tal como resina Sephadex o papel de filtro, utilizando métodos conocidos en la técnica. A continuación, el anticuerpo inmovilizado se pone en contacto con una muestra que contiene el *fused* de vertebrado a purificar, y a continuación, se lava el soporte con un disolvente adecuado que eliminará sustancialmente todo el material en la muestra a excepción del *fused* de vertebrado, que está unido al anticuerpo inmovilizado. Finalmente, el soporte se lava con otro disolvente adecuado que liberará el *fused* de vertebrado del anticuerpo.

H. Antagonistas de *Fused*

[0125] Se pueden utilizar de manera adecuada varias estrategias para crear compuestos antagonistas y agonistas de *fused*. Es adecuada cualquier estrategia en la que la molécula antagonista se puede dirigir al interior de la célula,

lo cual interfiere o evita la operación normal de *fused* de tipo salvaje. Por ejemplo, los inhibidores competitivos, que incluyen *fused* mutante, tal como mutante dominante negativo identificado en los ejemplos, que evita que *fused* se una correctamente a otras proteínas necesarias para la señalización de Hh. Las propiedades adicionales de dichas moléculas antagonistas o agonistas son determinables fácilmente por un experto en la materia, tal como el tamaño, carga e hidrofobicidad adecuada para el transporte transmembrana.

[0126] Cuando deben identificarse o evaluarse miméticos u otros homólogos de mamífero, las células se exponen al compuesto de prueba y se comparan con controles positivos que están expuestos solos a *fused* humano y a controles negativos que no se expusieron ni al compuesto ni al ligando natural. Cuando deben identificarse o evaluarse antagonistas o agonistas de la modulación de señal de *fused*, las células se exponen al compuesto de la presente invención en presencia del ligando natural y se comparan con controles que no están expuestos al compuesto de prueba.

[0127] Se pueden utilizar ensayos de detección como cribado primario para evaluar la actividad de inhibición/aumento de fosfatasa de los compuestos antagonistas/agonistas de la presente invención. Los ensayos también se pueden utilizar para evaluar la potencia relativa de un compuesto analizando un intervalo de concentraciones en un intervalo de 100 mM a 1 μ M, por ejemplo, y computar la concentración a la que la cantidad de fosforilación o transducción de señal se reduce o aumenta en un 50% (IC50) en comparación con los controles.

[0128] Los ensayos se pueden realizar para identificar compuestos que afectan a la fosforilación de sustrato de *fused*. Específicamente, los ensayos se pueden realizar para identificar compuestos que incrementan la actividad de fosforilación de *fused* o se pueden realizar ensayos para identificar compuestos que disminuyen la fosforilación de sustratos de *fused*. Estos ensayos se pueden realizar en las propias células completas o en extractos celulares. Los ensayos se pueden realizar en una serie de formatos, incluyendo ensayos de unión proteína-proteína, ensayos de cribado bioquímico, inmunoensayos, ensayos de base celular, etc. Dichos formatos de ensayo son conocidos en la técnica.

[0129] Los ensayos de cribado de la presente invención son susceptibles de cribado de alto rendimiento de bibliotecas químicas y son particularmente adecuados para identificar candidatos de fármacos de molécula pequeña.

(1) Moléculas antagonistas y agonistas

[0130] Para cribar los antagonistas y/o agonistas de la señalización de *fused*, la mezcla de ensayo se incuba bajo condiciones mediante las cuales, excepto por la presencia del agente farmacológico candidato, *fused* induce la señalización de hedgehog con una actividad de referencia. Los componentes de la mezcla se pueden añadir en cualquier orden que proporcione la actividad de hedgehog necesaria. La incubación se puede realizar a cualquier temperatura que facilite una unión óptima, habitualmente entre 4°C y 40°C, más habitualmente entre aproximadamente 15°C y 40°C. Los periodos de incubación se seleccionan asimismo para la unión óptima, pero también se minimizan para un cribado rápido de alto rendimiento, y están habitualmente aproximadamente 0,1 y 10 horas, preferiblemente menos de 5 horas, más preferiblemente menos de 2 horas. Después de la incubación, el efecto del agente farmacológico candidato en la señalización de *fused* se determina de cualquier manera conveniente. Para ensayos de tipo unión libres de células, a menudo se utiliza una etapa de separación para separar los componentes unidos y no unidos. La separación se puede realizar, por ejemplo, mediante precipitación (por ejemplo, precipitación con TCA, inmunoprecipitación, etc.), inmovilización (por ejemplo, en un sustrato sólido), seguido de lavado. La proteína unida se detecta de forma conveniente aprovechando el marcador detectable unido a la misma, por ejemplo, midiendo la emisión radioactiva, densidad óptica o electrónica, o mediante detección indirecta utilizando, por ejemplo, conjugados de anticuerpo.

[0131] Por ejemplo, un método de cribado de antagonistas y/o agonistas de *fused* adecuadas podría implicar la aplicación de agentes presentes en el ensayo informador de Gli que activa *fused* descrito en los ejemplos. Dicho ensayo de cribado podría comparar la hibridación in situ en presencia y ausencia del antagonista y/o agonista candidato en un tejido que expresa *fused*, así como la confirmación o ausencia de desarrollo celular modulado por *fused*. Habitualmente, estos métodos implican la exposición de un *fused* inmovilizado a una molécula sospechosa de unirse al mismo y la determinación de la unión o fosforilación de la molécula al *fused* inmovilizado y/o la evaluación de si la molécula activa o no (o bloquea la activación de) *fused*. A efectos de identificar dichos ligandos de unión a *fused*, se puede expresar *fused* en la superficie de una célula y utilizarse para cribar bibliotecas de compuestos candidatos o compuestos naturales (por ejemplo, de fuentes endógenas, tales como suero o células).

[0133] Las moléculas adecuadas que afectan a la interacción proteína-proteína de *fused* y sus proteínas de unión incluyen fragmentos de ésta última o moléculas pequeñas, por ejemplo, peptidomiméticos, que evitarán la interacción y formación correcta de complejos. Dichas moléculas pequeñas, que normalmente tienen un peso

molecular inferior a 10 K, son preferibles como agentes terapéuticos, ya que es más probable que sean permeables a las células, sean menos susceptibles a la degradación por diversos mecanismos celulares, y no sean aptas para obtener una respuesta inmune como proteínas. Las moléculas pequeñas incluyen, pero sin limitación, compuestos orgánicos o inorgánicos sintéticos. Muchas compañías farmacéuticas tienen bibliotecas extensas de dichas moléculas que se pueden cribar convenientemente mediante la utilización de los ensayos de la presente invención. Entre los ejemplos no limitantes se incluyen proteínas, péptidos, glicoproteínas, glicopéptidos, glicolípidos, polisacáridos, oligosacáridos, ácidos nucleicos, moléculas bioorgánicas, péptidomiméticos, agentes farmacológicos y sus metabolitos, secuencias de control transcripcional y traduccional, y similares.

[0134] Una técnica preferida para identificar moléculas que se unen a *fused* utiliza un sustrato químico (por ejemplo, *fused* etiquetado con epítipo o inmunoadhesina de *fused*) unido a una fase sólida, tal como el pocillo de una placa de ensayo. Se puede medir la unión de las moléculas candidatas, que están opcionalmente marcadas (por ejemplo, marcadas radioactivamente), al receptor inmovilizado. Alternativamente, se puede medir la competición por la activación de Gli. En el cribado de antagonistas y/o agonistas, se puede exponer *fused* a un sustrato de *fused* seguido del supuesto antagonista y/o agonista, o se pueden añadir simultáneamente la proteína de unión a *fused* y el antagonista y/o agonista y se puede evaluar la capacidad del antagonista y/o agonista para bloquear la activación de *fused*.

(2) Ensayos de detección

[0135] Los polipéptidos *fused* son útiles en ensayos para identificar compuestos de partida para agentes terapéuticamente activos que modulan la señalización de hedgehog por *fused*. Específicamente, se pueden identificar convenientemente compuestos de partida que evitan la formación de los complejos de señalización de *fused* o evitan o atenúan la señalización de hedgehog modulada por *fused* (por ejemplo, la unión al popio *fused* o a un sustrato).

[0136] Se pueden utilizar varios procedimientos conocidos en la técnica para identificar, evaluar o analizar la inhibición de actividad de proteínas *fused* de la presente invención. Dado que se cree que *fused* actúa de una forma similar a las otras quinasas, se pueden emplear con la presente invención también técnicas conocidas para utilizar con la identificación de moduladores de quinasa/fosfatasa. En general, dichos ensayos implican exponer las células diana en el cultivo con los compuestos y (a) analizar bioquímicamente los lisados celulares para evaluar el nivel y/o identidad de la fosforilación, o (b) valorar los cambios fenotípicos o funcionales en células tratadas en comparación con células de control que no se expusieron a la sustancia de prueba. Dichos ensayos de cribado se describen en Patente de Estados Unidos 5.602171, Patente de Estados Unidos 5,710,173, WO 96/35124 y WO 96/40276.

(a) Técnicas de detección bioquímica

[0137] Se pueden evaluar las técnicas de análisis bioquímico mediante una serie de técnicas. Una mezcla de ensayo típica que se puede utilizar con la presente invención contiene *fused* y una proteína con la que *fused* está normalmente asociada (por ejemplo, Gli), normalmente en una forma aislada, parcialmente pura o pura. Uno o ambos de estos componentes se pueden fusionar a otro péptido o polipéptido, que, por ejemplo, proporcionan o aumentan la unión proteína-proteína, mejoran la estabilidad bajo las condiciones de ensayo, etc. Además, uno de los componentes comprende normalmente o está acoplado a un marcador detectable. El marcador puede proporcionar la detección directa midiendo la radioactividad, luminiscencia, densidad óptica o electrónica, etc. o detección indirecta, tal como un epítipo etiqueta, una enzima, etc. La mezcla de ensayo puede comprender adicionalmente un agente farmacológico candidato, y opcionalmente una serie de componentes, tales como sales, tampones, proteínas portadoras, por ejemplo, albúmina, detergentes, inhibidores de proteasa, inhibidores de nucleasa, agentes antimicrobianos, etc., que facilitan la unión, incrementan la estabilidad, reducen las interacciones no específicas o de base, o en cualquier caso, mejoran la eficiencia o sensibilidad del ensayo.

[0138] Los siguientes métodos de detección también se pueden utilizar en un sistema libre de células en el que el lisado celular que contiene la molécula sustrato de transducción de señal y *fused* se mezclan con un compuesto de la presente invención. El sustrato se fosforila por la iniciación de la reacción de quinasa mediante la adición de adenosín trifosfato (ATP). Para evaluar la actividad del compuesto, la mezcla de reacción puede analizarse por la técnica de SDS-PAGE o puede añadirse al anticuerpo de anclaje específico de sustrato unido a un soporte sólido, y se realiza un procedimiento de detección tal como se ha descrito anteriormente en el sustrato separado o capturado para evaluar la presencia o ausencia de pSer/Thr. Los resultados se comparan con los obtenidos con las mezclas de reacción a las que no se añade el compuesto. El sistema libre de células no requiere el ligando natural o el conocimiento de su identidad. El sistema libre de células no requiere mezclas a las que no se añade el compuesto. El sistema libre de células no requiere el ligando natural o el conocimiento de su identidad. Por ejemplo, Posner et al. (Patente de Estados Unidos 5,155,031 describe la utilización de un receptor de insulina como sustrato y

adipocitos de rata como células diana para demostrar la capacidad de pervanadato de inhibir la actividad de PTP. Otro ejemplo, Burke et al., Biochem. Biophys. Res. Comm. 204: 129-134 (1994) describe la utilización de receptor de insulina autofosforilado y PTIB recombinante en la valoración de la actividad inhibidora de un mimético fosfotirosilo.

5 (i) *Detección de células completas*

[0139] Una técnica habitual implica incubar células con *fused* de vertebrado y fosfato radiomarcado, lisar las células, separar los componentes proteicos celulares del lisado utilizando una técnica de gel de SDS-poliacrilamida (SDS-PAGE), en una o dos dimensiones, y detectar la presencia de proteínas fosforiladas mediante la exposición a películas de rayos X. La detección también se puede llevar a cabo sin utilizar el marcaje radioactivo. En dicha técnica, los componentes proteicos (por ejemplo, separados mediante SDS-PAGE) se transfieren a una membrana de nitrocelulosa donde se detecta la presencia de serina/treoninas fosforiladas utilizando un anticuerpo anti-fosfoserina/treonina (anti-pS/T).

[0140] Alternativamente, el anti-pS/T se puede conjugar con una enzima, tal como peroxidasa de rábano picante, y detectar mediante la adición posterior de un sustrato colorimétrico para la enzima. Una alternativa adicional implica detectar el anti-PS/T mediante la reacción con un segundo anticuerpo que reconoce el anti-PS/T, estando este segundo anticuerpo marcado con un grupo radioactivo o una enzima tal como se ha descrito previamente. Ejemplos de estos y técnicas similares se describen en Hansen et al., Electrophoresis 14: 112-126 (1993); Campbell et al., J. Biol. Chem. 268:7427-7434 (1993); Donato et al., Cell Growth Diff. 3 258-268 (1992); Katagiri et al., J. Immunol. 150: 585-593 (1993). Adicionalmente, el anti-pS/T se puede detectar marcándolo con una sustancia radioactiva, seguido del rastreo de la nitrocelulosa marcada para detectar la radioactividad o la exposición de películas de rayos X.

(ii) *Ensayos de quinasa*

[0141] Cuando se llevan a cabo métodos de cribado de la presente invención para antagonistas de *fused* como un ensayo ex vivo, la quinasa diana (por ejemplo, *fused*) puede ser un polipéptido sustancialmente purificado. El sustrato quinasa (por ejemplo, MBP, Gli) es un sustrato sustancialmente purificado, que en el ensayo está fosforilado en una reacción con una fuente de fosfato sustancialmente purificada que está catalizada por la quinasa. El grado de fosforilación se determina midiendo la cantidad de sustrato fosforilado en la reacción. Se puede utilizar una variedad de sustratos posibles, incluyendo la quinasa en sí en cuyo caso la reacción de fosforilación medida es la autofosforilación. También se pueden utilizar sustratos exógenos, incluyendo sustratos de proteína estándar, tales como proteína básica de mielina (MBP); sustratos de proteína de levadura; sustratos de péptidos sintéticos, y sustratos de polímero. De estos, MBP y otros sustratos de proteínas estándar se pueden considerar como preferidos (véase el ejemplo 10). Sin embargo, se pueden identificar otras sustancias que son superiores mediante la afinidad para la quinasa, la perturbación mínima de la cinética de reacción, la posesión de sitios de reacción únicos u homogéneos, facilidad de manipulación y recuperación después de la reacción, potencial para una generación de señal fuerte, y resistencia o no reacción con los compuestos de prueba.

[0142] La medición de la cantidad de sustrato fosforilado en el ensayo ex vivo de la presente invención se puede llevar a cabo mediante inmunoensayo, radioensayo u otros métodos bien conocidos. En la medición de un inmunoensayo, se puede utilizar un anticuerpo (tal como un anticuerpo anti-fosfoserina/treonina de cabra o ratón) que es específico para grupos fosforilados formados durante la reacción. Utilizando técnicas eLISA bien conocidas, el complejo de anticuerpos de fosfoserina/treonina se detectaría por un anticuerpo adicional unido a un marcador capaz de desarrollar una señal medible (como por ejemplo, un marcador fluorescente o radioactivo). Adicionalmente, se pueden utilizar ensayos de tipo ELISA en placas de microtitulación para analizar sustratos purificados. Peraldi et al., J. Biochem. 285: 71-78 (1992); Schraag et al., Anal. Biochem. 211: 233-239 (1993); Cleavland, Anal. Biochem. 190: 249-253 (1990); Farley, Anal. Biochem. 203: 151-157 (1992) y Lozano, Anal. Biochem. 192: 257-261 (1991).

[0143] Por ejemplo, esquemas de detección pueden medir el agotamiento de sustrato durante la reacción de quinasa. Inicialmente, la fuente de fosfato se puede marcar radioactivamente con un isótopo, tal como ^{32}P o ^{33}P , y se puede medir la cantidad de fosforilación de sustrato determinando la cantidad de marcador radioactivo incorporado en el sustrato durante la reacción. La detección se puede realizar mediante: (a) placas que contienen agentes centelleantes disponibles comercialmente y partículas que utilizan un beta-contador, después de la adsorción a un filtro o a una superficie de pocillo de microtitulación, o (b) medios fotométricos después de la unión a una partícula de ensayo de proximidad de centelleo o placa con agente centelleante. Weernik and Kijken, J. Biochem. Biophys. Methods 31: 49, 1996; Braunwalder et al., Anal Biochem. 234: 23 (1996); Kentrup et al., J. Biol. Chem. 271: 3488 (1996) and Rusken et al., Meth. Enzymol. 200: 98 (1991).

[0144] Preferiblemente, el sustrato se une a una superficie de soporte sólido mediante una unión no específica, o preferiblemente, específica. Dicha unión permite la separación del sustrato fosforilado de la fuente de fosfato

marcado no incorporado (tal como adenosín trifosfato antes de la detección de la señal). En una realización, el sustrato se puede inmovilizar físicamente antes de la reacción, como a través de la utilización de la placa de unión elevada a proteínas Nunc™ (Hanke et al., J. Biol. Chem. 271: 695 (1996)) o placas Wallac ScintiStrip™ (Braunwalder et al., Anal. Biochem. 234: 23 (1996)). El sustrato también se puede inmovilizar después de la reacción mediante precipitación con TCA sobre papel Whatman™ 3MM, Tiganis et al., Arch. Biochem. Biophys. 325: 289 (1996); Morawetz et al., Mol. Gen. Genet. 250: 17 (1996); Budde et al. Int J. Pharmacognosy 33: 27 (1995) y Casnellie, Meth. Enz. 200: 115 (1991). Otra posibilidad es la unión del sustrato a la superficie del soporte, como mediante la conjugación con compañeros de unión, tales como glutatión y estreptavidina (en el caso de GST y biotina, respectivamente) que se han unido al soporte, o mediante anticuerpos específicos para las etiquetas que así mismo están unidas al soporte.

[0145] Se pueden desarrollar métodos de detección adicionales que son preferidos a los descritos anteriormente. Especialmente para utilizar en relación con el cribado de rendimiento elevado, se espera que dichos métodos exhiban una buena sensibilidad y especificidad, intervalo lineal extenso, señal de fondo baja, fluctuación mínima, compatibilidad con otros reactivos y compatibilidad con sistemas de manipulación automática.

[0146] La eficacia *in vivo* del tratamiento de la presente invención se puede estudiar contra tumores inducidos químicamente en varios modelos de roedores. Las líneas celulares de tumores propagadas en cultivos celulares *in vitro* se pueden introducir en roedores experimentales, por ejemplo, ratones por inyección, por ejemplo, mediante la ruta subcutánea. Las técnicas para la inducción química de tumores en animales experimentales es bien conocida en la técnica.

(b) Técnicas de detección biológica:

[0147] La capacidad de los compuestos antagonistas de la presente invención para modular la actividad de *fused*, que en sí misma modula la señalización de hedgehog, también se puede medir valorando los cambios morfológico o funcionales asociados con la unión a ligando. Se puede aplicar cualquier técnica cualitativa o cuantitativa conocida en el sector para observar y medir procesos celulares que se encuentran bajo el control de *fused*. La actividad de los compuestos de la presente invención también se puede evaluar en animales utilizando modelos experimentales de trastornos causados por o relacionados con la señalización disfuncional de hedgehog. Por ejemplo, la señalización inefectiva de hedgehog *DHh* en ratones conduce a ratones viables pero estériles. Los efectos de *fused* mutante (*hfused-DN*) también afectan al desarrollo de los intestinos, que está regulado por la expresión de *IHh*. Adicionalmente, la señalización de *SHh* es crítica para el desarrollo embrionario murino en el notocordio y la placa del piso, tubo neural, estructuras de extremidades distales, columna espinal y costillas. Una señalización incorrecta de *SHh* es también correlativa con la ciclopía. Se podría evaluar y cuantificar cualquiera de estas propiedades fenotípicas en un ensayo de cribado para antagonistas y/o agonistas de *fused*. Los estados patológicos asociados con la sobreexpresión de hedgehog están asociados con el carcinoma de células basales, mientras que una señalización de Sonic hedgehog inactiva conduce a un desarrollo neural incorrecto.

[0148] Los datos obtenidos a partir de estos ensayos de cultivos celulares y estudios de animales se pueden utilizar en la formulación de un intervalo de dosis para utilizar en humanos. La dosis de los compuestos de la presente invención debe encontrarse en un intervalo de concentraciones circulantes con poca o nula toxicidad. La dosis puede variar en este intervalo dependiendo de la forma de dosificación empleada y la ruta de administración.

(2) Nucleótidos antisentido

[0149] Otra clase preferida de antagonistas implica la utilización de técnicas de terapia génica, incluyendo la administración de nucleótidos antisentido. Entre las técnicas de terapia génica aplicables se incluyen administraciones individuales o múltiples de ADN o ARNm terapéuticamente eficaz. Los ARN y ADN antisentido se pueden utilizar como agentes terapéuticos para bloquear la expresión de ciertos genes *in vivo*. Los oligonucleótidos antisentido cortos se pueden importar en células en las que actúan como inhibidores, a pesar de sus concentraciones intracelulares bajas causadas por una captación limitada por la membrana celular, Zamecnik et al., Proc. Natl. Acad. Sci. USA 83: 4143-4146 (1986). Los oligonucleótidos se pueden modificar para aumentar su captación, por ejemplo, sustituyendo sus grupos fosfodiéster cargados negativamente por grupos no cargados.

[0150] Existe una variedad de técnicas conocidas para introducir ácidos nucleicos en células viables. Las técnicas varían dependiendo de si el ácido nucleico se transfiere a células cultivadas *in vitro*, *ex vivo*, o *in vivo* en las células del huésped pretendido. Las técnicas adecuadas para la transferencia de ácido nucleico en células de mamífero *in vitro* incluyen la utilización de liposomas, electroporación, microinyección, fusión celular, DEAE-dextrano, el método de precipitación con fosfato de calcio, etc. Las técnicas de transferencia de genes *in vivo* actualmente preferidas incluyen la transfección con vectores virales (habitualmente retrovirales) y transfección mediante liposomas –

proteína de cubierta viral, Dzau et al., Trends Biotech. 11: 205-210 (1993). En algunas situaciones, es deseable proporcionar la fuente de ácido nucleico con un agente que reconoce las células diana, tal como un anticuerpo específico para una proteína de membrana de la superficie celular asociada con endocitosis, se puede utilizar para reconocer y/o facilitar la captación, por ejemplo proteínas de la cápside o fragmentos de las mismas trópicas para un tipo de células particulares, anticuerpos para proteínas que experimentan internalización en el ciclado, y proteínas que reconocen la localización intracelular y aumentan la vida media intercelular. La técnica de endocitosis mediada por receptor se describe, por ejemplo, por Wu et al., J. Biol. Chem. 262: 4429-4432 (1987); Wagner et al., Proc. Natl. Acad. Sci. USA 87: 3410-3414 (1990). Para una revisión del marcaje de genes y protocolos de terapia génica conocidos, véase Anderson et al., Science 256: 808-813 (1992).

[0151] En una realización, las moléculas antagonistas de *fused* se pueden utilizar para unirse a ligando endógeno en la célula; provocando así que la célula sea insensible a *fused* de tipo salvaje, especialmente cuando los niveles de *fused* en la célula superan los niveles fisiológicos normales. Además, puede ser beneficioso unir sustratos de *fused* endógenos o agentes complejantes que están activando respuestas celulares no deseadas (tales como la proliferación de células tumorales).

[0152] En una realización de la presente invención, la expresión de *fused* se puede reducir mediante la disposición de células que expresan *fused* con una cantidad de ARN o ADN antisentido de *fused* eficaz para reducir la expresión de la proteína *fused*.

I. Usos diagnósticos

[0153] Otro uso de los compuestos de la presente invención (por ejemplo, anticuerpos anti-*fused* de vertebrado) descritos en la presente invención es ayudar a diagnosticar si un trastorno está impulsado, en cierto grado, por la señalización de *fused* o hedgehog. Por ejemplo, las células de carcinoma de células basales están asociadas con la señalización activa de hedgehog.

[0154] Se puede llevar a cabo un ensayo de diagnóstico para determinar si un trastorno particular está impulsado por la señalización de hedgehog utilizando las siguientes etapas: (a) cultivar células o tejidos de prueba; (2) administrar un compuesto que puede inhibir la señalización de hedgehog mediada por *fused*; y (3) medir el grado de atenuación de quinasa en el sustrato de *fused* en lisados celulares o efectos fenotípicos mediados por hedgehog en las células de prueba. Las etapas se pueden llevar a cabo utilizando técnicas estándar en vista de la presente descripción. Por ejemplo, se pueden utilizar técnicas estándar para aislar células o tejidos y cultivar *in vivo*.

[0155] Los compuestos de grado variante de selectividad son útiles para el diagnóstico del papel de *fused*. Por ejemplo, los compuestos que inhiben *fused* además de otra forma de quinasa se pueden utilizar como compuesto de prueba inicial para determinar si una de las diversas serina/treonina quinasas impulsa el trastorno. Los compuestos selectivos se pueden utilizar a continuación para eliminar adicionalmente el posible papel de las otras serina/treonina quinasas en el impulso del trastorno. Los compuestos de prueba deben ser más potentes en la inhibición de actividad de serina/treonina quinasa que en la realización del efecto citotóxico (por ejemplo, una IC50/LD50 superior a uno). La IC50 y LD50 de un compuesto debe tenerse en cuenta en la evaluación del ensayo de diagnóstico. En general, cuanto mayor es la proporción, más relativa es la proporción. Los controles apropiados que tienen en cuenta el posible efecto citotóxico de un compuesto, tal como el tratamiento de células no asociadas con un trastorno proliferativo celular (por ejemplo, células de control) con un compuesto de prueba, también se pueden utilizar como parte del ensayo de diagnóstico. Los métodos de diagnóstico de la presente invención implican el cribado de agentes que modulan los efectos de *fused* después de la señalización de hedgehog. Las técnicas de detección de ejemplo incluye marcaje radioactivo e inmunoprecipitación (Patente de Estados Unidos 5,385,915).

[0156] Los siguientes ejemplos se ofrecen únicamente con fines ilustrativos y no pretenden, de ningún modo, limitar el alcance de la presente invención.

EJEMPLOS

[0157] Los reactivos disponibles comercialmente referidos en los ejemplos se utilizaron de acuerdo con las instrucciones del fabricante salvo que se indique lo contrario. La fuente de las células se han identificado en los ejemplos siguientes, y a lo largo de la memoria, por los números de acceso ATCC que corresponden al American Type Culture Collection, Rockville, Maryland.

EJEMPLO 1

Aislamiento de clones de ADNc de *fused* humano

[0158] Se buscó en una base de datos de marcadores de secuencia expresada (EST) (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) un homólogo humano del gen *fused* de polaridad de segmento de *Drosophila* (SEQ ID NO 26) (Preat et al., Nature 347: 87-9 (1990)). Se identificó el EST Incyte #2515662 (Figura 2) (SEQ ID NO. 3) como un potencial candidato. A efectos de identificar bibliotecas de ADNc humano que contiene clones de *fused* humanos, las bibliotecas de ADNc humanos en pRK5 se cribaron primero mediante PCR utilizando los siguientes cebadores:

h-*FUSED*.f (SEQ ID NO. 8) 5'-CAATACAATGGTGCTGACATCCATCAAAGGCA-3'
h-*FUSED*.r (SEQ ID NO. 9) 5'-GAAGGGAGGGGTGCCTACTGCCA-3'

[0159] Se seleccionó una biblioteca de pulmón fetal y se enriqueció en clones de ADNc de *fused* mediante la extensión de ADN de cadena sencilla de bibliotecas de plásmidos desarrollados en un huésped dug'/bung' utilizando el cebador h-*FUSED*.f (SEQ ID NO. 8) en una reacción que contenía 0 µl de 10x tampón PCR (Perkin Elmer). Después de inicio en caliente se añadieron 1 µl dNTP (20 mM), 1 µl de ADN de biblioteca (200 ng), 0,5 ml de cebador, 86,5 µl H₂O y 1 µl de Amplitaq® (Perkin Elmer). La reacción se desnaturizó durante 1 min a 95°C, se hibridó durante 1 min a 60°C, a continuación se extendió durante 20 min a 72°C. El ADN se extrajo con fenol/CHCl₃, se precipitó con etanol, a continuación se transformó mediante electroporación en bacterias huésped DH10B. Las colonias de cada transformación se emplacaron y se elevaron en una membrana de nylon y se cribaron con una sonda de oligonucleótidos derivada de la secuencia EST de la siguiente secuencia:

h-*FUSED*.p (SEQ ID NO. 10) 5'-CTCCAGCTCTGGAGACATATAGAGTGGTGTGCCTTTGA-3'

[0160] Se marcó la sonda de oligonucleótidos con [γ -³²P]-ATP y T4 polinucleótido quinasa. Se hibridaron los filtros durante la noche a 42°C en formamida al 50%, 5xSSC, 10x solución de Denhardt, fosfato de sodio 0,05 M (pH 6,5), pirofosfato sódico al 0,1%, 50 µg/ml de ADN de esperma de salmón sonificado. A continuación, se enjuagaron los filtros en 2xSSC y se lavaron en 0,1x SSC, SDS al 0,1% y a continuación se expusieron a películas de rayos X Kodak®. Se aislaron y secuenciaron dos clones positivos (DNA28494 (SEQ ID NO. 6) y DNA28495 (SEQ ID NO. 4) - Figuras. 4 y 5) que contenían una inserción de aproximadamente 5 kb. La secuencia del clon DNA28495 (SEQ ID NO. 4) contiene una potencial metionina de iniciación en la posición 116 seguido de un marco de lectura abierto de 1944 pb (figura 4). Sin embargo, este marco de lectura abierto (ORF) codifica una proteína que sólo tiene 648 aminoácidos de largo, algo más corta que la secuencia de 795 aminoácidos de *fused* de *Drosophila*. De forma destacada, está presente un segundo marco de lectura abierto en la región 3' del ADNc, desde el nucleótido 2295 a 4349 (figura 4), lo cual sugiere que el ADNc puede haberse empalmado de forma incorrecta y que el intrón permanece entre los 2 ORF, o corresponden a una variante de corte y empalme alternativo de *fused*. La secuencia del clon DNA28494 (SEQ ID NO. 6) es muy similar. Existe una diferencia de un nucleótido entre el clon DNA28495 (SEQ ID NO. 4) y el clon DNA28494 (SEQ ID NO. 6) localizado en el primer ORF en la posición 1863 del clon 28495 (SEQ ID NO. 4) (A vs. G) que cambia la secuencia codificante de una Gln a una Arg en la posición 583. (figura 4). Este cambio es probablemente debido a una variación alélica. El primer marco de lectura de DNA28494 (SEQ ID NO. 6) empieza en el residuo 115 y está seguido de un marco de lectura abierto largo de 647 aminoácidos. Las secuencias son idénticas a excepción de un cambio descrito anteriormente en la posición 583 y durante los últimos 9 residuos en el primer marco de lectura abierto.

EJEMPLO 2

Expresión de clones de *fused*

[0161] A efectos de determinar el tamaño de la proteína expresada a partir del ADNc correspondiente a DNA28495 (SEQ ID NO. 4) y DNA28494 (SEQ ID NO. 6), se inseró un epítipo etiqueta de HA en el extremo N-terminal de la proteína mediante PCR utilizando los siguientes cebadores:

Hfus.Cla-HA.F: (SEQ ID NO. 11)
5'-CCATCGATGTACCCATACGACGTCCCAGACTACGCTGAAAAGTACCACGTGTTGGAGATG-3'

y hFus.Xba.R: (SEQ ID NO. 12)
5'-GCTCTAGACTAAGGGGCAGGTCCTGTGTTCTG-3'.

[0162] El producto PCR se purificó, se digirió con ClaI-SmaI y se subclonó en los plásmidos pRK5 que contenían DNA28494 (SEQ ID NO. 6) y DNA28495 (SEQ ID NO. 4). El ADN de cada una de las construcciones se transfeció durante la noche en células 293 utilizando el método de CaPO₄ (Sambrook et al. supra; Ausubel et al., supra). Después de aproximadamente 24 h. a 48 h., después de la transfección, se recogieron las células se lisó el residuo

celular en 1 ml de tampón de lisina (Tris 50 mM pH 8,0, NaCl 150 mM, EDTA 1 mM, NP40 al 1%, Aprotinina, Leupeptina, -PMSF, NaF 1 mM y vanadato de sodio 1 mM) durante 20 min a 4°C. El extracto se centrifugó durante 10 minutos a 10 K, a continuación se transfirió el sobrenadante a un nuevo tubo y se prepurificó con 20 µl de Proteína A Sefarosa durante 1 h. La proteína A sefarosa se centrifugó y se añadió 1 µl de anticuerpo anti-HA (5 µg, Boehringer) a cada tubo. Después de incubación durante la noche a 4°C, se añadieron 30 µl de Proteína G Sefarosa y los tubos se incubaron a 4°C durante 1 hora. A continuación, las partículas de proteína G se centrifugaron durante 1 min, se lavaron 3 veces con tampón de lisis, se resuspendieron en 20 µl de tampón de Laemmli en presencia de β-mercapto etanol. Las muestras se desnaturalizaron durante 5 min. a 100°C, a continuación se cargaron gel de poliacrilamida al 6%. A continuación, las proteínas se transfirieron a nitrocelulosa y se analizaron mediante transferencia Western utilizando el mismo anticuerpo anti-Ha durante la noche a 1 µg/ml en tampón de bloqueo (PBS 0,5%, Tween 5%, leche en polvo descremada, suero de cabra al 3% seguido de HRP anti-ratón). Se utilizó ECL para la detección y la membrana se expuso durante 90 segundos a películas de rayos X. Se detectó una banda específica de 150 kDa en el residuo celular de células transfectadas con la construcción correspondiente al clon DNA28494 (SEQ ID NO. 6) y podría detectarse una banda específica de aproximadamente 100 kDa para el clon DNA28495 (SEQ ID NO. 4) (Figura 6). Estas bandas no estaban presentes en el control transfectado por simulación. La presencia de la banda de 150 kDa sugiere que los dos marcos de lectura abiertos de DNA28494 (SEQ ID NO. 6) se pueden empalmar juntos para dirigir la síntesis de una proteína grande de 150 kDa. La ausencia de esta banda para DNA28495 (SEQ ID NO. 4) sugirió que este clon aparentemente no se emplama de forma correcta. El corte y empalme alternativo del gen de *fused* parece conducir a la producción de varios productos diferentes y puede ser un mecanismo de regulación de la actividad de *fused*. Se sabe que las regiones específicas en el extremo C-terminal de la proteína *fused* de *Drosophila* son necesarias para la actividad de la molécula, Therond et al., Genetics 142: 1181-1198 (1996); Robbins et al., Cell 90: 225-234 (1997). Las moléculas de *fused* más cortas truncadas en el extremo C-terminal pueden corresponder por tanto a formas inactivas o negativas dominantes de la molécula.

EJEMPLO 3

Transferencias Northern

[0163] A efectos de determinar la mayor fuente de tejidos para aislar más ADNc de *fused* y para identificar un transcrito que codifica una molécula *fused* de 150 kDa de longitud completa, se sondaron "blots" northern de múltiples tejidos humanos 1.11 y "blot" fetal de Clontech con un fragmento Clak-AccI de 1,6 kb derivado del clon DNA28494 (SEQ ID NO. 6) marcado mediante cebadores al azar. Los "blots" se hibridaron en formamida al 50%, 5 x SSC, 10 x solución de Denhardt, fosfato sódico 0,05 M (pH 6,5), pirofosfato sódico al 0,1%, ADN de esperma de salmón sonificado 50 mg/ml, todo en presencia de 1x10⁶ cpm/ml de sonda marcada con ³²P a 42°C durante la noche. Los "blots" se lavaron en 2 x SSC a temperatura ambiente durante 10 minutos y se lavaron en 0,2 x SSC/SDS al 0,1% a 42°C durante 30 minutos, a continuación se expusieron a una película de rayos X durante la noche. La figura 7 muestra que el mensaje de *fused* se expresa a niveles elevados en los testículos y en niveles bajos en la mayoría de otros tejidos, incluyendo tejidos fetales. (Figura 7).

EJEMPLO 4

PCR en diferentes tejidos para identificar la forma de corte y empalme ("splice") correcta

[0164] A efectos de aislar un ADNc donde los 2 ORF potenciales se empalmaron correctamente, se diseñaron los siguientes cebadores que flanqueaban el potencial intrón y se amplificaron las plantillas de varios tejidos incluyendo el cerebro fetal humano, queratinocito, testículos, ovario, hígado fetal y pulmón.

F1 (SEQ ID NO. 13) 5'-CTGACGACACAGCAGGTTGTC-3'
R4 (SEQ ID NO. 14) 5'-CAGATGCTTCAGGATGGACAT-3'

[0165] Se utilizaron dos microlitros de cada biblioteca de ADNc como molde y se realizó la PCR con polimerasa Klentaq®. La PCR se realizó durante 45 ciclos de amplificación con desnaturalización a 94°C durante 1 min., hibridación a 55°C durante 1 min., y extensiones a 68°C durante 2 min. Se cargó un quinto de la reacción en gel de agarosa al 1% y se sometió a transferencia Southern. El "blot" se hibridó durante la noche con la sonda de *fused* de longitud completa marcada mediante cebadores al azar tal como se describe para la transferencia Northern.

[0166] Se identificó un fragmento de PCT de 1 kb en cerebro fetal, testículos y ovario. Este fragmento se purificó por gel y se sometió a una secuencia de por PCR directa utilizando ambos cebadores F1 y R4 (SEQ ID NOS. 13 y 14) identificados anteriormente, así como los siguientes cebadores:

hfl6 (SEQ ID NO. 15) 5'-AGAGTAGCAACGTCCTACTGC-3'

hf8 (SEQ ID NO. 16) 5'-CCTCACTGACAAGGCAGCAGG-3'
 hf19 (SEQ ID NO. 17) 5'-CCCGAGGAGGCATCTGCACAG-3'

[0167] La secuencia de este fragmento de 1 kb reveló que estaban ausentes las secuencias de intrón y que los 2 ORF estaban conectados juntos en el mismo marco de lectura. La secuencia de la secuencia empalmada ("spliced") correctamente se muestra en la figura 1 (SEC ID No: 1). El iniciador ATG está presente en la posición 161 y está seguido por un ORF de 3945 nucleótidos que codifica una proteína larga de 1315 aminoácidos con un peso molecular predicho de 144 kDa.

[0168] La similitud global con *fused* de *Drosophila* (SEC ID NO. 23) es del 28% (Figura 2). El dominio N-terminal de 263 aminoácidos de la proteína que contiene el dominio quinasa tiene un 55% de homología con el dominio quinasa de *fused* de *Drosophila*. La parte restante de 1052 aminoácidos de la proteína no es homóloga de manera apreciable con otras proteínas conocidas y, de manera interesante, no es homóloga con la correspondiente región en *fused* de *Drosophila*. De manera destacada, esta región de no homología incluye el extremo C-terminal de la proteína de la mosca que parece ser necesaria para la actividad, Robbins et al., Cell 90: 225-34 (1997); Therond et al., Genetics 142: 1181-98 (1996). Los ADNc empalmados de forma incorrecta descritos anteriormente pueden reflejar el corte y empalme alternativo del gen de *fused* que conduce a la producción de una molécula con un extremo C-terminal truncado y puede ser un mecanismo para regular la actividad de *fused*.

EJEMPLO 5

Reconstitución del *fused* humano de longitud completa empalmado correctamente

[0169] El clon de *fused* DNA28495 (SEQ ID NO. 4) se subclonó a partir del plásmido pRK5B en pRK5.tkneo utilizando ClaI-HindIII. Se realizó la PCR utilizando ADNc de testículos humanos como molde y los cebadores hf3 (SEQ ID NO. 18) (CAGAACTTCAGGTCCTAAAGG) y R4 (véase la secuencia anterior, Ejemplo 4). Las condiciones de PCR fueron 45 ciclos de 94°C, 1 min. a un gradiente de temperatura de 46°C a 68°C, hibridación durante 1 min. y 68°C. 4 min). El fragmento de PCR se digirió con Accl y se ligó en plásmido pRK5.tkneo.*fused*, se cortó con Accl a efectos de sustituir la región que contenía el intrón por la forma empalmada correcta. Se secuenciaron dos subclones entre los dos sitios Accl y tenían la misma secuencia correcta.

EJEMPLO 6

Hibridación *in situ*

[0170] Los embriones de ratón E11.3 y E13.5 se fijaron por inmersión durante la noche a 4°C en paraformaldehído al 4%, se crioprotegieron durante la noche a sacarosa al 15%, se embebieron en O.T.C. y se congelaron en nitrógeno líquido. Los cerebros de ratón adulto se congelaron frescos con hielo seco en polvo. Los cerebros de ratón P1, los testículos de ratón adulto y las médulas espinales de rata adulta se embebieron en O.T.C. y se congelaron en nitrógeno líquido. Las secciones se cortaron a 16 µm, y se procesaron para hibridación *in situ* para *fused* mediante el método de Phillips et al., Science 250: 290-294 (1990). Las sondas de ARN se marcaron con ³³P-UTP tal como se ha descrito por Melton et al., Nucleic Acids Res. 12: 7035-7052 (1984). Se sintetizaron sondas de sentido y antisentido a partir de un fragmento de ADN de *fused* de ratón utilizando T3 y T7, respectivamente, correspondiente a la región que codifica los residuos de aminoácidos 317-486 de la secuencia humana.

[0171] La figura 8 revela que el ARNm de *fused* de ratón se distribuye ampliamente en los tejidos sensibles a *Shh*, incluyendo el tubo neural, el mesodermo presomítico, somitos, esbozos de extremidades en desarrollo y piel. Los transcritos para *fused* también se hallaron en los intestinos, testículos, cartílago y músculo embrionario – Tejidos que están expuestos a otros miembros de la familia de la proteína Hh; Desert e Indian. En el sistema nervioso del ratón E11-5, se detectaron niveles elevados de transcritos de *fused* a lo largo del prosencéfalo, mesencéfalo, rombencéfalo y médula espinal. Estos niveles elevados de expresión se mantuvieron en el día embrionario 13,5. En ambos días embrionarios 11,5 y 13,5, se detectó el ARNm de *fused* principalmente en el aspecto ventral del tubo neural en regiones que están probablemente expuestas a la *Shh* derivada de la línea media ventral. Antes del día después del nacimiento -1, aún se mantiene la expresión extendida de *fused* a lo largo del cerebro con niveles elevados de transcritos detectados en el córtex, hipocampo, epéndima y plexos coroideos. En el adulto, se detectan niveles bajos de expresión de *fused* a lo largo del cerebro con niveles más elevados limitados a la epéndima.

[0172] La distribución en el tejido de *fused* y los componentes del receptor de Hh. Smo y Ptch muestran un solapamiento considerable. Todos ellos se expresan inicialmente a través del tubo neural, así como en los otros tejidos sensibles a Hh. Sin embargo, mientras que el ARNm de Smo se distribuía de forma uniforme a lo largo del eje dorso-ventral, los ARNm de Ptch y *fused* se hallan a niveles elevados centralmente, sugiriendo que se pueden

sobrerregular mediante Hh. Además, mientras que antes del día E12, la expresión de Smo y Ptch se halla principalmente en células que están próximas a la zona ventricular, el ARNm de *fused* aún se expresa de manera amplia y sus niveles sólo disminuyen después. En la expresión adulta de Smo y *fused* se limita a la epéndima donde continúa la neurogénesis.

[0173] También se realizó un análisis detallado de la expresión de *fused* en testículos de adulto mediante hibridación in situ (figura 9). Se observó que se expresaba *fused* a niveles elevados en células germinales de las fases I y II en los túbulos seminíferos. Los niveles de *fused* varían en túbulos seminíferos diferentes, sugiriendo que su expresión es regulada según el estado celular terminal de diferenciación.

EJEMPLO 7

Ensayo de luciferasa con Gli

[0174] Dada la baja homología entre *dfused* y *hfused*, fue prudente determinar si de hecho el *hfused* aislado es un mediador de la señalización de Hh. Se desarrolló el siguiente ensayo para medir la activación del factor de transcripción GLI, el homólogo mamífero del *cubitus interruptus* (Ci) de *Drosophila*. Se ha observado que GLI es un factor de transcripción activado tras la estimulación con Ssh de las células.

[0175] Se introdujeron nueve (9) copias de un sitio de unión a GLI presente en el potenciador HNF3 β , (Sasaki et al., Development 124: 1313-1323 (1997)), delante de un promotor mínimo de timidina quinasa que dirige el gen informador de luciferasa en el plásmido pGL3 (Promega). La secuencia de la secuencia de unión a GLI fue: TCGACAAGCAGGGAACACCCAAGTAGAAGCTC (p9XGliLuc) (SEQ ID NO. 19), mientras que la secuencia de control negativo fue: TCGACAAGCAGGGAAGTGGGAAGTAGAAGCTC (p9XmGliLuc) (SEQ ID NO. 20). Estas construcciones se cotransfectaron con la construcción de *fused* de longitud completa o con un plásmido que codifica *sonic hedgehog* en células C3H10T1/2 desarrolladas en F12: DMEM (50:50), FCS al 10%, e inactivadas con calor. El día previo a la transfección, se inocularon 1×10^5 células por pocillo en 6 placas con pocillos, en 2 ml de medio. El día siguiente, se contranfectó 1 μ g de cada construcción por duplicado con 0,025 mg de plásmido de luciferasa ptkRenilla utilizando lipofectamina (Gibco-BRL) en 100 μ l de OptiMem (con GlutaMAX) según las instrucciones del fabricante durante 3 horas a 37°C. A continuación, se añadió suero (20%, 1 ml) a cada pocillo y las células se incubaron durante 3 horas más a 37°C. A continuación, las células se lavaron dos veces con PBS, a continuación se incubaron durante 48 horas a 37°C en 2 ml de medio. A continuación, cada pocillo se lavó con PBS, y las células se lisaron en 0,5 ml de Tampón de Lisis Pasivo (Promega) durante 15 min a temperatura ambiente en un agitador. El lisado se transfirió en tubos de eppendorf en hielo, se centrifugaron en una centrífuga refrigerada durante 30 segundos y el sobrenadante se guardó en hielo. Para cada medición, se añadieron 20 μ l del lisado celular se añadieron a 100 μ l de LARII (reactivo de ensayo de luciferasa, Promega) en un tubo de polipropileno y se midió la actividad de luz de la luciferasa. La reacción se detuvo mediante la adición de tampón de Stop and Glow (Promega) mezclando pipeteando arriba y debajo de 3 a 5 veces y se midió la actividad de luz de luciferasa de Renilla en el luminómetro.

[0176] Tal como se muestra en la Figura 6, *fused* puede inducir actividad de GLI (9,5 veces) de una manera similar a *SHh* (5,5 veces). Este resultado sugiere que el gen de *fused* aislado es un mediador de la señalización de *SHh*. La serina-treonina quinasa, Akt, irrelevante no era activo en este ensayo (datos no mostrados). La actividad de *fused* es dependiente del dominio quinasa intacto, ya que las moléculas con delección de esta región (C-terminal de *fused* (SEC ID No: 27) o la mutación de un residuo de lisina conservado en aproximadamente el posición de aminoácido 33 en el sitio de unión a ATP (*fused*-DN (SEC ID NO. 25)) no eran capaces de activar GLI. De manera similar, la cola C-terminal de la proteína no es necesaria para esta actividad ya que el dominio quinasa solo no era activo en este ensayo (KD de *fused*) (SEC ID NO. 24). La expresión de cada proteína se verificó mediante transferencia Western utilizando una etiqueta HA insertada en el extremo N-terminal de la molécula (datos no mostrados). Estos resultados sustentan la conclusión de que el homólogo del *dfused* aislado por los solicitantes es de hecho *hfused*. Además, estos resultados indican que *fused* es capaz de y suficiente para la activación de Gli, la principal diana de señalización de *SHh*, y, de este modo, es probable que sea un mediador directo de la señal de *SHh* en vertebrados.

EJEMPLO 8

Ciclopía inducida en embriones de ranas

Introducción:

[0177] Con el objetivo de demostrar que el gen *fused* humano no es sólo capaz sino también requerido para transducir la señal *Shh* en vertebrados, se creó una versión mutante de *fused* conocido como *fused*-DN (dominante negativo) que tenía una mutación de la lisina en la posición 33 en el sitio de unión a ATP (SEC ID N° 25). Este residuo se conserva entre todas las quinasas y es necesario para la actividad de quinasas (Hanks et al. Methods Enzymol. 200: 38-62 (1991) y su conversión a cualquier otro residuo en la mayoría de casos da lugar a la creación de mutantes negativos dominantes.

Métodos:

Construcción de plásmido:

[0178] El ADNc de *fused* de tipo salvaje (SEC ID N°1) con una etiqueta HA insertada en el extremo carboxilo se subclonó en pRK5 y se generó una forma negativa dominante mediante la conversión de la lisina en el positivo 33 en una arginina. Se preparó el ADN de plásmido superenrollado por Qiagen y se utilizó para la inyección en embrión de *Xenopus laevis*.

Manipulación de embriones de *Xenopus*:

[0179] Se estimularon ranas hembra adultas con 200 U.I. de suero de yegua embarazada 3 días antes de su uso y con 800 U.I. de gonadotropina coriónica humana la noche antes de la inyección. Se extrajeron oocitos nuevos de las ranas hembra la mañana siguiente y se realizó la fertilización *in vitro* de los oocitos mediante la mezcla de oocitos con testículos molidos de ranas macho sacrificadas. Se mantuvieron y organizaron en fases los embriones en desarrollo de acuerdo con Nieuwkoop y Faber, Normal Table of *Xenopus laevis*, N.-H. P. Co., ed. (Amsterdam, 1967).

[0180] Se desgelatinizaron los huevos fertilizados con cisteína al 2% (pH 7,8) durante 10 minutos, se lavaron una vez con agua destilada y se transfirieron a 0,1 x MBS con Ficoll al 5%. Se recubrieron bandejas de inyección con los huevos fertilizados en 0,1 x MBS con Ficoll al 5%. Se inyectaron embriones de *Xenopus* en desarrollo en la fase de dos células con 200 pg de pRK5 que contenía *fused* de tipo salvaje (WT (SEC ID N° 1)) o *fused* negativo dominante (DH (SEC ID NO. 25)). Se guardaron los embriones inyectados en bandejas durante otras 6 horas, tras las cuales se transfirieron a 0,1 x MBS con gentamicina 50 mg/ml durante 3 días hasta alcanzar la fase 35 de Nieuwkoop en el que el desarrollo del ojo es completo.

Resultados:

[0181] Para evaluar si el gen *fused* humano actúa como un transductor de señales de la señalización de Hedgehog, se inyectaron la forma de tipo salvaje o negativa dominante de *fused* humano en embriones de rana en desarrollo. Los embriones inyectados con 120 pg de ADN se dividieron normalmente en la fase de blástula y gastrulación. Mientras el desarrollo del ojo fue normal en el tipo salvaje, los embriones inyectados con *fused* (SEC ID N° 2) e inyectados con mock (simulación), aproximadamente el 30% (Tabla 1) de los embriones que se inyectaron con *fused*-DN mostraron estructura de ojos fusionados o dos ojos conectados por tejido de retina algo pigmentado (Figura 11A). En la Tabla 1, se liberaron 200 pg de ADN de plásmido al grupo de animales de embriones de la fase de 2 células. Cada muestra representa los resultados de por lo menos 3 experimentos independientes. Se valoraron los embriones visualmente por sus defectos de ciclopía.

TABLA I

Ciclopía inducida por fusión-DN en embriones de <i>Xenopus</i>			
ADN inyectado	Normal	Cíclope	n
Hu- <i>fused</i> (SEC ID N° 2)	45	0	45
Dominio de kinasa (SEC ID N° 24)	43	0	43
C-terminal (SEC ID N° 27)	53	1	54
<i>fused</i> DN (SEC ID N° 25)	32	15	47
No inyectado	61	0	61

[0182] El fenotipo de ciclopía observado es notablemente similar a uno de los embriones de ratón deficientes en *Shh* (Chiang et al., Nature 383: 407-13 (1996) y de los embriones de pez zebra donde la señalización de *Shh* ha sido bloqueada por la sobreexpresión de una PKA activa constitutiva. Hammerschmidt et al., Genes Dev. 10: 647-58 (1996); Ungar and Moon. Dev. Biol. 178:186-91 (1996). Además, tanto el desarrollo del cerebro (prosencéfalo) como

el del intestino parecieron normales en fases posteriores del desarrollo del renacuajo en los embriones inyectados con *fused*-DN (SEQ ID NO. 25) (Figura 11B). En cambio, los embriones que sobreexpresaban *fused* de tipo salvaje (SEC ID N° 2) o mutantes de truncamiento terminal N o C-terminales (SEC ID N° 27 y 24, respectivamente) no presentaron anomalía alguna.

[0183] Durante el desarrollo normal del ojo de *Xenopus*, el primordio del ojo comienza como un campo único que expresa el factor de transcripción Pax-6, que es un homólogo de vertebrado de *Drosophila* sin ojos, Li et al., Development, 124: 603-15 (1997). En la fase de neurula, este campo visual se separa en dos primordios de ojo debido a una señal inhibidora del mesodermo precordial. Se ha demostrado adicionalmente que *SHh* es la señal derivada del mesodermo precordial que es responsable de la inhibición de la expresión de Pax-6 en la línea media del campo visual.

[0184] Para comprender adicionalmente como la sobreexpresión de *fused*-DN (SEC ID N° 25) inducía un ojo fusionado en embriones de *Xenopus*, se realizó una hibridación in situ de todo el sorrote con el objetivo de determinar el patrón de expresión de Pax-6 en embriones inyectados. Tal y como se muestra en la Figura 11C, la expresión de Pax-6 en embriones inyectados con *fused*-DN (SEC ID N° 25) permanece como un único campo (Figura 11D). De este modo, *fused*-DN (SEC ID N° 25) induce un fenotipo de ciclopía muy probablemente evitando que *SHh* inhiba la expresión de Pax-6 en la línea media del campo visual.

EJEMPLO 9

Rescate de embriones de *Xenopus* inyectados con *fused*-DN (SEC ID NO. 25) por Gli

[0185] La expresión de células de placa del piso tempranas es inducida por *SHh* producida por los notocordios. Para analizar si la expresión de *SHh* en la placa del piso también se inhibirá cuando se bloquea la señalización de *SHh*, se tiñeron embriones en la fase de neurula temprana inyectadas con construcciones de *fused*-DN o de tipo salvaje para la expresión de *SHh* (véase el ejemplo 8 para el procedimiento). La expresión de *SHh* en células de placa del piso o embriones en la fase de neurula temprana se suprimió completamente en 26 de 28 embriones inyectados cuando se sobreexpresa en *fused* mutado (tabla 2, figura 11C, embrión izquierdo), mientras que la expresión de *SHh* no se vio afectada en los embriones de control (figura 6E, embrión derecho). La tabla 2 representa los datos evaluados para tres experimentos independientes. Se inyectaron 100 µg de *fused*-DN, 100 µg de *fused*-wt ó 50 µg de plásmido de Gli-1 en embriones en la etapa celular 2. Los embriones se recogieron en la fase de neurula temprana para la tinción de *SHh*.

TABLA 2

Expresión de <i>SHh</i> en el rescate <i>fused</i> de tipo salvaje y Gli en placa de piso cuando se coexpresa con <i>fused</i> -DN		
	Tinción de <i>SHh</i>	porcentaje
<i>fused</i> -DN	2/28	7%
<i>fused</i> -DN + <i>fused</i> WT	20/24	83%
<i>fused</i> -DN + Gli	36/36	100

[0186] Para confirmar que este fenotipo era debido a la inhibición específica del mecanismo de señalización de *SHh* en la placa del piso, se intentó rescatar el fenotipo mediante la coinfección de ARN de *fused* wt con ARN de *fused*-DN en una proporción 1:1. La tabla 2 muestra que más de un 80% de los embriones coinfectados con ARN de *fused* wt y *fused*-DN muestran una tinción de *SHh* normal en la placa del piso. Esto demuestra que la expresión de *SHh* en los embriones inyectados con *fused*-DN es bloqueada específicamente por la inhibición de la actividad de *fused* endógena.

[0187] Para demostrar adicionalmente que el fenotipo observado de *fused*-DN es debido a la alteración de la cascada de señal de *SHh* y para confirmar que *hfused* actúa en cascada arriba de Gli en este mecanismo, nos preguntamos si la sobreexpresión de Gli también puede rescatar el fenotipo de embriones de *Xenopus* inyectados con *fused*-DN. Tal como se muestra en Tabla 2, el rescate de la expresión de *SHh* en la placa del piso de embriones inyectados con *fused*-DN es completo cuando se sobreexpresa Gli. Conjuntamente, estos hallazgos concuerdan con la hipótesis de los solicitantes de que las funciones del *fused* de vertebrado en el mecanismo de *SHh* y que es un mediador necesario en el mecanismo de transducción de la señal de *SHh*, que actúa cascada arriba de Gli.

EJEMPLO 10

Inmunoprecipitaciones y ensayo de quinasa *in vitro*

[0188] Para determinar directamente si *hfused* tiene actividad quinasa, se marcaron los ADNc de *fused* (SEQ ID NO. 2), *fused*-DN (SEQ ID NO. 25) y *fused*-kd (SEQ ID NO. 24) con el epítipo etiqueta de HA de gripe y se transfectaron transitoriamente en células 293. Se analizaron los inmunoprecipitados para la actividad de quinasa en presencia de proteína básica de mielina (MBP) y γ -³²P-ATP. La cantidad de ³²P incorporado en MBP se determinó después de SDS-PAGE y se encontró que era aproximadamente 3 veces superior a los extractos que contenían *fused*-KD (SEQ ID NO. 24) y 2 veces superior en los extractos que contenían *fused* wt (SEQ ID NO. 2) en comparación con los controles, mientras que la mutación de Lys33 en Arg (*fused*-DN (SEQ ID NO. 25)) neutraliza la actividad (Figura 12).

[0189] Para los experimentos de inmunoprecipitación, se transfectaron transitoriamente células 293 de riñón embrionario humano con los diversos plásmidos de expresión. Después de 24 horas, se recogieron las células transfectadas y se lisaron durante 20 min a 4°C en 1 ml de tampón de lisis (Tris 50 mM, pH 8,0), NaCl 150 mM, EDTA 1 mM, fluoruro de sodio 1 mM, ortovanadato de sodio 1 mM, PMSF 1 mM e inhibidores de proteasa (Complete. Boehringer Mannheim) que contenían NP-40 al 1%, ácido desoxicólico al 0,5%. Se extrajo el residuo celular mediante centrifugación durante 10 min a 10.000 rpm y se incrementó la concentración de cloruro de sodio de los lisados celulares hasta 250 mM. El sobrenadante se prepurificó durante 1 hora con 20 µl de Proteína A Sefarosa (Pharmacia). Los lisados se inmunoprecipitaron utilizando anticuerpos anti-HA seguido de Proteína A Sefarosa. Las partículas se lavaron dos veces con tampón de lisis que contenía cloruro de sodio 250 mM, dos veces con tampón de lisis que contenía cloruro de sodio 1 M y a continuación dos veces con tampón de ensayo de quinasa (HEPES 20 mM, pH 7,6), DTT 1 mM, NaF 1 mM y ortovanadato de sodio 1 mM). Después del último lavado, las partículas se resuspendieron en 20 µl de tampón de ensayo de quinasa suplementado con 10 mCi de γ -³²P-ATP, 20 mM de β-glicerofosfato, 20 mM de PNPP, 20 mM de MgCl₂, 1 mM de EGTA, 100 µM de ATP frío y 0,5 mg/ml de Proteína Básica de Mielina (Sigma) y se incubaron durante 20 min. a 37°C. Las reacciones se detuvieron con 20 µl de tampón-SDS, se desarrollaron en un gel de poliacrilamida de SDS al 4-20% desnaturizante y se analizaron mediante fosfoimager.

EJEMPLO 11

Expresión de *fused* en *E. coli*

[0190] La secuencia de ADN que codifica el *fused* humano se amplificó inicialmente con cebadores de PCR seleccionados. Los cebadores deben contener sitios para enzimas de restricción que corresponden con los sitios para enzimas de restricción en el vector de expresión seleccionado. Pueden utilizarse diversos vectores de expresión. Un ejemplo de un vector adecuado es pBR322 (derivado de *E. coli*; véase Bolivar et al., Gene, 2:95 (1977)) que contiene genes para la resistencia a ampicilina y tetraciclina. El vector se digiere con la enzima de restricción y se desfosforila. Las secuencias de PCR amplificadas se unen a continuación en el vector. El vector incluirá preferentemente secuencias que codifican un gen de resistencia a un antibiótico, un promotor *trp*, una secuencia líder poli-His (incluyendo los 6 primeros codones STII, la secuencia poli-His, y un sitio de división por enterocinasa), la región codificante de *fused* de vertebrado, el terminador transcripcional de lambda y un gen *argU*.

[0191] A continuación, la mezcla de unión se utiliza para transformar una cepa seleccionada de *E. coli* utilizando los procedimientos descritos en Sambrook et. al., *supra*. Los transformantes se identifican por su capacidad para crecer en placas de LB y, a continuación, se seleccionan las colonias resistentes a los antibióticos. El ADN plasmídico se puede aislar y confirmar mediante el análisis de restricción y la secuenciación del ADN.

[0192] Los clones seleccionados se pueden desarrollar durante toda la noche en un medio de cultivo líquido tal como el caldo LB suplementado con antibióticos. El cultivo de toda la noche se puede utilizar posteriormente para inocular un cultivo a mayor escala. A continuación, las células se desarrollan hasta una densidad óptica deseada, durante la cual el promotor de la expresión se activa.

[0193] Después de cultivar las células durante muchas más horas, las células se pueden recoger mediante centrifugación. El residuo de células obtenido mediante la centrifugación se puede solubilizar utilizando diversos agentes conocidos en la técnica, y, a continuación, la proteína *fused* de vertebrado solubilizada se puede purificar utilizando una columna quelante de metal en condiciones que permiten la unión fuerte de la proteína.

EJEMPLO 12

Expresión de *fused* en células de mamífero

[0194] El vector pRK5 (véase, EP 307.247 publicada el 15 de marzo de 1989) se utiliza como vector de expresión. Opcionalmente, el ADN de *fused* de vertebrado se liga en el pRK5 con enzimas de restricción seleccionadas para

permitir la inserción del ADN de *fused* de vertebrado utilizando procedimientos de unión tales como los descritos en Sambrook et. al., *supra*. El vector resultante se denomina pRK5-*fused*.

[0195] En una realización, las células huésped seleccionadas pueden ser células 293. Las células 293 humanas (ATCC CCL 1573) se desarrollan hasta la confluencia en placas de cultivo de tejidos en un medio tal como DMEM suplementado con suero de ternera fetal y opcionalmente, componentes nutricionales y/o antibióticos. Se mezclan aproximadamente 10 µg de ADN de pRK5-*fused* con aproximadamente 1 µg de ADN que codifica el gen de ARN VA [Thimmappaya et. al., *Cell*, 31:543 (1982)] y se disuelve en 500 µl de Tris-HCl 1 mM, EDTA 0,1 mM, CaCl₂ 0,227 M. A esta mezcla se le añade, gota a gota, 500 µl de HEPES 50 mM (pH 7,35), NaCl 280 mM, NaPO₄ 1,5 mM, y se deja formar un precipitado durante 10 minutos a 25°C. El precipitado se suspende y se añade a las células 293 y se deja reposar durante aproximadamente cuatro horas a 37°C. El medio de cultivo se aspira y se añaden 2 ml de glicerol al 20% en PBS durante 30 segundos. A continuación, las células 293 se lavan con medio sin suero, se añade medio fresco y las células se incuban durante aproximadamente 5 días.

[0196] Aproximadamente 24 horas después de las transfecciones, el medio de cultivo se extrae y se reemplaza por medio de cultivo (solo) o medio de cultivo que contiene 200 µCi/ml de ³⁵S-cisteína y 200 µCi/ml de ³⁵S- metionina. Tras una incubación de 12 horas, se recoge el medio condicionado, se concentra en un filtro de centrifugación y se carga en un gel SDS al 15%. El gel procesado puede secarse y exponerse a una película durante un período de tiempo concreto para revelar la presencia del polipéptido *fused* de vertebrado. Los cultivos que contienen células transfectadas pueden experimentar una incubación adicional (en medio sin suero) y el medio se examina en bioensayos concretos.

[0197] En una técnica alternativa, se puede introducir *fused* de vertebrado en células 293 transitoriamente utilizando el procedimiento del sulfato de dextrano descrito por Somparyrac et. al., *Proc. Natl. Acad. Sci.*, 12: 1575 (1981). Las células 293 se desarrollan hasta la máxima densidad en un matraz giratorio y se añaden 700 µg de ADN de pRK5-*fused*. En primer lugar, las células se concentran a partir del matraz giratorio mediante centrifugación y se lavan con PBS. El precipitado de ADN-dextrano se incuba en el residuo celular durante cuatro horas. Las células se tratan con glicerol al 20% durante 90 segundos, se lavan con medio de cultivo de tejido, y se reintroducen en el matraz giratorio que contiene el medio de cultivo de tejidos, 5 µg/ml de insulina bovina y 0,1 µg/ml de transferrina bovina. Después de aproximadamente cuatro días, el medio condicionado se centrifuga y se filtra para eliminar las células y restos celulares. La muestra que contiene el *fused* de vertebrado expresado se puede a continuación concentrar y purificar mediante cualquier procedimiento seleccionado, tal como la diálisis y/o cromatografía en columna.

[0198] En otra realización, puede expresarse *fused* de vertebrado en células CHO. El vector pSUi-*fused* puede transfectarse en células CHO utilizando reactivos conocidos, tales como CaPO₄ o DEAE-dextrano. Tal y como se ha descrito anteriormente, los cultivos celulares pueden incubarse, y el medio puede sustituirse por medio de cultivo (solo) o medio que contiene un marcador radioactivo, tal como ³⁵S-metionina. Después de determinar la presencia del polipéptido *fused* de vertebrado, el medio de cultivo puede sustituirse por medio sin suero. Preferiblemente, los cultivos se incuban durante aproximadamente 6 días y, a continuación, se recoge el medio condicionado. A continuación, el medio que contiene el *fused* de vertebrado expresado puede concentrarse y purificarse mediante cualquier procedimiento seleccionado.

[0199] El *fused* de vertebrado etiquetado con epítipo puede expresarse también en células CHO huéspedes. El *fused* de vertebrado puede subclonarse fuera del vector pRK5. El inserto del subclón puede someterse a PCR para fusionarse en el marco con una etiqueta epítipo seleccionada, tal como una etiqueta de poli-His en un vector de expresiones. El inserto de *fused* de vertebrado etiquetado con poli-His puede subclonarse a continuación en un vector dirigido por SV40 que contiene un marcador de selección, tal como DHFR, para seleccionar clones estables. Finalmente, las células CHO pueden transfectarse (tal como se ha descrito anteriormente) con el vector dirigido por SV40. El marcaje puede realizarse, tal como se ha descrito anteriormente, para verificar la expresión. El medio de cultivo que contiene el *fused* de vertebrado etiquetado con poli-His expresado puede a continuación concentrarse y purificarse mediante cualquier procedimiento seleccionado, tal como mediante cromatografía de afinidad de quelato con Ni²⁺.

EJEMPLO 13

Expresión de *fused* de vertebrado en levadura

[0200] El siguiente método describe la expresión recombinante de *fused* de vertebrado en levadura.

[0201] En primer lugar, los vectores de expresión de levadura se construyen para la producción o secreción intracelular de *fused* de vertebrado a partir del promotor ADH2/GAPDH. El ADN que codifica el *fused* de vertebrado,

un péptido señal seleccionado y el promotor se insertan en los sitios para enzimas de restricción adecuados en el plásmido seleccionado para dirigir la expresión intracelular de *fused* de vertebrado. Para la secreción, el ADN que codifica el *fused* de vertebrado puede clonarse en el plásmido seleccionado, junto con el ADN que codifica el promotor ADH2/GAPDH, la secuencia señal/líder secretora del factor alfa o invertasa de levadura, y secuencias adaptadoras (si se necesitan) para la expresión de *fused* de vertebrado.

[0202] Las células de levadura, tales como la cepa AB110 de la levadura, pueden a continuación transformarse con los plásmidos de expresión descritos anteriormente y cultivarse en medios de fermentación seleccionados. Los sobrenadantes de levadura transformados pueden analizarse mediante precipitación con ácido tricloroacético al 10% y una separación mediante SDS-PAGE, seguido de la tinción de los geles con azul de Coomassie.

[0203] El *fused* de vertebrado recombinante puede aislarse posteriormente y purificarse mediante la extracción de las células de levadura del medio de fermentación por centrifugación y, a continuación, la concentración del medio utilizando filtros de cartucho específicos. El concentrado que contiene *fused* de vertebrado puede purificarse adicionalmente utilizando resinas de cromatografía en columna seleccionadas.

EJEMPLO 14

Expresión de *fused* de vertebrado en células de insecto infectadas por Baculovirus

[0204] El siguiente método describe la expresión recombinante de *fused* de vertebrado en células de insecto infectadas por Baculovirus.

[0205] La secuencia que codifica *fused* de vertebrado se fusiona en dirección 5' de un epítipo etiqueta contenido en un vector de expresión de baculovirus. Dichas epítipo etiquetas incluyen etiquetas de poli-His y etiquetas de inmunoglobulina (como las regiones Fc de IgG). Pueden utilizarse un conjunto de plásmidos, incluyendo plásmidos derivados de los plásmidos disponibles comercialmente, tales como pVL1393 (Novagen). Brevemente, la secuencia que codifica *fused* de vertebrado o la parte deseada de la secuencia codificante de *fused* de vertebrado (tal como la secuencia que codifica el dominio extracelular de una proteína transmembrana) se amplifican mediante PCR con cebadores complementarios a las regiones 5' y 3'. El cebador 5' puede incorporar sitios para enzimas de restricción flanqueantes (seleccionados). El producto se digiere a continuación con las enzimas de restricción seleccionadas y se subclona en el vector de expresión.

[0206] El baculovirus recombinante se genera mediante la cotransfección del plásmido anterior y el ADN del virus BaculoGold™ (Pharmingen) en células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711) utilizando lipofectina (disponible comercialmente de GIBCO-BRL). Después de 4-5 días de incubación a 28°C, los virus liberados se recogen y se utilizan para amplificaciones adicionales. La infección viral y la expresión de la proteína se realizan tal y como se describe en O'Reilly et. al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

[0207] A continuación, el *fused* de vertebrado etiquetado con poli-His expresado puede purificarse, por ejemplo, mediante cromatografía de afinidad de Ni²⁺-quelato tal y como se indica a continuación. Los extractos se preparan a partir de las células recombinantes de Sf9 infectadas del virus tal y como se ha descrito por Rupert et. al., *Nature*, 362: 175-179 (1993). Brevemente, las células Sf9 se lavan, se resuspenden en el tampón de sonicación (25 ml de HEPES, pH 7,9; MgCl₂ 12,5 mM; EDTA 0,1 mM; glicerol al 10%; NP-40 al 0,1%; KCl 0,4 M), y se sonicán dos veces durante 20 segundos en hielo. Los sonicados se depuran por centrifugación, y el sobrenadante se diluye 50 veces en el tampón de carga (fosfato 50 mM, NaCl 300 mM, glicerol al 10%, pH 7,8) y se filtra a través de un filtro de 0,45 µm. Se prepara una columna de agarosa Ni²⁺-NTA (comercialmente disponible de Qiagen) con un volumen de lecho de 5 ml, se lava con 25 ml de agua y se equilibra con 25 ml del tampón de carga. El extracto celular filtrado se carga en la columna a 0,5 ml por minuto. La columna se lava hasta la línea base a A₂₈₀ con el tampón de carga, en cuyo punto se inicia la recogida de la fracción. A continuación, la columna se lava con un tampón de lavado secundario (fosfato 50 mM; NaCl 300 mM, glicerol al 10%, pH 6,0), que eluye la proteína no unida específicamente. Después de alcanzar la línea base a A₂₈₀ de nuevo, la columna se desarrolla con un gradiente de 0 a 500 mM de imidazol en el tampón de lavado secundario. Se recogen fracciones de un ml y se analizan mediante SDS-PAGE y tinción con plata o transferencia Western con Ni²⁺-NTA-conjugado a fosfatasa alcalina (Qiagen). Las fracciones que contienen el *fused* de vertebrado etiquetado con His₁₀ eluido se agrupan y se dializan contra el tampón de carga. Alternativamente, la purificación del *fused* de vertebrado etiquetado con IgG (o con Fc) puede realizarse usando técnicas de cromatografía conocidas, incluyendo, por ejemplo, cromatografía en columna con Proteína A o proteína G.

EJEMPLO 15

Preparación de anticuerpos que se unen a *fused* de vertebrado

[0208] Este ejemplo ilustra la preparación de anticuerpos monoclonales que se pueden unir específicamente a *fused* de vertebrado.

[0209] Las técnicas para producir los anticuerpos monoclonales son conocidas en el sector y están descritas, por ejemplo, en Goding, *supra*. Entre los inmunógenos que se pueden utilizar se incluyen *fused* de vertebrado purificado, proteínas de fusión que contienen *fused* de vertebrado y células que expresan *fused* de vertebrado recombinante en la superficie celular. La selección del inmunógeno puede realizarse según el técnico en la materia sin una gran experimentación.

[0210] Los ratones, tales como Balb/c, se inmunizan con el inmunógeno de *fused* de vertebrado emulsionado en adyuvante completo de Freund y se inyecta subcutáneamente o intraperitonealmente en una cantidad de 1–100 microgramos. Alternativamente, el inmunógeno se emulsiona en el adyuvante de MPL-TDM (Ribi Immunochemical Research, Hamilton, MT) y se inyecta en las bases de las patas traseras del animal. A continuación, los ratones inmunizados se refuerzan 10 a 12 días después con inmunógeno adicional emulsionado en el adyuvante seleccionado. A continuación, durante diversas semanas, los ratones también se pueden reforzar con inyecciones de inmunización adicionales. Las muestras de suero se pueden obtener periódicamente de los ratones mediante muestras de sangre retro-orbital para ser analizadas en ensayos ELISA para detectar anticuerpos de *fused* de vertebrado.

[0211] Después de detectar un título de anticuerpo adecuado, a los animales "positivos" para anticuerpos se les puede inyectar una inyección intravenosa final de *fused* de vertebrado. De tres a cuatro días más tarde, los ratones se sacrifican y se recogen las células del bazo. A continuación, las células del bazo se fusionan (usando polietilenglicol al 35%) a una línea celular de mieloma murino seleccionada, tal como la P3X63AgU.1, disponible de ATCC, No. CRL 1597. Las fusiones generan células de hibridoma que se pueden colocar a continuación en placas de cultivo de tejido de 96 pocillos que contienen un medio HAT (hipoxantina, aminopterina y timidina) para inhibir la proliferación de células no fusionadas, híbridos de mieloma e híbridos de células de bazo.

[0212] Las células de hibridomas se cribarán en un ELISA para la reactividad contra *fused* de vertebrado. La determinación de células de hibridomas "positivas" que secretan los anticuerpos monoclonales deseados contra *fused* de vertebrado está dentro de la técnica.

[0213] Las células de hibridomas positivas se pueden inyectar intraperitonealmente en ratones singéneos Balb/c para producir fluidos ascíticos que contienen los anticuerpos monoclonales anti-*fused* de vertebrado. Alternativamente, las células de hibridoma pueden desarrollarse en matraces o en botellas en rodillo de cultivos de tejidos. La purificación de los anticuerpos monoclonales producidos en los fluidos ascíticos se puede realizar usando precipitación con sulfato de amonio, seguido por cromatografía de exclusión en gel. Alternativamente, puede usarse la cromatografía por afinidad basada en la unión del anticuerpo a la proteína A o la proteína G.

Depósito de Material

[0214] Los siguientes materiales se han depositado con la American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Designación	ATCC Dep. No.	Data del depósito
PRK5tkneo.h <i>Fused</i> .1271	209637	19/2/1998

[0215] Este depósito se realizó según lo estipulado en el Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los fines del Procedimiento en Materia de Patentes y el Reglamento bajo el mismo (Tratado de Budapest). Esto asegura el mantenimiento de un cultivo viable del depósito durante 30 años a partir de la fecha del depósito. El depósito estará disponible mediante la ATCC según los términos del Tratado de Budapest, y está sujeto a un acuerdo entre Genentech, Inc. y ATCC, que asegura la disponibilidad permanente y sin restricción de la progenie del cultivo del depósito al uso público tras la publicación de la respectiva patente estadounidense o tras ponerse abierta a la inspección pública de cualquier solicitud de patente estadounidense o extranjera, la que sea primera, y asegura la disponibilidad de la progenie para alguien determinado por la U.S. Commissioner of Patents and Trademarks para tener el derecho a la misma de acuerdo con 35 USC § 122 y las normas de la Commissioner según las mismas (incluyendo 37 CFR § 1.14 con referencia concreta a 886 OG 638).

[0216] El cesionario de la presente solicitud ha acordado que si un cultivo de los materiales en el depósito muriera o se perdiera o se destruyera cuando se cultiva en las condiciones adecuadas, los materiales serán inmediatamente reemplazados en una notificación por otros iguales. La disponibilidad del material depositado no se interpreta como una licencia para realizar la invención contraviniendo los derechos concedidos bajo la autoridad de cualquier gobierno de acuerdo con sus leyes de patente.

[0217] La memoria escrita anterior se considera que es suficiente para permitir a un experto en la materia realizar la invención. La presente invención no se limita en su alcance por la construcción depositada, ya que la realización depositada pretende ser una ilustración individual de ciertos aspectos de la presente invención y otras construcciones que son funcionalmente equivalentes están dentro del alcance de la presente invención. El depósito del material de la presente invención no constituye una admisión de que la descripción escrita contenida en la presente invención sea inadecuada para permitir la práctica de cualquier aspecto de la invención, incluyendo el modo óptimo de la misma, ni se interpreta como limitante del alcance de las reivindicaciones a las ilustraciones específicas que representa.

LISTADO DE SECUENCIAS

[0218]

<110> Genentech, Inc.

<120> *Fused*

<130> P1272PCT

<141> 1999-02-25

<150> US 09/031,563

<151> 1998-02-26

<160> 27

<210> 1

<211> 4880

<212> DNA

<213> Homo sapiens

<220>

<221> N desconocido

<222> 4160

<223> desconocido

<220>

<221> N desconocido

<222> 4243

<223> desconocido

<220>

<221> N desconocido

<222> 4361

<223> desconocido

<400> 1

cccggggatc ctctagagat ccctcgacct cgaccacgc gtccgcccac 50

gcgtccgccc acgcgtccgg ggcgtcccag atgttggtga actgtccctg 100

gatctatagc tcttcaccgt ctctactttc ttccttctaa gagatcctga 150

aacctctgtc atg gaa aag tac cac gtg ttg gag atg att 190

Met Glu Lys Tyr His Val Leu Glu Met Ile
1 5 10

ES 2 371 853 T3

5	gga Gly	gaa Glu	ggc Gly	tct Ser	ttt Phe 15	ggg Gly	agg Arg	gtg Val	tac Tyr	aag Lys 20	ggt Gly	cga Arg	aga Arg	229
	aaa Lys	tac Tyr 25	agt Ser	gct Ala	cag Gln	gtc Val	gtg Val 30	gcc Ala	ctg Leu	aag Lys	ttc Phe	atc Ile 35	cca Pro	268
10	aaa Lys	ttg Leu	ggg Gly	cgc Arg 40	tca Ser	gag Glu	aag Lys	gag Glu	ctg Leu 45	agg Arg	aat Asn	ttg Leu	caa Gln	307
15	cga Arg 50	gag Glu	att Ile	gaa Glu	ata Ile	atg Met 55	cgg Arg	ggt Gly	ctg Leu	cgg Arg	cat His 60	ccc Pro	aac Asn	346
20	att Ile	gtg Val	cat His 65	atg Met	ctt Leu	gac Asp	agc Ser	ttt Phe 70	gaa Glu	act Thr	gat Asp	aaa Lys	gag Glu 75	385
25	gtg Val	gtg Val	gtg Val	gtg Val	aca Thr 80	gac Asp	tat Tyr	gct Ala	gag Glu	gga Gly 85	gag Glu	ctc Leu	ttt Phe	424
	cag Gln	atc Ile 90	cta Leu	gaa Glu	gat Asp	gac Asp	gga Gly 95	aaa Lys	ctt Leu	cct Pro	gaa Glu	gac Asp 100	cag Gln	463
30	gtt Val	cag Gln	gcc Ala	att Ile 105	gct Ala	gcc Ala	cag Gln	ttg Leu	gtg Val 110	tca Ser	gcc Ala	ctg Leu	tac Tyr	502
35	tat Tyr 115	ctg Leu	cat His	tcc Ser	cac His	cgc Arg 120	atc Ile	cta Leu	cac His	cga Arg	gat Asp 125	atg Met	aag Lys	541
40	cct Pro	cag Gln	aac Asn 130	atc Ile	ctc Leu	ctc Leu	gcc Ala	aag Lys 135	ggt Gly	ggt Gly	ggc Gly	atc Ile	aag Lys 140	580
45	ctc Leu	tgt Cys	gac Asp	ttt Phe	gga Gly 145	ttt Phe	gcc Ala	cgg Arg	gct Ala	atg Met 150	agc Ser	acc Thr	aat Asn	619
	aca Thr	atg Met 155	gtg Val	ctg Leu	aca Thr	tcc Ser	atc Ile 160	aaa Lys	ggc Gly	aca Thr	cca Pro	ctc Leu 165	tat Tyr	658
50	atg Met	tct Ser	cca Pro	gag Glu 170	ctg Leu	gtg Val	gag Glu	gag Glu	cga Arg 175	cca Pro	tac Tyr	gac Asp	cac His	697
55	aca Thr 180	gcg Ala	gac Asp	ctc Leu	tgg Trp	tct Ser 185	gtt Val	ggc Gly	tgc Cys	ata Ile	cta Leu 190	tat Tyr	gaa Glu	736
60	ctg Leu	gca Ala	gta Val 195	ggc Gly	acc Thr	cct Pro	ccc Pro	ttc Phe 200	tat Tyr	gct Ala	aca Thr	agc Ser	atc Ile 205	775
	ttt Phe	cag Gln	ctg Leu	gtc Val	agc Ser 210	ctc Leu	att Ile	ctc Leu	aag Lys	gac Asp 215	cct Pro	gtg Val	cgc Arg	814
65	tgg Trp	ccc Pro 220	tca Ser	acc Thr	atc Ile	agt Ser	ccc Pro 225	tgc Cys	ttt Phe	aag Lys	aac Asn	ttc Phe 230	ctg Leu	853
70	cag	gga	ctg	ctc	acc	aaa	gac	cca	cgg	cag	cga	ctg	tcc	892

ES 2 371 853 T3

	Gln	Gly	Leu	Leu	Thr	Lys	Asp	Pro	Arg	Gln	Arg	Leu	Ser	
			235						240					
5	tgg	cca	gac	ctc	tta	tat	cac	ccc	ttt	att	gct	ggt	cat	931
	Trp	Pro	Asp	Leu	Leu	Tyr	His	Pro	Phe	Ile	Ala	Gly	His	
	245				250					255				
10	gtc	acc	ata	ata	act	gag	cca	gca	ggc	cca	gat	ttg	ggg	970
	Val	Thr	Ile	Ile	Thr	Glu	Pro	Ala	Gly	Pro	Asp	Leu	Gly	
			260					265					270	
15	acc	cca	ttc	acc	agc	cgc	cta	ccc	cca	gaa	ctt	cag	gtc	1009
	Thr	Pro	Phe	Thr	Ser	Arg	Leu	Pro	Pro	Glu	Leu	Gln	Val	
					275					280				
20	cta	aag	gac	gaa	cag	gcc	cat	cgg	ttg	gcc	ccc	aag	ggt	1048
	Leu	Lys	Asp	Glu	Gln	Ala	His	Arg	Leu	Ala	Pro	Lys	Gly	
		285					290					295		
25	aat	cag	tct	cgc	atc	ttg	act	cag	gcc	tat	aaa	cgc	atg	1087
	Asn	Gln	Ser	Arg	Ile	Leu	Thr	Gln	Ala	Tyr	Lys	Arg	Met	
				300					305					
30	gct	gag	gag	gcc	atg	cag	aag	aaa	cat	cag	aac	aca	gga	1126
	Ala	Glu	Glu	Ala	Met	Gln	Lys	Lys	His	Gln	Asn	Thr	Gly	
	310				315						320			
35	cct	gcc	ctt	gag	caa	gag	gac	aag	acc	agc	aag	gtg	gct	1165
	Pro	Ala	Leu	Glu	Gln	Glu	Asp	Lys	Thr	Ser	Lys	Val	Ala	
			325					330					335	
40	cct	ggc	aca	gcc	cct	ctg	ccc	aga	ctc	ggg	gcc	act	cct	1204
	Pro	Gly	Thr	Ala	Pro	Leu	Pro	Arg	Leu	Gly	Ala	Thr	Pro	
				340						345				
45	cag	gaa	tca	agc	ctc	ctg	gcc	ggg	atc	tta	gcc	tca	gaa	1243
	Gln	Glu	Ser	Ser	Leu	Leu	Ala	Gly	Ile	Leu	Ala	Ser	Glu	
		350					355					360		
50	ttg	aag	agc	agc	tgg	gct	aaa	tca	ggg	act	gga	gag	gtg	1282
	Leu	Lys	Ser	Ser	Trp	Ala	Lys	Ser	Gly	Thr	Gly	Glu	Val	
				365					370					
55	ccc	tct	gca	cct	cgg	gaa	aac	cgg	acc	acc	cca	gat	tgt	1321
	Pro	Ser	Ala	Pro	Arg	Glu	Asn	Arg	Thr	Thr	Pro	Asp	Cys	
					380						385			
60	gaa	cga	gca	ttc	cca	gag	gag	agg	cca	gag	gtg	ctg	ggc	1360
	Glu	Arg	Ala	Phe	Pro	Glu	Glu	Arg	Pro	Glu	Val	Leu	Gly	
			390					395					400	
65	cag	cgg	agc	act	gat	gta	gtg	gac	ctg	gaa	aat	gag	gag	1399
	Gln	Arg	Ser	Thr	Asp	Val	Val	Asp	Leu	Glu	Asn	Glu	Glu	
				405						410				
70	cca	gac	agt	gac	aat	gag	tgg	cag	cac	ctg	cta	gag	acc	1438
	Pro	Asp	Ser	Asp	Asn	Glu	Trp	Gln	His	Leu	Leu	Glu	Thr	
							420					425		
75	act	gag	cct	gtg	cct	att	caa	ctg	aag	gct	cct	ctc	acc	1477
	Thr	Glu	Pro	Val	Pro	Ile	Gln	Leu	Lys	Ala	Pro	Leu	Thr	
				430					435					
80	ttg	ctg	tgt	aat	cct	gac	ttc	tgc	cag	cgc	atc	cag	agt	1516
	Leu	Leu	Cys	Asn	Pro	Asp	Phe	Cys	Gln	Arg	Ile	Gln	Ser	
					445						450			
85	cag	ctg	cat	gaa	gct	gga	ggg	cag	atc	ctg	aaa	ggc	atc	1555
	Gln	Leu	His	Glu	Ala	Gly	Gly	Gln	Ile	Leu	Lys	Gly	Ile	
				455				460					465	

ES 2 371 853 T3

	ttg	gag	ggg	gct	tcc	cac	atc	ctg	cct	gca	ttc	cgg	gtc	1594
	Leu	Glu	Gly	Ala	Ser	His	Ile	Leu	Pro	Ala	Phe	Arg	Val	
5					470					475				
	ctg	agc	agt	ctt	ctc	tcc	agc	tgc	agt	gat	tct	gtt	gcc	1633
	Leu	Ser	Ser	Leu	Leu	Ser	Ser	Cys	Ser	Asp	Ser	Val	Ala	
		480					485					490		
10	ttg	tat	tcc	ttc	tgc	cgg	gag	gca	ggg	ctt	cct	ggg	ctg	1672
	Leu	Tyr	Ser	Phe	Cys	Arg	Glu	Ala	Gly	Leu	Pro	Gly	Leu	
				495					500					
15	ctg	ctg	agt	cta	ctc	agg	cac	agt	cag	gag	agc	aac	agc	1711
	Leu	Leu	Ser	Leu	Leu	Arg	His	Ser	Gln	Glu	Ser	Asn	Ser	
	505					510					515			
20	ctc	cag	cag	caa	tct	tgg	tat	ggg	acc	ttc	tta	cag	gac	1750
	Leu	Gln	Gln	Gln	Ser	Trp	Tyr	Gly	Thr	Phe	Leu	Gln	Asp	
			520					525					530	
	ctg	atg	gct	gtg	att	cag	gcc	tac	ttt	gcc	tgt	acc	ttc	1789
	Leu	Met	Ala	Val	Ile	Gln	Ala	Tyr	Phe	Ala	Cys	Thr	Phe	
					535					540				
25	aat	ctg	gag	agg	agc	cag	aca	agt	gac	agc	ctg	cag	gtg	1828
	Asn	Leu	Glu	Arg	Ser	Gln	Thr	Ser	Asp	Ser	Leu	Gln	Val	
		545					550					555		
30	ttt	cag	gag	gct	gcc	aac	ctt	ttt	ctg	gac	ctg	ttg	ggg	1867
	Phe	Gln	Glu	Ala	Ala	Asn	Leu	Phe	Leu	Asp	Leu	Leu	Gly	
				560					565					
35	aaa	ctg	ctg	gcc	caa	cca	gat	gac	tct	gag	cag	act	ttg	1906
	Lys	Leu	Leu	Ala	Gln	Pro	Asp	Asp	Ser	Glu	Gln	Thr	Leu	
	570					575					580			
40	cgg	agg	gac	agc	ctt	atg	tgc	ttt	act	gtc	ctg	tgc	gaa	1945
	Arg	Arg	Asp	Ser	Leu	Met	Cys	Phe	Thr	Val	Leu	Cys	Glu	
			585					590					595	
	gcc	atg	gat	ggg	aac	agc	cgg	gcc	atc	tcc	aaa	gcc	ttt	1984
	Ala	Met	Asp	Gly	Asn	Ser	Arg	Ala	Ile	Ser	Lys	Ala	Phe	
					600					605				
45	tac	tcc	agc	ttg	ctg	acg	aca	cag	cag	gtt	gtc	ttg	gat	2023
	Tyr	Ser	Ser	Leu	Leu	Thr	Thr	Gln	Gln	Val	Val	Leu	Asp	
		610					615					620		
50	ggg	ctc	ctt	cat	ggc	ttg	aca	gtt	cca	cag	ctc	cct	gtc	2062
	Gly	Leu	Leu	His	Gly	Leu	Thr	Val	Pro	Gln	Leu	Pro	Val	
				625					630					
55	cac	act	ccc	caa	gga	gcc	ccg	caa	gtg	agc	cag	cca	ctg	2101
	His	Thr	Pro	Gln	Gly	Ala	Pro	Gln	Val	Ser	Gln	Pro	Leu	
	635					640					645			
60	cga	gag	cag	agt	gag	gat	ata	cct	gga	gcc	att	tcc	tct	2140
	Arg	Glu	Gln	Ser	Glu	Asp	Ile	Pro	Gly	Ala	Ile	Ser	Ser	
			650					655					660	
	gcc	ctg	gca	gcc	ata	tgc	act	gct	cct	gtg	gga	ctg	ccc	2179
	Ala	Leu	Ala	Ala	Ile	Cys	Thr	Ala	Pro	Val	Gly	Leu	Pro	
					665					670				
65	gac	tgc	tgg	gat	gcc	aag	gag	cag	gtc	tgt	tgg	cat	ttg	2218
	Asp	Cys	Trp	Asp	Ala	Lys	Glu	Gln	Val	Cys	Trp	His	Leu	
		675					680					685		
70	gca	aat	cag	cta	act	gaa	gac	agc	agc	cag	ctc	agg	cca	2257

ES 2 371 853 T3

	Ala	Asn	Gln	Leu	Thr	Glu	Asp	Ser	Ser	Gln	Leu	Arg	Pro	
				690					695					
5	tcc	ctc	atc	tct	ggc	ctg	cag	cat	ccc	atc	ctg	tgc	ctg	2296
	Ser	Leu	Ile	Ser	Gly	Leu	Gln	His	Pro	Ile	Leu	Cys	Leu	
	700				705					710				
10	cac	ctt	ctc	aag	gtt	cta	tac	tcc	tgc	tgc	ctt	gtc	agt	2335
	His	Leu	Leu	Lys	Val	Leu	Tyr	Ser	Cys	Cys	Leu	Val	Ser	
			715					720					725	
15	gag	ggc	ctg	tgc	cgt	ctt	ctg	ggg	cag	gag	ccc	ctg	gcc	2374
	Glu	Gly	Leu	Cys	Arg	Leu	Leu	Gly	Gln	Glu	Pro	Leu	Ala	
					730					735				
20	ttg	gaa	tcc	ctg	ttt	atg	ttg	att	cag	ggc	aag	gta	aaa	2413
	Leu	Glu	Ser	Leu	Phe	Met	Leu	Ile	Gln	Gly	Lys	Val	Lys	
		740					745					750		
25	gta	gta	gat	tgg	gaa	gag	tct	act	gaa	gtg	aca	ctc	tac	2452
	Val	Val	Asp	Trp	Glu	Glu	Ser	Thr	Glu	Val	Thr	Leu	Tyr	
				755					760					
30	ttc	ctc	tcc	ctt	ctt	gtc	ttt	cgg	ctc	caa	aac	ctg	cct	2491
	Phe	Leu	Ser	Leu	Leu	Val	Phe	Arg	Leu	Gln	Asn	Leu	Pro	
						770					775			
35	tgt	gga	atg	gag	aag	cta	ggc	agt	gac	gtt	gct	act	ctc	2530
	Cys	Gly	Met	Glu	Lys	Leu	Gly	Ser	Asp	Val	Ala	Thr	Leu	
			780					785					790	
40	ttt	acc	cat	tcg	cat	gtc	gtc	tct	ctt	gtg	agt	gca	gca	2569
	Phe	Thr	His	Ser	His	Val	Val	Ser	Leu	Val	Ser	Ala	Ala	
					795					800				
45	gcc	tgt	cta	ttg	gga	cag	ctt	ggt	cag	caa	ggg	gtg	acc	2608
	Ala	Cys	Leu	Leu	Gly	Gln	Gly	Gln	Gln	Gln	Gly	Val	Thr	
		805					810					815		
50	ttt	gac	ctc	cag	ccc	atg	gaa	tgg	atg	gct	gca	gcc	aca	2647
	Phe	Asp	Leu	Gln	Pro	Met	Glu	Trp	Met	Ala	Ala	Ala	Thr	
				820					825					
55	cat	gcc	ttg	tct	gcc	cct	gca	gag	gtt	cgg	ttg	act	cca	2686
	His	Ala	Leu	Ser	Ala	Pro	Ala	Glu	Val	Arg	Leu	Thr	Pro	
						835					840			
60	cca	ggt	agt	tgt	gga	ttc	tat	gat	ggc	ctc	ctt	atc	ctt	2725
	Pro	Gly	Ser	Cys	Gly	Phe	Tyr	Asp	Gly	Leu	Leu	Ile	Leu	
			845					850					855	
65	ctg	ttg	cag	ctc	ctc	act	gag	cag	ggg	aag	gct	agc	cta	2764
	Leu	Leu	Gln	Leu	Leu	Thr	Glu	Gln	Gly	Lys	Ala	Ser	Leu	
					860					865				
70	atc	agg	gat	atg	tcc	agt	tca	gaa	atg	tgg	acc	gtt	ttg	2803
	Ile	Arg	Asp	Met	Ser	Ser	Ser	Glu	Met	Trp	Thr	Val	Leu	
						875						880		
75	tgg	cac	cgc	ttc	tcc	atg	gtc	ctg	agg	ctc	ccc	gag	gag	2842
	Trp	His	Arg	Phe	Ser	Met	Val	Leu	Arg	Leu	Pro	Glu	Glu	
				885					890					
80	gca	tct	gca	cag	gaa	ggg	gag	ctt	tcg	cta	tcc	agt	cca	2881
	Ala	Ser	Ala	Gln	Glu	Gly	Glu	Leu	Ser	Leu	Ser	Ser	Pro	
						900					905			
85	cca	agc	cct	gag	cca	gac	tgg	aca	ctg	att	tct	ccc	cag	2920
	Pro	Ser	Pro	Glu	Pro	Asp	Trp	Thr	Leu	Ile	Ser	Pro	Gln	
				910					915				920	

ES 2 371 853 T3

5	ggc Gly	atg Met	gca Ala	gcc Ala	ctg Leu 925	ctg Leu	agc Ser	ctg Leu	gcc Ala	atg Met 930	gcc Ala	acc Thr	ttt Phe	2959
	acc Thr	cag Gln 935	gag Glu	ccc Pro	cag Gln	tta Leu	tgc Cys 940	ctg Leu	agc Ser	tgc Cys	ctg Leu	tcc Ser 945	cag Gln	2998
10	cat His	gga Gly	agt Ser	atc Ile 950	ctc Leu	atg Met	tcc Ser	atc Ile	ctg Leu 955	aag Lys	cat His	ctg Leu	ctt Leu	3037
15	tgc Cys 960	ccc Pro	agc Ser	ttc Phe	ctg Leu	aat Asn 965	caa Gln	ctg Leu	cgc Arg	cag Gln	gcg Ala 970	cct Pro	cat His	3076
20	ggg Gly	tct Ser	gag Glu 975	ttt Phe	ctc Leu	cct Pro	gtc Val	gtg Val 980	gtg Val	ctc Leu	tct Ser	gtc Val	tgc Cys 985	3115
25	cag Gln	ctc Leu	ctt Leu	tgc Cys	ttc Phe 990	ccc Pro	ttt Phe	gcg Ala	ctg Leu	gac Asp 995	atg Met	gat Asp	gct Ala	3154
	gac Asp 1000	ctc Leu	ctt Leu	ata Ile	gtt Val	gtc Val	ttg Leu 1005	gcc Ala	gac Asp	ctc Leu	agg Arg 1010	gac Asp	tca Ser	3193
30	gaa Glu	gtt Val	gca Ala	gcc Ala 1015	cat His	ctg Leu	ctg Leu	cag Gln	gtc Val 1020	tgc Cys	tgc Cys	tac Tyr	cat His	3232
35	ctt Leu 1025	ccg Pro	ttg Leu	atg Met	caa Gln	gtg Val 1030	gag Glu	ctg Leu	ccc Pro	atc Ile 1035	agc Ser	ctt Leu	ctc Leu	3271
40	aca Thr	cgc Arg	ctg Leu 1040	gcc Ala	ctc Leu	atg Met	gat Asp	ccc Pro 1045	acc Thr	tct Ser	ctc Leu	aac Asn 1050	cag Gln	3310
45	ttt Phe	gtg Val	aac Asn	aca Thr 1055	gtg Val	tct Ser	gcc Ala	tcc Ser	cct Pro 1060	aga Arg	acc Thr	atc Ile	gtc Val	3349
	tcg Ser 1065	ttt Phe	ctc Leu	tca Ser	gtt Val	gcc Ala 1070	ctc Leu	ctg Leu	agt Ser	gac Asp	cag Gln 1075	cca Pro	ctg Leu	3388
50	ttg Leu	acc Thr	tcc Ser	gac Asp 1080	ctt Leu	ctc Leu	tct Ser	ctg Leu	ctg Leu 1085	gcc Ala	cat His	act Thr	gcc Ala	3427
55	agg Arg 1090	gtc Val	ctg Leu	tct Ser	ccc Pro	agc Ser 1095	cac His	ttg Leu	tcc Ser	ttt Phe 1100	atc Ile	caa Gln	gag Glu	3466
60	ctt Leu	ctg Leu	gct Ala 1105	ggc Gly	tct Ser	gat Asp	gaa Glu	tcc Ser 1110	tat Tyr	cgg Arg	ccc Pro	ctg Leu	cgc Arg 1115	3505
65	agc Ser	ctc Leu	ctg Leu	ggc Gly 1120	cac His	cca Pro	gag Glu	aat Asn	tct Ser 1125	gtg Val	cgg Arg	gca Ala	cac His	3544
	act Thr 1130	tat Tyr	agg Arg	ctc Leu	ctg Leu	gga Gly 1135	cac His	ttg Leu	ctc Leu	caa Gln	cac His 1140	agc Ser	atg Met	3583
70	gcc Gly	ctg Leu	cgt Leu	ggg Gly	gca Ala	ctg Leu	cag Glu	agc Ser	cag Glu	tct Ser	gga Glu	ctg Leu	ctc Leu	3622

ES 2 371 853 T3

Ala Leu Arg Gly Ala Leu Gln Ser Gln Ser Gly Leu Leu
1145 1150

5 agc ctt ctg ctg ctt ggg ctt gga gac aag gat cct gtt 3661
Ser Leu Leu Leu Leu Gly Leu Gly Asp Lys Asp Pro Val
1155 1160 1165

10 gtg cgg tgc agt gcc agc ttt gct gtg ggc aat gca gcc 3700
Val Arg Cys Ser Ala Ser Phe Ala Val Gly Asn Ala Ala
1170 1175 1180

15 tac cag gct ggt cct ctg gga cct gcc ctg gca gct gca 3739
Tyr Gln Ala Gly Pro Leu Gly Pro Ala Leu Ala Ala
1185 1190

20 gtg ccc agt atg acc cag ctg ctt gga gat cct cag gct 3778
Val Pro Ser Met Thr Gln Leu Leu Gly Asp Pro Gln Ala
1195 1200 1205

25 ggt atc cgg cgc aat gtt gca tca gct ctg ggc aac ttg 3817
Gly Ile Arg Arg Asn Val Ala Ser Ala Leu Gly Asn Leu
1210 1215

30 gga cct gaa ggt ttg gga gag gag ctg tta cag tgc gaa 3856
Gly Pro Glu Gly Leu Gly Glu Glu Leu Leu Gln Cys Glu
1220 1225 1230

35 gta ccc cag cgg ctc cta gaa atg gca tgt gga gac ccc 3895
Val Pro Gln Arg Leu Leu Glu Met Ala Cys Gly Asp Pro
1235 1240 1245

40 cag cca aat gtg aag gag gct gcc ctc att gcc ctc cgg 3934
Gln Pro Asn Val Lys Glu Ala Ala Leu Ile Ala Leu Arg
1250 1255

45 agc ctg caa cag gag cct ggc atc cat cag gta ctg gtg 3973
Ser Leu Gln Gln Glu Pro Gly Ile His Gln Val Leu Val
1260 1265 1270

50 tcc ctg ggt gcc agt gag aaa cta tcc ttg ctc tct ctg 4012
Ser Leu Gly Ala Ser Glu Lys Leu Ser Leu Leu Ser Leu
1275 1280

55 ggg aat cag tca ctg cca cac agc agt cct agg cct gcc 4051
Gly Asn Gln Ser Leu Pro His Ser Ser Pro Arg Pro Ala
1285 1290 1295

60 tct gcc aaa cac tgc agg aaa ctc att cac ctc ctg agg 4090
Ser Ala Lys His Cys Arg Lys Leu Ile His Leu Leu Arg
1300 1305 1310

65 cca gcc cat agc atg tgatt ccagattcct gcggtccagc 4130
Pro Ala His Ser Met
1315

70 ctccaacttt ggtgccagct ctttcttatt taatacacaa gcgccaaytc 4180
aactgagagc taaagagact agaaaagaga taagctgcca actcaactga 4230
gaacaggaaa ctngaagaga tttatatata aagcttcttc cttctcccag 4280
atgcaggatg ttttcaacca gtaaatttta ttgctgttg tgccagagaa 4330
gagtcctttt cttctctaca tccaggggcc nttttctcca ataattgtgcc 4380
tttaactcta gggacctgcc tcacggacct tagggaaaaa cctcaacctg 4430
aaagatctct tcctttcttg agctccttta atcttcccag cagggtttttg 4480
ccttagacgt gctggcccca ggacagtgat gaagacagag cctgtctcag 4530

ES 2 371 853 T3

ctctaggctg tggggatcaa tgccatcagt ccctgttatt gagggattat 4580
 5 cccttagcca acattcctat ctgtgggtgg gcgtggagag tgtatctttt 4630
 tttggggtgt gtgtgtatat gtgtgtgtgt atgtgtgtgt gtgtttaata 4680
 gttctgtttg taaactcttt taataaaagt tgtgcctcac cataactgaa 4730
 10 gctcccagga caaggggtga gaggtcaac ccctctttca gcttctatgt 4780
 ggtgttgag gtgtggtat cgtgttcaca caaaaaaaaa aaaaaaaaaa 4830
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4880
 15 <210> 2
 <211> 1315
 <212> PRT
 <213> Homo sapiens
 20 <400> 2
 Met Glu Lys Tyr His Val Leu Glu Met Ile Gly Glu Gly Ser Phe
 1 5 10 15
 25 Gly Arg Val Tyr Lys Gly Arg Arg Lys Tyr Ser Ala Gln Val Val
 20 25 30
 Ala Leu Lys Phe Ile Pro Lys Leu Gly Arg Ser Glu Lys Glu Leu
 35 35 40 45
 Arg Asn Leu Gln Arg Glu Ile Glu Ile Met Arg Gly Leu Arg His
 50 55 60
 Pro Asn Ile Val His Met Leu Asp Ser Phe Glu Thr Asp Lys Glu
 35 65 70 75
 Val Val Val Val Thr Asp Tyr Ala Glu Gly Glu Leu Phe Gln Ile
 80 85 90
 40 Leu Glu Asp Asp Gly Lys Leu Pro Glu Asp Gln Val Gln Ala Ile
 95 100 105
 Ala Ala Gln Leu Val Ser Ala Leu Tyr Tyr Leu His Ser His Arg
 110 115 120
 45 Ile Leu His Arg Asp Met Lys Pro Gln Asn Ile Leu Leu Ala Lys
 125 130 135
 Gly Gly Gly Ile Lys Leu Cys Asp Phe Gly Phe Ala Arg Ala Met
 50 140 145 150
 Ser Thr Asn Thr Met Val Leu Thr Ser Ile Lys Gly Thr Pro Leu
 155 160 165
 55 Tyr Met Ser Pro Glu Leu Val Glu Glu Arg Pro Tyr Asp His Thr
 170 175 180
 Ala Asp Leu Trp Ser Val Gly Cys Ile Leu Tyr Glu Leu Ala Val
 185 190 195
 60 Gly Thr Pro Pro Phe Tyr Ala Thr Ser Ile Phe Gln Leu Val Ser
 200 205 210
 Leu Ile Leu Lys Asp Pro Val Arg Trp Pro Ser Thr Ile Ser Pro
 215 220 225
 65 Cys Phe Lys Asn Phe Leu Gln Gly Leu Leu Thr Lys Asp Pro Arg
 230 235 240

ES 2 371 853 T3

	Gln Arg Leu Ser	Trp 245	Pro Asp Leu Leu	Tyr 250	His Pro Phe Ile	Ala 255
5	Gly His Val Thr	Ile 260	Ile Thr Glu Pro	Ala 265	Gly Pro Asp Leu	Gly 270
	Thr Pro Phe Thr	Ser 275	Arg Leu Pro Pro	Glu 280	Leu Gln Val Leu	Lys 285
10	Asp Glu Gln Ala	His 290	Arg Leu Ala Pro	Lys 295	Gly Asn Gln Ser	Arg 300
	Ile Leu Thr Gln	Ala 305	Tyr Lys Arg Met	Ala 310	Glu Glu Ala Met	Gln 315
15	Lys Lys His Gln	Asn 320	Thr Gly Pro Ala	Leu 325	Glu Gln Glu Asp	Lys 330
	Thr Ser Lys Val	Ala 335	Pro Gly Thr Ala	Pro 340	Leu Pro Arg Leu	Gly 345
20	Ala Thr Pro Gln	Glu 350	Ser Ser Leu Leu	Ala 355	Gly Ile Leu Ala	Ser 360
25	Glu Leu Lys Ser	Ser 365	Trp Ala Lys Ser	Gly 370	Thr Gly Glu Val	Pro 375
	Ser Ala Pro Arg	Glu 380	Asn Arg Thr Thr	Pro 385	Asp Cys Glu Arg	Ala 390
30	Phe Pro Glu Glu	Arg 395	Pro Glu Val Leu	Gly 400	Gln Arg Ser Thr	Asp 405
	Val Val Asp Leu	Glu 410	Asn Glu Glu Pro	Asp 415	Ser Asp Asn Glu	Trp 420
35	Gln His Leu Leu	Glu 425	Thr Thr Glu Pro	Val 430	Pro Ile Gln Leu	Lys 435
40	Ala Pro Leu Thr	Leu 440	Leu Cys Asn Pro	Asp 445	Phe Cys Gln Arg	Ile 450
	Gln Ser Gln Leu	His 455	Glu Ala Gly Gly	Gln 460	Ile Leu Lys Gly	Ile 465
45	Leu Glu Gly Ala	Ser 470	His Ile Leu Pro	Ala 475	Phe Arg Val Leu	Ser 480
50	Ser Leu Leu Ser	Ser 485	Cys Ser Asp Ser	Val 490	Ala Leu Tyr Ser	Phe 495
	Cys Arg Glu Ala	Gly 500	Leu Pro Gly Leu	Leu 505	Leu Ser Leu Leu	Arg 510
55	His Ser Gln Glu	Ser 515	Asn Ser Leu Gln	Gln 520	Gln Ser Trp Tyr	Gly 525
	Thr Phe Leu Gln	Asp 530	Leu Met Ala Val	Ile 535	Gln Ala Tyr Phe	Ala 540
60	Cys Thr Phe Asn	Leu 545	Glu Arg Ser Gln	Thr 550	Ser Asp Ser Leu	Gln 555
65	Val Phe Gln Glu	Ala 560	Ala Asn Leu Phe	Leu 565	Asp Leu Leu Gly	Lys 570
	Leu Leu Ala Gln	Pro 575	Asp Asp Ser Glu	Gln 580	Thr Leu Arg Arg	Asp 585
70	Ser Leu Met Cys	Phe Thr Val Leu	Cys Glu Ala Met	Asp Gly Asn		

ES 2 371 853 T3

		590		595		600
	Ser Arg Ala Ile	Ser 605	Lys Ala Phe Tyr	Ser 610	Ser Leu Leu Thr	Thr 615
5	Gln Gln Val Val	Leu 620	Asp Gly Leu Leu	His 625	Gly Leu Thr Val	Pro 630
10	Gln Leu Pro Val	His 635	Thr Pro Gln Gly	Ala 640	Pro Gln Val Ser	Gln 645
	Pro Leu Arg Glu	Gln 650	Ser Glu Asp Ile	Pro 655	Gly Ala Ile Ser	Ser 660
15	Ala Leu Ala Ala	Ile 665	Cys Thr Ala Pro	Val 670	Gly Leu Pro Asp	Cys 675
20	Trp Asp Ala Lys	Glu 680	Gln Val Cys Trp	His 685	Leu Ala Asn Gln	Leu 690
	Thr Glu Asp Ser	Ser 695	Gln Leu Arg Pro	Ser 700	Leu Ile Ser Gly	Leu 705
25	Gln His Pro Ile	Leu 710	Cys Leu His Leu	Leu 715	Lys Val Leu Tyr	Ser 720
	Cys Cys Leu Val	Ser 725	Glu Gly Leu Cys	Arg 730	Leu Leu Gly Gln	Glu 735
30	Pro Leu Ala Leu	Glu 740	Ser Leu Phe Met	Leu 745	Ile Gln Gly Lys	Val 750
35	Lys Val Val Asp	Trp 755	Glu Glu Ser Thr	Glu 760	Val Thr Leu Tyr	Phe 765
	Leu Ser Leu Leu	Val 770	Phe Arg Leu Gln	Asn 775	Leu Pro Cys Gly	Met 780
40	Glu Lys Leu Gly	Ser 785	Asp Val Ala Thr	Leu 790	Phe Thr His Ser	His 795
	Val Val Ser Leu	Val 800	Ser Ala Ala Ala	Cys 805	Leu Leu Gly Gln	Leu 810
45	Gly Gln Gln Gly	Val 815	Thr Phe Asp Leu	Gln 820	Pro Met Glu Trp	Met 825
50	Ala Ala Ala Thr	His 830	Ala Leu Ser Ala	Pro 835	Ala Glu Val Arg	Leu 840
	Thr Pro Pro Gly	Ser 845	Cys Gly Phe Tyr	Asp 850	Gly Leu Leu Ile	Leu 855
55	Leu Leu Gln Leu	Leu 860	Thr Glu Gln Gly	Lys 865	Ala Ser Leu Ile	Arg 870
	Asp Met Ser Ser	Ser 875	Glu Met Trp Thr	Val 880	Leu Trp His Arg	Phe 885
60	Ser Met Val Leu	Arg 890	Leu Pro Glu Glu	Ala 895	Ser Ala Gln Glu	Gly 900
65	Glu Leu Ser Leu	Ser 905	Ser Pro Pro Ser	Pro 910	Glu Pro Asp Trp	Thr 915
	Leu Ile Ser Pro	Gln 920	Gly Met Ala Ala	Leu 925	Leu Ser Leu Ala	Met 930
70	Ala Thr Phe Thr	Gln 935	Glu Pro Gln Leu	Cys 940	Leu Ser Cys Leu	Ser 945

ES 2 371 853 T3

	Gln His Gly Ser Ile Leu Met Ser Ile Leu Lys His Leu Leu Cys	950	955	960
5	Pro Ser Phe Leu Asn Gln Leu Arg Gln Ala Pro His Gly Ser Glu	965	970	975
	Phe Leu Pro Val Val Val Leu Ser Val Cys Gln Leu Leu Cys Phe	980	985	990
10	Pro Phe Ala Leu Asp Met Asp Ala Asp Leu Leu Ile Val Val Leu	995	1000	1005
	Ala Asp Leu Arg Asp Ser Glu Val Ala Ala His Leu Leu Gln Val	1010	1015	1020
15	Cys Cys Tyr His Leu Pro Leu Met Gln Val Glu Leu Pro Ile Ser	1025	1030	1035
20	Leu Leu Thr Arg Leu Ala Leu Met Asp Pro Thr Ser Leu Asn Gln	1040	1045	1050
	Phe Val Asn Thr Val Ser Ala Ser Pro Arg Thr Ile Val Ser Phe	1055	1060	1065
25	Leu Ser Val Ala Leu Leu Ser Asp Gln Pro Leu Leu Thr Ser Asp	1070	1075	1080
	Leu Leu Ser Leu Leu Ala His Thr Ala Arg Val Leu Ser Pro Ser	1085	1090	1095
30	His Leu Ser Phe Ile Gln Glu Leu Leu Ala Gly Ser Asp Glu Ser	1100	1105	1110
35	Tyr Arg Pro Leu Arg Ser Leu Leu Gly His Pro Glu Asn Ser Val	1115	1120	1125
	Arg Ala His Thr Tyr Arg Leu Leu Gly His Leu Leu Gln His Ser	1130	1135	1140
40	Met Ala Leu Arg Gly Ala Leu Gln Ser Gln Ser Gly Leu Leu Ser	1145	1150	1155
45	Leu Leu Leu Leu Gly Leu Gly Asp Lys Asp Pro Val Val Arg Cys	1160	1165	1170
	Ser Ala Ser Phe Ala Val Gly Asn Ala Ala Tyr Gln Ala Gly Pro	1175	1180	1185
50	Leu Gly Pro Ala Leu Ala Ala Ala Val Pro Ser Met Thr Gln Leu	1190	1195	1200
	Leu Gly Asp Pro Gln Ala Gly Ile Arg Arg Asn Val Ala Ser Ala	1205	1210	1215
55	Leu Gly Asn Leu Gly Pro Glu Gly Leu Gly Glu Glu Leu Leu Gln	1220	1225	1230
60	Cys Glu Val Pro Gln Arg Leu Leu Glu Met Ala Cys Gly Asp Pro	1235	1240	1245
	Gln Pro Asn Val Lys Glu Ala Ala Leu Ile Ala Leu Arg Ser Leu	1250	1255	1260
65	Gln Gln Glu Pro Gly Ile His Gln Val Leu Val Ser Leu Gly Ala	1265	1270	1275
	Ser Glu Lys Leu Ser Leu Leu Ser Leu Gly Asn Gln Ser Leu Pro	1280	1285	1290
70				

ES 2 371 853 T3

His Ser Ser Pro Arg Pro Ala Ser Ala Lys His Cys Arg Lys Leu
 1295 1300 1305
 Ile His Leu Leu Arg Pro Ala His Ser Met
 5 1310 1315
 <210> 3
 <211> 193
 <212> DNA
 10 <213> Homo sapiens
 <400> 3
 cccgggctat gagcaccaat acaatgggtgc tgacatccat caaaggcaca 50
 15 ccactctata tgtctccaga gctgggtggag gagcgaccat acgaccacac 100
 agcggacctc tggctctgttg gctgcatact atatgaactg gcagtaggca 150
 cccctccctt ctaatgctac aagcatcttt cagctgggtca gcc 193
 20
 <210> 4
 <211> 5125
 <212> DNA
 <213> Homo sapiens
 25
 <400> 4
 cccacgcgtc cgcccacgcg tccggggcgt cccagatggt gtggaactgt 50
 ccctggatct atagctcttc accgtctcta ctttcttcct tctaagagat 100
 30 cctgaaacct ctgtc atg gaa aag tac cac gtg ttg gag 139
 Met Glu Lys Tyr His Val Leu Glu
 1 5
 35 atg att gga gaa ggc tct ttt ggg agg gtg tac aag ggt 178
 Met Ile Gly Glu Gly Ser Phe Gly Arg Val Tyr Lys Gly
 10 15 20
 40 cga aga aaa tac agt gct cag gtc gtg gcc ctg aag ttc 217
 Arg Arg Lys Tyr Ser Ala Gln Val Val Ala Leu Lys Phe
 25 30
 45 atc cca aaa ttg ggg cgc tca gag aag gag ctg agg aat 256
 Ile Pro Lys Leu Gly Arg Ser Glu Lys Glu Leu Arg Asn
 35 40 45
 50 ttg caa cga gag att gaa ata atg cgg ggt ctg cgg cat 295
 Leu Gln Arg Glu Ile Glu Ile Met Arg Gly Leu Arg His
 50 55 60
 55 ccc aac att gtg cat atg ctt gac agc ttt gaa act gat 334
 Pro Asn Ile Val His Met Leu Asp Ser Phe Glu Thr Asp
 65 70
 60 aaa gag gtg gtg gtg gtg aca gac tat gct gag gga gag 373
 Lys Glu Val Val Val Val Thr Asp Tyr Ala Glu Gly Glu
 75 80 85
 65 ctc ttt cag atc cta gaa gat gac gga aaa ctt cct gaa 412
 Leu Phe Gln Ile Leu Glu Asp Asp Gly Lys Leu Pro Glu
 90 95
 70 gac cag gtt cag gcc att gct gcc cag ttg gtg tca gcc 451
 Asp Gln Val Gln Ala Ile Ala Ala Gln Leu Val Ser Ala
 100 105 110
 75 ctg tac tat ctg cat tcc cac cgc atc cta cac cga gat 490
 Leu Tyr Tyr Leu His Ser His Arg Ile Leu His Arg Asp

ES 2 371 853 T3

	115				120				125				
5	atg Met	aag Lys	cct Pro	cag Gln 130	aac Asn	atc Ile	ctc Leu	ctc Leu	gcc Ala 135	aag Lys 135	ggt Gly	ggt Gly	ggc Gly 529
10	atc Ile	aag Lys 140	ctc Leu	tgt Cys	gac Asp	ttt Phe	gga Gly 145	ttt Phe	gcc Ala	cgg Arg	gct Ala 150	atg Met	agc Ser 568
15	acc Thr	aat Asn	aca Thr	atg Met 155	gtg Val	ctg Leu	aca Thr	tcc Ser	atc Ile 160	aaa Lys	ggc Gly	aca Thr	cca Pro 607
20	ctc Leu 165	tat Tyr	atg Met	tct Ser	cca Pro	gag Glu 170	ctg Leu	gtg Val	gag Glu	gag Glu 175	cga Arg 175	cca Pro	tac Tyr 646
25	gac Asp	cac His	aca Thr 180	gcg Ala	gac Asp	ctc Leu	tgg Trp	tct Ser 185	gtt Val	ggc Gly	tgc Cys	ata Ile	cta Leu 190 685
30	tat Tyr	gaa Glu	ctg Leu	gca Ala 195	gta Val	ggc Gly	acc Thr	cct Pro	ccc Pro	ttc Phe 200	tat Tyr	gct Ala	aca Thr 724
35	agc Ser	atc Ile 205	ttt Phe	cag Gln	ctg Leu	gtc Val 210	agc Ser 210	ctc Leu	att Ile	ctc Leu	aag Lys	gac Asp 215	cct Pro 763
40	gtg Val	cgc Arg	tgg Trp	ccc Pro 220	tca Ser	acc Thr	atc Ile	agt Ser	ccc Pro 225	tgc Cys	ttt Phe	aag Lys	aac Asn 802
45	ttc Phe 230	ctg Leu	cag Gln	gga Gly	ctg Leu	ctc Leu 235	acc Thr	aaa Lys	gac Asp	cca Pro	cgg Arg 240	cag Gln	cga Arg 841
50	ctg Leu	tcc Ser	tgg Trp 245	cca Pro	gac Asp	ctc Leu	tta Leu	tat Tyr 250	cac His	ccc Pro	ttt Phe	att Ile	gct Ala 255 880
55	ggt Gly	cat His	gtc Val	acc Thr 260	ata Ile	ata Ile	act Thr	gag Glu	cca Pro	gca Ala 265	ggc Gly	cca Pro	gat Asp 919
60	ttg Leu	ggg Gly 270	acc Thr	cca Pro	ttc Phe	acc Thr	agc Ser 275	cgc Arg	cta Leu	ccc Pro	cca Pro	gaa Glu 280	ctt Leu 958
65	cag Gln	gtc Val	cta Leu	aag Lys 285	gac Asp	gaa Glu	cag Gln	gcc Ala	cat His 290	cgg Arg	ttg Leu	gcc Ala	ccc Pro 997
70	aag Lys 295	ggt Gly	aat Asn	cag Gln	tct Ser	cgc Arg 300	atc Ile	ttg Leu	act Thr	cag Gln	gcc Ala 305	tat Tyr	aaa Lys 1036
75	cgc Arg	atg Met	gct Ala 310	gag Glu	gag Glu	gcc Ala	atg Met	cag Gln 315	aag Lys	aaa Lys	cat His	cag Gln	aac Asn 1075
80	aca Thr	gga Gly	cct Pro	gcc Ala 325	ctt Leu	gag Glu	caa Gln	gag Glu	gac Asp	aag Lys 330	acc Thr	agc Ser	aag Lys 1114
85	gtg Val	gct Ala 335	cct Pro	ggc Gly	aca Thr	gcc Ala	cct Pro 340	ctg Leu	ccc Pro	aga Arg	ctc Leu	ggg Gly 345	gcc Ala 1153

ES 2 371 853 T3

	act	cct	cag	gaa	tca	agc	ctc	ctg	gcc	ggg	atc	tta	gcc	1192
	Thr	Pro	Gln	Glu	Ser	Ser	Leu	Leu	Ala	Gly	Ile	Leu	Ala	
				350					355					
5	tca	gaa	ttg	aag	agc	agc	tgg	gct	aaa	tca	ggg	act	gga	1231
	Ser	Glu	Leu	Lys	Ser	Ser	Trp	Ala	Lys	Ser	Gly	Thr	Gly	
	360				365						370			
10	gag	gtg	ccc	tct	gca	cct	cgg	gaa	aac	cgg	acc	acc	cca	1270
	Glu	Val	Pro	Ser	Ala	Pro	Arg	Glu	Asn	Arg	Thr	Thr	Pro	
			375					380					385	
15	gat	tgt	gaa	cga	gca	ttc	cca	gag	gag	agg	cca	gag	gtg	1309
	Asp	Cys	Glu	Arg	Ala	Phe	Pro	Glu	Glu	Arg	Pro	Glu	Val	
					390					395				
20	ctg	ggc	cag	cgg	agc	act	gat	gta	gtg	gac	ctg	gaa	aat	1348
	Leu	Gly	Gln	Arg	Ser	Thr	Asp	Val	Val	Asp	Leu	Glu	Asn	
		400					405					410		
25	gag	gag	cca	gac	agt	gac	aat	gag	tgg	cag	cac	ctg	cta	1387
	Glu	Glu	Pro	Asp	Ser	Asp	Asn	Glu	Trp	Gln	His	Leu	Leu	
				415					420					
30	gag	acc	act	gag	cct	gtg	cct	att	caa	ctg	aag	gct	cct	1426
	Glu	Thr	Thr	Glu	Pro	Val	Pro	Ile	Gln	Leu	Lys	Ala	Pro	
	425					430					435			
35	ctc	acc	ttg	ctg	tgt	aat	cct	gac	ttc	tgc	cag	cgc	atc	1465
	Leu	Thr	Leu	Leu	Cys	Asn	Pro	Asp	Phe	Cys	Gln	Arg	Ile	
			440					445					450	
40	cag	agt	cag	ctg	cat	gaa	gct	gga	ggg	cag	atc	ctg	aaa	1504
	Gln	Ser	Gln	Leu	His	Glu	Ala	Gly	Gly	Gln	Ile	Leu	Lys	
					455					460				
45	ggc	atc	ttg	gag	ggg	gct	tcc	cac	atc	ctg	cct	gca	ttc	1543
	Gly	Ile	Leu	Glu	Gly	Ala	Ser	His	Ile	Leu	Pro	Ala	Phe	
		465					470					475		
50	cgg	gtc	ctg	agc	agt	ctt	ctc	tcc	agc	tgc	agt	gat	tct	1582
	Arg	Val	Leu	Ser	Ser	Leu	Leu	Ser	Ser	Cys	Ser	Asp	Ser	
				480					485					
55	gtt	gcc	ttg	tat	tcc	ttc	tgc	cgg	gag	gca	ggg	ctt	cct	1621
	Val	Ala	Leu	Tyr	Ser	Phe	Cys	Arg	Glu	Ala	Gly	Leu	Pro	
	490					495					500			
60	ggg	ctg	ctg	ctg	agt	cta	ctc	agg	cac	agt	cag	gag	agc	1660
	Gly	Leu	Leu	Leu	Ser	Leu	Leu	Arg	His	Ser	Gln	Glu	Ser	
			505					510					515	
65	aac	agc	ctc	cag	cag	caa	tct	tgg	tat	ggg	acc	ttc	tta	1699
	Asn	Ser	Leu	Gln	Gln	Gln	Ser	Trp	Tyr	Gly	Thr	Phe	Leu	
					520					525				
70	cag	gac	ctg	atg	gct	gtg	att	cag	gcc	tac	ttt	gcc	tgt	1738
	Gln	Asp	Leu	Met	Ala	Val	Ile	Gln	Ala	Tyr	Phe	Ala	Cys	
		530					535					540		
75	acc	ttc	aat	ctg	gag	agg	agc	cag	aca	agt	gac	agc	ctg	1777
	Thr	Phe	Asn	Leu	Glu	Arg	Ser	Gln	Thr	Ser	Asp	Ser	Leu	
				545					550					
80	cag	gtg	ttt	cag	gag	gct	gcc	aac	ctt	ttt	ctg	gac	ctg	1816
	Gln	Val	Phe	Gln	Glu	Ala	Ala	Asn	Leu	Phe	Leu	Asp	Leu	
	555					560					565			
85	ttg	ggg	aaa	ctg	ctg	gcc	caa	cca	gat	gac	tct	gag	cag	1855
	Leu	Gly	Lys	Leu	Leu	Ala	Gln	Pro	Asp	Asp	Ser	Glu	Gln	

ES 2 371 853 T3

	570	575	580	
5	act ttg cag agg gac agc ctt atg tgc ttt act gtc ctg 1894			
	Thr Leu Gln Arg Asp Ser Leu Met Cys Phe Thr Val Leu			
		585	590	
10	tgc gaa gcc atg gat ggg aac agc cgg gcc atc tcc aaa 1933			
	Cys Glu Ala Met Asp Gly Asn Ser Arg Ala Ile Ser Lys			
		595	600	605
15	gcc ttt tac tcc agc ttg ctg acg aca cag cag gtt gtc 1972			
	Ala Phe Tyr Ser Ser Leu Leu Thr Thr Gln Gln Val Val			
		610	615	
20	ttg gat ggg ctc ctt cat ggc ttg aca gtt cca cag ctc 2011			
	Leu Asp Gly Leu Leu His Gly Leu Thr Val Pro Gln Leu			
		620	625	630
25	cct gtc cac act ccc caa ggt aac cag agt gga gaa ggg 2050			
	Pro Val His Thr Pro Gln Gly Asn Gln Ser Gly Glu Gly			
		635	640	645
30	agg ttc tct t gacttacttg ttgcataggt caggctccgc 2090			
	Arg Phe Ser			
		648		
35	tctttctatt gccatcacct agatcgcacc tggcatttag taggtgctca 2140			
40	ataaataact gtgaactgag agaatgaatg gggatctgag ggaacaaac 2190			
	agacctcatc ctgcattctt cccactccct taggttccct actcctgctg 2240			
	ccatgtcggt gagtactggt gctattgtct agggcaagag cctcaggcct 2290			
45	ttgg agt tac tct ttg ctt ttc tcc aca gga gcc ccg 2327			
	Ser Tyr Ser Leu Leu Phe Ser Thr Gly Ala Pro			
		1	5	10
50	caa gtg agc cag cca ctg cga gag cag agt gag gat ata 2366			
	Gln Val Ser Gln Pro Leu Arg Glu Gln Ser Glu Asp Ile			
		15	20	
55	cct gga gcc att tcc tct gcc ctg gca gcc ata tgc act 2405			
	Pro Gly Ala Ile Ser Ser Ala Leu Ala Ala Ile Cys Thr			
		25	30	35
60	gct cct gtg gga ctg ccc gac tgc tgg gat gcc aag gag 2444			
	Ala Pro Val Gly Leu Pro Asp Cys Trp Asp Ala Lys Glu			
		40	45	50
65	cag gtc tgt tgg cat ttg gca aat cag cta act gaa gac 2483			
	Gln Val Cys Trp His Leu Ala Asn Gln Leu Thr Glu Asp			
		55	60	
70	agc agc cag ctc agg cca tcc ctc atc tct ggc ctg cag 2522			
	Ser Ser Gln Leu Arg Pro Ser Leu Ile Ser Gly Leu Gln			
		65	70	75
75	cat ccc atc ctg tgc ctg cac ctt ctc aag gtt cta tac 2561			
	His Pro Ile Leu Cys Leu His Leu Leu Lys Val Leu Tyr			
		80	85	
80	tcc tgc tgc ctt gtc agt gag ggc ctg tgc cgt ctt ctg 2600			
	Ser Cys Cys Leu Val Ser Glu Gly Leu Cys Arg Leu Leu			
		90	95	100
85	ggg cag gag ccc ctg gcc ttg gaa tcc ctg ttt atg ttg 2639			
	Gly Gln Glu Pro Leu Ala Leu Glu Ser Leu Phe Met Leu			
		105	110	115

ES 2 371 853 T3

	att	cag	ggc	aag	gta	aaa	gta	gta	gat	tgg	gaa	gag	tct	2678
	Ile	Gln	Gly	Lys	Val	Lys	Val	Val	Asp	Trp	Glu	Glu	Ser	
					120					125				
5	act	gaa	gtg	aca	ctc	tac	ttc	ctc	tcc	ctt	ctt	gtc	ttt	2717
	Thr	Glu	Val	Thr	Leu	Tyr	Phe	Leu	Ser	Leu	Leu	Val	Phe	
		130					135					140		
10	cgg	ctc	caa	aac	ctg	cct	tgt	gga	atg	gag	aag	cta	ggc	2756
	Arg	Leu	Gln	Asn	Leu	Pro	Cys	Gly	Met	Glu	Lys	Leu	Gly	
				145					150					
15	agt	gac	gtt	gct	act	ctc	ttt	acc	cat	tcg	cat	gtc	gtc	2795
	Ser	Asp	Val	Ala	Thr	Leu	Phe	Thr	His	Ser	His	Val	Val	
	155					160					165			
20	tct	ctt	gtg	agt	gca	gca	gcc	tgt	cta	ttg	gga	cag	ctt	2834
	Ser	Leu	Val	Ser	Ala	Ala	Ala	Cys	Leu	Leu	Gly	Gln	Leu	
			170					175					180	
25	ggt	cag	caa	ggg	gtg	acc	ttt	gac	ctc	cag	ccc	atg	gaa	2873
	Gly	Gln	Gln	Gly	Val	Thr	Phe	Asp	Leu	Gln	Pro	Met	Glu	
					185					190				
30	tgg	atg	gct	gca	gcc	aca	cat	gcc	ttg	tct	gcc	cct	gca	2912
	Trp	Met	Ala	Ala	Ala	Thr	His	Ala	Leu	Ser	Ala	Pro	Ala	
		195					200					205		
35	gag	gtt	cgg	ttg	act	cca	cca	ggt	agt	tgt	gga	ttc	tat	2951
	Glu	Val	Arg	Leu	Thr	Pro	Pro	Gly	Ser	Cys	Gly	Phe	Tyr	
				210					215					
40	gat	ggc	ctc	ctt	atc	ctt	ctg	ttg	cag	ctc	ctc	act	gag	2990
	Asp	Gly	Leu	Leu	Ile	Leu	Leu	Leu	Gln	Leu	Leu	Thr	Glu	
	220					225					230			
45	cag	ggg	aag	gct	agc	cta	atc	agg	gat	atg	tcc	agt	tca	3029
	Gln	Gly	Lys	Ala	Ser	Leu	Ile	Arg	Asp	Met	Ser	Ser	Ser	
			235					240					245	
50	gaa	atg	tgg	acc	gtt	ttg	tgg	cac	cgc	ttc	tcc	atg	gtc	3068
	Glu	Met	Trp	Thr	Val	Leu	Trp	His	Arg	Phe	Ser	Met	Val	
					250					255				
55	ctg	agg	ctc	ccc	gag	gag	gca	tct	gca	cag	gaa	ggg	gag	3107
	Leu	Arg	Leu	Pro	Glu	Glu	Ala	Ser	Ala	Gln	Glu	Gly	Glu	
		260					265					270		
60	ctt	tcg	cta	tcc	agt	cca	cca	agc	cct	gag	cca	gac	tgg	3146
	Leu	Ser	Leu	Ser	Ser	Pro	Pro	Ser	Pro	Glu	Pro	Asp	Trp	
				275					280					
65	aca	ctg	att	tct	ccc	cag	ggc	atg	gca	gcc	ctg	ctg	agc	3185
	Thr	Leu	Ile	Ser	Pro	Gln	Gly	Met	Ala	Ala	Leu	Leu	Ser	
					290						295			
70	ctg	gcc	atg	gcc	acc	ttt	acc	cag	gag	ccc	cag	tta	tgc	3224
	Leu	Ala	Met	Ala	Thr	Phe	Thr	Gln	Glu	Pro	Gln	Leu	Cys	
			300					305					310	
75	ctg	agc	tgc	ctg	tcc	cag	cat	gga	agt	atc	ctc	atg	tcc	3263
	Leu	Ser	Cys	Leu	Ser	Gln	His	Gly	Ser	Ile	Leu	Met	Ser	
					315					320				
80	atc	ctg	aag	cat	ctg	ctt	tgc	ccc	agc	ttc	ctg	aat	caa	3302
	Ile	Leu	Lys	His	Leu	Leu	Cys	Pro	Ser	Phe	Leu	Asn	Gln	
			325				330					335		
85	ctg	cgc	cag	gcg	cct	cat	ggg	tct	gag	ttt	ctc	cct	gtc	3341
	Leu	Arg	Gln	Ala	Pro	His	Gly	Ser	Glu	Phe	Leu	Pro	Val	

ES 2 371 853 T3

[illegible]

ES 2 371 853 T3

ctt gga gat cct cag gct ggt atc cgg cgc aat gtt gca 4043
 Leu Gly Asp Pro Gln Ala Gly Ile Arg Arg Asn Val Ala
 575 580

5 tca gct ctg ggc aac ttg gga cct gaa ggt ttg gga gag 4082
 Ser Ala Leu Gly Asn Leu Gly Pro Glu Gly Leu Gly Glu
 585 590 595

10 gag ctg tta cag tgc gaa gta ccc cag cgg ctc cta gaa 4121
 Glu Leu Leu Gln Cys Glu Val Pro Gln Arg Leu Leu Glu
 600 605

15 atg gca tgt gga gac ccc cag cca aat gtg aag gag gct 4160
 Met Ala Cys Gly Asp Pro Gln Pro Asn Val Lys Glu Ala
 610 615 620

20 gcc ctc att gcc ctc cgg agc ctg caa cag gag cct ggc 4199
 Ala Leu Ile Ala Leu Arg Ser Leu Gln Gln Glu Pro Gly
 625 630 635

25 atc cat cag gta ctg gtg tcc ctg ggt gcc agt gag aaa 4238
 Ile His Gln Val Leu Val Ser Leu Gly Ala Ser Glu Lys
 640 645

30 cta tcc ttg ctc tct ctg ggg aat cag tca ctg cca cac 4277
 Leu Ser Leu Leu Ser Leu Gly Asn Gln Ser Leu Pro His
 650 655 660

35 agc agt cct agg cct gcc tct gcc aaa cac tgc agg aaa 4316
 Ser Ser Pro Arg Pro Ala Ser Ala Lys His Cys Arg Lys
 665 670

40 ctc att cac ctc ctg agg cca gcc cat agc atg t 4350
 Leu Ile His Leu Leu Arg Pro Ala His Ser Met
 675 680 685

45 gattccagat tcctgcggtc cagcctccaa ctttggttgc cagctctttc 4400
 ttattctact acacaagccg ccaactcaac tgagagctaa agagactaga 4450
 aaagagataa gctgccaaact caactgagaa caagaaacta gaagagattt 4500
 atatataaag cttcttcctt ctcccagatg caggatgttt tcaaccagta 4550
 aattttattg ctgttggtgc cagagaagag tcctttcttc tctacatcca 4600
 ggggcctttt ctccaataat gtgcctttaa ctctagggac ctgcctcacg 4650
 50 gaccttaggg aaaaacctca acctgaaaga tctcttcctt tctggagctc 4700
 ctttaatctt cccagcaggt ttttgcctta gacgtgctgg ccccaggaca 4750
 gtgatgaaga cagagcctgt ctcagctcta ggctgtgggg atcaatgcc 4800
 55 tcagtccctg ttattgaggg attatccctt agccaacatt cctatctgtg 4850
 ggtgggcgtg gagagtgtat ctttttttgg ggtgtgtgtg tatatgtgtg 4900
 60 tgtgtatgtg tgtgtgtgtt taatagttct gtttgtaaac tcttttaata 4950
 aaagtgtgtc ctcaccatac ttgaagctcc caggacaagg gttgagaggc 5000
 tcaaccctc tttcagcttc tatgtggtgt tggaggtgct ggtatcgtgt 5050
 65 tcacacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5100
 aaaaaaaaaa aaaaaaaaaa aaaaa 5125

<210> 5

70 <211> 648

ES 2 371 853 T3

<212> PRT
<213> Homo sapiens

<400> 5

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

ES 2 371 853 T3

	Thr Ser Lys Val	Ala 335	Pro Gly Thr Ala	Pro 340	Leu Pro Arg Leu	Gly 345
5	Ala Thr Pro Gln	Glu 350	Ser Ser Leu Leu	Ala 355	Gly Ile Leu Ala	Ser 360
	Glu Leu Lys Ser	Ser 365	Trp Ala Lys Ser	Gly 370	Thr Gly Glu Val	Pro 375
10	Ser Ala Pro Arg	Glu 380	Asn Arg Thr Thr	Pro 385	Asp Cys Glu Arg	Ala 390
	Phe Pro Glu Glu	Arg 395	Pro Glu Val Leu	Gly 400	Gln Arg Ser Thr	Asp 405
15	Val Val Asp Leu	Glu 410	Asn Glu Glu Pro	Asp 415	Ser Asp Asn Glu	Trp 420
20	Gln His Leu Leu	Glu 425	Thr Thr Glu Pro	Val 430	Pro Ile Gln Leu	Lys 435
	Ala Pro Leu Thr	Leu 440	Leu Cys Asn Pro	Asp 445	Phe Cys Gln Arg	Ile 450
25	Gln Ser Gln Leu	His 455	Glu Ala Gly Gly	Gln 460	Ile Leu Lys Gly	Ile 465
	Leu Glu Gly Ala	Ser 470	His Ile Leu Pro	Ala 475	Phe Arg Val Leu	Ser 480
30	Ser Leu Leu Ser	Ser 485	Cys Ser Asp Ser	Val 490	Ala Leu Tyr Ser	Phe 495
35	Cys Arg Glu Ala	Gly 500	Leu Pro Gly Leu	Leu 505	Leu Ser Leu Leu	Arg 510
	His Ser Gln Glu	Ser 515	Asn Ser Leu Gln	Gln 520	Gln Ser Trp Tyr	Gly 525
40	Thr Phe Leu Gln	Asp 530	Leu Met Ala Val	Ile 535	Gln Ala Tyr Phe	Ala 540
	Cys Thr Phe Asn	Leu 545	Glu Arg Ser Gln	Thr 550	Ser Asp Ser Leu	Gln 555
45	Val Phe Gln Glu	Ala 560	Ala Asn Leu Phe	Leu 565	Asp Leu Leu Gly	Lys 570
50	Leu Leu Ala Gln	Pro 575	Asp Asp Ser Glu	Gln 580	Thr Leu Gln Arg	Asp 585
	Ser Leu Met Cys	Phe 590	Thr Val Leu Cys	Glu 595	Ala Met Asp Gly	Asn 600
55	Ser Arg Ala Ile	Ser 605	Lys Ala Phe Tyr	Ser 610	Ser Leu Leu Thr	Thr 615
	Gln Gln Val Val	Leu 620	Asp Gly Leu Leu	His 625	Gly Leu Thr Val	Pro 630
60	Gln Leu Pro Val	His 635	Thr Pro Gln Gly	Asn 640	Gln Ser Gly Glu	Gly 645
65	Arg Phe Ser					

<210> 6
 <211> 5252
 <212> DNA

ES 2 371 853 T3

<213> Homo sapiens

<400> 6

```

5  ggagcttgga gctcctaggc tgggggcgctc ccagatgttg tggaactgtc 50
   cctggatcta tagctcttca ccgtctctac tttcttcctt ctaagagatc 100
   ctgaaacctc tgtc atg gaa aag tac cac gtg ttg gag atg 141
10      Met Glu Lys Tyr His Val Leu Glu Met
      1      5
   att gga gaa ggc tct ttt ggg agg gtg tac aag ggt cga 180
   Ile Gly Glu Gly Ser Phe Gly Arg Val Tyr Lys Gly Arg
      10      15      20
15  aga aaa tac agt gct cag gtc gtg gcc ctg aag ttc atc 219
   Arg Lys Tyr Ser Ala Gln Val Val Ala Leu Lys Phe Ile
      25      30      35
20  cca aaa ttg ggg cgc tca gag aag gag ctg agg aat ttg 258
   Pro Lys Leu Gly Arg Ser Glu Lys Glu Leu Arg Asn Leu
      40      45
25  caa cga gag att gaa ata atg cgg ggt ctg cgg cat ccc 297
   Gln Arg Glu Ile Glu Ile Met Arg Gly Leu Arg His Pro
      50      55      60
30  aac att gtg cat atg ctt gac agc ttt gaa act gat aaa 336
   Asn Ile Val His Met Leu Asp Ser Phe Glu Thr Asp Lys
      65      70
35  gag gtg gtg gtg gtg aca gac tat gct gag gga gag ctc 375
   Glu Val Val Val Val Thr Asp Tyr Ala Glu Gly Glu Leu
      75      80      85
40  ttt cag atc cta gaa gat gac gga aaa ctt cct gaa gac 414
   Phe Gln Ile Leu Glu Asp Asp Gly Lys Leu Pro Glu Asp
      90      95      100
45  cag gtt cag gcc att gct gcc cag ttg gtg tca gcc ctg 453
   Gln Val Gln Ala Ile Ala Ala Gln Leu Val Ser Ala Leu
      105      110
50  tac tat ctg cat tcc cac cgc atc cta cac cga gat atg 492
   Tyr Tyr Leu His Ser His Arg Ile Leu His Arg Asp Met
      115      120      125
55  aag cct cag aac atc ctc ctc gcc aag ggt ggt ggc atc 531
   Lys Pro Gln Asn Ile Leu Leu Ala Lys Gly Gly Gly Ile
      130      135
60  aag ctc tgt gac ttt gga ttt gcc cgg gct atg agc acc 570
   Lys Leu Cys Asp Phe Gly Phe Ala Arg Ala Met Ser Thr
      140      145      150
65  aat aca atg gtg ctg aca tcc atc aaa ggc aca cca ctc 609
   Asn Thr Met Val Leu Thr Ser Ile Lys Gly Thr Pro Leu
      155      160      165
70  tat atg tct cca gag ctg gtg gag gag cga cca tac gac 648
   Tyr Met Ser Pro Glu Leu Val Glu Glu Arg Pro Tyr Asp
      170      175
   cac aca gcg gac ctc tgg tct gtt ggc tgc ata cta tat 687
   His Thr Ala Asp Leu Trp Ser Val Gly Cys Ile Leu Tyr
      180      185      190
   gaa ctg gca gta ggc acc cct ccc ttc tat gct aca agc 726
   Glu Leu Ala Val Gly Thr Pro Pro Phe Tyr Ala Thr Ser
      195      200

```

ES 2 371 853 T3

5	atc Ile 205	ttt Phe	cag Gln	ctg Leu	gtc Val	agc Ser 210	ctc Leu	att Ile	ctc Leu	aag Lys	gac Asp 215	cct Pro	gtg Val	765
	cgc Arg	tgg Trp	ccc Pro 220	tca Ser	acc Thr	atc Ile	agt Ser 225	ccc Pro	tgc Cys	ttt Phe	aag Lys	aac Asn	ttc Phe 230	804
10	ctg Leu	cag Gln	gga Gly	ctg Leu	ctc Leu 235	acc Thr	aaa Lys	gac Asp	cca Pro	cgg Arg 240	cag Gln	cga Arg	ctg Leu	843
15	tcc Ser	tgg Trp 245	cca Pro	gac Asp	ctc Leu	tta Leu	tat Tyr 250	cac His	ccc Pro	ttt Phe	att Ile	gct Ala 255	ggt Gly	882
20	cat His	gtc Val	acc Thr	ata Ile 260	ata Ile	act Thr	gag Glu	cca Pro	gca Ala 265	ggc Gly	cca Pro	gat Asp	ttg Leu	921
25	ggg Gly 270	acc Thr	cca Pro	ttc Phe	acc Thr	agc Ser 275	cgc Arg	cta Leu	ccc Pro	cca Pro	gaa Glu 280	ctt Leu	cag Gln	960
	gtc Val	cta Leu	aag Lys 285	gac Asp	gaa Glu	cag Gln	gcc Ala	cat His 290	cgg Arg	ttg Leu	gcc Ala	ccc Pro	aag Lys 295	999
30	ggt Gly	aat Asn	cag Gln	tct Ser	cgc Arg 300	atc Ile	ttg Leu	act Thr	cag Gln	gcc Ala 305	tat Tyr	aaa Lys	cgc Arg	1038
35	atg Met 310	gct Ala	gag Glu	gag Glu	gcc Ala	atg Met	cag Gln 315	aag Lys	aaa Lys	cat His	cag Gln	aac Asn 320	aca Thr	1077
40	gga Gly	cct Pro	gcc Ala	ctt Leu 325	gag Glu	caa Gln	gag Glu	gac Asp	aag Lys 330	acc Thr	agc Ser	aag Lys	gtg Val	1116
45	gct Ala 335	cct Pro	ggc Gly	aca Thr	gcc Ala	cct Pro 340	ctg Leu	ccc Pro	aga Arg	ctc Leu	ggg Gly 345	gcc Ala	act Thr	1155
	cct Pro	cag Gln	gaa Glu 350	tca Ser	agc Ser	ctc Leu	ctg Leu	gcc Ala 355	ggg Gly	atc Ile	tta Leu	gcc Ala	tca Ser 360	1194
50	gaa Glu	ttg Leu	aag Lys	agc Ser	agc Ser 365	tgg Trp	gct Ala	aaa Lys	tca Ser	ggg Gly 370	act Thr	gga Gly	gag Glu	1233
55	gtg Val 375	ccc Pro	tct Ser	gca Ala	cct Pro	cgg Arg 380	gaa Glu	aac Asn	cgg Arg	acc Thr	acc Thr	cca Pro 385	gat Asp	1272
60	tgt Cys	gaa Glu	cga Arg	gca Ala 390	ttc Phe	cca Pro	gag Glu	gag Glu	agg Arg 395	cca Pro	gag Glu	gtg Val	ctg Leu	1311
65	ggc Gly 400	cag Gln	cgg Arg	agc Ser	act Thr	gat Asp 405	gta Val	gtg Val	gac Asp	ctg Leu	gaa Glu 410	aat Asn	gag Glu	1350
	gag Glu	cca Pro	gac Asp 415	agt Ser	gac Asp	aat Asn	gag Glu	tgg Trp 420	cag Gln	cac His	ctg Leu	cta Leu	gag Glu 425	1389
70	acc	act	gag	cct	gtg	cct	att	caa	ctg	aag	gct	cct	ctc	1428

ES 2 371 853 T3

	Thr	Thr	Glu	Pro	Val	Pro	Ile	Gln	Leu	Lys	Ala	Pro	Leu	
					430					435				
5	acc	ttg	ctg	tgt	aat	cct	gac	ttc	tgc	cag	cgc	atc	cag	1467
	Thr	Leu	Leu	Cys	Asn	Pro	Asp	Phe	Cys	Gln	Arg	Ile	Gln	
		440					445					450		
10	agt	cag	ctg	cat	gaa	gct	gga	ggg	cag	atc	ctg	aaa	ggc	1506
	Ser	Gln	Leu	His	Glu	Ala	Gly	Gly	Gln	Ile	Leu	Lys	Gly	
				455					460					
15	atc	ttg	gag	ggc	gct	tcc	cac	atc	ctg	cct	gca	ttc	cgg	1545
	Ile	Leu	Glu	Gly	Ala	Ser	His	Ile	Leu	Pro	Ala	Phe	Arg	
	465				470						475			
20	gtc	ctg	agc	agt	ctt	ctc	tcc	agc	tgc	agt	gat	tct	gtt	1584
	Val	Leu	Ser	Ser	Leu	Leu	Ser	Ser	Cys	Ser	Asp	Ser	Val	
			480					485					490	
25	gcc	ttg	tat	tcc	ttc	tgc	cgg	gag	gca	ggg	ctt	cct	ggg	1623
	Ala	Leu	Tyr	Ser	Phe	Cys	Arg	Glu	Ala	Gly	Leu	Pro	Gly	
					495					500				
30	ctg	ctg	ctg	agt	cta	ctc	agg	cac	agt	cag	gag	agc	aac	1662
	Leu	Leu	Leu	Ser	Leu	Leu	Arg	His	Ser	Gln	Glu	Ser	Asn	
		505					510					515		
35	agc	ctc	cag	cag	caa	tct	tgg	tat	ggg	acc	ttc	tta	cag	1701
	Ser	Leu	Gln	Gln	Gln	Ser	Trp	Tyr	Gly	Thr	Phe	Leu	Gln	
				520					525					
40	gac	ctg	atg	gct	gtg	att	cag	gcc	tac	ttt	gcc	tgt	acc	1740
	Asp	Leu	Met	Ala	Val	Ile	Gln	Ala	Tyr	Phe	Ala	Cys	Thr	
	530					535					540			
45	ttc	aat	ctg	gag	agg	agc	cag	aca	agt	gac	agc	ctg	cag	1779
	Phe	Asn	Leu	Glu	Arg	Ser	Gln	Thr	Ser	Asp	Ser	Leu	Gln	
			545					550					555	
50	gtg	ttt	cag	gag	gct	gcc	aac	ctt	ttt	ctg	gac	ctg	ttg	1818
	Val	Phe	Gln	Glu	Ala	Ala	Asn	Leu	Phe	Leu	Asp	Leu	Leu	
				560						565				
55	ggg	aaa	ctg	ctg	gcc	caa	cca	gat	gac	tct	gag	cag	act	1857
	Gly	Lys	Leu	Leu	Ala	Gln	Pro	Asp	Asp	Ser	Glu	Gln	Thr	
		570					575					580		
60	ttg	cgg	agg	gac	agc	ctt	atg	tgc	ttt	act	gtc	ctg	tgc	1896
	Leu	Arg	Arg	Asp	Ser	Leu	Met	Cys	Phe	Thr	Val	Leu	Cys	
				585					590					
65	gaa	gcc	atg	gat	ggg	aac	agc	cgg	gcc	atc	tcc	aaa	gcc	1935
	Glu	Ala	Met	Asp	Gly	Asn	Ser	Arg	Ala	Ile	Ser	Lys	Ala	
	595					600					605			
70	ttt	tac	tcc	agc	ttg	ctg	acg	aca	cag	cag	gtt	gtc	ttg	1974
	Phe	Tyr	Ser	Ser	Leu	Leu	Thr	Thr	Gln	Gln	Val	Val	Leu	
			610					615					620	
75	gat	ggg	ctc	ctt	cat	ggc	ttg	aca	gtt	cca	cag	ctc	cct	2013
	Asp	Gly	Leu	Leu	His	Gly	Leu	Thr	Val	Pro	Gln	Leu	Pro	
					625					630				
80	gtc	cac	act	ccc	caa	ggc	tcc	cta	ctc	ctg	ctg	cca	tgt	2052
	Val	His	Thr	Pro	Gln	Gly	Ser	Leu	Leu	Leu	Leu	Pro	Cys	
		635					640					645		
85	cgg	tga	g	t	act	ggc	gcta	ttgt	ctaggg	caagag	cctc			2090
	Arg	Xaa												
		648												

ES 2 371 853 T3

5 aggccttttg agt tac tct ttg ctt ttc tcc aca gga gcc 2130
 Ser Tyr Ser Leu Leu Phe Ser Thr Gly Ala 10
 10 ccg caa gtg agc cag cca ctg cga gag cag agt gag gat 2169
 Pro Gln Val Ser Gln Pro Leu Arg Glu Gln Ser Glu Asp 20
 15 ata cct gga gcc att tcc tct gcc ctg gca gcc ata tgc 2208
 Ile Pro Gly Ala Ile Ser Ser Ala Leu Ala Ala Ile Cys 35
 20 act gct cct gtg gga ctg ccc gac tgc tgg gat gcc aag 2247
 Thr Ala Pro Val Gly Leu Pro Asp Cys Trp Asp Ala Lys 45
 25 gag cag gtc tgt tgg cat ttg gca aat cag cta act gaa 2286
 Glu Gln Val Cys Trp His Leu Ala Asn Gln Leu Thr Glu 60
 30 gac agc agc cag ctc agg cca tcc ctc atc tct ggc ctg 2325
 Asp Ser Ser Gln Leu Arg Pro Ser Leu Ile Ser Gly Leu 75
 35 cag cat ccc atc ctg tgc ctg cac ctt ctc aag gtt cta 2364
 Gln His Pro Ile Leu Cys Leu His Leu Leu Lys Val Leu 85
 40 tac tcc tgc tgc ctt gtc agt gag ggc ctg tgc cgt ctt 2403
 Tyr Ser Cys Cys Leu Val Ser Glu Gly Leu Cys Arg Leu 100
 45 ctg ggg cag gag ccc ctg gcc ttg gaa tcc ctg ttt atg 2442
 Leu Gly Gln Glu Pro Leu Ala Leu Glu Ser Leu Phe Met 110
 50 ttg att cag ggc aag gta aaa gta gta gat tgg gaa gag 2481
 Leu Ile Gln Gly Lys Val Lys Val Val Asp Trp Glu Glu 125
 55 tct act gaa gtg aca ctc tac ttc ctc tcc ctt ctt gtc 2520
 Ser Thr Glu Val Thr Leu Tyr Phe Leu Ser Leu Leu Val 140
 60 ttt cgg ctc caa aac ctg cct tgt gga atg gag aag cta 2559
 Phe Arg Leu Gln Asn Leu Pro Cys Gly Met Glu Lys Leu 150
 65 ggc agt gac gtt gct act ctc ttt acc cat tcg cat gtc 2598
 Gly Ser Asp Val Ala Thr Leu Phe Thr His Ser His Val 165
 70 gtc tct ctt gtg agt gca gca gcc tgt cta ttg gga cag 2637
 Val Ser Leu Val Ser Ala Ala Ala Cys Leu Leu Gly Gln 175
 75 ctt ggt cag caa ggg gtg acc ttt gac ctc cag ccc atg 2676
 Leu Gly Gln Gln Gly Val Thr Phe Asp Leu Gln Pro Met 185
 80 gaa tgg atg gct gca gcc aca cat gcc ttg tct gcc cct 2715
 Glu Trp Met Ala Ala Ala Thr His Ala Leu Ser Ala Pro 205
 85 gca gag ctc ctc act gag gta cag atg gat ctt ggg atg 2754
 Ala Glu Leu Leu Thr Glu Val Gln Met Asp Leu Gly Met 215
 90 gat ggg aag taaagag agaggaactg ggcatttttg ggagcctctg 2800

ES 2 371 853 T3

Asp Gly Lys
220 221

```

5  gaccagagga atgaagaagc aaccacacagc cttccctctc aagctactgt 2850
   gcctgtgata gccttggaac ttcccccgcct gccctcagta ctgacccttt 2900
   gaaggaaacc attcgctgcg tcccctggga tccagtggga gataaaatga 2950
10 attccctggg tttcagcaga catacacatg agttgtgagg tcagaggggtt 3000
   aaggtttgat aagaaaatga aataagacga cagggaataa ctagggtggga 3050
   aagcgggaagg aattatttct gggacttcct ttacttgtaa gtcagggaca 3100
15 ggaatgaata aaagcatttg gattcctgac ttctgtcttt cccccgccc 3150
   tctttcactt ttatctctag cagggaagg ctagcctaata cagggatatg 3200
   tccagttcag aaatgtggac cgttttgtgg caccgcttct ccatggtcct 3250
   gaggctcccc gaggaggcat ctgcacagga aggggagctt tcgctatcca 3300
   gtccaccaag ccctgagcca gactggacac tgatttctcc ccagggcatg 3350
25 gcagccctgc tgagcctggc catggccacc tttaccag agccccagtt 3400
   atgcctgagc tgctgtccc agcatggaag taccctcatg tccatcctga 3450
   agcatctgct ttgccccagc ttctgaatc aactgcgcca ggcgcctcat 3500
   gggctctgagt ttctccctgt cgtggtgctc tctgtctgcc agctcctttg 3550
   cttccctttt gcgctggaca tggatgctga cctccttata ggtgtcttgg 3600
35 ccgacctcag ggactcagaa gttgcagccc atctgtgca ggtctgctgc 3650
   taccatcttc cgttgatgca agtggagctg cccatcagcc ttctcacacg 3700
   cctggccctc atggatccca cctctctcaa ccagtttgtg aacacagtgt 3750
   ctgcctcccc tagaaccatc gtctcgtttc tctcagttgc cctcctgagt 3800
   gaccagccac tgttgacctc cgaccttctc tctctgctgg ccatactgc 3850
45 cagggtcctg tctccagcc acttgctcctt tatccaagag cttctggctg 3900
   gctctgatga atcctatcgg cccctgcgca gcctcctggg ccaccagag 3950
50 aattctgtgc gggcacacac ttataggctc ctgggacact tgctccaaca 4000
   cagcatggcc ctgctgggg cactgcagag ccagtctgga ctgctcagcc 4050
   ttctgtgctg tgggcttggg gacaaggatc ctgttgctgc gtgcagtgcc 4100
55 agctttgctg tgggcaatgc agcctaccag gctggctcctc tgggacctgc 4150
   cctggcagct gcagtgccca gtatgacca gctgcttggg gatcctcagg 4200
60 ctggtatccg gcgcaatgtt gcatcagctc tgggcaactt gggacctgaa 4250
   ggtttgggag aggagctgtt acagtgcgaa gtacccagc ggctcctaga 4300
   aatggcatgt ggagaccccc agccaaatgt gaaggaggct gccctcattg 4350
65 ccctccggag cctgcaacag gagcctggca tccatcaggt actggtgtcc 4400
   ctgggtgcca gtgagaaact atccttgctc tctctgggga atcagtcact 4450
70 gccacacagc agtcctaggc ctgcctctgc caaactgc aggaaactca 4500

```

ES 2 371 853 T3

ttcacctcct gaggccagcc catagcatgt gattccagat tcctgcggtc 4550
 5 cagcctccaa ctttggttgc cagctctttc ttattctact acacaagccg 4600
 ccaactcaac tgagagctaa agagactaga aaagagataa gctgccaact 4650
 caactgagaa caagaaacta gaagagattt atatataaag cttcttcctt 4700
 10 ctcccagatg caggatgttt tcaaccagta aattttattg ctgttggtgc 4750
 cagagaagag tcctttcttc tctacatcca ggggcctttt ctccaataat 4800
 gtgcctttta ctctagggac ctgcctcacg gaccttaggg aaaaacctca 4850
 15 acctgaaaga tctcttcctt tctggagctc ctttaatctt cccagcaggt 4900
 ttttgcctta gacgtgctgg ccccaggaca gtgatgaaga cagagcctgt 4950
 20 ctcagctcta ggctgtgggg atcaatgcca tcagtccctg ttattgaggg 5000
 attatccctt agccaacatt cctatctgtg ggtgggcgtg gagagtgtat 5050
 ctttttttgg ggtgtgtgtg tatatgtgtg tgtgtatgtg tgtgtgtgtt 5100
 25 taatagttct gtttgtaaac tcttttaata aaagtgtgtc ctcaccatac 5150
 ttgaagctcc caggacaagg gttgagaggc tcaaccctc tttcagcttc 5200
 30 tatgtggtgt tggaggtgct ggtatcgtgt tcacacaaaa aaaaaaaaaa 5250
 aa 5252

<210> 7
 35 <211> 647
 <212> PRT
 <213> Homo sapiens

<400> 7
 40 Met Glu Lys Tyr His Val Leu Glu Met Ile Gly Glu Gly Ser Phe
 1 5 10 15
 Gly Arg Val Tyr Lys Gly Arg Arg Lys Tyr Ser Ala Gln Val Val
 20 25 30
 45 Ala Leu Lys Phe Ile Pro Lys Leu Gly Arg Ser Glu Lys Glu Leu
 35 40 45
 Arg Asn Leu Gln Arg Glu Ile Glu Ile Met Arg Gly Leu Arg His
 50 50 55 60
 Pro Asn Ile Val His Met Leu Asp Ser Phe Glu Thr Asp Lys Glu
 65 70 75
 55 Val Val Val Val Thr Asp Tyr Ala Glu Gly Glu Leu Phe Gln Ile
 80 85 90
 Leu Glu Asp Asp Gly Lys Leu Pro Glu Asp Gln Val Gln Ala Ile
 95 100 105
 60 Ala Ala Gln Leu Val Ser Ala Leu Tyr Tyr Leu His Ser His Arg
 110 115 120
 Ile Leu His Arg Asp Met Lys Pro Gln Asn Ile Leu Leu Ala Lys
 65 125 130 135
 Gly Gly Gly Ile Lys Leu Cys Asp Phe Gly Phe Ala Arg Ala Met
 140 145 150

ES 2 371 853 T3

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

ES 2 371 853 T3

	500	505	510
	His Ser Gln Glu Ser Asn Ser Leu Gln	Gln Gln Ser Trp Tyr	Gly
5	Thr Phe Leu Gln Asp Leu Met Ala Val	Ile Gln Ala Tyr Phe	Ala
	515 530	520 535	525 540
10	Cys Thr Phe Asn Leu Glu Arg Ser Gln	Thr Ser Asp Ser Leu	Gln
	545 560	550 565	555 570
15	Val Phe Gln Glu Ala Ala Asn Leu Phe	Leu Asp Leu Leu Gly	Lys
	575 590	580 595	585 600
20	Leu Leu Ala Gln Pro Asp Asp Ser Glu	Gln Thr Leu Arg Arg	Asp
	595 605	610 615	600 615
25	Ser Leu Met Cys Phe Thr Val Leu Cys	Glu Ala Met Asp Gly	Asn
	620 635	625 640	630 645
30	Gln Gln Val Val Leu Asp Gly Leu Leu	His Gly Leu Thr Val	Pro
	640 645	645 650	645 650
35	Gln Leu Pro Val His Thr Pro Gln Gly	Ser Leu Leu Leu Leu	Pro
	650 655	655 660	655 660
40	Cys Arg		

<210> 8

<211> 32

<212> DNA

<213> secuencia artificial

<220>

<223> secuencia artificial

<400> 8

caatacaatg gtgctgacat ccatcaaagg ca 32

<210> 9

<211> 23

<212> DNA

<213> secuencia artificial

<220>

<223> secuencia artificial

<400> 9

gaagggaggg gtcctactg cca 23

<210> 10

<211> 38

<212> DNA

<213> secuencia artificial

<220>

<223> secuencia artificial

<400> 10

ctccagctct ggagacatat agagtgggtg gcctttga 38

<210> 11
 <211> 60
 <212> DNA
 <213> secuencia artificial
 5
 <220>
 <223> secuencia artificial

 <400> 11
 10 ccatcgatgt acccatacga cgtcccagac tacgctgaaa agtaccacgt 50
 gttggagatg 60

 <210> 12
 15 <211> 32
 <212> DNA
 <213> secuencia artificial

 <220>
 20 <223> secuencia artificial

 <400> 12
 gctctagact aaggggcagg tcctgtgttc tg 32
 25
 <210> 13
 <211> 21
 <212> DNA
 <213> secuencia artificial
 30
 <220>
 <223> secuencia artificial

 <400> 13
 35 ctgacgacac agcaggttgt c 21

 <210> 14
 <211> 21
 <212> DNA
 <213> secuencia artificial
 40
 <220>
 <223> secuencia artificial

 <400> 14
 45 cagatgcttc aggatggaca t 21

 <210> 15
 <211> 19
 <212> DNA
 50 <213> secuencia artificial

 <220>
 <223> secuencia artificial

 <400> 15
 55 agagtagcaa cgtcactgc 19

 <210> 16
 <211> 21
 60 <212> DNA
 <213> secuencia artificial

<220>
 <223> secuencia artificial
 5 <400> 16
 cctcactgac aaggcagcag g 21
 <210> 17
 <211> 21
 10 <212> DNA
 <213> secuencia artificial
 <220>
 <223> secuencia artificial
 15 <400> 17
 cccgaggagg catctgcaca g 21
 <210> 18
 20 <211> 21
 <212> DNA
 <213> secuencia artificial
 <220>
 25 <223> secuencia artificial
 <400> 18
 cagaacttca ggtcctaaag g 21
 30 <210> 19
 <211> 32
 <212> DNA
 <213> secuencia artificial
 35 <220>
 <223> secuencia artificial
 <400> 19
 40 tcgacaagca gggaacaccc aagtagaagc tc 32
 <210> 20
 <211> 32
 <212> DNA
 <213> secuencia artificial
 45 <220>
 <223> secuencia artificial
 <400> 20
 50 tcgacaagca gggaagtggg aagtagaagc tc 32
 <210> 21
 <211> 685
 <212> PRT
 55 <213> Homo sapiens
 <400> 21
 Ser Tyr Ser Leu Leu Phe Ser Thr Gly Ala Pro Gln Val Ser Gln
 1 5 10 15
 60 Pro Leu Arg Glu Gln Ser Glu Asp Ile Pro Gly Ala Ile Ser Ser
 20 25 30

ES 2 371 853 T3

	Ala	Leu	Ala	Ala	Ile 35	Cys	Thr	Ala	Pro	Val 40	Gly	Leu	Pro	Asp	Cys 45
5	Trp	Asp	Ala	Lys	Glu 50	Gln	Val	Cys	Trp	His 55	Leu	Ala	Asn	Gln	Leu 60
	Thr	Glu	Asp	Ser	Ser 65	Gln	Leu	Arg	Pro	Ser 70	Leu	Ile	Ser	Gly	Leu 75
10	Gln	His	Pro	Ile	Leu 80	Cys	Leu	His	Leu	Leu 85	Lys	Val	Leu	Tyr	Ser 90
15	Cys	Cys	Leu	Val	Ser 95	Glu	Gly	Leu	Cys	Arg 100	Leu	Leu	Gly	Gln	Glu 105
	Pro	Leu	Ala	Leu	Glu 110	Ser	Leu	Phe	Met	Leu 115	Ile	Gln	Gly	Lys	Val 120
20	Lys	Val	Val	Asp	Trp 125	Glu	Glu	Ser	Thr	Glu 130	Val	Thr	Leu	Tyr	Phe 135
	Leu	Ser	Leu	Leu	Val 140	Phe	Arg	Leu	Gln	Asn 145	Leu	Pro	Cys	Gly	Met 150
25	Glu	Lys	Leu	Gly	Ser 155	Asp	Val	Ala	Thr	Leu 160	Phe	Thr	His	Ser	His 165
30	Val	Val	Ser	Leu	Val 170	Ser	Ala	Ala	Ala	Cys 175	Leu	Leu	Gly	Gln	Leu 180
	Gly	Gln	Gln	Gly	Val 185	Thr	Phe	Asp	Leu	Gln 190	Pro	Met	Glu	Trp	Met 195
35	Ala	Ala	Ala	Thr	His 200	Ala	Leu	Ser	Ala	Pro 205	Ala	Glu	Val	Arg	Leu 210
	Thr	Pro	Pro	Gly	Ser 215	Cys	Gly	Phe	Tyr	Asp 220	Gly	Leu	Leu	Ile	Leu 225
40	Leu	Leu	Gln	Leu	Leu 230	Thr	Glu	Gln	Gly	Lys 235	Ala	Ser	Leu	Ile	Arg 240
45	Asp	Met	Ser	Ser	Ser 245	Glu	Met	Trp	Thr	Val 250	Leu	Trp	His	Arg	Phe 255
	Ser	Met	Val	Leu	Arg 260	Leu	Pro	Glu	Glu	Ala 265	Ser	Ala	Gln	Glu	Gly 270
50	Glu	Leu	Ser	Leu	Ser 275	Ser	Pro	Pro	Ser	Pro 280	Glu	Pro	Asp	Trp	Thr 285
	Leu	Ile	Ser	Pro	Gln 290	Gly	Met	Ala	Ala	Leu 295	Leu	Ser	Leu	Ala	Met 300
55	Ala	Thr	Phe	Thr	Gln 305	Glu	Pro	Gln	Leu	Cys 310	Leu	Ser	Cys	Leu	Ser 315
60	Gln	His	Gly	Ser	Ile 320	Leu	Met	Ser	Ile	Leu 325	Lys	His	Leu	Leu	Cys 330
	Pro	Ser	Phe	Leu	Asn 335	Gln	Leu	Arg	Gln	Ala 340	Pro	His	Gly	Ser	Glu 345
65	Phe	Leu	Pro	Val	Val 350	Val	Leu	Ser	Val	Cys 355	Gln	Leu	Leu	Cys	Phe 360
70	Pro	Phe	Ala	Leu	Asp 365	Met	Asp	Ala	Asp	Leu 370	Leu	Ile	Val	Val	Leu 375

ES 2 371 853 T3

	Ala Asp Leu Arg	Asp 380	Ser Glu Val Ala	Ala 385	His Leu Leu Gln	Val 390
5	Cys Cys Tyr His	Leu 395	Pro Leu Met Gln	Val 400	Glu Leu Pro Ile	Ser 405
	Leu Leu Thr Arg	Leu 410	Ala Leu Met Asp	Pro 415	Thr Ser Leu Asn	Gln 420
10	Phe Val Asn Thr	Val 425	Ser Ala Ser Pro	Arg 430	Thr Ile Val Ser	Phe 435
	Leu Ser Val Ala	Leu 440	Leu Ser Asp Gln	Pro 445	Leu Leu Thr Ser	Asp 450
15	Leu Leu Ser Leu	Leu 455	Ala His Thr Ala	Arg 460	Val Leu Ser Pro	Ser 465
20	His Leu Ser Phe	Ile 470	Gln Glu Leu Leu	Ala 475	Gly Ser Asp Glu	Ser 480
	Tyr Arg Pro Leu	Arg 485	Ser Leu Leu Gly	His 490	Pro Glu Asn Ser	Val 495
25	Arg Ala His Thr	Tyr 500	Arg Leu Leu Gly	His 505	Leu Leu Gln His	Ser 510
	Met Ala Leu Arg	Gly 515	Ala Leu Gln Ser	Gln 520	Ser Gly Leu Leu	Ser 525
30	Leu Leu Leu Leu	Gly 530	Leu Gly Asp Lys	Asp 535	Pro Val Val Arg	Cys 540
35	Ser Ala Ser Phe	Ala 545	Val Gly Asn Ala	Ala 550	Tyr Gln Ala Gly	Pro 555
	Leu Gly Pro Ala	Leu 560	Ala Ala Ala Val	Pro 565	Ser Met Thr Gln	Leu 570
40	Leu Gly Asp Pro	Gln 575	Ala Gly Ile Arg	Arg 580	Asn Val Ala Ser	Ala 585
	Leu Gly Asn Leu	Gly 590	Pro Glu Gly Leu	Gly 595	Glu Glu Leu Leu	Gln 600
45	Cys Glu Val Pro	Gln 605	Arg Leu Leu Glu	Met 610	Ala Cys Gly Asp	Pro 615
50	Gln Pro Asn Val	Lys 620	Glu Ala Ala Leu	Ile 625	Ala Leu Arg Ser	Leu 630
	Gln Gln Glu Pro	Gly 635	Ile His Gln Val	Leu 640	Val Ser Leu Gly	Ala 645
55	Ser Glu Lys Leu	Ser 650	Leu Leu Ser Leu	Gly 655	Asn Gln Ser Leu	Pro 660
	His Ser Ser Pro	Arg 665	Pro Ala Ser Ala	Lys 670	His Cys Arg Lys	Leu 675
60	Ile His Leu Leu	Arg 680	Pro Ala His Ser	Met 685		
65	<210> 22					
	<211> 221					
	<212> PRT					
	<213> Homo sapiens					
	<400> 22					

ES 2 371 853 T3

Ser Tyr Ser Leu Leu Phe Ser Thr Gly Ala Pro Gln Val Ser Gln
 1 5 10 15
 5 Pro Leu Arg Glu Gln Ser Glu Asp Ile Pro Gly Ala Ile Ser Ser
 20 25 30
 Ala Leu Ala Ala Ile Cys Thr Ala Pro Val Gly Leu Pro Asp Cys
 35 40 45
 10 Trp Asp Ala Lys Glu Gln Val Cys Trp His Leu Ala Asn Gln Leu
 50 55 60
 Thr Glu Asp Ser Ser Gln Leu Arg Pro Ser Leu Ile Ser Gly Leu
 65 70 75
 15 Gln His Pro Ile Leu Cys Leu His Leu Leu Lys Val Leu Tyr Ser
 80 85 90
 20 Cys Cys Leu Val Ser Glu Gly Leu Cys Arg Leu Leu Gly Gln Glu
 95 100 105
 Pro Leu Ala Leu Glu Ser Leu Phe Met Leu Ile Gln Gly Lys Val
 110 115 120
 25 Lys Val Val Asp Trp Glu Glu Ser Thr Glu Val Thr Leu Tyr Phe
 125 130 135
 Leu Ser Leu Leu Val Phe Arg Leu Gln Asn Leu Pro Cys Gly Met
 140 145 150
 30 Glu Lys Leu Gly Ser Asp Val Ala Thr Leu Phe Thr His Ser His
 155 160 165
 Val Val Ser Leu Val Ser Ala Ala Ala Cys Leu Leu Gly Gln Leu
 170 175 180
 35 Gly Gln Gln Gly Val Thr Phe Asp Leu Gln Pro Met Glu Trp Met
 185 190 195
 40 Ala Ala Ala Thr His Ala Leu Ser Ala Pro Ala Glu Leu Leu Thr
 200 205 210
 Glu Val Gln Met Asp Leu Gly Met Asp Gly Lys
 215 220
 45
 <210> 23
 <211> 795
 <212> PRT
 <213> *Drosophila virilis*
 50
 <400> 23
 Met Asp Arg Tyr Ala Val Ser Ser Leu Val Gly Gln Gly Ser Phe
 1 5 10 15
 55 Gly Cys Val Tyr Lys Ala Gln Arg Arg Asp Asp Asp Lys Val Val
 20 25 30
 Ala Ile Lys Val Ile Ser Lys Arg Gly Arg Ser Asn Arg Glu Leu
 35 40 45
 60 Lys Asn Leu Arg Arg Glu Cys Asp Ile Gln Ala Arg Leu Lys His
 50 55 60
 Pro His Val Ile Glu Met Val Glu Ser Phe Glu Ser Lys Phe Asp
 65 70 75
 Leu Phe Val Val Thr Glu Phe Ala Leu Met Asp Leu His Arg Tyr
 80 85 90

ES 2 371 853 T3

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

ES 2 371 853 T3

				440					445				450
				Glu Met Leu Thr	Gln Gln Ser Gln Ala	Gln Gln Leu Gln Asp	Arg						
5				455					460				465
				Lys Thr Gln Gln	Leu Lys Gln Ser Met	His Ser Thr Asn Asp	Glu						
				470					475				480
10				Lys Leu Ser Ser	Asp Asn Ser Pro Pro	Cys Leu Leu Pro Gly	Trp						
				485					490				495
				Asp Ser Cys Asp	Glu Ser Gln Ser Pro	Pro Ile Glu Asn Asp	Glu						
				500					505				510
15				Trp Leu Ala Phe	Leu His Arg Ser Ile	Gln Glu Leu Leu Asp	Gly						
				515					520				525
				Glu Phe Asp Ser	Leu Lys Gln His Asn	Leu Val Ser Ile Ile	Val						
				530					535				540
20				Ala Pro Leu Arg	Asn Ser Lys Ala Ile	Pro Lys Val Leu Gln	Ser						
				545					550				555
				Val Ala Gln Leu	Leu Ser Leu Pro Phe	Val Leu Ala Glu Gln	His						
				560					565				570
25				Leu Val Ala Glu	Ala Ile Lys Gly Val	Tyr Ile Asp Val Lys	Leu						
				575					580				585
30				Val Pro Asn Leu	Met Tyr Ala Cys Lys	Leu Leu Leu Ser Gln	Arg						
				590					595				600
				His Leu Thr Asp	Ser Ala Ala Ser Leu	Pro Ala Gly Thr Gly	Val						
				605					610				615
35				Ser Leu Ser Arg	Thr Val Arg Ser Cys	Ser Asp Leu Ser Ala	Glu						
				620					625				630
40				Glu Met Ser Thr	Ala Cys Ser Leu Tyr	Glu Leu Val Cys His	Leu						
				635					640				645
				Val His Gln Gln	Gln Gln Phe Leu Thr	Gln Phe Cys Asp Ala	Val						
				650					655				660
45				Ala Ile Leu Ala	Val Asn Asp Met Phe	Ile Asn Phe Leu Thr	His						
				665					670				675
				Asp Phe Lys Asp	Ser Arg Pro Val Arg	Leu Ala Ser Cys Met	Leu						
				680					685				690
50				Ala Leu Phe Cys	Cys Val Leu Arg Glu	Leu Pro Glu Asn Ala	Glu						
				695					700				705
				Leu Val Glu Lys	Ile Val Phe Asp Ser	Arg Leu Gln Leu Ala	Val						
				710					715				720
55				Leu Leu Gln Ser	Arg His His Leu Leu	Arg Gln Arg Ala Cys	Gln						
				725					730				735
60				Met Leu Leu Leu	Leu Ala Arg Phe Ser	Leu Arg Gly Val Gln	Cys						
				740					745				750
				Ile Trp Ser Gly	Glu Leu Lys Ser Ala	Leu Gln Ala Trp Pro	Met						
				755					760				765
65				Gln Gln Thr Cys	Gln Ser Leu Arg Leu	Glu Ala Ala Gln Thr	Leu						
				770					775				780
70				Asp Glu Leu Ser	Gln Phe Ser Phe Phe	Val Ala Gln Ala Thr	Ala						
				785					790				795

ES 2 371 853 T3

```

<210> 24
<211> 260
<212> PRT
5  <213> Homo sapiens

<400> 24
Met Glu Lys Tyr His Val Leu Glu Met Ile Gly Glu Gly Ser Phe
 1          5          10          15
10 Gly Arg Val Tyr Lys Gly Arg Arg Lys Tyr Ser Ala Gln Val Val
   20          25          30
15 Ala Leu Lys Phe Ile Pro Lys Leu Gly Arg Ser Glu Lys Glu Leu
   35          40          45
   Arg Asn Leu Gln Arg Glu Ile Glu Ile Met Arg Gly Leu Arg His
   50          55          60
20 Pro Asn Ile Val His Met Leu Asp Ser Phe Glu Thr Asp Lys Glu
   65          70          75
   Val Val Val Val Thr Asp Tyr Ala Glu Gly Glu Leu Phe Gln Ile
   80          85          90
25 Leu Glu Asp Asp Gly Lys Leu Pro Glu Asp Gln Val Gln Ala Ile
   95          100          105
   Ala Ala Gln Leu Val Ser Ala Leu Tyr Tyr Leu His Ser His Arg
   110          115          120
   Ile Leu His Arg Asp Met Lys Pro Gln Asn Ile Leu Leu Ala Lys
   125          130          135
35 Gly Gly Gly Ile Lys Leu Cys Asp Phe Gly Phe Ala Arg Ala Met
   140          145          150
   Ser Thr Asn Thr Met Val Leu Thr Ser Ile Lys Gly Thr Pro Leu
   155          160          165
40 Tyr Met Ser Pro Glu Leu Val Glu Glu Arg Pro Tyr Asp His Thr
   170          175          180
   Ala Asp Leu Trp Ser Val Gly Cys Ile Leu Tyr Glu Leu Ala Val
   185          190          195
   Gly Thr Pro Pro Phe Tyr Ala Thr Ser Ile Phe Gln Leu Val Ser
   200          205          210
50 Leu Ile Leu Lys Asp Pro Val Arg Trp Pro Ser Thr Ile Ser Pro
   215          220          225
   Cys Phe Lys Asn Phe Leu Gln Gly Leu Leu Thr Lys Asp Pro Arg
   230          235          240
55 Gln Arg Leu Ser Trp Pro Asp Leu Leu Tyr His Pro Phe Ile Ala
   245          250          255
   Gly His Val Thr Ile
60          260

<210> 25
<211> 1315
<212> PRT
65 <213> secuencia artificial

<220>
<223> secuencia artificial

```

ES 2 371 853 T3

<400> 25
Met Glu Lys Tyr His Val Leu Glu Met Ile Gly Glu Gly Ser Phe
1 5 10 15
5 Gly Arg Val Tyr Lys Gly Arg Arg Lys Tyr Ser Ala Gln Val Val
20 25 30
10 Ala Leu Arg Phe Ile Pro Lys Leu Gly Arg Ser Glu Lys Glu Leu
35 40 45
Arg Asn Leu Gln Arg Glu Ile Glu Ile Met Arg Gly Leu Arg His
50 55 60
15 Pro Asn Ile Val His Met Leu Asp Ser Phe Glu Thr Asp Lys Glu
65 70 75
Val Val Val Val Thr Asp Tyr Ala Glu Gly Glu Leu Phe Gln Ile
80 85 90
20 Leu Glu Asp Asp Gly Lys Leu Pro Glu Asp Gln Val Gln Ala Ile
95 100 105
25 Ala Ala Gln Leu Val Ser Ala Leu Tyr Tyr Leu His Ser His Arg
110 115 120
Ile Leu His Arg Asp Met Lys Pro Gln Asn Ile Leu Leu Ala Lys
125 130 135
30 Gly Gly Gly Ile Lys Leu Cys Asp Phe Gly Phe Ala Arg Ala Met
140 145 150
Ser Thr Asn Thr Met Val Leu Thr Ser Ile Lys Gly Thr Pro Leu
155 160 165
35 Tyr Met Ser Pro Glu Leu Val Glu Glu Arg Pro Tyr Asp His Thr
170 175 180
40 Ala Asp Leu Trp Ser Val Gly Cys Ile Leu Tyr Glu Leu Ala Val
185 190 195
Gly Thr Pro Pro Phe Tyr Ala Thr Ser Ile Phe Gln Leu Val Ser
200 205 210
45 Leu Ile Leu Lys Asp Pro Val Arg Trp Pro Ser Thr Ile Ser Pro
215 220 225
Cys Phe Lys Asn Phe Leu Gln Gly Leu Leu Thr Lys Asp Pro Arg
230 235 240
50 Gln Arg Leu Ser Trp Pro Asp Leu Leu Tyr His Pro Phe Ile Ala
245 250 255
55 Gly His Val Thr Ile Ile Thr Glu Pro Ala Gly Pro Asp Leu Gly
260 265 270
Thr Pro Phe Thr Ser Arg Leu Pro Pro Glu Leu Gln Val Leu Lys
275 280 285
60 Asp Glu Gln Ala His Arg Leu Ala Pro Lys Gly Asn Gln Ser Arg
290 295 300
Ile Leu Thr Gln Ala Tyr Lys Arg Met Ala Glu Glu Ala Met Gln
305 310 315
65 Lys Lys His Gln Asn Thr Gly Pro Ala Leu Glu Gln Glu Asp Lys
320 325 330
70 Thr Ser Lys Val Ala Pro Gly Thr Ala Pro Leu Pro Arg Leu Gly
335 340 345

ES 2 371 853 T3

	Ala	Thr	Pro	Gln	Glu 350	Ser	Ser	Leu	Leu	Ala 355	Gly	Ile	Leu	Ala	Ser 360
5	Glu	Leu	Lys	Ser	Ser 365	Trp	Ala	Lys	Ser	Gly 370	Thr	Gly	Glu	Val	Pro 375
	Ser	Ala	Pro	Arg	Glu 380	Asn	Arg	Thr	Thr	Pro 385	Asp	Cys	Glu	Arg	Ala 390
10	Phe	Pro	Glu	Glu	Arg 395	Pro	Glu	Val	Leu	Gly 400	Gln	Arg	Ser	Thr	Asp 405
15	Val	Val	Asp	Leu	Glu 410	Asn	Glu	Glu	Pro	Asp 415	Ser	Asp	Asn	Glu	Trp 420
	Gln	His	Leu	Leu	Glu 425	Thr	Thr	Glu	Pro	Val 430	Pro	Ile	Gln	Leu	Lys 435
20	Ala	Pro	Leu	Thr	Leu 440	Leu	Cys	Asn	Pro	Asp 445	Phe	Cys	Gln	Arg	Ile 450
	Gln	Ser	Gln	Leu	His 455	Glu	Ala	Gly	Gly	Gln 460	Ile	Leu	Lys	Gly	Ile 465
25	Leu	Glu	Gly	Ala	Ser 470	His	Ile	Leu	Pro	Ala 475	Phe	Arg	Val	Leu	Ser 480
30	Ser	Leu	Leu	Ser	Ser 485	Cys	Ser	Asp	Ser	Val 490	Ala	Leu	Tyr	Ser	Phe 495
	Cys	Arg	Glu	Ala	Gly 500	Leu	Pro	Gly	Leu	Leu 505	Leu	Ser	Leu	Leu	Arg 510
35	His	Ser	Gln	Glu	Ser 515	Asn	Ser	Leu	Gln	Gln 520	Gln	Ser	Trp	Tyr	Gly 525
	Thr	Phe	Leu	Gln	Asp 530	Leu	Met	Ala	Val	Ile 535	Gln	Ala	Tyr	Phe	Ala 540
40	Cys	Thr	Phe	Asn	Leu 545	Glu	Arg	Ser	Gln	Thr 550	Ser	Asp	Ser	Leu	Gln 555
45	Val	Phe	Gln	Glu	Ala 560	Ala	Asn	Leu	Phe	Leu 565	Asp	Leu	Leu	Gly	Lys 570
	Leu	Leu	Ala	Gln	Pro 575	Asp	Asp	Ser	Glu	Gln 580	Thr	Leu	Arg	Arg	Asp 585
50	Ser	Leu	Met	Cys	Phe 590	Thr	Val	Leu	Cys	Glu 595	Ala	Met	Asp	Gly	Asn 600
	Ser	Arg	Ala	Ile	Ser 605	Lys	Ala	Phe	Tyr	Ser 610	Ser	Leu	Leu	Thr	Thr 615
55	Gln	Gln	Val	Val	Leu 620	Asp	Gly	Leu	Leu	His 625	Gly	Leu	Thr	Val	Pro 630
60	Gln	Leu	Pro	Val	His 635	Thr	Pro	Gln	Gly	Ala 640	Pro	Gln	Val	Ser	Gln 645
	Pro	Leu	Arg	Glu	Gln 650	Ser	Glu	Asp	Ile	Pro 655	Gly	Ala	Ile	Ser	Ser 660
65	Ala	Leu	Ala	Ala	Ile 665	Cys	Thr	Ala	Pro	Val 670	Gly	Leu	Pro	Asp	Cys 675
70	Trp	Asp	Ala	Lys	Glu 680	Gln	Val	Cys	Trp	His 685	Leu	Ala	Asn	Gln	Leu 690

ES 2 371 853 T3

	Thr	Glu	Asp	Ser	Ser 695	Gln	Leu	Arg	Pro	Ser 700	Leu	Ile	Ser	Gly	Leu 705
5	Gln	His	Pro	Ile	Leu 710	Cys	Leu	His	Leu	Leu 715	Lys	Val	Leu	Tyr	Ser 720
	Cys	Cys	Leu	Val	Ser 725	Glu	Gly	Leu	Cys	Arg 730	Leu	Leu	Gly	Gln	Glu 735
10	Pro	Leu	Ala	Leu	Glu 740	Ser	Leu	Phe	Met	Leu 745	Ile	Gln	Gly	Lys	Val 750
	Lys	Val	Val	Asp	Trp 755	Glu	Glu	Ser	Thr	Glu 760	Val	Thr	Leu	Tyr	Phe 765
15	Leu	Ser	Leu	Leu	Val 770	Phe	Arg	Leu	Gln	Asn 775	Leu	Pro	Cys	Gly	Met 780
20	Glu	Lys	Leu	Gly	Ser 785	Asp	Val	Ala	Thr	Leu 790	Phe	Thr	His	Ser	His 795
	Val	Val	Ser	Leu	Val 800	Ser	Ala	Ala	Ala	Cys 805	Leu	Leu	Gly	Gln	Leu 810
25	Gly	Gln	Gln	Gly	Val 815	Thr	Phe	Asp	Leu	Gln 820	Pro	Met	Glu	Trp	Met 825
	Ala	Ala	Ala	Thr	His 830	Ala	Leu	Ser	Ala	Pro 835	Ala	Glu	Val	Arg	Leu 840
30	Thr	Pro	Pro	Gly	Ser 845	Cys	Gly	Phe	Tyr	Asp 850	Gly	Leu	Leu	Ile	Leu 855
35	Leu	Leu	Gln	Leu	Leu 860	Thr	Glu	Gln	Gly	Lys 865	Ala	Ser	Leu	Ile	Arg 870
	Asp	Met	Ser	Ser	Ser 875	Glu	Met	Trp	Thr	Val 880	Leu	Trp	His	Arg	Phe 885
40	Ser	Met	Val	Leu	Arg 890	Leu	Pro	Glu	Glu	Ala 895	Ser	Ala	Gln	Glu	Gly 900
	Glu	Leu	Ser	Leu	Ser 905	Ser	Pro	Pro	Ser	Pro 910	Glu	Pro	Asp	Trp	Thr 915
45	Leu	Ile	Ser	Pro	Gln 920	Gly	Met	Ala	Ala	Leu 925	Leu	Ser	Leu	Ala	Met 930
50	Ala	Thr	Phe	Thr	Gln 935	Glu	Pro	Gln	Leu	Cys 940	Leu	Ser	Cys	Leu	Ser 945
	Gln	His	Gly	Ser	Ile 950	Leu	Met	Ser	Ile	Leu 955	Lys	His	Leu	Leu	Cys 960
55	Pro	Ser	Phe	Leu	Asn 965	Gln	Leu	Arg	Gln	Ala 970	Pro	His	Gly	Ser	Glu 975
	Phe	Leu	Pro	Val	Val 980	Val	Leu	Ser	Val	Cys 985	Gln	Leu	Leu	Cys	Phe 990
60	Pro	Phe	Ala	Leu	Asp 995	Met	Asp	Ala	Asp	Leu 1000	Leu	Ile	Val	Val	Leu 1005
65	Ala	Asp	Leu	Arg	Asp 1010	Ser	Glu	Val	Ala	Ala 1015	His	Leu	Leu	Gln	Val 1020
	Cys	Cys	Tyr	His	Leu 1025	Pro	Leu	Met	Gln	Val 1030	Glu	Leu	Pro	Ile	Ser 1035
70	Leu	Leu	Thr	Arg	Leu	Ala	Leu	Met	Asp	Pro	Thr	Ser	Leu	Asn	Gln

ES 2 371 853 T3

1040 1045 1050
 Phe Val Asn Thr Val Ser Ala Ser Pro Arg Thr Ile Val Ser Phe
 1055 1060 1065
 5 Leu Ser Val Ala Leu Leu Ser Asp Gln Pro Leu Leu Thr Ser Asp
 1070 1075 1080
 10 Leu Leu Ser Leu Leu Ala His Thr Ala Arg Val Leu Ser Pro Ser
 1085 1090 1095
 His Leu Ser Phe Ile Gln Glu Leu Leu Ala Gly Ser Asp Glu Ser
 1100 1105 1110
 15 Tyr Arg Pro Leu Arg Ser Leu Leu Gly His Pro Glu Asn Ser Val
 1115 1120 1125
 Arg Ala His Thr Tyr Arg Leu Leu Gly His Leu Leu Gln His Ser
 1130 1135 1140
 20 Met Ala Leu Arg Gly Ala Leu Gln Ser Gln Ser Gly Leu Leu Ser
 1145 1150 1155
 25 Leu Leu Leu Leu Gly Leu Gly Asp Lys Asp Pro Val Val Arg Cys
 1160 1165 1170
 Ser Ala Ser Phe Ala Val Gly Asn Ala Ala Tyr Gln Ala Gly Pro
 1175 1180 1185
 30 Leu Gly Pro Ala Leu Ala Ala Ala Val Pro Ser Met Thr Gln Leu
 1190 1195 1200
 Leu Gly Asp Pro Gln Ala Gly Ile Arg Arg Asn Val Ala Ser Ala
 1205 1210 1215
 35 Leu Gly Asn Leu Gly Pro Glu Gly Leu Gly Glu Glu Leu Leu Gln
 1220 1225 1230
 40 Cys Glu Val Pro Gln Arg Leu Leu Glu Met Ala Cys Gly Asp Pro
 1235 1240 1245
 Gln Pro Asn Val Lys Glu Ala Ala Leu Ile Ala Leu Arg Ser Leu
 1250 1255 1260
 45 Gln Gln Glu Pro Gly Ile His Gln Val Leu Val Ser Leu Gly Ala
 1265 1270 1275
 Ser Glu Lys Leu Ser Leu Leu Ser Leu Gly Asn Gln Ser Leu Pro
 1280 1285 1290
 50 His Ser Ser Pro Arg Pro Ala Ser Ala Lys His Cys Arg Lys Leu
 1295 1300 1305
 55 Ile His Leu Leu Arg Pro Ala His Ser Met
 1310 1315

<210> 26
 <211> 4586
 <212> DNA
 60 <213> *Drosophila virilis*

<400> 26
 tgcagagtct gggccatcgg ctagctctgt agatgtgtaa tagaggcatc 50
 65 ttcgcgcgca gcatcgattc gcgctccagt tggttgagat gccgaatgtt 100
 ggtgcgcgtc tcaaagatga cgccgacatg gtgcatatac agaaaaaaga 150
 aacccgagcc taatggccgc gtattatcgc tgaggcggcc ggcgattttc 200

aacaaatgct acttaccaat tagcgcgtgc gataaaatta cgtacaaatt 250
 5 ggcgtcgcgc atatttcttg cgtttggtt gcgtgtatct agttagtggg 300
 ctctcttatt cattatttac ttttggggcg tctgtttgat caaattagca 350
 gtgtctccta tgatatgcct gcagctttta cccgtaaaca aaattatttg 400
 10 ccacagctga tttattcggt gccatgtaga ttaatcagct gtttcgcaat 450
 ttttaaaacc aggtgacttt ttaaaattgt accagctgtg tgtatcgatg 500
 tgcaagcata ctattacgc catatgctgg tatatttata tcgaatataa 550
 15 acggttttggt tattttaata atcttaaaga agaaatagtt atgtgctgtg 600
 tatatgttca tcaagaactg ttcaaaatgt gcgccatact gatgttaatt 650
 20 ttgttttgct gggttttttt gggaaaataa attgacgtgt tgatgtctcc 700
 gaatatatcg atacaatagc tatcattcgg acaagatatc gatatgtgga 750
 gtgtgttcgg tattttgcct ttagtttttt gtttttaaatt tgcagtcaca 800
 25 ctgcggctta ttgaatttaa ggcacttcaa agcgcatttt actgtagaaa 850
 gttgagttct atttgcggtg acaatggacc gctacgcggt tagctctctg 900
 30 gtaggacaag gctcatttgg ctgtgtgtac aaggcccagc ggcgcgatga 950
 tgacaaagtg gtggccatca aagtcatatc aaaggtagc tcaattgcat 1000
 35 cccggcttag ctgaataaaa gagtattcta cgaattggcg tgttctttgt 1050
 ttgcagcgtg gtcgttccaa tcgcgagctt aagaacctgc gtcgcgaatg 1100
 tgacattcag gcgcgtctca agcatccgca tgttatagaa atggtggagt 1150
 40 cgttcgaatc caaatcgtat ttgttcgtgg tcaccgagtt cgctctaattg 1200
 gacttgcatac gatatttgtc ctttaattggc gccatgcccg aggagcacgc 1250
 45 acagcgtggt gtctgtcatt tgggtgcggc gctctattat ctgcactcga 1300
 atcgcatact gcatcgggat cttaaagccgc aaaatgtgct gttggacaaa 1350
 aacatgcacg ccaagctctg cgactttggg ctggcacgca acatgacgat 1400
 50 gggcacacat gtgttgactt ccataaaggg cacgccgctt tatatggcgc 1450
 cggagctgct ggctgagcag ccgtacgatc accaggcaga tatgtggctg 1500
 55 ctgggatgca ttgcctatga gagtatggcg ggccagccgc cgttctgcgc 1550
 aacctctata ctgcatctgg tgaagctgat caagcacgag gacgtcaaatt 1600
 ggccgagcac gctgagcagc gagtgccgtt cttttttgca gggcttgctc 1650
 60 gagaaggatc ctagcatgcg catctcatgg acgcagctgc tttgccatcc 1700
 ttttgtcgag ggcaagctat acatagccga ggtacaggca gcacaaactt 1750
 65 cgccctttat aaatccccag ctggccaagg acacaaaaa atcacagcaa 1800
 ttgagggtgcg tttataacgt gtactgtagc cagctccact tatcgttcaa 1850
 tttttatgta ggcatgtagg cgagatttg ggcgatgtct tggcagcggt 1900
 70 aaagttgagc gatgtggcca atgaaaactt gagcacatcg cgagatagta 1950

ES 2 371 853 T3

tcaatgccat tgcgccgagt gacattgagc agctggaaac cgatgttgag 2000
 5 gataatgtgc atcggcttat agtgccatth gcagatatth cctacagaga 2050
 gttgccatgc ggactgcag cagctgctcg tcgagctggt gccatgccac 2100
 tgattaattc gcaaacctgc tttgtaagtgc gcaactccaa tatgatactc 2150
 10 aatcatctga acgacaatth tgcaatcgaa gcgcctgctt cgagcgcaac 2200
 caagtccatg aagtcgaagc tgaagctggc tctcaatata aaacagtcgc 2250
 15 gtagcaagga tttggaaaag cgtaagctga gtcaaaatth ggataactth 2300
 tcgtgcgccc tgggacagag cattgacata gaagtgcagc gaaaacaac 2350
 tgagatgctc acgcagcaat cgcaggcaca acagctgcag gataggaaga 2400
 20 cacagcagct gaagcaatcg atgcattcca ccaacgacga gaaattgagc 2450
 agcgagtgcg taaatgcata catatthaaa agtgaagctc tctaaagcta 2500
 25 tttggthtat aatagcaatt cgccgccttg tctgttgccc ggttgggaca 2550
 gctgcgatga atctcagagc ccgcccattg agaatgacga gtggctggcg 2600
 tcttgcatc gctccataca ggagctgctg gacggcgaat ttgattcgct 2650
 30 gaagcagcac aatctagtca gcataattgt ggcgccattg cgaaactcca 2700
 aggcataacc caaggctgctg cagagcgctgg cgagctgct gtcgctgccc 2750
 35 tttgtgctgg ccgaacagca tttggtagcg gaggccataa aaggagthta 2800
 tattgatgtc aagctggctgc ccaacttaat gtacgcctgc aagctgcttc 2850
 tctgcgagcg ccaccttacc gattcggtg cttactgcc agccggcacg 2900
 40 ggcgtctccc tgagtcgaac cgtacgcagc tgctccgacc tgagtgcga 2950
 ggagatgagc accgcctgca gcctgtacga gctggctctgc catctggtcc 3000
 45 atcagcagca gcagttctc accagttct gtgacgctgt ggcaatactc 3050
 gccgtcaacg acatgttcat aaatthtctt acacatggtg agcagctggc 3100
 tggacacagt gtgagacgca agcttaacca ttccttgctt tgcatatth 3150
 50 aaggatagca ggccggtgcg actcgctagc tgcatgctgg cattgttctg 3200
 ttgcgtthtg cgtgaactac ccgagaacgc cgagctggtg gagaaaattg 3250
 55 tathtgactc gcgcctacag ctggccgtcc tgctgcagag ccgtcatcat 3300
 ttgttgctgc agcgcgctg tcaaatgctg ttgctattgg cacgctthg 3350
 cctgcgcggc gtacagtgc tctggagtgg ggagctgaag agtgcgctcc 3400
 60 aggcgtggcc gatgcagcaa acgtgtcaat cattgcgact ggaagccgcc 3450
 caaacgctgg atgagcttag ccagttcagc tctthtggc ctcaggcaac 3500
 65 tgcttagtct ttattaataa ttgtacttgt atthgtthaa taaatctth 3550
 tccttgctc gccgaacaga cttccaaat tgccttgaaa gtagtcgagc 3600
 agctcgcca gatagctgct aaagccatca aagcccaaaa gtagctacc 3650
 70 attacagtcc tgctcgata tctcgthg tthcgaaata tccttatccg 3700

ES 2 371 853 T3

acagccgcgc gctggcccag tgaggcatac ggatcctaata gataatagca 3750
 5 tgcattatta ttatttttca caatgtgtta ttcgtttaat acttataaaa 3800
 ccttaaattg tatgcatgta tgtatctatc ttatacctaa ttaatgaatg 3850
 aaattttatta acttgtctat ggatgtatgt gcatgtatgt atgtatgtat 3900
 10 gtatgcataa aaatgtatgt tcattttataa caaacgcaga caaagataac 3950
 gatctgctgc tctacttccc gaatctcata aattcaagta cgccccgcag 4000
 atttcacgag tacatcacaa gtgttttttt ttaacaagta atgttggtat 4050
 15 gtattttatgt atatatgtat ttaagtatgt atgtatttat gtatgtatgt 4100
 atgtatttat gtatgtatgt atttatgtat gtatgtattt atgtatgtat 4150
 20 gtattttatgt atgtatttat ttatgtatat atgtatttaa gtatgtatgt 4200
 atgtatttat gcatttatgt atttatgtat gtatgtataa gagtatgtgt 4250
 gtgtgtagat acatgtatgt atgtatgtat gtatgcgtgt atattttattt 4300
 25 atagtaaaca taccaacttt acttcccgcct gccttgcgaa tttaaaataa 4350
 cgtattttta aatgatgccc tactcctcga ttctcaaaca tttaaagtaag 4400
 30 ctctacaggt ttttccgatt tgattgtttt gtaaagttgt gttttttttt 4450
 ctgctcgatc tcttgtgtat tctctactct ttgtgtgcct ctcttttagtt 4500
 35 ttctctcctt ctctcttgct ctccccgttt ctctctctat ctctctccct 4550
 ccctctttcc acctatctca ttctctttct aagctt 4586

<210> 27

<211> 1055

40 <212> PRT

<213> Homo sapiens

<400> 27

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

ES 2 371 853 T3

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

ES 2 371 853 T3

	Ser	Leu	Phe	Met	Leu 485	Ile	Gln	Gly	Lys	Val 490	Lys	Val	Val	Asp	Trp 495
5	Glu	Glu	Ser	Thr	Glu 500	Val	Thr	Leu	Tyr	Phe 505	Leu	Ser	Leu	Leu	Val 510
	Phe	Arg	Leu	Gln	Asn 515	Leu	Pro	Cys	Gly	Met 520	Glu	Lys	Leu	Gly	Ser 525
10	Asp	Val	Ala	Thr	Leu 530	Phe	Thr	His	Ser	His 535	Val	Val	Ser	Leu	Val 540
	Ser	Ala	Ala	Ala	Cys 545	Leu	Leu	Gly	Gln	Leu 550	Gly	Gln	Gln	Gly	Val 555
15	Thr	Phe	Asp	Leu	Gln 560	Pro	Met	Glu	Trp	Met 565	Ala	Ala	Ala	Thr	His 570
20	Ala	Leu	Ser	Ala	Pro 575	Ala	Glu	Val	Arg	Leu 580	Thr	Pro	Pro	Gly	Ser 585
	Cys	Gly	Phe	Tyr	Asp 590	Gly	Leu	Leu	Ile	Leu 595	Leu	Leu	Gln	Leu	Leu 600
25	Thr	Glu	Gln	Gly	Lys 605	Ala	Ser	Leu	Ile	Arg 610	Asp	Met	Ser	Ser	Ser 615
	Glu	Met	Trp	Thr	Val 620	Leu	Trp	His	Arg	Phe 625	Ser	Met	Val	Leu	Arg 630
30	Leu	Pro	Glu	Glu	Ala 635	Ser	Ala	Gln	Glu	Gly 640	Glu	Leu	Ser	Leu	Ser 645
	Ser	Pro	Pro	Ser	Pro 650	Glu	Pro	Asp	Trp	Thr 655	Leu	Ile	Ser	Pro	Gln 660
35	Gly	Met	Ala	Ala	Leu 665	Leu	Ser	Leu	Ala	Met 670	Ala	Thr	Phe	Thr	Gln 675
40	Glu	Pro	Gln	Leu	Cys 680	Leu	Ser	Cys	Leu	Ser 685	Gln	His	Gly	Ser	Ile 690
	Leu	Met	Ser	Ile	Leu 695	Lys	His	Leu	Leu	Cys 700	Pro	Ser	Phe	Leu	Asn 705
45	Gln	Leu	Arg	Gln	Ala 710	Pro	His	Gly	Ser	Glu 715	Phe	Leu	Pro	Val	Val 720
	Val	Leu	Ser	Val	Cys 725	Gln	Leu	Leu	Cys	Phe 730	Pro	Phe	Ala	Leu	Asp 735
50	Met	Asp	Ala	Asp	Leu 740	Leu	Ile	Val	Val	Leu 745	Ala	Asp	Leu	Arg	Asp 750
55	Ser	Glu	Val	Ala	Ala 755	His	Leu	Leu	Gln	Val 760	Cys	Cys	Tyr	His	Leu 765
	Pro	Leu	Met	Gln	Val 770	Glu	Leu	Pro	Ile	Ser 775	Leu	Leu	Thr	Arg	Leu 780
60	Ala	Leu	Met	Asp	Pro 785	Thr	Ser	Leu	Asn	Gln 790	Phe	Val	Asn	Thr	Val 795
	Ser	Ala	Ser	Pro	Arg 800	Thr	Ile	Val	Ser	Phe 805	Leu	Ser	Val	Ala	Leu 810
65	Leu	Ser	Asp	Gln	Pro 815	Leu	Leu	Thr	Ser	Asp 820	Leu	Leu	Ser	Leu	Leu 825
70	Ala	His	Thr	Ala	Arg	Val	Leu	Ser	Pro	Ser	His	Leu	Ser	Phe	Ile

ES 2 371 853 T3

					830					835					840				
					Gln	Glu	Leu	Leu	Ala	Gly	Ser	Asp	Glu	Ser	Tyr	Arg	Pro	Leu	Arg
					845									850					855
5					Ser	Leu	Leu	Gly	His	Pro	Glu	Asn	Ser	Val	Arg	Ala	His	Thr	Tyr
					860									865					870
					Arg	Leu	Leu	Gly	His	Leu	Leu	Gln	His	Ser	Met	Ala	Leu	Arg	Gly
10					875									880					885
					Ala	Leu	Gln	Ser	Gln	Ser	Gly	Leu	Leu	Ser	Leu	Leu	Leu	Leu	Gly
					890									895					900
15					Leu	Gly	Asp	Lys	Asp	Pro	Val	Val	Arg	Cys	Ser	Ala	Ser	Phe	Ala
					905									910					915
					Val	Gly	Asn	Ala	Ala	Tyr	Gln	Ala	Gly	Pro	Leu	Gly	Pro	Ala	Leu
20					920									925					930
					Ala	Ala	Ala	Val	Pro	Ser	Met	Thr	Gln	Leu	Leu	Gly	Asp	Pro	Gln
					935									940					945
					Ala	Gly	Ile	Arg	Arg	Asn	Val	Ala	Ser	Ala	Leu	Gly	Asn	Leu	Gly
25					950									955					960
					Pro	Glu	Gly	Leu	Gly	Glu	Glu	Leu	Leu	Gln	Cys	Glu	Val	Pro	Gln
					965									970					975
30					Arg	Leu	Leu	Glu	Met	Ala	Cys	Gly	Asp	Pro	Gln	Pro	Asn	Val	Lys
					980									985					990
					Glu	Ala	Ala	Leu	Ile	Ala	Leu	Arg	Ser	Leu	Gln	Gln	Glu	Pro	Gly
35					995									1000					1005
					Ile	His	Gln	Val	Leu	Val	Ser	Leu	Gly	Ala	Ser	Glu	Lys	Leu	Ser
					1010									1015					1020
40					Leu	Leu	Ser	Leu	Gly	Asn	Gln	Ser	Leu	Pro	His	Ser	Ser	Pro	Arg
					1025									1030					1035
					Pro	Ala	Ser	Ala	Lys	His	Cys	Arg	Lys	Leu	Ile	His	Leu	Leu	Arg
					1040									1045					1050
45					Pro	Ala	His	Ser	Met										
					1055														

REIVINDICACIONES

1. Ácido nucleico aislado que comprende ADN que codifica:

(i) un polipéptido *fused* de vertebrado que tiene por lo menos un 80% de identidad en la secuencia de aminoácidos con la secuencia de aminoácidos 1 a 1315 de la Figura 1 (SEC ID NO:2); o
(ii) un fragmento de polipéptido *fused* de vertebrado que tiene los residuos de aminoácidos 1 a 260 de la Figura 1 (SEQ ID NO:24);
y en el que el ácido nucleico aislado codifica un polipéptido que tiene actividad biológica de *fused*, donde dicha actividad biológica de *fused* comprende una o más entre:

- unión a e influencia en la señalización de hedgehog;
- regulación de la patogénesis de carcinoma de células basales;
- fosforilación o modulación de la fosforilación de Gli;
- fosforilación de proteína básica de mielina, sustratos de proteínas de levadura, sustratos de péptidos sintéticos o sustratos de polímeros.

2. Ácido nucleico según la reivindicación 1(i), en el que la identidad de secuencia es de por lo menos un 85%.

3. Ácido nucleico según la reivindicación 2, en el que la identidad en la secuencia es de por lo menos un 90%.

4. Ácido nucleico según la reivindicación 3, en el que la identidad en la secuencia es de por lo menos un 95%.

5. Ácido nucleico aislado según la reivindicación 1, que comprende ADN que codifica un polipéptido *fused* de vertebrado que tiene una lisina en la posición de aminoácido 33.

6. Ácido nucleico aislado que comprende ADN que codifica el mismo polipéptido maduro codificado por el ADNc en el Depósito ATCC No. 209637, o una secuencia de nucleótidos que se hibrida al mismo bajo condiciones rigurosas y que codifica un polipéptido que tiene actividad biológica de *fused*, donde dicha actividad biológica de *fused* comprende una o más entre:

- unión a e influencia en la señalización de hedgehog;
- regulación de la patogénesis de carcinoma de células basales;
- fosforilación o modulación de la fosforilación de Gli;
- fosforilación de proteína básica de mielina, sustratos de proteínas de levadura, sustratos de péptidos sintéticos o sustratos de polímeros.

7. Vector que comprende el ácido nucleico según las reivindicaciones 1, 5, ó 6.

8. Vector según la reivindicación 7, en el que dicho ácido nucleico está unido operativamente a secuencias de control reconocidas por una célula huésped transformada con el vector.

9. Célula huésped transformada con el vector según la reivindicación 8.

10. Célula huésped según la reivindicación 9, que es de mamífero.

11. Célula huésped según la reivindicación 10, que es una célula CHO.

12. Célula huésped según la reivindicación 9, que es procariota.

13. Célula huésped según la reivindicación 12, que es *E. coli*.

14. Célula huésped según la reivindicación 9, que es una célula de levadura.

15. Célula huésped según la reivindicación 14, que es *Saccharomyces cerevisiae*.

16. Proceso para producir los polipéptidos *fused* de vertebrado que comprende cultivar la célula huésped según la reivindicación 9 bajo condiciones adecuadas para la expresión de *fused* de vertebrado y recuperar el *fused* de vertebrado del cultivo celular.

17. Polipéptido *fused* de vertebrado aislado que comprende un polipéptido que tiene por lo menos un 80% de

identidad en la secuencia de aminoácidos con los residuos de aminoácidos 1 a 1315 de la Figura 1 (SEQ ID NO:2) que tiene actividad biológica de *fused* tal como se define en la reivindicación 1.

18. Polipéptido *fused* de vertebrado según la reivindicación 17, que comprende un polipéptido que tiene por lo menos un 85% de identidad en la secuencia de aminoácidos con los residuos de aminoácidos 1 a 1315 de la figura 1 (SEQ ID NO:2) que tiene actividad biológica de *fused* tal como se define en la reivindicación 1.

19. Polipéptido *fused* de vertebrado según la reivindicación 18, que comprende un polipéptido que tiene por lo menos un 90% de identidad en la secuencia de aminoácidos con los residuos de aminoácidos 1 a 1315 de la figura 1 (SEQ ID NO:2) que tiene actividad biológica de *fused* tal como se define en la reivindicación 1.

20. Polipéptido *fused* de vertebrado según la reivindicación 19, que comprende un polipéptido que tiene por lo menos un 95% de identidad en la secuencia de aminoácidos con los residuos de aminoácidos 1 a 1315 de la figura 1 (SEQ ID NO:2) que tiene actividad biológica de *fused* tal como se define en la reivindicación 1.

21. Polipéptido *fused* de vertebrado aislado según la reivindicación 17, que comprende los residuos de aminoácidos 1 a 1315 de la figura 1 (SEQ ID NO:2).

22. Polipéptido *fused* de vertebrado aislado que comprende el polipéptido *fused* humano codificado por los nucleótidos depositados bajo el número de acceso ATCC 209637.

23. Molécula quimérica que comprende el polipéptido *fused* de vertebrado según la reivindicación 17, fusionado a una secuencia de aminoácidos heteróloga.

24. Molécula quimérica según la reivindicación 23, en la que dicha secuencia de aminoácidos heteróloga es una secuencia de epítipo etiqueta.

25. Molécula quimérica según la reivindicación 24, en la que dicha secuencia de aminoácidos heteróloga es una región constante de una inmunoglobulina.

26. Antagonista de *fused* de vertebrado que bloquea, evita, inhibe y/o neutraliza el funcionamiento normal de *fused* de vertebrado en el mecanismo de señalización de Hh, en la que dicho antagonista es un ácido nucleico antisentido, un anticuerpo anti-*fused* de vertebrado, o un mutante negativo dominante de *fused* de vertebrado; y en la que dicho *fused* de vertebrado es codificado por un ácido nucleico tal como se define en la reivindicación 1.

27. Método de cribado de antagonistas de la actividad biológica de *fused* de vertebrado tal como se define en la reivindicación 1, que comprende:

- (a) exponer las células diana que expresan *fused* en un cultivo a un compuesto candidato; y
- (b) analizar los lisados celulares para valorar el nivel y/o identidad de fosforilación; o
- (c) valorar los cambios fenotípicos o funcionales en las células tratadas;

y comparar los resultados con las células de control que no se expusieron al compuesto candidato; en el que dicho *fused* es codificado por un ácido nucleico según la reivindicación 1.

28. Método de diagnóstico para determinar si un trastorno particular está modulado por la señalización de hedgehog, que comprende:

- (a) cultivar células o tejidos de análisis;
- (b) administrar un compuesto según la reivindicación 26 que puede inhibir la señalización de hedgehog modulada por *fused*;
- (c) medir el grado de atenuación de quinasa en el sustrato de *fused* en lisados celulares o efectos fenotípicos mediados por hedgehog en las células de análisis;

en el que dicho *fused* es un polipéptido codificado por un ácido nucleico según la reivindicación 1.

Figura 1A

Longitud 480 pb (circular)

1 CCCGGGATC CTCTAGAGAT CCCTCGACCT CGACCCACGC GTCGCGCCAC GCGTCCGCC ACAGGTCCGG GCGTCCCG AGTTGTGGA ACTGTCCTG
GGGCCCCTAG GAGATCTTA GGGAGCTGGA GCTGGGTGG CAGCGGGTG CCGAGGCGG TCGCAGGCTC TACAACACCT TGACAGGAC

101 GATCTATAGC TCTTACCGT CTCTACTTTC TTCTTCTTA GAGATCCTGA AACCTCTGC ATGGAAAGT ACCACGTCTT GGAGATGATT GGAGAAAGCT
CTAGATATCG AGAAGTGGA GAGATGAAG AAGGAAGATT CTCTAGGACT TTGGAGACAG TACCTTTTCA TGGTGACAA CCTCTACTTA CCTCTTCCGA
M E K Y H V L E M I G E G S

201 CTTTGGGAG GGTGTACAG GGTGGAAGAA AATACAGTGC TCAGGTCTG GCCCTGAAGT TCATCCAAA ATTGGGGGC TCAGAGAAG AGCTGAGGAA
GAAACCTC CCACATGTTT CCAGCTTCTT TTATGTACG ATCCAGCAC CGGACTTCA AGTAGGTTT TAACCCCGG AGTCTCTTCC TCGACTCTT
15 F G R V Y K G R R K Y S A Q V V A L K F I P K L G R S E K E L R N

301 TTTGACGA GATTTGAA TAATGGGGG TCTGGGGCAT CCAACATTG TGCATATGCT TGACAGCTT GAAACTGATA AAGAGGTGT GGTGGTGA
AAACGTTGCT CTCTAATTT ATTACGCCCC AGACGCCGTA GGGTTGTAAC ACGTATACGA ACTGTGAAA CTTTGACTAT TTCTCCACCA CCACCACTGT
48 L Q R E I E I M R G L R H P N I V H M L D S F E T D K E V V V V T

401 GACTATGCT AGGGAGCT CTTTCAGATC CTAGAAGATG ACGGAAACT TCTGAAGAC CAGTTTCAGG CCATTGTGC CCAGTTGGTG TCAGCCCTGT
CTGATAGGAC TCCCTCTCGA GAAAGTCTAG GATCTTCTAC TGCTTTTGA AGGACTTCTG GTCCAAGTCC GGTAAAGGAG GTCAACCCAC AGTCGGGACA
81 D Y A E G E L F Q I L E D D G K L P E D Q V Q A I A A Q L V S A L Y

501 ACTATCTGA TTCCACCGC ATCTACACC GAGATATGAA GCTCAGAAC ATCTCTCTG CCAAGGGTG TGGCATCAAG CTCTGTGACT TTGGATTGCT
TGATAGACGT AAGGTGGCG TAGGATGTGG CTCTATACCT CGAGTCTTG TAGGAGGAGC GGTTCACAC ACCGTAGTTC GAGACACTGA AACCTAAACG
115 Y L H S H R I L H R D M K P Q N I L L A K G G I K L C D F G F A

601 CCGGGCTATG AGCACCATA CAATGGTGT GACATCATC AAAGGCACAC CACTCTATAT GTCTCCAGAG CTGGTGGAG AGGACCATTA CGACCACACA
GGCCCGATAC TCGTGGTAT GTTACCACGA CTGTAGTAG TTTCCGTGT GTGAGATATA CAGAGTCTC GACCACCTCC TCGCTGGTAT GCTGGTGT
148 R A M S T N T M V L T S I K G T P L Y M S P E L V E E R P Y D H T

701 GCGGACCTCT GGTCTGTGG CTCATACCTA TATGAAGTGG CAGTAGGCAC CCTCCCTTC TATGCTACAA GCATCTTCA GCTGGTCAAG CTCATTTCTCA
CGCCTGGAGA CCAGACRACC GAGGTATGAT ATACTTGACC GTCATCCGTG GGGAGGGAAG ATACGATGTT CGTAGAAGT GCACCAGTCG GAGTAAGAT
181 A D L W S V G C I L Y E L A V G T P P F Y A T S I F Q L V S L I L K

Figura 1B

801 AGGACCTGT GCGTGGGCC TCAACATCA GTCCCTGCTT TAAGACTTC CTGCAGGAC TGCTACCAA AGACCCACGG CAGCAGCTGT CCTGGCCAGA
 TCCCTGGACA CGCGACGGG AGTTGTAGT CAGGACGAA ATTCTGAAG GAGTCCTG ACAGTGGTT TCTGGTGCC GTCGTGACA GGACCGGTCT
 215 D P V R W P S T I S P C F K N F L O G L L T K D P R Q R L S W P D
 901 CCTCTATAT CACCCCTTTA TTGCTGTCA TGTCACATA ATAAGTACG CAGCAGGCC AGATTGGGG ACCCATTTCA CCAGCCGCTT ACCCCAGAA
 GGAGATATA GTGGGAAAT AACGACAGT ACAGTGAT TATTGACTG GTGCTCGGG TCTAACCCTT TGGGTAAAT GGTGGGCGGA TGGGGTCTT
 248 L L Y H P F I A G H V T I I T E P A G P D L G T P F T S R L P P E
 1001 CTTCAGGTCC TAAAGGACA ACAGGCCAT CGGTTGCCC CCAAGGTAA TCAGTCTGC ATCTTGACT AGGCTATATA ACATGCTT GAGGAGGCA
 GAAGTCCAGG ATTTCTGCT TGTCCGGTA GCCAACGGG GGTCCCAT AGTCAGGCG TAGAAGTAC TCCGATATT TGGGTACCGA CTCTCCCGT
 281 L Q V L K D E Q A H R L A P K G N Q S R I L T Q A Y K R M A E E A M
 1101 TGCAGRAGA ACATCAGAAC ACAGGACCTG CCCTTGAGCA AGAGGACAAG ACCAGCAAG TGGCTCTGG CACAGCCCTT CTGCCCCAGC TCGGGGCCAC
 ACGTCTCTT TGTAGTCTG TGTCTGGAC GGAAGTCTG TCTCTGTTT TGTCTGTTT ACCGAGGACC GTGTCGGGA GACGGGTCTG AGCCCCGTG
 315 Q K K H Q N T G P A L E Q E D K T S K V A P G T A P L P R L G A T
 1201 TCCTCAGGAA TCAAGCCTCC TGGCCGGAT CTTAGCCTCA GAATTGAAGA GCAGCTGGC TAAATCAGG ACTGGAGAGG TGCCCTCTGC ACCTCGGAA
 AGAGTCTCTT AGTTCGGAG ACCGGCCCTA GAATCGAGT CTTAAGTCT CTTGACCCG ATTAGTCTCC TGACCTCTCC ACGGGAGAGG TGGAGCCCTT
 348 P Q E S S L L A G I L A S E L K S S W A K S G T G E V P S A P R E
 1301 AACCGGACCA CCCAGATTG TGAACGAGCA TTCCAGAGG AGAGGCCAGA GGTGCTGGC CAGCGGAGCA CTGATGTAGT GGACCTGGAA AATGAGGAGC
 TTGGCTGGT GGGGTCTAAC ACTTGCTCGT AAGGTCTCC TCTCCGGTCT CACAGACCCG GTGCTCTGT GACTACATCA CCTGGACCTT TTAATCTCTG
 381 N R T T P D C E R A F P E E R P E V L G Q R S T D V V D L E N E E P
 1401 CAGACAGTGA CAATGAGTGG CAGCAGTGC TAGAGACCAC TGAGCCTGTG CCTATTCAAC TGAAGGCTCC TCTACCTTG CTGTGTAATC CTGACTTCTG
 GTCTGTCACT GTTACTACC GTCGTGAGC ATCTCTGTG ACTCGGACAC GGATAAGTGT ACTTCCGAGG AGAGTGGAAC GACACATTAG GACTGAAGAC
 415 D S D N E W Q H L L E T T E P V P I Q L K A P L T L L C N P D F C
 1501 CCAGCGCATC CAGAGTCAGC TGCATGAGC TGGAGGGCAG ATCTGAAAG GCATCTTGA GGTGCTTCC CACATCTGC CTGATTCGG GGTCTGAGC
 GGTGCGTAG GTCTCAGTC ACGTACTTC ACCTCCGTC TAGGACTTC CAGTAACTT CAGTAACTT CAGTAACTT CAGTAACTT CAGTAACTT CAGTAACTT
 448 Q R I Q S Q L H E A G G Q I L K G I L E G A S H I L P A F R V L S
 1601 AGTCTTCTCT CCAGTGCAG TCATTCTGTT GCCTTGATT CCTTCTGCG GGAGGAGGG CTTCTGGGC TGCTGCTAG TCTACTCAG CACAGTCAGG
 TCAGAGAGA GGTGAGCTC ACTAAGCAA CGGAACATAA GGAAGACGG CCTCCGTCCT GAAGACCCG ACAGGACTC AGATGAGTCC GTGTAGTCC
 481 S L L S S C S D S V A L Y S F C R E A G L P G L L L S L L R H S Q E

Figura 1C

1701 AGAGCAACAG CCTCCAGCAG CAATCTTGGT ATGGACCTT CTTACAGGAC CTGATGGCTG TGATTCAGGC CTACTTTGCC TGTACTTTCA ATCTGGAGAG
 TCTCGTTGTC GGRGGTCGTC GTTAGAACCA TACCTCGGAA GAATGTCCTG GACTACCGAC ACTAAGTCGG GATGAAACGG ACATGGAAGT TAGACCTCTC
 515 S N S L Q Q Q S W Y G T F L Q D L M A V I Q A Y F A C T F N L E R
 1801 GAGCCAGACA AGTGACAGCC TGCAGGTGTT TCAGGAGGCT GCCAACCTTT TTCTGGACCT GTTGGGAAA CTGCTGGCCC AACCAATGA CTCTGAGCAG
 CTCGGTCTGT TCACTGTGG ACGTCCACAA AGTCTCTCGA CGGTGGGAAA AAGACCTGGA CAACCCCTTT GAGCACCAGG TTGGTCTACT GAGACTCGTC
 548 S Q T S D S L Q V F Q E A A N L F L D L L G K L L A Q P D D S E Q
 1901 ACTTTGGGGA GGGACAGCCT TATGTGCTTT ACTGTCTCTG GCGAAGCCAT GGATGGGAAC AGCCGGGCCA TCTCCAAAGC CTTTACTCC AGCTTGCTGA
 TGAACGCTT CCTGTGCGA ATACACGAAA TGACAGGACA CGCTTCGGTA CCTACCTTGG TCGGCCCGGT AGAGTTTCG GAAATGAGG TCGAACGACT
 581 T L R R D S L M C F T V L C E A M D G N S R A I S K A F Y S S L L T
 2001 CGACACAGCA GGTGTCTTG GATGGGCTCC TTATGGCTT GACAGTTCCA CAGCTCCCTG TCACACTCC CCAAGGAGCC CCGCAACTGA GCCAGCCACT
 GCTGTGCTG CCAACAGAAC CTACCCGAGG AAGTACCGAA CTGTCAAGT GTCGAGGAC AGGTGTGAGG GGTTCTCTCG GGCCTTCACT CGGTGCGTGA
 615 T Q Q V V L D G L L H G L T V P Q L P V H T P Q G A P Q V S Q P L
 2101 GCGAGAGCAG AGTGAGGATA TACCTGGAGC CATTTCTCTT GCCCTGGCAG CCATATGCAC TGCTCTCTGT GGACTGCCCG ACTGCTGGGA TGCCAAAGAG
 CGTCTCTGTC TCACTCTCTAT ATGGACCTCG GTAAAGGAGA CGGGACCGTC GGTATACGTG ACAGGAGCAC CCTGACGGGC TGACGACCTT ACGGTTCTCTC
 648 R E Q S E D I P G A I S S A L A A I C T A P V G L P D C W D A K E
 2201 CAGGTCGTGT GGCATTGGC AAATCAGCTA ACTGAAGACA GCAGCCAGCT CAGGCCATCC CTATCTCTG GCCTGCAGCA TCCCATCTCTG TGCCTGCACC
 GTCCAGACAA CCGTAAACCG TTTAGTCGAT TGACTTCTGT CGTCGGTGA GTCCGGTAGG GAGTAGAGAC CGGACGTCGT AGGTAGGAC ACGGACGTGG
 681 Q V C W H L A N Q L T E D S S Q L R P S L I S G L Q H P I L C L H L
 2301 TTCTCAAGT TCTATACTCC TGCTGCCTTG TCAGTGAGGG CCTGTGGCGT CTCTGGGCG AGGAGCCCTT GGCCTTGGAA TCCCTGTTTA TGTGATTCA
 NAGAGTTCCA AGATATGAGG ACGACGGAAC AGTCACTCCC GGACACGCA GAAGACCCG TCCTCGGGGA CCGGRACCTT AGGACAAAT ACAACTAAGT
 715 L K V L Y S C C L V S E G L C R L L G Q E P L A L E S L F M L I Q
 2401 GGGCAAGTA AAGTAGTAG ATTGGGAAGA GTCTACTGAA GTGACACTCT ACTTCTCTCTC CTCTCTTGTG TTTCCGCTCC AAAACCTGCC TTGTGGAATG
 CCGTTCCTAT TTTTCATCATC TAACCTCTCT CAGATGACTT CACTGTGAGA TGAAGGAGAG GGAAGAACAG AAAGCCGAGG TTTTGGACGG AACACCTTAC
 748 G K V K V V D W E E S T E V T L Y F L S L L V F R L Q N L P C G M
 2501 GAGAAGCTAG GCAGTGACGT TGCTACTCTC TTACCCATT CGCATGTGCT CTTCTTTGTG AGTGAGCAG CCTGTCTATT GGGACACCTT GGTGAGCAG
 CTCTTCGATC CGTCACTGCA ACGATGAGAG AAATGGGTAA GCGTACAGCA GAGAGAACAC TCACGTCTGT GACAGATAA CCCTGTGAA CCAGTCGTTT
 781 E K L G S D V A T L F T H S H V V S L V S A A A C L L G Q L G Q Q G

Figura 1D

2601 GGGTGACCTT TGACCTCCAG CCCATGGGAT GGTGGCTGC AGCCACACAT GCCTTGTCTG CCCTGCAGA GGTTCGGTTG ACTCCACCAG GTAGTTGTGG
 CCCACTGGAA ACTGGAGGTC GGGTACCTTA CCTACCGACG TCGGTGTGTA CGGAACAGAC GGGGACGTCT CCAAGCCAAC TGAGGTGGTC CATCAACACC
 815 V T F D L Q P M E W M A A A T H A L S A P A E V R L T P P G S C G
 2701 ATTCTATGAT GGCCTCCCTTA TCCTTCTGTT GCAGTCTCTC ACTGACAGG GGAAGGCTAG CCTAATCAGG GATATGTCCA GTTCAGAAAT GTGGACCGTT
 TAAGATACTA CCGGAGGAAT AGGAAGACAA CGTCGAGGAG TGACTCGTCC CCTTCCGATC GGATTAGTCC CTATACAGGT CAAGTCTTTA CACCTGGCAA
 848 F Y D G L L I L L L Q L L T E Q G K A S L I R D M S S S E M W T V
 2801 TTGTGGCACC GCTTCTCCAT GGTCTGTAGG CTCGCCGAGG AGGCATCTGC ACAGGAAGGG GAGCTTTGCG TATCCAGTCC ACCAAGCCCT GAGCCAGACT
 AACACCGTGG CGAAGAGGTA CCAGGACTCC GAGGGGCTCC TCCGTAGACG TGTCTTCCCT CTCGAAAGCG ATAGGTGAGG TGGTTCGGGA CTCGGTCTGA
 881 L W H R F S M V L R L P E E A S A Q E G E L S L S S P P S P E P D W
 2901 GGACACTGAT TTCTCCCCAG GGCATGGCAG CCCTGCTGAG CCTGGCCATG GCCACCTTTA CCCAGGAGCC CCAGTTATGC CTGAGCTGCC TGTCCCAGCA
 CCTGTGACTA AAGAGGGGTC CCGTACCGTC GGGACGACTC GGCAGGCTAC CGGTGGAAT GGTCTCTCGG GGTCATACG GACTCGACGG ACAGGGTCTGT
 915 T L I S P Q G M A A L L S L A M A T F T Q E P Q L C L S C L S Q H
 3001 TGGNAGPATC CTCATGTCCA TCCTGAAGCA TCTGCTTTGC CCAGCTTCC TGAATCAACT GGGCCAGGG CCTCATGGGT CTGAGTTTCT CCCTGTCTGTG
 ACCTTCATAG GAGTACAGGT AGGACTTCGT AGACGAAGCG GGTGGAAGG ACTTAGTTGA CGCGTCCGC GGAGTACCCA GACTCAAGA GGGACAGCAC
 948 G S I L M S I L K H L L C P S F L N Q L R Q A P H G S E F L P V V
 3101 GTGCTCTCTG TCTGCCAGCT CCCTTGTCTC CCCTTTGCGC TGGACATGGA TGCTGACCTC CTTATAGTTG TCTTGGCCGA CCTCAGGGAC TCAGAAGTTG
 CACGAGAGAC AGACGGTCCA GGAACGAAG GGAACGCG ACCTGTACCT ACGACTGGAG GAATATCAAC AGAACCGGT GGAGTCCCTG AGTCTTCAAC
 981 V L S V C Q L L C F P F A L D M D A D L L I V V L A D L R D S E V A
 3201 CAGCCCATCT GCTGCAGGTC TGCTGCTACC ATCTTCGGT GATGCAAGTG GAGTGCCCA TCAGCCTTCT CACACGCCGT GCCCTCATGG ATCCCACCTC
 GTCGGGTAGA CGACGTCCAG ACGACGATGG TAGAAGCAA CTACGTTTAC CTCGACGGGT AGTCGGAAGA GTGTGGGAC CGGAGTACC TAGGGTGGAG
 1015 A H L L Q V C C Y H L P L M Q V E L P I S L L T R L A L M D P T S
 3301 TCTCAACCAG TTTGTGAACA CAGTGTCTGC CTCCTCTAGA ACCATCGTCT CGTTTCTCTC AGTTGCCCTC CTGAGTGACC AGCCACTGTT GACCTCCGAC
 AGAGTTGGTC AAACACTTGT GTCACAGACG GAGGGATCT TGGTAGCAGA GCAAGAGAG TCAACGGGAG GACTCACTGG TCGGTGACAA CTGGAGGCTG
 1048 L N Q F V N T V S A S P R T I V S F L S V A L L S D Q P L L T S D
 3401 CTCTCTCTC TGCTGGCCCA TACTGCCAGG GTCCTGTCTC CCAGCCACTT GTCCTTTATC CAAGAGCTTC TGGCTGGCTC TGATGAATCC TATCGGCCCC
 GAAGAGAGAG ACGACCGGT ATGACGGTCC CAGGACAGAG GGTCCGTGAA CAGGAATAG GTTCTCGAAG ACCGACCCGAG ACTACTTAGG ATAGCCGGGG
 1081 L L S L L A H T A R V L S P S H L S F I Q E L L A G S D E S Y R P L

Figura 1E

```

3501 TGGCGAGCCT CCTGGGCCAC CCAGAGRATT CTGTGGGGC ACACACTTAT AGGCTCCTGG GACACTTGCT CCAACRACGC ATGGCCCTGC GTGGGGCACT
    ACGCGTCGGA GGACTCGGTG GGTCTCTTAA GACACGCCCG TGTGTGAATA TCCGAGGACC CTGTGAACGA GGTGTGTGTCG TACCGGGACG CACCCCGTGA
1115 R S L L G H P E N S V R A H T Y R L L G H L L Q H S M A L R G A L
3601 GCAGAGCCAG TCTGGA CTGACTGC TCAGCCTTCT GCTGCTTGGG CTTGGAGACA AGGATCCTGT TGTCCGGTGC AGTCCAGCT TTGCTGTGGG CAAATCCAGCC
    CGTCTCGGTC AGACTGACG AGTCGGAGA CGACGRACCC GAACCTCTGT TCCTAGGACA ACACGCCACG TCACGGTGA AACACACCC GTTACGTCGG
1148 Q S Q S G L L S L L L L G L G D K D P V V R C S A S F A V G N A A
3701 TACCAGGCTG GTCCTCTGGG ACCTGCCCTG GCAGCTGCAG TGCCCACTAT GACCCAGCTG CTTGGAGATC CTCAGGCTGG TATCCGGCGC AATGTTGCAT
    ATGGTCCGAC CAGGAGACCC TGGACGGGAC CGTCGACGTC ACGGTCATA CTGGTTCGAC GAACCTCTAG GAGTCCGACC ATAGGCCGCG TTACAACGTA
1181 Y Q A G P L G P A L A A V P S M T Q L L G D P Q A G I R R N V A S
3801 CAGCTCTGGG CAACTTGGGA CCTGAAGGTT TGGGAGAGGA GCTGTTACAG TGCGAAGTAC CCCAGCGGCT CCTAGAAATG GCATGTGGAG ACCCCAGCC
    GTCGAGACCC GTTGAACCCCT GGACTTCCAA ACCCTCTCCT CGACAATGTC ACGTTTCATG GGATCGCGGA GGATCTTTAC CGTACACCTC TGGGGGTCCG
1215 A L G N L G P E G L G E E L L Q C E V P Q R L L E M A C G D P Q P
3901 AAATGTGAAG GAGGTGCC TCATTGCCCT CCGGAGCCTG CAACAGGAGC CTGGCATCCA TCAGGTACTG GTGTCCCTGG GTGCCAGTGA GAAACTATCC
    TTACACTTC CTCGACGGG AGTAACGGGA GGCTCGGAC GTGTCTCTCG GACCGTAGGT AGTCCATGAC CACAGGGACC CACGGTCACT CTTGATAGG
1248 N V K E A A L I A L R S L Q Q E P G I H Q V L V S L G A S E K L S
4001 TTGCTCTCTC TGGGGAATCA GTCAC TGCCA CACAGCAGTC CTAGGCCCTG CTCTGCCAAA CACTGCAGGA AACTCATTC AACTCTGAGG CCAGCCCAT
    AACGAGAGAG ACCCTTAGT CAGTGACGGT GTGTGCTCAG GATCCGGACG GAGACGGTTT GTGACGTCCT TTGAGTAAGT GGAGGACTCC GGTGCGGTAT
1281 L L S L G N Q S L P H S S P R P A S A K H C R K L I H L L R P A H S

```

Figura 1F

```

4101 GCATGTGATT CCAGATTCTT GCGGTCCAGC CTCCAACTTT GGTGCCAGCT CTTTCTTATN TAATACACAA GGGCCAAATC AACTGAGAGC TAAAGAGACT
CGTACACTAA GGTCTAAGGA CGCCAGGTCG GAGGTTGAAA CCACGGTCGA GAAAGATATN ATTATGTGTT CGCGGTTTRAG TTGACTCTCG ATTCTCTGA
1315 M O

4201 AGAAAGAGAGA TAAGCTGCCA ACTCAACTGA GAACAGGAAA CTNGAAGAGA TTTATATATA AAGCTTCTTC CTTCTCCAG ATGCAGGATG TTTTCAACCA
TCCTTTCTCT ATTGACGGT TGAGTTGACT CTGTCTCTTT GANCTTCTCT AATATATAT TCGAAGAGAG GAAGAGGGTC TACGTCCTAC AAAAGTTGCT
4301 GTAAATTTTA TTGCTGTTGG TGCCAGAGAA GAGTCCCTTT CTCTCTACA TCCAGGGGCC NTTTTCTCCA ATAATGTGCC TTTAACTCTA GGGACCTGCC
CATTTAAAT AACGACAACC ACGTCTCTT CTCAGGGAAA GAAGAGATGT AGGTCCCCG NAAAAGAGGT TATTACACGG AATTTGAGAT CCCTGGACGG
4401 TCACGGACCT TAGGGAAAA CCTCAACCTG AAAGATCTCT TCCTTTCTGG AGCTCCTTTA ATCTTCCAG CAGGTTTTTG CCTTAGACGT GCTGGCCCCA
AGTGCCCTGA ATCCCTTTT GGAGTTGGAC TTTCTAGAGA AGGAAAGACC TCGAGGAAAT TAGAAGGGTC GTCCAAAAAC GGAATCTGCA CGACCGGGT
4501 GGACACTGAT GAAGACAGAG CCTGTCTCAG CTCTAGGCTG TGGGGATCAA TGCCATCAGT CCTGTATT GAGGGATTAT CCCTTAGCCA ACATTCCTAT
CCTGTCTACTA CTTCTGTCTC GGACAGAGTC GAGATCCGAC ACCCTAGTT ACGGTAGTCA GGGACAATA CTCCCTAATA GGGAAATCGGT TGTAAGGATA
4601 CTGTGGGTGG GCGTGGAGAG TGTATCTTTT TTTGGGGTGT GTGTGTATAT GTGTGTGTGT ATGTGTGTGT GTGTTTAAATA GTTCTGTTTG TAAACTCTTT
GACACCCACC CGCACCTCTC ACATAGAAAA AAACCCACACA CACACATATA CACACACACA TACACACACA CACAAATTAT CAAGCAAAAC ATTTGAGAAA
4701 TAATAAAAGT TGTGCTCTAC CATCTTTGAA GCTCCAGGA CAAGGGTTGA GAGGCTCAAC CCTCTTTTCA GCTTCTATGT GGTGTTGGAG GTGCTGGTAT
ATTATTTTCA ACACGGAGTG GTATGAACTT CGAGGGTCCT GTTCCCACT CTCCGAGTTG GGGAGAAAGT CGAAGATACA CCACAACCTC CACGACCATTA
4801 CGTGTTCACA CAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA
GCACAGTGT GTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

```


Figura 2

**CCCGGGCTATGAGCACCAATACAATGGTGCTGACATCCATCAAAGGCACACCACTCTATA
TGTCTCCAGAGCTGGTGGAGGAGCGACCATACGACCACACAGCGGACCTCTGGTCTGTTG
GCTGCATACTATATGAACTGGCAGTAGGCACCCCTCCCTTCTAATGCTACAAGCATCTTT
CAGCTGGTCAGCC**

Figura 3A

hfused	1	M	E	K	Y	H	V	L	E	M	I	G	E	G	S	F	G	R	V	Y	K	G	R	R	K	Y	S	A	Q	V	V	A	L	K	F	I	P	K	L	G	R	S	E	K	E	L	R	N	L	Q	R
dfused	1	M	D	R	Y	A	V	S	S	L	V	G	Q	G	S	F	G	C	V	Y	K	A	Q	R	R	D	D	K	V	V	A	L	K	V	I	S	K	R	G	R	S	N	R	E	L	K	N	L	R	R	
hfused	51	E	I	E	I	M	R	G	L	R	H	P	N	I	V	H	M	L	D	S	F	E	T	D	K	E	V	V	V	V	T	D	Y	A	E	G	E	L	F	Q	I	L	E	D	D	G	K	L	P	E	D
dfused	51	E	C	D	I	Q	A	R	L	K	H	P	H	V	I	E	M	V	E	S	F	E	S	K	F	D	L	F	V	V	T	E	F	A	L	M	D	L	H	R	Y	L	S	F	N	G	A	M	P	E	E
hfused	101	Q	V	A	I	A	A	Q	L	V	S	A	L	Y	Y	L	H	S	H	R	I	L	H	R	D	M	K	P	Q	N	I	L	L	A	K	G	G	I	K	L	C	D	F	G	F	A	R	A	M		
dfused	101	H	A	Q	R	V	V	C	H	L	V	S	A	L	Y	Y	L	H	S	N	R	I	L	H	R	D	L	K	P	Q	N	V	L	L	D	K	N	M	H	A	K	L	C	D	F	G	L	A	R	N	M
hfused	151	S	T	N	T	M	V	L	T	S	I	K	G	T	P	L	Y	M	S	P	E	L	V	E	E	R	P	Y	D	H	T	A	D	L	W	S	V	G	C	I	L	Y	E	L	A	V	G	T	P	P	F
dfused	151	T	M	G	T	H	V	L	T	S	I	K	G	T	P	L	Y	M	A	P	E	L	L	A	E	Q	P	Y	D	H	Q	A	D	M	W	S	L	G	C	I	A	Y	E	S	M	A	G	Q	P	P	F
hfused	201	Y	A	T	S	I	F	Q	L	V	S	L	I	L	K	D	P	V	R	W	P	S	T	I	S	P	C	F	K	N	F	L	Q	G	L	L	T	K	D	P	R	Q	R	L	S	W	P	D	L	L	Y
dfused	201	C	A	T	S	I	L	H	L	V	K	L	I	K	H	E	D	V	K	W	P	S	T	L	S	S	E	C	R	S	F	L	Q	G	L	L	E	K	D	P	S	M	R	I	S	W	T	Q	L	L	C

Figura 3B

hfused	251	HPF	IAGHVT	I	I	T	E	P	A	G	P	D	L	G	T	P	F	T	S	R	L	P	P	E	L	Q	V	L	K	D	E	Q	A	H	R	L	A	P	K	G	N	Q	S	R							
dfused	251	HPF	VEGKL	Y	I	A	E	V	Q	A	A	Q	T	S	P	F	I	N	P	Q	L	A	K	D	T	K	.	.	K	S	Q	Q	L	R	H	V	G	A	D	L	G	D	V	.							
hfused	301	I	L	T	Q	A	Y	K	R	M	A	E	E	A	M	Q	K	K	H	Q	N	T	G	P	A	L	E	Q	E	D	K	T	S	K	V	A	P	G	T	A	P	L	P	R	L	G	A	T	P	Q	E
dfused	297	L	A	A	L	K	L	S	D	V	A	N	E	N	L	S	T	S	R	D	S	I	N	.	A	I	A	P	S	D	I	E	Q	L	E	T	D	V	E	D	N	V	H	R	L	.	I	V	P	.	.
hfused	351	S	S	L	L	A	G	I	L	A	S	E	L	K	S	S	W	A	K	S	G	T	G	E	V	P	S	A	P	R	E	N	R	T	T	P	D	C	.	E	R	A	F	P	E	E	R	P	E	V	L
dfused	343	.	.	.	F	A	D	I	S	Y	R	E	L	P	C	G	.	.	.	T	A	A	A	R	R	A	G	A	M	P	L	I	N	S	Q	T	C	F	V	S	G	N	S	N	M	I	L	N	H	L	
hfused	400	G	Q	R	S	T	D	V	V	D	L	E	N	E	E	P	D	S	D	N	E	W	Q	H	L	L	E	T	T	E	P	V	P	I	Q	L	K	A	P	L	T	L	L	C	N	P	D	F	C	Q	R
dfused	387	N	D	N	F	A	I	E	A	P	A	S	S	A	T	K	S	M	K	S	K	L	K	L	A	L	N	I	K	Q	S	R	S	K	D	L	E	K	R	K	L	S	O	N	L	D	N	F	S	L	R
hfused	450	I	Q	S	Q	L	H	E	A	G	G	I	L	K	G	I	L	E	G	A	S	H	I	L	P	A	F	R	V	L	S	S	L	L	S	S	C	S	D	S	V	A	L	Y	S	F	C	R	E	A	
dfused	437	L	G	Q	S	I	D	I	E	V	Q	R	K	T	T	E	M	L	T	Q	Q	S	Q	A	Q	Q	L	Q	D	R	K	T	Q	Q	L	K	Q	S	M	H	S	T	N	D	E	K	L	S	S	D	N

Figura 3C

hfused	500	GL	P	G	L	L	L	S	L	L	R	-	H	S	Q	E	S	N	S	L	Q	Q	S	W	Y	G	T	F	L	Q	D	L	M	A	V	I	Q	A	Y	F	A	C	T	F	N	L	E	-	R
dfused	487	SP	P	C	L	L	P	G	W	D	S	C	D	E	S	Q	S	P	P	I	E	N	D	E	W	L	A	F	L	H	R	S	I	Q	E	L	D	G	E	F	D	S	L	K	Q	H	N	L	V

hfused	548	S	Q	T	S	D	S	L	Q	V	F	Q	E	A	A	N	L	F	L	D	L	G	K	L	L	A	Q	P	D	D	S	E	Q	T	L	O	R	D	S	L	M	C	F	T	V	L	C	E	A	M
dfused	537	S	I	I	V	A	P	L	R	N	S	K	A	I	P	K	V	-	L	Q	S	V	A	Q	L	L	S	L	P	-	-	-	F	V	L	A	E	Q	H	L	V	A	E	A	I	-	-	K	G	V

hfused	598	D	G	N	S	R	A	I	S	K	A	F	Y	S	S	L	L	T	T	Q	Q	V	V	L	D	G	L	L	H	G	L	T	V	P	Q	L	P	V	H	T	P	Q	G	A	P	Q	V	S	Q	P	L
dfused	580	Y	I	D	V	K	L	V	P	N	L	M	Y	A	C	K	L	L	S	Q	R	H	L	T	D	-	-	-	S	A	A	S	L	P	A	G	T	G	V	S	L	S	R	T	V	R	S	C			

hfused	648	R	E	Q	S	E	D	I	P	G	A	I	S	A	L	A	A	I	C	T	A	P	V	G	L	P	D	C	W	D	A	K	E	Q	V	C	W	H	L	A	N	-	Q	L	T	E	D	S	S	Q
dfused	625	S	D	L	S	A	E	E	M	S	T	A	C	S	L	Y	E	L	V	C	H	L	V	H	Q	Q	Q	F	L	-	-	-	T	Q	F	C	D	A	V	A	I	L	A	V	N	D	M	F	I	N

hfused	697	L	R	P	S	L	I	S	G	L	Q	H	P	I	L	C	L	H	L	L	K	V	L	Y	S	C	C	L	V	S	E	-	-	G	L	C	R	L	L	G	Q	E	P	L	A	L	E	-	S	L	F
dfused	672	F	L	T	H	D	F	K	D	S	R	P	V	R	L	A	S	C	M	L	-	A	L	F	-	C	C	V	L	R	E	L	P	E	N	A	E	L	V	E	K	I	V	F	D	S	R	L	Q	L	A

Figura 3D

hfused	744	M	L	I	Q	G	K	V	K	V	D	W	E	E	S	T	E	V	T	L	Y	F	L	S	L	V	F	R	L	Q	N	L	P	C	G	-	M	E	K	L	G	S	D	V	A	T	L	F	T		
dfused	720	V	L	L	Q	S	R	H	H	L	-	R	Q	R	A	C	Q	M	-	-	-	L	L	L	L	L	A	R	F	S	L	R	G	V	Q	C	I	W	S	G	E	L	K	S	A	L	Q	A	W	P	M

hfused	793	H	S	H	V	V	S	L	V	S	A	A	C	L	L	G	Q	L	G	Q	Q	G	V	T	F	D	L	Q	P	M	E	W	M	A	A	A	T	H	A	L	S	A	P	A	E	V	R	L	T	P
dfused	766	Q	T	C	Q	S	L	R	L	E	A	A	Q	T	L	D	E	L	S	Q	F	S	F	-	F	V	A	Q	A	T	A																			

hfused	843	P	G	S	C	G	F	Y	D	G	L	L	L	L	Q	L	L	T	E	Q	G	K	A	S	L	I	R	D	M	S	S	S	E	M	W	T	V	L	W	H	R	F	S	M	V	L	R	L	P
--------	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

hfused	893	E	E	A	S	A	Q	E	G	E	L	S	L	S	P	P	S	P	E	P	D	W	T	L	I	S	P	Q	G	M	A	A	L	S	L	A	M	A	T	F	T	Q	E	P	Q	L	C	L	S
--------	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

hfused	943	C	L	S	Q	H	G	S	I	L	M	S	I	L	K	H	L	C	P	S	F	L	N	Q	L	R	Q	A	P	H	G	S	E	F	L	P	V	V	L	S	V	C	Q	L	C	F	P	F
--------	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

hfused	993	A	L	D	M	D	A	D	L	L	I	V	V	L	A	D	L	R	D	S	E	V	A	A	H	L	L	Q	V	C	C	Y	H	L	P	L	M	Q	V	E	L	P	I	S	L	L	T	R	L	A	L
--------	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Figura 3E

hfused 1043 MDPTSLNQFVNTVSASPRRTIVSFLSVALLSDQPLLTSDLLSLLAHTARVL
 hfused 1093 SPSHLSFIQELLAGSDESYRPLRSLLGHPENSVRATHYRLLGHLLQHSMA
 hfused 1143 LRGALQSQSGLLSLLLLGLGDKDPVVRCSASFVGNAA YQAGPLGPALAA
 hfused 1193 AVPSMTQLLGDPQAGIRRNVASALGNLGPEGLGEE LLQCEVPQR LLEMAC
 hfused 1243 GDPQPNVKEAALIALRSLQQEPG IHQV LVSLGASEKLSLLSLGNQSLPHS
 hfused 1293 SPRPASAKHCRKLIHLLRPAHSM

Figura 4A

> Longitud 5125 pb (circular)

```

1 CCCACGCGTC CGCCCCACGCG TCCGGGGCGT CCCAGATGTT GTGGAATGCT CCTGGATCT ATAGCTCTTC ACCGTCTCTA CTTTCTTCTCCT TCTAAGAGAT
GGGTGCGCAG GCGGGTGCGC AGGCCCCGCA GGGTCTACAA CACCTTGACA GGGACCTAGA TATCGAGAAG TGCGAGAGAT GAAAGAAGGA AGATTCTCTA

101 CCTGAAACCT CTGTCATGGA AAAGTACCAC GTGTTGGAGA TGATTGGAGA AGGCTCTTTT GGGAGGGTGT ACAAGGGTGC AAGAAATAC AGTGTCAAGG
GGACTTTGGA GACAGTACCT TTTCATGGTG CACAACCTCT ACTAACCTCT TCCGAGAAAA CCCTCCACACA TGTTCCTTATG TCACGAGTCC

1 ME K Y H V L E M I G E G S F G R V Y K G R R K Y S A Q V
Inicio de 1º ORF

201 TCGTGGCCCT GAAGTTCATC CCARAATTGG GCGGCTCAGA GAAGGAGCTG AGGAATTTGC AACGAGAGAT TGAATAATG CGGGGTCTGC GGCATCCCAA
AGCACCGGGA CTTCAGTAG GGTTTTAACC CCGCGAGTCT CTTCCTCGAC TCCTTAAACG TTGCTCTCTA ACTTTATTAC GCCCCAGACG CCGTAGGGTT

30 V A L K F I P K L G R S E K E L R N L Q R E I E I M R G L R H P N

301 CATTTGTGCAT ATGCTTGACA GCTTTGAAAC TGATAAAGAG GTGGTGGTGG TGACAGACTA TGCTGAGGGA GAGCTCTTTC AGATCCTAGA AGATGACGGA
GTAAACACGTA TACGAACTGT CGAACTTTG ACTATTCTC CACCACCACC ACTGTCTGAT ACGACTCCCT CTCGAGAAAG TCTAGGATCT TCTACTGCCCT

63 I V H M L D S F E T D K E V V V V T D Y A E G E L F Q I L E D D G

401 AAACCTTCTG AAGACCAGGT TCAGGCCATT GCTGCCCAGT TGGTGTGAGC CTTGTACTAT CTGCATTCCC ACCGCATCCT ACACCGAGAT ATGAAGCCTC
TTTGAAGGAC TTCTGGTCCA AGTCCGGTAA CGACGGGTCA ACCACAGTCG GGACATGATA GACGTAAGGG TGCGGTAGGA TGTGGCTCTA TACTTCGGAG

96 K L P E D Q V Q A I A A Q L V S A L Y Y L H S H R I L H R D M K P Q

501 AGAACATCCT CCTCGCCAAG GGTGGTGGCA TCAAGCTCTG TGACTTTGGA TTTGCCCGGG CTATGAGCAC CAATACAATG GTGCTGACAT CCATCAAAGG
TCTTGTAGGA GGAGCGGTTT CCACACCGT AGTTCAGAC ACTGAACCT AAACGGGCC GATCTCTG GTATGTTTAC CACGACTGTA GGTACTTTCC

130 N I L L A K G G I K L C D F G F A R A M S T N T M V L T S I K G

601 CACACCACCT TATATGCTC CAGAGCTGGT GGAGGAGCGA CCATACGACC ACACAGCGGA CCTCTGGTCT GTTGGCTGCA TACTATATGA ACTGCGAGTA
GTGTGGTGAG ATATACAGAG GTCTCGACCA CCTCCTCGCT GGTATGCTGG TGTGTCGCTT GGAGACCGA CAACCGACCT ATGATATACT TGACCGTCA

163 T P L Y M S P E L V E E R P Y D H T A D L W S V G C I L Y E L A V

```

Figura 4B

701 GGCACCCCTC CCTTCTATGC TACAAGCATC TTTCAGTGG TCAGCTCAT TCTCAAGGAC CCTGTGGCT GGCCTCAAC CATCAGTCCC TGCTTTAAGA
 CCGTGGGGAG GGAACATACG ATGTTCTGAG AAAGTCACC AGTCGAGTA AGAGTTCCTG GGACACCGA CCGGGAGTTG GTAGTCAGG ACMAATTTCT
 196 G T P P F Y A T S I F Q L V S L I L K D P V R W P S T I S P C F K N
 801 ACTTCTTGCA GGGACTGCTC ACCAAGACC CAGGCGAGC ACTGTCTCG CCAGACCTCT TATATCACCC CTTTATTGCT GGTCAATGCA CCATATTAAC
 TCAAGGACGT CCTGACGAG TGGTTTCTGG GTCCGTCGC TGACAGACC GGTCTGGAGA ATATAGTGG GAAATACGA CCAGTACAGT GGTATTATTG
 230 F L Q G L L T K D P R Q R L S W P D L L Y H P F I A G H V T I I T
 901 TGAGCCAGCA GGCACGATC TTGGGACCCC ATTCACGAG CGCCTACCC CAGAACTTCA GGTCTTAAAG GACGACAGG CCCATCGGT GGGCCCCAAG
 ACTCGTCTAA ACCCTGGGG TAAAGTGGCG GCGATGGG GTCTTGAAGT CCAGGATTC CTGCTTGTCC GGSTAGCCAA CCGGGGGTTC
 263 E P A G P D L G T P F T S R L P P E L Q V L K D E Q A H R L A P K
 1001 GGTATCAGT CTGCGATCTT GACTCAGGCC TATAACGCA TGGCTGAGGA GGCCATGCAG AAGAAACATC AGAACACAGG ACCTGCCCTT GAGCAAGAGG
 CCATTAGTCA GAGCOTAGAA CTGAGTCCGG ATATTTCGT ACCGACTCCT CCGTACGTC TTCTTTGAG TCTTGTGTC TGGACGGGAA CTCGTTCTCC
 296 G N Q S R I L T Q A Y K R M A E E A M Q K K H Q N T G P A L E Q E D
 1101 ACAAGACCAG CAAGTGGCT CCTGGCACAG CCCTCTGCC CAGACTCGG GCCACTCCTC AGGAATCAAG CCTCTGGCC GGGATCTTAG CCTCAGAATT
 TGTCTGGTC GTTCCACCGA GGACCGTGTG GGGGAGACGG GTCTAGCCCC CCGTGAGGAG TCCTTAGTTC GGAGGACCG CCCTAGAATC GGAGTCTTAA
 330 K T S K V A P G T A P L P R L G A T P Q E S S L L A G I L A S E L
 1201 GAAAGCAGC TGGGCTAAAT CAGGACTGG AGAGTGGCC TCTGCACCTC GGAACACCG GACCACCCA GATTGTGAAC GAGCATTCCT AGAGGAGAGG
 CTTCTGTCG ACCGATTTA GTCCCTGACC TCTCCACGG AGACGTGGAG CCCTTTTGGC CTGGTGGGT CTACACTTG CTCGTAAGG TCTCTCTCTCC
 363 K S S W A K S G T G E V P S A P R E N R T T P D C E R A F P E E R
 1301 CCAGAGGTC TGGGCCAGC GAGCACTGAT GTAGTGGACC TGGAAATGA GGAGCCAGAC ACTGACAATG AGTGGCAGCA CCTGCTAGAG ACCACTGAGC
 GGTCTCCAG ACCCGTCCG CTCGTGACTA CATCACCTGG ACCTTTTACT CCTCGGTCTG TCACGTGTAC TCACCCTGCT GGACGATCTC TGGTGACTCG
 396 P E V L G Q R S T D V V D L E N E E P D S D N E W Q H L L E T T E P
 1401 CTGTGCTAT TCAACTGAAG GCTCCTCTCA CCTTGTCTGT TAATCTGAC TTCTGCCAGC GCATCCAGAG TCAGCTGCAT GAAGCTGGAG GGCAGATCCT
 CACACGATA AGTTGACTTC CGAGGAGAGT GGAACGACAC ATTAGACTG AAGACGGTGC CGTAGGTCTC ACTCGAGCTA CTTCCGACCTC CCGTCTAGGA
 430 V P I Q L K A P L T L L C N P D F C Q R I Q S Q L H E A G G Q I L
 1501 GAAAGCATC TTGAGGGTG CTTCCCCACAT CCTGCCCTGCA TTCCGGTCC TCTCTCCAGC TCCAGTGCAT CTGTTGCCCT GTATTCTCTC
 CTTTCCGTAG AACCTCCAC GAAGGTGTA GGACGGAGCT AAGGCCAGG ACTCGTCAGA AGAGAGGTGC AGTCACTAA GACAACGGAA CATAAGGAGAG
 463 K G I L E G A S H I L P A F R V L S S L L S S C S D S V A L Y S F

Figura 4C

1601 TGCCGGGAGG CAGGGCTTCC TTGGCTGCTG CTGAGTCTAC TCAGGCACAG TCAGGAGAGC AACAGCCTCC AGCAGCAATC TTGGTATGGG ACCTTCTTAC
 496 C R E A G L P G L L L S L L R H S Q E S N S L Q Q Q S W Y G T F L Q
 1701 AGGACCTGAT GGCTGTGATT CAGGCCTACT TTGCCCTGTAC CTTCAATCTG GAGAGGAGCC AGACAAGTGA CAGCCTGCAG GTGTTTCAGG AGGTGSCAA
 530 D L M A V I Q A Y F A C T F N L E R S Q T S D S L Q V F Q E A A N
 1801 CCTTTTCTG GACCTGTTGG GGAACCTGCT GGCCCAACCA GATGACTCTG AGCAGACTTT GCAGAGGGAC AGCCTTATGT GCTTTACTGT CCTGTGCGAA
 563 L F L D L L G K L L A Q P D D S E Q T L Q R D S L M C F T V L C E
 1901 GCCATGGATG GGAACAGCCG GGCCATCTCC AAGCCCTTTT ACTCCAGCTT GCTGACGACA CAGCAGGTTG TCTTGGATGG GCTCCTTCAT GGCTTGACAG
 596 A M D G N S R A I S K A F Y S S L L T T Q Q V V L D G L L H G L T V
 2001 TTCCACAGCT CCTGTGCCAC ACTCCCAAG GTAACACAGG TGGAGAAGG AGTTCTCTT GACTTACTTG TTGCATAGGT CAGGCTCCGC TCTTTCTATT
 630 P Q L P V H T P Q G N Q S G E G R F S O
 Inicio de la secuencia de intrón
 2101 GCCATCACCT AGATCGCACC TGGCATTTAG TAGTGCTCA ATAAATAACT GTGAACCTGAG AGAATGAATG GGGATCTGAG GGAACAACAA AGACCTCATC
 CCGTAGTGGA TCTAGCGTGG ACCGTAAATC ATCCACGAGT TATTTATTGA CACTTGACTC TCTTACTTAC CCTAGACTC CCTTGTGTTG TCTGGAGTAG
 2201 CTGCATTCTT CCCACTCCCT TAGGTTCCCT ACTCCTGCTG CCATGTCCGT GAGTACTGGT GCTATTGTCT AGGGCAAGAG CCTCAGGCCT TTGGAGTTAC
 GACGTAAGAA GGGTGAGGA ATCCAAGGA TGAGGACGAC GGTACAGCCA CTCATGACCA CGATAACAGA TCCCGTCTC GGAGTCCGGA AACCTCAATG
 1 S Y
 2° ORF empieza aquí
 2301 TCTTTGCTTT TCTCCACAGG AGCCCGCAA GTGAGCCACG CACTGCGAGA GCAGAGTGAG GATATACCTG GAGCCATTTC CTCTGCCCTG GCAGCCATAT
 AGAAACGAAA AGAGTGTC TCGGGCGTT CACTCGGTG GTGACGCTCT CGTCTACTC CTATATGAC CTCGGTAAAG GAGACGGGAC CGTCGGTATA
 3 S L L F S T G A P Q V S Q P L R E Q S E D I P G A I S S A L A A I C
 2401 GCACCTGCTC TGTGGGACTG CCCGACTGCT GGGATGCCAA GGAGCAGGTC TGTGGCATT TGGCAATCA GCTAACTGAA GACACAGCC AGCTCAGGCC
 CGTGACGAGG ACACCTGAC GGGCTGACGA CCCTACGGTT CCTCGTCCAG ACAACGTA ACCGTTTAGT CGATTGACTT CTGTCTCGG TCGAGTCCGG
 37 T A P V G L P D C W D A K E Q V C W H L A N Q L T E D S S Q L R P

Figura 4D

2501 ATCCCTCATC TCTGGCCTGC AGCATCCCAT CCTGTGCCTG CACCTTCTCA AGTTTCTATA CTCCTGCTGC CTGTGACGT AGGGCTGTG CCGTCTTCTG
 TAGGGAGTAG AGACCGGACG TCGTAGGGTA GGCACGGAC GTGGAAGAT TCCAAGATAT GAGGACGACG GAACAGTCAC TCCCGGACAC GGCAGAAGAC
 70 S L I S G L Q H P I L C L H L L K V L Y S C C L V S E G L C R L L
 2601 GGGCAGGAGC CCGTGGCCTT GGAATCCCTG TTTATGTTGA TTCAAGGGA GGTAAAGTA GTAGATTGG AAGAGTCTAC TGAAGTGACA CTCTACTTCC
 CCCGTCTCTG GGGACCGGAA CCTTAGGGAC AAATACAAT AGTCCCGTT CCATTTCAT CATCTAACCC TTCTCAGATG ACTTCACTGT GAGATGAAGG
 103 G Q E P L A L E S L F M L I Q G K V K V V D W E E S T E V T L Y F L
 2701 TCTCCCTTCT TGTCTTTCGG CTCCTTTCGG AATGGAGAG CTAGGAGTG ACGTTGCTAC TCCTTTTACC CATTGCGATG TCCTCTCTCT
 AGAGGGAAGA ACAGAAAGCC GAGTTTTCGG ACGGAACACC TTACCTCTTC GATCCGTCAC TGCAAGCATG AGAGAAATGG GTAAGCGTAC AGCAGAGAGA
 137 S L L V F R L Q N L P C G M E K L G S D V A T L F T H S H V V S L
 2801 TGTGAGTGCA GCAGCTGTC TATTGGGACA GCTTGGTCAG CAAGGGGTGA CCTTTGACCT CCAGCCCATG GAATGGATGG CTGCAGCCAC ACATGCTTGT
 AACTCACGT CGTCGGACAG ATAACCTCTG CGAACCACTG GTTCCCCACT GGAACACTGA GGTCCGGTAC CTTACCTTACC GACGTGCGTG TGTACGGAAC
 170 V S A A A C L L G Q L G Q Q G V T F D L Q P M E W M A A A T H A L
 2901 TCTGCCCTG CAGAGTTTCG GTTGACTCCA CCAGGTAGTT GTGATTTCTA TGATGGCTC CTTATCTTC TGTGACGT CCTACTGAG CAGGGAAGG
 AGACGGGAC GTCTCCAAGC CAACTGAGGT GTTCCATCAA CACCTAAGAT ACTACCGAG GAATAGGAAG ACAACGTGA GGAGTACTC GTCCCTTCC
 203 S A P A E V R L T P P G S C G F Y D G L L I L L L Q L L T E Q G K A
 3001 CTAGCCTAAT CAGGATATG TCCAGTTTCAG AAATGTGGAC CGTTTGTGG CACCGTTCT CCATGCTCCT GAGGTCCTCC GAGGAGGCAT CTGCACAGGA
 GATCGGATTA GTCCCTATAC AGGTCAAGTC TTTACACCTG GCATAACACC GTGGCGAAGA GGTACCAAGA CTCCGAGGGG CTCCTCCGTA GACGTGCTCT
 237 S L I R D M S S S E M W T V L W H R F S M V L R L P E E A S A Q E
 3101 AGGGAGCTT TCGTATCCA GTCCACCAG CCTGAGCCA GACTGGACAC TGATTTTCTC CCAGGCGATG GCAGCCCTGC TGAGCCTGGC CATGGCCACC
 TCCCTCGAA AGCATAGGT CAGTGGTTC GGGACTCGT CTGACCTGTG ACTAAGAGG GTTCCGTCAC CGTCCGGACG ACTCGGACCG GTACCGGTGG
 270 G E L S L S S P P S P E P D W T L I S P Q G M A A L L S L A M A T
 3201 TTTACCCAG AGCCCAAGTT ATGCTGACC TGCTGTGCC AGCATGGAAG TATCTCATG TCCATCTGAT AGCATCTGAT TTGCCCCAGC TTCCTGAATC
 AAATGGGTCC TCGGGGTCAA TACGACTCG ACGACACAGG TCGTACCTTC ATAGAGTAC AGGTAGACT TCGTAGACGA AACGGGTGCG AAGGACTTAG
 303 F T Q E P Q L C L S C L S Q H G S I L M S I L K H L L C P S F L N Q
 3301 AACTGCGCCA GGCCTCAT GGGTCTGAGT TTCTCCCTGT CTRGTGCTC TCTGTGCTC AGTCTCTTTG CTTCCCTTTT GCGCTGGACA TGGATGCTGA
 TTGACGCGGT CCGCGGAGTA CCCAGACTCA AAGAGGGACA GCACACAGG TCGAGGAGG GAAGGGAAA CCGACCTGT ACCTAGACT
 337 L R Q A P H G S E F L P V V V L S V C Q L L C F P F A L D M D A D

Figura 4E

3401 CCTCCTTATA GTTGCTTTGG CCGACCTCAG GGACTCAGAA GTTGCAGCCC ATCTGCTGCA GGTCTGCTGC TACATCTTC CGTTGATGCA AGTGAGAGCTG
 GGAGGAATAT CAACAGAAC GGCTGGAGTC CCTGAGTCTT CAACGTGGG TAGAGGACGT CCAGACGACG ATGGTAGAG GCAACTACGT TCACCTCGAC
 370 L L I V V L A D L R D S E V A A H L L Q V C C Y H L P L M Q V E L
 3501 CCCATCAGCC TTCTCACACG CCTGGCCCTC ATGGAATCCCA CCTCTCTCA CCAGTTTGTG AACACAGTGT CTGCTCCTCC TAGAACCATC GTCTCGTTTC
 GGGTAGTCGG AAGAGTGTG GACCCGGGAG TACCTAGGGT GGAGAGATT GGTCAACAC TTGTGTCA GACGAGGGG ATCTTGGTAG CAGAGCAAG
 403 P I S L L T R L A L M D P T S L N Q F V N T V S A S P R T I V S F L
 3601 TCTCAGTTGC CCTCCTGAGT GACCAGCCAC TGTGACCTC CGACCTTCTC TCTCTGCTGG CCATATCTGC CAGGTCCTG TCTCCAGCC ACTTGTCTT
 AGAGTCAACG GGAGGACTCA CTGGTCGGTG ACAACTGGAG GCTGGAAGAG AGAGACGACC GGTATGACG GTCCAGGAC AGAGGGTCGG TGAACAGGAA
 437 S V A L L S D Q P L L T S D L L S L L A H T A R V L S P S H L S F
 3701 TATCCAAGAG CTCTGGCTG GCTCTGATGA ATCCTATCGG CCCCTGCGCA GCCTCTTGG CCACCCAGAG AATTCTGTGC GGCACACAC TTATAGSCTC
 ATAGGTTCTC GAAGACCGAC CGAGACTACT TAGGATAGCC GGGACGCGT CGGAGGACCC GTGGGTCTC TTAAGACACG CCCGTGTGTG AATATCCGAG
 470 I Q E L L A G S D E S Y R P L R S L L G H P E N S V R A H T Y R L
 3801 CTGGGACACT TGCTCCAACA CAGCATGGCC CTGCGTGGG CACTGCAGAG CCAGTCTGGA CTGCTCAGCC TTCTGCTGCT TGGGCTTGA GACAAGGATC
 GACCTGTGA ACGAGGTGT GTCTACCGG GACGCACCC GTGACGTCTC GGTACAGCTT GACGAGTCGG AAGACACGA ACCCGAACCT CTGTTCTTAG
 503 L G H L L Q H S M A L R G A L Q S Q S G L L S L L L L G L G D K D P
 3901 CTGTTGTGCG GTGCAGTGCC AGCTTTGCTG TGGGCAATGC AGCTACCAG GCTGGTCTC TGGGACCTGC CCTGGCAGCT GCAGTGCCCA GTATGACCCA
 GACAACACGC CACGTACCG TCGAAACGAC ACCCGTTACG TCGGATGGTC CGACCCAGGAG ACCCTGGACG GGACCCGCGA CGTCACGGGT CATACTGGGT
 537 V V R C S A S F A V G N A A Y Q A G P L G P A L A A A V P S M T Q
 4001 GCTGCTTGA GATCCTCAGG CTGATATCCG GCGCAATGT GCATCAGCTC TGGGCAACTT GGGACCTGAA GGTTTGGGAG AGGAGCTGTT ACAGTGCGAA
 CGACGAACCT CTAGGAGTCC GACCATAGGC CGCGTTACAA CGTACTCGAG ACCCGTTGAA CCCTGGACTT CCAPACCTC TCCTCGACAA TGTCAACGCTT
 570 L L G D P Q A G I R R N V A S A L G N L G P E G L G E L L Q C E
 4101 GTACCCAGC GGCTCCTAGA AATGGCATGT GGAGACCCCC AGCCAAATGT GAAGGAGGCT GCCCTCATTTG CCCTCCGAG CCTGCAACAG GAGCCTGGCA
 CATGGGTG CCGAGGATCT TTACCGTACA CCTCTGGGG TCGGTTTACA CTCTCTCCA CGGAGTAAC GGGAGGCTC GGACGTTGTC CTCGGACCGT
 603 V P Q R L L E N A C G D P Q P N V K E A A L I A L R S L Q Q E P G I

Figura 4F

```

4201 TCCATCAGGT ACTGGTGTC CTGGTGCCA GTGAGAAACT ATCCTGTCTC TCTCTGGGA ATCAGTCACT GCCACACAGC AGTCTAGGC CTGCTCTGTC
    AGGTAGTCCA TGACCACACAGG GACCCACGGT CACTCTTTGA TAGAACGAG AGAGACCCCT TAGTCAGTGA CCGTGTGTCG TCAGGATCCG GACGGAGACG
637 H Q V L V S L G A S E K L S L L S L G N Q S L P H S S P R P A S A
4301 CAAACACTGC AGGAACTCA TTCACCTCCT GAGGCCAGCC CATAGCATGT GATTCAGAT TCCTGCGGTC CAGCCTCCAA CTTTGGTTGC CAGCTCTTTC
    GTTGTGACG TCCTTTGAGT AAGTGGAGGA CTCGGGTCCG GTATCGTACA CTAAGGTCTA AGGACGCCAG GTCGGAGGTT GAAACCAACG GTCGAGAAAAG
670 K H C R K L I H L L R P A H S M O
4401 TTATTCTACT ACACAAGCCG CCAACTCAAC TGAGAGCTAA AGAGACTAGA AAAGAGATAA GCTGCCAACT CAACTGAGAA CAAGAAACTA GAAGAGATTT
    AATAAGATGA TGTGTTCCGC GGTGAGTTG ACTCTGATCT TCTCTGATCT TTCTCTATT CCACGGTTGA GTTGACTCTT GTTCTTTGAT CTTCCTTAAA
4501 ATATATAAAG CTTCTTCCTT CTCCAGATG CAGGATGTTT TCAACCAGTA AATTTTATTG CTGTTGGTGC CAGAGAAGAG TCCTTTCTTC TCTACATCCA
    TATATATTTC GAAGAAGGAA GAGGTCTAC GTCCTACAAA AGTTGGTCAAT TTAAATAAAC GACAACCACG GTCTCTTCTC AGGAAAGAAAG AGATGTAGGT
4601 GGGGCCCTTT CTCCAATAAT GTGCTTTAA CTCTAGGAC CTGCTTCACG GACCTTAGGG AAAAACTCA ACCTGAAAAG TCCTTTCTCT TCTGGAGCTC
    CCCCCGAAA GAGGTATTTA CACGGAATTT GAGATCCCTG GACGGAGTGC CTGGAATCCC TTTTGGAGT TGGACTTTCT AGAGAAGGAA AGACCTCGAG
4701 CTTTAACTCT CCCAGCAGGT TTTTGCCTTA GACGTGCTGG CCCCAGGACA GTGATGAAGA CAGAGCCTGT CTCAGCTCTA GGCTGTGGGG ATCAATGCCA
    GAAATTAGAA GGGTCGTCCA AAAACGGAAT CTGCACGACC GGGTCTCTGT CACTACTTCT GTCTCGGACA GAGTCGAGAT CCGACACCCC TAGTTACGGT
4801 TCAGTCCCTG TTATTGAGGG ATTATCCCTT AGCCAAACAT CCATCTGTG GGTGGCGTG GAGAGTGAT CTTTTTTGG GGTGTGTGTG TATATGTGTG
    AGTCAGGGAC AATACTCCC TATATAGGAA TCGGTTGTAA GGATAGACAC CCACCCGCAC CTCTCACATA GAAAAAACC CCACACACAC ATATACACAC
4901 TGTGTATGTG TGTGTGTGTT TAATAGTTCT GTTTGTAAAC TCTTTTAATA AAAGTTGTGC CTCACCATA CTCAACATAC TTGAAGCTCC CAGGACAAGG GTTGAGAGGC
    ACACATACAC ACACACACAA ATTATCAAGA CAAACATTTG AGAAAAATTAT TTTCAACACG GAGTGGTATG AACTTCGAGG GTCCGTGTTCC CAACTCTCCG
5001 TCAACCCCTC TTTACGCTC TATGTGGTGT TGGAGGTGCT GGTATCGTGT TCACACAAAA AAAAAAATAA AAAAAAATAA AAAAAAATAA
    AGTGGGGAG AAAGTCGAAG ATACACCACA ACCTCCACGA CCATAGCACA AGTGTGTTTT TTTTGTTTTT TTTTGTTTTT TTTTGTTTTT
5101 AAAAAAATAA AAAAAAATAA AAAAA
    TTTTTTTTTT TTTTTTTTTT TTTTT

```

Figura 5A

> Longitud 5252 pb (circular)

```

1 GGAGCTTGGG GCTCCTAGGC TGGGGGGGCTC CCAGATGTTG TGGAACTGTC CCTGGATCTA TAGCTCTTCA CCGTCTCTAC TTTCTTCCCTT CTAAGAGATC
CCTCGAACCT CGAGGATCCG ACCCCCGCAG GGTCTACAAC ACCTGACAG GGACCTAGAT ATCGAGAAGT GGCAGAGATG AAAGAAGGAA GATTCTCTAG

101 CTGAACCTC TGTCATGGAA AGTACCACG TGTGGAGAT GATTGGAGAA GGCTCTTTTG GGAGGTGTA CAAGGTGCGA AGAAAATACA GTGCTCAGGT
GACTTTGGAG ACAGTACCTT TTCATGGTGC ACACCTCTA CTRACCTCTT CCGAGAAAAC CCTCCACAT GTTCCAGCT TCTTTTATGT CACGAGTCCA

1 MEKYHVLEMI GEGEIGELVYKGR R K Y S A Q V
~^~ Traducción de ATG empieza aquí

201 CGTGGCCCTG AAGTTCATCC CAAAATTGGG GCGCTCAGAG AAGGAGCTGA GGAATTGCA ACGAGAGATT GAAATAATGC GGGGTCTGCG GCATCCCAAC
GCACGGGAC TTCAAGTAGG GTTTTACCC CCGAGTCTC TTCTCTGACT CCTTAAACGT TGCTCTCTAA CTTTATTACG CCCACAGCGC CGTAGGTTG

30 V A L K F I P K L G R S E K E L R N L Q R E I E I M R G L R H P N

301 ATTGTGCATA TGCTTGACAG CTTTGAACT GATAAAGAG TGGTGGTGGT GACAGACTAT GCTGAGGGAG AGCTCTTTCA GATCCTAGAA GATGACGAA
TAACACGTAT ACGAACTGTC GAAACTTTGA CTATTCTCC ACCACCACCA CTGTCTGATA CGACTCCCTC TCGAGAAAGT CTAGGATCTT CTACTGCTT

63 I V H M L D S F E T D K E V V V V T D Y A E G E L F Q I L E D D G K

401 AACTTCCTGA AGACCAGGTT CAGGCCATTG CTGCCCATTT GGTGTCAGCC CTGTACTATC TGCATTCCCA CCGCATCCTA CACCGAGATA TGAAGCCTCA
TTGAAGGACT TCTGGTCCAA GTCCGGTAA CACGGGTCAA CCACAGTCGG GACATGATAG ACGTAAGGT GCGTAGGAT GTGGCTCTAT ACTTCGGAGT

97 L P E D Q V Q A I A A Q L V S A L Y Y L H S H R I L H R D M K P Q

501 GAACATCCTC CTCGCCAAGG GTGGTGGCAT CAAGCTCTGT GACTTTGGAT TTGCCCGGGC TATGAGCACC AATACAATGG TGCTACATC CATCAAGGC
CTTGAGGAG GAGCGTTCC CACCACCGTA GTTCGAGACA CTGAAACCTA AACGGGCCG ATACTCGTGG TTATGTTACC ACGACTGTAG GTAGTTTCCG

130 N I L L A K G G I K L C D F G F A R A M S T N T M V L T S I K G

601 ACACCACCTCT ATATGTCTCC AGAGCTGGTG GAGGAGCGAC CATACGACCA CACAGCGGAC CTCTGGTCTG TTGGCTGCACT ACTATATGAA CTGGCAGTAG
TGTTGGTGAGA TATACAGAGG TCCGACCCAC CTCCTCGCTG GTATGCTGGT GTGTCGCTG GAGACCAGAC AACCGACGTA TGATATACCT GACCGTCATC

163 T P L Y M S P E L V E E R P Y D H T A D L W S V G C I L Y E L A V G

```

Figura 5B

701 GCACCCCTCC CTTCTATGCT ACAAGCATCT TTACGCTGGT CAGCCCTCATTT CTCAAGGACC CTGTGGCGTG GCCTCAACC ATCAGTCCCT GCTTTAAGAA
 CGTGGGAGG GAAGATACGA TGTTCGTAGA AAGTCGACCA GTCCGAGTAA GAGTTCCTGG GACACGCGAC CGGAGTTGG TAGTCAGGGA CGAAATTTCTT
 197 T P P F Y A T S I F Q L V S L I L K D P V R W P S T I S P C F K N
 801 CTTCTGCGAG GCACTGCTCA CCAAGACCC ACGGACGGA CTGTCTCTGGC CAGACCTCTT ATATCACCCC TTATTATTGCTG GTCATGTTCAC CATAATAACT
 GAAGGACGTC CCTGACGAGT GGTTCCTGGG TGCCCTCGCT GACAGGACCG GTCTGGAGAA TATAGTGGG AAATAACGAC CAGTACAGTG GTATTATTGA
 230 F L Q G L L T K D P R Q R L S W P D L L Y H P F I A G H V T I I T
 901 GAGCCAGCAG GCCCAGATTT GGGGACCCCA TTCACCAGCC GCCTACCCCC AGRACCTCAG GTCCTAAGG ACGAACAGGC CCATCGGTTG GCCCCCAAGG
 CTCGGTCGTC CGGGTCTAAA CCCCTGGGT AAGTGGTCGG CGGATGGGG TCTTGAAGTC CAGGATTTCC TGCTGTCCG GGTAGCCAAAC CGGGGTTCC
 263 E P A G P D L G T P F T S R L P P E L Q V L K D E Q A H R L A P K G
 1001 GTAATCAGTC TCGCATCTTG ACTCAGGCCT ATAAACGCAT GGCTGAGGAG GCCATGCAGA AGAAACATCA GAACACAGGA CCTGCCCTTG AGCAAGAGGA
 CATTAGTCAG AGCGTAGAAC TGAGTCCGGA TATTGGCGTA CCGACTCCTC CGGTACGCTCT TCTTTGTAGT CTGTGTCTCT GGACGGGAAC TCGTTCTCCT
 297 N Q S R I L T Q A Y K R M A E E A M Q K K H Q N T G P A L E Q E D
 1101 CAAGACCAGC AAGTGGCTC CTGGCACAGC CCCTCTGCC AGACTGGGG CCACCTCTCA GGAATCAAGC CTCCTGGCGG GGATCTTAGC CTCAGAAATTG
 GTTCTGGTCG TTCCACCGAG GACCGTGTG GGGAGACGGG TCTGAGCCCC GGTGAGGAGT CCTTAGTTCG GAGGACCGC CCTAGAATCG GAGTCTTAAC
 330 K T S K V A P G T A P L P R L G A T P Q E S S L L A G I L A S E L
 1201 AAGACAGCT GGGCTAAATC AGGACTGGA GAGTGGCCT CTGCACCTCG GGAAACCCG ACCACCCAG ATTGTGAAG AGCATTCCCA GAGGAGAGG
 TTCTCGTCGA CCCGATTTAG TCCCTGACCT CTCACGGGA GACGTGGAGC CCTTTTGGCC TGGTGGGTC TAACACTTGC TCGTAAGGGT CTCCCTCTCCG
 363 K S S W A K S G T G E V P S A P R E N R T T P D C E R A F P E E R P
 1301 CAGAGGTGCT GGGCCAGCG AGCACTGATG TAGTGGACCT GGAAATGAG GAGCCAGACA GTGACATGA GTGGCAGCAC CTGCTAGAGA CCACTGAGCC
 GTCTCCACGA CCCGGTCGCC TCGTGACTAC ATCACCCTGA CCTTTTACTC CTCGGTCTGT CACTGTACT CACCGTCGTG GACGATCTCT GGTGACTCGG
 397 E V L G Q R S T D V V D L E N E E P D S D N E W Q H L L E T T E P
 1401 TGTGCTATT CAACTGAAGG CTCCTCTCAC CTTGCTGTGT AATCCTGACT TCTGCCAGCG CATCCAGAGT CAGCTGCATG AAGCTGGAGG GCAGATCCTG
 ACACGATAA GTTGACTTCC GAGGAGAGTG GAACGACACA TTAGGACTGA AGACGTCGC GTAGGTCCTCA GTCGACGTAC TTCGACCTCC CGTCTAGGAC
 430 V P I Q L K A P L T L L C N P D F C Q R I Q S Q L H E A G G Q I L
 1501 AAAGGCATCT TGGAGGTCG TTCCCACATC CTGCTGCAT TCCGGTCTCT GAGCAGTCTT CTCTCCAGCT GCAGTGATC TGTTCCTTG TATTCTTCT
 TTTCGCTAGA ACCTCCACG AAGGGTGTAG GACGACGTA AGGCCACGGA CTCGTAGAA GAGAGTGA CGTCACCTAG ACAACGGAAC ATAAGGAAGA
 463 K G I L E G A S H I L P A F R V L S S L L S S C S D S V A L Y S F C

Figura 5C

1601 GCCGGGAGC AGGCTTCTT GGGCTGCTGC TGAGTCTACT CAGCAGACAGT CAGGAGAGCA ACAGCCTCCA AGCCTGATCT TGGTATGGGA CCTTCTTTACA
 CGGCCCTCCG TCCGGAAGGA CCGGACGACG ACTCAGATGA GTCCCTCTCGT TGTGGAGGT CCGTCTTTAGA ACCATACCCCT GGAAGAAATGT
 497 R E A G L P G L L L S L L R H S Q E S N S L Q Q Q S W Y G T F L Q
 1701 GGACCTGATG CTTGTGATTC AGGCTTACTT TGCCTGTACC TTCAATCTGG AGAGGAGCCA GACAAGTGAC AGCCTGACGG TGTTTCAGGA GGCTGCCAAC
 CCTGGACTAC CGACACTAAG TCCGGATGAA ACGGACATGG AAGTAGACC TCTCCTCGGT CTGTTCACATG TCGGACCTCC ACAAGTCTCT CCGACGGTTG
 530 D L M A V I Q A Y F A C T F N L E R S Q T S D S L Q V F Q E A A N
 1801 CTTTTCCTGG ACCTGTGGG GAAACTGCTG GCCCAACCAG ATGACTCTGA GCAGACTTTG CGGAGGGACA GCCTTATGTG CTTTACTGTC CTGTGCGAAG
 GAAAGAGACC TGGACAACCC CTTTGACGAC CGGTTGGTC TACTGAGACT CGTCTGAAC GCCTCCCTGT CGGAATACAC GAAATGACAG GACACGCTTC
 563 L F L D L L G K L L A Q P D D S E Q T L R R D S L M C F T V L C E A
 1901 CCATGGATGG GAACAGCCGG GCCATCTCCA AAGCCTTTTA CTCCAGCTTG CTGACGACAC AGCAGTTGT CTTGGATGGG CTCCTTCATG GCTTGACAGT
 GGTACCTACC CTTGTGGGC CGGTAGAGGT TTCGAAAAT GAGTCTGAC GACTGCTGTG TCGTCCAACA GAACCTACCC GAGGAAGTAC CGAACTGTCA
 597 M D G N S R A I S K A F Y S S L L T T Q Q V V L D G L L H G L T V
 2001 TCCACAGTTC CTTGTCCACA CTCCCAAGG TTCCCTACTC CTGCTGCCAT GTCGGTGTG TTGTCTAGGG CAAGAGCCTC AGSCCTTTGG
 AGGTGTCGAG GGACAGGTGT GAGGGTTCC AAGGATGAG GACGACGTA CAGCCACTCA TGACCACGAT AACAGATCCC GTTCTCGGAG TCCGGAAACC
 630 P Q L P V H T P Q G S L L L L P C R O
 2101 AGTTACTCTT TGCCTTCTC CACAGAGCC CCGCAAGTGA GCCAGCCACT CGGAGAGCAG AGTGAGGATA TACCTGGAGC CATTTCTCTT GCCTTGGCAG
 TCAATGAGAA ACGAAAGAG GTGTCTCTCG GGCCTTCACT CGGTCTGTA CGCTCTGTC TCACTCTCTAT ATGAGCTCG GTAAAGGAGA CCGGACCGTC
 1 S Y S L L F S T G A P Q V S Q P L R E Q S E D I P G A I S S A L A A
 ^2nd ORF starts from here
 2201 CCATATGCAC TGCTCTGTG GGAATGCCG ACTGCTGGA TGCCAGGAG CAGGTCTGTT GGCATTTGGC AAATCAGCTA ACTGAAGACA GCAGCCAGCT
 GGTATACGTG ACGAGGACAC CCTGACGGC TGACGACCTT ACGGTTCTCTC GTCCAGACAA CCGTAAACCG TTATGTCGAT TGACTTCTCT CCGTGGTCCA
 35 I C T A P V G L P D C W D A K E Q V C W H L A N Q L T E D S S Q L
 2301 CAGGCCATCC CTCATCTCTG GCCTGCAGCA TCCCATCTCTG TGCCCTGCACC TTCTCAAGT TCTATACTCC TGCTGCTTG TCACTGAGGG CCTGTGCCGT
 GTCCGGTAGG GAGTAGAGAC CGGACGTCGT AGGATAGGAC ACGGACGTCG AAGAGTTCCA AGATATGAGG ACGACGGAAC AGTCACCTCC GGACACGGCA
 68 R P S L I S G L Q H P I L C L H L L K V L Y S C C L V S E G L C R
 2401 CTTCTGGGGC AGGAGCCCTT GGCCTTGGAA TCCCTGTTTA TGTGATTTCA GGGCAAGGTA AAAGTAGTAG ATTGGGAAGA GTCTACTGAA GTGACACTCT
 GAAGACCCCG TCCTCGGGGA CCGGAACCTT AGGACAAT ACACATAGT CCGGTTCCAT TTTCATCATC TAACCTTCT CAGATGACTT CACTGTGAGA
 101 L L G Q E P L A L E S L F M L I Q G K V K V V D W E E S T E V T L Y

Figura 5D

2501 ACTTCCTCTC CCTTCTTGTC TTTCGGCTCC AAAACCTGCC TTGTGAATG GAGAGCTAG GCAGTGACGT TGCTACTCTC TTTACCCATT CCATCTCGT
 TGAAGGAGAG GGAAGACAG AAAGCCGAGG TTTTGGAGG AACACCTTAC CTCCTCGATC CGTCACTGCA ACCATGAGAG AAATGGGTAA GCGTACAGCA
 135 F L S L L V F R L Q N L P C G M E K L G S D V A T L F T H S H V V
 2601 CTCCTTTGTG AGTCAGCAG CCTGTCTATT GGCACAGCTT GGTACCAAG GGTGACCTT TGACCTCCAG CCCATGGAAT GGATGGCTGC AGCCACACAT
 GAGAGAACHC TCACGTGTC GGACAGATAA CCCTGTGAA CCAGTCGTTT CCCACTGGAA ACTGGAGGTC GGTACCTTGA CTTACCCGACG TCGGTGTGTA
 168 S L V S A A C L L G Q L G Q Q G V T F D L Q P M E W M A A A T H
 2701 GCCTTGTCTG CCCCTGCAGA GCTCCTCACT GAGGTACAGA TGGATCTTGG GATGGATGG AAGTAAAGAG AGAGGAAGTG GGCATTTTGG GGAGCCTCTG
 CGGAACAGAC GGGGACGTCT CGAGGAGTGA CTCCTATGCT ACCTAGAACC CTACCTACCC TTGATTTCTC TCTCCTTGAC CCGTAAACC CCTCGGAGAC
 201 A L S A P A E L L T E V Q M D L G M D G K O
 2801 GACCAGAGGA ATGAAGRAG AACCCACAGC CTTCCCTCTC AAGCTACTGT GCCTGTGATA GCCTTGGAAC TTCCCCCGCT GGCCTCAGTA CTGACCCCTT
 CTGCTCTCCT TACTTCTTCG TTGGGTGTCG GAAAGGAGAG TTCGATGACA CGGACACTAT CGGAACCTTG AAGGGCGGA CGGAGTCAT GACTGGGAAA
 2901 GAAGGAACC ATTTCGCTGC TCCCTCTGGA TCCAGTGGGA GATAAATGA ATTCCTCTGG TTTTCAGCAGA CATACACATG AGTTGTGAGG TCAGAGGGTT
 CTTCCCTTTGG TAAGCGACG AGGGACCCCT AGGTCACCCCT CTATTTTACT TAAGGGACCC AAAGTCGTCT GTATGTGTAC TCAACACTCC AGTCTCCCAA
 3001 AAGGTTTGTAT AAGAAATGA AATAAGACGA CAGGGAATA CTAGTGGGA AAGCGGAAGG AATTATTTCT GGGACTTCCT TTACTTGTAA GTCAGGGACA
 TTCCAAACTA TTCTTTTACT TTATTTCTGCT GTCCCTTTAT GATCCACCC TTTCCCTTCC TTAAATAAGA CCTGAAGGA AATGAACATT CAGTCCCTGT
 3101 GGRATGRATA AAAGCATTTG GATTCCTGAC TTCTGTCTTT CCCCCGCC TCCTTCACTT TTATCTCTAG CAGGGGAAGG CTAGCCTAAT CAGGATATG
 CCTTACTTAT TTTCTGTAAC CTAAGGACTG AAGACAGAAA GGGGGCGGG AGAAAGTGAA AATAGAGATC GTCCCTCTCC GATCGGATTA GTCCCTATAC
 3201 TCCAGTTTTCAG AAATGTGGAC CGTTTGTGCG CACCGCTTCT CCATGGTCTT CCATGGTCTT GAGGCTCCCG GAGGAGGCAT CTGCACAGGA AGGGGAGCTT TCGCTATCCA
 AGGTCAAGTC TTTACACCTG GCAAAACACC GTGGCGAAGA GGTACCAGGA CTCCGAGGG CTCTCCCGTA GACGTCTCCT TCCCTTCGAA AGCGATAGGT
 3301 GTCCACCAG CCCTGAGCCA GACTGGACAC TGATTTCTCC CCAGGCGATG GCAGCCCTGC TGAGCCCTGC CATGGCCACC TTTACCCAGG AGCCCCAGTT
 CAGGTGGTTC GGGACTCGGT CTGACCTGTG ACTAAAGAGG GGTCCCGTAC CGTCGGGAGG ACTCGGACCG GTACCGGTGG AAATGGTCC TCGGGGTCAA
 3401 ATGCCCTGAG TCCCTGTCCC AGCATGGAAG TATCTCTATG TCCATCTCATG AGCATCTGCT TTGCCCCCAGC TTCCCTGAATC AACTGCGCCA GGCCTCTCAT
 TACGGACTCG ACGGACAGGG TCGTACCTTC ATAGGAGTAC AGGTAGGACT TCGTAGACGA AACGGGTGCG AAGGACTTAG TTGACCGGT TTGACCGGT CCGGGAGTA
 3501 GGTCTGAGT TTCTCCCTGT CGTGTGCTC TCTGTCTGCC AGCTCCTTTG CTTCCCTTTT GCGCTGGACA TGGATCTGA CCTCCTTATA GGTGTCTTGG
 CCCAGACTCA AAGAGGGACA GCACCAGGAG AGACAGACGG TCGAGGAAC GAAGGGGAAA CCGGACCTGT ACCTAGACT GGAGGAATAT CCACAGAAC

Figura 5E

3601 CCGACCTCAG GGACTCAGAA GTTGACGCC ATCTCTCTGCA GGTCGTCTGC TACCATCTTC CGTTGATGCA AGTGGAGCTG CCCATCAGCC TTCTCACACG
 GGCCTGGAGTC CCTGAGTCTT CAACCTCGG TAGACACAGT CCAGACGACG ATGGTAGAAG GCAACTACGT TCACCTCGAC GGGTAGTCGG AAGAGTGTGC
 3701 CCTGGCCCTC ATGATCCCA CCTCTCTCAA CCAGTTTGTG AACACAGTGT CTGCCTCCCC TAGAACCATC GTCTCGTTTC TCTCAGTTGC CCTCTGAGT
 GGACCGGGAG TACCTAGGCT GGAGAGAGTT GGTCAACAC TTGTCTACA GACGAGGGG ATCTTGGTAG CAGACAAAG AGAGTCAACG GGAGGACTCA
 3801 GACCAGCCAC TGTGACCTC CGACCTTCTC TCTCTGCTGG CCCATCTGC CAGGTCTCTG TCTCCAGCC ACTTGTCCTT TATCCAAGAG CTTCTGGCTG
 CTGGTCGGTG ACACTGGAG GCTGAAGAG AGAGACGACC GGGTATGACG GTCCACGGAC AGAGGTCTGG TGAACAGGAA ATAGGTCTC GAAGACCGAC
 3901 GCTCTGATGA ATCTATCGG CCCCTGCGCA GCCTCTGGG CCACCCAGAG AATTCTGTGC GGGACACAC TTATAGGCTC CTGGGACACT TGCTCCAACA
 CGAGACTACT TAGGATAGCC GGGGACCGCT CGGAGGACCC GGTGGTCTC TTAAGACACG CCCGTGTGTG AATATCCGAG GACCTGTGTA ACGAGGTGT
 4001 CAGCATGGCC CTGGTGGG CACTGCAGAG CCAGTCTGGA CTGCTCAGCC TTCTGCTGCT TGGGTTTGA GACAAGGATC CTGTTCTGCG GTGCAGTGCC
 GTCGTACCGG GACGACCCC GTGACGTCTC GGTCAAGCT GACGAGTCGG AAGACGACGA ACCCGAACCT CTGTTCTCTAG GACAACACGC CACGTCAACG
 4101 AGCTTTGCTG TGGCAATGC AGCTTACCAG GCTGTCTCTC TGGGACCTGC CCTGGCAGCT GCAGTCCCCA GTATGACCCA GCTGCTTGA GATCCTCAGG
 TCGAAACGAC ACCGTTACG TCGGATGGTC CGACCAAGAG ACCCTGGACG GGACCTCGA CGTCACGGGT CATACTGGGT CGACGAACCT CTAGGAGTCC
 4201 CTGGTATCCG GCGCAATGTT GCATCAGCTC TGGCAACTT GGGCAACTT GGAOCTGAA GGTTTGGGAG AGGACTGTT ACAGTCCGAA GTACCCACG GCCTCCTAGA
 GACCATAGGC CGCTTACRA CGTAGTCGAG ACCCTTGAA CCTTGGACTT CCAAACCCCTC TCCTCGACAA TGTCACGCTT CATGGGTCG CCGAGGATCT
 4301 AATGGCATGT GGACACCCC AGCCAAATGT GAAGGAGGCT GCCCTCATTG CCCTCCGGAG CCTGCAACAG GAGCCTGGCA TCCATCAGGT ACTGGTGTCC
 TTACCGTACA CCTCTGGGG TCGGTTTACA CTTCCTCCGA CGGAGTAAAC GGGAGGCTC GGACGTTGTC CTCGGACCGT AGGTAGTCCA TGACCAACAG
 4401 CTGGGTGCCA GTGAGAACT ATCCTTGGT TCTCTGGGA ATCAGTCACT GCCACACAG AGTCTTAGGC CTGCCCTCTGC CAAACACTGC AGGAACTCA
 GACCCACGGT CACTCTTTGA TAGGAACGAG AGAGACCCCT TAGTCAGTGA CGGTGTGTG TCCAGGATCCG GACGGAGACG GTTTGTGACG TCCTTTGAGT
 4501 TTCACCTCCT GAGGCCAGCC CATAGCATGT GATTCAGAT TCCTGGGTC CAGCTCCAA CTTTGGTTGC CAGCTCTTTC TTATCTTACT ACACAAGCCG
 AAGTGGAGGA CTCGGTCCG GTATCGTACA CTAAGTCTA AGGACCCAG AGGACCCAG GTCGGAGGTT GAAACCAACG GTCGAGAAAG AATAAGATGA TGTGTTCGGC
 4601 CCAACTCAAC TGAGAGCTAA AGAGACTAGA AAAGAGATAA GCTGCCAAT CCACTGAGAA CAAGAACTA GAAGAGATTT ATATATAAAG CTTCTTCCTT
 GGTGAGTTG ACTCTCGATT TCTCTGATCT TCTCTTAT TTTCTCTATT CGACGGTTGA GTTGACTCTT GTTCTCTAAA TATATATTTC GAAGAAGGAA

Figura 5F

```

4701 CTCCAGATG CAGGATGTTT TCAACCAGTA AATTTTATTG CTGTTGGTGC CAGAGAAGAG TCCTTTCTTC TCTACATCCA GGGGCTTTT CTCCAATAAT
      GAGGCTCTAC GTCTACAAA AGTTGGTCAT TTAAAATAAC GACAACCACG GTCTCTTCTC AGGAAAGAAG AGATGTAGGT CCCCGGAAAA GAGTTATTA
4801 GTGCCTTTAA CTCTAGGGAC CTGCCTCAGC GACCTTAGGG AAAAACCTCA ACCTGAAAGA TCCTTTCTCT TCTGGAGCTC CTTTAATCTT CCCAGCAGGT
      CACGGAAAT GAGATCCCTG GACGGAGTGC CTGGAATCCC TTTTGGAGT TGGACTTTCT AGAGAAGGAA AGACCTCGAG GAAATTAGAA GGGTCGTCCA
4901 TTTTGCCTTA GACGTGCTGG CCCCAGGACA GTGATGAAGA CAGAGCCTGT CTCAGCTCTA GGCTGTGGGG ATCAATGCCA TCAGTCCCCTG TTATTGAGGG
      AAAACGGAAAT CTGCACGACC GGGGTCTGT CACTACTTCT GTCTCGGACA GAGTCGAGAT CCGACACCCC TAGTTACGGT AGTCAGGGAC AATAACTCCC
5001 ATTATCCCTT AGCCAACATT CCTATCTGTG GGTGGGCGTG GAGAGTGAT CTTTTTTTGG GGTGTGTGTG TATATGTGTG TGTGTATGTG TGTGTGTGT
      TAATAGGGAA TCGGTTGTAA GGATAGACAC CCACCCGCAC CTCTCACATA GAAAAAAACC CCACACACAC ATATACACAC ACACACACAA
5101 TAATAGTTCT GTTTGTAAAC TCTTTTAATA AAAGTTGTGC CTCACCATAC TTGAAGCTCC CAGGACAAGG GTTGAGAGGC TCAACCCCTC TTTCAGCTTC
      ATTATCAGA CAACATTTG AGAAATTAAT TTTCACACAG GAGTGGTATG AACTTCGAGG GTCTGTCTCC CAACTCTCCG AGTTGGGGAG AAAGTCGAAG
5201 TATGTGTGTG TGGAGGTGCT GGTATCGTGT TCACACAAAA AAAAAAAAAA AA
      ATACACCACA ACCTCCACGA CCATAGCACA AGTGTGTTT TTTTTTTTTT TT

```

Figura 6

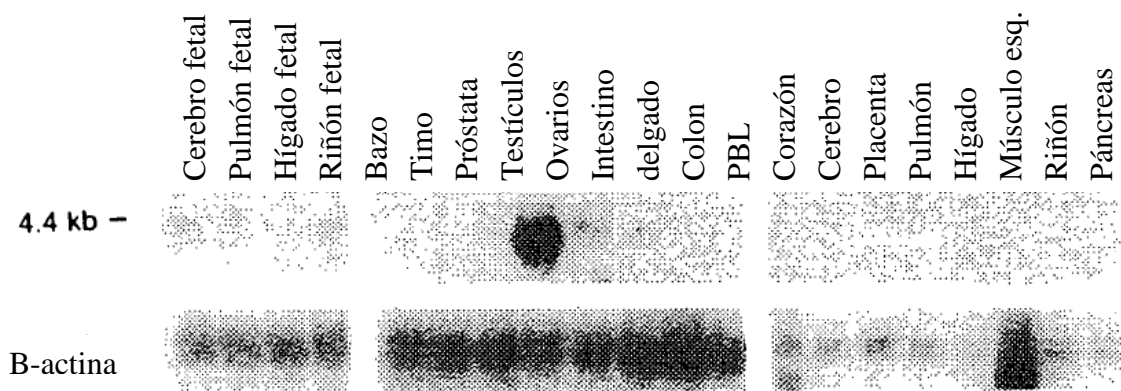
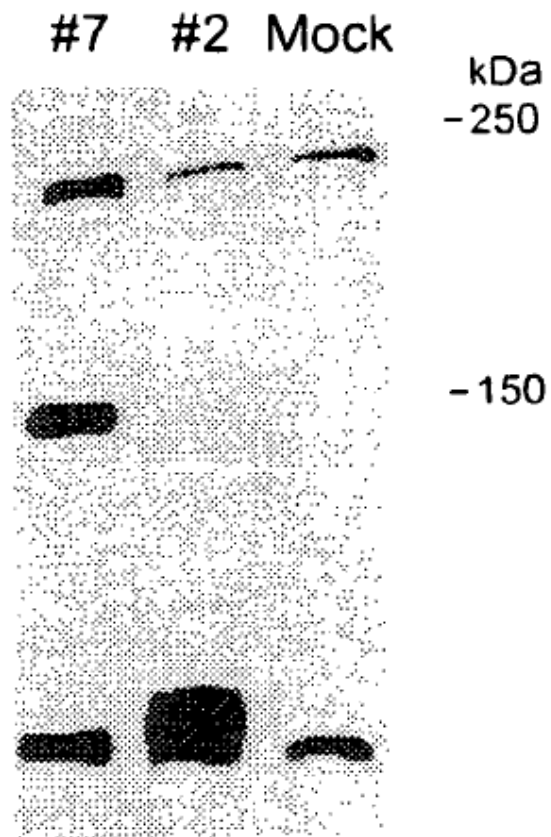


FIG. 7

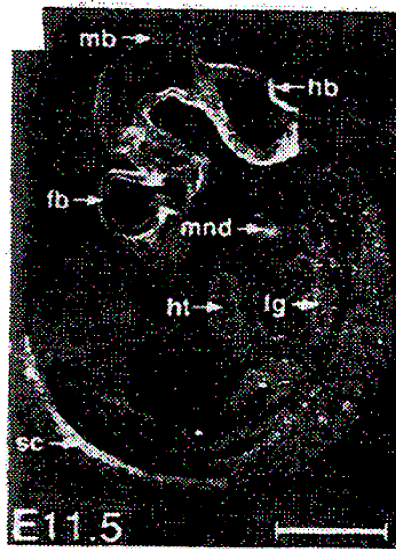


FIG. 8A

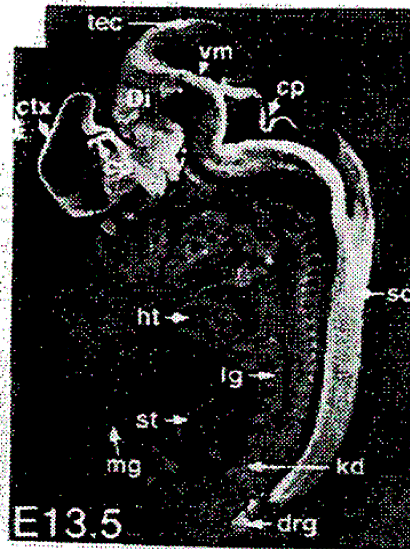


FIG. 8B



FIG. 8C

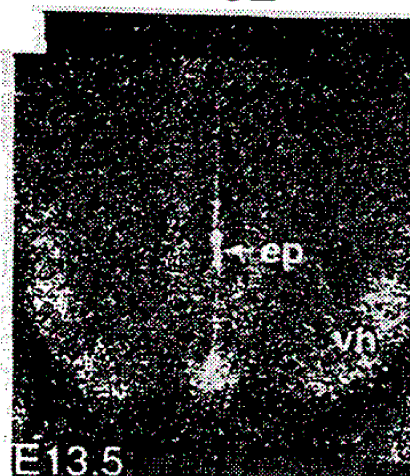


FIG. 8D

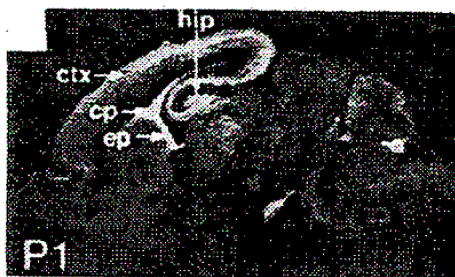
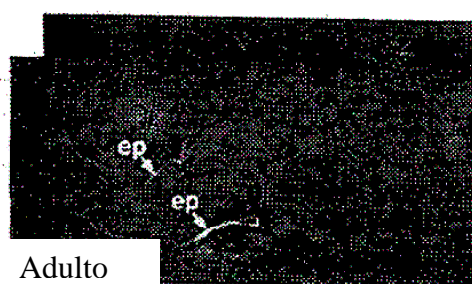


FIG. 8E



Adulto

FIG. 8F

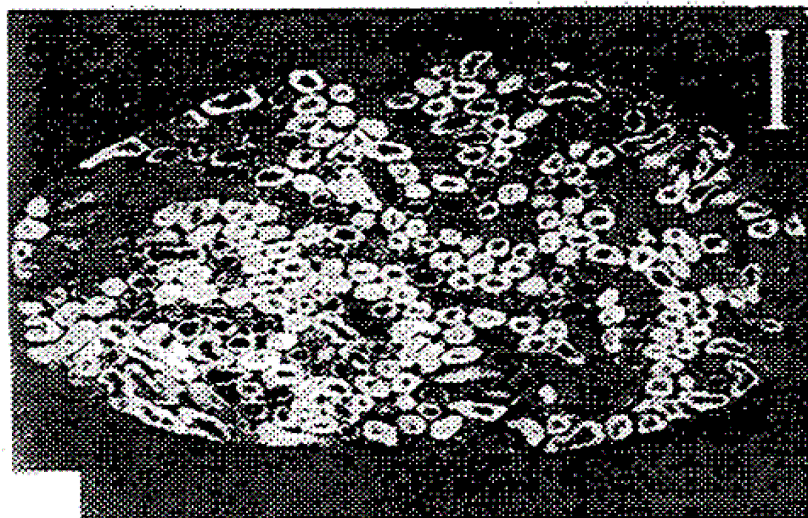


FIG. 9A

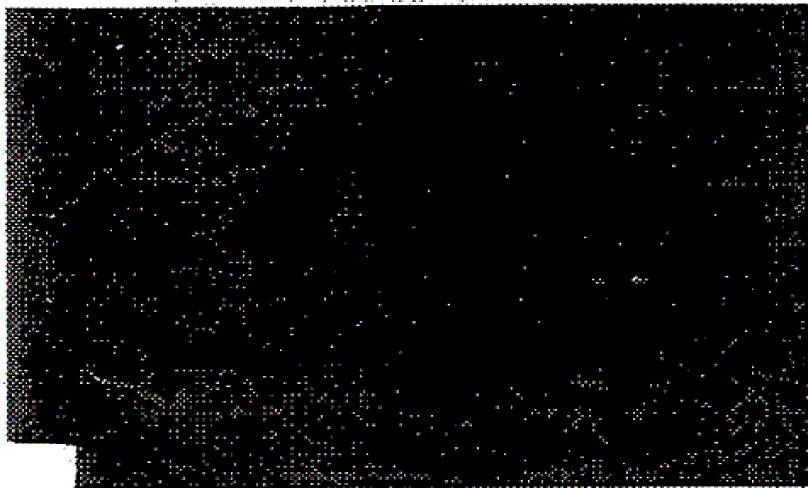


FIG. 9B

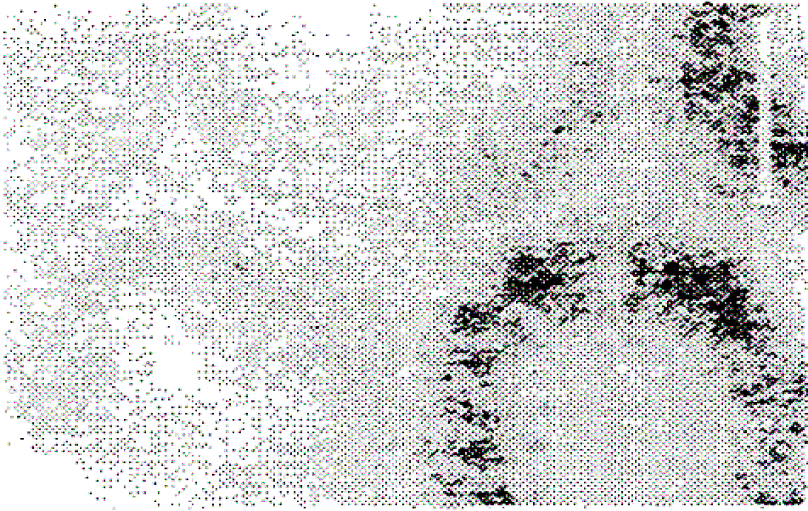


FIG. 9C

Figura 10A

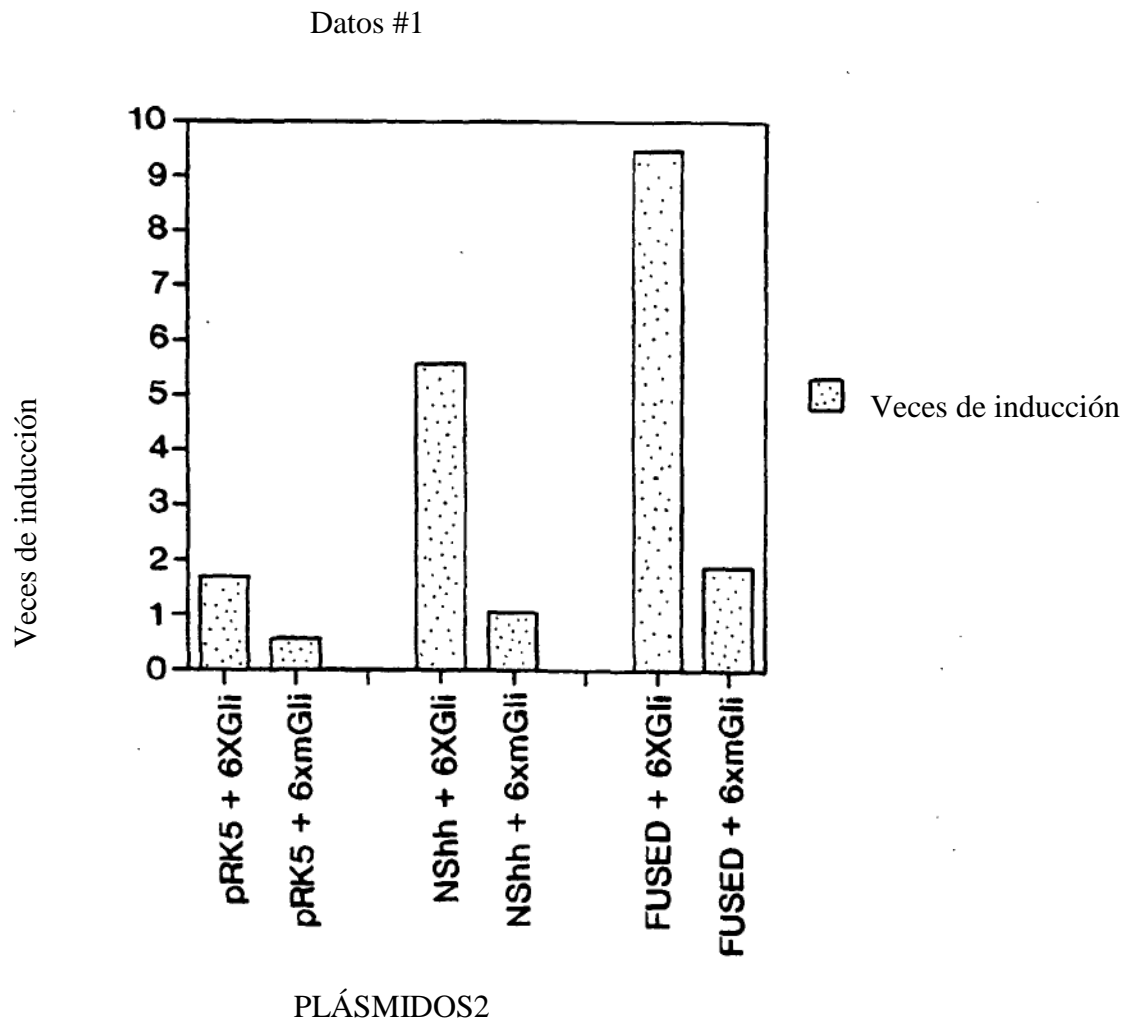


Figura 10B

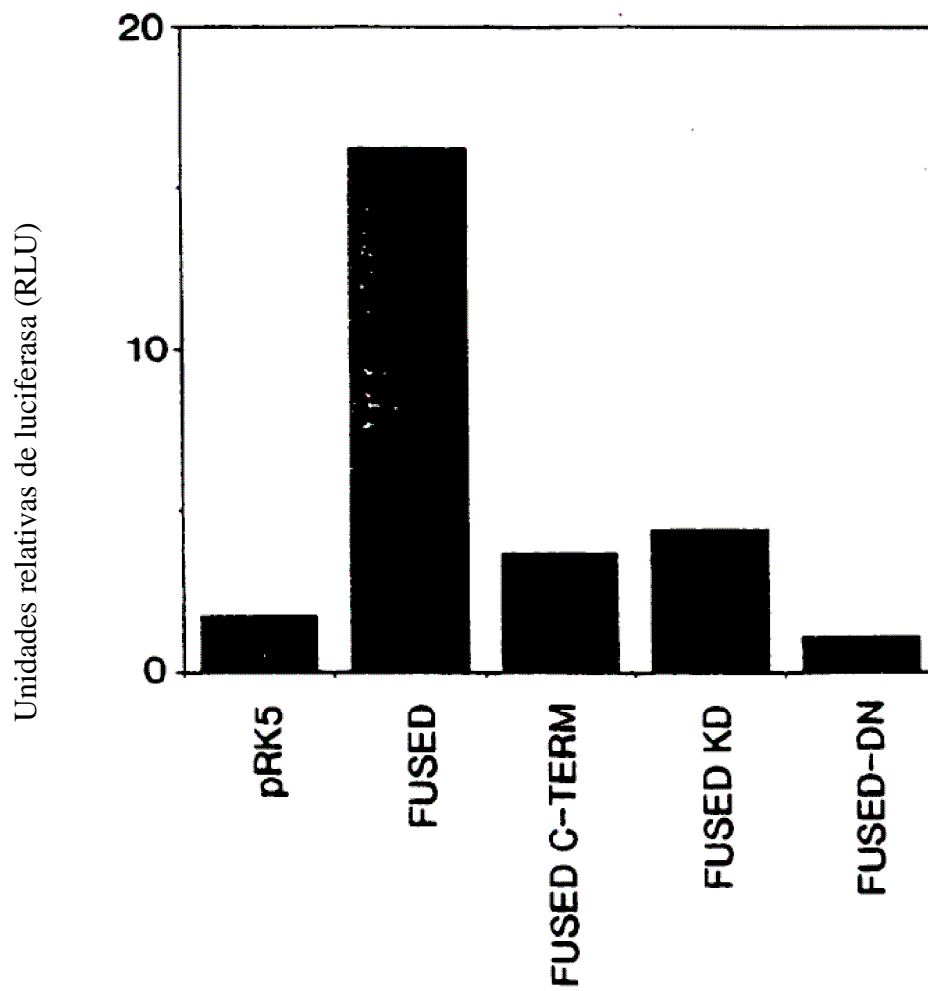




FIG. 11A

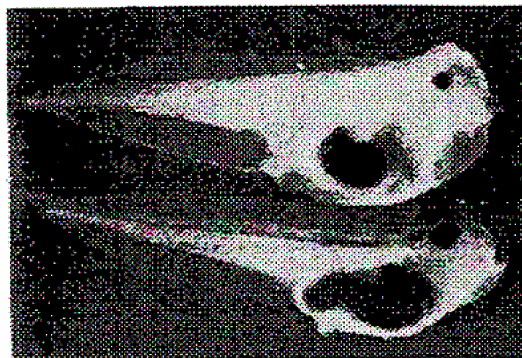


FIG. 11B

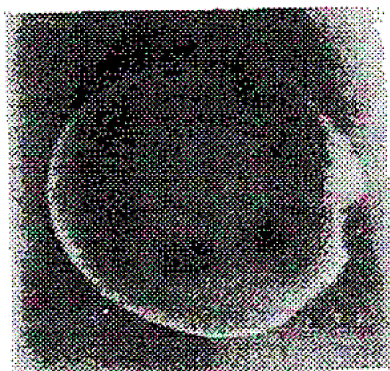


FIG. 11C

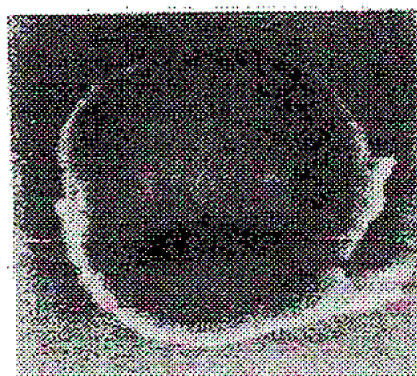


FIG. 11D

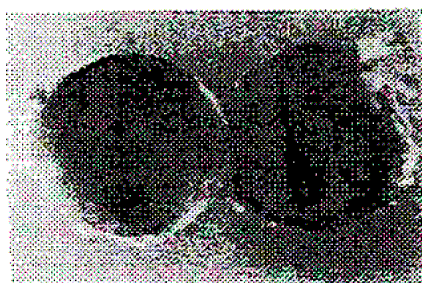


FIG. 11E

Figura 12

