(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 355 185**

(51) Int. Cl.:
***C07K 14/525*** (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **98926105 .2**
(96) Fecha de presentación : **28.05.1998**
(97) Número de publicación de la solicitud: **1015477**
(97) Fecha de publicación de la solicitud: **05.07.2000**

(54) Título: **32 proteínas humanas secretadas.**

(30) Prioridad: **30.05.1997 US 44039 P**
**30.05.1997 US 48093 P**
**30.05.1997 US 48190 P**
**30.05.1997 US 50935 P**
**30.05.1997 US 48101 P**
**30.05.1997 US 48356 P**
**29.08.1997 US 56250 P**
**29.08.1997 US 56296 P**
**29.08.1997 US 56293 P**

(45) Fecha de publicación de la mención BOPI:
**23.03.2011**

(45) Fecha de la publicación del folleto de la patente:
**23.03.2011**

(73) Titular/es: **HUMAN GENOME SCIENCES, Inc.**
**9410 Key West Avenue**
**Rockville, Maryland 20850, US**

(72) Inventor/es: **Ruben, Steven, M.;**
**Rosen, Craig, A.;**
**Carter, Kenneth, C.;**
**Dillon, Patrick, J.;**
**Endress, Gregory, A.;**
**Yu, Guo-Liang;**
**Ni, Jian y**
**Feng, Ping**

(74) Agente: **Arias Sanz, Juan**

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).



Venta de fascículos: Oficina Española de Patentes y Marcas. Pº de la Castellana, 75 – 28071 Madrid

## DESCRIPCIÓN

32 proteínas humanas secretadas.

### Campo de la invención

Esta invención se refiere a polinucleótidos recién identificados y a los polipéptidos codificados por estos polinucleótidos, a usos de tales polinucleótidos y polipéptidos y a su producción.

### Antecedentes de la invención

A diferencia de las bacterias, que existen como un único compartimento rodeado por una membrana, las células humanas y otros eucariotas están subdivididos por membranas en muchos compartimentos funcionalmente distintos. Cada compartimento delimitado por una membrana, u orgánulo, contiene diferentes proteínas esenciales para la función del orgánulo. La célula usa "señales de clasificación", que son motivos de aminoácidos ubicados dentro de la proteína, para dirigir las proteínas a orgánulos celulares particulares.

Un tipo de señal de clasificación, denominada secuencia señal, péptido señal o secuencia líder, dirige una clase de proteínas a un orgánulo denominado retículo endoplasmático (RE). El RE separa las proteínas unidas a la membrana de todos los otros tipos de proteínas. Una vez localizadas en el RE, ambos grupos de proteínas pueden dirigirse adicionalmente a otro orgánulo denominado aparato de Golgi. En este caso, el aparato de Golgi distribuye las proteínas a vesículas, incluyendo vesículas secretoras, la membrana celular, lisosomas y los otros orgánulos.

Las proteínas dirigidas al RE mediante una secuencia señal puede liberarse al espacio extracelular como una proteína secretada. Por ejemplo, pueden fusionarse vesículas que contienen proteínas secretadas con la membrana celular y liberar su contenido al espacio extracelular, un proceso denominado exocitosis. La exocitosis puede producirse de manera constitutiva o tras recibir una señal desencadenante. En este último caso, las proteínas se almacenan en vesículas secretoras (o gránulos secretores) hasta que se desencadena la exocitosis. De manera similar, pueden secretarse también proteínas que residen en la membrana celular al espacio extracelular mediante escisión proteolítica de un "ligador" que sujeta la proteína a la membrana.

A pesar del gran avance realizado en los últimos años, sólo se ha identificado un pequeño número de genes que codifican para proteínas humanas secretadas. Estas proteínas secretadas incluyen insulina humana, interferón, factor VIII, hormona de crecimiento humana, activador de plasminógeno tisular y eritropoyetina comercialmente valiosas. Por tanto, en vista del papel predominante de las proteínas secretadas en la fisiología humana, existe una necesidad de identificar y caracterizar proteínas humanas secretadas novedosas y los genes que las codifican. Este conocimiento permitirá detectar, tratar y prevenir trastornos médicos usando proteínas secretadas o los genes que las codifican.

### Sumario de la invención

La presente invención se refiere a polinucleótidos novedosos y a los polipéptidos codificados. Además, la presente invención se refiere a vectores, células huésped, anticuerpos y métodos recombinantes para producir los polipéptidos y polinucleótidos. Se proporcionan también métodos de diagnóstico para detectar trastornos relacionados con los polipéptidos, y composiciones farmacéuticas. La invención se refiere además a métodos de selección para identificar parejas de unión de los polipéptidos.

Específicamente, la presente invención se refiere a un polinucleótido seleccionado del grupo que consiste en:

(a) un polinucleótido que tiene toda la secuencia de nucleótidos de SEQ ID NO: 23;

(b) un polinucleótido que tiene toda la secuencia de nucleótidos de la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23;

(c) un fragmento de polinucleótido que comprende 150 nucleótidos contiguos de SEQ ID NO: 23;

(d) un fragmento de polinucleótido que comprende 150 nucleótidos de la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23;

(e) un polinucleótido que codifica un fragmento de polipéptido, en el que dicho fragmento de polipéptido comprende 40 ó 50 restos de aminoácido contiguos de longitud de SEQ ID NO: 60;

(f) un polinucleótido que codifica un fragmento de polipéptido codificado por la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23, en el que dicho fragmento de polipéptido comprende 40 ó 50 restos de aminoácido contiguos de SEQ ID NO:60;

(g) un polinucleótido que codifica un fragmento de polipéptido de SEQ ID NO: 60 o un fragmento de polipéptido del mismo codificado por la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, en el que dicho fragmento de polipéptido comprende 40 ó 50 restos de aminoácido contiguos de longitud y tiene actividad biológica;

(h) un polinucleótido que codifica un polipéptido de longitud completa de SEQ ID NO: 60 o que está codificado por la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23;

(i) un polinucleótido que es al menos el 95% idéntico a al menos 150 nucleótidos contiguos de un polinucleótido tal como se definió en uno cualquiera de (a) a (h);

(j) un polinucleótido que codifica un polipéptido que tiene una secuencia de aminoácidos al menos el 95% idéntica a la secuencia de aminoácidos de un polipéptido codificado por un polinucleótido de uno cualquiera de (a) a (i); y

o la hebra complementaria del polinucleótido (a) a (j).

La presente invención se refiere también al gen correspondiente a la secuencia de ADNc de SEQ ID NO: 23 o que codifica el polipéptido que tiene la secuencia de aminoácidos de SEQ ID NO: 60.

La presente invención se refiere además a un vector que comprende el polinucleótido de la invención.

La presente invención se refiere también a un método de preparación de una célula huésped recombinante que comprende introducir el polinucleótido o el vector de la invención.

La presente invención se refiere además a una célula huésped recombinante que contiene el polinucleótido, o el vector de la invención o producido mediante el método mencionado anteriormente de preparación de una célula huésped recombinante.

La presente invención se refiere también a un método de preparación de un polipéptido codificado por el polinucleótido de la invención que comprende;

(a) cultivar la célula huésped recombinante mencionada anteriormente que expresa dicho polipéptido en condiciones tales que se exprese dicho polipéptido; y

(b) recuperar dicho polipéptido.

La presente invención se refiere además a un polipéptido codificado por el polinucleótido de la invención o que puede obtenerse mediante el método de preparación de un polipéptido mencionado anteriormente.

La presente invención se refiere también a un anticuerpo que se une específicamente al polipéptido de la invención.

La presente invención se refiere adicionalmente a un antagonista que comprende un polinucleótido antisentido con respecto al polinucleótido de la invención que puede unirse a dicho polinucleótido y bloquear su traducción en el polipéptido codificado.

La presente invención se refiere también a un método para identificar antagonistas o agonistas del polipéptido de la invención que comprende:

(a) producir células que expresan dicho polipéptido;

(b) poner en contacto el polipéptido producido en la etapa (a) con una muestra de prueba que contiene posiblemente un antagonista o agonista; y

(c) identificar un antagonista o agonista observando la unión y la inhibición o estimulación de la actividad de dicho polipéptido.

La presente invención se refiere además a un método *in vitro* de diagnóstico de un estado patológico o una susceptibilidad a un estado patológico en un sujeto que comprende:

(a) determinar la presencia o ausencia de una mutación en el polinucleótido o el gen de la invención; y

(b) diagnosticar un estado patológico o una susceptibilidad a un estado patológico basándose en la presencia o ausencia de dicha mutación.

La presente invención se refiere también a un método *in vitro* de diagnóstico de un estado patológico o una susceptibilidad a un estado patológico en un sujeto que comprende:

(a) determinar la presencia o cantidad de expresión del polipéptido de la invención en una muestra biológica; y

(b) diagnosticar un estado patológico o una susceptibilidad a un estado patológico basándose en la presencia o cantidad de expresión del polipéptido.

La presente invención se refiere además a un método para identificar una pareja de unión al polipéptido de la invención que comprende:

(a) poner en contacto dicho polipéptido con una pareja de unión; y

(b) determinar si la pareja de unión efectúa una actividad del polipéptido.

Además, la presente invención se refiere a una composición farmacéutica que comprende el polinucleótido de la invención, el polipéptido de la invención, un ADN que codifica y que puede expresar dicho polipéptido *in vivo*, el anticuerpo de la invención, el antagonista de la invención y opcionalmente un vehículo farmacéuticamente aceptable.

La presente invención se refiere también a una composición de diagnóstico que comprende el polinucleótido o el anticuerpo de la invención.

Finalmente, la presente invención se refiere también a un método para la producción de una composición farmacéutica que comprende las etapas del método mencionado anteriormente para identificar antagonistas o agonistas y (d) formular el compuesto identificado en la etapa (c) en una forma farmacéuticamente aceptable.

**Descripción detallada**

*Definiciones*

Las siguientes definiciones se proporcionan para facilitar la comprensión de ciertos términos usados en toda esta memoria descriptiva.

En la presente invención, "aislado" se refiere a material sacado de su entorno original (por ejemplo, el entorno natural si se produce de manera natural), y por tanto se altera "por la mano del hombre" con respecto a su estado natural. Por ejemplo, un polinucleótido aislado podría ser parte de un vector o una composición de materia, o podría estar contenido dentro de una célula, y todavía estar "aislado" porque el vector, la composición de materia o la célula particular no es el entorno original del polinucleótido.

En la presente invención, una proteína "secretada" se refiere a aquellas proteínas que pueden dirigirse al RE, vesículas secretoras o el espacio extracelular como resultado de una secuencia señal, así como a aquellas proteínas liberadas al espacio extracelular sin contener necesariamente una secuencia señal. Si la proteína secretada se libera al espacio extracelular, la proteína secretada puede experimentar procesamiento extracelular para producir una proteína "madura". La liberación al espacio extracelular puede producirse mediante muchos mecanismos, incluyendo exocitosis y escisión proteolítica.

Tal como se usa en el presente documento, un "polinucleótido" se refiere a una molécula que tiene una secuencia de ácido nucleico contenida en SEQ ID NO:23 o el ADNc contenido dentro del clon depositado con el depósito de la ATCC n.º 209075 o variantes del mismo tal como se definió en de (a) a (j), anteriormente. Por ejemplo, el polinucleótido puede contener la secuencia de nucleótidos de la secuencia de ADNc de longitud completa, incluyendo las secuencias no traducidas en 5' y 3', la región codificante, con o sin la secuencia señal, la región codificante de proteína secretada, así como fragmentos, epítopos, dominios y variantes de la secuencia de ácido nucleico según las definiciones en de (a) a (j), anteriormente. Además, tal como se usa en el presente documento, un "polipéptido" se refiere a una molécula que tiene la secuencia de aminoácidos no traducida generada a partir del polinucleótido tal como se definió anteriormente de manera amplia.

En la presente descripción, la secuencia de longitud completa identificada como SEQ ID NO:X se generó a menudo solapando secuencias contenidas en múltiples clones (análisis de cóntigos). Se depositó un clon representativo que contenía toda o la mayor parte de la secuencia para SEQ ID NO:X en la Colección Americana de Cultivos Tipo (*American Type Culture Collection*) ("ATCC"). Tal como se muestra en la tabla 1, cada clon se identifica mediante un ID (identificador) de ADNc y el número de depósito de la ATCC. La ATCC se ubica en 10801 University Boulevard, Manassas, Virginia 20110-2209, EE.UU. El depósito de la ATCC se realizó siguiendo los términos del Tratado de Budapest sobre el reconocimiento internacional del depósito de microorganismos a los fines del procedimiento en materia de patentes.

Un "polinucleótido" de la presente invención incluye también aquellos polinucleótidos que pueden hibridar, en condiciones de hibridación rigurosas, con secuencias contenidas en SEQ ID NO: 23, el complemento de las mismas, o el ADNc dentro del clon depositado con el depósito de la ATCC n.º 209075, siempre que la secuencia hibridante tenga

al menos una identidad de secuencia del 95% con al menos 150 nucleótidos contiguos de un polinucleótido definido en una cualquiera de las reivindicaciones 1(a) a (h) o codifique para un polipéptido que tiene una secuencia de aminoácidos al menos el 95% idéntica a la secuencia de aminoácidos de un polipéptido codificado por un polinucleótido de una cualquiera de las reivindicaciones 1(a) a (i). "Condiciones de hibridación rigurosas" se refiere a una incubación durante la noche a 42ºC en una disolución que comprende formamida al 50%, 5x SSC (NaCl 750 mM, citrato de sodio 75 mM), fosfato de sodio 50 mM (pH 7,6), 5x solución de Denhardt, sulfato de dextrano al 10% y ADN de esperma de salmón desnaturalizado, fragmentado 20 $\mu$g/ml, seguido por lavado de los filtros en 0,1x SSC a aproximadamente 65ºC.

Se describen también moléculas de ácido nucleico que hibridan con los polinucleótidos de la presente invención en condiciones de hibridación de rigurosidad inferior. Los cambios en la rigurosidad de la hibridación y la detección de la señal se logran principalmente a través de la manipulación de la concentración de formamida (porcentajes inferiores de formamida dan como resultado una disminución de la rigurosidad); las condiciones de sal, o la temperatura. Por ejemplo, las condiciones de rigurosidad inferior incluyen una incubación durante la noche a 37ºC en una disolución que comprende 6X SSPE (20X SSPE = NaCl 3 M; $NaH_2PO_4$ 0,2 M; EDTA 0,02 M, pH 7,4), SDS al 0,5%, formamida al 30%, ADN bloqueante de esperma de salmón 100 ug/ml; seguido por lavados a 50ºC con 1XSSPE, SDS al 0,1%. Además, para lograr una rigurosidad incluso inferior, pueden hacerse lavados realizados tras la hibridación rigurosa a concentraciones de sal superiores (por ejemplo 5X SSC).

Obsérvese que pueden lograrse variaciones en las condiciones anteriores a través de la inclusión y/o sustitución de reactivos bloqueantes alternativos usados para suprimir el fondo en experimentos de hibridación. Los reactivos bloqueantes típicos incluyen reactivo de Denhardt, BLOTTO, heparina, ADN de esperma de salmón desnaturalizado y formulaciones registradas comercialmente disponibles. La inclusión de reactivos bloqueantes específicos puede requerir la modificación de las condiciones de hibridación descritas anteriormente, debido a problemas con la compatibilidad.

Por supuesto, un polinucleótido que hibrida sólo con secuencias de poliA+ (tales como cualquier extensión de poliA+ terminal en 3' de un ADNc mostrado en la lista de secuencias) o con un tramo complementario de restos de T (o U), no estaría incluido en la definición de "polinucleótido", puesto que un polinucleótido de este tipo hibridaría con cualquier molécula de ácido nucleico que contuviese un tramo de poli (A) o el complemento del mismo (por ejemplo, prácticamente cualquier clon de ADNc bicatenario).

El polinucleótido de la presente invención puede estar compuesto por cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ADN o ARN no modificado o ADN o ARN modificado. Por ejemplo, los polinucleótidos pueden estar compuestos por ADN mono y bicatenario, ADN que es una mezcla de regiones mono y bicatenarias, ARN mono y bicatenario y ARN que es una mezcla de regiones mono y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser monocatenarias o, más normalmente, bicatenarias o una mezcla de regiones mono y bicatenarias. Además, los polinucleótidos pueden estar compuestos por regiones tricatenarias que comprenden ARN o ADN o tanto ARN como ADN. Un polinucleótido también puede contener una o más bases modificadas o estructuras principales de ADN o ARN modificadas para lograr estabilidad o por otros motivos. Las bases "modificadas" incluyen, por ejemplo, bases tritiladas y bases poco comunes tales como inosina. Puede hacerse una variedad de modificaciones en el ADN y ARN; por tanto, "polinucleótido" abarca formas química, enzimática o metabólicamente modificadas.

El polipéptido de la presente invención puede estar compuesto por aminoácidos unidos entre sí mediante enlaces peptídicos o enlaces peptídicos modificados, es decir, isoésteres peptídicos, y puede contener aminoácidos distintos de los 20 aminoácidos codificados por genes. Los polipéptidos pueden modificarse o bien mediante procesos naturales, tales como procesamiento postraduccional, o bien mediante técnicas de modificación química que se conocen bien en la técnica. Tales modificaciones se describen bien en textos básicos y en monografías más detalladas, así como en una voluminosa bibliografía de investigación. Pueden producirse modificaciones en cualquier sitio en un polipéptido, incluyendo la estructura principal peptídica, las cadenas laterales de los aminoácidos y los extremos amino o carboxilo terminal. Se apreciará que puede estar presente el mismo tipo de modificación en grados iguales o variables en varios sitios de un polipéptido dado. Además, un polipéptido dado puede contener muchos tipos de modificaciones. Los polipéptidos pueden estar ramificados, por ejemplo, como resultado de la ubiquitinación, y pueden ser cíclicos, con o sin ramificación. Los polipéptidos cíclicos, ramificados y cíclicos ramificados pueden resultar de procesos naturales tras la traducción o pueden prepararse mediante métodos sintéticos. Las modificaciones incluyen acetilación, acilación, ADP-ribosilación, amidación, unión covalente de flavina, unión covalente de un resto hemo, unión covalente de un nucleótido o derivado de nucleótido, unión covalente de un lípido o derivado de lípido, unión covalente de fosfatidilinositol, reticulación, ciclación, formación de enlaces disulfuro, desmetilación, formación de reticulaciones covalentes, formación de cisteína, formación de piroglutamato, formilación, gamma-carboxilación, glicosilación, formación de ancla de GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, pegilación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, sulfatación, adición mediada por ARN de transferencia de aminoácidos a proteínas tal como arginilación y ubiquitinación. (Véase, por ejemplo, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2ª ed., T. E. Creighton, W. H. Freeman and Company, Nueva York (1993); POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, Nueva York, págs. 1-12 (1983); Seifter *et al*., Meth Enzymol 182:626-646 (1990); Rattan *et al*., Ann NY Acad Sci 663:48-62 (1992)).

"SEQ ID NO:X" se refiere a una secuencia de polinucleótido mientras que "SEQ ID NO:Y" se refiere a una secuencia de polipéptido, identificándose ambas secuencias mediante un número entero especificado en la tabla 1.

"Un polipéptido que tiene actividad biológica" se refiere a polipéptidos que presentan actividad similar, pero no necesariamente idéntica a, una actividad de un polipéptido de la presente invención, incluyendo formas maduras, tal como se mide en un ensayo biológico particular, con o sin dependencia de la dosis. En el caso en el que exista dependencia de la dosis, no es necesario que sea idéntica a la del polipéptido, sino en su lugar sustancialmente similar a la dependencia de la dosis en una actividad dada en comparación con el polipéptido de la presente invención (es decir, el polipéptido candidato presentará una mayor actividad o no más de aproximadamente 25 veces menos y, preferiblemente, no más de aproximadamente diez veces menos actividad, y lo más preferiblemente, no más de aproximadamente tres veces menos actividad en relación con el polipéptido de la presente invención).

*Polinucleótidos y polipéptidos de la invención o descritos en el presente documento*

*Características de la proteína codificada por el gen n.º: 1*

Este gen se mapea en el cromosoma 3 y por tanto los polinucleótidos descritos en el presente documento pueden usarse en un análisis de ligamiento como marcador para el cromosoma 3.

Este gen se expresa en varios tejidos fetales incluyendo cerebro, hígado y pulmón, y en menor grado en tejidos adultos, particularmente la piel.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, una variedad de cánceres, particularmente del cerebro, hígado y pulmón. De manera similar, polipéptidos y los anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema nervioso central, el sistema hepático y el sistema hepático, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, cerebro y otros tejidos del sistema nervioso, hígado, pulmón y piel, y tejidos cancerosos o dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles como diana para una variedad de agentes bloqueantes, ya que es probable que estén implicados en la promoción de una variedad de cánceres.

*Características de la proteína codificada por el gen n.º 2*

Los polipéptidos descritos en el presente documento comprenden la secuencia: MSVPAFIDISEEDQAAELRAYL KSKGAEISEENSEGGLHVDLAQIIEACDVCLEEDDKDVESVMNSVVSLLLILEPDKQEAUESLCEKLVKFREG ERPSLRLQLLSNLFHGMDKNTPVRYTVYCSLIKVAASCGAIQYIPTELDQVRKWISDWNLTTEKKHTLLRLL YEALVDCKKSDAASKVMVELLGSYTEDNASQARVDAHRCIVRALKDPNAFLFDHLLTLKPVKFLEGELLID LLTIFVSAKLASYVKFYQNNKDFIDSLGLLHEQNMAKMRLLTFMGMAVENKEISFDTMQQELQIGADDVEA FVIDAVRTKMVYCKIDQTQRKVVVSHSTHRTFGKQQWQQLYDTLNAWKQNLNKVKNSLLSLSDT (SEQ ID NO:85), MSVPAFIDISEED (SEQ ID NO:86), QAAELRAYLKSKG AE (SEQ ID NO:87), ISEENSEGGLHVD LAQI (SEQ ID NO: 88), IEACDVCLKED DKDVESV (SEQ ID NO:89), VARPSSLFRSAWSCEW (SEQ ID NO:90), LRLQLLS NLFHG (SEQ ID NO:91), KDVESVMNSWSLLUL (SEQ ID NO:92), DAASKVMVELLGSYTED NASQARVDA (SEQ ID NO:93), y/o VEAFVIDAVR (SEQ ID NO:94). También se dan a conocer polinucleótidos que codifica estos polipéptidos.

Este gen se expresa en hueso y en menor grado en cerebro, pulmón, células T, músculo, piel, testículos, bazo y macrófagos.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, cáncer de huesos, artrosis y enfermedades autoinmunitarias. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario y el sistema esquelético, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, cerebro y otros tejidos del sistema nervioso, células T y otras células y tejido del sistema inmunitario, pulmón, músculo, piel y testículos y otro tejido reproductor, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos

incluyen los que comprenden una secuencia mostrada en SEQ ID NO:49 como restos: Arg-31 a Ser-37, Met-50 a Val-56, Glu-80 a Trp-87, Thr-94 a His-99, Tyr-129 a Ser-135, Tyr-193 a Phe-199, Ser-274 a Gln-285 y/o Ala-293 a Lys-302.

*Características de la proteína codificada por el gen n.º: 3*

El producto de traducción de este gen comparte homología de secuencia con diversas cinasas. El homólogo más cercano es TIF1 de ratón que es una proteína nuclear de ratón. TIF1 potencia RXR y RAR AF-2 en levaduras e interacciona de una manera dependiente de ligando con varios receptores nucleares en células de mamíferos y levaduras, así como *in vitro*. De manera notable, estas interacciones requieren los aminoácidos que constituyen el dominio de activación AF-2 conservado en todas las NR activas. Además, el antagonista del receptor de estrógenos (ER) AF-2, hidroxitamoxifeno, no puede promover la interacción ER-TIF1. Se propone que TIF1, que contiene varios dominios conservados encontrados en proteínas reguladoras de la transcripción, es un mediador de AF-2 dependiente de ligando. De manera interesante, el resto N-terminal de TIF1 está fusionado con B-raf en la oncoproteína de ratón T18.

Este gen se expresa principalmente en células T y en menor grado en otros diversos tejidos incluyendo testículos y cerebro.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, enfermedades autoinmunitarias, SIDA, leucemias y otros diversos cánceres. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, células T y otras células y tejido del sistema inmunitario, testículos y otro tejido reproductor, y cerebro y otro tejido del sistema nervioso, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:50 como restos: Ala-31 a Glu-36.

La distribución tisular y la homología con TIF indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para la modulación de la interacción de ligandos y receptores nucleares en diversos trastornos inmunitarios.

*Características de la proteína codificada por el gen n.º: 4*

Este gen se mapea en el cromosoma 11. Por consiguiente, los polinucleótidos descritos en el presente documento pueden usarse en un análisis de ligamiento como marcador para el cromosoma 11. En realizaciones específicas, los polipéptidos descritos comprenden la secuencia:

MSEIYLRCQDEQQYARWMAGCRLASKGRTMADSSY (SEQ ID NO:95), LVAPRF QRKFKAKQLTPRILEA HQNVAQLSLAEAQLRFIQAWQSL (SEQ ID NO:96), VGD VVKTWRFSNMRQWNVNWDIR (SEQ ID NO:97), EEIDCTEEEMMVFAALQYH INKLSQS (SEQ ID NO:98) y/o EEIDCTEEEMMVFAALQYHINKLSQS (SEQ ID NO:99). También se dan a conocer polinucleótidos que codifican para estos polipéptidos.

Este gen se expresa principalmente en varios tipos de glóbulos blancos incluyendo monocitos, células T y neutrófilos y en menor grado en un número limitado de otros tejidos incluyendo vena umbilical e hígado.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, diversas enfermedades del sistema inmunitario incluyendo SIDA, enfermedades de inmunodeficiencia y trastornos autoinmunitarios. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, células sanguíneas, hígado y tejido vascular, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:51 como restos: Ser-3 a Pro-9, Leu-17 a Leu-29, Asp-64 a Pro-69, Ile-105 a Gln-110, Thr-183 a Gln-200, Cys-239 a Arg-247, Ser-256 a Met-261, Gln-280 a Ala-296, Arg-310 a Thr-321, Lys-363 a Asp-368, Ser-395 a Trp-400 y/o Thr-443 a Asp-453.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para la terapia de sustitución en una variedad de trastornos del sistema inmunitario.

*Características de la proteína codificada por el gen n.º: 5*

Este gen se expresa principalmente en cerebro y poco o nada en absoluto en cualquier otro tejido.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos del estado de ánimo, esquizofrenia y enfermedades relacionadas, trastorno bipolar y depresión unipolar. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema nervioso central, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, cerebro y otro tejido del sistema nervioso, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:52 como restos: Met-1 a Gly-8, Pro-10 a Arg-17, Pro-45 a Ser-55 y/o Gly-63 a Tyr-74.

La distribución tisular de este gen principalmente en el cerebro indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para la detección/el tratamiento de estados patológicos neurodegenerativos y trastornos del comportamiento tales como enfermedad de Alzheimer, enfermedad de Parkinson, enfermedad de Huntington, esquizofrenia, manía, demencia, paranoia, trastorno obsesivo-compulsivo y trastornos de pánico. También dada la expresión específica de cerebro de este gen, la región promotora de este gen contiene un elemento específico de cerebro que podría usarse para dirigir la expresión de sistemas de vector al cerebro en la terapia de sustitución génica.

*Características de la proteína codificada por el gen n.º: 6*

Este gen se mapea en el cromosoma 1 y por tanto los polinucleótidos descritos en el presente documento pueden usarse en un análisis de ligamiento como marcador para el cromosoma 1.

Este gen se expresa de manera abundante en rabdomiosarcoma, se expresa en un alto nivel y en diferentes regiones del cerebro y la hipófisis y en menor grado en una variedad de otros tejidos.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos neurológicos y trastornos musculares. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejidos o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del cerebro, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, músculo liso, cerebro y otro tejido del sistema nervioso, e hipófisis, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La abundante expresión de este gen en rabdomiosarcoma indica un papel para el producto de proteína o bien en la detección y/o bien el tratamiento de trastornos del músculo esquelético incluyendo degeneración muscular, atrofia muscular y rabdomiolisis. Además, la expresión en el cerebro indica un papel para el producto de proteína de este gen en la detección/el tratamiento de estados patológicos neurodegenerativos y trastornos del comportamiento tales como enfermedad de Alzheimer, enfermedad de Parkinson, enfermedad de Huntington, esquizofrenia, manía, demencia, paranoia, trastornos obsesivo-compulsivo y trastorno de pánico.

*Características de la proteína codificada por el gen n.º: 7*

El producto de traducción de este gen comparte homología de secuencia con el gen TDAG51 que se piensa que es importante en la mediación de la apoptosis y la muerte celular acoplando la estimulación de TCR con la expresión de Fas. En realizaciones específicas, los polipéptidos de la invención comprenden la secuencia: KELSFARIKAVEC VESTGRHIYFTLV (SEQ ID NO: 100) y/o GWNAQITLGLVKFKNQQ (SEQ ID NO:101).

Este gen se expresa en diversos tejidos humanos incluyendo macrófagos.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos inmunitarios. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, macrófagos y otras células sanguíneas, y tejidos

cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:54 como restos: Met-1 a Pro-9, Gln-43 a Glu-49 y/o Phe-95 a Arg-102.

La distribución tisular y la homología con el gen TDAG51 indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para el diagnóstico y la intervención de trastornos inmunitarios, tales como inmunodeficiencia, alergia, infección, inflamación, trasplante de tejidos/órganos.

*Características de la proteína codificada por el gen n.º: 8*

Este gen se expresa en tejido mamario y células amnióticas, y en menor grado en músculo liso, células T y cerebro de lactante.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, síndrome de sufrimiento fetal y atrofia embrionaria. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema reproductor femenino, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, tejido mamario, células amnióticas, músculo liso, cerebro y otro tejido del sistema nervioso, y células T y otras células y tejido del sistema inmunitario, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

*Características de la proteína codificada por el gen n.º: 9*

En realizaciones específicas, los polipéptidos descritos en el presente documento comprenden la secuencia: LVL GLSXLNNSYNFSF (SEQ ID NO:102), HVVIGSQAEEGQYSLNF (SEQ ID NO:103), HNCNNSVPGKEHPFDITV M (SEQ ID NO:104), FIKYVLSD KEKKVFGIV (SEQ ID NO:105), IPMQVLANVAYII (SEQ ID NO:106), IPMQV L ANVAYII (SEQ ID NO:107), DGKVAVNLAKLKLFR (SEQ ID NO:108) y/o IREKNPDGFLSAA (SEQ ID NO: 109). También se dan a conocer polinucleótidos que codifican para estos polipéptidos.

Este gen se expresa principalmente en el hígado, el bazo y la hipófisis fetales, y en menor grado en múltiples tejidos.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos inmunitarios y cáncer. De manera similar, polipéptidos y anticuerpos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente de los sistemas hepático, inmunitario y hematopoyético, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, hígado, bazo e hipófisis, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:56 como restos: Ser-62 a Cys-71, Thr-78 a Leu-86, Ser-104 a Lys-109, Ser-130 a Ala-135 y/o Gln-168 a Asp-174.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para el diagnóstico y tratamiento de trastornos hepáticos, y trastornos de los sistemas inmunitario y hematopoyético, tales como insuficiencia hepática, hepatitis, enfermedades alcohólicas del hígado, hipertensión portal, lesión tóxica del hígado, trasplante de hígado y neoplasma del hígado. La expresión en el bazo e hígado fetales indica también su función en la hematopoyesis, y por tanto el gen puede ser útil en trastornos hematopoyéticos incluyendo anemia, leucemia o cáncer, radioterapia/quimioterapia. La expresión en la hipófisis puede indicar su uso en trastornos endocrinos con manifestaciones sistémicas o específicas.

*Características de la proteína codificada por el gen n.º: 10*

El producto de traducción de este gen comparte homología de secuencia con una proteína de unión a ADN de pollo que se piensa que es importante en la regulación transcripcional de la expresión génica. Los polipéptidos descritos en el presente documento comprenden la secuencia: MMFGGYETI (SEQ ID NO:110), YRDESSSELSVDSE VEFQLYSQIH (SEQ ID NO: 111), YAQDLDDVIREEEHEEKNSGNSESSSSKPNQKKLIVLSDSEVIQLSDGSE VITLSDEDSIYRCKGKNVRVQAQENAHGLSSSLQSNELVDKKCKSDIEKPKSEERSGVIREVMIIEVSSSEEEE

STISEGDNVESW (SEQ ID NO: 112), MLLGCEVDDKDDDILLNLVGCENSVTEGEDGINWSIS (SEQ ID NO: 113), DKDIEAQI ANNRTPGRWT (SEQ ID NO: 114), QRYYSANKNIICRNCDKRGHLSKNCPLP RKV (SEQ ID NO:115) y/o RRCFLCSRRGHLLYSCPAPLCEYCPVPKMLDHSCLFRHSWDKQCDRCHMLGHYTDACTEDNR QYHLTTKPGPPKKPKTPSRPSALAYCYHCAQKGHYGHECPEREVYDPSPVSPFICYYXDKYEIQEREKRLKQ KIKV XKKNGVIPEPSKLPYIKAANENPHHDIRKORASWKSNRWPQ (SEQ ID NO:116). También se dan a conocer polinucleótidos que codifican para estos polipéptidos.

Este gen se expresa en amígdalas y médula ósea.

Por tanto, los polinucleótidos y polipéptidos descritos son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos de los sistemas inmunitario, hematopoyético y linfático. De manera similar, polipéptidos y anticuerpos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente de los sistemas inmunitario, hematopoyético y linfático, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, amígdalas y médula ósea, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular y la homología con una proteína de unión a ADN indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para el tratamiento y diagnóstico de trastornos en los sistemas inmunitario, hematopoyético y linfático.

*Características de la proteína codificada por el gen n.º: 11*

Este gen se expresa en células T y dendríticas.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos del sistema inmunitario. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos de células (por ejemplo, células dendríticas, y células T y otras células y tejido del sistema inmunitario, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los productos de proteína de este gen son útiles para el tratamiento y diagnóstico de trastornos del sistema inmunitario, particularmente los que implican a células T o dendríticas tales como inflamación.

*Características de la proteína codificada por el gen n.º: 12*

Este gen se expresa en neutrófilos activados, células endoteliales, células T y en menor grado en cerebro y hígado.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, SIDA, trastornos inmunitarios y susceptibilidad a enfermedad infecciosa. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario y la piel, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, neutrófilos y otras células sanguíneas, células endoteliales, células T y otras células y tejido del sistema inmunitario, cerebro y otro tejido del sistema nervioso, e hígado, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:59 como restos: Glu-41 a Val-46.

Este producto génico es útil para el diagnóstico y/o el tratamiento de una variedad de trastornos, incluyendo trastornos hematopoyéticos, trastornos neurológicos, enfermedad hepática y trastornos que implican angiogénesis.

*Características de la proteína codificada por el gen n.º: 13*

Este gen se expresa en queratinocitos y en menor grado en células endoteliales y placenta.

Por tanto, los polinucleótidos y polipéptidos de la invención son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, cicatrización de heridas alterada. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente de la piel, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos de células (por ejemplo, queratinocitos y otras células de la piel, células endoteliales y placenta, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:60 como restos: Pro-35 a Trp-42, Ala-53 a Asp-62 y/o Arg-103 a Pro-113.

La distribución tisular indica que los productos de proteína de este gen son útiles para el tratamiento de la deficiencia de cicatrización de heridas y trastornos de la piel.

*Características de la proteína codificada por el gen n.º: 14*

Este gen se expresa en riñón y en menor grado en tejidos embrionarios.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, insuficiencia renal. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del riñón, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, riñón, tejido embrionario y otro en desarrollo rápido (por ejemplo, en división), y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

*Características de la proteína codificada por el gen n.º: 15*

Este gen se expresa principalmente en cerebro y en menor grado en hígado.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, depresión, depresión maníaca y otras enfermedades mentales. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema nervioso central, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, cerebro y otro tejido del sistema nervioso, y hígado, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los productos de proteína de este gen son útiles para el tratamiento de trastornos del sistema nervioso central tales como depresión y otras enfermedades mentales.

*Características de la proteína codificada por el gen n.º: 16*

Este gen se expresa en cerebro fetal y en menor grado en placenta, células endoteliales, pulmón fetal y células T.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, reestenosis, defectos de nacimiento y trastornos inmunitarios. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema cardiovascular, y el proceso del desarrollo, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, placenta, células endoteliales, pulmón, y células T y otras células y tejido del sistema inmunitario, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero,

plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:63 como restos: Gln-36 a Lys-42 y/o Glu-89 a Arg-104.

La distribución tisular indica que los productos de proteína de este gen son útiles para el desarrollo de agonistas y/o antagonistas para el tratamiento de trastornos del sistema nervioso y el desarrollo fetal.

*Características de la proteína codificada por el gen n.º: 17*

Este gen se expresa en hemangiopericitoma y en menor grado en tejidos fetales.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, hemangiopericitomas y otros cánceres, así como trastornos del desarrollo. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, tejido vascular, tejido pericítico y tejido en desarrollo, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:64 como restos: Glu-43 a Pro-51, Gly-71 a Arg-82, Pro-96 a Arg-103 y/o Thr-130 a Gly-140.

Se cree que los polinucleótidos y polipéptidos relacionados con este gen son útiles para el tratamiento y diagnóstico de tumores, particularmente hemangiopericitomas, y para el tratamiento de trastornos del desarrollo.

*Características de la proteína codificada por el gen n.º: 18*

Este gen se expresa en hígado fetal y en menor grado en cerebro y células T.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos fetales, desarrollo fetal y trastornos inmunitarios. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema hepático, sistema nervioso y sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, hígado, cerebro y otro tejido del sistema nervioso y células T y otras células y tejido del sistema inmunitario, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los productos de proteína de este gen son útiles para la identificación de agonistas y/o antagonistas para el tratamiento de enfermedades mentales tales como esquizofrenia y depresión. El producto génico también puede ser útil para monitorizar el desarrollo fetal durante el embarazo.

*Características de la proteína codificada por el gen n.º: 19*

Este gen se expresa en células T y en menor grado en cerebro.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, enfermedades nerviosas centrales y trastornos inmunitarios. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema nervioso central y sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, células T y otras células y tejido del sistema inmunitario, y cerebro y otro tejido del sistema nervioso, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:66 como restos: Lys-69 a Leu-74, Ser-92 a Phe-97, Asp-109 a Leu-117, Leu-142 a Ser-159, Thr-166 a Glu-183, Ala-191 a Glu-205 y/o Pro-213 a Glu-220.

La distribución tisular indica que los productos de proteína de este gen son útiles para el desarrollo de fármacos para el tratamiento de trastornos que afectan al sistema nervioso central y al sistema inmunitario.

*Características de la proteína codificada por el gen n.º: 20*

El producto de traducción de este gen comparte homología de secuencia con un ORF de *C. elegans* que parece ser una proteína transmembrana. (Véase el número de registro de GenBank 790406). Este cóntigo tiene dos posibles desplazamientos del marco entre los marcos +2 y +3 basándose en la homología con el gen de *C. elegans*. Este desplazamiento del marco puede resolverse fácilmente secuenciando el clon depositado. Además, este gen se mapea en el cromosoma 8, y por tanto puede usarse como marcador en un análisis de ligamiento para el cromosoma 8.

Este gen se expresa de manera ubicua, incluyendo células T y núcleo amigdalino.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, cáncer. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario y sistema endocrino, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, células T y otras células y tejido del sistema inmunitario, núcleo amigdalino y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular ubicua y la homología con una proteína de tipo transmembrana de *C. elegans* indica que el producto de proteína de este gen desempeña un papel importante tanto en vertebrados como invertebrados y es útil para el diagnóstico o tratamiento de trastornos relacionados con este gen.

*Características de la proteína codificada por el gen n.º: 21*

Este gen se expresa principalmente en tejido embrionario y testículos y en menor grado en ovario, hepatoma, riñón, células endoteliales y de músculo liso.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastorno metabólico, desarrollo embrionario anómalo y tumor. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente de los tejidos vascular o embrionario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, ovario y otro tejido reproductor, riñón, células endoteliales, y células de músculo liso, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular y la homología con la NADH deshidrogenasa indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para el diagnóstico y/o tratamiento de trastornos metabólicos, particularmente que implican a tejidos embrionario y vascular.

*Características de la proteína codificada por el gen n.º: 22*

El producto de traducción de este gen comparte homología de secuencia con el receptor adrenérgico alfa 1C que se piensa que es importante en la transmisión de señales neuronales.

Este gen se expresa principalmente en el ganglio linfático mamario y en menor grado en cáncer uterino y tumor testicular.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos neurológicos. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del neurológico, ganglio linfático mamario, cáncer uterino y testículos, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por

ejemplo, tejido mamario, tejido linfático, tejido uterino, y testículos y otro tejido reproductor, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular y la homología con el receptor adrenérgico alfa 1C indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para transmitir señales a las neuronas.

*Características de la proteína codificada por el gen n.º: 23*

El producto de traducción de este gen comparte homología de secuencia con un receptor acoplado a proteína G que se piensa que es importante en la mediación de una amplia variedad de funciones fisiológicas y pertenece a una superfamilia génica con miembros que oscilan desde receptor de quimiocinas hasta receptor de bradicinina. Otro grupo ha clonado también recientemente este gen, denominando el gen homólogo del receptor de activación de plaquetas. (Véase el número de registro de GenBank 2580588). Los fragmentos de polipéptido preferidos comprenden la secuencia de aminoácidos: LSIIFLAFVSIDRCLQL (SEQ ID NO:117) y GSCFAIWAFIQENTNHRCVSIY LINLL TADFLLTLALPVKIVVDLGVAPWKLKIFHCQVTACLIYIN (SEQ ID NO:118). Se prefieren también fragmentos de polinucleótido que codifican para estos fragmentos de polipéptido.

Este gen se expresa principalmente en células inmunitarias, particularmente linfocitos.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos de linfocitos y otras células inmunitarias. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, linfocitos y otras células y tejido del sistema inmunitario, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:70 como restos: Asp-59 a Asn-65, Lys-72 a Trp-79, Tyr-110 a Val-121 y/o Ala-204 a Asn-215.

La distribución tisular y la homología con el receptor acoplado a proteína G indica que los polinucleótidos y polipéptidos que corresponden a este son útiles como receptor de quimiocinas en linfocitos que regulan la respuesta inmunitaria.

*Características de la proteína codificada por el gen n.º: 24*

El producto de traducción de este gen comparte homología de secuencia con proteína disulfuro isomerasa que se piensa que es importante en el plegamiento de proteínas y la interacción proteína-proteína. Este gen comparte también homología con genes que tienen dominios tiorredoxina. (Véase el número de registro 1943817). Este gen también se mapea en el cromosoma 9, y por tanto puede ser útil en un análisis de ligamiento como marcador para el cromosoma 9.

Este gen se expresa principalmente en tejidos tumorales y en menor grado en una amplia variedad de tejidos normales.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos debidos a una interacción proteína-proteína y un plegamiento de proteínas inapropiados. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del proceso tumorigénico, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Los epítopos preferidos incluyen los que comprenden una secuencia mostrada en SEQ ID NO:71 como restos: Glu-78 a Asn-83, Asp-91 a Gln-100, Glu-122 a Ser-128, Arg-137 a Pro-143, Asp-157 a Asn-162, Glu-168 a Asn-174, Ser-199 a Gly-206, Pro-213 a Ala-218, Glu-251 a Thr-257, Ser-353 a His-361, Gly-363 a Ala-375, Pro-382 a Phe-387 y/o Arg-401 a Leu-406.

La distribución tisular y la homología con proteína disulfuro isomerasa indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para regular el plegamiento de proteínas y la interacción proteína-proteína en tejidos tumorales.

*Características de la proteína codificada por el gen n.º: 25*

Este gen se expresa principalmente en leucocitos implicados en la defensa inmunitaria, incluyendo células T, macrófagos, neutrófilos y en menor grado en membrana sinovial, tumor de glándulas suprarrenales, tejido adiposo y placenta.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, defectos o trastornos en leucocitos. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente de los sistemas inmunitario y de defensa, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, leucocitos y otras células y tejidos del sistema inmunitario, membrana sinovial, glándula suprarrenal, tejido adiposo y placenta, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para regular la función leucocitaria y pueden usarse para el diagnóstico y tratamiento de trastornos en los sistemas inmunitario y de defensa.

*Características de la proteína codificada por el gen n.º: 26*

Este gen se expresa en una variedad de tejidos y tipos celulares, incluyendo cáncer de colon, cáncer de mama, neutrófilos, células T, médula espinal, fibroblastos y células endoteliales vasculares.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, cáncer, trastorno y anomalías en leucocitos y otros tejidos. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente las células implicadas en tumorigénesis y los sistemas de defensa inmunitaria, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, colon, tejido mamario, neutrófilos, células T y otras células sanguíneas, médula espinal y otro tejido del sistema nervioso, células endoteliales, tejido vascular, y fibroblastos, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para el diagnóstico y tratamiento de cáncer o trastornos del sistema inmunitario.

*Características de la proteína codificada por el gen n.º: 27*

El producto de traducción de este gen comparte homología de secuencia con polipéptido pancreático de ratón. (Véase el número de registro de GenBank 200464). Por tanto, es probable que este gen tenga una actividad similar a la del polipéptido pancreático de ratón. Los fragmentos de polipéptido preferidos comprenden la secuencia de aminoácidos: APLETMQNKPRAPQKRALPFPELELRDYASVLTRYSLGLRNKEPSLGHRWGTQKLGRSPC (SEQ ID NO:119). Se prefieren también fragmentos de polinucleótido que codifican para este fragmento de polipéptido.

Este gen se expresa principalmente en neutrófilos y en menor grado en células endoteliales inducidas.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos en la adhesión neutrófilos o leucocitos. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, neutrófilos y otras células sanguíneas, y células endoteliales, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra

de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para la regulación de la adhesión de neutrófilos o leucocitos a células endoteliales. Pueden usarse para diagnosticar o tratar trastornos asociados con neutrófilos y células endoteliales vasculares.

*Características de la proteína codificada por el gen n.º: 28*

Este gen se expresa principalmente en BPH prostática.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, hipertrofia benigna de la próstata. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema urogenital masculino, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, próstata, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los polinucleótidos y polipéptidos que corresponden a este gen son útiles para el diagnóstico y tratamiento de hipertrofia benigna de la próstata o cáncer de próstata.

*Características de la proteína codificada por el gen n.º 29*

El producto de traducción de este gen comparte homología de secuencia con C16C10.7, un gen de *C. elegans* similar a una proteína con dedos de zinc, una proteína implicada en la unión a ADN. Por tanto, se espera que esta proteína comparta ciertas actividades biológicas con C16C10.7 incluyendo actividades de unión a ADN.

Este gen se expresa principalmente en células T activadas y en menor grado en cerebro fetal, células amnióticas inducidas por TNF y epidídimo.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos inmunitarios o neurodegenerativos. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente de los sistemas nervioso central e inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, células T y otras células y tejido del sistema inmunitario, cerebro y otro tejido del sistema nervioso, células amnióticas, y epidídimo y otro tejido reproductor, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los productos de proteína de este gen son útiles para el diagnóstico y tratamiento de trastornos inmunitarios y/o neurodegenerativos y la promoción de la supervivencia y diferenciación de neuronas.

*Características de la proteína codificada por el gen n.º: 30*

Este gen se expresa principalmente en células T y en menor grado en médula ósea.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, trastornos inmunológicos incluyendo enfermedad autoinmunitaria. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema inmunitario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos y tipos celulares (por ejemplo, células T y otras células y tejido del sistema inmunitario, y médula ósea, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo,

en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno. Se cree que este gen se mapea en el cromosoma 4: Mapa de transcritos: WI-11395, Chr.4, D4S395-D4S414; mapa Whitehead: WI-11395, Chr.4, 498,0 cR; entradas de dbSTS: G21269.

La distribución tisular indica que los productos de proteína de este gen son útiles para el diagnóstico y tratamiento de trastornos mediados de manera inmunológica ya que se piensa que desempeñan un papel en la proliferación, supervivencia, diferenciación y/o activación de una variedad de células hematopoyéticas, incluyendo células madre hematopoyéticas o progenitoras tempranas.

*Características de la proteína codificada por el gen n.º: 31*

Este gen se expresa principalmente en piel humana.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, cicatrización de heridas y cánceres de piel. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema integumentario, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, piel y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular indica que los productos de proteína de este gen son útiles para el diagnóstico y tratamiento de cánceres de piel y cicatrización de heridas.

*Características de la proteína codificada por el gen n.º: 32*

El producto de traducción de este gen comparte homología de secuencia con la prealbúmina lacrimal humana (número de registro de GenBank gil307518) y proteína de unión a moléculas de olor de rata (número de registro de GenBank gil207551), ambas de las cuales se piensa que son importantes en el transporte y la unión a moléculas.

Este gen se expresa principalmente en tumor endometrial.

Por tanto, los polinucleótidos y polipéptidos descritos en el presente documento son útiles como reactivos para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es) presente(s) en una muestra biológica y para el diagnóstico de enfermedades y estados que incluyen, pero no se limitan a, cánceres del endometrio, piel y sistema hematopoyético. De manera similar, polipéptidos y anticuerpos dirigidos frente a estos polipéptidos son útiles para proporcionar sondas inmunológicas para la identificación diferencial del/de los tejido(s) o tipo(s) celular(es). Para varios trastornos de los tejidos o células anteriores, particularmente del sistema hematopoyético, la expresión de este gen a niveles significativamente superiores o inferiores puede detectarse de manera rutinaria en ciertos tejidos (por ejemplo, células y tejido del sistema inmunitario, y endometrio y otro tejido del sistema reproductor, y tejidos cancerosos y dañados) o líquidos corporales (por ejemplo, suero, plasma, orina, líquido sinovial o líquido cefalorraquídeo) u otra muestra de células o tejidos tomada de un individuo que tiene un trastorno de este tipo, en relación con el nivel de expresión génica convencional, es decir, el nivel de expresión en tejido sano o líquido corporal de un individuo que no tiene el trastorno.

La distribución tisular y la homología con la familia génica de transporte y unión a moléculas indica que los productos de proteína de este gen son útiles para el diagnóstico y tratamiento de cánceres del endometrio y sistema hematopoyético así como para el tratamiento de trastornos autoinmunitarios tales como inflamación.

| Gen n.° | ID de clon de ADNc | N.° de depósito de la ATCC y fecha | Vector | NT SEQ ID NO: X | Sec. de NT total | NT en 5′ de la sec. del clon | NT en 3′ de la sec. del clon | NT en 5′ del codón de inic. | NT en 5′ de los primeros AA de la señal Pep | AA SEQ ID NO: Y | Primer AA de pép. señal | Último AA de pép. señal | Primer AA de la parte secretada | Último AA de ORF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | HSVBZ80 | 209075 22/05/97 | Uni-ZAP XR | 11 | 1169 | 64 | 1060 | 162 | 162 | 48 | 1 | 38 | 39 | 145 |
| 2 | HTAAU21 | 209075 22/05/97 | Uni-ZAP XR | 12 | 1310 | 1 | 1310 | 283 | 283 | 49 | 1 | 18 | 19 | 311 |
| 3 | HTLEK16 | 209075 22/05/97 | Uni-ZAP XR | 13 | 1139 | 19 | 1111 | | 251 | 50 | 1 | 21 | 22 | 46 |
| 4 | HUSIR91 | 209075 22/05/97 | pSport1 | 14 | 2271 | 743 | 2271 | 59 | 59 | 51 | 1 | 23 | 24 | 466 |
| 4 | HUSIR91 | 209075 22/05/97 | pSport1 | 43 | 2581 | 1035 | 2164 | 1148 | 1148 | 80 | 1 | 27 | 28 | 207 |
| 5 | HADMC21 | 209075 22/05/97 | pBluescript | 15 | 626 | 60 | 479 | 91 | 91 | 52 | 1 | 51 | 52 | 82 |
| 6 | HAGFM45 | 209075 22/05/97 | Uni-ZAP XR | 16 | 2118 | 1170 | 2058 | 1248 | 1248 | 53 | 1 | 16 | 17 | 62 |
| 7 | HAIBE65 | 209075 22/05/97 | Uni-ZAP XR | 17 | 1076 | 396 | 993 | 528 | 528 | 54 | 1 | 31 | 32 | 123 |
| 8 | HAQBH57 | 209075 22/05/97 | Uni-ZAP XR | 18 | 1379 | 420 | 1306 | 618 | 618 | 55 | 1 | 25 | 26 | 179 |

| Gen n.° | ID de clon de ADNc | N.° de depósito de la ATCC y fecha | Vector | NT SEQ ID NO: X | Sec. de NT total | NT en 5′ de la sec. del clon | NT en 3′ de la sec. del clon | NT en 5′ del codón de inic. | NT en 5′ de los prime-ros AA de la señal Pep | AA SEQ ID NO: Y | Primer AA de pép. señal | Último AA de pép. señal | Primer AA de la parte secre-tada | Último AA de ORF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | HATCX80 | 209075 22/05/97 | Uni-ZAP XR | 19 | 1337 | 47 | 1337 | 199 | 199 | 56 | 1 | 18 | 19 | 286 |
| 10 | HCFLQ84 | 209075 22/05/97 | pSport1 | 20 | 1390 | 237 | 1390 | 410 | 410 | 57 | 1 | 20 | 21 | 33 |
| 11 | HCFLS78 | 209075 22/05/97 | pSport1 | 21 | 1431 | 178 | 981 | 420 | 420 | 58 | 1 | 21 | 22 | 23 |
| 12 | HTADI12 | 209075 22/05/97 | Uni-ZAP XR | 22 | 2539 | 69 | 2539 | 104 | 104 | 59 | 1 | 27 | 28 | 46 |
| 13 | HEMCM42 | 209075 22/05/97 | Uni-ZAP XR | 23 | 1041 | 48 | 1007 | 58 | 58 | 60 | 1 | 29 | 30 | 113 |
| 14 | HEONP72 | 209075 22/05/97 | pSport1 | 24 | 1962 | 1 | 1947 | 181 | 181 | 61 | 1 | 19 | 20 | 31 |
| 15 | HFCDW34 | 209075 22/05/97 | Uni-ZAP XR | 25 | 1228 | 321 | 1228 | 525 | 525 | 62 | 1 | 24 | 25 | 80 |
| 16 | HTTEU91 | 209075 22/05/97 | Uni-ZAP XR | 26 | 1340 | 325 | 1340 | 15 | 15 | 63 | 1 | 18 | 19 | 103 |
| 17 | HHGBF89 | 209075 22/05/97 | Lambda ZAP II | 27 | 806 | 31 | 806 | 77 | 77 | 64 | 1 | 19 | 20 | 145 |

| Gen n.° | ID de clon de ADNc | N.° de depósito de la ATCC y fecha | Vector | NT SEQ ID NO: X | Sec. de NT total | NT en 5′ de la sec. del clon | NT en 3′ de la sec. del clon | NT en 5′ del codón de inic. | NT en 5′ de los primeros AA de la señal Pep | AA SEQ ID NO: Y | Primer AA de pép. señal | Último AA de pép. señal | Primer AA de la parte secretada | Último AA de ORF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 17 | HHGBF89 | 209075 22/05/97 | LambdaZAP II | 45 | 796 | 31 | 796 | 77 | 77 | 82 | 1 | 25 | 26 | 145 |
| 18 | HKIYQ65 | 209075 22/05/97 | pBluescript | 28 | 696 | 1 | 684 | 98 | 98 | 65 | 1 | 17 | 18 | 30 |
| 19 | HKMLN27 | 209075 22/05/97 | pBluescript | 29 | 1007 | 71 | 963 | 129 | 129 | 66 | 1 | 23 | 24 | 259 |
| 20 | HKIAC30 | 209022 08/05/97 | Uni-ZAP XR | 30 | 2017 | 126 | 2007 | 161 | 161 | 67 | 1 | | | 22 |
| 21 | HKIXB95 | 209022 08/05/97 | pBluescript | 31 | 699 | 196 | 699 | 230 | 230 | 68 | 1 | 22 | 23 | 26 |
| 22 | HLMIY86 | 209022 08/05/97 | Lambda ZAP II | 32 | 1264 | 1 | 1264 | | 342 | 69 | 1 | 16 | 17 | 28 |
| 23 | HLYAZ61 | 209022 08/05/97 | pSport1 | 33 | 997 | 74 | 997 | 205 | 205 | 70 | 1 | 20 | 21 | 215 |
| 24 | HMQDT36 | 209022 08/05/97 | Uni-ZAP XR | 34 | 1914 | 37 | 1897 | 192 | 192 | 71 | 1 | 32 | 33 | 406 |
| 25 | HNEDF25 | 209022 08/05/97 | Uni-ZAP XR | 35 | 1020 | 11 | 1020 | | 211 | 72 | 1 | | | 8 |

| Gen n.º | ID de clon de ADNc | N.º de depósito de la ATCC y fecha | Vector | NT SEQ ID NO: X | Sec. de NT total | NT en 5′ de la sec. del clon | NT en 3′ de la sec. del clon | NT en 5′ del codón de inic. | NT en 5′ de los prime-ros AA de la señal Pep | AA SEQ ID NO: Y | Primer AA de pép. señal | Último AA de pép. señal | Primer AA de la parte secre-tada | Último AA de ORF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 26 | HNFET17 | 209022 08/05/97 | Uni-ZAP XR | 36 | 781 | 31 | 781 | 100 | 100 | 73 | 1 | | | 33 |
| 27 | HNHCR46 | 209022 08/05/97 | Uni-ZAP XR | 37 | 966 | 507 | 948 | | 576 | 74 | 1 | 29 | 30 | 56 |
| 28 | HPWAS91 | 209022 08/05/97 | Uni-ZAP XR | 38 | 416 | 1 | 416 | 95 | 95 | 75 | 1 | 24 | 25 | 25 |
| 29 | HWTAW41 | 209022 08/05/97 | Uni-ZAP XR | 39 | 1114 | 804 | 1114 | 843 | 843 | 76 | 1 | | | 14 |
| 30 | HBMUT52 | 209022 08/05/97 | Uni-ZAP XR | 40 | 602 | 142 | 602 | 204 | 204 | 77 | 1 | 26 | 27 | 32 |
| 31 | HERAG83 | 209022 08/05/97 | Uni-ZAP XR | 41 | 970 | 1 | 970 | 110 | 110 | 78 | 1 | 22 | 23 | 22 |
| 32 | HETFI51 | 209022 08/05/97 | Uni-ZAP XR | 42 | 1002 | 1 | 1002 | 43 | 43 | 79 | 1 | 21 | 22 | 172 |
| 32 | HETFI51 | 209022 08/05/97 | Uni-ZAP XR | 47 | 981 | 1 | 981 | 23 | 23 | 84 | 1 | 17 | 18 | 30 |

La tabla 1 resume la información correspondiente a cada "gen n.º" descrito anteriormente. La secuencia de nucleótidos identificada como "NT SEQ ID NO:X" se ensambló a partir de secuencias parcialmente homólogas ("solapantes") obtenidas del "ID de clon de ADNc" identificado en la tabla 1 y, en algunos casos, de clones de ADN relacionados adicionales. Las secuencias solapantes se ensamblaron en una única secuencia contigua de alta redundancia (habitualmente de tres a cinco secuencias solapantes en cada posición de nucleótido), dando como resultado una secuencia final identificada como SEQ ID NO:X.

El ID de clon de ADNc se depositó en la fecha y se le dio el correspondiente número de depósito enumerado en "n.º: Z de depósito de la ATCC y fecha". Algunos de los depósitos contienen múltiples clones diferentes correspondientes al mismo gen. "Vector" se refiere al tipo de vector contenido en el ID de clon de ADNc.

"Sec. de NT total" se refiere al número total de nucleótidos en el cóntigo identificado por "gen n.º". El clon depositado puede contener todas o la mayoría de estas secuencias, reflejado por la posición de nucleótidos indicada como "NT en 5' de la sec. del clon" y el "NT en 3' de la sec. del clon" de SEQ ID NO:X. La posición de nucleótidos de SEQ ID NO:X del supuesto codón de iniciación (metionina) se identifica como "NT en 5' del codón de iniciación". De manera similar, la posición de nucleótidos de SEQ ID NO:X de la secuencia señal pronosticada se identifica como "NT en 5' de los primeros AA de pép. señal".

La secuencia de aminoácidos traducida, que comienza con la metionina, se identifica como "AA SEQ ID NO:Y", aunque también pueden traducirse fácilmente otros marcos de lectura usando técnicas de biología molecular conocidas.

La primera y última posición de aminoácidos de SEQ ID NO:Y del péptido señal pronosticado se identifica como "primer AA de pép. señal" y "último AA de pép. señal". La primera posición de aminoácidos pronosticada de SEQ ID NO:Y de la parte secretada se identifica como "primer AA pronosticado de la parte secretada". Finalmente, la posición de aminoácidos de SEQ ID NO:Y del último aminoácido en el marco de lectura abierto se identifica como "último AA de ORF".

SEQ ID NO:X y la SEQ ID NO:Y traducida son suficientemente exactas y por lo demás adecuadas para una variedad de usos bien conocidos en la técnica y descritos adicionalmente a continuación. Por ejemplo, SEQ ID NO:X es útil para diseñar sondas de hibridación de ácido nucleico que detectarán secuencias de ácido nucleico contenidas en SEQ ID NO:X o el ADNc contenido en el clon depositado. Estas sondas se hibridarán también con moléculas de ácido nucleico en muestras biológicas, permitiendo de ese modo una variedad de métodos forenses y de diagnóstico de la invención. De manera similar, pueden usarse polipéptidos identificados a partir de SEQ ID NO:Y para generar anticuerpos que se unen específicamente a las proteínas secretadas codificadas por los clones de ADNc identificados en la tabla 1.

No obstante, las secuencias de ADN generadas mediante reacciones de secuenciación pueden contener errores de secuenciación. Los errores existen como nucleótidos mal identificados, o como inserciones o deleciones de nucleótidos en la secuencia de ADN generada. Los nucleótidos delecionados o insertados erróneamente provocan desplazamientos del marco en los marcos de lectura de la secuencia de aminoácidos pronosticada. En estos casos, la secuencia de aminoácidos pronosticada diverge de la secuencia de aminoácidos real, aún cuando la secuencia de ADN generada puede ser mayor del 99,9% idéntica a la secuencia de ADN real (por ejemplo, inserción o deleción de una base en un marco de lectura abierto de más de 1000 bases).

Por consiguiente, para aquellas aplicaciones que requieren precisión en la secuencia de nucleótidos o la secuencia de aminoácidos, la presente invención proporciona no sólo la secuencia de nucleótidos generada identificada como SEQ ID NO:23 y la secuencia de aminoácidos traducida pronosticada identificada como SEQ ID NO:60, sino también una muestra de ADN de plásmido que contiene un ADNc humano de la invención depositado en la ATCC, tal como se expone en la tabla 1. La secuencia de nucleótidos de cada clon depositado puede determinarse fácilmente secuenciando el clon depositado según métodos conocidos. La secuencia de aminoácidos pronosticada puede verificarse entonces a partir de tales depósitos. Además, la secuencia de aminoácidos de la proteína codificada por un clon particular puede determinarse también directamente secuenciando péptidos o expresando la proteína en una célula huésped adecuada que contiene el ADNc humano depositado, recogiendo la proteína y determinando su secuencia.

La presente invención se refiere también a los genes correspondientes a SEQ ID NO:23, SEQ ID NO:60, o el clon depositado. El gen correspondiente puede aislarse según métodos conocidos usando la información de secuencia dada a conocer en el presente documento. Tales métodos incluyen preparar sondas o cebadores a partir de la secuencia dada a conocer e identificar o amplificar el gen correspondiente a partir de fuentes apropiadas de material genómico. También se proporcionan en la presente invención homólogos de especie siempre que se encuentren dentro de la reivindicación 1. Pueden aislarse e identificarse homólogos de especie preparando sondas o cebadores adecuados a partir de las secuencias proporcionadas en el presente documento y seleccionando una fuente de ácido nucleico adecuada para el homólogo deseado.

Los polipéptidos de la invención pueden prepararse de cualquier manera adecuada. Tales polipéptidos incluyen polipéptidos aislados que se producen de manera natural, polipéptidos producidos de manera recombinante, polipéptidos producidos de manera sintética o polipéptidos producidos mediante una combinación de estos métodos. Se comprenden bien en la técnica medios para preparar tales polipéptidos.

Los polipéptidos pueden estar en forma de la proteína secretada, incluyendo la forma madura, o pueden ser una parte de una proteína mayor, tal como una proteína de fusión (véase a continuación). A menudo es ventajoso incluir una secuencia de aminoácidos adicional que contiene secuencias líder o secretoras, pro-secuencias, secuencias que ayudan en la purificación, tales como múltiples restos de histidina, o una secuencia adicional para lograr estabilidad durante la producción recombinante.

Los polipéptidos de la presente invención se proporcionan preferiblemente en una forma aislada, y preferiblemente están sustancialmente purificados. Una versión producida de manera recombinante de un polipéptido, incluyendo el polipéptido secretado, puede purificarse sustancialmente mediante el método de una etapa descrito en Smith y Johnson, Gene 67:31-40 (1988). También pueden purificarse polipéptidos de la invención a partir de fuentes naturales o recombinantes usando anticuerpos de la invención generados contra la proteína secretada en métodos que se conocen bien en la técnica.

*Secuencias señal*

Están disponibles métodos para pronosticar si una proteína tiene una secuencia señal, así como el punto de escisión para esa secuencia. Por ejemplo, el método de McGeoch, Virus Res. 3:271-286 (1985), usa la información de una región cargada N-terminal corta y una región no cargada posterior de la proteína completa (no escindida). El método de von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) usa la información de los restos que rodean el sitio de escisión, normalmente los restos -13 a +2, en los que +1 indica el extremo amino-terminal de la proteína secretada. La exactitud del pronóstico de los puntos de escisión de proteínas secretoras de mamíferos conocidas para cada uno de estos métodos está en el intervalo del 75-80% (von Heinje, citado anteriormente). Sin embargo, los dos métodos no siempre producen el/los mismo(s) punto(s) de escisión pronosticado(s) para una proteína dada.

En el presente caso, la secuencia de aminoácidos deducida del polipéptido secretado se analizó mediante un programa informático denominado SignalP (Henrik Nielsen *et al*., Protein Engineering 10:1-6 (1997)), que pronostica la ubicación celular de una proteína basándose en la secuencia de aminoácidos. Como parte de este pronóstico informático de la localización, pueden usarse los métodos de McGeoch y von Heinje. El análisis de la secuencia de aminoácidos de las proteínas secretadas descritas en el presente documento mediante este programa proporcionó los resultados mostrados en la tabla 1.

Sin embargo, tal como apreciaría un experto habitual, los sitios de escisión algunas veces varían de organismo a organismo y no pueden pronosticarse con absoluta seguridad. Por consiguiente, la presente invención proporciona polipéptidos secretados que tienen una secuencia mostrada en SEQ ID NO:60 que tienen un extremo N-terminal que comienza a 5 restos (es decir, + o -5 restos) del punto de escisión pronosticado. De manera similar, se reconoce también que, en algunos casos, la escisión de la secuencia señal de una proteína secretada no es totalmente uniforme, dando como resultado más de una especie secretada. Estos polipéptidos, y los polinucleótidos que codifican para tales polipéptidos, se contemplan por la presente invención.

Además, la secuencia señal identificada mediante el análisis anterior puede no pronosticar necesariamente la secuencia señal que se produce de manera natural. Por ejemplo, la secuencia señal que se produce de manera natural puede estar más en el sentido de 5' con respecto a la secuencia señal pronosticada. Sin embargo, es probable que la secuencia señal pronosticada pueda dirigir la proteína secretada al RE. Estos polipéptidos, y los polinucleótidos que codifican para tales polipéptidos, se contemplan por la presente invención.

*Variantes de polipéptidos y polinucleótidos*

"Variante" se refiere a un polinucleótido o polipéptido que difiere del polinucleótido o polipéptido de la presente invención, pero que conserva propiedades esenciales de los mismos. Generalmente, las variantes son estrechamente similares globalmente y, en muchas regiones, idénticas al polinucleótido o polipéptido de la presente invención.

Mediante un polinucleótido que tiene una secuencia de nucleótidos al menos, por ejemplo, el 95% "idéntica" a una secuencia de nucleótidos de referencia de la presente invención, se pretende que la secuencia de nucleótidos del polinucleótido sea idéntica a la secuencia de referencia excepto porque la secuencia de polinucleótido puede incluir hasta cinco mutaciones puntuales por cada 100 nucleótidos de la secuencia de nucleótidos de referencia que codifica el polipéptido. En otras palabras, para obtener un polinucleótido que tiene una secuencia de nucleótidos al menos el 95% idéntica a una secuencia de nucleótidos de referencia, hasta el 5% de los nucleótidos en la secuencia de referencia pueden delecionarse o sustituirse por otro nucleótido, o varios nucleótidos hasta el 5% de los nucleótidos totales en la secuencia de referencia pueden insertarse en la secuencia de referencia. La secuencia problema puede ser una secuencia completa mostrada en la tabla 1, el ORF (marco de lectura abierto) o cualquier fragmento especificado tal como se describe en el presente documento.

Como cuestión práctica, puede determinarse de manera convencional si cualquier polipéptido o molécula de ácido nucleico particular es al menos el 90%, 95%, 96%, 97%, 98% o 99% idéntico a una secuencia de nucleótidos de la presente invención usando programas informáticos conocidos. Un método preferido para determinar el mejor apareamiento global entre una secuencia problema (una secuencia de la presente invención) y una secuencia sujeto, también denominado alineación de secuencias global, puede determinarse usando el programa informático FASTDB basado en el algoritmo de Brutlag *et al*. (Comp. App. Biosci. (1990) 6:237-245). En una alineación de secuencias las secuencias

sujeto y problema son ambas secuencias de ADN. Puede compararse una secuencia de ARN convirtiendo U en T. El resultado de dicha alineación de secuencias global es en porcentaje de identidad. Parámetros preferidos usados en una alineación FASTDB de secuencias de ADN para calcular el porcentaje de identidad son:

Matriz = unitaria, k-tupla = 4, penalización por apareamiento erróneo = 1, penalización por unión = 30, longitud del grupo de aleatorización = 0, puntuación de punto de corte = 1, penalización por hueco = 5, penalización por tamaño de hueco 0,05, tamaño de ventana = 500 o la longitud de la secuencia de nucleótidos sujeto, la que sea más corta.

Si la secuencia sujeto es más corta que la secuencia problema debido a deleciones en 5' o 3', no debido a deleciones internas, debe hacerse una corrección manual en los resultados. Esto se debe a que el programa FASTDB no tiene en cuenta truncamientos en 5' y 3' de la secuencia sujeto cuando calcula el porcentaje de identidad. Para secuencias sujeto truncadas en los extremos 5' o 3', en relación con la secuencia problema, el porcentaje de identidad se corrige calculando el número de bases de la secuencia problema que están en 5' y 3' de la secuencia sujeto, que no están apareadas/alineadas, como un porcentaje de las bases totales de la secuencia problema. Si un nucleótido está apareado/alineado se determina mediante los resultados de la alineación de secuencias FASTDB. Este porcentaje se resta entonces del porcentaje de identidad, calculado mediante el programa FASTDB anterior usando los parámetros especificados, para llegar a una puntuación de porcentaje de identidad final. Esta puntuación corregida es la que se usa para los fines de la presente invención. Sólo las bases fuera de las bases en 5' y 3' de la secuencia sujeto, tal como se representa mediante la alineación FASTDB, que no están apareadas/alineadas con la secuencia problema, se calculan para los fines de ajustar manualmente la puntuación de porcentaje de identidad.

Por ejemplo, una secuencia sujeto de 90 bases se alinea con una secuencia problema de 100 bases para determinar el porcentaje de identidad. Las deleciones se producen en el extremo 5' de la secuencia sujeto y, por tanto, la alineación FASTDB no muestra un apareamiento/alineación de las primeras 10 bases en el extremo 5'. Las 10 bases no apareadas representan el 10% de la secuencia (número de bases en los extremos 5' y 3' no apareadas/número total de bases en la secuencia problema) de modo que se resta un 10% de la puntuación de porcentaje de identidad calculada mediante el programa FASTDB. Si las 90 bases restantes se apareasen perfectamente, el porcentaje de identidad final sería del 90%. En otro ejemplo, se compara una secuencia sujeto de 90 bases con una secuencia problema de 100 bases. Esta vez, las deleciones son deleciones internas de modo que no hay ninguna base en el extremo 5' o 3' de la secuencia sujeto que no esté apareada/alineada con la secuencia problema. En este caso, el porcentaje de identidad calculado mediante FASTDB no se corrige manualmente. Una vez más, sólo las bases en 5' y 3' de la secuencia sujeto que no están apareadas/alineadas con la secuencia problema se corrigen manualmente. No debe hacerse ninguna otra corrección manual para los fines de la presente invención.

Mediante un polipéptido que tiene una secuencia de aminoácidos al menos, por ejemplo, el 95% "idéntica" a una secuencia de aminoácidos problema de la presente invención, se pretende que la secuencia de aminoácidos del polipéptido sujeto sea idéntica a la secuencia problema excepto porque la secuencia de polipéptido sujeto puede incluir hasta cinco alteraciones de aminoácidos por cada 100 aminoácidos de la secuencia de aminoácidos problema. En otras palabras, para obtener un polipéptido que tiene una secuencia de aminoácidos al menos el 95% idéntica a una secuencia de aminoácidos problema, hasta el 5% de los restos de aminoácido en la secuencia sujeto pueden insertarse, delecionarse (indels) o sustituirse por otro aminoácido. Estas alteraciones de la secuencia de referencia pueden producirse en las posiciones amino o carboxilo terminales de la secuencia de aminoácidos de referencia o en cualquier sitio entre esas posiciones terminales, intercaladas o bien individualmente entre restos en la secuencia de referencia o bien en uno o más grupos contiguos dentro de la secuencia de referencia.

Como cuestión práctica, puede determinarse convencionalmente si cualquier polipéptido particular es al menos el 90%, 95%, 96%, 97%, 98% o 99% idéntico a, por ejemplo, la secuencia de aminoácidos mostrada en la tabla 1 o a la secuencia de aminoácidos codificada por el clon de ADN depositado usando programas informáticos conocidos. Un método preferido para determinar el mejor apareamiento global entre una secuencia problema (una secuencia de la presente invención) y una secuencia sujeto, también denominado alineación de secuencias global, puede determinarse usando el programa informático FASTDB basado en el algoritmo de Brutlag *et al*. (Comp. App. Biosci. (1990) 6:237-245). En una alineación de secuencias, las secuencias sujeto y problema son o bien ambas secuencias de nucleótidos o bien ambas secuencias de aminoácidos. El resultado de dicha alineación de secuencias global es en porcentaje de identidad. Parámetros preferidos usados en una alineación de aminoácidos FASTDB son: Matriz = PAM 0, k-tupla = 2, penalización por apareamiento erróneo = 1, penalización por unión = 20, longitud de grupo de aleatorización = 0, puntuación de punto de corte = 1, tamaño de ventana = longitud de secuencia, penalización por hueco = 5, penalización por tamaño de hueco = 0,05, tamaño de ventana = 500 o la longitud de la secuencia de aminoácidos sujeto, la que sea más corta.

Si la secuencia sujeto es más corta que la secuencia problema debido a deleciones N- o C-terminales, no debido a deleciones internas, debe hacerse una corrección manual en los resultados. Esto se debe a que el programa FASTDB no tiene en cuenta truncamientos en N- y C-terminales de la secuencia sujeto cuando calcula el porcentaje de identidad global. Para secuencias sujeto truncadas en los extremos N- y C-terminales, en relación con la secuencia problema, el porcentaje de identidad se corrige calculando el número de restos de la secuencia problema que son N- y C-terminales con respecto a la secuencia sujeto, que no están apareados/alineados con un resto sujeto correspondiente, como un porcentaje de las bases totales de la secuencia problema. Si un resto está apareado/alineado se determina mediante los resultados de la alineación de secuencias FASTDB. Este porcentaje se resta entonces del porcentaje de identidad, calculado mediante el programa FASTDB anterior usando los parámetros especificados, para llegar a una puntuación

de porcentaje de identidad final. Esta puntuación de porcentaje de identidad final es la que se usa para los fines de la presente invención. Sólo los restos en los extremos N- y C-terminales de la secuencia sujeto, que no están apareados/alineados con la secuencia problema, se consideran para los fines de ajustar manualmente la puntuación de porcentaje de identidad. Es decir, sólo las posiciones de restos problema fuera de los restos N- y C-terminales más lejanos de la secuencia sujeto.

Por ejemplo, una secuencia sujeto de 90 restos de aminoácido se alinea con una secuencia problema de 100 restos para determinar el porcentaje de identidad. La deleción se producen en el extremo N-terminal de la secuencia sujeto y, por tanto, la alineación FASTDB no muestra un apareamiento/alineación de los primeros 10 restos en el extremo N-terminal. Los 10 restos no apareados representan el 10% de la secuencia (número de restos en los extremos N- y C-terminales no apareados/número total de restos en la secuencia problema) de modo que se resta un 10% de la puntuación de porcentaje de identidad calculada mediante el programa FASTDB. Si los 90 restos restantes se apareasen perfectamente, el porcentaje de identidad final sería del 90%. En otro ejemplo, se compara una secuencia sujeto de 90 restos con una secuencia problema de 100 restos. Esta vez, las deleciones son deleciones internas de modo que no hay ningún resto en los extremos N- o C-terminales de la secuencia sujeto que no esté apareado/alineado con la secuencia problema. En este caso, el porcentaje de identidad calculado mediante FASTDB no se corrige manualmente. Una vez más, sólo las posiciones de restos fuera de los extremos N- y C-terminales de la secuencia sujeto, tal como se representa en la alineación FASTDB, que no están apareadas/alineadas con la secuencia problema se corrigen manualmente. No debe hacerse ninguna otra corrección manual para los fines de la presente invención.

Las variantes pueden contener alteraciones en las regiones codificantes, las regiones no codificantes o ambas. Se prefieren especialmente variantes de polinucleótidos que contienen alteraciones que producen deleciones, adiciones o sustituciones silenciosas, pero que no alteran las propiedades o actividades del polipéptido codificado. Se prefieren variantes de nucleótidos producidas por sustituciones silenciosas debidas a la degeneración del código genético. Además, se prefieren también variantes en las que se sustituyen, delecionan o añaden 5-10, 1-5 ó 1-2 aminoácidos. Pueden producirse variantes de polinucleótidos por una variedad de motivos, por ejemplo, para optimizar la expresión de codones para un huésped particular (cambio de codones en el ARNm humano a los preferidos por un huésped bacteriano tal como *E. coli*).

Las variantes que se producen de manera natural se denominan "variantes alélicas", y hacen referencia a una de varias formas alternativas de un gen que ocupa un locus dado en un cromosoma de un organismo. (Genes II, Lewin, B., ed., John Wiley & Sons, Nueva York (1985)). Estas variantes alélicas pueden variar o bien al nivel de polinucleótido y/o bien al de polipéptido. Alternativamente, pueden producirse variantes que se producen de manera no natural mediante técnicas de mutagénesis o mediante síntesis directa.

Usando métodos conocidos de tecnología de ADN recombinante e ingeniería de proteínas, pueden generarse variantes para mejorar o alterar las características de los polipéptidos de la presente invención. Por ejemplo, pueden delecionarse uno o más aminoácidos del extremo N-terminal o C-terminal de la proteína secretada sin pérdida sustancial de la función biológica. Los autores de Ron *et al*., J. Biol. Chem. 268: 2984-2988 (1993), notificaron proteínas KGF variantes que tenían actividad de unión a heparina incluso tras delecionar 3, 8 ó 27 restos de aminoácido amino-terminales. De manera similar, el interferón gamma presentaba una actividad hasta diez veces superior tras delecionar 8-10 restos de aminoácido del extremo carboxilo-terminal de esta proteína. (Dobeli *et al*., J. Biotechnology 7:199-216 (1988)).

Además, numerosas pruebas demuestran que las variantes conservan una actividad biológica similar a la de la proteína que se produce de manera natural. Por ejemplo, Gayle y colaboradores (J. Biol. Chem 268:22105-22111 (1993)) realizaron un análisis mutacional extenso de la citocina humana IL-1a. Usaron mutagénesis al azar para generar más de 3.500 mutantes de IL-1a individuales que promediaban 2,5 cambios de aminoácidos por variante a lo largo de toda la longitud de la molécula. Se examinaron mutaciones múltiples en cada posible posición de aminoácido. Los investigadores encontraron que "la mayor parte de la molécula podría alterarse con poco efecto sobre la actividad o bien biológica o bien de unión". (Véase el resumen). De hecho, sólo 23 secuencias de aminoácidos únicas, de más de 3.500 secuencias de nucleótidos examinadas, produjeron una proteína que difería significativamente en su actividad con respecto al tipo natural.

Además, aunque delecionar uno o más aminoácidos del extremo N-terminal o C-terminal de un polipéptido dé como resultado una modificación o pérdida de una o más funciones biológicas, pueden conservarse todavía otras actividades biológicas. Por ejemplo, la capacidad de una variante de deleción para inducir y/o unirse a anticuerpos que reconocen la forma secretada se conservará probablemente cuando se eliminan menos de la mayoría de los restos de la forma secretada del extremo N-terminal o C-terminal. Puede determinarse fácilmente si un polipéptido particular que carece de restos N-terminales o C-terminales de una proteína conserva tales actividades inmunogénicas mediante métodos de rutina descritos en el presente documento o conocidos de otra forma en la técnica.

Por tanto, la invención incluye además variantes de polipéptidos dentro de los ámbitos de las reivindicaciones que muestran una actividad biológica sustancial. Tales variantes incluyen deleciones, inserciones, inversiones, repeticiones y sustituciones seleccionadas según normas generales conocidas en la técnica de modo que tengan poco efecto sobre la actividad. Por ejemplo, se proporcionan directrices referentes a cómo realizar sustituciones de aminoácidos fenotípicamente silenciosas en Bowie, J. U. *et al*., Science 247:1306-1310 (1990), en el que los autores indican que hay dos estrategias principales para estudiar la tolerancia de una secuencia de aminoácidos a los cambios.

La primera estrategia se aprovecha de la tolerancia de sustituciones de aminoácidos por selección natural durante el proceso de evolución. Comparando secuencias de aminoácidos en diferentes especies, pueden identificarse aminoácidos conservados. Estos aminoácidos conservados probablemente son importantes para la función de la proteína. En cambio, las posiciones de aminoácidos en las que se han tolerado sustituciones por selección natural indican que estas posiciones no son críticas para la función de la proteína. Por tanto, podrían modificarse posiciones que toleran la sustitución de aminoácidos mientras que se mantiene la actividad biológica de la proteína.

La segunda estrategia usa la ingeniería genética para introducir cambios de aminoácidos en posiciones específicas de un gen clonado para identificar regiones críticas para la función de la proteína. Por ejemplo, pueden usarse mutagénesis dirigida al sitio o mutagénesis de barrido de alanina (introducción de mutaciones de alanina únicas en cada resto de la molécula). (Cunningham y Wells, Science 244:1081-1085 (1989)). Entonces, las moléculas mutantes resultantes pueden someterse a prueba para determinar su actividad biológica.

Tal como establecen los autores, estas dos estrategias han revelado que las proteínas son sorprendentemente tolerantes de sustituciones de aminoácidos. Los autores indican además que es probable que los cambios de aminoácidos sean permisivos en ciertas posiciones de aminoácidos en la proteína. Por ejemplo, los restos de aminoácido más enterrados (dentro de la estructura terciaria de la proteína) requieren cadenas laterales no polares, mientras que se conservan generalmente pocas características de cadenas laterales de superficie. Además, las sustituciones de aminoácidos conservativas toleradas implican el reemplazo de los aminoácidos alifáticos o hidrófobos Ala, Val, Leu y Ile; el reemplazo de los restos de hidroxilo Ser y Thr; el reemplazo de los restos ácidos Asp y Glu; el reemplazo de los restos de amida Asn y Gln, el reemplazo de los restos básicos Lys, Arg e His; el reemplazo de los restos aromáticos Phe, Tyr y Trp, y el reemplazo de los aminoácidos de pequeño tamaño Ala, Ser, Thr, Met y Gly. Además de la sustitución de aminoácidos conservativa, las variantes de la presente invención tal como se definen en las reivindicaciones incluyen (i) sustituciones por uno o más de los restos de aminoácido no conservados, en las que los restos de aminoácido sustituidos pueden estar codificados o no por el código genético, o (ii) la sustitución por uno o más de restos de aminoácido que tienen un grupo sustituyente, o (iii) la fusión del polipéptido maduro con otro compuesto, tal como un compuesto para aumentar la estabilidad y/o solubilidad del polipéptido (por ejemplo, polietilenglicol), o (iv) la fusión del polipéptido con aminoácidos adicionales, tales como un péptido de la región de fusión Fc de IgG, o una secuencia líder o secretora, o una secuencia que facilita la purificación. Se considera que tales polipéptidos variantes están dentro del alcance de los expertos en la técnica a partir de las enseñanzas en el presente documento.

Por ejemplo, variantes de polipéptidos que contienen sustituciones de aminoácidos de aminoácidos cargados por otros aminoácidos cargados o neutros pueden producir proteínas con características mejoradas, tales como menos agregación. La agregación de formulaciones farmacéuticas tanto reduce la actividad como aumenta el aclaramiento debido a la actividad inmunogénica del agregado. (Pinckard *et al*., Clin. Exp. Immunol. 2:331-340 (1967); Robbins *et al*., Diabetes 36: 838-845 (1987); Cleland *et al*., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993)).

*Fragmentos de polipéptidos y polinucleótidos*

En la presente descripción, un "fragmento de polinucleótido" se refiere a un polinucleótido corto que tiene una secuencia de ácido nucleico contenida en el clon depositado o mostrada en SEQ ID NO:23. Los fragmentos de nucleótidos cortos tienen al menos aproximadamente 15 nt, al menos aproximadamente 20 nt, al menos aproximadamente 30 nt, o al menos aproximadamente 40 nt de longitud. Un fragmento de "al menos 20 nt de longitud", por ejemplo, pretende incluir 20 o más bases contiguas de la secuencia de ADNc contenida en el clon depositado o la secuencia de nucleótidos mostrada en SEQ ID NO:23. Estos fragmentos de nucleótidos son útiles como sondas de diagnóstico y cebadores tal como se trata en el presente documento. Fragmentos más grandes (por ejemplo, 150, 500, 600 nucleótidos) son parte de la invención.

Además, los ejemplos representativos de fragmentos de polinucleótidos de la descripción incluyen, por ejemplo, fragmentos que tienen una secuencia de desde aproximadamente el número de nucleótido 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950 ó 951-1000 de SEQ ID NO:23 o el ADNc contenido en el clon depositado. En este contexto, "aproximadamente" incluye los intervalos particularmente mencionados, más grandes o más pequeños en varios (5, 4, 3, 2 ó 1) nucleótidos, en cualquier extremo terminal o en ambos extremos terminales. Preferiblemente, estos fragmentos codifican para un polipéptido que tiene actividad biológica. Más preferiblemente, estos polinucleótidos pueden usarse como sondas o cebadores tal como se trata en el presente documento.

En la presente descripción, un "fragmento de polipéptido" se refiere a una secuencia de aminoácidos corta contenida en SEQ ID NO:60 o codificada por el ADNc contenido en el clon depositado. Los fragmentos de proteína pueden ser "independientes", o estar comprendidos dentro de un polipéptido más grande del que el fragmento forma una parte o región, lo más preferiblemente como una única región continua. Los ejemplos representativos de fragmentos de polipéptidos de la descripción incluyen, por ejemplo, fragmentos de desde aproximadamente el número de aminoácido 1-20, 21-40, 41-60, 61-80 u 81-100. Además, los fragmentos de polipéptidos de la invención pueden tener aproximadamente 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140 ó 150 aminoácidos de longitud. En este contexto, "aproximadamente" incluye los intervalos particularmente mencionados, más grandes o más pequeños en varios (5, 4, 3, 2 ó 1) aminoácidos, en cualquier extremo o en ambos extremos.

Los fragmentos de polipéptidos preferidos incluyen la proteína secretada así como la forma madura. Los fragmentos de polipéptidos preferidos adicionales incluyen la proteína secretada o la forma madura que tienen una serie continua de restos delecionados de los extremos amino o carboxilo terminales, o ambos. Por ejemplo, puede delecionarse cualquier número de aminoácidos, que oscila desde 1-60, del extremo amino terminal de o bien el polipéptido secretado o bien la forma madura. De manera similar, puede delecionarse cualquier número de aminoácidos, que oscila desde 1-30, del extremo carboxilo terminal de la proteína secretada o la forma madura. Además, se prefiere cualquier combinación de las deleciones en los extremos amino y carboxilo terminales anteriores. De manera similar, se prefieren también fragmentos de polinucleótidos que codifican para estos fragmentos de polipéptidos.

Se prefieren también fragmentos de polinucleótidos y polipéptidos caracterizados por dominios estructurales o funcionales, tales como fragmentos que comprenden hélice alfa y regiones que forman hélice alfa, lámina beta y regiones que forman lámina beta, giros y regiones que forman giros, espirales y regiones que forman espirales, regiones hidrófilas, regiones hidrófobas, regiones alfa anfipáticas, regiones beta anfipáticas, regiones flexibles, regiones que forman superficies, región de unión a sustrato y regiones de alto índice antigénico. Los fragmentos de polipéptidos de SEQ ID NO:60 que se encuentran dentro de dominios conservados se contemplan específicamente por la presente invención. Además, se contemplan también fragmentos de polinucleótidos que codifican para estos dominios.

Otros fragmentos preferidos son fragmentos biológicamente activos. Fragmentos biológicamente activos son los que presentan actividad similar, pero no necesariamente idéntica, a una actividad del polipéptido de la presente invención. La actividad biológica de los fragmentos puede incluir una actividad deseada mejorada, o una actividad no deseada disminuida.

*Epítopos y anticuerpos*

En la presente descripción, "epítopos" se refiere a fragmentos de polipéptidos que tienen actividad antigénica o inmunogénica en un animal, especialmente en un ser humano. Una realización preferida de un fragmento de polipéptido de la presente invención se refiere a fragmentos que comprenden un epítopo. En consecuencia, se prefieren asimismo polinucleótidos de la presente invención que codifican para tales fragmentos. Una región de una molécula de proteína a la que un anticuerpo puede unirse se define como "epítopo antigénico". En cambio, un "epítopo inmunogénico" se define como una parte de una proteína que provoca una respuesta de anticuerpos. (Véase, por ejemplo, Geysen *et al*., Proc. Natl. Acad. Sci. EE.UU 81:3998-4002 (1983)).

Pueden producirse fragmentos que funcionan como epítopos mediante cualquier medio convencional. (Véase, por ejemplo, Houghten, R. A., Proc. Natl. Acad. Sci. EE.UU 82:5131-5135 (1985) descrito adicionalmente en la patente estadounidense n.º 4.631.211).

En la presente descripción, los epítopos antigénicos contienen preferiblemente una secuencia de al menos siete, más preferiblemente al menos nueve y lo más preferiblemente entre aproximadamente 15 y aproximadamente 30 aminoácidos. Los péptidos antigénicos son útiles para generar anticuerpos, incluyendo anticuerpos monoclonales, que se unen específicamente al epítopo. (Véase, por ejemplo, Wilson *et al*., Cell 37: 767-778 (1984); Sutcliffe, J. G. *et al*., Science 219:660-666 (1983)).

De manera similar, pueden usarse epítopos inmunogénicos para inducir anticuerpos según métodos bien conocidos en la técnica. (Véase, por ejemplo, Sutcliffe *et al*., citado anteriormente; Wilson *et al*., citado anteriormente; Chow, M. *et al*., Proc. Natl. Acad. Sci. USA 82:910-914; y Bittle, F. J. *et al*., J. Gen. Virol. 66:2347-2354 (1985)). Un epítopo inmunogénico preferido incluye la proteína secretada. Los epítopos inmunogénicos pueden presentarse junto con una proteína portadora, tal como una albúmina, a un sistema animal (tal como conejo o ratón) o, si son suficientemente largos (al menos aproximadamente 25 aminoácidos), sin un portador. Sin embargo, se ha demostrado que epítopos inmunogénicos que comprenden tan sólo de 8 a 10 aminoácidos son suficientes para generar anticuerpos que pueden unirse a, como mínimo, epítopos lineales en un polipéptido desnaturalizado (por ejemplo, en inmunotransferencias de tipo Western).

Tal como se usa en el presente documento, el término "anticuerpo" (Ab) o "anticuerpo monoclonal" (AcM) pretende incluir moléculas intactas así como fragmentos de anticuerpos (tales como, por ejemplo, fragmentos Fab y F (ab')2) que pueden unirse específicamente a una proteína. Los fragmentos Fab y F(ab')2 carecen del fragmento Fc del anticuerpo intacto, se aclaran más rápidamente de la circulación y pueden tener menos unión tisular no específica que un anticuerpo intacto. (Wahl *et al*., J. Nucl. Med. 24:316-325 (1983)). Por tanto, se prefieren estos fragmentos, así como los productos de un FAB u otra biblioteca de expresión de inmunoglobulinas. Además, los anticuerpos de la presente invención incluyen anticuerpos quiméricos, de cadena sencilla y humanizados.

*Proteínas de fusión*

Cualquier polipéptido de la presente invención puede usarse para generar proteínas de fusión. Por ejemplo, el polipéptido de la presente invención, cuando se fusiona con una segunda proteína, puede usarse como etiqueta antigénica. Pueden usarse anticuerpos generados contra el polipéptido de la presente invención para detectar indirectamente la segunda proteína mediante la unión al polipéptido. Además, debido a que las proteínas secretadas seleccionan como diana ubicaciones celulares basándose en señales de comunicación, los polipéptidos de la presente invención pueden usarse como moléculas de direccionamiento una vez fusionados con otras proteínas.

Los ejemplos de dominios que pueden fusionarse con polipéptidos de la presente invención incluyen no sólo secuencias señal heterólogas, sino también otras regiones funcionales heterólogas. No se necesita necesariamente que la fusión sea directa, sino que puede producirse a través de secuencias de ligador.

Además, también pueden modificarse por ingeniería genética proteínas de fusión para mejorar características del polipéptido de la presente invención. Por ejemplo, puede añadirse una región de aminoácidos adicionales, particularmente aminoácidos cargados, al extremo N-terminal del polipéptido para mejorar la estabilidad y persistencia durante la purificación a partir de la célula huésped o el almacenamiento y manejo posterior. Además, pueden añadirse restos peptídicos al polipéptido para facilitar la purificación. Tales regiones pueden eliminarse antes de la preparación final del polipéptido. La adición de restos peptídicos para facilitar el manejo de los polipéptidos son técnicas familiares y de rutina en la técnica.

Además, pueden combinarse polipéptidos de la presente invención, incluyendo fragmentos, y específicamente epítopos, con partes del dominio constante de inmunoglobulinas (IgG), dando como resultado polipéptidos quiméricos. Estas proteínas de fusión facilitan la purificación y muestran una semivida aumentada *in vivo*. Un ejemplo notificado describe proteína quiméricas que consisten en los primeros dos dominios del polipéptido CD4 humano y diversos dominios de las regiones constantes de las cadenas pesada o ligera de inmunoglobulinas de mamífero. (Documento EP A 394.827; Traunecker *et al*., Nature 331:84-86 (1988)). Proteínas de fusión que tienen estructuras diméricas unidas por disulfuros (debido a la IgG) pueden ser también más eficaces en la unión a y la neutralización de otras moléculas, que la proteína o el fragmento de proteína secretada monomérica sola. (Fountoulakis *et al*., J. Biochem. 270:3958-3964 (1995)).

De manera similar, el documento EP-A-O 464 533 (homólogo canadiense 2045869) da a conocer proteínas de fusión que comprenden diversas partes de una región constante de moléculas de inmunoglobulina junto con otra proteína humana o parte de la misma. En muchos casos, la parte Fc en una proteína de fusión es beneficiosa en la terapia y el diagnóstico, y por tanto puede dar como resultado, por ejemplo, propiedades farmacocinéticas mejoradas. (Documento EP-A 0232 262). Alternativamente, se desearía la deleción de la parte Fc tras haberse expresado, detectado y purificado la proteína de fusión. Por ejemplo, la parte Fc puede dificultar la terapia y el diagnóstico si la proteína de fusión se usa como antígeno para inmunizaciones. En el descubrimiento de fármacos, por ejemplo, se han fusionado proteínas humanas, tales como hIL-5, con partes Fc para el fin de realizar ensayos de selección de alto rendimiento para identificar antagonistas de hIL-5. (Véase, D. Bennett *et al*., J. Molecular Recognition 8:52-58 (1995); K. Johanson *et al*., J. Biol. Chem. 270:9459-9471 (1995)).

Además, los polipéptidos de la presente invención pueden fusionarse con secuencias marcadoras, tal como un péptido que facilita la purificación del polipéptido fusionado. En realizaciones preferidas, la secuencia de aminoácidos marcadora es un péptido de hexa-histidina, tal como la etiqueta proporcionada en un vector pQE (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), entre otros, muchos de los cuales están comercialmente disponibles. Tal como se describe en Gentz *et al*., Proc. Natl. Acad. Sci. EE.UU 86:821-824 (1989), por ejemplo, la hexa-histidina proporciona una purificación conveniente de la proteína de fusión. Otra etiqueta peptídica útil para la purificación, la etiqueta "HA", corresponde a un epítopo derivado de la proteína hemaglutinina del virus de la gripe. (Wilson *et al*., Cell 37:767 (1984)).

Por tanto, cualquiera de estas fusiones anteriores puede modificarse por ingeniería genética usando los polinucleótidos o los polipéptidos de la presente invención.

*Vectores, células huésped y producción de proteínas*

La presente invención se refiere también a vectores que contienen el polinucleótido de la presente invención, a células huésped y a la producción de polipéptidos por técnicas recombinantes. El vector puede ser, por ejemplo, un fago, plásmido, vector viral o retroviral. Los vectores retrovirales pueden ser de replicación competente o de replicación defectuosa. En el último caso, la propagación viral se producirá generalmente sólo en células huésped de complementación.

Los polinucleótidos pueden unirse a un vector que contiene un marcador seleccionable para su propagación en un huésped. Generalmente, se introduce un vector de plásmido en un precipitado, tal como un precipitado de fosfato de calcio, o en un complejo con un lípido cargado. Si el vector es un virus, puede empaquetarse *in vitro* usando una línea celular de empaquetamiento apropiada y luego transducirse al interior de células huésped.

El inserto de polinucleótido debe estar operativamente unido a un promotor apropiado, tal como el promotor PL del fago lambda, los promotores lac, trp, phoA y tac de *E. coli*, los promotores tardío y temprano de SV40 y promotores de LTR retrovirales, por nombrar unos cuantos. El experto conocerá otros promotores adecuados. Los constructos de expresión contendrán además sitios para la iniciación y terminación de la transcripción y, en la región transcrita, un sitio de unión a ribosomas para la traducción. La parte codificante de los transcritos expresados por los constructos incluirán preferiblemente un codón de iniciación de la traducción al principio y un codón de terminación (UAA, UGA o UAG) apropiadamente colocado al final del polipéptido que va a traducirse.

Tal como se indica, los vectores de expresión incluirán preferiblemente al menos un marcador seleccionable. Tales marcadores incluyen dihidrofolato reductasa, G418 o genes de resistencia a neomicina para el cultivo de células

eucariotas y genes de resistencia a tetraciclina, kanamicina o ampicilina para cultivar en *E. coli* y otras bacterias. Los ejemplos representativos de huéspedes apropiados incluyen, pero no se limitan a, células bacterianas, tales como células de *E. coli*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como células de levadura; células de insecto tales como células S2 de *Drosophila* y células Sf9 de *Spodoptera*; células animales tales como células CHO, COS, 293 y de melanoma de Bowes; y células vegetales. Se conocen en la técnica condiciones y medios de cultivo apropiados para las células huésped descritas anteriormente.

Entre los vectores preferidos para su uso en bacterias se incluyen pQE70, pQE60 y pQE-9, disponibles de QIAGEN, Inc.; vectores pBluescript, vectores Phagescript, pNH8A, pNH16a, pNH18A, pNH46A, disponibles de Stratagene Cloning Systems, Inc.; y ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 disponibles de Pharmacia Biotech, Inc. Entre los vectores eucariotas preferidos están pWLNEO, pSV2CAT, pOG44, pXT1 y pSG disponibles de Stratagene; y pSVK3, pBPV, pMSG y pSVL disponibles de Pharmacia. Otros vectores adecuados resultarán fácilmente evidentes para el experto en la técnica.

La introducción del constructo en la célula huésped puede efectuarse mediante transfección con fosfato de calcio, transfección mediada por DEAE-dextrano, transfección mediada por lípidos catiónicos, electroporación, transducción, infección u otros métodos. Tales métodos se describen en muchos manuales de laboratorio convencionales, tales como Davis *et al*., Basic Methods in Molecular Biology (1986). Se contempla específicamente que los polipéptidos de la presente invención puedan expresarse de hecho por una célula huésped que carece de un vector recombinante.

Un polipéptido de esta invención puede recuperarse y purificarse a partir de cultivos celulares recombinantes mediante métodos bien conocidos incluyendo precipitación con sulfato de amonio o etanol, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía en fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía de hidroxilapatita y cromatografía con lectinas. Lo más preferiblemente, se emplea cromatografía líquida de alta resolución ("HPLC") para la purificación.

Los polipéptidos de la presente invención, y preferiblemente la forma secretada, pueden recuperarse también de: productos purificados a partir de fuentes naturales, incluyendo líquidos corporales, tejidos y células, ya estén directamente aislados o cultivados; productos de procedimientos sintéticos químicos; y productos producidos por técnicas recombinantes a partir de un huésped eucariota o procariota, incluyendo, por ejemplo, células bacterianas, de levadura, de plantas superiores, de insecto y de mamífero. Dependiendo del huésped empleado en un procedimiento de producción recombinante, los polipéptidos de la presente invención pueden estar glicosilado o pueden estar no glicosilados. Además, los polipéptidos de la invención pueden incluir también un resto de metionina modificado inicial, en algunos casos como resultado de procesos mediados por el huésped. Por tanto, se conoce bien en la técnica que la metionina N-terminal codificada por el codón de iniciación de la traducción se elimina generalmente con alta eficacia de cualquier proteína tras su traducción en todas las células eucariotas. Aunque la metionina N-terminal en la mayoría de las proteínas también se elimina eficazmente en la mayoría de los procariotas, para algunas proteínas, este proceso de eliminación en procariotas es ineficaz, dependiendo de la naturaleza del aminoácido al que está unida covalentemente la metionina N-terminal.

*Usos de los polinucleótidos*

Cada uno de los polinucleótidos identificados en el presente documento puede usarse de numerosas formas como reactivo. La siguiente descripción debe considerarse a modo de ejemplo y utiliza técnicas conocidas.

Los polinucleótidos de la presente invención son útiles para la identificación de cromosomas. Existe una necesidad creciente de identificar nuevos marcadores de cromosomas, puesto que están disponibles en la actualidad pocos reactivos de marcaje de cromosomas, basándose en datos de secuencia reales (polimorfismos de repetición). Cada polinucleótido de la presente invención puede usarse como marcador de cromosomas.

En resumen, pueden mapearse secuencias en cromosomas preparando cebadores de PCR (preferiblemente de 15-25 pb) a partir de la secuencia mostrada en SEQ ID NO:23. Pueden seleccionarse cebadores usando análisis informático de modo que los cebadores no abarquen más de un exón pronosticado en el ADN genómico. Estos cebadores se usan entonces para la selección por PCR de híbridos de células somáticas que contienen cromosomas humanos individuales. Sólo los híbridos que contienen el gen humano correspondiente a la SEQ ID NO:23 producirán un fragmento amplificado.

De manera similar, híbridos somáticos proporcionan un método rápido de mapeo por PCR de los polinucleótidos en cromosomas particulares. Pueden asignarse tres o más clones al día usando un único termociclador. Además, puede lograrse la sublocalización de los polinucleótidos con paneles de fragmentos de cromosomas específicos. Otras estrategias de mapeo génico que pueden usarse incluyen hibridación *in situ*, selección previa con cromosomas marcados clasificados por flujo y preselección mediante hibridación para construir bibliotecas de ADNc específico de cromosoma.

Puede lograrse también una ubicación cromosómica precisa de los polinucleótidos usando hibridación por fluorescencia *in situ* (FISH) de una extensión cromosómica en metafase. Esta técnica usa polinucleótidos de tan sólo 500 ó 600 bases; sin embargo, se prefieren polinucleótidos de 2.000-4.000 pb. Para una revisión de esta técnica, véase Verma *et al*., "Human Chromosomes: a Manual of Basic Techniques", Pergamon Press, Nueva York (1988).

Para el mapeo de cromosomas, los polinucleótidos pueden usarse individualmente (para marcar un único cromosoma o un único sitio en ese cromosoma) o en paneles (para marcar múltiples sitios y/o múltiples cromosomas). Polinucleótidos preferidos corresponden a las regiones no codificantes de los ADNc porque es más probable que las secuencias codificantes estén conservadas dentro de las familias génicas, aumentando así la posibilidad de hibridación cruzada durante el mapeo cromosómico.

Una vez que se ha mapeado un polinucleótido en una ubicación cromosómica precisa, puede usarse la posición física del polinucleótido en un análisis de ligamiento. El análisis de ligamiento establece la herencia conjunta entre una ubicación cromosómica y la presentación de una enfermedad particular. (Se encuentran datos de mapeo de enfermedades, por ejemplo, en V. McKusick, Mendelian Inheritance in Man (disponible online a través de la Biblioteca Médica Welch de la Universidad Johns Hopkins)). Suponiendo una resolución de mapeo de 1 megabase y un gen por 20 kb, un ADNc localizado de manera precisa en una región cromosómica asociada con la enfermedad podría ser uno de 50-500 posibles genes causantes.

Por tanto, una vez establecida la herencia conjunta, pueden examinarse diferencias en el polinucleótido y el gen correspondiente entre individuos afectados y no afectados. En primer lugar, se examinan alteraciones estructurales visibles en los cromosomas, tales como deleciones o translocaciones, en extensiones de cromosomas o mediante PCR. Si no existen alteraciones estructurales, se determina la presencia de mutaciones puntuales. Las mutaciones observadas en algunos o todos los individuos afectados, pero no en individuos normales, indica que la mutación puede provocar la enfermedad. Sin embargo, se requiere la secuenciación completa del polipéptido y el gen correspondiente de varios individuos normales para distinguir la mutación de un polimorfismo. Si se identifica un nuevo polimorfismo, puede usarse este polipéptido polimórfico para un análisis de ligamiento adicional.

Además, puede evaluarse la expresión aumentada o disminuida del gen en individuos afectados en comparación con individuos no afectados usando polinucleótidos de la presente invención. Cualquiera de estas alteraciones (expresión alterada, redisposición cromosómica o mutación) puede usarse como marcador de diagnóstico o pronóstico.

Además de lo anterior, puede usarse un polinucleótido puede controlar la expresión génica a través de la formación de triple hélice o ARN o ADN antisentido. Ambos métodos se basan en la unión del polinucleótido a ADN o ARN. Para estas técnicas, polinucleótidos preferidos tienen habitualmente de 20 a 40 bases de longitud y son complementarios a o bien la región del gen implicado en la transcripción (triple hélice - véase Lee *et al*., Nucl. Acids Res. 6:3073 (1979); Cooney *et al*., Science 241:456 (1988); y Dervan *et al*., Science 251:1360 (1991)) o bien al propio ARNm (antisentido - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). La formación de triple hélice da como resultado de manera óptima una interrupción de la transcripción de ARN a partir de ADN, mientras que la hibridación con ARN antisentido bloquea la traducción de una molécula de ARNm en un polipéptido. Ambas técnicas son eficaces en sistemas modelo, y la información dada a conocer en el presente documento puede usarse para diseñar polinucleótidos de triple hélice o antisentido en un esfuerzo por tratar la enfermedad.

Los polinucleótidos de la presente invención también son útiles en terapia génica. Un objetivo de la terapia génica es insertar un gen normal en un organismo que tiene un gen defectuoso, en un esfuerzo por corregir el defecto genético. Los polinucleótidos descritos en la presente invención ofrecen un medio para seleccionar como diana tales defectos genéticos de una manera sumamente precisa. Otro objetivo es insertar un nuevo gen que no estaba presente en el genoma del huésped, produciendo de ese modo un nuevo rasgo en la célula huésped.

Los polinucleótidos también son útiles para identificar individuos a partir de muestras biológicas mínimas. El ejército de los Estados Unidos, por ejemplo, está considerando el uso del polimorfismo de longitud de fragmentos de restricción (RFLP) para la identificación de su personal. En esta técnica, se digiere el ADN genómico de un individuo con una o más enzimas de restricción, y se estudia con sonda en una transferencia de tipo Southern para producir bandas únicas para identificar al personal. Este método no se ve afectado por las limitaciones actuales de las "placas de identificación" que pueden perderse, cambiarse o robarse, lo que dificulta una identificación positiva. Los polinucleótidos de la presente invención pueden usarse como marcadores de ADN adicionales para RFLP.

Los polinucleótidos de la presente invención pueden usarse también como una alternativa a RFLP, determinando la secuencia de ADN base por base real de partes seleccionadas del genoma de un individuo. Estas secuencias pueden usarse para preparar cebadores de PCR para amplificar y aislar tal ADN seleccionado, que puede entonces secuenciarse. Usando esta técnica, pueden identificarse individuos porque cada individuo tendrá un único conjunto de secuencias de ADN. Una vez que se establece una base de datos de ID única para un individuo, puede hacerse una identificación positiva de ese individuo, vivo o muerto, a partir de muestras de tejido extremadamente pequeñas.

La biología forense se beneficia también del uso de técnicas de identificación basadas en ADN tal como se da a conocer en el presente documento. Pueden amplificarse secuencias de ADN tomadas de muestras biológicas muy pequeñas tales como tejidos, por ejemplo, cabello o piel, o líquidos corporales, por ejemplo, sangre, saliva, semen, etc., usando PCR. En una técnica de la técnica anterior, se usan secuencias génicas amplificadas a partir de locus polimórficos, tales como el gen de HLA clase II DQa, en biología forense para identificar individuos. (Erlich, H., PCR Technology, Freeman y Co. (1992)). Una vez amplificados estos locus polimórficos específicos, se digieren con una o más enzimas de restricción, produciendo un conjunto de bandas de identificación en una

transferencia de tipo Southern estudiada con sonda con ADN correspondiente al gen de HLA clase II DQa. De manera similar, pueden usarse polinucleótidos de la presente invención como marcadores polimórficos para fines forenses.

Existe también una necesidad de reactivos que puedan identificar la fuente de un tejido particular. Tal necesidad surge, por ejemplo, en medicina forense cuando se presenta un tejido de origen desconocido. Reactivos apropiados pueden comprender, por ejemplo, sondas o cebadores de ADN específicos para un tejido particular preparado a partir de las secuencias de la presente invención. Paneles de tales reactivos pueden identificar el tejido por la especie y/o por el tipo de órgano. De un modo similar, estos reactivos pueden usarse para investigar cultivos tisulares para detectar contaminación.

Por último, los polinucleótidos de la presente invención pueden usarse como marcadores de peso molecular en geles de Southern, como sondas de diagnóstico para detectar la presencia de un ARNm específico en un tipo celular particular, como sonda para "restar" secuencias conocidas en el proceso de descubrimiento de polinucleótidos novedosos, para seleccionar y preparar oligómeros para su unión a un "chip génico" u otro soporte, para generar anticuerpos anti-ADN usando técnicas de inmunización con ADN, y como antígeno para provocar una respuesta inmunitaria.

*Usos de los polipéptidos*

Cada uno de los polipéptidos identificados en el presente documento puede usarse de numerosos modos. La siguiente descripción debe considerarse a modo de ejemplo y utiliza técnicas conocidas.

Un polipéptido de la presente invención puede usarse para someter a ensayo los niveles de proteína en una muestra biológica usando técnicas basadas en anticuerpos. Por ejemplo, puede estudiarse la expresión de proteínas en tejidos con métodos inmunohistológicos clásicos. (Jalkanen, M., *et al*., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., *et al*., J. Cell. Biol. 105:3087-3096 (1987)). Otros métodos basados en anticuerpos útiles para detectar la expresión génica de proteínas incluyen inmunoensayos, tales como el ensayo inmunoabsorbente ligado a enzimas (ELISA) y el radioinmunoensayo (RIA). Se conocen en la técnica marcadores de ensayo de anticuerpos adecuados e incluyen marcadores enzimáticos, tales como, glucosa oxidasa, y radioisótopos, tales como yodo (125I, 121I), carbono (14C), azufre (35S), tritio (3H), indio (112In) y tecnecio (99mTc), y marcadores fluorescentes, tales como fluoresceína y rodamina, y biotina.

Además de someter a ensayo los niveles de proteínas secretadas en una muestra biológica, también pueden detectarse proteínas *in vivo* mediante obtención de imágenes. Los marcadores de anticuerpos o marcadores para la obtención de imágenes *in vivo* de proteínas incluyen aquellos detectables mediante radiografía X, RMN o REE. Para la radiografía X, los marcadores adecuados incluyen radioisótopos tales como bario o cesio, que emiten radiación detectable pero que no son abiertamente perjudiciales para el sujeto. Los marcadores adecuados para RMN y REE incluyen aquellos con un espín detectable característico, tales como deuterio, que puede incorporarse en el anticuerpo mediante marcaje de nutrientes para el hibridoma relevante.

Un fragmento de anticuerpo o anticuerpo específico de proteína que se ha marcado con un resto de obtención de imágenes detectable apropiado, tal como un radioisótopo (por ejemplo, 131I, 112In, 99mTc), una sustancia radioopaca, o un material detectable mediante resonancia magnética nuclear, se introduce (por ejemplo, por vía parenteral, por vía subcutánea o por vía intraperitoneal) en el mamífero. Se entenderá en la técnica que el tamaño del sujeto y el sistema de obtención de imágenes usado determinarán la cantidad de resto de obtención de imágenes necesario para producir imágenes de diagnóstico. En el caso de un resto de radioisótopo, para un sujeto humano, la cantidad de radiactividad inyectada oscilará normalmente desde aproximadamente 5 hasta 20 milicuries de 99mTc. El fragmento de anticuerpo o anticuerpo marcado se acumulará entonces preferentemente en la ubicación de células que contienen la proteína específica. La obtención de imágenes de tumores *in vivo* se describe en S.W. Burchiel *et al*., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments". (Capítulo 13 en Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel y B. A. Rhodes, eds., Masson Publishing Inc. (1982)).

Por tanto, la invención proporciona un método de diagnóstico de un trastorno, que implica (a) someter a ensayo la expresión de un polipéptido de la presente invención en células o líquido corporal de un individuo; (b) comparar el nivel de expresión génica con un nivel de expresión génica convencional, mediante lo cual un aumento o descenso en el nivel de expresión génica del polipéptido sometido a ensayo en comparación con el nivel de expresión convencional es indicativo de un trastorno.

Además, los polipéptidos de la presente invención pueden usarse para tratar una enfermedad. Por ejemplo, puede administrarse a pacientes un polipéptido de la presente invención en un esfuerzo por reemplazar niveles ausentes o disminuidos del polipéptido (por ejemplo, insulina), para complementar niveles ausentes o disminuidos de un polipéptido diferente (por ejemplo, hemoglobina S para hemoglobina B), para inhibir la actividad de un polipéptido (por ejemplo, un oncogén), para activar la actividad de un polipéptido (por ejemplo, mediante unión a un receptor), para reducir la actividad de un receptor unido a membrana compitiendo con el mismo por ligando libre (por ejemplo, receptores de TNF soluble usados para reducir la inflamación), o para provocar una respuesta deseada (por ejemplo, crecimiento de vasos sanguíneos).

De manera similar, pueden usarse también anticuerpos dirigidos frente a un polipéptido de la presente invención para tratar una enfermedad. Por ejemplo, la administración de un anticuerpo dirigido frente a un polipéptido de la presente invención puede unirse y reducir la sobreproducción del polipéptido. De manera similar, la administración de un anticuerpo puede activar el polipéptido, tal como mediante unión a un polipéptido unido a una membrana (receptor).

Por último, los polipéptidos de la presente invención pueden usarse como marcadores de peso molecular en geles de SDS-PAGE o en columnas de filtración en gel de tamiz molecular usando métodos bien conocidos para los expertos en la técnica. Los polipéptidos pueden usarse también para generar anticuerpos, que a su vez se usan para medir la expresión de proteínas a partir de una célula recombinante, como un modo de evaluar la transformación de la célula huésped. Además, los polipéptidos de la presente invención pueden usarse para someter a prueba las siguientes actividades biológicas.

*Actividades biológicas*

Los polinucleótidos y polipéptidos de la presente invención pueden usarse en ensayos para someter a prueba una o más actividades biológicas. Si estos polinucleótidos y polipéptidos presentan actividad en un ensayo particular, es probable que estas moléculas estén implicadas en las enfermedades asociadas con la actividad biológica. Por tanto, los polinucleótidos y polipéptidos podrían usarse para tratar la enfermedad asociada.

*Actividad inmunitaria*

Un polipéptido o polinucleótido de la presente descripción puede ser útil en el tratamiento de deficiencias o trastornos del sistema inmunitario, activando o inhibiendo la proliferación, diferenciación o movilización (quimiotaxis) de células inmunitarias. Las células inmunitarias se desarrollan a través de un proceso denominado hematopoyesis, que produce células mieloides (plaquetas, glóbulos rojos, neutrófilos y macrófagos) y linfoides (linfocitos B y T) a partir de células madre pluripotentes. La etiología de estos trastornos o deficiencias inmunitarios puede ser genética, somática, tal como cáncer o algunos trastornos autoinmunitarios, adquiridos (por ejemplo, por quimioterapia o toxinas) o infecciosos. Además, un polinucleótido o polipéptido de la presente descripción puede usarse como marcador o detector de un trastorno o una enfermedad del sistema inmunitario particular.

Un polinucleótido o polipéptido de la presente descripción puede ser útil en el tratamiento o la detección de deficiencias o trastornos de células hematopoyéticas. El polipéptido o polinucleótido podría usarse para aumentar la diferenciación y proliferación de células hematopoyéticas, incluyendo las células madre pluripotentes, en un esfuerzo por tratar aquellos trastornos asociados con una disminución en ciertos (o muchos) tipos de células hematopoyéticas. Los ejemplos de síndromes de deficiencia inmunológica incluyen, pero no se limitan a: trastornos de proteínas sanguíneas (por ejemplo agammaglobulinemia, disgammaglobulinemia), ataxia telangiectasia, inmunodeficiencia variable común, síndrome de Digeorge, infección por VIH, infección por VLTH-VLB, síndrome de deficiencia de adhesión leucocitaria, linfopenia, disfunción de fagocito bactericida, inmunodeficiencia combinada grave (SCID), trastorno de Wiskott-Aldrich, anemia, trombocitopenia o hemoglobinuria.

Además, un polipéptido o polinucleótido de la presente descripción podría usarse para modular la actividad hemostática (la detención de la hemorragia) o trombolítica (formación del coágulo). Por ejemplo, aumentando la actividad hemostática o trombolítica, el polinucleótido o polipéptido podría usarse para tratar trastornos de la coagulación sanguínea (por ejemplo, afibrinogenemia, deficiencias de factores), trastornos de plaquetas sanguíneas (por ejemplo tromboctopenia), o heridas que resultan de traumatismos, cirugía u otras causas. Alternativamente, un polinucleótido o polipéptido de la presente descripción que pueda disminuir la actividad hemostática o trombolítica podría usarse para inhibir o disolver la coagulación. Estas moléculas podrían ser importantes en el tratamiento de ataques al corazón (infarto), accidentes cerebrovasculares o cicatrización.

Un polinucleótido o polipéptido de la presente descripción puede ser útil también en el tratamiento o detección de trastornos autoinmunitarios. Muchos trastornos autoinmunitarios resultan del reconocimiento inapropiado de material propio como foráneo por células inmunitarias. Este reconocimiento inapropiado da como resultado una respuesta inmunitaria que conduce a la destrucción del tejido del huésped. Por tanto, la administración de un polipéptido o polinucleótido de la presente descripción que inhibe una respuesta inmunitaria, particularmente la proliferación, diferenciación o qumiotaxis de células T, puede ser una terapia eficaz en la prevención de trastornos autoinmunitarios.

Los ejemplos de trastornos autoinmunitarios que pueden tratarse o detectarse incluyen, pero no se limitan a: enfermedad de Addison, anemia hemolítica, síndrome antifosfolipídico, artritis reumatoide, dermatitis, encefalomielitis alérgica, glomerulonefritis, síndrome de Goodpasture, enfermedad de Graves, esclerosis múltiple, miastenia grave, neuritis, oftalmia, penfigoide ampolloso, pénfigo, poliendocrinopatías, púrpura, enfermedad de Reiter, síndrome de Stiff-Man, tiroiditis autoinmunitaria, lupus eritematoso sistémico, inflamación pulmonar autoinmunitaria, síndrome de Guillain-Barre, diabetes mellitus insulinodependiente y enfermedad ocular inflamatoria autoinmunitaria.

De manera similar, también pueden tratarse estados y reacciones alérgicas, tales como asma (particularmente asma alérgico) u otros problemas respiratorios, mediante un polipéptido o polinucleótido de la presente descripción. Además, estas moléculas pueden usarse para tratar anafilaxia, hipersensibilidad a una molécula antigénica o incompatibilidad de grupos sanguíneos.

Un polinucleótido o polipéptido de la presente descripción puede usarse también para tratar y/o prevenir el rechazo de órganos o la enfermedad de injerto contra huésped (EICH). El rechazo de órganos se produce mediante la destrucción por células inmunitarias del huésped del tejido trasplantado a través de una respuesta inmunitaria. De manera similar, una respuesta inmunitaria también está implicada en EICH aunque, en este caso, las células inmunitarias trasplantadas foráneas destruyen los tejidos del huésped. La administración de un polipéptido o polinucleótido de la presente descripción que inhibe una respuesta inmunitaria, particularmente la proliferación, diferenciación o quimiotaxis de células T, puede ser una terapia eficaz en la prevención del rechazo de órganos o EICH.

De manera similar, un polipéptido o polinucleótido de la presente descripción puede usarse también para modular la inflamación. Por ejemplo, el polipéptido o polinucleótido puede inhibir la proliferación y diferenciación de células implicadas en una respuesta inflamatoria. Estas moléculas pueden usarse para tratar estados inflamatorios, tanto estados crónicos como agudos, incluyendo inflamación asociada con infección (por ejemplo, choque septicémico, septicemia o síndrome de respuesta inflamatoria sistémica (SRIS)), lesión por isquemia-reperfusión, letalidad por endotoxinas, artritis, rechazo hiperagudo mediado por el complemento, nefritis, lesión pulmonar inducida por citocinas o quimiocinas, enfermedad inflamatoria del intestino, enfermedad de Crohn, o que resulta de una sobreproducción de citocinas (por ejemplo, TNF o IL-1).

*Trastornos hiperproliferativos*

Un polipéptido o polinucleótido de la presente descripción puede usarse para tratar o detectar trastornos hiperproliferativos, incluyendo neoplasmas. El polipéptido o polinucleótido puede inhibir la proliferación del trastorno a través de interacciones directas o indirectas. Alternativamente, el polipéptido o polinucleótido puede hacer que proliferen otras células que pueden inhibir el trastorno hiperproliferativo.

Por ejemplo, aumentando una respuesta inmunitaria, particularmente aumentando las cualidades antigénicas del trastorno hiperproliferativo o haciendo que proliferen, se diferencien o se movilicen células T, pueden tratarse trastornos hiperproliferativos. Esta respuesta inmunitaria puede aumentarse o bien potenciando una respuesta inmunitaria existente, o bien iniciando una nueva respuesta inmunitaria. Alternativamente, la disminución de una respuesta inmunitaria puede ser también un método de tratamiento de trastornos hiperproliferativos, tal como un agente quimioterápico.

Los ejemplos de trastornos hiperproliferativos que pueden tratarse o detectarse mediante un polinucleótido o polipéptido de la presente invención incluyen, pero no se limitan a neoplasmas ubicados en el: abdomen, hueso, mama, sistema digestivo, hígado, páncreas, peritoneo, glándulas endocrinas (suprarrenal, paratiroidea, hipofisaria, testículos, ovarios, timo, tiroides), ojo, cabeza y cuello, sistema nervioso (central y periférico), sistema linfático, pelvis, piel, tejido blando, bazo, tórax y sistema urogenital.

De manera similar, otros trastornos hiperproliferativos pueden tratarse o detectarse mediante un polinucleótido o polipéptido de la presente descripción. Los ejemplos de tales trastornos hiperproliferativos incluyen, pero no se limitan a: hipergammaglobulinemia, trastornos linfoproliferativos, paraproteinemias, púrpura, sarcoidosis, síndrome de Sezary, macroglobulinemia de Waldenstron, enfermedad de Gaucher, histiocitosis y cualquier otra enfermedad hiperproliferativa, además de neoplasia, ubicada en un sistema de órganos enumerado anteriormente.

*Enfermedades infecciosas*

Un polipéptido o polinucleótido de la presente descripción puede usarse para tratar o detectar agentes infecciosos. Por ejemplo, aumentando la respuesta inmunitaria, particularmente aumentando la proliferación y diferenciación de células B y/o T, pueden tratarse enfermedades infecciosas. La respuesta inmunitaria puede aumentarse o bien potenciando una respuesta inmunitaria existente, o bien iniciando una nueva respuesta inmunitaria. Alternativamente, el polipéptido o polinucleótido puede inhibir también directamente el agente infeccioso, sin provocar necesariamente una respuesta inmunitaria.

Los virus son un ejemplo de un agente infeccioso que puede provocar una enfermedad o síntomas que pueden tratarse o detectarse mediante un polinucleótido o polipéptido de la presente descripción. Los ejemplos de virus incluyen, pero no se limitan a las siguientes familias de virus de ADN y ARN: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (hepatitis), Herpesviridae (tal como, citomegalovirus, herpes simple, herpes zóster), Mononegavirus (por ejemplo, Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (por ejemplo, Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (tal como viruela o Vaccinia), Reoviridae (por ejemplo, Rotavirus), Retroviridae (VLTH-I, VLTH-II, Lentivirus) y Togaviridae (por ejemplo, Rubivirus). Los virus que se encuentran dentro de estas familias pueden provocar una variedad de enfermedades o síntomas, incluyendo, pero sin limitarse a: artritis, bronquiolitis, encefalitis, infecciones oculares (por ejemplo, conjuntivitis, queratitis), síndrome de fatiga crónica, hepatitis (A, B, C, E, activa crónica, delta), meningitis, infecciones oportunistas (por ejemplo, SIDA), neumonía, linfoma de Burkitt, varicela, fiebre hemorrágica, sarampión, paperas, parainfluenza, rabia, resfriado común, poliomielitis, leucemia, rubéola, enfermedades de transmisión sexual, enfermedades cutáneas (por ejemplo, sarcoma de Kaposi, verrugas) y viremia. Un polipéptido o polinucleótido de la presente descripción puede usarse para tratar o detectar cualquiera de estos síntomas o enfermedades.

De manera similar, los agentes bacterianos o fúngicos que pueden provocar una enfermedad o síntomas y que pueden tratarse o detectarse mediante un polinucleótido o polipéptido de la presente descripción incluyen, pero no se limitan a, las siguientes familias bacterianas Gram-negativas y Gram-positivas y hongos: Actinomycetales (por ejemplo, *Corynebacterium*, *Mycobacterium*, *Norcardia*), aspergilosis, Bacillaceae (por ejemplo, *Anthrax*, *Clostridium*), Bacteroidaceae, blastomicosis, *Bordetella*, *Borrelia*, brucelosis, candidiasis, *Campylobacter*, coccidioidomicosis, criptococosis, dermatomicosis, Enterobacteriaceae (*Klebsiella*, *Salmonella*, *Serratia*, *Yersinia*), *Erysipelothrix*, *Helicobacter*, legionelosis, leptospirosis, *Listeria*, Mycoplasmatales, Neisseriaceae (por ejemplo, *Acinetobacter*, *Gonorrhea*, meningococos), infecciones por Pasteurellacea (por ejemplo, *Actinobacillus*, *Heamophilus*, *Pasteurella*), *Pseudomonas*, Rickettsiaceae, Chlamydiaceae, sífilis y estafilococos. Estas familias bacterianas o fúngicas pueden provocar las siguientes enfermedades o síntomas, incluyendo, pero sin limitarse a: bacteremia, endocarditis, infecciones oculares (conjuntivitis, tuberculosis, uveitis), gingivitis, infecciones oportunistas (por ejemplo, infecciones relacionadas con el SIDA), paroniquia, infecciones relacionadas con prótesis, enfermedad de Reiter, infecciones de las vías respiratorias, tales como tos ferina o empiema, septicemia, enfermedad de Lyme, enfermedad por arañazo de gato, disentería, fiebre paratifoidea, envenenamiento alimentario, fiebre tifoidea, neumonía, gonorrea, meningitis, clamidiosis, sífilis, difteria, lepra, paratuberculosis, tuberculosis, lupus, botulismo, gangrena, tétanos, impétigo, fiebre reumática, escarlatina, enfermedades de transmisión sexual, enfermedades cutáneas (por ejemplo, celulitis, dermatomicosis), toxemia, infecciones de las vías urinarias, infecciones de heridas. Un polipéptido o polinucleótido de la presente descripción puede usarse para tratar o detectar cualquiera de estos síntomas o enfermedades.

Además, los agentes parasitarios que provocan una enfermedad o síntomas que pueden tratarse o detectarse mediante un polinucleótido o polipéptido de la presente descripción incluyen, pero no se limitan a, las siguientes familias: amebiasis, babesiosis, coccidiosis, criptosporidiosis, dientamoebiasis, durina, ectoparasitosis, giardiasis, helmintiasis, leishmaniosis, theileriosis, toxoplasmosis, tripanosomiasis y *Trichomonas*. Estos parásitos pueden provocar una variedad de enfermedades o síntomas, incluyendo, pero sin limitarse a: sarna, trombiculiasis, infecciones oculares, enfermedad intestinal (por ejemplo, disenteria, giardiasis), enfermedad hepática, enfermedad pulmonar, infecciones oportunistas (por ejemplo, relacionadas con el SIDA), malaria, complicaciones del embarazo y toxoplasmosis. Un polipéptido o polinucleótido de la presente descripción puede usarse para tratar o detectar cualquiera de estos síntomas o enfermedades.

Preferiblemente, el tratamiento usando un polipéptido o polinucleótido de la presente descripción podría ser o bien administrando una cantidad eficaz del polipéptido al paciente, o bien extrayendo células del paciente, administrando a las células el polinucleótido y devolviendo las células modificadas por ingeniería genética al paciente (terapia *ex vivo*). Además, el polipéptido o polinucleótido puede usarse como antígeno en una vacuna para generar una respuesta inmunitaria contra una enfermedad infecciosa.

*Regeneración*

Un polinucleótido o polipéptido de la presente descripción puede usarse para hacer que se diferencien, proliferen o se atraigan células, conduciendo a la regeneración de tejidos. (Véase, Science 276:59-87 (1997)). La regeneración de tejidos podría usarse para reparar, reemplazar o proteger tejido dañado por defectos congénitos, traumatismos (heridas, quemaduras, incisiones o úlceras), la edad, una enfermedad (por ejemplo osteoporosis, artrosis, enfermedad periodontal, insuficiencia hepática), cirugía, incluyendo cirugía plástica cosmética, fibrosis, lesión por reperfusión o daño sistémico por citocinas.

Los tejidos que podrían regenerarse usando la presente invención incluyen órganos (por ejemplo, páncreas, hígado, intestino, riñón, piel, endotelio), músculo (liso, esquelético o cardiaco), vascular (incluyendo endotelio vascular), tejido nervioso, hematopoyético y esquelético (hueso, cartílago, tendones y ligamentos). Preferiblemente, la regeneración se produce sin cicatrización o con disminución de la cicatrización. La regeneración también puede incluir angiogénesis.

Además, un polinucleótido o polipéptido de la presente descripción puede aumentar la regeneración de tejidos difíciles de curar. Por ejemplo, un aumento de la regeneración de tendones/ligamentos aceleraría el tiempo de recuperación tras el daño. El polinucleótido o polipéptido podría usarse también de manera profiláctica en un esfuerzo por evitar el daño. Las enfermedades específicas que podrían tratarse incluyen tendinitis, síndrome del túnel carpiano y otros defectos de tendones o ligamentos. Un ejemplo adicional de regeneración de tejido de heridas que no cicatrizan incluye úlceras de decúbito, úlceras asociadas con insuficiencia vascular, quirúrgica y heridas traumáticas.

De manera similar, también podría regenerarse tejido cerebral y nervioso usando un polinucleótido o polipéptido de la presente descripción para hacer que proliferen y se diferencien células nerviosas. Las enfermedades que podrían tratarse usando este método incluyen enfermedades del sistema nervioso central y periférico, neuropatías, o trastornos traumáticos y mecánicos (por ejemplo, trastornos de la médula espinal, traumatismo craneal, enfermedad cerebrovascular y accidente cerebrovascular). Específicamente, enfermedades asociadas con lesiones de nervios periféricos, neuropatía periférica (por ejemplo, que resulta de quimioterapia u otras terapias médicas), neuropatías localizadas y enfermedades del sistema nervioso central (por ejemplo, enfermedad de Alzheimer, enfermedad de Parkinson, enfermedad de Huntington, esclerosis lateral amiotrófica y síndrome de Shy-Drager), podrían tratarse todas usando el polinucleótido o polipéptido de la presente descripción.

*Quimiotaxis*

Un polinucleótido o polipéptido de la presente descripción puede tener actividad de quimiotaxis. Una molécula quimiotáctica atrae o moviliza células (por ejemplo, monocitos, fibroblastos, neutrófilos, células T, mastocitos, eosinófilos, células epiteliales y/o endoteliales) a un sitio particular en el organismo, tal como inflamación, infección o sitio de hiperproliferación. Las células movilizadas pueden entonces combatir y/o curar la anomalía o traumatismo particular.

Un polinucleótido o polipéptido de la presente descripción puede aumentar la actividad quimiotáctica de células particulares. Estas moléculas quimiotácticas pueden usarse entonces para tratar la inflamación, infección o trastornos hiperproliferativos, o cualquier trastorno del sistema inmunitario aumentando el número de células dirigidas a una ubicación particular en el organismo. Por ejemplo, pueden usarse moléculas quimiotácticas para tratar heridas u otro traumatismo en tejidos atrayendo células inmunitarias a la ubicación lesionada. Las moléculas quimiotácticas de la presente descripción también pueden atraer fibroblastos, que pueden usarse para tratar heridas.

Se contempla también que un polinucleótido o polipéptido de la presente descripción pueda inhibir la actividad quimiotáctica. Estas moléculas podrían usarse también para tratar trastornos. Por tanto, el polinucleótido o polipéptido podría usarse como inhibidor de la qumiotaxis.

*Actividad de unión*

Un polipéptido de la presente descripción puede usarse para seleccionar moléculas que se unen al polipéptido o moléculas a las que se une el polipéptido. La unión del polipéptido y la molécula puede activar (agonista), aumentar, inhibir (antagonista) o disminuir la actividad del polipéptido o la molécula unida. Los ejemplos de tales moléculas incluyen anticuerpos, oligonucleótidos, proteínas (por ejemplo, receptores) o moléculas pequeñas.

Preferiblemente, la molécula está estrechamente relacionada con el ligando natural del polipéptido, por ejemplo, un fragmento del ligando, o un sustrato natural, un ligando, un mimético estructural o funcional. (Véase, Coligan *et al*., Current Protocols in Immunology 1(2): capítulo 5 (1991)). De manera similar, la molécula puede estar estrechamente relacionada con el receptor natural al que se une el polipéptido, o al menos, un fragmento del receptor al que se une el polipéptido (por ejemplo, sitio activo). En cualquier caso, la molécula puede diseñarse racionalmente usando técnicas conocidas.

Preferiblemente, la selección de estas moléculas implica producir células apropiadas que expresan el polipéptido, o bien como una proteína secretada o bien en la membrana celular. Las células preferidas incluyen células de mamíferos, levaduras, *Drosophila* o *E. coli*. Las células que expresan el polipéptido (o la membrana celular que contiene el polipéptido expresado) se ponen en contacto entonces preferiblemente con un compuesto de prueba que contiene posiblemente la molécula para observar la unión, estimulación o inhibición de la actividad de o bien el polipéptido o bien la molécula.

El ensayo puede simplemente someter a prueba la unión de un compuesto candidato al polipéptido, en el que se detecta la unión mediante un marcador, o en un ensayo que implica competición con un competidor marcado. Además, el ensayo puede someter a prueba si el compuesto candidato da como resultado una señal generada por la unión al polipéptido.

Alternativamente, el ensayo puede llevarse a cabo usando preparaciones libres de células, polipéptidos/moléculas fijados a un soporte sólido, bibliotecas químicas o mezclas de productos naturales. El ensayo también puede comprender simplemente las etapas de mezclar un compuesto candidato con una disolución que contiene un polipéptido, medir la actividad o unión del polipéptido/molécula y comparar la actividad o unión del polipéptido/molécula con un patrón.

Preferiblemente, un ensayo ELISA puede medir la actividad o el nivel del polipéptido en una muestra (por ejemplo, una muestra biológica) usando un anticuerpo monoclonal o policlonal. El anticuerpo puede medir la actividad o el nivel del polipéptido o bien uniéndose, directa o indirectamente, al polipéptido o bien compitiendo con el polipéptido por un sustrato.

Todos estos ensayos anteriores pueden usarse como marcadores de diagnóstico o pronóstico. Las moléculas descubiertas usando estos ensayos pueden usarse para tratar una enfermedad o para ocasionar un resultado particular en un paciente (por ejemplo, crecimiento de vasos sanguíneos) activando o inhibiendo el polipéptido/molécula. Además, los ensayos pueden descubrir agentes que pueden inhibir o potenciar la producción del polipéptido a partir de tejidos o células adecuadamente manipulados.

Por tanto, la invención incluye un método de identificación de compuestos que se unen a un polipéptido de la invención que comprende las etapas de: (a) incubar un compuesto de unión candidato con un polipéptido de la invención; y (b) determinar si se ha producido unión. Además, la descripción incluye un método de identificación de agonistas/antagonistas que comprende las etapas de: (a) incubar un compuesto candidato con un polipéptido de la invención, (b) someter a ensayo una actividad biológica, y (b) determinar si se ha alterado una actividad biológica del polipéptido.

*Otras actividades*

Un polipéptido o polinucleótido de la presente invención puede aumentar o disminuir la diferenciación o proliferación de células madre embrionarias, además, tal como se trató anteriormente, del linaje hematopoyético.

Un polipéptido o polinucleótido de la presente invención pude usarse también para modular características de mamíferos, tales como la altura, el peso corporal, el color del pelo, el color de los ojos, la piel, el porcentaje de tejido adiposo, la pigmentación, el tamaño y la forma (por ejemplo, cirugía cosmética). De manera similar, un polipéptido o polinucleótido de la presente invención puede usarse para modular el metabolismo de mamíferos afectando al catabolismo, anabolismo, procesamiento, utilización y almacenamiento de energía.

Un polipéptido o polinucleótido de la presente invención puede usarse para cambiar el estado físico o el estado mental de un mamífero influyendo en los biorritmos, ritmos circadianos, depresión (incluyendo trastornos depresivos), tendencia a la violencia, tolerancia al dolor, capacidades reproductoras (preferiblemente mediante actividad de tipo activina o inhibina), niveles hormonales o endocrinos, apetito, líbido, memoria, estrés u otras cualidades cognitivas.

Un polipéptido o polinucleótido de la presente invención puede usarse también como conservante o aditivo alimentario, tal como para aumentar o disminuir las capacidades de almacenamiento, el contenido en grasa, lípidos, proteínas, hidratos de carbono, vitaminas, minerales, cofactores u otros componentes nutricionales.

*Otras cuestiones*

Se describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos aproximadamente 50 nucleótidos contiguos en la secuencia de nucleótidos de SEQ ID NO:23 tal como se define en la tabla 1.

También se describe una molécula de ácido nucleico en la que dicha secuencia de nucleótidos contiguos está incluida en la secuencia de nucleótidos de SEQ ID NO:23 en el intervalo de posiciones que comienza con el nucleótido en aproximadamente la posición del nucleótido en 5' de la secuencia del clon y que acaba con el nucleótido en aproximadamente la posición del nucleótido en 3' de la secuencia del clon tal como se define para SEQ ID NO:23 en la tabla 1.

También se describe una molécula de ácido nucleico en la que dicha secuencia de nucleótidos contiguos está incluida en la secuencia de nucleótidos de SEQ ID NO:23 en el intervalo de posiciones que comienza con el nucleótido en aproximadamente la posición del nucleótido en 5' del codón de iniciación y que acaba con el nucleótido en aproximadamente la posición del nucleótido en 3' de la secuencia del clon tal como se define para SEQ ID NO:23 en la tabla 1.

De manera similar se describe una molécula de ácido nucleico en la que dicha secuencia de nucleótidos contiguos está incluida en la secuencia de nucleótidos de SEQ ID NO:23 en el intervalo de posiciones que comienza con el nucleótido en aproximadamente la posición del nucleótido en 5' del primer aminoácido del péptido señal y que acaba con el nucleótido en aproximadamente la posición del nucleótido en 3' de la secuencia del clon tal como se define para SEQ ID NO:23 en la tabla 1.

Una realización preferida de la presente invención es una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos aproximadamente 150 nucleótidos contiguos en la secuencia de nucleótidos de SEQ ID NO:23.

Se describe además una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos aproximadamente 500 nucleótidos contiguos en la secuencia de nucleótidos de SEQ ID NO:23.

Se describe además una molécula de ácido nucleico que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a la secuencia de nucleótidos de SEQ ID NO:23 que comienza con el nucleótido en aproximadamente la posición del nucleótido en 5' del primer aminoácido del péptido señal y que acaba con el nucleótido en aproximadamente la posición del nucleótido en 3' de la secuencia del clon tal como se define para SEQ ID NO:23 en la tabla 1.

Se describe además una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a la secuencia de nucleótidos completa de SEQ ID NO:23.

También se describe una molécula de ácido nucleico aislada que hibrida en condiciones de hibridación rigurosas con una molécula de ácido nucleico, en la que dicha molécula de ácido nucleico que hibrida no hibrida en condiciones de hibridación rigurosas con una molécula de ácido nucleico que tiene una secuencia de nucleótidos que consiste sólo en restos de A o sólo en restos de T.

También se prefiere una molécula de ADN que comprende un clon de ADNc humano identificado como el depósito de la ATCC n.º 209075, molécula de ADN que está contenida en el material depositado en la Colección Americana de Cultivos Tipo y al que se le asignó número de depósito de la ATCC mostrado en la tabla 1 para dicho identificador de clon de ADNc.

También se describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en la secuencia de nucleótidos de un clon de ADNc humano, molécula de ADN que está contenida en depósito al que se le asignó el número de depósito de la ATCC 209075.

También se describe una molécula de ácido nucleico aislada, estando incluida dicha secuencia de al menos 50 nucleótidos contiguos en la secuencia de nucleótidos de la secuencia completa de marco de lectura abierto codificada por dicho clon de ADNc humano.

También se prefiere una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica una secuencia de al menos 150 nucleótidos contiguos en la secuencia de nucleótidos codificada por dicho clon de ADNc humano.

Se describe además una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica una secuencia de al menos 500 nucleótidos contiguos en la secuencia de nucleótidos codificada por dicho clon de ADNc humano.

Se describe además una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a la secuencia de nucleótidos completa codificada por dicho clon de ADNc humano.

Se describe además un método para detectar en una muestra biológica una molécula de ácido nucleico que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de nucleótidos de SEQ ID NO:23 tal como se define en la tabla 1; y una secuencia de nucleótidos codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075; método que comprende una etapa de comparar una secuencia de nucleótidos de al menos una molécula de ácido nucleico en dicha muestra con una secuencia seleccionada de dicho grupo y determinar si la secuencia de dicha molécula de ácido nucleico en dicha muestra es al menos el 95% idéntica a dicha secuencia seleccionada.

También se describe el método anterior en el que dicha etapa de comparar secuencias comprende determinar el grado de hibridación de ácido nucleico entre moléculas de ácido nucleico en dicha muestra y una molécula de ácido nucleico que comprende dicha secuencia seleccionada de dicho grupo. De manera similar, también se describe el método anterior en el que dicha etapa de comparar secuencias se realiza comparando la secuencia de nucleótidos determinada a partir de una molécula de ácido nucleico en dicha muestra con dicha secuencia seleccionada de dicho grupo. Las moléculas de ácido nucleico pueden comprender moléculas de ADN o moléculas de ARN.

Se describe además un método para identificar la especie, el tejido o el tipo de célula de una muestra biológica, método que comprende una etapa de detectar moléculas de ácido nucleico en dicha muestra, si hay alguna, que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de nucleótidos de SEQ ID NO:23 tal como se define en la tabla 1; y una secuencia de nucleótidos codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

El método para identificar la especie, el tejido o el tipo de célula de una muestra biológica puede comprender una etapa de detectar moléculas de ácido nucleico que comprenden una secuencia de nucleótidos en un panel de al menos dos secuencias de nucleótidos, siendo al menos una secuencia en dicho panel al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en una secuencia seleccionada de dicho grupo.

También se prefiere un método para diagnosticar en un sujeto un estado patológico asociado con una expresión o estructura anómala de un gen que codifica una proteína secretada identificada en la tabla 1, método que comprende una etapa de detectar en una muestra biológica obtenida de dicho sujeto moléculas de ácido nucleico, si hay alguna, que comprenden una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de nucleótidos de SEQ ID NO:23; y una secuencia de nucleótidos codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

El método para diagnosticar un estado patológico puede comprender una etapa de detectar moléculas de ácido nucleico que comprenden una secuencia de nucleótidos en un panel de al menos dos secuencias de nucleótidos, siendo al menos una secuencia en dicho panel al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en una secuencia seleccionada de dicho grupo.

También se describe una composición de materia que comprende moléculas de ácido nucleico aisladas, comprendiendo las secuencias de nucleótidos de dichas moléculas de ácido nucleico un panel de al menos dos secuencias de nucleótidos, siendo al menos una secuencia en dicho panel al menos el 95% idéntica a una secuencia de al menos 50 nucleótidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de nucleótidos de SEQ ID NO:23 tal como se define en la tabla 1; y una secuencia de nucleótidos codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075. Las moléculas de ácido nucleico pueden comprender moléculas de ADN o moléculas de ARN.

También se describe un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 90% idéntica a una secuencia de al menos aproximadamente 10 aminoácidos contiguos en la secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1.

También se describe un polipéptido, en el que dicha secuencia de aminoácidos contiguos está incluida en la secuencia de aminoácidos de SEQ ID NO:60 en el intervalo de posiciones que comienza con el resto en aproximadamente la posición del primer aminoácido de la parte secretada y que acaba con el resto en aproximadamente el último aminoácido del marco de lectura abierto tal como se expone para SEQ ID NO:60 en la tabla 1.

También se describe un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 95% idéntica a una secuencia de al menos aproximadamente 30 aminoácidos contiguos en la secuencia de aminoácidos de SEQ ID NO:60.

Se describe además un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 95% idéntica a una secuencia de al menos aproximadamente 100 aminoácidos contiguos en la secuencia de aminoácidos de SEQ ID NO:60.

Se prefiere adicionalmente un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 95% idéntica a la secuencia de aminoácidos completa de SEQ ID NO:60.

Se describe además un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 90% idéntica a una secuencia de al menos aproximadamente 10 aminoácidos contiguos en la secuencia de aminoácidos completa de una proteína secretada codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

También se describe un polipéptido en el que dicha secuencia de aminoácidos contiguos está incluida en la secuencia de aminoácidos de una parte secretada de la proteína secretada codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

También se describe un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 95% idéntica a una secuencia de al menos aproximadamente 30 aminoácidos contiguos en la secuencia de aminoácidos de la parte secretada de la proteína codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

También se describe un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 95% idéntica a una secuencia de al menos aproximadamente 100 aminoácidos contiguos en la secuencia de aminoácidos de la parte secretada de la proteína codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

Se prefiere un polipéptido aislado que comprende una secuencia de aminoácidos al menos el 95% idéntica a la secuencia de aminoácidos de la parte secretada de la proteína codificada por el clon de ADNc humano contenido en el depósito con el número de depósito de la ATCC 209075.

Se describe además un anticuerpo aislado que se une específicamente a un polipéptido que comprende una secuencia de aminoácidos que es al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1; y una secuencia de aminoácidos completa de una proteína codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

Se describe además un método para detectar en una muestra biológica un polipéptido que comprende una secuencia de aminoácidos que es al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1; y una secuencia de aminoácidos completa de una proteína codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075; método que comprende una etapa de comparar una secuencia de aminoácidos de al menos una molécula de polipéptido en dicha muestra con una secuencia seleccionada de dicho grupo y determinar si la secuencia de dicha molécula de polipéptido en dicha muestra es al menos el 90% idéntica a dicha secuencia de al menos 10 aminoácidos contiguos.

También se describe el método anterior en el que dicha etapa de comparar una secuencia de aminoácidos de al menos una molécula de polipéptido en dicha muestra con una secuencia seleccionada de dicho grupo comprende determinar el grado de unión específica de polipéptidos en dicha muestra a un anticuerpo que se une específicamente a un polipéptido que comprende una secuencia de aminoácidos que es al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1; y una secuencia de aminoácidos completa de una proteína codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

También se describe el método anterior en el que dicha etapa de comparar secuencias se realiza comparando la secuencia de aminoácidos determinada a partir de una molécula de polipéptido en dicha muestra con dicha secuencia seleccionada de dicho grupo.

También se describe un método para identificar la especie, el tejido o el tipo de célula de una muestra biológica, método que comprende una etapa de detectar moléculas de polipéptido en dicha muestra, si hay alguna, que comprenden una secuencia de aminoácidos que es al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1; y una secuencia de aminoácidos completa de una proteína secretada codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

También se describe el método anterior para identificar la especie, el tejido o el tipo de célula de una muestra biológica, método que comprende una etapa de detectar moléculas de polipéptido que comprenden una secuencia de aminoácidos en un panel de al menos dos secuencias de aminoácidos, siendo al menos una secuencia en dicho panel al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo anterior.

También se prefiere un método para diagnosticar en un sujeto un estado patológico asociado con una expresión o estructura anómala de un gen que codifica una proteína secretada identificada en la tabla 1, método que comprende una etapa de detectar en una muestra biológica obtenida de dicho sujeto moléculas de polipéptido que comprenden una secuencia de aminoácidos en un panel de al menos dos secuencias de aminoácidos, siendo al menos una secuencia en dicho panel al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1; y una secuencia de aminoácidos completa de una proteína secretada codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

En cualquiera de estos métodos, la etapa de detectar dichas moléculas de polipéptido incluye usar un anticuerpo.

También se describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que es al menos el 95% idéntica a una secuencia de nucleótidos que codifica un polipéptido comprendiendo dicho polipéptido una secuencia de aminoácidos que es al menos el 90% idéntica a una secuencia de al menos 10 aminoácidos contiguos en una secuencia seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60; y una secuencia de aminoácidos completa de una proteína secretada codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

También se describe una molécula de ácido nucleico aislada, en la que dicha secuencia de nucleótidos que codifica un polipéptido se ha optimizado para la expresión de dicho polipéptido en un huésped procariota.

También se describe una molécula de ácido nucleico aislada, en la que dicho polipéptido comprende una secuencia de aminoácidos seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 tal como se define en la tabla 1; y una secuencia de aminoácidos completa de una proteína secretada codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075.

Se describe además un método de preparación de un vector recombinante que comprende insertar cualquiera de las moléculas de ácido nucleico aisladas anteriores en un vector. También se prefiere el vector recombinante producido mediante este método. También se prefiere un método de preparación de una célula huésped recombinante que comprende introducir el vector en una célula huésped, así como la célula huésped recombinante producida mediante este método.

También se prefiere un método de preparación de un polipéptido aislado que comprende cultivar esta célula huésped recombinante en condiciones tales que se exprese dicho polipéptido y recuperar dicho polipéptido. También se prefiere este método de preparación de un polipéptido aislado, en el que dicha célula huésped recombinante es una célula eucariota y dicho polipéptido es una parte secretada de una proteína secretada humana que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en: una secuencia de aminoácidos de SEQ ID NO:60 que comienza con el resto en la posición del primer aminoácido de la parte secretada de SEQ ID NO:60 expuesta en la tabla 1 y

dicha posición del primer aminoácido de la parte secretada de SEQ ID NO:60 se define en la tabla 1; y una secuencia de aminoácidos de una parte secretada de una proteína codificada por un clon de ADNc humano identificado por un identificador de clon de ADNc en la tabla 1 y contenido en el depósito con el número de depósito de la ATCC 209075. También se prefiere el polipéptido aislado producido mediante este método.

También se describe un método de tratamiento de un individuo que necesita un nivel aumentado de una actividad de proteína secretada, método que comprende administrar a un individuo de este tipo una composición farmacéutica que comprende una cantidad de un polipéptido, polinucleótido o anticuerpo aislado de la invención reivindicada eficaz para aumentar el nivel de dicha actividad de proteína en dicho individuo.

Habiendo descrito en general la invención, la misma se entenderá más fácilmente mediante referencia a los ejemplos siguientes, que se proporcionan a modo de ilustración y no pretenden ser limitativos.

**Ejemplos**

Ejemplo 1

*Aislamiento de un clon de ADNc seleccionado a partir de la muestra depositada*

Cada clon de ADNc en un depósito de ATCC citado está contenido en un vector de plásmido. La tabla 1 identifica los vectores usados para construir la biblioteca de ADNc a partir de la cual se aisló cada clon. En muchos casos, el vector usado para construir la biblioteca es un vector de fago del que se ha escindido un plásmido. La tabla inmediatamente a continuación correlaciona inmediatamente el plásmido relacionado para cada vector de fago usado en la construcción de la biblioteca de ADNc. Por ejemplo, cuando se identifica un clon particular en la tabla 1 como aislado en el vector "Lambda Zap", el clon depositado correspondiente está en "pBluescript".

| Vector usado para construir la biblioteca | Plásmido depositado correspondiente |
|---|---|
| Lambda Zap | pBluescript (pBS) |
| Uni-Zap XR | pBluescript (pBS) |
| Zap Express | pBK |
| lafmid BA | plafmid BA |
| pSport1 | pSport1 |
| pCMVSport 2.0 | pCMVSport 2.0 |
| pCMVSport 3.0 | pCMVSport 3.0 |
| pCR®2.1 | pCR®2.1 |

Vectores Lambda Zap (patentes estadounidenses n.os 5.128.256 y 5.286.636), Uni-Zap XR (patentes estadounidenses n.os 5.128,256 y 5.286.636), Zap Express (patentes estadounidenses n.os 5.128.256 y 5.286.636), pBluescript (pBS) (Short, J. M. *et al*., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. y Short, J. M., Nucleic Acids Res. 17:9494 (1989)) y pBK (Alting-Mees, M. A. *et al*., Strategies 5:58-61 (1992)) están disponibles comercialmente de Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contiene un gen de resistencia a ampicilina y pBK contiene un gen de resistencia a neomicina. Ambos pueden transformarse en la cepa XL-1 Blue de *E. coli*, también disponible de Stratagene. pBS viene en 4 formas SK+, SK-, KS+ y KS. S y K se refieren a la orientación del poliligador con respecto a las secuencias de cebador T7 y T3 que flanquean la región de poliligador ("S" representa SacI y "K" representa KpnI que son los primeros sitios en cada extremo respectivo del ligador). "+" o "-" se refieren a la orientación del origen f1 de replicación ("ori"), de manera que en una orientación, el rescate monocatenario iniciado a partir del ori f1 genera ADN de cadena sentido, y en la otra, antisentido.

Se obtuvieron los vectores pSport1, pCMVSport 2.0 y pCMVSport 3.0, de Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. Todos los vectores Sport contienen un gen de resistencia a ampicilina y pueden transformarse en la cepa DH10B de *E. coli*, también disponible de Life Technologies. (Véanse, por ejemplo, Gruber, C. E., *et al*., Focus 15:59 (1993)). El vector lafmid BA (Bento Soares, Universidad de Columbia, NY) contiene un gen de resistencia a ampicilina y puede transformarse en la cepa XL-1 Blue de *E. coli*. El vector pCR®2.1, que está disponible de Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contiene un gen de resistencia a ampicilina y puede transformarse en la cepa DH10B de *E. coli*, disponible de Life Technologies. (Véanse, por ejemplo, Clark, J. M., Nuc.

Acids Res. 16:9677-9686 (1988) y Mead, D. *et al*., Bio/Technology 9: (1991)). Preferiblemente, un polinucleótido de la presente invención no comprende las secuencias de vector de fago identificadas para el clon particular en la tabla 1, así como las secuencias de vector de plásmido correspondientes designadas anteriormente.

El material depositado en la muestra al que se le asignó el número de depósito de ATCC citado en la tabla 1 para cualquier clon de ADNc dado también puede contener uno o más plásmidos adicionales, comprendiendo cada uno un clon de ADNc diferente de ese clon dado. Por tanto, los depósitos que comparten el mismo número de depósito de ATCC contienen al menos un plásmido para cada clon de ADNc identificado en la tabla 1. Normalmente, cada muestra de depósito de ATCC citada en la tabla 1 comprende una mezcla de cuantidades aproximadamente iguales (en peso) de aproximadamente 50 ADN de plásmido, conteniendo cada una un clon de ADNc diferente; pero una muestra de depósito de este tipo puede incluir plásmidos para más o menos de 50 clones de ADNc, hasta aproximadamente 500 clones de ADNc.

Pueden usarse dos enfoques para aislar un clon particular a partir de la muestra depositada de los ADN de plásmido citada para ese clon en la tabla 1. En primer lugar, se aísla directamente un plásmido seleccionando los clones usando una sonda polinucleotídica correspondiente a la SEQ ID NO:X.

Particularmente, se sintetiza un polinucleótido específico con 30-40 nucleótidos usando un sintetizador de ADN de Applied Biosystems según la secuencia notificada. Se marca el oligonucleótido, por ejemplo, con $^{32}$P-$\gamma$-ATP usando polinucleótido cinasa de T4 y se purifica según métodos habituales. (Por ejemplo, Maniatis *et al*., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982)). Se transforma la mezcla de plásmido en un huésped adecuado, tal como se indicó anteriormente (tal como XL-1 Blue (Stratagene)) usando técnicas conocidas por los expertos en la técnica, tales como las proporcionadas por el proveedor del vector o en publicaciones relacionadas o patentes citadas anteriormente. Se siembran en placa los transformantes en placas con agar al 1,5% (que contienen el agente de selección apropiado, por ejemplo, ampicilina) a una densidad de aproximadamente 150 transformantes (colonias) por placa. Se seleccionan estas placas usando membranas de nailon según métodos habituales para la selección de colonias bacterianas (por ejemplo, Sambrook *et al*., Molecular Cloning: A Laboratory Manual, 2ª Edición, (1989), Cold Spring Harbor Laboratory Press, páginas 1.93 hasta 1.104), u otras técnicas conocidas por los expertos en la técnica.

Alternativamente, se sintetizan dos cebadores de 17-20 nucleótidos derivados de ambos extremos de la SEQ ID NO:X (es decir, dentro de la región de la SEQ ID NO:X unida al NT en 5' y al NT en 3' del clon definido en la tabla 1) y se usan para amplificar el ADNc deseado usando el plásmido de ADNc depositado como molde. Se lleva a cabo la reacción en cadena de la polimerasa en condiciones habituales, por ejemplo, en 25 $\mu$l de mezcla de reacción con 0,5 ug del molde de ADNc anterior. Una mezcla de reacción conveniente es $MgCl_2$ 1,5-5 mM, gelatina al 0,01% (p/v), 20 $\mu$M de cada uno de dATP, dCTP, dGTP, dTTP, 25 pmoles de cada cebador y 0,25 unidades de Taq polimerasa. Se realizan treinta y cinco ciclos de PCR (desnaturalización a 94ºC durante 1 min.; apareamiento a 55ºC durante 1 min.; alargamiento a 72ºC durante 1 min.) con un ciclador térmico automatizado Cetus de Perkin-Elmer. Se analiza el producto amplificado mediante electroforesis en gel de agarosa y se corta la banda de ADN con el peso molecular esperado y se purifica. Se verifica que el producto de PCR es la secuencia seleccionada subclonando y secuenciando el producto de ADN.

Se encuentran disponibles varios métodos para la identificación de las partes no codificantes en 5' o 3' de un gen que pueden no estar presentes en el clon depositado. Estos métodos incluyen pero no se limitan a, estudiar con sondas filtro, enriquecimiento de clones usando sondas específicas, y protocolos similares o idénticos a los protocolos de "RACE" en 5' y 3' que se conocen bien en la técnica. Por ejemplo, un método similar a RACE en 5' está disponible para generar el extremo 5' que falta de un transcrito de longitud completa deseado. (Fromont-Racine *et al*., Nucleic Acids Res. 21(7):1683-1684 (1993)).

En resumen, se liga un oligonucleótido de ARN específico a los extremos 5' de una población de ARN que contiene supuestamente transcritos de ARN de gen de longitud completa. Se usa un conjunto de cebadores que contiene un cebador específico para el oligonucleótido de ARN ligado y un cebador específico para una secuencia conocida del gen de interés para amplificar por PCR la parte en 5' del gen de longitud completa deseado. Entonces este producto amplificado puede secuenciarse y usarse para generar el gen de longitud completa.

Este método anterior empieza con ARN total aislado de la fuente deseada, aunque puede usarse ARN poli-A+. Entonces puede tratarse la preparación de ARN con fosfatasa si es necesario eliminar grupos fosfato en 5' en ARN degradado o dañado que pueden interferir con la etapa de ARN ligasa tardía. Entonces la fosfatasa debe inactivarse y tratarse el ARN con pirofosfatasa ácida del tabaco con el fin de eliminar la estructura de caperuza presente en los extremos 5' de los ARN mensajeros. Esta reacción deja un grupo fosfato en 5' en el extremo 5' del ARN escindido de la caperuza que entonces puede ligarse a un oligonucleótido de ARN usando ARN ligasa de T4.

Se usa esta preparación de ARN modificado como molde para las síntesis de ADNc de la primera cadena usando un oligonucleótido específico de gen. Se usa la reacción de síntesis de la primera cadena como molde para la amplificación por PCR del extremo 5' deseado usando un cebador específico para el oligonucleótido de ARN ligado y un cebador específico para la secuencia conocida del gen de interés. Entonces se secuencia el producto resultante y se analiza para confirmar que la secuencia del extremo 5' pertenece al gen deseado.

Ejemplo 2

*Aislamiento de clones genómicos correspondientes a un polinucleótido*

Se selecciona una biblioteca P1 genómica humana (Genomic Systems, Inc.) mediante PCR usando cebadores seleccionados para la secuencia de ADNc correspondiente a la SEQ ID NO:X., según el método descrito en el ejemplo 1. (Véase también, Sambrook).

Ejemplo 3

*Distribución tisular de polipéptido*

Se determina la distribución tisular de la expresión de ARNm de polinucleótidos de la presente invención usando protocolos para análisis de transferencia de tipo Northern, descrito por, entre otros, Sambrook *et al*. Por ejemplo, se marca una sonda de ADNc producida mediante el método descrito en el ejemplo 1 con $P^{32}$ usando el sistema de marcaje de ADN rediprime™ (Amersham Life Science), según las instrucciones del fabricante. Tras el marcaje, se purifica la sonda usando una columna CHROMA SPIN-100™ (Clontech Laboratories, Inc.), según el número de protocolo del fabricante PT1200-1. Entonces se usa la sonda marcada purificada para examinar varios tejidos humanos para determinar la expresión de ARNm.

Se examinan las transferencias de tipo Northern de múltiples tejidos (MTN) que contienen varios tejidos humanos (H) o tejidos del sistema inmunitario humano (IM) (Clontech) con la sonda marcada usando disolución de hibridación ExpressHyb™ (Clontech) según el número de protocolo del fabricante PT1190-1. Tras la hibridación y el lavado, se montan las transferencias y exponen películas a -70ºC durante la noche, y se revelan las películas según procedimientos convencionales.

Ejemplo 4

*Mapeo cromosómico de los polinucleótidos*

Se diseña un conjunto de cebadores de oligonucleótido según la secuencia en el extremo 5' de la SEQ ID NO:X. Este cebador preferiblemente abarca aproximadamente 100 nucleótidos. Entonces se usa este conjunto de cebadores en una reacción en cadena de la polimerasa en el siguiente conjunto de condiciones: 30 segundos, 95ºC; 1 minuto, 56ºC; 1 minuto, 70ºC. Se repite este ciclo 32 veces seguido por un ciclo de 5 minutos a 70ºC. Se usa ADN humano, de ratón, y de hámster como molde además de un panel de híbridos de células somáticas que contiene cromosomas individuales o fragmentos de cromosoma (Bios, Inc). Se analizan las reacciones o bien en geles de poliacrilamida al 8% o bien en de geles de agarosa al 3,5%. Se determina el mapeo cromosómico mediante la presencia de un fragmento de PCR de aproximadamente 100 pb en el híbrido de células somáticas particular.

Ejemplo 5

*Expresión bacteriana de un polipéptido*

Se amplifica un polinucleótido que codifica un polipéptido de la presente invención usando cebadores de oligonucleótido de PCR correspondientes a los extremos 5' y 3' de la secuencia de ADN, tal como se resumió en el ejemplo 1, para sintetizar fragmentos de inserción. Los cebadores usados para amplificar el inserto de ADNc deben contener preferiblemente sitios de restricción, tales como BamHI y XbaI, en el extremo 5' de los cebadores con el fin de clonar el producto amplificado en el vector de expresión. Por ejemplo, BamHI y XbaI corresponden a los sitios de enzima de restricción en el vector de expresión bacteriano pQE-9. (Qiagen, Inc., Chatsworth, CA). Este vector de plásmido codifica la resistencia a antibiótico ($Amp^r$), un origen bacteriano de replicación (ori), un promotor/operador (P/O) regulable por IPTG, un sitio de unión de ribosoma (RBS), una etiqueta de 6-histidina (6-His), y sitios de clonación de enzima de restricción.

Se digiere el vector pQE-9 con BamHI y XbaI y se liga el fragmento amplificado en el vector pQE-9 manteniendo el marco de lectura iniciado en el RBS bacteriano. Entonces se usa la mezcla de ligación para transformar la cepa M15/rep4 de *E. coli* (Qiagen, Inc.) que contiene múltiples copias del plásmido pREP4, que expresa el represor lacI y también confiere resistencia a kanamicina ($Kan^r$). Se identifican los transformantes por su capacidad para crecer en placas con LB y se seleccionan colonias resistentes a ampicilina/kanamicina. Se aísla ADN de plásmido y se confirma mediante análisis de restricción.

Se hacen crecer clones que contienen los constructos deseados durante la noche (O/N) en cultivo líquido en medios LB complementados tanto con Amp (100 ug/ml) como con Kan (25 ug/ml). Se usa el cultivo O/N para inocular un gran cultivo a una razón de 1:100 a 1:250. Se hacen crecer las células hasta una densidad óptica de 600 ($D.O._{600}$) de entre 0,4 y 0,6. Entonces se añade IPTG (isopropil-B-D-tiogalacto-piranósido) hasta una concentración final de 1 mM. IPTG induce, inactivando el represor lacI, el aclaramiento del P/O llevando a una expresión génica aumentada.

Se hacen crecer las células durante de 3 a 4 horas adicionales. Entonces se recogen las células mediante centrifugación (20 min. a 6000Xg). Se solubiliza el sedimento celular en el agente caotrópico Guanidina HCl 6 Molar agitando durante 3-4 horas a 4ºC. Se eliminan el resto celular mediante centrifugación, y se carga el sobrenadante que contiene el polipéptido en una columna de resina de afinidad de níquel-nitrilo-ácido tri-acético ("Ni-NTA") (disponible de QIAGEN, Inc., citado anteriormente). Las proteínas con una etiqueta de 6 x His se unen a la resina de Ni-NTA con alta afinidad y pueden purificarse en un procedimiento de una sola etapa sencilla (para más detalles véase: The QIAexpressionist (1995) QIAGEN, Inc., citado anteriormente).

En resumen, se carga el sobrenadante en la columna en guanidina-HCl 6 M, pH 8, en primer lugar se lava la columna con 10 volúmenes de guanidina-HCl 6 M, pH 8, entonces se lava con 10 volúmenes de guanidina-HCl 6 M, pH 6, y finalmente se eluye el polipéptido con guanidina-HCl 6 M, pH 5.

Entonces se renaturaliza la proteína purificada dializándola frente a solución salina tamponada con fosfato (PBS) o acetato de Na 50 mM, tampón de pH 6 más NaCl 200 mM. Alternativamente, la proteína puede volver a plegarse satisfactoriamente mientras está inmovilizada en la columna de Ni-NTA. Las condiciones recomendadas son las siguientes: renaturalización usando un gradiente de urea lineal 6 M - 1 M en NaCl 500 mM, glicerol al 20%, Tris/HCl 20 mM pH 7,4, que contenía inhibidores de proteasas. La renaturalización debe realizarse a lo largo de un periodo de 1,5 horas o más. Tras la renaturalización, se eluyen las proteínas mediante la adición de imidazol 250 mM. Se elimina el imidazol mediante una etapa de diálisis final frente a PBS o tampón acetato de sodio 50 mM pH 6 más NaCl 200 mM. Se almacena la proteína purificada a 4ºC o se congela a -80ºC.

Además del vector de expresión anterior, la presente invención incluye además un vector de expresión que comprende elementos promotores y operadores de fagos operativamente unidos a un polinucleótido de la presente invención, denominado pHE4a. (Número de registro de la ATCC 209645, depositado el 25 de febrero de 1998). Este vector contiene: 1) un gen de neomicinfosfotransferasa como marcador de selección, 2) un origen de replicación de *E. coli*, 3) una secuencia promotora de fago T5, 4) dos secuencias de operador lac, 5) una secuencia Shine-Dalgarno, y 6) el gen represor del operón lactosa (lacIq). El origen de replicación (oriC) se deriva de pUC19 (LTI, Gaithersburg, MD). La secuencia promotora y las secuencias operadoras se preparan de manera sintética.

Puede insertarse ADN en el pHEa mediante restricción del vector con NdeI y XbaI, BamHI, XhoI o Asp718, haciendo pasar el producto sometido a restricción sobre un gel, y aislando el fragmento mayor (el fragmento de relleno debe tener aproximadamente 310 pares de bases). El inserto de ADN se genera según el protocolo de PCR descrito en el ejemplo 1, usando cebadores de PCR que tienen sitios de restricción para NdeI (cebador en 5') y XbaI, BamHI, XhoI o Asp718 (cebador en 3'). Se purifica en gel el inserto de PCR y se somete a restricción con enzimas compatibles. Se ligan el inserto y el vector según protocolos convencionales.

El vector diseñado mediante ingeniería genética podría sustituirse fácilmente en el protocolo anterior para expresar proteína en un sistema bacteriano.

Ejemplo 6

*Purificación de un polipéptido a partir de un cuerpo de inclusión*

El siguiente método alternativo puede usarse para purificar un polipéptido expresado en *E. coli* cuando está presente en forma de cuerpos de inclusión. A menos que se especifique lo contrario, todas las etapas siguientes se realizan a 4-10ºC.

Tras finalizar la fase de producción de la fermentación con *E. coli*, se enfría el cultivo celular hasta 4-10ºC y se recogen las células mediante centrifugación continua a 15.000 rpm (Heraeus Sepatech). Basándose en el rendimiento esperado de proteína por unidad de peso de la pasta celular y la cantidad de proteína purificada requerida, se suspende una cantidad apropiada de pasta celular, en peso, en una disolución tampón que contiene Tris 100 mM, EDTA 50 mM, pH 7,4. Se dispersan las células para dar una suspensión homogénea una mezcladora de alta cizalladura.

Entonces se lisan las células haciendo pasar la disolución a través de un microfluidizador (Microfuidics, Corp. o APV Gaulin, Inc.) dos veces a 4000-6000 psi. Entonces se mezcla el homogeneizado con disolución de NaCl hasta una concentración final de NaCl 0,5 M, seguido por centrifugación a 7000 xg durante 15 min. Se lava de nuevo el sedimento resultante usando NaCl 0,5 M, Tris 100 mM, EDTA 50 mM, pH 7,4.

Se solubilizan los cuerpos de inclusión lavados resultantes con clorhidrato de guanidina (GuHCl) 1,5 M durante 2-4 horas. Tras la centrifugación a 7000 xg durante 15 min., se desecha el sedimento y se incuba el sobrenadante que contiene el polipéptido a 4ºC durante la noche para permitir una extracción de GuHCl adicional.

Tras la centrifugación a alta velocidad (30.000 xg) para eliminar las partículas insolubles, se vuelve a plegar la proteína solubilizada en GuHCl mezclando rápidamente el extracto de GuHCl con 20 volúmenes de tampón que contiene sodio 50 mM, pH 4,5, NaCl 150 mM, EDTA 2 mM con agitación vigorosa. Se mantiene la disolución de proteína diluida replegada a 4ºC sin mezclar durante 12 horas antes de etapas de purificación adicionales.

Para clarificar la disolución de polipéptido replegado, se emplea una unidad de filtración tangencial preparada previamente con un filtro de membrana de 0,16 $\mu$m con un área superficial apropiada (por ejemplo, Filtron), equilibrada con acetato de sodio 40 mM, pH 6,0. Se carga la muestra filtrada en una resina de intercambio catiónico (por ejemplo, Poros HS-50, Perseptive Biosystems). Se lava la columna con acetato de sodio 40 mM, pH 6,0 y se eluye con NaCl 250 mM, 500 mM, 1000 mM y 1500 mM en el mismo tampón, por etapas. Se monitoriza continuamente la absorbancia a 280 nm del efluente. Se recogen fracciones y se analizan adicionalmente mediante SDS-PAGE.

Entonces se reúnen las fracciones que contienen el polipéptido y se mezclan con 4 volúmenes de agua. Entonces se carga la muestra diluida en un conjunto preparado previamente de columnas en serie de resinas de intercambio aniónico fuerte (Poros HQ-50, Perseptive Biosystems) y aniónico débil (Poros CM-20, Perseptive Biosystems). Se equilibran las columnas con acetato de sodio 40 mM, pH 6,0. Se lavan ambas columnas con acetato de sodio 40 mM, pH 6,0, NaCl 200 mM. Entonces se eluye la columna CM-20 usando un gradiente lineal de 10 volúmenes de columna que oscila desde NaCl 0,2 M, acetato de sodio 50 mM, pH 6,0 hasta NaCl 1,0 M, acetato de sodio 50 mM, pH 6,5. Se recogen fracciones bajo monitorización de la $A_{280}$ constante del efluente. Entonces se reúnen las fracciones que contienen el polipéptido (determinado, por ejemplo, mediante SDS-PAGE al 16%).

El polipéptido resultante debe presentar una pureza superior al 95% tras las etapas de replegamiento y purificación anteriores. No deben observarse bandas contaminantes importantes procedentes del gel de SDS-PAGE al 16% teñido con azul de coomassie cuando se cargan 5 $\mu$g de proteína purificada. La proteína purificada puede someterse a prueba también para determinar la contaminación por endotoxinas/LPS, y normalmente el contenido en LPS es inferior a 0,1 ng/ml según ensayos LAL.

Ejemplo 7

*Clonación y expresión de un polipéptido en un sistema de expresión de baculovirus*

En este ejemplo, se usa el vector pA2 lanzadera de plásmido para insertar un polinucleótido en un baculovirus para expresar un polipéptido. Este vector de expresión contiene el promotor fuerte de polihedrina del virus de la polihedrosis nuclear de *Autographa californica* (AcMNPV) seguido por sitios de restricción convenientes tales como BamHI, Xba I y Asp718. El sitio de poliadenilación del virus de simio 40 ("SV40") se usa para lograr una poliadenilación eficaz. Para la fácil selección de virus recombinante, el plásmido contiene el gen de beta-galactosidasa de *E. coli* bajo el control de un promotor débil de *Drosophila* en la misma orientación, seguido por la señal de poliadenilación del gen de polihedrina. Los genes insertados están flanqueados en ambos lados por secuencias virales para la recombinación homóloga mediada por células con ADN viral de tipo natural para generar un virus viable que exprese el polinucleótido clonado.

Pueden usarse muchos otros vectores de baculovirus en lugar del vector anterior, tales como pAc373, pVL941 y pAcIM1, tal como apreciaría fácilmente un experto en la técnica, siempre que el constructo proporcione señales ubicadas correctamente para la transcripción, traducción, secreción y similares, incluyendo un péptido señal y un AUG en marco según se requiera. Tales vectores se describen, por ejemplo, en Luckow *et al*., Virology 170:31-39 (1989).

Específicamente, la secuencia de ADNc contenida en el clon depositado, que incluye el codón de iniciación AUG y la secuencia líder asociada de manera natura identificada en la tabla 1, se amplifica usando el protocolo de PCR descrito en el ejemplo 1. Si se usa la secuencia señal que se produce de manera natural para producir la proteína secretada, el vector pA2 no necesita un segundo péptido señal. Alternativamente, el vector puede modificarse (pA2 GP) para incluir una secuencia líder de baculovirus, usando los métodos convencionales descritos en Summers *et al*., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures", Texas Agricultural Experimental Station Bulletin N.º 1555 (1987).

Se aísla el fragmento amplificado a partir de un gel de agarosa al 1% usando a un kit comercialmente disponible ("Geneclean", BIO 101 Inc., La Jolla, Ca.). Entonces se digiere el fragmento con enzimas de restricción apropiadas y se purifica de nuevo en un gel de agarosa al 1%.

Se digiere el plásmido con las enzimas de restricción correspondientes y, opcionalmente, puede desfosforilarse usando fosfatasa intestinal de ternero, usando procedimientos de rutina conocidos en la técnica. Entonces se aísla el ADN en un gel de agarosa al 1% usando un kit comercialmente disponible ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

Se ligan entre sí el fragmento y el plásmido desfosforilado con ADN ligasa de T4. Se transforman células HB101 de *E. coli* u otros huéspedes de *E. coli* adecuados tales como XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) con la mezcla de ligación y se propagan en placas de cultivo. Se identifican las bacterias que contienen el plásmido digiriendo el ADN de colonias individuales y analizando el producto de digestión mediante electroforesis en gel. La secuencia del fragmento clonado se confirma mediante secuenciación de ADN.

Se transfectan conjuntamente cinco $\mu$g de un plásmido que contiene el polinucleótido con 1,0 $\mu$g de un ADN de baculovirus linealizado comercialmente disponible (ADN de baculovirus "BaculoGold™", Pharmingen, San Diego,

CA), usando el método de lipofección descrito por Felgner *et al*., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). Se mezclan un $\mu$g de ADN de virus BaculoGold™ y 5 $\mu$g del plásmido en un pocillo estéril de una placa de microtitulación que contiene 50 $\mu$l de medio Grace libre de suero (Life Technologies Inc., Gaithersburg, MD). Después de eso, se añadieron 10 $\mu$l de lipofectina más 90 $\mu$l de medio Grace, se mezclaron e incubaron durante 15 minutos a temperatura ambiente. Entonces se añadió la mezcla de transfección gota a gota a células de insecto Sf9 (ATCC CRL 1711) sembradas en una placa de cultivo tisular de 35 mm con 1 ml de medio Grace sin suero. Entonces se incuba la placa durante 5 horas a 27ºC. Se elimina entonces la disolución de transfección de la placa y se añade 1 ml de medio Grace para insectos complementado con suero de ternero fetal al 10%. Entonces se continúa el cultivo a 27ºC durante cuatro días.

Después de cuatro días se recoge el sobrenadante y se realiza un ensayo de placas, tal como describen Summers y Smith, citado anteriormente. Se usa un gel de agarosa con "Blue Gal" (Life Technologies Inc., Gaithersburg) para permitir una identificación y un aislamiento fáciles de clones que expresan gal, que producen placas teñidas de azul. (Puede encontrarse también una descripción detallada de un "ensayo de placas" de este tipo en la guía del usuario para el cultivo de células de insecto y baculovirología distribuida por Life Technologies Inc., Gaithersburg, páginas 9-10). Tras una incubación apropiada, se recogen las placas teñidas de azul con la punta de una micropipeta automática (por ejemplo, Eppendorf). Entonces se resuspende el agar que contiene los virus recombinantes en tubo de microcentrífuga que contiene 200 $\mu$l de medio Grace y se usa la suspensión que contiene el baculovirus recombinante para infectar células Sf9 sembradas en placas de 35 mm. Cuatro días más tarde se recogen los sobrenadantes de estas placas de cultivo y entonces se almacenan a 4ºC.

Para verificar la expresión del polipéptido, se hacen crecer células Sf9 en medio Grace complementado con FBS inactivado por calor al 10%. Se infectan las células con el baculovirus recombinante que contiene el polinucleótido a una multiplicidad de infección ("MOI") de aproximadamente 2. Si se desean proteínas radiomarcadas, 6 horas más tarde se retira el medio y se sustituye por medio SF900 II menos metionina y cisteína (disponible de Life Technologies Inc., Rockville, MD). Tras 42 horas, se añaden 5 $\mu$Ci de $^{35}$S-metionina y 5 $\mu$Ci de $^{35}$S-cisteína (disponibles de Amersham). Se incuban adicionalmente las células durante 16 horas y entonces se recogen mediante centrifugación. Se analizan las proteínas en el sobrenadante así como las proteínas intracelulares mediante SDS-PAGE seguido por autorradiografía (si estaban radiomarcadas).

Puede usarse microsecuenciación de la secuencia de aminoácidos del extremo amino terminal de la proteína purificada para determinar la secuencia amino terminal de la proteína producida.

Ejemplo 8

*Expresión de un polipéptido en células de mamífero*

El polipéptido de la presente invención puede expresarse en una célula de mamífero. Un vector de expresión de mamífero típico contiene un elemento promotor, que media la iniciación de la transcripción de ARNm, una secuencia que codifica proteína y señales requeridas para la terminación de transcripción y poliadenilación del transcrito. Los elementos adicionales incluyen potenciadores, secuencias Kozak y secuencias intercaladas flanqueadas por sitios donadores y aceptores para el corte y empalme del ARN. Se consigue una transcripción altamente eficaz con los promotores temprano y tardío de SV40, las repeticiones terminales largas (LTR) de retrovirus, por ejemplo, VSR, VLTHI, VIHI y el promotor temprano del citomegalovirus (CMV). Sin embargo, también pueden usarse elementos celulares (por ejemplo, el promotor de actina humano).

Los vectores de expresión adecuados para su uso para poner en práctica la presente invención incluyen, por ejemplo, vectores tales como pSVL y pMSG (Pharmacia, Uppsala, Suecia), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0 y pCMVSport 3.0. Células huésped de mamífero que pueden usarse incluyen, células Hela, 293, H9 y Jurkat humanas, células NIH3T3 y C127 de ratón, Cos 1, Cos 7 y CV1, células QC1-3 de codorniz, células L de ratón y células de ovario de hámster chino (CHO).

Alternativamente, el polipéptido puede expresarse en líneas celulares estables que contienen el polinucleótido integrado en un cromosoma. La transfección conjunta con un marcador seleccionable tal como dhfr, gpt, neomicina, higromicina permite la identificación y el aislamiento de las células transfectadas.

El gen transfectado puede amplificarse también para expresar grandes cantidades de la proteína codificada. El marcador DHFR (dihidrofolato reductasa) es útil para desarrollar líneas celulares que portan varios cientos o incluso varios miles de copias del gen de interés. (Véase, por ejemplo, Alt, F. W., *et al*., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. y Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. y Sydenham, M. A., Biotechnology 9:64-68 (1991)). Otro marcador de selección útil es la enzima glutamina sintasa (GS) (Murphy *et al*., Biochem J. 227:277-279 (1991); Bebbington *et al*., Bio/Technology 10:169-175 (1992). Usando estos marcadores, se hacen crecer las células de mamífero en medio selectivo y se seleccionan las células con la mayor resistencia. Estas líneas celulares contienen el/los gen(es) amplificado(s) integrado(s) en un cromosoma. A menudo se usan células de ovario de hámster chino (CHO) y NSO para la producción de proteínas.

Los derivados del plásmido pSV2-dhfr (número de registro de la ATCC 37146), los vectores de expresión pC4 (número de registro de la ATCC 209646) y pC6 (número de registro de la ATCC 209647) contienen el promotor fuerte (LTR) del virus del sarcoma de Rous (Cullen *et al*., Molecular and Cellular Biology, 438-447 (marzo de 1985)) más un fragmento del potenciador de CMV (Boshart *et al*., Cell 41:521-530 (1985)). Múltiples sitios de clonación, por ejemplo, con los sitios de escisión de enzimas de restricción BamHI, XbaI y Asp718, facilitan la clonación del gen de interés. Los vectores también contienen el intrón en 3', la señal de poliadenilación y terminación del gen de preproinsulina de rata y el gen de DHFR de ratón bajo el control del promotor temprano de SV40.

Específicamente, el plásmido pC6, por ejemplo, se digiere con enzimas de restricción apropiadas y entonces se desfosforila usando fosfatos intestinales de ternero mediante procedimientos conocidos en la técnica. Entonces se aísla el vector a partir de un gel de agarosa al 1%.

Se amplifica un polinucleótido de la presente invención según el protocolo resumido en el ejemplo 1. Si se usa la secuencia señal que se produce de manera natural para producir la proteína secretada, el vector no necesita un segundo péptido señal. Alternativamente, si no se usa la secuencia señal que se produce de manera natural, el vector puede modificarse para incluir una secuencia señal heteróloga. (Véase, por ejemplo, el documento WO 96/34891).

Se aísla el fragmento amplificado a partir de un gel de agarosa al 1% usando un kit comercialmente disponible ("Geneclean", BIO 101 Inc., La Jolla, Ca.). Entonces se digiere el fragmento con enzimas de restricción apropiadas y se purifica de nuevo en un gel de agarosa al 1%.

Entonces se digiere el fragmento amplificado con la misma enzima de restricción y se purifica en un gel de agarosa al 1%. Entonces se ligan el fragmento aislado y el vector desfosforilado con ADN ligasa de T4. Se transforman entonces células XL-1 Blue o HB 101 de *E. coli* y se identifican las bacterias que contienen el fragmento insertado en el plásmido pC6 usando, por ejemplo, análisis de enzimas de restricción.

Se usan para la transfección células de ovario de hámster chino que carecen de DHFR activa. Se transfectan conjuntamente cinco $\mu$g del plásmido de expresión pC6 con 0,5 $\mu$g del plásmido pSVneo usando lipofectina (Felgner *et al*., citado anteriormente). El plásmido pSV2-neo contiene un marcador seleccionable dominante, el gen *neo* de Tn5 que codifica una enzima que confiere resistencia a un grupo de antibióticos incluyendo G418. Se siembran las células en medio alfa menos MEM complementado con G418 1 mg/ml. Tras 2 días, se tripsinizan las células y se siembran en placas de clonación de hibridomas (Greiner, Alemania) en medio alfa menos MEM complementado con 10, 25 ó 50 ng/ml de metotrexato más G418 1 mg/ml. Después de aproximadamente 10-14 días, se tripsinizan clones individuales y entonces se siembran en placas Petri de 6 pocillos o frascos de 10 ml usando diferentes concentraciones de metotrexato (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Entonces se transfieren los clones que crecen a las mayores concentraciones de metotrexato a nuevas placas de 6 pocillos que contienen concentraciones incluso mayores de metotrexato (1 $\mu$M, 2 $\mu$M, 5 $\mu$M, 10 mM, 20 mM). Se repite el mismo procedimiento hasta que se obtienen clones que crecen a una concentración de 100 - 200 $\mu$M. Se analiza la expresión del producto génico deseado, por ejemplo, mediante SDS-PAGE e inmunotransferencia de tipo Western o mediante análisis por HPLC en fase inversa.

Ejemplo 9

*Fusiones de proteínas*

Los polipéptidos de la presente invención se fusionan preferiblemente con otras proteínas. Estas proteínas de fusión pueden usarse para una variedad de aplicaciones. Por ejemplo, la fusión de los presentes polipéptidos con etiqueta de His, etiqueta de HA, proteína A, dominios de IgG, y proteína de unión a maltosa facilita la purificación. (Véase el ejemplo 5; véase también el documento EP A 394.827; Traunecker, *et al*., Nature 331:84-86 (1988)). De manera similar, la fusión con IgG-1, IgG-3 y albúmina aumenta el tiempo de semivida *in vivo*. Señales de localización nuclear fusionadas con los polipéptidos de la presente invención pueden dirigir la proteína a una localización subcelular específica, mientras que homodímeros o heterodímeros covalentes pueden aumentar o reducir la actividad de una proteína de fusión. Las proteínas de fusión pueden crear también moléculas quiméricas que tienen más de una función. Finalmente, las proteínas de fusión pueden aumentar la solubilidad y/o la estabilidad de la proteína fusionada en comparación con la proteína no fusionada. Todos los tipos de proteínas de fusión descritos anteriormente pueden prepararse modificando el siguiente protocolo, que resume la fusión de un polipéptido con una molécula de IgG, o el protocolo descrito en el ejemplo 5.

En resumen, la parte Fc humana de la molécula de IgG puede amplificarse mediante PCR, usando cebadores que abarcan los extremos 5' y 3' de la secuencia descrita más adelante. Estos cebadores deben tener también sitios de enzimas de restricción convenientes que facilitarán la clonación en un vector de expresión, preferiblemente un vector de expresión de mamífero.

Por ejemplo, si se usa pC4 (número de registro 209646), la parte Fc humana puede ligarse en el sitio de clonación BamHI. Obsérvese que el sitio BamHI en 3' debe destruirse. A continuación, se vuelve a someter a restricción con BamHI el vector que contiene la parte Fc humana, linealizando el vector, y se liga en este sitio BamHI un polinu-

cleótido de la presente invención, aislado mediante el protocolo de PCR descrito en el ejemplo 1. Obsérvese que el polinucleótido se clona sin un codón de parada, de lo contrario no se producirá una proteína de fusión.

Si se usa la secuencia señal que se produce de manera natural para producir la proteína secretada, pC4 no necesita un segundo péptido señal. Alternativamente, si no se usa la secuencia señal que se produce de manera natural, el vector puede modificarse para incluir una secuencia señal heteróloga. (Véase, por ejemplo, el documento WO 96/34891).

Región Fc de IgG humana:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGCC
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAACC
CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA

GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Ejemplo 10

*Producción de un anticuerpo a partir de un polipéptido*

Los anticuerpos de la presente invención pueden prepararse mediante una variedad de métodos. (Véase, Current Protocols, capítulo 2). Por ejemplo, se administran células que expresan un polipéptido de la presente invención a un animal para inducir la producción de sueros que contienen anticuerpos policlonales. En un método preferido, se prepara una preparación de la proteína secretada y se purifica para hacer que esté sustancialmente libre de contaminantes naturales. Entonces se introduce una preparación de este tipo en un animal con el fin de producir antisueros policlonales de mayor actividad específica.

En el método más preferido, los anticuerpos de la presente invención son anticuerpos monoclonales (o fragmentos de unión a proteína de los mismos). Tales anticuerpos monoclonales pueden prepararse usando tecnología de hibridomas. (Köhler *et al*., Nature 256:495 (1975); Köhler *et al*., Eur. J. Immunol. 6:511 (1976); Köhler *et al*., Eur. J. Immunol. 6:292 (1976); Hammerling *et al*., en: Monoclonal Antibody and T-Cell Hybridomas, Elsevier, N.Y., págs. 563-681 (1981)). En general, tales procedimientos implican inmunizar a un animal (preferiblemente un ratón) con polipéptido o, más preferiblemente, con una célula que expresa polipéptido secretado. Tales células pueden cultivarse en cualquier medio de cultivo tisular adecuado; sin embargo, es preferible cultivar las células en un medio de Eagle modificado por Earle complementado con suero bovino fetal al 10% (inactivado a aproximadamente 56ºC), y complementado con aproximadamente 10 g/l de aminoácidos no esenciales, aproximadamente 1.000 U/ml de penicilina y aproximadamente 100 $\mu$g/ml de estreptomicina.

Se extraen los esplenocitos de tales ratones y se fusionan con una línea celular de mieloma adecuada. Puede emplearse cualquier línea celular de mieloma adecuada según la presente invención; sin embargo, es preferible emplear la línea celular de mieloma original (SP2O), disponible de la ATCC. Tras la fusión, se mantienen en medio HAT de manera selectiva las células de hibridoma resultantes, y entonces se clonan mediante dilución limitante tal como describen Wands *et al*. (Gastroenterology 80:225-232 (1981)). Las células de hibridoma obtenidas a través de una selección de este tipo se someten entonces a ensayo para identificar clones que secretan anticuerpos que pueden unirse al polipéptido.

Alternativamente, pueden producirse en un procedimiento de dos etapas anticuerpos adicionales que pueden unirse al polipéptido usando anticuerpos anti-idiotípicos. Un método de este tipo utiliza el hecho de que los anticuerpos son

por sí mismos antígenos y, por tanto, es posible obtener un anticuerpo que se une a un segundo anticuerpo. Según este método, se usan anticuerpos específicos de proteína para inmunizar a un animal, preferiblemente un ratón. Los esplenocitos de un animal de este tipo se usan entonces para producir células de hibridoma, y se seleccionan las células de hibridoma para identificar clones que producen un anticuerpo cuya capacidad para unirse al anticuerpo específico de proteína puede bloquearse mediante el polipéptido. Tales anticuerpos comprenden anticuerpos anti-idiotípicos frente al anticuerpo específico de proteína y pueden usarse para inmunizar a un animal para inducir la formación de anticuerpos específicos de proteína adicionales.

Se apreciará que Fab y F(ab')2 y otros fragmentos de los anticuerpos de la presente invención pueden usarse según los métodos descritos en el presente documento. Tales fragmentos se producen normalmente mediante escisión proteolítica, usando enzimas tales como papaína (para producir fragmentos Fab) o pepsina (para producir fragmentos F(ab')2). Alternativamente, pueden producirse fragmentos de unión a proteína secretada a través de la aplicación de tecnología de ADN recombinante o a través de química sintética.

Para el uso *in vivo* de anticuerpos en seres humanos, puede ser preferible usar anticuerpos monoclonales quiméricos "humanizados". Tales anticuerpos pueden producirse usando constructos genéticos derivados de células de hibridoma que producen los anticuerpos monoclonales descritos anteriormente. En la técnica se conocen métodos para producir anticuerpos quiméricos. (Véase, para revisión, Morrison, Science 229:1202 (1985); Oi *et al*., BioTechniques 4:214 (1986); Cabilly *et al*., patente estadounidense n.º 4.816.567; Taniguchi *et al*., documento EP 171496; Morrison *et al*., documento EP 173494; Neuberger *et al*., documento WO 8601533; Robinson *et al*., documento WO 8702671; Boulianne *et al*., Nature 312:643 (1984); Neuberger *et al*., Nature 314:268 (1985)).

Ejemplo 11

*Producción de proteína secretada para ensayos de selección de alto rendimiento*

El siguiente protocolo produce un sobrenadante que contiene un polipéptido que va a someterse a prueba. Este sobrenadante puede usarse entonces en los ensayos de selección descritos en los ejemplos 13-20.

En primer lugar, diluir disolución madre de poli-D-lisina (644 587 Boehringer-Mannheim) (1 mg/ml en PBS) 1:20 en PBS (sin calcio o magnesio 17-516F Biowhittaker) para una disolución de trabajo de 50 ug/ml. Añadir 200 ul de esta disolución a cada pocillo (placas de 24 pocillos) e incubar a TA durante 20 minutos. Garantizar que la disolución se distribuye por cada pocillo (nota: puede usarse una pipeta automática de 12 canales con puntas en canales alternos). Eliminar mediante aspiración la disolución de poli-D-lisina y enjuagar con 1 ml de PBS (solución salina tamponada con fosfato). La PBS debe permanecer en el pocillo hasta justo antes de sembrar en placa las células y las placas pueden cubrirse con poli-lisina de antemano durante hasta dos semanas.

Sembrar en placa células 293T (no portar células más allá de P+20) a 2 x $10^5$ células/pocillo en 5 ml de medio DMEM (medio de Eagle modificado por Dulbecco) (con 4,5 g/l de glucosa y L-glutamina (12-604F Biowhittaker))/FBS inactivado por calor al 10% (14-503F Biowhittaker)/1x Penstrep (17-602E Biowhittaker). Dejar crecer las células durante la noche.

Al día siguiente, mezclar entre sí en un recipiente en un recipiente de disolución estéril: 300 ul de Lipofectamine (18324-012 Gibco/BRL) y 5 ml de Optimem I (31985070 Gibco/BRL)/placa de 96 pocillos. Con una pipeta multicanal de volúmenes pequeños, colocar alícuotas de aproximadamente 2 ug de un vector de expresión que contiene un inserto de polinucleótido, producido mediante los métodos descritos en los ejemplos 8 ó 9, en una placa de fondo redondo de 96 pocillos marcada de manera apropiada. Con una pipeta multicanal, añadir 50 ul de la mezcla Lipofectamine/Optimem I a cada pocillo. Pipetear suavemente de arriba abajo para mezclar. Incubar a TA durante 15-45 minutos. Tras aproximadamente 20 minutos, usar una pipeta multicanal para añadir 150 ul de Optimem I a cada pocillo. Como control debe transfectarse una placa de ADN de vector que carece de un inserto con cada conjunto de transfecciones.

Preferiblemente, la transfección debe realizarse trabajando en equipo las siguientes tareas. Trabajando en equipo, el tiempo de trabajo experimental se reduce a la mitad, y las células no pasan demasiado tiempo en PBS. En primer lugar, la persona A elimina por aspiración los medios de cuatro placas de 24 pocillos de células, y luego la persona B enjuaga cada pocillo con 0,5-1 ml de PBS. La persona A elimina entonces por aspiración el enjuague de PBS, y la persona B, usando una pipeta de 12 canales con puntas en canales alternos, añade los 200 ul del complejo de ADN/Lipofectamine/Optimem I en primer lugar a los pocillos impares, luego a los pocillos pares, en cada fila de las placas de 24 pocillos. Incubar a 37ºC durante 6 horas.

Mientras se están incubando las células, preparar medios apropiados, o bien BSA al 1% en DMEM con 1x penstrep, o bien medios CHO-5 (116,6 mg/l de $CaCl_2$ (anhid.); $CuSO_4$-$5H_2O$ 0,00130 mg/l; 0,050 mg/l de $Fe(NO_3)_3$-$9H_2O$; 0,417 mg/l de $FeSO_4$-$7H_2O$; 311,80 mg/l de KCl; 28,64 mg/l de $MgCl_2$; 48,84 mg/l de $MgSO_4$; 6995,50 mg/l de NaCl; 2400,0 mg/l de $NaHCO_3$; 62,50 mg/l de $NaH_2PO_4$-$H_2O$; 71,02 mg/l de $Na_2HPO_4$; 0,4320 mg/l de $ZnSO_4$-$7H_2O$; 0,002 mg/l de ácido araquidónico; 1,022 mg/l de colesterol; 0,070 mg/l de acetato de DL-alfa-tocoferol; 0,0520 mg/l de ácido linoleico; 0,010 mg/l de ácido linolénico; 0,010 mg/l de ácido mirístico; 0,010 mg/l de ácido oleico; 0,010 mg/l de ácido palmítrico; 0,010 mg/l de ácido palmítico; 100 mg/l de Pluronic F-68; 0,010 mg/l de ácido esteárico;

2,20 mg/l de Tween 80; 4551 mg/l de D-glucosa; 130,85 mg/ml de L-alanina; 147,50 mg/ml de L-arginina-HCl; 7,50 mg/ml de L-asparagina-$H_2O$; 6,65 mg/ml de L-ácido aspártico; 29,56 mg/ml de L-cistina-2HCl-$H_2O$; 31,29 mg/ml de L-cistina-2HCl; 7,35 mg/ml de L-ácido glutámico; 365,0 mg/ml de L-glutamina; 18,75 mg/ml de glicina; 52,48 mg/ml de L-histidina-HCl-$H_2O$; 106,97 mg/ml de L-isoleucina; 111,45 mg/ml de L-leucina; 163,75 mg/ml de L-lisina HCl; 32,34 mg/ml de L-metionina; 68,48 mg/ml de L-fenilalainina; 40,0 mg/ml de L-prolina; 26,25 mg/ml de L-serina; 101,05 mg/ml de L-treonina; 19,22 mg/ml de L-triptófano; 91,79 mg/ml de L-trirosina-2Na-2$H_2O$; 99,65 mg/ml de L-valina; 0,0035 mg/l de biotina; 3,24 mg/l de D-pantotenato de Ca; 11,78 mg/l de cloruro de colina; 4,65 mg/l de ácido fólico; 15,60 mg/l de i-inositol; 3,02 mg/l de niacinamida; 3,00 mg/l de piridoxal HCl; 0,031 mg/l de piridoxina HCl; 0,319 mg/l de riboflavina; 3,17 mg/l de tiamina HCl; 0,365 mg/l de timidina; y 0,680 mg/l de vitamina $B_{12}$; 25 mM de tampón HEPES; 2,39 mg/l de hipoxantina de Na; 0,105 mg/l de ácido lipoico; 0,081 mg/l de putrescina de sodio-2HCl; 55,0 mg/l de piruvato de sodio; 0,0067 mg/l de selenito de sodio; 20 uM de etanolamina; 0,122 mg/l de citrato férrico; 41,70 mg/l de metil-B-ciclodextrina complejada con ácido linoleico; 33,33 mg/l de metil-B-ciclodextrina complejada con ácido oleico; y 10 mg/l de metil-B-ciclodextrina complejada con Retinal) con 2 mm de glutamina y 1x penstrep. (100 g de BSA (81-068-3 Bayer) disueltos en 1 l de DMEM para una disolución madre de BSA al 10%). Filtrar los medios y recoger 50 ul para el ensayo de endotoxinas en un tubo cónico de poliestireno de 15 ml.

Se termina la reacción de transfección, preferiblemente trabajando en equipo, al final del periodo de incubación. La persona A elimina por aspiración los medios de transfección, mientras que la persona B añade 1,5 ml de los medios apropiados a cada pocillo. Incubar a 37ºC durante 45 ó 72 horas dependiendo de los medios usados: BSA al 1% durante 45 horas o CHO-5 durante 72 horas.

En el día cuatro, usando una pipeta multicanal de 300 ul, colocar alícuotas de 600 ul en una placa de pocillos de 1 ml de profundidad y el sobrenadante restante en un pocillo de 2 ml de profundidad. Los sobrenadantes de cada pocillo pueden usarse entonces en los ensayos descritos en los ejemplos 13-20.

Se entiende específicamente que cuando se obtiene actividad en cualquiera de los ensayos descritos a continuación usando un sobrenadante, la actividad se origina o bien a partir del polipéptido directamente (por ejemplo, como una proteína secretada) o bien por el polipéptido que induce la expresión de otras proteínas, que entonces se secretan en el sobrenadante. Por tanto, la invención proporciona además un método de identificación de la proteína en el sobrenadante caracterizado por una actividad en un ensayo particular.

Ejemplo 12

*Construcción del constructo indicador de GAS*

Una ruta de transducción de señales implicada en la diferenciación y proliferación de células se denomina la ruta Jak-STAT. Las proteínas activadas en la ruta Jak-STAT se unen a elementos del sitio de activación gamma “GAS” o elemento de respuesta sensible a interferón (“ISRE”), localizados en el promotor de muchos genes. La unión de una proteína a estos elementos altera la expresión del gen asociado.

Los elementos GAS e ISRE se reconocen por una clase de factores de transcripción denominados transductores de señales y activadores de la transcripción, o “STAT”. Existen seis miembros de la familia de los STAT. Stat1 y Stat3 están presentes en muchos tipos de células, al igual que Stat2 (puesto que la respuesta a IFN-alfa está extendida). Stat4 está más restringido y no se encuentra en muchos tipos de células, aunque se ha encontrado en células T auxiliares de clase I tras tratamiento con IL-12. Stat5 de denominó originalmente factor de crecimiento mamario, pero se ha encontrado a mayores concentraciones en otras células incluyendo células mieloides. Puede activarse en células de cultivo tisular por muchas citocinas.

Los STAT se activan para translocarse desde el citoplasma hasta el núcleo tras la fosforilación de tirosinas mediante un conjunto de cinasas conocido como la familia de cinasas Janus (“Jak”). Jak representa una familia diferenciada de tirosina cinasas solubles e incluye Tyk2, Jak1, Jak2 y Jak3. Estas cinasas presentan una similitud de secuencia significativa y generalmente son catalíticamente inactivas en células en reposo.

Las Jak se activan mediante una amplia variedad de receptores que se recogen en la tabla a continuación. (Adaptado de la revisión de Schidler y Darnell, Ann. Rev. Biochem. 64:621-51 (1995)). Una familia de receptores de citocinas, que puede activar Jak, se divide en dos grupos: (a) la clase 1 incluye receptores para IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF y trombopoyetina; y (b) la clase 2 incluye IFN-a, IFN-g e IL-10. Los receptores de clase 1 comparten un motivo de cisteína conservado (un conjunto de cuatro cisteínas conservadas y un triptófano) y un motivo WSXWS (una región proximal de membrana que codifica Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Por tanto, con la unión de un ligando a un receptor, se activan Jak, que a su vez activan STAT, que entonces se translocan y se unen a elementos de GAS. Este proceso completo está englobado en la ruta de transducción de señales Jak-STAT.

Por tanto, la activación de la ruta Jak-STAT, reflejada mediante la unión del elemento GAS o el elemento ISRE, puede usarse para indicar proteínas implicadas en la proliferación y la diferenciación de células. Por ejemplo, se sabe que factores de crecimiento y citocinas activan la ruta Jak-STAT. (Véase la tabla a continuación). Por tanto, usando elementos GAS unidos a moléculas indicadoras, pueden identificarse activadores de la ruta Jak-STAT.

| | | JAK | | | STAT | GAS (elementos) o ISRE |
|---|---|---|---|---|---|---|
| **Ligando** | **tyk2** | **Jak1** | **Jak2** | **Jak3** | | |
| Familia IFN | | | | | | |
| IFN-a/B | + | + | - | - | 1,2,3 | ISRE |
| IFN-g | | + | + | - | 1 | GAS (IRFI>Lys6>IFP) |
| IL-10 | + | ? | ? | - | 1,3 | |
| Familia gp130 | | | | | | |
| IL-6 (pleiotrópica) | + | + | + | ? | 1,3 | GAS (IRF1>Lys6>IFP) |
| IL-11 (pleiotrópica) | ? | + | ? | ? | 1,3 | |
| OnM (pleiotrópica) | ? | + | + | ? | 1,3 | |
| LIF (pleiotrópico) | ? | + | + | ? | 1,3 | |
| CNTF (pleiotrópico) | -/+ | + | + | ? | 1,3 | |
| G-CSF (pleiotrópico) | ? | + | ? | ? | 1,3 | |
| IL-12 (pleiotrópica) | + | - | + | + | 1,3 | |
| Familia g-C | | | | | | |
| IL-2 (linfocitos) | - | + | - | + | 1,3,5 | GAS |
| IL-4 (linf./mieloide) | - | + | - | + | 6 | GAS (IRFI = IFP>>Ly6)(IgH) |
| IL-7 (linfocitos) | - | + | - | + | 5 | GAS |
| IL-9 (linfocitos) | - | + | - | + | 5 | GAS |
| IL-13 (linfocitos) | - | + | ? | ? | 6 | GAS |
| IL-15 | ? | + | ? | + | 5 | GAS |
| Familia gp140 | | | | | | |
| IL-3 (mieloide) | - | - | + | - | 5 | GAS (IRFI>IFP»Ly6) |
| IL-5 (mieloide) | - | - | + | - | 5 | GAS |
| GM-CSF (mieloide) | - | - | + | - | 5 | GAS |
| Familia de hormonas del crecimiento | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GH | ? | - | + | - | 5 | |
| PRL | ? | +/- | + | - | 1,3,5 | |
| EPO | ? | - | + | - | 5 | GAS (B-CAS>IRFI=IFP>>Ly6) |
| Receptores tirosina cinasa | | | | | | |
| EGF | ? | + | + | - | 1,3 | GAS (IRF1) |
| PDGF | ? | + | + | - | 1,3 | |
| CSF-1 | ? | + | + | - | 1,3 | GAS(no IRF1) |

Para construir un elemento promotor que contiene GAS sintético, que se usa en los ensayos biológicos descritos en los ejemplos 13-14, se emplea una estrategia basada en PCR para generar una secuencia promotora de GAS-SV40. El cebador en 5' contiene cuatro copias en tándem del sitio de unión de GAS encontrado en el promotor IRF1 y previamente se demostró que se unía a STAT tras la inducción con una variedad de citocinas (Rothman *et al*., Immunity 1:457-468 (1994)), aunque pueden usarse en su lugar otros elementos GAS o ISRE. El cebador en 5' también contiene 18 pb de secuencia complementaria a la secuencia de promotor temprano de SV40 y está flanqueado con un sitio XhoI. La secuencia del cebador en 5' es:

5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCG
AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

El cebador en el sentido de 3' es complementario al promotor de SV40 y está flanqueado con un sitio Hind III:

5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4).

La amplificación por PCR se realiza usando el molde de promotor de SV40 presente en plásmido B-gal:promotor obtenido de Clontech. Se digiere el fragmento de PCR resultante con XhoI/Hind III y se subclona en BLSK2-. (Stratagene). La secuenciación con cebadores directos e inversos confirma que el inserto contiene la siguiente secuencia:

5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
TGCAAAAAGCTT:3' (SEQ ID NO:5)

Con este elemento promotor GAS ligado al promotor de SV40, se diseña a continuación mediante ingeniería genética un constructo indicador GAS:SEAP2. En este caso, la molécula indicadora es una fosfatasa alcalina secretada, o "SEAP". Claramente, sin embargo, puede insertarse cualquier molécula indicadora en lugar de SEAP, en este o en cualquiera de los demás ejemplos. Las moléculas indicadoras bien conocidas que pueden usarse en lugar de SEAP incluyen cloranfenicol acetiltransferasa (CAT), luciferasa, fosfatasa alcalina, B-galactosidasa, proteína fluorescente verde (GFP), o cualquier proteína detectable por un anticuerpo.

La secuencia anterior confirmaba que el elemento promotor GAS-SV40 sintético se subclona en el vector pSEAP-promotor obtenido de Clontech usando HindIII y XhoI, sustituyendo eficazmente el promotor de SV40 por el elemento promotor GAS:SV40 amplificado, para crear el vector GAS-SEAP. Sin embargo, este vector no contiene un gen de resistencia a neomicina, y por tanto, no se prefiere para sistemas de expresión de mamífero.

Por tanto, con el fin de generar líneas celulares estables de mamífero que expresan el indicador de GAS-SEAP, se elimina el casete de GAS-SEAP del vector GAS-SEAP usando SalI y NotI, y se inserta en un vector de estructura principal que contiene el gen de resistencia a neomicina, tal como pGFP-1 (Clontech), usando estos sitios de restricción en el sitio de clonación múltiple, para crear el vector GAS-SEAP/Neo. Una vez que se transfecta este vector en células de mamífero, puede usarse entonces este vector como molécula indicadora para la unión a GAS tal como se describe en los ejemplos 13-14.

Pueden prepararse otros constructos usando la descripción anterior y sustituyendo GAS por una secuencia promotora diferente. Por ejemplo, la construcción de moléculas indicadoras que contienen secuencias promotoras de NFK-B y EGR se describen en los ejemplos 15 y 16. Sin embargo, pueden sustituirse muchos otros promotores usando los protocolos descritos en estos ejemplos. Por ejemplo, pueden sustituirse promotores de SRE, IL-2, NFAT u osteocalci-

na, solos o en combinación (por ejemplo, GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, o NF-KB/GAS). De manera similar, pueden usarse otras líneas celulares para someter a prueba la actividad de constructos indicadores, tales como HELA (epitelial), HUVEC (endotelial), Reh (célula B), Saos-2 (osteoblasto), HUVAC (aórtica), o cardiomiocito.

Ejemplo 13

*Ensayo de selección de alto rendimiento para determinar la actividad de las células T*

Se usa el siguiente protocolo para evaluar la actividad de las células T mediante la identificación de factores, tales como factores de crecimiento y citocinas, que pueden hacer proliferar o diferenciar células T. La actividad de las células T se evalúa usando el constructo GAS/SEAP/Neo producido en el ejemplo 12. Por tanto, los factores que aumentan la actividad de SEAP indican la capacidad para activar la ruta de transducción de señales Jak-STAT. Las células T usadas en este ensayo son células T Jurkat (número de registro de la ATCC TIB-152), aunque también pueden usarse células Molt-3 (número de registro de la ATCC CRL-1552) y células Molt-4 (número de registro de la ATCC CRL-1582).

Las células T Jurkat son células auxiliares CD4+ Th1 linfoblásticas. Con el fin de generar líneas celulares estables, se transfectan aproximadamente 2 millones de células Jurkat con el vector GAS-SEAP/neo usando DMRIE-C (Life Technologies) (procedimiento de transfección descrito más adelante). Se siembran las células transfectadas hasta una densidad de aproximadamente 20.000 células por pocillo y se seleccionan los transfectantes resistentes a genticina 1 mg/ml. Se expanden las colonias resistentes y luego se someten a prueba para determinar su respuesta a concentraciones crecientes de interferón gamma. Se demuestra la respuesta a la dosis de un clon seleccionado.

Específicamente, el siguiente protocolo producirá células suficientes para 75 pocillos que contienen 200 ul de células. Por tanto, o bien se amplía a escala, o bien se realiza múltiples veces para generar células suficientes para múltiples placas de 96 pocillos. Las células Jurkat se mantienen en RPMI + 10% de suero con 1% de Pen-Strep (penicilina-estreptomicina). Combinar 2,5 ml de OPTI-MEM (Life Technologies) con 10 ug de ADN de plásmido en un frasco T25. Añadir 2,5 ml de OPTI-MEM que contiene 50 ul de DMRIE-C e incubar a temperatura ambiente durante 15-45 min.

Durante el periodo de incubación, determinar la concentración de células, centrifugar el número de células requerido ($10^7$ por transfección) y resuspender en OPTI-MEM hasta una concentración final de $10^7$ células/ml. Entonces añadir 1 ml de 1 x $10^7$ células en OPTI-MEM a un frasco T25 e incubar a 37ºC durante 6 h. Tras la incubación, añadir 10 ml de RPMI + 15% de suero.

Las líneas indicadoras estables Jurkat:GAS-SEAP se mantienen en RPMI + 10% de suero, 1 mg/ml de genticina, y 1% de Pen-Strep. Estas células se tratan con sobrenadantes que contienen un polipéptido producido por el protocolo descrito en el ejemplo 11.

En el día del tratamiento con el sobrenadante, las células deben lavarse y resuspenderse en RPMI + 10% de suero nuevo hasta una densidad de 500.000 células por ml. El número exacto de células requerido dependerá del número de sobrenadantes que se está seleccionando. Para una placa de 96 pocillos se requieren aproximadamente 10 millones de células (para 10 placas, 100 millones de células).

Transferir las células a una navecilla de depósito triangular, con el fin de dispensar las células a una placa de 96 pocillos, usando una pipeta de 12 canales. Usando una pipeta de 12 canales, transferir 200 ul de células a cada pocillo (por tanto añadiendo 100.000 células por pocillo).

Una vez sembradas todas las placas, se transfieren 50 ul de los sobrenadantes directamente desde la placa de 96 pocillos que contiene los sobrenadantes a cada pocillo usando una pipeta de 12 canales. Además, se añade una dosis de interferón gamma exógeno (0,1, 1,0, 10 ng) a los pocillos H9, H10 y H11 para servir como controles positivos adicionales para el ensayo.

Las placas de 96 pocillos que contienen células Jurkat tratadas con sobrenadantes se colocan en un incubador durante 48 h (nota: este tiempo es variable entre 48-72 h). Entonces se transfieren muestras de 35 ul de cada pocillo a una placa de 96 pocillos opaca usando una pipeta de 12 canales. Las placas opacas deben cubrirse (usando cubiertas de Sellophene) y almacenarse a -20ºC hasta que se realizan los ensayos de SEAP según el ejemplo 17. Las placas que contienen las células tratadas restantes se colocan a 4ºC y sirven como una fuente de material para repetir el ensayo en un pocillo especifico si se desea.

Como control positivo, pueden usarse 100 unidades/ml de interferón gamma que se sabe que activa las células T Jurkat. Normalmente se observa una inducción de más de 30 veces en los pocillos de control positivo.

Ejemplo 14

*Ensayo de selección de alto rendimiento que identifica la actividad mieloide*

Se usa el siguiente protocolo para evaluar la actividad mieloide mediante la identificación de factores, tales como factores de crecimiento y citocinas, que pueden hacer proliferar o diferenciar células mieloides. La actividad de las células mieloides se evalúa usando el constructo GAS/SEAP/Neo producido en el ejemplo 12. Por tanto, los factores que

aumentan la actividad de SEAP indican la capacidad para activar la ruta de transducción de señales Jak-STAT. La célula mieloide usada en este ensayo es U937, una línea celular premonocítica, aunque puede usarse TF-1, HL60 o KG1.

Para transfectar de manera transitoria células U937 con el constructo GAS/SEAP/Neo producido en el ejemplo 12, se usa un método de DEAE-dextrano (Kharbanda *et. al.*, 1994, Cell Growth & Differentiation, 5:259-265). En primer lugar, recoger 2 x $10e^7$ células U937 y lavar con PBS. Las células U937 se hacen crecer habitualmente en medio RPMI 1640 que contiene 10% de suero bovino fetal inactivado por calor (FBS) complementado con 100 unidades/ml de penicilina y 100 mg/ml de estreptomicina.

A continuación, suspender las células en 1 ml de tampón Tris-HCl 20 mM (pH 7,4) que contiene 0,5 mg/ml de DEAE-dextrano, 8 ug de ADN de plásmido GAS-SEAP2, NaCl 140 mM, KCl 5 mM, $Na_2HPO_4.7H_2O$ 375 uM, $MgCl_2$ 1 mM y $CaCl_2$ 675 uM. Incubar a 37ºC durante 45 min.

Lavar las células con medio RPMI 1640 que contiene 10% de FBS y luego resuspender en 10 ml de medio completo e incubar a 37ºC durante 36 h.

Las células estables GAS-SEAP/U937 se obtienen haciendo crecer las células en G418 400 ug/ml. Se usa el medio libre de G418 para el crecimiento de rutina, pero cada uno a dos meses las células deben hacerse crecer de nuevo en G418 400 ug/ml en un par de pases.

Estas células se someten a prueba recogiendo 1 x $10^8$ células (esto es suficiente para un ensayo de diez placas de 96 pocillos) y se lavan con PBS. Suspender las células en 200 ml del medio de crecimiento descrito anteriormente, con una densidad final de 5 x $10^5$ células/ml. Sembrar en placa 200 ul de células por pocillo en la placa de 96 pocillos (o 1 x $10^5$ células/pocillo).

Añadir 50 ul del sobrenadante preparado mediante el protocolo descrito en el ejemplo 11. Incubar a 37ºC durante de 48 a 72 h. Como control positivo, pueden usarse 100 unidades/ml de interferón gamma que se sabe que activa las células U937. Normalmente se observa una inducción de más de 30 veces en los pocillos de control positivo. Someter el sobrenadante al ensayo de SEAP según el protocolo descrito en el ejemplo 17.

Ejemplo 15

*Ensayo de selección de alto rendimiento que identifica la actividad neuronal*

Cuando las células experimentan diferenciación y proliferación, se activa un grupo de genes a través de muchas rutas de transducción de señales diferentes. Uno de estos genes, EGR1 (gen de respuesta temprana de crecimiento 1), se induce en diversos tejidos y tipos celulares con la activación. El promotor de EGR1 es responsable de tal inducción. Usando el promotor de EGR1 unido a moléculas indicadoras, puede evaluarse la activación de las células.

Particularmente, se usa el siguiente protocolo para evaluar la actividad neuronal en líneas celulares PC12. Se sabe que las células PC12 (células de feocromocitoma de rata) proliferan y/o se diferencian mediante la activación con varios mitógenos, tales como TPA (acetato de tetradecanoilforbol), NGF (factor de crecimiento nervioso) y EGF (factor de crecimiento epidérmico). La expresión del gen EGR1 se activa durante este tratamiento. Por tanto, transfectando de manera estable células PC12 con un constructo que contiene un promotor de EGR unido al indicador de SEAP, puede evaluarse la activación de las células PC12.

El constructo indicador EGR/SEAP puede ensamblarse mediante el siguiente protocolo. La secuencia del promotor de EGR-1 (de -633 a +1) (Sakamoto K *et al.*, Oncogene 6:867-871 (1991)) puede amplificarse mediante PCR a partir de ADN genómico humano usando los siguientes cebadores:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG-3' (SEQ ID NO: 6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO: 7)

Usando el vector GAS:SEAP/Neo producido en el ejemplo 12, puede insertarse entonces el producto amplificado de EGR1 en este vector. Linealizar el vector GAS:SEAP/Neo usando las enzimas de restricción XhoI/HindIII, eliminando el fragmento de relleno GAS/SV40. Someter a restricción el producto amplificado de EGR1 con estas mismas enzimas. Ligar el vector y el promotor de EGR1.

Para preparar placas de 96 pocillos para el cultivo celular, se añaden dos ml de una disolución de recubrimiento (dilución de colágeno tipo I 1:30 (Upstate Biotech Inc. n.º de catálogo 08-115) en etanol al 30% (filtro esterilizado) por una placa de 10 cm o 50 ml por pocillo de la placa de 96 pocillos, y se deja secar al aire durante 2 h.

Las células PC12 se hacen crecer rutinariamente en medio RPMI-1640 (Bio Whittaker) que contiene de suero de caballo al 10% (JRH BIOSCIENCES, n.º de catálogo 12449-78P), suero bovino fetal inactivado por calor (FBS) al 5% complementado con penicilina 100 unidades/ml y estreptomicina 100 ug/ml en una placa de cultivo tisular de 10 cm previamente recubierta. Se realizan de una a cuatro divisiones cada tres a cuatro días. Se extraen las células de las placas mediante raspado y se resuspenden pipeteando de arriba abajo más de 15 veces.

Transfectar el constructo EGR/SEAP/Neo en las células PC12 usando el protocolo de Lipofectamine descrito en el ejemplo 11. Se obtienen células estables EGR-SEAP/PC12 haciendo crecer las células en G418 300 ug/ml. Se usa el medio libre de G418 para el crecimiento de rutina, pero cada uno a dos meses las células deben hacerse crecer de nuevo en G418 300 ug/ml en un par de pases.

Para evaluar la actividad neuronal, se selecciona una placa de 10 cm con células confluentes en aproximadamente del 70 al 80% eliminando el medio antiguo. Lavar las células una vez con PBS (solución salina tamponada con fosfato). Luego, privar las células de nutrientes en medio bajo en suero (RPMI-1640 que contiene suero de caballo al 1% y FBS al 0,5% con antibióticos) durante la noche.

A la mañana siguiente, extraer el medio y lavar las células con PBS. Extraer por raspado las células de la placa, suspender las células bien en 2 ml de medio bajo en suero. Contar el número de células y añadir más medio bajo en suero hasta alcanzar la densidad de células final de 5 x $10^5$ células/ml.

Añadir 200 ul de la suspensión celular a cada pocillo de la placa de 96 pocillos (equivalente a 1 x $10^5$ células/pocillo). Añadir 50 ul de sobrenadante producido por el ejemplo 11, a 37ºC durante de 48 a 72 h. Como control positivo, puede usarse un factor de crecimiento que se sabe que activa las células PC12 a través de EGR, tal como 50 ng/ul de factor de crecimiento neuronal (NGF). Normalmente se observa una inducción de más de cincuenta veces de SEAP en los pocillos de control positivo. Someter el sobrenadante al ensayo de SEAP según el ejemplo 17.

Ejemplo 16

*Ensayo de selección de alto rendimiento para determinar la actividad de las células T*

NF-κB (factor nuclear κB) es un factor de transcripción activado por una amplia variedad de agentes, incluyendo las citocinas inflamatorias IL-1 y TNF, CD30 y CD40, linfotoxina-alfa y linfotoxina-beta, mediante la exposición a LPS o trombina, y mediante la expresión de ciertos productos génicos virales. Como factor de transcripción, NF-κB regula la expresión de genes implicados en la activación de las células inmunitarias, en el control de la apoptosis (NF-κB parece proteger a las células de la apoptosis), en el desarrollo de las células B y T, en las respuestas antivirales y antimicrobianas y en múltiples respuestas de estrés.

En condiciones no estimuladas, NF-κB queda retenido en el citoplasma con I-κB (inhibidor de κB). Sin embargo, con estimulación, I-κB se fosforila y se degrada, haciendo que NF-κB se transporte al núcleo, activando así la transcripción de genes diana. Los genes diana activados por NF-κB incluyen IL-2, IL-6, GM-CSF, ICAM-1 y MHC de clase 1.

Debido a su papel fundamental y a su capacidad para responder a una variedad de estímulos, se usan constructos indicadores que utilizan el elemento promotor de NF-κB para seleccionar los sobrenadantes producidos en el ejemplo 11. Activadores o inhibidores de NF-κB serían útiles en el tratamiento de enfermedades. Por ejemplo, podrían usarse inhibidores de NF-κB para tratar aquellas enfermedades relacionadas con la activación aguda o crónica de NF-κB, tales como la artritis reumatoide.

Para construir un vector que contenga el elemento promotor de NF-κB, se emplea una estrategia basada en PCR. El cebador en el sentido de 5' contiene cuatro copias en tándem del sitio de unión a NF-κB (GGGGACTTTCCC) (SEQ ID NO: 8), 18 pb de secuencia complementaria al extremo 5' de la secuencia del promotor temprano de SV40, y está flanqueado con un sitio XhoI:

5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

El cebador en el sentido de 3' es complementario al extremo 3' del promotor de SV40 y está flanqueado con un sitio HindIII:

5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO: 4)

Se lleva a cabo la amplificación por PCR usando el molde del promotor de SV40 presente en el plásmido pB-gal:promotor obtenido de Clontech. El fragmento de PCR resultante se digiere con XhoI e HindIII y se subclona en BLSK2- (Stratagene). La secuenciación con los cebadores de T7 y T3 confirma que el inserto contiene la siguiente secuencia:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA
TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT
AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
3' (SEQ ID NO:10)

A continuación, reemplazar el elemento promotor mínimo de SV40 presente en el plásmido pSEAP2-promotor (Clontech) por este fragmento NF-κB/SV40 usando XhoI y HindIII. Sin embargo, este vector no contiene un gen de resistencia a neomicina, y por tanto, no se prefiere para sistemas de expresión en mamíferos.

Con el fin de generar líneas celulares de mamífero estables, se elimina el casete NF-κB/SV40/SEAP del vector NF-κB/SEAP anterior usando las enzimas de restricción SalI y NotI, y se inserta en un vector que contiene resistencia a neomicina. Particularmente, se insertó el casete NF-κB/SV40/SEAP en pGFP-1 (Clontech), reemplazando el gen GFP, tras la restricción de pGFP-1 con SalI y NotI.

Una vez que se crea el vector NF-κB/SV40/SEAP/Neo, se crean células T Jurkat estables y se mantienen según el protocolo descrito en el ejemplo 13. De manera similar, el método para someter a ensayo sobrenadantes con estas células T Jurkat estables también se describe en el ejemplo 13. Como control positivo, se añade TNF alfa exógeno (0,1, 1, 10 ng) a los pocillos H9, H10 y H11, observándose normalmente una activación de 5-10 veces.

Ejemplo 17

*Ensayo para determinar la actividad de SEAP*

Como molécula indicadora para los ensayos descritos en los ejemplos 13-16, se somete a ensayo la actividad de SEAP usando el kit Phospho-light de Tropix (n.º de catálogo BP-400) según el siguiente procedimiento general. El kit Phospho-light de Tropix suministra los tampones de dilución, ensayo y reacción usados más adelante.

Cebar un dispensador con el tampón de dilución 2,5x y dispensar 15 μl de tampón de dilución 2,5x en placas Optiplate que contienen 35 μl de un sobrenadante. Sellar las placas con un sellador de plástico e incubar a 65ºC durante 30 min. Separar las placas Optiplate para evitar el calentamiento desigual.

Enfriar las muestras hasta temperatura ambiente durante 15 minutos. Vaciar el dispensador y cebar con el tampón de ensayo. Añadir 50 μl de tampón de ensayo e incubar a temperatura ambiente 5 min. Vaciar el dispensador y cebar con el tampón de reacción (véase la tabla a continuación). Añadir 50 μl de tampón de reacción e incubar a temperatura ambiente durante 20 minutos. Puesto que la intensidad de la señal quimioluminiscente depende del tiempo y se tarda aproximadamente 10 minutos en leer 5 placas en el luminómetro, deben tratarse 5 placas cada vez y comenzar con el segundo juego 10 minutos más tarde.

Leer la unidad relativa de luz en el luminómetro. Fijar H12 como blanco e imprimir los resultados. Un aumento en la quimioluminiscencia indica actividad indicadora.

*Formulación del tampón de reacción*

| N.º de placas | Diluyente de tampón de reacción (ml) | CSPD (ml) |
|---|---|---|
| 10 | 60 | 3 |
| 11 | 65 | 3,25 |
| 12 | 70 | 3,5 |
| 13 | 75 | 3,75 |
| 14 | 80 | 4 |
| 15 | 85 | 4,25 |
| 16 | 90 | 4,5 |
| 17 | 95 | 4,75 |
| 18 | 100 | 5 |
| 19 | 105 | 5,25 |
| 20 | 110 | 5,5 |
| 21 | 115 | 5,75 |

| | | |
|---|---|---|
| 22 | 120 | 6 |
| 23 | 125 | 6,25 |
| 24 | 130 | 6,5 |
| 25 | 135 | 6,75 |
| 26 | 140 | 7 |
| 27 | 145 | 7,25 |
| 28 | 150 | 7,5 |
| 29 | 155 | 7,75 |
| 30 | 160 | 8 |
| 31 | 165 | 8,25 |
| 32 | 170 | 8,5 |
| 33 | 175 | 8,75 |
| 34 | 180 | 9 |
| 35 | 185 | 9,25 |
| 36 | 190 | 9,5 |
| 37 | 195 | 9,75 |
| 38 | 200 | 10 |
| 39 | 205 | 10,25 |
| 40 | 210 | 10,5 |
| 41 | 215 | 10,75 |
| 42 | 220 | 11 |
| 43 | 225 | 11,25 |
| 44 | 230 | 11,5 |
| 45 | 235 | 11,75 |
| 46 | 240 | 12 |
| 47 | 245 | 12,25 |
| 48 | 250 | 12,5 |
| 49 | 255 | 12,75 |
| 50 | 260 | 13 |

Ejemplo 18

*Ensayo de selección de alto rendimiento que identifica cambios en la concentración de moléculas pequeñas y en la permeabilidad de la membrana*

Se sabe que la unión de un ligando a un receptor altera los niveles intracelulares de moléculas pequeñas, tales como calcio, potasio, sodio, y el pH, así como que altera el potencial de membrana. Estas alteraciones pueden medirse en un ensayo para identificar sobrenadantes que se unen a receptores de una célula particular. Aunque el siguiente protocolo describe un ensayo para el calcio, este protocolo puede modificarse fácilmente para detectar cambios en potasio, sodio, pH, potencial de membrana o cualquier otra molécula pequeña que pueda detectarse mediante una sonda fluorescente.

El siguiente ensayo usa un lector de placas fluorométrico de obtención de imágenes ("FLIPR") para medir cambios en moléculas fluorescentes (Molecular Probes) que se unen a moléculas pequeñas. Claramente, puede usarse cualquier molécula fluorescente que detecte una molécula pequeña en lugar de la molécula fluorescente de calcio, fluo-3, usada en este caso.

Para las células adherentes, sembrar las células a 10.000 - 20.000 células/pocillo en una placa negra Co-star de 96 pocillos con fondo transparente. La placa se incuba en un incubador de $CO_2$ durante 20 horas. Las células adherentes se lavan dos veces en el lavador Biotek con 200 ul de HBSS (solución salina equilibrada de Hank) dejando 100 ul de tampón tras el lavado final.

Se prepara una disolución madre de 1 mg/ml de fluo-3 en 10% de ácido plurónico - DMSO. Para cargar las células con fluo-3, se añaden 50 ul de fluo-3 12 ug/ml a cada pocillo. La placa se incuba a 37ºC en un incubador de $CO_2$ durante 60 min. La placa se lava cuatro veces en el lavador Biotek con HBSS dejando 100 ul de tampón.

Para las células no adherentes, se centrifugan las células a partir de los medios de cultivo. Se resuspenden las células hasta 2-5 x $10^6$ células/ml con HBSS en un tubo cónico de 50 ml. Se añaden 4 ul de una disolución de fluo-3 1 mg/ml en 10% de ácido plurónico - DMSO a cada ml de suspensión celular. Entonces de coloca el tubo en un baño de agua a 37ºC durante 30-60 min. Se lavan las células dos veces con HBSS, se resuspenden hasta 1 x $10^6$ células/ml, y se dispensan en una microplaca, 100 ul/pocillo. Se centrifuga la placa a 1000 rpm durante 5 min. Entonces se lava la placa una vez en Denley CellWash (sistema de lavado celular Denley) con 200 ul, seguido por una etapa de aspiración hasta 100 ul de volumen final.

Para un ensayo sin células, cada pocillo contiene una molécula fluorescente, tal como fluo-3. Se añade el sobrenadante al pocillo y se detecta un cambio en la fluorescencia.

Para medir la fluorescencia del calcio intracelular, se fija el FLIPR para los siguientes parámetros: (1) la ganancia del sistema es 300-800 mW; (2) el tiempo de exposición es de 0,4 segundos; (3) F/stop (apertura relativa) de la cámara es F/2; (4) la excitación es de 488 nm; (5) la emisión es de 530 nm; y (6) la adición de la muestra es de 50 ul. El aumento de la emisión a 530 nm indica un acontecimiento de señalización extracelular que ha dado como resultado un aumento en la concentración intracelular de $Ca^{++}$.

Ejemplo 19

*Ensayo de selección de alto rendimiento que identifica la actividad de la tirosina cinasa*

Las proteínas tirosina cinasa (PTK) representan un grupo diverso de cinasas transmembrana y citoplasmáticas. Dentro del grupo de receptores de proteína tirosina cinasa (RPTK) hay receptores para una variedad de factores de crecimiento mitogénicos y metabólicos incluyendo las subfamilias de receptores de PDGF, FGF, EGF, NGF, HGF e insulina. Además, hay una gran familia de RPTK para la que se desconoce el ligando correspondiente. Los ligandos para los RPTK incluyen principalmente proteínas pequeñas secretadas, pero también proteínas unidas a la membrana y de la matriz extracelular.

La activación de RPTK por ligandos implica la dimerización del receptor mediada por ligando, dando como resultado la transfosforilación de las subunidades del receptor y la activación de las tirosina cinasas citoplasmáticas. Las tirosina cinasas citoplasmáticas incluyen tirosina cinasas asociadas a receptor de la familia src (por ejemplo, src, yes, lck, lyn, fyn) y proteínas tirosina cinasas no unidas a receptor y citosólicas, tales como la familia Jak, cuyos miembros median la transducción de señales desencadenada por la superfamila de receptores de citocina (por ejemplo, las interleucinas, interferones, GM-CSF y leptina).

Debido a la amplia variedad de factores conocidos que pueden estimular la actividad de la tirosina cinasa, es de interés la identificación de proteínas secretadas humanas novedosas que puedan activar rutas de transducción de señales de la tirosina cinasa. Por tanto, el siguiente protocolo está diseñado para identificar esas proteínas secretadas humanas novedosas que pueden activar las rutas de transducción de señales de la tirosina cinasa.

Sembrar las células diana (por ejemplo, queratinocitos primarios) a una densidad de aproximadamente 25.000 células por pocillo en una placa de selección silenciosa Loprodyne de 96 pocillos adquirida de Nalge Nunc (Naperville, IL). Se esterilizan las placas con dos enjuagues de 30 minutos con etanol al 100%, se enjuagan con agua y se secan durante la noche. Se recubren algunas placas durante 2 h con 100 ml de colágeno de tipo I de calidad para cultivo celular (50 mg/ml), gelatina (2%) o polilisina (50 mg/ml), pudiendo adquirirse todo ello de Sigma Chemicals (St. Louis, MO) o Matrigel al 10% adquirido de Becton Dickinson (Bedford, MA), o suero de ternero, enjuagado con PBS y almacenado a 4ºC. Se somete a ensayo el crecimiento celular en estas placas sembrando 5.000 células/pocillo en medio de crecimiento y la cuantificación indirecta del número de células a través del uso de alamarBlue, tal como describe el fabricante Alamar Biosciences, Inc. (Sacramento, CA) tras 48 h. Se usan cubiertas de placa Falcon n.º 3071 de Becton Dickinson (Bedford, MA) para cubrir las placas de selección silenciosa Loprodyne. También pueden usarse placas de cultivo celular Falcon Microtest III en algunos experimentos de proliferación.

Para preparar los extractos, se siembran células A431 en las membranas de nailon de las placas Loprodyne (20.000/200 ml/pocillo) y se cultivan durante la noche en medio completo. Las células se hacen quiescentes mediante

la incubación en medio basal libre de suero durante 24 h. Tras de 5-20 minutos de tratamiento con EGF (60 ng/ml) o 50 ul del sobrenadante producido en el ejemplo 11, se extrajo el medio y se añaden a cada pocillo 100 ml de tampón de extracción (HEPES 20 mM pH 7,5, NaCl 0,15 M, Triton X-100 al 1%, SDS al 0,1%, $Na_3VO_4$ 2 mM, $Na_4P_2O_7$ 2 mM y un cóctel de inhibidores de proteasa (n.º 1836170) obtenido de Boeheringer Mannheim (Indianápolis, IN) y se agita la placa en un agitador rotatorio durante 5 minutos a 4ºC. Entonces se coloca la placa en un colector de transferencia de vacío y se filtra el extracto a través de los fondos de membrana de 0,45 mm de cada pocillo usando una aspiradora doméstica. Se recogen los extractos en una placa de captura/ensayo de 96 pocillos en el fondo del colector de vacío y se colocan inmediatamente en hielo. Para obtener los extractos clarificados mediante centrifugación, se extrae el contenido de cada pocillo, tras solubilización con detergente durante 5 minutos, y se centrifuga durante 15 minutos a 4ºC a 16.000 x g.

Someter a prueba los extractos filtrados para determinar los niveles de actividad de la tirosina cinasa. Aunque se conocen muchos métodos de detección de la actividad de la tirosina cinasa, en el presente documento se describe un método.

Generalmente, se evalúa la actividad de la tirosina cinasa de un sobrenadante determinando su capacidad para fosforilar un resto de tirosina en un sustrato específico (un péptido biotinilado). Los péptidos biotinilados que pueden usarse para este fin incluyen PSK1 (que corresponde a los aminoácidos 6-20 de la cinasa de la división celular cdc2-p34) y PSK2 (que corresponde a los aminoácidos 1-17 de la gastrina). Ambos péptidos son sustratos de una variedad de tirosina cinasas y están disponibles de Boehringer Mannheim.

Se establece la reacción de la tirosina cinasa añadiendo los siguientes componentes en orden. En primer lugar, añadir 10 ul de péptido biotinilado 5 uM, luego 10 ul de ATP/$Mg_{2+}$ (ATP 5 mM/$MgCl_2$ 50 mM), luego 10 ul de 5x tampón de ensayo (clorhidrato de imidazol 40 mM, pH 7,3, beta-glicerofosfato 40 mM, EGTA 1 mM, $MgCl_2$ 100 mM, $MnCl_2$ 5 mM, BSA 0,5 mg/ml), luego 5 ul de vanadato de sodio (1 mM), y luego 5 ul de agua. Mezclar los componentes suavemente y preincubar la mezcla de reacción a 30ºC durante 2 min. Iniciar la reacción añadiendo 10 ul de enzima control o el sobrenadante filtrado.

Entonces se termina la reacción de ensayo de la tirosina cinasa añadiendo 10 ul de EDTA 120 mM y se colocan las reacciones en hielo.

Se determina la actividad de la tirosina cinasa transfiriendo una alícuota de 50 ul de mezcla de reacción a un módulo de placas de microtitulación (MTP) e incubando a 37ºC durante 20 min. Esto permite que la placa de 96 pocillos recubierta con estreptavidina se asocie con el péptido biotinilado. Lavar el módulo de MTP con 300 ul/pocillo de PBS cuatro veces. A continuación añadir 75 ul de anticuerpo anti-fosfotirosina conjugado con peroxidasa del rábano picante (anti-P-Tyr-POD (0,5 u/ml)) a cada pocillo e incubar a 37ºC durante una hora. Lavar el pocillo como anteriormente.

A continuación añadir 100 ul de disolución de sustrato de peroxidase (Boehringer Mannheim) e incubar a temperatura ambiente durante al menos 5 min. (hasta 30 min.). Medir la absorbancia de la muestra a 405 nm usando un lector de ELISA. Se cuantifica el nivel de actividad de la peroxidasa unida usando un lector de ELISA y refleja el nivel de actividad de la tirosina cinasa.

Ejemplo 20

*Ensayo de selección de alto rendimiento que identifica la actividad de fosforilación*

Como posible alternativa y/o complemento al ensayo de la actividad de la proteína tirosina cinasa descrito en el ejemplo 19, también puede usarse un ensayo que detecte la activación (fosforilación) de productos intermedios de la transducción de señales intracelulares principales. Por ejemplo, tal como se describe más adelante, un ensayo particular puede detectar la fosforilación de la tirosina de las Erk-1 y Erk-2 cinasas. Sin embargo, la fosforilación de otras moléculas, tales como Raf, JNK, p38 MAP, Map cinasa (MEK), MEK cinasa, Src, cinasa específica del músculo (MuSK), IRAK, Tec y Janus, así como cualquier otra molécula de fosfoserina, fosfotirosina o fosfotreonina, puede detectarse sustituyendo estas moléculas por Erk-1 o Erk-2 en el siguiente ensayo.

Específicamente, se preparan placas de ensayo recubriendo los pocillos de una placa de ELISA de 96 pocillos con 0,1 ml de proteína G (1 ug/ml) durante 2 h a temperatura ambiente (TA). Entonces se enjuagan las placas con PBS y se bloquean con 3% de BSA/PBS durante 1 h a TA. Se tratan entonces las placas con proteína G con 2 anticuerpos monoclonales comerciales (100 ng/pocillo) frente a Erk-1 y Erk-2 (1 h a TA) (Santa Cruz Biotechnology). (Para detectar otras moléculas, puede modificarse fácilmente esta etapa sustituyendo por un anticuerpo monoclonal que detecte cualquiera de las moléculas descritas anteriormente). Tras 3-5 enjuagues con PBS, se almacenan las placas a 4ºC hasta su uso.

Se siembran células A431 a 20.000/pocillo en una placa de filtro de 96 pocillos Loprodyne y se cultivan durante la noche en medio de crecimiento. Entonces se privan las células de nutrientes durante 48 h en medio basal (DMEM) y luego se tratan con EGF (6 ng/pocillo) o 50 ul de los sobrenadantes obtenidos en el ejemplo 11 durante 5-20 minutos. Entonces se solubilizan las células y se filtran los extractos directamente en la placa de ensayo.

Tras la incubación con el extracto durante 1 h a TA, se enjuagan de nuevo los pocillos. Como control positivo, se usa una preparación comercial de MAP cinasa (10 ng/pocillo) en lugar del extracto de A431. Entonces se tratan las placas con un anticuerpo policlonal comercial (de conejo) (1 ug/ml) que reconoce específicamente el epítopo fosforilado de las Erk-1 y Erk-2 cinasas (1 h a TA). Se biotiniliza este anticuerpo mediante procedimientos convencionales. Entonces se cuantifica el anticuerpo policlonal unido mediante incubaciones sucesivas con europio-estreptavidina y reactivo que mejora la fluorescencia del europio en el instrumento DELFIA de Wallac (fluorescencia de resolución temporal). Una señal fluorescente aumentada con respecto al nivel inicial indica una fosforilación.

Ejemplo 21

*Método de determinación de alteraciones en un gen que corresponde a un polinucleótido*

Ha de aislarse ARN aislado de familias enteras o pacientes individuales que se presentan un fenotipo de interés (tal como una enfermedad). Entonces se genera ADNc a partir de estas muestras de ARN usando protocolos conocidos en la técnica. (Véase, Sambrook). Entonces se usa el ADNc como molde para PCR, empleando cebadores que rodean las regiones de interés en la SEQ ID NO: X. Las condiciones de PCR sugeridas consisten en 35 ciclos a 95ºC durante 30 segundos; 60-120 segundos a 52-58ºC; y 60-120 segundos a 70ºC, usando disoluciones tampón descritas en Sidransky, D., *et al.*, Science 252:706 (1991).

Entonces se secuencian los productos de la PCR usando cebadores marcados en su extremo 5' con polinucleótido cinasa de T4, empleando la polimerasa SequiTherm. (Epicentre Technologies). También se determinan los límites intrón-exón de exones seleccionados y se analizan los productos genómicos de la PCR para confirmar los resultados. Entonces se clonan los productos de la PCR que albergan mutaciones sospechadas y se secuencian para validar los resultados de la secuenciación directa.

Los productos de la PCR se clonan en vectores de cola de T tal como se describe en Holton, T.A. y Graham, M.W., Nucleic Acids Research, 19:1156 (1991) y se secuencian con polimerasa de T7 (United States Biochemical). Se identifican los individuos afectados por las mutaciones no presentes en los individuos no afectados.

También se observan transposiciones genómicas como método de determinación de alteraciones en un gen que corresponde a un polinucleótido. Clones genómicos aislados según el ejemplo 2 se someten a traslado de mella con digoxigenin deoxiuridina 5'-trifosfato (Boehringer Manheim), y FISH realizada tal como se describe en Johnson, Cg. *et al.*, Methods Cell Biol. 35: 73-99 (1991). Se lleva a cabo la hibridación con la sonda marcada usando un enorme exceso de ADN cot-1 humano para la hibridación específica en el locus genómico correspondiente.

Se contratiñen cromosomas con 4,6-diamino-2-fenilidol y yoduro de propidio, produciendo una combinación de bandas C y R. Se obtienen imágenes alineadas para el mapeo preciso usando un equipo de filtro de triple banda (Chroma Technology, Brattleboro, VT) en combinación con una cámara con dispositivo de carga acoplada con enfriamiento (Photometrics, Tucson, AZ) y filtros de longitud de onda de excitación variable. (Johnson, Cv. *et al.*, Genet. Anal. Tech. Appl., 8:75 (1991)). La recogida de imágenes, el análisis y las mediciones de la longitud fraccional cromosómica se realizan usando el sistema ISee Graphical Program. (Inovision Corporation, Durham, NC). Las alteraciones cromosómicas de la región genómica hibridada por la sonda se identifican como inserciones, deleciones y translocaciones. Estas alteraciones se usan como marcador de diagnóstico para una enfermedad asociada.

Ejemplo 22

*Método de detección de niveles anómalos de un polipéptido en una muestra biológica*

Puede detectarse un polipéptido de la presente invención en una muestra biológica, y si se detecta un nivel aumentado o disminuido del polipéptido, este polipéptido es un marcador para un fenotipo particular. Los métodos de detección son numerosos, y por tanto, se entiende que un experto en la técnica puede modificar el siguiente ensayo para adaptarlo a sus necesidades particulares.

Por ejemplo, se usan ELISA *sándwich* con anticuerpos para detectar polipéptidos en una muestra, preferiblemente una muestra biológica. Se recubren los pocillos de una placa de microtitulación con anticuerpos específicos, a una concentración final de 0,2 a 10 ug/ml. Los anticuerpos son o bien monoclonales o bien policlonales y se producen mediante el método descrito en el ejemplo 10. Se bloquean los pocillos de modo que se reduzca la unión no específica del polipéptido al pocillo.

Entonces se incuban los pocillos recubiertos durante > 2 horas a TA con una muestra que contiene el polipéptido. Preferiblemente, deben usarse diluciones en serie de la muestra para validar los resultados. Entonces se lavan las placas tres veces con agua desionizada o destilada para eliminar el polipéptido no unido.

A continuación, se añaden 50 ul de conjugado de anticuerpo específico-fosfatasa alcalina, a una concentración de 25-400 ng, y se incuban durante 2 horas a temperatura ambiente. Las placas se lavan de nuevo tres veces con agua desionizada o destilada para eliminar el conjugado no unido.

Añadir 75 ul de disolución de sustrato de 4-metilumbeliferil fosfato (MUP) o p-nitrofenil fosfato (NPP) a cada pocillo e incubar 1 hora a temperatura ambiente. Medir la reacción mediante un lector de placas de microtitulación. Preparar una curva patrón, usando diluciones en serie de una muestra control, y representar gráficamente la concentración de polipéptido en el eje X (escala logarítmica) y la fluorescencia o la absorbancia en el eje Y (escala lineal). Interpolar la concentración del polipéptido en la muestra usando la curva patrón.

Ejemplo 23

*Formulación de un polipéptido*

Se formulará y se dosificará la composición de polipéptido secretada de una manera consecuente con la práctica médica, teniendo en cuenta el estado clínico del paciente individual (especialmente los efectos secundarios del tratamiento con el polipéptido secretado solo), el sitio de administración, el método de administración, el programa de administración y otros factores conocidos por los médicos. Por tanto, se determina la "cantidad eficaz" para los fines del presente documento según tales consideraciones.

Como propuesta general, la cantidad farmacéuticamente eficaz total de polipéptido secretado administrada por vía parenteral por dosis estará en el intervalo de aproximadamente 1 $\mu$g/kg/día a 10 mg/kg/día de peso corporal del paciente, aunque, tal como se observó anteriormente, esto se someterá al criterio terapéutico. Más preferiblemente, esta dosis es de al menos 0,01 mg/kg/día, y lo más preferiblemente para seres humanos entre aproximadamente 0,01 y 1 mg/kg/día para la hormona. Si se administra de manera continua, el polipéptido secretado normalmente se administra a una tasa de dosis de aproximadamente 1 $\mu$g/kg/hora a aproximadamente 50 $\mu$g/kg/hora, o bien mediante 1-4 inyecciones al día o bien mediante infusiones subcutáneas continuas, por ejemplo, usando una minibomba. También puede emplearse una disolución en bolsa intravenosa. La duración del tratamiento necesaria para observar cambios y el intervalo tras el tratamiento para que se produzcan respuestas parece variar dependiendo del efecto deseado.

Las composiciones farmacéuticas que contienen la proteína secretada de la invención se administran por vía oral, rectal, parenteral, intracisternal, intravaginal, intraperitoneal, tópica (tal como mediante polvos, pomadas, geles, gotas o parche transdérmico), bucal, o como un aerosol oral o nasal. "Vehículo farmacéuticamente aceptable" se refiere a una formulación auxiliar, carga líquida, semisólida o sólida, diluyente, material de encapsulación de cualquier tipo, no tóxicos. El término "parenteral", tal como se usa en el presente documento, se refiere a modos de administración que incluyen inyección e infusión intravenosa, intramuscular, intraperitoneal, intraesternal, subcutánea e intraarticular.

El polipéptido secretado también se administra de manera adecuada mediante sistemas de liberación sostenida. Los ejemplos adecuados de composiciones de liberación sostenida incluyen matrices de polímero semipermeables en forma de artículos conformados, por ejemplo, películas o microcápsulas. Las matrices de liberación sostenida incluyen polilactidas (patente estadounidense n.º 3.773.919, documento EP 58.481), copolímeros de ácido L-glutámico y gamma-etil-L-glutamato (Sidman, U. *et al.*, Biopolymers 22:547-556 (1983)), poli(metacrilato de 2-hidroxietilo) (R. Langer *et al.*, J. Biomed. Mater. Res. 15:167-277 (1981), y R. Langer, Chem. Tech. 12:98-105 (1982)), acetato de etilenvinilo (R. Langer *et al.*) o poli(ácido D-(-)-3-hidroxibutírico) (documento EP 133.988). Las composiciones de liberación sostenida también incluyen polipéptidos atrapados en liposomas. Los liposomas que contienen el polipéptido secretado se preparan mediante métodos conocidos *per se*: documento DE 3.218.121; Epstein *et al.*, Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); documentos EP 52.322; EP 36.676; EP 88.046; EP 143.949; EP 142.641; solicitud de patente japonesa 83-118008; patentes estadounidenses n.$^{os}$ 4.485.045 y 4.544.545; y documento EP 102.324. Habitualmente, los liposomas son del tipo unilamelar pequeño (aproximadamente 200-800 Ángstroms) en el que el contenido en lípidos es superior a aproximadamente 30 por ciento en moles de colesterol, ajustándose la proporción seleccionada para la terapia óptima con el polipéptido secretado.

Para la administración parenteral, en una realización, el polipéptido secretado se formula generalmente mezclándolo con el grado de pureza deseado, en una forma farmacéutica unitaria inyectable (disolución, suspensión o emulsión), con un vehículo farmacéuticamente aceptable, es decir, uno que no sea tóxico para los receptores a las dosificaciones y concentraciones empleadas y que sea compatible con otros componentes de la formulación. Por ejemplo, la formulación preferiblemente no incluye agentes oxidantes y otros compuestos que se sabe que son perjudiciales para los polipéptidos.

Generalmente, las formulaciones se preparan poniendo en contacto el polipéptido de manera uniforme e íntima con vehículos líquidos o vehículos sólidos finamente divididos o ambos. Luego, si es necesario, se da forma al producto en la formulación deseada. Preferiblemente, el vehículo es un vehículo parenteral, más preferiblemente una disolución que es isotónica con la sangre del receptor. Ejemplos de tales vehículos incluyen agua, solución salina, solución de Ringer, y solución de dextrosa. También son útiles en este caso vehículos no acuosos tales como aceites fijos y oleato de etilo, así como liposomas.

El vehículo contiene de manera adecuada cantidades minoritarias de aditivos tales como sustancias que potencian la isotonicidad y la estabilidad química. Tales materiales son no tóxicos para los receptores a las dosificaciones y concentraciones empleadas, e incluyen tampones tales como fosfato, citrato, succinato, ácido acético y otros ácidos orgánicos o sus sales; antioxidantes tales como ácido ascórbico; polipéptidos de bajo peso molecular (inferior a aproximadamente diez restos), por ejemplo, poliarginina o tripéptidos; proteínas, tales como albúmina sérica, gelatina o inmunoglobulinas; polímeros hidrófilos tales como polivinilpirrolidona; aminoácidos, tales como glicina, ácido glu-

támico, ácido aspártico o arginina; monosacáridos, disacáridos y otros hidratos de carbono incluyendo celulosa o sus derivados, glucosa, manosa o dextrinas; agentes quelantes tales como EDTA; alcoholes de azúcar tales como manitol o sorbitol; contraiones tales como sodio; y/o tensioactivos no iónicos tales como polisorbatos, poloxámeros o PEG.

El polipéptido secretado normalmente se formula en tales vehículos a una concentración de aproximadamente 0,1 mg/ml a 100 mg/ml, preferiblemente 1-10 mg/ml, a un pH de aproximadamente 3 a 8. Se entenderá que el uso de algunos de los excipientes, vehículos o estabilizadores anteriores dará como resultado la formación de sales de polipéptido.

Cualquier polipéptido que va a usarse para la administración terapéutica puede hacerse estéril. La esterilidad se obtiene fácilmente mediante filtración a través de membranas de filtración estériles (por ejemplo, membranas de 0,2 micras). Las composiciones de polipéptido terapéuticas generalmente se colocan en un recipiente que tiene un orificio de acceso estéril, por ejemplo, una bolsa o vial de disolución intravenosa que tiene un tapón que puede perforarse mediante una aguja de inyección hipodérmica.

Los polipéptidos se almacenarán habitualmente en recipientes de dosis unitaria o múltiple, por ejemplo, ampollas o viales sellados, como disolución acuosa o como formulación liofilizada para su reconstitución. Como ejemplo de formulación liofilizada, se llenan viales de 10 ml con 5 ml de disolución acuosa de polipéptido al 1% (p/v) esterilizada mediante filtración, y se liofiliza la mezcla resultante. La disolución de infusión se prepara reconstituyendo el polipéptido liofilizado usando agua para inyección bacteriostática.

La invención también proporciona un paquete o kit farmacéutico que comprende uno o más recipientes llenos con uno o más de los componentes de las composiciones farmacéuticas de la invención. Asociado con tal(es) recipiente(s) puede haber un aviso en la forma recomendada por el organismo que regula la fabricación, el uso o la venta de productos farmacéuticos o biológicos, aviso que refleja la aprobación por el organismo de la fabricación, el uso o la venta para la administración en seres humanos. Además, los polipéptidos de la presente invención pueden emplearse conjuntamente con otros compuestos terapéuticos.

Ejemplo 24

*Método de tratamiento de niveles disminuidos del polipéptido*

Se apreciará que los estados producidos por una disminución en el nivel de expresión convencional o normal de una proteína secretada en un individuo pueden tratarse mediante la administración del polipéptido de la presente invención, preferiblemente en la forma secretada. Por tanto, la invención también proporciona un método de tratamiento de un individuo que necesita un nivel aumentado del polipéptido que comprende administrar a tal individuo una composición farmacéutica que comprende una cantidad del polipéptido para aumentar el nivel de actividad del polipéptido en tal individuo.

Por ejemplo, un paciente con niveles disminuidos de un polipéptido recibe una dosis diaria de 0,1-100 ug/kg del polipéptido durante seis días consecutivos. Preferiblemente, el polipéptido está en la forma secretada. Los detalles exactos del esquema de dosificación, basado en la administración y la formulación, se facilitan en el ejemplo 23.

Ejemplo 25

*Método de tratamiento de niveles aumentados del polipéptido*

Se usa la tecnología antisentido para inhibir la producción de un polipéptido de la presente invención. Esta tecnología es un ejemplo de un método de disminución de los niveles de un polipéptido, preferiblemente una forma secretada, debido a una variedad de etiologías, tales como cáncer.

Por ejemplo, a un paciente al que se ha diagnosticado niveles anómalamente aumentados de un polipéptido, se le administran por vía intravenosa polinucleótidos antisentido a 0,5, 1,0, 1,5, 2,0 y 3,0 mg/kg al día durante 21 días. Este tratamiento se repite tras un periodo de descanso de 7 días si el tratamiento se toleró bien. La formulación del polinucleótido antisentido se proporciona en el ejemplo 23.

Ejemplo 26

*Método de tratamiento usando terapia génica*

Un método de terapia génica es mediante el trasplante de fibroblastos, que pueden expresar un polipéptido, en un paciente. Generalmente, los fibroblastos se obtienen de un sujeto mediante biopsia cutánea. El tejido resultante se coloca en un medio de cultivo tisular y se separa en pequeñas porciones. Los pequeños trozos de tejido se colocan sobre una superficie húmeda de un frasco de cultivo tisular, aproximadamente se colocan diez porciones en cada matraz. El matraz se coloca del revés, se cierra herméticamente y se deja a temperatura ambiente durante la noche. Tras 24 horas a temperatura ambiente, se invierte el matraz y los trozos de tejido se quedan fijados en el fondo del matraz y se añaden medios nuevos (por ejemplo, medios F12 de Ham, con 10% de FBS, penicilina y estreptomicina). Entonces se incuban los frascos a 37ºC durante aproximadamente una semana.

En este momento, se añaden medios nuevos y posteriormente se cambian cada varios días. Tras dos semanas adicionales en cultivo, aparece una monocapa de fibroblastos. La monocapa se trata con tripsina y amplía a escala en frascos más grandes.

Se digiere con EcoRI y HindIII el vector pMV-7 (Kirschmeier, P.T. *et al.*, DNA, 7:219-25 (1988)), flanqueado por las repeticiones terminales largas del virus del sarcoma murino de Moloney y posteriormente se trata con fosfatasa intestinal de ternero. El vector lineal se fracciona en gel de agarosa y se purifica, usando perlas de vidrio.

El ADNc que codifica un polipéptido de la presente invención puede amplificarse usando cebadores de PCR que corresponden a las secuencias de los extremos 5' y 3', respectivamente, tal como se expone en el ejemplo 1. Preferiblemente, el cebador en 5' contiene un sitio EcoRI y el cebador en 3' incluye un sitio HindIII. Se añaden juntas cantidades iguales de la estructura principal lineal del virus del sarcoma murino de Moloney y el fragmento amplificado con EcoRI e HindIII, en presencia de ADN ligasa de T4. La mezcla resultante se mantiene en condiciones apropiadas para la ligación de los dos fragmentos. Entonces se usa la mezcla de ligación para transformar bacterias HB101, que luego se siembran en placa sobre agar que contiene kanamicina para confirmar que el vector tiene el gen de interés insertado correctamente.

Se hacen crecer las células de empaquetamiento GP+am12 o pA317 anfotróficas en cultivo tisular hasta la densidad confluente en medio de Eagle modificado por Dulbecco (DMEM) con 10% de suero de ternero (ST), penicilina y estreptomicina. Entonces se añade el vector del VSM que contiene el gen a los medios y se transducen las células de empaquetamiento con el vector. Las células de empaquetamiento producen ahora partículas virales infecciosas que contienen el gen (las células de empaquetamiento se denominan ahora células productoras).

Se añaden medios nuevos a las células productoras transducidas y posteriormente se recogen los medios de una placa de 10 cm de células productoras confluentes. Se filtran los medios gastados, que contienen las partículas virales infecciosas, a través de un filtro Millipore para eliminar las células productoras separadas y estos medios se usan luego para infectar fibroblastos. Se eliminan los medios de una placa subconfluente de fibroblastos y se sustituyen rápidamente por los medios de las células productoras. Estos medios se eliminan y se sustituyen por medios nuevos. Si el título de virus es alto, entonces prácticamente todos los fibroblastos resultarán infectados y no se requiere selección. Si el título es muy bajo, entonces es necesario usar un vector retroviral que tenga un marcador seleccionable, tal como neo o his. Una vez que los fibroblastos se han infectado eficazmente, se analizan los fibroblastos para determinar si produce proteína.

Entonces se trasplantan los fibroblastos modificados por ingeniería genética en el huésped, o bien solos o bien tras haberse hecho crecer hasta la confluencia en perlas de microsoporte Cytodex 3.

## REIVINDICACIONES

1. Un polinucleótido seleccionado del grupo que consiste en:

(a) un polinucleótido que tiene toda la secuencia de nucleótidos de SEQ ID NO: 23;

(b) un polinucleótido que tiene toda la secuencia de nucleótidos de la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23;

(c) un fragmento de polinucleótido que comprende 150 nucleótidos contiguos de SEQ ID NO: 23;

(d) un fragmento de polinucleótido que comprende 150 nucleótidos de la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23;

(e) un polinucleótido que codifica un fragmento de polipéptido, en el que dicho fragmento de polipéptido comprende 40 ó 50 restos de aminoácido contiguos de longitud de SEQ ID NO: 60;

(f) un polinucleótido que codifica un fragmento de polipéptido codificado por la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23, en el que dicho fragmento de polipéptido comprende 40 ó 50 restos de aminoácido contiguos de SEQ ID NO:60;

(g) un polinucleótido que codifica un fragmento de polipéptido de SEQ ID NO: 60 o un fragmento de polipéptido del mismo codificado por la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, en el que dicho fragmento de polipéptido comprende 40 ó 50 restos de aminoácido contiguos de longitud y tiene actividad biológica;

(h) un polinucleótido que codifica un polipéptido de longitud completa de SEQ ID NO: 60 o que está codificado por la secuencia de ADNc incluida en el depósito de la ATCC n.º 209075, que puede hibridarse en condiciones rigurosas con SEQ ID NO: 23;

(i) un polinucleótido que es al menos el 95% idéntico a al menos 150 nucleótidos contiguos de un polinucleótido como el definido en uno cualquiera de (a) a (h); y

(j) un polinucleótido que codifica un polipéptido que tiene una secuencia de aminoácidos al menos 95% idéntica a la secuencia de aminoácidos de un polipéptido codificado por un polinucleótido de uno cualquiera de (a) a (i);

o la hebra complementaria del polinucleótido (a) a (j).

2. Polinucleótido de la reivindicación 1, que es ADN o ARN.

3. El gen correspondiente a la secuencia de ADNc de SEQ ID NO: 23 o que codifica el polipéptido que tiene la secuencia de aminoácidos de SEQ ID NO: 60.

4. Un vector que comprende el polinucleótido de la reivindicación 1 ó 2.

5. Un método de preparación de una célula huésped recombinante que comprende introducir el polinucleótido de la reivindicación 1 ó 2 o el vector de la reivindicación 4.

6. Una célula huésped recombinante que contiene el polinucleótido de la reivindicación 1 ó 2 o el vector de la reivindicación 4 o producido mediante el método de la reivindicación 5.

7. Célula huésped recombinante de la reivindicación 6, que expresa un polipéptido codificado por el polinucleótido de la reivindicación 1 ó 2.

8. Un método de preparación de un polipéptido codificado por el polinucleótido de la reivindicación 1 ó 2, que comprende;

(a) cultivar la célula huésped recombinante de la reivindicación 7 en condiciones tales que se exprese dicho polipéptido; y

(b) recuperar dicho polipéptido.

9. Un polipéptido codificado por el polinucleótido de la reivindicación 1 ó 2 o que puede obtenerse mediante el método de la reivindicación 8.

10. Un anticuerpo que se une específicamente al polipéptido de la reivindicación 9.

11. Un antagonista que comprende un polinucleótido antisentido con respecto al polinucleótido de la reivindicación 1 ó 2 que puede unirse a dicho polinucleótido y bloquear su traducción en el polipéptido codificado.

12. Un método para identificar antagonistas o agonistas del polipéptido de la reivindicación 9 que comprende:

(a) producir células que expresan el polipéptido de la reivindicación 9;

(b) poner en contacto el polipéptido producido en la etapa (a) con una muestra de prueba que contiene potencialmente un antagonista o agonista; e

(c) identificar un antagonista o agonista observando la unión y la inhibición o estimulación de la actividad de dicho polipéptido.

13. Un método *in vitro* de diagnóstico de un estado patológico o una susceptibilidad a un estado patológico en un sujeto que comprende:

(a) determinar la presencia o ausencia de una mutación en el polinucleótido de la reivindicación 1 ó 2 o el gen de la reivindicación 3; y

(b) diagnosticar un estado patológico o una susceptibilidad a un estado patológico basado en la presencia o ausencia de dicha mutación.

14. Un método *in vitro* de diagnóstico de un estado patológico o una susceptibilidad a un estado patológico en un sujeto que comprende:

(a) determinar la presencia o cantidad de expresión del polipéptido de la reivindicación 9 en una muestra biológica; y

(b) diagnosticar un estado patológico o una susceptibilidad a un estado patológico basado en la presencia o cantidad de expresión del polipéptido.

15. Un método para identificar una pareja de unión al polipéptido de la reivindicación 9 que comprende:

(a) poner en contacto el polipéptido de la reivindicación 9 con una pareja de unión; y

(b) determinar si la pareja de unión efectúa una actividad del polipéptido.

16. Una composición farmacéutica que comprende el polinucleótido de la reivindicación 1 ó 2, el polipéptido de la reivindicación 9, un ADNc que codifica y puede expresar dicho polipéptido *in vivo*, el anticuerpo de la reivindicación 10, o el antagonista de la reivindicación 11, y opcionalmente un vehículo farmacéuticamente aceptable.

17. Una composición de diagnóstico que comprende el polinucleótido de la reivindicación 1 ó 2 o el anticuerpo de la reivindicación 10.

18. Un método para la producción de una composición farmacéutica que comprende las etapas del método de la reivindicación 12 y (d) formular el compuesto identificado en la etapa (c) en una forma farmacéuticamente aceptable.

**LISTA DE SECUENCIAS**

(1) INFORMACIÓN GENERAL:

(i) SOLICITANTE: Human Genome Sciences, Inc., *et al.*

(ii) TÍTULO DE LA INVENCIÓN: 32 Proteínas humanas secretadas

(iii) NÚMERO DE SECUENCIAS: 120

(iv) DIRECCIÓN DE CORRESPONDENCIA:

(A) DESTINATARIO: Human Genome Sciences, Inc.

(B) CALLE: 9410 Key West Avenue

(C) CIUDAD: Rockville

(D) ESTADO: Maryland

(E) PAÍS: EE. UU.

(F) CÓDIGO POSTAL: 20850

(v) FORMA LEGIBLE EN ORDENADOR:

(A) TIPO DE MEDIO: Disquete, 3,50 pulgadas, almacenamiento de 1,4 Mb

(B) ORDENADOR: HP Vectra 486/33

(C) SISTEMA OPERATIVO: MSDOS versión 6.2

(D) SOFTWARE: Texto ASCII

(vi) DATOS DE LA SOLICITUD ACTUAL:

(A) NÚMERO DE SOLICITUD:

(B) FECHA DE PRESENTACIÓN: 27 de mayo de 1998

(C) CLASIFICACIÓN:

(vii) DATOS DE LA SOLICITUD ANTERIOR:

(A) NÚMERO DE SOLICITUD:

(B) FECHA DE PRESENTACIÓN:

(viii) INFORMACIÓN DEL ABOGADO/AGENTE:

(A) NOMBRE: A. Anders Brookes

(B) NÚMERO DE REGISTRO: 36.373

(C) NÚMERO DE REFERENCIA/EXPEDIENTE: PZ006PCT

(vi) INFORMACIÓN DE TELECOMUNICACIÓN:

(A) TELÉFONO: (301) 309-8504

(B) FAX: (301) 309-8439

(2) INFORMACIÓN PARA SEQ ID NO: 1:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 733 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 1:

```
GGGATCCGGA GCCCAAATCT TCTGACAAAA CTCACACATG CCCACCGTGC CCAGCACCTG      60

AATTCGAGGG TGCACCGTCA GTCTTCCTCT TCCCCCCAAA ACCCAAGGAC ACCCTCATGA     120

TCTCCCGGAC TCCTGAGGTC ACATGCGTGG TGGTGGACGT AAGCCACGAA GACCCTGAGG     180

TCAAGTTCAA CTGGTACGTG GACGGCGTGG AGGTGCATAA TGCCAAGACA AAGCCGCGGG     240

AGGAGCAGTA CAACAGCACG TACCGTGTGG TCAGCGTCCT CACCGTCCTG CACCAGGACT     300

GGCTGAATGG CAAGGAGTAC AAGTGCAAGG TCTCCAACAA AGCCCTCCCA ACCCCCATCG     360

AGAAAACCAT CTCCAAAGCC AAAGGGCAGC CCCGAGAACC ACAGGTGTAC ACCCTGCCCC     420

CATCCCGGGA TGAGCTGACC AAGAACCAGG TCAGCCTGAC CTGCCTGGTC AAAGGCTTCT     480

ATCCAAGCGA CATCGCCGTG GAGTGGGAGA GCAATGGGCA GCCGGAGAAC AACTACAAGA     540

CCACGCCTCC CGTGCTGGAC TCCGACGGCT CCTTCTTCCT CTACAGCAAG CTCACCGTGG     600

ACAAGAGCAG GTGGCAGCAG GGGAACGTCT TCTCATGCTC CGTGATGCAT GAGGCTCTGC     660

ACAACCACTA CACGCAGAAG AGCCTCTCCC TGTCTCCGGG TAAATGAGTG CGACGGCCGC     720

GACTCTAGAG GAT                                                        733
```

(2) INFORMACIÓN PARA SEQ ID NO: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 5 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 2:

```
Trp Ser Xaa Trp Ser
 1               5
```

(2) INFORMACIÓN PARA SEQ ID NO: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 86 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 3:

```
GCGCCTCGAG ATTTCCCCGA AATCTAGATT TCCCCGAAAT GATTTCCCCG AAATGATTTC      60

CCCGAAATAT CTGCCATCTC AATTAG                                           86
```

(2) INFORMACIÓN PARA SEQ ID NO: 4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 27 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 4:

GCGGCAAGCT TTTTGCAAAG CCTAGGC 27

(2) INFORMACIÓN PARA SEQ ID NO: 5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 271 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 5:

CTCGAGATTT CCCCGAAATC TAGATTTCCC CGAAATGATT TCCCCGAAAT GATTTCCCCG 60

AAATATCTGC CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC 120

GCCCCTAACT CCGCCCAGTT CCGCCCATTC TCCGCCCCAT GGCTGACTAA TTTTTTTTAT 180

TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT 240

TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT T 271

(2) INFORMACIÓN PARA SEQ ID NO: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 32 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 6:

GCGCTCGAGG GATGACAGCG ATAGAACCCC GG 32

(2) INFORMACIÓN PARA SEQ ID NO: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 31 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 7:

GCGAAGCTTC GCGACTCCCC GGATCCGCCT C 31

(2) INFORMACIÓN PARA SEQ ID NO: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 12 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 8:

GGGGACTTTC CC 12

(2) INFORMACIÓN PARA SEQ ID NO: 9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 73 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 9:

GCGGCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCGGGACT TTCCATCCTG 60

CCATCTCAAT TAG 73

(2) INFORMACIÓN PARA SEQ ID NO: 10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 256 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 10:

CTCGAGGGGA CTTTCCCGGG GACTTTCCGG GGACTTTCCG GGACTTTCCA TCTGCCATCT 60

CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC 120

CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA 180

GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG 240

CTTTTGCAAA AAGCTT 256

(2) INFORMACIÓN PARA SEQ ID NO: 11:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 1169 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 11:

```
GGGGCGCAAA TAGGGTCAGT GGGCCGCTTG GCGKTGTTCG TTGCGGTACC AGGTCCGCGT      60
GAGGGGTTCG GGGGTTCTGG GCAGGCACAA TGGCGTCTCG AGCAGGCCCG CGAGCGGCCG     120
RCACCGACGC AGCGAGCTTT CAGCACCGGG AGCGCGTCGC CATGCACTAC CAGATGAGTG     180
TGACCCTCAA GTATGAAATC AAGAAGCTGA TCTACGTACA TCTGGTCATA TGGCTGCTGC     240
TGGTTGCTAA GATGAGCGTG GGACACCTGA GGCTCTTGTC ACATGATCAG GTGGCCATGC     300
CCTATCAGTG GGAATACCCG TATTTGCTGA GCATTTTGCC CTCTCTCTTG GGCCTTCTCT     360
CCTTTCCCCG CAACAACATT AGCTACCTGG TGCTCTCCAT GATCAGCATG GGACTCTTTT     420
CCATCGCTCC ACTCATTTAT GGCAGCATGG AGATGTTCCC TGCTGCACAG CCTTCTACCG     480
CCATGGCAAG GCCTACCGTT TCCTCTTTGG TTTTTCTGCC GTTTCCATCA TGTACCTGGT     540
GTTGGTGTTG GCAGTGCAAG TGCATGCCTG GCAGTTGTAC TACAGCAAGA AGCTCCTAGA     600
CTCTTGGTTC ACCAGCACAC AGGAGAAGAA GCATAAATGA AGCCTCTTTG GGGTGAAGCC     660
TGGACATCCC ATCGAATGAA AGGACACTAG TACAGCGGTT CCAAAATCCC TTCTGGTGAT     720
TTTAGCAGCT GTGATGTTGG TACCTGGTGC AGACCCAGGC CAAAGTTCTG GAAAGCTCCT     780
TTTGCCATCT GCTGAGGTGG CAAAACTATA ATTTATTCCT GGTTGGCTAG AACTGGGTGA     840
CCAACAGCTA TGAAACAAAT TTCAGCTGTT TGAAGTTGAA CTTTGAGGTT TTTCTTTAAG     900
AATGAGCTTC GTCCTTGCCT CTACTCGGTC ATTCTCCCCA TTTCCATCCA TTACCCCTTA     960
GCCATTGAGA CTAAAGGAAA TAGGGAATAA ATCAAATTAC TTCATCTCTA GGTCACGGGT    1020
CAGGAAACAT TTGGGCAGCT GCTCCCTTGG CAGCTGTGGT CTCCTCTGCA AAGCATTTTA    1080
ATTAAAAACC TCAATAAAGA TGCCCTGCCC ACAAAAAAAA AAAAAAAAAA AATTCGGGGG    1140

GGGGCCCGGG NAACCAATTN GCCCCTANA                                      1169
```

(2) INFORMACIÓN PARA SEQ ID NO: 12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1310 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 12:

```
AATTCGGCAC GAGGCAGCGT CGCGCGGCCC AGTTCCCTTT TCCGGTCGGC GTGGTCTTGC      60
GAGTGGAGTG TCCGCTGTGC CCGGGCCTGC ACCATGAGCG TCCCGGCCTT CATCGACATC     120
AGTGAAGAAG ATCAGGCTGC TGAGCTTCGT GCTTATCTGA AATCTAAAGG AGCTGAGATT     180
TCAGAAGAGA ACTCGGAAGG TGGACTTCAT GTTGATTTAG CTCAAATTAT TGAAGCCTGT     240
GATGTGTGTC TGAAGGAGGA TGATAAAGAT GTTGAAAGTG TGATGAACAG TGTGGTATCC     300
CTACTCTTGA TCCTGGAACC AGACAAGCAA GAAGCTTTGA TTGAAAGCCT ATGTGAAAAG     360
CTGGTCAAAT TTCGCGAAGG TGAACGCCCG TCTCTGAGAC TGCAGTTGTT AAGCAACCTT     420
TTCCACGGGA TGGATAAGAA TACTCCTGTA AGATACACAG TGTATTGCAG CCTTATTAAA     480
GTGGCAGCAT CTTGTGGGGC CATCCAGTAC ATCCCAACTG AGCTGGATCA AGTTAGAAAA     540
TGGATTTCTG ACTGGAATCT CACCACTGAA AAAAAGCACA CCCTTTTAAG ACTACTTTAT     600
GAGGCACTTG TGGATTGTAA GAAGAGTGAT GCTGCTTCAA AAGTCATGGT GGAATTGCTC     660
GGAAGTTACA CAGAGGACAA TGCTTCCCAG GCTCGAGTTG ATGCCCACAG GTGTATTGTA     720
CGAGCATTGA AAGATCCAAA TGCATTTCTT TTTGACCACC TTCTTACTTT AAAACCAGTC     780
AAGTTTTTGG AAGGCGAGCT TATTCATGAT CTTTTAACCA TTTTTGTGAG TGCTAAATTG     840
GCATCATATG TCAAGTTTTA TCAGAATAAT AAAGACTTCA TTGATTCACT TGGCCTGTTA     900
CATGAACAGA ATATGGCAAA AATGAGACTA CTTACTTTTA TGGGAATGGC AGTAGAAAAT     960
AAGGAAATTT CTTTTGACAC AATGCAGCAA GAACTTCAGA TTGGAGCTGA TGATGTTGAA    1020
GCATTTGTTA TTGACGCCGT AAGAACTAAA ATGGTCTACT GCAAAATTGA TCAGACCCAG    1080
AGAAAAGTAG TTGTCAGTCA TAGCACACAT CGGACATTTG GAAAACAGCA GTGGCAACAA    1140
CTGTATGACA CACTTAATGC CTGGAAACAA AATCTGAACA AAGTGAAAAA CAGCCTTTTG    1200
AGTCTTTCTG ATACCTGAGT TTTTATGCTT ATAATTTTTG TTCTTTGAAA AAAAAGCCCT    1260
AAATCATAGT AAAACATTAT AAACTAAAAA AAAAAAAAAA AAAAAAAAAA               1310
```

(2) INFORMACIÓN PARA SEQ ID NO: 13:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1139 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 13:

```
AGGGCANACT TACAGAGATA TCATATGAGA TCACCCCTCG CATTCGTGTC TGGCGCCAGA      60
CCCTCGAGCG GTGCCGGAGC GCASCCAGGT GTGCTTGTGC CTGGGCCAGC TGGAGAGGTC     120
CATTGCCTGG GANGAAGTCT GTCAACAAAG TGACATGTCT AGTCTGCCGG AAGGGTGACA     180
ATGATGAGTT TCTTCTGCTT TGTGATGGGT GTRACCGTGG CTGCCACATT TACTGCCATC     240
GTCCCAAGAT GGAGGCTGTC CCAGAAGGAG ATTGGTTCTG TACTGTCTGT TTGGCTCAGC     300
AGGTGGAGGG AGAATTCACT CAGAAGCCTG GTTTCCCAAA GCGTGGCCAG AAGCGGAAAA     360
GTGGTTATTC GCTGAACTTC TCAGAGGGTG ATGGCCGCCG ACGCCGGGTA CTGTTGAGGG     420
GCCGAGAAAG CCCAGCAGCA GGGCCTCGGT ACTCGGAAGA AGGGCTCTCC CCCTCCAAGC     480
GGCGGCGACT CTCTATGCGG AACCACCACA GTGATCTCAC ATTTTGCGAG ATTATCCTGA     540
TGGAGATGGA GTCCCATGAT GCAGCCTGGC CTTTCCTAGA GCCTGTGAAC CCACGTTTGG     600
TGAGTGGGTA CCGGCGCATC ATCAAAAATC CTATGGATTT TTCCACCATG CGGGAGCGGC     660
TGCTCAGGGG AGGGTACACC AGCTCAGAGG AGTTTGCGGC TGATGCCCTC CTGGTATTTG     720
ACAACTGCCA GACTTTCAAC GAGGATGACT CTGAAGTAGG CAAGGCTGGG CACATCATGC     780
GCCGCTTCTT CGAGAGCCGC TGGGAGGAGT TTTATCAGGG AAAACAGGCC AATCTGTGAG     840
GCAAGGGAGG TGGGGAGTCA CCTTGTGGCA TCTCCCCCCA CCTTCCAAAC AAAAACCTGC     900
CATTTTCACC TGCTGATGCT GCCCTGGGTC CAGACTCAAG TCAGATACAA CCCTGATTTT     960
TGACCTTNCC CTTGGCAGTG CCCCACATCC TCTTATTCCT ACATCCCTTT CTCCCTTCCC    1020
TCCTCTTGCT CCTCAAGTAA GAGGTGCAGA GATGAGGTCC TTCTGGACTA AAAGCCAAAA    1080
AAAGAAAGAA AAAAAWAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAN     1139
```

(2) INFORMACIÓN PARA SEQ ID NO: 14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 2271 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 14:

```
GTTCCGGGGG ATGCCAGCTC ACTTCTCCGA CAGCGCCCAG ACTGAGGCCT GCTACCACAT      60
GCTGAGCCGG CCCCAGCCGC CACCCGACCC CCTCCTGCTC CAGCGTCTGC CACGGCCCAG     120
CTCCCTGTCA GACAAGACCC AGCTCCACAG CAGGTGGCTG GACTCGTCGC GGTGTCTCAT     180
GCAGCAGGGC ATCAAGGCCG GGGACGCACT CTGGCTGCGC TTCAAGTACT ACAGCTTCTT     240
CGATTTGGAT CCCAAGACAG ACCCCGTGCG GCTGACACAG CTGTATGAGC AGGCCCGGTG     300
GGACCTGCTG CTGGAGGAGA TTGACTGCAC CGAGGAGGAG ATGATGGTGT TTGCCGCCCT     360
GCAGTACCAC ATCAACAAGC TGTCCCAGAG CGGGGAGGTG GGGGAGCCGG CTGGCACAGA     420
CCCAGGGCTG GACGACCTGG ATGTGGCCCT GAGCAACCTG GAGGTGAAGC TGGAGGGGTC     480
GGCGCCCACA GATGTGCTGG ACAGCCTCAC CACCATCCCA GAGCTCAAGG ACCATCTCCG     540
AATCTTTCGG CCCCGGAAGC TGACCCTGAA GGGCTACCGC CAACACTGGG TGGTGTTCAA     600
GGAGACCACA CTGTCCTACT ACAAGAGCCA GGACGAGGCC CCTGGGGACC CCATTCAGCA     660
GCTCAACCTC AAGGGCTGTG AGGTGGTTCC CGATGTTAAC GTCTCCGGCC AGAAGTTCTG     720
CATTAAACTC CTAGTGCCCT CCCCTGAGGG CATGAGTGAG ATCTACCTGC GGTGCCAGGA     780
TGAGCAGCAG TATGCCCGCT GGATGGCTGG CTGCCGCCTG GCCTCCAAAG GCCGCACCAT     840
GGCCGACAGC AGCTACACCA GCGAGGTGCA GGCCATCCTG GCCTTCCTCA GCCTGCAGCG     900
CACGGGCAGT GGGGGCCCGG GCAACCACCC CCACGGCCCT GATGCCTCTG CCGAGGGCCT     960
CAACCCCTAC GGCCTCGTTG CCCCCCGTTT CCAGCGAAAG TTCAAGGCCA AGCAGCTCAC    1020
CCCACGGATC CTGGAAGCCC ACCAGAATGT GGCCCAGTTG TCGCTGGCAG AGGCCCAGCT    1080
GCGCTTCATC CAGGCCTGGC AGTCCCTGCC CGACTTCGGC ATCTCCTATG TCATGGTCAG    1140
GTTCAAGGGC AGCAGGAAAG ACGAGATCCT GGGCATCGCC AACAACCGAC TGATCCGCAT    1200
CGACTTGGCC GTGGGCGACG TGGTCAAGAC CTGGCGTTTC AGCAACATGC GCCAGTGGAA    1260
TGTCAACTGG GACATCCGGC AGGTGGCCAT CGAGTTTGAT GAACACATCA ATGTGGCCTT    1320
CAGCTGCGTG TCTGCCAGCT GCCGAATTGT ACACGAGTAT ATCGGGGGCT ACATTTTCCT    1380
GTCGACGCGG GAGCGGGCCC GTGGGGAGGA GCTGGATGAA GACCTCTTCC TGCAGCTCAC    1440
CGGGGGCCAT GAGGCCTTCT GAGGGCTGTC TGATTGCCCC TGCCCTGCTC ACCACCCTGT    1500
CACAGCCACT CCCAAGCCCA CACCCACAGG GGCTCACTGC CCCACACCCG CTCCAGGCAG    1560
GCACCCAGCT GGGCATTTCA CCTGCTGTCA CTGACTTTGT GCAGGCCAAG GACCTGGCAG    1620
GGCCAGACGC TGTACCATCA CCCAGGCCAG GGATGGGGGT GGGGGTCCCT GAGCTCATGT    1680
```

```
GGTGCCCCCT TTCCTTGTCT GAGTGGCTGA GGCTGATACC CCTGACCTAT CTGCAGTCCC    1740

CCAGCACACA AGGAAGACCA GATGTAGCTA CAGGATGATG AAACATGGTT TCAAACGAGT    1800

TCTTTCTTGT TACTTTTTAA AATTTCTTTT TTATAAATTA ATATTTTATT GTTGGATCCT    1860

CCTCCTTTCT CTGGAGCTGT GCTTGGGGCT ACTCTGACAC TCTGTCTCTT CATCACCAGC    1920

CAAGGAAAGG GGCTTTCCTG ATAAAGACAA GAGTTGGTTA GAGAAAGGGA CACCTAAGTC    1980

AGTCTAGGGT TGGAAGCTAG GAGAGAGGTG AGGGCAGAAG GGCACAGCTT TCAGGAACAA    2040

GGAATAGGGG CTGGGGTKGT KGTTCTCACG GGTAGGCGTA CCTGCAGGGC CTCCTTGAAG    2100

TACTTGGGAA GGAGGAAGCC ATCAGTATTC CCTGGAGTCA GAATCACCCC ATTGGCAGAG    2160

CGGAAGAAGG GTATTCCATC TGCTGACAGA GCCAGAGATG TGACTCATGC CCTCCCCGAA    2220

GGCAAAGTCA GCTCCTGCTT TGTCCAGACT CACCTGCCAG AGCCAGGGGT C              2271
```

(2) INFORMACIÓN PARA SEQ ID NO: 15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 626 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 15:

```
ACAACAAACA TCGAAAATCG ANTATGTGCC CCGAAAAGTC GGAACGCAGG CAATCAGTCC      60

GCACGMGCGC AAGTTCAACA TGAAGATGAT ATGAGGCCGG GGCGGGGGGC AGGGACCCCC     120

GGGCGGCCGG GCAGGGGAAG GGGCCTGGCC GCCACCTGCT CACTCTCCAG TCCTTCCCAC     180

CTCCTCCCTA CCCTTCTACA CACGTTCTCT TTCTCCCTCC CGCCTCCGTC CCCTGCTGCC     240

CCCCGCCAGC CCTCACCACC TGCCCTCCTT CTACCAGGAC CTCAGAAGCC CAGACCTGGG     300

GACCCCACCT ACACAGGGGC ATTGACAGAC TGGAGTTGAA AGCCGACGAA CCGACACGCG     360

GCAGAGTCAA TAATTCAATA AAAAAGTTAC GAACTTTCTC TGTAACTTGG GTTTCAATAA     420

TTATGGATTT TTATGAAAAC TTGAAATAAT AAAAAGAGAA AAAAACTATT TCCTATAGCT     480

AGTCGGAATG CAAACTTTTG ACGTCCTGAT TGCTCCAGGG CCCTCTTTCC AACTCAGTTT     540

CTTGTTTTTC CTCTTCCTCC TCCTCCTCTT CTTCCTCCTT TCTTTCTCTT NCCCCATGGG     600

GGAGGGGTTC ATTCAGGGAA AACAGG                                          626
```

(2) INFORMACIÓN PARA SEQ ID NO: 16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 2118 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 16:

TTTTCCAGCC ATGTCACTAA TTGTGAATTC CTACCAACTA TTGACAGAAT ACAGAGTTGA 60

TTTTTTAATA AAAAGTTATA TATAATTATC CCTTTAATTA AAGGGAGCAA AGGGGCGTTC 120

CACATGGACA GAGGCTTGGA CCGAGGCCTG GTCACAGCAG CGAGCATCCA GGGTTTGCAG 180

GGACGATGTT ACAGACTCTG TTTTCTGCCT GGCGTTTCAC TTGTGTCTGC TCCTAGCCTG 240

TGCTCTGCCA GCAGCACAGA CATCTGCTCC ATCAGACCTC TTCCATTTTG CACAGGGAGT 300

GCAGGAGGTG AATGTTCACT TTCTGTTCTC CAGTGTCACT GTTCTGTTTC CACGGGATGG 360

AAAGCGCATG GGCCTGTGTC CATTGTAGAT TTCCTTCTAG ATTTCTGTGT ACACACACTT 420

GATTGTTCTG GATGAATGTC TTTTTTAATA CTCCGAAAAT TTCATCATCT AAGAAAATGA 480

TTCCATACAA ATAACTCAGC ACACAAGTGA CCCAGGACAT ATGCCTGCCA AAGGGATGTG 540

TTAGAAGGCT GCCTTCTCAT GCGCATTGTC ACTTGGATCT TGTGGTGAGG ACGGCCCCAT 600

CTTTCTTGCC ACAGATTGAG GCCACTTTTG AGCAAGGGAG ATCCTGGAGT TAAGACAGGT 660

GTTGGGGGCA GCCTGTATTT TACCCTAGGG GCAGGTCTGC ATGGTGACCC CACATYGCAC 720

TGGTAAACCA TTTGAGTCCC ACTCTTCATC CTGGAAGTGG GAACTGGAGT CCCACCCACA 780

GTGCATTCAG AAAGCATGCT GTGTGGGGGC TGCTTCTCAG GAGGCCAGGC CCTTCTGAGC 840

GGAACCGTCC TGGAGAGAGC CTGCCCTCGT TTCCAGGCTG CAGCCGTAAC GCACTTTCTC 900

CCAGGCTGAG GGCGGGTGTT CTGGGGTGTC TGCCCTCTGT CGGCCCTGCT TCCTGCCAGG 960

ACGTGGCCTC TTCCGATCCT TTTCTCTCAG ACACTGGAGG TCTCTTCTGC CATTGTGCTG 1020

GTCCCATCCC AAGAATTGTA GGACAGAGAC CACACTGGGT CGGCGGACAC AAAGTCCATC 1080

CAGGACCCAG GCCGCAGAGG GAGCAGGAAG AGATGCTGAT AGTTTGATCT AGAAACCAGC 1140

AGCTACTGGC TCAAATTCAG GTTCTGGCGT CAAATAGCGA CATTTCCAGT TTCTCTTAAA 1200

AACCGTGTTT GGTTTCAGTT GGGATAGGCT TGTTTTGTCT GTTGAAAATG TTTCTAGTTT 1260

TTTTTCTTTC ATTTTTCTCT CATTCCATTT CTGCCTTAAC TTTAGTTTGT TCACAGGGAG 1320

GCAAAGCTGA CATGAACCTT TTGTCGTGGG ACTTCAGGCC ACATTGGCTT GAAGGCATTC 1380

GTTTCCTTCT GGGGTGGGGA CAGGCCCTCA TGGCAGGCTT GTTCCCGTGG CTCTGAGCGA 1440

GGCCTCTTCC TGCTGGGCTC CCAGACTCCT GCATCCAGGC CCCCACCTTC TCGGCTTCTG 1500

GTTTTTCTTT CTTTTTGGTA GAACACAACA TCTACCATTC AGTTAAACCT TCTTTATCTC 1560

```
CTCCTYTGGC ATCCATTTTT CCAAAGAAGA GTCGAGTCCT CTGAGGTCTG TGCTTGAAAR    1620

CCGTCCGAAG GCATTCTTGT TAGCTTTGCT TTTCTCCCCA TATCCCAAGG CGAAGCGCTG    1680

AGATTCTTCC ATCTAAAAAA CCCTCGACCC GAAACCCTCA CCAGATAAAC TACAGTTTGT    1740

TTAGGAGGCC CTGACCTTCA TGGTGTCTTT GAAGCCCAAC CACTCGGTTT CCTTCGGATT    1800

TTCCTCCCTT TGTTCGGGGT TTGGTTTGGC TCCTCTGTGT GTGTCCGTAT CTTGTTCGGT    1860

GTCCTCGAGG TTGAGCTTCA CTCCACTGCG GCAGAGGCAG CGTGCACACT CGGATTTGCT    1920

ACGTTTCTAT ATATCTTGAA GCTAAATGTA TATATGAGTA GTTTGCCATG AGATAACACA    1980

GTGTAAACAG TAGACACCCA GAAATCGTGA CTTCTGTGTT CTCTCCATTT GAGTATTTTG    2040

TAATTTTTTT GAAATATTTG TGGACATAAA TAAAACCAAG CTACACTACA AAAAAAAAAA    2100

AAAAAAACTG GAGACTAG                                                  2118
```

(2) INFORMACIÓN PARA SEQ ID NO: 17:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1076 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 17:

```
GCCCAAGGAG CTCAGCTTCG CCCGCATCAA GGCCGTTGAG TGCGTGGAGA GCACCGGGCG      60

CCACATCTAC TTCACGCTGG TGACCGAAGG GWGCGGCGAG ATCGACTTCC GCTGCCCCCT     120

GGAAGATCCC GGCTGGAACG CCCAGATCAC CCTAGGCCTG GTCAAGTTCA AGAACCAGCA     180

GGCCATCCAG ACAGTGCGGG CCCGGCAGAG CCTCGGGACC GGGACCCTCG TGTCCTAAAC     240

CACCGGGCGC ACCATCTTTC CTTCATGCTA CCCACCACCT CAGTGCTGAG GTCAAGGCAG     300

CTTCGTTGTT CCCTCTGGCT TGTGGGGGCA CGGCTGTSYT CCATGTGGCA AGGTGGAAGG     360

CATGGACGTG TGGAGGAGGC GCTGGAGCTG AAGGAATGGA CGAGCCCTGG GAGGAGGGCA     420

GAAGGCTACG CAGGGCTGAG GATGAAGATG CAGCCCCTGG ATGGTCCCAG ACTCTCAGGA     480

CATGCCCAGC TCAGGGGCTT CGAGCCACAG GCCTGGCCTC ATATGGCATG AGGGGGAGCT     540

GGCATAGGAG CCCCCTCCCT GCTGTGGTCC TGCCCTCTGT CCTGCAGACT GCTCTTAGCC     600

CCCTGGCTTT GTGCCAGGCC TGGAGGAGGG CAGTCCCCCA TGGGGTGCCG AGCCAACGCC     660

TCAGGAATCA GGAGGCCAGC CTGGTACCAA AAGGAGTACC CAGGGCCTGG TACCCAGGCC     720

CACTCCAGAA TGGCCTCTGG ACTCACCTTG AGAAGGGGGA GCTGCTGGGC CTAAAGCCCA     780

CTCCTGGGGG TCTCCTGCTG CTTAGGTCCT TTTGGGACCC CCACCCATCC AGGCCCTTTC     840
```

```
TTTGCACACT TCTTCCCCCA CCTCTAYGCA TCTTCCCCCC ACTGCGGTGT TCGGCCTGAA      900

GGTGGTGGGG GTGAGGGGGG GTTTGGCCAT TAGCATTTCA TGTCTTTCCC CAAATGAAGA      960

TGCCCTGCAA AGGGCAGTNA ACCACAAAAA AAAAAAAAAA AAAAACNTGG GGGGGGGGCC     1020

CCGTTAACCA TTTTGGCCTN ATAGGGGGGN GGTTTTTAAA AATTAATTGG GCCCCG         1076
```

(2) INFORMACIÓN PARA SEQ ID NO: 18:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1379 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 18:

```
GGCACGAGCA CCCTCCCACA CCTCCCTGAA CTTCCATCTG ATCGACTTCA ACTTGCTGAT       60

GGTGACCACC ATCGTTCTGG GCCGCCGCTT CATTGGGTCC ATCGTGAAGG AGGCCTCTCA      120

GAGGGGGAAG GTCTCCCTCT TTCGCTCCAT CCTGCTGTTC CTCACTCGCT TCACCGTTCT      180

CACGGCAACA GGCTGGAGTC TGTGCCGATC CCTCATCCAC CTCTTCAGGA CCTACTCCTT      240

CCTGAACCTC CTGTTCCTCT GCTATCCGTT TGGGATGTAC ATTCCGTTCC TGCARCTGAA      300

TTKCGAMCTY CGSAAGACAA GCCTCTTCAA CCACATGGCC TCCATGGGGC CCCGGGAGGC      360

GGTCAGTGGC CTGGCAAAGA GCCGGGACTA CCTCCTGACA CTGCGGGAGA CGTGGAAGCA      420

GCACASAAGA CAGCTGTATG GCCCGGACGC CATGCCCACC CATGCCTGCT GCCTGTCGCC      480

CAGCCTCATC CGCAGTGAGG TGGAGTTCCT CAAGATGGAC TTCAACTGGC GCATGAAGGA      540

AGTGCTCGTS AGCTCCATGC TGAGCGCCTA CTATGTGGCC TTTGTGCCTG TYTGGTTCGT      600

GAAGAACACA CATTACTATG ACAAGCGCTG GTCCTGTGNA ACTCTTCCTG CTGGTGTCCA      660

TCAGCACCTC CGTGATCCTC ATGCAGCACC TGCTGCNTGC CAGCTACTGT GACCTGCTGC      720

ACAAGGCCGC CGCCCATCTG GGCTGTTGGC AGAAGGTGGA CCCAGCGCTG TGCTCCAACG      780

TGCTGCAGCA CCCGTGGACT GAAGAATGCA TGTGGCCGCA GGGCGTGCTG GTGAAGCACA      840

GCAAGAACGT CTACAAAGCC GTAGGCCAMW ACAAMGTGGC TATCCCCTCT GACGTCTCCC      900

ACTTCCGCTT CCAKTTCTTT TTCAGCAAAC CCCTGCGGAT CCTCAACATC CTCCTGCTGC      960

TGGAGGGCGC TGTCATTGTC TATCAGCTGT ACTCCCTAAT GTCCTCTGAA AAGTGGCACC     1020

AGACCATCTC GCTGGCCCTC ATCCTCTTCA GCAACTACTA TGCCTTCTTC AAGCTGCTCC     1080

GGGACCGCTT GGTATTGGGC AAGGCCTACT CATACTCTGC TAGCCCCCAG AGAGACCTGG     1140
```

ACCACCGTTT CTCCTGAGCC CTGGGGTCAC CTCAGGGACA GCGTCCAGGC TTCAGCAAGG 1200

GCTCCCTGGC AAGGGGCTGT TGGGTAGAAG TGGTGGTGGG GGGGACAAAA GACAAAAAAA 1260

TCCACCAGAG CTTTGTATTT TTGTTACGTA CTGTTTCTTT GATAATTGAT GTGATAAGGA 1320

AAAAAGTCCT ATTTTTATAC TCCCAANMAA AAAAAAAAAA NAAAAAGCGG CCGAAAGCT 1379

(2) INFORMACIÓN PARA SEQ ID NO: 19:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1337 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 19:

CTGGTGTTGG GCCTGAGCCN CCTCAACAAC TCCTACAACT TCAGTTTCCA CGTGGTGATC 60

GGCTCTCAGG CGGAAGAAGG CCAGTACAGC CTGAACTTCC ACAACTGCAA CAATTCAGTG 120

CCAGGAAAGG AGCATCCATT CGACATCACG GTGATGATCC GGGAGAAGAA CCCCGATGGC 180

TTCCTGTCGG CAGCGGAGAT GCCCCTTTTC AAGCTCTACA TGGTCATGTC CGCCTGCTTC 240

CTGGCCGCTG GCATCTTCTG GGTGTCCATC CTCTGCAGGA ACACGTACAG CGTCTTCAAG 300

ATCCACTGGC TCATGGCGGC CTTGGCCTTC ACCAAGAGCA TCTCTCTCCT CTTCCACAGC 360

ATCAACTACT ACTTCATCAA CAGCCAGGGG CCACCCCATC GAAGGCCTTG CCGKCATGTA 420

CTACATCGCA CACCTGCTGA AGGGCGCCCT CCTCTTCATC ACCATCGCCC TGATTGGCTC 480

AGGCTGGGCT TCATCAAGTA CGTCCTGTCG GATAAGGAGA AGAAGGTCTT TGGGATCGTG 540

ATCCCCATGC AGGTCCTGGC CAACGTGGCC TACATCATCA TCGAGTCCCG CGAGGAAGGC 600

GCCACGAACT ACGTGCTGTG GAAGGAGATT TTGTTCCTGG TGGACCTCAT CTGCTGTGGT 660

GCCATCCTGT TCCCCGTAGT CTGGTCCATC CGGCATCTCC AGGATGCGTC TGGCACAGAC 720

GGGAAGGTGG CAGTGAACCT GGCCAAGCTG AAGCTGTTCC GGCATTACTA TGTCATGGTC 780

ATCTGCTACG TCTACTTCAC CCGCATCATC GCCATCCTGC TGCAGGTGGC TGTGCCCTTT 840

CAGTGGCAGT GGCTGTACMA GCTCTTGGTG GARGGCTCCA CCCTGGCCTT CTTCGTGCTC 900

ACGGGCTACA AGTTCCAGCC CACAGGGAAC AACCCGTACC TGCAGCTGCC CCAGGAGGAC 960

GAGGAGGATG TTCAGATGGA GCAAGTAATG ACGGACTCTG GGTTCCGGGA AGGCCTCTCC 1020

AAAGTCAACA AAACAGCCAG CGGGCGGGAA CTGTTATGAT CACCTCCACA TCTCAGACCA 1080

AAGGGTCGTC CTCCCCCAGC ATTTCTCACT CCTGCCCTTC TTCCACAGCG TATGTGGGGA 1140

GGTGGAGGGG TCCATGTGGA CCAGGCGCCC AGCTCCCCGG ACSCCGGTTC CCGGACAAGC 1200

```
CCATTTGGAA GAAGAGTCCC TTCCTCCCCC CAAATATTGG GCAGCCCTGT CCTTACCCCG    1260

GGACCACCCC TCCCTTCCAG CTATGTGTAC AATAATGACC AATCTGTTTG GCTAAAAAAA    1320

AAAAAAAAAA AACTCGA                                                   1337
```

(2) INFORMACIÓN PARA SEQ ID NO: 20:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1390 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 20:

```
GCCGTTTTGG TTCCCGGTTG GTGCTTCCTG TTCGCAGCTG CGGCACTTCA AGGTTACTGA     60

CTTTTTATGA TGTTTGGTGG CTATGAGACT ATAGAWGCRT RSGRRGATGA TYTTTATCGA    120

GATGAGTCAT CTAGTGAACT GAGTGTTGAT AGTGAGGTGG AATTTCAACT CTATAGCCAA    180

ATTCATTATG CCCAAGATCT TGATGATGTC ATCAGGGAGG AAGAGCATGA AGAAAAGAAC    240

TCTGGGAATT CGGAATCTTC GAGTAGTAAA CCAAATCAGA AGAAGCTAAT CGTCCTTTCA    300

GATAGTGAGG TCATCCAGCT GTCAGATGGG TCAGAGGTCA TCACTTTGTC TGATGAAGAC    360

AGTATTTATA GATGTAAAGG AAAGAATGTT AGAGTTCAAG CACAAGAAAA TGCCCATGGT    420

CTTTCTTCTT CTCTTCAATC TAATGAGCTG GTTGATAAGA AATGCAAGAG TGATATTGAG    480

AAGCCTAAAT CTGAAGAGAG ATCAGGTGTA ATCCGAGAGG TCATGATTAT AGAGGTCAGT    540

TCAAGTGAAG AGGAAGAGAG CACCATTTCA GAAGGTGATA ATGTGGAAAG CTGGATGCTA    600

CTGGGATGTG AAGTAGATGA TAAAGATGAT GATATCCTTC TCAACCTTGT GGGATGTGAA    660

AACTCTGTTA CTGAAGGAGA AGATGGTATA AACTGGTCCA TCAGTGACAA AGACATTGAG    720

GCCCAGATAG CTAATAACCG AACACCTGGA AGATGGACCC AGCGGTACTA TTCAGCCAAC    780

AAAAACATTA TCTGTAGAAA TTGTGACAAA CGTGGTCATT TATCAAAAAA CTGCCCCTTA    840

CCACGAAAAG TTCGTCGCTG CTTCCTGTGC TCCAGGAGAG GACATCTCCT GTATTCCTGT    900

CCAGCCCCCC TTTGCGAATA CTGTCCTGTG CCTAAGATGT TGGACCACTC ATGTCTTTTC    960

AGACATTCCT GGGATAAACA GTGTGACCGA TGTCATATGC TAGGCCACTA TACAGATGCT   1020

TGCACAGAAA TCTGGAGGCA GTATCACCTA ACGACCAAAC CTGGACCACC CAAAAAGCCG   1080

AAGACCCCTT CAAGACCATC AGCCTTAGCA TATTGCTATC ACTGCGCGCA AAAAGGCCAT   1140

TATGGACACG AATGTCCAGA AAGAGAAGTG TATGACCCGT CTCCAGTATC TCCATTCATC   1200
```

```
TGCTACTATG RTGACAAATA TGAAATTCAG GAGAGAGAAA AGAGACTAAA ACAAAAAATA    1260

AAAGTANTCA AGAAAAATGG GGTTATCCCA GAGCCATCCA AGCTACCTTA TATAAAAGCA    1320

GCAAATGAGA ACCCCCACCA TGATATAAGG AAGGGCCGTG CCTCATGGAA AAGCAACAGG    1380

TGGCCTCAAG                                                           1390
```

(2) INFORMACIÓN PARA SEQ ID NO: 21:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1431 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 21:

```
GCCTGCAGTC GACACTAGTG GATCCAAAGA ATTCGGCCTG TGCGAGTAGG CGCTTGGGCA     60

CTCAGTCTCC CTGGCGAGCG ACGGGCAGAA ATCTCGAACC AGTGGAGCGC ACTCGTAACC    120

TGGATCCCAG AAGGTCGCGA AGGCAGTACC GTTTCCTCAG CGGCGGACTG CTGCAGTAAG    180

AATGTCTTTT CCACCTCATT TGAATCGCCC TCCCATGGGA ATCCCAGCAC TCCCACCAGG    240

GATCCCACCC CCGCAGTTTC CAGGATTTCC TCCACCTGTA CCTCCAGGGA CCCCAATGAT    300

TCCTGTACCA ATGAGCATTA TGGCTCCTGC TCCAACTGTC TTAGTACCCA CTGTGTCTAT    360

GGTTGGAAAG CATTTGGGCG CAAGAAAGGA TCATCCAGGC TTAAAGGCTA AAGAAAATGA    420

TGAAAATTGT GGTCCTACTA CCACTGTTTT TGTTGGCAAC ATTTCCGAGA AAGCTTCAGA    480

CATGCTTATA AGACAACTCT TAGCTAAATG TGGTTTGGTT TTGAGCTGGA AGAGAGTACA    540

AGGTGCTTCC GGAAAGCTTC AACCCTTCGG ATTCTGTGAG TACAAGGAGC CAGAATCTAC    600

CCTCCGTGCA CTCAGATTAT TACATGACCT GCAAATTGGA GAGAAAAAGC TACTCGTTAA    660

AGTTGATGCA AAGACAAAGG CACAGCTGGA TGAATGGAAA GCAAAGAAGA AAGCTTCTAA    720

TGGGAATGCA AGGCCAGAAA CTGTCACTAA TGACGATGAA GAAGCCTTGG ATGAAGAAAC    780

AAAGAGGAGA GATCAGATGA TTAAAGGGGC TATTGAAGTT TTAATTCGTG AATACTCCAG    840

TGAGCTAAAT GCCCCCTCAC AGGAATCTGA TTCTCACCCC AGGAAGAAGA AGAAGGAAAA    900

GAAGGAGGAC ATTTTCCGCA GATTTCCAGT GGCCCCACTG ATCCCTTATC CACTCATCAC    960

TAAGGAGGAT ATAAATGCTA TAGAAATGGA AGAAGACAAA AGAGACCTGA TATCTCGAGA   1020

GATCAGCAAA TTCAGAGACA CACATAAGAA ACTGGAAGAA GAGAAAGGCA AAAAGGAAAA   1080

AGAAAGACAG GAAATTGAGA AAGAACGGAG AGAAAGAGAG AGGGAGCGTG AAAGGGAACG   1140

AGAAAGGCGA GAACGGGAAC GAGAAAGGGA AAGAGAACGT GAACGAGAAA AGGAGAAAGA   1200
```

```
ACGGGAGCGG GAACGAGAAC GGGATAGGGA CCGTGACCGG ACAAAAGAGA GAGACCGAGA    1260

TCGGGATCGA GAGAGAGATC GTGACCGGGA TAGAGAAAGG AGCTCAGATC GTAATAAGGA    1320

TCGCATTCGA TCAAGAGAAA AAAGCAGAGA TCGTGAAAGG GAACGAGAGC GGGAAAGAGA    1380

GAGAGAGAGA GAACGAGAGC GAGAACGAGA ACGGGAGCGA GAGAGAGAAG C             1431
```

(2) INFORMACIÓN PARA SEQ ID NO: 22:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 2539 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 22:

```
GGGTGCAGGA GTGCCACCCC CAGGGCCCTG TCAACCTCTC TTTTCTCCTC CATGGCTGTC      60

TGCCTGCGTA TCTGTCTCTG AGAATCCTCG GGGCGGTCAG GGGATGTCAG GAGGGGAAGG     120

AGCCGCCCTC CCTATCTTGC TGCTCCTCTT GGCACTCAGG GGCACCTTCC ATGGAGCCAG     180

ACCGGGTGGA GGGGCTTCTG GGATTTGGTG TCTGCTGCTG CCAGAGCAGG AACCCCCAGT     240

CTAGGACTTG GGCATTTTAA CAGGGAGAAA GTAGTGGCTT CCCTTTTCTC TCTCTCCTCC     300

TTTTTCCCTT TAAGCCCACA GATTCAGGTC ATGCCAAAAG CTCTCTGGTT GTAACCTGGA     360

GACATGTGGA GGGGAATGGC GATGGGATTA TAGGACTCTC CCCATCTCGG GCCCTGACCC     420

TGACCCTTGC CACCAACCCA AAGACAGCTG GTGGGTTTCC CCTTGGAGAM AATCCTGCGT     480

TTGCCTGGGC CGGCCCTGGC TGCCCTCAGC TTTCGCTGAT CTGCCCGGCC TGGAGCCTCC     540

CATCACCCCG CTTCTTGTTG GGCCTCAGGC ACTGGTTACC AGAAGGGGGT CTGGGTCTGC     600

TCAGGAATCA TGTTTTGTAG CACCTCCTGT TGGAGGGGTG GAGGGATGTT CCCCTGAGCC     660

AGGCTGAGAC TAGAACCCCA TCTTCCCTGA GCCAGGCTGA GACTAGAACC CCATCTTCCC     720

CACCACGCCA CCCCTGTGST KGCTACAGGA GCACAGTAGT GAAGGCCTGA GCTCCAGGTT     780

TGAAAGACCC AACTGGAGCG TGGGGCGGGC AGGCAGGGGT TAGTGAAAGG ACACTTCCAG     840

GGTTAGGACA GAGCATTTAG CCTTCTGGAA GAACCCCTGC CTGGGGTGGG ACTGTGCAGG     900

CCAGAGAAGG TGGCATGGGC CTGAACCCAC CTGGACTGAC TTCTGCACTG AAGCCACAGA     960

TGGAGGGTAG GCTGGTGGGT GGGGGTGGTT CGTTCTCTAG CCGGGGCAGA CACCCAGCTG    1020

GCTGGGTCCT TCCTCAGCCT TGCCTCCTCC TGTCCCCAAC CCTTTCCTTT CCTCCTGCTT    1080

GCGGACTGCT GGTCCCCTCT CCTTCCCTCC TTCCAGCTGT TTCTAGTTAC CACCTACCCC    1140
```

```
TGGGCCGTGG ACTGATCAGA CCAGCATTCA AAATAAAAGT TTGTTCCAAG TTGACAGTGT    1200

GGTGCTCCCT GCCCAGCCCC TCCAGGTGGA GGTGCTGCCA CGGGAACGCA GTTGCTCTGC    1260

CTGCCCTGGG CCCCTGGCGA CANTGGGAGC AGGGCAGTGC TGTGAGGAGC CCAGCTTTCC    1320

CAGTCAGGCA GGCATGGCTT CCGTGTTCAG GCTCCCTCAC CAGCTGGTGA CACGGGACAA    1380

GCTTACAAAC CTTCTCTGAA CCTCAGTTTT CTCATTTACA AGAGGCAAAG CATCCATCAC    1440

CTTGTGTGGA TTCARAGAAT GTRAGGCCCT GGGGTGTCCT ACACAAGGGA AAGGCTTGCT    1500

CAGTGAGCGG TCTGCACACC GTTAGCCACC CTGCCACCTC TGTGCCCTGG GCAGGCTCCA    1560

AAGGAAAGCT CTGGCTGGGA CTGCCRGGAG TCTCACACGC TCCTGTTGAC ATTCCCAGCA    1620

GCYGCCCCTG AGGTCGATGT TTGTTCTGTT TTTCTTTTTC TTTTTTGAGA CGGAGTCTCG    1680

CTGTGTTGCC AGGCTGGAGT GCAGTGGTGT GATCTCTGCT CACTGCAACC TCCGCCTGCC    1740

AGTTTCAAGT GATTCTCTGC CTCAGCCTTC TGAGTAGCTG GGACTACAGG TGCACGCCAC    1800

CACGCCCAGC TAACTTTTTG TATTTWAGTA GAGACAGGGT TTCGCCATGT CGGCCAGGGT    1860

GGTCTTGATC TCCTGACCTC ATGATCCACC CGCCTCAGCC TCCCAAAGTG CTGGGATTAC    1920

AGGTATGAGC CACCGCACCG GGCCTGTTCT ATTTTTCTAG TTAAGGGAAC TGAAGCTCAG    1980

ARAGGTGTCA CCAGCARGTG TTCATTCCCA TGCCAGCCTT GCCCCCCGGC TTTTCCCAGG    2040

CAGGCTCCTG CGTGCCCACT GGCTCCAGCC TGGTCCTCTG TCTCTTGGCT GCTTCACTCC    2100

TGCTCTTTGT CCCGACTCTG GCCCTGCTTA CAGGGGCCAC TACCTGCTGG TGCCTCCATA    2160

ACAAGCGTCT GGCGTTGAGA CCCCTGGCAT GGCAGGGGCT TTGGGGTCTG GTTTCCACAA    2220

GGCTTAGCCA TGGCAGAACC TCGTTTTATT TTAACTCTTT GCCCCTACAA ACAAACAGCA    2280

GTACTTGCCA GAACCATTCT TGGGATTCAG GAGCTCGGGC GACTGCCTTG GCCTCTGGCC    2340

GCACCCAGGA GGGTGGGGTT GGATCTGTGT AGTTGCCAGG CCCACACCTG CCAGCAGGGG    2400

GCTGACTGGA TCCATGCTTT ACTGTGTTTA ATGGGGGTAA CAGGGGTCCC TACAGCCCTC    2460

CCAGYTAAAM ATTTGGAACA AAACACCAGC CCTTTTGTAG TGGATGCAGA ATAAAATTGT    2520

TAATCCAATC AAAAAAAAA                                                 2539
```

(2) INFORMACIÓN PARA SEQ ID NO: 23:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1041 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 23:

```
TCGACCCACG CGTCCGCCCA CGCGTCCGCC CACGCGTCCG GGCGCAGGAC GTGCACTATG      60
GCTCGGGGCT CGCTGCGCCG GTTGCTGCGG CTCCTCGTGC TGGGGCTCTG GCTGGCGTTG     120
CTGCGCTCCG TGGCCGGGGA GCAAGCGCCA GGCACCGCCC CCTGCTCCCG CGGCAGCTCC     180
TGGAGCGCGG ACCTGGACAA GTGCATGGAC TGCGCGTCTT GCAGGGCGCG ACCGCACAGC     240
GACTTCTGCC TGGGCTGCGC TGCAGCACCT CCTGCCCCCT TCCGGCTGCT TTGGCCCATC     300
CTTGGGGGCG CTCTGAGCCT GACCTTCGTG CTGGGGCTGC TTTCTGGCTT TTTGGTCTGG     360
AGACGATGCC GCAGAGAGAG AAGTTCACCA CCCCCATAGA GGAGACCGGC GGAGAGGGCT     420
GCCCAGCTGT GGCGCTGATC CAGTGACAAT GTGCCCCCTG CCAGCCGGGG CTCGCCCACT     480
CATCATTCAT TCATCCATTC TAGAGCCAGT CTCTGCCTCC CAGACGCGGC GGGAGCAAGC     540
TCCTCCAACC ACAAGGGGGG TGGGGGGCGG TGAATCACCT CYGAGGCCTG GGCCCAGGGT     600
TCAGGGGAAC TTCCAAGGTG TCTGGTTGCC CTGCCTCTGG CTCCAGAACA GAAAGGGAGC     660
CTCACGCTGG CTCACACAAA ACAGCTGACA CTGACTAAGG AACTGCAGCA TTTGCACAGG     720
GGAGGGGGGT GCCCTCCTTC CTAGAGGCCC TGGGGGCCAG GCTGACTTGG GGGGCAGACT     780
TGACACTAGG CCCCACTCAC TCAGATGTCC TGAAATTCCA CCACGGGGGT CACCCTGGGG     840
GGTTAGGGAC CTATTTTTAA CACTAGGGGG CTGGCCCACT AGGAGGGCTG GCCCTAAGAT     900
ACAGACCCCC CCAACTCCCC AAAGCGGGGA GGAGATATTT ATTTTGGGGA GAGTTTGGAG     960
GGGAGGGAGA ATTTATTAAT AAAAGAATCT TTAACTTTAA AAAAAAAAAA AAAAAAGGGC    1020
GGCCGCTCTA GAGGATCCCT C                                              1041
```

(2) INFORMACIÓN PARA SEQ ID NO: 24:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1962 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 24:

```
ACCCACGCGT CCGGTACAAA ACACAGTTTT ATTCTATGAA AATTTTGAGA TTATTAGAAA      60
CATTAGATTT AGGGTTGCAT ATTAAAAACT ATATCCATTT TGCCTTATTA TTTAGTGTCT     120
CACTCAGGAT ATAACACACT ATAATAGAAA ATGTAGACTT CAGAATCAGG TATATTTGAG     180
ATGGTTTGTA TACTGGTTCT GACACTTGTT AGCTATTCAT CTTTGGTAAA TTCCCCATTA     240
CCCTTTGTKC ACCTATWTGT GGGGATCAGT GCATAGTGTG TGTWAAGCAT TTAATACCTG     300
```

GCAAGTGTTC AGCAAATTTT TTGTTCTATA TATTTATTAT TTGATTATTG GCCCTGAGGA 360

GTAGGTGTTT GTTTGTTTGT TTGTTTGTTT AGTTTTATTT CTCATCTCCT CAGGAACACA 420

AATGAAACTT GGATATTGTT ATGGTGCTTT TNATAATATA TTTATTATTT TCAGCAATTN 480

ATTCTTGTTA AAACAATTTC TTATGACAAG TTACTCATCT TCAATGGTGA GAAGAAATCT 540

AGCTCAGAAT AATATATTTT TAGTGTTTGT ATCTCTGGAT ACTCATTTTG CTCATTGCCA 600

CGTAAAGTAA AAAAATACAT AAATTAGCTT ATTCCAATGT AATATCTTCA GGATAGTCAT 660

GGGCAAGGAA TTAATCACAT TAAGAGATAA CTGCAACTAA GCACTATTTG AGGTGACTTC 720

TGTGGAAAAA AAATTAATYC TTTACCATTG CAGCGTTCTG CCCTAGGTCC AAATGTTACC 780

AAAATCACTC TAGAATCTTT TCTTGCCTGG AAGAAAAGGA AAAGACAAGA AAAGATTGAT 840

AAACTTGAAC AAGATATGGA AAGAAGGAAA GCTGACTTCA AAGCAGGGAA AGCACTAGTG 900

ATCAGTGGTC GTGAAGTGTT TGAATTTCGT CCTGAACTGG TCAATGATGA TGATGAGGAA 960

GCAGATGATA CCCGCTACAC CCAGGGAACA GGTGGTGATG AGGTTGATGA TTCAGTGAGT 1020

GTAAATGACA TAGATTTAAG CCTGTACATC CCAAGAGATG TAGATGAAAC AGGTATTACT 1080

GTAGCCAGTC TTGAAAGATT CAGCACATAT ACTTCAGATA AAGATGAAAA CAAATTAAGT 1140

GAAGCTTCTG GAGGTAGGGC TGAAAATGGT GAAAGAAGTG ACTTGGAAGA GGACAACGAG 1200

AGGGAGGGAA CGGAAAATGG AGCCATTGAT GCTGTTCCTG TTGATGAAAA TCTTTTCACT 1260

GGAGAGGATT TGGATGAACT AGAAGAAGAA TTAAATACAC TTGATTTAGA AGAATGACAC 1320

CAAACACATC GCTGAAAAAA TTAAGTCAGC TCAGCACGAG TTGAAATTGA CTACATTAAT 1380

TTCTTTCCAC CTAGAATCAA CAGGATGTTT ATTTCCTATG CTGATTCTGG AGGAGTTAAC 1440

CTCCTGCAAA AAAGGCATCT TGTCCCTACA TCTTCTCTTC TGACTTTGGC TACATCTCAT 1500

AGTAAGTTCA GAGTAGTTCA TGATAAATTG AAAATATAAT GGTCATTGCA GAAAATGATT 1560

GATGTTGTAA CTGTCCACCC AAGTAAGAAG TGTATCTGCC TTTCCATCTT TTGGTTTTCA 1620

TTTGGGCATG TGCTATTACC AGAAACAACA AACTTATATT TAAAATACCC TTCATTTGAC 1680

ACAGTTTTTA ATGAGTGATT TAATTTCCTC TGTATTTGTA TGTTTAGAAG ACTGCCTAAA 1740

ACATGAGCAC TGTACTTCAT AAAGGAAACG CGTATGCAGA TTCAGTATTG TGTATCTTTG 1800

GACAATTAGA TGGACATTTA AAATGGAACT TCTTTTATCT GACAGGATCA GCTACAATGC 1860

CCTGTGTTAA ATTGTTTAAA AGTTTCCCTT TTCTTTTTTG CCAATAAAGT TGTAAATAAA 1920

GACCATCATA CATTAAAATC CAAAAAAAAA AAAAAAAAAA AA 1962

(2) INFORMACIÓN PARA SEQ ID NO: 25:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1228 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 25:

```
GGCTGCCCAG GCCCCGCACT GGAAGAGCCT CCAGCAGCAA GATGTGACCG YTGTGCCGAT      60
GAGCCCCAGC AGCCACTCCC CAGAGGGGAG GCCTCCACCT CTGCTGCCTG GGGGTCCAGT     120
GTGTAAGGCA GCTGCATCTG CACCGAGCTC CCTCCTGGAC CAGCCGTGCC TCTGCCCCGC     180
ACCCTCTGTC CGCACCGCTG TTGCCCTGAC AACGCCGGAT ATCACATTGG TTCTGCCCCC     240
TGACATCATC CAACAGGAAG CGTCACCCTG AGGGAGGAGA CAGAAGCCTG GGCCAGGTGA     300
ACAGTGGTAT AGCAGCCACT CCAGCCTCTG CTGCAGCAGC CACCCTGGAT GTGGCTGTTC     360
GGAGAGGCCT GTCCCACGGA GCCCAGAGGC TGCTGTGCGT GGCCCTGGGA CAGCTGGACC     420
GGCCTCCAGA CCTCGCCCAT GACGGGAGGA GTCTGTGGCT GAACATCAGG GGCAAGGAGG     480
CGGCTGCCCT ATCCATGTTC CATGTCTCCA CGCCACTGCC AGTGATGACC GGTGGTTTCC     540
TGAGCTGCAT CTTGGGCTTG GTGCTGCCCC TGGCCTATGN TTCCAGCCTG ACCTGGTGCT     600
GGTGGCGCTG GGGCCTGCCA NTGCCTGCAG GGCCCCCACG CTGCACTCCT GGCTGCAATG     660
CTTCGGGGGC TGGCAGGGGG CCGAGTCCTG GCCCTCCTGG AGGAGAACTC CACACCCCAG     720
CTAGCAGGGA TCCTGGCCCG GGTGCTGAAT GGAGAGGCAC CTCCTAGCCT AGGCCCTTCC     780
TCTGTGGCCT CCCCAGAGGA CGTCCAGGCC CTGATGTACC TGAGAGGGCA GCTGGAGCCT     840
CAGTGGAAGA TGTTGCAGTG CCATCCTCAC CTGGTGGCTT GAAATCGGCC AAGGTGGGAG     900
CATTTACACC GCAGAAATGA CACCGCACGC CAGCGCCCCG CGGCCGCGAT CCGGACCCCA     960
AGCCCACGGC TCCCTCGACT CTGGGGCACG GAACCCCGCC CACTCCCAAT CCCCGCGCCC    1020
CGCCCTCTCC CACCCGTGCT TCCCCCGCTC CACCCCTCAC CTCACCTCGC CCCSGCCCCA    1080
CCCATCGCGC CCCGGCCCGT CCCATCGAGG CCCATGCAAC CCACGCTCGG TYCCGTTCCG    1140
GCCCCTGCGC TCKCGCTKNS TTCGCTCCCC GCCCTTGCGC CGTTAGTAAA CATCGCTCAA    1200
ACGAAAAAAA AAAAAAAAAA AAACTCGA                                       1228
```

(2) INFORMACIÓN PARA SEQ ID NO: 26:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1340 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 26:

```
AATTCGGCAG AGAGATGGCC GCCCCCGTGG ATCTAGAGCT GAAGAAGGCC TTCACAGAGC      60
TTCAAGCCAA AGTTATTGAC ACTCAACAGA AGGTGAAGCT CGCAGACATA CAGATTGAAC     120
AGCTAAACAG AACGAAAAAG CATGCACATC TTACAGATAC AGAGATCATG ACTTTGGTAG     180
ATGAGACTAA CATGTATGAA GGTGTAGGAA GAATGTTTAT TCTTCAGTCC AAGGAAGCAA     240
TTCACAGTCA GCTGTTAGAG AAGCAGAAAA TAGCAGAAGA AAAAATTAAA GAACTAGAAC     300
AGAAAAAGTC CTACCTGGAG CGACGTTAAA GGAAGCTGAG GACAACATCC GGGAGATGCT     360
GATGGCACGA AGGGCCCAGT AGGGAGCCTC TCTGGGAAGC TCTTCCTCCT GCCCCTCCCA     420
TTCCTGGTGG GGGCAGAGGA GTGTCTGCAG GGAAACAGCT TCTCCTCTGC CCCGATGGAT     480
GCTTTATTTG GATGGCCTGG CAACATCACA TTTTCTGCAT CACCCTGAGC CCCATTTGCT     540
TCCCAGCCCT GGAGTTTTTA CCCGGCTTTG CTGCCACCTC TGCCCAGGAC ACKCTTCCCT     600
CTCGGGATGT GTGATGAACT CCCAGGAGAG GGAAGATGGG AGCCAGGGCA AGATAGGAAG     660
CTCTGCCTGA GCTTTCCACT AGGCACGCCA GCCAGACCAA TAAAAAGCGT CTGTCCCACT     720
CTGCTAAGCC TGGTTTTCTT GAGCAGAGGG ATGGAACAGA GGGTGAGAGA GGCAGTGGCC     780
GTCTCCACCT CAGCTCCTGC TCCCTCTGCA TCAGAGCCCT TCCTTTCTTG GGGGATGGGC     840
CTTGCCNTCT TCTCTTTTCC CTTCCTGTAC CTTTGACTAA CGCTCAGCTT CCGGGCCTGC     900
ATGCAGTAGA CAGAAGAGGA AGAAAGAACA GATGTTCACA GCTGAATCTC AGTGAACAGA     960
ATAGCAGTCC CTGGATGGCA GTCTGCCTAA AGATTCCTTT CCCTGCCTTC TCCCATACAT    1020
TCCAAAAGGA AGTTCAACAG TAAGCAGCAC CTCCAAGACT GTCTCCTTTY GGCCARTATC    1080
ATAAGATGGA CGCCATAATC CTGAGGCCTC CTAGAGGCTG AGGGGGCAAC GGTGTGATCC    1140
AGCTGGCTCA TCCCAGCCAG GTGGGCCAAT TATTCAATTT TCAAGAATTT TGTTGCAAGC    1200
CAGTTGTCAA ACACAGCCAT TATAATTATG TAAATTTGCA AATTATGTTA AAAACAAGGA    1260
CAATAAATAT TCAAAATGCA TCCCTAAWWA AAAAAAAAAA AANGGGNGGC CGCNCTAGGG    1320
GATCCAAGCT TACGTACGCG                                                1340
```

(2) INFORMACIÓN PARA SEQ ID NO: 27:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 806 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 27:

```
ACCTTCTTCC ATGTTTAGTC CCTTGGGCTC TGCTACCCTC CTGCTGGAGG TGAGAGCATC      60
CTGTGTGCAA CCAGAGATGC CCTCTGGCTT TCAGACCTGC CTGCTTTTCA CCCTCAGCCC     120
TTTCTCACTC AGCAAAATTG TGGGGGTCCC TAGTCAGCAG CTCCCTGGGC AGCTCTCTGA     180
GCAAGGTGGT CTCTGTGGTC ATGAAGGAGA GCCGGCTAGG ACAGTGCCGG AAACTCAGCT     240
GCCTCTCCCC TTCAACTCAG CTGGCCCCCC GCACCTGAAG TGCACAGGAG CCGGGAAGAG     300
AGTCTGGAGC CCACCCCGGA GGGCAGCACA GGAGGTGTCT CTGCAGCTGG TGTCCTGCCA     360
CCCCTGCAGG CAGCACACGT CCCGGGCATT CTCCTTAGCC ACAGACAGAA CAGCCAGTGC     420
CAGAGTCTGC TGTCGTTCCC CTTTAAGCAC ACTCATTCAC CACACCCGAG GAGGCCAGAG     480
GTGCAGGGAG CATGGGCTGT CGCTTCCCCT TTAAGCACAC TCATTCACCA CACCCGAGGA     540
GGCCAGAAGT GCAGGGAGCA TGGGCTGGGT GCACCTCCGC AGGAGAGAAG GCTGAGCCAC     600
CGCCGTCCCG GGAGCCCGGC TCCCAGGCCT CTCGTTTTCC CCTACCTCCC TAAGACTTTT     660
CTGTCACTCT CTGGCCATTG AAAGGCTTCT GTTCCTTAAA GTGCTGTTAC ACTCTCCTTT     720
CCCAGGATGC AGCAAGCCAA AACAGTACCA CTGCACGTCA GCCTGGGTGA CAGAGTGAGA     780
CCCTATCTTA AAAAAAAAAA AAAAAA                                          806
```

(2) INFORMACIÓN PARA SEQ ID NO: 28:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 696 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 28:

```
GAGTTCCCNA CGCGGTGGCG NCCGTTTTAG AAATTAGTGG ATCCCCCCGG GCTGGCAGGG      60
AATTCGGCAC GAGCACAGAG GAAAGCGGGT GCCCGGCATG GCCATCCTGA TGTTGCTGGC     120
GGGATCCCCA TGCACCTTGT CCTTCTCCAC TGATACTGGC AGCTCGGCTC CTGGACCCAA     180
GATCCCTTGA GTGGAATTCT GCAGTGCAAG AGCCCTTCGT GGGAGCTGTC CCATGTTTCC     240
ATGGTCCCCA GTCTCCCCTC CACTTGGTGG GGTCACCAAC TACTCACCAG AAGGGGGCTT     300
ACCAAGAAAG CCCTAAAAAG CTGTTGACTT ATCTGCGCTT GTTCCAACTC TTATGCCCCC     360
AACCTGCCCT ACCACCACCA CGCGCTCAGC CTGATGTGTT TACATGGTAC TGTATGTATG     420
GGAGAGCAGA CTGCACCCTC CAGCAACAAC AGATGAAAGC CAGTGAGCCT ACTAACCGTG     480
```

```
CCATCTTGCA AACTACACTT TAAAAAAAAC TCATTGCTTT GTATTGTAGT AACCAATATG      540
TGCAGTATAC GTTGAATGTA TATGAACATA CTTTCCTATT TCTGTTCTTT GAAAATGTCA      600
GAAATATTTT TTTCTTTCTC ATTTTATGTT GAACTAAAAA GGATTAAAAA AAAAATCTCC      660
AGAMAAAAAA AAAAAAAAAA AAATTACTGC GGTCCG                                696
```

(2) INFORMACIÓN PARA SEQ ID NO: 29:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 1007 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: doble
- (D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 29:

```
AATTCGGCAC GAGGAAAAAA TACCATTTGT GTATGATACC CAATTTGGAT CTCAATTTGG      60
ATAGAGATTT GGTGCTTCCA GATGTRAGTT ATCAGGTGGA ATCCAGTGAG GAGGATCAGT     120
CTCAGACTAT GGATCCTCAA GGACAAACTC TGCTGCTTTT TCTCTTTGTG GATTTCCACA     180
GTGCATTTCC AGTCCAGCAA ATGGAAATCT GGGGAGTCTA TACTTTGCTC ACAACTCATC     240
TCAATGCCAT CCTTGTGGAG AGCCACAGTG TAGTGCAAGG TTCCATCCAA TTCACTGTGG     300
ACAAGGTCTT GGAGCAACAT CACCAGGCTG CCAAGGCTCA GCAGAAACTA CAGGCCTCAC     360
TCTCAGTGGC TGTGAACTCC ATCATGAGTA TTCTGACTGG AAGCACTAGG AGCAGCTTCC     420
GAAAGATGTG TCTCCAGACC CTTCAAGCAG CTGACACACA AGAGTTCAGG ACCAAACTGC     480
ACAAAGTATT TCGTGAGATC ACCCAACACC AATTTCTTCA CCACTGCTCA TGTGAGGTGA     540
AGCAGCTAAC CCTAGAAAAA AAGGACTCAG CCCAGGGCAC TGAGGACGCA CCTGATAACA     600
GCAGCCTGGA GCTCCTAGCA GATACCAGCG GGCAAGCAGA AAACAAGAGG CTCAAGAGGG     660
GCAGCCCCCG CATAGAGGAG ATGCGAGCTC TGCGCTCTGC CAGGGCCCCG AGCCCGTCAG     720
AGGCCGCCCC GCGCCGCCCG GAAGCCACCG CGGCCCCCCT CACTCCTAGA GGAAGGGAGC     780
ACCGCGAGGC TCACGGCAGG GCCCTGGCGC CGGCCAGGGC GAGCCTCCGA AGCCGCCTGG     840
AGGACGTGCT GTGGCTGCAG GAGGTCTCCA ACCTGTCAGA GTGGCTGAGT CCCAGCCCTG     900
GGCCCTGAGC CGGGTCCCCT TNCGCAAGCG CCCACCGATC CGGARGCTGC GGGCAGCCGT     960
TATCCCGTGG TTTAATAAAG TGCCGCGCGC TCACCAAAAA AAAAAAA                   1007
```

(2) INFORMACIÓN PARA SEQ ID NO: 30:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 2017 pares de bases
- (B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 30:

```
AATTCGGCAC GAGCGGATCC GTTGCGGCTG CAGCTCTGCA GTCGGGCCGT TCCTTCGCCG      60
CCGCCAGGGG TAGCGGTGTA GCTGCGCACG TCGCGCGCGC TACCGCACCC AGGTTCGGCC     120
CGTAGCGTCT GGCAGCCCGG CGCCATCTTC ATCGAGCGCC ATGGCCGCAG CCTGCGGGCC     180
GGGAGCGGCG GGTACTGCTT GCTCCTCGGC TTGCATTTGT TTCTGCTGAC CGCGGGCCCT     240
GCCTGGGCTG GAACGACCCT GACAGAATGT TGCTGCGGGA TGTAAAAGCT CTTACCCTCC     300
ACTATGACCG CTATACCACC TCCCGCAGCT GGATCCCATC CCACAGTTGA AATGTGTTGG     360
AGGCACAGCT GGTTGTGATT CTTATACCCC AAAAGTCATA CAGTGTCAGA ACAAAGGCTG     420
GGATGGGTAT GATGTACAGT GGGAATGTAA GACGGACTTA GATATTGCAT ACAAATTTGG     480
AAAAACTGTG GTGAGCTGTG AAGGCTATGA GTCCTCTGAA GACCAGTATG TACTAAGAGG     540
TTCTTGTGGC TTGGAGTATA ATTTAGATTA TACAGAACTT GGCCTGCAGA AACTGAAGGA     600
GTCTGGAAAG CAGCACGGCT TTGCCTCTTT CTCTGATTAT TATTATAAGT GGTCCTCGGC     660
GGATTCCTGT AACATGAGTG GATTGATTAC CATCGTGGTA CTCCTTGGGA TCGCCTTTGT     720
AGTCTATAAG CTGTTCCTGA GTGACGGGCA GTATTCTCCT CCACCGTACT CTGAGTATCC     780
TCCATTTTCC CACCGTTACC AGAGATTCAC CAACTCAGCA GGACCTCCTC CCCCAGGCTT     840
TAAGTCTGAG TTCACAGGAC CACAGAATAC TGGCCATGGT GCAACTTCTG GTTTTGGCAG     900
TGCTTTTACA GGACAACAAG GATATGAAAA TTCAGGACCA GGGTTCTGGA CAGGCTTGGG     960
AACTGGTGGA ATACTAGGAT ATTTGTTTGG CAGCAATAGA GCGGCAACAC CCTTCTCAGA    1020
CTCGTGGTAC TACCCGTCCT ATCCTCCCTC CTACCCTGGC ACGTGGAATA GGGCTTACTC    1080
ACCCCTTCAT GGAGGCTCGG GCAGCTATTC GGTATGTTCA AACTCAGACA CGAAAACCAG    1140
AACTGCATCA GGATATGGTG GTACCAGGAG ACGATAAAGT AGAAAGTTGG AGTCAAACAC    1200
TGGATGCAGA AATTTTGGAT TTTTCATCAC TTTCTCTTTA GAAAAAAAGT ACTACCTGTT    1260
AACAATTGGG AAAAGGGGAT ATTCAAAAGT TCTGTGGTGT TATGTCCAGT GTAGCTTTTT    1320
GTATTCTATT ATTTGAGGCT AAAAGTTGAT GTGTGACAAA ATACTTATGT GTTGTATGTC    1380
AGTGTAACAT GCAGATGTAT ATTGCAGTTT TTGAAAGTGA TCATTACTGT GGAATGCTAA    1440
AAATACATTA ATTTCTAAAA CCTGTGATGC CCTAAGAAGC ATTAAGAATG AAGGTGTTGT    1500
ACTAATAGAA ACTAAGTACA GAAAATTTCA GTTTTAGGTG GTTGTAGCTG ATGAGTTATT    1560
```

```
ACCTCATAGA GACTATAATA TTCTATTTGG TATTATATTA TTTGATGTTT GCTGTTCTTC    1620

AAACATTTAA ATCAAGCTTT GGACTAATTA TGCTAATTTG TGAGTTCTGA TCACTTTTGA    1680

GCTCTGAAGC TTTGAATCAT TCAGTGGTGG AGATGGCCTT CTGGTAACTG AATATTACCT    1740

TCTGTAGGAA AAGGTGGAAA ATAAGCATCT AGAAGGTTGT TGTGAATGAC TCTGTGCTGG    1800

CAAAAATGCT TGAAACCTCT ATATTTCTTT CGTTCATAAG AGGTAAAGGT CAAATTTTTC    1860

AACAAAAGTC TTTTAATAAC AAAAGCATGC AGTTCTCTGT GAAATCTCAA ATATTGTTGT    1920

AATAGTCTGT TTCAATCTTA AAAAGAATCA ATAAAAACAA ACAAGGGAAA AAAAAAAAAA    1980

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA                             2017
```

(2) INFORMACIÓN PARA SEQ ID NO: 31:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 699 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 31:

```
GNGTTTTTTC CAGCCAGGAA GTGACCGNTA CTGCAGCACG AGANAGATTG GTTGGGTTGG     60

TTGRAAATGA CYCTGAACAT TTATTTCCAT TGCAATTTCT GTGGCTGAGG AGACTTAAAC    120

TTTACAAGTA TTATCCTTTT AAGATCATTT TAATTTTAGT TGAGTGCAGA GGGCTTTTAT    180

AACAAACGTG CAGAAATTTT GGAGGGCTGT GATTTTTCCA GTATTAAACA TGCATGCATT    240

AATCTTGCAG TTTATTTTCT CATTGTGTAT GTATATATCG CTTTTCTCTG CAGCACGATT    300

TCTCTTTTGA TAAWKCCCTT TAGGGCACAA CTAGTTATCA GTAACTGAAT GTATCTTAAT    360

CATTATGGCT GCTTCTGTTT TTTCATTAAC AAAGGTTATT CATATGTTAG CATATAGTTT    420

CTTTGCACCC ACTATTTATG TCTGAATCAT TTGTCACAAG AGAGTGTGTG CTGATGAGAT    480

TGTAAGTTTG TGTGTTTAAA CTTTTTTTTG AGCGAGGGAA GAAAAAGCTG TATGCATTTC    540

ATTGCTGTCT ACAGGTTTCT TTCAGATTAT GTTCATGGGT TTGTGTGTAT ACAATATGAA    600

GAATGATCTG AAGTAATTGT GCTGTATTTA TGTTTATTCA CCAGTCTTTG ATTAAATAAA    660

AAGGAAAACC AGAAAAAAAA AAAAAAAAAA AAAAAAAAA                           699
```

(2) INFORMACIÓN PARA SEQ ID NO: 32:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1264 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 32:

```
GGCACGAGGG CACTGTTTCC TCAGTCCATG GCTGAGTACA TCACCGGTGT TTTCTCTCTT      60
ATTCCTCCCA TCAAGCCTAA AAGGAATCTC TATTGGAGAT ACTGCCATTA GTGTTCCTTT     120
TATAGGTGAG GAACTGAGGC ATAKAGGGTT CCCCAGTTGA ACCAACTGAT AAATAGTAGA     180
ACTTGGATTT TAATTCAGTC TTGATGCCAG GGATAAGGCT CTTACTTTCT ACCTTAGGCT     240
ATTTCTAGGA AACGCAGGAG AGTGTTGAAG GGGCAGAGAA AGGGATCCAG TTCCTTTCTG     300
TCCCGCATCC TAGTCCCTGA GAAGCAAAGA ARAATGTGTG GCTTCTTTTG CTTTGCTTTT     360
GTTGTCATCC CACACATCTC CAGGGGAMCT GGGCTCTTGA TCTTGGSCTC TTCCCCTTTA     420
ACTGTTAAGT GGGAGCARGT AAGGGGGTAC AGTAGGGCTG GCCTGGAGTT AGAGGCTTGG     480
ATGCCTTAGC TCCTCTGTCT GCACTCCAGA ACTGCCTGAC TTCATTTCGT ATGTTGTCCT     540
TTGTTTTGAC AATTGATCCA TGTCCCAGTC CGTCTCTTCT TCCTTCTTGA TACTTACACT     600
GCTTCTTTCT GTTGGTTTCC AGTGTTTAAC ACTGTATACA ACAGTGACGA CAACGTGTTT     660
GTGGGGGCCC CCACGGGCAG CGGGAAGACT ATTTGTGCAG AGTTTGCCAT CCTGCGAATG     720
CTGCTGCAGA GCTCGGAGGG GCGCTGTGYS TWCWTCACCM CCATGGAGGC CCTGGCCAGA     780
RCAGGTATGA CGTGGCGCTG TGTCATGTGA ATTTCCCAAG AAGCATTTCA TCTGTGATTC     840
CGTATGAAGG CTTTCTAAGC CCTGAAATTT GCAGGGTCAT TTCCTCAGTT TGTGTATTAA     900
AGAAAAGCTG CCCCAGCCAA GCGTGGTGGC TCACGCCTGT AATCCCAGCA CTTTGGGAGG     960
CCGAGGCGGG CAGATCTCCG GAGATCAGGA GTTCGAGACC AGCCTGGCCA ACATGGTGRA    1020
ACCCTGTCTC TACTAAAAWT ACAGAAATTA GCTGGGNGTG GTGGTGTGCG CCTGTAATCC    1080
CAGCTACTTG GAAGGCTGAG GCAGGAGAAT CGCTTGAACC CGGGAGGCGG AGGTTGCAGT    1140
GAGCCAAGTT CGCACCACTG CACTCCAGCC TGGGCAACAA GAGCGAGACT TCATCTCAAA    1200
AAAAAAAAAA AAAAACTCGA GGGGGGGCCC GGTACCCAAT TCGCCCTATA GTGATCGTAT    1260
TACA                                                                1264
```

(2) INFORMACIÓN PARA SEQ ID NO: 33:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 997 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 33:

```
ATTGGAAGTT GTTTTGCAAC CTGGGCTTTT ATACAGAAGA ATACGAATCA CAGGTGTGTG   60
AGCATCTACT TAATTAATTT GCTTACAGCC GATTTCCTGC TTACTCTGGC ATTACCAGTG  120
AAAATTGTTG TTGACTTGGG TGTGGCACCT TGGAAGCTGA AGATATTCCA CTGCCAAGTA  180
ACAGCCTGCC TCATCTATAT CAATATGTAT TTATCAATTA TCTTCTTAGC ATTTGTCAGC  240
ATTGACCGCT GTCTTCAGCT GACACACAGC TGCAAGATCT ACCGAATACA AGAACCCGGA  300
TTTGCCAAAA TGATATCAAC CGTTGTGTGG CTAATGGTCC TTCTTATAAT GGTGCCAAAT  360
ATGATGATTC CCATCAAAGA CATCAAGGAA AAGTCAAATG TGGGTTGTAT GGAGTTTAAA  420
AAGGAATTTG GAAGAAATTG GCATTTGCTG ACAAATTTCA TATGTGTAGC AATATTTTTA  480
AATTTCTCAG CCATCATTTT AATATCCAAT TGCCTTGTAA TTCGACAGCT CTACAGAAAC  540
AAAGATAATG AAAATTACCC AAATGTGAAA AAGGCTCTCA TCAACATACT TTTAGTGACC  600
ACGGGCTACA TCATATGCTT TGTTCCTTAC CACATTGTCC GAATCCCGTA TACCCTCAGC  660
CAGACAGAAG TCATAACTGA TTGCTCAACC AGGATTTCAC TCTTCAAAGC CAAAGAGGCT  720
ACACTGCTCC TGGCTGTGTC GAACCTGTGC TTTGATCCTA TCCTGTACTA TCACCTCTCA  780
AAAGCATTCC GCTCAAAGGT CACTGAGACT TTTGCCTCMC CTAAAGAGAC CAAGGTYAGA  840
AAGAAAAATT AAGANGTGGA AATAATGGCT AAAAGACAGG NTTTTTGTGG TACCAATTCT  900
GGGCTTTATG GGACCNTAAA GTTATTATAG CTTGGAAGGT AAAAAAAAAA AAAGGGNGGG  960
CGCTCTAGAG GTTCCCCGAG GGGCCAGCTT AGGGTGC                            997
```

(2) INFORMACIÓN PARA SEQ ID NO: 34:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1914 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 34:

```
GTGTGAGAGG CCTCTCTGGA AGTTGTCCCG GGTGTTCGCC GCTGGAGCCC GGGTCGAGAG   60
GACGAGGTGC CGCTGCCTGG AGAATCCTCC GCTGCCGTCG GCTCCCGGAG CCCAGCCCTT  120
TCCTAACCCA ACCCAACCTA GCCCAGTCCC AGCCGCCAGC GCCTGTCCCT GTCACGGACC  180
CCAGCGTTAC CATGCATCCT GCCGTCTTCC TATCCTTACC CGACCTCAGA TGCTCCCTTC  240
TGCTCCTGGT AACTTGGGTT TTTACTCCTG TAACAACTGA AATAACAAGT CTTGATACAG  300
```

AGAATATAGA TGAAATTTTA AACAATGCTG ATGTTGCTTT AGTAAATTTT TATGCTGACT 360

GGTGTCGTTT CAGTCAGATG TTGCATCCAA TTTTTGAGGA AGCTTCCGAT GTCATTAAGG 420

AAGAATTTCC AAATGAAAAT CAAGTAGTGT TTGCCAGAGT TGATTGTGAT CAGCACTCTG 480

ACATAGCCCA GAGATACAGG ATAAGCAAAT ACCCAACCCT CAAATTGTTT CGTAATGGGA 540

TGATGATGAA GAGAGAATAC AGGGGTCAGC GATCAGTGAA AGCATTGGCA GATTACATCA 600

GGCAACAAAA AAGTGACCCC ATTCAAGAAA TTCGGGACTT AGCAGAAATC ACCACTCTTG 660

ATCGCAGCAA AAGAAATATC ATTGGATATT TTGAGCAAAA GGACTCGGAC AACTATAGAG 720

TTTTTGAACG AGTAGCGAAT ATTTTGCATG ATGACTGTGC CTTTCTTTCT GCATTTGGGG 780

ATGTTTCAAA ACCGGAAAGA TATAGTGGCG ACAACATAAT CTACAAACCA CCAGGGCATT 840

CTGCTCCGGA TATGGTGTAC TTGGGAGCTA TGACAAATTT TGATGTGACT TACAATTGGA 900

TTCAAGATAA ATGTGTTCCT CTTGTCCGAG AAATAACATT TGAAAATGGA GAGGAATTGA 960

CAGAAGAAGG ACTGCCTTTT CTCATACTCT TTCACATGAA AGAAGATACA GAAAGTTTAG 1020

AAATATTCCA GAATGAAGTA GCTCGGCAAT TAATAAGTGA AAAAGGTACA ATAAACTTTT 1080

TACATGCCGA TTGTGACAAA TTTAGACATC CTCTTCTGCA CATACAGAAA ACTCCAGCAG 1140

ATTGTCCTGT AATCGCTATT GACAGCTTTA GGCATATGTA TGTGTTTGGA GACTTCAAAG 1200

ATGTATTAAT TCCTGGAAAA CTCAAGCAAT TCGTATTTGA CTTACATTCT GGAAAACTGC 1260

ACAGAGAATT CCATCATGGA CCTGACCCAA CTGATACAGC CCCAGGAGAG CAAGCCCAAG 1320

ATGTAGCAAG CAGTCCACCT GAGAGCTCCT TCCAGAAACT AGCACCCAGT GAATATAGGT 1380

ATACTCTATT GAGGGATCGA GATGAGCTTT AAAAACTTGA AAAACAGTTT GTAAGCCTTT 1440

CAACAGCAGC ATCAACCTAC GTGGTGGAAA TAGTAAACCT ATATTTTCAT AATTCTATGT 1500

GTATTTTTAT TTTGAATAAA CAGAAAGAAA TTTTGGGTTT TTAATTTTTT TCTCCCCGAC 1560

TCAAAATGCA TTGTCATTTA ATATAGTAGC CTCTTAAAAA AAAAAAAAAC CTGCTAGGAT 1620

TTAAAAATAA AAATCAGAGG CCTATCTCCA CTTTAAATCT GTCCTGTAAA AGTTTTATAA 1680

ATCAAATGAA AGGTGACATT GCCAGAAACT TACCATTAAC TTGCACTACT AGGGTAGGGA 1740

GGACTTAGGG ATGTTTCCTG TGTCGTATGT GCTTTTCTTT CTTTCATATG ATCAATTCTG 1800

TTGGTATTTT CAGTATCTCA TTTCTCAAAG CTAAAGAGAT ATACATTCTG GATACTTGGG 1860

AGGGGAATAA ATTAAAGTTT TCACACTGNA AAAAAAAAAA AAAAAAAAAC TCGA 1914

(2) INFORMACIÓN PARA SEQ ID NO: 35:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1020 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 35:

```
CCNTTNNTTT TTTTTTTTTG CAAGACAAAA TATACTTTAT TGTGACAGCA AATGCACATA      60
GTGCTGTAGG TAAGGCATGC TACTAGGAAT CTGCATATAA TCAAAAGCCA GTATGGAAAT     120
GAATGGAAAT GAATGCTGTT GTTCTCAGAT TGAGTCCATG GTGGAGAAAG GATAGTTTGT     180
GTCCACTTAT TTCAAATGCA GTATCATACC TACTTAATCA GTTACCTATG CTTCTAACCA     240
ACAGCCCAGT GGCAAATAGG AGGAACTTAA CTGTACTCAG AAGTCACTTT TAATATCAAC     300
GACAGAAATA TTTCACTAAT TCAACTGAGG CAAATTTCCT TTCTAGACAA AGGACCTAGA     360
AATTGAGCAT GCAAAACATC CATCCATTCA TTCATTCAAA TAATTAGCCA ATTTTACCGT     420
CATTTAATTC CACCAGAAGC AAATACTAGA ATATCTAGAA GTAGTTTGGG TAAAGAAACA     480
TTTACATTTT AATATTGTGT AATGTCATAA ATTTGGGGCT AAAATAACAC CAGGTCAAAT     540
TTGATCCCTT TGTATGTGAG GGTACAAAGT ACAGTTTTCG TTTCAACAGC TGAACTTCTG     600
AGAGAAGAGC TGAAAAAAAT GCTAAATAAG AGATCTAGGC CTTTGATGGA AACTATTAGG     660
CTCTACAGAC TTGTCAAAAA ATCAATGCAA AACTGAGGGG GAAAGGCTGA AATGCTTTGT     720
AAAGCAGTAT TTTTAGACAA GTTGCTTCAT TTCCCCCTTT TCTAAAACAG ATGCAGATTA     780
AATGTTTTTT TGCATGAATG CACATTGACA TTCTGTTCAA CTGTTTTCTA AATGCAACAC     840
TGCGGGTTTC AACAGTATGC TTTCATTTAA ACAAAGAATA TTATATGCAT GGTCAATTTA     900
GTTTAAGAGA TGAAAAAAAA CTTTACTACT ATGAAAATTG CTTATCAAAT ACTCTCCTCT     960
TTTATAAGGT GTTTTTARGC AACACAGGAC CGGTNGAACC GANCAAATTT ATAATTATAC    1020
```

(2) INFORMACIÓN PARA SEQ ID NO: 36:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 781 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 36:

```
AACTCCTGAC CTCAAGTGCT CCACCTGCGT TGGCTTCCCA AAGTGCTGGG ATACAGGAGT      60
RAGCCACTGC GCCTGGCTGA TCCCAGCACT TTTMAAATGA TGCCGCTCAA AGCCGTGACT     120
TGGCCTACTT TGAACAGCAA ACTTGTTGCT GCTGTTGTCA ACCTGAAGGC CTCTCAAATG     180
```

```
CCAGCTTCAA GCAGGGTGTG AATTGGCCAG TGTCAGATCT CAGGAGTCCT GTGTTGAGAG      240

TGTGGCTTTC AGCTGCGGGG AGCTGCACTT GGTGGGGAAA GCCAGGCAGG TCACCCTCAC      300

AGCCAGATAA TGTGGAGGTC AGAACCCAAG GAAGGGAGTG AGACCTCCAC TCCCAGTGGG      360

GGACCTGGCC ACCCATCCTT GGGGACCTGA GAAAGCGTAC TTCACCTTGG GGTGAAGGCT      420

GGGTGGGGCC AGAGGGACCA GTGCCCTCCT CAGTGCTTAG GGGCAGAGCC ACCTGCAGCA      480

ATGGTATCTG CATATTAGCC CCTCTCCACC TTCTTTCTCC CGCTGAATCA TTTCCCTCAA      540

AGCCCAAGAG CTGTCACTGC TTCTTTCTCC CTGGGAAGAA TGCGTGGACT CTGCCTGGTG      600

ATAGACTGAA GCCAGAACAG TGCCACACCC TCGCCTTAAT TCCTTGCTAG GTGTTCTCAG      660

ATTTATGAGA CTTCTTAGTC AAATATGAGG GAGGTTGGAT GTGGTGGCTT GTGCCTGTAA      720

TCCCAGCATT TTGGGAAGCC GAGGTGGGAG GATCCCTTGA AGCCAGGAGT TTGAGACAAG      780

C                                                                      781
```

(2) INFORMACIÓN PARA SEQ ID NO: 37:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 966 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 37:

```
GGCACGAGGA AGCAGCTGGG GGCTGATCAG GGGGAGCACG CAGCCCTCCG ATTGCAGGGC       60

TGCCTATTTG AGTGGCAGCT CCTCTTGAAA CAATGCAGAA CAAGCCCAGG GCCCCACAGA      120

AAAGGGCACT GCCCTTCCCA GAACTTGAGC TCCGGGACTA CGCATCTGTT CTCACCAGAT      180

ACAGCTTGGG GCTGAGGAAC AAAGAGCCTT CCCTGGGCCA CAGGTGGGGG ACCCAGAAGC      240

TGGGCAGGAG CCCCTGTTCT GAAGGGTCCC AGGGCCACAC CACAGATGCT GCTGACGTGC      300

AGAACCACTC TAAAGAAGAA CAGAGAGACG CAGGAGCACA GAGGARGTGC GGCCAGGGGA      360

GGCACACCTG GGCGTACAGG NGAGGGGCGC AGGACACTTC GAGGCTGACA GGAGACCCAC      420

GTGGTGGGGA AAGGAGCCCC CCAAAGTGTC AGAGCATGAA GCAGCAGGAA GGAGCTCCCT      480

CGGGCCACTG CTGGGATCAG TGGTGCCATG GAGCAAGCGA GGTTGTTTGG CCTGAAAGCC      540

GGAAGCGTGC CCAAATCTTT SCATCACCAT GTAGGCAGTC ACCTCGCTCC TCAGCACTCG      600

GGGCAGGACA GAAGCTTGCT GTCTGCTCAC CAGACATCCT GTGCTGCCCT ACAGACACCT      660

TGCTCGCCAG CCATCCCCAC TCACTTCTGA CCGGGACCCA ATTCTCTGGC CAAACCCAGG      720

CTCTAGCACC GTCTTGGTGT GCTTGAGAAA CATCTAGTTT AAGTCAAAAT CCAATGTCTT      780
```

TTTAATATAT AGACTATATG TACCTATGGA CTAGAGGTGA ATATATATAC ATCATATCAA 840

ATTCAAGTGA CCCAGTATTT CGGGAGAACC CACTATGTCC CCAGCCTGCA TGGGAAGCTG 900

GGGATTCTGG CATGAACTGC ACCTTATCTT CCTCGAGGGG GGGCCGGTAC CAATTGCCNA 960

TAGTGG 966

(2) INFORMACIÓN PARA SEQ ID NO: 38:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 416 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 38:

GAATTCGGCA CGAGGTAATA GGAGCCCTCG TACCTCTTGT GTTCCTTACA AACATTCTCA 60

TCAGTAGCTC TACGCGTTGA CTGGGTGGTT TGARATGGCT GGTATACACA GGGCTTTCTT 120

GGTGTTCTGT CTCTGGGGCT TARCTTTGTG TGTGGTTGGA GGGCCCTGGT GAGATTGGAA 180

GTACCAGAGA GTGCTGTGTC AGGGGCAGAG GGGCCTGTCG CTGGAGCTGG AGGGTGCCTG 240

CCTTTGTGTC TGACTCARTC TCCTGTCTGC CTTGCCCCCT CAGGGTCTCG CCAGCCCAGC 300

CTCTGTGGGA ATCTAAAAGG ARTGGATGTG GACGTKTGAC CAAGCACATC TCAGCTTTTA 360

ATACCTGGGC TATTTATAGA CCTTTGGGGG GAATNGCTTG TGGAACAACA AGGGTT 416

(2) INFORMACIÓN PARA SEQ ID NO: 39:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1114 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 39:

TGTGTATTTG GGGGGACTGA AGGGTACGTG GGGCGAAACA AAACCGGCCA TGGCAGCAGC 60

GGAGGAGGAG GACGGGGGCC CCGAAGCCAA AATCGCGAGC GGGGCGGGGC GGGCGCGACC 120

TTCGAATGTA ATATATGTTT GGAGACTGCT CGGGAAGCTG TGGTCAGTGT GTGTGGCCAC 180

CTGTACTGTT GGCCATGTCT TCATCAGTGG CTGGAGACAC GGCCAGAACG GCAAGAGTGT 240

CCAGTATGTA AAGCTGGGAT CAGCAGAGAG AAGGTTGTCC CGCTTTATGG GCGAGGGAGC 300

CAGAAGCCCC AGGATCCCAG ATTAAAAACT CCACCCCGCC CCCAGGGCCA GAGACCAGCT 360

CCGGAGAGCA GAGGGGGATT CCAGCCATTT GGTGATACCG GGGGCTTCCA CTTCTCATTT 420

GGTGTTGGTG CTTTTCCCTT TGGCTTTTTC ACCACCGTCT TCAATGCCCA TGAGCCTTTC 480

CGCCGGGGTA CAGGTGTGGA TCTGGGACAG GGTCACCCAG CCTCCAGCTG GCAGGATTCC 540

CTCTTCCTGT TTCTCGCCAT CTTCTTCTTT TTTTGGCTGC TCAGTATTTG AGCTATGTCT 600

GCTTCCTGCC CACCTCCAGC CAGAGAAGAA TCAGTATTGA GGGTCCCTGC TGACCCTTCC 660

GTACTCCTGG ACCCCCTTGA CCCCTCTATT TCTGTTGGCT AAGGCCAGCC CTGGACATTG 720

TCCAGGAAGG CCTGGGGAGG AGGAGTGAAG TCTGTGCATA GATGGGAGAG CCTTCTGCTC 780

AGAGGCTCAC TCAGTAACGT TGTTTAATTC TCTGCCCTGG GGAAGGAGGA TGGATTGAGA 840

GAATGTCTTT CTCCTCTCCT AAGTCTTTGC TTTCCCTGAT TTCTTGATTT GATCTTCAAA 900

GGTGGGCAAA GTTCCCTCTG ACTCTTCCCC CACTCCCCAT CTTACTGATT TAATTTAATT 960

TTTCACTCCC CAGAGTCTAA TATGGATTCT GACTCTTAAG TGCTTCCGCC CCCTCACTAC 1020

CTCCTTTAAT ACAAATTCAA TAAAAAAGGT GAAATATAAA AAAAAAAAAA AAAAAACYCG 1080

GGGGGGGCCC CGGTCCCCAT TCCCTTTGGG GGGT 1114

(2) INFORMACIÓN PARA SEQ ID NO: 40:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 602 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 40:

GGGTCGACCC ACGCGTCCGT CCCAGGCCAC AAGACATTTC CTGCTCGGAA CCTTGTTTAC 60

TAATTGTCTC TGTGGCACAT TTTGTTTCCC GTGCCTTGGG TGTCAAGTTG CAGCTGATAT 120

GAATGAATGC TGTCTGTGTG GAACAAGCGT CGCAATGAGG ACTCTCTACA GGACCCGATA 180

TGGCATCCCT GGATCTATTT GTGATGACTA TATGGCAACT CTTTGCTGTC CTCATTGTAC 240

TCTTTGCCAA ATCAAGAGAG ATATCAACAG AAGGAGAGCC ATGCGTACTT TCTAAAAACT 300

GATGGTGAAA AGCTCTTACC GAAGCAACAA AATTCAGCAG ACACCTCTTC AGCTTGAGTT 360

CTTCACCATC TTTTGCAACT GAAATATGAT GGATATGCTT AAGTACAACT GATGGCATGA 420

AAAAAATCAA ATTTTTGATT TATTATAAAT GAATGTTGTC CCTGAACTTA GCTAAATGGT 480

GCAACTTAGT TTCTCCTTGC TTTCATATTA TCGAATTTCC TGGCTTATAA ACTTTTTAAA 540

TTACATTTGA AATATAAACC AAATGAAATA TTTTACTGAA AAAAAAAAAA AAAAAANCCC 600

CA 602

(2) INFORMACIÓN PARA SEQ ID NO: 41:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 970 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 41:

```
GGCAGAGCTT AGGAGAACAG CTCCCTTTGG ATCCCTNTCA AAGGTGATAC CATTGGCTCC     60
CAGCTTAGAG TAAGAAGCTC TGAGAAGTTG AATGAAGGGT GAGATAGAGA TGCTGAACCC    120
ATTCTTSCAG CTTCTTCTAG. TGTTGTTATT TCCAGAATGG CCAACACCCC TACATTGATA    180
CATAAACACA TTCCAAGGCC TTGTGTAATA CAAAGTTCAC CGTCCTCCTG GAATAGGAGC    240
CCTGGGTTCT AGTTCTCACT CTGCCACTGG GGGAAAATCC AATTAAAGTC TGGTTTAGTC    300
AGCTTGGGTC ACCATAGACT GGGTGGCTTA AACAGCAGAC ATTTATTTCT GGTAGTTTCT    360
GGAGGCTACA AATCTAAGAG CAAGGTGCCA GCATGGTCAC ATTCTGGTGA GGGSCCTCTT    420
CCTGGCTTGT AGACGGCTGC YTTCTCACCG TGTGCTCACA TAGCCTTTCG TGTGTGTGTG    480
TGTGTGTGTG TGCGTKCGTG CAAGCTTCCK GATGTCTCTT CTTAGAAGGA CACCAACCCC    540
ATCATGAGAG CCCTACTCTC ATGACTTAGC CTAACCCTAA TTACCCTCCA AAGGCCCCAT    600
CTCCAAATGC CATCACATTG GAGGGTAGAG CTTCAACATA GGGATTTTGG GGGACACAAA    660
CATTCAGTCC ATAACAAAGG CTGTAGTCCT TARTTTCCTT GTCTGTGAAA TGAGAGTGTT    720
GAGATTCTTT CTAGCCTTTA TCATTTATAA TTCTGTGAGA TGTAGATTTG CATTATTTTC    780
GAGTTCGAGT TATATGAAAT GTTTCCCTCT ACATTTTCTT GGGCAACTGA GAACTGAATA    840
GGGCTAGGTT TAAATAGAGT TAGGCAGTTA GGCTTATTCT TTTATTTAAT AAGCATTTTT    900
GGAGCATCTA CGGTGTTCCA GGAACTGAAC TGTTGTAAAC ATTGGAGCTG TAACAGAGAA    960
CAAAAGAGAC                                                           970
```

(2) INFORMACIÓN PARA SEQ ID NO: 42:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1002 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 42:

```
GAATTCGGCA CGAGCCGAGG TCGGCAGCAC AGAGCTCTGG AGATGAAGAC CCTGTTCCTG      60
GGTGTCACGC TCGGMCTGGC CGCTGCCCTG TCCTTMACCC TGGRGGAGGA GGATATCACA     120
GGGACCTGGT ACGTGAAGGC CATGGTGGTC GATAAGACTT TCCGGAGACA GGAGGCCCAG     180
AAGGTGTCCC CAGTGAAGGT GACAGCCCTG GGCGGTGGGA AGTTGGAAGC CACGTTCACC     240
TTCATGAGGG AGGATCGGTG CATCCAGAAG AAAATCCTGR TGCGGAAGAC GGAGGAGCCT     300
GGCAAATACA GCGCCTGTGA GCCCCTCCCC CAYTCCCACC CCCACCYTCC CCCACCGCCA     360
ACCCCAGTGC ACCAGCCTCC ACAGGTAGAG AGTGCCCAGG CTGCCCTTTT GCCAGGGCCC     420
CAGCTCTGCC CACCTCCAAG GAGGGGCTGG CCTCTCCTTC CTGGGGGGCT GGTGGCCCTG     480
ACATCAGACA CCGGGTGTGA CAGGCTTGTC CGCAGTCGAG ATGGACCAGA TCACGCCTGC     540
CCTCTGGGAG GCCCTAGCCA TTGACACATT GAGGAAGCTG AGGATTCGGA CAAGGAGGCC     600
AAGGATTAGA TGGGGGCAGG AAGCTCATGT ACCTGCAGGA GCTGCCCAGG AGGGACCAYT     660
ACATCTTTTA CTGCAAAGAC CAGCACCATG GGGGCSTGCT CCACATGGGA AAGCTTGTGG     720
GTAGGAATTC TGATACCAAC CGGGAGGCCC TGGAAGAATT TAAGAAATTG GTGCAGCGCA     780
AGGGACTCTC GGAGGAGGAC ATTTTCACGC CCCTGCAGAC GGGAAGCTGC GTTCCCGAAC     840
ACTAGGCAGC CCCCGGGTCT GCACCTCCAG AGCCCACCCT ACCACCAGAC ACAGAGCCCG     900
GACCACCTGG ACCTACCCTC CAGCCATGAC CCTTCCCTGC TCCCACCCAC CTGACTCCAA     960
ATAAAGTCCT TCTCCCCCAA AAAAAAAAAA AAAAAAACTC GA                       1002
```

(2) INFORMACIÓN PARA SEQ ID NO: 43:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 2581 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 43:

```
TGCAAAACCA CTGGACACTG GACAAGTACG GGATCCTGGS CGACGCACGC CTCTTCTTTG      60
GGCCCCAGCA CCGGSCCGTC ATCCTTCGGT TGTCCAACCG CCGCGCACTG CGCCTCCGTG     120
CCAGCTTCTC CCAGCCCCTC TTCCAGGCTG TGGSTGCCAT CTGCCGCCTC CTCAGCATCC     180
GGCACCCCGA GGAGCTGTCC CTGCTCCGGG CTCCTGAGAA GAAGGAGAAG AAGAAGAAAG     240
AGAAGGAGCC AGAGGAAGAG CTCTATGACT TGAGCAAGGT TGTCTTGGCT GGGGGCGTGG     300
CACCTGCACT GTTCCGGGGG ATGCCAGCTC ACTTCTCGGA CAGCGCCCAG ACTGAGGCCT     360
```

GCTACCACAT GCTGAGCCGG CCCCAGCCGC CACCCGACCC CCTCCTGCTC CAGCGTCTGC 420

CACGGCCCAG CTCCCTGTCA GACAAGACCC AGCTCCACAG CAGGTGGCTG GACTCGTCGC 480

GGTGTCTCAT GCAGCAGGGC ATCAAGGCCG GGGACGCACT CTGGCTGCGC TTCAAGTACT 540

ACAGCTTCTT CGATTTGGAT CCCAAGACAG ACCCCGTGCG GCTGACACAG CTGTATGAGC 600

AGGCCCCGTG GCACCTGCTG CTGGAGGAGA TTGACTGCAC CGAGGAGGAG ATGATGGTGT 660

TTGCCGCCCT GCAGTACCAC ATCAACAAGC TGTCCCAGAG CGGGGAGGTG GGGGAGCCGG 720

CTGGCACAGA CCCAGGGCTG GACGACCTGG ATGTGGCCCT GAGCAACCTG GAGGTGAAGC 780

TGGAGGGGTC GGCGCCCACA GATGTGCTGG ACAGCCTCAC CACCATCCCA GAGCTCAAGG 840

ACCATCTCCG AATCTTTCGG CCCCGGAAGC TGACCCTGAA GGGCTACCGC CAACACTGGG 900

TGGTGTTCAA GGAGACCACA CTGTCCTACT ACAAGAGCCA GGACGAGGCC CCTGGGGACC 960

CCATTCAGCA GCTCAACCTC AAGGGCTGTG AGGTGGTTCC CGATGTTAAC GTCTCCGGCC 1020

AGAAGTTCTG CATTAAACTC CTAGTGCCCT CCCCTGAGGC ATGAGTGAGA TCTACCTGCG 1080

GTGCCAGGAT GAGCAGCAGT ATGCCCGCTG GATGGCTGGC TGCCGCCTGG CCTCCAAAGG 1140

CCGCACCATG GCCGACAGCA GCTACACCAG CGAGGTGCAG GCCATCCTGG CYTTCCTCAG 1200

CCTGCAGCGC ACGGGCAGTG GGGGCCCGGG CAACCACCCC CACGGCCCTG ATGCCTCTGC 1260

CGAGGGCCTC AACCCCTACG GCCTCGTTGC CCCCCGTTTC CAGCGAAAGT TCAAGGCCAA 1320

GCAGCTCACC CCACGGATCC TGGAAGCCCA CCAGAATGTG GCCCAGTTGT CGCTGGCAGA 1380

GGCCCAGCTG CGCTTCATCC AGGCCTGGCA GTCCCTGCCC GACTTCGGCA TCTCCTATGT 1440

CATGGTCAGG TTCAAGGGCA GCAGGAAAGA CGAGATCCTG GGCATCGCCA ACAACCGACT 1500

GATCCGCATC GACTTGGCCG TGGGCGACGT GGTCAAGACC TGGCGTTTCA GCAACATGCG 1560

CCAGTGGAAT GTCAACTGGG ACATCCGGCA NGTGGCCATC GAGTTTGATG AACACATCAA 1620

TGTGGCCTTC AGCTGCGTGT CTGCCAGCTG CCGAATTGTA CACGAGTATA TCGGGGGCTA 1680

CATTTTCCTG TCGACGCGGG AGNGGGCCCG TGGGGAGGAG CTGGATGAAG ACCTCTTCCT 1740

GCAGCTCACC GGGGGCCATG AGGCCTTCTG AGCGCTGTCT GATTGCCCCT GCCCTGCTCA 1800

CCACCCTGTC ACAGCCACTC CCAAGCCCAC ACCCACAGGG GCTCACTGCC CCACACCCGC 1860

TCCAGGCAGG CACCCAGCTG GGCATTTCAC CTGCTGTCAC TGACTTTGTG CAGGCCAAGG 1920

ACCTGGCAGG GCCAGACGCT GTACCATCAC CCAGGCCAGG GATGGGGGTG GGGGTCCCTG 1980

AGCTCATGTG GTGCCCCCTT TCCTTGTCTG AGTGGCTGAG GCTGATACCC CTGACCTATC 2040

TGCAGTCCCC CAGCACACAA GGAAGACCAG ATGTAGCTAC AGGATGATGA AACATGGTTT 2100

CAAACGAGTT CTTTCTTGTT ACTTTTTAAA ATTTCTTTTT TATAAATTAA TATTTTATTG 2160

```
TTGGATCCTC CTCCTTTCTC TGGAGCTGTG CTTGGGGCTA CTCTGACACT CTGTCTCTTC    2220

ATCACCAGCC AAGGAAAGGG GCTTTCCTGA TAAAGACAAG AGTTGGTTAG AGAAAGGGAC    2280

ACCTAAGTCA GTCTAGGGTT GGAAGCTAGG AGAGAGGTGA GGGCAGAAGG GCACAGCTTT    2340

CAGGAACAAG GAATAGGGGC TGGGGTKGTK GTTCTCACGG GTAGGCGGTA CCTGCAGGGC    2400

CTCCTTGAAG TACTTGGGAA GGAGGAAGCC ATCAGTATTC CCTGGAGTCA GAATCACCCC    2460

ATTGGCAGAG CGGAAGAAGG GTATTCCATC TGCTGACAGA GCCAGAGATG TGACTCATGC    2520

CCTCCCCGAA GGCAAAGTCA GCTCCTGCTT TGTCCAGACT CACCTGCCAG AGCCAGGGGT    2580

C                                                                    2581
```

(2) INFORMACIÓN PARA SEQ ID NO: 44:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1764 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 44:

```
GAATTCGGCA CGAGGATGAT ATTCCTACTA TTCCTCACCC CACTCTGGCT GCAAAAAGGA     60

AGTGCAGGGA AAATGAGTGG GGAGTTCCTG TATGCCAGTC TGTTTCAATG GAACTATTTT    120

TGGAGGAATA AAAAAGTATG CTAGATTATA TTGGTACGAT AGGCATTTTC TTACATTGCA    180

TATAGTCTGC TTTGGCTTTT ACCTGTTGAG GGGAAGAATG AGGAGAGGAT AAAAATCATT    240

GTATCCCCTA GAGAAGGAAT ATCAAAATCC ATTTAATAAA AAAACTCATA CTAAGAATAA    300

AATTGCATAG TGTTTTATTC TCCTTTGTTC ATAATTAAAC ACAAGATATT TTAAATTGTC    360

AAATCAGTTT CTTTATGAAA AAATATGACC TGTATGCCTT TATTCTCTCC TTTCCTTCTT    420

CCCACCCGTC GCTTCTTTTC TTCTCTTCCT TTTTTTCTTT CCTTGTCCTC TGACTAAATG    480

AAGAACAAAC ATTTGATAAA AGCCACTGCC AATTCATGAT AAAAATTCAC AGCAAAGTTG    540

GTACAGAAAA GAACTTTCTC TGCGTGTTAA AGGGTGCCTC TCCCATGCTC TCAGCAAATA    600

TTTAATGATG AAATCTTATT AATAATCACT GTAGAACCAA GAATTAAACT AGTATACCCA    660

CTGTCTTGGC TTGTAATCAA CAATATACAG GTGGTTCTAG CCAGTGCAAT AAGACAAGAG    720

AAACAAAAAT GTTATAAGGC CTGGAAAAGA TGAAACAAAC TGTTATTCAC AAAATACTGT    780

CTATACAGAA TGCTCAGTGT CTTTTTTTCT TTTCTTTTTT TTAAACTTTA GTGAGATACC    840

CTTCTGCCCT ATCTTAAAAT CACGTGGTGG GGGGTGGTGT CTGCACTTGA AACAGGACAC    900
```

```
TTGGTTCCTG GGTTTAGCAT TGACCTTGCC AGCTTGGTYT GGCAGCTGAG TTGTTGGACT      960
AGGAAGCGTC CYTGCAGGTT GTGKTCTGKT ACCTCTCTGT AAAGCCTGAA AGCATCCTAC     1020
SATTGCATTT GCTAGKTCTC AGTAGAGCTA TTTAACAAGA ATCTGGAAAC ATTTTYCCTG     1080
AGGGCTCTCT TTAGACAGCA GTAAAATGTA GCTGGAGACA TATTGAGTAA ATGGAAAAGA     1140
AAAATCTAAT GAGGCCAGGA ATTTTTTTAA TCTTCTATTC TCACAGAAGG CCTCAAGGAG     1200
AACACCATAA TTCATATTTT ACTCAKGTGG GTTAGGCATA AAGCCTCCCC CATAGATCCA     1260
ATAACCTGTA RGTGTYCTGG TTTTGAAATT GCACCTGCTT ACATKGCTGG ATCNTAGCAC     1320
TAAWTCACAC RGCAACGGCT TCTGGTTCAA TKGTTCATTA CTTGGGAATG TCAGATTGCC     1380
AGAGAGCAGC CTGATGTTTA CATCCAATCG GCAATGCCTT AGGAAATCAG TTTTAATTAC     1440
AATCTCACGT AGCAGCACTG CACTCAACCT TCAGAGAGGC TGGGATTTGT GTTGAACCTA     1500
CATCTTATAG CTGTGCAGAA AATGCCTGTC CGACTGGGTC ATGCAAAATG GACAGCAAAG     1560
TCAGCAGAAC CTTAGAAAAG ATGACACAGC AAGTGGAACA CAGCTGGATC ATCCCCCGTC     1620
CTGTCAAGCG TGCAGTGCTC TCTGGCCCCT TTTTAAAACA AGGGAACCCA GTTGGCGTTT     1680
GCCTTTCAGC TTCCCCATTC TGATATAAAA ATCTGTGACC CAGCAGCTTT AACCATAAAA     1740
AAAAAAAAAA AAAAAAAAAC TCGA                                            1764
```

(2) INFORMACIÓN PARA SEQ ID NO: 45:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 796 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 45:

```
ACCTTCTTCC ATGTTTAGTC CCTTGGGCTC TGCTACCCTC CTGCTGGAGG TGAGAGCATC      60
CTGTGTGCAA CCAGAGATGC CCTCTGGCTT TCAGACCTGC CTGCTTTTCA CCCTCAGCCC     120
TTTCTCACTC AGCAAAATTG TGGGGGTCCC TAGTCAGCAG CTCCCTGGGC AGCTCTCTGA     180
GCAAGGTGGT CTCTGTGGTC ATGAAGGAGA GCCGGCTAGG ACAGTGCCGG AAACTCAGCT     240
GCCTCTCCCC TTCAACTCAG CTGGCCCCCC GCACCTGAAG TGCACAGGAG CCGGGAAGAG     300
AGTCTGGAGC CCACCCCGGA GGGCAGCACA GGAGGTGTCT YTGCAGCTGG TGTCCTGCMA     360
CCCYTGCAGG CAGMACACGT CCCGGGCATT YTCYTTAGCC ACAGACAGAA CAGCCAGTGC     420
CAGAGTCTGC TGTCGYTTCC CCTTTAAGCA CACTCATTCA CCACACCCGA GGAGGCCAGA     480
GGTGCAGGGA GCATGGGCTG TCGTTCCCCT TTAAGCACAC TCATTCACCA CACCCGAGGA     540
```

```
GGCCAGAAGT GCAGGGAGCA TGGGCTGGGT GCACCTCCGC AGGAGAGAAG GCTGAGCCAC      600

CGCCGTCCCG GGAGCCCGGC TCCCAGGCCT CTCGTTTTCC CCTACCTCCC TAAGACTTTT      660

CTGTCACTCT CTGGCCATTG AAAGGCTTCT GTTCCTTAAA GTGCTGTTAC ACTCTCCTTT      720

CCCAGGATGC AGCAAGCCAA AACAGTACCA CTGCACGTCA GCCTGGGTGA CAGAGTGAGA      780

CCCTATCTTA AAAAAA                                                      796
```

(2) INFORMACIÓN PARA SEQ ID NO: 46:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1705 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 46:

```
TGGCCATGGA AGCGCTAGAA GGTTTAGATT TTGAAACAGC AAAGAAGGAT TTCCTTGGAT       60

CTGGAGACCC CAAAGAAACA AAGATGCTAA TCACCAAACA GGCTGACTGG GCCAGAAATA      120

TCAAGGAGCC CAAAGCCGCC GTGGAGATGT ACATCTCAGC AGGAGAGCAC GTCAAGGCCA      180

TCGAGATCTG TGGTGACCAT GGCTGGGTTG ACATGTTGAT CGACATCGCC CGCAAACTGG      240

ACAAGGCTGA GCGCGAGCCC CTGCTGCTGT GCGCTACCTA CCTCAAGAAG CTGGACAGCC      300

CTGGCTATGC TGCTGAGACC TACCTGAAGA TGGGTGACCT CAAGTCCCTG GTGCAGCTGC      360

AGTGGAGACC CAGCGCTGGG ATGAGGCCTT TGCTTTGGGT GAGAAGCATC CTGAGTTTAA      420

GGATGACATC TACATGCCGT ATGCTCAGTG GCTAGCAGAG AACGATCGCT TTGAGGAAGC      480

CCAGAAAGCG TTCCACAAGG CTGGGCGACA GAGAGAAGCG GTCCAGGTGC TGGAGCAGCT      540

CACAAACAAT GCCGTGGCGG AGAGCAGGTT TAATGATGCT GCCTATTATT ACTGGATGCT      600

GTCCATGCAG TGCCTCGATA TAGCTCAAGA TCCTGCCCAG AAGGACACAA TGCTTGGCAA      660

GTTCTACCAC TTCCAGCGTT TGGCAGAGCT GTACCATGGT TACCATGCCA TCCATCGCCA      720

CACGGAAGAT CCGTTCAGTG TCCATCGTCC TGAAACTCTT TTCAACATCT CCAGGTTCCT      780

GCTGCACAGC CTGCCCAAGG ACACCCCCTC GGGCATCTCT AAAGTGAAAA TACTCTTCAC      840

CTTGGCCAAG CAGAGCAAGG CCCTCGGTGC CTACAGGCTG GCCCGGCACG CCTATGACAA      900

GCTGCGTGGC CTGTACATCC CTGCCAGATT CCAAAAGTCC ATTGAGCTGG GTACCCTGAC      960

CATCCGCGCC AAGCCCTTCC ACGACAGTGA GGAGTTGGTG CCCTTGTGCT ACCGCTGCTC     1020

CACCAACAAC CCGCTGCTCA ACAACCTGGG CAACGTCTGC ATCAACTGCC GCCAGCCCTT     1080
```

```
CATCTTCTCC GCCTCTTCCT ACGACGTGCT ACACCTGGTT GAGTTCTACC TGGAGGAAGG    1140

GATCACTGAT GAAGAAGCCA TCTCCCTCAT CGACCTGGAG GTGCTGAGAC CCAAGCGGGA    1200

TGACAGACAG CTAGAGATTT GCAAACAACA GCTCCCAGAT TCTTGCGGCT AGTGGGAGAC    1260

CAAGGGACTC CATCGGAGAT NAGGACCCGT TCACAGCTAA GCTRAGCTTT GAGCAAGGTG    1320

GCTCARAGTT CGTGCCAGTG GTGGTGAGCC GGCTGGTGCT GCGCTCCATG AGCCGCCGGG    1380

ATGTCCTCAT CAAGCGATGG CCCCCACCCC TGAGGTGGCA ATACTTCCGC TCACTGCTGC    1440

CTGACGCCTC CATTACCATG TGCCCCTCCT GCTTCCAGAT GTTCCATTCT GAGGACTATG    1500

AGTTGCTGGT GCTTCAGCAT GGCTGCTGCC CCTACTGCCG CAGGTGCAAG GATGACCCTG    1560

GCCCATGACC AGCATCCTGG GGACGGCCTG CACCCTCTGC CCGCCTTGGG GTCTGCTGGG    1620

CTGTGAAGGA GAATAAAGAG TTAAACTGTC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA    1680

AAAAAAAAAA AAAAAAAAAA AAANA                                          1705
```

(2) INFORMACIÓN PARA SEQ ID NO: 47:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 981 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: doble

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 47:

```
TCGGCAGCAC AGAGCTCTGG AGATGAAGAC CCTGTTCCTG GGTGTCACGC TCGGCCTGGC     60

GCTGCCCTGT CCTTCACCCT GGRGGAGGAG GATATCACAG GGACCTGGTA CGTGAAGGCC    120

ATGGTGGTCG ATAAGACTTT CCGGAGACAG GAGGCCCAGA AGGTGTCCCC AGTGAAGGTG    180

ACAGCCCTGG GCGGTGGGAA GTTGGAAGCC ACGTTCACCT TCATGAGGGA GGATCGGTGC    240

ATCCAGAAGA AAATCCTGRT GCGGAAGACG GAGGAGCCTG GCAAATACAG CGCCTGTGAG    300

CCCCTCCCCC AYTCCCACCC CCACCYTCCC CCACCGCCAA CCCCAGTGCA CCAGCCTCCA    360

CAGGTAGAGA GTGCCCAGGC TGCCCTTTTG CCAGGGCCCC AGCTCTGCCC ACCTCCAAGG    420

AGGGGCTGGC CTCTCCTTCC TGGGGGGCTG GTGGCCCTGA CATCAGACAC CGGGTGTGAC    480

AGGCTTGTCC GCAGTCGAGA TGGACCAGAT CACGCCTGCC CTCTGGGAGG CCCTAGCCAT    540

TGACACATTG AGGAAGCTGA GGATTGGGAC AAGGAGGCCA AGGATTAGAT GGGGGCAGGA    600

AGCTCATGTA CCTGCAGGAG CTGCCCAGGA GGGACCAYTA CATCTTTTAC TGCAAAGACC    660

AGCACCATGG GGGCSTGCTC CACATGGGAA AGCTTGTGGG TAGGAATTCT GATACCAACC    720

GGGAGGCCCT GGAAGAATTT AAGAAATTGG TGCAGCGCAA GGGACTCTCG GAGGAGGACA    780
```

```
TTTTCACGCC CCTGCAGACG GGAAGCTGCR TTCCCGAACA CTAGGCAGCC CCCGGGTCTG      840

CACCTCCAGA GCCCACCCTA CCACCAGACA CAGAGCCCGG ACCACCTGGA CCTACCCTCC      900

AGCCATGACC CTTCCCTGCT CCCACCCACC TGACTCCAAA TAAAGTCCTT CTCCCCCAAA      960

AAAAAAAAAA AAAAAACTCG A                                                981
```

(2) INFORMACIÓN PARA SEQ ID NO: 48:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 146 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 48:

```
Met His Tyr Gln Met Ser Val Thr Leu Lys Tyr Glu Ile Lys Lys Leu
 1               5                  10                  15

Ile Tyr Val His Leu Val Ile Trp Leu Leu Leu Val Ala Lys Met Ser
            20                  25                  30

Val Gly His Leu Arg Leu Leu Ser His Asp Gln Val Ala Met Pro Tyr
        35                  40                  45

Gln Trp Glu Tyr Pro Tyr Leu Leu Ser Ile Leu Pro Ser Leu Leu Gly
    50                  55                  60

Leu Leu Ser Phe Pro Arg Asn Asn Ile Ser Tyr Leu Val Leu Ser Met
65                  70                  75                  80

Ile Ser Met Gly Leu Phe Ser Ile Ala Pro Leu Ile Tyr Gly Ser Met
                85                  90                  95

Glu Met Phe Pro Ala Ala Gln Pro Ser Thr Ala Met Ala Arg Pro Thr
            100                 105                 110

Val Ser Ser Leu Val Phe Leu Pro Phe Pro Ser Cys Thr Trp Cys Trp
        115                 120                 125

Cys Trp Gln Cys Lys Cys Met Pro Gly Ser Cys Thr Thr Ala Arg Ser
    130                 135                 140

Ser Xaa
145
```

(2) INFORMACIÓN PARA SEQ ID NO: 49:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 312 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 49:

```
Met Asn Ser Val Val Ser Leu Leu Leu Ile Leu Glu Pro Asp Lys Gln
 1               5                  10                  15

Glu Ala Leu Ile Glu Ser Leu Cys Glu Lys Leu Val Lys Phe Arg Glu
            20                  25                  30

Gly Glu Arg Pro Ser Leu Arg Leu Gln Leu Leu Ser Asn Leu Phe His
        35                  40                  45

Gly Met Asp Lys Asn Thr Pro Val Arg Tyr Thr Val Tyr Cys Ser Leu
    50                  55                  60

Ile Lys Val Ala Ala Ser Cys Gly Ala Ile Gln Tyr Ile Pro Thr Glu
65                  70                  75                  80

Leu Asp Gln Val Arg Lys Trp Ile Ser Asp Trp Asn Leu Thr Thr Glu
                85                  90                  95

Lys Lys His Thr Leu Leu Arg Leu Leu Tyr Glu Ala Leu Val Asp Cys
            100                 105                 110

Lys Lys Ser Asp Ala Ala Ser Lys Val Met Val Glu Leu Leu Gly Ser
        115                 120                 125

Tyr Thr Glu Asp Asn Ala Ser Gln Ala Arg Val Asp Ala His Arg Cys
    130                 135                 140

Ile Val Arg Ala Leu Lys Asp Pro Asn Ala Phe Leu Phe Asp His Leu
145                 150                 155                 160

Leu Thr Leu Lys Pro Val Lys Phe Leu Glu Gly Glu Leu Ile His Asp
                165                 170                 175

Leu Leu Thr Ile Phe Val Ser Ala Lys Leu Ala Ser Tyr Val Lys Phe
            180                 185                 190

Tyr Gln Asn Asn Lys Asp Phe Ile Asp Ser Leu Gly Leu Leu His Glu
        195                 200                 205

Gln Asn Met Ala Lys Met Arg Leu Leu Thr Phe Met Gly Met Ala Val
    210                 215                 220

Glu Asn Lys Glu Ile Ser Phe Asp Thr Met Gln Gln Glu Leu Gln Ile
225                 230                 235                 240

Gly Ala Asp Asp Val Glu Ala Phe Val Ile Asp Ala Val Arg Thr Lys
                245                 250                 255

Met Val Tyr Cys Lys Ile Asp Gln Thr Gln Arg Lys Val Val Val Ser
            260                 265                 270

His Ser Thr His Arg Thr Phe Gly Lys Gln Gln Trp Gln Gln Leu Tyr
        275                 280                 285

Asp Thr Leu Asn Ala Trp Lys Gln Asn Leu Asn Lys Val Lys Asn Ser
    290                 295                 300

Leu Leu Ser Leu Ser Asp Thr Xaa
305                 310
```

(2) INFORMACIÓN PARA SEQ ID NO: 50:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 47 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 50:

```
Gly Gly Cys Pro Arg Arg Arg Leu Val Leu Tyr Cys Leu Phe Gly Ser
 1               5                  10                  15

Ala Gly Gly Gly Arg Ile His Ser Glu Ala Trp Phe Pro Lys Ala Trp
            20                  25                  30

Pro Glu Ala Glu Lys Trp Leu Phe Ala Glu Leu Leu Arg Gly Xaa
        35                  40                  45
```

(2) INFORMACIÓN PARA SEQ ID NO: 51:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 467 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 51:

```
Met Leu Ser Arg Pro Gln Pro Pro Pro Asp Pro Leu Leu Leu Gln Arg
 1               5                  10                  15

Leu Pro Arg Pro Ser Ser Leu Ser Asp Lys Thr Gln Leu His Ser Arg
            20                  25                  30

Trp Leu Asp Ser Ser Arg Cys Leu Met Gln Gln Gly Ile Lys Ala Gly
        35                  40                  45

Asp Ala Leu Trp Leu Arg Phe Lys Tyr Tyr Ser Phe Phe Asp Leu Asp
    50                  55                  60

Pro Lys Thr Asp Pro Val Arg Leu Thr Gln Leu Tyr Glu Gln Ala Arg
65                  70                  75                  80

Trp Asp Leu Leu Leu Glu Glu Ile Asp Cys Thr Glu Glu Glu Met Met
                85                  90                  95

Val Phe Ala Ala Leu Gln Tyr His Ile Asn Lys Leu Ser Gln Ser Gly
            100                 105                 110

Glu Val Gly Glu Pro Ala Gly Thr Asp Pro Gly Leu Asp Asp Leu Asp
        115                 120                 125

Val Ala Leu Ser Asn Leu Glu Val Lys Leu Glu Gly Ser Ala Pro Thr
    130                 135                 140

Asp Val Leu Asp Ser Leu Thr Thr Ile Pro Glu Leu Lys Asp His Leu
145                 150                 155                 160

Arg Ile Phe Arg Pro Arg Lys Leu Thr Leu Lys Gly Tyr Arg Gln His
```

165 170 175

Trp Val Val Phe Lys Glu Thr Thr Leu Ser Tyr Tyr Lys Ser Gln Asp
180 185 190

Glu Ala Pro Gly Asp Pro Ile Gln Gln Leu Asn Leu Lys Gly Cys Glu
195 200 205

Val Val Pro Asp Val Asn Val Ser Gly Gln Lys Phe Cys Ile Lys Leu
210 215 220

Leu Val Pro Ser Pro Glu Gly Met Ser Glu Ile Tyr Leu Arg Cys Gln
225 230 235 240

Asp Glu Gln Gln Tyr Ala Arg Trp Met Ala Gly Cys Arg Leu Ala Ser
245 250 255

Lys Gly Arg Thr Met Ala Asp Ser Ser Tyr Thr Ser Glu Val Gln Ala
260 265 270

Ile Leu Ala Phe Leu Ser Leu Gln Arg Thr Gly Ser Gly Gly Pro Gly
275 280 285

Asn His Pro His Gly Pro Asp Ala Ser Ala Glu Gly Leu Asn Pro Tyr
290 295 300

Gly Leu Val Ala Pro Arg Phe Gln Arg Lys Phe Lys Ala Lys Gln Leu
305 310 315 320

Thr Pro Arg Ile Leu Glu Ala His Gln Asn Val Ala Gln Leu Ser Leu
325 330 335

Ala Glu Ala Gln Leu Arg Phe Ile Gln Ala Trp Gln Ser Leu Pro Asp
340 345 350

Phe Gly Ile Ser Tyr Val Met Val Arg Phe Lys Gly Ser Arg Lys Asp
355 360 365

Glu Ile Leu Gly Ile Ala Asn Asn Arg Leu Ile Arg Ile Asp Leu Ala
370 375 380

Val Gly Asp Val Val Lys Thr Trp Arg Phe Ser Asn Met Arg Gln Trp
385 390 395 400

Asn Val Asn Trp Asp Ile Arg Gln Val Ala Ile Glu Phe Asp Glu His
405 410 415

Ile Asn Val Ala Phe Ser Cys Val Ser Ala Ser Cys Arg Ile Val His
420 425 430

Glu Tyr Ile Gly Gly Tyr Ile Phe Leu Ser Thr Arg Glu Arg Ala Arg
435 440 445

Gly Glu Glu Leu Asp Glu Asp Leu Phe Leu Gln Leu Thr Gly Gly His
450 455 460

Glu Ala Phe
465

(2) INFORMACIÓN PARA SEQ ID NO: 52:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 83 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 52:

```
Met Arg Pro Gly Arg Gly Ala Gly Thr Pro Gly Arg Pro Gly Arg Gly
 1               5                  10                  15

Arg Gly Leu Ala Ala Thr Cys Ser Leu Ser Ser Pro Ser His Leu Leu
            20                  25                  30

Pro Thr Leu Leu His Thr Phe Ser Phe Ser Leu Pro Pro Pro Ser Pro
        35                  40                  45

Ala Ala Pro Arg Gln Pro Ser Pro Pro Ala Leu Leu Leu Pro Gly Pro
    50                  55                  60

Gln Lys Pro Arg Pro Gly Asp Pro Thr Tyr Thr Gly Ala Leu Thr Asp
65                  70                  75                  80

Trp Ser Xaa
```

(2) INFORMACIÓN PARA SEQ ID NO: 53:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 63 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 53:

```
Met Phe Leu Val Phe Phe Leu Ser Phe Phe Ser His Ser Ile Ser Ala
 1               5                  10                  15

Leu Thr Leu Val Cys Ser Gln Gly Gly Lys Ala Asp Met Asn Leu Leu
            20                  25                  30

Ser Trp Asp Phe Arg Pro His Trp Leu Glu Gly Ile Arg Phe Leu Leu
        35                  40                  45

Gly Trp Gly Gln Ala Leu Met Ala Gly Leu Phe Pro Trp Leu Xaa
    50                  55                  60
```

(2) INFORMACIÓN PARA SEQ ID NO: 54:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 124 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 54:

```
Met Arg Gly Ser Trp His Arg Ser Pro Leu Pro Ala Val Val Leu Pro

 1               5                  10                  15

Ser Val Leu Gln Thr Ala Leu Ser Pro Leu Ala Leu Cys Gln Ala Trp
            20                  25                  30

Arg Arg Ala Val Pro His Gly Val Pro Ser Gln Arg Leu Arg Asn Gln
        35                  40                  45

Glu Ala Ser Leu Val Pro Lys Gly Val Pro Arg Ala Trp Tyr Pro Gly
    50                  55                  60

Pro Leu Gln Asn Gly Leu Trp Thr His Leu Glu Lys Gly Glu Leu Leu
65                  70                  75                  80

Gly Leu Lys Pro Thr Pro Gly Gly Leu Leu Leu Arg Ser Phe Trp
                85                  90                  95

Asp Pro His Pro Ser Arg Pro Phe Leu Cys Thr Leu Leu Pro Pro Pro
            100                 105                 110

Leu Xaa Ile Phe Pro Pro Leu Arg Cys Ser Ala Xaa
        115                 120
```

(2) INFORMACIÓN PARA SEQ ID NO: 55:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 180 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 55:

```
Met Thr Ser Ala Gly Pro Val Xaa Leu Phe Leu Leu Val Ser Ile Ser
 1               5                  10                  15

Thr Ser Val Ile Leu Met Gln His Leu Leu Xaa Ala Ser Tyr Cys Asp
            20                  25                  30

Leu Leu His Lys Ala Ala Ala His Leu Gly Cys Trp Gln Lys Val Asp
        35                  40                  45

Pro Ala Leu Cys Ser Asn Val Leu Gln His Pro Trp Thr Glu Glu Cys
    50                  55                  60

Met Trp Pro Gln Gly Val Leu Val Lys His Ser Lys Asn Val Tyr Lys
65                  70                  75                  80

Ala Val Gly Xaa Xaa Xaa Val Ala Ile Pro Ser Asp Val Ser His Phe
                85                  90                  95

Arg Phe Xaa Phe Phe Phe Ser Lys Pro Leu Arg Ile Leu Asn Ile Leu
            100                 105                 110

Leu Leu Leu Glu Gly Ala Val Ile Val Tyr Gln Leu Tyr Ser Leu Met
        115                 120                 125

Ser Ser Glu Lys Trp His Gln Thr Ile Ser Leu Ala Leu Ile Leu Phe
    130                 135                 140

Ser Asn Tyr Tyr Ala Phe Phe Lys Leu Leu Arg Asp Arg Leu Val Leu
145                 150                 155                 160

Gly Lys Ala Tyr Ser Tyr Ser Ala Ser Pro Gln Arg Asp Leu Asp His
                165                 170                 175

Arg Phe Ser Xaa
            180
```

(2) INFORMACIÓN PARA SEQ ID NO: 56:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 287 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 56:

```
Met Pro Leu Phe Lys Leu Tyr Met Val Met Ser Ala Cys Phe Leu Ala
 1               5                  10                  15

Ala Gly Ile Phe Trp Val Ser Ile Leu Cys Arg Asn Thr Tyr Ser Val
            20                  25                  30

Phe Lys Ile His Trp Leu Met Ala Ala Leu Ala Phe Thr Lys Ser Ile
        35                  40                  45

Ser Leu Leu Phe His Ser Ile Asn Tyr Tyr Phe Ile Asn Ser Gln Gly
    50                  55                  60

Pro Pro His Arg Arg Pro Cys Arg His Val Leu His Arg Thr Pro Ala
65                  70                  75                  80

Glu Gly Arg Pro Pro Leu His His His Arg Pro Asp Trp Leu Arg Leu
                85                  90                  95

Gly Phe Ile Lys Tyr Val Leu Ser Asp Lys Glu Lys Lys Val Phe Gly
            100                 105                 110

Ile Val Ile Pro Met Gln Val Leu Ala Asn Val Ala Tyr Ile Ile Ile
        115                 120                 125

Glu Ser Arg Glu Glu Gly Ala Thr Asn Tyr Val Leu Trp Lys Glu Ile
    130                 135                 140

Leu Phe Leu Val Asp Leu Ile Cys Cys Gly Ala Ile Leu Phe Pro Val
145                 150                 155                 160

Val Trp Ser Ile Arg His Leu Gln Asp Ala Ser Gly Thr Asp Gly Lys
                165                 170                 175

Val Ala Val Asn Leu Ala Lys Leu Lys Leu Phe Arg His Tyr Tyr Val
            180                 185                 190

Met Val Ile Cys Tyr Val Tyr Phe Thr Arg Ile Ile Ala Ile Leu Leu
        195                 200                 205

Gln Val Ala Val Pro Phe Gln Trp Gln Trp Leu Tyr Xaa Leu Leu Val
    210                 215                 220

Glu Gly Ser Thr Leu Ala Phe Phe Val Leu Thr Gly Tyr Lys Phe Gln
225                 230                 235                 240

Pro Thr Gly Asn Asn Pro Tyr Leu Gln Leu Pro Gln Glu Asp Glu Glu
                245                 250                 255

Asp Val Gln Met Glu Gln Val Met Thr Asp Ser Gly Phe Arg Glu Gly
            260                 265                 270

Leu Ser Lys Val Asn Lys Thr Ala Ser Gly Arg Glu Leu Leu Xaa
        275                 280                 285
```

(2) INFORMACIÓN PARA SEQ ID NO: 57:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 34 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 57:

```
Met Pro Met Val Phe Leu Leu Leu Phe Asn Leu Met Ser Trp Leu Ile
 1               5                  10                  15

Arg Asn Ala Arg Val Ile Leu Arg Ser Leu Asn Leu Lys Arg Asp Gln
            20                  25                  30

Val Xaa
```

(2) INFORMACIÓN PARA SEQ ID NO: 58:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 58:

```
Met Lys Ile Val Val Leu Leu Pro Leu Phe Leu Leu Ala Thr Phe Pro
 1               5                  10                  15

Arg Lys Leu Gln Thr Cys Leu Xaa
            20
```

(2) INFORMACIÓN PARA SEQ ID NO: 59:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 47 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 59:

```
Met Ser Gly Gly Glu Gly Ala Ala Leu Pro Ile Leu Leu Leu Leu Leu
 1               5                  10                  15

Ala Leu Arg Gly Thr Phe His Gly Ala Arg Pro Gly Gly Gly Ala Ser
            20                  25                  30

Gly Ile Trp Cys Leu Leu Leu Pro Glu Gln Glu Pro Pro Val Xaa
        35                  40                  45
```

(2) INFORMACIÓN PARA SEQ ID NO: 60:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 114 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 60:

```
Met Ala Arg Gly Ser Leu Arg Arg Leu Leu Arg Leu Leu Val Leu Gly
 1               5                  10                  15

Leu Trp Leu Ala Leu Leu Arg Ser Val Ala Gly Glu Gln Ala Pro Gly
            20                  25                  30

Thr Ala Pro Cys Ser Arg Gly Ser Ser Trp Ser Ala Asp Leu Asp Lys
        35                  40                  45

Cys Met Asp Cys Ala Ser Cys Arg Ala Arg Pro His Ser Asp Phe Cys
    50                  55                  60

Leu Gly Cys Ala Ala Ala Pro Pro Ala Pro Phe Arg Leu Leu Trp Pro
65                  70                  75                  80

Ile Leu Gly Gly Ala Leu Ser Leu Thr Phe Val Leu Gly Leu Leu Ser
                85                  90                  95

Gly Phe Leu Val Trp Arg Arg Cys Arg Arg Glu Arg Ser Ser Pro Pro
            100                 105                 110

Pro Xaa
```

(2) INFORMACIÓN PARA SEQ ID NO: 61:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 32 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 61:

```
Met Val Cys Ile Leu Val Leu Thr Leu Val Ser Tyr Ser Ser Leu Val
 1               5                  10                  15

Asn Ser Pro Leu Pro Phe Val His Leu Xaa Val Gly Ile Ser Ala Xaa
            20                  25                  30
```

(2) INFORMACIÓN PARA SEQ ID NO: 62:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 81 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 62:

```
Met Thr Gly Gly Phe Leu Ser Cys Ile Leu Gly Leu Val Leu Pro Leu
 1               5                  10                  15

Ala Tyr Xaa Ser Ser Leu Thr Trp Cys Trp Trp Arg Trp Gly Leu Pro
            20                  25                  30

Xaa Pro Ala Gly Pro Pro Arg Cys Thr Pro Gly Cys Asn Ala Ser Gly
        35                  40                  45

Ala Gly Arg Gly Pro Ser Pro Gly Pro Pro Gly Gly Glu Leu His Thr
    50                  55                  60

Pro Ala Ser Arg Asp Pro Gly Pro Gly Ala Glu Trp Arg Gly Thr Ser
65                  70                  75                  80

Xaa
```

(2) INFORMACIÓN PARA SEQ ID NO: 63:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 104 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 63:

```
Met Ala Ala Pro Val Asp Leu Glu Leu Lys Lys Ala Phe Thr Glu Leu
 1               5                  10                  15

Gln Ala Lys Val Ile Asp Thr Gln Gln Lys Val Lys Leu Ala Asp Ile
            20                  25                  30

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

Thr Glu Ile Met Thr Leu Val Asp Glu Thr Asn Met Tyr Glu Gly Val
    50                  55                  60

Gly Arg Met Phe Ile Leu Gln Ser Lys Glu Ala Ile His Ser Gln Leu
65                  70                  75                  80

Leu Glu Lys Gln Lys Ile Ala Glu Glu Lys Ile Lys Glu Leu Glu Gln
                85                  90                  95

Lys Lys Ser Tyr Leu Glu Arg Arg
            100
```

(2) INFORMACIÓN PARA SEQ ID NO: 64:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 146 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 64:

```
Met Pro Ser Gly Phe Gln Thr Cys Leu Leu Phe Thr Leu Ser Pro Phe
 1               5                  10                  15

Ser Leu Ser Lys Ile Val Gly Val Pro Ser Gln Gln Leu Pro Gly Gln
            20                  25                  30

Leu Ser Glu Gln Gly Gly Leu Cys Gly His Glu Gly Glu Pro Ala Arg
        35                  40                  45

Thr Val Pro Glu Thr Gln Leu Pro Leu Pro Phe Asn Ser Ala Gly Pro
    50                  55                  60

Pro His Leu Lys Cys Thr Gly Ala Gly Lys Arg Val Trp Ser Pro Pro
65                  70                  75                  80

Arg Arg Ala Ala Gln Glu Val Ser Leu Gln Leu Val Ser Cys His Pro
                85                  90                  95

Cys Arg Gln His Thr Ser Arg Ala Phe Ser Leu Ala Thr Asp Arg Thr
            100                 105                 110

Ala Ser Ala Arg Val Cys Cys Arg Ser Pro Leu Ser Thr Leu Ile His
        115                 120                 125

His Thr Arg Gly Gly Gln Arg Cys Arg Glu His Gly Leu Ser Leu Pro
    130                 135                 140

Leu Xaa
145
```

(2) INFORMACIÓN PARA SEQ ID NO: 65:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 31 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 65:

```
Met Ala Ile Leu Met Leu Leu Ala Gly Ser Pro Cys Thr Leu Ser Phe
 1               5                  10                  15

Ser Thr Asp Thr Gly Ser Ser Ala Pro Gly Pro Lys Ile Pro Xaa
            20                  25                  30
```

(2) INFORMACIÓN PARA SEQ ID NO: 66:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 260 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 66:

```
Met Asp Pro Gln Gly Gln Thr Leu Leu Leu Phe Leu Phe Val Asp Phe
 1               5                  10                  15

His Ser Ala Phe Pro Val Gln Gln Met Glu Ile Trp Gly Val Tyr Thr
            20                  25                  30

Leu Leu Thr Thr His Leu Asn Ala Ile Leu Val Glu Ser His Ser Val
        35                  40                  45

Val Gln Gly Ser Ile Gln Phe Thr Val Asp Lys Val Leu Glu Gln His
    50                  55                  60

His Gln Ala Ala Lys Ala Gln Gln Lys Leu Gln Ala Ser Leu Ser Val
65                  70                  75                  80

Ala Val Asn Ser Ile Met Ser Ile Leu Thr Gly Ser Thr Arg Ser Ser
                85                  90                  95

Phe Arg Lys Met Cys Leu Gln Thr Leu Gln Ala Ala Asp Thr Gln Glu
            100                 105                 110

Phe Arg Thr Lys Leu His Lys Val Phe Arg Glu Ile Thr Gln His Gln
        115                 120                 125

Phe Leu His His Cys Ser Cys Glu Val Lys Gln Leu Thr Leu Glu Lys
    130                 135                 140

Lys Asp Ser Ala Gln Gly Thr Glu Asp Ala Pro Asp Asn Ser Ser Leu
145                 150                 155                 160

Glu Leu Leu Ala Asp Thr Ser Gly Gln Ala Glu Asn Lys Arg Leu Lys
                165                 170                 175

Arg Gly Ser Pro Arg Ile Glu Glu Met Arg Ala Leu Arg Ser Ala Arg
            180                 185                 190

Ala Pro Ser Pro Ser Glu Ala Ala Pro Arg Arg Pro Glu Ala Thr Ala
        195                 200                 205

Ala Pro Leu Thr Pro Arg Gly Arg Glu His Arg Glu Ala His Gly Arg
    210                 215                 220

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

Leu Trp Leu Gln Glu Val Ser Asn Leu Ser Glu Trp Leu Ser Pro Ser
                245                 250                 255

Pro Gly Pro Xaa
            260
```

(2) INFORMACIÓN PARA SEQ ID NO: 67:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 23 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 67:

```
Met Ala Ala Ala Cys Gly Pro Gly Ala Ala Gly Thr Ala Cys Ser Ser
  1               5                  10                  15

Ala Cys Ile Cys Phe Cys Xaa
            20
```

(2) INFORMACIÓN PARA SEQ ID NO: 68:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 68:

```
Met His Ala Leu Ile Leu Gln Phe Ile Phe Ser Leu Cys Met Tyr Ile
  1               5                  10                  15

Ser Leu Phe Ser Ala Ala Arg Phe Leu Phe Xaa
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 69:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 29 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 69:

```
Leu Leu Leu Leu Cys Phe Cys Cys His Pro Thr His Leu Gln Gly Xaa
  1               5                  10                  15

Trp Ala Leu Asp Leu Gly Leu Phe Pro Phe Asn Cys Xaa
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 70:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 216 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 70:

```
Met Tyr Leu Ser Ile Ile Phe Leu Ala Phe Val Ser Ile Asp Arg Cys
  1               5                  10                  15

Leu Gln Leu Thr His Ser Cys Lys Ile Tyr Arg Ile Gln Glu Pro Gly
            20                  25                  30
```

```
Phe Ala Lys Met Ile Ser Thr Val Val Trp Leu Met Val Leu Leu Ile
          35                  40                  45

Met Val Pro Asn Met Met Ile Pro Ile Lys Asp Ile Lys Glu Lys Ser
      50                  55                  60

Asn Val Gly Cys Met Glu Phe Lys Lys Glu Phe Gly Arg Asn Trp His
 65                  70                  75                  80

Leu Leu Thr Asn Phe Ile Cys Val Ala Ile Phe Leu Asn Phe Ser Ala
                 85                  90                  95

Ile Ile Leu Ile Ser Asn Cys Leu Val Ile Arg Gln Leu Tyr Arg Asn
            100                 105                 110

Lys Asp Asn Glu Asn Tyr Pro Asn Val Lys Lys Ala Leu Ile Asn Ile
        115                 120                 125

Leu Leu Val Thr Thr Gly Tyr Ile Ile Cys Phe Val Pro Tyr His Ile
    130                 135                 140

Val Arg Ile Pro Tyr Thr Leu Ser Gln Thr Glu Val Ile Thr Asp Cys
145                 150                 155                 160

Ser Thr Arg Ile Ser Leu Phe Lys Ala Lys Glu Ala Thr Leu Leu Leu
                165                 170                 175

Ala Val Ser Asn Leu Cys Phe Asp Pro Ile Leu Tyr Tyr His Leu Ser
            180                 185                 190

Lys Ala Phe Arg Ser Lys Val Thr Glu Thr Phe Ala Ser Pro Lys Glu
        195                 200                 205

Thr Lys Val Arg Lys Lys Asn Xaa
    210                 215
```

(2) INFORMACIÓN PARA SEQ ID NO: 71:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 407 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 71:

```
Met His Pro Ala Val Phe Leu Ser Leu Pro Asp Leu Arg Cys Ser Leu
 1               5                  10                  15

Leu Leu Leu Val Thr Trp Val Phe Thr Pro Val Thr Thr Glu Ile Thr
            20                  25                  30

Ser Leu Asp Thr Glu Asn Ile Asp Glu Ile Leu Asn Asn Ala Asp Val
        35                  40                  45

Ala Leu Val Asn Phe Tyr Ala Asp Trp Cys Arg Phe Ser Gln Met Leu
    50                  55                  60

His Pro Ile Phe Glu Glu Ala Ser Asp Val Ile Lys Glu Glu Phe Pro
65                  70                  75                  80
```

```
Asn Glu Asn Gln Val Val Phe Ala Arg Val Asp Cys Asp Gln His Ser
                85                  90                  95

Asp Ile Ala Gln Arg Tyr Arg Ile Ser Lys Tyr Pro Thr Leu Lys Leu
            100                 105                 110

Phe Arg Asn Gly Met Met Met Lys Arg Glu Tyr Arg Gly Gln Arg Ser
        115                 120                 125

Val Lys Ala Leu Ala Asp Tyr Ile Arg Gln Gln Lys Ser Asp Pro Ile
    130                 135                 140

Gln Glu Ile Arg Asp Leu Ala Glu Ile Thr Thr Leu Asp Arg Ser Lys
145                 150                 155                 160

Arg Asn Ile Ile Gly Tyr Phe Glu Gln Lys Asp Ser Asp Asn Tyr Arg
                165                 170                 175

Val Phe Glu Arg Val Ala Asn Ile Leu His Asp Asp Cys Ala Phe Leu
            180                 185                 190

Ser Ala Phe Gly Asp Val Ser Lys Pro Glu Arg Tyr Ser Gly Asp Asn
        195                 200                 205

Ile Ile Tyr Lys Pro Pro Gly His Ser Ala Pro Asp Met Val Tyr Leu
    210                 215                 220

Gly Ala Met Thr Asn Phe Asp Val Thr Tyr Asn Trp Ile Gln Asp Lys
225                 230                 235                 240

Cys Val Pro Leu Val Arg Glu Ile Thr Phe Glu Asn Gly Glu Glu Leu
                245                 250                 255

Thr Glu Glu Gly Leu Pro Phe Leu Ile Leu Phe His Met Lys Glu Asp
            260                 265                 270

Thr Glu Ser Leu Glu Ile Phe Gln Asn Glu Val Ala Arg Gln Leu Ile
        275                 280                 285

Ser Glu Lys Gly Thr Ile Asn Phe Leu His Ala Asp Cys Asp Lys Phe
    290                 295                 300

Arg His Pro Leu Leu His Ile Gln Lys Thr Pro Ala Asp Cys Pro Val
305                 310                 315                 320

Ile Ala Ile Asp Ser Phe Arg His Met Tyr Val Phe Gly Asp Phe Lys
                325                 330                 335

Asp Val Leu Ile Pro Gly Lys Leu Lys Gln Phe Val Phe Asp Leu His
            340                 345                 350

Ser Gly Lys Leu His Arg Glu Phe His His Gly Pro Asp Pro Thr Asp
        355                 360                 365

Thr Ala Pro Gly Glu Gln Ala Gln Asp Val Ala Ser Ser Pro Pro Glu
    370                 375                 380

Ser Ser Phe Gln Lys Leu Ala Pro Ser Glu Tyr Arg Tyr Thr Leu Leu
385                 390                 395                 400

Arg Asp Arg Asp Glu Leu Xaa
                405
```

(2) INFORMACIÓN PARA SEQ ID NO: 72:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 9 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 72:

```
Tyr Leu Ile Ser Tyr Leu Cys Phe Xaa
 1               5
```

(2) INFORMACIÓN PARA SEQ ID NO: 73:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 34 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 73:

```
Met Pro Leu Lys Ala Val Thr Trp Pro Thr Leu Asn Ser Lys Leu Val
 1               5                  10                  15

Ala Ala Val Val Asn Leu Lys Ala Ser Gln Met Pro Ala Ser Ser Arg
            20                  25                  30

Val Xaa
```

(2) INFORMACIÓN PARA SEQ ID NO: 74:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 57 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 74:

```
Gln Ser Pro Arg Ser Ser Ala Leu Gly Ala Gly Gln Lys Leu Ala Val
 1               5                  10                  15

Cys Ser Pro Asp Ile Leu Cys Cys Pro Thr Asp Thr Leu Leu Ala Ser
            20                  25                  30

His Pro His Ser Leu Leu Thr Gly Thr Gln Phe Ser Gly Gln Thr Gln
        35                  40                  45

Ala Leu Ala Pro Ser Trp Cys Ala Xaa
    50                  55
```

(2) INFORMACIÓN PARA SEQ ID NO: 75:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 75:

Met Ala Gly Ile His Arg Ala Phe Leu Val Phe Cys Leu Trp Gly Leu
1 5 10 15

Xaa Leu Cys Val Val Gly Gly Pro Trp Xaa
20 25

(2) INFORMACIÓN PARA SEQ ID NO: 76:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 15 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 76:

Met Ser Phe Ser Ser Pro Lys Ser Leu Leu Ser Leu Ile Ser Xaa
1 5 10 15

(2) INFORMACIÓN PARA SEQ ID NO: 77:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 33 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 77:

Met Thr Ile Trp Gln Leu Phe Ala Val Leu Ile Val Leu Phe Ala Lys
1 5 10 15

Ser Arg Glu Ile Ser Thr Glu Gly Glu Pro Cys Val Leu Ser Lys Asn
20 25 30

Xaa

(2) INFORMACIÓN PARA SEQ ID NO: 78:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 23 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 78:

Met Leu Asn Pro Phe Xaa Gln Leu Leu Leu Val Leu Leu Phe Pro Glu
1 5 10 15

Trp Pro Thr Pro Leu His Xaa

20

(2) INFORMACIÓN PARA SEQ ID NO: 79:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 173 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 79:

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1               5                  10                  15

Ser Xaa Thr Leu Xaa Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
            20                  25                  30

Ala Met Val Val Asp Lys Thr Phe Arg Arg Gln Glu Ala Gln Lys Val
        35                  40                  45

Ser Pro Val Lys Val Thr Ala Leu Gly Gly Gly Lys Leu Glu Ala Thr
    50                  55                  60

Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Xaa
65                  70                  75                  80

Arg Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Glu Pro Leu Pro
                85                  90                  95

His Ser His Pro His Xaa Pro Pro Pro Pro Thr Pro Val His Gln Pro
            100                 105                 110

Pro Gln Val Glu Ser Ala Gln Ala Ala Leu Leu Pro Gly Pro Gln Leu
        115                 120                 125

Cys Pro Pro Pro Arg Arg Gly Trp Pro Leu Leu Pro Gly Gly Leu Val
    130                 135                 140

Ala Leu Thr Ser Asp Thr Gly Cys Asp Arg Leu Val Arg Ser Arg Asp
145                 150                 155                 160

Gly Pro Asp His Ala Cys Pro Leu Gly Gly Pro Ser His
                165                 170
```

(2) INFORMACIÓN PARA SEQ ID NO: 80:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 208 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 80:

```
Met Ala Asp Ser Ser Tyr Thr Ser Glu Val Gln Ala Ile Leu Ala Phe
 1               5                  10                  15

Leu Ser Leu Gln Arg Thr Gly Ser Gly Gly Pro Gly Asn His Pro His
            20                  25                  30
```

```
Gly Pro Asp Ala Ser Ala Glu Gly Leu Asn Pro Tyr Gly Leu Val Ala
         35                  40                  45

Pro Arg Phe Gln Arg Lys Phe Lys Ala Lys Gln Leu Thr Pro Arg Ile
     50                  55                  60

Leu Glu Ala His Gln Asn Val Ala Gln Leu Ser Leu Ala Glu Ala Gln
 65                  70                  75                  80

Leu Arg Phe Ile Gln Ala Trp Gln Ser Leu Pro Asp Phe Gly Ile Ser
                 85                  90                  95

Tyr Val Met Val Arg Phe Lys Gly Ser Arg Lys Asp Glu Ile Leu Gly
            100                 105                 110

Ile Ala Asn Asn Arg Leu Ile Arg Ile Asp Leu Ala Val Gly Asp Val
        115                 120                 125

Val Lys Thr Trp Arg Phe Ser Asn Met Arg Gln Trp Asn Val Asn Trp
    130                 135                 140

Asp Ile Arg Xaa Val Ala Ile Glu Phe Asp Glu His Ile Asn Val Ala
145                 150                 155                 160

Phe Ser Cys Val Ser Ala Ser Cys Arg Ile Val His Glu Tyr Ile Gly
                165                 170                 175

Gly Tyr Ile Phe Leu Ser Thr Arg Glu Xaa Ala Arg Gly Glu Glu Leu
            180                 185                 190

Asp Glu Asp Leu Phe Leu Gln Leu Thr Gly Gly His Glu Ala Phe Xaa
        195                 200                 205
```

(2) INFORMACIÓN PARA SEQ ID NO: 81:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 43 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 81:

```
Met Ile Phe Leu Leu Phe Leu Thr Pro Leu Trp Leu Gln Lys Gly Ser
 1               5                  10                  15

Ala Gly Lys Met Ser Gly Glu Phe Leu Tyr Ala Ser Leu Phe Gln Trp
            20                  25                  30

Asn Tyr Phe Trp Arg Asn Lys Lys Val Cys Xaa
        35                  40
```

(2) INFORMACIÓN PARA SEQ ID NO: 82:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 146 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 82:

```
Met Pro Ser Gly Phe Gln Thr Cys Leu Leu Phe Thr Leu Ser Pro Phe
 1               5                  10                  15

Ser Leu Ser Lys Ile Val Gly Val Pro Ser Gln Gln Leu Pro Gly Gln
            20                  25                  30

Leu Ser Glu Gln Gly Gly Leu Cys Gly His Glu Gly Glu Pro Ala Arg
        35                  40                  45

Thr Val Pro Glu Thr Gln Leu Pro Leu Pro Phe Asn Ser Ala Gly Pro
    50                  55                  60

Pro His Leu Lys Cys Thr Gly Ala Gly Lys Arg Val Trp Ser Pro Pro
65                  70                  75                  80

Arg Arg Ala Ala Gln Glu Val Ser Leu Gln Leu Val Ser Cys Xaa Pro
                85                  90                  95

Cys Arg Gln Xaa Thr Ser Arg Ala Phe Ser Leu Ala Thr Asp Arg Thr
            100                 105                 110

Ala Ser Ala Arg Val Cys Cys Arg Phe Pro Phe Lys His Thr His Ser
        115                 120                 125

Pro His Pro Arg Arg Pro Glu Val Gln Gly Ala Trp Ala Val Val Pro
    130                 135                 140

Leu Xaa
145
```

(2) INFORMACIÓN PARA SEQ ID NO: 83:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 25 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 83:

```
Met Pro Trp Arg Arg Ala Gly Leu Met Met Leu Pro Ile Ile Thr Gly
 1               5                  10                  15

Cys Cys Pro Cys Ser Ala Ser Ile Xaa
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 84:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 31 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 84:

```
Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Leu Pro Cys
 1               5                  10                  15

Pro Ser Pro Trp Xaa Arg Arg Ile Ser Gln Gly Pro Gly Thr Xaa
            20                  25                  30
```

(2) INFORMACIÓN PARA SEQ ID NO: 85:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 374 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 85:

```
Met Ser Val Pro Ala Phe Ile Asp Ile Ser Glu Glu Asp Gln Ala Ala
 1               5                  10                  15

Glu Leu Arg Ala Tyr Leu Lys Ser Lys Gly Ala Glu Ile Ser Glu Glu
            20                  25                  30

Asn Ser Glu Gly Gly Leu His Val Asp Leu Ala Gln Ile Ile Glu Ala
        35                  40                  45

Cys Asp Val Cys Leu Lys Glu Asp Asp Lys Asp Val Glu Ser Val Met
    50                  55                  60

Asn Ser Val Val Ser Leu Leu Leu Ile Leu Glu Pro Asp Lys Gln Glu
65                  70                  75                  80

Ala Leu Ile Glu Ser Leu Cys Glu Lys Leu Val Lys Phe Arg Glu Gly
                85                  90                  95

Glu Arg Pro Ser Leu Arg Leu Gln Leu Leu Ser Asn Leu Phe His Gly
            100                 105                 110

Met Asp Lys Asn Thr Pro Val Arg Tyr Thr Val Tyr Cys Ser Leu Ile
        115                 120                 125

Lys Val Ala Ala Ser Cys Gly Ala Ile Gln Tyr Ile Pro Thr Glu Leu
    130                 135                 140

Asp Gln Val Arg Lys Trp Ile Ser Asp Trp Asn Leu Thr Thr Glu Lys
145                 150                 155                 160

Lys His Thr Leu Leu Arg Leu Leu Tyr Glu Ala Leu Val Asp Cys Lys
                165                 170                 175

Lys Ser Asp Ala Ala Ser Lys Val Met Val Glu Leu Leu Gly Ser Tyr
            180                 185                 190

Thr Glu Asp Asn Ala Ser Gln Ala Arg Val Asp Ala His Arg Cys Ile
        195                 200                 205

Val Arg Ala Leu Lys Asp Pro Asn Ala Phe Leu Phe Asp His Leu Leu
    210                 215                 220
```

```
Thr Leu Lys Pro Val Lys Phe Leu Glu Gly Glu Leu Ile His Asp Leu
225                 230                 235                 240

Leu Thr Ile Phe Val Ser Ala Lys Leu Ala Ser Tyr Val Lys Phe Tyr
                245                 250                 255

Gln Asn Asn Lys Asp Phe Ile Asp Ser Leu Gly Leu Leu His Glu Gln
            260                 265                 270

Asn Met Ala Lys Met Arg Leu Leu Thr Phe Met Gly Met Ala Val Glu
        275                 280                 285

Asn Lys Glu Ile Ser Phe Asp Thr Met Gln Gln Glu Leu Gln Ile Gly
    290                 295                 300

Ala Asp Asp Val Glu Ala Phe Val Ile Asp Ala Val Arg Thr Lys Met
305                 310                 315                 320

Val Tyr Cys Lys Ile Asp Gln Thr Gln Arg Lys Val Val Val Ser His
                325                 330                 335

Ser Thr His Arg Thr Phe Gly Lys Gln Gln Trp Gln Gln Leu Tyr Asp
            340                 345                 350

Thr Leu Asn Ala Trp Lys Gln Asn Leu Asn Lys Val Lys Asn Ser Leu
        355                 360                 365

Leu Ser Leu Ser Asp Thr
    370
```

(2) INFORMACIÓN PARA SEQ ID NO: 86:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 13 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 86:

```
Met Ser Val Pro Ala Phe Ile Asp Ile Ser Glu Glu Asp
  1               5                  10
```

(2) INFORMACIÓN PARA SEQ ID NO: 87:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 15 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 87:

```
Gln Ala Ala Glu Leu Arg Ala Tyr Leu Lys Ser Lys Gly Ala Glu
  1               5                  10                  15
```

(2) INFORMACIÓN PARA SEQ ID NO: 88:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 17 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 88:

```
Ile Ser Glu Glu Asn Ser Glu Gly Gly Leu His Val Asp Leu Ala Gln
 1               5                  10                  15

Ile
```

(2) INFORMACIÓN PARA SEQ ID NO: 89:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 18 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 89:

```
Ile Glu Ala Cys Asp Val Cys Leu Lys Glu Asp Asp Lys Asp Val Glu
 1               5                  10                  15

Ser Val
```

(2) INFORMACIÓN PARA SEQ ID NO: 90:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 16 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 90:

```
Val Ala Arg Pro Ser Ser Leu Phe Arg Ser Ala Trp Ser Cys Glu Trp
 1               5                  10                  15
```

(2) INFORMACIÓN PARA SEQ ID NO: 91:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 12 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 91:

```
Leu Arg Leu Gln Leu Leu Ser Asn Leu Phe His Gly
 1               5                  10
```

(2) INFORMACIÓN PARA SEQ ID NO: 92:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 17 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 92:

```
Lys Asp Val Glu Ser Val Met Asn Ser Val Val Ser Leu Leu Leu Ile
 1               5                  10                  15

Leu
```

(2) INFORMACIÓN PARA SEQ ID NO: 93:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 93:

```
Asp Ala Ala Ser Lys Val Met Val Glu Leu Leu Gly Ser Tyr Thr Glu
 1               5                  10                  15

Asp Asn Ala Ser Gln Ala Arg Val Asp Ala
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 94:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 10 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 94:

```
Val Glu Ala Phe Val Ile Asp Ala Val Arg
 1               5                  10
```

(2) INFORMACIÓN PARA SEQ ID NO: 95:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 35 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 95:

```
Met Ser Glu Ile Tyr Leu Arg Cys Gln Asp Glu Gln Gln Tyr Ala Arg
 1               5                  10                  15

Trp Met Ala Gly Cys Arg Leu Ala Ser Lys Gly Arg Thr Met Ala Asp
            20                  25                  30

Ser Ser Tyr
        35
```

(2) INFORMACIÓN PARA SEQ ID NO: 96:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 45 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 96:

```
Leu Val Ala Pro Arg Phe Gln Arg Lys Phe Lys Ala Lys Gln Leu Thr
 1               5                  10                  15

Pro Arg Ile Leu Glu Ala His Gln Asn Val Ala Gln Leu Ser Leu Ala
            20                  25                  30

Glu Ala Gln Leu Arg Phe Ile Gln Ala Trp Gln Ser Leu
        35                  40                  45
```

(2) INFORMACIÓN PARA SEQ ID NO: 97:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 23 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 97:

```
Val Gly Asp Val Val Lys Thr Trp Arg Phe Ser Asn Met Arg Gln Trp
 1               5                  10                  15

Asn Val Asn Trp Asp Ile Arg
            20
```

(2) INFORMACIÓN PARA SEQ ID NO: 98:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 98:

```
Glu Glu Ile Asp Cys Thr Glu Glu Glu Met Met Val Phe Ala Ala Leu
  1               5                  10                  15

Gln Tyr His Ile Asn Lys Leu Ser Gln Ser
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 99:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 99:

```
Glu Glu Ile Asp Cys Thr Glu Glu Glu Met Met Val Phe Ala Ala Leu
  1               5                  10                  15

Gln Tyr His Ile Asn Lys Leu Ser Gln Ser
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 100:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 100:

```
Lys Glu Leu Ser Phe Ala Arg Ile Lys Ala Val Glu Cys Val Glu Ser
  1               5                  10                  15

Thr Gly Arg His Ile Tyr Phe Thr Leu Val
            20                  25
```

(2) INFORMACIÓN PARA SEQ ID NO: 101:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 17 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 101:

```
Gly Trp Asn Ala Gln Ile Thr Leu Gly Leu Val Lys Phe Lys Asn Gln
  1               5                  10                  15

Gln
```

(2) INFORMACIÓN PARA SEQ ID NO: 102:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 16 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 102:

```
Leu Val Leu Gly Leu Ser Xaa Leu Asn Asn Ser Tyr Asn Phe Ser Phe
 1               5                  10                  15
```

(2) INFORMACIÓN PARA SEQ ID NO: 103:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 17 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 103:

```
His Val Val Ile Gly Ser Gln Ala Glu Glu Gly Gln Tyr Ser Leu Asn
 1               5                  10                  15

Phe
```

(2) INFORMACIÓN PARA SEQ ID NO: 104:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 19 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 104:

```
His Asn Cys Asn Asn Ser Val Pro Gly Lys Glu His Pro Phe Asp Ile
 1               5                  10                  15

Thr Val Met
```

(2) INFORMACIÓN PARA SEQ ID NO: 105:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 17 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 105:

```
Phe Ile Lys Tyr Val Leu Ser Asp Lys Glu Lys Lys Val Phe Gly Ile
 1               5                  10                  15

Val
```

(2) INFORMACIÓN PARA SEQ ID NO: 106:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 13 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 106:

Ile Pro Met Gln Val Leu Ala Asn Val Ala Tyr Ile Ile
1 5 10

(2) INFORMACIÓN PARA SEQ ID NO: 107:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 13 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 107:

Ile Pro Met Gln Val Leu Ala Asn Val Ala Tyr Ile Ile
1 5 10

(2) INFORMACIÓN PARA SEQ ID NO: 108:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 15 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 108:

Asp Gly Lys Val Ala Val Asn Leu Ala Lys Leu Lys Leu Phe Arg
1 5 10 15

(2) INFORMACIÓN PARA SEQ ID NO: 109:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 13 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 109:

Ile Arg Glu Lys Asn Pro Asp Gly Phe Leu Ser Ala Ala
1 5 10

(2) INFORMACIÓN PARA SEQ ID NO: 110:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 9 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 110:

```
Met Met Phe Gly Gly Tyr Glu Thr Ile
 1               5
```

(2) INFORMACIÓN PARA SEQ ID NO: 111:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 111:

```
Tyr Arg Asp Glu Ser Ser Ser Glu Leu Ser Val Asp Ser Glu Val Glu
 1               5                  10                  15

Phe Gln Leu Tyr Ser Gln Ile His
            20
```

(2) INFORMACIÓN PARA SEQ ID NO: 112:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 136 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 112:

```
Tyr Ala Gln Asp Leu Asp Asp Val Ile Arg Glu Glu Glu His Glu Glu
 1               5                  10                  15

Lys Asn Ser Gly Asn Ser Glu Ser Ser Ser Ser Lys Pro Asn Gln Lys
            20                  25                  30

Lys Leu Ile Val Leu Ser Asp Ser Glu Val Ile Gln Leu Ser Asp Gly
        35                  40                  45

Ser Glu Val Ile Thr Leu Ser Asp Glu Asp Ser Ile Tyr Arg Cys Lys
    50                  55                  60

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

Ser Ser Leu Gln Ser Asn Glu Leu Val Asp Lys Lys Cys Lys Ser Asp
                85                  90                  95

Ile Glu Lys Pro Lys Ser Glu Glu Arg Ser Gly Val Ile Arg Glu Val
            100                 105                 110

Met Ile Ile Glu Val Ser Ser Ser Glu Glu Glu Glu Ser Thr Ile Ser
        115                 120                 125

Glu Gly Asp Asn Val Glu Ser Trp
    130                 135
```

(2) INFORMACIÓN PARA SEQ ID NO: 113:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 37 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 113:

```
Met Leu Leu Gly Cys Glu Val Asp Asp Lys Asp Asp Asp Ile Leu Leu
 1               5                  10                  15

Asn Leu Val Gly Cys Glu Asn Ser Val Thr Glu Gly Glu Asp Gly Ile
            20                  25                  30

Asn Trp Ser Ile Ser
        35
```

(2) INFORMACIÓN PARA SEQ ID NO: 114:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 18 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 114:

```
Asp Lys Asp Ile Glu Ala Gln Ile Ala Asn Asn Arg Thr Pro Gly Arg
 1               5                  10                  15

Trp Thr
```

(2) INFORMACIÓN PARA SEQ ID NO: 115:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 31 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 115:

```
Gln Arg Tyr Tyr Ser Ala Asn Lys Asn Ile Ile Cys Arg Asn Cys Asp
 1               5                  10                  15

Lys Arg Gly His Leu Ser Lys Asn Cys Pro Leu Pro Arg Lys Val
            20                  25                  30
```

(2) INFORMACIÓN PARA SEQ ID NO: 116:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 179 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 116:

```
Arg Arg Cys Phe Leu Cys Ser Arg Arg Gly His Leu Leu Tyr Ser Cys
 1               5                  10                  15

Pro Ala Pro Leu Cys Glu Tyr Cys Pro Val Pro Lys Met Leu Asp His
            20                  25                  30

Ser Cys Leu Phe Arg His Ser Trp Asp Lys Gln Cys Asp Arg Cys His
        35                  40                  45

Met Leu Gly His Tyr Thr Asp Ala Cys Thr Glu Ile Trp Arg Gln Tyr
    50                  55                  60

His Leu Thr Thr Lys Pro Gly Pro Pro Lys Lys Pro Lys Thr Pro Ser
65                  70                  75                  80

Arg Pro Ser Ala Leu Ala Tyr Cys Tyr His Cys Ala Gln Lys Gly His
                85                  90                  95

Tyr Gly His Glu Cys Pro Glu Arg Glu Val Tyr Asp Pro Ser Pro Val
            100                 105                 110

Ser Pro Phe Ile Cys Tyr Tyr Xaa Asp Lys Tyr Glu Ile Gln Glu Arg
        115                 120                 125

Glu Lys Arg Leu Lys Gln Lys Ile Lys Val Xaa Lys Lys Asn Gly Val
    130                 135                 140

Ile Pro Glu Pro Ser Lys Leu Pro Tyr Ile Lys Ala Ala Asn Glu Asn
145                 150                 155                 160

Pro His His Asp Ile Arg Lys Gly Arg Ala Ser Trp Lys Ser Asn Arg
                165                 170                 175

Trp Pro Gln
```

(2) INFORMACIÓN PARA SEQ ID NO: 117:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 17 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 117:

```
Leu Ser Ile Ile Phe Leu Ala Phe Val Ser Ile Asp Arg Cys Leu Gln
 1               5                  10                  15

Leu
```

(2) INFORMACIÓN PARA SEQ ID NO: 118:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 67 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 118:

```
Gly Ser Cys Phe Ala Thr Trp Ala Phe Ile Gln Lys Asn Thr Asn His
 1               5                  10                  15

Arg Cys Val Ser Ile Tyr Leu Ile Asn Leu Leu Thr Ala Asp Phe Leu
            20                  25                  30

Leu Thr Leu Ala Leu Pro Val Lys Ile Val Val Asp Leu Gly Val Ala
        35                  40                  45

Pro Trp Lys Leu Lys Ile Phe His Cys Gln Val Thr Ala Cys Leu Ile
    50                  55                  60

Tyr Ile Asn
65
```

(2) INFORMACIÓN PARA SEQ ID NO: 119:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 60 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 119:

```
Ala Pro Leu Glu Thr Met Gln Asn Lys Pro Arg Ala Pro Gln Lys Arg
 1               5                  10                  15

Ala Leu Pro Phe Pro Glu Leu Glu Leu Arg Asp Tyr Ala Ser Val Leu
            20                  25                  30

Thr Arg Tyr Ser Leu Gly Leu Arg Asn Lys Glu Pro Ser Leu Gly His
        35                  40                  45

Arg Trp Gly Thr Gln Lys Leu Gly Arg Ser Pro Cys
    50                  55                  60
```

(2) INFORMACIÓN PARA SEQ ID NO: 120:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 166 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 120:

```
Asn Arg Glu Arg Gly Gly Ala Gly Ala Thr Phe Glu Cys Asn Ile Cys
 1               5                  10                  15

Leu Glu Thr Ala Arg Glu Ala Val Val Ser Val Cys Gly His Leu Tyr
            20                  25                  30

Cys Trp Pro Cys Leu His Gln Trp Leu Glu Thr Arg Pro Glu Arg Gln
        35                  40                  45

Glu Cys Pro Val Cys Lys Ala Gly Ile Ser Arg Glu Lys Val Val Pro
    50                  55                  60

Leu Tyr Gly Arg Gly Ser Gln Lys Pro Gln Asp Pro Arg Leu Lys Thr
65                  70                  75                  80

Pro Pro Arg Pro Gln Gly Gln Arg Pro Ala Pro Glu Ser Arg Gly Gly
                85                  90                  95

Phe Gln Pro Phe Gly Asp Thr Gly Gly Phe His Phe Ser Phe Gly Val
            100                 105                 110

Gly Ala Phe Pro Phe Gly Phe Phe Thr Thr Val Phe Asn Ala His Glu
        115                 120                 125

Pro Phe Arg Arg Gly Thr Gly Val Asp Leu Gly Gln Gly His Pro Ala
    130                 135                 140

Ser Ser Trp Gln Asp Ser Leu Phe Leu Phe Leu Ala Ile Phe Phe Phe
145                 150                 155                 160

Phe Trp Leu Leu Ser Ile
                165
```