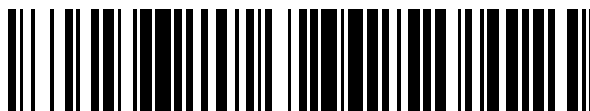


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 670 894**

51 Int. Cl.:

C12N 15/12 (2006.01)
C12N 15/13 (2006.01)
C12N 15/63 (2006.01)
C12N 15/85 (2006.01)
C12N 5/10 (2006.01)
C07K 14/47 (2006.01)
C07K 16/18 (2006.01)
C12P 21/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **27.11.2012 PCT/JP2012/080532**
- 87 Fecha y número de publicación internacional: **06.06.2013 WO13080934**
- 96 Fecha de presentación y número de la solicitud europea: **27.11.2012 E 12853220 (7)**
- 97 Fecha y número de publicación de la concesión europea: **21.03.2018 EP 2787080**

54 Título: **Promotor derivado de un gen humano**

30 Prioridad:

28.11.2011 JP 2011258724

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

01.06.2018

73 Titular/es:

**DAIICHI SANKYO COMPANY, LIMITED (100.0%)
3-5-1, Nihonbashi Honcho Chuo-ku
Tokyo 103-8426, JP**

72 Inventor/es:

MURAKAMI KENJI

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 670 894 T3

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

DESCRIPCIÓN

Promotor derivado de un gen humano

Campo técnico

5 La presente invención se refiere a un polinucleótido promotor así como a vectores y unidades de expresión de genes extraños que comprenden el promotor, así como a células transformadas cuya actividad transcripcional de proteínas extrañas se ha mejorado mediante la utilización del vector de expresión de genes extraños que tiene el promotor que procede de un gen humano y a un procedimiento para producir la proteína extraña utilizando la célula huésped.

Técnica antecedente

10 Debido al desarrollo de técnicas de recombinación genética, se ha expandido rápidamente el mercado de productos farmacéuticos proteicos tales como proteínas terapéuticas y fármacos con anticuerpos. En particular, los fármacos con anticuerpos pueden tener alta especificidad sin causar una inmunorreacción adversa cuando se administran al cuerpo humano y, por lo tanto, se ha buscado activamente el desarrollo de los mismos.

15 Como un hospedador mediante el cual se produce un producto proteico farmacéutico tipificado mediante un fármaco con anticuerpos, se puede utilizar un microorganismo, una levadura, un insecto, una célula animal o vegetal, una célula animal o vegetal transgénica o similares. Para que el producto proteico farmacéutico tenga actividad biológica o inmunogenicidad, es esencial una modificación postraducciona tal como plegamiento o glucosilación. Por lo tanto, un microorganismo, con el que no se puede realizar una modificación postraducciona complicada, o una planta, que tiene una estructura de glucano diferente, no es adecuado como el hospedador. El uso de una célula de mamífero cultivada tal como una célula CHO (ovario de hámster chino), que proviene de una especie estrechamente relacionada con los seres humanos, es el estándar actual considerando que dicha célula tiene una estructura de glucano similar a las de los seres humanos y es segura, y se puede realizar la modificación postraducciona utilizando dicha célula.

20 En los casos en los que se utiliza una célula de mamífero cultivada como la huésped, existe el problema de que la tasa de crecimiento es baja, la productividad es baja, el coste es alto, etc., en comparación con un microorganismo o similar (NPL 1). Además, para utilizar un producto proteico farmacéutico clínicamente, es necesario administrar una gran cantidad del producto. Por lo tanto, la falta de capacidad de producción del mismo es otro problema en el mundo. Cuando se produce un producto proteico farmacéutico en un sistema de expresión de células de mamífero cultivadas, el coste de producción es alto en comparación con un producto farmacéutico sintético de peso molecular bajo. En consecuencia, se han realizado intentos para reducir el coste de producción mejorando las fases de producción respectivas. La mejora de la cantidad de producción en el sistema de expresión de células de mamífero cultivadas es un procedimiento eficaz para reducir el coste de producción (NPL 2 y NPL 3). Por consiguiente, para mejorar la productividad de un gen extraño en una célula de mamífero cultivada, se han investigado diversos enfoques basados en promotores, potenciadores, marcadores de selección de antibióticos, amplificación de genes, técnicas de ingeniería de cultivo y similares. En los casos en los que una célula CHO se utiliza como una célula huésped para expresar un gen extraño, es decir, para producir un producto proteico farmacéutico, generalmente se utiliza un promotor temprano inmediato principal de citomegalovirus humano derivado de virus (en lo sucesivo denominado "promotor de CMV") (NPL 4, NPL 5 y NPL 6). Además, se sabe que un polinucleótido cadena arriba del sitio de inicio de la transcripción de un gen de proteína ribosómica humana tal como RPL32 o RPS11 puede utilizarse como un elemento de ADN para la expresión de proteínas en una célula CHO, en combinación con otro promotor heterólogo (NPL 7 y PLT 1).

45 Los números de acceso de base de datos AC108488, AC018836 y AC084209 revelan secuencias de varios genes que codifican proteínas, sin mencionar su función. Annilo y col. (1995) desvela el gen que codifica la proteína S7 ribosómica humana pero no revela ninguna actividad promotora significativa. El documento WO 2009/155950 desvela construcciones de indicadores lineales activos de expresión transcripcional y/o postranscripcionalmente controlables.

Listado de citas

Referencias de Patente

PTL 1: documento WO 2006/123097

Referencias de No Patente

50 NPL 1: Florian M. Wurm., Nat. Biotechnol. 22(11): 1393-1398, 2004
 NPL 2: Farid SS., J Chromatogr B Analyt Technol Biomed Life Sci. 848(1): 8-18, 2007
 NPL 3: Werner RG. Economic aspects of commercial manufacture of biopharmaceuticals. J Biotechnol. 113 (1-3): 171-182, 2004
 NPL 4: Durocher Y y col., Curr Opin Biotechnol. 20(6): 700-707, 2009
 55 NPL 5: Boshart M y col., Cell. 41(2): 521-530, 1985

NPL 6: Foecking MK y col., Gene. 45(1): 101-105, 1986

NPL 7: Hoeksema F. y col., Biotechnology Research International, Volumen 2011, Artículo ID 492875, 11 páginas

Sumario de la invención

Problema técnico

5 Un objeto de la invención es proporcionar un procedimiento para aumentar la producción de una proteína extraña para utilizarse en un producto proteico farmacéutico, utilizando un promotor que tiene una alta actividad para mejorar la expresión de genes extraños en una célula huésped tal como una célula de mamífero cultivada. Mediante la identificación de un promotor que tiene una actividad promotora equivalente a o mayor que la de un promotor de CMV en una célula CHO o similar, se proporciona un procedimiento para lograr establemente una alta expresión de genes extraños en una célula de mamífero, y un procedimiento para contribuir a la mejora de niveles de producción, en otras palabras, se puede proporcionar una reducción en los costes de producción de un producto proteico farmacéutico en un sistema de expresión de células de mamífero cultivadas.

Solución al problema

15 Los presentes inventores realizaron estudios intensivos para resolver los problemas anteriores y encontraron que un polinucleótido, que partía de un nucleótido situado aproximadamente 2 kpb cadena arriba del lugar de inicio de la transcripción y terminaba en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde al codón de inicio de un gen de proteína ribosómica humana, tiene una alta actividad promotora. Encontraron que la actividad promotora puede mejorar significativamente la producción de una proteína extraña que se va a expresar en una célula de mamífero cultivada y, así, completaron la invención. La invención incluye los siguientes aspectos.

- 20 (1) Un polinucleótido promotor que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 1 en el listado de secuencias.
- (2) Una unidad de expresión de genes extraños que comprende el polinucleótido promotor de acuerdo con el punto anterior (1).
- 25 (3) La unidad de expresión de genes extraños de acuerdo con el punto anterior (2), en la que el gen extraño es un gen que codifica una proteína multimérica o una proteína heteromultimérica.
- (4) La unidad de expresión de genes extraños de acuerdo con el punto anterior (2), en la que el gen extraño es un gen que codifica un anticuerpo o un fragmento funcional del mismo.
- 30 (5) Un vector de expresión de genes extraños que comprende la unidad de expresión de genes extraños de acuerdo con uno cualquiera de los puntos anteriores (2) a (4).
- (6) Un vector de expresión de genes extraños que comprende la unidad de expresión de genes extraños de acuerdo con uno cualquiera de los puntos anteriores (2) a (4) y uno o más polinucleótidos seleccionados a partir de los polinucleótidos descritos en (a) a (i) en el siguiente grupo A:

Grupo A

- 35 (a) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 10 en el Listado de Secuencias;
- (b) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 11 en el Listado de Secuencias;
- 40 (c) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 12 en el Listado de Secuencias;
- (d) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 13 en el Listado de Secuencias;
- (e) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 14 en el Listado de Secuencias;
- 45 (f) un polinucleótido que comprende al menos 3000 nucleótidos consecutivos de una secuencia de nucleótidos representada por una cualquiera de las SEQ ID NOS: 10 a 14 en el Listado de secuencias;
- (g) un polinucleótido que comprende al menos 2000 nucleótidos consecutivos de una secuencia de nucleótidos representada por una cualquiera de las SEQ ID NOS: 10 a 14 en el Listado de secuencias;
- 50 (h) un polinucleótido que comprende una secuencia polinucleotídica que tiene una identidad de 95 % o más con la secuencia de nucleótidos del polinucleótido de acuerdo con uno cualquiera de los puntos anteriores (a) a (g) y que tiene la actividad de mejorar la expresión de genes extraños; y
- (i) un polinucleótido que comprende una secuencia de nucleótidos que tiene una identidad de 99 % o más con la secuencia de nucleótidos del polinucleótido de acuerdo con uno cualquiera de los puntos anteriores (a) a (g) y que tiene la actividad de mejorar la expresión de genes extraños.
- 55 (7) Una célula transformada en la que se ha introducido el vector de expresión de genes extraños de acuerdo con el punto anterior (5) o (6).
- (8) Una célula transformada en la que se han introducido el vector de expresión de genes extraños de acuerdo con el punto anterior (5) o (6) y un vector de elemento.

(9) La célula transformada de acuerdo con el punto anterior (7) u (8), en la que la célula es una célula cultivada que procede de un mamífero.

(10) La célula transformada de acuerdo con el punto anterior (9), en la que la célula cultivada que procede de un mamífero es una célula COS-1, una célula 293 o una célula CHO.

5 (11) Un procedimiento para producir una proteína caracterizado porque comprende cultivar la célula transformada de acuerdo con uno cualquiera de los puntos anteriores (7) a (10) y obtener una proteína que procede de un gen extraño a partir del producto de cultivo resultante.

(12) Uso de una secuencia polinucleotídica promotora de acuerdo con el punto anterior (1) para expresar un gen extraño en una célula transformada.

10 (13) Uso de una secuencia polinucleotídica promotora que tiene una identidad de 95 % o más con la secuencia de nucleótidos del polinucleótido de acuerdo con la reivindicación 1 para expresar un gen extraño en una célula transformada.

(14) Uso de una secuencia polinucleotídica promotora que tiene una identidad de 99% o más con la secuencia de nucleótidos del polinucleótido de acuerdo con la reivindicación 1 para expresar un gen extraño en una célula transformada.

15 (15) Uso del vector de expresión de genes extraños de acuerdo con el punto anterior (5) o (6) para expresar un gen extraño en una célula transformada.

Efectos ventajosos de la invención

20 Mediante la introducción de un vector de expresión de genes extraños utilizando un promotor que procede de un gen humano de la invención en una célula huésped de mamífero, se puede mejorar significativamente la expresión de un gen extraño de una proteína terapéutica, un anticuerpo o similar. Además, mediante la utilización del promotor de la invención en combinación con un elemento de ADN, se puede mejorar adicionalmente la expresión de un gen extraño de una proteína terapéutica, un anticuerpo o similar.

Breve descripción de los dibujos

25 [FIG. 1] La figura 1 muestra un gráfico en el que se evaluó la actividad de los promotores mediante la utilización de la actividad de SEAP como un índice en células policlonales CHO-K1 transfectadas. El gráfico muestra la actividad de SEAP para cada promotor, con el valor para un promotor de CMV normalizado a 1. Se muestran los resultados de dos experimentos diferentes (n = 3, media ± SD).

30 [FIG. 2] La figura 2 muestra un gráfico en el que se evaluó la actividad de promotores truncados mediante la utilización de la actividad de SEAP como un índice en células policlonales CHO-K1 transfectadas. El gráfico muestra la actividad de cada promotor, con el valor para un promotor de CMV normalizado a 1 (n = 3, media ± SD).

[FIG. 3] La figura 3 muestra un gráfico en el que se confirmó mediante la amplificación de una región GAPDH que una muestra sometida a ChIP-on-chip fue inmunoprecipitada con cromatina específicamente con un anticuerpo anti histona H3 acetilada.

35 [FIG. 4] La figura 4 es una vista esquemática de un vector de expresión de SEAP en el que se ha insertado un elemento de ADN.

[FIG. 5] La figura 5 muestra un gráfico en el que los efectos de mejora de la expresión de los elementos de ADN A2, A7, A18, B5 y C14 se confirmaron mediante la utilización de la actividad de SEAP expresada mediante un promotor de CMV como un índice en una línea celular CHO transfectada.

40 [FIG. 6] La figura 6 muestra gráficos en los que se confirmaron los efectos de mejora de la expresión de los elementos de ADN A2 y A7 mediante la utilización de la actividad de SEAP expresada mediante un EF-1α o un promotor de SV40 como un índice en una línea celular CHO transfectada.

[FIG. 7] La figura 7 es una vista esquemática de un vector de expresión de anticuerpos (coexpresión de cadena ligera y cadena pesada del gen X de anticuerpos) en el que se ha insertado un elemento de ADN.

45 [FIG. 8] La figura 8 muestra gráficos en los que se confirmó el efecto de mejora de la expresión del elemento de ADN A7 mediante la utilización del nivel de producción (medido por el procedimiento ELISA) de un anticuerpo expresado mediante un CMV o un promotor EF-1α como un índice en una línea celular CHO transfectada.

[FIG. 9] La figura 9 es una tabla que muestra las longitudes de secuencia del elemento de ADN A2 y secuencias relacionadas.

50 [Fig. 10] La figura 10 muestra gráficos en los que se confirmaron los efectos de mejora de la expresión del elemento de ADN A2 y secuencias relacionadas mediante la utilización de la actividad de SEAP como un índice en una línea celular CHO transfectada.

[FIG. 11] La figura 11 es una tabla que muestra las longitudes de secuencia del elemento de ADN A7 y secuencias relacionadas.

55 [FIG. 12] La figura 12 muestra gráficos en los que se confirmaron los efectos de mejora de la expresión del elemento de ADN A7 y secuencias relacionadas mediante la utilización de la actividad de SEAP como un índice en una línea celular CHO transfectada.

[FIG. 13] La figura 13 es una tabla que muestra las longitudes de secuencia del elemento de ADN A18 y secuencias relacionadas.

60 [FIG. 14] La figura 14 muestra un gráfico en el que se confirmaron los efectos de mejora de la expresión del elemento de ADN A18 y secuencias relacionadas mediante la utilización de la actividad de SEAP como un índice en una línea celular CHO transfectada.

[FIG. 15] La figura 15 es una tabla que muestra las longitudes de secuencia del elemento de ADN B5 y secuencias relacionadas.

[FIG. 16] La figura 16 muestra un gráfico en el que se confirmaron los efectos de mejora de la expresión del elemento de ADN B5 y secuencias relacionadas mediante la utilización de la actividad de SEAP como un índice en una línea celular CHO transfectada.

[FIG. 17] La figura 17 es una tabla que muestra las longitudes de secuencia del elemento de ADN C14 y secuencias relacionadas.

[FIG. 18] La figura 18 muestra gráficos en los que se confirmaron los efectos de mejora de la expresión del elemento de ADN C14 y secuencias relacionadas mediante la utilización de la actividad de SEAP como un índice en una línea celular CHO transfectada.

[FIG. 19] La figura 19 muestra un gráfico en el que se confirmaron los efectos de mejora de la expresión de los elementos de ADN A2, A7, A18, B5 y C14 mediante la utilización de la actividad de SEAP como un índice en una línea celular HEK293 transfectada.

[FIG. 20] La figura 20 es una tabla que muestra nucleótidos en los puntos de inicio y final basándose en la secuencia de longitud completa del elemento de ADN A2, A7 o A18.

[FIG. 21] La figura 21 es una tabla que muestra nucleótidos en los puntos de inicio y final basándose en la secuencia de longitud completa del elemento de ADN B5 o C14.

Descripción de las realizaciones

En lo sucesivo en el presente documento, la invención se describirá específicamente.

El término "gen" como se usa en el presente documento se refiere a un segmento que se transcribe en un ARNm y, a continuación, se traduce en una proteína e incluye no solo un ADN, sino también un ARNm del mismo, ADNc del mismo y un ARN del mismo.

El término "polinucleótido" como se usa en el presente documento se utiliza con el mismo significado que ácido nucleico y también incluye ADN, ARN, sonda, oligonucleótido y cebador.

Los términos "polipéptido" y "proteína" como se usa en el presente documento se utilizan sin distinción.

La expresión "expresión genética" como se usa en el presente documento se refiere a un fenómeno en el que un ARNm se transcribe a partir de un gen y/o un fenómeno en el que una proteína se traduce a partir del ARNm.

La expresión "gen extraño" como se usa en el presente documento se refiere a un gen que se introduce artificialmente en una célula huésped.

La expresión "proteína extraña" como se usa en el presente documento se refiere a una proteína codificada mediante un gen extraño.

La expresión "unidad de expresión de genes" como se usa en el presente documento se refiere a un polinucleótido que tiene, en la dirección del marco de lectura de la transcripción, al menos una región promotora, un gen extraño y una región de terminación de la transcripción (señal de adición de poli(A)).

La expresión "actividad para mejorar la expresión de genes extraños" como se usa en el presente documento se refiere a la actividad para mejorar la producción de una proteína extraña en una célula huésped mediante la creación de un medio ventajoso para la transcripción en cualquier ADN alrededor de la unidad de expresión de genes que contiene un gen extraño y mejorar significativamente la eficacia de la transcripción.

El término "promotor" como se usa en el presente documento se refiere a una región a la que puede unirse un factor de transcripción implicado en el inicio de la transcripción de ADN en ARN y, a veces, se denomina "región promotora" en esta descripción. Los ejemplos del promotor incluyen un polinucleótido que empieza en un nucleótido localizado aproximadamente 2 kpb cadena arriba de un sitio de inicio de la transcripción y termina en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde al codón de inicio y el promotor puede contener un 5' UTR y un intrón.

La expresión "actividad promotora" como se usa en el presente documento se refiere a una actividad en la que un factor de transcripción se une a un promotor e inicia la transcripción para producir una proteína codificada mediante un gen. Se puede analizar utilizando la actividad de una proteína codificada mediante un gen indicador tal como fosfatasa alcalina secretora (SEAP) como un índice.

La frase "que tiene una actividad promotora" como se usa en el presente documento se refiere a que tiene la actividad de expresar SEAP equivalente a o mayor que la de un promotor de CMV en las mismas condiciones que aquellas descritas a continuación (Ejemplo 3) para evaluar una actividad promotora mediante la utilización del nivel de expresión de SEAP como un índice.

La expresión "elemento de ADN" como se usa en el presente documento se refiere a un polinucleótido que tiene la actividad de mejorar la expresión de genes extraños en los casos en los que el polinucleótido se localiza en la

proximidad de una unidad de expresión de genes o en un vector de expresión de genes extraños que contiene una unidad de expresión de genes.

5 La expresión "fragmento funcional de un anticuerpo" como se usa en el presente documento se refiere a un fragmento parcial de un anticuerpo que tiene una actividad de unión a antígeno e incluye Fab, F(ab')₂, y similares. Sin embargo, la expresión no se limita a estas moléculas siempre que el fragmento tenga una afinidad de unión por un antígeno.

10 El término "identidad" como se usa en el presente documento se refiere a una relación entre las secuencias de dos o más secuencias de nucleótidos o secuencias de aminoácidos determinadas mediante la comparación de las secuencias, como se conoce en la técnica. En la técnica, el término "identidad" también puede referirse al grado de coincidencia entre cadenas de dos o más secuencias de nucleótidos o dos o más secuencias de aminoácidos. La "identidad" puede evaluarse mediante el cálculo del porcentaje de coincidencias idénticas entre la más pequeña de las dos o más secuencias con alineaciones vacías (si las hay) direccionadas mediante un modelo matemático específico o programa informático (es decir, "algoritmos"). Específicamente, la identidad se puede evaluar mediante la utilización de un software tal como Clustal W2 proporcionado por European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI), pero el software no se limita al mismo y se puede utilizar cualquiera siempre que sea utilizado por los expertos en la materia.

20 La frase "hibridado en condiciones rigurosas" como se usa en el presente documento se refiere a la hibridación en condiciones en las que se forma un denominado híbrido específico pero no se forma un híbrido no específico. Los ejemplos de las condiciones incluyen condiciones en las que una cadena complementaria de un ácido nucleico que comprende una secuencia de nucleótidos que tiene una identidad del 80 % o más, preferentemente el 90 % o más, más preferentemente el 95% o más, lo más preferentemente el 99% o más con otro ácido nucleico se hibrida y una cadena complementaria de un ácido nucleico que comprende una secuencia de nucleótidos que tiene una identidad menor no se hibrida. De manera más específica, significa que la hibridación se efectúa a 68 °C en una solución de hibridación disponible comercialmente ExpressHyb Hybridization Solution (fabricada por Clontech, Inc.) o la hibridación se efectúa en condiciones tales que la hibridación se realiza a 68 °C en presencia de NaCl 0,7 a 1,0 M utilizando un filtro que tiene ADN inmovilizado sobre el mismo, seguido de lavado a 68 °C utilizando solución SSC 0,1 a 2 x (la solución SSC 1 x se compone de NaCl 150 mM y citrato de sodio 15 mM) o en condiciones equivalentes a estas.

30 1. Promotor a utilizar para mejorar la expresión de genes extraños

35 Como un promotor que procede de un gen humano de la invención (en lo sucesivo en el presente documento a veces también denominado "promotor de la invención"), se prefiere un polinucleótido que empieza en un nucleótido localizado aproximadamente 2 kpb cadena arriba del sitio de inicio de la transcripción y termina en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde al codón de inicio de un gen de proteína ribosómica humana. El promotor que procede de un gen humano puede ser un polinucleótido que empieza en un nucleótido localizado aproximadamente 1 kpb o aproximadamente 0,5 kpb cadena arriba del sitio de inicio de la transcripción y termina en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde a la secuencia del codón de inicio de un gen de proteína ribosómica humana.

40 El gen de proteína ribosómica humana es preferentemente un gen de proteína S7 ribosómica humana (en lo sucesivo en el presente documento denominado "RPS7"), un gen de proteína L32 ribosómica humana (en lo sucesivo en el presente documento denominado "RPL32") o un gen de proteína L34 ribosómica humana (en lo sucesivo en el presente documento denominado "RPL34").

45 El promotor de la invención comprende un polinucleótido representado por la SEQ ID NO: 1 en el listado de secuencias y el uso del mismo incluye un polinucleótido que tiene una identidad del 95 % o más, lo más preferentemente el 99% o más con la SEQ ID NO: 1.

50 Las secuencias de nucleótidos de SEQ ID NOS: 1, 2 y 3 son secuencias que empiezan en un nucleótido localizado aproximadamente 2 kpb cadena arriba del sitio de inicio de la transcripción y terminan en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde al codón de inicio de RPS7, RPL32 y RPL34, respectivamente. Las secuencias de nucleótidos de SEQ ID NOS: 4, 6 y 8 son secuencias que empiezan en un nucleótido localizado aproximadamente 1 kpb cadena arriba del sitio de inicio de la transcripción y terminan en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde al codón de inicio de RPS7, RPL32 y RPL34, respectivamente. Las secuencias de nucleótidos de SEQ ID NOS: 5, 7 y 9 son secuencias que empiezan en un nucleótido localizado aproximadamente 0,5 kpb cadena arriba del sitio de inicio de la transcripción y terminan en un nucleótido inmediatamente cadena arriba de una secuencia de nucleótidos que corresponde al codón de inicio de RPS7, RPL32 y RPL34, respectivamente.

55 Además, el promotor desvelado en el presente documento puede ser un polinucleótido que comprende una secuencia de nucleótidos que tiene una identidad del 95 % o más, más preferentemente el 99 % o más con la secuencia de nucleótidos de SEQ ID NO: 1 y tiene una actividad promotora.

El promotor desvelado en el presente documento puede hibridarse con un polinucleótido que comprende una secuencia de nucleótidos complementaria a un polinucleótido representado por la SEQ ID NO: 1 en condiciones rigurosas y tiene una actividad promotora.

5 El promotor desvelado en el presente documento puede ser un polinucleótido que es un polinucleótido mutado que comprende una secuencia de nucleótidos en la que uno o más, por ejemplo, 1 a 300 o 1 a 30 nucleótidos se han eliminado, sustituido y/o añadido en la secuencia de nucleótidos representada por la SEQ ID NO: 1 y tiene actividad promotora.

10 La introducción de una mutación (eliminación, sustitución y/o adición) en la secuencia de nucleótidos anteriormente mencionada se puede realizar mediante un procedimiento conocido en la técnica tal como un procedimiento de Kunkel o un procedimiento de dúplex con huecos (gapped duplex) o un procedimiento equivalente. Por ejemplo, se puede utilizar un kit de introducción de mutaciones que utiliza un procedimiento de mutagénesis dirigida al sitio tal como Mutant-K (fabricado por TaKaRa Bio, Inc.) o Mutant-G (fabricado por TaKaRa Bio, Inc.), un kit de la serie de mutagénesis in vitro LA PCR (fabricado por TaKaRa Bio, Inc.). Dicho polinucleótido mutado también se puede utilizar como el promotor de la invención.

15 La actividad del promotor de la invención para mejorar la expresión de genes extraños se puede analizar mediante la utilización, como un índice, de la actividad de una proteína codificada por un gen indicador tal como SEAP. En los casos en los que la actividad de una proteína indicadora, cuando se utiliza el promotor de la invención, es equivalente a o mayor que cuando se utiliza un promotor de CMV, preferentemente, la actividad aumenta 1,2 veces o más, más preferentemente 1,5 veces o más, se puede considerar que el promotor tiene la actividad de mejorar la expresión de genes extraños. Incluso en los casos en los que la actividad se aumenta aproximadamente 1,2 veces o más, se espera que esto reduzca la escala de cultivo celular, el tiempo de cultivo celular y el paso de purificación, haciendo posible aumentar el rendimiento y reducir el coste del cultivo celular. Si el rendimiento aumenta, entonces es posible suministrar de manera estable una proteína extraña para utilizarse como un producto farmacéutico. Además, si se reduce el coste del cultivo celular, se reduce el coste de la proteína extraña a utilizar como un producto farmacéutico.

Además, también se puede utilizar el promotor de la invención para mejorar la expresión de un gen endógeno de una célula huésped mediante la introducción del promotor en la célula huésped utilizando un procedimiento bien conocido por los expertos en la materia.

2. Unidad de expresión de genes extraños

30 La unidad de expresión de genes extraños de la invención (en lo sucesivo en el presente documento a veces también denominado "unidad de expresión de genes de la invención") tiene, opcionalmente, en la dirección del marco de lectura de la transcripción, al menos el promotor de la invención descrito en el anterior punto "1", un gen extraño y una región de terminación de la transcripción (señal de adición de poli(A)).

35 Además, la secuencia de adición de poli(A) puede ser una secuencia que tiene la actividad para causar la terminación de la transcripción para la transcripción a partir del promotor y puede ser una secuencia procedente de un gen idéntico a o diferente al del promotor.

3. Elemento de ADN a utilizar para mejorar la expresión de genes extraños

40 Mediante la utilización de la unidad de expresión de genes de la invención descrita en el punto anterior "2" y en combinación con un elemento de ADN, se puede mejorar además la expresión de un gen extraño. Se puede obtener el elemento de ADN para utilizarse en combinación mediante la utilización, como un índice, de la interacción entre la histona H3 acetilada y el elemento como se describe en el ejemplo 6. En general, se dice que la acetilación de histonas (H3 y H4) está asociada con la activación de la transcripción y se han defendido dos teorías principales. Una teoría es que la acetilación de histonas está asociada con un cambio en la conformación del nucleosoma de tal manera que las colas de las histonas se acetilan para ser neutralizadas eléctricamente, debilitando las interacciones ADN-histona (Mellor J. (2006) Dynamic nucleosomes and gene transcription. Trends Genet. 22 (6): 320-329). La otra teoría es que la acetilación de histonas está asociada con el reclutamiento de diversos factores de transcripción (Nakatani Y. (2001) Histone acetylases-versatile players. Genes Cells. 6(2): 79-86). De acuerdo con cualquiera de las dos teorías, hay una alta posibilidad de que la acetilación de histonas esté asociada con la activación de la transcripción y mediante la realización de inmunoprecipitación con cromatina (ChIP) utilizando un anticuerpo anti histona H3 acetilada, es posible concentrar un elemento de ADN que interactúa con la histona H3 acetilada.

A2 es un ejemplo del elemento de ADN para utilizarse en combinación con el promotor de la invención para mejorar la expresión de genes extraños. A2 se localiza en la región de 80966429 a 80974878 del cromosoma 15 humano y es un polinucleótido de 8450 pb que tiene un contenido de AT del 62,2 %. La secuencia nucleótidos de A2 está representada por la SEQ ID NO: 10 en el listado de secuencias.

55 A7, A18, B5 y C14 son ejemplos de elementos de ADN similares. A7 se localiza en la región de 88992123 a 89000542 del cromosoma 11 humano y es un polinucleótido de 8420 pb que tiene un contenido de AT del 64,52 %. La secuencia nucleótidos de A7 está representada por la SEQ ID NO: 11 en el listado de secuencias.

A18 se localiza en la región de 111275976 a 111284450 del cromosoma 4 humano y es un polinucleótido de 8475 pb que tiene un contenido de AT del 62,54 %. La secuencia nucleótidos de A18 está representada por la SEQ ID NO: 12 en el listado de secuencias.

5 B5 se localiza en la región de 143034684 a 143043084 del cromosoma 1 humano y es un polinucleótido de 8401 pb que tiene un contenido de AT del 66,37 %. La secuencia nucleótidos de B5 está representada por la SEQ ID NO: 13 en el listado de secuencias.

Finalmente, C14 se localiza en la región de 46089056 a 46097482 del cromosoma 11 humano y es un polinucleótido de 8427 pb que tiene un contenido de AT del 63,81 %. La secuencia nucleótidos de C14 está representada por la SEQ ID NO: 14 en el listado de secuencias.

10 La actividad de mejorar la expresión de genes extraños del elemento de ADN para utilizarse en combinación con el promotor de la invención se puede analizar mediante la utilización, como un índice, de la actividad de una proteína codificada por un gen indicador tal como SEAP.

15 En los casos en los que el elemento de ADN se utiliza en combinación con el promotor de la invención, puede utilizarse uno cualquiera de los elementos de ADN anteriores solo o pueden utilizarse dos o más copias de un tipo del elemento de ADN. Como alternativa, pueden utilizarse en combinación dos o más tipos diferentes de los elementos de ADN anteriores.

A2, A7, A18, B5 y C14 son ejemplos preferidos del elemento de ADN para utilizarse en combinación con el promotor de la invención.

20 El elemento de ADN para utilizarse en la invención puede ser una secuencia de nucleótidos que comprende una secuencia de nucleótidos que tiene una identidad del 95 % o más, más preferentemente del 99 % o más con cualquiera de las secuencias de nucleótidos representadas por las SEQ ID NOS: 10 a 14 y tiene la actividad de mejorar la expresión de genes extraños. La búsqueda de homología de secuencia de nucleótidos se puede realizar contra, por ejemplo, el banco de datos de ADN de Japón o similar utilizando un programa tal como FASTA o BLAST.

25 El elemento de ADN para utilizarse en combinación con el promotor de la invención puede hibridarse con un polinucleótido que comprende una secuencia de nucleótidos complementaria a un polinucleótido representado por las SEQ ID NOS: 10 a 14 en condiciones rigurosas y tiene la actividad de mejorar la expresión de genes extraños.

30 Un experto en la materia puede obtener fácilmente dicho gen homólogo con referencia a la clonación molecular (Sambrook, J. y col., Molecular Cloning: a Laboratory Manual 2ª ed., Cold Spring Harbor Laboratory Press, 10 Skyline Drive Plainview, N.Y. (1989)) o similares. Además, la identidad de la secuencia de nucleótidos mencionada anteriormente se puede determinar mediante una búsqueda FASTA o búsqueda BLAST de la misma forma.

35 La introducción de una mutación (eliminación, sustitución y/o adición) en el polinucleótido mencionado anteriormente se puede realizar mediante un procedimiento conocido en la técnica tal como un procedimiento de Kunkel o un procedimiento de dúplex con huecos (gapped duplex) o un procedimiento equivalente. Por ejemplo, se puede utilizar un kit de introducción de mutaciones que utiliza un procedimiento de mutagénesis dirigida al sitio tal como Mutant-K (fabricado por TaKaRa Bio, Inc.), Mutant-G (fabricado por TaKaRa Bio, Inc.), o un kit de series de mutagénesis in vitro LA PCR (fabricado por TaKaRa Bio, Inc.) o similares. También se puede utilizar dicho polinucleótido mutado como el elemento de ADN de la invención.

40 Como el elemento de ADN para utilizarse en combinación con el promotor de la invención, se puede utilizar un fragmento parcial que comprende al menos 3000 o al menos 2000 nucleótidos consecutivos de una secuencia de nucleótidos representada por una cualquiera de las SEQ ID NOS: 10 a 14 en el listado de secuencias. Los ejemplos de dicho fragmento parcial incluyen: A2-1 a A2-17 que son fragmentos parciales de A2; A7-1 a A7-18 que son fragmentos parciales de A7; A18-1 a A18-4 que son fragmentos parciales de A18; B5-1 a B5-6 que son fragmentos parciales de B5; y C14-1 a C14-14 que son fragmentos parciales de C14. Sin embargo, el elemento de ADN no se limita a estos fragmentos parciales siempre que tenga la actividad de mejorar la expresión de genes extraños.

45 En la invención, uno cualquiera de los fragmentos parciales anteriores puede utilizarse solo y también pueden utilizarse dos o más copias de un tipo del fragmento parcial. Como alternativa, pueden utilizarse en combinación dos o más tipos diferentes de los fragmentos parciales. Además, pueden utilizarse en combinación una secuencia de longitud completa y un fragmento parcial de cualquiera de los elementos de ADN mencionados anteriormente. En la combinación anterior, la secuencia de longitud completa y el fragmento parcial pueden proceder del mismo elemento de ADN o de diferentes elementos de ADN.

55 En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de A2, A2-1 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 3000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-2 corresponde a la secuencia polinucleotídica de los nucleótidos 2801 a 5800 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-3 corresponde a la secuencia polinucleotídica de los nucleótidos 5401 a 8450 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-4 corresponde a la secuencia polinucleotídica de los nucleótidos 701 a 2700 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-5 corresponde a la secuencia polinucleotídica de los nucleótidos

ID NO: 14 en el Listado de Secuencias; C14-11 corresponde a la secuencia polinucleotídica de los nucleótidos 1987 a 7119 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-12 corresponde a la secuencia polinucleotídica de los nucleótidos 2994 a 8141 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-13 corresponde a la secuencia polinucleotídica de los nucleótidos 960 a 7119 de la SEQ ID NO: 14 en el Listado de Secuencias; y C14-14 corresponde a la secuencia polinucleotídica de los nucleótidos 1987 a 8141 de la SEQ ID NO: 14 en el listado de secuencias.

4. Adquisición de polinucleótidos

En la invención, un polinucleótido que contiene un gen extraño que codifica una proteína extraña, cuya producción debe ser aumentada, que se describirá más adelante, se puede obtener mediante procedimientos comunes como se describe a continuación. Por ejemplo, se puede aislar dicho polinucleótido mediante la identificación de una biblioteca de ADNc que procede de células o tejidos que expresan el gen extraño utilizando una sonda de ADN sintetizada a partir de un fragmento del gen extraño. El ARNm para los mismos se puede preparar mediante procedimientos utilizados normalmente en la técnica. Por ejemplo, se tratan células o tejidos con un reactivo de guanidina, un reactivo de fenol, etc., obteniendo, de este modo, ARN total y, a continuación, se obtiene poli (A) + ARN (ARNm) mediante un procedimiento de columna de afinidad utilizando una columna de oligo (dT) celulosa o una columna de poli U-Sepharose que contiene Sepharose 2B o similar, como un vehículo o mediante un procedimiento discontinuo. Además, el poli(A) + ARN puede fraccionarse adicionalmente mediante centrifugación en gradiente de densidad de sacarosa o similar. Después, se sintetiza ADNc de cadena simple utilizando el ARNm así obtenido como un molde, cebadores oligo (dT) y una transcriptasa inversa. A partir del ADNc de cadena simple así obtenido, se sintetiza ADNc de doble cadena utilizando ADN polimerasa I, ADN ligasa, RNasa H y similares. El ADNc de doble cadena así sintetizado se recorta utilizando la ADN polimerasa T4, seguido de ligadura a un adaptador (tal como el adaptador de EcoRI), fosforilación y similares, y el ADN resultante se incorpora en un fago lambda tal como λ gt11 para lograr el empaquetado in vivo, con lo que se prepara una biblioteca de ADNc. También es posible preparar una biblioteca de ADNc utilizando un vector plasmídico en lugar del fago lambda. A continuación, puede seleccionarse un clon que contiene el ADN diana (un clon positivo) a partir de la biblioteca de ADNc.

En los casos en los que el promotor mencionado anteriormente, un polinucleótido que contiene una región de terminación, el elemento de ADN mencionado anteriormente o un polinucleótido que contiene un gen extraño a utilizar para producir una proteína, se aísla del ADN genómico, de acuerdo con un procedimiento común (Molecular Cloning (1989), Methods in Enzymology 194 (1991)), se extrae ADN genómico de una línea celular de un organismo para utilizarse como una fuente de recogida, y se selecciona y aísla el polinucleótido. La extracción de ADN genómico se puede realizar de acuerdo con, por ejemplo, el procedimiento de Cryer y col. (Methods in Cell Biology, 12, 39-44 (1975)) o el procedimiento de P. Philippsen y col. (Methods Enzymol., 194, 169-182 (1991)).

El promotor diana, elemento de ADN o polinucleótido que contiene un gen extraño también se puede obtener mediante, por ejemplo, el procedimiento de PCR (PCR Technology. Henry A. Erlich, Atockton press (1989)). En la amplificación de un polinucleótido utilizando el procedimiento de PCR, se utilizan ADN de cadena simple sintéticos de 20 a 30 meros como cebadores y se utiliza ADN genómico como un molde. El gen amplificado se utiliza después de que se confirma la secuencia polinucleotídica del gen. Como el molde para PCR, se puede utilizar una biblioteca de ADN genómico tal como una biblioteca de cromosomas artificiales de bacteria (BAC).

Por otra parte, se puede obtener el polinucleótido que contiene un gen extraño cuya secuencia no se conoce mediante (a) la preparación de una biblioteca de genes de acuerdo con un procedimiento común y (b) la selección de un polinucleótido deseado a partir de la biblioteca de genes preparada y la amplificación del polinucleótido. Se puede preparar la biblioteca de genes mediante la digestión parcial del ADN cromosómico obtenido mediante un procedimiento común a partir de una línea celular de un organismo para utilizarse como una fuente de recogida utilizando una enzima de restricción apropiada para fragmentar el ADN cromosómico, uniendo los fragmentos obtenidos a un vector apropiado e introduciendo el vector en un hospedador apropiado. También se puede preparar la biblioteca de genes mediante la extracción de ARNm de las células, la síntesis de ADNc a partir del ARNm, la unión del ADNc a un vector apropiado y la introducción del vector en un hospedador apropiado. Como el vector para utilizarse en dicha preparación, se puede utilizar un plásmido normalmente conocido como un vector para la preparación de la biblioteca de genes y también se pueden utilizar un vector fago, un cósmido o similares. Como el hospedador para transformarse o transfectarse, puede utilizarse un hospedador adecuado para el tipo de vector mencionado anteriormente. El polinucleótido que contiene el gen extraño se selecciona a partir de la biblioteca de genes mencionada anteriormente mediante un procedimiento de hibridación de colonias, un procedimiento de hibridación de placas o similares utilizando una sonda marcada que contiene una secuencia específica para el gen extraño.

Además, el polinucleótido que contiene el gen extraño también puede producirse mediante síntesis química total. Por ejemplo, se puede sintetizar el gen mediante un procedimiento en el que se preparan y alinean dos pares de oligonucleótidos complementarios, un procedimiento en el que se ligan varias cadenas de ADN hibridadas mediante una ADN ligasa, un procedimiento en el que se preparan diversos oligonucleótidos parcialmente complementarios y se rellenan espacios mediante PCR o similares.

La determinación de una secuencia polinucleotídica se puede realizar mediante una técnica convencional, por ejemplo, un procedimiento dideoxi (Sanger y col., Proc. Natl. Acad. Sci., USA, 74, 5463-5467 (1977)), o similar. Además, la anterior determinación de una secuencia polinucleotídica también se puede realizar fácilmente utilizando un kit de secuenciación comercialmente disponible o similares.

5 5. Vector de expresión de genes extraños, Vector de elemento

Como un vector de expresión de genes extraños de la invención, se proporciona un vector que contiene la unidad de expresión de genes extraños descrita en el anterior punto "2" que contiene el promotor de la invención descrito en el anterior punto "1". El vector de expresión de genes extraños de la invención puede contener, opcionalmente, un tipo de los elementos de ADN descritos en el anterior punto "3", dos o más copias de uno de los tipos de los elementos de ADN mencionados anteriormente o dos o más tipos diferentes de los elementos de ADN mencionados anteriormente en combinación. Cuando se expresa un gen extraño en una célula huésped utilizando el vector de expresión de genes extraños mencionado anteriormente, el elemento de ADN puede localizarse inmediatamente cadena arriba o cadena abajo de la unidad de expresión de genes o puede localizarse en una posición lejos de la unidad de expresión de genes. Además, puede utilizarse un vector de expresión de genes extraños que contiene una pluralidad de dichos elementos de ADN. En relación con esto, el elemento de ADN se puede insertar en orientación directa o inversa con respecto a la unidad de expresión de genes.

Además, como el vector para utilizarse en la invención, también se incluye un vector que contiene un tipo de los elementos de ADN mencionados anteriormente, dos o más copias de un tipo de los elementos mencionados anteriormente o dos o más tipos diferentes de los elementos de ADN mencionados anteriormente en combinación y que no contiene unidad de expresión de genes (en lo sucesivo en el presente documento denominado "vector de elemento"). Dicho vector de elemento se puede utilizar en combinación con el vector de expresión de genes extraños mencionado anteriormente que contiene el elemento de ADN o un vector de expresión de genes extraños que no contiene el elemento de ADN y que contiene solo la unidad de expresión de genes extraños. Al permitir que el vector de elemento coexista con el vector de expresión de genes extraños, la expresión del gen extraño se mejora en comparación con los casos en los que se utiliza el vector de expresión de genes extraños solo y, por lo tanto, la combinación de los vectores mencionados anteriormente también se incluye dentro del vector de expresión de genes extraños de la invención.

El gen extraño no está particularmente limitado, pero los ejemplos del mismo incluyen genes indicadores tales como los genes de fosfatasa alcalina secretora (SEAP), una proteína fluorescente verde (GFT) y luciferasa; diversos genes enzimáticos tales como un gen α -amilasa y un gen α -galactosidasa; genes de diversos interferones que son proteínas farmacéuticamente útiles y fisiológicamente activas tales como interferón α e interferón γ ; genes de diversas interleucinas tales como IL-1 e IL-2; diversos genes de citocina tales como un gen de eritropoyetina (EPO) y un gen del factor estimulante de colonias de granulocitos (G-CSF); un gen del factor de crecimiento; y un gen que codifica una proteína multimérica tal como un gen que codifica un heteromultímero que es un anticuerpo o un fragmento funcional del mismo. Estos genes pueden obtenerse por cualquier procedimiento.

El "fragmento funcional de un anticuerpo" se refiere a un fragmento parcial de un anticuerpo que tiene una actividad de unión a antígeno e incluye Fab, F(ab')₂, Fv, scFv, diacuerpos, anticuerpos lineales, anticuerpos poliespecíficos formados a partir de fragmentos de anticuerpos y similares. El fragmento funcional de un anticuerpo también incluye Fab' que es un fragmento monovalente en una región variable de un anticuerpo obtenido tratando F(ab')₂ en condiciones reductoras. Sin embargo, el fragmento funcional no está limitado a estas moléculas siempre que el fragmento tenga una afinidad de unión por un antígeno. Además, estos fragmentos funcionales incluyen no solamente un fragmento obtenido tratando una molécula de longitud completa de una proteína de anticuerpo con una enzima apropiada, sino también una proteína producida en una célula huésped apropiada utilizando un gen de anticuerpo genéticamente modificado.

Además, el vector de expresión de genes extraños y el vector de elemento de la invención pueden contener cada uno un marcador de selección para seleccionar un transformante. Mediante la utilización de, por ejemplo, un marcador resistente a antibiótico que imparte resistencia a un antibiótico tal como cerulenina, aureobasidina, Zeocina, canavanina, cicloheximida, higromicina, puromicina, blastidina, tetraciclina, kanamicina, ampicilina o neomicina, se puede seleccionar un transformante. Además, también se puede seleccionar un transformante cuando se utiliza como un marcador un gen que imparte resistencia a un disolvente tal como etanol, resistencia a la presión osmótica de glicerol, una sal o similares, resistencia a un ion metálico tal como un ion de cobre o similares.

El vector de expresión de genes extraños y el vector de elemento de la invención pueden ser cada uno un vector que no se incorpora en el ADN cromosómico. En general, el vector de expresión de genes extraños transfecta a una célula huésped y, por lo tanto, se incorpora aleatoriamente en el cromosoma. Sin embargo, mediante la utilización de un componente constituyente que procede de un virus de mamífero tal como el virus 40 de simio (SV40), un papilomavirus (VPB, VPH) o VEB, el vector se puede utilizar como un vector episómico que es autorreplicable en la célula huésped transfectada. Por ejemplo, se utilizan ampliamente un vector que contiene un origen de replicación que procede de SV40 y una secuencia que codifica un antígeno T grande de SV40 que es un factor de acción trans, un vector que contiene un oriP que procede de VEB y una secuencia que codifica EBNA-1 y similares. El elemento de ADN puede exhibir eficazmente la actividad de mejorar la expresión de genes extraños independientemente del

tipo de vector o la presencia o ausencia de incorporación del mismo en el cromosoma.

6. Célula transformada

La célula transformada de la invención es una célula transformada en la que se ha introducido el vector de expresión de genes extraños descrito en el anterior punto "5". Como el vector de expresión de genes extraños, (A) puede introducirse solo un vector de expresión de genes extraños que no contiene elemento de ADN o (B) puede introducirse en combinación un vector de expresión de genes extraños que no contiene elemento de ADN y un vector de elemento. Como alternativa, (C) puede introducirse un vector de expresión de genes extraños que contiene un elemento de ADN o (D) puede introducirse en combinación un vector de expresión de genes extraños que contiene un elemento de ADN y un vector de elemento.

La expresión de un gen extraño en una célula huésped mediante la combinación descrita en el anterior (B) o (D) se puede realizar de acuerdo con, por ejemplo, el procedimiento de Girod y col. (Biotechnology and Bioengineering, 91, 2-11 (2005)) y el procedimiento de Otte y col. (Biotechnol. Prog., 2007, 23, 801-807 (2007)).

Los ejemplos de la célula huésped para ser transformada incluyen una célula eucariota, los ejemplos preferentes de la misma incluyen una célula de mamífero y los ejemplos más preferentes incluyen una célula que procede de seres humanos, ratones, ratas, hámsteres, monos o ganado. Los ejemplos de dicha célula de mamífero incluyen una célula COS-1, una célula 293 y una célula CHO (CHO-K1, DG44, CHO dhfr-, CHO-S), pero la célula huésped no se limita a estas.

En la invención, puede utilizarse cualquier procedimiento para introducir el vector de expresión en la célula huésped siempre que el procedimiento permita que el gen introducido esté presente de manera estable en la célula huésped y se exprese adecuadamente en la misma. Los ejemplos del procedimiento que se utiliza normalmente incluyen un procedimiento de fosfato de calcio (Ito y col., (1984) Agric. Biol. Chem., 48, 341), un procedimiento de electroporación (Becker, D. M. y col., 1990; Methods. Enzymol., 194, 182-187), un procedimiento de esferoplastos (Creggh y col., Mol. Cell. Biol., 5, 3376 (1985)), un procedimiento de acetato de litio (Ito, H. (1983) J. Bacteriol. 153, 163-168), y un procedimiento de lipofección.

7. Procedimiento para la producción de proteínas extrañas

En la invención, se puede producir una proteína extraña mediante el cultivo de la célula transformada descrita en el anterior punto "6", en la que se ha introducido un gen que codifica la proteína extraña, mediante un procedimiento conocido, la recolección de la proteína del producto de cultivo resultante, seguida de la purificación de la proteína. La expresión "producto de cultivo" como se utiliza en el presente documento se refiere a células cultivadas o un homogeneizado de células además de un sobrenadante de cultivo. En relación con esto, como la proteína extraña que se puede producir utilizando la célula transformada descrita en el anterior punto "6", se puede seleccionar no solo una proteína monomérica, sino también una proteína multimérica. En los casos en los que se produce una proteína heteromultimérica formada de una pluralidad de diferentes subunidades, es necesario introducir una pluralidad de genes que codifican estas subunidades en la célula huésped descrita en el anterior punto "6", respectivamente.

El procedimiento para cultivar la célula transformada se puede realizar de acuerdo con procedimientos convencionales para el cultivo de células huésped.

En los casos en los que la célula transformada es una célula de mamífero, se cultiva la célula en condiciones de, por ejemplo, 37 °C y CO₂ al 5 % u 8 % durante un tiempo de cultivo de aproximadamente 24 a 1000 horas. El cultivo se puede realizar a través de un cultivo discontinuo, un cultivo discontinuo alimentado, un cultivo continuo o similares en condiciones estáticas, de sacudida, de agitación o de aireación.

Se puede realizar la confirmación del producto de expresión del gen que codifica la proteína extraña a partir del producto de cultivo mencionado anteriormente (solución de cultivo) mediante SDS-PAGE, un análisis Western, ELISA o similares. Para aislar y purificar la proteína producida, puede utilizarse un procedimiento convencional de aislamiento y purificación de proteínas. Después de completarse el cultivo, en los casos en los que se produce la proteína diana en las células, se homogeneizan las células utilizando un homogeneizador ultrasónico, una prensa francesa, un homogeneizador Manton-Gaulin, DYNOMILL o similares, obteniendo de este modo la proteína diana. Además, en los casos en los que se produce la proteína diana fuera de las células, la solución de cultivo se utiliza como tal o se eliminan las células mediante centrifugación o similar. A continuación, se recoge la proteína diana mediante extracción o similar utilizando un disolvente orgánico y, a continuación, puede aislarse y purificarse la proteína diana recogida mediante la utilización de técnicas tales como diversas técnicas de cromatografía (cromatografía hidrófoba, cromatografía de fase inversa, cromatografía de afinidad, cromatografía de intercambio iónico, etc.), filtración en gel utilizando un tamiz molecular o electroforesis utilizando un gel de poliacrilamida o similares, solas o en combinación de acuerdo con las necesidades.

Los procedimientos de cultivo y procedimientos de purificación mencionados anteriormente son solo ejemplos y los procedimientos no se limitan a ellos. La secuencia de aminoácidos del producto génico purificado se puede confirmar mediante una técnica de análisis de aminoácidos conocida tal como secuenciación de aminoácidos

automatizada utilizando el procedimiento de degradación de Edman.

8. Procedimiento para la producción de proteínas de anticuerpo

Se puede ejemplificar como la proteína heteromultimérica a producir utilizando el procedimiento de producción descrito en el anterior punto "7", una proteína de anticuerpo. La proteína de anticuerpo es una proteína tetramérica que comprende dos moléculas de polipéptidos de cadena pesada y dos moléculas de polipéptidos de cadena ligera. Por consiguiente, para obtener dicha proteína de anticuerpo en un estado de mantenimiento de una afinidad de unión a antígeno, es necesario introducir los genes de cadena pesada y ligera en la célula transformada descrita en el anterior punto "6". En este caso, las unidades de expresión de genes de cadena pesada y ligera pueden presentarse en el mismo vector de expresión o diferentes vectores de expresión.

Se puede ejemplificar como el anticuerpo para producirse en realizaciones de la invención, un anticuerpo preparado mediante inmunización de un animal experimental tal como un conejo, un ratón o una rata, con un antígeno deseado. Además, también se puede ejemplificar como el anticuerpo para producirse en la invención, un anticuerpo quimérico y un anticuerpo humanizado obtenidos mediante la utilización del anticuerpo mencionado anteriormente como material de inicio. Además, también se incluye un anticuerpo humano obtenido utilizando un animal modificado genéticamente o un procedimiento de presentación en fagos, en el anticuerpo a producir en realizaciones de la invención.

El gen de anticuerpo a utilizar para la producción del anticuerpo no se limita a un gen de anticuerpo que tiene una secuencia polinucleotídica específica siempre que la combinación del polinucleótido de cadena pesada y el polinucleótido de cadena ligera para transcribirse y traducirse a partir del gen de anticuerpo tengan la actividad de unión a una proteína de antígeno dada.

Además, no es necesario que el gen de anticuerpo codifique la molécula de longitud completa del anticuerpo y se puede utilizar un gen que codifica un fragmento funcional del anticuerpo. Se puede obtener dicho gen que codifica un fragmento funcional del mismo mediante modificación genética de un gen que codifica la molécula de longitud completa de una proteína de anticuerpo.

9. Procedimiento de producción para otras proteínas extrañas

Los ejemplos de la proteína extraña a producir utilizando el procedimiento de producción de la invención incluyen, además de los anticuerpos mencionados anteriormente, diversas proteínas que proceden de seres humanos o no humanos, fragmentos funcionales de las mismas y productos modificados de las mismas. Los ejemplos de dichas proteínas y similares incluyen hormonas peptídicas tales como péptido natriurético auricular (ANP), péptido natriurético cerebral (BNP), péptido natriurético de tipo C (CNP), vasopresina, somatostatina, hormona del crecimiento (GH), insulina, oxitocina, ghrelina, leptina, adiponectina, renina, calcitonina, osteoprotegerina y factor de crecimiento similar a la insulina (IGF); citoquinas tales como interleuquina, quimiocina, interferón, factores de necrosis tumoral (tal como TNF- α , TNF- β y la superfamilia de TNF), factores de crecimiento nervioso (tales como NGF), factores de crecimiento celular (tales como EGF, FGF, PDGF, HGF y TGF), factores de crecimiento hematopoyéticos (tales como CSF, G-CSF y eritropoyetina) y adipocina; receptores tales como receptores de TNF; enzimas tales como lisozima, proteasa, proteinasa y peptidasa; fragmentos funcionales de las mismas (fragmentos que tienen parte o toda la actividad biológica de la proteína original) y proteínas de fusión que comprenden cualquiera de estas proteínas. Sin embargo, las proteínas no se limitan a éstas.

Ejemplos

10. Ejemplos

En lo sucesivo en el presente documento, la invención se describirá específicamente con referencia a los ejemplos. Sin embargo, estos ejemplos no limitan el ámbito técnico de la invención. Son productos comercialmente disponibles y se pueden utilizar de acuerdo con procedimientos comunes, los plásmidos, enzimas de restricción, enzimas de modificación de ADN y similares para utilizarse en los ejemplos de la invención. Además, también son bien conocidos por los expertos en la materia o se pueden encontrar en las referencias los procedimientos utilizados para clonar ADN, secuenciar polinucleótidos, transformar una célula huésped, cultivar una célula huésped transformada, recoger una proteína del producto de cultivo resultante, purificar una proteína y similares.

(Ejemplo 1) Construcción del vector CMV/pSeapIRESpuo para su uso en evaluación de actividad promotora

La evaluación de actividad promotora se realizó mediante la utilización de la expresión de SEAP como un índice y se construyó un vector para su uso en la evaluación.

1-1) Amplificación de ADNc de SEAP mediante PCR y adición de sitio de enzima de restricción

Se amplificó el ADNc de SEAP mediante PCR utilizando cebadores en los que se añadió un sitio NheI inmediatamente cadena arriba del codón de inicio ATG y se añadió un sitio BglII inmediatamente cadena abajo del codón de terminación y KOD -Plus- (TOYOBO). Como molde, se utilizó el control pSEAP2 (Clontech). Se digirió el

fragmento obtenido con NheI y BglII y, a continuación, se purificó utilizando un kit de reacción MinElute (Qiagen).

Los cebadores utilizados:

SEAPF: AAAGCTAGCATGCTGCTGCTGCTGCTGCTGCTGGGCC

SEAPR: AAAAGATCTTCATGTCTGCTCGAAGCGGCCGCGC

5 1-2) Construcción de CMV/pSeapIRESpuro

Después de digerirse un vector pIRESpuro3 (Clontech) con NheI y BamHI, se integró el fragmento de SEAP preparado en 1-1) en el mismo mediante una reacción de ligadura. El plásmido obtenido se nombró "CMV/pSeapIRESpuro".

(Ejemplo 2) Clonación de regiones promotoras de RPS7, RPL32 y RPL34

10 Se seleccionaron como genes humanos que se consideraba que contenían un promotor que tiene una alta actividad transcripcional, EEF2, YBX1, PPIA, PSAP, RAN, RPL32, RPL34, RPLP1, RPS7, RPS24, TMSB4X, UBC, YWHAE, ARPC2 y SERBP1 mediante la utilización, como un índice, del nivel de ARNm y se realizó la clonación de la región promotora de cada gen. Se utilizaron los plásmidos obtenidos para la evaluación de la actividad promotora en el ejemplo 3.

15 2-1) Clonación de la región promotora de RPS7

Se utilizó como la región promotora de RPS7, con referencia a la secuencia de ARNm registrada con el número de referencia NM_001011.3 en GenBank, una secuencia que empieza en un nucleótido localizado aproximadamente 2 kpb cadena arriba del sitio de inicio de la transcripción y termina en un nucleótido inmediatamente cadena arriba de la secuencia de nucleótidos que corresponde a la secuencia del codón de inicio de RPS7.

20 Se amplificó la región promotora de RPS7 mediante PCR utilizando el clon RP11-644P19 del cromosoma artificial de E. coli (GenoTechs) como un molde y también utilizando el siguiente conjunto de cebadores y KOD -Plus- (TOYOBO) y, a continuación, se purificó utilizando el kit de reacción MinElute (Qiagen). Después se digirió CMV/pSeapIRESpuro con SpeI y NheI y se eliminó el promotor de CMV, se integró la región promotora de RPS7 al sitio SpeI-NheI utilizando un kit de clonación de la PCR In-Fusion Advantage (Clontech), con lo que se construyó RPS7/pSeapIRESpuro. La secuencia de nucleótidos de la región promotora clonada de RPS7 se representa mediante la SEQ ID NO: 1 en el listado de secuencias.

Conjunto de cebadores para RPS7:

RPS7-F: TTGATTATTGACTAGTATTTATGTATATTAACAGCACATTAACAGC

RPS7-R: GCAGCAGCATGCTAGCGGCTTTCTCCTGGGAGAACTGAAGGCACAGCGG

30 2-2) Clonación de la región promotora de RPL32

Como la región promotora de RPL32, con referencia a la secuencia de ARNm registrada en el número de referencia NM_000994.3 en GenBank, se utilizó una secuencia que empieza en un nucleótido localizado aproximadamente 2 kpb cadena arriba del sitio de inicio de la transcripción y termina en el nucleótido inmediatamente cadena arriba de la secuencia de nucleótidos que corresponde a la secuencia del codón de inicio de RPL32.

35 Se amplificó la región promotora de RPL32 mediante PCR utilizando el clon RP11-767C1 del cromosoma artificial de E. coli (GenoTechs) como un molde y también utilizando el siguiente conjunto de cebadores y KOD -Plus- (TOYOBO) y, a continuación, se purificó utilizando un kit de reacción MinElute (Qiagen). Después se digirió CMV/pSeapIRESpuro con SpeI y NheI y se eliminó el promotor de CMV, se integró la región promotora de RPL32 al sitio SpeI-NheI utilizando un kit de clonación de la PCR In-Fusion Advantage (Clontech), con lo que se construyó RPL32/pSeapIRESpuro. La secuencia de nucleótidos de la región promotora clonada de RPL32 se representa mediante la SEQ ID NO: 2 en el listado de secuencias.

Conjunto de cebadores para RPL32:

RPL32-F: TTGATTATTGACTAGTCTAAAGTGATTCCTAAAGAATTCTTCCC

RPL32-R: GCAGCAGCATGCTAGCGATGCCTTTTGGGGAAGAAGCGGCCCC

45 2-3) Clonación de la región promotora de RPL34

Como la región promotora de RPL34, con referencia a la secuencia de ARNm registrada en el número de referencia NM_033625.2 en GenBank, se utilizó una secuencia que empieza en un nucleótido localizado aproximadamente 2 kpb cadena arriba del sitio de inicio de la transcripción y termina en el nucleótido inmediatamente cadena arriba de la secuencia de nucleótidos que corresponde a la secuencia del codón de inicio de RPL34.

50

Se amplificó la región promotora de RPL34 mediante PCR utilizando el clon RP11-462C24 del cromosoma artificial de *E. coli* (GenoTechs) como un molde y también utilizando el siguiente conjunto de cebadores y KOD -Plus- (TOYOBO) y, a continuación, se purificó utilizando un kit de reacción MinElute (Qiagen). Después se digirió CMV/pSeapIRESpuo con SpeI y NheI y se eliminó el promotor de CMV, se integró la región promotora de RPL34 al sitio SpeI-NheI utilizando un kit de clonación de la PCR In-Fusion Advantage (Clontech), con lo que se construyó RPL34/pSeapIRESpuo. La secuencia de nucleótidos de la región promotora clonada de RPL34 se representa mediante la SEQ ID NO: 3 en el listado de secuencias.

Conjunto de cebadores para RPL34:

RPL34-F: TTGATTATTGACTAGTATGGTGGCACAATCATGGTTCCTACTGCAGCC
RPL34-R: GCAGCAGCATGCTAGCTCTGAGTGCCTAAATTAAGAATAGAGTAACATC

2-4) Clonación de regiones promotoras de otros genes humanos

La clonación de cada una de las regiones promotoras de EEF2, YBX1, PPIA, PSAP, RAN, RPLP1, RPS24, TMSB4X, UBC, YWHA, ARPC2 y SERBP1 se realizó de acuerdo con el procedimiento descrito en el punto anterior 2-1) mediante el cual se construyó pSeapIRESpuo que contiene el polinucleótido clonado.

(Ejemplo 3) Evaluación de actividad promotora utilizando el nivel de expresión de SEAP en células policlonales CHO-K1 transfectadas como índice 3-1) Transfección

Se subcultivaron células CHO-K1 (ATCC) en CO₂ al 5 % a 37 °C utilizando el medio de mezcla de nutrientes F-12 (GIBCO) que contiene FBS de IgG ultrabaja (GIBCO) al 10 %.

Se sembraron las células CHO-K1 en una placa de 6 pocillos (IWAKI) con 5 x 10⁵ células/pocillo. Al día siguiente, se introdujeron por transfección 2 µg de cada CMV/pSeapIRESpuo, RPS7/pSeapIRESpuo, RPL32/pSeapIRESpuo, RPL34/pSeapIRESpuo o similares construidos en los ejemplos 1) y 2) utilizando Lipofectamine 2000 (Invitrogen).

3-2) Selección de antibiótico con puromicina

Dos días después de la transfección, se recogieron las células de la placa de 6 pocillos mediante un tratamiento de tripsina, se sembró la cantidad total de las células recogidas en una placa de 6 cm (Nunc) y se añadió también puromicina (Clontech) al medio a una concentración final de 8 µg/ml para iniciar la selección con antibióticos.

3-3) Evaluación utilizando una línea celular policlonal transfectada

Después de 11 días del inicio de la selección de antibióticos, se recogió la línea celular policlonal transfectada con tripsina y se contó el número de células. Después, se sembraron las células en una placa de 24 pocillos (IWAKI) con 1 x 10⁵ células/ml/pocillo. Después de 24 horas, se recogió el sobrenadante del cultivo y se midió la actividad de SEAP en el sobrenadante del cultivo utilizando el ensayo indicador de fosfatasa alcalina secretada pNPP SensoLyte™ (ANASPEC). La actividad de SEAP fue mayor bajo el control de cada una de las regiones promotoras de RPS7, RPL32 y RPL34 que bajo el control del promotor de CMV (CMV/pSeapIRESpuo) que sirve como el control, y la actividad de SEAP fue 1,7 veces o más, 2,0 veces o más y 2,5 veces o más mayor que la del control, respectivamente (figura 1). Paralelamente, la actividad de SEAP fue menor bajo el control de cada una de las regiones promotoras de EEF2, YBX1, PPIA, PSAP, RAN, RPLP1, RPS24, TMSB4X, UBC, YWHA, ARPC2 y SERBP1 que bajo el control del promotor de CMV.

(Ejemplo 4) Clonación de un promotor truncado

Se realizó la clonación de los promotores truncados utilizando como los promotores truncados de RPS7, RPL32 y RPL34, una secuencia de nucleótidos (T1) que empieza en un nucleótido localizado aproximadamente 1 kpb cadena arriba del sitio de inicio de la transcripción y termina en el nucleótido inmediatamente cadena arriba de la secuencia de nucleótidos que corresponde al codón de inicio de cada gen, y una secuencia de nucleótidos (T2) que empieza en un nucleótido localizado aproximadamente 0,5 kpb cadena arriba del sitio de inicio de la transcripción y termina en el nucleótido inmediatamente cadena arriba de la secuencia de nucleótidos que corresponde al codón de inicio de cada gen.

4-1) Clonación de RPS7T1 y RPS7T2

Se amplificaron RPS7T1 y RPS7T2 mediante PCR utilizando RPS7/pSeapIRESpuo construido en 2-1) como un molde y utilizando también el siguiente conjunto de cebadores y KOD -Plus- (TOYOBO) y, a continuación, se purificaron utilizando un kit de reacción MinElute (Qiagen). Después se digirió CMV/pSeapIRESpuo con SpeI y NheI y se eliminó el promotor de CMV, se integró cada una de las regiones promotoras de RPS7T1 y RPS7T2 al sitio SpeI-NheI utilizando un kit de clonación de la PCR de In-Fusion Advantage (Clontech), con el que se construyeron RPS7T1/pSeapIRESpuo y RPS7T2/pSeapIRESpuo. Las secuencias de nucleótidos de las regiones promotoras clonadas de RPS7T1 y RPS7T2 se representan mediante las SEQ ID NOS: 4 y 5 en el Listado de secuencias, respectivamente.

Conjunto de cebadores para RPS7T1

RPS7-T1: TTGATTATTGACTAGTCCTAGTGTGGCTTCTGCATTTTTTCACAGTGC
RPS7-R: GCAGCAGCATGCTAGCGGCTTTCTCCTGGGAGAACTGAAGGCACAGCGG

Conjunto de cebadores para RPS7T2

5 RPS7-T2: TTGATTATTGACTAGTCCTCGGCTCACGGCAGCCTCGACCTTTTCGGC
RPS7-R: GCAGCAGCATGCTAGCGGCTTTCTCCTGGGAGAACTGAAGGCACAGCGG

4-2) Clonación de RPL32T1 y RPL32T2

10 Se amplificaron RPL32T1 y RPL32T2 mediante PCR utilizando RPL32/pSeapIRESpuro construido en 2-2) como un molde y utilizando también el siguiente conjunto de cebadores y KOD -Plus- (TOYOBO) y, a continuación, se purificaron utilizando un kit de reacción MinElute (Qiagen). Después se digirió CMV/pSeapIRESpuro con SpeI y NheI y se eliminó el promotor de CMV, se integró cada una de las regiones promotoras de RPL32T1 y RPL32T2 al sitio SpeI-NheI utilizando un kit de clonación de la PCR de In-Fusion Advantage (Clontech), con el que se construyeron RPL32T1/pSeapIRESpuro y RPL32T2/pSeapIRESpuro. Las secuencias de nucleótidos de las regiones promotoras clonadas de RPL32T1 y RPL32T2 se representan mediante las SEQ ID NOS: 6 y 7 en el Listado de secuencias, respectivamente.

Conjunto de cebadores para RPL32T1

RPL32T1: TTGATTATTGACTAGTCCTCTCGAGTAACTGGGACTACAGGCATGC
RPL32-R: GCAGCAGCATGCTAGCGATGCCTTTTGGGGAAGAAGCGGCCCC

Conjunto de cebadores para RPL32T2

20 RPL32T2: TTGATTATTGACTAGTGCAGTTTCGCCAGTGGTTAGAAGCGTGG
RPL32-R: GCAGCAGCATGCTAGCGATGCCTTTTGGGGAAGAAGCGGCCCC

4-3) Clonación de RPL34T1 y RPL34T2

25 Se amplificaron RPL34T1 y RPL34T2 mediante PCR utilizando RPL34/pSeapIRESpuro construido en 2-3) como un molde y utilizando también el siguiente conjunto de cebadores y KOD -Plus- (TOYOBO) y, a continuación, se purificaron utilizando un kit de reacción MinElute (Qiagen). Después se digirió CMV/pSeapIRESpuro con SpeI y NheI y se eliminó el promotor de CMV, se integró cada una de las regiones promotoras de RPL34T1 y RPL34T2 al sitio SpeI-NheI utilizando un kit de clonación de la PCR de In-Fusion Advantage (Clontech), con el que se construyeron RPL34T1/pSeapIRESpuro y RPL34T2/pSeapIRESpuro. Las secuencias de nucleótidos de las regiones promotoras clonadas de RPL34T1 y RPL34T2 se representan mediante las SEQ ID NOS: 8 y 9 en el Listado de secuencias, respectivamente.

Conjunto de cebadores para RPL34T1

RPL34T1: TTGATTATTGACTAGTCTTCCTGGAGGTGCATTCTAAGAGCGCTCCCC
RPL34-R: GCAGCAGCATGCTAGCTCTGAGTGCCTAAATTAAGAAATAGAGTAACATC

Conjunto de cebadores para RPL34T2

35 RPL34T2: TTGATTATTGACTAGTGTAAGCTTGTGCTCTGAATAAATGACAAGG
RPL34-R: GCAGCAGCATGCTAGCTCTGAGTGCCTAAATTAAGAAATAGAGTAACATC

(Ejemplo 5) Evaluación de actividad de promotor truncado utilizando el nivel de expresión de SEAP en células policlonales CHO-K1 transfectadas como índice

5-1) Transfección

40 Se subcultivaron células CHO-K1 (ATCC) en CO₂ al 5 % a 37 °C utilizando el medio de mezcla de nutrientes F-12 (GIBCO) que contiene FBS de IgG ultrabaja (GIBCO) al 10 %.

45 Se sembraron las células CHO-K1 en una placa de 6 pocillos (IWAKI) con 2 x 10⁵ células/pocillo. Al día siguiente, se introdujeron por transfección 2 µg de cada CMV/pSeapIRESpuro, RPS7/pSeapIRESpuro, RPS7T1/pSeapIRESpuro, RPS7T2/pSeapIRESpuro, RPL32/pSeapIRESpuro, RPL32T1/pSeapIRESpuro, RPL32T2/pSeapIRESpuro, RPL34/pSeapIRESpuro, RPL34T1/pSeapIRESpuro y RPL34T2/pSeapIRESpuro construidos en los ejemplos 1), 2) y 4) utilizando Fugene 6 (Roche Applied Science).

5-2) Selección de antibiótico con puromicina

Dos días después de la transfección, se recogieron las células de la placa de 6 pocillos mediante un tratamiento con tripsina y se sembró la cantidad total de células recogidas en una placa de 6 cm (Nunc) y se añadió también

puromicina al medio a una concentración final de 8 µg/ml para iniciar la selección de antibióticos.

5-3) Evaluación utilizando una línea celular policlonal transfectada

Después de 11 días del inicio de la selección de antibióticos, se recogió cada línea celular policlonal transfectada con tripsina y se contó el número de células. Después, se sembraron las células en una placa de 24 pocillos (IWAKI) con 1×10^5 células/ml/pocillo. Después de 24 horas, se recogió el sobrenadante del cultivo y se midió la actividad de SEAP en el sobrenadante del cultivo utilizando el ensayo de indicador de fosfatasa alcalina secretada pNPP SensoLyte (marca registrada) (ANASPEC). Los resultados de la medición se muestran en la figura 2. La actividad de SEAP fue mayor bajo el control de cada uno de los promotores truncados que bajo el control del promotor de CMV (CMV/pSeapIRESpuro) que sirve como el control y, de este modo, se demostró que estos promotores tienen una actividad promotora mayor que el promotor de CMV.

(Ejemplo 6) Extracción de elemento de ADN

(6-1) Inmunoprecipitación con cromatina utilizando el anticuerpo anti histona H3 acetilada

Se realizó ChIP utilizando un anticuerpo anti histona acetilada que utiliza EZ ChIP (Upstate) de acuerdo con el siguiente procedimiento. En relación con esto, a menos que se indique otra cosa, se utilizaron productos de Upstate como los anticuerpos, tampones o similares en el siguiente procedimiento.

En primer lugar, se cultivaron células 293F (Invitrogen) utilizando el medio FreeStyle™ 293 GIBCO (marca registrada) (Invitrogen) en condiciones de 37 °C y CO₂ al 8 %, seguido de centrifugación (1000 rpm, 5 min a temperatura ambiente), con lo que se recogieron células en la fase de crecimiento. Después se agitaron 2×10^7 células en un medio que contenía formaldehído al 1 % durante 10 minutos, se añadió glicina 10x al mismo, seguido de agitación a temperatura ambiente durante 5 minutos. Después de la centrifugación (3000 rpm, 5 min, 4°C), se eliminó el sobrenadante y se añadió PBS al sedimento celular para suspender las células. Después, se centrifugó la suspensión celular de nuevo para eliminar PBS y, a continuación, se añadió un tampón de lisis de SDS al sedimento celular para suspender y lisar las células. Cada muestra obtenida mediante lisis celular se sometió a fragmentación de ADN utilizando un homogeneizador ultrasónico (BRANSON) mientras se refrigeraba la muestra con hielo y se añadió a la misma un tampón de dilución que contenía un cóctel de inhibidores de proteasas y agarosa inmovilizada con proteína G. La mezcla resultante se agitó a 4 °C durante 1 hora, seguido de centrifugación y, a continuación, se recogió el sobrenadante. Posteriormente, se añadieron a la misma 10 µg de IgG de conejo normal o un anticuerpo de α-acetil histona H3, seguido de agitación durante la noche a 4 °C. A la solución resultante, se le añadió agarosa inmovilizada con proteína G y la mezcla resultante se agitó a 4 °C durante 1 hora, seguido de centrifugación y, a continuación, se recogió el sedimento. Se lavó el sedimento así obtenido dos veces con el tampón de lavado Low Salt Immune Complex, dos veces con el tampón de lavado High Salt Immune Complex, dos veces con el tampón de lavado LiCl Immune Complex y, finalmente, cuatro veces con tampón TE. A continuación, se añadió al mismo un tampón de elución (que contenía 20 µl de hidrogenocarbonato de sodio 1 M, 10 µl de SDS y 170 µl de agua estéril). Después de 30 minutos, se centrifugó la mezcla y se recogió el sobrenadante.

Posteriormente, se añadió cloruro de sodio 5 M al sobrenadante y se calentó la mezcla resultante durante la noche a 65 °C. A continuación, se añadió a la misma RNasa A y se incubó la mezcla resultante a 37 °C durante 30 minutos. A continuación, se añadieron a la misma EDTA 0,5 M, Tris-HCl 1 M y proteinasa K, y se incubó la mezcla resultante a 45 °C durante 2 horas. Finalmente, se añadieron a la misma reactivos A, B y C en una cantidad 5 veces superior que la de la solución obtenida mediante el tratamiento con proteinasa K, seguido de centrifugación (10000 rpm, 30 segundos a temperatura ambiente) utilizando un filtro de centrifugación, con lo que se purificó ADN inmunoprecipitado con cromatina.

(6-2) Análisis de micromatriz

Mediante la utilización de un kit de amplificación genómica completa (WGA) Genomeplex (Sigma), se amplificó cada muestra ChIP obtenida en (6-1). El procedimiento fue de acuerdo con el protocolo de Sigma que acompaña al kit.

Para confirmar ChIP, mediante la utilización de 320 ng de cada ADN amplificado mediante WGA como un molde y utilizando también los siguientes cebadores y SYBR (marca registrada) Premix Ex Taq™ (Perfect Real Time) (TAKARA), se amplificó internamente un gen de la gliceraldehído-3-fosfato deshidrogenasa (GAPDH) mediante el procedimiento de PCR (95 °C durante 5 segundos y 60 °C durante 20 segundos x 45 ciclos). En relación con esto, GAPDH es un gen constitutivo para utilizarse como un control positivo para confirmar si un elemento de ADN se enriquece mediante ChIP y se realizó el procedimiento PCR utilizando cebadores unidos a un EZ ChIP (Upstate).

5'-TACTAGCGGTTTTACGGGCG-3'
5'-TCGAACAGGAGGAGCAGAGAGCGA-3'

Como resultado, se confirmó que GAPDH se amplificaba especialmente en la muestra sometida a inmunoprecipitación con un anticuerpo anti histona H3 acetilada (figura 3). Se sometió cada una de las muestras de ADN amplificadas mediante WGA a análisis de micromatriz (NimbleGen) para realizar inmunoprecipitación con cromatina en chips (ChIP-on-chip). "ChIP-on-chip" es una técnica para identificar cada elemento de ADN sometiendo

ADN enriquecido en (6-1) a análisis de micromatriz.

(6-3) Extracción de un elemento de ADN

Basándose en los resultados del análisis ChIP-on-chip obtenido en (6-2), se extrajeron 5 secuencias que tenían un contenido de AT del 62% o más.

- 5 A2: cromosoma 15 (80966429 a 80974878)
- A7: cromosoma 11 (88992123 a 89000542)
- A18: cromosoma 4 (111275976 a 111284450)
- B5: cromosoma 1 (143034684 a 143043084)
- C14: cromosoma 11 (46089056 a 46097482)

10 **(Ejemplo 7)**

Efecto del elemento de ADN utilizando la expresión de la fosfatasa alcalina secretora (SEAP) como índice (7-1)
Construcción del vector de expresión SEAP

15 Mediante la utilización del control pSEAP2 (Clontech) como un molde, se amplificó el gen SEAP mediante el procedimiento de PCR (94 °C durante 30 segundos y 68 °C durante 2 min x 40 ciclos) utilizando los siguientes cebadores y KOD -Plus- (TOYOBO).

5'-AAAGCTAGCATGCTGCTGCTGCTGCTGCTGCTGGGCC-3'
5'-AAAAGATCTTCATGTCTGCTCGAAGCGGCCGGCCGC-3'

20 Posteriormente, se separó el fragmento SEAP amplificado mediante electroforesis en gel de agarosa y se cortó del gel, seguido de purificación utilizando un kit de extracción de gel QIAquick (Qiagen). El fragmento de ADN así obtenido se utilizó como un inserto. Se digirió el inserto con las enzimas de restricción NheI y BglII y se digirió un vector pRES hyg3 (Clontech) con las enzimas de restricción NheI y BamHI. Los fragmentos de ADN resultantes se sometieron a electroforesis en gel de agarosa para separar los fragmentos diana, respectivamente, y se cortaron los fragmentos diana del gel, seguido de purificación. Después, se realizó una reacción de ligadura y transformación. La
25 reacción de ligadura se realizó utilizando el sistema de ligadura de ADN rápido LigaFast (Promega). La transformación se realizó del siguiente modo. En primer lugar, se descongelaron células JM109 competentes congeladas (TAKARA), se añadieron 10 µl de una solución obtenida después de la reacción de ligadura a una solución de las células descongeladas y la mezcla resultante se dejó en reposo en hielo durante 30 minutos. A continuación, se aplicó un choque térmico a la mezcla (42 °C, 45 segundos) y la mezcla se enfrió en hielo durante 5 minutos. A esta suspensión celular, se le añadió 1 ml de medio LB y la mezcla resultante se agitó a 37 °C durante 1
30 hora. Después, la mezcla se colocó en una placa LB que contenía 0,1 mg/ml de ampicilina y la placa se incubó a 37 °C durante 14 a 16 horas. A continuación, mediante lisis alcalina, se recogió un plásmido diana de colonias cultivadas en la placa LB. Finalmente, se determinó la secuencia polinucleotídica de SEAP en el plásmido obtenido mediante lisis alcalina, con lo que se construyó pCMV/SEAP ires Hygro.

(7-2) Clonación de elemento de ADN

35 Posteriormente, cada uno de los elementos de ADN extraídos en el ejemplo 6 se clonó a partir de un cromosoma artificial bacteriano (BAC) que contenía una secuencia polinucleotídica que corresponde al elemento de ADN en el vector de expresión de SEAP obtenido en (7-1) utilizando un kit BAC SUBCLONING (Gene Bridges).

40 En primer lugar, se digirió pCMV/SEAP ires Hygro obtenido en (7-1) con la enzima de restricción SpeI durante varias horas, seguido de precipitación con etanol y el precipitado se disolvió en agua estéril. Mediante la utilización del vector digerido con SpeI como un molde, se realizó el procedimiento de PCR (94 °C durante 15 segundos, 55 °C durante 30 segundos y 68 °C durante 10 minutos x 30 ciclos) utilizando los siguientes cebadores y KOD -Plus- (TOYOBO).

A2D:

5' -GGAAATTGAGAAGTATCATTCACAACAGTACCACAAACATGAAATAAATGTGGAT
CCTATTAATAGTAATCAATTACG-3'

45 A2R:

5' -CTCATTCTGTGGGTTGTCATTTCACTTCCTTGATGCTATCCTTTCAAGCAAATC
CTAGTCAATAATCAATGTCAACG-3'

A7D:

5' -CTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAATAAAAACTTGGATC
CTATTAATAGTAATCAATTACG-3'

A7R:

5' -CTCTTCCCATTCTCATTGGAATCTACTTCAAAGGTTTACCATACTAAGACCTAG
TCAATAATCAATGTCAACG-3'

5 A18D:

5' -CGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGTGGATCACCTGAGGTCGA
TCCTATTAATAGTAATCAATTACG-3'

A18R:

5' -CATACAGAAGCCAGTTTGAAGTACTGAGACCTCACTCCATTTCTTACAAGTTATGCC
TAGTCAATAATCAATGTCAACG-3'

B5D:

10 5' -ACCGTTTTATATTGTTTAAGCATTTCCTAGACATATTTGGCTACAAATCTAGATC
CTATTAATAGTAATCAATTACG-3'

B5R:

5' -GATCTTAGGGGGCTGATTATATAAAACAATAGAAATGTAGTCTTAGATGAAACC
TAGTCAATAATCAATGTCAACG-3'

C14D:

5' -CACAAAGTTCCTGTCAAGGCCAGGTGATGAGGCCACACATGCCCGGACCTTGA
TCCTATTAATAGTAATCAATTACG-3'

15 C14R:

5' -CAAACCTCATCTCTACTGAAAATAGAAAATTAGCTGGGCGTGGTGGCAGGTGCC
CTAGTCAATAATCAATGTCAACG-3'

Después de que la amplificación se confirmara mediante electroforesis en gel de agarosa utilizando una porción de la solución de reacción, el resto de la solución de reacción se sometió a precipitación con etanol. El precipitado se disolvió en agua estéril y la solución resultante se utilizó como ADN para transformación.

20 Posteriormente, se realizó una preparación de *Escherichia coli* para transformación.

Los clones de BAC correspondientes a las 5 secuencias extraídas en el Ejemplo 6 son los siguientes.

[Tabla 1]

Secuencia extraída	Clon de BAC correspondiente
A2	RP11-152F13
A7	RP11-643G5
A18	RP11-115A14
B5	RP11-640M9
C14	RP11-702F3

Se inocularon 10 µl del clon BAC mencionado anteriormente (Advanced GenoTechs Co.), que había sido descongelado, en 1 ml de un medio (que contenía cloranfenicol a una concentración final de 15 µg/ml) y se incubó durante la noche a 37 °C. Se transfirieron 30 µl de la solución de cultivo a 1,4 ml de un medio (que contenía cloranfenicol a una concentración final de 15 µg/ml) y se incubó a 37 °C durante 2 horas. Se repitieron centrifugación y lavado con agua estéril dos veces y se suspendieron las células en 20 µl de agua estéril. A una cubeta enfriada (0,1 cm), se le añadieron 1 µl de pRED/ET (Gene Bridges) y *Escherichia coli*, seguido de electroporación (1350 V, 10 µF). Después, a esto se le añadió 1 ml de medio SOC, y la mezcla resultante se incubó a 30 °C durante 70 minutos. 100 µl de la solución de cultivo se colocaron en una placa LB (que contenía tetraciclina y cloranfenicol a concentraciones finales de 3 µg/ml y 15 µg/ml, respectivamente) y se incubaron durante la noche a 30 °C. Al día siguiente, cada colonia así obtenida se inoculó en 1 ml de un medio (que contenía tetraciclina y cloranfenicol a concentraciones finales de 3 µg/ml y 15 µg/ml, respectivamente) y se incubó durante la noche a 30 °C. Se transfirieron 30 µl de la solución de cultivo a 1,4 ml de un medio (que contenía tetraciclina y cloranfenicol a concentraciones finales de 3 µg/ml y 15 µg/ml, respectivamente) y se incubaron a 30 °C durante 2 horas. Después, a esto se le añadieron 50 µl de L-arabinosa al 10 % y se realizó incubación adicional a 37 °C durante 1 hora. A continuación, se repitió el lavado con agua estéril dos veces y se añadieron a una cubeta enfriada (0,1 cm) *Escherichia coli*, que estaba suspendida en 30 µl de agua estéril, y 1 µl del ADN para la transfección, seguido de electroporación (1350 V, 10 µF). Después, a esto se le añadió 1 ml de medio SOC, y la mezcla resultante se incubó a 37 °C durante 90 minutos. La cantidad total de la solución de cultivo se colocó en una placa LB (que contenía 100 µg/ml de ampicilina) y se incubó la placa. A continuación, se obtuvo un plásmido diana mediante lisis alcalina. Finalmente, se confirmaron la secuencia del plásmido obtenido y los sitios de la enzima de restricción del mismo, con lo que se construyó un plásmido diana (figura 4).

(7-3) Evaluación utilizando la expresión de SEAP como índice

Cada plásmido construido en (7-2) se evaluó utilizando la célula huésped CHO-K1 (ATCC) y el reactivo de transfección Lipofectamine 2000 (Invitrogen).

Se realizó una selección con antibióticos con higromicina a 800 µg/ml durante aproximadamente 2 semanas empezando 2 días después de la transfección, con lo que se estabilizó una línea celular policlonal que expresaba de manera estable. La línea celular así establecida se sometió a reemplazo de medio el día antes de la medición, y se sembraron un número dado de células en una placa de 24 pocillos (IWAKI). A las 24 horas del cultivo en placas de las células, se recogió el sobrenadante del cultivo y se midió la actividad de SEAP. La actividad de SEAP en el sobrenadante del cultivo se midió utilizando un ensayo indicador de fosfatasa alcalina secretada pNPP SensoLyte™ (ANASPEC).

Los resultados de la medición se muestran en la figura 5. Cuando la actividad de SEAP en el control sin ningún elemento se normalizó a 1, la actividad de SEAP en el sobrenadante del cultivo de la línea celular CHO que expresaba de manera estable que tenía el elemento de ADN A2, A7, A18, B5 o C14 mostró un valor numérico cinco veces o más mayor que el del control. Según estos resultados, se confirmó que los 5 tipos de elementos de ADN mejoran espectacularmente la expresión de SEAP. En relación con esto, las secuencias de polinucleótidos de los anteriores 5 tipos de elementos de ADN se representan mediante las SEQ ID NOS: 10 a 14 en el listado de secuencias, respectivamente.

40 (Ejemplo 8) Generalidad del promotor a utilizar en combinación

El promotor del vector utilizado en la evaluación de los elementos de ADN en el ejemplo 7 es un promotor de CMV y, de ese modo, se estudió el uso de elementos de ADN en combinación con otros promotores generales.

(8-1) Construcción del vector de expresión SEAP utilizando promotores SV40 y EF-1α

Mediante la utilización del control pSEAP2 (Clontech) como un molde, se amplificó el gen SEAP mediante el procedimiento de PCR (94 °C durante 30 segundos y 68 °C durante 2 min x 40 ciclos) utilizando los cebadores descritos en (7-1) y KOD -Plus-. Se preparó el gen SEAP amplificado como un inserto de la misma forma que en (7-1). Se digirió el inserto con las enzimas de restricción NheI y BglII y se digirió un vector pIRES puro3 (Clontech) con

ES 2 670 894 T3

las enzimas de restricción NheI y BamHI y se construyó pCMV/SEAP ires Puro de la misma forma que en (7-1).

Posteriormente, mediante la utilización de pEF1/V5-His A (Invitrogen) como un molde, se amplificó un promotor EF-1 α mediante el procedimiento de PCR (94 °C durante 15 segundos, 60 °C durante 30 segundos y 68 °C durante 2 min x 30 ciclos) utilizando los siguientes cebadores y KOD -Plus-.

- 5 5'-AAAAGTATCCCTAGCCAGCTTGGGTGGTACCAAGC-3'
 5'-AAAAGTATCCCTAGCCAGCTTGGGTGGTACCAAGC-3'

Mediante la utilización de pCMV/SEAP ires Puro construido anteriormente como un vector, se realizó la digestión con las enzimas de restricción SpeI y EcoRV para el vector y el promotor, y se construyó pEF/SEAP ires Puro de acuerdo con el procedimiento descrito en (7-1).

- 10 De manera análoga, mediante la utilización de pcDNA3.1+ (Invitrogen) como un molde, se amplificó un promotor SV40 mediante el procedimiento de PCR (94 °C durante 15 segundos, 60 °C durante 30 segundos y 68 °C durante 1 min x 30 ciclos) utilizando los siguientes cebadores y KOD -Plus-.

 5'-AAAAGTATCCCTAGCCAGCTTGGGTGGTACCAAGC-3'
 5'-AAAAGTATCCCTAGCCAGCTTGGGTGGTACCAAGC-3'

- 15 Mediante la utilización de pCMV/SEAP ires Puro construido anteriormente como un vector, se realizó la digestión con las enzimas de restricción SpeI y EcoRV para el vector y el promotor y se construyó pSV40/SEAP ires Puro de acuerdo con el procedimiento descrito en (7-1).

(8-2) Clonación del elemento de ADN A2 o A7

- 20 Posteriormente, se realizó la clonación del elemento de ADN A2 o A7 utilizando los pEF/SEAP ires Puro y pSV40/SEAP ires Puro construidos en (8-1) como estructuras básicas.

- 25 En primer lugar, se digirieron pEF/SEAP ires Puro y pSV40/SEAP ires Puro con la enzima de restricción SpeI durante varias horas, seguido de precipitación con etanol y el precipitado se disolvió en agua estéril. Mediante la utilización de los respectivos vectores digeridos con SpeI como moldes, se preparó ADN para transfección mediante el procedimiento de PCR (94 °C durante 15 segundos, 55 °C durante 30 segundos y 68 °C durante 10 min x 30 ciclos) utilizando los siguientes cebadores y KOD -Plus-.

A2 (EF/D):

 5' -GGAAATTGAGAAGTATCATTCACAACAGTACCACAAACATGAAATAAATGTGCTA
 GTCAGAGAGGAATCTTTGCAGC-3'

A2 (SV40/D):

 5' -GGAAATTGAGAAGTATCATTCACAACAGTACCACAAACATGAAATAAATGTGCTA
 GTCTGTGGAATGTGTGTCAGTTAG-3'

- 30 A2 (EF y SV40/R):

 5' -CTCATTCTGTGGGTGTGTCATTTCACTTCCTTGATGCTATCCTTTCAAGCAAAATT
 TTAAAACCTTTATCCATCTTTGCA-3'

A7 (EF/D):

 5' -CTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAAACCTTGCTAG
 TCAGAGAGGAATCTTTGCAGC-3'

A7 (SV40/D):

 5' -CTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAAACCTTGCTAG
 TCTGTGGAATGTGTGTCAGTTAG-3'

- 35

A7 (EF y SV40/R):

5' -CTCTTCCCATTCTCATTGGAATCTACTTCAAAGGTTTACCATACTAAGAAGCTAG

TTTTAAACTTTTATCCATCTTTGCA-3'

Mediante la utilización del ADN así preparado para transfección y un BAC transfectado con pRed/ET, se clonó el elemento de ADN A2 o A7 en el vector descrito en (8-1). En relación con esto, se realizó el procedimiento de acuerdo con el procedimiento descrito en (7-2).

(8-3) Evaluación utilizando la expresión de SEAP como índice

Cada plásmido construido en (8-2) se evaluó utilizando la célula huésped CHO-K1 (ATCC) y el reactivo de transfección Lipofectamine 2000 (Invitrogen).

Se realizó selección de antibióticos con puromicina a 8 µg/ml durante aproximadamente 2 semanas empezando 2 días después de la transfección, con lo que se estabilizó una línea celular policlonal que expresaba de manera estable. La línea celular así establecida se sometió a reemplazo de medio el día antes de la medición y se sembraron un número dado de células en una placa de 24 pocillos. A las 24 horas del cultivo en placas de las células, se recogió el sobrenadante del cultivo y se midió la actividad de SEAP. La actividad de SEAP en el sobrenadante del cultivo se midió utilizando el ensayo indicador de fosfatasa alcalina secretada pNPP SensoLyte™ (ANASPEC).

Los resultados de la medición se muestran en la figura 6. Cuando la actividad de SEAP en el control sin ningún elemento se normalizó a 1, el elemento de ADN A2 o A7 exhibió un efecto de mejora de la expresión de tal manera que la actividad de SEAP fue dos veces o más mayor en el caso del uso con el promotor EF-1α y cuatro veces o más mayor en el caso del uso con el promotor SV40 que la del control. Según estos resultados, se confirmó que estos elementos de ADN exhiben el efecto de mejorar la expresión del gen extraño cuando se utilizan en combinación con un promotor general.

(Ejemplo 9) Evaluación utilizando la expresión de anticuerpos como índice

(9-1) Construcción del vector de expresión de cadena ligera humano pEF6KCL

Mediante la utilización del plásmido pEF6/V5-HisB (Invitrogen) como un molde, un fragmento de ADN entre la posición 2174 (inmediatamente corriente abajo de BGHpA) y la posición 2958 (SmaI) (un fragmento de ADN que contiene un origen f1 de replicación y un promotor SV40 y origen de replicación, denominado en lo sucesivo "fragmento A", viniendo representada la secuencia polinucleotídica del fragmento A por la SEQ ID NO: 15 en el listado de secuencias) se obtuvo mediante el procedimiento de PCR utilizando los siguientes cebadores y KOD -Plus-

5'-CCACGCGCCCTGTAGCGGCGCATTAAAGC-3'
5'-AAACCCGGGAGCTTTTGC AAAAGCCTAGG-3'

El fragmento A obtenido y un fragmento de ADN que contiene una secuencia de ADN que codifica una señal secretora de cadena k humana, una región constante de cadena k humana, y una señal de adición de poli(A) humana (en lo sucesivo en el presente documento denominado "fragmento B") se ligaron mediante PCR solapadas. Se digirió el fragmento de ADN así obtenido, en el que el fragmento A y el fragmento B se ligaron, con las enzimas de restricción KpnI y SmaI y el fragmento resultante se ligó al plásmido pEF6/V5-HisB (Invitrogen) que se digirió con las enzimas de restricción KpnI y SmaI, con lo que se construyó un vector de expresión de cadena ligera humano pEF6KCL que tenía una secuencia señal, un sitio de clonación, una región constante de cadena k humana y una secuencia señal de adición de poli(A) humana corriente abajo del promotor EF-1α.

Se ligó un fragmento de ADN, obtenido mediante la digestión del pEF6KCL preparado mediante el procedimiento mencionado anteriormente con las enzimas de restricción KpnI y SmaI, a pEF1/myc-HisB (Invitrogen) que se digirió con KpnI y SmaI, seguido de transformación, lisis alcalina y secuenciación, con lo que se construyó el plásmido pEF1KCL.

(9-2) Construcción del vector de expresión de cadena pesada humano pEF1FCCU

Un fragmento de ADN (la secuencia polinucleotídica de este fragmento de ADN se representa por la SEQ ID NO: 16 en el listado de secuencias) que comprende una secuencia de ADN que codifica una secuencia señal IgG1 humana y una secuencia de aminoácidos de región constante se digirió con las enzimas de restricción NheI y PmeI y el fragmento resultante se ligó al plásmido pEF1KCL que se digirió con NheI y PmeI, con lo que se construyó el vector de expresión de cadena pesada humano pEF1FCCU que tenía una secuencia señal, un sitio de clonación, una región constante de cadena pesada humana y una secuencia señal de adición de poli(A) humana corriente abajo del promotor EF-1α.

(9-3) Construcción de un vector de expresión de gen X de anticuerpo humanizado único (Gen de anticuerpo humanizado X/pEF_LHN#)

Ligando el vector de expresión de cadena H o cadena L construidos en (9-1) o (9-2), se construyó un vector de expresión de anticuerpo único (pEF_LHN (que carece de una región variable)).

5 Se añadió un sitio de enzima de restricción Sall mediante el procedimiento de PCR a ambos extremos de la unidad de expresión de genes: uno cadena arriba del promotor y el otro cadena abajo del poli(A) de pEF1KCL. Se realizó, a continuación, electroforesis en gel de agarosa, corte de un fragmento de ADN deseado del gel y purificación del fragmento de ADN, con lo que se preparó un inserto. Mediante la digestión del vector pEF1FCCU construido en (9-2) con la enzima de restricción Sall, se linealizó el vector en el sitio Sall localizado cadena arriba de la unidad de expresión de genes. Después, se ligó el vector linealizado al inserto anterior, seguido de transformación, lisis alcalina y secuenciación, con lo que se construyó un vector de expresión de anticuerpo humanizado único (pEF_LHN (que carece de región variable)).

Posteriormente, se introdujeron los siguientes oligonucleótidos en un sitio AatII del vector pEF_LHN (que carece de una región variable).

15 5'-CGCGGCCGCGACTAGTGACGT-3'
5'-CACTAGTGCGGCCGCGACGT-3'

Se diluyeron los oligonucleótidos respectivos a 5 pmol y mediante la utilización de quinasa polinucleotídica T4 (TAKARA), se dejó que la reacción continuara a 37 °C durante 1 hora. Después, se añadió a ello tampón H 10x (TAKARA) y se realizó hibridación mediante una reacción a 96 °C durante 1 minuto y, a continuación, a temperatura ambiente durante 30 minutos. Se ligaron estos oligonucleótidos y el vector pEF_LHN que se digirió con la enzima de restricción AatII, seguido de transformación, lisis alcalina y secuenciación, con lo que se construyó pEF_LHN# (que carece de una región variable).

Mediante la integración de una región variable de gen X de anticuerpo humanizado en el vector universal construido anteriormente (pEF_LHN# (que carece de una región variable)), se completó la construcción de un vector de expresión del gen X de anticuerpo humanizado único (gen de anticuerpo humanizado X/pEF_LHN#).

En primer lugar, mediante la utilización de los siguientes cebadores y KOD -Plus-, se amplificó una región variable de cadena L del gen X de anticuerpo humanizado mediante el procedimiento de PCR (94 °C durante 15 segundos, 55 °C durante 30 segundos y 68 °C durante 1 min x 30 ciclos). Región variable de cadena L:

30 5'-AAACATATGGCGACATCCAGATGAC-3'
5'-AAACGTACGCTTGATCTCCACCTTGG-3'

El fragmento de región variable de cadena L amplificado y el vector universal (pEF_LHN# (que carece de una región variable)) se digirieron con las enzimas de restricción NdeI y BsiWI, seguido de electroforesis en gel de agarosa, corte de un fragmento deseado del gel, purificación, reacción de ligadura, transformación, lisis alcalina y secuenciación, con lo que se integró la región variable de cadena L en el vector. De la misma manera, mediante la utilización de los siguientes cebadores y KOD -Plus-, se amplificó una región variable de cadena H de gen X de anticuerpo humanizado mediante el procedimiento de PCR (94 °C durante 15 segundos, 55 °C durante 30 segundos y 68 °C durante 1 min x 30 ciclos).

Región variable de cadena H:

40 5'-AAAGCTGAGCCAGGTGCAGCTGCAGG-3'
5'-AAAGCTGAGCTCACGGTCACCAGGGTTC-3'

Se digirieron el fragmento de región variable de cadena H amplificado y el vector que tiene la región variable de cadena L insertada en el mismo con la enzima de restricción BlnI, seguido de electroforesis en gel de agarosa, corte de un fragmento deseado del gel, purificación, reacción de ligadura, transformación, lisis alcalina y secuenciación, con lo que se integró la región variable de cadena H en el vector y se construyó un vector de expresión de gen X de anticuerpo humanizado único (gen de anticuerpo humanizado X/pEF_LHN#).

(9-4) Construcción de un vector de expresión de gen X de anticuerpo humanizado único (gen de anticuerpo humanizado X/pCMV_LHN#)

Mediante la utilización del vector de expresión de gen X de anticuerpo humanizado único (gen de anticuerpo humanizado X/pEF_LHN#) construido en (9-3) como una estructura de vector básica, se construyó otro vector de expresión de gen X de anticuerpo humanizado único (gen de anticuerpo humanizado X/pCMV_LHN#) mediante reemplazo del promotor de acuerdo con el siguiente proceso.

Mediante la utilización de pIRES puro3 como un molde, se amplificó un fragmento de promotor de CMV mediante el procedimiento de PCR (94 °C durante 30 segundos y 68 °C durante 3 min x 40 ciclos) utilizando los siguientes cebadores y KOD -Plus-.

Cadena arriba de cadena H:

5'-CTTTTGCAAAAAGCTTCGCGTTACATAACTTACGGTAAATGGCC-3'
5'-TTCATGGTGGCGCTAGCCCGCAGATATCGATCCGAGCTCGGTA-3'

Cadena arriba de cadena L:

5 5'-TGACGTCGACAAGCTTCGCGTTACATAACTTACGGTAAATGGCC-3'
5' -CTGGATGTCGCCATATGCGCCGGAGATCCACAGCAGCAGGGAGATGAACACCTGG
GTCTGCAGCACCATGGTGGCGCTAGCCCGCAGATATCGATCCGAGCTCGGTA-3'

A la solución de reacción de PCR, se añadió la enzima de restricción DpnI y se dejó que la reacción continuara a 37 °C durante 1 hora, seguida de purificación utilizando un kit MinElute reaction Cleanup (Qiagen), con el que se preparó una muestra para su uso en In-Fusion. Paralelamente, se digirió el gen de anticuerpo humanizado X/pEF_LHN# con las enzimas de restricción HindIII, NheI, NdeI y FseI, seguido de electroforesis en gel de agarosa, con lo que se separaron dos fragmentos grandes entre los fragmentos resultantes. Cada uno de los fragmentos se cortó del gel y se extrajo el ADN del gel, con el que se preparó una muestra para su uso en In-Fusion. Todas las muestras para su uso en In-Fusion se colocaron juntas y se realizó la clonación utilizando un kit de clonación de la PCR In-Fusion™ Advantage (TAKARA), seguido de transformación, lisis alcalina y secuenciación, con lo que se construyó un vector de expresión de gen X de anticuerpo humanizado (gen de anticuerpo humanizado X/pCMV_LHN#).

(9-5) Clonación del elemento de ADN A7

Se seleccionó A7 a partir de los 5 tipos de elemento de ADN que, según se había confirmado, tenían el efecto de mejorar la expresión de SEAP y se clonaron en un vector de expresión de anticuerpo.

20 De la misma forma que en (7-2), mediante la utilización de cada uno de los vectores de expresión de gen X de anticuerpo humanizado único (gen de anticuerpo humanizado X/pEF_LHN# y gen de anticuerpo humanizado X/pCMV_LHN#) digeridos con la enzima de restricción NotI como un molde, se preparó ADN para transfección mediante el procedimiento de PCR (94 °C durante 15 segundos, 55 °C durante 30 segundos y 68 °C durante 11 min x 30 ciclos) utilizando los siguientes cebadores y KOD -Plus-.

25 Gen de anticuerpo humanizado X/pEF_LHN# D:

5' -CTCTTCCCATTTCTCATTTGAATCTACTTCAAAGGTTTACCATACTAAGACTCGA
GGCACTAGTGACGTCAGGTGGCACT-3'

Gen de anticuerpo humanizado X/pEF_LHN# R:

5' -CTCTTCCCATTTCTCATTTGAATCTACTTCAAAGGTTTACCATACTAAGAGCACT
AGTGACGTCAGGTGGCACTTTTCGG-3'

Gen de anticuerpo humanizado X/pCMV_LHN# D:

30 Se utilizó gen de anticuerpo humanizado X/pEF_LHN# D.

Gen de anticuerpo humanizado X/pCMV_LHN# R:

Se utilizó gen de anticuerpo humanizado X/pEF_LHN# R.

35 Mediante la utilización del ADN preparado anteriormente para transfección y un BAC transfectado con pRed/ET, se clonó el elemento de ADN A7 en el vector de expresión de gen X de anticuerpo humanizado único descrito en (9-3) y (9-4). Una vista esquemática del vector de construcción se muestra en la Figura 7. En relación con esto, se realizó el procedimiento de acuerdo con el procedimiento descrito en (7-2).

(9-6) Evaluación utilizando la expresión de anticuerpo como índice

Cada plásmido construido en (9-5) se evaluó utilizando la célula huésped CHO-K1 (ATCC) y el reactivo de transfección Lipofectamine 2000 (Invitrogen).

40 Se realizó selección de antibióticos con Geneticin (Roche) a 800 µg/ml durante aproximadamente 2 semanas empezando 2 días después de la transfección, con lo que se estabilizó una línea celular policlonal que expresaba de

manera estable. La línea celular así establecida se sometió a reemplazo de medio el día antes de la medición y se sembraron un número dado de células en una placa de 24 pocillos. A las 24 horas del cultivo en placas de las células, se recogió el sobrenadante del cultivo y se midió el nivel de expresión del anticuerpo en el sobrenadante del cultivo mediante el procedimiento ELISA. En relación con esto, se realizó ELISA del siguiente modo. En una placa de 5 96 pocillos revestida con anticuerpo anti-cadena ligera kappa a 50 ng/pocillo, se añadió a cada pocillo 100 µl del sobrenadante del cultivo libre de células y se incubó la placa a 37 °C durante 1 hora. Posteriormente, se eliminó la muestra (sobrenadante de cultivo) y se lavó cada pocillo con 200 µl de PBS-Tween (0,05%). Después, se añadió a cada pocillo 100 µl de anti-IgG humana (Fc) marcada con HRP y se incubó la placa a 37 °C durante 1 hora más. A 10 continuación, se eliminó el anticuerpo anti-IgG humana (Fc) marcado con HRP y se lavó cada pocillo con PBS-Tween (0,05%). Después, se desarrolló un color utilizando un kit POD Substrate ABTS (Nacalai) y se midió la absorbancia a una longitud de onda de medición de 405 nm. Para la dilución del anticuerpo anti-cadena ligera kappa, se utilizó el anticuerpo anti-IgG humana (Fc) y la muestra, PBS-Tween (0,05%). Mediante la utilización de IgG humana diluida en serie a 12 ng, 6 ng, 3 ng, 1,5 ng, 0,75 ng, 0,375 ng y 0,1875 ng como un patrón, se calculó la concentración de la muestra.

15 Los resultados se muestran en la figura 8. Se confirmó que la muestra que tiene el elemento de ADN A7 tiene un efecto mayor de mejora de la producción de anticuerpos en comparación con un control sin ningún elemento cuando se utilizaron el promotor EF-1α o el promotor de CMV en el vector de expresión de anticuerpos.

(Ejemplo 10) Longitud de secuencia que exhibe actividad de mejora de la expresión de genes extraños

(10-1) Clonación de elementos de ADN que tienen diferentes longitudes de secuencia

20 Basándose en la longitud de la secuencia utilizada en el ejemplo 7, se construyeron vectores que contienen cada uno de los elementos de ADN pero que tienen diferentes longitudes de secuencia.

Los detalles de los elementos de ADN que tienen diferentes longitudes de secuencia que se diseñaron según la longitud completa de cada uno de los elementos de ADN A2, A7, A18, B5 y C14 se muestran en las figuras 9, 11, 13, 15 y 17, respectivamente. Se digirió el pCMV/SEAP ires Hygro descrito en (7-1) con la enzima de restricción SpeI durante varias horas, seguido de precipitación con etanol y el precipitado se disolvió en agua estéril. Mediante la utilización del vector digerido con SpeI como un molde, se preparó ADN para transfección mediante el procedimiento de PCR (94 °C durante 15 segundos, 55 °C durante 30 segundos y 68 °C durante 10 min x 30 ciclos) utilizando los siguientes cebadores y KOD -Plus-. Mediante la utilización del ADN así preparado para transfección y el BAC correspondiente transfectado con pRed/ET, se clonó cada elemento de ADN que tiene una longitud de secuencia 25 diferente en el pCMV/SEAP ires Hygro descrito en (7-1). En relación con esto, se realizó el procedimiento de acuerdo con el procedimiento descrito en (7-2).

A2-1D:

5' -CATGCACAGATTAGCCATTTAGTACTTACTAAATCAAACCTCAATTTCTGAAGTCT
AGTTATTAATAGTAATCAATTACG-3'

A2-1R:

5' -CTCATTCTGTGGGTTGTCATTTCACTTCCTTGATGCTATCCTTTCAAGCAAATT
35 CAATAATCAATGTCAACGCGTATAT-3'

A2-2D:

5' -ACACTGGTCAAAGGGACAGGTCATTGTTATGCTGGCAATGCAGGCTGCTGAAAAC
TAGTTATTAATAGTAATCAATTACG-3'

A2-2R:

5' -ACTGTAGCTTCTTATTTTTTACCTGCAGTGCATTCCTGTAAAAGTAGTGTGGAGT
40 CAATAATCAATGTCAACGCGTATAT-3'

ES 2 670 894 T3

A2-3D:

5' -CTGGAAATTGAGAAGTATCATTACACAACAGTACCACAAACATGAAATAAATGTGC
TAGTTATTAATAGTAATCAATTACG-3'

A2-3R:

5' -CCAAGCTTGTCCAACCGCGGCTGCAGGCTGCATGCAGCCTGTGAAGGCTTTGAT
CAATAATCAATGTCAACGCGTATAT-3'

5 A2-4D:

5' -TCAATCATTTATCAATTTTATCTTCAAAGTCCCTCACTTCAGGGAGATGATATAC
TAGTTATTAATAGTAATCAATTACG-3'

A2-4R:

5' -ATATATAAAAGTTCATGTATATATAAAATCATGCAATACACGGCCTTTTGTGACT
CAATAATCAATGTCAACGCGTATAT-3'

A2-5D:

10 5' -CGCATAAAAGGAAAAGCATCCTTAAAATAAACACCATCAATGGCTCCTCGGTGGC
TAGTTATTAATAGTAATCAATTACG-3'

A2-5R:

se utilizó A2-4R.

A2-6D:

15 5' -GGGAGGCTACAGCTTGCCTCTCTAACCCTAAAAGGCATGACCCTCCTCAAAGCT
AGTTATTAATAGTAATCAATTACG-3'

A2-6R:

se utilizó A2-4R.

A2-7D:

5' -TCTGGCTTCCCTGGGCCACGCTGGAAGAAGAATTGTCTTGCGCCACACATAAAAC
TAGTTATTAATAGTAATCAATTACG-3'

A2-7R:

20 5' -AGCTGATTTTTACGTTAAATGTAACATGTAAAGAAATATATGTGTGTTTTAGAT
CAATAATCAATGTCAACGCGTATAT-3'

A2-8D:

5' -GTGAAGAGGAGGAGATGTCAAATTCAAAGTCTTAAATGATGTAGTTTTAAGTAC
TAGTTATTAATAGTAATCAATTACG-3'

ES 2 670 894 T3

A2-8R:

5' -ATGACACTTGATATTGTTGTTTATATTGCTGGTTAGTATGTGCCTTCATTTACCT

CAATAATCAATGTCAACGCGTATAT - 3'

A2-9D:

se utilizó A2-6D.

5

A2-9R:

se utilizó A2R.

A2-10D:

se utilizó A2-2D.

A2-10R:

10

se utilizó A2-7R.

A2-11D:

se utilizó A2-8D.

A2-11R:

se utilizó A2-2R.

15

A2-12D:

se utilizó A2-2D.

A2-12R:

se utilizó A2-4R.

A2-13D:

20

se utilizó A2-8D.

A2-13R:

se utilizó A2-7R.

A2-14D:

se utilizó A2D.

25

A2-14R:

se utilizó A2-2R.

A2-15D:

se utilizó A2-2D.

A2-15R:

30

se utilizó A2R.

A2-16D:

se utilizó A2-8D.

A2-16R:

se utilizó A2-4R.

ES 2 670 894 T3

A2-17D:

se utilizó A2D.

A2-17R:

se utilizó A2-7R.

5

A7-1D:

5' -AAAAACAAAACCTGGAGTAAACAAGATGAATTGTTTTAATAGAGGCACTGTATTAC

TAGTTATTAATAGTAATCAATTACG-3'

A7-1R:

5' -ATACAATGTTCCATGTATTCTGTGCCTGAACCTATGCAGCTGATGTAGCTGAAGT

CAATAATCAATGTCAACGCGTATAT-3'

A7-2D:

5' -GATCTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAATAAAAACCTTGC

10

TAGTTATTAATAGTAATCAATTACG-3'

A7-2R:

5' -TGTTGTTTTTCAGCCACTAAGTTTGAGGTGATTTGTTCTGGCAGTCCTAGGAACT

CAATAATCAATGTCAACGCGTATAT-3'

A7-3D:

se utilizó A7-2D.

15

A7-3R:

5' -AGCCTACACTACCCTTTGCAGCCTTTGGTAACTATCCTTCTGCTGTCTACCTCCT

CAATAATCAATGTCAACGCGTATAT-3'

A7-4D:

5' -AGGAGCTCCTGAATGAAGGACATCACTCAGCTGTGTTAAGTATCTGGAACAATAC

TAGTTATTAATAGTAATCAATTACG-3'

A7-4R:

5' -GACATAAAATGTAAGATATGATATGCTATGTAAGATATGATACCTGCCTTAAAT

20

CAATAATCAATGTCAACGCGTATAT-3'

A7-5D:

5' -CACTGCTTGATACTTACTGTGGACTTTGAAAATTATGAATGTGTGTGTGTGTGTC

TAGTTATTAATAGTAATCAATTACG-3'

ES 2 670 894 T3

A7-5R:

5' -CAATTACATTCCAGTGATCTGCTACTTAGAATGCATGACTGAACTCCTGGGTGGT
CAATAATCAATGTCAACGCGTATAT-3'

A7-6D:

5' -TTATTTTGAAGAGAACTCCTGGTTCCTTAAATCCTTTCTTGTTTCCAAGC
TAGTTATTAATAGTAATCAATTACG-3'

5

A7-6R:

5' -AAGCAGTGTGTGTTTACCTGCATGTGTATGTGAATTAACCTCTGTTCCCTGAGGCAT
CAATAATCAATGTCAACGCGTATAT-3'

A7-7D:

5' -ATTGCATGTTCTCATTATTTGTGGGATGTAAAAATCAAACAATAGAACGTATC
TAGTTATTAATAGTAATCAATTACG-3'

A7-7R:

5' -TTGGGAGGCCGCAGCTGGTAGATCACTTGAGGCCACGAATTTGACACCAGCAGGT
CAATAATCAATGTCAACGCGTATAT-3'

10

A7-8D:

se utilizó A7-1D.

A7-8R:

se utilizó A7R.

15

A7-9D:

se utilizó A7-7D.

A7-9R:

se utilizó A7-5R.

A7-10D:

20

se utilizó A7-4D.

A7-10R:

se utilizó A7-7R.

A7-11D:

se utilizó A7-6D.

25

A7-11R:

se utilizó A7-4R.

A7-12D:

se utilizó A7-2D.

A7-12R:

se utilizó A7-6R.

A7-13D:

se utilizó A7-7D.

5

A7-13R:

se utilizó A7R.

A7-14D:

se utilizó A7-4D.

A7-14R:

10

se utilizó A7-5R.

A7-15D:

se utilizó A7-6D.

A7-15R:

se utilizó A7-7R.

15

A7-16D:

se utilizó A7-2D.

A7-16R:

se utilizó A7-4R.

A7-17D:

20

se utilizó A7-4D.

A7-17R:

se utilizó A7R.

A7-18D:

se utilizó A7-6D.

25

A7-18R

se utilizó A7-5R.

A18-1:

5' -ATCCCCTGCTCTGCTAAAAAGAATGGATGTTGACTCTCAGGCCCTAGTTCTTGA
TCCTATTAATAGTAATCAATTACG-3'

A18-1R:

30

se utilizó A18R.

A18-2D:

5' -CTAAAGTGCTGGGATTACAGGCATAAGCCACCGTGCCCGGCTGGAGCATTGGGAT
CCTATTAATAGTAATCAATTACG-3'

ES 2 670 894 T3

A18-2R:

5' -ACTACTTACACATTTTCGAGTTTTAAATAAGGCGTTCAATATAGAGTGAACACCTA
GTCAATAATCAATGTCAACG-3'

A18-3D:

5' -CAGGCATAAGCCACCGCACCCGGCCACCCCTTACTAATTTTTAGTAACGTCGATC
CTATTAATAGTAATCAATTACG-3'

5

A18-3R:

5' -CTGATTGACTTTGACCTCTGCTTTCCAACCTTGCCCCAAGAAAGTTAGTCACCT
AGTCAATAATCAATGTCAACG-3'

A18-4D:

se utilizó A18-3D.

A18-4R:

5' -TTCAATGAAACAAGCTCTGTGAGGCTCATTTGTACCCATTTTGTTCAGTACTGCC
TAGTCAATAATCAATGTCAACG-3'

10

B5-1D:

5' -ACATACCCAGAGACACTGAGAGAGACAGACAGACAGTAAACAGAGGAGCACGATC
CTATTAATAGTAATCAATTACG-3'

B5-1R:

se utilizó B5R.

15

B5-2D:

5' -GCTCAATTGTATCTTATGAAAACAATTTTTCAAATAAAACAAGAGATATGATCC
TATTAATAGTAATCAATTACG-3'

B5-2R:

se utilizó B5R.

B5-3D:

5' -CCTGTGCTGAATACCGTCTGCATATGTATAGGAAAGGGTAACTCAGCAGGGATC
CTATTAATAGTAATCAATTACG-3'

20

B5-3R:

5' -TATGTGAATGGAATAAAATAATCAAGCTTGTTAGAATTGTGTTTCATAATGACCC
TAGTCAATAATCAATGTCAACG-3'

B5-4D:

se utilizó B5D.

ES 2 670 894 T3

B5-4R:

5' -GAAAGTCTACAATTTTTTCAGTTTAAAATGGTATTTATTTGTAACATGTACCCTA
GTCAATAATCAATGTCAACG-3'

B5-5D:

se utilizó B5-1D.

5

B5-5R:

5' -CAAAGATGAAGGATGAGAGTGACTTCTGCCTTCATTATGTTATGTGTTTCATATCC
TAGTCAATAATCAATGTCAACG-3'

B5-6D:

5' -CAGTGAATTATTCACCTTTGTCTTAGTTAAGTAAAAATAAAATCTGACTGTGATCC
TATTAATAGTAATCAATTACG-3'

B5-6R:

10

5' -GAACAGACAGGTGAATGAGCACAGAGGTCATTTGTAAACCGTTTGTGGTTAGCCT
AGTCAATAATCAATGTCAACG-3'

C14-1D:

5' -CTTTTTGGCTTCTGTGTTTAAGTTATTTTTCCCCTAGGCCACAAACAGAGTCGA
TCCTATTAATAGTAATCAATTACG-3'

C14-1R:

5' -AACCTTGGAAAAATTCTGTTGTGTTTAGAAGCATGTACCAATCTATCACTCCTAG
TCAATAATCAATGTCAACG-3'

15

C14-2D:

5' -CTATTCACTGTCTGTAGGATGAAAAGTTAATAACACCCTGAGAGGTTTCGATCC
TATTAATAGTAATCAATTACG-3'

C14-2R:

5' -CCTTAGATTAGTTTATTGTATTTTTTATCAGCTACTATAAGGTTTACACACCCTA
GTCAATAATCAATGTCAACG-3'

C14-3D:

20

5' -CAAGACCCTCAAATTCAAAAATTCCTTTATCTTGCTGTAGCACCTCCTGCGAT
CCTATTAATAGTAATCAATTACG-3'

C14-3R:

5' -GGAGGGGATAGGAAGGGGATGAGGCCTAACAGGTTGATGATCTAGGCTTTACCTA
GTCAATAATCAATGTCAACG-3'

C14-4D:

5' -CTCAAAAAGGAGATAATTCCAGCCCCTCGCCTTAAAGAATCCCTATCAAGTGATC
CTATTAATAGTAATCAATTACG-3'

5

C14-4R:

se utilizó C14-1R.

C14-5D:

5' -CGCTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCCGAGATCACGCCGTTGGATCC
TATTAATAGTAATCAATTACG-3'

C14-5R:

10

se utilizó C14-1R.

C14-6D:

se utilizó C14-4D.

C14-6R:

5' -TTAACTTTTTTCATCCTACAGACAGTGAATAGTAAAGCTTCTGTGAAGACATACC
CTAGTCAATAATCAATGTCAACG-3'

15

C14-7D:

se utilizó C14-2D.

C14-7R:

se utilizó C14-1R.

C14-8D:

20

se utilizó C14-3D.

C14-8R:

5' -AAATTATTTCTGCTGGTGGGCAATATTAGAATATGGGGAATGTTTGCTTCTGAGCCT
AGTCAATAATCAATGTCAACG-3'

C14-9D:

se utilizó C14-4D.

25

C14-9R:

se utilizó C14-3R.

C14-10D:

se utilizó C14-2D.

C14-10R:

se utilizó C14R.

C14-11D:

se utilizó C14-3D.

5 C14-11R:

se utilizó C14-2R.

C14-12D:

se utilizó C14-4D.

C14-12R:

10 se utilizó C14-8R.

C14-13D:

se utilizó C14-3D.

C14-13R:

se utilizó C14-1R.

15 C14-14D:

se utilizó C14-4D.

C14-14R:

se utilizó C14-2R.

20 En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de A2, A2-1 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 3000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-2 corresponde a la secuencia polinucleotídica de los nucleótidos 2801 a 5800 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-3 corresponde a la secuencia polinucleotídica de los nucleótidos 5401 a 8450 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-4 corresponde a la secuencia polinucleotídica de los nucleótidos 701 a 2700 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-5 corresponde a la secuencia polinucleotídica de los nucleótidos 701 a 2200 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-6 corresponde a la secuencia polinucleotídica de los nucleótidos 701 a 3700 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-7 corresponde a la secuencia polinucleotídica de los nucleótidos 2001 a 5000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-8 corresponde a la secuencia polinucleotídica de los nucleótidos 4001 a 7000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-9 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 3700 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-10 corresponde a la secuencia polinucleotídica de los nucleótidos 2001 a 5800 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-11 corresponde a la secuencia polinucleotídica de los nucleótidos 2801 a 7000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-12 corresponde a la secuencia polinucleotídica de los nucleótidos 701 a 5800 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-13 corresponde a la secuencia polinucleotídica de los nucleótidos 2001 a 7000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-14 corresponde a la secuencia polinucleotídica de los nucleótidos 2801 a 8450 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-15 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 5800 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-16 corresponde a la secuencia polinucleotídica de los nucleótidos 701 a 7000 de la SEQ ID NO: 10 en el Listado de Secuencias; A2-17 corresponde a la secuencia polinucleotídica de los nucleótidos 2001 a 8450 de la SEQ ID NO: 10 en el listado de secuencias.

40 En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de A7, A7-1 corresponde a la secuencia polinucleotídica de los nucleótidos 601 a 3600 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-2 corresponde a la secuencia polinucleotídica de los nucleótidos 3601 a 8420 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-3 corresponde a la secuencia polinucleotídica de los nucleótidos 5401 a 8420 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-4 corresponde a la secuencia polinucleotídica de los nucleótidos 3401 a 6400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-5 corresponde a la secuencia polinucleotídica de los nucleótidos 1501 a 4500 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-6 corresponde a la secuencia polinucleotídica de los nucleótidos 4401 a 7400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-7 corresponde a la secuencia polinucleotídica de los nucleótidos 2401 a 5400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-8 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 3600 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-9 corresponde a la secuencia polinucleotídica de los nucleótidos 1501 a 5400 de la SEQ ID NO: 11 en el Listado de

Secuencias; A7-10 corresponde a la secuencia polinucleotídica de los nucleótidos 2401 a 6400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-11 corresponde a la secuencia polinucleotídica de los nucleótidos 3401 a 7400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-12 corresponde a la secuencia polinucleotídica de los nucleótidos 4401 a 8420 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-13 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 5400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-14 corresponde a la secuencia polinucleotídica de los nucleótidos 1501 a 6400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-15 corresponde a la secuencia polinucleotídica de los nucleótidos 2401 a 7400 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-16 corresponde a la secuencia polinucleotídica de los nucleótidos 3401 a 8420 de la SEQ ID NO: 11 en el Listado de Secuencias; A7-17 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 6400 de la SEQ ID NO: 11 en el Listado de Secuencias; y A7-18 corresponde a la secuencia polinucleotídica de los nucleótidos 1501 a 7400 de la SEQ ID NO: 11 en el listado de secuencias.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de A18, A18-1 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 5040 de la SEQ ID NO: 12 en el Listado de Secuencias; A18-2 corresponde a la secuencia polinucleotídica de los nucleótidos 1001 a 6002 de la SEQ ID NO: 12 en el Listado de Secuencias; A18-3 corresponde a la secuencia polinucleotídica de los nucleótidos 2001 a 7000 de la SEQ ID NO: 12 en el Listado de Secuencias; y A18-4 corresponde a la secuencia polinucleotídica de los nucleótidos 3000 a 7000 de la SEQ ID NO: 12 en el listado de secuencias.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de B5, B5-1 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 4001 de la SEQ ID NO: 13 en el Listado de Secuencias; B5-2 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 3200 de la SEQ ID NO: 13 en el Listado de Secuencias; B5-3 corresponde a la secuencia polinucleotídica de los nucleótidos 2491 a 5601 de la SEQ ID NO: 13 en el Listado de Secuencias; B5-4 corresponde a la secuencia polinucleotídica de los nucleótidos 5373 a 8401 de la SEQ ID NO: 13 en el Listado de Secuencias; B5-5 corresponde a la secuencia polinucleotídica de los nucleótidos 901 a 4001 de la SEQ ID NO: 13 en el Listado de Secuencias; B5-6 corresponde a la secuencia polinucleotídica de los nucleótidos 4001 a 7000 de la SEQ ID NO: 13 en el listado de secuencias.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de C14, C14-1 corresponde a la secuencia polinucleotídica de los nucleótidos 960 a 4015 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-2 corresponde a la secuencia polinucleotídica de los nucleótidos 1987 a 5014 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-3 corresponde a la secuencia polinucleotídica de los nucleótidos 4020 a 7119 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-4 corresponde a la secuencia polinucleotídica de los nucleótidos 960 a 8141 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-5 corresponde a la secuencia polinucleotídica de los nucleótidos 960 a 6011 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-6 corresponde a la secuencia polinucleotídica de los nucleótidos 4939 a 8141 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-7 corresponde a la secuencia polinucleotídica de los nucleótidos 960 a 5014 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-8 corresponde a la secuencia polinucleotídica de los nucleótidos 2994 a 7119 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-9 corresponde a la secuencia polinucleotídica de los nucleótidos 4020 a 8141 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-10 corresponde a la secuencia polinucleotídica de los nucleótidos 1 a 5014 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-11 corresponde a la secuencia polinucleotídica de los nucleótidos 1987 a 7119 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-12 corresponde a la secuencia polinucleotídica de los nucleótidos 2994 a 8141 de la SEQ ID NO: 14 en el Listado de Secuencias; C14-13 corresponde a la secuencia polinucleotídica de los nucleótidos 960 a 7119 de la SEQ ID NO: 14 en el Listado de Secuencias; y C14-14 corresponde a la secuencia polinucleotídica de los nucleótidos 1987 a 8141 de la SEQ ID NO: 14 en el listado de secuencias.

Los puntos de inicio y final de los respectivos fragmentos en la secuencia de longitud completa también se muestran en las figuras 20 y 21.

(10-2) Evaluación de elementos de ADN que tienen diferentes longitudes de secuencia

Se evaluó cada plásmido construido en (10-1) utilizando la célula huésped CHO-K1 (ATCC) y el reactivo de transfección Lipofectamine 2000.

De la misma forma que en (7-3), se realizó selección de antibióticos con higromicina después de la transfección, con lo que se estabilizó una línea celular policlonal que expresaba de manera estable. La línea celular así establecida se sometió a reemplazo de medio el día antes de la medición y se sembraron un número dado de células en una placa de 24 pocillos. A las 24 horas del cultivo en placas de las células, se recogió el sobrenadante del cultivo y se midió la actividad de SEAP.

Los resultados de la medición se muestran en las figuras 10, 12, 14, 16 y 18. Se confirmó que no solo los elementos de ADN de longitud completa, sino también clones que tienen una longitud de secuencia más corta que los de longitud completa, tienen en efecto en la mejora de expresión. Según estos resultados, se confirmó que los elementos de ADN A2, A7, A18, B5 y C14 tienen la actividad de mejorar la expresión de genes extraños incluso en los casos en los que tienen una longitud de secuencia más corta que los de longitud completa. Sin embargo, exhiben el mayor efecto cuando la longitud de secuencia es la longitud completa.

(Ejemplo 11) Efecto utilizando células huésped distintas de la línea celular CHO

Se utilizó la línea celular CHO como la línea celular en la evaluación en los ejemplos 7 a 10. Sin embargo, en el Ejemplo 11, se seleccionó una línea celular HEK293 como una línea celular diferente de la línea celular CHO. La línea celular HEK293 se sometió a cultivo estático a 37 °C en presencia de CO₂ al 5 % utilizando el medio DMEM (Invitrogen) que contiene FCS al 10 % y se sembraron un número dado de células en una placa de 6 pocillos el día antes de la transfección. Para evaluar el vector de expresión SEAP que contenía cada elemento de ADN construido en (8-2), se realizó transfección utilizando cada plásmido y el reactivo de transfección Lipofectamine 2000 (Invitrogen). Se realizó la selección de antibióticos con higromicina durante aproximadamente 2 semanas empezando 2 días después de la transfección, con lo que se estabilizó una línea celular policlonal que expresaba de manera estable. La línea celular así establecida se sometió a reemplazo de medio el día antes de la medición y se sembraron un número dado de células en una placa de 24 pocillos. A las 24 horas del cultivo en placas de las células, se recogió el sobrenadante del cultivo y se midió la actividad de SEAP. Se midió la actividad de SEAP en el sobrenadante del cultivo utilizando el ensayo indicador de fosfatasa alcalina pNPP Sensolyte™ (ANASPEC).

Los resultados de la medición se muestran en la figura 19. De la misma manera que en el ejemplo 3, se confirmó que cada elemento de ADN es también altamente efectivo en la mejora de la expresión de un gen extraño (SEAP) en la línea celular HEK293.

Aplicabilidad industrial

Mediante la introducción de la unidad de expresión de genes extraños utilizando un promotor de acuerdo con la invención o el vector de expresión de genes extraños de acuerdo con la invención, en células huésped de mamífero, se hace posible mejorar la producción de un gen extraño de una proteína terapéutica, un anticuerpo o similar.

Listado de secuencias

- <110> DAIICHI SANKYO COMPANY, LIMITED
- <120> Promotores de genes humanos
- <130> DSPCT-FP1232
- <160> 16
- <170> Versión 3.4 de PatentIn
- <210> 1
- <211> 2348
- <212> ADN
- <213> *Homo sapiens*
- <400> 1

ES 2 670 894 T3

atttatgtat attaacagca cattaacagc taaaaagaaa aactcacata atcatattag 60
 ttcatagaaa caatgcattt gacatactcc aacatccatt catgatttta ttttattttg 120
 tatttatata ttttttttga gatggagtct cgctgtcacc caggctggag tgcaatggct 180
 cgatctcggc tcaactgcagg ctccgcccc ggcggttcac gccattctcc tgcctcagcc 240
 tcccgagtag ctgggactac aggggccccg cacctgccc agctaatttt ttgtattttt 300
 agtagagacg gggtttcacc gtgtagcca ggatggctc gatctcctga cctcgtgatc 360
 caccgcctc ggcctcccaa agtgctggga ttacaggcgt gagccaccgt gcccggccca 420
 tgattttaaa aaaacctctc agaaatagaa acagagggaa cttcctgaat ttgattaaaa 480
 atacctgcaa aatcctagag ctaatattat acttaatggt gaaagactga atggttttcc 540
 cctaagatgg agaacaaggc acggatgtcc ttgctcacca ctctattca acaaaggact 600
 ggaaaccaga aataacttta attttttttt ttctcaagt ataggttcac agagaaaaag 660
 ctttactgga tgaactttta gattactact tttatagagc agcagagata aaagccaggt 720
 cgaaaagtgc atgtggagta aggaaatgga cctagttcga caaaagggt cagaacgact 780
 gccagatga gattgtagac gcagctgtag tttactttct atctggaaga aacttcaagt 840
 tctcttaat tttccaagga gacagtcact taccttttaa aaaacattat tagagaagca 900
 ccgggcggga gcatatctag cattaanaat gtgggatgaa taccatctct gcttggtaaa 960
 ggtggttggg aatcctgaga gagggcacct agtgtggctt ctgcattttt cacagtgcct 1020
 ggaccacggc tgaaagtaac tcttgcatga catttgacag aagaggaaac cgaagctcag 1080
 attaaattcc ctgtccacag cgggatcatt cggcacgagt tcctccctgt ctggaatgct 1140
 cttccccagc aatagatcct ggggtcttc cgtgtctcag tctaattgtca ttccgttcca 1200
 ggattcccga cttcttaaag cataaataat cctccccac cctctcattg tactgttatg 1260
 taacttatta caatatgtca ttatatattt agtcatactg ctttaggtaa tgtcttctcc 1320
 actgaactgt aagctccatg agggcaagag ttcagtcggt tttacttaat aattagcacc 1380

ES 2 670 894 T3

tagtacagta ctagcataga atgaaggcct cgcaatTTTT tttaaattta tttttagaca 1440
 gggctcttgcg ctgtcgccca ggctggagtg cagtgggtgca acctcggctc acggcagcct 1500
 cgacctttcg gctccagcga tcctcccgcg tggcctccg gggtagctgg gactgcaggc 1560
 gcgcaccacc atgactggct aaaaaaaaa tttttttttt tgtagacatg gggctctcgcc 1620
 atgttgccca ggctggttcc tgagctcaag tgatcctcct gcctcggcct cccaaagtgc 1680
 tgggattaca ggcgtgagcc tcagcgccca gccaaagttag ctttttttaa acgtcctgtc 1740
 tccggagggtt gccgaagttg gttttcttcg gcctccttct ctctcccagg cccagggtg 1800
 ggacgaggcc ggttcccgcg tgcaacctgc actgaagacg ggaaccttgg gagccggtac 1860
 cggaacgctc ggaaacggca ccaaagtacg aatcctaggg cggaaaagcg ttaccaagac 1920
 actcgtcccc agagccgctt cctgggactc tctagcctcc taccgcttct cagtgatggt 1980
 ccggtttccg ccctcctcct cgcgctgttt ccgcctcttg ccttcggacg ccggattttg 2040
 acgtgctctc gcgagatttg ggtctcttcc taagccggcg ctccggcaagg taggttggcg 2100
 gcctgctctc cgacagaact tttcttcttg ggttgaggaa aacgcctttt ggagtcaggc 2160
 cctggagggg cgagccttgc tcacaggggtg gggatacagc cgattaccocg ccctgtgctt 2220
 tccgatggct tctgcggggc gagcggggcc tggccggggg gtgcggggcg gagggcgagc 2280
 cagcggcgcc tgcagcccgg gccgcgtaac gctgaccgct gtgccttcag ttctcccagg 2340
 agaaagcc 2348

<210> 2
 <211> 3345
 <212> ADN
 <213> *Homo sapiens*

5

<400> 2

ctaaagtgat tcctaaagaa ttcttccctt ttatcacttc cagtaggcct ctgtgaaacc 60
 aaatctacct ccgcttacia gaaagatgct gggctggcct tctctcaaag tctttccaaa 120
 cttttcttgg cattgactta gacaccctag gaatctaact tgagaaaatg ttttcattaa 180
 aaaaaatctc aggaagttaa acctcctgaa tgattactga gttgacataa atcttatgtg 240
 tatattctta tcagaaaaaa agtatcttca ttttgtggga caccaattca tgtattatta 300
 ttattttgag acaaagtttc gctcttggtg cccaggctgg agtgcaatgg cgcgatctcc 360
 acttactgca acctccacct cctgagttca agtgattctc gtgcctcagc ctccctagta 420
 gctggaacta caggcatgtg ccaccacacc cagctaattt tttgtaactt tagtagagat 480
 ggggtttcac catgttggcc aggatggact cgaactcctg accacagggtg atctgcccac 540
 ctgagcttcc caaagtgtg ggattacagg catgagccac cgcgcccagt cgctgggtct 600
 tacagtaact ttatgtttaa cattttgagg aaatgctatt cttttccaaa gtgactgcac 660

ES 2 670 894 T3

catttcatat ttgcactagc actgtacgga cattoccatt tctctgtcct agtgagtgtg 720
 aaatggatc tcaactgcagt tccagtttgt atttccctga tggctaataga tgtggatcat 780
 ttcattgtgtt cattggccac agagaaatgt ctatattggat tctttacca tttttcaatt 840
 gggttatttg tctttatagg tttgttggtg ttgagacaga gtcttgctct gtcactcagg 900
 ctggagtgca gtggcattat cacagctaac tgcagtctag aactgctggg ctcacgtgat 960
 catcccagct cagcctctcg agtaactggg actacaggca tgcgccacca gccccagcta 1020
 attattttat tttttgtaga gacaggtct tactatggtg cctaggccgg tcttgaactc 1080
 ctgggctcaa gcaaatctcc cacctcagac tcccaaagta ttggaattat aggtgtgaac 1140
 catagtgtc agccaatttg cacaataatc ttaaatacaa aagctaagca aaacaaatca 1200
 agagcatctt taaaaactag gcagtctggg aggcaggggc tgccgtgagc cgtgagatgg 1260
 cacctttgca ttccagccta ggtgacagag ggaggccctg tctaaaaaaaa accaaaaacc 1320
 aaaaaacaaa acaaaaacaaa aaacatctag gcagtagctc gtgcccgtaa tcccagctac 1380
 tcaggaggct gaggcgagag aatcgtttga gcccaggagt tcaagaccag cctgggcaac 1440
 agagtgagc cccatttcta aaaaatgaac aaagaaaaac taggcagttt cgcccagtg 1500
 ttagaagcgt ggagtttga gtcaagtctc caaatttcat cttccacata tgcaaatgg 1560
 agacaataat aggggtacgt tatagaattg tggtaggcat agtgaactcc atcgcatgtt 1620
 agctgttttc gttactattt actgtctaaa ttcggtgatg aaattattag gaagtctctg 1680
 tcttgttctc ttctgaccac taagaggcgc acttcggagt agaagaaacg cgggcggaaa 1740
 tagcccaaaa gcggattggc ttcgacttct ggcggaagta aattcctccc tccaccaggt 1800
 cttattagct cagaaagaat tccaaatttc tacgtagtcc caaggatagg tagaatacat 1860
 ttctcagtcc tattcctagt tattattgtc tattaacaa tgtatactca gaatttttgc 1920
 ggcattattt tttgacgtgt ctttatttta tttaaaagag ccggagccgg aagtgttgc 1980
 cttttccct gctaggacc aggggttacg acccatcagc ccttgccgc caccgtccct 2040
 tctctctcc tcggcgctgc ctacggagggt ggcagccatc tccttctcg taagtgttaa 2100
 tccgtggcaa tccgcattcc tgcgggattc atctggcccc gtcgccaggt ggtgcccagg 2160
 cctcccctc agcgcggtag tgtctgtggg tattgttatt gtcagcttac tggagcgtgt 2220
 acaggaacag aacgaagccg ccgagttgat agggctttgc gtcccagagc ctctgcct 2280
 ccgcctgat tcagagctgc gggctgcttg tttgttctt ggcggtggag ggtgctagtt 2340
 gaggccagc ttcggggtct cctgggggccc gtgggacgac caggggtggc ccagcttgac 2400
 agctttcagc tgggatctgt ggatcccagc gctcaccaat gtcggccac gtgtattcgt 2460
 tcatgccatg gccggcttct tccgctgcag tctctggccc gagggctgct gctgcccggac 2520

ES 2 670 894 T3

cgccaaggaa agacgagctg taggtcggct ggtccagctg caggcagaaa ttctggtagt 2580
atctctggga atatgaagat gcaactgccc ccaccttgcc ttcgaggata tcatgggcca 2640
gaaggcagag tcgttttgaa tacgtgggtc attgagtacc cactctgggc cagttgatgg 2700
ctgccaagag agcagaaggg gtgctgctgt aggaaatcaa tggctcggaa gaccacactg 2760
aggaaggtgt gagttgatac tgggaagatct ccaggtttga ggcacattca gaggtatatg 2820
gtggttttgt gtgtgttgag ggtgtggtag cgcagcagct ccctagggaa ttagaaggtt 2880
ttattgaaca ttaccctgt gacaggcact gcaggcattc agcgcgcagt gtcacattca 2940
ttttacaggt gaggaaaaga ctcaggttca agtagatggt caaggccagt actaccggaa 3000
ggaccatctg ggggttcgga cactggtggg gtgggatttg ctgccccttg caaattgaga 3060
gtgtcttggg gtcagttttg atttgctcag ctggtggcat tctttgggct ctgagtgggt 3120
gaggtgacct ttgacctct gggatcgcat ctggagagtg cctagtattc tgccagcttc 3180
ggaaagggag ggaaagcaag cctggcagag gcaccattc cattcccagc ttgtccgta 3240
gctggcgatt ggaagacact ctgogacagt gttcagtccc tgggcaggaa agcctccttc 3300
caggattctt cctcacctgg ggccgcttct tccccaaaag gcac 3345

<210> 3
<211> 3337
<212> ADN
<213> *Homo sapiens*

5

<400> 3
atggtggcac aatcatggtt cactgcagcc tcaaacttct aggctcaagt gatcctccag 60
cctcagcctc ctgagtaggt gggactaaag gtgtgtgcta caacacctgg ctaattaa 120
aaaaaaaaatt tgtagagatg ggcgtctcgc tatggtgccc aggctgatct ttaactacag 180
gtctcagtga tctttctgta tcagcctccc gaagtgctag gattaacagg catgggccac 240
tgtacctggc tgtctttgaa gtttttaata agattgccat atctctctca agactgatca 300
agaaaaaatg caaatacaaa tttaaagaga gggaaaaata gtattatgga actgggttaa 360
gtaccttaag ggaaaactac aacaatgaga ccactattct tatctacttg ggtgggggaa 420
ggcacggaca gagacgtaaa caacagcata gtgtgccaag ggctcccatc tcctgttctc 480
ttctgctctg caaagtctct cttaaaatat agagaacata ttgcaagtaa atacattcat 540
tagtgcattt attcaaaaac cacatttgtc tctctatgca agcttccatc tcaatgcctg 600
gcacagagac gaaaaggatt tcagaggaca aaaatcaaag gactggaaag gagggaagta 660
gctgatgaca aggtcggggg agaagcagct atttgagagt gagcagaggg aaggatataa 720
ccagagcctt ccatctggtg ttgtgacaaa gatggcagtg atgtgaggct tgatagaatc 780
aataggcctc acagttcaag acagactaca tgccctcagt cacttgctctg ctttcttgggt 840

ES 2 670 894 T3

ttgctagatg gattagaaat actagaagga aagttgagtg ggtactggga tatectatcc 900
 aaatcctctc ttctggggca atgcagctgc ctcacagcaa ccacgttggt gatcacctga 960
 ccccgccctt ctcagcttcc tggaggtgca ttctaagagc gctccccata agccaacact 1020
 caagtctctt cgaggtgtgt ttccaggag cctgatctaa gattaaaaac ttaaaggccc 1080
 cttttagaat tgtcttttac agcccaggca aagagttctt tactttttta gaggtcatga 1140
 atcccttga caaggtgatg aaagccaggg accctcttct cagaaaaata cacacgtgcg 1200
 cttaatgtaa aattttacat actgttttaa aggttaaact ttaccacca tgaaaagcct 1260
 caagtagttc tttgtacata atcataagat tagcaacat ttactgagca cttcctctct 1320
 ataagaccct gtgctaagtg ctttaacttc ataatgcctt tgatcatgac ataacatggc 1380
 acggtagttc ctaatatctt ccttttagag atttaagggc ttgctcaaag taacatagct 1440
 aataattagc agacactgga ttaaaatccc aatttgttg taaagcttgt gctctgaata 1500
 aatgacaagg aaagagaagg gaaggtgaa gaaggaagg ttcttgaagg tccttgggcc 1560
 ttgaaggtcc cttaacata aatgtcaaga gttgggattt aaaccaggt ctaacgccag 1620
 agctggcgcc ctttagatta aaagtgcagt gtccatgaca acgaaagaag ttgattttgt 1680
 cccacctttg ctctttgcgg cttttcattt gcgtttggtt ccacagcgat ttccaataga 1740
 tttctgctg gccttgacac agacagctag tgtgaatccc cgcccacaga gggcgggcac 1800
 gttggttgc gtacaacgtg gtggttccct gcatctctgc ccacgtcgga gaggtgctc 1860
 ggcttccgta caacacggat actctctctc tgacgcaact tcctgtcctg cgcaattcta 1920
 tttgacctt gaactggcaa aggctttttt cttcctctc cggggacgtt gtctgcaggt 1980
 atggatgttg ttctcttttc cctgtcttta tttccttacc aatcggctgc catccgagga 2040
 gctgaggaag cctagagctc tcagaagcag tcctttgagc tgggtgtagg gtaaggggca 2100
 caacaggag gttggtggtg aggaagtcc ttactttgat ctttgaaat cccttgttcc 2160
 tgggtggacc tccaaagccg tgagtagcca cagctacca cccgggactt tgctgcattc 2220
 caagtgtagc gtttgagac taacgagttg tggtttggcg gtttgagtct ggaaaatcgc 2280
 caaacgttt catattttac acccacgtt tcacagcacg cctgtacgtg tccttagtct 2340
 ttgggagggc agggccggc gagttcgggt ggtttcgcta tttggcttct gcgtccaagg 2400
 cccatgtcaa ggaagagaaa aatgtgttag aagtttctgt cttgcttttg gagatgcaaa 2460
 cagaataatg gcttcataaa tcactcgac tggttttacg tgtcaagttt tgggtgtctgg 2520
 taattctgtt ttagtttaaat tttagtgaga ggcttgtgac aacaaatgag gtggttacia 2580
 ggggtggaat ggaagatta aattagtca agtattgatt atgttttacg ttgggtagtt 2640
 cccttaacga agttgctcgt atgcatatct gtataaccga tttgctaaat aacatcacga 2700
 tgtttccaga agtgggaaga aagcaggtgc cataacccaa agaaacttgt gtaatatcaa 2760

ES 2 670 894 T3

aattagtatt aaagggatg cctttacgca ggtgggtgctt tagggcaaga cattgaaccc 2820
tgatatgtgc caggcattgt gttgggacgg atagcccacg tcgtttaatc ctaatgacag 2880
ctgtataaag tagacagaat tcccatgtta gagataagga ggctagctcc ttgcccttta 2940
tattcccagt aaatggcata gctaaggatt cgacttcaga gctcactttt tgtgctcttt 3000
gtttaaagcg gtgtttctcc aactgggctc gtggcacttt tctggacacc actccaaaca 3060
aaattagact ctgagtaagg agcctggtca tcagaacggt aaggaagtgc cacgtttgat 3120
taccatcagg aaagctaaca ttcttggcct cttgtttatc agtcaccttt aaatacaagt 3180
agttttaaaa tgtggaataa tacatcttaa ttttaagggg ttacatacaa ggatatgtat 3240
gtgaatgaaa tagaccacat gatactgttt tgagatttta tttactttta caatggaaag 3300
at ttgatgtt actctattct taatttaggc actcaga 3337

<210> 4
<211> 1361
<212> ADN
<213> *Homo sapiens*
<400> 4

5

ES 2 670 894 T3

cctagtgtgg cttctgcatt tttcacagtg cctggaccac ggctgaaagt aactcttgca 60
 tgacatttga cagaagagga aaccgaagct cagattaaat tcctgtcca cagcgggatc 120
 attcggcacg agttcctccc tgtctggaat gctcttcccc agcaatagat cctgcggctc 180
 ttccgtgtct cagtctaatag tcattccggt ccaggattcc cgacttctta aagcataaat 240
 aatccctccc caccctctca ttgtactggt atgtaactta ttacaatatg tcattatata 300
 tttagtcata ctgcttttagg taatgtcttc tccactgaac tgtaagctcc atgagggcaa 360
 gagttcagtc ggttttactt aataattagc acctagtaca gtactagcat agaatgaagg 420
 cctcgcaatt ttttttaaata ttatttttag acagggctct gcgctgtcgc ccaggctgga 480
 gtgcagtggt gcaacctcgg ctcacggcag cctcgacctt tcggctccag cgatcctccc 540
 gcgtcggcct ccggggtagc tgggactgca ggcgcgcacc accatgactg gctaattttt 600
 tttttttttt tttttagtagc atggggctct gccatgttgc ccaggctggt tcctgagctc 660
 aagtgatcct cctgcctcgg cctcccaaag tgctgggatt acaggcgtga gcctcagcgc 720
 ccagccaagt tagccttttt taaacgtcct gtctccggag gttgccgaag ttgggtttct 780
 tcggcctcct tctctctccc aggccaggg ctgggacgag gccggttccc gcctgcaacc 840
 tgcactgaag acgggaacct tgggagccgg taccggaacg ctcggaacg gcaccaaagt 900
 acgaatccta gggcggaaaa gcgttaccaa gacactcgtc ccagagccg cttcctggga 960
 ctctctagcc tcctaccgct tctcagtgat gttccggttt ccgccctcct cctcgcgctg 1020
 tttccgcctc ttgccttcgg acgccggatt ttgacgtgct ctcgcgagat ttgggtctct 1080
 tcctaagccg gcgctcggca aggtaggttg gcggcctgct ctccgacaga acttttcttc 1140
 ttgggttgag gaaaacgcct tttggagtca ggccctggag gggcgcgcct tgctcacagg 1200
 gtggggatac agccgattac ccgccctgtg ctttccgatg gcttctgcgg ggcgagcggg 1260
 gcctggccgg ggggtgcggg cgggagggcg agccagcggc gcctgcagcc cgggcccgt 1320
 aacgctgacc gctgtgcctt cagttctccc aggagaaagc c 1361

<210> 5
 <211> 867
 <212> ADN
 <213> *Homo sapiens*

5

<400> 5

ES 2 670 894 T3

cctcggctca cggcagcctc gacctttcgg ctccagcgat cctcccgcgt cggcctccgg 60
 ggtagctggg actgcaggcg cgcaccacca tgactggcta attttttttt tttttttttt 120
 gtagacatgg ggtctcgcca tgttgcccag gctggttcct gagctcaagt gatcctcctg 180
 cctcggcctc ccaaagtgct gggattacag gcgtgagcct cagcgcaccag ccaagttagc 240
 cttttttaaa cgtcctgtct ccggagggtg ccgaagttgg ttttcttgg cctccttctc 300
 tctcccaggc ccagggtggt gacgaggcgg gttcccgcct gcaacctgca ctgaagacgg 360
 gaaccttggg agccggtacc ggaacgctcg gaaacggcac caaagtacga atcctagggc 420
 ggaaaagcgt taccaagaca ctcgtcccca gagccgcttc ctgggactct ctagcctcct 480
 accgcttctc agtgatgttc cggtttcggc cctcctcctc gcgctgtttc cgcctcttgc 540
 cttcggacgc cggattttga cgtgctctcg cgagatttgg gtctcttctt aagccggcgc 600
 tcggcaaggt aggttggcgg cctgctctcc gacagaactt ttcttcttgg gttgaggaaa 660
 acgccttttg gagtcaggcc ctggaggggg gagccttgct cacaggggtg ggatacagcc 720
 gattaccggc cctgtgcttt ccgatggctt ctgcggggcg agcggggcct ggccgggggg 780
 tgcgggcggg agggcgagcc agcgggcgct gcagcccggg ccgcgtaacg ctgaccgctg 840
 tgccttcagt tctcccagga gaaagcc 867

<210> 6
 <211> 2372
 <212> ADN
 <213> *Homo sapiens*

5

<400> 6
 cctctcgagt aactgggact acaggcatgc gccaccagcc ccagctaatt attttatttt 60
 ttgtagagac agggctttac tatgttgctt aggccggtct tgaactcctg ggctcaagca 120
 aatctcccac ctcagactcc caaagtattg gaattatagg tgtgaacat agtgctcagc 180
 caatttgcac aataatctta aatacaaaaag ctaagcaaaa caaatcaaga gcctctttaa 240

ES 2 670 894 T3

aaactaggca gtctgggagg caggggctgc cgtgagccgt gagatggcac ctttgcatte 300
 cagcctaggt gacagagggg ggcctgtct aaaaaaaccc aaaaaccaa aaacaaaaca 360
 aaacaaaaaa catctaggca gtagctcgtg cccgtaatcc cagctactca ggaggctgag 420
 gcgagagaat cgtttgagcc caggagtca agaccagcct gggcaacaga gtgagacccc 480
 atttctaaaa aatgaacaaa gaaaaactag gcagtttcgc ccagtggta gaagcgtgga 540
 gtttgagatc aagtctccaa atttcatctt ccacatatgc aaaatggaga caataatagg 600
 ggtacgttat agaattgtgg taggcatagt gaactccatc gcatgttagc tgttttcggt 660
 actatttact gtctaaattc ggtgatgaaa ttattaggaa gtctctgtct tgttctcttc 720
 tgaccactaa gaggcgcact tcggagtaga agaaacgcgg gcggaaatag cccaaaagcg 780
 gattggcttc gacttctggc ggaagtaaat tcctccctcc accaggtctt attagctcag 840
 aaagaattcc aaatttctac gtagtcccaa ggataggtag aatacatttc tcagtcctat 900
 tcctagttat tattgtctat taaaacatgt atactcagaa tttttgcggc attatttttt 960
 gacgtgtctt tattttattt aaaagagccg gagccggaag tgcttgcctt tttccctgct 1020
 aggaccagc ggttacgacc catcagccct tgcgcgccac cgtcccttct ctcttcctcg 1080
 gcgctgccta cggaggtggc agccatctcc ttctcggtaa gtgttaatcc gtggcaatcc 1140
 gcattcctgc gggattcatc tggccccgtc gcccagtggt gcggagccct ccccttcagc 1200
 gcggtagtgt ctgtgggtat tgttattgtc agcttactgg agcgtgtaca ggaacagaac 1260
 gaagccgccg agttgatagg gctttgcgtc ccagagcctc ctgccctccg cctgtattca 1320
 gagctgcggg ctgcttgttt gttccttggc ggtggagggt gctagttgag gccagacttc 1380
 ggggtctcct gggggccgtg ggacgaccag ggggtggcca gcttgacagc tttcagctgg 1440
 gatctgtgga tcccagcgtc caccaatgtc ggcccacgtg tattcgttca tgccatggcc 1500
 ggcttcttcc gctgcagtct ctggcccag ggtgctgtct gcgggaccgc caaggaaaga 1560
 cgagctgtag gtccgctggt ccagctgcag gcagaaatc tggtagtacc tctgggaata 1620
 tgaagatgca actgccccca ccttgccttc gaggatatca tggccagaa ggcagagtcg 1680
 ttttgaatac gtggttcatt gagtaccac tctgggccag ttgatggctg cgaagagagc 1740
 agaaggggtg ctgctgtagg aatcaatgg ctccggaagac cactctgagg aaggtgtgag 1800
 ttgatactgg aagatctcca ggtttgaggc atcttcagag gtatatggtg gttttgtgtg 1860
 tgttgagggt gtggtagcgc agcagctccc tagggaatta gaaggtttta ttgaacattt 1920
 accctgtgac aggcaactgca ggcattcagc gcgcagtgtc atcttcattt tacaggtgag 1980
 gaaaagactc aggttcaagt agatgtcaa ggccagtact accggaagga ccatctgggg 2040
 gttcggacac tgggtggggtg ggatttgcgt ccccttgcaa attgagagtg tcttggggtc 2100
 agttttgatt tgctcagctg ttggcattct ttgggctctg agtgggtgag gtgacccttg 2160

ES 2 670 894 T3

acctcctggg atcgcatctg gagagtgcct agtattctgc cagcttcgga aaggagggga 2220
 aagcaagcct ggcagaggca cccattccat tcccagcttg ctccgtagct ggcgattgga 2280
 agacactctg cgacagtgtt cagtccctgg gcaggaaagc ctccctccag gattcttctt 2340
 cacctggggc cgcttcttcc ccaaaaggca tc 2372

<210> 7
 <211> 1862
 <212> ADN
 <213> *Homo sapiens*

5

<400> 7
 gcagtttcgc ccagtgggta gaagcgtgga gtttgagtc aagtctccaa atttcatctt 60
 ccacatatgc aaaatggaga caataatagg ggtacgttat agaattgtgg taggcatagt 120
 gaactccatc gcatgttagc tgttttcggt actatctact gtctaaattc ggtgatgaaa 180
 ttattaggaa gtctctgtct tgttctcttc tgaccactaa gaggcgcaact tccgagtaga 240
 agaaacgcgg gcggaaatag cccaaaagcg gattggcttc gacttctggc ggaagtaaat 300
 tcctccctcc accaggtctt attagctcag aaagaattcc aaatttctac gtagtcccaa 360
 ggataggtag aatacatttc tcagtcctat tcctagttat tattgtctat taaaacatgt 420
 atactcagaa tttttgcggc attatctttt gacgtgtctt tattttatatt aaaagagccg 480
 gagccggaag tgcttgcctt tttccctgct aggaccagcagg ggttacgacc catcagccct 540
 tgccgcgccac cgtcccttct ctcttctctg gcgctgccta cggagggtggc agccatctcc 600
 ttctcggtaa gtgttaatcc gtggcaatcc gcattcctgc gggattcacc tggccccgtc 660
 gcccagtggt gcggaggcct ccccttcagc gcggtagtgt ctgtgggtat tgttattgtc 720
 agcttactgg agcgtgtaca ggaacagAAC gaagccgcgg agttgatagg gctttgcgtc 780
 ccagagcctc ctgccctccg cctgtattca gagctgcggg ctgcttgttt gttccttggc 840
 ggtggagggt gctagttgag gccagacttc ggggtctcct gggggccgtg ggacgaccag 900
 gggtgggcca gcttgacagc tttcagctgg gatctgtgga tcccagcgtc caccaatgtc 960
 ggcccacgtg tattcgttca tgccatggcc ggcttcttcc gctgcagtct ctggcccag 1020
 ggctgctgct gcgggaccgc caaggaaaga cgagctgtag gtcggctggc ccagctgcag 1080
 gcagaaattc tggtagtata tctgggaata tgaagatgca actgccccca ccttgccttc 1140
 gaggatatca tgggcccagaa ggcagagtgc ttttgaatac gtggttcatt gagtaccac 1200
 tctgggcccag ttgatggctg cgaagagagc agaaggggtg ctgctgtagg aatcaatgg 1260
 ctccgaagac cacactgagg aagggtgtgag ttgatactgg aagatctcca ggtttgaggc 1320
 atcttcagag gtatatggtg gttttgtgtg tgttgagggt gtggtagcgc agcagctccc 1380
 tagggaatta gaaggtttta ttgaacattt accctgtgac aggcactgca ggcattcagc 1440

ES 2 670 894 T3

gcgcagtgtc atcttcattt tacaggtgag gaaaagactc aggttcaagt agatgggtcaa 1500
 gccagtact accggaagga ccatctgggg gttcggacac tgggtggggtg ggatttgctg 1560
 ccccttgcaa attgagagtg tcttgggggtc agttttgatt tgctcagctg ttggcattct 1620
 ttgggctctg agtgggtgag gtgacccttg acctcctggg atcgcacatctg gagagtgcct 1680
 agtattctgc cagcttcgga aagggagggga aagcaagcct ggcagaggca cccattccat 1740
 tcccagcttg ctccgtagct ggcgattgga agacactctg cgacagtgtt cagtccctgg 1800
 gcaggaaagc ctccttcag gattcttctt cacctggggc cgcttcttcc ccaaaaggca 1860
 tc 1862

<210> 8
 <211> 2363
 <212> ADN
 <213> *Homo sapiens*

5

<400> 8
 gcttcctgga ggtgcattct aagagcgctc cccataagcc aacactcaag tctcttcgag 60
 gtgtgtttcc agggagcctg atctaagatt aaaaacttaa aggcccttt tagaattgtc 120
 ttttacagcc caggcaaaga gttctttact tttttagagg tcatgaatcc ctttgacaag 180
 gtgatgaaag ccagggaccc tcttctcaga aaaatacaca cgtgcgctta atgtaaaatt 240
 ttacatactg ttttaaaggt taaactttac caccatgaa aagcctcaag tagttctttg 300
 tacataatca taagattagc aaccatttac tgagcacttc ctctctataa gaccctgtgc 360
 taagtgcttt aacttcataa tgctttgat catgacataa catggcacgg tagttcctaa 420
 tatcttcctt ttagagattt aagggcttgc tcaaagtaac atagctaata attagcagac 480
 actggattaa aatcccaatt tgtttgtaa gcttgtgtc tgaataaatg acaaggaaag 540
 agaaggaag gttgaagaag ggaaggttct tgaaggtccc tggtccttga aggtccctta 600
 accataaatg tcaagagttg ggatttaaac ccaggtctaa cgccagagct ggcgcccttt 660
 agattaaaag tgcagtgtcc atgacaacga aagaagttga ttttgtccca cctttgctct 720
 ttgcggtctt tcatttgogt ttgtttccac agcgatttcc aatagatttc tgcgtggcct 780
 tgacacagac agctagtgtg aatccccgcc cacagagggg cggcacgttg gttgccgtac 840
 aacgtggtgg ttccctgcat ctctgcccac gtcggagagg tgcgtcggct tccgtacaac 900
 acggatactc tctctctgac gcaacttct gtctcgcga attctatttg accttgaac 960
 tggcaaaggc tttttcttc ctcttcggg gacgttgtct gcaggtatgg atgttgttct 1020
 cttttccctg tctttatttc cttaccaatc ggctgccatc cgaggagctg aggaagccta 1080
 gagctctcag aagcagtcct ttgagctggt gtaggggtaa ggggcacaac agggaggtg 1140
 gtggtgagga agttccttac tttgatcttt ggaaatcctt tgttcttggg ggcacctca 1200

ES 2 670 894 T3

aagccgtgag tagccacagc tcaccaccog ggactttgct gcattccaag tgtagcgttt 1260
 ggagactaac gagttgtggt ttggcgggtt gagtctggaa aatcgccaaa cgttttcata 1320
 ttttacaccc acgttttcac agcacgcctg tacgtgtcct tagtctttgg gagggcaggg 1380
 tccggcgagt tccgggtggt tcgctatttg gcttctgcgt ccaaggccca tgtcaaggaa 1440
 gagaaaaatg tgttagaagt ttctgtcttg cttttggaga tgcaaacaga ataatggctt 1500
 cataaatcac tcgcactggt tttacgtgtc aagttttggt gtctggtaat tctgttttag 1560
 ttttaatttta gtgagaggct tgtgacaaca aatgagggtg ttacaagggg tggaatggga 1620
 agattaaatt agttcaagta ttgattatgt tttacgttgg gtagttccct taacgaagtt 1680
 gctcgtatgc atatctgtat aaccgatttg ctaaataaca tcacgatggt tccagaagtg 1740
 ggaagaaagc aggtgccata acccaaagaa acttgtgtaa tatcaaaatt agtattaaag 1800
 ggtatgcctt tacgcaggtg gtgctttagg gcaagacatt gaaccctgat atgtgccagg 1860
 cattgtgttg ggacggatag cccacgtcgt ttaatcctaa tgacagctgt ataaagtaga 1920
 cagaattccc atgttagaga taaggaggct agctccttgc cctttatatt cccagtaaatt 1980
 ggcatagcta aggattcgac ttcagagctc actttttgtg ctctttgttt aaagcgggtg 2040
 ttctccaact gggctcgtgg cacttttctg gacaccactc caaacaaaat tagactctga 2100
 gtaaggagcc tggatcatcag aacgttaagg aagtgccacg tttgattacc atcaggaaag 2160
 ctaacattct tggcctcttg tttatcagtc acctttaaat acaagtagtt ttaaaatgtg 2220
 gaataataca tcttaattta aggggtgttac atacaaggat atgtatgtga atgaaataga 2280
 ccacatgata ctgttttgag attttattta cttttacaat ggaaagattt gatgttactc 2340
 tattcttaat ttaggcactc aga 2363

<210> 9
 <211> 1858
 <212> ADN
 <213> *Homo sapiens*

5

<400> 9

gtaaagcttg tgctctgaat aatgacaag gaaagagaag ggaagggtga agaaggggaag 60
 gttcttgaag gtccctggtc cttgaaggtc ccttaaccat aatgtcaag agttgggatt 120
 taaaccaggg tctaaccgca gagctggcgc cctttagatt aaaagtgcag tgtccatgac 180
 aacgaaagaa gttgattttg tcccacctt gctctttgcg gcttttcatt tgcgtttgtt 240
 tccacagcga tttccaatag atttctgcgt ggccttgaca cagacagcta gtgtgaatcc 300
 ccgcccacag aggggocggca cgttggttgc cgtacaacgt ggtggttccc tgcattctctg 360
 cccacgtcgg agaggtgcgt cggcttccgt acaacacgga tactctctct ctgacgcaac 420
 ttctgtcct gcgcaattct atttgacct tgaactggca aaggcttttt tcttctctt 480

ES 2 670 894 T3

ccggggacgt tgtctgcagg tatggatggt gttctctttt ccctgtcttt atttccttac 540
 caatcggctg ccatccgagg agctgaggaa gcctagagct ctcagaagca gtcctttgag 600
 ctggtgtagg ggtaaggggc acaacagggg ggttggtggt gaggaagttc cttactttga 660
 tctttgaaa tcccttggtc ctggtggcac ctocaaagcc gtgagtagcc acagctcacc 720
 acccgggact ttgctgcatt ccaagtgtag cgtttggaga ctaacgagtt gtggtttggc 780
 ggtttgagtc tggaaaatcg ccaaacgttt tcatatttta caccacggtt ttcacagcac 840
 gcctgtacgt gtccttagtc tttgggaggg caggggtccgg cgagttccggg tggtttcgct 900
 atttggcttc tgcgtccaag gcccatgtca aggaagagaa aaatgtgtta gaagtttctg 960
 tcttgctttt ggagatgcaa acagaataat ggcttcataa atcactcgca ctggttttac 1020
 gtgtcaagtt ttggtgtctg gtaattctgt tttagtttaa ttttagtgag aggcttgtga 1080
 caacaaatga ggtggttaca aggggtggaa tgggaagatt aaattagttc aagtattgat 1140
 tatgttttac gttgggtagt tcccttaacg aagttgctcg tatgcatatc tgtataaccg 1200
 atttgctaaa taacatcacg atgtttccag aagtgggaag aaagcaggtg ccataaccca 1260
 aagaaacttg tgtaatatca aaattagtat taaagggat gcctttacgc aggtggtgct 1320
 ttagggcaag acattgaacc ctgatatgtg ccaggcattg tgttgggacg gatagcccac 1380
 gtcgtttaat cctaatgaca gctgtataaa gtagacagaa ttcccatggt agagataagg 1440
 aggctagctc cttgcccttt atattcccag taaatggcat agctaaggat tgcacttcag 1500
 agctcacttt ttgtgctctt tgtttaaagc ggtgtttctc caactgggct cgtggcactt 1560
 ttctggacac cactccaaac aaaattagac tctgagtaag gagcctggtc atcagaacgt 1620
 taaggaagtg ccacgtttga ttaccatcag gaaagctaac attcttggcc tcttgtttat 1680
 cagtcacctt taaatacaag tagttttaaa atgtggaata atacatctta atttaagggt 1740
 gttacataca aggatatgta tgtgaatgaa atagaccaca tgatactggt ttgagatttt 1800
 atttactttt acaatggaaa gatttgatgt tactctattc ttaatttagg cactcaga 1858

<210> 10
 <211> 8450
 <212> ADN
 <213> *Homo sapiens*

5

<400> 10
 attttgcttg aaaggatagc atcaaggaag tgaaatgaca acccacagaa tgagagataa 60
 tttttgcaaa tcatgtatct gataagggac ctgtagtcag aatatgcaaa gaacccttac 120
 aattcaataa gacaacccaa tttaaaaaca ggcaaaggat gtgaataggc atttctccaa 180
 agatacggaa aaacggccaa taagcacata aaaagatgct caaaatcatt tgccatttgg 240
 gaaatgcaat caaaaccaca atgaggtatc acttcacgcc cattaggggtg gctatagatc 300

ES 2 670 894 T3

agaaagtcag ataacatgtg ttggcaagca catggaaaca ctgaagtcct tacacactgc 360
 tggtaggaat gtaaaatggt gcagccactg tggaaaacag ttttccaatt tctcaaaatg 420
 ttaaacacag ttatcataca cccaagcaat tctactctta ggtatatacc caagagaaat 480
 gaaaacatat gtcttcacca gaacttgctg ttcacagcag cattatgcat aatagaccaa 540
 aagtggaaac aactcaactg cccatcaact ggtgaatgga taagtaaaat gtgatgtaac 600
 cagtcattgg actgtcattc attaataaaa agaacaaggt actgattcat gttctaacat 660
 gagtgaatct tgaaaacact atgctaaatt aaagaagcca gtcacaaaag gccgtgtatt 720
 gcatgatfff atatatacat gaacttttat atatataata ttatatatat tatatataat 780
 tttatatata taaatttcta tatataaata tataaaatca tatatatgat atatattttt 840
 tcatatacat catatatatt tacaaaaatt atatataata tatcatatga tatatgagat 900
 atatatacat atatatatga tatatgatat atatcatatg agatatatga tatcatgaga 960
 tatatgatat catatgatat atatgatata gatatacatat gatatatata taatatatat 1020
 atgatagata tattatatat gatagatatg atagatatca tattatatat gatagatatg 1080
 atagatatca tattatatat gatagatatata gatatacatat tatatatgat agatatgata 1140
 gatatacatat tatatatgat agatatgata gatatacatat tatatatgat agatatgata 1200
 gatatacatat tatatatgat agatatgata gatatacatat tatatatgat agatatgata 1260
 gatatacatat tatatatgat agatatgata gatatacatat tatatatgat agatatgata 1320
 gatatacatat tatatatgat agatatgata gatatacatat tatatatgat agatatgata 1380
 gatatacatat tatatatgat agatatgata gatatacatat tatatatgat agatatgata 1440
 gatatacatat tatatatgat atcatatata taccacatac atcatatata catcatatat 1500
 acatcatata tatcatacat atatatgaac tttccagaat aggtatatca ataaagacag 1560
 gaagtataca agtggttggc acagcctgag aggagcaggg aatggtgagt gactgctaata 1620
 ggatatggca ctttttttgg ggggtgatga aaatgttctg gtcagacaat ggcaattaca 1680
 aaactgtata cacacgaaaa accaaagaat cacacacttt aaaagggagg atttagctcg 1740
 gcatggtggc atgcgcctgt actcccagtt actcgggagg ctgaagcagg actgcttaga 1800
 gccaggact tcaaggctgc agcgagctat gatcgctcca ctgcactcca acaaggatga 1860
 cagtgcgaga cccgttttct aaataataat aataataata ataataaata acccaaggta 1920
 cccagttcac atgcaaaacc actggtaac ataaattatc tccaagtaat ctagaagaa 1980
 aatgagcaca taagacgtct tctaaaaaca cacatatatt tctttacatg ttacatttaa 2040
 cgtaaaaatc agctatgcag aagttacatg aacattttat gttggaaagg taaatgacta 2100
 ttattaatac agaatggtta agtacattta tgtttttatg tacaacgca taaaaggaaa 2160

ES 2 670 894 T3

agcatcctta aaataaacac catcaatggc tcctcgggtg tcacaaaaca aaatcctcac 2220
 acctttgtct tccttcacaa ttgagcttta tccacctttt caggcttatc tccattatt 2280
 acctgacaca aacttgggtg ggccagagtt tccactgacc atccccgac tattcatcca 2340
 aactatggt cactgcctcc cattcctgac catttgcctt ttgtcttcaa ctaattctgg 2400
 ggacgttttg tccaaataaa tgatccatat tcttgaaggc tggaatcaag tcctattaca 2460
 aatatatttt ctacacctct ccagagcata gcaaccagc atctactggc ctctcacagc 2520
 tctaaccatc cacaacccta agctggcttc tcatcaaacg ggtacttttc accacccaaa 2580
 ttcaattaat tcaactctac aataatgaag aatagtcgcc tacagcctac cttttccagc 2640
 cttgattcaa tcatttatca attttatctt caaagtccct cacttcaggg agatgatata 2700
 tcagctttca cccagagtcc taaagaaaac agcactcttg ccaatgacat agtgcacct 2760
 agtggcaaca taaggtaaat cacagtggca gtagaaggat ctccacacta cttttacagg 2820
 aatgcactgc aggtaaaaaa taagaagcta cagtactgtt tggcaggaca atttgtttca 2880
 tacgtgcata ctatcgccct gactaaatta actogcaagt cttacaggta ttatttgttt 2940
 tcagttccat gcacagatta gccatttagt acttactaaa tcaaactcaa tttctgaagt 3000
 gtcttacacc aatatattca tgcacatatg gttaaaattt tccttgagga tctatcatgt 3060
 gagagtgtgg cttattataa caagtaaca gaacaaataa atacaaaatg aaaagaaatc 3120
 gtatgattta ctogcatata agggagcttg ttgtggatta agtttcatga cccaggacac 3180
 tgaaacagaa atggaataaa tgagaataaa attaaaagtt gtcacaaaa atatagaagc 3240
 catctaaaga cctaggtgtc aagcatagct ctatgagtac aatcccgtgc ctgagattac 3300
 catatgcca gctgtatgct atacactaag agatttagga aggaagcggg gtcagggatt 3360
 gacccagac tccatctttt caagtgggga agaaagatct tccgattgaa aaataaaggc 3420
 aaaaaaggct tcaccgtcac agaagtttca acaaccaaca ggatatttaa aacagttatc 3480
 aaagcaaac cattgtatgt tcacttacat ttttcatag tcctcaaac tcacaaaatg 3540
 ctgtttactc agggacttct tccggtctta ctaggagacc tggaaagtga cgggaggatt 3600
 gcaagggacc actagaacce tcttctcaa ttccccctct ctgagaaggg aggctacagc 3660
 ttgcctctct aaccactaaa aggcagacc ctccctcaaag ttaatagccg gattccctga 3720
 tagatatttt cactaaatga attctcataa aactctcact aagatttaga gaaggcttcc 3780
 agggttgaat tcctgaacat taagaacagc atgtttttta aaagttaac ttggtgattg 3840
 gaccaggact tcatctaggc tatgaatgct cagaatggta ggtcctttac caaacagctt 3900
 gagtttgtgt ataaagtgat ctcatcctct taagagtcag agaaacagaa ccaagcgact 3960
 tcactataat ttgatctgag gaagtttctt actcacaata ggtaaatgaa ggcacatact 4020
 aaccagcaat ataaacaaca atatcaagtg tcattcacac atgcaaaaaa cagacaaaat 4080

ES 2 670 894 T3

cccaaactct gtgttctaac aaatcgcaaa aacctcacta acaataaatt gaaatgacca 4140
 aatgtttgga ctgaaaagca atgccttggg agcctagcca tgcctaactc aaataacaga 4200
 accatctcga tgttaaaatc ctcacagatc aagctgtgta tgtctcgggt caagacttcg 4260
 ccaaaaagca gtgagcacac acttaagagg gaaaaaatct acctcagcct cctaaatgca 4320
 atcatctcta cacgagttgc aggccccaag cttcaacgtg ttctgctgga caacgcagta 4380
 gaaagctgac aagcaggtgg ccttcccaca ctgactgaac cacctccatg cccatgtcca 4440
 ttcattttct tgccaccccc atgtgctata acagacctcc tggctcaggg cactctttcc 4500
 ttctgactg ccttcaacta atgactttgt acttttaggt gcaaaaatta tctgcagaaa 4560
 tccacactga aaaccaagct tgagaaaggc agcaataacc aacattttta caagaagaac 4620
 aaggtcaata tcaagcccat cagattcaaa tagcaagcat ggatgaaaat gaaagattga 4680
 aaggcttgag tgccttctta atgtattaaa tatccattta atttacaatt aagctcactg 4740
 tgctcactgg ccttttaatc agctttccag gtcctgctca gacttgccca ggacatggga 4800
 atgaaagaac ctatacattt atggaccaat ctaccttaac taacttgtca agtgttcctg 4860
 catcaagcag aagaaacatc agtgaaactg atacaggaat taacccttg ttaatccata 4920
 aaacttaaag gagcgggatc caatcttctg gcttccctgg gccacgctgg aagaagaatt 4980
 gtcttgcgcc acacataaaa tacacgaaca ctaataatag ctgctaagct ttaaaaaaat 5040
 tgcaaaaaag gaaaatctca taattttttg tttgttgtga ggtggagcct cactctgtca 5100
 cccaggccgg agtgacgtgg caccatcttg gctcactgca acctctgcct cctgggttca 5160
 agccattctc ctgcctcagc ctcccagagta gctgggatga taggcgtgtg ccaccatgcc 5220
 cagctaattt tcgtatTTTT agtagagacg gggtttcacc atggtggcca ggctggtctc 5280
 aaactcctga cctcaggtga tccaccacc tggcctccc aaagtgctgg gattacaggt 5340
 gtgagccacc gtgcccggcc aatgttttaa gaacgtttac gaatttgtat tgggccacat 5400
 tcaaagcctt cacaggctgc atgcagcctg caggccggg ttggacaagc ttggattaga 5460
 gaaatctaca gagacaaact agtgacttag tagccctctg atagctcatg atttgcaaga 5520
 aacttaggat gactatgtgt aaagaccaca aacatcaatt taactgaatg gttcccgcca 5580
 cactggaatg aggaagctga gcaaactcag aggactctaa gaaagggctg atgtcatctg 5640
 aactgttcgg aattataaac tcctcctaac atgtttcaaa gccagaactt gtaggagttg 5700
 ttctgatata cggattaaaa gagggatgac aaagtgtctg tccccacac tgggtcaaagg 5760
 gacaggtcat tgttatgctg gcaatgcagg ctgctgaaaa gaatgtatct gtcaaaagta 5820
 atcaaagtaa tgaccaccaga aggctccaga aacagactgg taaattcagg ttgctttcag 5880
 acttcacaaa tgctggcaca caaggggaaa gacaaaacta acatttacag agcattatat 5940

ES 2 670 894 T3

ttgatattac atttaatccc cattaaaaag atactatttc ccgtttcaact agtgaaaaag 6000
 ttgatctttc aaaggttaaa ttatttaaca ccaagggtcaa agggtaagtt ggagagacca 6060
 gattcaaacc cagtctgaca ttaaaacatg tgttttcccc ccacatcgtc tcctgctaata 6120
 aacctcaaat ctaaaaaactg acttgcctta caccttgagc cccatcctac aaactctccc 6180
 tgacggttatt aattcagctg tcaactgtgca cctacaacgt gccagacacc atactcctca 6240
 aactctgta ggcacagaag gaacagataa aaatccctac ctcatagat attattctag 6300
 gggtaacaca ggtaaataaa acattaaaat agttttcaca tagtagcaaa ttccatatag 6360
 caaaataaaa cagaagaagg aatagcaaat gagggagatg ccctcttaa catggtgctg 6420
 agggaaggcc tccttgagaa agatatcatt taccocaaaa ataaaaaagc aagtaataga 6480
 aaaaacaggt aaaaggtggt ctgacactt aaacctgcca cattgagaac tcagggttct 6540
 gatgcaaaac ctgctgcat agaatgcatt aacttatttt tatacattta aacaaacaaa 6600
 ctctacttaa gaactgtggt ctaaaggaag gagcatatta caggaaggca atttttggtc 6660
 agagtagaca cacttaaaaa ctaaacctat tgaaagacca agaacaactg aaagtctttg 6720
 ctttgtcaga tttttgacca aaaggaaaat taaagaaaca caccgtgcc atccaatgat 6780
 ttcaccaagg aattttaaga gagaaaatcc tacttcttcc tcaccagta gccagtgaaa 6840
 tgactgagca aattcacaag ttcactgggg ctgctttcat gtaacacagg gacaacacat 6900
 gacagacaca gtggaacct acaggttgcc tagtatttga aagactgtga agaggaggag 6960
 atgtcaaaat tcaaagtctt aatgatgta gttttaagta tgttcagcaa tttcaccact 7020
 cagtagtaaa gccagctaca gttgaaagca atcagaaatt tgaggggtgt gaaataagca 7080
 gaagcacaga agttaaggat ttgtattctt cccacatttt ccactttatt ttatactgct 7140
 gagaaaaaac aaatttaata gttttctgct gtataagaga gacacattca ctttatgtca 7200
 cagtaagagt cactcaattt taatacaact atctcaatgt ataaattaac attctcccc 7260
 ctgcccacac atagtaagtc tcttatgatg ttgctgatta gagaagcaaa agttgcccgt 7320
 acaattctct tcctgcattt taatataaac aatcatcagt cttttcttca tagagtgcag 7380
 tgtgggcaact atcatcagaa tgtaccagca ctgggtgtgc aaagtttaca aagattagca 7440
 agagcaaaag tgttgagatt tttgaaattc atgctgctgc aaagaagtat gtaaaaactc 7500
 actcaccata gaggaccaca cagaaactca ggcataagat tatatggctg tgtgagtggt 7560
 ttgggagaag gaacggaaag cacttcacc aacctatatg cctgagcaaa ttaatgcaaa 7620
 acctcagaag ctacaaaaaa gtttatctac ctaaattaa attggtgtcc acagcagtag 7680
 ccagcaaaat gcctgcgaag cgcaaagtgg taaatatttt agggctctgta ggtcatatgg 7740
 tctctgttaa acaatatgta aatgaatggg tgtggctgtg ttccaataaa acttcattta 7800
 taaaagagg cagcatggta catccagtca gcaagctata atgtaccaac ccccggtcta 7860

ES 2 670 894 T3

aactaacca aatacctctt aataagccaa agaaactgtg tcctcttagg ccggaagcgg 7920
 tggctcacac ctataatccc agcattttgg gaggccgagg cggggagatc acctgaggtc 7980
 aggagtttga gaccatcctg gccaacatgg tgaaacctta tttctactaa aaatacaaaa 8040
 attagccagg cgtgctggcg ggcgcctgta atgccaaacta ctggggaggc tgaagcacga 8100
 gaatcgcttg aaccagaggag gcagagggtg cagcgagcct agatcacgcc attgcactcc 8160
 agcctgggca acaagagaga aactccgtct caaaaaaaaa aaaggaaata aaagtataca 8220
 aagtgaaaac aaagaaatta aactgccctt atttgccagt gacattactg tctatgcaca 8280
 aaattccaaa aatctacaaa aaagcttcta gtactaaaaa tgagtttagc aaggttgtag 8340
 aatccaaggt cagcatataa cataaaatca ccttcctata tactagcaat caccaactgg 8400
 aaattgagaa gtatcattca caacagtacc acaaacatga aataaatgtg 8450

<210> 11
 <211> 8420
 <212> ADN
 <213> *Homo sapiens*

5

<400> 11
 tcttagtatg gtaaaccctt tgaagtagat tcaaatgaga atgggaagag agaaaagggga 60
 gagaagcaac ataagaaatc tcttttaagg aattttatat agagagaaac agaggaatca 120
 gttgatagtt ggaaattatt ttaaagaaaa tgggttattt taaagaaaa aggtattaca 180
 acatgtttgc actattgtgg gaataatcaa gttgagacag aaaattattt ttaaggaag 240
 agtctaattg ctgaagtgaa agagaatgaa tgagaccctg tgcataagtg tgatcagata 300
 ggagcatgta cagctcaagt aagaacagga agaaagagac aataaacatg tacagatagg 360
 atgggctggt cgatgtggtg gtgaaaagac atgcgagtta ttactgatta cttctatttc 420
 cccagtgaaa taggaagcca ggttcataaa ccaaaatgaa gaggagcgag gcagtattgg 480
 aagttcagga aaagtaatag gtgtaaaaat atgtaaagta gaattaccag ggagtatgaa 540
 gatacatttc caattaagga tgaagaattt aaagtgaggc cagccaatac ccctgctttg 600
 cttcagctac atcagctgca taggttcagg cacagaatac atggaacatt gtatttaaat 660
 agggcctgga ttttacaaaa gtaacacaat gaagaagaga gatgcaaggc tatttgaggg 720
 tgtttggtggg agagattgta aaatattagc taagtaagaa ggggactgca aatttttagtg 780
 gtataaagga atgaggaaaa gtgtaaatac agtggggtca aagaatgttt ggagccaagg 840
 cactagaggc aattagctga aaatgtaggt gattattggt gagtgacatg gtttaaatga 900
 aaagtataga aggttacaat tatccatcat gaaaagttct aggttacaac taagatctga 960
 gtagctgaag tagaatgaaa gtagaatgga cctttccata tccagccagg ttcagtgaca 1020
 gaaggttagg aaacaaatta taaaccactt gagagaacat atcccctaag ttgtttttgc 1080

ES 2 670 894 T3

tatTTTTctt tcagcatata tttgTTggaa tgccaactat gttcagttca attaatatgg 1140
gcttcttaaa taagggctcc agcactggat aatcctgccca tttatTTTga tacattccat 1200
cctgctgctc agatctattg gcatctacag gatgtctttt gagaagatgg gcattcacat 1260
ccctatgtcc tagcaaatTT ccaactcaga aaaccacatt aggcttctct atatatcttc 1320
caactatTTc aatggaaaat acaattctct gatttcttcc tatgatattt atcaaagaga 1380
atggtgcctg ccagttctag ggtgggggaa ctcaatacaa atcaccaacc tttagatgac 1440
accctgtctt caaagtgtt tcaaagtctg gcagaaaaaa agtaccaggt ggctataaga 1500
ccaccagga gttcagtcatt gcattctaag tagcagatca ctggaatgta attggctagt 1560
gagttcattt tactcttctc ttcttggTca catgttaccg cccttgtacc ctgcacgttc 1620
tctttcccag acttacaag catgttctct tgaattcgtt ctctttttaa attcacacag 1680
tcttaatgat tcttctttca caagagtctt tcaactcttac aattcagttc aagtcattcca 1740
catgcttatt atgagcaagg gtctgggact taggggaaaa ggggaataaaa agatgaatga 1800
aatgtgatcc ctgcagtcca agagcttgct gtgaaaaagg aagtttggct tacattgcct 1860
ccctaattccc ttggctaggc cagaacagaa tattgtctaa aacctcctca cgtcagcagt 1920
cctctggggT ggtgactgga agtagaattt aaacaaaaat ataattgaca cataataatt 1980
gtgcatactt atagggTaca atctgatgtt togatatgtg tttaaatggg Tgcattgtgt 2040
aatgatcaaa ttgaggtaat ttatccacca ccttgaagag agatttttca atattctcat 2100
tgcaagaag caggaatttt tagcagacaa ctgagatgct tcttgttcac actaagtcat 2160
tctgacgatg gatttacata acttgttgtt ttttttTgtgt gtgtgttttt gagacagagt 2220
cttactttgt cgactaggct gaagtgcagt ggcacaatct cggctcactg caacctccac 2280
ctcccgggtt caaacgattc tcttgcctca gcctcctgag tagctgggat tacaggtgca 2340
tgcaactagg cctggctaat ttttatattt ttaatacaga tgggatttca ccatgttggc 2400
cctgctggtg tcaaattcgt ggcctcaagt gatctaccag ctgcggcctc ccaaagtgca 2460
gggattacag gtgtgagaca ccaagcctgg tacatttaca tttcttatct ggatctttcc 2520
tttagtaagt gctaaggaat cctacttccc ccaatatttt ttctatttc aatgttttag 2580
catgatcat gtactactt Tgcagacatt tgattttccc ctttgtttac Tgtaaagtat 2640
atTTTTatag cctttgtaat agaagtattc taaaatctgc ctgcaaccta tctttctgac 2700
tctgcatttt agggaataat tctctgttgt ggaatgaaaa aaaaaacaga gcctgtggag 2760
tcagagatct catttcaaat tatagttatc cctaggaata aatctgagtg acaggtagta 2820
tagtataata ataagtataa agctatgggt aaggaaaact caacaacctt atctgtaaat 2880
tgggatgaca acagcctacg tcaaaaaaat gtgaaggtaa atgagataat gtaaggctga 2940

ES 2 670 894 T3

tacttagtaa gcaattttaa aacacccaaa aaactattgc catgattact ctacttactc 3000
tatttctcta tgctccaggc aatgaacta ctaatgaccc aggggtcctt cccattctc 3060
ttcttcacaa ggaaatattc tctctctgtg tgctgtttat taaaatctac tgcccctttt 3120
agaagccttt ccagatcatc ccatggccaa gaacgatcgc tgcttcctct tctttacata 3180
cagatgtttt tctcctgctt gacaattatt tttgtgcaat tattttcctt ttgattgtgt 3240
ttttaatgtc cccccaccc cacaattttc cagactgttt gctccacgag agaggagacc 3300
atcatctctg tgctcaccgt tgtatgacca gtatcctgag gagtggctgt tacataatta 3360
catcaggcac tcaataaaaa tttgatgaat aaacactgga ttttaaggca ggtatcatat 3420
cttacatagc atatcatatc ttacatttta tgtccctcac ataaatacca cagagtgaag 3480
tatatgacag ataaggtcat ttctcttgat aagtacatag tccagtctga aacagatag 3540
ccaaaaaaaa acaaaactgg agtaacaag atgaattggt ttaatagagg cactgtatta 3600
gtttcctagg actgccagaa caaatcacct caaacttagt ggctgaaaac acaaaaaatt 3660
tattgtctca cagttataga tgttagaagt ataaaattaa ggtgtcagtg ggattggttc 3720
cttctggggg ctgtggaaga gaatctgtcc caagccttca cactgtaaag tacagtactg 3780
gagggatagg acttcaactt gctctatctc agatagagag gagccatttg ttgtgaattg 3840
agaagagggg tatgttgaat ccataataag cacataaaaa ctgggctggt tcataggaga 3900
agtaacatgt ttccagctct agtaaaaaac aaattgaagt ggcctataaa aaggtacaga 3960
gtacgacaga atgaaaaata aatgaacaag aatacagaga ggatgtggta aattatcatg 4020
tttcccta atgttattgg aactaaatg gtattagaat tatttatcaa taataattct 4080
aaactgttgc aattgaaaga atatattaag tgggtgttata tgagaagtgc cagggcattc 4140
tcatttctgt ccaatgggag aacattttc gtttgagacc tccgtgaata atacagtctt 4200
ttagttagga gagctgcatt ttgagtgtg caggcagaat ggcgatctct caccacaca 4260
aacactaaga tagagagaga cagagacaga gacagagaca gcagagagag acagagaaag 4320
gaagtacagg tactcagata gagataagcc atttcttgac attaagaaat aaagtagaat 4380
ccattggagg gaaataaaac tgcctcagga acagagttaa ttcacataca catgcaggta 4440
aacacacact gcttgatact tactgtggac tttgaaaatt atgaatgtgt gtgtgtgtgt 4500
gtgtgtacat tcagccctcc atatccatgg attttgcatt cacagattca accaaccatg 4560
aattaaaaac atttggaaat acaaacatt aaaatataac aatacaaca taaaaataat 4620
acaataaaaa aatatagtgt acaactgtt tacatagcat gtatgttcta ttaagtagta 4680
taaatctaga gattacttaa tgtataccag aggatgcata ggctatatgc aaatactatg 4740
ccactttaa ctgataagaa cagatactaa acttcatctt agccaaaagt cagagaaaca 4800
atataactat gccattttac ataaggact tgagctgagc atcctcagat ttcagtatct 4860

ES 2 670 894 T3

ttggagttcc tggaaacaat tccttgtttt atatatatat atgtgtgtgt atatatatat 4920
 atatatatac acacatatat atatatatat atatatgata gctactgagt gacaggtgat 4980
 attataccat accacttgtc actcagtagc tgtatatgca tatgtatata tatacatata 5040
 catatatgtg tgtatgtgta tgtgtgtgtg tgtgtgtgtg tgtgtgtatg ctgtctttcc 5100
 tcggtatcac aggaattgg agatatatat attcttttca gtacaaaaaa aattgaacac 5160
 agatgggtat ggtaccagaa cagaaggtaa agacacatga aaaaaatttg caacaacatg 5220
 aatggaactg gagatcatta tttgaggaga aataatccag gcacagaaaa acaagcattt 5280
 tattatttta ggtgaaagac aaacatttta ttttaggtga aataatccag gcacagaaag 5340
 acaaacattg catgttctca tttatttggt ggatgtaaaa atcaaaaca tagaacgtat 5400
 ggaggtagac agcagaagga tagttaccaa aggctgcaaa gggtagtgta ggctttgagg 5460
 gtgaggtggg gatggttatt gggtaaaaa aatagttaga aagaataaat aatatctagt 5520
 atttaatagc acaacaggtt gactatagtc aaaataacat aattgtacaa tttaaatatg 5580
 aaattaaata tatatacaag actagaacac caagttgaat gactccagct tgcgaaaccc 5640
 acattgatca ccatgcttgc cccaagggaa gctgtacaat gtctggctcg tccagaaccc 5700
 catcatttat cactagcaat ctattgtcca taatcatggt taaattaata gcattttaaa 5760
 ggtacaaata ttttttaaaa aacaaataat tatttaattc gccttttaaa agctttttaa 5820
 aaacgttttt aaaaactttt ttaaagtctt gaggactatt ttcttttaag tgctcagtta 5880
 cagagctcca tatattgggc tatgatagcc ttacctgatt cttgccaaga atctagtgcc 5940
 cagaaaatgc aaatacaaaag taagcaactg aaaaataaac aaataagttg gaggtatgct 6000
 acctgttgaa atatgaccta gcgcaaacac ctatgccact tgcttatgaa atcatatagg 6060
 ttttcggtgt gcagttttga ctgaatgagg gagtttacgc tggaccacaa gggggcccct 6120
 ctgtcaataa cgtaactccat ttgtgtatta agtcaaaaat gaaatggaag agaaaagaaa 6180
 catcgatgac cccaagtctc ttttaattgaa tggaggtaaa agggaaacaa cgaatgagaa 6240
 aagtactctg cccttttaag aatcttgcct tcacattcct gatgaagtta ttttctctcc 6300
 tctcactgat tcccatttca ctctattaca tagcaccgtg ttccccagga gctcctgaat 6360
 gaaggacatc actcagctgt gttaagtatc tggacaata aatatactag tttcaatgtc 6420
 taggctatgg gtattccttt ttactgaagg tatgacatat agctgccag gcctgactaa 6480
 attaatagta ataataatta ataatggcaa atttttattc tattaagtta cttggcttga 6540
 cttgtagaaa tagcaacatt catctgaaat gccccctcct acacttatgt ctaaggacaa 6600
 atcccacata caccacagat aacttcattt tacatgtttt attctgttac caaactaaat 6660
 ttttatcata tagtctgttg ctcaactgaac tcttcagtaa ttctcaacat accatgtaaa 6720

ES 2 670 894 T3

gcattaagca cagttccaac acagagcaaa tgagcaataa ctgttagtta ttataacatt 6780
attatgtggtt ttcagtgcat taaacctactg gtctgatacc tagcccaaca ttctattaaa 6840
ccacataatc cagttgaata atatatgata atataataaa atggcgataa gtgctaaata 6900
tccagataga aacacagatg gaatcagaca gctttcccaa gaaatagaga aaatagtaga 6960
taggcgatct aggcctaagc actctaagca gaagctaagt tatcacagga tatcttgcca 7020
atctgtggca cgtgaaccct tttcttctgg agtctggaac tatggtgcaa ctctcacttt 7080
ctccctatct agagactcag tttgttccct tgtgattatc agcagttgag aaatccttag 7140
accttctgaa aggactactt tttaaattta tatatataat atttaaaata catatcttta 7200
tatataatat atatttaaat atataatatt taaattaata tatatttaa tatataatat 7260
ttaaattaat atatatataa ataaataaat ttatatataa atatataata attaaaatat 7320
atatttaagt aacagagagt aaaggattat tttgaagaga aactcctggt tcccacttaa 7380
aatcctttct tgtttccaag tttttcaaat ggagccctct ttaccagctt gccccctcag 7440
agataagctg ttccctact tattcagatc tgagatctga aaacattcct tttcctgtga 7500
gttcagctag gacaaagatg gagctttttg ataaaatttg gcaaacacat tttttaaaga 7560
tgaaaatttt taaaaattga aaaaaaaca tttatagaaa gagacttcta atccaaattt 7620
aacttctcaa actatgtttt gaccggctag cataatggtt cagtctttct ggagaatgcc 7680
ccttgaaact gttttcttct acacaacttc ctcccttctt ttgactttcc tgctctggaa 7740
gggaagaaca ggaagaggac agatcaaatt actcaagagg aaggacaaga aataaggaac 7800
caaattatca acaattggag aaagaaagct gatgtcagta tcatttcata tatgattatg 7860
tcagagtcag gtggataagc caatcctggt gaatagcata cttttcctgc tactcctgaa 7920
gggtaaagag gtctttctct taaaaagccg tcctagctag taatcttaca ggtgcaaaaa 7980
gcttgttttc atgttatttc ttagtaactc aaaatacctc taaagttata catattatga 8040
aagtactaca gtcacagtgc tgagaaaagg agtaaataag acaatgtata taaaaaact 8100
tggctcagcc cctggctctg tggttgataa atattaagtt agtattcatt attattataa 8160
tttccaaaga gtccattaaa agatatagaa gaagggaggc agcaataaca ctaagagaaa 8220
attccattat ctccaactat ttatcctcta gcccaaaata attgccatta gaaagagcaa 8280
ctttaacaaa aattttaagt tgcaatagat gttcaacttt aaatccatcc cagaaaaatt 8340
tctaacaaa ggagcataga agatttgatc ttattttcta agtagtatag acttaattgt 8400
gagaacaaaa taaaaacttg 8420

<210> 12
<211> 8475
<212> ADN
<213> *Homo sapiens*

ES 2 670 894 T3

<400> 12

gcataacttg taagaaatgg agtgaggtct cagttcaaac tggcttctgt atgacttcaa 60
agccaaagtc agcaacttag aaggcaaaaa ttataattta gttggcaaat acgagaaaag 120
gtcagaaaca catgaaatga agctcaatag gaacacttac agggtagcag ggtagtagcc 180
tagggaaaaa agtcagacac taaaattggt taaataggta agttcaaggg acaggtaaag 240
accttagtgg gtaagaagcc aatcagcaga cgaactgcaa gcaagcactg tctctctttc 300
ccttctgtct cctctttagt taactgacca caattaaggc tgcctagggg aataatgaag 360
taatcctcct attatcagca atggtctgat ccagtgccag gcaccacaga caacttgggtg 420
ttcagagaag atccttcaag atgaacaaag ggtcaaaata aaaaattcta gaagagagaa 480
gactgatcac aatttaatgt aaggcttggg aggaactgat ctctaccttc cttaacatct 540
caagaacttc ctcagattca ttggatgttg agtgtgtgtg agtctagtag aaaaatgaat 600
ttttgtttct taacttggat atgtgattag gatgttaata attaagtctg ggctaataatt 660
gaaggtatct tatgatgggc ttcttaaagc attgatcaca aagactgcat gttcataaac 720
tgagctgcac ttgttaggat tctagatggt tgaattttct tgtgttattt tggctctcaga 780
tttctagaca aattttctca aattcctatt tcactttttg acatatcatg agtgactcaa 840
atgtttgccc ttgagtcgga aaacaccag cattaggaat aggcacataa acataatact 900
tcaagcttca gatttaagct caattataaa gtgtttaaag gctgtgctga tagttcttct 960
gagtagaatt cctacaacta tgggtttgtc tataataaaa tgttcaactct atattgaacg 1020
ccttatttaa aactcgaaat gtgtaagtag taataaagaa aatatgtcct cctgtaacca 1080
aagctaggac cgattacatg ttcacttgac tgacagatac aatcacctat attaggagca 1140
atcagcactt ccttacaaac taacaacttg agatgtagtg ttcccattgg ctatgaagat 1200
tttctttatt tactcagaat agtctgtagg atctgccagc tgcccctgat tataccagct 1260
gcacccaatg atcacagtga acattatfff acattctaaa taactgggtgc aaggtgagcc 1320
atggttttct gagtttccta tcacctttgt gtttcaggtc ctcaaagtt aatttgtaaa 1380
gctgctgttt caggcaaac taacaaaatt agcatctaata caataacat actatgtcca 1440
cccatatcct ataacacaga agtaggggaa gagtgagaaa ggtggaagtg gagaaataga 1500
ggcccaaaaa gaaagtttta tcacaggaat atctagatgt cttctgggat tgtctgtaaa 1560
agagctgtga cactcatata aatgcagaat tactctcttt cttccttgggt ggttagaagg 1620
ccaaggtgac catggttaata ctaccaaacata tatatcaaag cttggcagga aaaatggtac 1680
cttcagaaat tttataatct gatatcaaat aggtcaagaa atataataaa actagtttct 1740
ttggtttcct tagaaacctg gaaaacttta aattagaaac ttagaaagct ttaaactcaga 1800
ctttgtagtt aaaaaaggaa attttagttc cttccagcat tagaattccg tgattctctg 1860

ES 2 670 894 T3

actctgagcc tggattaaat ctagcccagc tgagtggaaa cttaagtaac tagctggttg 1920
 ccttttagtga tcttccactt tatggctgct tccgctaag aagttcatca tCGTgactta 1980
 ctttctttgg ggcaaagtCG tgactaactt tctttggggc aaagttggaa agcagaggtc 2040
 aaagtcaatc agaaatggga caaactcact tcctactgcc tggTgaaggg gccattttca 2100
 gtagcccctt ttcaagatta gtttcattca agatttgata agctgTTTTg actttactat 2160
 agatcttatt atccatgtca gTtaagTtta tgcttccact aaatctatct gaattcaaaa 2220
 ggtaaaaagc taatgctcag tcttatcaga tttatcttat ttattaatag aatgtggatt 2280
 tttttaagca tataacaata atagtaatga taggaccata aatgtggatg gctctttaca 2340
 agtcactaac attacataaa ttctcaaca acacactctg aggcataac aaacttttag 2400
 aaataacaca attggctacg gaactccagc catctagctt catgggctcc cactttaatt 2460
 tcaaaaacac agaactgtgc acattcattt acatgattag ggcagagctt aactgtatct 2520
 catgtagcac ctacatcatt cttcagacaa acttattgcc ttttacagac aagaaaactg 2580
 gggctcaaaa aaggacttgc ttataactgg ctaataaaga ggaactctgg gttcaaagtg 2640
 agtccaattc tttcttccac ccacagcttc tgctaaagtc attacagaaa tgcatagagc 2700
 agttcttcca cgttattgct taggtttcta aagagcagtg acctaataca acatgctcta 2760
 taatttatta ctgatttaac tatttcaacta aggattcact tttactttt aacttgtaaa 2820
 tatgtctaataaacaccact gaaatagcaa cctctttctt catggccttg tggTtgtaaa 2880
 gcaagctagt aatatatgtc tgtggatttg tgctaataaa gttctataca cctcattaat 2940
 tccacaaatc ctactgggta tttcttatct gccagatcct acgctaggta ctggatacac 3000
 agtactgaac aaaatgggta caaatgagcc tcacagagct tgtttcattg aaaagcagag 3060
 agatacacac taatcaacaa attaatagta acacactacg atgtgTTTTg aaggaaaatt 3120
 agagcatcaa agagacggtg ttagcaggtg gaggggagct cttttagatg gagaatgaga 3180
 atgcctccct aaagacatgg gaataaattg agatcacaaa aaatgagaaa tagccagcct 3240
 tgagaagagc agaaggaaga acattcaaag gaaaagaaag tgcatactgg aaagcctgaa 3300
 cactagagtt tggTgtatgt aaggagctga gcaatggTca cttgtgtgat aagatgtgtg 3360
 gatgtggggT ggggggcagg ggtgagtccc acgcagctct taagtgtgtc ctCagactcc 3420
 tgtggTTTTcc atcagccaca acctgaataa ctgtgtggta atccaaaaat gattacagat 3480
 taaacatata aaaatatcat tacaccata gtacctaagc caaggacaca gtattctatc 3540
 tttcaatga agatctgcat gaagtaaaat tattatatat aattttaggt attgatatag 3600
 atacatcagt ggatagatat agatatgtgt ctctggTata gaaaaaagtt ttaaagggat 3660
 attaaaagtt cttatcttgc agggTtgaag attgtggcaa ctttcatttc tttttaattt 3720

ES 2 670 894 T3

taagaaaaaa gtggtattat gggggattag catgtttgtg ggtatatgta tatttttaat 3780
 taaaaaataa acaacaaaat gaaaacgttt ttcttctatg aaagcctaata aagaagaaat 3840
 ttcagctggt ttaacttagg gagctaaaaa catcaaatcc aagaatgttc tctggaactg 3900
 agctcaatac atttttattt gagtaagaat tggatacatt tccatcccct tggggctcca 3960
 gtctgtcaat attttacttt tcagcgataa aaagacacat gtagataatc acagtgcct 4020
 cagtaacttt ccttctctta ttttaagttta ttttatttct atcgtagttt tcctgttaa 4080
 agattttttc tttttgctta catatataat tttagagaat aacaatgcac acacaaaaaa 4140
 ttctctgtg tctgctagac ctggactttt tctctaataat atatctccat tttttgtctt 4200
 ttttcagacg tatttttgaa gcaaaggaga gaattgctat atagctgact tcctcttctc 4260
 atcaacagtg ttttaacagt ttttaagcaa aagtcagctt tgtttatcta agattttttt 4320
 tgctggcatt taacctacc ctgcctccc tttcccaagt ccacttcagc caacctctca 4380
 ttcgacaggt accaccctct aacataactg aaataatgtc taccattact ggatcttgct 4440
 agcaaagaat ctcaaatttt ccacttgggt tgtaaattat tttgtaatct ctagtgttta 4500
 aggtgcgctt gtccatctta atcccctccc tggcaggaca ccttacagaa cctaccctt 4560
 aactagtc ttaagcacca tcagggacgg atggctgtgt cactggtctg tttggtatte 4620
 cctactgatc ctaccatgtg gtgattatct atgacttccc taatccctgg ctgccttagc 4680
 tgggactggc tgacatgctt ctcaggttgc cgctggcttt acagtccttt actgccatg 4740
 ccactttgga gataggcagg gctagtactt ttctatataa gccccaaac ttgactttgt 4800
 gtttcacagt aggtgaaaaa gttgggtctc ttttctttta ctttctttc cacaagatga 4860
 taaagctagg ggaagcctgt ggacatggtt tatttctgca actgcaatga ttgattggtg 4920
 cttctgctg cttacttctt aaactttgtg ctcagtgctca gatccctagc agtttctatc 4980
 ccctgctctg ctaaaaaaga atggatgttg actctcaggc cctagtctt ttttaattaaa 5040
 ttgtattttt gttatcatta ttattattat tattttgaga tggggcttta ctctgtcgcc 5100
 caggctgaag tgcagtggtg caatcacagc tcaactgtttt agcctcctga gtagctggga 5160
 ctacaagcgt catgccacca tgcttctttt taatttttta aatggtttt ctgccttcaa 5220
 ttctaagcac ttctcaattg taaccaagag ataatacttt ttatgaattc ttaaagttat 5280
 caacagatac tcaaagtttt agcaaagtct aatgatatt aagcttgtcc ttattgcca 5340
 agtgacttca atgactattt gtttaattgca accaagggtc attttttaa tgaatatata 5400
 ttattattat atatataata ttaaggtcct caaatacctt aaagtttagc aaaatctaaa 5460
 taatattgtg catattcttt tattactgta ttagtccgtt ttcatgttgc tgataaagac 5520
 ataccaaga ctgggcaatt tacaaaagaa agaggttcac tggactcaca gttccacgtg 5580
 gctggggagg cctcacaatc acggcagctt acgggattgt tgagaaatga cacttctcaa 5640

ES 2 670 894 T3

gctggggcta aactatctct gtggtagttg ttctgattca agtattgaat tggttttttt 5700
tgtttttttt gagatggagt ttcgttcttg ttgccaggc tggagtgcaa tggcacgac 5760
tcagctcacc gcaacctctg cctccccggg tcaagtgatt ctctgcttc agcctcccaa 5820
gtagctggga ctacaggcat gagccaccac acccagctaa ttttgtattt ttagtagaga 5880
catggtttct ccatgttggg caggctgggc tcaaactccc aacctcaggt gatccacctg 5940
ccttggcctc cttaaagtgc gggattacag gcataagcca ccgtgcccg ctggagcatt 6000
ggtatataaa agctgcctag gtaactctaa ctttggccc catacatctg aaggatacct 6060
acaatgcacc tgaaaaatgc aactgaaaca gtagttccct gggaccacac actcagaaag 6120
ggggtgtatc aggagatcta gggaccagga ggggtggaaga cctaaggcag cactacagat 6180
gatggagaaa aaccactgg ggaggggcca tccaaactt gagaatcact gagatcatgc 6240
agaagtattt gatcctacag cattaatatt gtattgtatt gtattagtat atatatatag 6300
tgtatatata tagtattagt atatatattg tattgtatta gcatatata actaattgta 6360
ttgtattgta tttatatata tagtattgta ttagtatata tatacagtat atattgtatat 6420
ataactaac aatgtactaa tacaatacaa taccatata atatacacta acacaataca 6480
attagtatat atatatatat atataactaat acaatacaat actatatata tactaataca 6540
atatacat atatactcac caagacatat tagtggctg atgtctggct gccacactca 6600
tcttctacct tcagctctgc tctaccaaat atcatttgtt tctgggatct ttgcagtcca 6660
aggaacttca tccttgatat cccaccctt actaactttt ttttttttt ttttttttga 6720
gacggagtct cgctgtgtca cccaggctgg agtgcagtgg tgtgatctcg gctcactgca 6780
agctccacct cctgggatca caccattctc ctgcctcagc ctcccaagta gctgggacta 6840
caggtgcccg ccaccacacc aggctaagt tttaccgtgt tagcaaggat ggtctcgatc 6900
tcctgacctc atgatccatc cgccttggcc tcctaaagtg ctgggattac aggcataagc 6960
caccgcacc gccaccctt tactaatttt tagtaacgtc caaggattaa aggaaatttg 7020
ccttacctat ttaacaggaa tcaacagggt taatctcact ccctttctaa aaataattta 7080
taaacattgc agacaatctc atctatccct gtctaaactg tgtggaatta ctgccattta 7140
atgtaatcag tctactcatt tagtttgcct aaggaatttt tgaaaaaaca gttaatgaa 7200
tgacttaatg gaataaccag gaagtgaag tctccaatag taagaatgaa ctcttgctct 7260
ctggataatc aatgggtcc ttctccttc aggtagatca tgccatttcc tcacttacac 7320
tgaacaggta aacaacataa ttactgactt caacttctag ttaattcctt cttttatcac 7380
tgagtatcct ttggctggga gttttgttg ctatgctgcc attttttcta gttatcacag 7440
tcctataaca taccaatcct tcaatataac tcatctttaa attgtggttt taccttctca 7500

ES 2 670 894 T3

agaagttatt aattatgcca gtgctaaatc ttctaaaatg attgttgact tgttgattag 7560
 cccccatgca attcccctct cccgtccctc agcacgtaag gaatggccct ttgcttactt 7620
 ccacagatcc ttaaactctac cagttagaag ctaatagcct acctctctac caggaaggaa 7680
 ctgtgggctg gaacataata catgttgact tataatctct tagaaaattg tgtgagaaac 7740
 atcaaactcc tgattccagg atatgccaaa gacacatcat taaaaagcaa aacaaaacaa 7800
 aacaaacctc atttgacggt gctagtagtg gcatatctca tcaagatcag ctcaaataaa 7860
 tagaagtgag attttcacac aaattagact gtagtgcttt tttttttaac ttatctttac 7920
 catatgattt ttaacggtaa aaaaaatcgt ttgagatatt agatgtataa tatttatcat 7980
 ccaattactt cattagttca atcttttttc aatggcgctc ctgcatctga gaataaggctc 8040
 agaaaatttc atgttctgat ttcattgctga ttttcagaag aaaaatgtta gttttgtata 8100
 gaataacca tcctaagaaa tacatttctt attatatttc ttatcttata tttcttagga 8160
 caatgagcta ttcaaagggt gatgataacc agcaccatca gtcagcatta tctaagaata 8220
 agaatctgtg tttctacata cagacctcct aaaaaggaac ctacacttaa caggattccc 8280
 caggcaattt ggatgcacat taaagcttga gcaacactgc attagaaagt tagttttcca 8340
 tcacaaaaac agtaacaaaa ggaatataaa gtaagttact ttaataatat aagaagaggg 8400
 gcaggccggg cgcagtggct cagcctgta atcccagcac tttgggaggc tgaggccgggt 8460
 ggatcacctg aggtc 8475

<210> 13
 <211> 8401
 <212> ADN
 <213> *Homo sapiens*

5

<400> 13

tttcatctaa gactacattt ctattgtttt atataatcag cccccctaag atcaacatgt 60
 ccacatcttt tggcaaagac aaagcctact gatttcagga tcattatctt cctttttcaa 120
 aagcacaac ccaaactgag aaataaatca agagaaattc tccttttttc tatgctaatt 180
 tagaagtaga gtctttatct cttttcaaac ccaaagagaa tcagacatac aatatgaatt 240
 tatctacttt cgcttgctca gactgagagg aaagattaat attttcaggc tgttagtcaa 300
 aactgttcat tcaaatatta ttttaataaaa tccaagaacc agctaaaaag tcgcttaagc 360
 taagaaacct tcaccagcct catgggaaat tgtgtacagt tttccactag aatagcctat 420
 aaatgcttac tgaaaatgtc taagttcata tcttggtaac taacatctta attcaatctg 480
 cagaataata tatgcttctt tagtgctaag atatgaatat tagaggcatt ctttcttaaa 540
 atttctatct agttatactt tcacaaataa ctatataata ttaaaattct gcatgtggca 600
 taaaacatat tttaatggag aaggtaatgt gtagggagtt tatttctggt tgctattaga 660

ES 2 670 894 T3

acttgtgttt attccttggt aaaaaaactg cagattacaa catagaaaa aacaaaagta 720
 tgttgatat ctcttacagt agaagataaa gagtagttct aaatttagaa aggaaaaata 780
 aatatacaca gtgaaaatat gtgtcagtga gatgttaatc aaagatcaac tattgctgag 840
 accagcaata ttaaatccct gcacaattac tcatattata atgagaattt taaaaagaaa 900
 atatgaacac ataacataat gaaggcagaa gtcactctca tccttcatct ttgtattccc 960
 aattcaggaa gctggtatag tatcttcatt ataattacta ttcaacaaac atttgtaaaa 1020
 tgaatgaata aggaatgaat gatgagaaaa atgataaaca tctccctctg tctcctggga 1080
 gttaactgca ctactttctt ttaaatttaa ttaatcctca atgtccttgt aaaatagcca 1140
 aagggaaaat gtatttacct tactctaaat attgatgcaa tctacaaaa gtgttaaaca 1200
 acttcctcaa agtaaataaa atgttcacaa tccagctagg ataaaaggat ttaaatcatt 1260
 tcctaggtag agggctttca attagagccc ctgctgcatt aacctggga actcatctca 1320
 ctctctcat gatggagccc tgagtgttgc tgctaactctg tactctacca ttctaagtct 1380
 ttttaggttc cttttcagcc cttcctcctc gtaatccaca aatactgaga ccaaggcatt 1440
 ttttgggtca gtccataatt caagcattct atcctgcctt ccccaaatga actcacactt 1500
 attagaccat atgttctat attagttcag gaagggggaa aaaatgttaa tcacacttgt 1560
 atataagaga tcatagaaaa acagtttact aacctgtgaa aataccattc attctctggt 1620
 tactctggt ccacagctaa gcaatcagca ggatataaat gtaccctatg ttactattc 1680
 agtattcata agtatactac ttatgaattg gaaatctgac acaacattta catgacctaa 1740
 ttttgaat ttaaaatagt gtaaggcccc taggcttaat ttacagggg aaagattaaa 1800
 gggacacaag caaacatata ttctctctct gtgctgtggg aactggtaa ttttttgact 1860
 taaaatattt gatacttaaa atgccaaact tctacatttc tgcagtaaca aggcagttat 1920
 catattgaat accatttctt tctctccagt aagtagagtt aatattagca catgaactga 1980
 aatattaag tgattataaa aacgtccaaa taaattcatt aaaatttagc ttggcaaaaat 2040
 gttagttca tgttcttggg agaagtcctt ttatatttat attcaaatga aatgaacaat 2100
 ttacaagcaa aggaaatggc atcaaatatt tcacaccctg cctcccaagg tgtattgatt 2160
 catgcttttt gctcagatct aggtttctcc actcaggaaa agaggagaat gtaccatac 2220
 ttgggaaaac aagtttccga tggcacagct ttgatcaaac agcaaaattc tatccatcta 2280
 tgtattgcca tctgacagta tgacaaatgg tcccatgtgc gatattcaca ctgcattgca 2340
 gtcaaactg taagtcaaag gatatgaaat aatagtaact atacattaag cacagaagaa 2400
 aatgaacaa acaaaaaggt ttaaaccaa ccaaaaatat gtcttatttt ggatgttcta 2460
 tatgttctta cattctctca ggtcttttgt gtcattatga acacaattct aacaagcttg 2520
 attatttat ttccattcac atattacagg caacaagctg aaaaagtaga acgggggtgta 2580

ES 2 670 894 T3

gagagacagg acaaagtaca gattagggct tgaagtgcc ctgaccagtc gacagcaacc 2640
acatggaata atgactcatg tgcattaatg atcacactaa atgatatttg tttttttacc 2700
tagtccttca actgacagct taaagaactt caggttgttc tgattcttga gcctcctcta 2760
cagcttcaga gaggactttc attttatttt ggatcaaatg ctccacaact agttgaaact 2820
ggaattaaat tttatatgaa gttcctagat gatttaaagc tgtaagaaga agaataatga 2880
atcataagaa aacttgctgc tacagatatc aaaaaggaat gttacatcc ctcatgctaa 2940
tccttttcat tttaaataaa caggatctaa aaaaaataat gctgggaagt cctaaccaca 3000
tcaagaatgc ctcatgctgc tgaccaggg aaccttccag aatggatgaa atagacccaa 3060
agctgaattc acctaatttt agggccaaaa acccaaaaaa caaaacaaga ccaaaaaaat 3120
cttcagatac tgggagaaca aatctcaatt gctcaattgt atcttatgaa aacaattttt 3180
caaaataaaa caagagatat ttaagattca ttaagttctt gtcatttcaa attttaagaa 3240
aaatattttc taatggaatt acatatattt atatgattct tctagttata tccatggtaa 3300
taaataactct tttcagttgg aaataaaacc catttgtgct atattattag ggaaaatctc 3360
tacataaatt agtttttaat ttaactaaag tctatctttt gaattcataa gcataaaatt 3420
ttaaccactt gcaaaattta taacacactt aaggtagtca gatgccttgt caagtagttt 3480
aacaaaagtg attttcacct gtttgtttta ataacagtgc atcgatttta tgaaaatcag 3540
gcatgccctc gggctcctaac aaagtatacg aagctgaatg gatctatgcc aaatatgcca 3600
gattttactt tctgagtctg attttatact tctgtcctct ttcttaccac atggcttcca 3660
gtatcactta cagactaacc cttcaaaagg agaaggctaa gttactaaca tttggaaggc 3720
ttatgaaagt gaagcatagt tatgagccag caatgttttt atttagggaa tgtgtgcaaa 3780
ccatacactt aagcaagctc tggggaatga gagttggggg gaatcaactc ttttatttgc 3840
taattggtat ttcctttaa agatagagtt cttccagatt ttaactgtgt taatagttac 3900
tctagaaaaa ttggagattt gtgtgcatat attttatggt gtaaacagac acatacccag 3960
agacactgag agagacagac agacagtaaa cagaggagca ctaaccacaa acggtttaca 4020
aatgacctct gtgctcattc acctgtctgt tccccacct gccttttata gcaactatag 4080
caacagccat gagagtcatt gtggaaagaa ataaaataaa attaaaaaat cctggaagct 4140
tgtaaagaat gtgagcaaag gggaggaagt tgtgaaaaaa atgaataaag ggcaccgatc 4200
cagagtattg aagaaggcag agtggagagc ctagtaatga gtatctggta cccagtatc 4260
ctctcccaca gaatctgtac agctctccgt ttatgacagt ttaaacttaa tttaaattat 4320
caaacagaca ctttcctcaa acatataaat gatgaggcag ttcattcagg ctgtatgtat 4380
aaagttgttc cagccacctt tttctaattg cttctctata tcttttacct ggagacaatg 4440

ES 2 670 894 T3

agagatttgc ttaggacaat ttgactgtaa tttagaagta ggaaatggga agtattttgta 4500
 tcttctttgc ctaactcaca ttagttactc aagtaagcat ttcttccgtt attgcatttt 4560
 cctgattaca agttttatgt tttctctaaa acacatatca aaagaaatgt cctaagcact 4620
 atgcaggggg aagccatgac atttatccac cactgtcagc aaaaacatga acttagccct 4680
 caacagaata tttcacttca ttctagtgtc acctctgcgt cacctgcact ggagtcacca 4740
 cttgcctggt gggtaagacc aggatgcacc gctgaaataa aaaggggtca gacaatacaa 4800
 gaaaagccag tagaaattgc caaatgtatc agaatacaca caggctttct aaggatatgg 4860
 cccaagagga aggctctaga gccaccctg aacaggatt tttgacttca cagataaatt 4920
 atttaatttt caataacaca attcaattaa agaaagggaa atacaaggct aaacaaataa 4980
 gaaatgaaga caaaaaccca acctttcaa tctaaagaaa ataactgtt ttaaagacac 5040
 agatgaagat caggaaccca aacagaaga aaggaaaggc aattaacgct ggcatctgat 5100
 aacaacgaaa agtatggagt ctggagaatc gctagactct aaaaattata aaggtttaga 5160
 cttggacttt gtacactgaa gaaaagaaaa ctgcatgcat ttatactgac caatgtacac 5220
 tattgctgct ttttaacttt tgtgtatatg tagggtagat ttttttttaa gtgaaagcaa 5280
 gcttattaag aaagtaaaag aataaaaagg tggcttctcc ataggcagaa aactagcgta 5340
 gttttttat tagaaattgt tattcaataa tagtacatgt tacaataaa taccatttta 5400
 aactgaaaa attgtagact ttcaaatcag ttaggggtgt caccctaaaa aagggcattt 5460
 tttcccctta gtctccttgt tcatgttgct cacaacaaga aatgggctaa tgctatgaat 5520
 aataataaca aacactgcct tctgtcaggc cctgtgctga ataccgtctg catatgtata 5580
 ggaaaggggt aactcagcag gtcttgttg cccagactct gtacatttcc aagaaaggtc 5640
 tgcctttagc actggtcctt ggccagctcc tggagaatga gctctcagct tttagaaaat 5700
 tctatctgct aagaatagtt ttgcatgtct caggcttgg gccacaaaat atcagtttaa 5760
 tcagatggtt tatgttaaca agtatgatt atggcaaaca tagatctcta atctccattt 5820
 ctctctcata tatctatatt tatctatcca tatatatgta cctatatata tcaaatatga 5880
 agatatgttt atagcaattg catataaata gagagatagt atgtagtagg aagagagaca 5940
 tagatattat tcttcatttt agaatgttat cttgggtatgt ttaaaggaa aaacttaaga 6000
 tgtgttgcaa ttgcagtatg agtttcaggt atgtacatgt tatgtgtgtg tgtgagagac 6060
 acacacaaac acatttcaa catgttttat gtttaagctc aatattcaa cacagaaata 6120
 taacatctat tcttaatatg ttttatgtaa gtacagcagc agcattatta aatactgtat 6180
 ttctatggtg attgaaaatt agtaggcaga gaatttttgt aatggttctt aataattttt 6240
 gtaatagtaa atgattactt tttgtttagt atagttttat aatctataca tgaataaagt 6300
 ggatatttct attcatatag aatgtgatt tactctcatg tacttatcta catgctaaaa 6360

ES 2 670 894 T3

ccataagtta tcaatthtttag ttctgtgcca aggcactttt actgaataaa aataatcagc 6420
 taatthttata thttctctgat tcaatthttat atgcccctgt aatgthccgg ggtthttttt 6480
 thtaatthct gtaaatcaga atatthcagat gthgaaaaag thttthgcctt cagatthtaa 6540
 agatacctth gaaatgtagc atathcccaa atgcaacca gaggctggca atgtcaacat 6600
 thttctgtht taaaaaacct cttatgaaaa ctatthgcat actaaattht thactthgctg 6660
 atgactthaca gctggaaaag atthctgtaca tataagacat caaatatthg ggatactgga 6720
 actthttaaht taatggcaaa gaaagtcaac aaaggaagth catatgaaht caaactagta 6780
 atathgattac aaaaaaaaa gthttaaahh thttctthggcc ccagthctth cattthctgag 6840
 ccaaatacaa thctatcgaa atcacctgaa actgaahtca ccattctagg ctggtthttcc 6900
 cataaagath gactgctcca aaaagaggaa tcaagaaaga atthggctca cagthgaaht 6960
 thcactthgt cthagthtaag taaaaataaa atctgactgt thactacaga aatcattthca 7020
 aatthctgthg thgataataaa gthaatgacca cthttcagct ggagggacta actthcttht 7080
 thttthtgct gcatatathg ctgthggtaca thttaatgthg aatgathgac thcatcagct 7140
 thathccathg gagcagatth tagcattcag cthgggtctc ccagthcaata thctacgagthc 7200
 thctthtaag gagathgathg acacagathc atacagacta acaaatgthg thaccaataht 7260
 caagaatthca thcagthtaag atththgccc atgaththcca cacaagaaac thagaatthta 7320
 thgaththctt gthgcctgthg ggctccactc atthccctga atcacaaaag thacagagth 7380
 thtagathgaa aatathcctha thctthaacath gaaccattht aatathathgth atthctgthgth 7440
 ccacaggagth acactthtaaa gthgggactth cactctthcaa thctctccaath cacgthgthac 7500
 thaaagthggc atgthgthtcc thaaagctth ataaactgaca thgctthtaaa aaaggggtht 7560
 gthtcccgac thaatgthgaa aaagthctgaa aatgaththt aatctththca thaaatthct 7620
 caththgthca cthggaggaa aatgaththca ccaaatagath actctcathh atthththtaht 7680
 gthaatthath aaagaahtga aataththgaa thaahtccag atthccccca ccatgagctth 7740
 thccgaaagth atactccathc acagactgct cacthaagaa ctctactgca gthcaagthgaa 7800
 ccgaatthta ggggacathh thgactactthc thgtacacag aaacaththath catctctaac 7860
 actthccctath gagathggaag acgactthct aatcagththc cagagagggc thtgccaact 7920
 thcagggcttht gathgaahtag aatgththgag agcgtctathc ataahtgaaht thcagthataac 7980
 thgathgagaa agthgagagaa ccagagaaht aatctctcath gthagaaahh thaggggthathg 8040
 aatgthccaaa thgccagthta ccaagcttht cththgthcata aagcaactthc thataaaahh 8100
 gthgaaaata aatthctthcath ggctcaathgth gaathcagthaa ththccathctc thathactg 8160
 thgththacc aaaaactath ththaatgact aagactcagaa gththgcccagaa gthgthththca 8220

ES 2 670 894 T3

caaaacaact gttttgagat actccagatc tgtaatcaag taagtctgaa aaaccccaaa 8280
 tacctcactc acctcttggga tatgcataaa gcacactaat atataacggt ctaaaaagcc 8340
 aatcattaata accgttttat attgtttaag catttcctag acatatttgg ctacaaatct 8400
 a 8401

<210> 14
 <211> 8427
 <212> ADN
 <213> *Homo sapiens*

5

<400> 14
 gcacctgcca ccacgcccag ctaatthttct atthttcagta gagatgaggt tttgccatgt 60
 tggccaggct ggtctcgaac tcttgacctc aggtgatcca cccgcctcag cctaccaaaag 120
 agctgggatt acaggcgtga gccaccgcgc ctggccatat taacaaattt taaatcacia 180
 ctatgtgggg ggggaggcta gtattattac agcagattgg tttgctatat aaacaagtac 240
 tttaaaaaat atthttctggg ccaggcgtgg tggctcacgc ctgtaatccc agcactttgg 300
 gaggccgagg tgggcagatc acttgaggcc aagagttaag agaccagcct ggccaacatg 360
 gtgaaacccc atctctacta aaaatataac aattagccag gcatggagggt gcatgcctgt 420
 aattccagct gctcgagagg ctgaggcatg agaactgctg gatcctggga ggcagagggt 480
 gcagtgagct gatattgcgc cactgcactc catatccagc ctgggcaaca tggcaagact 540
 ccgtctcaaa taaataaata aataaataaa taaaactaaa ggcagagttt tcttaaataa 600
 acatggtagc cctcagcaac aatattgtaa gaactcctcg caagagaaaa agctggaata 660
 agatactggc taagcaagta agaaaggcac tggcctgctt ctgcatacat tcaaactaag 720
 acatatacat tgcagcttac acttacattt tccaatatcc ccaggcatcc ctttcccttc 780
 tcaaacagcc aaaaggaacc agccatgcaa ataaaaatac aagttcaaga gcctaaaaga 840
 agtcagtgtc ctaaaagaga aaattaatgt aaagaattaa gattttttga aactacactt 900
 tctttctggg gctgtttact ggccctccat acatcaatcc tgtaaacactg tgaactacag 960
 tgatagattg gtacatgctt ctaaacacia cagaatthtt ccaaggttac atacactgta 1020
 acaaaagggg cattttgcag catcttattt tccttaatca actagtttgg atattctaac 1080
 agtgcaaaca ttgtaaacia taaatthttca ttaccttttg aactthttctga agtcaacia 1140
 aggcttgtgg tatggatgca atgagtacta gacaggcaga gctgaatact agtcaaaata 1200
 ttcagttact ggtgtgatag tcctthttggg ggcatacatc acttagggag aaactgaggt 1260
 gcaaggacat tttacacaca gcaaaaacat tctcaggaat ttgtcacatc attaccataa 1320
 gccaaaaatc tcaaggtctt agaacagcct gagcttctga tcaaattata ttgtaaaaag 1380
 agaggaaaaa aatgtgaagc gtgctatthtt ttaaataaac agtaactact actactgctg 1440

ES 2 670 894 T3

ctgctgctaa ttctaaacgt ttactgagcc cttattatgt gccaaagcacc gtgctaggta 1500
 cggatcataga ttttaacaat taatccctgt aacaaccctc tgatattagt taataaaatt 1560
 aaagtagaat cctcaccaaa aaaatttaaa ctttccaaat aaaaatataa ataaattatt 1620
 aaagacattt cacctctttc tctgcctcag actacatttt caagtattaa atttactacta 1680
 aaaccacatt tattttcagg aattccagtt aaagcgtaca gatattcaag atgttgacaa 1740
 ttattacaga agaatcacag aactctgaaa ttaaatactg gcacagaaaa ctttccatcc 1800
 aaccttacgg aacaactatc cccattttta aaaaaaagga acagcatata tatcaggctt 1860
 gataataaga ggcttctcat gcccacacta gcaatgaatg atgccataat tataaagaga 1920
 cctgtatcgc cacatgcata aaaataattt acatctgcta agtcaagttt tcaatatatt 1980
 atttgtgtg taaaccttat agtagctgat aaaaaataca ataaactaat ctaaggtaaa 2040
 ctaaaacact aggttgtttc tgaagactca ctttagaatt tgagcagcat aataatcata 2100
 atattagtaa tcaaactact tagcagaaaag ttcttagagg gctgggaagc tgtgtataat 2160
 aaaatggagc agacaagaag gaagggtttt ccgtagctgt taaatcaact acagggtcca 2220
 gcatgcagtg ctctaactctg aagttaagca aaaactgcaa tgcatactgg gacttgtagt 2280
 aagtaaacca cgttatcaca gcaagtttca agaaagtctg aactatctag cacaatttga 2340
 ctatatctta ttatcagagt ctaatcaaat ttaaatacaa tttgtatgtt ctctgatgtg 2400
 gcacacagtt tctctagcac ataccgaaa aagtatcaat atttagacca acattttcac 2460
 attagaaaaa tcttacgtag gagaagcaca gaaaaaatg ctgaaaaagc aaaaaaactt 2520
 gatgaataaa aaatataatt tttgaaatag ttttttaaag tttgaatgga tccatttcaa 2580
 cattctctaa tcctcccca caaaaagttt aattgttttg gccgggcgcg gtggctcacg 2640
 cctgtaatcc caaaccttta ggaggctgag gcggtggaat tacgagatca agagatcgag 2700
 accatcctgg ccaacatggt gaaacctct ctactaaaaa taaaaaatt agttgggcgt 2760
 ggtggcgcac gcctgtagtc ccagctactc aggaggctga gacaggagaa ttgcttgaac 2820
 ctgggaggtg gaggctgcag tgagctaata tgcaccact gcactccagc ctggtgacag 2880
 tgtgagattc attctcaaaa aaaaaaaaaa aaaaagttta attgtttta caggttgctt 2940
 ttaacaatt attcaagatg tattttataa ataatttttc ttgaagaaa ttctcagaag 3000
 caaacattcc ccatattcta atattgcca ccaggaaata attttttag taatacgcac 3060
 acacccatc acaaaaacaa acaaaaaaca ctgaagttct gcttttgcac agtccttact 3120
 caatatttat gccctcatt cctcacctct aattccctac acacacacac acacacgcac 3180
 acatccccac acacacacgc ttctacaaag aacacttaga aaacagtat tccaactaca 3240
 agcccacttc tctcatccac tgacctctc tgaaaacaca aaagatttt taagctatca 3300
 gtaacacgtc caaacacaag ctgataagtt tgagctagaa tttacatata tacagttgct 3360

ES 2 670 894 T3

acacaccctc ctatctctg caagtctgtg gaaggaggct gggaaagaac taagtgcaat 3420
 ctgcatcagg aggcctaaca caggtggtgg gttatcttca ggcaacagca ccttcacaaa 3480
 catgctttgg aatatagtcc aagaaattcc taacaaggaa agataagctg gcacacaaat 3540
 ttaacgcaat ccagctaaaa atcatctgca acacatgcta ctacatttca ccataaaagt 3600
 gacgggctac tataaaggat ttgaagcttc gtcaatacaa catactgtcc ataaggccag 3660
 agatagcagt tgccatgggt actataccca cttttatcag gaaattactg tcattacccc 3720
 aaagctttgg gtacttattt aaaatttaa aaaaacacac acaatttagg gttctgactg 3780
 ttaattgagt gaaataatca actactgttt gatttgtaag tatgtcgctt tggagatgca 3840
 catggttaac aatacttggg tctgcagcag aaaaaaatc aattcctttc tgctgctcct 3900
 tctcctcaag tactgacagt ttgtattctc aatgcagcca aaacaataaa acaaaacca 3960
 tctttttggc ttctgtgttt aagttatttt tcccctaggg ccacaaacag agtcaaaata 4020
 aagcctagat catcaacctg ttaggcctca tccccttct atcccctcca tactgggttca 4080
 ctttctgac tacttagaaa aggcagaaaa ctttctgta actgattcca aagtatagaa 4140
 aagaatagtt gccttcaact gagatatttt caccaaagtc ttttttattt actttttttt 4200
 taaggcaggg agaggggaga gacttgcagg gtactgaaag ggagaagtgg aggagtattc 4260
 aaattgccac acaagtctag tgtaagaaag ttgctttaga agagtccaaa ggatggctga 4320
 acctcacata taatttctaa aagctttgga agagttcacc ataattttta gactgaattg 4380
 aggacaagt aatagaaaag ttattcataa agtctacttc aacattttta caaaagataa 4440
 ctattcaaaa atttaacaca catataagaa ttatacgaaa gcctacaaaa tagtatggcc 4500
 acatatacac acaaacatac aaagtagaaa acataagcta ttttaagaaat aattatctac 4560
 aataaattca atgcaatggt aacatattat ctctttttta aaaaatcgca aagcagcaaa 4620
 aacatacacc tgagaaaatt aatgtgatca aaacgttaaa gaattcttag gcctataaaa 4680
 aaagcccatg tacaaaagct cctgagaagt caacataaat cattaatatt tcccagcaca 4740
 aaataatatg aaaattcaaa catgcttcaa gaaatcagtt ctagatatag atataaaaga 4800
 attccattaa aggtcagaga cctaaaactt taattccttc ctttctctgt ttgaatagta 4860
 attaaataca aaagccttca gcaataaaat actaaggata caaaatttaa aagcacatta 4920
 atataagctt aacttcagta tgtcttcaca gaaagcttta ctattcactg tctgtaggat 4980
 gaaaaagtta ataaccctt gagaggtttc atttttatct aaacagttaa gtgtttttct 5040
 caccgttcac agaagcaagt ttctatattt actttctaaa gggggcaatt tcaaaagaat 5100
 agtcacttct aaaatttaag atactatacc ttttgatagg ctcataaaca cagggttctt 5160
 aattatctat attttacttt aaaatgtttc tattccaaat ttgtgagcag agtttataag 5220

ES 2 670 894 T3

aaagctgaaa ctcaaggctt taaacttttg ggttatTTTT acacaaaaat atttcagtgc 5280
actcctctag atttgagtag tcatttcctt gtgcacCctt ctaaaataga aaaacaaaaa 5340
tgatatatcc atatatacct aatactaaca catacagata tacatctttt tcaactgtgaa 5400
acaagcttga aagcttttagg cagtaagaat ttttcagaaa gttagcagag tcagtcaaaa 5460
cattcaaaac ttgaaccatg acatctgtta ctctgtcaat aagagtctat agaagaatca 5520
gggaacttac atactcacta aatcaacta ctatcacatc acatcaatgg agaaatgaag 5580
aaaaactgta ataggggaca tacaattcac aggatcttca aaagggaaaa tgatcttttt 5640
ttttttttta aattatgaga aactgactag gcagcatttt ttcaaaagca gcttcaaaac 5700
tataacaaag acatTTTTtg taaccacagc agtatTTaaa aaacaaaaat ttaggccggg 5760
cgtggtggct cacgcctata atcccagcac tttgggaggc caaggcaggt ggatcacctg 5820
agtcaggagt tcaagaccag cctgaccaac atggtgatac cccgtctcta ctcaaaatac 5880
aaaacttagc cgggcgtagt ggcgacacc tctataatca cagctactca ggaggctgag 5940
aggcaggaga atcgcttgaa cctgggaggc agaggttgca gtgagccgag atcacgccgt 6000
tgcactccag cctgggaaac agagcgagac tccgtctcaa aaaataaaaa aataaaaaaa 6060
ctatagtgtc cagggtgcac tttaaatgta ttactttctc aactgatatg gaaaaagtta 6120
gcatttaaag acagaagctt ctgtccatgt attaattagt tacctatctc aacaacttaa 6180
tatctgcatg ctttcttacc atttatgaag aacttttata tgtattatct catttggctc 6240
tactgagaaa acagtatTTt gcctacaaaa tagacaaaat tcaaagcaga tttatcaaac 6300
tttctagcat ccccaaattt ttaaaacttc gacacaaaac tttacaagca accacagtgg 6360
catgatattt tcagtgataa tcaattcacc taacactaac agagtTtcaa aggaccatgt 6420
gctataaatg ctatgaaact gttaaagtag ctatattcat ctttatgcag ttaactgttac 6480
atcaacaatg acctaccact gatacaactt gacttacagt tcaagaatct cagtctttgc 6540
aggctaactt aagtacatca accatatgta tttataaagc cgagtgccta aaaattgatc 6600
tatattagaa tcatagtctg taaatccgag gggaaaaaac tacaagaagt ctaaaatttt 6660
ttcaacacac tatacccctt tccaaaatct caactactct atatcctatt tgtattaata 6720
ttatagggat gataacaagg cttaaagccc taaatcatac caactacttt tgtttataac 6780
aattacaaat aatTTTTtaa aatacatgct caacatccca ctcatcaaca caagactaat 6840
tccccttcca aataaaataa ttctaacag tgctctgtac caagggccag aatccttata 6900
ctatccgcaa tcgcacatct actttgtaca gtcaaagact tcactttcaa gtagcaaaca 6960
ttatttatga atggaatttt taaatggact tactcaaaat ctttctggaa ctttaaggtg 7020
ttaatcctgt tgcttagctg aagctaagca gagctgtaat aagtagcaag accctcaaaa 7080
ttcaaaaatt tcctttatct tgctgtagca cctcctgctg gatagcattt agagatcttc 7140

ES 2 670 894 T3

atgtaagcag aagaagagta tttcagagggc agctccttcc agaagactga ataggaaaaa 7200
 ggatggaccc ttcaaagcta aaagaaatag gccccatcca tcacttatac cttctaaaaa 7260
 tacaatttag cccaggtagg tgtctttttc atctattact actccagttc cacaaagact 7320
 tgcctcagtc caaaatacaa catgcttaaa taaagcctgc aaaattgtct aaaaactaag 7380
 ttaaaaagca ttcaatagca cccaagcaaa acactttatt atgggcagcc aagcaatgtc 7440
 agtcaaactg taaatactat tatgttacca aaagcaaaag tctgatgtta aaaaaaaaaa 7500
 aaaaaaagcc cctggaatat tcgtaacatg ttagccagat gtttgtgttt tgagaacttt 7560
 gtgcactatt actatgctct tcacttaagg atagttgtac atctacaaac gttttaagta 7620
 cagaaatfff ttataaaca ttagcataac tgtacacaaa atttcctctt tgccatgaaa 7680
 agataggtcc tgggatttga aaatgtatff ttcagacatt tttaatgacc ccctaaaata 7740
 aactagttff aagcccacaa caccgattcc ataaacaagt aaagacagaa gaagagaata 7800
 agaaggaact taccaaaatt aaaatgaata atagtatffc cagtaaaaat gtagtaacag 7860
 tttccaacaa tgctgtaaac caaataaatt gtgaaactta aaaaaggaag gagggggcca 7920
 gtcttcaag accaaaagca aagctgacct atttatffct attgcttaga gtgaacacca 7980
 gatgtaaaca aatatcataa acactgaaaa gtacgcttac atggtttagc ctcaatffca 8040
 gtacccttac caggccctca ataaagctac agatgttggg gagaactcgc tcaaaaagga 8100
 gataattcca gcccctcgcc ttaaagaatc cctatcaagt gaacctgtga aaagactffc 8160
 ttcccagagt gcacaactgc tttaaaaaaa aaaaactffc atcagcccaa attaactctga 8220
 ttctaattatt caactatcca ttatfftatat ataaatgttc ttccctctct aactffccca 8280
 gctcgagcat ctacattcct gacaccgact attagcaaaa atgcacaact cttcccccag 8340
 ctatggggca aatctffgaa atctgaaaca cagccacaaa gttcactgtc aaggccaggt 8400
 gatgaggccc acacatgccc ggacctt 8427

<210> 15
 <211> 1704
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Fragmento A

<400> 15

ES 2 670 894 T3

ggtaccaccc aagctggcta ggtaagcttg ctagcgccac catggtgctg cagaccagg 60
 tgttcatctc cctgctgctg tggatctccg gcgcatatgg cgatatcgtg atgattaaac 120
 gtacgggtggc cgccccctcc gtgttcatct tccccctc cgacgagcag ctgaagtccg 180
 gcaccgcctc cgtggtgtgc ctgctgaata acttctaccc cagagaggcc aaggtgcagt 240
 ggaaggtgga caacgcctg cagtccggga actcccagga gagcgtgacc gagcaggaca 300
 gcaaggacag cacctacagc ctgagcagca ccctgaccct gagcaaagcc gactacgaga 360
 agcacaaggt gtacgcctgc gaggtgacc accagggcct gagctcccc gtcaccaaga 420
 gcttcaacag gggggagtgt taggggccg tttaaacggg tggcatccct gtgaccctc 480
 cccagtgcct ctctggccc tggaaattgc cactccagtg cccaccagcc ttgtcctaata 540
 aaaattaagt tgcatactt tgtctgacta ggtgtccttc tataatatta tggggtggag 600
 gggggtggta tggagcaagg ggcaagttgg gaagacaacc tgtagggcct gcggggtcta 660
 ttgggaacca agctggagtg cagtggcaca atcttggctc actgcaatct ccgcctcctg 720
 ggttcaagcg attctcctgc ctacgcctcc cgagttgttg ggattccagg catgcatgac 780
 caggctcacc taatTTTTgt tTTTTggta gagacggggt ttcaccatat tggccaggct 840
 ggtctccaac tcctaactc aggtgatcta cccaccttgg cctcccaaat tgctgggatt 900
 acaggcgtga accactgctc cacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg 960
 tggttacgcg cagcgtgacc gctacacttg ccagcgcctc agcgcctcct cctttcgctt 1020
 tcttcccttc ctttctcgcc acgttcgccc gctttccccg tcaagctcta aatcgggggc 1080
 tccctttagg gttcogattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg 1140
 gtgatggttc acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg 1200
 agtccacggt ctttaatatg ggactcttgt tccaaacttg aacaacactc aaccctatct 1260
 cggctctattc ttttgattta taagggattt tgccgatttc ggcctatttg ttaaaaaatg 1320
 agctgattta acaaaaattt aacgcgaatt aattctgtgg aatgtgtgtc agttaggggtg 1380
 tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 1440
 agcaaccagg tgtggaaagt ccccaggctc cccagcaggc agaagtatgc aaagcatgca 1500
 tctcaattag tcagcaacca tagtcccgcc cctaactccg cccatcccgc ccctaactcc 1560
 gccagttcc gccattctc cgccccatgg ctgactaatt tttttattt atgcagagggc 1620
 cgaggccgcc tctgcctctg agctattcca gaagtagtga ggaggctttt ttggaggcct 1680
 aggcttttgc aaaaagctcc cggg 1704

<210> 16
 <211> 1120
 <212> ADN
 <213> Secuencia artificial

ES 2 670 894 T3

<220>

<223> señal IgG1 humana + constante IgG1 humana

<400> 16

tgctagcgcc	accatgaaac	acctgtggtt	cttcctcctg	ctggtggcag	ctcccagatg	60
ggtgctgagc	caggtgcaat	tgtgcaggcg	gttagctcag	cctccaccaa	gggcccaagc	120
gtcttcccc	tggcaccctc	ctccaagagc	acctctggcg	gcacagccgc	cctgggctgc	180
ctggtcaagg	actacttccc	cgaaccctg	accgtgagct	ggaactcagg	cgccctgacc	240
agcggcgtgc	acaccttccc	cgctgtcctg	cagtctcag	gactctactc	cctcagcagc	300
gtggtgaccg	tgccctccag	cagcttgggc	accagacct	acatctgcaa	cgtgaatcac	360
aagcccagca	acaccaaggt	ggacaagaga	gttgagccca	aatcttgtga	caaaactcac	420
acatgcccac	cctgcccagc	acctgaactc	ctggggggac	cctcagtctt	cctcttcccc	480
ccaaaacca	aggacaccct	catgatctcc	cggaccctg	aggtcacatg	cgtggtggtg	540
gacgtgagcc	acgaagacc	tgaggtcaag	ttcaactgg	acgtggacgg	cgtggaggtg	600
cataatgcca	agacaaagcc	ccgggaggag	cagtacaaca	gcacgtaccg	ggtggtcagc	660
gtcctcaccg	tcctgcacca	ggactggctg	aatggcaagg	agtacaagtg	caaggtctcc	720
aacaaagccc	tcccagcccc	catcgagaaa	accatctcca	aagccaaagg	ccagccccgg	780
gaaccacagg	tgtacaccct	gccccatcc	cgggaggaga	tgaccaagaa	ccaggtcagc	840
ctgacctgcc	tgggtcaaagg	cttctatccc	agcgacatcg	ccgtggagtg	ggagagcaat	900
ggccagcccc	agaacaacta	caagaccacc	cctcccgtgc	tggactccga	cggtccttc	960
ttcctctaca	gcaagctcac	cgtggacaag	agcaggtggc	agcagggcaa	cgtcttctca	1020
tgctccgtga	tgcatgaggc	tctgcacaac	cactacaccc	agaagagcct	ctccctgtct	1080
cccggcaaat	gagatatcgg	gcccgtttaa	acgggtggca			1120

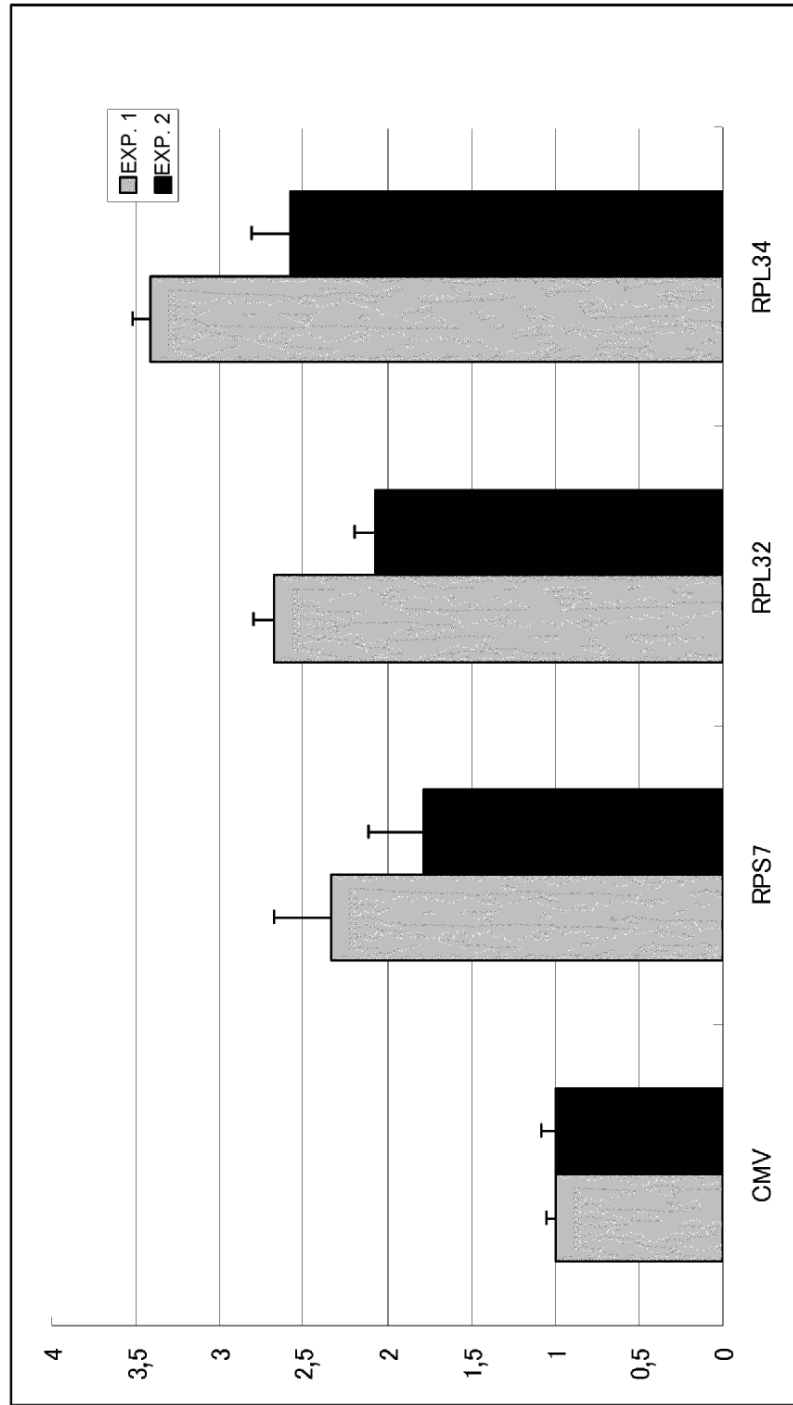
REIVINDICACIONES

1. Un polinucleótido promotor que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 1 en el listado de secuencias.
- 5 2. Una unidad de expresión de genes extraños que comprende el polinucleótido promotor de acuerdo con la reivindicación 1.
3. La unidad de expresión de genes extraños de acuerdo con la reivindicación 2, en la que el gen extraño es un gen que codifica una proteína multimérica o una proteína heteromultimérica.
4. La unidad de expresión de genes extraños de acuerdo con la reivindicación 2, en la que el gen extraño es un gen que codifica un anticuerpo o un fragmento funcional del mismo.
- 10 5. Un vector de expresión de genes extraños que comprende la unidad de expresión de genes extraños de acuerdo con una cualquiera de las reivindicaciones 2 a 4.
6. Un vector de expresión de genes extraños que comprende la unidad de expresión de genes extraños de acuerdo con una cualquiera de las reivindicaciones 2 a 4, y uno o más polinucleótidos seleccionados a partir de los polinucleótidos descritos en (1) a (9) en el siguiente grupo A:
- 15 Grupo A
 - (1) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 10 en el Listado de Secuencias;
 - (2) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 11 en el
 - 20 (3) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 12 en el Listado de Secuencias;
 - (4) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 13 en el Listado de Secuencias;
 - 25 (5) un polinucleótido que comprende una secuencia de nucleótidos representada por la SEQ ID NO: 14 en el Listado de Secuencias;
 - (6) un polinucleótido que comprende al menos 3000 nucleótidos consecutivos de una secuencia de nucleótidos representada por una cualquiera de las SEQ ID NOS: 10 a 14 en el Listado de secuencias;
 - (7) un polinucleótido que comprende al menos 2000 nucleótidos consecutivos de una secuencia de nucleótidos representada por una cualquiera de las SEQ ID NOS: 10 a 14 en el Listado de secuencias;
 - 30 (8) un polinucleótido que comprende una secuencia polinucleotídica que tiene una identidad del 95 % o más con la secuencia de nucleótidos del polinucleótido de acuerdo con uno cualquiera de los puntos anteriores (1) a (7) y que tiene la actividad de mejorar la expresión de genes extraños; y
 - (9) un polinucleótido que comprende una secuencia de nucleótidos que tiene una identidad del 99 % o más con la secuencia de nucleótidos del polinucleótido de acuerdo con uno cualquiera de los puntos anteriores (1) a (7) y que tiene la actividad de mejorar la expresión de genes extraños.
 - 35
7. Una célula transformada en la que se ha introducido el vector de expresión de genes extraños de acuerdo con la reivindicación 5 o 6.
8. Una célula transformada en la que se han introducido el vector de expresión de genes extraños de acuerdo con la reivindicación 5 o 6 y un vector de elemento.
- 40 9. La célula transformada de acuerdo con la reivindicación 7 u 8, en la que la célula es una célula cultivada que procede de un mamífero.
10. La célula transformada de acuerdo con la reivindicación 9, en la que la célula es una célula COS-1, una célula 293 o una célula CHO.
- 45 11. Un procedimiento para producir una proteína **caracterizado porque** comprende cultivar la célula transformada de acuerdo con una cualquiera de las reivindicaciones 7 a 10 y obtener una proteína que procede de un gen extraño a partir del producto de cultivo resultante.
12. Uso de la secuencia polinucleotídica promotora de acuerdo con cualquier reivindicación 1 para expresar un gen extraño en una célula transformada.
- 50 13. Uso de una secuencia polinucleotídica promotora que tiene una identidad del 95 % o más con la secuencia de nucleótidos del polinucleótido de acuerdo con la reivindicación 1 para expresar un gen extraño en una célula transformada.
14. Uso de una secuencia polinucleotídica promotora que tiene una identidad del 99% o más con la secuencia de nucleótidos del polinucleótido de acuerdo con la reivindicación 1 para expresar un gen extraño en una célula

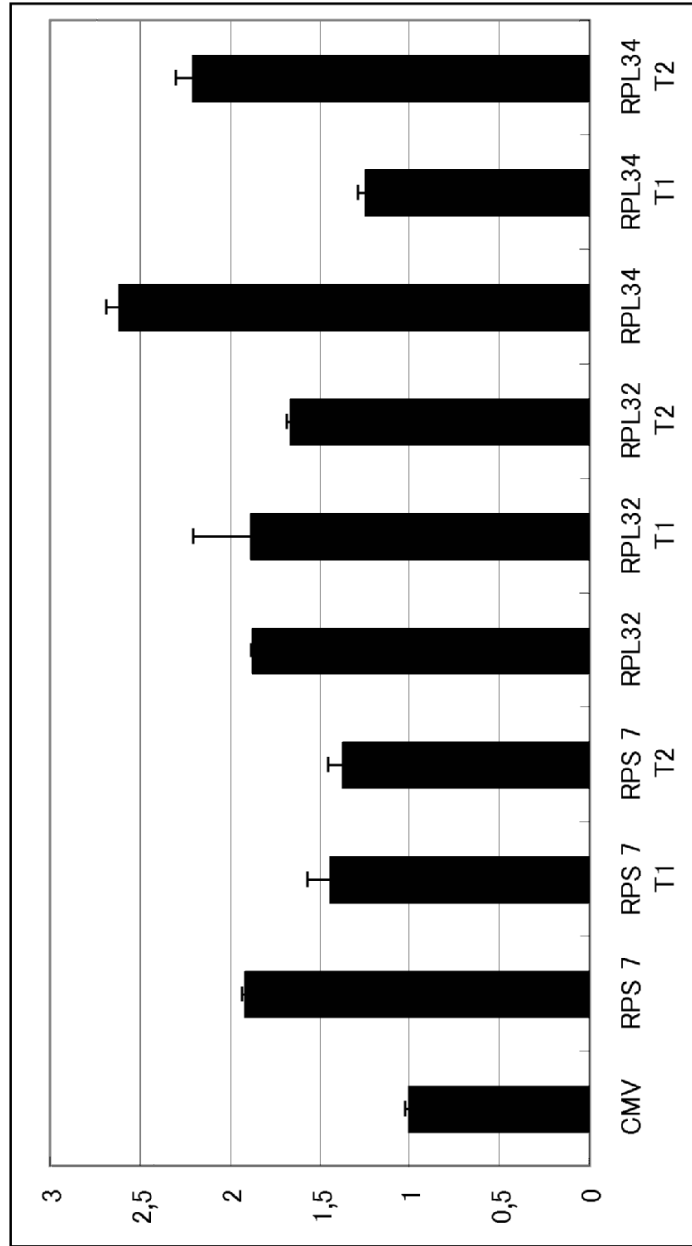
transformada.

15. Uso del vector de expresión de genes extraños de acuerdo con la reivindicación 5 o 6 para expresar un gen extraño en una célula transformada.

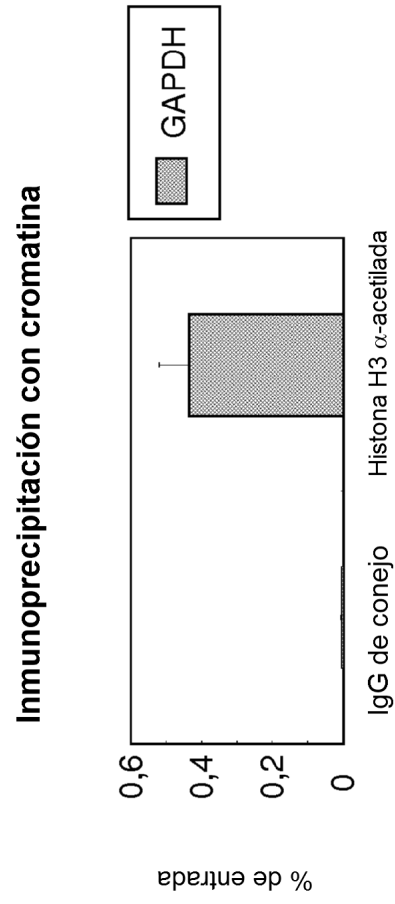
[FIG. 1]



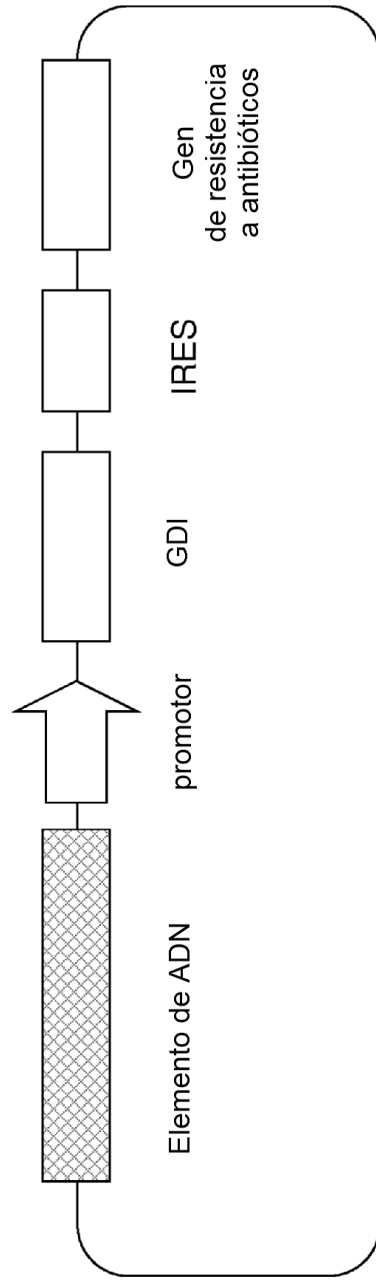
[FIG. 2]



[FIG. 3]

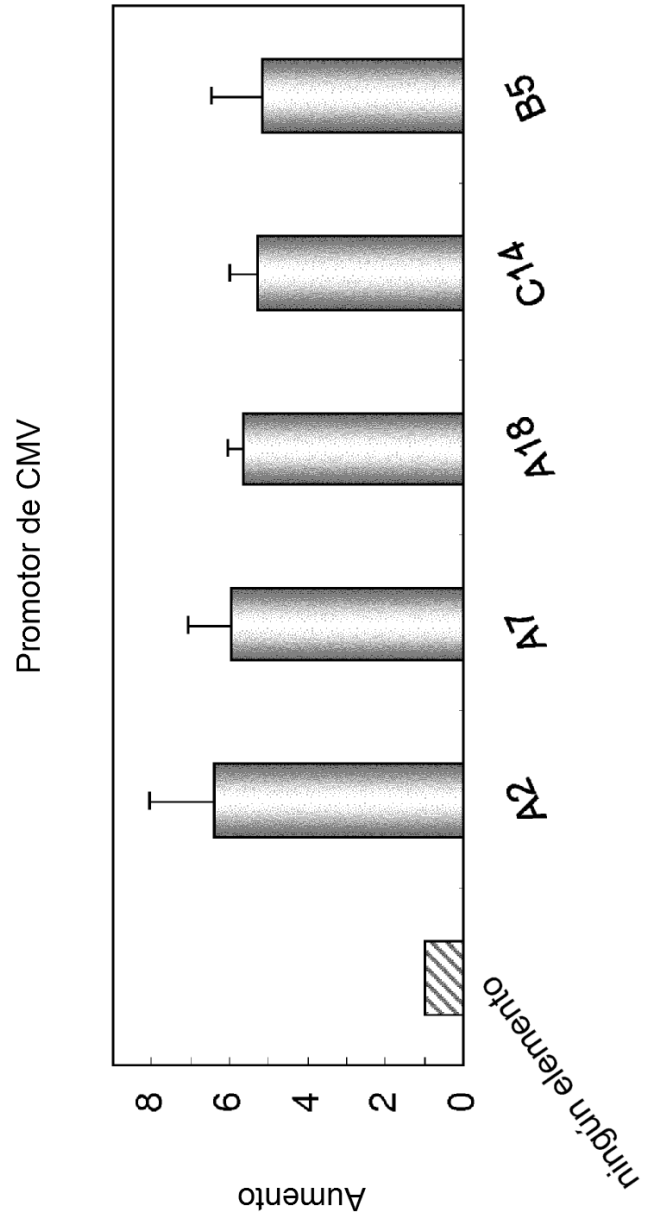


[FIG. 4]

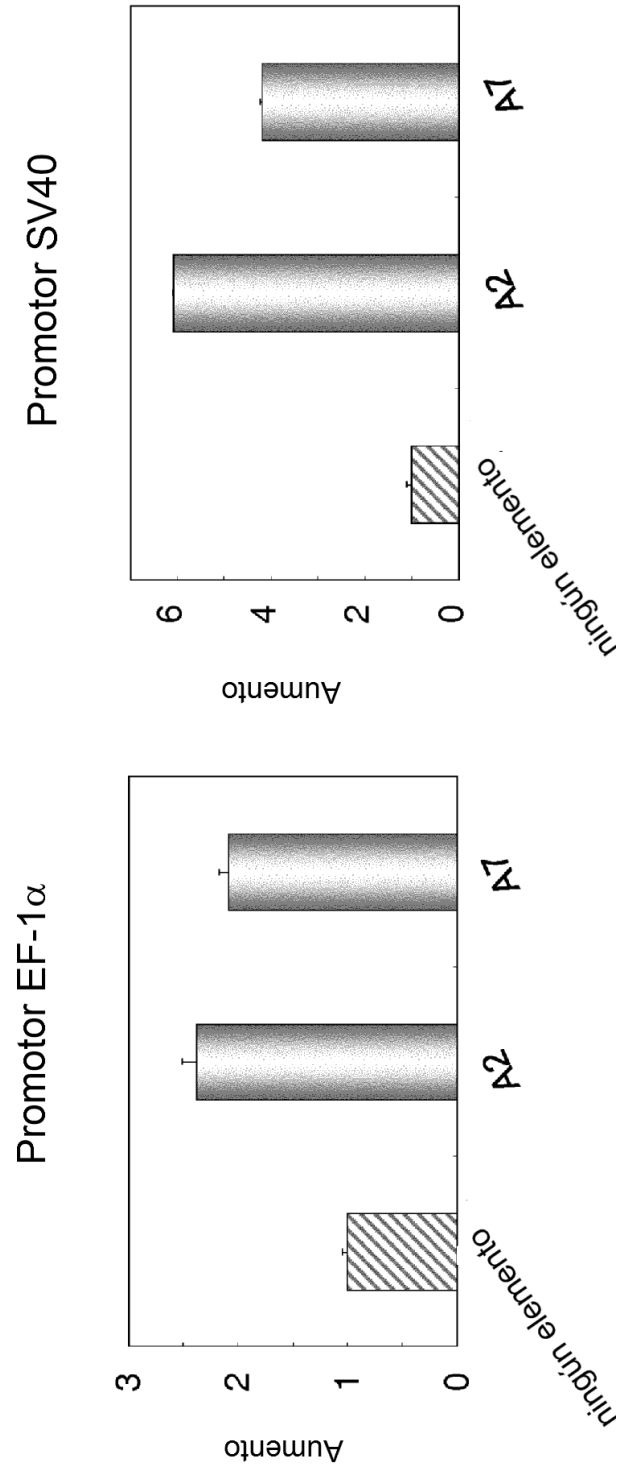


GDI: Gen de interés (SEAP)

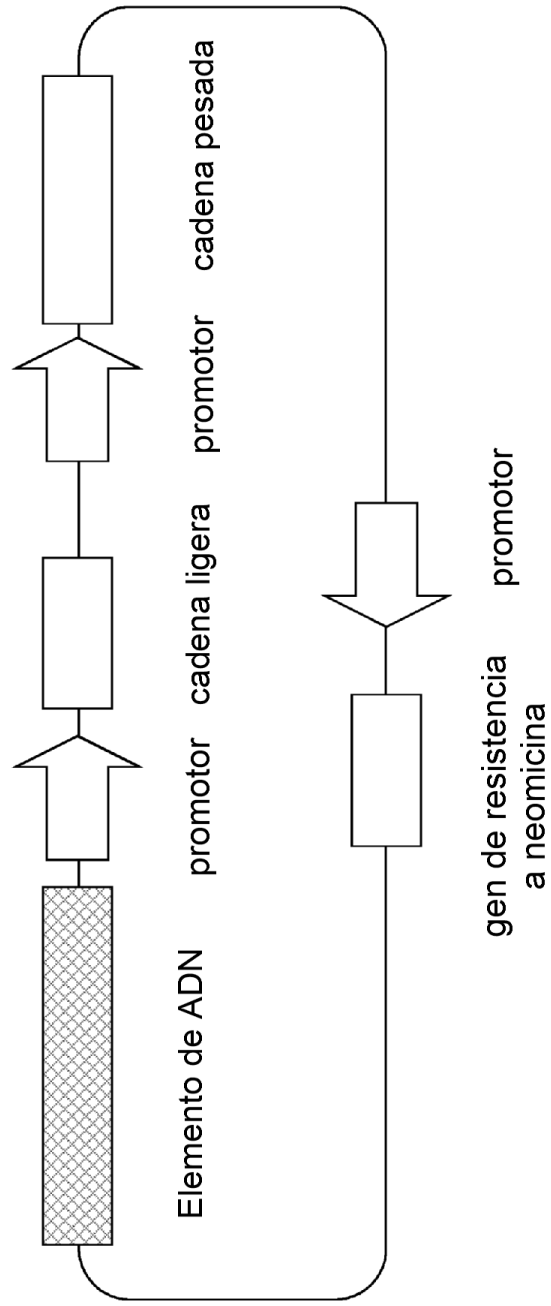
[FIG. 5]



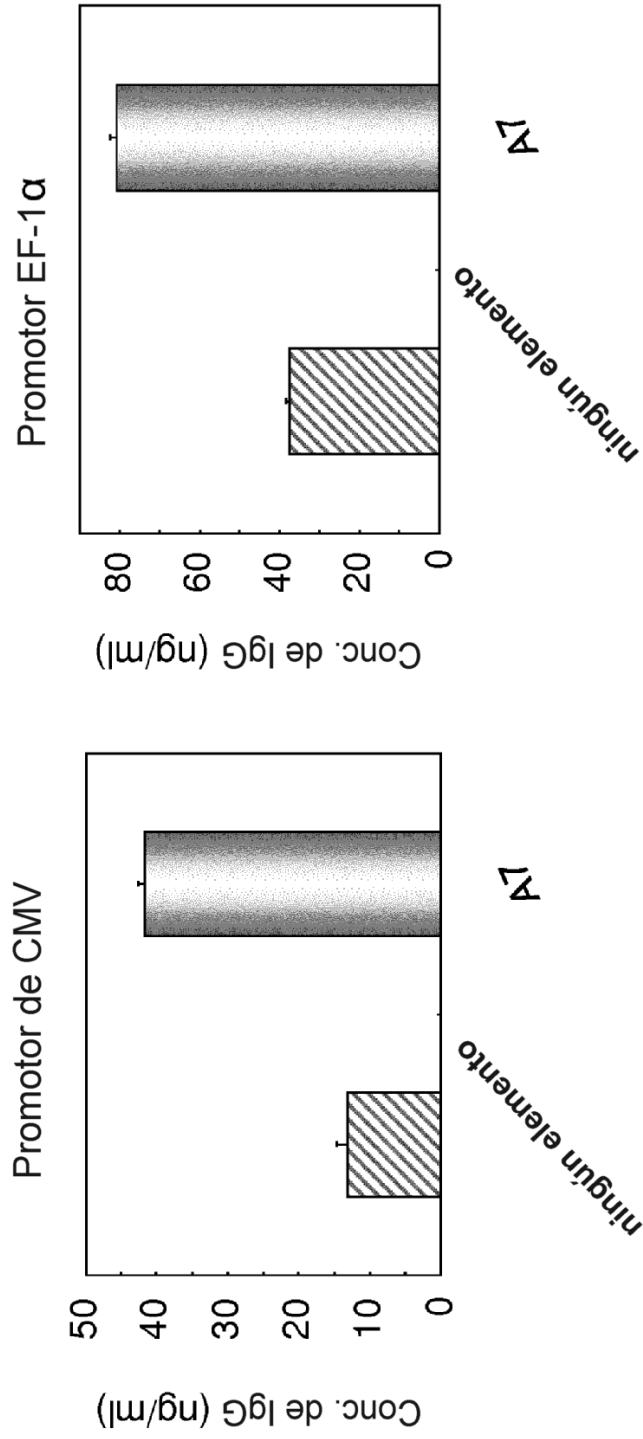
[FIG. 6]



[FIG. 7]



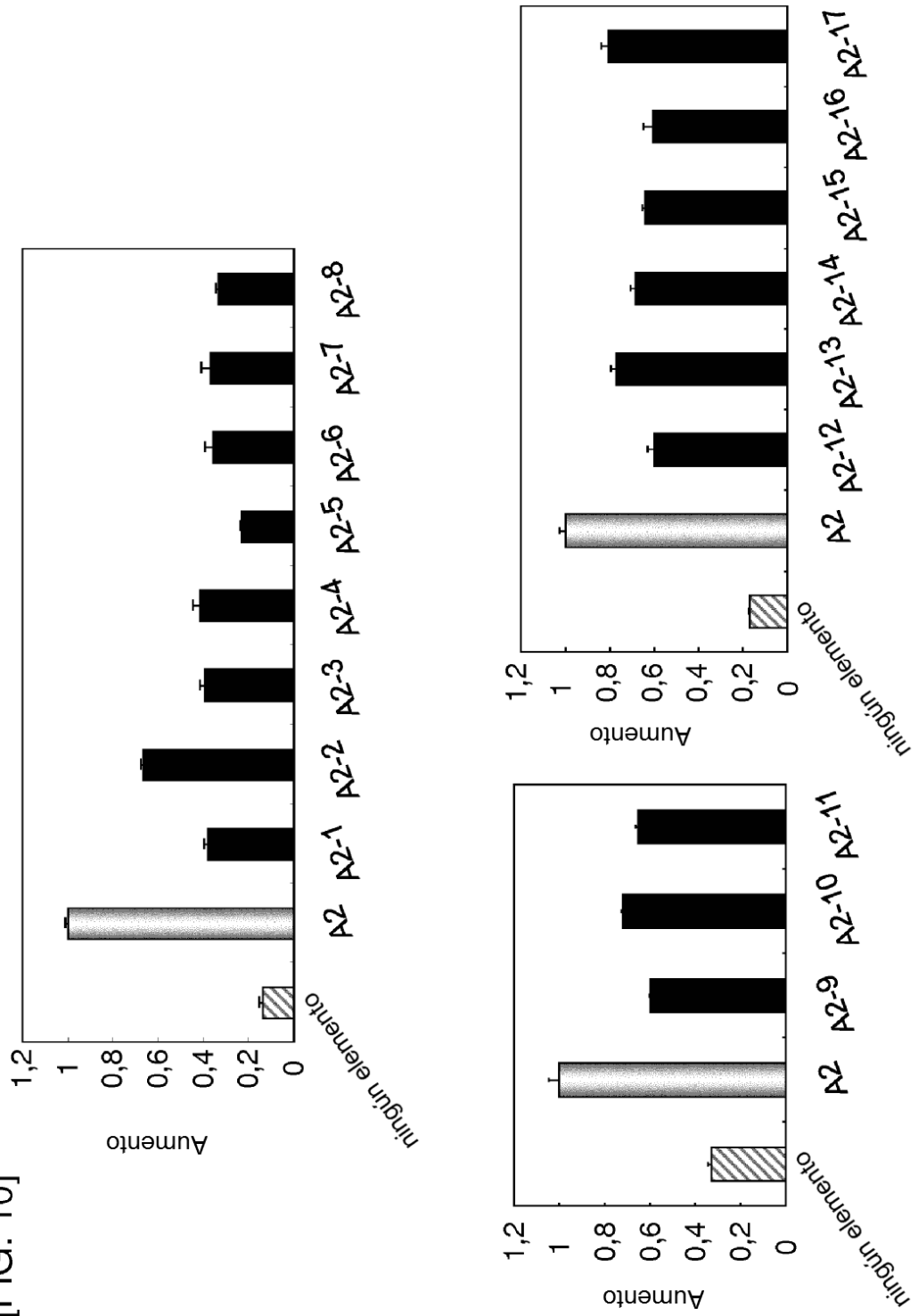
[FIG. 8]



[FIG. 9]

ID	Localización	Localización	Longitud (pb)
A2	80966429	80974878	8450
A2-1	80966429	80969428	3000
A2-2	80969229	80972228	3000
A2-3	80971829	80974878	3050
A2-4	80967129	80969128	2000
A2-5	80967129	80968628	1500
A2-6	80967129	80970128	3000
A2-7	80968429	80971428	3000
A2-8	80970429	80973428	3000
A2-9	80966429	80970128	3700
A2-10	80968429	80972228	3800
A2-11	80969229	80973428	4200
A2-12	80967129	80972228	5100
A2-13	80968429	80973428	5000
A2-14	80969229	80974878	5650
A2-15	80966429	80972228	5800
A2-16	80967129	80973428	6300
A2-17	80968429	80974878	6450

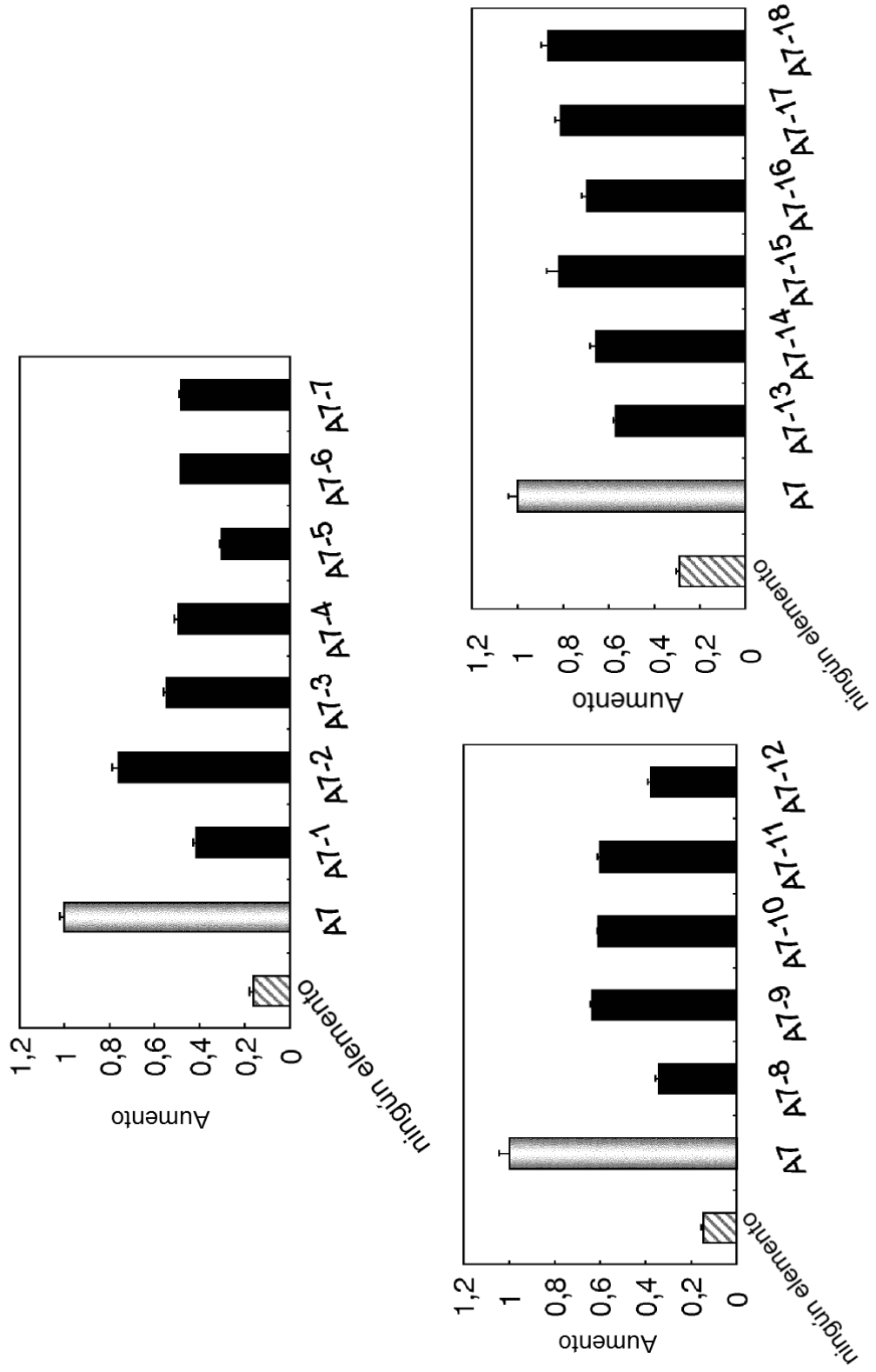
[FIG. 10]



[FIG. 11]

ID	Localización	Localización	Longitud (pb)
A7	88992123	89000542	8420
A7-1	88992723	88995722	3000
A7-2	88995723	89000542	4820
A7-3	88997523	89000542	3020
A7-4	88995523	88998522	3000
A7-5	88993623	88996622	3000
A7-6	88996523	88999522	3000
A7-7	88994523	88997522	3000
A7-8	88992123	88995722	3600
A7-9	88993623	88997522	3900
A7-10	88994523	88998522	4000
A7-11	88995523	88999522	4000
A7-12	88996523	89000542	4020
A7-13	88992123	88997522	5400
A7-14	88993623	88998522	4900
A7-15	88994523	88999522	5000
A7-16	88995523	89000542	5020
A7-17	88992123	88998522	6400
A7-18	88993623	88999522	5900

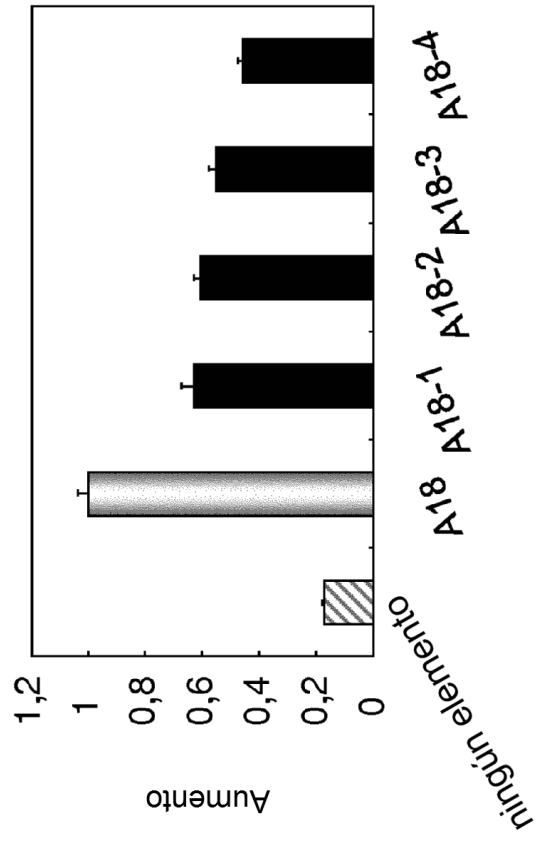
[FIG. 12]



[FIG. 13]

ID	Localización	Localización	Localización	Longitud (pb)
A18	111275976	111284450	111284450	8475
A18-1	111275976	111281015	111281015	5040
A18-2	111276976	111281977	111281977	5002
A18-3	111277976	111282975	111282975	5000
A18-4	111278975	111282975	111282975	4001

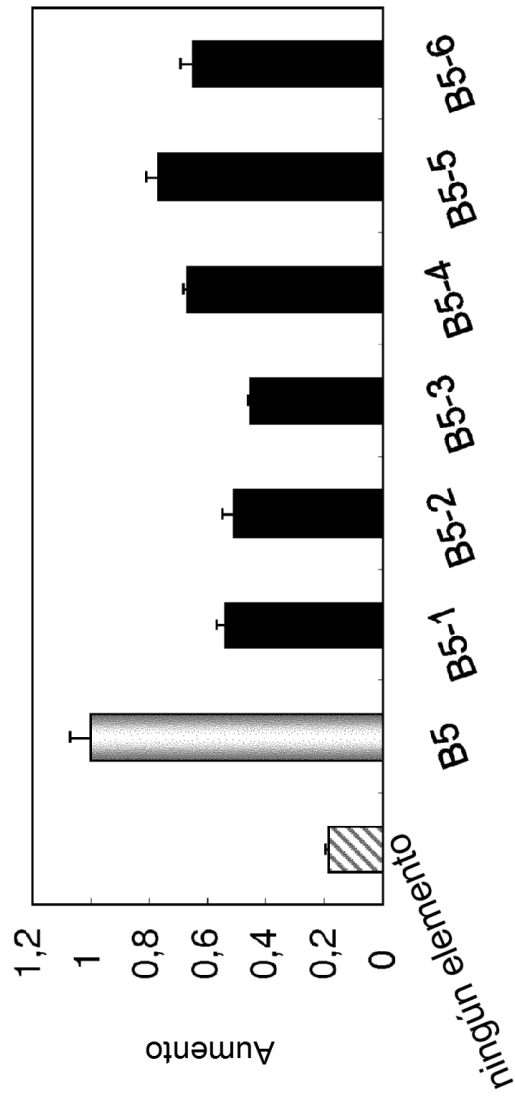
[FIG. 14]



[FIG. 15]

ID	Localización	Localización	Localización	Longitud (pb)
B5	143034684	143043084	143043084	8401
B5-1	143034684	143038684	143038684	4001
B5-2	143034684	143037883	143037883	3200
B5-3	143037174	143040284	143040284	3111
B5-4	143040056	143043084	143043084	3029
B5-5	143035584	143038684	143038684	3101
B5-6	143038684	143041683	143041683	3000

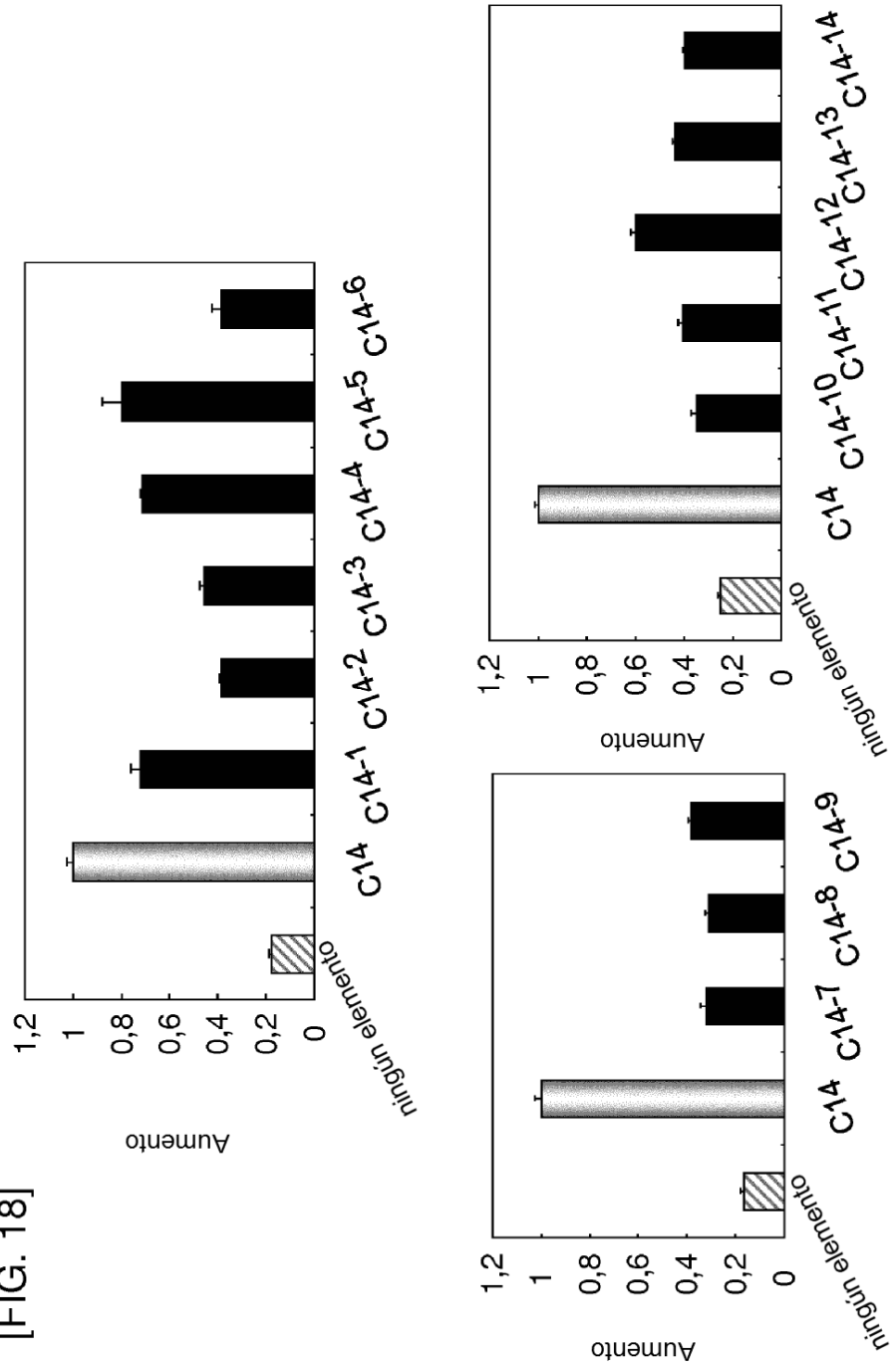
[FIG. 16]

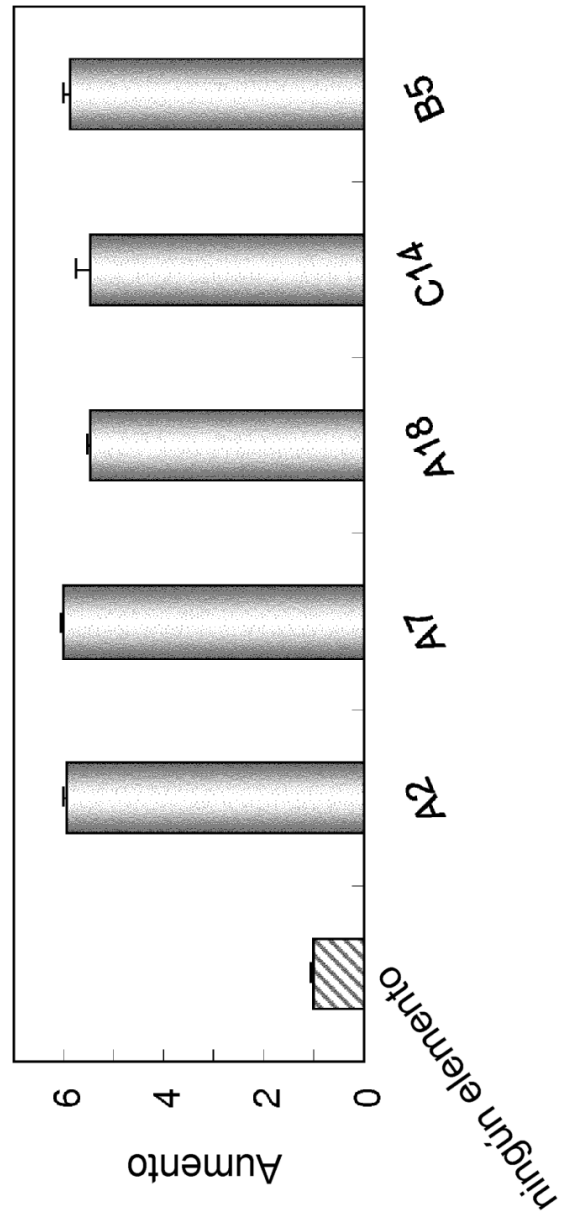


[FIG. 17]

ID	Localización	Localización	Longitud (pb)
C14	46089056	46097482	8427
C14-1	46090015	46093070	3056
C14-2	46091042	46094069	3028
C14-3	46093075	46096174	3100
C14-4	46090015	46097196	7182
C14-5	46090015	46095066	5052
C14-6	46093994	46097196	3203
C14-7	46090015	46094069	4055
C14-8	46092049	46096174	4126
C14-9	46093075	46097196	4122
C14-10	46089056	46094069	5014
C14-11	46091042	46096174	5133
C14-12	46092049	46097196	5148
C14-13	46090015	46096174	6160
C14-14	46091042	46097196	6155

[FIG. 18]





[FIG. 19]

[FIG. 20]

Puntos de inicio y final basándose en A2			Puntos de inicio y final basándose en A7			Puntos de inicio y final basándose en A18		
	Punto de inicio	Punto final		Punto de inicio	Punto final		Punto de inicio	Punto final
A2	1	8450	A7	1	8420	A18	1	8475
A2-1	1	3000	A7-1	601	3600	A18-1	1	5040
A2-2	2801	5800	A7-2	3601	8420	A18-2	1001	6002
A2-3	5401	8450	A7-3	5401	8420	A18-3	2001	7000
A2-4	701	2700	A7-4	3401	6400	A18-4	3000	7000
A2-5	701	2200	A7-5	1501	4500			
A2-6	701	3700	A7-6	4401	7400			
A2-7	2001	5000	A7-7	2401	5400			
A2-8	4001	7000	A7-8	1	3600			
A2-9	1	3700	A7-9	1501	5400			
A2-10	2001	5800	A7-10	2401	6400			
A2-11	2801	7000	A7-11	3401	7400			
A2-12	701	5800	A7-12	4401	8420			
A2-13	2001	7000	A7-13	1	5400			
A2-14	2801	8450	A7-14	1501	6400			
A2-15	1	5800	A7-15	2401	7400			
A2-16	701	7000	A7-16	3401	8420			
A2-17	2001	8450	A7-17	1	6400			
			A7-18	1501	7400			

[FIG. 21]

Puntos de inicio y final
basándose en B5

	Punto de inicio	Punto final
B5	1	8401
B5-1	1	4001
B5-2	1	3200
B5-3	2491	5601
B5-4	5373	8401
B5-5	901	4001
B5-6	4001	7000

Puntos de inicio y final
basándose en C14

	Punto de inicio	Punto final
C14	1	8427
C14-1	960	4015
C14-2	1987	5014
C14-3	4020	7119
C14-4	960	8141
C14-5	960	6011
C14-6	4939	8141
C14-7	960	5014
C14-8	2994	7119
C14-9	4020	8141
C14-10	1	5014
C14-11	1987	7119
C14-12	2994	8141
C14-13	960	7119
C14-14	1987	8141