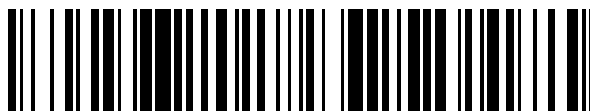


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 511 694**

51 Int. Cl.:

C07K 14/47 (2006.01)
C12N 15/12 (2006.01)
A61K 38/17 (2006.01)
C12N 5/10 (2006.01)
C12Q 1/68 (2006.01)
C07K 16/18 (2006.01)
C12N 15/11 (2006.01)
A61K 39/395 (2006.01)
A61P 9/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.03.2003 E 03727609 (4)**

97 Fecha y número de publicación de la concesión europea: **09.07.2014 EP 1483289**

54 Título: **Genes reguladores de la angiogénesis, preparaciones farmacéuticas que los contienen y sus aplicaciones**

30 Prioridad:

22.03.2002 FR 0203655

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

23.10.2014

73 Titular/es:

**GENE SIGNAL INTERNATIONAL SA (100.0%)
Parc Scientifique EPFL, PSE-A
1015 Lausanne VD, CH**

72 Inventor/es:

**AL-MAHMOOD, SALMAN;
COLIN, SYLVIE y
SCHNEIDER, CHRISTOPHE**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 511 694 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

Genes reguladores de la angiogénesis, preparaciones farmacéuticas que los contienen y sus aplicaciones

La presente invención pertenece al dominio de las composiciones farmacéuticas útiles para el tratamiento de patologías que resultan de una desregulación del mecanismo de la angiogénesis.

5 Se refiere a composiciones farmacéuticas que comprenden por una parte secuencias de nuevos genes, cuya función no ha sido identificada hasta la fecha y cuya implicación en el mecanismo de la angiogénesis ha sido demostrada por primera vez por el solicitante y por otra parte secuencias de genes, de los cuales al menos una de sus funciones ha sido identificada previamente, pero cuya implicación como genes que codifican los constituyentes celulares de la célula endotelial en el mecanismo de la angiogénesis, ha sido demostrada por primera vez en los trabajos realizados por el solicitante en el marco de la presente invención. Estos genes se identifican por sus secuencias nucleotídicas en la lista de secuencias adjunta. La presente invención se refiere también a las secuencias polipeptídicas de los factores codificados por dichos genes que encuentran aplicación en el estudio clínico del proceso de la angiogénesis, en el pronóstico, diagnóstico y tratamiento de patologías ligadas a este proceso, así como en la realización de ensayos farmacológicos, farmacogenómicos, o farmacosignálicos.

15 La angiogénesis es un proceso fundamental por el cual se forman nuevos vasos sanguíneos. Este proceso es esencial en varios fenómenos fisiológicos normales tales como la reproducción, el desarrollo o incluso la cicatrización. En estos fenómenos biológicos normales, la angiogénesis está bajo estricto control, es decir, que se desencadena durante un breve período, algunos días, y después se inhibe completamente. Sin embargo, diversas patologías están relacionadas con una angiogénesis invasiva e incontrolada. La artritis, por ejemplo, es una patología debida al daño causado a los cartílagos por los neovasos invasivos. En la retinopatía diabética, la invasión de la retina por los neovasos lleva a la ceguera de los enfermos; la neovascularización del aparato ocular es la principal causa de ceguera y esta neovascularización es predominante en una veintena de enfermedades oculares. Por último, el crecimiento y la metástasis de los tumores están directamente relacionados con la neovascularización y son dependientes de la angiogénesis. El tumor estimula el crecimiento de los nuevos vasos sanguíneos para su propio crecimiento. Además, estos neovasos son las vías de escape de los tumores para alcanzar el sistema circulatorio sanguíneo y provocar las metástasis en sitios remotos tales como el hígado, el pulmón o los huesos.

20 En otras patologías tales como las enfermedades cardiovasculares, las enfermedades de las arterias periféricas y las lesiones vasculares o cerebrales, la angiogénesis puede presentar una base terapéutica importante. En efecto, la promoción de la angiogénesis en las áreas dañadas puede llevar a la formación de nuevos vasos sanguíneos laterales y alternativos a los vasos dañados, suministrando de este modo la sangre y por lo tanto el oxígeno y otros factores nutrientes y biológicos necesarios para la supervivencia de los tejidos afectados.

25 La formación de nuevos vasos sanguíneos por las células endoteliales implica la migración, el crecimiento, y la diferenciación de las células endoteliales. La regulación de estos fenómenos biológicos está directamente relacionada con la expresión genética. En términos de angiogénesis, un número creciente de estudios demuestran que la regulación de la angiogénesis se realiza a través de un equilibrio entre los factores que actúan directamente sobre las células endoteliales. Estos factores pueden ser estimulantes, por una parte, tales como, entre otros, el VEGF (factor de crecimiento endotelial vascular), el FGF (factor de crecimiento de fibroblastos), la IL-8 (interleucina 8), el HGF/SF (factor de crecimiento de hepatocitos/factor disperso), el PDGF (factor de crecimiento derivado de plaquetas). Los factores también pueden ser inhibidores, tales como, entre otros, IL-10, IL-12, gro- α y gro- β , el factor plaquetario 4, la angiostatina, el inhibidor derivado de condrocitos humanos, la trombospondina, el factor inhibidor de la leucemia. (Jensen, 1998, Surg. Neural., 49, 189-195; Tamatani *et al.*, 1999, Carcinogenesis, 20, 957-962; Tanaka *et al.*, 1998, Cancer Res., 58, 3362-3369; Ghe *et al.*, 1997, Cancer Res., 57, 3733-3740; Kawahara *et al.*, 1998, Hepatology, 28, 1512-1517; Chandhuni *et al.*, 1997, Cancer Res., 57, 1814-1819; Jendraschak and Sage, 1996, Semin. Cancer Biol., 7, 139-146; Majewski *et al.*, 1996, J. Invest. Dermatol., 106, 1114-1119).

35 La regulación de la angiogénesis tal como se describe actualmente se realiza a través de un equilibrio de dos tipos de factores:

- los factores angiogénicos (polipéptidos extracelulares, en su mayoría mitógenos) que actúan directamente sobre las células endoteliales induciendo la angiogénesis.
- los factores angiostáticos (polipéptidos extracelulares, también en su mayoría mitógenos o que actúan sobre la mitogénesis) que actúan también directamente sobre las células endoteliales para inhibir la angiogénesis.

40 El equilibrio entre estos dos tipos de factores extracelulares regula la angiogénesis. Se observa en este punto, que el control de la angiogénesis se realiza por la producción de factores angiogénicos y factores angiostáticos. Por ejemplo, ya se ha demostrado que la estimulación de las células endoteliales por un factor angiogénico induce la expresión 1) del activador del plasminógeno tipo urocinasa (uPA) y de su inhibidor PAI-I (Pepper *et al.*, 1990, J. Cell Biol. 111(2), 743-55; Pepper *et al.*, 1996, Enzyme Protein, 49(1-3), 138-62), 2) de las metaloproteinasas de la matriz (las MMP) y de inhibidores fisiológicos de la actividad de estas MMP (TIMPs) (Cornelius *et al.*, 1995, J. Invest. Dermatol., 105(2), 170-6; Jackson and Nguyen, 1997, Int. J. Biochem. Cell Biol. 29(10), 1167-1177), 3) de los

inhibidores tales como la angiopoyetina-2 (Ang-2) o la trombospondina-1 (TSP-1) (Mandriota and Pepper, 1998, *Circ. Res.* 83, 852-859; Oh *et al.*, 1999, *J. Biol. Chem.*, 274(22), 15732-9; Suzuma *et al.*, 1999, *American Journal of Pathology*, 154, 343-354). Por lo tanto, parece que normalmente una célula endotelial en el estado angiogénico produce factores angiogénicos, pero también factores angiostáticos. La producción de estos últimos permite controlar la angiogénesis.

En paralelo a esta operación, las células endoteliales estimuladas por un factor angiostático producen factores angiostáticos pero también factores angiogénicos para controlar el estado angiostático. Este fenómeno ha sido ya descrito en otros tipos de células productoras de factores implicados en la angiogénesis, que cuando son estimuladas por un factor angiostático tal como el interferón-gamma producen factores angiogénicos y factores angiostáticos (Kobayashi *et al.*, 1995, *Immunopharmacology* 31(1), 93-101; Arkins *et al.*, 1995, *Mol. Endocrinol.*, 9(3), 350-60; Kodelja *et al.*, 1997, *Immunobiology*, 197(5), 478-93).

En la presente invención, el solicitante ha demostrado que la estimulación de las células endoteliales por un factor angiostático lleva a la expresión de genes que codifican los constituyentes celulares implicados en la promoción de la angiogénesis.

Por lo tanto, la estimulación de las células endoteliales por un factor angiogénico o por un factor angiostático puede inducir tanto una como otra, la expresión de reguladores positivos y negativos de la angiogénesis.

El control de la angiogénesis representa por tanto un eje estratégico, a la vez que una investigación fundamental, con el fin de mejorar la comprensión de numerosos fenómenos patológicos asociados con la angiogénesis, es también una base para el desarrollo de nuevas terapias destinadas a tratar las patologías asociadas con la angiogénesis.

Con el fin de controlar la angiogénesis, varios grupos farmacéuticos han desarrollado estrategias terapéuticas basadas directamente en el uso de señales paracrinas, los factores estimuladores e inhibidores, como agentes para promover o inhibir la angiogénesis. Estas estrategias se basan esencialmente en el uso de estos factores bajo su forma polipeptídica como agentes estimuladores o inhibidores de la angiogénesis, o incluso más recientemente, bajo la forma de vectores de expresión que codifican los factores seleccionados.

Se ha desarrollado un procedimiento que permite la identificación de nuevos genes implicados en la regulación de la angiogénesis. Dicho procedimiento ha sido objeto de una solicitud de patente francesa publicada con el número FR2798674 y de una solicitud de patente internacional PCT publicada con el número WO01/21831. Este método tiene la particularidad de traducir fielmente el mecanismo íntimo que regula la angiogénesis teniendo en cuenta todos los factores extracelulares descritos como agentes reguladores de la angiogénesis, es decir, los factores angiogénicos, los factores angiostáticos, así como los diferentes componentes de la matriz extracelular. Esta metodología consiste en emplear estos diferentes factores extracelulares a través de cuatro condiciones experimentales bien definidas. Las células endoteliales se cultivan sobre un componente y/o una mezcla bien definida de varios componentes de la matriz extracelular y se ponen bajo las cuatro condiciones experimentales, a saber:

Una condición de control en la que las células endoteliales no son estimuladas.

Una condición angiogénica en la que las células endoteliales son estimuladas por uno o más factores angiogénicos.

Una condición de inhibición de la angiogénesis, en la que las células endoteliales son estimuladas por uno o más factores angiogénicos y puestas en presencia de uno o más factores angiostáticos.

Otra condición de control en la que las células endoteliales son estimuladas por uno o más factores angiostáticos.

Estas cuatro condiciones permiten obtener preparaciones de ARNm específicas de la angiogénesis a saber, del estado angiogénico y/o la inhibición de la angiogénesis y permiten detectar los genes que codifican los constituyentes celulares implicados en la regulación de la angiogénesis, incluyendo los reguladores positivos y los reguladores negativos. Así pues, el procedimiento descrito antes permite el cribado sistemático de todos los factores angiogénicos y angiostáticos así como de los diferentes componentes de la matriz extracelular con el fin de revelar e identificar los genes que codifican los constituyentes celulares implicados en la regulación de la angiogénesis. Además, dado que la expresión génica se puede analizar a lo largo de la cinética de la formación de nuevos vasos sanguíneos por las células endoteliales, este enfoque constituye una metodología *in vitro* que permite relacionar la expresión génica con los parámetros funcionales biológicos de la angiogénesis.

La identificación de los cinco genes descritos más adelante se ha llevado a cabo según la metodología descrita anteriormente, utilizando los factores angiogénicos y angiostáticos, así como el colágeno tipo I como componente de la matriz extracelular para reproducir las cuatro condiciones experimentales.

El solicitante ha demostrado la implicación de estos cinco nuevos genes identificados por las secuencias SEQ ID N° 1 a SEQ ID N° 5 de la lista de secuencias adjunta, en el mecanismo de regulación de la angiogénesis.

Según un modo particular de realización, la composición farmacéutica de la invención comprende como compuesto activo al menos una secuencia nucleotídica tal como la descrita en la reivindicación 1.

5 Según la presente invención, se deben considerar como secuencias equivalentes, las secuencias nucleotídicas que presentan modificaciones estructurales menores, que no modifican su función, tales como deleciones, mutaciones o adiciones de bases, cuya identidad es al menos del 90 % con las secuencias nucleotídicas identificadas con los números SEQ ID N° 1 a SEQ ID N° 5 y SEQ ID N° 27 a SEQ ID N° 30 en la lista de secuencias adjunta.

Según otro modo particular de realización, la composición farmacéutica, reguladora de la angiogénesis, comprende al menos una secuencia inhibidora de la angiogénesis.

10 Según un modo particular de realización, la composición farmacéutica de la invención comprende una o más secuencias inhibidoras de la angiogénesis tal como la descrita en la reivindicación 2.

Preferiblemente, la composición farmacéutica de la invención comprende una secuencia antisentido seleccionada de la SEQ ID N° 11, en la lista de secuencias adjunta.

15 La invención tiene también por objeto una composición farmacéutica destinada al diagnóstico, al pronóstico y/o al tratamiento de patologías asociadas con la angiogénesis, caracterizada porque comprende al menos una secuencia polipeptídica tal como la descrita en la reivindicación 1.

20 Según la presente invención, se deben considerar como secuencias equivalentes, las secuencias polipeptídicas que presentan modificaciones estructurales menores, que no modifican su función, tales como deleciones, mutaciones o adiciones de residuos de aminoácidos, cuya identidad es al menos del 85 %, preferiblemente al menos del 90 % con las secuencias polipeptídicas identificadas con los números SEQ ID N° 6 a SEQ ID N°10 o con las secuencias polipeptídicas identificadas con los números SEQ ID N° 31 a SEQ ID N° 34 en la lista de secuencias adjunta.

Otro objeto de la invención se refiere a las secuencias nucleotídicas antisentido de las secuencias nucleotídicas identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 25 en la lista de secuencias adjunta.

25 En el marco de la presente invención, se comprende por secuencia antisentido: toda secuencia de ADN de al menos 10 meros, complementaria de un ARNm que al hibridarse con una porción de la secuencia de ARNm inhibe su expresión, es decir, su traducción en una proteína.

Muy preferiblemente, la invención tiene por objeto las secuencias antisentido que tienen una identidad al menos del 85 %, preferiblemente al menos del 95 % con una secuencia seleccionada entre las secuencias identificadas con los números SEQ ID N° 11 a SEQ ID N° 14 en la lista de secuencias adjunta.

30 La invención tiene también por objeto un vector de expresión de mamífero que comprende al menos una secuencia antisentido definida anteriormente.

Según modos de realización preferidos, dicho vector de expresión de mamífero comprende al menos una secuencia antisentido de al menos una de las secuencias identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta, así como un promotor que permite la expresión de dicho ADN antisentido.

35 Más particularmente, dicho vector se selecciona entre el grupo de vectores GS-V1, identificados por su secuencia que lleva el número SEQ ID N° 15 en la lista de secuencias adjunta.

Además, la introducción de dicha secuencia antisentido identificada con el número SEQ ID N° 11 en la lista de secuencias adjunta, o de uno de sus homólogos en vectores de expresión de mamífero y la inserción posterior de dichos vectores en las células de mamífero permite la obtención de líneas celulares que sub-expresan los genes que intervienen en el mecanismo de la angiogénesis.

40 También, para la construcción de estos vectores, se diseñan cebadores específicos para cada una de las secuencias identificadas. Estos cebadores comprenden en sus extremos los sitios de restricción no contenidos en el fragmento clonado y presentes en la región de multisitios del vector de expresión.

Los cebadores particularmente preferidos se indican en la Tabla I (véase Material y métodos) y se identifican con los números de secuencias SEQ ID N° 19 a SEQ ID N° 26 en la lista de secuencias adjunta.

45 La invención tiene también por objeto un vector de expresión de mamífero que comprende al menos una secuencia nucleotídica seleccionada del conjunto de secuencias identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta, o uno de sus fragmentos.

50 Estas construcciones son útiles, por una parte para preparar composiciones terapéuticas destinadas al tratamiento, por terapia celular, de trastornos angiogénicos, y por otra parte para verificar la eficacia de un tratamiento de un trastorno angiogénico en un mamífero, especialmente en un ser humano, o incluso para verificar la funcionalidad de los genes eventualmente implicados en el mecanismo de la angiogénesis, en dicho mamífero.

Otro objeto de la invención es un método de preparación de una proteína codificada por al menos uno de los genes cuyas secuencias se identifican con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta o uno de sus fragmentos.

5 Estas proteínas identificadas por las secuencias SEQ ID N° 6, SEQ ID N° 31 y SEQ ID N° 32 en la lista de secuencias adjunta, o sus fragmentos, se pueden producir bajo la forma de proteínas recombinantes *in vitro*, introduciendo en un hospedante adecuado un vector de expresión correspondiente adecuado, y después se purifican y se utilizan a continuación como un agente terapéutico.

Dicho método de preparación de una proteína recombinante comprende las etapas de:

10 a) la construcción de un vector de expresión que comprende al menos una secuencia entre las identificadas con los números SEQ ID N° 1 a SEQ ID N° 5 y SEQ ID N° 27 a SEQ ID N° 30 en la lista de secuencias adjunta o uno de sus fragmentos.

b) la introducción de dicho vector en un hospedante celular,

c) el cultivo de dichas células en un medio adecuado,

d) la purificación de las proteínas expresadas o de uno de sus fragmentos.

15 La invención se refiere también a una proteína recombinante obtenida por el método anterior.

La invención tiene también por objeto una composición terapéutica que comprende dicha proteína recombinante.

A título de ejemplo, los sistemas de expresión de proteínas recombinantes en bacterias tales como *E. coli* se pueden utilizar para expresar proteínas o polipéptidos no glucosilados.

20 La secuencia codificante o una secuencia parcial del gen de interés se amplifica por PCR utilizando cebadores específicos de este gen que tienen en los extremos sitios de enzimas de restricción con preferencia diferentes para permitir la orientación del gen en el vector de expresión. El ADN amplificado se purifica y después se digiere por las enzimas de restricción y a continuación se inserta mediante unión al vector de expresión previamente digerido por las mismas enzimas de restricción. Se pueden utilizar un gran número de vectores, tales como el vector pBR322 (Bolivar *et al.*, Gene 2 (1977) 95-113) o sus derivados, que contienen, por ejemplo, el promotor de la ARN polimerasa del bacteriófago T7 para un alto nivel expresión, o como el plásmido pET3a (Studier and Moffatt, 1986, J. Mol. Biol., 189(1): 113-30), que contiene preferiblemente las secuencias que codifican los marcadores de selección (resistencia a los antibióticos), un sitio múltiple de clonaciones que contiene los sitios de enzimas de restricción adecuados para la inserción del ADN, el sistema de célula/hospedante es con preferencia un sistema inducible tal como el utilizado para el radiomarcado *in vivo* del factor de crecimiento FGF2 (Colin *et al.*, 1997, Eur. J. Biochem., 249, 473-480), y previamente descrito por Patry *et al.* (1994, FEBS Lett, 349 (1): 23-8), y puede contener también una región que codifica una cola de polihistidina en el extremo del polipéptido de interés para facilitar la purificación.

35 El ADN amplificado se liga en el plásmido que se transforma en la bacteria según el método descrito por Sambrook *et al.* (1989, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.). Las células transformadas se extienden sobre el medio agar LB que contiene los antibióticos, las colonias resistentes a los antibióticos se controlan entonces por PCR y después se analizan en gel. El ADN plasmídico se puede aislar y después secuenciar para confirmar la construcción. La producción y la purificación de la proteína recombinante se pueden realizar como se ha descrito (Patry *et al.*, 1994, FEBS Lett, 349(1): 23-8), 473-480). Brevemente, una colonia aislada se inocula en el medio de cultivo líquido tal como el medio caldo LB con la adición de antibióticos. Después de una noche de incubación, se puede utilizar el pre-cultivo para sembrar un cultivo de mayor volumen. Se induce entonces la expresión del polipéptido, se desarrollan las células durante unas horas y después se recogen por centrifugación. El sedimento celular se puede lisar por agentes químicos conocidos en la técnica o mecánicamente, por ejemplo, por sonicación. La proteína se puede purificar gracias a sus propiedades fisicoquímicas como se describe para la purificación de FGF2 recombinante (Colin *et al.* 1997 Eur. J. Biochem., 249, 473-480) o si la proteína está marcada con una cola de polihistidina, se puede purificar a través de esta cola por 45 inmovilización sobre un soporte quelante de iones metálicos como se ha descrito previamente (Tang *et al.*, Protein Expr. Purif. 1997 Dec. 11(3): 279-83).

A título de ejemplo, se pueden utilizar sistemas de expresión de proteínas recombinantes para expresar polipéptidos que tienen modificaciones post-traduccionales como la glucosilación, tales como sistemas eucariotas (levaduras, plantas, insectos).

50 Así pues, la proteína recombinante se puede producir por ejemplo en la levadura *Pichia pastoris* como ha sido descrito por Sreerikshna *et al.* (1988, J. Basic Microbiol. 28(4): 265-78). El ADN amplificado será introducido de la misma forma después de digestión y ligación en un vector de expresión de *Pichia pastoris*, que contiene preferiblemente una secuencia que codifica un marcador de selección (Scorer *et al.*, Biotechnology (NY) 1994 Feb; 12(2): 181-4). La proteína puede ser o bien intracelular o bien secretada si el vector contiene introducida en el extremo del gen una secuencia que codifica una secuencia señal de secreción tal como por ejemplo el factor

prepropéptido de *Saccharomyces cerevisiae* (Cregg *et al.*, 1993; Scorer *et al.* 1993). Se puede añadir también una cola de histidina a uno de los extremos de la proteína recombinante para facilitar la purificación (Mozley *et al.*, 1997, Photochem Photobiol, 66(5): 710-5).

5 Preferiblemente, dicho hospedante se selecciona entre: una bacteria, una levadura, una célula de insecto y una célula de mamífero, una célula vegetal.

La administración de composiciones terapéuticas que comprenden dichas proteínas se puede hacer por ejemplo por vía tópica, oral, intradérmica, transdérmica, intravenosa.

10 Los fragmentos de dichas proteínas se pueden utilizar como antagonistas de la proteína de la que se derivan. Así pues, se recomienda la administración adecuada a un animal de una composición terapéutica que comprende dichos fragmentos, para inducir una disminución en la actividad de dicha proteína en el mecanismo de la angiogénesis en una patología dada.

15 La presente invención se refiere también a un método de diagnóstico y/o de pronóstico de una patología angiogénica en un mamífero, particularmente en un ser humano, que consiste en detectar en las células de dicho mamífero la sobreexpresión o la subexpresión de una o más secuencias nucleotídicas identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 25 en la lista de secuencias adjunta.

Dicho método de diagnóstico y/o de pronóstico comprende las siguientes etapas:

- la detección de la expresión de una o más de dichas secuencias nucleotídicas SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 por una población celular de un mamífero,
- 20 - la detección de la expresión de la misma o mismas secuencias por una población celular de referencia cuyo estado angiogénico es conocido,
- la identificación de posibles diferencias en el nivel de expresión de la misma o mismas secuencias por las dos poblaciones celulares.

25 La presente invención se refiere también a un método de diagnóstico y de pronóstico de una patología angiogénica en un mamífero, particularmente en un ser humano, que consiste en detectar en las células de dicho mamífero la sobreexpresión o la subexpresión de una o más secuencias polipeptídicas identificadas con los números SEQ ID N° 6, o con las secuencias polipeptídicas identificadas con los números SEQ ID N° 31 y SEQ ID N° 32 en la lista de secuencias adjunta.

Según un modo de realización preferido, dicho método comprende las siguientes etapas:

- 30 a) la detección de la expresión de una o más de dichas secuencias polipeptídicas SEQ ID N° 6 o con las secuencias polipeptídicas identificadas con los números SEQ ID N° 31 y SEQ ID N° 32 por una población celular de un mamífero,
- b) la detección de la expresión de la misma o mismas secuencias polipeptídicas por una población celular de referencia cuyo estado angiogénico es conocido,
- 35 c) la identificación de las posibles diferencias en el nivel de expresión de la misma o mismas secuencias polipeptídicas por las dos poblaciones celulares.

Según un modo particular de realización, en el método de diagnóstico y de pronóstico de la invención, la detección de la expresión de las secuencias se lleva a cabo después de haber puesto en contacto las células endoteliales con un fluido biológico obtenido de un paciente.

40 La presente invención se refiere también a un método de verificación de la eficacia terapéutica de un tratamiento angiogénico en un mamífero, especialmente en un ser humano, por la identificación de una población celular en dicho mamífero, capaz de sobreexpresar o de subexpresar una o más secuencias nucleotídicas identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta.

Dicho método de verificación de la eficacia terapéutica comprende las siguientes etapas:

- 45 - la detección de la expresión de una o más de dichas secuencias nucleotídicas SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 por una población celular de un mamífero al que se ha administrado una composición terapéutica destinada a tratar un trastorno angiogénico,
- la detección de la expresión de la misma o mismas secuencias nucleotídicas por una población celular de referencia cuyo estado angiogénico es conocido,
- 50 - la identificación de posibles diferencias en el nivel de expresión de la misma o mismas secuencias nucleotídicas por las dos poblaciones celulares.

Según modos de realización preferidos, el método de verificación se lleva a cabo sobre una población celular de un mamífero *in vivo*, *ex vivo*, o bien sobre una población celular aislada de dicho mamífero *in vitro*.

5 Según un modo particular de realización, en el método de verificación de la invención, la detección de la expresión de las secuencias se realiza después de haber puesto en contacto las células endoteliales con un fluido biológico obtenido de un paciente.

La presente invención se refiere también a un método de cribado de compuestos útiles para el tratamiento angiogénico de un mamífero, especialmente de un ser humano.

Según un modo de realización preferido, dicho método de cribado comprende las siguientes etapas:

10 - la detección de la expresión de una o más de dichas secuencias nucleotídicas SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 por una población celular puesta en contacto con un compuesto capaz de tener un efecto terapéutico sobre un trastorno angiogénico,

- la detección de la expresión de la misma o mismas secuencias nucleotídicas por una población celular de referencia cuyo estado angiogénico es conocido,

15 - la identificación de posibles diferencias en el nivel de expresión de la misma o mismas secuencias nucleotídicas por las dos poblaciones celulares.

Según otro modo de realización preferido, dicho método de cribado comprende también las siguientes etapas:

20 - la detección de la expresión de una o más de dichas secuencias polipeptídicas identificadas con los números SEQ ID N° 6 o con las secuencias polipeptídicas identificadas con los números SEQ ID N° 31 y SEQ ID N° 32 en la lista de de secuencias adjunta, por una población celular puesta en contacto con un compuesto capaz de tener un efecto terapéutico sobre un trastorno angiogénico,

- la detección de la expresión de la misma o mismas secuencias polipeptídicas por una población celular de referencia cuyo estado angiogénico es conocido,

- la identificación de posibles diferencias en el nivel de expresión de la misma o mismas secuencias polipeptídicas por las dos poblaciones celulares.

25 Según un modo de realización particular, en el método de cribado de la invención, la detección de la expresión de las secuencias se realiza después de haber puesto en contacto las células endoteliales con un fluido biológico obtenido de un paciente.

30 Entre los trastornos angiogénicos capaces de ser diagnosticados o tratados con las composiciones farmacéuticas de la invención se pueden citar: vascularización de tumores, retinopatías, artritis reumatoide, enfermedad de Crohn, aterosclerosis, hiperestimulación ovárica, psoriasis, endometriosis asociada con neovascularización, reestenosis debida a la angioplastia con globo, hiperproducción tisular debida a cicatrización, enfermedad vascular periférica, hipertensión, inflamación vascular, la enfermedad y el fenómeno de Raynaud, aneurisma, reestenosis arterial, tromboflebitis, linfagitis, linfedema, cicatrización y reparación de tejidos, isquemia, angina de pecho, infarto de miocardio, cardiopatía crónica, insuficiencias cardiacas tales como insuficiencia cardíaca congestiva, degeneración macular asociada a la edad y osteoporosis.

35 En el marco de la presente invención se entiende por sonda cualquier fragmento de ADN de una sola hebra cuya secuencia es complementaria de una secuencia que se busca: esta secuencia puede ser detectada así por hibridación con la sonda marcada (por incorporación de átomos radiactivos o de grupos fluorescentes), que desempeña el papel de un "anzuelo" molecular.

40 Según los modos preferidos de realización, el soporte de dicho dispositivo se selecciona entre una membrana de vidrio, una membrana metálica, una membrana polimérica, una membrana de sílice.

Tales dispositivos son por ejemplo chips de ADN que comprenden una o más secuencias nucleotídicas identificadas con los números SEQ ID N° 1 a SEQ ID N° 5 y SEQ ID N° 27 a SEQ ID N° 30 en la lista de secuencias adjunta.

45 La verificación de la implicación de los cinco genes identificados y de sus homólogos en el mecanismo de la angiogénesis se ha realizado según la metodología descrita en la sección Material y métodos.

Esta implicación se ilustra con ayuda de la figura 1 adjunta, en la cual:

La figura 1 muestra los resultados obtenidos sobre la formación de tubos capilares sobre células endoteliales humanas bajo el efecto de la expresión de los vectores GS-V1, GS-V2, GS-V3, V4-GS, en la cual:

50 - la figura 1A muestra que la formación de los tubos capilares se inhibe en las células endoteliales transfectadas con el vector GS-V1 que codifica el transcrito antisentido específico de GS-N1,

- la figura 1B muestra que la formación de los tubos capilares se inhibe en las células endoteliales transfectadas con el vector GS-V2 que codifica el transcrito antisentido específico de GS-N2,
- la figura 1C muestra que la formación de los tubos capilares se inhibe en las células endoteliales transfectadas con el vector GS-V3 que codifica el transcrito antisentido específico de GS-N3,
- 5 - la figura 1D muestra que la formación de los tubos capilares se inhibe en las células endoteliales transfectadas con el vector GS-V4 que codifica el transcrito antisentido específico de GS-N4 y de su homólogo GS-N5 y
- la figura 1F muestra que la formación de los tubos capilares no se modifica en las células endoteliales transfectadas con el vector vacío (control).

Material y métodos

10 1. Cultivo de las células y ensayo de angiogénesis

Las células endoteliales de venas umbilicales (HUVEC) en las cuatro condiciones de cultivo mencionadas se utilizan en este caso para identificar los genes que codifican los constituyentes celulares implicados en la regulación de la angiogénesis.

Las células endoteliales se mantienen en medio completo (EGM-2-MV de Clonetics).

15 Para la identificación de los genes implicados en la angiogénesis en los ensayos *in vitro* de la angiogénesis según el modelo de Montesano *et al.* (1986, Proc. Natl. Acad. Sci. U.S.A., 83(19), 7297-301), se siembran primero las células en un gel de colágeno tipo I en medio completo hasta confluencia. Después, las células HUVEC de referencia se cultivan en medio empobrecido en suero y sin factores de crecimiento: se añaden en las condiciones de ensayo EBM-2-MV + 2 % de suero y diferentes factores.

20 El FGF2: a concentraciones comprendidas entre 5 ng/ml y 60 ng/ml, preferiblemente entre 10 y 40 ng/ml; el VEGF: a concentraciones comprendidas entre 10 ng/ml y 60 ng/ml, preferiblemente comprendidas entre 30 ng/ml y 50 ng/ml; el PF4: a concentraciones comprendidas entre 0,1 y 5 μ g/ml, preferiblemente comprendidas entre 0,5 μ g/ml y 1 μ g/ml; el TNF-alfa a concentraciones comprendidas entre 20 ng/ml y 100 ng/ml, preferiblemente comprendidas entre 30 ng/ml y 60 ng/ml; el IFN-gamma: a concentraciones comprendidas entre 50 ng/ml y 200 ng/ml, preferiblemente comprendidas entre 80 ng/ml y 120 ng/ml.

Las células endoteliales humanas puestas en las cuatro condiciones de cultivo previamente mencionadas se utilizan entonces para identificar los genes que codifican los constituyentes celulares implicados en la regulación de la angiogénesis.

2. Factores angiogénicos y angiostáticos.

30 Los factores angiogénicos y angiostáticos que tienen efecto sobre la expresión de los genes identificados en correlación con la formación de nuevos vasos sanguíneos o con la inhibición de nuevos vasos sanguíneos respectivamente, utilizados a modo de ejemplo en el marco de la presente invención, se describen a continuación.

VEGF = factor de crecimiento endotelial vascular.

PF4 = factor plaquetario 4.

35 TNF - α = factor alfa de necrosis tumoral

El TNF-alfa que es un regulador de la angiogénesis, puede inducir la angiogénesis *in vivo* pero también puede inhibir la formación de los vasos sanguíneos *in vitro* (Frater-Schroder *et al.*, 1987, Proc. Natl. Acad. Sci. U.S.A., 84(15), 5277 -81; Sato *et al.*, 1987, J. Natl. Cancer Inst. 79(6), 1383-1391; Fajardo *et al.*, 1992, Am. J. Pathol. Mar, 140(3), 539-44; Niida *et al.*, 1995, Neurol. Med. Chir. (Tokyo), 35(4), 209-14. En el modelo de angiogénesis *in vitro* utilizado aquí, el TNF-alfa se utiliza en condiciones de inhibición de la angiogénesis.

3. Comparación de las expresiones génicas.

Las expresiones génicas se pueden comparar entonces utilizando los chips de ADN, el SAGE (análisis seriado de la expresión génica), una reacción de amplificación por PCR cuantitativa, los vectores virales para la construcción de bancos sustractivos, o también el análisis por visualización diferencial.

45 En el marco de los estudios experimentales que han llevado a la presente invención, el solicitante ha utilizado preferentemente la técnica de visualización diferencial para la identificación de dichos genes.

Visualización diferencial

Los ARN totales se preparan a partir de las células HUVEC cultivadas sobre un gel de colágeno en presencia de los diferentes factores utilizados, según el método RNeasy Mini Kit (Qiagen) mediante la integración de una etapa de digestión con ADNasa I según el protocolo recomendado por el fabricante.

5 La visualización diferencial a partir de los ARN totales se realiza según el método descrito por Liang and Pardee (1992, Science, 14: 257 (5072) 967-7) utilizando alfa_P33-ATP en dilución isotópica a lo largo de la amplificación por PCR para la visualización de las bandas por autorradiografía de los geles de electroforesis.

Por lo tanto, los fragmentos de ADN diferencialmente presentes sobre el gel en función de las condiciones de cultivo analizadas, se cortan, se reamplifican, se clonan en un plásmido vector PGEM easy (Promega), se secuencian y se identifican mediante las preguntas del banco BLAST.

10 4. Verificación de la implicación de los genes identificados en el mecanismo de la angiogénesis.

Ensayo de funcionalidad de los genes

En una segunda etapa, se ensaya la funcionalidad de cada secuencia identificada sobre el modelo de angiogénesis *in vitro* con las células endoteliales transfectadas con un vector de expresión que comprende un oligonucleótido antisentido de dicha secuencia.

15 Para la construcción de estos vectores, se diseñan cebadores específicos para cada una de las secuencias identificadas. Estos cebadores se muestran en la Tabla I que sigue, y se identifican con los números de secuencias SEQ ID N° 19 a SEQ ID N° 26 en la lista de secuencias adjunta.

Tabla I

| SEQ ID. del gen identificado | Nombre del cebador |
|------------------------------|--------------------|
| SEQ ID N° 1 (GS-N1) | GV1-1 |
| | GV1-2 |
| SEQ ID N° 2 (GS-N2) | GV2-1 |
| | GV2-2 |
| SEQ ID N° 3 (GS-N3) | GV3-1 |
| | GV3-2 |
| SEQ ID N° 4 (GS-N4) | GV4-1 |
| | GV4-2 |
| SEQ ID N° 5 (GS-N5) | GV4-1 |
| | GV4-2 |

20 Estos cebadores contienen en cada uno de sus extremos, un sitio de una enzima de restricción diferente (Sall: GTCGAC o MluI: ACGCGT).

Los fragmentos amplificados de cada gen se obtienen por PCR a partir de los plásmidos bacterianos que contienen el fragmento del gen identificado utilizando dichos cebadores.

25 Estos fragmentos se purifican, se digieren con las enzimas de restricción Sall y MluI y se insertan en un vector de expresión en mamíferos del tipo vector PCI-neo (Promega), que ha sido digerido él mismo por estas dos enzimas de restricción.

Cada fragmento se introduce en la orientación antisentido.

30 De una forma general, los vectores que se pueden utilizar para demostrar la funcionalidad de los genes identificados en la presente invención en el mecanismo de la angiogénesis, comprenden cualquier sistema de vectores de expresión en mamíferos que comprende un promotor que permite la expresión de un gen clonado, por ejemplo se puede citar el promotor "fuerte" del citomegalovirus humano (CMV).

Los vectores de expresión constitutivos o inducibles que se pueden utilizar, se indican en la lista no exhaustiva detallada a continuación:

35 Los vectores, pCI Mammalian Expression Vector, cuyo número de acceso es U47120 en la base de datos EMBL. Expression Vector System cloning vector pALTER(R)*-MAX (Promega), vectores pcDNA3.1, -/hygro, -/Zeo, pcDNA4/HisMax, -E, pBudCE4, pRcRSV, pRcCMV2, pSecTag2 - vectores de secreción -/hygro, los vectores pEBVHis A, B y C) (Invitrogen) los vectores de expresión en el mamífero pIRES, pIRES-EYFP pIRES2-EGFP, pCMV-Myc y pCMV-HA. La secuencia de CMV cuyo número de acceso es ABV26916 o AI953818 se describe en la

solicitud de patente WO01/60860. Epitope-Tagged pTRE, los vectores VP16 Minimal Domain (pTA 2, pTA 3 y pTA 4), los vectores de expresión Tet bidireccionales (pBI, pBI-EGFP, pBI-G, pBI-L), pRevTRE, pTRE2, pLEGFP-N1 vector retroviral pLEGFP-C1, los sistemas de expresión adenovirales Adeno-X, PCMS-EGFP, pd1EGFP-N1, pd2EGFP-N1, pd2EYFP-N1, pEGFP (-C1, -C2, -C3, -N1, -N2, -N3), pEYFP-C1, -N1 (Clontech).

5 Cada vector que comprende dicho fragmento antisentido se produce entonces en *E. coli*, se extrae, se purifica y se cuantifica. Se incuba 1 μg de cada vector en presencia de un agente transfectante (effectene, Qiagen) siguiendo el protocolo recomendado por el fabricante con las células endoteliales. Veinticuatro horas después de la transfección, las células endoteliales se tripsinan y se extienden sobre la matriz extracelular que contiene los factores angiogénicos, en este caso Matrigel, según el modelo descrito por Grant *et al.*, (1989, Cell, 58(5), 933-43). Después
10 de 24 horas de incubación, se observa la formación de vasos sanguíneos y se compara con las células control transfectadas con el vector de expresión de mamífero vacío.

5. Establecimiento del banco de líneas estables que expresan los vectores de expresión que contienen las secuencias de los genes o fragmentos de las mismas o las secuencias antisentido.

15 Los sistemas de expresión pueden comprender un marcador de selección para un antibiótico (un gen de resistencia a un antibiótico) para seleccionar las células transfectadas que expresan de forma estable el vector que comprende el ácido nucleico clonado en dicho vector, ya sea en el mismo vector o ya sea en un segundo vector cotransfectado.

Este vector de expresión puede ser un sistema de expresión constitutivo o inducible.

20 En el ejemplo particular descrito más adelante, las líneas estables para la expresión del oligonucleótido antisentido correspondiente a cada gen identificado se han obtenido con un vector de expresión constitutivo y después de selección en presencia de antibiótico.

25 Para ello, 24 horas después de la transfección efectuada en las condiciones descritas antes, las células endoteliales BAEC se tripsinan y se siembran a razón de 80.000 células/pocillo en placas de seis pocillos en presencia de 700 $\mu\text{g}/\text{ml}$ del antibiótico G418 (Promega). Se siembra un pocillo control con células no transfectadas. El medio se cambia cada tres días con una recarga de antibiótico. Se eliminan las células control después de 8 a 10 días, las células resistentes al antibiótico se recogen en la confluencia (después de dos a tres semanas) después se transfieren a frascos de cultivo siempre en presencia del antibiótico. Las líneas celulares estables se ensayan a continuación para determinar su capacidad para formar o no vasos sanguíneos en el ensayo de angiogénesis *in vitro*.

6. Resultados

30 6.1 Identificación de los genes.

35 La secuencia de ácidos nucleicos denominados GS-N1, y la proteína codificada por dicha secuencia de ácidos nucleicos GS-N1 identificada con el número SEQ ID N° 6 denominada: angiopartnerina, no se ha identificado con anterioridad que tengan alguna función biológica, y todavía menos una función en el proceso de la angiogénesis o en la diferenciación de las células endoteliales en tubos capilares. Las secuencias de ácidos nucleicos denominados GS-N2 a GS-N5 identificadas con los números SEQ ID N° 2 a SEQ ID N° 5 en la lista de secuencias adjunta y respectivamente las proteínas codificadas por dichos ácidos nucleicos, identificadas con los números SEQ ID N° 7 a SEQ ID N° 10 en la lista de secuencias adjunta, no se ha identificado previamente que tengan una función biológica en el proceso de la angiogénesis o en la diferenciación de las células endoteliales en tubos capilares. Estas secuencias se describen más adelante.

40 El método de visualización diferencial descrito antes ha permitido la identificación de los siguientes ARNm:

- GS-N1: un ARNm de 6160 pares de bases (pb) identificado por la secuencia SEQ ID N° 1 en la lista de secuencias adjunta. Una búsqueda utilizando BLAST sobre la base de secuencias GENBANK permite identificarlo con el número de acceso AB037857.

45 La secuencia de este ARNm tiene una secuencia codificante parcial del nucleótido 1 al nucleótido 2777. Ha sido por lo tanto identificada una proteína GS-P1, resultante de la traducción de este ARNm. Esta proteína se compone de 924 aminoácidos (aa), se identifica con el número SEQ ID N° 6 en la lista de secuencias adjunta y se llama angiopartnerina.

50 La angiopartnerina es homóloga con el regulador negativo del receptor de la prostaglandina F2 (PTGFRN) (número de acceso XM_040709, secuencia nucleica: 5975 pb, secuencia proteica: 560 aa) PTGFRN cuyo número de acceso es también BC022235 en la base de datos EMBL, está descrito en la patente US 7.747.660, con la proteína 6 asociada con la célula del músculo liso humano (SMAP6) (Número de acceso AB014734, secuencia nucleica: 2197 pb, secuencia proteica parcial: 186 aa), identificadas con los números SEQ ID N° 27 y SEQ ID N° 28, respectivamente, en la lista de secuencias adjunta, siendo ella misma similar a la proteína reguladora del receptor de la prostaglandina F2 alfa (FPRP) llamada indistintamente CD9P-1-F y renombrada EW1-F por Stipp *et al.* (2001, J.

Biol. Chem., 276, 44, 40545 a 40554.). La secuencia GS-N1 comprende las secuencias SEQ ID N° 27 y SEQ ID N° 28, presentando el 99 % de homología con ellas.

La FPRP es una glucoproteína transmembranal de tipo 1 que contiene 6 dominios de inmunoglobulinas extracelulares. Originalmente se identificó y caracterizó por su capacidad de combinarse con la prostaglandina F2 alfa e inhibir la unión de esta última con su receptor. También se puede combinar con otros receptores acoplados a la proteína G que contienen 7 dominios transmembranales, y reducir la interacción ligando-receptor (Orlicky and Nordeen, 1996, Prostaglandins Leukot. Essent. Fatty Acids 55: 261-268; Orlicky *et al.*, 1998 J. Lip. Res., Vol. 39, 1152-1161). El aumento de la expresión de la FPLP ya ha sido asociado con la diferenciación celular, especialmente la de los adipocitos (Orlicky *et al.*, 1998, J. Lip. Res., Vol. 39, 1152-1161). Diferentes estudios han demostrado que CD9P-1 o FPLP es un importante socio de dos miembros de la familia de las tetraspaninas (llamada también TM4SF), CD9 y CD81 en complejos proteicos, combinándose específicamente o bien con CD81 o bien con CD9 y CD81 (Charrin S *et al.*, 2001, J. Biol. Chem; Stipp *et al.* (2001, J. Biol. Chem., 276, 7, 4853-4862). Los tetraspaninas han sido implicadas en muchas funciones celulares tales como la adhesión, la migración, la co-estimulación, la transducción de señales y la diferenciación, pudiendo estar ligadas las diversas funciones atribuidas a las tetraspaninas a su combinación específica con las moléculas asociadas (partenaires) específicas; (Le Naour *et al.*, 2000, Science, 287, 319-321).

La proteína CD9 tiene una amplia distribución tisular, se ha encontrado especialmente en varios tipos de tumores (Si and Hersey, 1993 Int J. Cancer, 54: 37-43; Miyake *et al.*, 1996, Cancer Res. 56: 1244-1249) así como en los vasos sanguíneos formados por las células endoteliales (Zola *et al.*, 1989, Immunol Cell Biol., 67: 63-70). Esta proteína ha sido implicada en funciones tales como la transducción de señales, la adhesión celular, la motilidad, la evolución de los tumores (Ozaki *et al.*, 1995, J. Biol. Chem. 270: 15.119-15.124, Forsyth, 1991 Immunology, 72: 292-296; Anton *et al.*, 1995, J. Neurosci., 15: 584-595; Shaw *et al.*, 1995, J. Biol. Chem., 270: 24092-24099; Ikeyama *et al.*, 1993, J. Exp. Med. 177: 1231-1237) y especialmente en la adhesión y migración de las células endoteliales a lo largo de la angiogénesis (Klein-Soyer *et al.*, 2000, Arterioscler. Thromb. Vasc. Biol., 20: 360-9). La sobreexpresión de CD9 en células de adenocarcinoma suprime su motilidad y potencial metastásico (Ikeyama *et al.*, 1993, J. Exp. Med., 177: 1231-1237), su expresión está inversamente correlacionada con los tumores primarios y con la aparición de metástasis en melanomas, y cánceres de pulmón, de colon y de mama (Si and Hersey, 1993 Int. J. Cancer, 54, 37-43; Miyake *et al.*, 1995, Cancer Res., 55. 4127-4131; Adachi *et al.*, 1998; J. Clin. Oncol., 16, 1397-1406; Mori *et al.*, 1998, Clin. Cancer Res., 4, 1507-1510).

La proteína CD9P-1 (o FPLP) ha sido identificada como la principal molécula asociada de CD9 en las líneas cancerosas (Serru *et al.*, 1999, Biochem. J., 340, 103-111).

También se ha descrito que la proteína CD81 está implicada en diversas funciones tales como la señalización celular y la activación de los linfocitos B (Fearon and Carter, 1995), la regulación de la proliferación de los linfocitos T (Miyazaki *et al.*, 1997, EMBO J., 16, 4217-4225), también puede desempeñar un papel en el cáncer porque la CD81 es un posible receptor para el virus de la hepatitis C, una causa importante del carcinoma hepático (Pileri *et al.*, 1998, Science, 282, 938-941).

Aunque el papel exacto de CD9P-1 o FPRP aún no está definido, su asociación con CD9 o CD81 puede dar a entender un papel en las funciones de regulación de los receptores CD9 o CD81. Sin embargo, hasta la fecha, no se ha descrito ningún papel en la regulación de la angiogénesis ni para la proteína llamada angiopartnerina, identificada con el número SEQ ID N° 6 en la lista de secuencias adjunta, ni para las proteínas PTGFRN, CD9P-1/FPRP.

- GS-N2: un ARNm de 3266 pb, identificado por la secuencia SEQ ID N° 2 en la lista de secuencias adjunta. Una búsqueda utilizando BLAST sobre la base de secuencias GENBANK permite identificarlo con el número de acceso AF247704.

La secuencia de este ARNm tiene una secuencia codificante del nucleótido 49 al nucleótido 753. Ha sido por tanto identificada una proteína GS-P2, resultante de la traducción de este ARNm. Esta proteína se compone de 234 aa, se identifica con el número SEQ ID N° 7 en la lista de secuencias adjunta y se llama proteína homeobox NKX3.1

La proteína NKX3.1 es un miembro de la clase NK de las proteínas homeobox, estrechamente ligada a la proteína NK-3 de la Drosophila (Kim, Y and Nirenberg, 1989, Proc. Natl. Acad. Sci. USA, 86, 7716-7720; He *et al.*, 1997, Genomics, 43, 69-77). Los estudios en el ratón, han permitido demostrar la expresión del gen *Nkx3.1* en el feto y en el embrión en desarrollo en una variedad de tipos de tejidos tales como el mesodermo, el músculo liso vascular, el epitelio y regiones del sistema nervioso central (Kos *et al.*, 1998, Mech. Dev., 70, 25-34; Tanaka *et al.*, 1999, Mech. Dev., 85, 179-182; Bhatia-Gaur *et al.*, 1999, Genes Dev., 13, 966-977). En el adulto, la proteína NKX3.1 está localizada de forma predominante en la próstata, más particularmente en las células epiteliales y su expresión está regulada por los andrógenos (He *et al.*, 1997, Genomics, 43, 69-77; Prescott *et al.*, 1998, Prostate, 35, 71-80; Scivolino *et al.*, 1997, Dev. Dyn., 209, 127-138; Ornstein *et al.*, 2001, J. Urol., 165(4): 1329-34). Parece jugar un papel esencial en la función de la próstata y regula la proliferación de las células epiteliales de la próstata; se ha sugerido que el gen NKX3.1 es un gen supresor de los tumores específicos de la próstata (Bhatia-Gaur *et al.*, 1999, Genes Dev., 13 (8): 966-977). La pérdida de expresión de NKX3.1 en los cánceres humanos de próstata ha sido correlacionada recientemente con la evolución de los tumores (Bowen *et al.*, 2000, Cancer Res., 60 (21): 6111-5).

Además, ya se ha informado de que las proteínas homeobox están implicadas en la regulación de la angiogénesis (revisión: Gorski and Walsh, 2000, *Circulation Research*, 87: 865-872).

Sin embargo, no se ha descrito hasta ahora, ningún papel para la proteína homeobox NKX3.1 en la regulación de la angiogénesis.

- 5 - GS-N3: un ARNm de 6771 pb, identificado por la secuencia SEQ ID N° 3 en la lista de secuencias adjunta. Una búsqueda utilizando BLAST sobre la base de secuencias GENBANK permite identificarlo con el número de acceso U91543.

10 La secuencia de este ARNm tiene una secuencia codificante del nucleótido 151 al nucleótido 6153. Por lo tanto, se ha identificado una proteína GS-P3 resultante de la traducción de este ARNm. Esta proteína se compone de 2000 aa, se identifica con el número SEQ ID N° 8 en la lista de secuencias adjunta, y se llama helicasa con dedo de zinc.

15 La proteína hZFH (helicasa con dedo de zinc, humana) pertenece a la familia de las helicasas tipo Snf2, conocidas por actuar como reguladores transcripcionales para múltiples genes (Aubry *et al.*, 1998, *Eur. J. Biochem.* 254(3): 558-64). Contiene también un cromodominio y es homóloga con la proteína CHD3 ("chromodomain helicase DNA binding protein 3") identificada por la secuencia SEQ ID N° 29 y SEQ ID N° 30 en la lista de secuencias adjunta. Las proteínas CHD3 están descritas como capaces de regular la expresión de los genes reprimiendo la transcripción a través de una alteración de la estructura de la cromatina (Zhang *et al.*, 1998, *Cell*, 95, 279-289; Kehle *et al.*, 1998, *Science*, 282, 1897-1900). Hasta la fecha, no se ha descrito ningún papel en la regulación de la expresión de los genes implicados en la angiogénesis ni para la proteína hZFH ni para la proteína CHD3.

- 20 - GS-N4: un ARNm de 3041 pb identificado por la secuencia SEQ ID N° 4 en la lista de secuencias adjunta. Una búsqueda utilizando BLAST sobre la base de secuencias GENBANK permite identificarlo con el número de acceso BC001571.

25 La secuencia de este ARNm tiene una secuencia codificante del nucleótido 67 al nucleótido 2808. Por lo tanto, se ha identificado una proteína GS-P4, resultante de la traducción de este ARNm. Esta proteína se compone de 913 aa, se identifica con el número SEQ ID N° 9 en la lista de secuencias adjunta y se llama factor de iniciación de la traducción eucariota, subunidad 8 (110 kD) (EIF3S8).

Esta secuencia de GS-N4 presenta homología con la siguiente secuencia:

- 30 - GS-N5: un ARNm de 1507 pb identificado por la secuencia SEQ ID N° 5 en la lista de secuencias adjunta. Una búsqueda utilizando BLAST sobre la base de secuencias GENBANK permite identificarlo con el número de acceso BC000533.

La secuencia de este ARNm tiene una secuencia codificante del nucleótido 407 al nucleótido 1384. Por lo tanto, se ha identificado una proteína GS-P5 resultante de la traducción de este ARNm. Esta proteína se compone de 325 aa, se identifica con el número SEQ ID N° 10 en la lista de secuencias adjunta y se llama proteína similar al factor de iniciación de la traducción eucariota, subunidad 8 (110 kD).

35 El factor 3 de iniciación de la traducción eucariota (EIF3), el factor más grande de iniciación de la síntesis proteica con un tamaño de 650 kDa se compone de al menos nueve subunidades peptídicas (Hershey *et al.*, 1996, *Biochimie*, 78, 903-907), incluyendo la subunidad 8 (p110) (Asano *et al.*, 1997, *J. Biol. Chem.*, 272, 1101-1109). El EIF3 desempeña un papel central en el proceso de iniciación de la biosíntesis proteica especialmente en la unión del iniciador metionil-ARNt y el ARNm en la subunidad ribosómica 40S para formar un complejo de iniciación 40S (Merrick and Hershey, 1996, *The pathway and mechanism of eukaryotic protein synthesis*. In: Hershey JWB, Mathews MB, Sonenberg N, eds *Translational Control*. Cold Spring Harbor, NY: Cold Spring Harbor Press, 1996: 31-67). El EIF3 parece jugar un papel central en la iniciación por interacción con otros muchos componentes del sistema de traducción (Vornlocher *et al.*, 1999, *J. Biol. Chem.*, Vol. 274, Issue 24, 16802-16812). Las funciones de cada subunidad son todavía poco conocidas. Se detecta en los tumores un alto nivel de expresión de ciertas subunidades tales como la p150, la p170; se sugiere que la p170 desempeña otro papel además de sus funciones en la iniciación de la traducción (Lin *et al.*, 2001, *J. Cell Biochem.*, 80(4): 483-90; Pincheira *et al.*, 2000, *Eur. J. Cell Biol.*, 80(6): 410-8). La sobreexpresión de la subunidad 8 (p110) también ha sido demostrada en un tumor de las células germinales por Roche *et al.* (2000, *American Journal of Pathology*, 157: 1597-1604), lo que sugiere un papel de esta subunidad en el desarrollo del tumor aumentando la traducción y en general, llevando a un aumento del crecimiento y de la división celular.

50 Hasta la fecha, no se ha descrito nunca ninguna implicación ni de la subunidad 8 del EIF3, ni de su proteína similar, en la regulación de la angiogénesis.

La expresión de los ARNm identificados antes se observa en las células endoteliales humanas cuando la formación de tubos capilares, sinónimo de la angiogénesis, es inhibida debido a la acción de un factor angiostático.

55 El solicitante ha demostrado que el aumento de la expresión del gen correspondiente a cada uno de estos ARNm acompaña a la inhibición de la formación de nuevos vasos sanguíneos por las células endoteliales.

5 En efecto, las células endoteliales humanas que forman nuevos vasos sanguíneos como consecuencia de la estimulación por un factor angiogénico muestran una expresión muy baja de este ARNm, mientras que las mismas células endoteliales humanas estimuladas por el mismo factor angiogénico y puestas en contacto con un factor angiostático (donde la angiogénesis se inhibe) y/o incluso las mismas células endoteliales estimuladas únicamente por el factor angiostático, muestran una expresión elevada de este gen (como se indica en la tabla II).

Estos resultados prueban la existencia de una correlación directa entre la expresión de cada uno de estos genes y el estado angiostático (es decir, la inhibición de la angiogénesis) de las células endoteliales humanas.

Tabla II

| SEQ. ID | Inductores de la expresión |
|---------------------|----------------------------|
| SEQ ID N° 1 (GS-N1) | TNF-alfa |
| SEQ ID N° 2 (GS-N2) | TNF-alfa |
| SEQ ID N° 3 (GS-N3) | TNF-alfa |
| SEQ ID N° 4 (GS-N4) | PF4 |
| SEQ ID N° 5 (GS-N5) | PF4 |

6.2 Verificación del papel de los genes identificados en la regulación de la angiogénesis.

10 Además, se ha demostrado también el papel funcional de estos genes en la formación de nuevos vasos sanguíneos por las células endoteliales humanas.

En efecto, una secuencia nucleotídica específica de cada uno de los genes identificados, seleccionada entre las secuencias nucleotídicas identificadas por las secuencias SEQ ID N° 11 a SEQ ID: 14 en la lista de secuencias adjunta, se ha introducido en el vector de expresión Vector pCI-neo en la orientación antisentido.

15 Los vectores resultantes denominados GS-V1 a GS-V4 identificados por sus secuencias SEQ ID N° 15 a SEQ ID N° 18 en la lista de secuencias adjunta, han sido utilizados para reprimir la expresión del gen que codifica este ARNm en las células endoteliales humanas después de la transfección de estas últimas por este vector.

20 Las células endoteliales humanas han sido estimuladas entonces por factores angiogénicos. Los resultados obtenidos, para cada una de las secuencias GS-N1 a GS-N5 utilizando las secuencias antisentido y los vectores correspondientes, demuestran que:

la represión de la expresión de los genes identificados con los números de secuencias SEQ ID N° 1 a SEQ ID N° 5, inhibe la formación de nuevos vasos sanguíneos por las células endoteliales humanas.

Los resultados obtenidos, para cada una de las secuencias, utilizando las secuencias antisentido y los vectores correspondientes, indicados en la tabla III que sigue, se ilustran en la figura 1 (Figuras 1A a 1E) adjunta.

25

Tabla III

| NOMBRE | Genes SEQ. ID | Proteína SEQ. ID | Secuencias antisentido | Vector con el antisentido insertado | Fig. | Fig. control |
|--------|---------------------|---|------------------------|-------------------------------------|------|--------------|
| 1 | SEQ ID N° 1 (GS-N1) | SEQ ID N° 6 (GS-P1) angiopartnerina | SEQ ID N° 11 (392 pb) | SEQ ID N° 15 (GS-V1) | 1A | 1F |
| 2 | SEQ ID N° 2 (GS-N2) | SEQ ID N° 7 (GS-P2) Nkx3.1 homeobox | SEQ ID N° 12 (250 pb) | SEQ ID N° 16 (GS-V2) | 1B | 1F |
| 3 | SEQ ID N° 3 (GS-N3) | SEQ ID N° 8 (GS-P3) helicasa dedo de zinc | SEQ ID N° 13 (pb) | SEQ ID N° 17 (GS-V3) | 1C | 2C |
| 4 | SEQ ID N° 4 (GS-N4) | SEQ ID N° 9 (GS-P4) factor de iniciación de la traducción, subunidad 8 e | SEQ ID N° 14 (167 pb) | SEQ ID N° 18 (GS-V4) | 1D | 1F |
| 5 | SEQ ID N° 5 (GS-N5) | SEQ ID N° 10 (GS-P5) factor de iniciación de la traducción eucariota, subunidad 8 | SEQ ID N° 14 (167 pb) | SEQ ID N° 18 (GS-V4) | | |

LISTA DE SECUENCIAS

- <110> GENE SIGNAL Salman Al-Mahmood Colin, Sylvie Schneider, Christophe Al-Mahmood, Salman
- <120> Genes reguladores de la angiogénesis, preparaciones farmacéuticas que los contienen y sus aplicaciones
- <130> 25599 GENE SIGNAL
- 5 <140> PCT/FR03/XXXXX
- <141> 21-03-2003
- <150> FR02/03655
- <151> 22-03-2002
- <160> 34
- 10 <170> PatentIn versión 3.1
- <210> 1
- <211> 6160
- <212> ADN
- <213> Homo sapiens
- 15 <220>
- <221> característica_miscelánea
- <223> secuencia llamada GS-N1
- <300>
- <308> GenBank/AB037857
- 20 <309> 14-03-2000
- <313> (1)..(6160)
- <400> 1

```

ctcgcgagga gagcggagca ggcgcgcgcc ccaggcggag gagcgccgac tctggagcag      60
ccggagctgg aagaggagga ggaggagagg cggcggggaa ggaggaggag ggggagagtc      120
gctcccgcog ggcgagcatg ggcgcctgg cctcgaggcc getgctgctg gcgctcctgt      180
cgttggtctt ttgcccaggc cgtgtggtga gactccccac agcgaccctg gttcgagtgg      240
tgggactga gctggtcatc cctgcaacg tcagtgacta tgatggcccc agcgagcaaa      300
actttgactg gagcttctca tctttgggga gcagctttgt ggagcttgca agcacctggg      360
agggtggggtt cccagcccag ctgtaccagg agcggctgca gaggggcgag atcctgttaa      420
ggcggactgc caacgacgcc gtggagctcc acataaagaa cgtccagcct tcagaccaag      480
gccactacaa atgttaacc cccagcacag atgccactgt ccagggaac tatgaggaca      540
cagtgcaggt taaagtgtct gccgactccc tgcacgtggg ccccgccgcg cggccccgcg      600
cgagcctgag cctgcgggag ggggagccct tcgagctgcg ctgcaactgc gcctccgctt      660
cgccgctgca cacgacctg gcgctgctgt gggaggtgca ccgcgcccg gccagccgga      720
gcgtcctcgc cctgaccac gagggcaggt tccaccggg cctggggtac gagcagcgtt      780
accacagtgg ggacgtgcgc ctgcacaccg tgggcagcga cgcctaccgc ctctcagtgt      840
cccgggctct gtctgcccag cagggtcctt acaggtgtat cgtcagcgag tggatcgccg      900
agcagggcaa atggcaggaa atccaagaaa aggcggtgga agttgccacc gtggtgatcc      960
agccatcagt tctgcgagca gctgtgcccc agaattgtgtc tgtggctgaa ggaaaggaac     1020
tggacctgac ctgtaacatc acaacagacc gagccgatga cgtccggccc gaggtgacgt     1080
ggtccttcag caggatgcct gacagcacc tacctggctc ccgctgtgtg gcgcggttg     1140
accgtgattc cctggtgcac agctgcctc atgttgcttt gagtcattgt gatgcacgct     1200
cctaccattt actggttcgg gatgttagca aagaaaactc tggctactat tactgccacg     1260
tgtcctgtg ggcaccgga cacaacagga gctggcaca agtggcagag gccgtgtctt     1320
cccagctgg tgtgggtgtg acctggctag aaccagacta ccaggtgtac ctgaatgctt     1380
ccaaggtccc cgggtttgcg gatgacccca cagagctggc atgcccgggtg gtggacacga     1440
agagtgggga ggcgaatgtc cgttcacgg tttcgtggta ctacaggatg aaccggcgca     1500
gcgacaatgt ggtgaccagc gagctgcttg cagtcatgga cggggactgg acgctaaaat     1560
atggagagag gagcaagcag cgggcccagg atggagactt tattttttct aaggaacata     1620
cagacacgtt caatttccgg atccaaagga ctacagagga agacagaggc aattattact     1680
gtgtgtgtc tgcctggacc aacagcggga acaacagctg ggtgaaaagc aaggatgtct     1740
tctccaagcc tgttaacata ttttgggcat tagaagattc cgtgctgtgt gtgaaggcga     1800
ggcagccaaa gcctttcttt gctgccggaa atacatttga gatgacttgc aaagtatctt     1860
ccaagaatat taagtgcgca cgtactctt ttctcatcat ggctgagaag cctgtcggcg     1920
acctctcag tcccaatgaa acgaagtaca tcatctctct ggaccaggat tctgtggtga     1980
agctggagaa ttggacagat gcatcacggg tggatggcgt tgttttagaa aaagtgcagg     2040

```

ES 2 511 694 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| aggatgagtt | ccgctatoga | atgtaccaga | ctcaggtctc | agacgcaggg | ctgtaccgct | 2100 |
| gcatgggtgac | agcctggtct | cctgtcaggg | gcagcctttg | gcgagaagca | gcaaccagtc | 2160 |
| tctccaatcc | tattgagata | gacttccaaa | cctcaggtcc | tataatthaat | gcttctgtgc | 2220 |
| attcagacac | accatcagta | attcggggag | atctgatcaa | attgttctgt | atcatcactg | 2280 |
| tcgagggagc | agcactggat | ccagatgaca | tggcctttga | tgtgtcctgg | tttgcgggtgc | 2340 |
| actctttttg | cctggacaag | gctcctgtgc | tctgtcttc | cctggatcgg | aagggcatcg | 2400 |
| tgaccacctc | aagagggac | tggaagagcg | aaactcagcct | ggagcgctg | agtgtgctgg | 2460 |
| aattcttgct | gcaagtgc | ggctccgagg | accaggactt | tggcaactac | tactgttccg | 2520 |
| tgactccatg | ggtgaagtca | ccaacaggtt | cctggcagaa | ggaggcagag | atccactcca | 2580 |
| agcccgtttt | tataactgtg | aagatggatg | tgctgaacgc | cttcaagtat | cccttgctga | 2640 |
| tcggcgctcg | tctgtccacg | gtcatcgggc | tctgtcctg | tctcatcggg | tactgcagct | 2700 |
| cccactggtg | ttgtaagaag | gaggttcagg | agacacggcg | cgagcgccgc | aggctcatgt | 2760 |
| cgatggagat | ggactaggct | ggcccgggag | gggagtga | gagggacggt | ctaggagcaa | 2820 |
| ttggggcaag | aagagagacag | tgatatttta | aaacaaaagtg | tgttacacta | aaaaccagtc | 2880 |
| ctctctaate | tcaggtggga | cttggcgctc | tctcttttct | gcatgtcaag | ttctgagcgc | 2940 |
| ggacatgttt | accagcacac | ggctcttctt | cccacggcac | tttctgatgt | acaatcgag | 3000 |
| tgtgtgtttt | cccactgca | gctttttaat | ggttaaccct | catctaattt | tttttctccc | 3060 |
| actggtttat | agatcctctg | acttgtgtgt | gtttatagct | tttgtttcgc | ggggttgtgg | 3120 |
| tgaggaaggg | gtgatggcat | gcggagttct | ttatctcag | tgagaatgtg | cctgcccgc | 3180 |
| tgaggagcag | cttcccggtt | ggaggcacgt | gttcagagag | ctgctgagcg | ccacctcta | 3240 |
| cccggctgac | agacaacaca | gacctgtgcc | gaaggtaaat | ttgtggcttt | tacgacctta | 3300 |
| ccccaccccc | tgttttcagg | ggttttagact | acatttgaaa | tccaaacttg | gagtataata | 3360 |
| cttcttattg | agcccaactg | cttttttttt | tttttttttt | gcttctctgc | cccttttcca | 3420 |
| tttcttttgt | atttgttttc | tgtgagagca | ctgaaatggc | agccctggaa | tctacaattt | 3480 |
| ggctctccac | tgagcacctt | atcttgccac | cttagcctta | agaatgaata | tgaagaaaaa | 3540 |
| tacacagcca | cctctgtcca | gggcagtaag | aagggctgca | aggaagggga | ggatggggac | 3600 |
| aaggaaagga | tcagatacct | gctccagtag | ttgtgaggcc | actgtgtctc | aggggactcc | 3660 |
| aggaggagca | gaagagggat | cccacgaagt | tattcttacg | cagctggggc | caggagggtc | 3720 |
| agagtgggtc | caggtgcaag | ttaggctaaa | gaagccacca | ctattcctct | ctcttgccca | 3780 |
| ttgtgggggg | caaaggcatt | ggtcaccaag | agtcttgag | ggggaccac | agatatgcca | 3840 |
| tgtccttcac | acgtgcttgg | gctccttaac | ctgaaggcaa | attgctactt | gcaagactga | 3900 |
| ctgacttcaa | ggaatcagaa | attacctaga | agcaccatgt | tttttctatg | accttttcag | 3960 |
| tccttcaggt | cattttaagg | tccactgcag | ggggttagtg | agaaaaggta | tactttgtgg | 4020 |
| tatgttttgc | tttctaata | gggacatgaa | ggaaaccag | caatttgctg | ttatgtgaat | 4080 |
| ggcctgtaga | gcagagtcaa | gagcgggtgtg | ctttgcccg | ctgctcccat | caggaatagg | 4140 |
| agagtagaca | gagatcttcc | acatcccagg | cttctgctgc | tgctttaaaa | gctctgtcct | 4200 |
| tggagcctcc | cgctccctga | agtgtctcgc | cccctgcaca | gcactggcct | ttcggagca | 4260 |
| tcccagtagg | gttttctgag | gctcgtggt | gactcatgcc | ctaattgcaa | tcctctgctt | 4320 |
| ttactttgac | ttgaaggat | ctaacactgc | tctctcttcc | aaaggggaaa | aaaagattca | 4380 |
| ttgttttgc | gcaataaaact | aatacaaaat | gatggccatt | catgtgcagc | tctttgtcac | 4440 |
| catgggccgg | atgagttgtg | ctcctcctgg | ctcaccattt | ccccctgctc | ccccacagcc | 4500 |
| ggttctgcac | ttatcacoga | gtcgcctctg | gaagcagatt | cccattgagt | tttccccacc | 4560 |
| aaggggacca | tgcatatggt | agaaacatta | gattctgcat | tgacagtagc | ctttccttgg | 4620 |
| cccggcctg | tgttgggaag | acgggcaaca | agtatacccc | accagggcct | gagtgcactag | 4680 |
| aggaagagga | cgaggccttg | ttggcactag | atttgggtat | tttctgcatg | tcataacata | 4740 |
| tcctaactgc | tattbcagaa | gaggcagctt | gtaggtgatt | gtacaagtga | gaattaaaga | 4800 |
| gagaacagat | atttaaacag | gtgctgtatt | agtaacagcc | agtgcctttt | cagcccttgc | 4860 |
| atctattaaa | aggagattca | ggattttatt | ggcacaggcc | cttcttagta | ggaagaaagg | 4920 |
| gtgcttagct | ttggacctga | ccgggtgtgt | gtaaaacat | ggactgagtc | acagcagaca | 4980 |
| ctcgatggtg | gtaaatgtga | tgggtgctta | cacactgtac | cttttctctt | catactgatg | 5040 |
| ctgcagtcca | gggctggagt | tgttaaggca | ttgacctcca | cccacctgcc | ccatgtccac | 5100 |
| tgggctgccc | aagctgcatg | tcacctgagg | gctggcagga | aggggcgaga | aatcccaggg | 5160 |
| cattgtacca | aggacctagt | tcttcttagg | gatataaatt | tccaggaatg | tgtattttta | 5220 |
| atgtggtgag | atgcaactct | ttgttgtacc | aaatagggct | ccccaccca | ccctgogac | 5280 |
| aagtgtctct | ctagaacagg | ttcctaccag | cagcactggt | gtgaatgaaa | gagagacca | 5340 |
| gccgctctc | acacaggtgg | aattgcactt | cttaacaaaa | aggaacttta | taaaagtttg | 5400 |
| ggattttttt | tcctaatacat | aaaaatagcc | ccagaaagag | cctaagctat | gttcagatag | 5460 |
| aagcctcgaa | attcctgtaa | attgtttact | ttatgatgtt | tacatacacg | tttacttttg | 5520 |
| aaaaaaatg | caactcgact | tttaacaac | tgtttgagatg | tttcatggga | cagtagaact | 5580 |
| ctgactcacc | aactgggcta | aatttttaatt | taaaaaatgta | tttattttgag | tgtctttccc | 5640 |
| cccctcacc | tcaccatctg | aggggctccc | tgagatcttg | gtagaggagg | cccctcctgc | 5700 |
| ccagaccttc | gtttgtttcc | ccgggtggccc | ttgcttcttg | ctttgcagac | tgctcgcagc | 5760 |
| catgattttg | tcaactgacat | ctgtgagcca | aagactgagc | ctttttggca | ggaataataa | 5820 |
| gcaatactac | acaacttgct | actttcagaa | aacttttttt | tagcttcacc | gatgacaaca | 5880 |
| gaggaagaag | ggaactggga | tttgggtaag | ttctctcca | ctgtttgacc | aaattctcag | 5940 |
| tgataaatat | gtgtgcagat | cctagaaga | gaaaacgctg | actttctttt | taagtgtggc | 6000 |
| acataaggat | ctgcagaatt | ttccgtagac | aaagaaagga | tcttgtgtat | ttttgtccat | 6060 |

ES 2 511 694 T3

```
atccaatggt atatgaacta attgtattgt tttatactgt gaccacaaat attatgcaat 6120  
gcaccatttg ttttttattt cattaaagga agttaaattt 6160
```

```
<210> 2  
<211> 3266  
<212> ADN
```

5 <213> Homo sapiens

```
<220>  
<221> característica_miscelánea  
<223> secuencia llamada GS-N2
```

```
<300>  
10 <308> GenBank/AF247704  
<309> 02-09-2000  
<313> (1)..(3266)
```

```
<400> 2
```


ES 2 511 694 T3

```

gCGGTGCGGG cCGGGCGGGT gCattcagGc caaggCGGGG cGccCGGgAT gctcagggTt 60
cGgGagccGc gGccCGGGgA gGcGaaagcG gAgggGgGccG cGccGccGac cccGtccaag 120
ccGctcaGct ccttcctcat ccagGacatc ctGcgggacG gCgCgCagcG gcaaggcGgC 180
cGcagcagGca gCcaGagaca gCgCgacccG gAgccGgagc cagagccaga gCcaGagggA 240
gGagcGagcc gCgCcgGGGc gCagaacGac cagctGagca cCGggccCG cGccGcGccG 300
gagGagGcGg agacGctGgC agagaccGag ccagaaagGc acttGggGtc ttatctgTtG 360
gactctGaaa acacttcagc cGcccttcca gGcttcccc aaaccctaa gCagccGcag 420
aagcGctccc gagctGcctt ctcccacact caggtGatcG agttGgagag gaagttcagc 480
catcagaagT acctGtcagc ccctGaacGg gccacctGg ccaagacct caagctcacG 540
gagaccCaag tgaagatagT gttccagaac agacGctata agactaagcG aaagcagctc 600
tcctcGgagc tgggagactt ggagaagcac tcctctttGc cggccctgaa agaggagggc 660
ttctccGgGg cctccGtGgT ctccGtGtat aaagcGtatc ctactacc atacctgtac 720
tgCtGggGca cTtggagcct agctttttGg taatGccagc tcaggtGaca accattatga 780
tcaaaaaactG ccttccccag ggtGtctcat atGaaaagca caaggGgcca aggtcagggA 840
gcaagaggtG tGcacaccaa aactatGga gaattGcGtG gaaatcttca gattcttcac 900
tggtGagaca atGaaacaac agagacagTg aaagTtttaa tacctaagtc attccccag 960
tgCatactgt agcGtcaagT ttttGcttct ggctacctgt ttgaagggA gagagggaaa 1020
atcaagTggt gTattttacag actttGtatg attttGtatg agctgtacac ccaaggattc 1080
tgTtctGcaa ctccatcctc ctgtGtcact gaaatcaac tctGaaagag caaacctaac 1140
aggagaaagG acaaccagga tgaggatGtC accaactgaa ttaaacttaa gtccagaagc 1200
ctcctGttGg ccttGgaata tggccaagGc tctctctGtC cctGtaaag agaggggcaa 1260
atagTctcca aagagaacGc cctcatGtC agcacatatt tgcatGgaag ggggagatGg 1320
gtgggagGag atGaaaatat cagcttttct tattcctttt tattcctttt aaaatggTat 1380
gccaacttaa gTattttacag ggtGgcccaa atagaacaag atGcactcGc tGtgatttta 1440
agacaagctG tataaacaga actccactGc aagagggagG gccggccag gagaatctcc 1500
gcttGtccaa gacagggGcc taaggaggtG ctccacactG ctGctagggG ctgttGcatt 1560
ttttattag tagaaagTgG aaaggcctct tctcaacttt tttcccttGg gctggagaat 1620
ttagaatcag aagTttcctGg gagTttcag gctatcatat atactgtatc ctGaaaggca 1680
acataattct tccttcctc cttttaaaat tttgtgttcc tttttGcagc aattactcac 1740
taaaggctt cattttagtc cagattttta gTctggctGc acctaactta tgctcGctt 1800
atttagccGg agatccGtGc tttttttttt tttttttttc cgtctcccca aagctttatc 1860
tgtcttGact ttttaaaaaa gtttggggGc agattctGaa ttggctaaaa gacatGcatt 1920
tttaaaacta gcaactctta tttcttctc ttaaaaatac atagcattaa atcccaaatc 1980
ctatttaaag acctGacagc ttgagaaggg tcaactactGc atttatagga ccttctGgtG 2040
gttctGctgt tacgTttgaa gTctGacaat ccttGagaat ctttGcatGc agaggaggtA 2100
agagTtattg gattttcaca gagggaagaa cacagcGcag aatgaaggcG caggcttact 2160
gagctGtcca gtggagggct catgggtggg acatGgaaaa gaaggcagcc taggccctGg 2220
ggagccGagT ccactGagca agcaagggac tgagTgagcc ttttGcagga aaaggctaag 2280
aaaaagGaaa accattctaa aacacaacaa gaaactGtcc aaatGctttg ggaactGtGt 2340
ttattGccta taatgggGtc cccaaaatGg gtaacctaga cttcagagag aatgagcaga 2400
gagcaaaGga gaaatctggG ctGtcttcc attttcattc tGttatctca ggtgagctGg 2460
tagagggGag acattagaa aaaatGaaac acaaaaacaa ttactaatga ggtacGctga 2520
ggcctggGag tctcttGagc tccacgactt caaaaattaaa atgagccatg agtcaaacca 2580
ctGcaatcca gcttggGcaa cGagcaagac ccagTctcta ctGttGgtGg caaaattGcc 2640
aacataagTt aatagaaagT tggccaattt caccctattt tctGtGgttt gggctccaca 2700
ttGcaatGtt caatGccagc tGctGctGac accgaccGga gTactagcca gcaaaaaagG 2760
cagggtagcc tgaattGctt tctGctcttt acatttcttt taaaataagc atttagTgct 2820
cagTccctac tGagTactct ttctctccc tctctGaat ttactcttt caacttGcaa 2880
tttGcaagG ttacacattt cactGtGatg tatattGtGt tGcaaaaaa aaaaagTgTc 2940
tttGtttaa attacttGgt ttgtGaatcc atcttGcttt ccccatGga actagTcatt 3000
aaccatctc tgaactGgta gaaaaacatc tgaagagcta gTctatcagc atctGacagG 3060
tgaattGgat gTtctcaga accattcac ccagacagcc tGtttctatc ctGtttaata 3120
aattagTttg gTtctctac atGcataaca aaccctGtC caatctGtca cataaaagTc 3180
tGtgacttga agTttagTca gCacccccac caaactttat ttttctatgt gttttttGca 3240
acatatgagT gTtttGaaaa taaagT 3266

```

<210> 3
 <211> 6771
 5 <212> ADN
 <213> Homo sapiens

<220>
 <221> característica_misclánea
 <223> secuencia llamada GS-N3

10 <300>
 <308> GenBank/U91543
 <309> 04-08-1998

ES 2 511 694 T3

<313> (1)..(6771)

<400> 3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|------------|------|
| atatggaggt | gaaggggtgag | atcgggaaac | aaagggtagt | gccccctagt | tcccaaaggg | 60 |
| agcaggggaga | tgggaataga | attgaaggta | ggtttttaggc | tacttggggag | gaggaatatt | 120 |
| taggtaattg | tggagacttt | ctcctgtgtg | atgaaggcgg | cagacactgt | gatcctgtgg | 180 |
| gcaagaagta | aaaatgacca | gctgaggatt | tcttttcctc | caggactgtg | ttggggtgac | 240 |
| aggatgcctg | ataaggatga | cattcggctg | ctgccgtcag | cattgggtgt | gaagaagaga | 300 |
| aaacgaggac | ccaagaagca | gaaggagaac | aagccaggaa | aaccccgaaa | acgcaagaag | 360 |
| cgtgacagtg | aggaggaatt | tggttctgag | cgagatgagt | accggggagaa | gtcagagagt | 420 |
| gggggcagtg | aatatggaac | cggaccgggt | cggaaacgaa | gaaggaagca | ccgagaaaaa | 480 |
| aaggagaaga | agacaaagcg | gcggaaaaag | ggggaggggag | atggggggca | aaagcaagtg | 540 |
| gaacagaagt | catcagcaac | tctgcttctg | acctggggcc | tggaggatgt | ggagcatgtg | 600 |
| ttctctgagg | aggattacca | cacgctcacc | aactacaaa | ccttcagcca | gttcatgagg | 660 |
| cccctaattg | ctaagaagaa | tcctaagatc | ccaatgtcta | agatgatgac | catccttggg | 720 |
| gccaaatgga | gagagttcag | tgccaacaac | cccttcaagg | ggtcagcagc | tgctgtggcg | 780 |
| gcggcagcgg | cagcagcagc | agcagctgta | gctgagcagg | tgtcagctgc | tgtctcgtcg | 840 |
| gccaccccca | tagcaccctc | cggacccccc | gcccttccac | caccccctgc | tgctgataac | 900 |
| cagcccccac | ccatccgaag | agccaaaacc | aaagagggca | aaggctccagg | ccataagagg | 960 |
| cgagtaaga | gcccccgagt | gcctgatgga | cgcaagaagc | ttcggggaaa | gaaaatggca | 1020 |
| ccactcaaaa | taaaactag | gcttctgggt | ggcaagagga | agaagggagg | ctcgtatggt | 1080 |
| tttcagagcg | acgaaggtcc | tgaaccagag | gctgaggaat | cagacctgga | cagtggcagt | 1140 |
| gtccacagtg | cctcaggccg | gcctgatggc | cctgtccgca | ccaagaaact | aaagagaggc | 1200 |
| cgccacaggaa | ggaagaagaa | gaaggtcctg | ggctgtcctg | cagtggcccg | ggagaggagg | 1260 |
| gttgatggct | acgagacgga | tcaccaggat | tactgtgagg | tgtgccagca | gggtggggaa | 1320 |
| attatctctg | gtgacacctg | ccctcgtgcc | taccacctcg | tctgccttga | tcctgagctt | 1380 |
| gaccgggctc | cagagggcaa | atggagctgc | cctcactgtg | agaagggagg | ggccaagtgg | 1440 |
| gaggccaagg | aggaagaaga | agaatacga | gaggaggggag | aggaagaagg | ggagaaggag | 1500 |
| gaggaggatg | atcacatgga | gtactgccgc | gtatgcaagg | acggcgggga | gctcctgtgc | 1560 |
| tgtgacgcgt | gcattctctc | ctaccacatt | cattgtctaa | accctcccct | gcctgacatt | 1620 |
| cccaatggtg | aatggctgtg | tccccgatgc | acatgccccg | tgtgaaggg | tcgagtgcag | 1680 |
| aagatcctac | attggcggtg | gggggagcca | cctgtagcag | tgccagcccc | tcaacaggca | 1740 |
| gatggaaatc | cagatgtccc | accccccctg | cctcttcaag | gcagatcaga | gcgagagttc | 1800 |
| tttgtcaagt | gggtaggact | atcctactgg | cactgctcct | gggccaagga | gcttcagctg | 1860 |
| gaaatcttcc | atttggttat | gtatcgaaac | taccagcggg | agaatgacat | ggatgagccc | 1920 |
| ccaccctctg | actatggctc | cggcggaggat | gatgggaaga | gcgacaagcg | taaagtgaaa | 1980 |
| gaccgcact | atgctgagat | ggaggagaag | tactatcgtt | ttggcatcaa | gccagagtgg | 2040 |
| atgaccgtcc | accgcatcat | caaccacagt | gtggataaaa | aggggaatta | ccactatcta | 2100 |
| gtaaaatgga | gggacttacc | atatgaccag | tccacgtggg | aggaagatga | aatgaatata | 2160 |
| cctgaatacg | aagaacataa | gcaaagctac | tggagacacc | gagaactaat | tatgggggaa | 2220 |
| gaccctgccc | agccccgcaa | gtataagaag | aagaagaagg | agctacaggg | tgatgggcct | 2280 |
| cccagttctc | ccactaatga | tcttaccgtg | aaatatgaga | ctcagccacg | gtttatcaca | 2340 |
| gccactggag | gcaccctgca | catgtatcag | ttggaagggc | tgaactggct | acgcttctcc | 2400 |
| tggggccagg | gcactgacac | cattctagct | gatgagatgg | ggctaggcaa | gaccatacaa | 2460 |
| accatcgtct | tcctctactc | actctacaag | gagggccaca | caaaaggtcc | cttctctggg | 2520 |
| agtgccccac | tctctaccat | cattaactgg | gagcgggagt | tccagatgtg | ggcaccctaa | 2580 |
| ttctatgtgg | tgacatacac | gggtgacaag | gacagccggg | ccatcattcg | tgagaatgaa | 2640 |
| ttctcctttg | aggacaatgc | catcaaaggg | ggcaagaag | cttttaagat | gaagagggag | 2700 |
| gcacaggtga | agttccatgt | tctcctgaca | tcgtagagc | tgatcaccat | tgatcaggca | 2760 |
| gcacttgggt | ccatccgctg | ggcctgtctt | gtggtagatg | aggcccatcg | actcaagaac | 2820 |
| aaccagtcca | agttttctcag | ggttctcaat | ggttacaaga | tagatcataa | gttgctgctg | 2880 |
| acaggaaccc | cattgcagaa | taatctggag | gagctcttcc | atctcctgaa | cttctctacc | 2940 |

ES 2 511 694 T3

ccagagagat ttaacaactt ggagggcttc ctggaggagt ttgctgacat atccaaagag 3000
gaccagatca agaaactgca tgatttgctg gggccacaca tgctgaggag actcaaggca 3060
gatgtcttta agaacatgcc agccaagaca gagctcatcg ttcgggtgga gctaagcccc 3120
atgcagaaga aatactacaa atacatcctg actcgaattt ttgaggcctt gaattcacga 3180
gggtggtggga accaggtgtc gctgcttaat atcatgatgg atcttaagaa gtgctgcaac 3240
catccatacc tttttcccggt ggctgctatg gactccccca aactccccag tggggcttat 3300
gagggtgagg cacttattaa gtcgtctggg aagctcatgc tgctccagaa gatgctgcca 3360
aagctgaagg agcaaggaca ccgagtgtc atcttctcgc agatgaccaa aatggttagac 3420
ttgcttgagg acttcttaga ctatgaaggc tacaagtatg agcgcacga tgggtgtatc 3480
acgggtgccc tgaggcagga ggccatcgat cgggtttaatg ctcctggggc ccaacaattc 3540
tgcttctccc tgtccaccgg agctgggggc ctgggcatca atctggccac tgctgacact 3600
gtcatcatct ttgattctga ctggaacccc cataatgaca tccaggcctt tagccgggct 3660
catcggattg gccaggccaa caaagtatg atttaccggt ttgtgactcg cgcgtcagtg 3720
gaagagcgaa tcacacaagt ggccaagaga aagatgatgc tgacacacct ggttgtgcgg 3780
cctgggctgg gctccaaggc aggctccatg tccaagcagg agcttgacga cattctcaa 3840
tttggcactg aagagctatt caaggatgaa aacgaggggg agaacaagga ggaggacagc 3900
agtgtgatcc attatgacaa tgaggccatc gctcggctgt tggaccggaa ccaggatgca 3960
actgaggaca ctgacgtgca gaacatgaat gactatctca gctccttcaa ggtggcacag 4020
tacgtcgtgc gccaggaaga caagattgag gaaattgagc gagagatcat caagcaggag 4080
gagaatgtgg accctgacta ctgggagaag ctgctgaggc atcactatga gcaacagcag 4140
gaagacctag cccggaatct aggcaagggc aagcgggttc gcaagcaagt taactacaat 4200
gatgctgctc aggaagacca agacaaccag tcagagtact cgttgggttc agaggaggag 4260
gatgaagact tcgatgaacg tcctgaaggg cgtagacagt caaagaggca gctccggaat 4320
gagaaagata gcccactgcc tccactgctg gccagagtcg ggggcaacat tgagggtctg 4380
ggcttaaca cccgtcagcg gaaggcttct ccaatgctg tgatgcgctg ggggatgcca 4440
ccacaggatg ccttaccac acagtggctg gtgcccggacc tgaggggcaa gactgagaag 4500
gagtttaagg cctatgtgtc tttgttcatg cgccatctgt gtgagcctgg ggcagacggc 4560
tctgaaacct ttgccgatgg ggtcccctcg gagggactga gtcgccagca ggtgttgacc 4620
cgcattggag tcattgtctc cgtcaaaaag aagggtgcagg agtttgagca catcaatggg 4680
ccgttgcaaa tggcggaaact gccagcgcgg attctaagcg ctctccaga 4740
gcctcctctc ctacaaaac gtctcccacc actcctgagg cttctgctac caacagtccc 4800
tgcacctcta aacctgctac tccagctcca agtgagaaag gagaaggcat aaggacacct 4860
cttgagaagg aggaagctga aaaccaggag gaaaagccag agaagaacag cagaattggg 4920
gagaagatgg agacagaggc tgatgcccc agcccagccc catcacttgg ggagcggctg 4980
gagccaagga agattcctct agaggatgag gtgccagggg tgccctggaga gatggagcct 5040
gaacctgggt accgtgggga cagagagaag tcagccacag agtcgacgcc aggagaaaag 5100
ggggaggaga agccgttggg tggacaggaa cacagggaga ggccggaggg ggaacaggg 5160
gatttgggca agagagaaga tgtaaagggt gaccgggagc ttcgaccagg gcctcgagat 5220
gagccacggg ccaatgggcy acgagaggaa aagacagaga accccgggtt catgttcaat 5280
atcgccgatg gtggcttcac agagcttcc acactgtggc agaatgagga acgggcagct 5340
atctcctcgg gaaactcaa tgagatctgg cacagaagac atgactattg gcttctggct 5400
gggatgtcc tccatggcta tgcaaggttg caggacatcc agaatgatgc tcaatttggc 5460
attatcaacg agccatttaa aactgaagcc aataagggga acttcttggg gatgaaaaat 5520
aagttcctgg cccggaggtt caagctcctg gagcaggcgc tgggtattga ggagcagctg 5580
cgggcggcgg cctacctgaa cctgtcgcag gagccggcgc accccgccat ggccctccac 5640
gcccgttctg ccgaggccga gtgctggcc gagagccacc agcacctctc caaggagtctg 5700
ctggcgggga acaagccggc caacgccgtc ctgcacaagg ttctgaacca gctggaggag 5760
ttgctgagcg acatgaaggc ggacgtgacc cgcctgccag ccacgctgtc ccgaataacc 5820
cccatcgcag cccgccttca gatgtccgag cgcagcatcc tcagccggct ggccagcaag 5880
ggcacggagc ctaccccc accggcctac ccgcccgggt cctacgctac acctccgggg 5940
tacggggcgg ccttcagcgc cgcaccgta ggggccctgg ccgcccagc gcceaattac 6000
agccagatgc ctgcagggtc ctatcatcaca gccgccacca acggccctcc agtgcttgtg 6060
aagaaggaga aggaaatggt gggggcattg gtgtcagac ggctggatcg gaaggagccc 6120
cgagccgggg aggtgatctg tatagacgac tgactggatc ccaggcctgc ccttaccaca 6180
ggccccgtcc ccgaggccga cccccagctc aagcgttggg gcctgctgct agccctccac 6240
cttccccacc ccttgggcca tcaactgggt aggaaccctt ttgccctct ctgcagctcc 6300
tctcttcaag aagggccctt tgtcttctc cactcccaca caccttctcc accaagcctt 6360
gaagactgtg ctgggtgagaa gaagtctggg tgggagatgg ctggcagggt cttccaagta 6420
ccttctccc ccttaccgaa gtatacaca cttcccagta aatggttgtg gggaggaag 6480
aggtggagcc tccccagcy tttccctgca gaatcagctc tgtctcatgt ggaagtggag 6540
aatcagcctt gcctggcctt taggaacttt tgtggggaag agagcttga agagaggagg 6600
gggacttag agaggatga aaatgagccc tgggaggag gaaggagcga ggagggtgg 6660
ctgcatgtta ccgtccccta cctctcccga cgtggagggt ggagcagtta tgaggaggga 6720
agtcaactgc tgttcagcct cagaataaag gtgccgttca ctggctcagt t 6771

<210> 4
<211> 3041
<212> ADN

ES 2 511 694 T3

<213> Homo sapiens

<220>

<221> característica_misclánea

<223> secuencia llamada GS-N4

5

<300>

<308> GenBank/BC001571

<309> 29-10-2001

<313> (1)..(3041)

<400> 4

```

gtcagggccc tgcggtgtga ctgcgggct cagctggtcc ggccgtagca cctccgcgcc 60
gtcgcgatgt cgcgggtttt caccaccggt teggacagcg agtccgagtc gtccttgtcc 120
ggggaggagc tcgtcaccaa acctgtcggg ggaactatg gcaaacagcc attgttgctg 180
agcgaaggat aagaagatac caagagagtt gtccgcagtg ccaaggacaa gaggtttgag 240
gagctgacca accttatccg gaccatccgt aatgccatga agattcgtga tgtcaccaag 300
tgcctggaag agtttgagct cctgggaaaa gcatatggga aggcacaaaag cattgtggac 360
aaagaagggt tcccccggtt ctatatccgc atcctggctg acctagagga ctatcttaat 420
gagctttggg aagataagga agggaagaag aagatgaaca agaacaatgc caaggctctg 480
agcaccttgc gtcagaagat ccgaaaatac aaccgtgatt tcgagtccca taccacaagc 540
tacaagcaga accccgagca gtctgcggat gaagatgctg agaaaaatga ggaggattca 600
gaaggctctt cagatgagga tgaggatgag gacggagtca gtgctgcaac tttcttgaag 660
aagaaatcag aagctccttc tggggagagt cgcaagttcc tcaaaaagat ggatgatgaa 720
gatgaggact cagaagattc cgaagatgat gaagactggg acacaggttc cacatcttcc 780
gattccgact cagaggagga agaagggaaa caaacccgctc tggcctcaag atttcttaaa 840
aaggcaccca ccacagatga ggacaagaag gcagccgaga agaaacggga ggacaaagct 900
aagaagaagc acgacaggaa atccaagcgc ctggatgagg aggaggagga caatgaaggc 960
ggggagtggg aaagggtccg gggcggagtg ccgttggtta aggagaagcc aaaaatgttt 1020
gccaagggaa ctgagatcac ccatgctgtt gtatcaaga aactgaatga gatcctacag 1080
gcacgaggca agaagggaaac tgatcgtgct gccagattg agctgctgca actgctgggt 1140
cagattgcag cggaaaacaa cctgggagag ggcgtcattg tcaagatcaa gttcaatatt 1200
atcgctctc tctatgacta caacccaac ctggcaacct acatgaagcc agagatgtgg 1260
gggaaagtgc ccaaggagct tctggccag ggctgctgc tgtttgcaaa tccaacatt 1320
tttgttgag agaatattct ggaagagagt gagaacctgc acaacgctga ccagccactg 1380
cgtgtccgtg gctgcatcct aactctggtg gaacgaatg atgaagaatt taccaaaata 1440
atgcaaaata ctgacctca ctcccaagag tacgtggagc acttgaagga tgaggccag 1500
gtgtgtgcca tcatcgagcg tgtgcagcgc tactggagg agaagggcac taccgaggag 1560
gtctgccgca tctacctgac ggcacacct cacacctact acaagtttga ttacaaggcc 1620
catcacgac agctgacccc cctgagggc ctctcaaagt ctgagcaaga ccaggcagaa 1680
aatgagggcg aggactcggc tgtgttgatg gagagactgt gcaagtacat ctacgccaag 1740
gaccgcacag accggatccg cacatgtgcc atcctctgcc acatctacca ccatgctctg 1800
cactcgcgct gttaccaggc ccgcgacctc atgctcatga gccacttgca ggacaacatt 1860
cagcatgcag acccgccagt gcagatcctt tacaaccgca ccatggtgca gctgggcatc 1920
tgtgccttcc gccaaaggct gaccaaggac gcacacaacg cctgctgga catccagtgc 1980
agtggccgag ccaaggagct tctggccag ggctgctgc tgcgcagcct gcaggagcgc 2040
aaccaggagc aggagaaggt ggagcggcgc cgtcaggtcc ccttccacct gcacatcaac 2100
ctggagctgc tggagtgtgt ctacctggtg tctgccatgc tctggagat ccctacatg 2160
gccgccatg agagcgatgc ccgccagcgc atgatcagca agcagttcca ccaccagctg 2220
cgcgtgggcy agcgacagcc cctgctgggt cccctgagt ccatgcggga acatgtggtc 2280
gctgcctoca aggccatgaa gatgggtgac tggaaacct gtcacagttt tatcatcaat 2340
gagaagatga atgggaaagt gtgggacctt ttccccgagg ctgacaaagt ccgaccatg 2400
ctggttagga agatccagga agagtactg aggacctacc tcttacctta cagcagtgtc 2460
tatgactoca tcagcatgga gacgctgtca gacatgtttg agctggatct gccactgtg 2520
cactccatca tcagcaaaat gatcattaat gaggagctga tggcctccct ggaccagcca 2580
acacagacag tggatgatgca ccgactgag cccactgccc agcagaacct ggctctgcag 2640
ctggccgaga agctgggagc cctgggtggag aacaacgaac ggggtgttga ccacaagcag 2700
ggcacctacg ggggctactt ccgagaccag aaggcagcct accgcaaaaa cgagggtac 2760
atgcgccgcy gtggctaccg ccagcagcag tctcagacgg cctactgagc tctccactct 2820
gtttcccgcc tgggcatcc aaccttgaag tcttaacca cacctcagtc actaaaggtc 2880
tgtttaaagt tgttctggtt gattgcttgt tgccacaaag gtgtgttgag cccattcatt 2940
tgctgattat atttctgagc tgctaccaca tgacagacac aaggctcaga tgaagagggt 3000
aattaaacac gttcctgacc tcaaaaaaaaa aaaaaaaaaa a 3041

```

10

<210> 5

<211> 1507

<212> ADN

ES 2 511 694 T3

<213> Homo sapiens

<220>

<221> característica_miscelánea

<223> secuencia llamada GS-N5

5 <300>

<308> GenBank/BC000533

<309> 12-07-2001

<313> (1)..(1507)

<400> 5

```

ggcacgaggg ggccctgcgg tgtgactcgc gggctcagct ggtccggccg tagcacctcc 60
gcgccgtcgc catgtcgcgg tttttacca cgggttcgga cagcgagtcc gagtcgtcct 120
tgtccgggga ggagctcgtc accaaacctg tcggaggcaa ctatggcaaa cagccattgt 180
tgctgagcga ggatgaagaa gataccaaga gatttgccg cagtgcgaag gacaagaggt 240
ttgaggagct gaccaacctt atccggacca tccgtaatgc catgaagatt cgtgatgtca 300
ccaagtgcct ggaagagttt gagctcctgg gaaaagcaca tgtgccatcc tctgccacat 360
ctaccaccat gctctgcact cgcgctggta ccaggcccgc gacctcatgc tcatgagcca 420
cttgacggac aacattcagc atgcagacct gccagtgcag atcctttaca accgcacat 480
ggtgcagctg ggcatctgtg ctttccgcca agcctgacc aaggacgcac acaacgcct 540
gctggacatc cagtcgagtg gccgagccaa ggagcttctg ggccagggcc tgctgctgcg 600
cagcctgcag gagcgcaacc aggagcagga gaaggtggag cggcgccgctc aggtcccctt 660
ccacctgcac atcaacctgg agctgctgga gtgtgtctac ctggtgtctg ccatgctcct 720
ggagatcccc tacatggccg cccatgagag cgatgcccg cgcagcatga tcagcaagca 780
gttccaccac cagctgcgcg tgggcgagcg acagcccctg ctgggtcccc ctgagtccat 840
gcgggaaacat gtggctcgtg cctccaaggc catgaagatg ggtgactgga agacctgtca 900
cagttttatc atcaatgaga agatgaatgg gaaagtgtg gaccttttcc ccgaggctga 960
caaagtccgc accatgctgg ttaggaagat ccaggaagag tactgagga cctacctctt 1020
cacctacagc agtgtctatg actccatcag catggagacg ctgtcagaca tgtttgagct 1080
ggatctgccc actgtgcact ccatcatcag caaatgatc ataatgagg agctgatggc 1140
ctccctggac cagccaacac agacagtggg gatgcaccgc actgagccca ctgcccagca 1200
gaacctggct ctgcagctgg ccgagaagct gggcagcctg gtggagaaca acgaacgggt 1260
gtttgaccac aagcagggca cctacggggg ctacttccga gaccagaagg acggctaccg 1320
caaaaacgag ggctacatgc gccgcggtgg ctaccgccag cagcagtctc agacggccta 1380
ctgagctctc cactctgttt cccgcctggg ccatccaacc ttgaagtctt aaaccacacc 1440
tcagtcaacta aaggtctgtt taaagttaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1500
aaaaaaaa 1507

```

10

<210> 6

<211> 924

<212> PRT

<213> Homo sapiens

15 <220>

<221> característica_miscelánea

<223> CDS parcial de la secuencia GS-N1:

<1..2777

<220>

20 <221>

característica_miscelánea

<223> secuencia llamada GS-P1: Angipartnerina

<400> 6

ES 2 511 694 T3

Arg Glu Glu Ser Gly Ala Gly Ala Arg Pro Arg Arg Arg Ser Ala Asp
1 5 10 15
Ser Gly Ala Ala Gly Ala Gly Arg Gly Gly Gly Gly Glu Ala Ala Gly
20 25 30
Lys Glu Glu Glu Gly Glu Ser Arg Ser Arg Arg Ala Ser Met Gly Arg
35 40 45
Leu Ala Ser Arg Pro Leu Leu Leu Ala Leu Leu Ser Leu Ala Leu Cys
50 55 60
Arg Gly Arg Val Val Arg Val Pro Thr Ala Thr Leu Val Arg Val Val
65 70 75 80
Gly Thr Glu Leu Val Ile Pro Cys Asn Val Ser Asp Tyr Asp Gly Pro
85 90 95

ES 2 511 694 T3

Ser Glu Gln Asn Phe Asp Trp Ser Phe Ser Ser Leu Gly Ser Ser Phe
100 105 110

Val Glu Leu Ala Ser Thr Trp Glu Val Gly Phe Pro Ala Gln Leu Tyr
115 120 125

Gln Glu Arg Leu Gln Arg Gly Glu Ile Leu Leu Arg Arg Thr Ala Asn
130 135 140

Asp Ala Val Glu Leu His Ile Lys Asn Val Gln Pro Ser Asp Gln Gly
145 150 155 160

His Tyr Lys Cys Ser Thr Pro Ser Thr Asp Ala Thr Val Gln Gly Asn
165 170 175

Tyr Glu Asp Thr Val Gln Val Lys Val Leu Ala Asp Ser Leu His Val
180 185 190

Gly Pro Ser Ala Arg Pro Pro Pro Ser Leu Ser Leu Arg Glu Gly Glu
195 200 205

Pro Phe Glu Leu Arg Cys Thr Ala Ala Ser Ala Ser Pro Leu His Thr
210 215 220

His Leu Ala Leu Leu Trp Glu Val His Arg Gly Pro Ala Arg Arg Ser
225 230 235 240

Val Leu Ala Leu Thr His Glu Gly Arg Phe His Pro Gly Leu Gly Tyr
245 250 255

Glu Gln Arg Tyr His Ser Gly Asp Val Arg Leu Asp Thr Val Gly Ser
260 265 270

Asp Ala Tyr Arg Leu Ser Val Ser Arg Ala Leu Ser Ala Asp Gln Gly
275 280 285

Ser Tyr Arg Cys Ile Val Ser Glu Trp Ile Ala Glu Gln Gly Asn Trp
290 295 300

Gln Glu Ile Gln Glu Lys Ala Val Glu Val Ala Thr Val Val Ile Gln
305 310 315 320

Pro Ser Val Leu Arg Ala Ala Val Pro Lys Asn Val Ser Val Ala Glu
325 330 335

Gly Lys Glu Leu Asp Leu Thr Cys Asn Ile Thr Thr Asp Arg Ala Asp
340 345 350

Asp Val Arg Pro Glu Val Thr Trp Ser Phe Ser Arg Met Pro Asp Ser
355 360 365

Thr Leu Pro Gly Ser Arg Val Leu Ala Arg Leu Asp Arg Asp Ser Leu
370 375 380

Val His Ser Ser Pro His Val Ala Leu Ser His Val Asp Ala Arg Ser
385 390 395 400

Tyr His Leu Leu Val Arg Asp Val Ser Lys Glu Asn Ser Gly Tyr Tyr
405 410 415

Tyr Cys His Val Ser Leu Trp Ala Pro Gly His Asn Arg Ser Trp His
420 425 430

Lys Val Ala Glu Ala Val Ser Ser Pro Ala Gly Val Gly Val Thr Trp
435 440 445

ES 2 511 694 T3

Leu Glu Pro Asp Tyr Gln Val Tyr Leu Asn Ala Ser Lys Val Pro Gly
 450 455 460
 Phe Ala Asp Asp Pro Thr Glu Leu Ala Cys Arg Val Val Asp Thr Lys
 465 470 475 480
 Ser Gly Glu Ala Asn Val Arg Phe Thr Val Ser Trp Tyr Tyr Arg Met
 485 490 495
 Asn Arg Arg Ser Asp Asn Val Val Thr Ser Glu Leu Leu Ala Val Met
 500 505 510
 Asp Gly Asp Trp Thr Leu Lys Tyr Gly Glu Arg Ser Lys Gln Arg Ala
 515 520 525
 Gln Asp Gly Asp Phe Ile Phe Ser Lys Glu His Thr Asp Thr Phe Asn
 530 535 540
 Phe Arg Ile Gln Arg Thr Thr Glu Glu Asp Arg Gly Asn Tyr Tyr Cys
 545 550 555 560
 Val Val Ser Ala Trp Thr Lys Gln Arg Asn Asn Ser Trp Val Lys Ser
 565 570 575
 Lys Asp Val Phe Ser Lys Pro Val Asn Ile Phe Trp Ala Leu Glu Asp
 580 585 590
 Ser Val Leu Val Val Lys Ala Arg Gln Pro Lys Pro Phe Phe Ala Ala
 595 600 605
 Gly Asn Thr Phe Glu Met Thr Cys Lys Val Ser Ser Lys Asn Ile Lys
 610 615 620
 Ser Pro Arg Tyr Ser Val Leu Ile Met Ala Glu Lys Pro Val Gly Asp
 625 630 635 640
 Leu Ser Ser Pro Asn Glu Thr Lys Tyr Ile Ile Ser Leu Asp Gln Asp
 645 650 655
 Ser Val Val Lys Leu Glu Asn Trp Thr Asp Ala Ser Arg Val Asp Gly
 660 665 670
 Val Val Leu Glu Lys Val Gln Glu Asp Glu Phe Arg Tyr Arg Met Tyr
 675 680 685
 Gln Thr Gln Val Ser Asp Ala Gly Leu Tyr Arg Cys Met Val Thr Ala
 690 695 700
 Trp Ser Pro Val Arg Gly Ser Leu Trp Arg Glu Ala Ala Thr Ser Leu
 705 710 715 720
 Ser Asn Pro Ile Glu Ile Asp Phe Gln Thr Ser Gly Pro Ile Phe Asn
 725 730 735
 Ala Ser Val His Ser Asp Thr Pro Ser Val Ile Arg Gly Asp Leu Ile
 740 745 750
 Lys Leu Phe Cys Ile Ile Thr Val Glu Gly Ala Ala Leu Asp Pro Asp
 755 760 765
 Asp Met Ala Phe Asp Val Ser Trp Phe Ala Val His Ser Phe Gly Leu
 770 775 780
 Asp Lys Ala Pro Val Leu Leu Ser Ser Leu Asp Arg Lys Gly Ile Val
 785 790 795 800
 Thr Thr Ser Arg Arg Asp Trp Lys Ser Asp Leu Ser Leu Glu Arg Val

ES 2 511 694 T3

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| | 805 | | 810 | | 815 | | | | | | | | | | | | | | |
| Ser | Val | Leu | Glu | Phe | Leu | Leu | Gln | Val | His | Gly | Ser | Glu | Asp | Gln | Asp | | | | |
| | | | 820 | | | | | 825 | | | | | 830 | | | | | | |
| Phe | Gly | Asn | Tyr | Tyr | Cys | Ser | Val | Thr | Pro | Trp | Val | Lys | Ser | Pro | Thr | | | | |
| | | 835 | | | | | 840 | | | | | 845 | | | | | | | |
| Gly | Ser | Trp | Gln | Lys | Glu | Ala | Glu | Ile | His | Ser | Lys | Pro | Val | Phe | Ile | | | | |
| | 850 | | | | | 855 | | | | | 860 | | | | | | | | |
| Thr | Val | Lys | Met | Asp | Val | Leu | Asn | Ala | Phe | Lys | Tyr | Pro | Leu | Leu | Ile | | | | |
| 865 | | | | | 870 | | | | | 875 | | | | | 880 | | | | |
| Gly | Val | Gly | Leu | Ser | Thr | Val | Ile | Gly | Leu | Leu | Ser | Cys | Leu | Ile | Gly | | | | |
| | | | | 885 | | | | | 890 | | | | | | 895 | | | | |
| Tyr | Cys | Ser | Ser | His | Trp | Cys | Cys | Lys | Lys | Glu | Val | Gln | Glu | Thr | Arg | | | | |
| | | | 900 | | | | | 905 | | | | | 910 | | | | | | |
| Arg | Glu | Arg | Arg | Arg | Leu | Met | Ser | Met | Glu | Met | Asp | | | | | | | | |
| | | 915 | | | | | 920 | | | | | | | | | | | | |

<210> 7
 <211> 234
 <212> PRT
 5 <213> Homo sapiens

<220>
 <221> característica_misclánea
 <223> CDS de la secuencia GS-N2: 49..753

10 <220>
 <221> característica_misclánea
 <223> secuencia llamada GSP2: proteína homeobox NKX3.1

<400> 7

ES 2 511 694 T3

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

5 <210> 8
 <211> 2000
 <212> PRT
 <213> Homo sapiens

<220>
 <221> característica_misclánea
 <223> CDS de la secuencia GS-N3: 151..6153

10 <220>
 <221> característica_misclánea
 <223> secuencia llamada GS-P3: helicasa con dedo de Zinc

<400> 8

ES 2 511 694 T3

Met Lys Ala Ala Asp Thr Val Ile Leu Trp Ala Arg Ser Lys Asn Asp
1 5 10 15

Gln Leu Arg Ile Ser Phe Pro Pro Gly Leu Cys Trp Gly Asp Arg Met
20 25 30

Pro Asp Lys Asp Asp Ile Arg Leu Leu Pro Ser Ala Leu Gly Val Lys
35 40 45

Lys Arg Lys Arg Gly Pro Lys Lys Gln Lys Glu Asn Lys Pro Gly Lys
50 55 60

Pro Arg Lys Arg Lys Lys Arg Asp Ser Glu Glu Glu Phe Gly Ser Glu
65 70 75 80

Arg Asp Glu Tyr Arg Glu Lys Ser Glu Ser Gly Gly Ser Glu Tyr Gly
85 90 95

Thr Gly Pro Gly Arg Lys Arg Arg Arg Lys His Arg Glu Lys Lys Glu
100 105 110

Lys Lys Thr Lys Arg Arg Lys Lys Gly Glu Gly Asp Gly Gly Gln Lys
115 120 125

Gln Val Glu Gln Lys Ser Ser Ala Thr Leu Leu Leu Thr Trp Gly Leu
130 135 140

Glu Asp Val Glu His Val Phe Ser Glu Glu Asp Tyr His Thr Leu Thr
145 150 155 160

Asn Tyr Lys Ala Phe Ser Gln Phe Met Arg Pro Leu Ile Ala Lys Lys
165 170 175

Asn Pro Lys Ile Pro Met Ser Lys Met Met Thr Ile Leu Gly Ala Lys
180 185 190

Trp Arg Glu Phe Ser Ala Asn Asn Pro Phe Lys Gly Ser Ala Ala Ala
195 200 205

Val Ala Ala Ala Ala Ala Ala Ala Ala Ala Val Ala Glu Gln Val

ES 2 511 694 T3

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 210 | | | | | | 215 | | | | | | | | | | 220 |
| Ser | Ala | Ala | Val | Ser | Ser | Ala | Thr | Pro | Ile | Ala | Pro | Ser | Gly | Pro | Pro | |
| 225 | | | | | 230 | | | | | 235 | | | | | | 240 |
| Ala | Leu | Pro | Pro | Pro | Pro | Ala | Ala | Asp | Ile | Gln | Pro | Pro | Pro | Ile | Arg | |
| | | | | 245 | | | | | 250 | | | | | 255 | | |
| Arg | Ala | Lys | Thr | Lys | Glu | Gly | Lys | Gly | Pro | Gly | His | Lys | Arg | Arg | Ser | |
| | | | 260 | | | | | 265 | | | | | 270 | | | |
| Lys | Ser | Pro | Arg | Val | Pro | Asp | Gly | Arg | Lys | Lys | Leu | Arg | Gly | Lys | Lys | |
| | | 275 | | | | | 280 | | | | | 285 | | | | |
| Met | Ala | Pro | Leu | Lys | Ile | Lys | Leu | Gly | Leu | Leu | Gly | Gly | Lys | Arg | Lys | |
| | 290 | | | | | 295 | | | | | 300 | | | | | |
| Lys | Gly | Gly | Ser | Tyr | Val | Phe | Gln | Ser | Asp | Glu | Gly | Pro | Glu | Pro | Glu | |
| 305 | | | | | 310 | | | | | 315 | | | | | | 320 |
| Ala | Glu | Glu | Ser | Asp | Leu | Asp | Ser | Gly | Ser | Val | His | Ser | Ala | Ser | Gly | |
| | | | | 325 | | | | | 330 | | | | | 335 | | |
| Arg | Pro | Asp | Gly | Pro | Val | Arg | Thr | Lys | Lys | Leu | Lys | Arg | Gly | Arg | Pro | |
| | | | 340 | | | | | 345 | | | | | 350 | | | |
| Gly | Arg | Lys | Lys | Lys | Lys | Val | Leu | Gly | Cys | Pro | Ala | Val | Ala | Gly | Glu | |
| | | 355 | | | | | 360 | | | | | 365 | | | | |
| Glu | Glu | Val | Asp | Gly | Tyr | Glu | Thr | Asp | His | Gln | Asp | Tyr | Cys | Glu | Val | |
| | | 370 | | | | 375 | | | | | 380 | | | | | |
| Cys | Gln | Gln | Gly | Gly | Glu | Ile | Ile | Leu | Cys | Asp | Thr | Cys | Pro | Arg | Ala | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | |
| Tyr | His | Leu | Val | Cys | Leu | Asp | Pro | Glu | Leu | Asp | Arg | Ala | Pro | Glu | Gly | |
| | | | | 405 | | | | | 410 | | | | | 415 | | |
| Lys | Trp | Ser | Cys | Pro | His | Cys | Glu | Lys | Glu | Gly | Val | Gln | Trp | Glu | Ala | |
| | | | 420 | | | | | 425 | | | | | 430 | | | |
| Lys | Glu | Glu | Glu | Glu | Glu | Tyr | Glu | Glu | Glu | Gly | Glu | Glu | Glu | Gly | Glu | |
| | | 435 | | | | | 440 | | | | | 445 | | | | |
| Lys | Glu | Glu | Glu | Asp | Asp | His | Met | Glu | Tyr | Cys | Arg | Val | Cys | Lys | Asp | |
| | 450 | | | | | 455 | | | | | 460 | | | | | |
| Gly | Gly | Glu | Leu | Leu | Cys | Cys | Asp | Ala | Cys | Ile | Ser | Ser | Tyr | His | Ile | |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 | |
| His | Cys | Leu | Asn | Pro | Pro | Leu | Pro | Asp | Ile | Pro | Asn | Gly | Glu | Trp | Leu | |
| | | | | 485 | | | | | 490 | | | | | 495 | | |
| Cys | Pro | Arg | Cys | Thr | Cys | Pro | Val | Leu | Lys | Gly | Arg | Val | Gln | Lys | Ile | |
| | | | 500 | | | | | 505 | | | | | 510 | | | |
| Leu | His | Trp | Arg | Trp | Gly | Glu | Pro | Pro | Val | Ala | Val | Pro | Ala | Pro | Gln | |
| | | 515 | | | | | 520 | | | | | 525 | | | | |
| Gln | Ala | Asp | Gly | Asn | Pro | Asp | Val | Pro | Pro | Pro | Arg | Pro | Leu | Gln | Gly | |
| | 530 | | | | | 535 | | | | | 540 | | | | | |
| Arg | Ser | Glu | Arg | Glu | Phe | Phe | Val | Lys | Trp | Val | Gly | Leu | Ser | Tyr | Trp | |
| 545 | | | | | 550 | | | | | 555 | | | | | 560 | |
| His | Cys | Ser | Trp | Ala | Lys | Glu | Leu | Gln | Leu | Glu | Ile | Phe | His | Leu | Val | |
| | | | | 565 | | | | | 570 | | | | | 575 | | |

ES 2 511 694 T3

Met Tyr Arg Asn Tyr Gln Arg Lys Asn Asp Met Asp Glu Pro Pro Pro
580 585 590

Leu Asp Tyr Gly Ser Gly Glu Asp Asp Gly Lys Ser Asp Lys Arg Lys
595 600 605

Val Lys Asp Pro His Tyr Ala Glu Met Glu Glu Lys Tyr Tyr Arg Phe
610 615 620

Gly Ile Lys Pro Glu Trp Met Thr Val His Arg Ile Ile Asn His Ser
625 630 635 640

Val Asp Lys Lys Gly Asn Tyr His Tyr Leu Val Lys Trp Arg Asp Leu
645 650 655

Pro Tyr Asp Gln Ser Thr Trp Glu Glu Asp Glu Met Asn Ile Pro Glu
660 665 670

Tyr Glu Glu His Lys Gln Ser Tyr Trp Arg His Arg Glu Leu Ile Met
675 680 685

Gly Glu Asp Pro Ala Gln Pro Arg Lys Tyr Lys Lys Lys Lys Glu
690 695 700

Leu Gln Gly Asp Gly Pro Pro Ser Ser Pro Thr Asn Asp Pro Thr Val
705 710 715 720

Lys Tyr Glu Thr Gln Pro Arg Phe Ile Thr Ala Thr Gly Gly Thr Leu
725 730 735

His Met Tyr Gln Leu Glu Gly Leu Asn Trp Leu Arg Phe Ser Trp Ala
740 745 750

Gln Gly Thr Asp Thr Ile Leu Ala Asp Glu Met Gly Leu Gly Lys Thr
755 760 765

Ile Gln Thr Ile Val Phe Leu Tyr Ser Leu Tyr Lys Glu Gly His Thr
770 775 780

Lys Gly Pro Phe Leu Val Ser Ala Pro Leu Ser Thr Ile Ile Asn Trp
785 790 795 800

Glu Arg Glu Phe Gln Met Trp Ala Pro Lys Phe Tyr Val Val Thr Tyr
805 810 815

Thr Gly Asp Lys Asp Ser Arg Ala Ile Ile Arg Glu Asn Glu Phe Ser
820 825 830

Phe Glu Asp Asn Ala Ile Lys Gly Gly Lys Lys Ala Phe Lys Met Lys
835 840 845

Arg Glu Ala Gln Val Lys Phe His Val Leu Leu Thr Ser Tyr Glu Leu
850 855 860

Ile Thr Ile Asp Gln Ala Ala Leu Gly Ser Ile Arg Trp Ala Cys Leu
865 870 875 880

Val Val Asp Glu Ala His Arg Leu Lys Asn Asn Gln Ser Lys Phe Phe
885 890 895

Arg Val Leu Asn Gly Tyr Lys Ile Asp His Lys Leu Leu Leu Thr Gly
900 905 910

Thr Pro Leu Gln Asn Asn Leu Glu Glu Leu Phe His Leu Leu Asn Phe
915 920 925

ES 2 511 694 T3

Leu Thr Pro Glu Arg Phe Asn Asn Leu Glu Gly Phe Leu Glu Glu Phe
 930 935 940
 Ala Asp Ile Ser Lys Glu Asp Gln Ile Lys Lys Leu His Asp Leu Leu
 945 950 955 960
 Gly Pro His Met Leu Arg Arg Leu Lys Ala Asp Val Phe Lys Asn Met
 965 970 975
 Pro Ala Lys Thr Glu Leu Ile Val Arg Val Glu Leu Ser Pro Met Gln
 980 985 990
 Lys Lys Tyr Tyr Lys Tyr Ile Leu Thr Arg Asn Phe Glu Ala Leu Asn
 995 1000 1005
 Ser Arg Gly Gly Gly Asn Gln Val Ser Leu Leu Asn Ile Met Met
 1010 1015 1020
 Asp Leu Lys Lys Cys Cys Asn His Pro Tyr Leu Phe Pro Val Ala
 1025 1030 1035
 Ala Met Glu Ser Pro Lys Leu Pro Ser Gly Ala Tyr Glu Gly Gly
 1040 1045 1050
 Ala Leu Ile Lys Ser Ser Gly Lys Leu Met Leu Leu Gln Lys Met
 1055 1060 1065
 Leu Arg Lys Leu Lys Glu Gln Gly His Arg Val Leu Ile Phe Ser
 1070 1075 1080
 Gln Met Thr Lys Met Leu Asp Leu Leu Glu Asp Phe Leu Asp Tyr
 1085 1090 1095
 Glu Gly Tyr Lys Tyr Glu Arg Ile Asp Gly Gly Ile Thr Gly Ala
 1100 1105 1110
 Leu Arg Gln Glu Ala Ile Asp Arg Phe Asn Ala Pro Gly Ala Gln
 1115 1120 1125
 Gln Phe Cys Phe Leu Leu Ser Thr Arg Ala Gly Gly Leu Gly Ile
 1130 1135 1140
 Asn Leu Ala Thr Ala Asp Thr Val Ile Ile Phe Asp Ser Asp Trp
 1145 1150 1155
 Asn Pro His Asn Asp Ile Gln Ala Phe Ser Arg Ala His Arg Ile
 1160 1165 1170
 Gly Gln Ala Asn Lys Val Met Ile Tyr Arg Phe Val Thr Arg Ala
 1175 1180 1185
 Ser Val Glu Glu Arg Ile Thr Gln Val Ala Lys Arg Lys Met Met
 1190 1195 1200
 Leu Thr His Leu Val Val Arg Pro Gly Leu Gly Ser Lys Ala Gly
 1205 1210 1215
 Ser Met Ser Lys Gln Glu Leu Asp Asp Ile Leu Lys Phe Gly Thr
 1220 1225 1230
 Glu Glu Leu Phe Lys Asp Glu Asn Glu Gly Glu Asn Lys Glu Glu
 1235 1240 1245
 Asp Ser Ser Val Ile His Tyr Asp Asn Glu Ala Ile Ala Arg Leu
 1250 1255 1260
 Leu Asp Arg Asn Gln Asp Ala Thr Glu Asp Thr Asp Val Gln Asn

ES 2 511 694 T3

| | | | | | | | | | | | | | | | | |
|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|--|------|
| 1265 | | | | | | 1270 | | | | | | | | | | 1275 |
| Met | Asn | Glu | Tyr | Leu | Ser | Ser | Phe | Lys | Val | Ala | Gln | Tyr | Val | Val | | |
| 1280 | | | | | | 1285 | | | | | 1290 | | | | | |
| Arg | Glu | Glu | Asp | Lys | Ile | Glu | Glu | Ile | Glu | Arg | Glu | Ile | Ile | Lys | | |
| 1295 | | | | | | 1300 | | | | | 1305 | | | | | |
| Gln | Glu | Glu | Asn | Val | Asp | Pro | Asp | Tyr | Trp | Glu | Lys | Leu | Leu | Arg | | |
| 1310 | | | | | | 1315 | | | | | 1320 | | | | | |
| His | His | Tyr | Glu | Gln | Gln | Gln | Glu | Asp | Leu | Ala | Arg | Asn | Leu | Gly | | |
| 1325 | | | | | | 1330 | | | | | 1335 | | | | | |
| Lys | Gly | Lys | Arg | Val | Arg | Lys | Gln | Val | Asn | Tyr | Asn | Asp | Ala | Ala | | |
| 1340 | | | | | | 1345 | | | | | 1350 | | | | | |
| Gln | Glu | Asp | Gln | Asp | Asn | Gln | Ser | Glu | Tyr | Ser | Val | Gly | Ser | Glu | | |
| 1355 | | | | | | 1360 | | | | | 1365 | | | | | |
| Glu | Glu | Asp | Glu | Asp | Phe | Asp | Glu | Arg | Pro | Glu | Gly | Arg | Arg | Gln | | |
| 1370 | | | | | | 1375 | | | | | 1380 | | | | | |
| Ser | Lys | Arg | Gln | Leu | Arg | Asn | Glu | Lys | Asp | Lys | Pro | Leu | Pro | Pro | | |
| 1385 | | | | | | 1390 | | | | | 1395 | | | | | |
| Leu | Leu | Ala | Arg | Val | Gly | Gly | Asn | Ile | Glu | Val | Leu | Gly | Phe | Asn | | |
| 1400 | | | | | | 1405 | | | | | 1410 | | | | | |
| Thr | Arg | Gln | Arg | Lys | Ala | Phe | Leu | Asn | Ala | Val | Met | Arg | Trp | Gly | | |
| 1415 | | | | | | 1420 | | | | | 1425 | | | | | |
| Met | Pro | Pro | Gln | Asp | Ala | Phe | Thr | Thr | Gln | Trp | Leu | Val | Arg | Asp | | |
| 1430 | | | | | | 1435 | | | | | 1440 | | | | | |
| Leu | Arg | Gly | Lys | Thr | Glu | Lys | Glu | Phe | Lys | Ala | Tyr | Val | Ser | Leu | | |
| 1445 | | | | | | 1450 | | | | | 1455 | | | | | |
| Phe | Met | Arg | His | Leu | Cys | Glu | Pro | Gly | Ala | Asp | Gly | Ser | Glu | Thr | | |
| 1460 | | | | | | 1465 | | | | | 1470 | | | | | |
| Phe | Ala | Asp | Gly | Val | Pro | Arg | Glu | Gly | Leu | Ser | Arg | Gln | Gln | Val | | |
| 1475 | | | | | | 1480 | | | | | 1485 | | | | | |
| Leu | Thr | Arg | Ile | Gly | Val | Met | Ser | Leu | Val | Lys | Lys | Lys | Val | Gln | | |
| 1490 | | | | | | 1495 | | | | | 1500 | | | | | |
| Glu | Phe | Glu | His | Ile | Asn | Gly | Arg | Trp | Ser | Met | Pro | Glu | Leu | Met | | |
| 1505 | | | | | | 1510 | | | | | 1515 | | | | | |
| Pro | Asp | Pro | Ser | Ala | Asp | Ser | Lys | Arg | Ser | Ser | Arg | Ala | Ser | Ser | | |
| 1520 | | | | | | 1525 | | | | | 1530 | | | | | |
| Pro | Thr | Lys | Thr | Ser | Pro | Thr | Thr | Pro | Glu | Ala | Ser | Ala | Thr | Asn | | |
| 1535 | | | | | | 1540 | | | | | 1545 | | | | | |
| Ser | Pro | Cys | Thr | Ser | Lys | Pro | Ala | Thr | Pro | Ala | Pro | Ser | Glu | Lys | | |
| 1550 | | | | | | 1555 | | | | | 1560 | | | | | |
| Gly | Glu | Gly | Ile | Arg | Thr | Pro | Leu | Glu | Lys | Glu | Glu | Ala | Glu | Asn | | |
| 1565 | | | | | | 1570 | | | | | 1575 | | | | | |
| Gln | Glu | Glu | Lys | Pro | Glu | Lys | Asn | Ser | Arg | Ile | Gly | Glu | Lys | Met | | |
| 1580 | | | | | | 1585 | | | | | 1590 | | | | | |
| Glu | Thr | Glu | Ala | Asp | Ala | Pro | Ser | Pro | Ala | Pro | Ser | Leu | Gly | Glu | | |
| 1595 | | | | | | 1600 | | | | | 1605 | | | | | |

ES 2 511 694 T3

Arg Leu Glu Pro Arg Lys Ile Pro Leu Glu Asp Glu Val Pro Gly
 1610 1615 1620
 Val Pro Gly Glu Met Glu Pro Glu Pro Gly Tyr Arg Gly Asp Arg
 1625 1630 1635
 Glu Lys Ser Ala Thr Glu Ser Thr Pro Gly Glu Arg Gly Glu Glu
 1640 1645 1650
 Lys Pro Leu Asp Gly Gln Glu His Arg Glu Arg Pro Glu Gly Glu
 1655 1660 1665
 Thr Gly Asp Leu Gly Lys Arg Glu Asp Val Lys Gly Asp Arg Glu
 1670 1675 1680
 Leu Arg Pro Gly Pro Arg Asp Glu Pro Arg Ser Asn Gly Arg Arg
 1685 1690 1695
 Glu Glu Lys Thr Glu Asn Pro Arg Phe Met Phe Asn Ile Ala Asp
 1700 1705 1710
 Gly Gly Phe Thr Glu Leu His Thr Leu Trp Gln Asn Glu Glu Arg
 1715 1720 1725
 Ala Ala Ile Ser Ser Gly Lys Leu Asn Glu Ile Trp His Arg Arg
 1730 1735 1740
 His Asp Tyr Trp Leu Leu Ala Gly Ile Val Leu His Gly Tyr Ala
 1745 1750 1755
 Arg Trp Gln Asp Ile Gln Asn Asp Ala Gln Phe Ala Ile Ile Asn
 1760 1765 1770
 Glu Pro Phe Lys Thr Glu Ala Asn Lys Gly Asn Phe Leu Glu Met
 1775 1780 1785
 Lys Asn Lys Phe Leu Ala Arg Arg Phe Lys Leu Leu Glu Gln Ala
 1790 1795 1800
 Leu Val Ile Glu Glu Gln Leu Arg Arg Ala Ala Tyr Leu Asn Leu
 1805 1810 1815
 Ser Gln Glu Pro Ala His Pro Ala Met Ala Leu His Ala Arg Phe
 1820 1825 1830
 Ala Glu Ala Glu Cys Leu Ala Glu Ser His Gln His Leu Ser Lys
 1835 1840 1845
 Glu Ser Leu Ala Gly Asn Lys Pro Ala Asn Ala Val Leu His Lys
 1850 1855 1860
 Val Leu Asn Gln Leu Glu Glu Leu Leu Ser Asp Met Lys Ala Asp
 1865 1870 1875
 Val Thr Arg Leu Pro Ala Thr Leu Ser Arg Ile Pro Pro Ile Ala
 1880 1885 1890
 Ala Arg Leu Gln Met Ser Glu Arg Ser Ile Leu Ser Arg Leu Ala
 1895 1900 1905
 Ser Lys Gly Thr Glu Pro His Pro Thr Pro Ala Tyr Pro Pro Gly
 1910 1915 1920
 Pro Tyr Ala Thr Pro Pro Gly Tyr Gly Ala Ala Phe Ser Ala Ala
 1925 1930 1935

ES 2 511 694 T3

Pro Val Gly Ala Leu Ala Ala Ala Gly Ala Asn Tyr Ser Gln Met
1940 1945 1950

Pro Ala Gly Ser Phe Ile Thr Ala Ala Thr Asn Gly Pro Pro Val
1955 1960 1965

Leu Val Lys Lys Glu Lys Glu Met Val Gly Ala Leu Val Ser Asp
1970 1975 1980

Gly Leu Asp Arg Lys Glu Pro Arg Ala Gly Glu Val Ile Cys Ile
1985 1990 1995

Asp Asp
2000

<210> 9

<211> 913

<212> PRT

5 <213> Homo sapiens

<220>

<221> característica_miscelánea

<223> CDS de la secuencia GS-N4:67..2808

<220>

10 <221> característica_miscelánea

<223> secuencia llamada GS-P4: factor de iniciación de la traducción eucariota, subunidad 8 (110 kD) (EIF3S8).

<400> 9

ES 2 511 694 T3

Met Ser Arg Phe Phe Thr Thr Gly Ser Asp Ser Glu Ser Glu Ser Ser
 1 5 10 15

Leu Ser Gly Glu Glu Leu Val Thr Lys Pro Val Gly Gly Asn Tyr Gly
 20 25 30

Lys Gln Pro Leu Leu Leu Ser Glu Asp Glu Glu Asp Thr Lys Arg Val
 35 40 45

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

Arg Thr Ile Arg Asn Ala Met Lys Ile Arg Asp Val Thr Lys Cys Leu
 65 70 75 80

Glu Glu Phe Glu Leu Leu Gly Lys Ala Tyr Gly Lys Ala Lys Ser Ile
 85 90 95

Val Asp Lys Glu Gly Val Pro Arg Phe Tyr Ile Arg Ile Leu Ala Asp
 100 105 110

Leu Glu Asp Tyr Leu Asn Glu Leu Trp Glu Asp Lys Glu Gly Lys Lys
 115 120 125

Lys Met Asn Lys Asn Asn Ala Lys Ala Leu Ser Thr Leu Arg Gln Lys
 130 135 140

Ile Arg Lys Tyr Asn Arg Asp Phe Glu Ser His Ile Thr Ser Tyr Lys
 145 150 155 160

Gln Asn Pro Glu Gln Ser Ala Asp Glu Asp Ala Glu Lys Asn Glu Glu
 165 170 175

Asp Ser Glu Gly Ser Ser Asp Glu Asp Glu Asp Glu Asp Gly Val Ser
 180 185 190

Ala Ala Thr Phe Leu Lys Lys Lys Ser Glu Ala Pro Ser Gly Glu Ser
 195 200 205

ES 2 511 694 T3

Arg Lys Phe Leu Lys Lys Met Asp Asp Glu Asp Glu Asp Ser Glu Asp
 210 215 220
 Ser Glu Asp Asp Glu Asp Trp Asp Thr Gly Ser Thr Ser Ser Asp Ser
 225 230 235 240
 Asp Ser Glu Glu Glu Glu Gly Lys Gln Thr Ala Leu Ala Ser Arg Phe
 245 250 255
 Leu Lys Lys Ala Pro Thr Thr Asp Glu Asp Lys Lys Ala Ala Glu Lys
 260 265 270
 Lys Arg Glu Asp Lys Ala Lys Lys Lys His Asp Arg Lys Ser Lys Arg
 275 280 285
 Leu Asp Glu Glu Glu Glu Asp Asn Glu Gly Gly Glu Trp Glu Arg Val
 290 295 300
 Arg Gly Gly Val Pro Leu Val Lys Glu Lys Pro Lys Met Phe Ala Lys
 305 310 315 320
 Gly Thr Glu Ile Thr His Ala Val Val Ile Lys Lys Leu Asn Glu Ile
 325 330 335
 Leu Gln Ala Arg Gly Lys Lys Gly Thr Asp Arg Ala Ala Gln Ile Glu
 340 345 350
 Leu Leu Gln Leu Leu Val Gln Ile Ala Ala Glu Asn Asn Leu Gly Glu
 355 360 365
 Gly Val Ile Val Lys Ile Lys Phe Asn Ile Ile Ala Ser Leu Tyr Asp
 370 375 380
 Tyr Asn Pro Asn Leu Ala Thr Tyr Met Lys Pro Glu Met Trp Gly Lys
 385 390 395 400
 Cys Leu Asp Cys Ile Asn Glu Leu Met Asp Ile Leu Phe Ala Asn Pro
 405 410 415
 Asn Ile Phe Val Gly Glu Asn Ile Leu Glu Glu Ser Glu Asn Leu His
 420 425 430
 Asn Ala Asp Gln Pro Leu Arg Val Arg Gly Cys Ile Leu Thr Leu Val
 435 440 445
 Glu Arg Met Asp Glu Glu Phe Thr Lys Ile Met Gln Asn Thr Asp Pro
 450 455 460
 His Ser Gln Glu Tyr Val Glu His Leu Lys Asp Glu Ala Gln Val Cys
 465 470 475 480
 Ala Ile Ile Glu Arg Val Gln Arg Tyr Leu Glu Glu Lys Gly Thr Thr
 485 490 495
 Glu Glu Val Cys Arg Ile Tyr Leu Leu Arg Ile Leu His Thr Tyr Tyr
 500 505 510
 Lys Phe Asp Tyr Lys Ala His Gln Arg Gln Leu Thr Pro Pro Glu Gly
 515 520 525
 Ser Ser Lys Ser Glu Gln Asp Gln Ala Glu Asn Glu Gly Glu Asp Ser
 530 535 540
 Ala Val Leu Met Glu Arg Leu Cys Lys Tyr Ile Tyr Ala Lys Asp Arg
 545 550 555 560

ES 2 511 694 T3

Thr Asp Arg Ile Arg Thr Cys Ala Ile Leu Cys His Ile Tyr His His
565 570 575

Ala Leu His Ser Arg Trp Tyr Gln Ala Arg Asp Leu Met Leu Met Ser
580 585 590

His Leu Gln Asp Asn Ile Gln His Ala Asp Pro Pro Val Gln Ile Leu
595 600 605

Tyr Asn Arg Thr Met Val Gln Leu Gly Ile Cys Ala Phe Arg Gln Gly
610 615 620

Leu Thr Lys Asp Ala His Asn Ala Leu Leu Asp Ile Gln Ser Ser Gly
625 630 635 640

Arg Ala Lys Glu Leu Leu Gly Gln Gly Leu Leu Leu Arg Ser Leu Gln
645 650 655

Glu Arg Asn Gln Glu Gln Glu Lys Val Glu Arg Arg Arg Gln Val Pro
660 665 670

Phe His Leu His Ile Asn Leu Glu Leu Leu Glu Cys Val Tyr Leu Val
675 680 685

Ser Ala Met Leu Leu Glu Ile Pro Tyr Met Ala Ala His Glu Ser Asp
690 695 700

Ala Arg Arg Arg Met Ile Ser Lys Gln Phe His His Gln Leu Arg Val
705 710 715 720

Gly Glu Arg Gln Pro Leu Leu Gly Pro Pro Glu Ser Met Arg Glu His
725 730 735

Val Val Ala Ala Ser Lys Ala Met Lys Met Gly Asp Trp Lys Thr Cys
740 745 750

His Ser Phe Ile Ile Asn Glu Lys Met Asn Gly Lys Val Trp Asp Leu
755 760 765

Phe Pro Glu Ala Asp Lys Val Arg Thr Met Leu Val Arg Lys Ile Gln
770 775 780

Glu Glu Ser Leu Arg Thr Tyr Leu Phe Thr Tyr Ser Ser Val Tyr Asp
785 790 795 800

Ser Ile Ser Met Glu Thr Leu Ser Asp Met Phe Glu Leu Asp Leu Pro
805 810 815

Thr Val His Ser Ile Ile Ser Lys Met Ile Ile Asn Glu Glu Leu Met
820 825 830

Ala Ser Leu Asp Gln Pro Thr Gln Thr Val Val Met His Arg Thr Glu
835 840 845

Pro Thr Ala Gln Gln Asn Leu Ala Leu Gln Leu Ala Glu Lys Leu Gly
850 855 860

Ser Leu Val Glu Asn Asn Glu Arg Val Phe Asp His Lys Gln Gly Thr
865 870 875 880

Tyr Gly Gly Tyr Phe Arg Asp Gln Lys Asp Gly Tyr Arg Lys Asn Glu
885 890 895

Gly Tyr Met Arg Arg Gly Gly Tyr Arg Gln Gln Gln Ser Gln Thr Ala
900 905 910

Tyr

ES 2 511 694 T3

<210> 10
 <211> 325
 <212> PRT
 <213> Homo sapiens

5 <220>
 <221> característica_miscelánea
 <223> CDS de la secuencia GS-N5:407..1384

<220>
 <221> característica_miscelánea
 10 <223> secuencia llamada GS-P5: proteína similar al factor de iniciación de la traducción eucariota, subunidad 8 (110 kD).

<400> 10

```

Met Leu Met Ser His Leu Gln Asp Asn Ile Gln His Ala Asp Pro Pro
1                               5                               10 15

Val Gln Ile Leu Tyr Asn Arg Thr Met Val Gln Leu Gly Ile Cys Ala
                20                               25                               30

Phe Arg Gln Gly Leu Thr Lys Asp Ala His Asn Ala Leu Leu Asp Ile
          35                               40                               45

Gln Ser Ser Gly Arg Ala Lys Glu Leu Leu Gly Gln Gly Leu Leu Leu
          50                               55                               60

Arg Ser Leu Gln Glu Arg Asn Gln Glu Gln Glu Lys Val Glu Arg Arg
65                               70                               75                               80

Arg Gln Val Pro Phe His Leu His Ile Asn Leu Glu Leu Leu Glu Cys
          85                               90                               95

Val Tyr Leu Val Ser Ala Met Leu Leu Glu Ile Pro Tyr Met Ala Ala
          100                              105                              110

His Glu Ser Asp Ala Arg Arg Arg Met Ile Ser Lys Gln Phe His His
          115                              120                              125

Gln Leu Arg Val Gly Glu Arg Gln Pro Leu Leu Gly Pro Pro Glu Ser
          130                              135                              140

Met Arg Glu His Val Val Ala Ala Ser Lys Ala Met Lys Met Gly Asp
145                              150                              155                              160

Trp Lys Thr Cys His Ser Phe Ile Ile Asn Glu Lys Met Asn Gly Lys
          165                              170                              175

Val Trp Asp Leu Phe Pro Glu Ala Asp Lys Val Arg Thr Met Leu Val
          180                              185                              190

Arg Lys Ile Gln Glu Glu Ser Leu Arg Thr Tyr Leu Phe Thr Tyr Ser
          195                              200                              205

Ser Val Tyr Asp Ser Ile Ser Met Glu Thr Leu Ser Asp Met Phe Glu
210                              215                              220

Leu Asp Leu Pro Thr Val His Ser Ile Ile Ser Lys Met Ile Ile Asn
225                              230                              235                              240

Glu Glu Leu Met Ala Ser Leu Asp Gln Pro Thr Gln Thr Val Val Met
          245                              250                              255

His Arg Thr Glu Pro Thr Ala Gln Gln Asn Leu Ala Leu Gln Leu Ala
          260                              265                              270
  
```

ES 2 511 694 T3

Glu Lys Leu Gly Ser Leu Val Glu Asn Asn Glu Arg Val Phe Asp His
275 280 285

Lys Gln Gly Thr Tyr Gly Gly Tyr Phe Arg Asp Gln Lys Asp Gly Tyr
290 295 300

Arg Lys Asn Glu Gly Tyr Met Arg Arg Gly Gly Tyr Arg Gln Gln Gln
305 310 315 320

Ser Gln Thr Ala Tyr
325

<210> 11
<211> 392
<212> ADN
5 <213> Homo sapiens

<220>
<221> característica_misclánea
<223> secuencia nucleica antisentido de GS-N1:4369... 3978

<400> 11
ttcccctttg gaagagagag cagtgttaga tccttcaaag tcaagataaa agcagaggat 60
tgcaattagg gcatgagtca ccagcgagcc tcagaaaaacc ctactgggat gcttccgaaa 120
ggccagtgtc gtgcaggggg cgagacactt cagggagcgg gaggctccaa ggacagagct 180
tttaaagcag cagcagaagc ctgggatgtg gaagatctct gtctactctc ctattcctga 240
tgggagcagt cgggcaaagc acaccgctct tgactctgct ctacaggcca ttcacataac 300
agcaaattgc tgggtttcct tcatgtccct attaggaaaag caaaacatac cacaaagtat 360
10 accctttctc actaaccctc tgcagtggac ct 392

<210> 12
<211> 250
<212> ADN
<213> Homo sapiens

15 <220>
<221> característica_misclánea
<223> secuencia nucleica antisentido de GS-N2: 2859...2619

<400> 12
ttcagaggag gggagagaaa gagtactcag tagggactga gcactaaatg cttatitaa 60
aagaaatgta aagagcagaa agcaattcag gctaccctgc cttttgtgct ggctagcact 120
ccggtcggtg tcagcagcac gtggcattga acattgcaat gtggagccca aaccacagaa 180
aatggggtga aattggccaa ctttctatta acttatgttg gcaatitgct caccaacagt 240
aagctggccc 250

20 <210> 13
<211> 359
<212> ADN
<213> Homo sapiens

25 <220>
<221> característica_misclánea
<223> secuencia nucleica antisentido de GS-N3: 3980...3622

<400> 13
tgcacgtcag tgtcctcagt tgcacctcgg ttccggtcgg acagccgagc gatggcctca 60
ttgtcataat gaatcacact gctgtcctcc tccttgttct ccccctcgtt ttcaccttg 120
aatagctctt cagtgccaaa tttgagaatg tcgtcaagct cctgcttggga catggagcct 180
gccttggagc ccagcccagg ccgcacaacc aggtgtgtca gcatcatctt tctcttggcc 240
acttgtgtga ttgcctcttc cactgacgcg cgagtcacaa accggtaaatt catcactttg 300
ttggcctggc caatccgatg agcccggcta aaggcctgga tgtcattatg ggggttcca 359

30 <210> 14
<211> 167

ES 2 511 694 T3

<212> ADN
<213> Homo sapiens

<220>
<221> característica_misclánea
5 <223> secuencia nucleica antisentido de GS-N4: 2832...2666

<220>
<221> característica_misclánea
<223> secuencia nucleica antisentido de GS-N5:1408...1242

<400> 14

caggcgggaa acagagtgga gagctcagta ggccgtctga gactgctgct ggcggtagcc 60
10 accgcggcgc atgtagccct cgtttttgcg gtagccgtcc ttctggtctc ggaagtagcc 120
cccgtagggtg cctgcttgt ggtcaaacac ccgttcggtg ttctcca 167

<210> 15
<211> 5852
<212> ADN
<213> Homo sapiens

15 <220>
<221> característica_misclánea
<223> vector de expresión llamado GS-V1 que contiene la secuencia antisentido GS-N1; antisentido:1108...1499

<400> 15

ES 2 511 694 T3

| | | | | | | |
|------------|-------------|-------------|-------------|------------|------------|------|
| tcaatattgg | ccattagcca | tattattcat | tggttatata | gcataaatca | atattggcta | 60 |
| ttggccattg | catacgttgt | atctatatca | taatatgtac | atztatattg | gctcatgtcc | 120 |
| aatatgaccg | ccatgttggc | attgattatt | gactagttat | taatagtaat | caattacggg | 180 |
| gtcattagtt | catagcccat | atatggagtt | ccgcgttaca | taacttacgg | taaatggccc | 240 |
| gcctggctga | ccgccaacg | acccccgcc | attgacgtca | ataatgacgt | atgttcccat | 300 |
| agtaacgcca | atagggactt | tccattgacg | tcaatgggtg | gagtatttac | ggtaaacctg | 360 |
| ccacttggca | gtacatcaag | tgtatcatat | gccaagtccg | ccccctattg | acgtcaatga | 420 |
| cggtaaatgg | cccgcctggc | attatgcccc | gtacatgacc | ttacgggact | ttcctacttg | 480 |
| gcagtacatc | tacgtattag | tcatcgctat | taccatgggtg | atgcggtttt | ggcagtacac | 540 |
| caatgggctg | ggatagcggg | ttgactcacg | gggatttcca | agtctccacc | ccattgacgt | 600 |
| caatgggagt | ttgttttggc | acaaaaatca | acgggacttt | ccaaaatgtc | gtaacaactg | 660 |
| cgatcgcccc | cccogttgac | gcaaatgggg | ggtaggcgtg | tacggtggga | ggtctatata | 720 |
| agcagagctc | gtttagttaa | ccgtcagatc | actagaagct | ttattgcggt | agtttatcac | 780 |
| agttaaattg | ctaacgcagt | cagtgcctct | gacacaacag | tctcgaactt | aagctgcagt | 840 |
| gactctctta | aggtagcctt | gcagaagttg | gtcgtgaggc | actgggcagg | taagtatcaa | 900 |
| ggttacaaga | caggtttaag | gagaccaata | gaaactgggc | ttgtcagagc | agagaagact | 960 |
| cttgcgtttc | tgataggcac | ctattggctt | tactgacatc | cactttgcct | ttctctccac | 1020 |
| aggtgtccac | tcccagttca | attacagctc | ttaaggctag | agtacttaat | acgactcact | 1080 |
| ataggctagc | ctcgagaatt | cacgcgtttc | ccctttggaa | gagagagcag | tgttagatcc | 1140 |
| ttcaaagtca | agataaaagc | agaggattgc | aattagggca | tgagtcacca | gcgagcctca | 1200 |
| gaaaacccta | ctgggatgct | tccgaaaggc | cagtgctgtg | cagggggcga | gacacttcag | 1260 |
| ggagcgggag | gctccaagga | cagagctttt | aaagcagcag | cagaagcctg | ggatgtggaa | 1320 |
| gatctctgtc | tactctccta | ttcctgatgg | gagcagtcgg | gcaaagcaca | ccgctcttga | 1380 |
| ctctgctcta | aacgccattc | acataacagc | aaattgctgg | gtttccttca | tgtccctatt | 1440 |
| aggaaagcaa | aacataccac | aaagtatacc | ctttctcact | aaccctctgc | agtggacctg | 1500 |
| tcgaccgggg | cggccgcttc | ccttttagtga | gggttaatgc | ttcgagcaga | catgataaga | 1560 |
| tacattgatg | agtttggaca | aaccacaact | agaatgcagt | gaaaaaaatg | ctttatattg | 1620 |
| gaaatttgtg | atgctattgc | tttatttcta | accattataa | gctgcaataa | acaagttaac | 1680 |
| aacaacaatt | gcattcattt | tatgtttcag | gttcaggggg | agatgtggga | ggttttttaa | 1740 |
| agcaagtaaa | acctctacaa | atgtggtaaa | atccgataag | gatcgatccg | ggctggcgtg | 1800 |
| atagcgaaga | ggcccgcacc | gatcgcctct | cccaacagtt | gcgagcctg | aatggcgaat | 1860 |
| ggacgcgccc | tgtagcggcg | cattaagcgc | ggcgggtgtg | gtggttacgc | gcagcgtgac | 1920 |
| cgctacactt | gccagegccc | tagcgcgccg | tcctttcgct | ttcttccctt | cctttctcgc | 1980 |
| cacgttcgcc | ggctttcccc | gtcaagctct | aaatoggggg | ctccctttag | ggttccgatt | 2040 |
| tagtgcctta | cggcacctcg | acccccaaaa | acttgattag | ggtgatggtt | cacgtagtgg | 2100 |
| gccatcgccc | tgatagacgg | tttttcgccc | tttgacgttg | gagtccacgt | tctttaatag | 2160 |
| tggactcttg | ttccaaaactg | gaacaacact | caaccctatc | tcggtctatt | cttttgattt | 2220 |
| ataagggatt | ttgccgattt | cggcctattg | gttaaaaaat | gagctgattt | aacaaaaatt | 2280 |
| taacgcgaat | tttaacaaaa | tattaacgct | tacaatttcc | tgatgcggta | ttttctcctt | 2340 |
| acgcactctg | gcggtatttc | acaccgcata | cgcggatctg | cgcagcacca | tggcctgaaa | 2400 |
| taacctctga | aagaggaact | tggttaggtg | ccttctgagg | cggaaagaac | cagctgtgga | 2460 |
| atgtgtgtca | gttaggggtg | ggaaagtccc | caggctcccc | agcaggcaga | agtatgcaaa | 2520 |
| gcatgcatct | caattagtca | gcaaccaggt | gtggaaagtc | cccaggctcc | ccagcaggca | 2580 |
| gaagtatgca | aagcatgcat | ctcaattagt | cagcaacat | agtcccggcc | ctaactccgc | 2640 |
| ccatcccgcc | cctaactccg | cccagttccg | ccattctcc | gccccatggc | tgactaattt | 2700 |
| tttttattta | tgcagaggcc | gaggccgctt | cggcctctga | gctattccag | aagtagtgag | 2760 |
| gaggcttttt | tggaggccta | ggcttttga | aaaagcttga | ttcttctgac | acaacagtct | 2820 |

ES 2 511 694 T3

cgaacttaag gctagagcca ccatgattga acaagatgga ttgcacgcag gttctccggc 2880
cgcttggggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga 2940
tgccgcogtg ttccggctgt cagcgcaggg ggcgccgggt ctttttgtca agaccgacct 3000
gtccgggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc tggccacgac 3060
ggggttctct tgcgcagctg tgctcgacgt tgtcactgaa gcggaaggg actggctgct 3120
attggcgcaa gtgccggggc aggatctcct gtcactcac ctgctcctg ccgagaaagt 3180
atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta cctgccatt 3240
cgaccaccaa gcgaaacatc gcacgcagcg agcacgtact cggatggaag ccggtcttgt 3300
cgatcaggat gatctggacg aagagcatca ggggctcggc ccagccgaac tgttcgccag 3360
gctcaaggcg cgcagtcggc acggcgagga tctcgtcgtg acccatggcg atgcctgctt 3420
gccgaatatc atgggtgaaa atggccgctt ttctggattc atcgactgtg gccggctggg 3480
tgtggcggac cgctatcagc acatagcgtt ggctaaccgt gatattgctg aagagcttgg 3540
cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg attcgcagcg 3600
catcgccttc tatcgccttc ttgacgagtt cttctgagcg ggactctggg gttcgaatg 3660
accgaccaag cgacgcccaa cctgccatca cgatggcggc aataaaatat ctttattttc 3720
attacatctg tgtgttgggt ttttgtgtga atcgatagcg ataaggatcc gcgatgggtg 3780
cactctcagt acaatctgct ctgatgccgc atagttaagc cagcccgcac acccgccaac 3840
accgctgac gcgccctgct gggcttgtct gctcccggca tccgcttaca gacaagctgt 3900
gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag 3960
acgaaagggc ctcgtgatcc gcctattttt ataggttaat gtcagataa taatggtttc 4020
ttagacgtca ggtggcactt ttccgggaaa tgtgcgcgga acccctattt gtttattttt 4080
ctaaatacat tcaaatatgt atccgctcat gagacaataa ccttgataaa tgcttcaata 4140
atattgaaaa aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt 4200
tgccggcattt tgccttctctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc 4260
tgaagatcag ttgggtgcac gagggtggtta catcgaactg gatctcaaca gcggtaagat 4320
ccttgagagt ttccgcccgg aagaacggtt tccaatgatg agcactttta aagttctgct 4380
atgtggcggc gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca 4440
ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg 4500
catgacagta agagaattat gcagtgtctgc cataaccatg agtgataaca ctgcggccaa 4560
cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg 4620
ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga 4680
cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg 4740
cgaactactt actctagctt cccggcaaca ataatagac tggatggagg cggataaagt 4800
tgcaggacca ctctcgcgct cggcccttcc ggctggctgg ttattgctg ataaatctgg 4860
agccgggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc 4920
ccgatcgtgta gttatctaca cgacggggag tcaggcaact atggatgaa gaaatagaca 4980
gatcgcctgag ataggtgcct cactgattaa gcatttgtaa ctgtcagacc aagtttactc 5040
atataactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagt 5100
cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttc actgagcgtc 5160
agaccocgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc cgtaatctg 5220
ctgcttgcaa acaaaaaaac caccgctacc agcgggtggt tgtttgcccg atcaagagct 5280
accaactctt tttccgaagg taactggctt cagcagagcg cagataccea atactgttct 5340
tctagtgtag ccgtagttag gccaccactt caagaactct gtgacccgc ctacatacct 5400
cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg 5460
gttggactca agacgatagt taccggataa ggcgagcgg tcgggctgaa cgggggggtc 5520
gtgcacacag ccagccttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga 5580
gctatgagaa agcggcacgc ttcccgaagg gagaaagggc gacaggatc cggtaagcgg 5640
cagggtcggg acaggagagc gcacgagggg gcttccaagg ggaacgcct ggtatcttta 5700
tagtctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg 5760
ggggcgggac ctatggaaa acgccagcaa cgcggccttt ttacggttcc tggccttttt 5820
ctggcctttt gctcacatg ctccacagat ct 5852

- <210> 16
- <211> 5710
- <212> ADN
- 5 <213> Homo sapiens

- <220>
- <221> característica_misclánea
- <222> (1).. (5710)
- <223> vector de expresión llamado GS-V2 que contiene la secuencia antisentido de GS-N2;
- 10 antisentido:2859...2619

<400> 16
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120
aatatgaccg ccatgttggc attgattatt gactagtatt taatagtaat caattacggg 180
gtcattagtt catagcccat atatggagtt ccgctgtaca taacttacgg taaatggccc 240

ES 2 511 694 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gcctggctga | ccgcccaacg | acccccgcc | attgacgtca | ataatgacgt | atgttcccat | 300 |
| agtaacgcc | atagggactt | tccattgacg | tcaatgggtg | gagtatattac | ggtaaacctgc | 360 |
| ccacttgcca | gtacatcaag | tgtatcatat | gccaagtccg | ccccctattg | acgtcaatga | 420 |
| cggtaaatgg | cccgcctggc | attatgccca | gtacatgacc | ttacgggact | ttcctacttg | 480 |
| gcagtacatc | tacgtattag | tcacgcctat | taccatgggtg | atgcggtttt | ggcagtacac | 540 |
| caatgggcgt | ggatagcggg | ttgactcacg | gggatttcca | agtctccaacc | ccattgacgt | 600 |
| caatggggagt | ttgttttggc | acaaaaatca | acgggacttt | ccaaaaatgtc | gtaacaactg | 660 |
| cgatcgcccc | ccccgttgac | gcaaatgggc | ggtaggcgtg | tacggtgagg | ggcttatata | 720 |
| agcagagctc | gttttagtgaa | ccgtcagatc | actagaagct | ttattgcggg | agtttatcac | 780 |
| agttaaattg | ctaacgcagt | cagtgttctc | gacacaacag | tctcgaactt | aagctgcagt | 840 |
| gactctctta | aggtagcctt | gcagaagttg | gtcgtgaggg | actgggcagg | taagtatcaa | 900 |
| ggttacaaga | caggtttaag | gagaccaata | gaaactgggc | ttgtcggagc | agagaagact | 960 |
| cttgcgtttc | tgataggcac | ctattgggtc | tactgacatc | cactttgcct | ttctctccac | 1020 |
| aggtgtccac | tcccagttca | attacagctc | ttaaggctag | agtacttaat | acgactcact | 1080 |
| ataggctagc | ctcggagaatt | cacgcgtttc | agaggagggg | agagaaagag | tactcagtag | 1140 |
| ggactgagca | ctaaatgctt | attttaaag | aaatgtaaag | agcagaaagc | aattcaggct | 1200 |
| accctgcctt | ttgtgctggc | tagcaactccg | gtcgtgtca | gcagcacgtg | gcattgaaca | 1260 |
| ttgcaatgtg | gagcccaaac | cacagaaaat | ggggtgaaat | tggccaactt | tctattaact | 1320 |
| tatgttggca | attttgccac | caacagtaag | ctggcccgtc | gacccggcgg | gcccgttccc | 1380 |
| tttagtgagg | gttaatgctt | cgagcagaca | tgataagata | cattgatgag | tttggacaaa | 1440 |
| ccacaactag | aatgcagtga | aaaaaatgct | ttatttgtga | aatttgtgat | gctattgctt | 1500 |
| tatttghtaac | cattataaagc | tgcaataaac | aagttaaaca | caacaattgc | attcatttta | 1560 |
| tgtttcagg | tcaggggggag | atgtggggag | ttttttaaag | caagtaaaac | ctctacaaat | 1620 |
| gtggtaaaat | ccgataagga | tcgatccggg | ctggcgtaat | agcgaagagg | cccgcaccga | 1680 |
| tcgcccttcc | caacagttgc | gcagcctgaa | tggcgaatgg | acgcgccctg | tagcggcgca | 1740 |
| ttaagcggcg | cgggtgtggt | ggttacgcgc | agcgtgaccg | ctacacttgc | cagcgcctca | 1800 |
| gcgcccgctc | ctttcgcttt | cttcccttcc | tttctcgcca | cgttcgccgg | ctttcccctg | 1860 |
| caagctctaa | atcgggggct | cccttaggg | ttccgattta | gtgctttacg | gcacctcgac | 1920 |
| cccaaaaaac | ttgattaggg | tgatggttca | cgtagtgggc | catcgccctg | atagacgggt | 1980 |
| tttccgctt | tgacgttggg | gtccacgttc | tttaatagtg | gactcttgtt | ccaactgga | 2040 |
| acaacactca | accctatctc | ggtctattct | tttgatttat | aagggatttt | gccgatttcg | 2100 |
| gctatgtgt | taaaaaatga | gctgatttaa | caaaaattta | acgcgaattt | taacaaaata | 2160 |
| ttaacgctta | caatttccgt | atgcgggtatt | ttctccttac | gcatctgtgc | ggtatttcac | 2220 |
| accgcatacg | cggatctgcg | cagcaccatg | gcctgaaata | acctctgaaa | gaggaacttg | 2280 |
| gtaggttacc | ttctgaggcg | gaaagaacca | gctgtggaat | gtgtgtcagt | tagggtgtgg | 2340 |
| aaagtcccca | ggctccccag | caggcagaag | tatgcaaagc | atgcatctca | attagtccgc | 2400 |
| aaccaggtgt | ggaaagtccc | caggctcccc | agcaggcaga | agtatgcaaa | gcatgcatct | 2460 |
| caattagtca | gcaaccatag | tcccgccctc | aactccgcc | atcccgcccc | taactccgcc | 2520 |
| cagttccgcc | catttctccgc | ccatggctg | actaatttt | tttatttatg | cagaggccga | 2580 |
| ggccgctcg | gcctctgagc | tattccagaa | gtagtgagga | ggcttttttg | gaggcctagg | 2640 |
| cttttgcaaa | aagcttgatt | cttctgacac | aacagtctcg | aacttaaggc | tagagccacc | 2700 |
| atgattgaac | aagatggatt | gcacgcagg | tctccggccg | cttgggtgga | gaggctattc | 2760 |
| ggctatgact | gggcacaaca | gacaatcggc | tgctctgatg | ccgcctgtgt | ccggctgtca | 2820 |
| gcgcaggggc | gcccggttct | ttttgtcaag | accgacctgt | ccgggtccct | gaatgaactg | 2880 |
| caggacgagg | cagcggcggc | atcgtggctg | gccacgacgg | gcgttccctg | cgcagctgtg | 2940 |
| ctcgcagttg | tcactgaagc | gggaagggac | tggctgctat | tgggcgaagt | gcccggggcag | 3000 |
| gatctcctgt | catctcacct | tgctcctgcc | gagaaagtat | ccatcatggc | tgatgcaatg | 3060 |
| cgccggctgc | atacgtctga | tccggctacc | tgcccattcg | accaccaagc | gaaacatcgc | 3120 |
| atcgaagcag | cacgtactcg | gatggaagcc | ggtcttgtcg | atcaggatga | tctggacgaa | 3180 |
| gagcatcagg | ggctcgcgcc | agccgaactg | ttcgccaggc | tcaaggcggc | catgcccagc | 3240 |
| ggcgaaggatc | tcgtcgtgac | ccatggcgat | gcctgcttgc | cgaatatcat | ggtggaaaat | 3300 |
| ggccgctttt | ctggattcat | cgactgtggc | cggctgggtg | tggcggaccg | ctatcaggac | 3360 |
| atagcgttgg | ctaccctgta | tattgtctgaa | gagcttggcg | gcgaatgggc | tgaccgcttc | 3420 |
| ctcgtgcttt | acgggtatcg | cgctcccgat | tcgcagcgca | tcgccttcta | tcgccttctt | 3480 |
| gacgagttct | tctgagcggg | actctgggg | tcgaaatgac | cgaccaagcg | acgcccaccc | 3540 |
| tgccatcacg | atggccgcaa | taaaatatct | ttattttcat | tacatctgtg | tgttggtttt | 3600 |
| ttgtgtgaat | cgatagcgat | aaggatccgc | gtatggtgca | ctctcagtac | aatctgctct | 3660 |
| gatgcccgat | agttaaagcca | gccccgacac | ccgccaacac | ccgctgacgc | gcctgacgg | 3720 |
| gcttgcctgc | tcccggcctc | cgcttacaga | caagctgtga | ccgtctccgg | gagctgcatg | 3780 |
| tgctcagagg | tttccaccgtc | atcaccgaaa | cgccgagac | gaaagggcct | cgtgatacgc | 3840 |
| ctatttttat | aggttaatgt | catgataata | atggtttctt | agacgtcagg | tggcactttt | 3900 |
| cggggaaatg | tgccggaagc | ccctatttgt | ttattttctt | aaatacattc | aaatatgtat | 3960 |
| ccgctcatga | gacaataacc | ctgataaatg | cttcaataat | attgaaaaag | gaagagtatg | 4020 |
| agtattcaac | atttccgtgt | cgcccttatt | ccctttttg | cggcattttg | ccttctctgt | 4080 |
| tttctcacc | cagaaaagct | ggtgaaagta | aaagatgctg | aagatcagtt | gggtgcacga | 4140 |
| gtgggttaca | tcgaaactgga | tctcaacagc | ggtaagatcc | ttgagagttt | tcgccccgaa | 4200 |
| gaacgttttc | caatgatgag | cacttttaaa | gttctgctat | gtggcggcgg | attatcccgt | 4260 |

ES 2 511 694 T3

attgacgccg ggcaagagca actcggctgc cgcatacact attctcagaa tgacttgggt 4320
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 4380
agtgctgccca taaccatgag tgataacact gcggccaact tacttctgac aacgatcggg 4440
ggaccgaag agctaacccg ttttttgac aacatggggg atcatgtaac tcgccttgat 4500
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 4560
gtagcaatgg caacaacgtt gcgcaacta ttaactggcg aactacttac tctagcttcc 4620
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 4680
gcccttccgg ctggctggtt tattgctgat aaatctggag ccgggtgagcg tgggtctcgc 4740
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 4800
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 4860
ctgattaagc attggttaact gtcagaccaa gtttactcat atatacttta gattgattta 4920
aaacttcatt ttttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 4980
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 5040
ggatcttctt gagatccttt tttctcgcg gtaatctgct gcttgcaaac aaaaaacca 5100
ccgctaccag cgggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 5160
actggcttca gcagagcgca gataccaaat actgttcttc tagttagacc gtagttaggc 5220
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaact cctgttacca 5280
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 5340
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 5400
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgcacgctt 5460
cccgaaggga gaaaggcggg caggtatccg gtaagcggca gggctcggaa aggagagcgc 5520
acgagggagc ttccaggggg aaacgcctgg tatctttata gtctgtcgg gtttcgccac 5580
ctctgacttg agcgtcgatt tttgtgatgc tcgctcagggt ggcggaacct atggaaaaac 5640
gccagcaacg cgccttttt acggttctcg gccttttgc ggccttttgc tcacatggct 5700
cgacagatct 5710

- <210> 17
- <211> 5819
- <212> ADN
- <213> Homo sapiens

5

- <220>
- <221> característica_misclánea
- <223> vector de expresión llamado GS-V3 que contiene la secuencia antisentido de GS-N3; antisentido:1108...1466

10 <400> 17

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240
gcctggctga ccgccaacg acccccgcc attgacgtca ataatgacgt atgttcccat 300
agtaaagcca ataggactt tcattgacg tcaatgggtg gagtattac ggtaaacctg 360
ccacttgcca gtacatcaag tgtatcata gccaaagtcg ccccctattg acgtcaatga 420
cggtaaatgg cccgcctggc attatgcca gtacatgacc ttacgggact ttcctacttg 480
gcagtacatc tacgtattag tcatcgctat taccatgggt atgctggttt ggcatgacac 540
caatggcgtt ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt 600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactg 660
cgatcgcccg ccccgttgac gcaaatgggc ggtagcgtg tacgggtggg ggtctatata 720
agcagagctc gtttagtgaa ccgtcagatc actagaagct ttattgoggt agtttatcac 780
agttaaattg ctaacgcagt cagtgttct gacacaacag tctcgaactt aagctgcagt 840
gactctotta aggtagcctt gcagaagttg gtcgtgaggg actgggcagg taagtatcaa 900
ggttacaaga caggtttaag gagaccaata gaaactgggc ttgtcgagac agagaagact 960
cttgcttttc tgataggcac ctattggtct tactgacatc cactttgcct ttctctccac 1020
aggtgtocac toccagttca attacagctc ttaaggctag agtacttaac acgactcact 1080
ataggctagc ctcgagaatt cacgcgttgc acgtcagtg cctcagttgc atcctggttc 1140
cggctcgaca gccgagcgat ggctcattg tcataatgaa tcacactgct gtctcctcc 1200
ttgttctccc cctcgttttc atccttgaat agctcttcag tgccaaattt gagaatgtcg 1260
tcaagctcct gcttggacat ggagcctgcc ttggagccca gcccgggcg cacaaccagg 1320
tgtgtcagca tcatcttct cttggccact ctgtgtattc gctcttccac tgacgcgcga 1380
gtcacaacc ggtaaatcat cactttgttg gcctggccaa tccgatgagc cgggctaaag 1440
gcctggatgt cattatgggg gttccagctg acccggggcg ccgcttccct ttagtgaggg 1500
ttaatgcttc gagcagacat gataagatac attgatgagt ttggacaac cacaactaga 1560
atgacagtga aaaaatgctt tattttgtgaa atttgtgat ctattgcttt atttgtaacc 1620
attataagct gcaataaac agttaacaac aacaattgca ttcattttat gtttcagggt 1680
cagggggaga tgtggagggt ttttaaacg aagtaaaacc tctacaatg ttgtaaaatc 1740
cgataaggat cgtaccgggc tggcgtaata gcgaagaggg ccgcaccgat cgccttccc 1800
aacagttgcg cagcctgaat ggcaatgga cgcgcctgt agcggcgcat taagcggcg 1860

ES 2 511 694 T3

ggggtgtggg gttacgcgca gcgtgaccgc tacacttgcc agcgccctag cgcccgtcc 1920
 tttcgtttt ttcccttctt ttctcgccac gttcgcggc tttcccgtc aagctctaaa 1980
 tcgggggctc ccttttaggt tccgatttag tgctttacgg cacctcgacc ccaaaaaact 2040
 tgattagggg gatggttcac gtagtggggc atcgcctga tagacggttt ttcgccctt 2100
 gacgttgag tccacgttct ttaatagtg actcttgctc caaactggaa caaactcaa 2160
 ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg cctattgggt 2220
 aaaaaatgag ctgatttaac aaaaatttaa cgcgaatfff aacaaaatat taacgcttac 2280
 aatttctga tgcggtatit tctccttac catctgtgcg gtatttcaca ccgcatacgc 2340
 ggatctgcgc agcaccatgg cctgaaataa cctctgaaag aggaacttgg ttaggtacct 2400
 tctgaggcgg aaagaaccag ctgtggaatg tgtgtcagtt aggggtgtga aagtcccag 2460
 gctcccagc aggcagaagt atgcaaaagca tgcctctcaa ttagtcagca accagggtg 2520
 gaaagtcccc aggcctccca gcaggcagaa gtatgcaaaq catgcatctc aattagtcag 2580
 caaccatagt cccgccccta actccgcca tcccgcctt aactccgccc agtccgccc 2640
 atttccgccc ccatggctga ctaattttt ttatttatgc agaggccgag gccgcctcg 2700
 cctctgagct attccagaag tagtgaggag gcttttttgg aggcctaggg ttttgcaaaa 2760
 agcttgattc ttctgacaca acagtctga acttaaggct agagccacca tgattgaaca 2820
 agatggattg cacgcaggtt ctccggccgc ttgggtggag aggctattcg gctatgactg 2880
 ggcacaacag acaatcggct gctctgatgc cgccgtgttc cggctgtcag cgcagggcg 2940
 cccggttctt tttgtcaaga ccgacctgtc cggtgccctg aatgaactgc aggacgaggc 3000
 agcgcggcta tcgtggctgg ccacgacggg cgctccttgc gcagctgtgc tcgacgttgc 3060
 cactgaagcg ggaaggatg ggctgctatt gggcgaagtg ccggggcagg atctcctgtc 3120
 atctaacctt gctcctgacg agaaagtatc catcatggct gatgcaatgc ggcggctgca 3180
 tacgcttgat ccggctacct gccattcga ccaccaagcg aaacatcgca tcgagcgagc 3240
 acgtactcgg atggaagccg gtcttgcga tcaggatgat ctggacgaag agcatcaggg 3300
 gctcgcgcca gccgaactgt tcgccaggct caaggcgcgc atgcccagc gcgaggatct 3360
 cgtcgtgacc catggcgatg cctgcttgcc gaatatcatg gtggaaaaatg gccgctttc 3420
 tggattcatc gactgtggcc ggctgggtgt ggcggaccgc taccaggaca tagcgttggc 3480
 taccogtgat attgctgaag agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta 3540
 cggtatcgcg gctcccgatt cgcagcgcac cgccttctat cgccttcttg acgagttctt 3600
 ctgagcggga ctctggggtt cgaatgacc gaccaagcga cgcccacct gccatcacga 3660
 tggccgcaat aaaatatctt tattttcatt acatctgtgt gttggttttt tgtgtgaatc 3720
 gatagcgata aggatcccg gctgtgacac tctcagtaca atctgctctg atgccgata 3780
 gttaagccag ccccgacacc cgccaacacc cctgacgcg ccctgacggg ctgtctgct 3840
 cccggcatcc gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagagggt 3900
 ttaccogtca tcaccgaac gcgcgagacg aaaggcctc gtgatacgc tatttttata 3960
 ggtaaatgct atgataataa tggtttctta gacgtcagggt ggcacttttc ggggaaatgt 4020
 gcgcggaacc cctatttggtt tatttttcta aatacattca aatatgtatc cgtcatgag 4080
 acaataaccc tgataaatg ttcaataata ttgaaaaagg aagagtatga gtattcaaca 4140
 tttccgtgct gcccttattc ccttttttgc ggcattttgc cttcctgttt ttgctcacc 4200
 agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat 4260
 cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgag aacgttttcc 4320
 aatgatgagc acttttaag ttctgctatg tggcgcggta ttatcccgt ttagcgcgg 4380
 gcaagagcaa ctccgtcgcc gcatacacta ttctcagaat gacttggtg agtactcacc 4440
 agtcacagaa aagcatctta cggatggcat gcagtaaga gaattatgca gtgctgccat 4500
 aacctgagtg gataaacctg cggccaactt acttctgaca acgatcggag gaccgaagga 4560
 gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc gttgggaacc 4620
 ggagctgaat gaagccatac caacgacga gcgtgacacc acgatgcctg tagcaatggc 4680
 aacaacgttg cgcaaacctat taactggcga actacttact ctagcttccc ggcaacaatt 4740
 aatgactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg cccttccggc 4800
 tggctgggtt attgctgata aatctggagc cgggtgagct ggggtctcgg gtatcattgc 4860
 agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtca 4920
 ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca 4980
 ttggttaact tcagaccaag ttactcata tatactttag attgatttaa aacttcatt 5040
 ttaatttaa aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta 5100
 acgtgagttt tcgctccact gagcgtcaga ccccgtagaa aagatcaaaq gatcttctg 5160
 agatcctttt tttctgcgcg taatctgctg cttgcaaaaca aaaaaaccac cgtaccagc 5220
 ggtggtttgt ttgccggatc aagagctacc aactctttt cgaaggtaa ctggcttcag 5280
 cagagcgcag ataccaaata ctgttcttct agttagcgg tagttaggcc accacttcaa 5340
 gaactctgta gcaccgcta catacctcgc tctgctaate ctgttaccag tggctgctgc 5400
 cagtggcgat aagctgtgct ttaccgggtt ggaactcaaga cgatagttac cggataaggc 5460
 gcagcggctg ggctgaacgg ggggttcgtg cacacagccc agcttggagc gaacgacct 5520
 caccgaactg agatacctac agcgtgagct atgagaaag gccacgctc ccgaaggag 5580
 aaaggcggac aggtatccgg taagcggcag ggtcggaaaca ggagagcga cgaggagct 5640
 tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgcacc tctgacttga 5700
 gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc 5760
 ggccttttta cggttcctgg ccttttgctg gcctttgctc cacatggctc gacagatct 5819

<210> 18
 <211> 5627

ES 2 511 694 T3

<212> ADN
 <213> Homo sapiens

<220>
 <221> característica_misclánea

5 <223> vector de expresión llamado GS-V4 que contiene la secuencia antisentido de GS-N4;
 antisentido:1108...2606

<220>
 <221> característica_misclánea

10 <223> vector de expresión llamado GS-V4 que contiene la secuencia antisentido de GS-N5;
 antisentido:1108...2606

<400> 18

| | | | | | | |
|------------|-------------|------------|-------------|-------------|-------------|------|
| tcaatattgg | ccattagcca | tattattcat | tggttatata | gcataaatca | atattggcta | 60 |
| ttggccattg | catacgttgt | atctatatca | taatatgtac | atttatattg | gctcatgtcc | 120 |
| aatatgaccg | ccatgttggc | attgattatt | gactagttat | taatagtaat | caattacggg | 180 |
| gtcattagtt | catagcccat | atatggagtt | ccgcgttaca | taacttacgg | taaatggccc | 240 |
| gcctggctga | ccgcccacg | acccccgcc | attgacgtca | ataatgacgt | atgttcccat | 300 |
| agtaaagcca | atagggactt | tccattgacg | tcaatgggtg | gagtatttac | ggtaaaactgc | 360 |
| ccacttggca | gtacatcaag | tgtatcatat | gccaaagtccg | cccctattg | acgtcaatga | 420 |
| cggtaaatgg | ccgcctggc | attatgcccc | gtacatgacc | ttacgggact | ttcctacttg | 480 |
| gcagtacatc | tacgtattag | tcatcgctat | taccatgggtg | atgcccgtttt | ggcagtacac | 540 |
| caatggcggt | ggatagcgg | ttgactcacg | gggatttcca | agtctccacc | ccattgacgt | 600 |
| caatgggagt | ttgttttggc | acaaaaatca | acgggacttt | ccaaaatgtc | gtaacaactg | 660 |
| cgatcgccc | cccogttgac | gcaaatgggc | ggtaggcgtg | tacggtgga | ggtctatata | 720 |
| agcagagctc | gtttagtga | ccgtcagatc | actagaagct | ttattgcggt | agtttatcac | 780 |
| agttaaattg | ctaacgcagt | cagtgttct | gacacaacag | tctcgaactt | aagctgcagt | 840 |
| gactctctta | aggtagcctt | gcagaagttg | gtcgtgaggc | actgggcagg | taagatcaa | 900 |
| ggttacaaga | caggttttaag | gagaccaata | gaaactgggc | ttgtcagac | agagaagact | 960 |
| cttgcgtttc | tgataggcac | ctattggct | tactgacatc | cactttgcct | ttctctccac | 1020 |
| aggtgtccac | tcccagttca | attacagctc | ttaaggctag | agtactaat | acgactcact | 1080 |
| ataggttagc | ctcgagaatt | cacgcgtcag | gcgggaaca | gagtggagag | ctcagtaggc | 1140 |
| cgtctgagac | tgctgctggc | ggtagccacc | gcggcgcag | tagccctcgt | ttttgcggta | 1200 |
| gccgtccttc | tggtctcgg | agtagcccc | gtaggtgcc | tgcttgggt | caaacaccg | 1260 |
| ttcgttgttc | tccagtcgac | ccggcggcc | gcttccctt | agtgggggt | aatgcttcca | 1320 |
| gcagacatga | taagatacat | tgatgagttt | ggacaaacca | caactagaat | gcagtgaaaa | 1380 |
| aaatgcttta | tttgtgaaat | ttgtgatgct | attgctttat | ttgtaaccat | tataagctgc | 1440 |
| aataaacaag | ttaacaacaa | caattgcatt | cattttatgt | ttcaggttca | gggggagatg | 1500 |
| tgggaggttt | tttaaagcaa | gtaaacctc | tacaaatgtg | gtaaaatccg | ataaggatcg | 1560 |
| atccgggctg | gcgtaaatgc | gaagaggccc | gcaccgatcg | cccttcccaa | cagttgcgca | 1620 |
| gcctgaatgg | cgaatggacg | cgccctgtag | cggcgcatta | agcggggcgg | gtgtggtggt | 1680 |
| tacgtcgcag | tgacccgcta | cacttgccag | gcgcctagcg | cccgtcctt | tcgctttctt | 1740 |
| cccttctttt | ctcgccacgt | tcgcccgtt | tcccgtcaa | gctctaaatc | gggggctccc | 1800 |
| tttagggttc | cgatttagtg | ctttacggca | cctcgacc | aaaaaacttg | attaggggtga | 1860 |
| tggttcacgt | agtgggcat | cgccctgata | gacggttttt | cgcccttga | cgttggagtc | 1920 |
| cacgttcttt | aatagtgagc | tcttgttcca | aactggaaca | acactcaacc | ctatctcgg | 1980 |
| ctattctttt | gatttataag | ggattttgcc | gatttcggcc | tattggttaa | aaaatgagct | 2040 |
| gatttaacaa | aaatttaacg | cgaattttaa | caaaatatta | acgcttaca | tttctgatg | 2100 |
| cggatatttc | tccttacgca | tctgtgoggt | atttcacacc | gcatacgcgg | atctgcgcag | 2160 |
| caccatggcc | tgaataaacc | tctgaaagag | gaacttggtt | aggtaccttc | tgaggcggaa | 2220 |
| agaaccagct | gtggaatgtg | tgtaagttag | ggtgtggaaa | gtccccagc | tccccagcag | 2280 |
| gcagaagtat | gcaaagcatg | catctcaatt | agtcagcaac | caggtgtgga | aagtccccag | 2340 |
| gctccccagc | aggcagaagt | atgcaaagca | tgcatctcaa | ttagtacgca | accatagtcc | 2400 |
| cgcccctaac | tccgcccata | cggcccctaa | ctccgcccag | ttccgccc | tctccgccc | 2460 |
| atggctgact | aatttttttt | atztatgcag | aggccgaggc | cgccctggcc | tctgagctat | 2520 |
| tccagaagta | gtgaggaggc | ttttttggag | gcctaggctt | ttgcaaaaag | cttgattctt | 2580 |
| ctgacacaac | agtctcgaac | ttaaggctag | agccaccatg | attgaacaag | atggattgca | 2640 |
| cgcaggttct | ccggccgctt | gggtggagag | gctattcggc | tatgactggg | cacaacagac | 2700 |
| aatcgctg | tctgatgccc | ccgtgttccg | gctgtcagcg | cagggggcgc | cggttctttt | 2760 |
| tgtcaagacc | gacctgtccg | gtgcccgtga | tgaactgcag | gacgaggcag | cgcggtatc | 2820 |
| gtggctggcc | acgacggg | ttccttgcgc | agctgtgctc | gacgttgtca | ctgaagcggg | 2880 |
| aagggactgg | ctgctattgg | gcgaagtgcc | ggggcaggat | ctcctgtcat | ctcaccttgc | 2940 |
| tcctgcccag | aaagtatcca | tcatggctga | tgcaatgcgg | cggctgcata | cgcttgatcc | 3000 |
| ggctacctgc | ccattcgacc | accaagcgaa | acatcgcatc | gagcagcac | gtactcggat | 3060 |
| ggaagccggt | cttgtcagatc | aggatgatct | ggacgaagag | catcaggggc | tcgcccagc | 3120 |

ES 2 511 694 T3

cgaactgttc gccaggctca aggcgcgcat gcccgacggc gaggatctcg tcgtgaccca 3180
 tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg gattcatcga 3240
 ctggtggccg ctgggtgtgg cggaccgcta tcaggacata gcggttgcta cccgtgatat 3300
 tgctgaagag ctgggcggcg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc 3360
 tcccgattcg cagcgcacgc ccttctatcg ccttcttgac gagttcttct gagcgggact 3420
 ctggggttcg aatgaccga ccaagcgacg cccaacctgc catcacgatg gccgcaataa 3480
 aatatcttta ttttcattac atctgtgtgt tggttttttg tgtgaatcga tagcgataag 3540
 gatccgcgta tgggtgcactc tcagtacaat ctgctctgat gccgcatagt taagccagcc 3600
 ccgacacccg ccaacaccog ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc 3660
 ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcatc 3720
 accgaacgc gcgagacgaa agggcctcgt gatacgccta tttttatagg ttaatgtcat 3780
 gataataatg gtttcttaga cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc 3840
 tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg 3900
 ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc 3960
 ccttattccc ttttttgccg cattttgcct tcctgttttt gctcaccag aaacgtggt 4020
 gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct 4080
 caacagcggg aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac 4140
 ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgcggggc aagagcaact 4200
 cggtcgcccgc atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa 4260
 gcatcttacg gatggcatga cagttaagaga attatgcagt gctgccataa ccatgagtga 4320
 taacactgcg gccaaactac ttctgacaac gatcggagga ccgaaggagc taaccgcttt 4380
 tttgcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga 4440
 agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg 4500
 caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat 4560
 ggaggcggat aaagttgcag gaccacttct gcgctcggcc ctccggctg gctggtttat 4620
 tgctgataaa tctggagccg gtgagcgtgg gtctcgggt atcattgcag cactggggcc 4680
 agatggtaag ccctcccgta tcgtagtatt ctacacgacg gggagtcagg caactatgga 4740
 tgaacgaaat agacagatcg ctgagatagg tgccctactg attaagcatt ggtaactgct 4800
 agaccaagtt tactcatata tactttagat tgatttaaaa cttcattttt aatttaaaag 4860
 gatctagggtg aagatccttt ttgataatct catgaccaa atccctaac gtgagttttc 4920
 gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt 4980
 tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt 5040
 gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat 5100
 accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc 5160
 accgcctaca tacctcgtc tgctaatacct gttaccagtg gctgctgcca gtggcgataa 5220
 gtcgtgtctt accgggttg actcaagacg atagttaccg gataaggcgc agcggtcggg 5280
 ctgaacgggg gttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag 5340
 atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggggagaa aggcggacag 5400
 gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa 5460
 cgctcgggat ctttatagtc ctgtcggggt tcgccacctc tgacttgagc gtcgattttt 5520
 gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg 5580
 gttcctggcc ttttgctgac cttttgctca catggctcga cagatct 5627

- <210> 19
- <211> 29
- <212> ADN
- 5 <213> Homo sapiens

- <220>
- <221> característica_misclánea
- <223> cebador sentido: GV1-1

<400> 19

10 cgggtcgaca ggtccactgc agggggtta 29

- <210> 20
- <211> 31
- <212> ADN
- <213> Homo sapiens

- 15 <220>
- <221> característica_misclánea
- <223> cebador inverso: GV1-2

<400> 20

cgcacgcggt tccccttgg aagagagac a 31

ES 2 511 694 T3

<210> 21
 <211> 29
 <212> ADN
 <213> Homo sapiens

5 <220>
 <221> característica_misclánea
 <223> cebador sentido: GV2-1
 <400> 21
 cgggtcgacg ggccagctta ctgttggtg 29

10 <210> 22
 <211> 32
 <212> ADN
 <213> Homo sapiens
 <220>
 <221> característica_misclánea
 <223> cebador inverso: GV2-2
 <400> 22
 cgcacgcggt tcagaggagg ggagagaaag ag 32

20 <210> 23
 <211> 31
 <212> ADN
 <213> Homo sapiens
 <220>
 <221> característica_misclánea
 <223> cebador sentido: GV3-1
 <400> 23
 cgggtcgact ggaaccccca taatgacatc c 31

30 <210> 24
 <211> 31
 <212> ADN
 <213> Homo sapiens
 <220>
 <221> característica_misclánea
 <223> cebador inverso: GV3-2
 <400> 24
 cgcacgcggt gcacgtcagt gtctcagtt g 31

40 <210> 25
 <211> 30
 <212> ADN
 <213> Homo sapiens
 <220>
 <221> característica_misclánea
 <223> cebador sentido: GV4-1
 <400> 25
 cgggtcgact ggagaacaac gaacgggtgt 30

45 <210> 26
 <211> 29
 <212> ADN
 <213> Homo sapiens

ES 2 511 694 T3

<220>

<221> característica_misclánea

<223> cebador inverso: GV4-2

<400> 26

5 cgcacgcgctc aggcgggaaa cagagtgga

29

<210> 27

<211> 5975

<212> ADN

<213> Homo sapiens

10 <300>

<308> GenBank/XM_040709

<309> 07-02-2002

<313> (1)..(5975)

<400> 27

ES 2 511 694 T3

| | | | | | | |
|------------|-------------|------------|------------|------------|-------------|------|
| gctctttgcc | gagggcgtgt | ggtgagagtc | cccacagcga | ccctggttcg | agtggggggc | 60 |
| actgagctgg | tcatcccctg | caacgtcagt | gactatgatg | gccccagcga | gcaaaacttt | 120 |
| gactggagct | tctcatcttt | ggggagcagc | tttgtggagc | ttgcaagcac | ctgggaggtg | 180 |
| gggttcccag | cccagctgta | ccaggagcgg | ctgcagaggg | gcgagatcct | gttaaggcgg | 240 |
| actgccaacg | acgccgtgga | gtccacata | aagaacgtcc | agccttcaga | ccaaggccac | 300 |
| tacaaatggt | caacccccag | cacagatgcc | actgtccagg | gaaactatga | ggacacagtg | 360 |
| caggttaaag | tgctggccga | ctccctgcac | gtgggcccc | gcgcgcggcc | cccgccgagc | 420 |
| ctgagcctgc | gggaggggga | gcccttcgag | ctgcgctgca | ccgccgcctc | cgctcgcgg | 480 |
| ctgcacacgc | acctggcgct | gctgtgggag | gtgcaccgcg | gcccggccag | gcggagcgtc | 540 |
| ctgcgcctga | cccacgaggg | caggttccac | ccgggctgg | ggtacgagca | gcgctaccac | 600 |
| agtggggagc | tgccctcga | caccgtgggc | agcgacgcct | accgcctctc | agtgtcccgg | 660 |
| gctctgtctg | ccgaccaggg | ctcctacagg | tgtatcgtca | gcgagtggat | cgccgagcag | 720 |
| ggcaactggc | aggaaatcca | agaaaaggcc | gtggaagtgg | ccaccgtggt | gatccagcca | 780 |
| tcagtctctg | gagcagctgt | gcccaagaat | gtgtctgtgg | ctgaaggaaa | ggaactggac | 840 |
| ctgacctgta | acatcacaac | agaccgagcc | gatgacgtcc | ggcccagagt | gacgtgggtc | 900 |
| ttcagcagga | tgccctgacag | caccctacct | ggctcccgg | tggtggcgg | gcttgaccgt | 960 |
| gattccctgg | tgccacagctc | gcctcatggt | gctttgagtc | atgtggatgc | acgctcctac | 1020 |
| catttactgg | ttcgggatgt | tagcaaagaa | aactctggct | actattactg | ccacgtgtcc | 1080 |
| ctgtgggcac | ccggacacaa | caggagctgg | cacaaagtgg | cagaggccgt | gtcttcccca | 1140 |
| gctgtgtgtg | gtgtgacctg | gctagaacca | gactaccagg | tgtacctgaa | tgctccaag | 1200 |
| gtcccgggtg | ttgcggtatg | ccccacagag | ctggcatgcc | gggtgggtga | cacgaagagt | 1260 |
| ggggagggca | atgtccgatt | cacggtttcg | tggtactaca | ggatgaaccg | gcgcagcgac | 1320 |
| aatctgttga | ccagcagctt | gcttgacgtc | atggacgggg | actggacgct | aaaatatgga | 1380 |
| gagaggagca | agcagcgggc | ccaggatgga | gactttatct | tttctaagga | acatacagac | 1440 |
| acgttcaatt | tccggatcca | aaggactaca | gaggaagaca | gaggcaatta | ttactgtgtt | 1500 |
| gtgtctgcct | ggaccaaaaca | gcggaacaac | agctgggtga | aaagcaagga | tgtcttctcc | 1560 |
| aagcctgtta | acatattttg | ggcattagaa | gattccgtgc | ttgtgggtga | ggcgaggcag | 1620 |
| ccaaagcctt | tctttgctgc | cgaaaataca | tttgagatga | cttgcaaagt | atcttccaag | 1680 |
| aatattaagt | cgccacgcta | ctctgttctc | atcatggctg | agaagcctgt | cgccgacctc | 1740 |
| tccagtccca | atgaaacgaa | gtacatcadc | tctctggacc | aggattctgt | ggtgaagctg | 1800 |
| gagaattgga | cagatgcadc | acgggtggat | ggcgttgttt | tagaaaaagt | gcaggaggat | 1860 |
| gagttccgct | atcgaatgta | ccagactcag | gtctcagacg | cagggctgta | ccgctgcagt | 1920 |
| gtgacagcct | ggtctcctgt | caggggcagc | ctttggcgag | aagcagcaac | cagtctctcc | 1980 |
| aatcctattg | agatagactt | ccaaacctca | ggtcctatat | ttaatgcttc | tgtgcattca | 2040 |
| gacacaccat | cagtaattcg | gggagatctg | atcaaattgt | tctgtatcat | cactgtcgag | 2100 |
| ggagcagcac | tggatccaga | tgacatggcc | tttgatgtgt | cctggtttgc | ggtgactctc | 2160 |
| tttggcctgg | acaaggctcc | tgtgctcctg | tcttccctgg | atcgggaagg | catcgtgacc | 2220 |
| acctcccgga | gggactggaa | gagcgcacct | agcctggagc | gcgtgagtgt | gctggaattc | 2280 |
| ttgctgcaag | tgcatggctc | cgaggaccag | gactttggca | actactactg | ttccgtgact | 2340 |
| ccatgggtga | agtcaccaac | agttcctgg | cagaaggagg | cagagatcca | ctccaagccc | 2400 |
| gtttttataa | ctgtgaagat | ggatgtgctg | aaoccttca | agtatccctt | gctgatcggc | 2460 |
| gtcggctctg | ccacggctcat | cggtcctctg | tctgtctca | tcgggtactg | cagctcccac | 2520 |
| tggtgttgta | agaaggaggt | tcaggagaca | cggcgcgagc | gcccagcgtt | catgtcagtg | 2580 |
| gagatggact | aggtggccc | gggaggggag | tgacagaggg | acgttctagg | agcaattggg | 2640 |
| gcaagaagag | gacagtgata | ttttaaaca | aagtgtgtta | caactaaaac | cagtccctctc | 2700 |
| taatctcagg | tgggacttgg | cgctctctct | tttctgcatg | tcaagttctg | agcgcggaca | 2760 |
| tgtttaccag | cacacggctc | ttctcccac | ggcactttct | gatgtaacaa | tcgagtgtgt | 2820 |
| gttttcccaa | ctgcagcttt | ttaatggtta | accttcatct | aatttttttt | ctcccactgg | 2880 |
| tttatagatc | ctctgacttg | tgtgtgttta | tagcttttgt | ttcgcggggt | tgtggtgagg | 2940 |
| aaggggtgat | ggcatgcgga | gttctttatc | ttcagtgaga | atgtgcctgc | ccgcctgaga | 3000 |
| gccagcttcc | gcgttgaggg | cacgtgttca | gagagctgct | gagcgcacc | ctctaccogg | 3060 |
| ctgacagaca | acacagacct | gtgccgaagg | ctaatttgg | gcttttacga | ccctacccca | 3120 |
| cccctgttt | tcaggggttt | agactacatt | tgaaatccaa | acttgagta | tataacttct | 3180 |
| tattgagccc | aactgctttt | ttttttttt | tttttgcttc | tctgcccctt | ttccatttct | 3240 |
| tttgattttg | ttttctgtga | gagcactgaa | atggcagccc | tggaatctac | aatttggtctc | 3300 |
| tccactgagc | accttatctt | gccaccttag | ccttaagaat | gaatatgaag | aaaaatacac | 3360 |
| agccacctct | gtccagggca | gtaagaagg | ctgcaaggaa | ggggaggtg | gggacaagga | 3420 |
| aaggatcaga | tacctgctcc | agttagttgt | aggccactgt | gtctcagggg | actccaggag | 3480 |
| gagcagaaga | gggatcccac | gaagtatttc | ttacgcagct | ggggccagga | gggtcagagt | 3540 |
| ggtgccaggt | gcaagttagg | ctaagaagc | caccactatt | cctctctctt | gcccattgtg | 3600 |
| gggggcaaa | gcattggtca | ccaagatct | tgcaggggga | cccacagata | tgccatgtcc | 3660 |

ES 2 511 694 T3

ttcacacgtg cttgggctcc ttaacctgaa ggcaaattgc tacttgcaag actgactgac 3720
 ttcaaggaat cagaaattac ctagaagcac catgtttttt ctatgacctt ttcagtcctt 3780
 caggctcattt taagggtccac tgcaggggggt tagtgagaaa gggataactt tgtgggatgt 3840
 tttgctttcc taatagggac atgaaggaaa cccagcaatt tgcgtttatg tgaatggcct 3900
 gtagagcaga gtcaagagcg gtgtgctttg cccgactgct cccatcagga ataggagagt 3960
 agacagagat cttccacatc ccaggcttct gctgctgctt taaaagctct gtccttggag 4020
 cctcccgcctc cctgaagtgt ctgccccct gcacagcact ggcctttcgg aagcatccca 4080
 gtagggtttt ctgaggctcg ctggtgactc atgccctaat tgcaatcctc tgcttttatc 4140
 ttgactttga aggatctaac actgctctct cttccaaagg ggaaaaaag attcatttgt 4200
 tttgagcaat aaactaatac aaaaatgatgg ccattcatgt gcagctcttt gtcaccatgg 4260
 gccggatgag ttgtgctcct cctggctcac cattedcccc tgctcccca cagccggttc 4320
 tgcacttatc accgagtcgc ccctggaagc agattcccat tgagttttcc ccaccaaggg 4380
 gaccatgcac atggtagaaa cattagattc tgcattgaca gtagcctttc cttggcccgg 4440
 gcctgtgggtg ggaagacggg caacaagtat accccaccag ggcctgagtg actagaggaa 4500
 gaggacgagg ccttgttggc actagatttg ggtattttct gcatgtcata acatatacta 4560
 actgctattt cagaagaggc agctttagag tgattgtaca agtgagaatt aaagagagaa 4620
 cagatattta aacagggtct gtattagtaa cagccagtcg cctttcagcc cttgcatcta 4680
 ttaaaaaggag attcaggatt ttattggcac aggcccttct tagtaggaag aaagggtgct 4740
 tagctttgga cctgaccggg tgtgtgtaaa accatggact gagtccacgc agacactcga 4800
 tgggtgtaaa tgtgatgggt gcttacacac tgtacctttt cctttcatac tgatgctgca 4860
 gttcagggtc ggagttgtta aggcattgac ctcccaccac ctgccccatg tccactgggc 4920
 tgcccagct gcattgtcacc tgagggtctg caggaagggg cgagaaatcc cagggcattg 4980
 tacciaaggc ctagtctcct ctagggatat aaatttccag gaatgtgtat ttttaatgtg 5040
 gtgagatgca ctcttttgtt gtaccaataa gggctcccca ccccaccct gcgacaagtg 5100
 ctcttctaga acaggttcct accagcagca ctgggtgtgaa tgaagagag acccagccgc 5160
 gtctcacaca ggtggaattg cacttcttaa caaaaaggaa cttataaaa gtttgggatt 5220
 tttttccta atcataaaaa tagccccaga aagagcctaa gctatgttca gatagaagcc 5280
 tcgaaatttc tgtaaattgt ttactttatg atgtttacat acacgtttca ctttgaaaaa 5340
 aaatgcaaat cgacttttta acaactgttg agatgtttca tgggacagta gaactctgac 5400
 tcaccaactg ggctaaattt taatttaaaa atgtatttat ttgagtgtct tccccccct 5460
 caccctcacc atctgagggg ctccctgaga tcttggtaga ggaggccct cctgcccaga 5520
 ccttcgtttg ttccccgggt ggcccttgct tcttgccttg cagactgcct gcagccatga 5580
 ttttgtcact gacatctgtg agccaaagac tgagcctttt tggcaggaat aataagcaat 5640
 actacacaac ttgctacttt cagaaaactt ttttttagct tcaccgatga caacagagga 5700
 agaagggaac tgggatttgg gtaagtctc ctccactggt tgaccaaaat ctcagtgata 5760
 aatatgtgtg cagatcccta gaagagaaaa cgctgacttt ctttttaagt gtggcacata 5820
 aggatctgca gaattttccg tagacaaaga aaggatcttg tgtattttt tccatatcca 5880
 atgttatatg aactaattgt attgttttat actgtgacca caaatattat gcaatgcacc 5940
 atttgttttt tatttcatta aaggaagttt aattt 5975

- <210> 28
- <211> 2197
- <212> ADN
- 5 <213> Homo sapiens

- <220>
- <221> característica_misclánea
- <222> (1037)..(1037)
- <223> n es un residuo que se puede seleccionar entre los nucleótidos a, t, c o g

- 10 <220>
- <221> característica_misclánea
- <222> (1236)..(1236)
- <223> n es un residuo que se puede seleccionar entre los nucleótidos a,t, c o g.

- <300>
- 15 <308> GenBank/AB014734
- <309> 06-01-2001
- <313> (1)..(2197)

<400> 28

gtgcattcag acacaccatc agtaattcgg ggagatctga tcaaattggt ctgtatcatc 60
 actgtcgagg gaggcagcact ggatccagat gacatggcct ttgatgtgtc ctggtttgcg 120
 gtgcactcct ttggcctgga caaggctcct gtgctcctgt cttccctgga tcggaagggc 180
 atcgtgacca cctcccggag ggactggaag agcgcactca gcttgagcgc cgtgagtgtg 240
 ctggaattct tgctgcaagt gcatggctcc gaggaccagg actttggcaa ctactactgt 300
 tccgtgactc catgggtgaa gtcaccaaca ggttcctggc agaaggaggc agagatccac 360

ES 2 511 694 T3

```
tccaagcccg tttttataac tgtgaagatg gatgtgctga acgccttcaa gtatcccttg 420
ctgatcggcg tcggtctgtc cacggtcatc gggctcctgt cctgtctcat cgggtactgc 480
agctcccact ggtgttgtaa gaaggagggt caggagacac ggcgcgagcg ccgcaggctc 540
atgtcgatgg agatggacta ggctggcccc ggaggggagt gacagaggga cgttctagga 600
gcaattgggg caagaagagg acagtgatat tttaaaacia agtgtgttac actaaaaacc 660
agtcctctct aatctcaggt gggacttggc gctctctctt ttctgcatgt caagttctga 720
gcgcggacat gtttaccagc acacggctct tcttcccacg gcactttctg atgtaacaat 780
cgagtgtgtg ttttcccaac tgcagctttt taatggttaa ccttcatcta attttttttc 840
tcccactggt ttatagatcc tctgacttgt gtgtgtttat agcttttgtt tcgcgggggtt 900
gtggtgagga aggggtgatg gcatgaggag ttctttatct tcagtgagaa tgtgcctgcc 960
cgcctgagag ccagcttccg cgttgagggc acgtgttcag agagctgctg agcgcacccc 1020
tctacccccc tgacagmcaa cacagacctg tgccgaaggc taatttggg cttttacgac 1080
cctaccccac ccctgtttt caggggttta gactacattt gaaatccaaa cttggagtat 1140
ataacttctt attgagccca actgcttttt tttttttttt ttttttgett ctctgccctt 1200
tttcatttc ttttgatatt gttttctgtg agagcncctga aatggcagcc ctggaatcta 1260
caatttggct ctccactgag caccttatct tgccacctta gccttaagaa tgaatatgaa 1320
gaaaaatata cagccacctc tgtccagggc agtaagaagg gctgcaagga aggggaggat 1380
ggggacaagg aaaggatcag atacctgctc cagtagtgtt gaggccactg tgtctcaggg 1440
gactccagga ggagcagaag agggatccca cgaagttatt cttacgcagc tggggccagg 1500
agggctcagag tggtgccagg tgcaagttag gctaaagaag ccaccactat tcctctctct 1560
tgcccattgt ggggggcaaa ggcatggctc accaagagtc ttgcaggggg acccacagat 1620
atgccatgtc cttcacacgt gcttgggctc cttaacctga aggcaaattg ctacttgcaa 1680
gactgactga cttcaaggaa tcagaaatta cctagaagca ccatgttttt tctatgacct 1740
tttcagctct tcaggtcatt ttaaggtcca ctgcaggggg ttagtgagaa agggatatact 1800
ttgtggtatg ttttgctttc ctaatagggg catgaaggaa acccagcaat ttgctgttat 1860
gtgaatggcc tgtagagcag agtcaagagc ggtgtgcttt gcccgactgc tcccatcagg 1920
aataggagag tagacagaga tcttccacat cccaggcttc tgctgctgct ttaaaagctc 1980
tgtccttgga gcctcccgtc ccctgaagtg tctcgcccc tgacacagcac tggcctttcg 2040
gaagcatccc agtaggggtt tctgaggctc gctgggtgact catgccctaa ttgcaatcct 2100
ctgcttttat cttgactttg aaggatctaa cactgctctc tcttccaaag gggaaaaaaaa 2160
gattcatttg ttttgagcaa taaactaata caaatg 2197
```

```
<210> 29
<211> 6331
<212> ADN
5 <213> Homo sapiens
```

```
<300>
<308> GenBank/AF006515
<309> 27-11-1997
<313> (1)..(6331)
```

```
10 <400> 29
```

ES 2 511 694 T3

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| atctttgttt | gggtctccca | tactgctgat | agatgaatgg | gtcaggatat | ctggaacaaa | 60 |
| atatggaggt | gaagggtag | atcggaac | aaaggtatg | gccccctagt | tcccaaaggg | 120 |
| agcagggaga | tgggaataga | attgaaggta | ggttttaggc | tacttgggag | gaggaatatt | 180 |
| taggtaattg | tggagacttt | ctcctgtgtg | atgaaggcgg | cagacactgt | gatcctgtgg | 240 |
| gcaagaagta | aaaatgacca | gctgaggatt | tcttttcctc | caggactgtg | ttggggtag | 300 |
| aggatgcctg | ataaggatga | cattcggtg | ctgccgtcag | cattgggtgt | gaagaagaga | 360 |
| aaacgaggac | ccaagaagca | gaaggagaac | aagccaggaa | aaccccgaaa | acgcaagaag | 420 |
| cgtgacagtg | aggaggaatt | tggttctgag | cgagatgagt | accgggagaa | gtcagagagt | 480 |
| gggggcagtg | aatatggaac | cggaccgggt | cggaaacgaa | gaaggaagca | ccgagaaaaa | 540 |
| aaggagaaga | agacaaagcg | gcgaaaaag | ggggaggag | atggggggca | aaagcaagtg | 600 |
| gaacagaagt | catcagcaac | tctgcttctg | acctggggcc | tggaggatgt | ggagcatgtg | 660 |
| ttctctgagg | aggattacca | caogctcacc | aactacaaag | ccttcagcca | gttcatgagg | 720 |
| cccctaattg | ctaagaagaa | tcctaagatc | ccaatgtcta | agatgatgac | catccttggg | 780 |
| gccaaatgga | gagagttcag | tgccaacaac | cccttcaagg | ggtcagcagc | tgctgtggcg | 840 |
| gcggcagcgg | cagcagcagc | agcagctgta | gctgagcagg | tgtcagctgc | tgtctcgtcg | 900 |
| gccaccccca | tagcaccctc | cggacccccc | gcccttcac | cacccctgc | tgctgatatc | 960 |
| cagccccac | ccatccgaag | agccaaaacc | aaagagggca | aaggtccagg | ccataagagg | 1020 |
| cggagtaaga | gccccgagt | gcctgatgga | cgcaagaagc | ttcggggaaa | gaaaatggca | 1080 |
| ccactcaaaa | taaaactagg | gcttctgggt | ggcaagagga | agaaaggagg | ctcgtatggt | 1140 |
| tttcagagcg | acgaaggctc | tgaaccagag | gctgaggaat | cagacctgga | cagtggcagt | 1200 |
| gtccacagtg | cctcaggccg | gcctgatggc | cctgtccgca | ccaagaaact | aaagagaggc | 1260 |
| cggccaggaa | ggaagaagaa | gaaggtcctg | ggctgtcctg | cagtggccgg | ggaggaggag | 1320 |
| gttgatggct | acgagacgga | tcaccaggat | tactgtgagg | tgtgccagca | gggtggggaa | 1380 |
| attattctgt | gtgacacctg | ccctcgtgcc | taccacctcg | tctgccttga | tcctgagctt | 1440 |
| gaccgggctc | cagagggcaa | atggagctgc | cctcactgtg | agaaggaggg | ggtccagtgg | 1500 |
| gaggccaagg | aggaagaaga | agaatacga | gaggaggag | aggaagaag | ggagaaggag | 1560 |

ES 2 511 694 T3

| | | | | | | |
|-------------|-------------|------------|------------|-------------|-------------|------|
| gaggaggatg | atcacatgga | gtactgccgc | gtatgcaagg | acggcgggga | gctcctgtgc | 1620 |
| tgtgacgcgt | gcatctcctc | ctaccacatt | cattgtctaa | accctcccct | gcctgacatt | 1680 |
| cccaatgggtg | aatggctgtg | tccccgatgc | acatgccccc | tgctgaaggg | tcgagtgca | 1740 |
| aagatcctac | attggcggtg | gggggagcca | cctgtagcag | tgccagcccc | tcaacaggca | 1800 |
| gatggaatc | agatgtccc | accccccg | cctcttcaag | gcagatcaga | gcgagagttc | 1860 |
| tttgcgaagt | gggtaggact | atcctactgg | cactgctcct | gggccaaagga | gcttcagctg | 1920 |
| gaaatcttcc | atthggttat | gtatcgaaac | taccagcgga | agaatgacat | ggatgagccc | 1980 |
| ccacccttgg | actatggctc | cgccgaggat | gatgggaaga | gcgacaagcg | taaagtgaaa | 2040 |
| gaccgcgact | atgctgagat | ggaggagaag | tactatcggt | ttggcatcaa | gccagagtg | 2100 |
| atgaccgtcc | accgcatcat | caaccacagt | gtggataaaa | aggggaatta | ccactatcta | 2160 |
| gtaaaatgga | gggacttacc | atatgaccag | tccaogtggg | aggaagatga | aatgaatatc | 2220 |
| cctgaataacg | aagaacataa | gcaaagctac | tggagacacc | gagaactaat | tatgggggaa | 2280 |
| gaccctgccc | agccccgcaa | gtataagaag | aagaagaagg | agctacaggg | tgatgggcct | 2340 |
| cccagttctc | ccactaatga | tctaccctg | aaatatgaga | ctcagccaag | gtttatcaca | 2400 |
| gccactggag | gcaccctgca | catgtatcag | ttggaagggc | tgaactggct | acgcttctcc | 2460 |
| tgggcccagg | gcaactgac | cattctagct | gatgagatgg | ggctaggcaa | gaccatacaa | 2520 |
| accatcgtct | tctctactc | actctacaag | gagggccaca | caaaaggtcc | cttctggtg | 2580 |
| agtgccccac | tctctaccat | cattaactgg | gagcgggagt | tccagatgtg | ggcaccctaa | 2640 |
| ttctatgtgg | tgacatacac | gggtgacaag | gacagccggg | ccatcattcg | tgagaatgaa | 2700 |
| ttctcctttg | aggacaatgc | catcaaagg | ggcaagaaag | cttttaagat | gaaaagggag | 2760 |
| gcacaggtga | agttccatgt | tctcctgaca | tcgtatgagc | tgatcacat | tgatcaggca | 2820 |
| gcacttgggt | ccatccgctg | ggcctgtcct | gtgtagatg | agggccatcg | actcaagaac | 2880 |
| aaccagtgcca | agtttttcag | ggttctcaat | ggttacaaga | tagatcataa | gttgctgctg | 2940 |
| acaggaaccc | cattgcagaa | taatctggag | gagctcttcc | atctcctgaa | cttctcacc | 3000 |
| ccagagagat | taacaactt | ggagggcttc | ctggaggagt | ttgctgacat | atccaaagag | 3060 |
| gaccagatca | agaaactgca | tgatttctg | gggccacaca | tgctgcggag | actcaaggca | 3120 |
| gatgtcttta | agaactatgc | agccaagaca | gagctcatcg | ttcgggtgga | gctaagcccc | 3180 |
| atgcagaaga | aatactacaa | atacatcctg | actcgaaatt | ttgaggcctt | gaattcacga | 3240 |
| gggtgtggga | accaggtgtc | gctgcttaat | atcatgatgg | atcttaagaa | gtgctgcaac | 3300 |
| catccatacc | tttttcccg | ggctgctatg | gagtccccca | aactccccag | tggggcttat | 3360 |
| gaggttgggg | cacttattaa | gtcgtctggg | aagctcatgc | tgctccagaa | gatgctgcga | 3420 |
| aagctgaagg | agcaaggaca | ccgagtgctc | atcttctcgc | agatgaccaa | aatgttagac | 3480 |
| ttgctbgagg | acttctttaga | ctatgaaggc | tacaagtatg | agcgcacoga | tggtgggtatc | 3540 |
| acgggtgccc | tgaggcagga | ggccatcgat | cggtttaatg | ctcctggggc | ccaacaattc | 3600 |
| tgcttctctc | tgtccaccgg | agctgggggc | ctgggcatca | atctggccac | tgctgacact | 3660 |
| gtcatcatct | tgattctgca | ctggaacccc | cataatgaca | tccaggcctt | tagccgggct | 3720 |
| catcggattg | gccaggccaa | caaagtgatg | atltaccggt | ttgtgactcg | cgctcagtg | 3780 |
| gaagagcgaa | tcacacaagt | ggccaagaga | aagatgatgc | tgacacacct | ggtgtgctcg | 3840 |
| cctgggctgg | gctccaaggc | aggctccatg | tccaagcagg | agcttgacga | cattctcaaa | 3900 |
| tttggcactg | aagagctatt | caaggatgaa | aacgaggggg | agaacaagga | ggaggacagc | 3960 |
| agtgtgattc | attatgacaa | tgaggccatc | gctcggctgt | tggaaccgga | ccaggatgca | 4020 |
| actgaggaca | ctgacgtgca | gaacatgaat | gagtatctca | gctcctcaa | ggtggcacag | 4080 |
| tacgtcgtgc | gggaagaaga | caagattgag | gaaattgagc | gagagatcat | caagcaggag | 4140 |
| gagaatgtgg | accctgacta | ctgggagaag | ctgctgaggc | atcactatga | gcaacagcag | 4200 |
| gaagacctag | cccggaatct | aggcaagggc | aagcgggttc | gcaagcaagt | taactacaat | 4260 |
| gatgctgctc | aggaagacca | agacaaccag | tcagagtact | cggtgggttc | agaggaggag | 4320 |
| gatgaagact | tcgatgaacg | tctgaaggg | cgtagacagt | caaagaggca | gctccggaat | 4380 |
| gagaaagata | agccactgcc | tccactgctg | gcccagctcg | ggggcaacat | tgaggtgctg | 4440 |
| ggcttcaaca | cccgtcagcg | gaaggctttc | ctcaatgctg | tgatgcgctg | gggatgcca | 4500 |
| ccacaggatg | ccttaccacc | acagtggctg | gtgcgggacc | tgaggggcaa | gactgagaag | 4560 |
| gagtttaagg | cctatgtgtc | ttgtttcatg | cgccatctgt | gtgagcctgg | ggcagacggc | 4620 |
| tctgaaacct | ttgccgatgg | ggtccctcgg | gagggactga | gtcgcagca | ggtgttgacc | 4680 |
| cgcattggag | tcatgtctct | cgtcaaaaag | aaggtgcagg | agtttgagca | catcaatggg | 4740 |
| cgttgttcaa | tgccggaact | gatgcctgac | cccagcggcg | attctaagcg | ctcctccaga | 4800 |
| gcctcctctc | ctacccaaac | gtctcccacc | actcctgagg | cttctgctac | caacagctcc | 4860 |
| tgacactcta | aacctgctac | tccagctcca | agtgagaag | gagaaggcat | aaggacacct | 4920 |
| cttgagaagg | aggaagctga | aaaccaggag | gaaaagccag | agaagaacag | cagaattggg | 4980 |
| gagaagatgg | agacagagcc | tgatgcccc | agcccagccc | catcacttgg | ggagcggctg | 5040 |
| gagccaagga | agattcctct | agaggatgag | gtcccagggg | tgccctggaga | gatggagcct | 5100 |
| gaacctgggt | accgtgggga | cagagagaag | tcagaagatg | taaaaggtga | ccgggagctt | 5160 |
| cgaccagggc | ctcgagatga | gccacggctc | aatgggagac | gagaggaaaa | gacagagaag | 5220 |
| ccccggttca | tgttcaatat | cgccgatgg | ggcttcacag | agcttcacac | actgtggcag | 5280 |
| aatgaggaac | ggcgagctat | ttcctcgggg | aaactcaatg | agatctggca | cagaagacat | 5340 |
| gactattggc | ttctggctgg | gattgtctct | catggctatg | cacgggtggca | ggacatccag | 5400 |
| aatgatgctc | aatttgccat | tatcaacgag | ccatttaaaa | ctgaagccaa | taaggggaa | 5460 |
| tttctgggga | tgaaaaataa | gttctgggcc | cggaggttca | agctcctgga | gcaggcgtg | 5520 |
| gtgattgagg | agcagctgcg | gcggggggcc | tacctgaacc | tgctgcagga | gccggcgcac | 5580 |

ES 2 511 694 T3

```

cccgccatgg ccctccacgc ccgcttcgcc gaggccgagt gcctggccga gagccaccag 5640
cacctctcca aggagtcgct ggcgggggaa aagccggcca acgccgtcct gcacaagggt 5700
ctgaaccagc tggaggagtt gctgagcgac atgaaggcgg acgtgaccgg cctgccagcc 5760
acgctgtccc gaataccccc catcgcagcc cgccttcaga tgtccgagcg cagcatcctc 5820
agccggctgg ccagcaaggg cacggagcct caccacacac cggcctaccg gccgggtccc 5880
tacgctacac ctccggggta cggggcggcc ttcagcgccg caccgtagg ggccctggcc 5940
gccgcaggcg ccaattacag ccagatgcct gcagggtcct tcatcacagc cgccaccaac 6000
ggcccaattc accgattttt taaaaaagtt ccagaaatcc agtgacgaat gtggtataca 6060
aaaaaatata taaattcttt caacttagaa taattaagtc ataaaataca tagggtacaa 6120
ataccacatt ccgttctaaa atgatattctt aggatcatca aaagaaaaag aggatttggg 6180
ttatgcaaaa aatgattcct atatatataa tcaattatct aactgacatt tttgcaaatc 6240
taccacaact tcgcctttta ttgcatatgc taaacaagca gatgctaagt ctgtaaactg 6300
tgaattaacc tcctttttta ttaattgttc g 6331

```

```

<210> 30
<211> 6331
<212> ADN

```

5 <213> Homo sapiens

```

<300>
<308> GenBank/NM_001272
<309> 03-02-2001
<313> (1)..(6331)

```

10 <400> 30

ES 2 511 694 T3

| | | | | | | |
|-------------|-------------|------------|------------|-------------|-------------|------|
| atctttgttt | gggtctccca | tactgcgtat | agatgaatgg | gtcaggatat | ctggaacaaa | 60 |
| atatggagggt | gaagggtgag | atcgggaaac | aaagggtatg | gccccctagt | tcccaaaggg | 120 |
| agcaggggaga | tgggaataga | attgaaggta | ggttttaggc | tacttgggag | gaggaatatt | 180 |
| taggtaattg | tggagacttt | ctcctgtgtg | atgaaggcgg | cagacactgt | gatcctgtgg | 240 |
| gcaagaagta | aaaatgacca | gctgaggatt | tcttttcctc | caggactgtg | ttgggggtgac | 300 |
| aggatgcctg | ataaggatga | cattcggctg | ctgccgtcag | cattgggtgt | gaagaagaga | 360 |
| aaacgaggac | ccaagaagca | gaaggagaac | aagccaggaa | aaccccgaaa | acgcaagaag | 420 |
| cgtgacagtg | aggaggaatt | tggttctgag | cgagatgagt | accggggagaa | gtcagagagt | 480 |
| gggggcagtg | aatatggaac | cggaccgggt | cgaaacgaa | gaaggaaagca | ccgagaaaaa | 540 |
| aaggagaaga | agacaaagcg | gcggaaaaag | ggggaggag | atggggggca | aaagcaagtg | 600 |
| gaacagaagt | catcagcaac | tctgcttctg | acctggggcc | tggaggatgt | ggagcatgtg | 660 |
| ttctctgagg | aggattacca | cacgctcacc | aactacaaag | ccttcagcca | gttcatgagg | 720 |
| cccctaattg | ctaagaagaa | tcctaagatc | ccaatgtcta | agatgatgac | catccttggg | 780 |
| gccaaatgga | gagagtccag | tgccaacaac | cccttcaagg | ggtcagcagc | tgctgtggcg | 840 |
| gcggcagcgg | cagcagcagc | agcagctgta | gctgagcagg | tgtcagctgc | tgctctogtcg | 900 |
| gccaccccca | tagcaccctc | cggaccccc | gccttccac | cacccctgc | tgctgatatc | 960 |
| cagccccac | ccatccgaag | agccaaaacc | aaaggggca | aaggtccagg | ccataagagg | 1020 |
| cggagttaaga | gcccccagtg | gcttgatgga | cgcaagaagc | ttcggggaaa | gaaaatggca | 1080 |
| ccactcaaaa | taaaactag | gcttctgggt | ggcaagagga | agaaaggagg | ctcgtatggt | 1140 |
| tttcagagcg | acgaaggtcc | tgaaccagag | gctgaggaat | cagacctgga | cagtggcagt | 1200 |
| gtccacagtg | cctcaggccg | gcctgatggc | cctgtccgca | ccaagaaact | aaagagaggc | 1260 |
| cggccaggaa | ggaagaagaa | gaaggtcctg | ggctgtcctg | cagtggccgg | ggaggaggag | 1320 |
| gttgatggct | acgagacgga | tcaccagat | tactgtgagg | tgtgccagca | gggtggggaa | 1380 |
| attattctgt | gtgacacctg | ccctcgtgcc | taccacctcg | tctgccttga | tcctgagctt | 1440 |
| gaccgggctc | cagagggcaa | atggagctgc | cctcactgtg | agaaggagg | ggtccagtgg | 1500 |
| gaggccaagg | aggaagaaga | agaatacga | gaggaggag | aggaagaagg | ggagaaggag | 1560 |
| gaggagatg | atcacatgga | gtactgccgc | gtatgcaagg | acggcgggga | gctcctgtgc | 1620 |
| tgtgacgct | gcattctcctc | ctaccacatt | cattgtctaa | accctcccct | gcctgacatt | 1680 |
| cccaatgggt | aatggctgtg | tcccgatgc | acatgcccc | tgctgaagg | tcgagtgcag | 1740 |
| aagatcctac | attggcgggtg | gggggagcca | cctgtagcag | tgccagcccc | tcaacaggca | 1800 |
| gatggaatc | cagatgtccc | acccccctg | cctcttcaag | gcagatcaga | gagagagttc | 1860 |
| tttgtaagt | ggtaggact | atcctactgg | cactgctcct | gggccaagga | gcttcagctg | 1920 |
| gaaatcttc | atttggttat | gtatcgaac | taccagcgg | agaatgacat | ggatgagccc | 1980 |
| ccaccctgg | actatggctc | cggcgaggat | gatgggaaga | gagacaagcg | taaagtgaaa | 2040 |
| gaccgcact | atgctgagat | ggaggagaag | tactatcgtt | ttggcatcaa | gccagagtgg | 2100 |
| atgaccgtcc | accgcatcat | caaccacagt | gtggataaaa | aggggaatta | ccactatcta | 2160 |
| gtaaaatgga | gggacttacc | atatgaccag | tccacgtggg | aggaagatga | aatgaatatac | 2220 |
| cctgaatacg | aagaacataa | gcaaagctac | tgagacacc | gagaactaat | tatgggggaa | 2280 |
| gaccctgccc | agccccgcaa | gtataagaag | aagaagaagg | agctacaggg | tgatgggcct | 2340 |
| cccagtctc | ccactaatga | tcctaccgtg | aaatatgaga | ctcagccacg | gtttatcaca | 2400 |
| gccactggag | gcaccctgca | catgtatcag | ttggaagggc | tgaactggct | acgcttctcc | 2460 |
| tgggccag | gactgacac | cattctagct | gatgagatgg | ggctaggcaa | gaccatacaa | 2520 |
| accatcgtct | tcctctactc | actctacaag | gagggccaca | caaaaggctc | cttctgggtg | 2580 |
| agtgcccac | tctctacat | cattaactgg | gagcgggagt | tccagatgtg | ggcaccctaaa | 2640 |

ES 2 511 694 T3

ttctatgtgg tgacatacac ggggtgacaag gacagccggg ccatcattcg tgagaatgaa 2700
 ttctcctttg aggacaatgc catcaaaggg ggcaagaaag cttttaagat gaaaggggag 2760
 gcacaggtga agtcccatgt tctcctgaca tCGtatgagc tgatcaccat tgatcaggca 2820
 gcacttggtt ccatccgctg ggcctgtctt gtggtagatg aggccatcg actcaagaac 2880
 aaccagtcca agtTTTTcag ggttctcaat ggttacaaga tagatcataa gttgctgctg 2940
 acaggaaccc cattgcagaa taatctggag gagctcttcc atctcctgaa ctctctcacc 3000
 ccagagagat ttaacaactt ggagggcttc ctggaggagt ttgctgacat atccaaagag 3060
 gaccagatca agaaactgca tgatttgctg gggccacaca tgctgcggag actcaaggca 3120
 gatgtcttta agaacatgcc agccaagaca gagctcatcg ttcgggtgga gctaagcccc 3180
 atgcagaaga aatactacaa atacatcctg actcgaaatt ttgaggcctt gaattcacga 3240
 ggtggtggga accaggtgtc gctgcttaat atcatgatgg atcttaagaa gtgctgcaac 3300
 catccaatac ttttcccggt ggctgctatg gagtccccca aactccccag tggggcttat 3360
 gaggggtggg cacttattaa gtctgtctgg aagctcatgc tgctccagaa gatgctgca 3420
 aagctgaagg agcaaggaca ccgagtgtc atcttctcgc agatgaccaa aatgtagac 3480
 ttgcttgagg acttcttaga ctatgaaggc tacaagtatg agcgcacgca tgggtggtatc 3540
 acgggtgccc tgaggcagga ggccatcgat cggtttaatg ctctggggc ccaacaattc 3600
 tgcttctctc tgtccaccgc agctgggggc ctgggcatca atctggccac tgctgacact 3660
 gtcatcatct ttgattctga ctggaacccc cataatgaca tccaggcctt tagccgggct 3720
 catcggattg gccaggccaa caaagtgatg atttaccggt ttgtgactcg cgcgtcagt 3780
 gaagagcgaa tcacacaagt ggccaagaga aagatgatgc tgacacacct ggttgtgctg 3840
 cctgggctgg gctccaaggc aggtccatg tccaagcagg agcttgacga cattctcaaa 3900
 tttggcactg aagagctatt caaggatgaa aacgaggggg agaacaagga ggaggacagc 3960
 agtgtgattc attatgacaa tgaggccatc gctcggctgt tggaccggaa ccaggatgca 4020
 actgaggaca ctgacgtgca gaacatgaat gagtatctca gctccttcaa ggtggcagag 4080
 tacgtcgtgc ggaagaaga caagattgag gaaattgagc gagagatcat caagcaggag 4140
 gagaatgtgg accctgacta ctgggagaag ctgctgaggc atcactatga gcaacagcag 4200
 gaagacctag cccggaatct aggcaagggc aagcgggttc gcaagcaagt taactacaat 4260
 gatgtgctc aggaagacca agacaaccag tcagagtact cgggtgggttc agaggaggag 4320
 gagaaagata agcactgccc tccactgctg gcccagctcg ggggcaacat tgagggtgctg 4380
 ggcttcaaca cccgtcagcg gaaggctttc ctcaatgctg tgatgcgctg ggggatgcca 4440
 ccacaggatg ccttcaccac acagtggctg gtgctggacc tgaggggcaa gactgagaag 4500
 gagtttaagg cctatgtgtc tttgttcatt cgccatctgt gtgagcctgg ggcagacggc 4560
 tctgaaacct ttgccgatgg ggtccctcgg gagggactga gtcgccagca ggtgttgacc 4620
 cgcatgggag taattgctct cgtcaaaaag aaggtgcagg agtttgagca catcaatggg 4680
 cgttgggtcaa tgccggaact gatgcctgac cccagcgcg attctaagcg ctctccaga 4740
 gcctcctctc ctacaaaac gtctcccacc actcctgagg cttctgctac caacagtccc 4800
 tgcacctcta aacctgctac tccagctcca agtgagaaag gagaaggcat aaggacacct 4860
 cttgagaagg aggaagctga aaaccaggag gaaaagccag agaagaacag cagaattggg 4920
 gagaagatgg agacagagggc tgatgcccc agcccagccc catcacttgg ggagcggctg 4980
 gagccaagga agagatgag agaggatgag gtgccagggg tgctcgagaga gatggagcct 5040
 gaacctgggt accgtgggga cagagagaag tcagaagatg taaaagggtga ccgggagctt 5100
 cgaccagggc ctcgagatga gccacggctc aatgggagc gagaggaaaa gacagagaag 5160
 ccccggttca tgttcaatat cgccgatggt ggcttcacag agcttcacac actgtggcag 5220
 aatgaggaac gggcagctat ttcctcgggg aaactcaatg agatctggca cagaagacat 5280
 gactattggc ttctggcttg gattgtcctc catggctatg cacggtggca ggacatccag 5340
 aatgatgctc aatttgccat tatcaacgag ccatttmeta ctgaagccaa taaggggaac 5400
 tttctggaga tgaataataa gttcctggcc cggagggttca agctcctgga gcaggcgtg 5460
 gtgattgagg agcagctgcg cggggcggcc tacctgaacc tgtcgcagga gccggcgcac 5520
 cccgccatgg cctccacgc ccgctcgcg gagggcagat gcctggccga gagccaccag 5580
 cacctctcca aggagtcgct ggccgggaac aagccggcca acgcccctc gcacaaggtt 5640
 ctgaaccagc tggaggagtt gctgagcgac atgaaggcgg acgtgaccg cctgccagcc 5700
 acgctgtccc gaataccccc catcgagcc cgccttcaga tgtccgagcg cagcatcctc 5760
 agccggctgg ccagcaaggg cacggagcct caccacacac cggcctacc gccgggtccc 5820
 tacgctacac ctccggggta cggggcggcc ttcagcgccg caccctagg gccctggcc 5880
 gccgcaggcg ccaattacag ccagatgcct gcagggtcct tcatcacagc cgccaccaac 5940
 ggcccaattc accgattttt taaaaaagtt ccagaaatcc agtgacgaat gtggtataca 6000
 aaaaaatata taattcttt caacttagaa taataagtc ataaaataca taggtataca 6060
 ataccacatt ccgttctaaa atgatatctt aggatcatca aaagaaaaag aggatttgg 6120
 ttatgcaaaa aatgattcct atatatataa tcaattatct aactgacatt tttgcaaatc 6180
 taccacaact tcgcttttta ttgcatatgc taacaagca gatgctaagt ctgtaactg 6240
 tgaattaacc tcctttttaa ttaattgttc g 6300
 6331

- <210> 31
- <211> 560
- <212> PRT
- <213> Homo sapiens

ES 2 511 694 T3

<220>
 <221> característica_miscelánea
 <223> CDS de la seq ID 27, n° de acceso GenBank/XM_040709:910..2592

5 <220>
 <221> característica_miscelánea
 <223> nombre: regulador negativo del receptor de la prostaglandina F2 (PTG FRN)

<400> 31

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

ES 2 511 694 T3

Arg Val Asp Gly Val Val Leu Glu Lys Val Gln Glu Asp Glu Phe Arg
 305 310 315 320

Tyr Arg Met Tyr Gln Thr Gln Val Ser Asp Ala Gly Leu Tyr Arg Cys
 325 330 335

Met Val Thr Ala Trp Ser Pro Val Arg Gly Ser Leu Trp Arg Glu Ala
 340 345 350

Ala Thr Ser Leu Ser Asn Pro Ile Glu Ile Asp Phe Gln Thr Ser Gly
 355 360 365

Pro Ile Phe Asn Ala Ser Val His Ser Asp Thr Pro Ser Val Ile Arg
 370 375 380

Gly Asp Leu Ile Lys Leu Phe Cys Ile Ile Thr Val Glu Gly Ala Ala
 385 390 395 400

Leu Asp Pro Asp Asp Met Ala Phe Asp Val Ser Trp Phe Ala Val His
 405 410 415

Ser Phe Gly Leu Asp Lys Ala Pro Val Leu Leu Ser Ser Leu Asp Arg
 420 425 430

Lys Gly Ile Val Thr Thr Ser Arg Arg Asp Trp Lys Ser Asp Leu Ser
 435 440 445

Leu Glu Arg Val Ser Val Leu Glu Phe Leu Leu Gln Val His Gly Ser
 450 455 460

Glu Asp Gln Asp Phe Gly Asn Tyr Tyr Cys Ser Val Thr Pro Trp Val
 465 470 475 480

Lys Ser Pro Thr Gly Ser Trp Gln Lys Glu Ala Glu Ile His Ser Lys
 485 490 495

Pro Val Phe Ile Thr Val Lys Met Asp Val Leu Asn Ala Phe Lys Tyr
 500 505 510

Pro Leu Leu Ile Gly Val Gly Leu Ser Thr Val Ile Gly Leu Leu Ser
 515 520 525

Cys Leu Ile Gly Tyr Cys Ser Ser His Trp Cys Cys Lys Lys Glu Val
 530 535 540

Gln Glu Thr Arg Arg Glu Arg Arg Arg Leu Met Ser Met Glu Met Asp
 545 550 555 560

<210> 32
 <211> 186
 <212> PRT
 5 <213> Homo sapiens

<220>
 <221> característica_misclánea
 <223> CDS parcial de la seq ID 28, nº de acceso GenBank/AB014734: <1.
 .561

10 <220>
 <221> característica_misclánea
 <223> nombre: proteína 6 asociada con la célula del músculo liso humano (SM AP-6)

<400> 32

Val His Ser Asp Thr Pro Ser Val Ile Arg Gly Asp Leu Ile Lys Leu
 1 5 10 15

ES 2 511 694 T3

Phe Cys Ile Ile Thr Val Glu Gly Ala Ala Leu Asp Pro Asp Asp Met
 20 25 30
 Ala Phe Asp Val Ser Trp Phe Ala Val His Ser Phe Gly Leu Asp Lys
 35 40 45
 Ala Pro Val Leu Leu Ser Ser Leu Asp Arg Lys Gly Ile Val Thr Thr
 50 55 60
 Ser Arg Arg Asp Trp Lys Ser Asp Leu Ser Leu Glu Arg Val Ser Val
 65 70 75 80
 Leu Glu Phe Leu Leu Gln Val His Gly Ser Glu Asp Gln Asp Phe Gly
 85 90 95
 Asn Tyr Tyr Cys Ser Val Thr Pro Trp Val Lys Ser Pro Thr Gly Ser
 100 105 110
 Trp Gln Lys Glu Ala Glu Ile His Ser Lys Pro Val Phe Ile Thr Val
 115 120 125
 Lys Met Asp Val Leu Asn Ala Phe Lys Tyr Pro Leu Leu Ile Gly Val
 130 135 140
 Gly Leu Ser Thr Val Ile Gly Leu Leu Ser Cys Leu Ile Gly Tyr Cys
 145 150 155 160
 Ser Ser His Trp Cys Cys Lys Lys Glu Val Gln Glu Thr Arg Arg Glu
 165 170 175
 Arg Arg Arg Leu Met Ser Met Glu Met Asp
 180 185

<210> 33
 <211> 1944
 <212> PRT
 5 <213> Homo sapiens

<220>
 <221> característica_misclánea
 <223> CDS de la seq ID 29, n° de acceso GenBank/AF006515: 211..6045

10 <220>
 <221> característica_misclánea
 <223> nombre: proteína CHD3

<400> 33

ES 2 511 694 T3

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

ES 2 511 694 T3

Lys Lys Thr Lys Arg Arg Lys Lys Gly Glu Gly Asp Gly Gly Gln Lys
115 120 125

Gln Val Glu Gln Lys Ser Ser Ala Thr Leu Leu Leu Thr Trp Gly Leu
130 135 140

Glu Asp Val Glu His Val Phe Ser Glu Glu Asp Tyr His Thr Leu Thr
145 150 155 160

Asn Tyr Lys Ala Phe Ser Gln Phe Met Arg Pro Leu Ile Ala Lys Lys
165 170 175

Asn Pro Lys Ile Pro Met Ser Lys Met Met Thr Ile Leu Gly Ala Lys
180 185 190

Trp Arg Glu Phe Ser Ala Asn Asn Pro Phe Lys Gly Ser Ala Ala Ala
195 200 205

Val Ala Ala Ala Ala Ala Ala Ala Ala Ala Val Ala Glu Gln Val
210 215 220

Ser Ala Ala Val Ser Ser Ala Thr Pro Ile Ala Pro Ser Gly Pro Pro
225 230 235 240

Ala Leu Pro Pro Pro Pro Ala Ala Asp Ile Gln Pro Pro Pro Ile Arg
245 250 255

Arg Ala Lys Thr Lys Glu Gly Lys Gly Pro Gly His Lys Arg Arg Ser
260 265 270

Lys Ser Pro Arg Val Pro Asp Gly Arg Lys Lys Leu Arg Gly Lys Lys
275 280 285

Met Ala Pro Leu Lys Ile Lys Leu Gly Leu Leu Gly Gly Lys Arg Lys
290 295 300

Lys Gly Gly Ser Tyr Val Phe Gln Ser Asp Glu Gly Pro Glu Pro Glu
305 310 315 320

Ala Glu Glu Ser Asp Leu Asp Ser Gly Ser Val His Ser Ala Ser Gly
325 330 335

Arg Pro Asp Gly Pro Val Arg Thr Lys Lys Leu Lys Arg Gly Arg Pro
340 345 350

Gly Arg Lys Lys Lys Lys Val Leu Gly Cys Pro Ala Val Ala Gly Glu
355 360 365

Glu Glu Val Asp Gly Tyr Glu Thr Asp His Gln Asp Tyr Cys Glu Val
370 375 380

Cys Gln Gln Gly Gly Glu Ile Ile Leu Cys Asp Thr Cys Pro Arg Ala
385 390 395 400

Tyr His Leu Val Cys Leu Asp Pro Glu Leu Asp Arg Ala Pro Glu Gly
405 410 415

Lys Trp Ser Cys Pro His Cys Glu Lys Glu Gly Val Gln Trp Glu Ala
420 425 430

Lys Glu Glu Glu Glu Tyr Glu Glu Glu Gly Glu Glu Glu Gly Glu
435 440 445

Lys Glu Glu Glu Asp Asp His Met Glu Tyr Cys Arg Val Cys Lys Asp
450 455 460

Gly Gly Glu Leu Leu Cys Cys Asp Ala Cys Ile Ser Ser Tyr His Ile

ES 2 511 694 T3

Phe Glu Asp Asn Ala Ile Lys Gly Gly Lys Lys Ala Phe Lys Met Lys
 835 840 845
 Arg Glu Ala Gln Val Lys Phe His Val Leu Leu Thr Ser Tyr Glu Leu
 850 855 860
 Ile Thr Ile Asp Gln Ala Ala Leu Gly Ser Ile Arg Trp Ala Cys Leu
 865 870 875 880
 Val Val Asp Glu Ala His Arg Leu Lys Asn Asn Gln Ser Lys Phe Phe
 885 890 895
 Arg Val Leu Asn Gly Tyr Lys Ile Asp His Lys Leu Leu Leu Thr Gly
 900 905 910
 Thr Pro Leu Gln Asn Asn Leu Glu Glu Leu Phe His Leu Leu Asn Phe
 915 920 925
 Leu Thr Pro Glu Arg Phe Asn Asn Leu Glu Gly Phe Leu Glu Glu Phe
 930 935 940
 Ala Asp Ile Ser Lys Glu Asp Gln Ile Lys Lys Leu His Asp Leu Leu
 945 950 955 960
 Gly Pro His Met Leu Arg Arg Leu Lys Ala Asp Val Phe Lys Asn Met
 965 970 975
 Pro Ala Lys Thr Glu Leu Ile Val Arg Val Glu Leu Ser Pro Met Gln
 980 985 990
 Lys Lys Tyr Tyr Lys Tyr Ile Leu Thr Arg Asn Phe Glu Ala Leu A
 995 1000 1005
 Ser Arg Gly Gly Gly Asn Gln Val Ser Leu Leu Asn Ile Met Met
 1010 1015 1020
 Asp Leu Lys Lys Cys Cys Asn His Pro Tyr Leu Phe Pro Val Ala
 1025 1030 1035
 Ala Met Glu Ser Pro Lys Leu Pro Ser Gly Ala Tyr Glu Gly Gly
 1040 1045 1050
 Ala Leu Ile Lys Ser Ser Gly Lys Leu Met Leu Leu Gln Lys Met
 1055 1060 1065
 Leu Arg Lys Leu Lys Glu Gln Gly His Arg Val Leu Ile Phe Ser
 1070 1075 1080
 Gln Met Thr Lys Met Leu Asp Leu Leu Glu Asp Phe Leu Asp Tyr
 1085 1090 1095
 Glu Gly Tyr Lys Tyr Glu Arg Ile Asp Gly Gly Ile Thr Gly Ala
 1100 1105 1110
 Leu Arg Gln Glu Ala Ile Asp Arg Phe Asn Ala Pro Gly Ala Gln
 1115 1120 1125
 Gln Phe Cys Phe Leu Leu Ser Thr Arg Ala Gly Gly Leu Gly Ile
 1130 1135 1140
 Asn Leu Ala Thr Ala Asp Thr Val Ile Ile Phe Asp Ser Asp Trp
 1145 1150 1155
 Asn Pro His Asn Asp Ile Gln Ala Phe Ser Arg Ala His Arg Ile
 1160 1165 1170

ES 2 511 694 T3

Gly Gln Ala Asn Lys Val Met Ile Tyr Arg Phe Val Thr Arg Ala
 1175 1180 1185

Ser Val Glu Glu Arg Ile Thr Gln Val Ala Lys Arg Lys Met Met
 1190 1195 1200

Leu Thr His Leu Val Val Arg Pro Gly Leu Gly Ser Lys Ala Gly
 1205 1210 1215

Ser Met Ser Lys Gln Glu Leu Asp Asp Ile Leu Lys Phe Gly Thr
 1220 1225 1230

Glu Glu Leu Phe Lys Asp Glu Asn Glu Gly Glu Asn Lys Glu Glu
 1235 1240 1245

Asp Ser Ser Val Ile His Tyr Asp Asn Glu Ala Ile Ala Arg Leu
 1250 1255 1260

Leu Asp Arg Asn Gln Asp Ala Thr Glu Asp Thr Asp Val Gln Asn
 1265 1270 1275

Met Asn Glu Tyr Leu Ser Ser Phe Lys Val Ala Gln Tyr Val Val
 1280 1285 1290

Arg Glu Glu Asp Lys Ile Glu Glu Ile Glu Arg Glu Ile Ile Lys
 1295 1300 1305

Gln Glu Glu Asn Val Asp Pro Asp Tyr Trp Glu Lys Leu Leu Arg
 1310 1315 1320

His His Tyr Glu Gln Gln Gln Glu Asp Leu Ala Arg Asn Leu Gly
 1325 1330 1335

Lys Gly Lys Arg Val Arg Lys Gln Val Asn Tyr Asn Asp Ala Ala
 1340 1345 1350

Gln Glu Asp Gln Asp Asn Gln Ser Glu Tyr Ser Val Gly Ser Glu
 1355 1360 1365

Glu Glu Asp Glu Asp Phe Asp Glu Arg Pro Glu Gly Arg Arg Gln
 1370 1375 1380

Ser Lys Arg Gln Leu Arg Asn Glu Lys Asp Lys Pro Leu Pro Pro
 1385 1390 1395

Leu Leu Ala Arg Val Gly Gly Asn Ile Glu Val Leu Gly Phe Asn
 1400 1405 1410

Thr Arg Gln Arg Lys Ala Phe Leu Asn Ala Val Met Arg Trp Gly
 1415 1420 1425

Met Pro Pro Gln Asp Ala Phe Thr Thr Gln Trp Leu Val Arg Asp
 1430 1435 1440

Leu Arg Gly Lys Thr Glu Lys Glu Phe Lys Ala Tyr Val Ser Leu
 1445 1450 1455

Phe Met Arg His Leu Cys Glu Pro Gly Ala Asp Gly Ser Glu Thr
 1460 1465 1470

Phe Ala Asp Gly Val Pro Arg Glu Gly Leu Ser Arg Gln Gln Val
 1475 1480 1485

Leu Thr Arg Ile Gly Val Met Ser Leu Val Lys Lys Lys Val Gln
 1490 1495 1500

Glu Phe Glu His Ile Asn Gly Arg Trp Ser Met Pro Glu Leu Met

ES 2 511 694 T3

| | | | | |
|-----------------|-----------------|---------------------|-----------------|------|
| 1505 | | 1510 | | 1515 |
| Pro Asp 1520 | Pro Ser Ala Asp | Ser Lys Arg Ser Ser | Arg Ala Ser Ser | 1530 |
| Pro Thr 1535 | Lys Thr Ser Pro | Thr Thr Pro Glu Ala | Ser Ala Thr Asn | 1545 |
| Ser Pro 1550 | Cys Thr Ser Lys | Pro Ala Thr Pro Ala | Pro Ser Glu Lys | 1560 |
| Gly Glu 1565 | Gly Ile Arg Thr | Pro Leu Glu Lys Glu | Glu Ala Glu Asn | 1575 |
| Gln Glu 1580 | Gly Lys Pro Glu | Lys Asn Ser Arg Ile | Gly Glu Lys Met | 1590 |
| Glu Thr 1595 | Gly Ala Asp Ala | Pro Ser Pro Ala Pro | Ser Leu Gly Glu | 1605 |
| Arg Leu 1610 | Gly Pro Arg Lys | Ile Pro Leu Glu Asp | Glu Val Pro Gly | 1620 |
| Val Pro 1625 | Gly Glu Met Glu | Pro Glu Pro Gly Tyr | Arg Gly Asp Arg | 1635 |
| Glu Lys 1640 | Ser Glu Asp Val | Lys Gly Asp Arg Glu | Leu Arg Pro Gly | 1650 |
| Pro Arg 1655 | Asp Glu Pro Arg | Ser Asn Gly Arg Arg | Glu Glu Lys Thr | 1665 |
| Glu Lys 1670 | Pro Arg Phe Met | Phe Asn Ile Ala Asp | Gly Gly Phe Thr | 1680 |
| Glu Leu 1685 | His Thr Leu Trp | Gln Asn Glu Glu Arg | Ala Ala Ile Ser | 1695 |
| Ser Gly 1700 | Lys Leu Asn Glu | Ile Trp His Arg Arg | His Asp Tyr Trp | 1710 |
| Leu Leu 1715 | Ala Gly Ile Val | Leu His Gly Tyr Ala | Arg Trp Gln Asp | 1725 |
| Ile Gln 1730 | Asn Asp Ala Gln | Phe Ala Ile Ile Asn | Glu Pro Phe Lys | 1740 |
| Thr Glu 1745 | Ala Asn Lys Gly | Asn Phe Leu Glu Met | Lys Asn Lys Phe | 1755 |
| Leu Ala 1760 | Arg Arg Phe Lys | Leu Leu Glu Gln Ala | Leu Val Ile Glu | 1770 |
| Glu Gln 1775 | Leu Arg Arg Ala | Ala Tyr Leu Asn Leu | Ser Gln Glu Pro | 1785 |
| Ala His 1790 | Pro Ala Met Ala | Leu His Ala Arg Phe | Ala Glu Ala Glu | 1800 |
| Cys Leu 1805 | Ala Glu Ser His | Gln His Leu Ser Lys | Glu Ser Leu Ala | 1815 |
| Gly Asn 1820 | Lys Pro Ala Asn | Ala Val Leu His Lys | Val Leu Asn Gln | 1830 |
| Leu Glu 1835 | Gly Leu Leu Ser | Asp Met Lys Ala Asp | Val Thr Arg Leu | 1845 |

ES 2 511 694 T3

Pro Ala Thr Leu Ser Arg Ile Pro Pro Ile Ala Ala Arg Leu Gln
1850 1855 1860

Met Ser Glu Arg Ser Ile Leu Ser Arg Leu Ala Ser Lys Gly Thr
1865 1870 1875

Glu Pro His Pro Thr Pro Ala Tyr Pro Pro Gly Pro Tyr Ala Thr
1880 1885 1890

Pro Pro Gly Tyr Gly Ala Ala Phe Ser Ala Ala Pro Val Gly Ala
1895 1900 1905

Leu Ala Ala Ala Gly Ala Asn Tyr Ser Gln Met Pro Ala Gly Ser
1910 1915 1920

Phe Ile Thr Ala Ala Thr Asn Gly Pro Ile His Arg Phe Phe Lys
1925 1930 1935

Lys Val Pro Glu Ile Gln
1940

<210> 34
<211> 1944
<212> PRT

5 <213> Homo sapiens

<220>
<221> característica_misclánea
<223> CDS de la seq ID 30, nº de acceso GenBank/NM_001272: 211..6045

10 <220>
<221> característica_misclánea
<223> nombre: proteína 3 que une la helicasa de ADN con el cromodominio CHD3

<400> 34

Met Lys Ala Ala Asp Thr Val Ile Leu Trp Ala Arg Ser Lys Asn Asp
1 5 10 15

Gln Leu Arg Ile Ser Phe Pro Pro Gly Leu Cys Trp Gly Asp Arg Met
20 25 30

Pro Asp Lys Asp Asp Ile Arg Leu Leu Pro Ser Ala Leu Gly Val Lys
35 40 45

Lys Arg Lys Arg Gly Pro Lys Lys Gln Lys Glu Asn Lys Pro Gly Lys
50 55 60

Pro Arg Lys Arg Lys Lys Arg Asp Ser Glu Glu Glu Phe Gly Ser Glu
65 70 75 80

Arg Asp Glu Tyr Arg Glu Lys Ser Glu Ser Gly Gly Ser Glu Tyr Gly
85 90 95

Thr Gly Pro Gly Arg Lys Arg Arg Arg Lys His Arg Glu Lys Lys Glu
100 105 110

Lys Lys Thr Lys Arg Arg Lys Lys Gly Glu Gly Asp Gly Gly Gln Lys
115 120 125

Gln Val Glu Gln Lys Ser Ser Ala Thr Leu Leu Leu Thr Trp Gly Leu
130 135 140

Glu Asp Val Glu His Val Phe Ser Glu Glu Asp Tyr His Thr Leu Thr
145 150 155 160

Asn Tyr Lys Ala Phe Ser Gln Phe Met Arg Pro Leu Ile Ala Lys Lys
165 170 175

ES 2 511 694 T3

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

ES 2 511 694 T3

Gln Ala Asp Gly Asn Pro Asp Val Pro Pro Pro Arg Pro Leu Gln Gly
530 535 540

Arg Ser Glu Arg Glu Phe Phe Val Lys Trp Val Gly Leu Ser Tyr Trp
545 550 555 560

His Cys Ser Trp Ala Lys Glu Leu Gln Leu Glu Ile Phe His Leu Val
565 570 575

Met Tyr Arg Asn Tyr Gln Arg Lys Asn Asp Met Asp Glu Pro Pro Pro
580 585 590

Leu Asp Tyr Gly Ser Gly Glu Asp Asp Gly Lys Ser Asp Lys Arg Lys
595 600 605

Val Lys Asp Pro His Tyr Ala Glu Met Glu Glu Lys Tyr Tyr Arg Phe
610 615 620

Gly Ile Lys Pro Glu Trp Met Thr Val His Arg Ile Ile Asn His Ser
625 630 635 640

Val Asp Lys Lys Gly Asn Tyr His Tyr Leu Val Lys Trp Arg Asp Leu
645 650 655

Pro Tyr Asp Gln Ser Thr Trp Glu Glu Asp Glu Met Asn Ile Pro Glu
660 665 670

Tyr Glu Glu His Lys Gln Ser Tyr Trp Arg His Arg Glu Leu Ile Met
675 680 685

Gly Glu Asp Pro Ala Gln Pro Arg Lys Tyr Lys Lys Lys Lys Glu
690 695 700

Leu Gln Gly Asp Gly Pro Pro Ser Ser Pro Thr Asn Asp Pro Thr Val
705 710 715 720

Lys Tyr Glu Thr Gln Pro Arg Phe Ile Thr Ala Thr Gly Gly Thr Leu
725 730 735

His Met Tyr Gln Leu Glu Gly Leu Asn Trp Leu Arg Phe Ser Trp Ala
740 745 750

Gln Gly Thr Asp Thr Ile Leu Ala Asp Glu Met Gly Leu Gly Lys Thr
755 760 765

Ile Gln Thr Ile Val Phe Leu Tyr Ser Leu Tyr Lys Glu Gly His Thr
770 775 780

Lys Gly Pro Phe Leu Val Ser Ala Pro Leu Ser Thr Ile Ile Asn Trp
785 790 795 800

Glu Arg Glu Phe Gln Met Trp Ala Pro Lys Phe Tyr Val Val Thr Tyr
805 810 815

Thr Gly Asp Lys Asp Ser Arg Ala Ile Ile Arg Glu Asn Glu Phe Ser
820 825 830

Phe Glu Asp Asn Ala Ile Lys Gly Gly Lys Lys Ala Phe Lys Met Lys
835 840 845

Arg Glu Ala Gln Val Lys Phe His Val Leu Leu Thr Ser Tyr Glu Leu
850 855 860

Ile Thr Ile Asp Gln Ala Ala Leu Gly Ser Ile Arg Trp Ala Cys Leu
865 870 875 880

Val Val Asp Glu Ala His Arg Leu Lys Asn Asn Gln Ser Lys Phe Phe

ES 2 511 694 T3

| 885 | 890 | 895 |
|--|-----|-----|
| Arg Val Leu Asn Gly Tyr Lys Ile Asp His Lys Leu Leu Leu Thr Gly 900 905 910 | | |
| Thr Pro Leu Gln Asn Asn Leu Glu Glu Leu Phe His Leu Leu Asn Phe 915 920 925 | | |
| Leu Thr Pro Glu Arg Phe Asn Asn Leu Glu Gly Phe Leu Glu Glu Phe 930 935 940 | | |
| Ala Asp Ile Ser Lys Glu Asp Gln Ile Lys Lys Leu His Asp Leu Leu 945 950 955 960 | | |
| Gly Pro His Met Leu Arg Arg Leu Lys Ala Asp Val Phe Lys Asn Met 965 970 975 | | |
| Pro Ala Lys Thr Glu Leu Ile Val Arg Val Glu Leu Ser Pro Met Gln 980 985 990 | | |
| Lys Lys Tyr Tyr Lys Tyr Ile Leu Thr Arg Asn Phe Glu Ala Leu Asn 995 1000 1005 | | |
| Ser Arg Gly Gly Gly Asn Gln Val Ser Leu Leu Asn Ile Met Met 1010 1015 1020 | | |
| Asp Leu Lys Lys Cys Cys Asn His Pro Tyr Leu Phe Pro Val Ala 1025 1030 1035 | | |
| Ala Met Glu Ser Pro Lys Leu Pro Ser Gly Ala Tyr Glu Gly Gly 1040 1045 1050 | | |
| Ala Leu Ile Lys Ser Ser Gly Lys Leu Met Leu Leu Gln Lys Met 1055 1060 1065 | | |
| Leu Arg Lys Leu Lys Glu Gln Gly His Arg Val Leu Ile Phe Ser 1070 1075 1080 | | |
| Gln Met Thr Lys Met Leu Asp Leu Leu Glu Asp Phe Leu Asp Tyr 1085 1090 1095 | | |
| Glu Gly Tyr Lys Tyr Glu Arg Ile Asp Gly Gly Ile Thr Gly Ala 1100 1105 1110 | | |
| Leu Arg Gln Glu Ala Ile Asp Arg Phe Asn Ala Pro Gly Ala Gln 1115 1120 1125 | | |
| Gln Phe Cys Phe Leu Leu Ser Thr Arg Ala Gly Gly Leu Gly Ile 1130 1135 1140 | | |
| Asn Leu Ala Thr Ala Asp Thr Val Ile Ile Phe Asp Ser Asp Trp 1145 1150 1155 | | |
| Asn Pro His Asn Asp Ile Gln Ala Phe Ser Arg Ala His Arg Ile 1160 1165 1170 | | |
| Gly Gln Ala Asn Lys Val Met Ile Tyr Arg Phe Val Thr Arg Ala 1175 1180 1185 | | |
| Ser Val Glu Glu Arg Ile Thr Gln Val Ala Lys Arg Lys Met Met 1190 1195 1200 | | |
| Leu Thr His Leu Val Val Arg Pro Gly Leu Gly Ser Lys Ala Gly 1205 1210 1215 | | |
| Ser Met Ser Lys Gln Glu Leu Asp Asp Ile Leu Lys Phe Gly Thr 1220 1225 1230 | | |

ES 2 511 694 T3

Glu Glu Leu Phe Lys Asp Glu Asn Glu Gly Glu Asn Lys Glu Glu
 1235 1240 1245
 Asp Ser Ser Val Ile His Tyr Asp Asn Glu Ala Ile Ala Arg Leu
 1250 1255 1260
 Leu Asp Arg Asn Gln Asp Ala Thr Glu Asp Thr Asp Val Gln Asn
 1265 1270 1275
 Met Asn Glu Tyr Leu Ser Ser Phe Lys Val Ala Gln Tyr Val Val
 1280 1285 1290
 Arg Glu Glu Asp Lys Ile Glu Glu Ile Glu Arg Glu Ile Ile Lys
 1295 1300 1305
 Gln Glu Glu Asn Val Asp Pro Asp Tyr Trp Glu Lys Leu Leu Arg
 1310 1315 1320
 His His Tyr Glu Gln Gln Gln Glu Asp Leu Ala Arg Asn Leu Gly
 1325 1330 1335
 Lys Gly Lys Arg Val Arg Lys Gln Val Asn Tyr Asn Asp Ala Ala
 1340 1345 1350
 Gln Glu Asp Gln Asp Asn Gln Ser Glu Tyr Ser Val Gly Ser Glu
 1355 1360 1365
 Glu Glu Asp Glu Asp Phe Asp Glu Arg Pro Glu Gly Arg Arg Gln
 1370 1375 1380
 Ser Lys Arg Gln Leu Arg Asn Glu Lys Asp Lys Pro Leu Pro Pro
 1385 1390 1395
 Leu Leu Ala Arg Val Gly Gly Asn Ile Glu Val Leu Gly Phe Asn
 1400 1405 1410
 Thr Arg Gln Arg Lys Ala Phe Leu Asn Ala Val Met Arg Trp Gly
 1415 1420 1425
 Met Pro Pro Gln Asp Ala Phe Thr Thr Gln Trp Leu Val Arg Asp
 1430 1435 1440
 Leu Arg Gly Lys Thr Glu Lys Glu Phe Lys Ala Tyr Val Ser Leu
 1445 1450 1455
 Phe Met Arg His Leu Cys Glu Pro Gly Ala Asp Gly Ser Glu Thr
 1460 1465 1470
 Phe Ala Asp Gly Val Pro Arg Glu Gly Leu Ser Arg Gln Gln Val
 1475 1480 1485
 Leu Thr Arg Ile Gly Val Met Ser Leu Val Lys Lys Lys Val Gln
 1490 1495 1500
 Glu Phe Glu His Ile Asn Gly Arg Trp Ser Met Pro Glu Leu Met
 1505 1510 1515
 Pro Asp Pro Ser Ala Asp Ser Lys Arg Ser Ser Arg Ala Ser Ser
 1520 1525 1530
 Pro Thr Lys Thr Ser Pro Thr Thr Pro Glu Ala Ser Ala Thr Asn
 1535 1540 1545
 Ser Pro Cys Thr Ser Lys Pro Ala Thr Pro Ala Pro Ser Glu Lys
 1550 1555 1560

ES 2 511 694 T3

Gly Glu Gly Ile Arg Thr Pro Leu Glu Lys Glu Glu Ala Glu Asn
 1565 1570 1575
 Gln Glu Glu Lys Pro Glu Lys Asn Ser Arg Ile Gly Glu Lys Met
 1580 1585 1590
 Glu Thr Glu Ala Asp Ala Pro Ser Pro Ala Pro Ser Leu Gly Glu
 1595 1600 1605
 Arg Leu Glu Pro Arg Lys Ile Pro Leu Glu Asp Glu Val Pro Gly
 1610 1615 1620
 Val Pro Gly Glu Met Glu Pro Glu Pro Gly Tyr Arg Gly Asp Arg
 1625 1630 1635
 Glu Lys Ser Glu Asp Val Lys Gly Asp Arg Glu Leu Arg Pro Gly
 1640 1645 1650
 Pro Arg Asp Glu Pro Arg Ser Asn Gly Arg Arg Glu Glu Lys Thr
 1655 1660 1665
 Glu Lys Pro Arg Phe Met Phe Asn Ile Ala Asp Gly Gly Phe Thr
 1670 1675 1680
 Glu Leu His Thr Leu Trp Gln Asn Glu Glu Arg Ala Ala Ile Ser
 1685 1690 1695
 Ser Gly Lys Leu Asn Glu Ile Trp His Arg Arg His Asp Tyr Trp
 1700 1705 1710
 Leu Leu Ala Gly Ile Val Leu His Gly Tyr Ala Arg Trp Gln Asp
 1715 1720 1725
 Ile Gln Asn Asp Ala Gln Phe Ala Ile Ile Asn Glu Pro Phe Lys
 1730 1735 1740
 Thr Glu Ala Asn Lys Gly Asn Phe Leu Glu Met Lys Asn Lys Phe
 1745 1750 1755
 Leu Ala Arg Arg Phe Lys Leu Leu Glu Gln Ala Leu Val Ile Glu
 1760 1765 1770
 Glu Gln Leu Arg Arg Ala Ala Tyr Leu Asn Leu Ser Gln Glu Pro
 1775 1780 1785
 Ala His Pro Ala Met Ala Leu His Ala Arg Phe Ala Glu Ala Glu
 1790 1795 1800
 Cys Leu Ala Glu Ser His Gln His Leu Ser Lys Glu Ser Leu Ala
 1805 1810 1815
 Gly Asn Lys Pro Ala Asn Ala Val Leu His Lys Val Leu Asn Gln
 1820 1825 1830
 Leu Glu Glu Leu Leu Ser Asp Met Lys Ala Asp Val Thr Arg Leu
 1835 1840 1845
 Pro Ala Thr Leu Ser Arg Ile Pro Pro Ile Ala Ala Arg Leu Gln
 1850 1855 1860
 Met Ser Glu Arg Ser Ile Leu Ser Arg Leu Ala Ser Lys Gly Thr
 1865 1870 1875
 Glu Pro His Pro Thr Pro Ala Tyr Pro Pro Gly Pro Tyr Ala Thr
 1880 1885 1890
 Pro Pro Gly Tyr Gly Ala Ala Phe Ser Ala Ala Pro Val Gly Ala

ES 2 511 694 T3

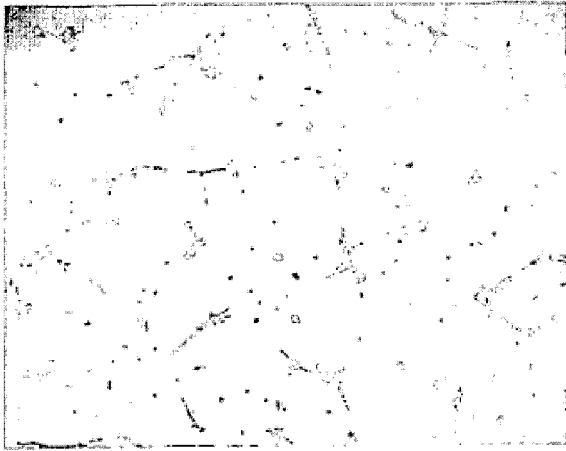
| 1895 | 1900 | | | | | 1905 | | | | | | | | |
|------|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|
| Leu | Ala | Ala | Ala | Gly | Ala | Asn | Tyr | Ser | Gln | Met | Pro | Ala | Gly | Ser |
| 1910 | | | | | | 1915 | | | | | 1920 | | | |
| Phe | Ile | Thr | Ala | Ala | Thr | Asn | Gly | Pro | Ile | His | Arg | Phe | Phe | Lys |
| 1925 | | | | | | 1930 | | | | | 1935 | | | |
| Lys | Val | Pro | Glu | Ile | Gln | | | | | | | | | |
| 1940 | | | | | | | | | | | | | | |

REIVINDICACIONES

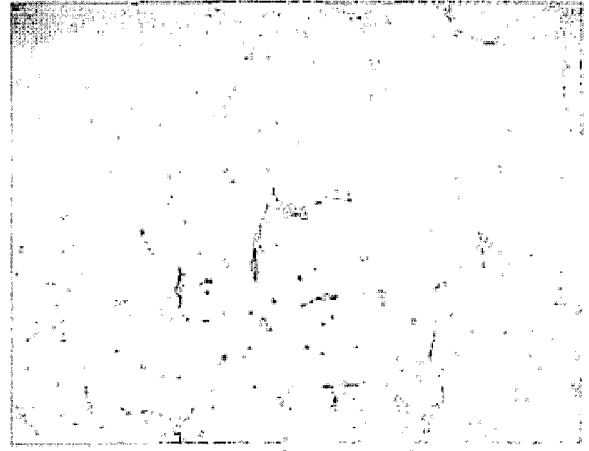
1. Una composición farmacéutica que comprende como agente activo al menos una sustancia seleccionada entre:
- (i) una molécula de ácido nucleico de un gen de una célula endotelial, de secuencia identificada con el número SEQ ID N° 1, SEQ ID N° 27 o SEQ ID N° 28 en la lista de secuencias adjunta,
- 5 (ii) una molécula polipeptídica codificada por dicha molécula de ácido nucleico identificada con el número SEQ ID N° 6, SEQ ID N° 31 o SEQ ID N° 32 en la lista de secuencias adjunta,
- para el diagnóstico *in vivo*, el pronóstico *in vivo* y/o el tratamiento de una patología angiogénica.
2. Una composición que comprende como agente activo al menos una molécula seleccionada entre una molécula de ácidos nucleicos antisentido de al menos una molécula seleccionada entre las SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta, para el diagnóstico *in vivo*, el pronóstico *in vivo* y/o el tratamiento de una patología angiogénica.
- 10 3. La composición farmacéutica para el diagnóstico *in vivo*, el pronóstico *in vivo* y/o el tratamiento de una patología angiogénica según la reivindicación 2, en donde dicha patología angiogénica comprende la enfermedad vascular periférica, cicatrización y reparación de tejidos, osteoporosis, vascularización de tumores, retinopatías, artritis reumatoide, enfermedad de Crohn, aterosclerosis, hiperestimulación ovárica, psoriasis, endometriosis asociada con neovascularización, reestenosis debida a la angioplastia con globo, hiperproducción tisular debida a cicatrización, hipertensión, inflamación vascular, la enfermedad y el fenómeno de Raynaud, aneurisma, reestenosis arterial, tromboflebitis, linfagitis, linfedema, isquemia, angina de pecho, infarto de miocardio, cardiopatía crónica, insuficiencias cardíacas tales como insuficiencia cardíaca congestiva, degeneración macular asociada a la edad.
- 15 4. Una composición para el diagnóstico *in vivo*, el pronóstico *in vivo* y/o el tratamiento de una patología angiogénica según la reivindicación 2 o 3, caracterizada porque comprende la molécula de ácidos nucleicos antisentido identificada con el número SEQ ID N° 11 en la lista de secuencias adjunta.
- 20 5. Una molécula de ácidos nucleicos antisentido identificada con el número SEQ ID N° 11 en la lista de secuencias adjunta o una molécula de ácidos nucleicos antisentido que tiene una identidad al menos del 85 %, preferiblemente al menos del 95 % con la secuencia identificada con el número SEQ ID N° 11 en la lista de secuencias adjunta.
- 25 6. Un vector de expresión de mamífero que comprende al menos una secuencia antisentido definida en la reivindicación 5.
7. El vector según la reivindicación 6, caracterizado porque es el vector GS-V1, identificado por la SEQ ID N° 15 en la lista de secuencias adjunta.
- 30 8. Un método de diagnóstico y/o de pronóstico de una patología angiogénica en un mamífero, particularmente en un ser humano, caracterizado porque se detecta *in vitro* en las células de dicho mamífero la sobreexpresión o la subexpresión de una o más secuencias nucleotídicas identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta, o la sobreexpresión o la subexpresión de una o más secuencias polipeptídicas identificadas con los números SEQ ID N° 6, SEQ ID N° 31 y SEQ ID N° 32 en la lista de secuencias adjunta,
- 35 en donde dicha patología angiogénica comprende la vascularización de tumores, retinopatías, artritis reumatoide, enfermedad de Crohn, aterosclerosis, hiperestimulación ovárica, psoriasis, endometriosis asociada con neovascularización, reestenosis debida a la angioplastia con globo, hiperproducción tisular debida a cicatrización, enfermedad vascular periférica, hipertensión, inflamación vascular, la enfermedad y el fenómeno de Raynaud, aneurisma, reestenosis arterial, tromboflebitis, linfagitis, linfedema, cicatrización y reparación de tejidos, isquemia, angina de pecho, infarto de miocardio, cardiopatía crónica, insuficiencias cardíacas tales como insuficiencia cardíaca congestiva, degeneración macular asociada a la edad y osteoporosis.
- 40 9. El método según la reivindicación 8, caracterizado porque comprende las siguientes etapas:
- 45 - la detección de la expresión de una o más de dichas secuencias nucleotídicas SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 o de una o más de dichas secuencias polipeptídicas SEQ ID N° 6, SEQ ID N° 31 y SEQ ID N° 32 por una población celular de un mamífero,
- la detección de la expresión de una o más de estas mismas secuencias nucleotídicas o polipeptídicas por una población celular de referencia cuyo estado angiogénico es conocido,
- 50 - la identificación de las posibles diferencias del nivel de expresión de estas mismas secuencias por las dos poblaciones celulares.

10. El método de diagnóstico y/o de pronóstico según una cualquiera de las reivindicaciones 8 a 9, caracterizado porque la detección de la expresión de las secuencias se realiza después de haber puesto en contacto las células endoteliales con un fluido biológico obtenido de un paciente.
- 5 11. Un método de verificación de la eficacia terapéutica de un tratamiento angiogénico en un mamífero, particularmente en un ser humano, caracterizado porque comprende la identificación *in vitro* en una población celular de dicho mamífero de la sobreexpresión o de la subexpresión de al menos un gen implicado en un trastorno angiogénico identificado por una de las secuencias nucleotídicas identificadas con los números SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 en la lista de secuencias adjunta.
- 10 12. El método de verificación de la eficacia terapéutica según la reivindicación 11, caracterizado porque comprende las siguientes etapas:
- la detección de la expresión de una o más de dichas secuencias nucleotídicas SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 por una población celular aislada de un mamífero al que se administra una composición terapéutica destinada a tratar un trastorno angiogénico que comprende la vascularización de tumores, retinopatías, artritis reumatoide, enfermedad de Crohn, aterosclerosis, hiperestimulación ovárica, psoriasis, endometriosis asociada con neovascularización, reestenosis debida a la angioplastia con globo, hiperproducción tisular debida a cicatrización, enfermedad vascular periférica, hipertensión, inflamación vascular, la enfermedad y el fenómeno de Raynaud, aneurisma, reestenosis arterial, tromboflebitis, linfagitis, linfedema, cicatrización y reparación de tejidos, isquemia, angina de pecho, infarto de miocardio, cardiopatía crónica, insuficiencias cardíacas tales como insuficiencia cardíaca congestiva, degeneración macular asociada a la edad y osteoporosis,
 - 15 - la detección de la expresión de estas mismas secuencias nucleotídicas por una población celular de referencia cuyo estado angiogénico es conocido,
 - la identificación de las posibles diferencias del nivel de expresión de estas mismas secuencias por las dos poblaciones celulares.
- 20 13. El método de verificación de la eficacia terapéutica según la reivindicación 12, caracterizado porque la detección de la expresión de una o más secuencias nucleotídicas se realiza sobre una población celular aislada de dicho mamífero.
- 25 14. El método de verificación según una cualquiera de las reivindicaciones 11 a 13, caracterizado porque la detección de la expresión de las secuencias se realiza después de haber puesto en contacto las células endoteliales con un fluido biológico obtenido de un paciente.
- 30 15. Un método de cribado *in vitro* de compuestos útiles para el tratamiento de un trastorno angiogénico en un mamífero, especialmente un ser humano, caracterizado porque comprende las siguientes etapas:
- a) la detección de la expresión de una o más de dichas secuencias nucleotídicas SEQ ID N° 1, SEQ ID N° 27 y SEQ ID N° 28 o de dichas secuencias polipeptídicas SEQ ID N° 6, SEQ ID N° 31 y SEQ ID N° 32 por una población celular aislada de un mamífero puesta en contacto con un compuesto capaz de tener un efecto terapéutico sobre un trastorno angiogénico que comprende la vascularización de tumores, retinopatías, artritis reumatoide, enfermedad de Crohn, aterosclerosis, hiperestimulación ovárica, psoriasis, endometriosis asociada con neovascularización, reestenosis debida a la angioplastia con globo, hiperproducción tisular debida a cicatrización, enfermedad vascular periférica, hipertensión, inflamación vascular, la enfermedad y el fenómeno de Raynaud, aneurisma, reestenosis arterial, tromboflebitis, linfagitis, linfedema, cicatrización y reparación de tejidos, isquemia, angina de pecho, infarto de miocardio, cardiopatía crónica, insuficiencias cardíacas tales como insuficiencia cardíaca congestiva, degeneración macular asociada a la edad y osteoporosis,
 - 35 de miocardio, cardiopatía crónica, insuficiencias cardíacas tales como insuficiencia cardíaca congestiva, degeneración macular asociada a la edad y osteoporosis,
 - 40 b) la detección de la expresión de estas mismas secuencias nucleotídicas o polipeptídicas por una población celular de referencia cuyo estado angiogénico es conocido,
 - c) la identificación de las posibles diferencias del nivel de expresión de estas mismas secuencias por las dos poblaciones celulares.
- 45 16. El método de cribado según la reivindicación 15, caracterizado porque la detección de la expresión de las secuencias se realiza después de haber puesto en contacto las células endoteliales con un fluido biológico obtenido de un paciente.

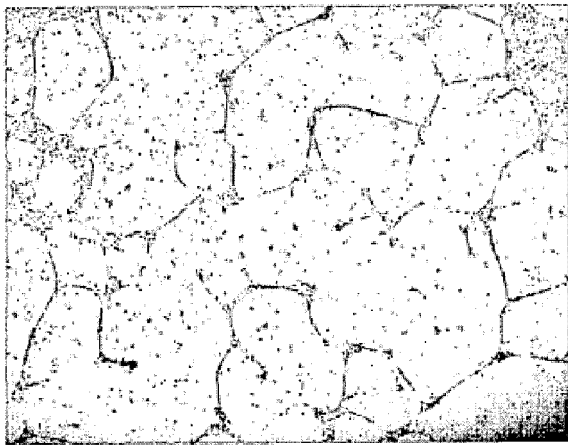
Figura 1



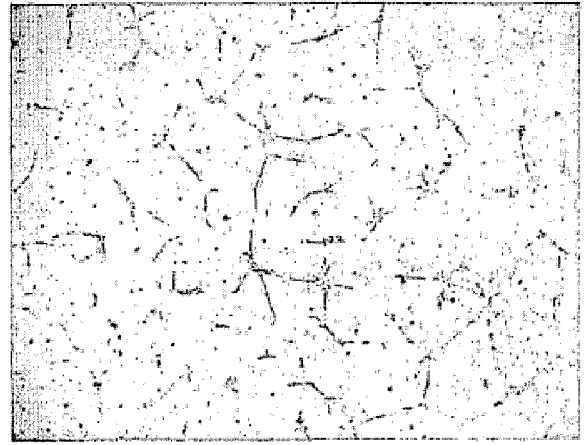
1A) GS-V1



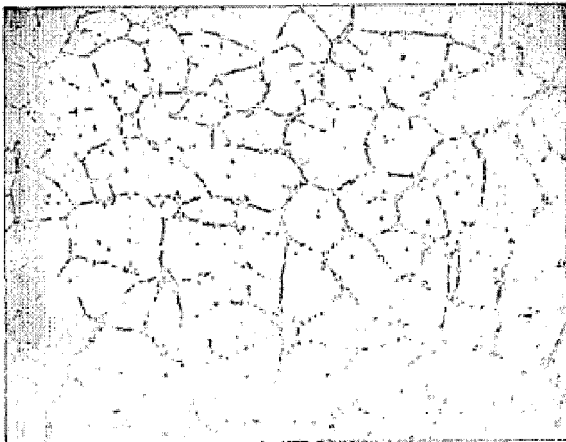
1B) GS-V2



1C) GS-V3



1D) GS-V4



1E) Control