

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 565 428**

51 Int. Cl.:

C12N 15/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.07.2011 E 13186031 (4)**

97 Fecha y número de publicación de la concesión europea: **20.01.2016 EP 2684961**

54 Título: **Elemento de ADN que tiene actividad de aumento de expresión de un gen exógeno**

30 Prioridad:

07.07.2010 JP 2010154782

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

04.04.2016

73 Titular/es:

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

72 Inventor/es:

**NISHIMYIA, DAISUKE y
INOUE, TATSUYA**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 565 428 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

Elemento de ADN que tiene actividad de aumento de expresión de un gen exógeno

Campo técnico

5 La presente invención se refiere a una célula hospedadora de mamífero cuya capacidad para segregar una proteína exógena se ha aumentado usando un vector de expresión de un gen exógeno que tiene un elemento de ADN y a un procedimiento para producir la proteína exógena usando la célula hospedadora.

Técnica antecedente

10 Debido al desarrollo de las técnicas de recombinación genética, se ha expandido rápidamente el mercado de productos farmacéuticos proteicos tales como proteínas terapéuticas y fármacos de anticuerpos. En particular, los fármacos de anticuerpos tienen una alta especificidad y no producen una reacción inmunitaria adversa incluso cuando se administran al cuerpo humano y por lo tanto, el desarrollo de los mismos se ha llevado a cabo activamente.

15 Como célula hospedadora en la que se produce una proteína farmacéutica tipificada por un fármaco de anticuerpo puede usarse un microorganismo, una célula de insecto, animal o vegetal, una célula de animal o vegetal transgénica, o similares. Con el fin de que la proteína farmacéutica tenga actividad biológica o inmunogenicidad, es esencial la modificación post-traducciona tal como el plegado o la glucosilación y por lo tanto no será adecuado un microorganismo en el que no se pueda llevar a cabo una complicada modificación post-traducciona o una planta que tenga una estructura de glucanos diferente como célula hospedadora que funcione como biorreactor. El uso de una célula de mamífero cultivada, tal como una célula CHO, que sea de una especie estrechamente relacionada con los seres humanos es convencional actualmente, considerando que dicha célula tenga una estructura de glucanos similar a la de los seres humanos y que sea segura, y que pueda llevarse a cabo la modificación post-traducciona usando dicha célula.

20 En los casos en los que se usa una célula de mamífero cultivada como célula hospedadora, existe los problemas de que la velocidad de crecimiento es baja, la productividad es baja, el coste es alto, etc., en comparación con un microorganismo o similar (véase el documento no de patente 1). Además, con el fin de usar un producto farmacéutico proteico en un ensayo clínico, es necesario administrar una gran cantidad del producto. Por lo tanto, la falta de capacidad de producción del mismo también es un problema en todo el mundo. Por consiguiente, con el fin de mejorar la productividad de un gen exógeno en una célula de mamífero en cultivo, se han llevado a cabo hasta ahora muchos estudios de promotores, amplificadores, marcadores de selección farmacológicos, técnicas de amplificación genética y de modificación de cultivos, y similares. Sin embargo, la situación actual es que no se ha establecido aún un sistema capaz de aumentar la expresión génica uniformemente. Como una de las causas de la baja productividad de una proteína exógena, se considera un "efecto de posición" (véase el documento no de patente 2). Cuando un gen exógeno se introduce en una célula hospedadora, se integra aleatoriamente en el genoma cromosómico del hospedador y la transcripción del gen exógeno se ve afectada en gran medida por el ADN que se encuentra alrededor de la región donde se ha integrado el gen exógeno. Un efecto de posición se ve afectado por factores tales como el sitio de inserción, el número de copias, la estructura, etc. del gen exógeno, sin embargo es muy difícil controlar el sitio de inserción en el cromosoma.

35 Con el fin de resolver el problema, se han identificado recientemente secuencias de polinucleótidos reguladoras (también conocidas como elementos de ADN) tales como la región de control del locus (LCR), la región de anclaje del armazón/matriz (S/MAR), un aislante, el elemento ubicuo de apertura de cromatina (UCOE), y un anti-represor (elemento STAR) (véanse los documentos no de patente 3 a 6). La LCR no es necesaria para abrir la estructura de cromatina en un locus génico endógeno. Sin embargo, la LCR es un elemento regulador de la transcripción que tiene la capacidad de abrir la estructura de cromatina alrededor del ADN donde se ha integrado el gen exógeno y de remodelar un amplio tramo de cromatina cuando se usa junto con la unidad de expresión de un gen exógeno y se dice que necesita una región rica en AT (véase el documento no de patente 7).

40 El elemento de ADN mencionado anteriormente tipificado por la LCR se usa a menudo en combinación con un promotor, se sabe que en los casos en los que se usa el elemento de ADN en combinación con un promotor, el nivel de expresión de un gen exógeno aumenta en comparación con los casos en los que solamente se usa el promotor. Sin embargo, hasta ahora se han comunicado muy pocos tipos de elementos de ADN, y los distintos mecanismos que contribuyen al aumento de la expresión de un gen exógeno son diferentes entre sí. Además, incluso si se usan en combinación un elemento de ADN y un promotor, no se producen cantidades suficientes de una proteína terapéutica bajo el control del elemento de ADN y el promotor. Por lo tanto, no se puede decir que se haya adquirido un conocimiento suficiente de un elemento de ADN que sea capaz de aumentar la productividad de una proteína exógena.

55 El documento no de patente 8 desvela ADN genómico de *Homo sapiens* de 135.637 nucleótidos de longitud.

Por consiguiente, un objetivo de la invención es proporcionar un procedimiento para aumentar la producción de una proteína exógena para su uso en un producto farmacéutico proteico usando un elemento de ADN que tenga una alta

actividad de aumento de la expresión de un gen exógeno en una célula hospedadora tal como una célula de mamífero en cultivo.

Listado de citas

Bibliografía no de patente

- 5 BNP 1: Florian M. Wurm. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol. 22(11): 1393-1398
- BNP 2: Ted H. J. Kwaks y Arie P. Otte. (2006) Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. TRENDS in Biotechnol. 24(3): 137-142
- 10 BNP 3: Pierre-Alain Girod, Duc-Quang Nguyen y col. (2007) Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nat. Methods. 4(9): 747-753
- BNP 4: Adam C. Bell, Adam G. West, Gary Felsenfeld (2001) Insulators and Boundaries: Versatile Regulatory Elements in the Eukaryotic Genome, Science 291: 447-450
- BNP 5: Steven Williams, Tracey Mustoe y col. (2005) CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnol. 5(17): 1-9
- 15 BNP 6: Arie P. Otte, Ted H. J. Kwaks y col. (2007) Various Expression-Augmenting DNA Elements Benefit from STAR-Select, a Novel High Stringency Selection System for Protein Expression. Biotechnol. Prog. 23: 801-807
- BNP 7: Qiliang Li, Kenneth R. Peterson, Xiangdong Fang y George Stamatoyannopoulos, (2002) Locus control regions, Blood 100(9): 3077-3086
- 20 BNP 8: BASE DE DATOS EMBL [en línea] 19 de marzo de 2001 (19-03-2001), "Homo sapiens genomic DNA, chromosome 11q, clone: RP11-643G5.", recuperado del número de registro EM_STD:AP003400 n.º de registro de la base de datos AP00340

Sumario de la invención

Problemas técnicos

- 25 Como se ha descrito anteriormente, aún no hay muchos tipos de elementos ADN que sean secuencias de polinucleótidos reguladoras y además, hay pocos elementos de ADN entre ellos que sean altamente eficaces en potenciar la expresión de un gen exógeno. Un objeto de la invención es proporcionar un procedimiento para conseguir una alta expresión de manera estable en una célula de mamífero usando un elemento de ADN que aumenta la activación de la transcripción estando acompañado por un cambio en la estructura de cromatina alrededor
- 30 de un locus génico en el que se ha introducido una unidad de expresión de un gen exógeno, etc.

Solución al problema

- Los presentes inventores han hecho estudios exhaustivos con el fin de resolver los problemas anteriores y como resultado, descubrieron que la productividad y secreción de una proteína exógena que se va a expresar puede mejorarse usando uno o más de los tipos específicos de elementos de ADN en una célula de mamífero en cultivo, y
- 35 de esta manera, completaron la invención.

Es decir, la invención incluye las siguientes invenciones.

- (1) Un polinucleótido que consiste en una secuencia de polinucleótido representada por la SEC ID N°: 1 en el listado de secuencias.
- 40 (2) Un polinucleótido que comprende al menos 3000 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias en el que el polinucleótido es un fragmento parcial de la SEC ID N°: 2 y tiene una actividad de potenciar la expresión de un gen exógeno.
- (3) Un polinucleótido que comprende al menos 2000 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias en el que el polinucleótido es un fragmento parcial de la SEC ID N°: 2 y tiene una actividad de potenciar la expresión de un gen exógeno.
- 45 (4) Un polinucleótido que comprende al menos 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias en que el polinucleótido es un fragmento parcial de la SEC ID N°: 2 y tiene una actividad de potenciar la expresión de un gen exógeno.
- (5) Un polinucleótido que consiste en una secuencia de polinucleótido que tiene una homología del 95 % o más respecto de la secuencia de polinucleótido de acuerdo con uno cualquiera de los puntos (1) a (4) y que tiene una actividad de potenciar la expresión de un gen exógeno.
- 50 (6) Un polinucleótido que consiste en una secuencia de polinucleótido que tiene una homología del 99 % o más respecto de la secuencia de polinucleótido de acuerdo con uno cualquiera de los puntos (1) a (4) y tiene una actividad de potenciar la expresión de un gen exógeno.
- (7) Un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótidos de acuerdo con uno cualquiera de los puntos (1) a (6) y que tiene una actividad de potenciar la expresión de un gen exógeno.
- 55 (8) Un polinucleótido que consiste en una secuencia de polinucleótido que contiene:

- al menos una de las secuencias de polinucleótidos de acuerdo con uno cualquiera de los puntos (1) a (6); y
- al menos uno de

- 5 (a) una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos;
- (b) una secuencia de polinucleótido representada por la SEC ID N°: 3 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos;
- 10 (c) una secuencia de polinucleótido representada por la SEC ID N°: 4 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos;
- (d) una secuencia de polinucleótido representada por la SEC ID N°: 5 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial;

- 15 o bien;
- y que tiene una actividad de potenciar la expresión de un gen exógeno.

(9) Un vector que comprende un elemento de ADN que tiene la secuencia de polinucleótido de:

- (i) el polinucleótido de acuerdo con uno cualquiera de los puntos (1) a (8);
- 20 (ii) un polinucleótido que comprende al menos 3000, 2000, o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
- (iii) un polinucleótido que comprende una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más respecto de la secuencia de polinucleótido del polinucleótido que consiste en al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno; o
- 25 (iv) un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótidos del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno.

(10) Un vector de acuerdo con el punto (9), en que el vector es un vector de expresión de un gen exógeno que comprende además un gen exógeno, y en el que el elemento de ADN está posicionado con respecto al gen exógeno de una manera funcionalmente eficaz.

(11) El vector de expresión de un gen exógeno de acuerdo con el punto (10), en el que la proteína codificada por el gen exógeno es una proteína multimérica, o una proteína hetero-multimérica.

35 (12) El vector de expresión de un gen exógeno de acuerdo con el punto (11), en el que la proteína codificada por el gen exógeno es una proteína hetero-multimérica y es un anticuerpo o un fragmento funcional del mismo.

(13) Una célula transformada dentro de la que:

- a) se han introducido el vector de acuerdo con el punto (9) y un gen exógeno, o
- b) se ha introducido el vector de expresión de un gen exógeno de acuerdo con uno cualquiera de los puntos (10) a (12).

40 (14) La célula transformada de acuerdo con el punto (13), en la que la célula es una célula en cultivo procedente de un mamífero.

(15) La célula transformada de acuerdo con el punto (14), en la que la célula cultivada procedente de un mamífero es una célula que se selecciona entre el grupo que consiste en células COS-1, células 293, y células CHO.

45 (16) La célula transformada de acuerdo con uno cualquiera de los puntos (13) a (15), en la que la proteína codificada por el gen exógeno es una proteína multimérica o una proteína hetero-multimérica.

(17) La célula transformada de acuerdo con el punto (16), en la que la proteína codificada por el gen exógeno es una proteína hetero-multimérica y es un anticuerpo o un fragmento funcional del mismo.

50 (18) Un procedimiento para producir una proteína caracterizado porque comprende cultivar la célula transformada de acuerdo con uno cualquiera de los puntos (13) a (17) y obtener la proteína codificada por el gen exógeno a partir del producto del cultivo resultante.

(19) Un procedimiento para potenciar la expresión de un gen exógeno en una célula transformada dentro de la que se ha introducido un gen exógeno o un vector de expresión de un gen exógeno, caracterizado porque usa:

- 55 (i) un polinucleótido de acuerdo con uno cualquiera de los puntos (1) a (8);
- (ii) un polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
- (iii) un polinucleótido que comprende una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más respecto de la secuencia de polinucleótido del polinucleótido que consiste en al menos

3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
 (iv) un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótidos del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno; o
 (v) un vector de acuerdo con el punto (9) o un vector de expresión de un gen exógeno de acuerdo con uno cualquiera de los puntos (10) a (12).

(20) El uso de:

- (i) el polinucleótido de acuerdo con uno cualquiera de los puntos (1) a (8);
- (ii) un polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
- (iii) un polinucleótido que consiste en una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más respecto de la secuencia de polinucleótido del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de aumento de la expresión de un gen exógeno;
- (iv) un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótidos del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno; o
- (v) un vector de acuerdo con el punto (9),

para potenciar la expresión de un gen exógeno en una célula transformada.

Efectos ventajosos de la invención

De acuerdo con la invención, al introducir un vector de expresión de un gen exógeno usando un elemento de ADN en una célula hospedadora de mamífero, se puede potenciar significativamente la expresión de un gen exógeno para una proteína terapéutica, un anticuerpo o similar.

Breve descripción de los dibujos

- [Fig. 1] La Fig. 1 muestra un gráfico en el que se confirma por amplificación de una región de GAPDH que en una muestra sometida a ChIP-on-chip se inmunoprecipitó la cromatina específicamente con un anticuerpo anti-histona H3 acetilada.
- [Fig. 2] La Fig. 2 es una vista esquemática de un vector de expresión SEAP en el que se ha insertado un elemento de ADN.
- [Fig. 3] La Fig. 3 es un gráfico que muestra la expresión de SEAP bajo el control de un promotor de CMV en una línea celular CHO que expresa establemente sin un elemento de ADN o con un elemento de ADN A2, A7, A18, B5 o C14. Se confirmaron los efectos de los elementos A2, A7, A18, B5 y C14 sobre el potenciamiento de la expresión.
- [Fig. 4] La Fig. 4 comprende dos gráficos que muestran la expresión de SEAP bajo el control de un promotor EF-1α o SV40 en una línea celular CHO que expresa establemente sin un elemento de ADN o con un elemento de ADN A2 o A7. Se confirmaron los efectos de los elementos A2 y A7 sobre la potenciación de la expresión.
- [Fig. 5] La Fig. 5 es una vista esquemática de un vector de expresión de anticuerpo (co-expresión del gen X de cadena pesada y cadena ligera de anticuerpo) en el que se ha insertado un elemento de ADN.
- [Fig.6] La Fig. 6 comprende dos gráficos que muestran los niveles de secreción (medidos mediante un procedimiento ELISA) de un anticuerpo bajo el control de un promotor CMV o EF-1α en una línea celular CHO que expresaba establemente sin un elemento de ADN o con un elemento de ADN A7. Se confirmó el efecto del elemento de ADN A7 sobre la potenciación de la expresión.
- [Fig. 7] La Fig. 7 es una tabla que muestra las longitudes de secuencia del elemento de ADN A2 y de secuencias relacionadas.
- [Fig.8] La Fig. 8 comprende tres gráficos que muestran la expresión de SEAP en una línea celular CHO que expresaba establemente sin un elemento de ADN o con un elemento de ADN A2 o una secuencia relacionada. Se confirmaron los efectos del elemento de ADN A2 y de las secuencias relacionadas sobre la potenciación de la expresión.
- [Fig. 9] La Fig. 9 es una tabla que muestra las longitudes de secuencia del elemento de ADN A7 y de secuencias relacionadas.
- [Fig. 10] La Fig. 10 comprende tres gráficos que muestran la expresión de SEAP en una línea celular CHO que expresaba establemente sin un elemento de ADN o con un elemento de ADN A7 o una secuencia relacionada. Se confirmaron los efectos del elemento de ADN A7 y secuencias relacionadas sobre la potenciación de la expresión.
- [Fig. 11] La Fig. 11 es una tabla que muestra las longitudes de secuencia de un elemento de ADN A18 y de secuencias relacionadas.

[Fig. 12] La Fig. 12 es un gráfico que muestra la expresión de SEAP en una línea celular CHO que expresaba establemente sin un elemento de ADN o con un elemento de ADN A18 o una secuencia relacionada. Se confirmaron los efectos del elemento de ADN A18 y secuencias relacionadas sobre la potenciación de la expresión.

5 [Fig. 13] La Fig. 13 es una tabla que muestra las longitudes de secuencia de un elemento de ADN B5 y de secuencias relacionadas.

[Fig. 14] La Fig. 14 es un gráfico que muestra la expresión de SEAP en una línea celular CHO que expresaba establemente sin un elemento de ADN o con un elemento de ADN B5 o una secuencia relacionada. Se confirmaron los efectos del elemento de ADN B5 y secuencias relacionadas sobre el aumento de la expresión.

10 [Fig. 15] La Fig. 15 es una tabla que muestra las longitudes de secuencia de un elemento de ADN C14 y de secuencias relacionadas.

[Fig. 16] La Fig. 16 es un gráfico que muestra la expresión de SEAP en una línea celular CHO que expresaba establemente sin un elemento de ADN o con un elemento de ADN C14 o una secuencia relacionada. Se confirmaron los efectos del elemento de ADN C14 y secuencias relacionadas sobre el aumento de la expresión.

15 [Fig. 17] La Fig. 17 es un gráfico que muestra la expresión de SEAP en una línea celular HEK293 que expresaba establemente sin un elemento de ADN o con un elemento de ADN A2, A7, A18, B5, o C14. Se confirmaron los efectos de los elementos de ADN A2, A7, A18, B5, y C14 sobre el aumento de la expresión en células HEK293.

[Fig. 18] La Fig. 18 es una vista que muestra los nucleótidos del punto de partida y final basándose en la secuencia de longitud completa de un elemento de ADN A2, A7, o A18.

20 [Fig. 19] La Fig. 19 es una vista que muestra los nucleótidos del punto de partida y final basándose en la secuencia de longitud completa de un elemento de ADN B5 o C14.

Descripción de las realizaciones

En lo sucesivo, la invención se describirá específicamente en referencia a los ejemplos. Sin embargo, estos ejemplos no limitan el ámbito técnico de la invención. Los plásmidos, enzimas de restricción, enzimas de modificación de ADN y similares que se van a usar en los ejemplos de la invención son productos disponibles en el comercio y pueden usarse de acuerdo con procedimientos comunes. Además, los procedimientos que se usan para la clonación de ADN, determinación de secuencias de polinucleótidos, transformación de una células hospedadoras, cultivo de una célula hospedadora transformada, aislamiento de un anticuerpo a partir de una solución de cultivo obtenida, purificación de un anticuerpo y similares también son bien conocidos por los expertos en la técnica o están disponibles en la bibliografía.

25 El término "gen" tal como se usa en el presente documento incluye no solo el ADN, sino también el ARNm del mismo, el ADNc y el ARN del mismo.

El término "polinucleótido" tal como se usa en el presente documento se usa con el mismo significado que ácido nucleico y también incluye ADN, ARN, sondas, oligonucleótidos y cebadores.

35 Los términos "polipéptido" y "proteína" tal como se usan en el presente documento se usan sin distinción.

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

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

40 La expresión "proteína exógena" tal como se usa en el presente documento se refiere a una proteína codificada por el gen exógeno.

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

45 La expresión "actividad de potenciar la expresión de un gen exógeno" tal como se usa en el presente documento se refiere a la actividad de potenciar la producción de una proteína exógena en una célula hospedadora creando un ambiente ventajoso para la transcripción y traducción del ADN alrededor de la unidad de expresión génica que contiene un gen exógeno y que mejora significativamente la eficacia de la transcripción y traducción.

50 La expresión "elemento de ADN" tal como se usa en el presente documento se refiere a un polinucleótido que tiene una actividad de potenciar la expresión de un gen exógeno en los casos en los que el polinucleótido se localiza en la vecindad de una unidad de expresión génica o en un vector de expresión de un gen exógeno que contiene una unidad de expresión génica.

55 La expresión "fragmento funcional de un anticuerpo" tal como se usa en el presente documento se refiere a un fragmento parcial de un anticuerpo que tiene una actividad de unión a antígeno y que 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.

1. Elemento de ADN que se usa para potenciar la expresión de un gen exógeno

Tal como se muestra en el ejemplo 1, se puede obtener un elemento de ADN de acuerdo con la invención usando la interacción entre la histona H3 acetilada y el ADN genómico. En general, se dice que la acetilación de histonas (H3 y H4) se asocia con la activación de la transcripción, y se han defendido dos teorías principales. Una teoría es que la acetilación de histonas se asocia con un cambio en la conformación del nucleosoma de tal modo que las colas de histonas se acetilan, neutralizándose eléctricamente de esta manera, lo que da como resultado un debilitamiento de 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 las histonas se asocia con el reclutamiento de varios factores de transcripción (Nakatani Y. (2001) Histone acetylases - versatile players. Genes Cells. 6(2): 79-86). En cualquiera de las dos teorías, hay una elevada posibilidad de que la acetilación de las histonas se asocie con la activación de la transcripción y llevando a cabo la inmunoprecipitación de cromatina (ChIP) usando un anticuerpo anti-histona H3 acetilada, es posible concentrar un elemento de ADN que interactúe con la histona H3 acetilada.

En la presente invención, A7 es un ejemplo de un elemento de ADN que se va a usar para aumentar la expresión de un gen exógeno. A7 se localiza en la región desde 88992123 a 89000542 del cromosoma 11 humano y es una secuencia de polinucleótido de 8420 pb que tiene un contenido de AT del 64,52 %. La secuencia de polinucleótido de A7 está representada por la SEC ID N°: 2 en el listado de secuencias.

A2, A18, B5 y C14 son ejemplos de elementos de ADN similares. El A2 se localiza en la región desde 80966429 a 80974878 del cromosoma 15 humano y es una secuencia de polinucleótido de 8450 pb, que tiene un contenido de AT del 62,2 %. La secuencia de polinucleótido de A2 está representada por la SEC ID N°: 1 en el listado de secuencias.

A18 está localizado en la región desde 111275976 a 111284450 del cromosoma 4 humano y es una secuencia de polinucleótido de 8475 pb, que tiene un contenido de AT del 62,54 %. La secuencia de polinucleótido de A18 está representada por la SEC ID N°: 3 en el listado de secuencias.

B5 se localiza en la región desde 143034684 a 143043084 del cromosoma 1 humano y es una secuencia de polinucleótido de 8401 pb, que tiene un contenido de AT del 66,37 %. La secuencia de polinucleótido de B5 está representada por la SEC ID N°: 4 en el listado de secuencias.

Finalmente, C14 se localiza en la región desde 46089056 a 46097482 del cromosoma 11 humano y es una secuencia de polinucleótido de 8427 pb, que tiene un contenido de AT del 63,81 %. La secuencia de polinucleótido de C14 está representada por la SEC ID N°: 5 en el listado de secuencias.

En la invención, la actividad de aumento de la expresión de un gen exógeno del elemento de ADN se puede ensayar usando la actividad de una proteína codificada por un gen indicador tal como SEAP como índice. En los casos en los que de la actividad de una proteína indicadora en presencia del elemento de ADN aumenta, preferentemente el doble o más, más preferentemente el cuádruple o más, aún más preferentemente el quíntuple o más en comparación con el caso en el que no está presente el elemento de ADN, se puede determinar que el elemento de ADN tiene una actividad de potenciar la expresión de un gen exógeno. Incluso en los casos en los que la actividad aumenta al doble o más, se espera esto reduzca la escala del cultivo celular y el tiempo de cultivo celular, y como resultado, es posible aumentar el rendimiento y reducir el coste del cultivo celular. Si aumenta el rendimiento, entonces es posible suministrar de manera estable una proteína exógena para su uso como agente farmacéutico. Además, si se reduce el coste del cultivo celular, se reduce el coste de la proteína exógena que se usa como agente farmacéutico, y también se reduce la carga financiera sobre los pacientes a quienes se va a administrar la proteína exógena.

En la invención, se puede usar el elemento de ADN A7 solo, y pueden usarse dos o más copias del elemento de ADN A7. Como alternativa, puede usarse al menos un elemento de ADN A7 en combinación con al menos uno de los diferentes tipos de elementos de ADN anteriores.

A2, A18, B5 y C14 son ejemplos preferidos de diferentes tipos de elementos de ADN para su uso en la invención.

El elemento de ADN que se va a usar tal como se ha descrito anteriormente puede ser una secuencia de polinucleótido que comprende una secuencia de polinucleótido que tiene una homología del 95 % o más respecto de cualquiera de las secuencias de polinucleótidos representadas por las SEC ID N°: 1 a 5 y que tiene una actividad de potenciar la expresión de un gen exógeno. La homología del 95 % o más es más preferentemente una homología del 99 % o más. La búsqueda de homología de secuencia de polinucleótido puede llevarse a cabo en, por ejemplo, la Base de datos de ADN de Japón o similares usando un programa tal como FASTA o BLAST.

El elemento de ADN que para su uso tal como se ha descrito anteriormente puede ser un elemento de ADN que hibride con un polinucleótido que consiste en una secuencia de polinucleótido complementaria a un polinucleótido que consiste en una secuencia de polinucleótido seleccionada entre el grupo que consiste en las secuencias de polinucleótidos representadas por las SEC ID N°: 1 a 5 en condiciones rigurosas y que tiene una actividad de potenciar la expresión de un gen exógeno.

La expresión "condiciones rigurosas" tal como se usa en el presente documento se refiere a condiciones en las que se forma lo que se denomina un híbrido específico pero no se forma un híbrido no específico. Por ejemplo, son condiciones rigurosas ejemplares las condiciones en las que hibrida una hebra complementaria de un ácido nucleico que consiste en una secuencia de polinucleótido que tiene una elevada homología, es decir, una secuencia de polinucleótido que tiene una homología del 95 % o más, más preferentemente del 99 % o más respecto de una secuencia de polinucleótido seleccionada entre el grupo que consiste en las secuencias nucleotídicas representadas por las SEC ID N°: 1 a 5 y en las que no hibrida una hebra complementaria de un ácido nucleico que comprende una secuencia de polinucleótido que tenga menor homología. Para ser más específicos, se pueden ejemplificar condiciones en las que la concentración de sal sódica sea de 15 a 750 mM, preferentemente de 50 a 750 mM, más preferentemente de 300 a 750 mM, la temperatura sea de 25 a 70 °C, preferentemente de 50 a 70 °C, más preferentemente de 55 a 65 °C, y la concentración de formamida sea del 0 al 50 %, preferentemente del 20 al 50 %, más preferentemente del 35 al 45 %. Además, se pueden ejemplificar como condiciones rigurosas, condiciones para el lavado de un filtro tras la hibridación en las que la concentración de sal de sodio sea generalmente de 15 a 600 mM, preferentemente de 50 a 600 mM, más preferentemente de 300 a 600 mM, y la temperatura sea de 50 a 70 °C, preferentemente de 55 a 70 °C, más preferentemente de 60 a 65 °C.

Un experto en la técnica puede obtener fácilmente un gen homólogo con referencia a Molecular Cloning (Sambrook, J. y col., Molecular Cloning: a Laboratory Manual 2ª ed., Cold Spring Harbor Laboratory Press, 10 Skyline Drive Plainview, NY (1989)) o similar. Además, la homología de la secuencia de polinucleótido mencionada anteriormente se puede determinar con una búsqueda con FASTA o una búsqueda con BLAST de la misma manera.

La introducción de una mutación (eliminación, sustitución y/o adición) en la secuencia de polinucleótido mencionada anteriormente se puede llevar a cabo mediante un procedimiento conocido en este campo técnico tal como un procedimiento Kunkel o un procedimiento dúplex con espacios, o basándose en este procedimiento. Por ejemplo, se puede usar un kit de introducción de mutación usando un procedimiento de mutagénesis de sitio dirigido (por ejemplo, Mutant-K (fabricado por TaKaRa Bio, Inc.), Mutant-G (fabricado por TaKaRa Bio, Inc.), o un kit LA PCR *in vitro* de la serie Mutagenesis (fabricado por TaKaRa Bio, Inc.)), o similares. Dicho polinucleótido mutado se puede usar también como el elemento de ADN de la invención.

Como el elemento de ADN de la invención, se puede usar un fragmento parcial que comprenda al menos 3000 o al menos 2000 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias. Los ejemplos de dichos fragmentos parciales incluyen: A7-1 a A7-18, que son fragmentos parciales de A7. También se desvelan en el presente documento A2-1 a A2-17; 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 está limitado a estos fragmentos parciales siempre que tenga una actividad de potenciar la expresión de un gen exógeno.

En la invención, puede usarse cualquiera de los fragmentos de A7 anteriores solo y también se pueden usar dos o más copias de un tipo de los fragmentos parciales de A7. Como alternativa, se pueden usar en combinación dos o más tipos diferentes de los fragmentos parciales de A7. Además, se puede usar en combinación una secuencia de longitud completa y/o un fragmento parcial del elemento de ADN A7 mencionado anteriormente. Además, se puede usar al menos una secuencia de longitud completa y/o un fragmento parcial del elemento de ADN A7 en combinación con al menos una secuencia de longitud completa y/o un fragmento parcial de diferentes elementos de ADN.

En cuanto a las secuencias de polinucleótido de los respectivos fragmentos de A2, A2-1 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 3000 de la SEC ID N°: 1 en el listado de secuencias; A2-2 se corresponde con la secuencia de polinucleótido de los nucleótidos 2801 a 5800 de la SEC ID N°: 1 en el listado de secuencias; A2-3 se corresponde con la secuencia de polinucleótido de los nucleótidos 5401 a 8450 de la SEC ID N°: 1 en el listado de secuencias; A2-4 se corresponde con la secuencia de polinucleótido de los nucleótidos 701 a 2700 de la SEC ID N°: 1 en el listado de secuencias; A2-5 se corresponde con la secuencia de polinucleótido de los nucleótidos 701 a 2200 de la SEC ID N°: 1 en el listado de secuencias; A2-6 se corresponde con la secuencia de polinucleótido de los nucleótidos 701 a 3700 de la SEC ID N°: 1 en el listado de secuencias; A2-7 se corresponde con la secuencia de polinucleótido de los nucleótidos 2001 a 5000 de la SEC ID N°: 1 en el listado de secuencias; A2-8 se corresponde con la secuencia de polinucleótido de los nucleótidos 4001 a 7000 de la SEC ID N°: 1 en el listado de secuencias; A2-9 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 3700 de la SEC ID N°: 1 en el listado de secuencias; A2-10 se corresponde con la secuencia de polinucleótido de los nucleótidos 2001 a 5800 de la SEC ID N°: 1 en el listado de secuencias; A2-11 se corresponde con la secuencia de polinucleótido de los nucleótidos 2801 a 7000 de la SEC ID N°: 1 en el listado de secuencias; A2-12 se corresponde con la secuencia de polinucleótido de los nucleótidos 701 a 5800 de la SEC ID N°: 1 en el listado de secuencias; A2-13 se corresponde con la secuencia de polinucleótido de los nucleótidos 2001 a 7000 de la SEC ID N°: 1 en el listado de secuencias; A2-14 se corresponde con la secuencia de polinucleótido de los nucleótidos 2801 a 8450 de la SEC ID N°: 1 en el listado de secuencias; A2-15 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 5800 de la SEC ID N°: 1 en el listado de secuencias; A2-16 se corresponde con la secuencia de polinucleótido de los nucleótidos 701 a 7000 de la SEC ID N°: 1 en el listado de secuencias; y A2-17 se corresponde con la secuencia de polinucleótido de los nucleótidos 2001 a 8450 de la SEC ID N°: 1 en el listado de secuencias.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de A7, A7-1 se corresponde con la secuencia de polinucleótido de los nucleótidos 601 a 3600 de la SEC ID N°: 2 en el listado de secuencias; A7-2 se corresponde con la secuencia de polinucleótido de los nucleótidos 3601 a 8420 de la SEC ID N°: 2 en el listado de secuencias; A7-3 se corresponde con la secuencia de polinucleótido de los nucleótidos 5401 a 8420 de la SEC ID N°: 2 en el listado de secuencias; A7-4 se corresponde con la secuencia de polinucleótido de los nucleótidos 3401 a 6400 de la SEC ID N°: 2 en el listado de secuencias; A7-5 se corresponde con la secuencia de polinucleótido de los nucleótidos 1501 a 4500 de la SEC ID N°: 2 en el listado de secuencias; A7-6 se corresponde con la secuencia de polinucleótido de los nucleótidos 4401 a 7400 de la SEC ID N°: 2 en el listado de secuencias; A7-7 se corresponde con la secuencia de polinucleótido de los nucleótidos 2401 a 5400 de la SEC ID N°: 2 en el listado de secuencias; A7-8 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 3600 de la SEC ID N°: 2 en el listado de secuencias; A7-9 se corresponde con la secuencia de polinucleótido de los nucleótidos 1501 a 5400 de la SEC ID N°: 2 en el listado de secuencias; A7-10 se corresponde con la secuencia de polinucleótido de los nucleótidos 2401 a 6400 de la SEC ID N°: 2 en el listado de secuencias; A7-11 se corresponde con la secuencia de polinucleótido de los nucleótidos 3401 a 7400 de la SEC ID N°: 2 en el listado de secuencias; A7-12 se corresponde con la secuencia de polinucleótido de los nucleótidos 4401 a 8420 de la SEC ID N°: 2 en el listado de secuencias; A7-13 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 5400 de la SEC ID N°: 2 en el listado de secuencias; A7-14 se corresponde con la secuencia de polinucleótido de los nucleótidos 1501 a 6400 de la SEC ID N°: 2 en el listado de secuencias; A7-15 se corresponde con la secuencia de polinucleótido de los nucleótidos 2401 a 7400 de la SEC ID N°: 2 en el listado de secuencias; A7-16 se corresponde con la secuencia de polinucleótido de los nucleótidos 3401 a 8420 de la SEC ID N°: 2 en el listado de secuencias; A7-17 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 6400 de la SEC ID N°: 2 en el listado de secuencias; y A7-18 se corresponde con la secuencia de polinucleótido de los nucleótidos 1501 a 7400 de la SEC ID N°: 2 en el listado de secuencias.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de A18, A18-1 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 5040 de la SEC ID N°: 3 en el listado de secuencias; A18-2 se corresponde con la secuencia de polinucleótido de los nucleótidos 1001 a 6002 de la SEC ID N°: 3 en el listado de secuencias; A18-3 se corresponde con la secuencia de polinucleótido de los nucleótidos 2001 a 7000 de la SEC ID N°: 3 en el listado de secuencias; y A18-4 se corresponde con la secuencia de polinucleótido de los nucleótidos 3000 a 7000 de la SEC ID N°: 3 en el listado de secuencias.

Los puntos de inicio y terminación de los respectivos fragmentos de A2, A7 y A18 se exponen también en la Figura 18.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de B5, B5-1 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 4001 de la SEC ID N°: 4 en el listado de secuencias; B5-2 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 3200 de la SEC ID N°: 4 en el listado de secuencias; B5-3 se corresponde con la secuencia de polinucleótido de los nucleótidos 2491 a 5601 de la SEC ID N°: 4 en el listado de secuencias; B5-4 se corresponde con la secuencia de polinucleótido de los nucleótidos 5373 a 8401 de la SEC ID N°: 4 en el listado de secuencias; B5-5 se corresponde con la secuencia de polinucleótido de los nucleótidos 901 a 4001 de la SEC ID N°: 4 en el listado de secuencias; y B5-6 se corresponde con la secuencia de polinucleótido de los nucleótidos 4001 a 7000 de la SEC ID N°: 4 en el listado de secuencias.

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de C14, C14-1 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 4015 de la SEC ID N°: 5 en el listado de secuencias; C14-2 se corresponde con la secuencia de polinucleótido de los nucleótidos 1987 a 5014 de la SEC ID N°: 5 en el listado de secuencias; C14-3 se corresponde con la secuencia de polinucleótido de los nucleótidos 4020 a 7119 de la SEC ID N°: 5 en el listado de secuencias; C14-4 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-5 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 6011 de la SEC ID N°: 5 en el listado de secuencias; C14-6 se corresponde con la secuencia de polinucleótido de los nucleótidos 4939 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-7 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 5014 de la SEC ID N°: 5 en el listado de secuencias; C14-8 se corresponde con la secuencia de polinucleótido de los nucleótidos 2994 a 7119 de la SEC ID N°: 5 en el listado de secuencias; C14-9 se corresponde con la secuencia de polinucleótido de los nucleótidos 4020 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-10 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 5014 de la SEC ID N°: 5 en el listado de secuencias; C14-11 se corresponde con la secuencia de polinucleótido de los nucleótidos 1987 a 7119 de la SEC ID N°: 5 en el listado de secuencias; C14-12 se corresponde con la secuencia de polinucleótido de los nucleótidos 2994 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-13 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 7119 de la SEC ID N°: 5 en el listado de secuencias; y C14-14 se corresponde con la secuencia de polinucleótido de los nucleótidos 1987 a 8141 de la SEC ID N°: 5 en el listado de secuencias.

Los puntos de inicio y terminación de los respectivos fragmentos de B5 y C14 también se exponen en la Figura 19.

2. Adquisición del polinucleótido

En la invención, se puede obtener un polinucleótido que contiene un gen exógeno que codifica una proteína exógena cuya producción va a aumentar, lo que se describirá más adelante, por procedimientos comunes como se

describen a continuación. Por ejemplo, dicho polinucleótido puede aislarse explorando una biblioteca de ADNc procedente de células o tejidos que expresan el gen exógeno usando una sonda de ADN que se sintetiza basándose en un fragmento del gen exógeno. El ARNm se puede preparar por procedimientos que se usan comúnmente en este campo técnico. Por ejemplo, las células o tejidos se tratan con un reactivo de guanidina, un reactivo de fenol, etc., obteniendo de esta manera el ARN total, y a partir de este, se obtiene el poli(A) + ARN (ARNm) mediante un procedimiento en columna de afinidad usando una columna de celulosa oligo (dT) o una columna de Sepharose-poli U que contiene Sepharose 2B como portador, o similares, o mediante un procedimiento discontinuo. Además, también se puede fraccionar el poli(A) + ARN por centrifugación en gradiente de densidad de sacarosa o similares. Luego, se sintetiza un ADNc monocatenario usando el ARNm obtenido de esta manera como molde, y también usando cebadores dT y una transcriptasa inversa. A partir del ADNc de cadena sencilla que se obtiene de esta manera, se sintetiza un ADNc de doble cadena usando una ADN polimerasa I, ADN ligasa, RNasa H, y similares. El ADNc de doble cadena sintetizado de esta manera se trunca usando ADN polimerasa T4, seguido de la unión a un adaptador (tal como el adaptador EcoRI), fosforilación, y similares, y el ADN resultante se incorpora en un fago lambda tal como el λ gt11 para conseguir un empaquetamiento in vivo, de esta manera se puede preparar una biblioteca de ADNc. También es posible usar una biblioteca de ADNc usando un vector plásmido distinto de fagos lambda. A partir de entonces, se puede seleccionar un clon que contenga el ADN diana (un clon positivo) a partir de la biblioteca de ADNc.

En los casos en los que el elemento de ADN mencionado anteriormente que se va a usar para aumentar la producción de una proteína o un polinucleótido que contiene un gen exógeno se aísla a partir de ADN genómico, o en el que se aísla un polinucleótido que contiene las regiones promotoras y de terminación a partir de ADN genómico, de acuerdo con un procedimiento común (Molecular Cloning (1989), Methods in Enzymology 194 (1991)), se extrae el ADN genómico de una línea celular de un organismo para usarse como fuente de recolección y se selecciona y aísla un polinucleótido. La extracción del ADN genómico se puede llevar a cabo 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 elemento de ADN diana o el polinucleótido que contiene un gen exógeno se puede obtener también, por ejemplo, por el procedimiento PCR (PCR Technology. Henry A. Erlich, Atocckton press (1989)). En la amplificación de un polinucleótido usando el procedimiento PCR, se usan 20 o 30meros de ADN monocatenarios sintéticos como cebadores y se usa el ADN genómico como molde. El gen amplificado se usa después de confirmarse la secuencia de polinucleótido del gen. Como molde para la PCR, se puede usar una biblioteca de ADN genómico tal como un cromosoma bacteriano artificial (BAC).

Por otra parte, el polinucleótido que contiene un gen exógeno cuya secuencia no se conoce se puede obtener (a) preparando una biblioteca génica de acuerdo con un procedimiento común, y (b) seleccionando un polinucleótido deseado de la biblioteca génica que se ha preparado y amplificando el polinucleótido. La biblioteca génica se puede preparar por digestión parcial del ADN cromosómico que se obtiene mediante un procedimiento común a partir de una línea celular de un organismo que se va a usar como fuente de recolección usando una enzima de restricción adecuada para el fragmento de ADN cromosómico, uniendo los fragmentos obtenidos en un vector adecuado, e introduciendo el vector en un hospedador adecuado. La biblioteca génica se puede preparar también por extracción del ARNm de las células, sintetizando el ADNc a partir de ARNm, uniendo el ADNc a un vector adecuado, e introduciendo el vector en un hospedador adecuado. Como vector para su uso en dicha preparación, también se puede usar un plásmido que se sepa en general que es un vector para la preparación de la biblioteca génica, un vector de fago, un cósmido o similares. Como hospedador que se va a transformar o transfectar se puede usar un hospedador adecuado para el tipo de vector mencionado anteriormente. El polinucleótido que contiene el gen exógeno se selecciona de la biblioteca génica mencionada anteriormente mediante un procedimiento de hibridación de colonias, un procedimiento de hibridación en placas, o similares, usando una sonda marcada que contiene una secuencia específica para el gen exógeno.

Además, el polinucleótido que contiene el gen exógeno también se puede producir mediante síntesis química total. Por ejemplo, el gen se puede sintetizar mediante un procedimiento en el que se preparan y se hibridan dos pares de oligonucleótidos complementarios, un procedimiento en el que se unen varias hebras de ADN hibridadas por una ADN ligasa, un procedimiento en el que se preparan varios polinucleótidos parcialmente complementarios y los huecos se rellenan por PCR o similares.

La determinación de una secuencia de polinucleótido se puede llevar a cabo por una técnica convencional, por ejemplo, un procedimiento didesoxi (Sanger y col., Proc. Natl. Acad. Sci., USA, 74, 5463-5467, (1977)), o similares. Además, la determinación anterior de una secuencia de polinucleótido también se puede llevar a cabo fácilmente usando un kit de secuenciación disponible en el comercio o similar.

3. Vector de expresión de un gen exógeno, vector de elemento

Como vector de expresión de un gen exógeno de la invención, se proporciona un vector que contiene un elemento de ADN del tipo A7 mencionado anteriormente o dos o más copias de un elemento de ADN del tipo A7 mencionado anteriormente en combinación, o al menos uno de elemento de ADN del tipo A7 mencionados anteriormente en combinación con al menos uno de los diferentes tipos de elementos de ADN y que contienen además una unidad de

expresión de un gen exógeno. Cuando se expresa un gen exógeno en una célula hospedadora usando el vector de expresión de un gen exógeno mencionado anteriormente, el elemento de ADN se puede situar inmediatamente cadena arriba o cadena abajo de la unidad de expresión génica, o se puede localizar en una posición lejana de la unidad de expresión génica. Además, se puede usar un vector de expresión de un gen exógeno que contenga una diversidad de dichos elementos de ADN. De hecho, el elemento de ADN se puede insertar tanto en orientación directa como inversa con respecto a la unidad de expresión génica.

Además, como un vector que se usa en la invención, también se incluye un vector que contiene un tipo del elemento de ADN A7 mencionado anteriormente, dos o más copias de un tipo del elemento de ADN A7 mencionado anteriormente o dos o más diferentes tipos del elemento de ADN A7 mencionado anteriormente en combinación, o al menos uno de los tipos de elemento de ADN A7 en combinación con al menos uno de los diferentes tipos de elementos de ADN diferentes y que no contenga una unidad de expresión génica (de aquí en adelante denominado también "vector de elemento"). Dicho vector de elemento se puede usar en combinación con el vector de expresión de un gen exógeno mencionado anteriormente que contiene el elemento de ADN o un vector de expresión de un gen exógeno que no contiene un elemento de ADN y solamente contiene la unidad de expresión de un gen exógeno. Permitiendo la coexistencia del vector de elemento con el vector de expresión de un gen exógeno, la expresión de un gen exógeno aumenta en comparación con los casos en los que se usa solo el vector de expresión de un gen exógeno, por lo tanto, la combinación de los vectores mencionados anteriormente también se incluyen en el vector de expresión de un gen exógeno de la invención.

El gen que codifica la proteína exógena no está limitado particularmente, sin embargo, los ejemplos del mismo incluyen los genes indicadores tales como el de la fosfatasa alcalina secretora (SEAP), la proteína verde fluorescente (GFP), y luciferasa; varios genes de enzimas tales como el gen de la α -amilasa, y el gen de la α -galactosidasa; genes de varios interferones que son útiles farmacéuticamente y proteínas fisiológicamente activas tales como el interferón α y el interferón γ ; genes de varias interleucinas tales como IL-1 e IL-2; varios genes de citocinas tales como un gen de eritropoyetina (EPO) y un gen de factor estimulante de colonias de granulocitos (G-CSF); genes de factores de crecimiento; y genes de anticuerpos. Estos genes se pueden obtener mediante cualquier procedimiento.

La invención es particularmente eficaz en relación con una proteína que sea altamente hidrófoba y una proteína que sea difícil de producir y secretar debido a su formación compuesta. Por lo tanto, las proteínas exógenas mencionadas anteriormente incluyen una proteína multimérica tal como un heteromultímero que es un anticuerpo o un fragmento funcional del mismo. El "fragmento funcional de un anticuerpo" se refiere a un fragmento parcial de un anticuerpo que tiene actividad de unión a antígeno y que incluye Fab, F(ab')₂, Fv, scFv, diacuerpos, anticuerpos lineales, anticuerpos poliespecíficos formados a partir de fragmentos de anticuerpo, y similares. El fragmento funcional de un anticuerpo también incluyen al Fab' que es un fragmento monovalente de una región variable de un anticuerpo que se obtiene tratando el F(ab')₂ en condiciones reductoras. Además, estos fragmentos funcionales incluyen no solamente un fragmento obtenido mediante el tratamiento de una molécula de una proteína de anticuerpo de longitud completa con una enzima adecuada, sino también una proteína producida en una célula hospedadora adecuada usando un gen de anticuerpo modificado genéticamente.

La unidad de expresión génica tiene, en la dirección de la fase de lectura de transcripción, al menos una región promotora, un gen exógeno, y una región de terminación de la transcripción (señal de adición de poli(A)). El promotor que se puede usar en el presente documento puede ser un promotor de expresión constitutivo o un promotor de expresión inducible. Los ejemplos de promotor de expresión constitutivo incluyen varios promotores naturales tales como un promotor temprano de SV40, un promotor de adenovirus E1A, un promotor de CMV (citomegalovirus), un promotor EF-1 α (factor-1 α de elongación humano), un promotor de HSP70, un promotor MT, un promotor de RSV, un promotor de UBC, y un promotor actina, y promotores artificiales (de fusión) tales como un promotor SR α y un promotor CAG. Además, la secuencia de adición de poli(A) puede ser una secuencia que tiene la actividad de provocar la terminación de la transcripción para la transcripción a partir del promotor, y puede ser una secuencia de un gen igual o diferente al del promotor.

Es necesario usar un promotor fuerte con el fin de aumentar la producción de una proteína exógena. Sin embargo, cuando se intenta producir una proteína que sea difícil que se pliegue o una proteína que sea difícil de secretar usando un promotor altamente activo, la proteína puede no llegar a secretarse. Esto es debido a que cuando la proteína se produce en una cantidad que excede la capacidad del ribosoma en el que se lleva a cabo la traducción y la del retículo endoplásmico donde se lleva a cabo el plegamiento y la secreción, la proteína producida en exceso se desnaturaliza, se acumula y se ubiquitina en las células, y luego se degrada en los proteosomas. Por consiguiente, es preferible que se seleccione adecuadamente un promotor que pueda alcanzar un nivel de expresión tal que la proteína no se desnaturalice o agregue o que la cantidad de la proteína resultante no exceda la capacidad de secreción. Como alternativa, el promotor se usa ajustando (por ejemplo, disminuyendo) la actividad del promotor. Entre las proteínas multiméricas, una molécula que forma un heteromultímero es susceptible del efecto descrito anteriormente y en particular una molécula, tal como un anticuerpo, que es un heterotetrámero. Un anticuerpo tiene dos moléculas de cadena pesada y dos moléculas de cadena ligera que se asocian entre sí y por lo tanto, con el fin de asociar adecuadamente las moléculas, el nivel de expresión de las mismas es un factor importante.

Además, el vector de expresión de un gen exógeno y el vector de elemento de la invención pueden contener cada uno un marcador de selección para seleccionar un transformante. Se puede seleccionar un transformante, por ejemplo, mediante el uso de un marcador de resistencia a un fármaco que dé lugar a resistencia a un fármaco tal como cerulenina, aureobasidina, zeocina, canavanina, cicloheximida, higromicina, puromicina, blasticidina, tetraciclina, kanamicina, ampicilina o neomicina. Además, también se puede seleccionar un transformante cuando se usa como marcador un gen que da lugar a resistencia a disolventes tales como etanol, resistencia a la presión osmótica de glicerol, de una sal, o similares, la resistencia a un ion metálico tal como el ion cobre, o similares.

El vector de expresión de un gen exógeno 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 un gen exógeno se transfecta en una célula hospedadora y entonces se incorpora aleatoriamente en el cromosoma. Sin embargo, usando un componente constituyente procedente de un virus de mamífero tal como un virus de simio 40 (SV40), un papilomavirus (BPV, HPV), o EBV, se puede usar el vector como un vector episómico que es auto-replicable en la célula hospedadora transfectada. Por ejemplo, se puede usar un vector que contiene un origen de replicación (oriP) procedente de SV40 y una secuencia que codifica un gran antígeno T de SV40 que es un factor trans-actuante, un vector que contiene un oriP procedente de EBV y una secuencia que codifica EBNA-1, o similares. El efecto del elemento de ADN se puede expresar por la actividad de aumento de la expresión del gen exógeno independientemente del tipo de vector o de la presencia o ausencia de la incorporación del mismo en el cromosoma.

4. Célula transformada

La célula transformada de la invención es una célula transformada en la que se ha introducido el vector de expresión de un gen exógeno descrito en el artículo "3" anterior que contiene el elemento de ADN descrito en el artículo "1". Como vector de expresión de un gen exógeno, se puede introducir solamente un vector de expresión de un gen exógeno que contenga un elemento de ADN (A), o se pueden introducir en combinación un vector de expresión de un gen exógeno que contiene un elemento de ADN y también un vector de elemento descrito en el artículo "3" anterior (B). Como alternativa, se pueden introducir en combinación un vector de expresión de un gen exógeno que no contiene un elemento de ADN y un vector elemento (C).

La expresión de un gen exógeno en una célula hospedadora usando la combinación anterior de (B) o (C) se puede llevar a cabo, por ejemplo, según 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 hospedadora que se va a transformar incluyen células eucariotas, incluyendo los ejemplos preferidos de las mismas células de mamífero, incluyendo los ejemplos más preferidos células procedentes de seres humanos, ratones, ratas, hámsteres, monos, o ganado. Los ejemplos de dichas células de mamífero incluyen las células COS-1, células 293, y células CHO (CHO-K1, DG44, CHO dhfr-, CHO-S), sin embargo, la célula hospedadora no se limita a estas.

En la invención, se puede usar cualquier procedimiento para introducir el vector de expresión en la célula hospedadora siempre que el procedimiento permita que gen introducido esté presente de manera estable en la célula hospedadora y se exprese en la misma de manera adecuada. Los ejemplos del método que se usa generalmente incluyen un procedimiento de fosfato cálcico (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 (Itoh, H. (1983) *J. Bacteriol.* 153, 163-168), y un procedimiento de lipofección.

5. Procedimiento para producir proteína exógena

En la invención, se puede producir una proteína exógena mediante el cultivo de la célula transformada descrita en el artículo "4" anterior, en la que se ha introducido un gen que codifica la proteína exógena usando el vector descrito en el artículo "3" anterior mediante un procedimiento conocido, la recogida de la proteína del producto del cultivo resultante, seguido de la purificación de la proteína. La expresión "producto del cultivo" tal como se usa en el presente documento se refiere a células cultivadas o a un homogenado celular además de a un sobrenadante del cultivo. De hecho, puede seleccionarse como proteína exógena que puede producirse usando la célula transformada descrita en el artículo "4" anterior, no solo una proteína monomérica, sino también una proteína multimérica. En los casos en los que se produce una proteína hetero-multimérica formada de una pluralidad de diferentes subunidades, es necesario introducir una pluralidad de genes que codifican estas subunidades en la célula hospedadora que se describe en el artículo "4" anterior, respectivamente.

El procedimiento para el cultivo de la célula transformada se puede llevar a cabo de acuerdo con procedimientos convencionales para cultivar células hospedadoras.

En los casos en que la célula transformada sea una célula de mamífero, la célula se cultiva en condiciones, por ejemplo, de 37 °C, y un 5 % o un 8 % de CO₂ durante un tiempo de cultivo de aproximadamente 24 a 1000 horas. El cultivo se puede llevar a cabo por medio de cultivo discontinuo, cultivo de alimentación discontinua, cultivo continuo y similares en condiciones estáticas, de agitación, removido o aireación.

La confirmación del producto de expresión del gen que codifica la proteína exógena a partir del producto de cultivo (solución de cultivo) mencionado anteriormente se puede llevar a cabo por SDS-PAGE, análisis de Western, ELISA o similares. Con el fin de aislar y purificar la proteína producida, se puede usar un procedimiento de purificación y aislamiento de proteínas convencional. Tras completar el cultivo, en los casos en los que la proteína diana se produce en células, se homogenizan las células usando un homogeneizador ultrasónico, una prensa de French, un homogeneizador de Manton-Gaulin, Dinomil, o similar, obteniendo de esta manera la proteína diana. Además, en los casos en los que se produce la proteína diana fuera de las células, la solución el cultivo se usa como tal, o se eliminan las células por centrifugación o similar. A continuación se recoge la proteína diana por extracción o similar usando un disolvente orgánico y luego se aísla y purifica la proteína diana usando técnicas tales como varias técnicas de cromatografía (cromatografía hidrófoba, cromatografía de fase inversa, cromatografía de afinidad, cromatografía de intercambio iónico, etc.), filtración en gel usando un tamiz molecular y electroforesis usando un gel de poliacrilamida o similar solo o en combinación según las necesidades.

Los procedimientos de cultivo mencionados anteriormente y los procedimientos de purificación son solo a modo de ejemplo, y los procedimientos no se limitan a estos. La secuencia de aminoácidos del producto genético purificado se puede confirmar por una técnica de análisis de aminoácidos conocida, tal como la determinación de secuencia de aminoácidos automatizada usando el procedimiento de degradación de Edman.

6. Procedimiento para producir una proteína de anticuerpo

Como proteína hetero-multimérica que se va a producir usando el procedimiento de producción descrito en el artículo "5" anterior, puede ejemplificarse una proteína de anticuerpo. La proteína de anticuerpo es una proteína tetramérica que comprende dos moléculas de polipéptido de cadena pesada y dos moléculas de polipéptido de cadena ligera. Por consiguiente, para obtener dicha proteína de anticuerpo en un estado que mantenga la afinidad de unión a antígeno, es necesario introducir ambos genes de cadena pesada y ligera en la célula transformada descrita en el artículo "4" anterior. En este caso, las unidades de expresión génica de la cadena pesada y ligera pueden estar presentes en el mismo vector de expresión o en diferentes vectores de expresión.

Como anticuerpo que se va a producir en la invención, puede ejemplificarse un anticuerpo preparado mediante la inmunización de un animal experimental, tal como un conejo, un ratón o una rata con un antígeno deseado. Además, también puede ejemplificarse como el anticuerpo que se va a producir en la invención un anticuerpo quimérico y un anticuerpo humanizado obtenido mediante el uso del anticuerpo anteriormente mencionado como material de partida. Además, también se incluye como anticuerpo que se va a producir en la invención un anticuerpo humano obtenido usando un animal modificado genéticamente o un procedimiento de presentación en fagos.

El gen de anticuerpo que se va a usar para la producción del anticuerpo no está limitado a un gen de anticuerpo que tenga una secuencia de polinucleótido específica, siempre que la combinación de polipéptido de cadena pesada y polipéptido de cadena ligera que se van a transcribir y traducir a partir del gen de anticuerpo tenga la actividad de unión a una proteína antigénica dada.

Además, no es necesario que el gen de anticuerpo codifique la molécula de longitud completa del anticuerpo, puede usarse un gen que codifica un fragmento funcional del anticuerpo. Dicho gen que codifica un fragmento funcional del mismo puede obtenerse modificando genéticamente un gen que codifica la molécula de longitud completa de la proteína de anticuerpo.

7. Procedimiento de producción de otras proteínas exógenas

Los ejemplos de proteínas exógenas que se van a producir usando el procedimiento de producción de la invención incluyen, además de los anticuerpos mencionados anteriormente, varias proteínas procedentes de seres humanos y 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 el péptido natriurético auricular (ANP), péptido natriurético cerebral (BNP), péptido natriurético tipo C (CNP), vasopresina, somatostatina, hormona de crecimiento (GH), insulina, oxitocina, grelina, leptina, adiponectina, renina, calcitonina, osteoprotegerina, y factor de crecimiento insulínico (IGF); citocinas tales como interleucinas, quimiocinas, interferón, factores de necrosis tumoral (tales como TNF- α , TNF- β , y la superfamilia de TNF), factores de crecimiento nerviosos (tales como NGF), factores de crecimiento celular (tal como EGF, FGF, PDGF, HGF, y TGF), factores de crecimiento hematopoyético (tales como CSF, G-CSF, y eritropoyetina), y adipocina; receptores tales como los receptores de TNF; enzimas tales como lisozima, proteasas, proteinasas, y peptidasas; fragmentos funcionales de las mismas (fragmentos que tienen la totalidad o parte de 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 estas.

Ejemplos

En lo sucesivo, la invención se describirá específicamente en referencia a los ejemplos. Sin embargo, estos ejemplos no limitan el ámbito técnico de la invención. Los plásmidos, enzimas de restricción, enzimas de modificación de ADN, y similares que se van a usar en los ejemplos de la invención son productos disponibles en el mercado y se pueden usar según procedimientos comunes. Además, los procedimientos que se usan para la clonación de ADN, la determinación de la secuencia de polinucleótido, la transformación de una célula hospedadora,

el cultivo de una célula hospedadora transformada, la recogida de una proteína a partir del producto de cultivo resultante, la purificación de una proteína y similares también son bien conocidos por los expertos en la técnica o se pueden encontrar en la bibliografía.

Ejemplo 1

5 Extracción del elemento de ADN

(1-1) Inmunoprecipitación de cromatina usando un anticuerpo anti-histona H3 acetilada

Se llevó a cabo una ChIP usando un anticuerpo anti-histona H3 acetilada usando EZ ChIP (Upstate) de acuerdo con el siguiente procedimiento. De hecho, a menos que se indique lo contrario, se usaron productos de Upstate como anticuerpos, tampones y similares en el siguiente procedimiento.

10 En primer lugar, se cultivaron células 293F (Invitrogen) usando Medio 293 FreeStyle™ GIBCO (marca registrada) (Invitrogen) en condiciones de 37 °C y un 8 % de CO₂, seguido de centrifugación (1000 rpm, 5 min, a temperatura ambiente), por la que se recogieron las células en fase de crecimiento. Después se fijaron 2 x 10⁷ células en un medio que contenía un 1 % de formaldehído durante 10 minutos, se añadió al mismo 10 x de glicina, seguido de incubación a temperatura ambiente durante 5 minutos. Tras la centrifugación (3000 rpm, 5 min, 4 °C), se retiró el sobrenadante, y se añadió PBS al sedimento celular para suspender las células. Entonces, la suspensión celular se centrifugó de nuevo para retirar el PBS, y a continuación se añadió un tampón de lisis SDS al sedimento celular para suspender y lisar las células. Cada muestra que se obtenía por lisis celular se sometió a fragmentación de ADN usando un homogeneizador ultrasónico (BRANSON) mientras se enfriaba la muestra con agua helada, y se añadió a la misma un tampón de dilución que contenía un cóctel inhibidor de proteasas y proteína G inmovilizada en agarosa.

15 La mezcla resultante se sometió a rotación a 4 °C durante 1 hora, seguido de centrifugación, y luego se recolectó el sobrenadante. Posteriormente, se añadieron al mismo 10 µg de IgG normal de conejo o de un anticuerpo α-acetil histona H3, seguido de rotación durante una noche a 4 °C. A la solución resultante, se añadió proteína G inmovilizada en agarosa y la mezcla resultante se sometió a rotación a 4 °C durante 1 hora, seguido de centrifugación, y luego se recolectó el sedimento. El sedimento obtenido de esta manera se lavó dos veces en tampón de lavado de complejo inmunitario bajo en sal, dos veces con tampón de lavado de complejo inmunitario alto en sal, dos veces con tampón de lavado de complejo inmunitario LiCl, y finalmente cuatro veces con tampón TE. Luego se añadió al mismo un tampón de elución (que contenía 20 µl de hidrógeno carbonato sódico 1M, 10 µl de SDS, y 170 µl de agua estéril). Tras 30 minutos, se centrifugó la mezcla, y se recogió el sobrenadante.

30 Posteriormente, se añadieron 5 M de cloruro sódico al sobrenadante y la mezcla resultante se calentó durante una noche a 65 °C. Después se añadió a la misma RNasa A, y la mezcla resultante se incubó a 37 °C durante 30 minutos. Luego se añadió a la misma 0,5 M de EDTA, 1 M de Tris-HCl y Proteinasa K, y la mezcla resultante se incubó a 45 °C durante 2 horas.

Finalmente se añadieron a la misma los Reactivos A, B y C en una cantidad 5 veces mayor que la de la solución obtenida por el tratamiento con Proteinasa K, seguido de centrifugación (10000 rpm, 30 seg, a temperatura ambiente) usando un filtro Spin. Por el que se purificó el ADN de la inmunoprecipitación de cromatina.

35 (1-2) Análisis de micromatriz

Mediante el uso del kit GenomePlex Complete Whole Genome Amplification (WGA, amplificación de genoma completo) (Sigma), se amplificó cada muestra de ChIP obtenida en (1-1). El procedimiento fue de acuerdo con el protocolo de Sigma adjunto en el kit.

40 Con el fin de confirmar la ChIP, usando 320 ng de cada ADN amplificado por WGA como molde, y usando también los siguientes cebadores y SYBR (marca registrada) Premix Ex Taq™ (Tiempo real perfecto) (TAKARA), se amplificó un gen interno de gliceraldehído-3-fosfato deshidrogenasa (GAPDH) por el procedimiento de PCR (95 °C durante 5 s y 60 °C durante 20 s x 45 ciclos). De hecho, el GAPDH es un gen constitutivo que se va a usar como control positivo para confirmar si se había enriquecido el elemento de ADN mediante ChIP y se llevó a cabo el procedimiento PCR usando cebadores unidos al EZ ChIP (Upstate).

45 5'-TACTAGCGGTTTTACGGGCG-3'
5'-TCGAACAGGAGGAGCAGAGAGCGA-3'

50 Tal como se muestra en la Fig. 1, se confirmó que GAPDH estaba amplificado específicamente en la muestra sometida a inmunoprecipitación con un anticuerpo anti-histona H3 acetilada. Cada una de las muestras de ADN amplificado por WGA se sometió a análisis de micromatriz (NimbleGen) para llevar a cabo la inmunoprecipitación de cromatina sobre chip (ChIP-on-chip). "ChIP-on-chip" es una técnica para identificar cada elemento de ADN sometiendo cada ADN enriquecido en (1-1) a análisis de micromatriz.

(1-3) Extracción del elemento de ADN

55 Basándose en los resultados del análisis ChIP-on-chip que se obtienen en (1-2), se extrajeron 5 secuencias que tienen un contenido en AT del 62 % o más.

A2: cromosoma 15 (80966429 to 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)

5

Ejemplo 2

Efecto del elemento de ADN usando la expresión de fosfatasa alcalina secretora (SEAP) como índice

(2-1) Construcción de un vector de expresión de SEAP

10 Utilizando el pSEAP2-control (Clontech) como molde, se amplificó el gen SEAP mediante el procedimiento PCR (94 °C durante 30 s y 68 °C durante 2 min x 40 ciclos) usando los siguientes cebadores y KOD-plus- (TOYOBO).
 5'-AAAGCTAGCATGCTGCTGCTGCTGCTGCTGCTGGGCC-3'
 5'-AAAAGATCTTCATGTCTGCTCGAAGCGGCCGCGCCG-3'

15 Posteriormente, el fragmento de SEAP amplificado se separó por electroforesis en gel de agarosa y se recortó del gel, seguido de purificación usando un kit de extracción de Gel QIAquick (Qiagen). El fragmento de ADN obtenido de esta manera se usó como inserción. La inserción se digirió con las enzimas de restricción NheI y BgIII, y se digirió el vector pIRES 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 los fragmentos diana se recortaron del gel, seguido de purificación. Luego, se llevó a cabo una reacción de ligadura y transformación. La reacción de ligadura se llevó a cabo usando el sistema de ligadura de ADN LigaFast Rapid (Promega). La transformación se llevó a cabo de la siguiente manera. En primer lugar, se descongelaron células JM109 competentes (TAKARA) congeladas, y se añadieron 10 µl de una solución obtenida tras 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 (42 °C, 45 seg) a la mezcla, y se enfrió la mezcla 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 hora. Posteriormente se emplacó la mezcla sobre una placa LB que contenía 0,1 mg/ml de ampicilina, y se incubó la placa a 37 °C durante 14 a 16 horas. A continuación, por lisis alcalina, se recolectó un plásmido diana de las colonias cultivadas en la placa LB. Finalmente, se determinó la secuencia de polinucleótido de SEAP en el plásmido que se obtuvo por lisis alcalina, de esta manera se construyó el pCMV/SEAP ires Hygro.

15
20
25

(2-2) Clonación del elemento de ADN

30 Posteriormente, se clonó cada uno de los elementos de ADN extraídos en el ejemplo 1 en el vector de expresión de SEAP que se obtuvo en (2-1) usando el kit BAC SUBCLONING (Gene Bridges) a partir de un cromosoma bacteriano artificial (BAC) que contenía una secuencia de polinucleótido que se correspondía con cada uno de los elementos de ADN.

35 En primer lugar, el pCMV/SEAP ires Hygro obtenido en (2-1) se digirió con la enzima de restricción SpeI durante varias horas, seguido de precipitación en etanol y el precipitado se disolvió en agua estéril. Utilizando el vector digerido con SpeI como molde, se llevó a cabo el procedimiento PCR (94 °C durante 15 s, 55 °C durante 30 s y 68 °C durante 10 min x 30 ciclos) usando los siguientes cebadores y KOD-plus- (TOYOBO).

A2D:
 5'-GGAAATTGAGAAGTATCATTACACAACAGTACCACAAACATGAAATAAATGTGGAT
 CCTATTAATAGTAATCAATTACG-3'

40

A2R:
 5'-CTCATTCTGTGGGTTGTCATTTCACTTCCTTGATGCTATCCTTTCAAGCAAAATC
 CTAGTCAATAATCAATGTCAACG-3'

A7D:
 5'-CTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAAACCTTGGATC
 CTATTAATAGTAATCAATTACG-3'

45

A7R:
 5'-CTCTTCCCATTCTCATTTGAATCTACTTCAAAGGTTTACCATACTAAGACCTAG
 TCAATAATCAATGTCAACG-3'

50

A18D:

5'-CGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGTGGATCACCTGAGGTCGA
TCCTATTAATAGTAATCAATTACG-3'

A18R:

5 5'-CATACAGAAGCCAGTTTGAAGTGAAGACCTCACTCCATTTCTTACAAGTTATGCC
TAGTCAATAATCAATGTCAACG-3'

B5D:

5'-ACCGTTTTATATTGTTTAAGCATTTCCTAGACATATTTGGCTACAAATCTAGATC
CTATTAATAGTAATCAATTACG-3'

10 B5R:

5'-GATCTTAGGGGGGCTGATTATATAAAACAATAGAAATGTAGTCTTAGATGAAACC
TAGTCAATAATCAATGTCAACG-3'

C14D:

15 5'-CACAAAGTTCAGTGTCAAGGCCAGGTGATGAGGCCACACATGCCCGGACCTTGA
TCCTATTAATAGTAATCAATTACG-3'

C14R:

5'-CAAAACCTCATCTCTACTGAAAATAGAAAATTAGCTGGGCGTGGTGGCAGGTGCC
CTAGTCAATAATCAATGTCAACG-3'

20 Tras confirmarse la amplificación por electroforesis en gel de agarosa usando una porción de la solución de reacción, el resto de la solución de reacción se sometió a precipitación en etanol. El precipitado se disolvió en agua estéril, y la solución resultante se usó para la transformación de ADN.

Posteriormente, se llevó a cabo la preparación de *Escherichia coli* para la transformación.

Los clones de BAC que correspondían a las 5 secuencias extraídas en el ejemplo 1 son los siguientes.

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

25 10 µl del BAC mencionado anteriormente (Avanced Geno Techs Co.) que se habían descongelado, se inocularon en 1 ml de un medio (que contenía cloranfenicol a una concentración final de 15 µg/ml) y se incubaron durante una 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 incubaron a 37 °C durante 2 horas. Se repitió dos veces una centrifugación y lavado con agua estéril, y se suspendieron las células en 20 µl de agua estéril. Se añadieron en una cubeta enfriada (0,1 cm) 1 µl de pRED/ET (Gene Bridges) y *Escherichia coli*, seguido de electroporación (1350 V, 10 µF). Después, se añadió a esto 1 ml de medio SOC, y la mezcla resultante se incubó a 30 °C durante 70 minutos. Se emplacaron 100 µl de la solución de cultivo en una placa LB (que contenía tetraciclina y cloranfenicol a concentraciones finales de 3 µg/ml y 15 µg/ml, respectivamente), y se incubó una 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. Luego, se añadieron a la misma 50 µl de L-arabinosa al 10 %, y se llevó a cabo una incubación adicional a 37 °C durante 1 hora. A continuación se repitió dos veces un lavado con agua estéril, y se añadieron las *Escherichia coli* que se suspendieron en 30 µl de agua estéril y 1 µl del ADN para la transformación a una cubeta enfriada (0,1 cm), seguido de electroporación 1350 V, 10 µF). Luego, se añadió a esto 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 emplacó en una placa LB (que contenía 100 µg/ml de ampicilina), y la placa se incubó. A continuación se obtuvo el plásmido diana por lisis alcalina. Finalmente, se confirmaron la secuencia del plásmido que se obtuvo y los sitios de la enzima de restricción del mismo, por lo que se construyó el

plásmido diana. El vector construido se muestra en la Fig. 2.

(2-3) Evaluación usando la expresión de SEAP como índice

Se evaluó cada plásmido construido en (2-2) usando la célula hospedadora CHO-K1 (ATCC) y el reactivo de transfección Lipofectamina 2000 (Invitrogen)

- 5 Se llevó a cabo la selección con antibióticos con higromicina a 800 µg/ml durante aproximadamente 2 semanas comenzando 2 días tras la transfección, por lo que se estableció una línea celular policlonal que expresaba establemente. La línea celular establecida de esta manera se sometió a un replazo de medio el día antes de la medición, y se sembraron un número determinado de células en una placa de 24 pocillos (IWAKI). A las 24 horas tras poner las células en las placas, se recolectó el sobrenadante, y se midió la actividad de SEAP. La actividad de SEAP en el sobrenadante del cultivo se midió usando el Ensayo de Fosfatasa Alcalina Segregada indicadora SensoLyte™ pNPP (ANASPEC).

10 Los resultados medidos se muestran en la Fig. 3. Cuando la actividad de SEAP del control sin elemento se normalizó a 1, la actividad SEAP en el sobrenadante del cultivo de la línea celular CHO que expresaba establemente que tenía el elemento de ADN A2, A7, A18, B5 o C14 mostraba un valor numérico cinco veces mayor o más que el del control. Basándose en los resultados, se confirmó que los 5 tipos de elementos de ADN aumentan drásticamente la expresión de SEAP. De hecho, las secuencias de polinucleótidos de los 5 tipos de elementos de ADN anteriores se representan en las SEC ID N°: 1 a 5 del listado de secuencias, respectivamente.

Ejemplo 3

Generalidad de promotores que se va a usar en combinación

- 20 El promotor para el vector que se usa en la evaluación de los elementos de ADN del ejemplo 2 era un promotor de CMV, y de esta manera se estudió en el ejemplo 3 el uso de elementos de ADN en combinación con otros promotores generales.

(3-1) Construcción del vector de expresión de SEAP usando promotores EF-1α y SV40

- 25 Utilizando el pSEAP2-control (Clontech) como molde, se amplificó el gen SEAP por el procedimiento PCR (94 °C durante 30 s y 68 °C durante 2 min x 40 ciclos) usando los cebadores descritos en (2-1) y KOD-plus-. El SEAP amplificado se preparó como una inserción de la misma manera que en (2-1). La inserción se digirió con las enzimas de restricción NheI y BglII, y se digirió un vector pIRES puro3 (Clontech) con las enzimas de restricción NheI y BamHI, y se construyó el pCMV/SEAP ires Puro de la misma manera que en (2-1).

- 30 Posteriormente, usando el PEF1/V5-His A (Invitrogen) como molde, se amplificó un promotor EF-1α por el procedimiento PCR (94 °C durante 15 seg, 60 °C durante 30 seg, y 68 1C durante 2 min x 30 ciclos) usando los siguientes cebadores y KOD-plus-.

5'-AAAAGTATGTCAGAGAGGAATCTTTGTCAGCTAATGGACC-3'
5'-AAAGATATCCCTAGCCAGCTTGGGTGGTACCAAGC-3'

- 35 Utilizando el pCMV/SEAP ires Puro construido anteriormente como vector, se llevó a cabo la digestión con las enzimas de restricción SpeI y EcoRV para el vector y el promotor, y se construyó el pEF/SEAP ires Puro de acuerdo con el procedimiento descrito en (2-1).

- De manera similar, usando el pcDNA3.1+ (Invitrogen) como molde, se amplificó un promotor de SV40 por el procedimiento PCR (94 °C durante 15 s, 60 °C durante 30 s, y 68 1C durante 1 min x 30 ciclos) usando los siguientes cebadores y KOD-plus.

- 40 5'-AAAAGTATGTCAGAGGAATCTTTGTCAGCTAATGGACC-3'
5'-AAAGATATCAGCTTTTTGCAAAAGCCTAGGCCTC-3'

Utilizando el pCMV/SEAP ires Puro construido anteriormente como vector, se llevó a cabo la digestión con las enzimas de restricción SpeI y EcoRV para el vector y el promotor, y se construyó el pSV40/SEAP ires Puro de acuerdo con el procedimiento descrito en (2-1).

- 45 (3-2) Clonación del elemento de ADN A2 o A7

Posteriormente, se llevó a cabo la clonación del elemento de ADN A2 o A7 usando el pEF/SEAP ires Puro y pSV40/SEAP ires Puro construidos en (3-1) como estructuras básicas.

- 50 Primero se digirieron el pEF/SEAP ires Puro y el pSV40/SEAP ires Puro con la enzima de restricción SpeI durante varias horas, seguido de precipitación en etanol, y el precipitado se disolvió en agua estéril. Utilizando los respectivos vectores digeridos con SpeI como moldes, se preparó el ADN para la transformación por el procedimiento PCR (94 °C durante 15 s, 55 °C durante 30 s, y 68 °C durante 10 min x 30 ciclos) usando los siguientes cebadores y KOD-plus.

A2 (EF/D):

5'-GGAAATTGAGAAGTATCATTACACAACAGTACCACAAACATGAAATAAATGTGCTA
GTCAGAGAGGAATCTTTGCAGC-3'

A2 (SV40/D):

5 5'-GGAAATTGAGAAGTATCATTACACAACAGTACCACAAACATGAAATAAATGTGCTA
GTCTGTGGAATGTGTGTCAGTTAG-3'

A2 (EF y SV40/R) :

5'-CTCATTCTGTGGGTTGTCATTTCACTTCCTTGATGCTATCCTTTCAAGCAAAAATT
TTAAAACCTTATCCATCTTTGCA-3'

10 A7 (EF/D):

5'-CTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAAACCTTGCTAG
TCAGAGAGGAATCTTTGCAGC-3'

A7 (SV40/D):

15 5'-CTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAAACCTTGCTAG
TCTGTGGAATGTGTGTCAGTTAG-3'

A7 (EF y SV40/R):

5'-CTCTTCCCATTCTCATTGGAATCTACTTCAAAGGTTTACCATACTAAGAAGCTAG
TTTTAAAACCTTATCCATCTTTGCA-3'

20 Utilizando el ADN preparado de esta manera para la transformación y el BAC transfectedo con pRed/ET, se clonó el elemento de ADN A2 o A7 en el vector descrito en (3-1). La construcción del vector se muestra en la Fig. 2. De hecho, el procedimiento se llevó a cabo de acuerdo con el procedimiento descrito en (2-2).

(3-3) Evaluación usando la expresión de SEAP como índice

Cada plásmido construido en (3-2) se evaluó usando la célula hospedadora CHO-K1 (ATCC) y el reactivo de transfección Lipofectamina 2000 (Invitrogen).

25 Se llevó a cabo una selección por antibiótico con puromicina a 8 µg/ml durante aproximadamente 2 semanas comenzando a los 2 días tras la transfección, de esta manera se estableció una línea celular policlonal que expresaba establemente. La línea celular establecida de esta manera se sometió a una sustitución del medio el día antes de la medición, y se sembró un número determinado de células en una placa de 24 pocillos. A las 24 horas tras colocar las células en la placa, se recolectó el sobrenadante, y se midió la actividad SEAP. La actividad de SEAP en el sobrenadante del cultivo se midió usando el Ensayo Indicador de Fosfatasa Alcalina segregada pNPP SensoLyte™ (ANASPEC).

30 Los resultados de la medición se muestran en la Fig. 4. Cuando la actividad de SEAP del control sin elemento se normalizaba a 1, el elemento de ADN A2 o A7 mostraba un efecto de aumento de la expresión tal que la actividad de SEAP era mayor en dos veces o más en el caso del uso del promotor EF-1α, y mayor en cuatro veces o más en el caso del uso del promotor SV40 que el del control. Basándose en los resultados, se confirmó que estos elementos de ADN muestran el efecto de aumento de expresión de un gen exógeno cuando se usa en combinación con un promotor general.

Ejemplo 4

Evaluación usando la expresión de anticuerpo como índice

40 (4-1) Construcción de un vector pEF6KCL de expresión de cadena ligera humana

Utilizando el plásmido pEF6/V5-HisB (Invitrogen) como molde, se obtuvo un fragmento de ADN entre la posición 2174 (inmediatamente cadena abajo de BGHPA) y la posición 2958 (SmaI) (un fragmento de ADN que contiene un origen de replicación f1 y un promotor y origen de SV40, de aquí en adelante denominado "fragmento A", estando la secuencia de polinucleótido del fragmento A representada por la SEC ID N°: 6 del listado de secuencias) por el procedimiento PCR usando los siguientes cebadores y KOD-plus-

5'-CCACGCGCCCTGTAGCGGCGCATTAAAGC-3'
5'-AAACCCGGGAGCTTTTTGCAAAAGCCTAGG-3'

El fragmento A obtenido y un fragmento de ADN que contenía una secuencia de ADN que codifica una señal secretora de cadena κ humana, una región constante de cadena κ humana, y una señal de adición de poli(A)

humana (de aquí en adelante denominado "fragmento B") se unieron por PCR de solapamiento. El fragmento de ADN que se obtuvo de esta manera en la que se ligaron el fragmento A y el fragmento B, se digirió con enzimas de restricción KpnI y SmaI, y el fragmento resultante se unió al plásmido pEF6/V5-HisB (Invitrogen) que se digirió con las enzimas de restricción KpnI y SmaI, por lo que se construyó un vector de expresión de cadena ligera humana pEF6KCL que tiene una secuencia de señal, un sitio de clonación, una región constante de cadena k humana, y una secuencia de señal de adición de poli(A) humana cadena abajo del promotor EF-1 α .

El fragmento de ADN que se obtiene por escisión del pEF6KCL obtenido por el procedimiento mencionado anteriormente con las enzimas de restricción KpnI y SmaI se unió al pEF1/myc-HisB (Invitrogen) que se digirió con KpnI y SmaI, seguido de la transformación por lisis alcalina, y su confirmación de secuencia, por lo que se construyó el plásmido pEF1KCL.

(4-2) Construcción del vector de expresión de cadena pesada humana pEF1FCCU

Un fragmento de ADN (la secuencia de polinucleótido de este fragmento de ADN está representada por la SEC ID N°: 7 en el listado de secuencias) que contenía una secuencia de ADN que codificaba una secuencia de señal de IgG1 humana y una secuencia de aminoácidos de una región constante se digirieron con las enzimas de restricción NheI y PmeI, y el fragmento resultante se unió a un plásmido pEF1KCL que se digirió con NheI y PmeI, por lo que se construyó un vector de expresión de cadena pesada humana pEF1FCCU que tenía una secuencia de señal, un sitio de clonación, una región constante de cadena pesada humana, y una secuencia de señal de adición de poli(A) humana cadena abajo del promotor EF-1 α .

(4-3) Construcción de un vector de expresión del gen X de anticuerpo humanizado sencillo (gen X de anticuerpo humanizado/pEF_LHN#)

Uniendo el vector de expresión de cadena L o cadena H construido en (1-1) o (1-2), se construyó un vector de expresión de anticuerpo humanizado sencillo (pEF_LHN (que carece de región variable)).

Se añadió un sitio de enzima de restricción Sall por el procedimiento de PCR a ambos extremos de la unidad de expresión génica desde cadena arriba del promotor a cadena abajo del poli(A) del pEF1KCL. Se llevaron a cabo entonces la electroforesis en gel, el recorte de un fragmento de ADN deseado del gel, y la purificación del fragmento de ADN, preparando de este modo la inserción. Digiriendo el pEF1FCCU construido en (1-2) con la enzima de restricción Sall, se linealizó el vector y el sitio de Sall localizado cadena arriba de la unidad de expresión génica. Entonces, el vector linealizado se unió con la inserción anterior, seguido de transformación, lisis alcalina, y confirmación de secuencias, construyendo de esta manera un vector de expresión de anticuerpo humanizado sencillo (pEF_LHN (que carece de región variable)).

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

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

Los respectivos oligonucleótidos se diluyeron hasta 5 pmol, y usando la T4 polinucleótido cinasa (TAKARA), se permitió que se produjera una reacción a 37 °C durante 1 hora. Luego, se añadió a esto 10 x de tampón (TAKARA), y se llevó a cabo la hibridación a 96 °C durante 1 min a temperatura ambiente. Estos oligonucleótidos y el vector pEF_LHN que se había digerido con la enzima de restricción AatII se unieron, seguido de transformación, lisis alcalina, y confirmación de secuencia, construyéndose así el pEF_LHN# (que carece de región variable).

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

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

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

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

Región variable de cadena H:

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

5 El fragmento de región variable de la cadena H y el vector que tenía insertada la región variable de cadena L en él se digirieron con la enzima de restricción BlnI, seguido de electroforesis en gel de agarosa, recorte de un fragmento deseado del gel, purificación, reacción de ligadura, transformación, lisis alcalina, y confirmación de secuencia, se integró de esta manera la región variable de la cadena H en el vector y se construyó un vector de expresión del gen X de anticuerpo humanizado sencillo (gen X de anticuerpo humanizado/pEF_LHN#).

10 (4-4) Construcción de un vector de expresión de un gen X de anticuerpo humanizado sencillo (gen X de anticuerpo humanizado/pCMV_LHN#)

Utilizando el vector de expresión del gen X de anticuerpo humanizado sencillo (gen X de anticuerpo humanizado/pEF_LHN#) construido en (4-3) como estructura básica de vector, se construyó otro vector de expresión del gen X de anticuerpo humanizado sencillo (gen X de anticuerpo humanizado/pCMV_LHN#) sustituyendo el promotor de acuerdo con el siguiente procedimiento.

15 Utilizando pIRES puro3 como molde, se amplificó un fragmento del promotor CMV por el procedimiento PCR (94 °C durante 30 s y 68 °C durante 3 min x 40 ciclos) usando los siguientes cebadores y KOD-plus-.

Cadena arriba de cadena H:

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

20 Cadena arriba de la cadena L:

5'-TGACGTCGACAAGCTTCGCGTTACATAACTTACGGTAAATGGCC-3'
5'-CTGGATGTCGCCATATGCGCCGAGATCCACAGCAGCAGGGAGATGAACACCTGG
GTCTGCAGCACCATGGTGGCGCTAGCCCGCAGATATCGATCCGAGCTCGGTA-3'

25 A la solución de reacción de PCR se le añadió la enzima de restricción DpnI, y se permitió que se produjera la reacción a 37 °C durante 1 hora, seguido de purificación usando un kit Cleanup reaction miniElute (Qiagen), por lo que se preparó una muestra para su uso en In-Fusion. Mientras tanto, se digirió el gen X de anticuerpo humanizado/pEF-LHN# con las enzimas de restricción HindIII, NheI, y FseI, seguido de electroforesis en gel de agarosa, separando de esta manera dos grandes fragmentos entre los fragmentos resultantes. Cada uno de los fragmentos se recortó del gel, y se extrajo el ADN del gel, preparando de esta manera una muestra para su uso en el kit de clonación PCR Advantage In-Fusión™ (TAKARA), seguido de la transformación, lisis alcalina, y confirmación de secuencia, construyendo de esta manera un vector de expresión del gen X de un anticuerpo humano sencillo (gen X de anticuerpo humanizado/pCMV_LHN#).

30 (4-5) Clonación de un elemento de ADN A7

35 Se seleccionó A7 de entre los 5 tipos de elementos de ADN que se confirmó que tenían un efecto de aumento de la expresión de SEAP, y se clonó en un vector de expresión de anticuerpo.

40 De la misma manera que en (2-2), usando cada uno de los vectores de expresión del gen X de anticuerpo humanizado sencillo (gen X de anticuerpo humanizado/pEF_LHN# y gen X de anticuerpo humanizado/pCMV_LHN#) digeridos con la enzima de restricción NotI como molde, se preparó el ADN para la transformación por el procedimiento de PCR (94 °C durante 15 s, 55 °C durante 30 s, y 68 °C durante 11 min x 30 ciclos) que se llevó a cabo usando los siguientes cebadores y KOD-plus-.

Gen X de anticuerpo humanizado/pEF_LHN# D:

5'-CTCTTCCCATTCTCATTGGAATCTACTTCAAAAAGGTTTACCATACTAAGACTCGA
GGCACTAGTGACGTCAGGTGGCACT-3'

Gen X de anticuerpo humanizado/pEF_LHN# R:

45 5'-CTCTTCCCATTCTCATTGGAATCTACTTCAAAAAGGTTTACCATACTAAGAGCACT
AGTGACGTCAGGTGGCACTTTTCGG-3'

Gen X de anticuerpo humanizado/pCMV_LHN# D:

Se usó el gen X de anticuerpo humanizado/pEF_LHN#D

Gen X de anticuerpo humanizado/pCMV_LHN# R:
Se usó el gen X de anticuerpo humanizado/pEF_LHN# R.

5 Utilizando el ADN preparado anteriormente para la transformación y el BAC transfectado con pRed/ET, se clonó el elemento de ADN A7 en los vectores de expresión del gen X de anticuerpo humanizado sencillo que se describen en (4-3) y (4-4). La construcción del vector se muestra en la Fig. 5. De hecho, se llevó a cabo el procedimiento de acuerdo con el procedimiento descrito en (2-2).

(4-6) Evaluación usando la expresión de anticuerpo como índice

Cada plásmido que se construyó en (1-5) se evaluó usando la célula hospedadora CHO-K1 (ATCC) y el reactivo de transfección Lipofectamina 2000 (Invitrogen).

10 Se llevó a cabo la selección por antibiótico con Geneticina (Roche) a 800 µg/ml durante aproximadamente 2 semanas comenzando 2 días después de la transfección, estableciendo de esta manera una línea celular policlonal que expresaba establemente. La línea celular establecida de esta manera se sometió a sustitución del medio el día antes de la medición, y se sembró un número determinado de células en una placa de 24 pocillos. A las 24 horas tras la colocación en placas de las células, se recolectó el sobrenadante del cultivo, y se midió el nivel de expresión del anticuerpo en el sobrenadante del cultivo por el procedimiento ELISA. De hecho, el ELISA se llevó a cabo de la siguiente manera. Se añadieron 100 µl del sobrenadante del cultivo libre de células a cada pocillo de una placa de 96 pocillos revestidos con un anti-cadena ligera kappa a 50 ng/pocillo, y se incubó la placa a 37 °C durante 1 hora. Posteriormente, se retiró la muestra (sobrenadante del cultivo), y cada pocillo se lavó con 200 µl de PBS-Tween (0,05 %). Luego, se añadieron 100 µl de IgG (Fc) anti-humano marcada con HRP a cada pocillo y la placa se incubó a 37 °C durante 1 hora adicional. A continuación, se retiró la IgG (Fc) anti-humana marcada con HRP, y se lavó cada pocillo con PBS-Tween (0,05 %). Luego se reveló el color usando un kit ABTS (Nacalai) de sustrato POD, y se midió la absorbancia a una longitud de onda determinada de 405 nm. Para la dilución de la anti-cadena ligera kappa, la IgG (Fc) anti-humano y la muestra, se usó PBS-Tween (0,05 %). Utilizando la IgG humana diluida en serie a 12 ng, 6 ng, 3 ng, 0,75 ng, 0,375 ng, y 0,1875 ng como referencia, se calculó la concentración de la muestra.

25 Los resultados se muestran en la Fig. 6. Se confirmó que la muestra que tenía el elemento de ADN A7 tenía un efecto mayor de aumento de producción de anticuerpo en comparación con un control sin elemento cuando se usaba el promotor EF-1α o el promotor CMV en el vector de expresión de anticuerpo.

Ejemplo 5

Longitud de la secuencia que muestra actividad de aumento de expresión de un gen exógeno

30 (5-1) Clonación de elementos de ADN que tienen diferentes longitudes de secuencia

Basándose en la longitud de secuencia que se usa en el ejemplo 2, se construyeron vectores que contenían cada uno de los elementos de ADN pero tenían diferentes longitudes de secuencia.

35 Los detalles de los elementos de ADN que tenían diferentes longitudes de secuencia que fueron diseñados basándose en la longitud completa de cada elemento de ADN A2, A7, A18, B5 y C14 se muestran en las Figs. 7, 9, 11, 13, 15, 18 y 19 respectivamente. El pCMV/SEAP ires Hygro descrito en (2-1) se digirió con la enzima de restricción SpeI durante varias horas, seguido de precipitación en etanol, y se disolvió el precipitado en agua estéril. Utilizando el vector digerido con SpeI como molde, se preparó el ADN para la transformación por el procedimiento PCR (94 °C durante 15 s, 55 °C durante 30 s, y 68 °C durante 10 min x 30 ciclos) usando los siguientes cebadores y KOD-plus-. Utilizando el ADN preparado de esta manera para la transformación y el correspondiente BAC transfectado con pRed/ET, cada elemento de ADN que tenía una diferente longitud de secuencia se clonó en el pCMV/SEAP ires Hygro descrito en (2-1). La construcción del vector se muestra en la Fig. 2. De hecho, el procedimiento se llevó a cabo de acuerdo con el procedimiento descrito en (2-2).

A2-1D:

45 5'-CATGCACAGATTAGCCATTTAGTACTTACTAAATCAAACCTCAATTTCTGAAGTCT
AGTTATTAATAGTAATCAATTACG-3'

A2-1R:

5'-CTCATTCTGTGGGTTGTCATTTCACTTCCTTGATGCTATCCTTTCAAGCAAATT
CAATAATCAATGTCAACGCGTATAT-3'

A2-2D:

50 5'-ACACTGGTCAAAGGGACAGGTCATTGTTATGCTGGCAATGCAGGCTGCTGAAAAC
TAGTTATTAATAGTAATCAATTACG-3'

ES 2 565 428 T3

A2-2R:

5'-ACTGTAGCTTCTTATTTTTTACCTGCAGTGCATTCCTGTAAAAGTAGTGTGGAGT
CAATAATCAATGTCAACGCGTATAT-3'

A2-3D:

5 5'-CTGGAAATTGAGAAGTATCATTCAACAACAGTACCACAAACATGAAATAAATGTGC
TAGTTATTAATAGTAATCAATTACG-3'

A2-3R:

5'-CCAAGCTTGTCCAACCGCGGCCTGCAGGCTGCATGCAGCCTGTGAAGGCTTTGAT
CAATAATCAATGTCAACGCGTATAT-3'

10 A2-4D:

5'-TCAATCATTTATCAATTTTATCTTCAAAGTCCCTCACTTCAGGGAGATGATATAC
TAGTTATTAATAGTAATCAATTACG-3'

A2-4R:

15 5'-ATATATAAAAGTTCATGTATATATAAAATCATGCAATACACGGCCTTTTGTGACT
CAATAATCAATGTCAACGCGTATAT-3'

A2-5D:

5'-CGCATAAAAGGAAAAGCATCCTTAAAATAAACACCATCAATGGCTCCTCGGTGGC
TAGTTATTAATAGTAATCAATTACG-3'

20 A2-5R:

Se usó A2-4R.

A2-6D:

5' -GGGAGGCTACAGCTTGCCTCTCTAACCACTAAAAGGCATGACCCTCCTCAAAGCT
AGTTATTAATAGTAATCAATTACG-3'

25 A2-6R:

Se usó A2-4R.

A2-7D:

5'-TCTGGCTTCCCTGGGCCACGCTGGAAGAAGAATTGTCTTGCGCCACACATAAAAC
TAGTTATTAATAGTAATCAATTACG-3'

30 A2-7R:

5'-AGCTGATTTTTACGTTAAATGTAACATGTAAAGAAATATATGTGTGTTTTAGAT
CAATAATCAATGTCAACGCGTATAT-3'

A2-8D:

5'-GTGAAGAGGAGGAGATGTCAAATTCAAAGTCTTAAATGATGTAGTTTTAAGTAC
TAGTTATTAATAGTAATCAATTACG-3'

35 A2-8R:

5'-ATGACACTTGATATTGTTGTTTATATTGCTGGTTAGTATGTGCCTTCATTTACCT
CAATAATCAATGTCAACGCGTATAT-3'

A2-9D:

Se usó A2-6D.

40 A2-9R:

Se usó A2R.

A2-10D:

Se usó A2-2D.

A2-10R:

Se usó A2-7R.

5

A2-11D:

Se usó A2-8D.

A2-11R:

Se usó A2-2R.

A2-12D:

10

Se usó A2-2D.

A2-12R:

Se usó A2-4R.

A2-13D:

Se usó A2-8D.

15

A2-13R:

Se usó A2-7R.

A2-14D:

Se usó A2D.

A2-14R:

20

Se usó A2-2R.

A2-15D:

Se usó A2-2D.

A2-15R:

Se usó A2R.

25

A2-16D:

Se usó A2-8D.

A2-16R:

Se usó A2-4R.

A2-17D:

30

Se usó A2D.

A2-17R:

Se usó A2-7R.

A7-1D:

35

5'-AAAAACAAAACCTGGAGTAAACAAGATGAATTGTTTTAATAGAGGCACTGTATTAC
TAGTTATTAATAGTAATCAATTACG-3'

A7-1R:

5'-ATACAATGTTCCATGTATTCTGTGCCTGAACCTATGCAGCTGATGTAGCTGAAGT
CAATAATCAATGTCAACGCGTATAT-3'

A7-2D:

5'-GATCTTATTTTCTAAGTAGTATAGACTTAATTGTGAGAACAAAATAAAAAGTTGC
TAGTTATTAATAGTAATCAATTACG-3'

A7-2R:

5 5'-TGTTGTTTTTCAGCCACTAAGTTTGAGGTGATTTGTTCTGGCAGTCCTAGGAACT
CAATAATCAATGTCAACGCGTATAT-3'

A7-3D:

Se usó A7-2D.

A7-3R:

10 5'-AGCCTACACTACCCTTTGCAGCCTTTGGTAACTATCCTTCTGCTGTCTACCTCT
CAATAATCAATGTCAACGCGTATAT-3'

A7-4D:

5'-AGGAGCTCCTGAATGAAGGACATCACTCAGCTGTGTTAAGTATCTGGAACAATAC
TAGTTATTAATAGTAATCAATTACG-3'

15 A7-4R:

5'-GACATAAAATGTAAGATATGATATGCTATGTAAGATATGATACCTGCCTTAAAT
CAATAATCAATGTCAACGCGTATAT-3'

A7-5D:

20 5'-CACTGCTTGATACTTACTGTGGACTTTGAAAATTATGAATGTGTGTGTGTGTGTC
TAGTTATTAATAGTAATCAATTACG-3'

A7-5R:

5'-CAATTACATTCCAGTGATCTGCTACTTAGAATGCATGACTGAACTCCTGGGTGGT
CAATAATCAATGTCAACGCGTATAT-3'

A7-6D:

25 5'-TTATTTTGAAGAGAACTCCTGGTTCCCACTTAAATCCTTTCTTGTTCCTCAAGC
TAGTTATTAATAGTAATCAATTACG-3'

A7-6R:

5'-AAGCAGTGTGTGTTTACCTGCATGTGTATGTGAATTAAGTCTGTTCCCTGAGGCAT
CAATAATCAATGTCAACGCGTATAT-3'

30 A7-7D:

5'-ATTGCATGTTCTCATTATTTGTGGGATGTAAAAATCAAACAATAGAACGTATC
TAGTTATTAATAGTAATCAATTACG-3'

A7-7R:

35 5'-TTGGGAGGCCGCGCAGCTGGTAGATCACTTGAGGCCACGAATTTGACACCAGCAGGT
CAATAATCAATGTCAACGCGTATAT-3'

A7-8D:

Se usó A7-1D.

A7-8R:

Se usó A7R.

40 A7-9D:

Se usó A7-7D.

A7-9R:

A7-5R Se usó.

A7-10D:

Se usó A7-4D.

5

A7-10R:

Se usó A7-7R.

A7-11D:

Se usó A7-6D.

A7-11R:

10

Se usó A7-4R.

A7-12D:

Se usó A7-2D.

A7-12R:

Se usó A7-6R.

15

A7-13D:

Se usó A7-7D.

A7-13R:

Se usó A7R.

A7-14D:

20

Se usó A7-4D.

A7-14R:

Se usó A7-5R.

A7-15D:

Se usó A7-6D.

25

A7-15R:

Se usó A7-7R.

A7-16D:

Se usó A7-2D.

A7-16R:

30

Se usó A7-4R.

A7-17D:

Se usó A7-4D.

A7-17R:

Se usó A7R.

35

A7-18D:

Se usó A7-6D.

A7-18R

Se usó A7-5R.

A18-1:

5 5'-ATCCCCTGCTCTGCTAAAAAGAATGGATGTTGACTCTCAGGCCCTAGTTCTTGA
TCCTATTAATAGTAATCAATTACG-3'

A18-1R:

Se usó A18R.

A18-2D:

10 5'-CTAAAGTGCTGGGATTACAGGCATAAGCCACCGTGCCCGGCTGGAGCATTGGGAT
CCTATTAATAGTAATCAATTACG-3'

A18-2R:

5'-ACTACTTACACATTTTCGAGTTTTAAATAAGGCGTTCAATATAGAGTGAACACCTA
GTCAATAATCAATGTCAACG-3'

A18-3D:

15 5'-CAGGCATAAGCCACCGCACCCGGCCACCCCTTACTAATTTTTAGTAACGTCGATC
CTATTAATAGTAATCAATTACG-3'

A18-3R:

5'-CTGATTGACTTTGACCTCTGCTTTCCAACCTTGCCCCAAAGAAAGTTAGTCACCT
AGTCAATAATCAATGTCAACG-3'

20 A18-4D:

A18-3D Se usó.

A18-4R:

5'-TTCAATGAAACAAGCTCTGTGAGGCTCATTTGTACCCATTTTGTTCAGTACTGCC
TAGTCAATAATCAATGTCAACG-3'

25 B5-1D:

5'-ACATACCCAGAGACACTGAGAGAGACAGACAGACAGTAAACAGAGGAGCACGATC
CTATTAATAGTAATCAATTACG-3'

B5-1R:

B5R Se usó.

30 B5-2D:

5'-GCTCAATTGTATCTTATGAAAACAATTTTTCAAATAAAACAAGAGATATGATCC
TATTAATAGTAATCAATTACG-3'

B5-2R:

Se usó B5R.

35 B5-3D:

5'-CCTGTGCTGAATACCGTCTGCATATGTATAGGAAAGGGTAACTCAGCAGGGATC
CTATTAATAGTAATCAATTACG-3'

B5-3R:

40 5'-TATGTGAATGGAAATAAAATAATCAAGCTTGTTAGAATTGTGTTTCATAATGACCC
TAGTCAATAATCAATGTCAACG-3'

ES 2 565 428 T3

B5-4D:

Se usó B5D.

B5-4R:

5 5'-GAAAGTCTACAATTTTTTCAGTTTAAAATGGTATTTATTTGTAACATGTACCCTA
GTCAATAATCAATGTCAACG-3'

B5-5D:

Se usó B5-1D.

B5-5R:

10 5'-CAAAGATGAAGGATGAGAGTGACTTCTGCCTTCATTATGTTATGTGTTTCATATCC
TAGTCAATAATCAATGTCAACG-3'

B5-6D:

5'-CAGTGAATTATTCACCTTTGTCTTAGTTAAGTAAAAATAAAATCTGACTGTGATCC
TATTAATAGTAATCAATTACG-3'

B5-6R:

15 5'-GAACAGACAGGTGAATGAGCACAGAGGTCATTTGTAAACCGTTTGTGGTTAGCCT
AGTCAATAATCAATGTCAACG-3'

C14-1D:

5'-CTTTTTGGCTTCTGTGTTTAAAGTTATTTTTCCCTAGGCCACAAAACAGAGTCGA
TCCTATTAATAGTAATCAATTACG-3'

20 C14-1R:

5'-AACCTTGAAAAATTCTGTTGTGTTTAGAAGCATGTACCAATCTATCACTCCTAG
TCAATAATCAATGTCAACG-3'

C14-2D:

25 5'-CTATTCAGTGTCTGTAGGATGAAAAAGTTAATAACACCCTGAGAGGTTTCGATCC
TATTAATAGTAATCAATTACG-3'

C14-2R:

5'-CCTTAGATTAGTTTATTGTATTTTTTATCAGCTACTATAAGGTTTACACACCCTA
GTCAATAATCAATGTCAACG-3'

C14-3D:

30 5'-CAAGACCCTCAAAATTCAAAAATTTCTTTATCTTGCTGTAGCACCTCCTGCGAT
CCTATTAATAGTAATCAATTACG-3'

C14-3R:

35 5'-GGAGGGGATAGGAAGGGGATGAGGCCTAACAGGTTGATGATCTAGGCTTTACCTA
GTCAATAATCAATGTCAACG-3'

C14-4D:

5'-CTCAAAAAGGAGATAATTCCAGCCCCTCGCCTTAAAGAATCCCTATCAAGTGATC
CTATTAATAGTAATCAATTACG-3'

C14-4R:

C14-1R Se usó.

40 C14-5D:

5'-CGCTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCCGAGATCACGCCGTTGGATCC
TATTAATAGTAATCAATTACG-3'

ES 2 565 428 T3

C14-5R:

Se usó C14-1R.

C14-6D:

Se usó C14-4D.

5

C14-6R:

5'-TTAACTTTTTTCATCCTACAGACAGTGAATAGTAAAGCTTTCTGTGAAGACATACC
CTAGTCAATAATCAATGTCAACG-3'

C14-7D:

Se usó C14-2D.

10

C14-7R:

Se usó C14-1R.

C14-8D:

Se usó C14-3D.

C14-8R:

15

5'-AAATTATTTTCCTGGTGGGCAATATTAGAATATGGGGAATGTTTGCTTCTGAGCCT
AGTCAATAATCAATGTCAACG-3'

C14-9D:

Se usó C14-4D.

C14-9R:

20

Se usó C14-3R.

C14-10D:

Se usó C14-2D.

C14-10R:

Se usó C14R.

25

C14-11D:

Se usó C14-3D.

C14-11R:

Se usó C14-2R.

C14-12D:

30

Se usó C14-4D.

C14-12R:

Se usó C14-8R.

C14-13D:

Se usó C14-3D.

35

C14-13R:

Se usó C14-1R.

C14-14D:

En cuanto a las secuencias de polinucleótidos de los respectivos fragmentos de C14, C14-1 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 4015 de la SEC ID N°: 5 en el listado de secuencias; C14-2 se corresponde con la secuencia de polinucleótido de los nucleótidos 1987 a 5014 de la SEC ID N°: 5 en el listado de secuencias; C14-3 se corresponde con la secuencia de polinucleótido de los nucleótidos 4020 a 7119 de la SEC ID N°: 5 en el listado de secuencias; C14-4 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-5 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 6011 de la SEC ID N°: 5 en el listado de secuencias; C14-6 se corresponde con la secuencia de polinucleótido de los nucleótidos 4939 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-7 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 5014 de la SEC ID N°: 5 en el listado de secuencias; C14-8 se corresponde con la secuencia de polinucleótido de los nucleótidos 2994 a 7119 de la SEC ID N°: 5 en el listado de secuencias; C14-9 se corresponde con la secuencia de polinucleótido de los nucleótidos 4020 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-10 se corresponde con la secuencia de polinucleótido de los nucleótidos 1 a 5014 de la SEC ID N°: 5 en el listado de secuencias; C14-11 se corresponde con la secuencia de polinucleótido de los nucleótidos 1987 a 7119 de la SEC ID N°: 5 en el listado de secuencias; C14-12 se corresponde con la secuencia de polinucleótido de los nucleótidos 2994 a 8141 de la SEC ID N°: 5 en el listado de secuencias; C14-13 se corresponde con la secuencia de polinucleótido de los nucleótidos 960 a 7119 de la SEC ID N°: 5 en el listado de secuencias; y C14-14 se corresponde con la secuencia de polinucleótido de los nucleótidos 1987 a 8141 de la SEC ID N°: 5 en el listado de secuencias.

Los puntos de inicio y final de los respectivos fragmentos de la secuencia de longitud completa se muestran también en las Figs. 18 y 19.

(5-2) Evaluación de los elementos de ADN que tienen diferentes longitudes de secuencia

Cada plásmido construido en (5-1) se evaluó usando la célula hospedadora CHO-K1 (ATCC) y el reactivo de transfección Lipofectamina 2000 (Invitrogen).

De la misma manera que en (2-3), se llevó a cabo una selección por antibiótico con higromicina tras la transfección, estableciendo de esta manera una línea celular policlonal que expresaba establemente. La línea celular establecida de esta manera se sometió a sustitución del medio el día antes de la medición, y se sembró un número determinado de células en una placa de 24 pocillos. A las 24 horas tras la colocación en placa de las células, se recolectó el sobrenadante del cultivo y se midió la actividad de SEAP.

Los resultados de la medición se muestran en las Fig. 8, 10, 12, 14 y 16. Se confirmó que no solo el elemento de ADN de longitud completa, sino también los clones que tenían una longitud de secuencia más corta que el de longitud completa tenían un efecto de aumento de la expresión. Basándose en los resultados, se confirmó que los elementos de ADN A2, A7, A18, B5, y C14 tienen una actividad de aumento de la expresión de un gen exógeno incluso en los casos en los que tienen una longitud de secuencia más corta que el de longitud completa. Sin embargo, muestran el mayor efecto cuando la longitud de secuencia es la longitud completa.

Ejemplo 6

Efecto del uso de células hospedadoras distintas a la línea celular CHO

Se usó una línea celular CHO como la línea celular en la evaluación de los Ejemplos 2 a 5. Sin embargo, en el ejemplo 6 se seleccionó una línea celular HEK293 como una línea celular distinta a la línea celular CHO. La línea celular HEK293 se sometió a cultivo estático a 37 °C en presencia de un 5 % de CO₂ usando un medio DMEM (Invitrogen) que contenía un 10 % de FCS, y se sembró un número determinado de células en una placa de 6 pocillos el día antes de la transfección. Con el fin de evaluar el vector de expresión de SEAP que contenía cada uno de los elementos de ADN que se construyeron en (3-2), se llevó a cabo la transfección usando cada plásmido y el reactivo de transfección Lipofectamina 2000 (Invitrogen). Se llevó a cabo la selección por antibiótico con higromicina durante aproximadamente 2 semanas comenzando 2 días tras la transfección, estableciendo una línea celular policlonal que expresaba establemente. La línea celular establecida de esta manera se sometió a sustitución del medio el día antes de la medición y se sembró un número determinado de células en placas de 24 pocillos. A las 24 horas tras la colocación de las células en placas, se recolectó el sobrenadante del cultivo, y se midió la actividad de SEAP. La actividad de SEAP en el sobrenadante del cultivo se midió usando el Ensayo Indicador de Fosfatasa Alcalina Segregada pNPP SensoLyte™ (ANASPEC).

Los resultados de la medición se muestran en la Fig. 17. De la misma manera que en el ejemplo 3, se confirmó que cada elemento era también altamente eficaz en el aumento de la expresión de un gen exógeno (SEAP) en la línea celular HEK293.

Aplicabilidad industrial

Al introducir un vector de expresión de un gen exógeno usando el elemento de ADN de acuerdo con la invención en células hospedadoras de mamífero, es posible mejorar la productividad de un gen exógeno de una proteína terapéutica, de un anticuerpo o similares.

LISTADO DE SECUENCIAS

<110> DAIICHI SANKYO COMPANY, LIMITED

5 <120> Elementos de ADN que tienen el efecto de aumentar la expresión de un transgén

<130> PN797479EPA

10 <150> JP2010-154782
<151> 07-07-2010

<150> EP11743651.9
<151> 2011-07-06

15 <160> 7

<170> PatentIn versión 3.4

20 <210> 1
<211> 8450
<212> ADN
<213> *Homo sapiens*

25 <400> 1

ES 2 565 428 T3

attttgcttg aaaggatagc atcaaggaag tgaatgaca acccacagaa tgagagataa 60
 tttttgcaaa tcatgtatct gataagggac ctgtagtcag aatatgcaaa gaacccttac 120
 aattcaataa gacaacccaa tttaaaaaca ggcaaaggat gtgaatagcc atttctccaa 180
 agatacggaa aaacggccaa taagcacata aaaagatgct caaaatcatt tgcattttg 240
 gaaatgcaat caaaaccaca atgaggatc acttcacgcc cattagggtg gctatagatc 300
 agaaagtcag ataacatgtg ttggcaagca catggaaaca ctgaagtcct tacacactgc 360
 tggtaggaat gtaaaatggt gcagccactg tggaaaacag tttccaatt tctcaaaatg 420
 ttaaacacag ttatcataca cccaagcaat tctactctta ggtatatacc caagagaaat 480
 gaaaacatct gctttacca gaacttgctg ttcacagcag cattatgcat aatagaccaa 540
 aagtgaaac aactcaactg cccatcaact ggtgaatgga taagtaaaat gtgalgtaac 600
 cagtcattgg actgtcattc attaataaaa agaacaaggt actgattcat gtictaacat 660
 gagtgaatct tgaaaacact atgctaaatt aaagaagcca gtcacaaaag gccgtgtatt 720
 gcatgatttt atatatatct gaacttttat atatatataa ttatatatat tatatataat 780
 tttatatata taaatttcta tatataaata tataaatca tatatatgat atatattttt 840
 tcatatacat catatatatt tacaanaatt atatatcata tatcatatga tatatgagat 900
 atatatcatg atatatatga tatatgatat atatcatatg agatatatga tatcatgaga 960
 tatatgatat catatgatat atatgatata gatatcatat gatatatata taatatatat 1020
 atgatagata tattatatat gatagataig atagatatca tattatatat gatagatatg 1080
 atagatatca tattatatat gatagatata gatatcatat tatatatgat agatatgata 1140
 gatatcatat tatatatgat agatatgata gatatcatat tatatatgat agatatgata 1200
 gatatcatat tatatatgat agatatgata gatatcatat tatatatgat agatatgata 1260
 gatatcatat tatatatgat agatatgata gatatcatat tatatatgat agatatgata 1320
 gatatcatat tatatatgat agatatgata gatatcatat tatatatgat agatatgata 1380
 gatatcatat tatatatgat atcatatata taccacatac atcatatata catcatatat 1500
 acatcatata tatcatatct atatatgaac ttccagaat aggtatatca ataaagacag 1560
 gaagtatata agtggttgcc acagcctgag aggagcaggg aatggtagt gactgctaat 1620
 ggatagcca ctttttttgg ggggtgatga aatgttctg gtcagacaat ggcaattca 1680

aaactgtata cacacgaaaa accaaagaat cacacacttt aaaagggagg atttagctcg 1740
 gcatggtggc atgcgcctgt actcccagtt actcgggagg ctgaagcagg actgcttaga 1800
 gcccaggact tcaaggctgc agcgagctat gatcgctcca cigcactcca acaaggatga 1860
 cagtgcgaga cccgttttct aaataataat aataataata ataataaata acccaaggtg 1920
 cccagttcac atgcaaaacc actggtaaac ataaattatc tccaagtaat ctagaagaa 1980
 aatgagcaca taagacgict tctaaaaaca cacatatatt tctttacatg ttacatttaa 2040
 cgtaaaaatc agctatgcag aagttacatg aacattttat gttgaaagg taaatgacta 2100
 ttattaatac agaatggita agtacctta tgtttttatg tacaacgca taaaaggaaa 2160
 agcatcctta aaataaacac catcaatggc tctcgggtg tcacaaaaca aaatcctcac 2220
 acctttgtct tcttcacaa ttgagcttta tccacctttt caggcttalc tcccattatt 2280
 acctgacaca aactgggtg gcccagagtt tccactgacc atccccgac tattcatcca 2340
 acactatgtt cactgcctcc cattcctgac catttgccit tigtcttcaa ctaattctgg 2400
 ggacgtttg tccaataaa tgatccatat tctgaaggc tggaaatcaag tctattaca 2460
 aatataatct ctcacctct ccagagcata gcaaccgagc atctactggc ctctcacagc 2520
 tctaaccatc cacaacctta agctggcttc tcatcaaacg ggtactttc accacccaaa 2580
 ttcaattaat tcactcttac aataatgaag aatagtcgcc tacagcctac cttttccagc 2640
 ctgattcaa tcaittatca atttiatctt caaagtcctt cacttcaggg agatgatata 2700
 tcagctttca cccagagtc taaagaaaac agcactcttg ccaatgacat agtgccacct 2760
 agtggcaaca taaggtaaat cacagtgga glagaaggat ctccacacta cttttacagg 2820
 aatgcactgc aggtaaaaaa taagaagcta cagtactgtt tggcaggaca atttgttca 2880
 tacgtgcata ctatcgccct gactaaaita actcgcaagt ctacaggta ttatttgitt 2940
 tcagttccat gcacagatta gccatttagt acttactaaa tcaaaactca tttctgaagt 3000
 gtcttacacc aatatattca tgcacatatg gttaaaaitt tccctgagga tctatcatgt 3060
 gagagtgtg cttattataa caagtaaca gaacaaataa atacaaaatg aaaagaaatc 3120
 gtatgattta ctccatata agggagcttg ttgtgatta agtttcatga cccaggacac 3180
 tgaacagaa atggaataaa tgagaataaa attaaaagtt gtcacaaaa atatagaagc 3240
 catctaaaga cctaggtgtc aagcatagct ctatgagtac aatcccgctc ctgagattac 3300
 catatgccc gctgtatgct atacactaag agatttagga aggaagcggg gtcaggatt 3360
 gacccagac tccatctttt caagtgggga agaaagatct tccgattgaa aaataaaggc 3420
 aaaaaaggct tcaccgtcac agaagtttca acaaccaaca ggatatttaa aacagttatc 3480
 aaagcaaac cattgtatgt tcacttacct tttcatag tccctcaac tcacaaaatg 3540
 ctgttactc agggacttct tccgcttcta ctaggagcc tggaaagta cgggaggatt 3600
 gcaagggacc actagaacc tcttctcaa tccccctct ctgagaaggg aggctacagc 3660
 ttgcctctct aaccactaaa aggcattgacc ctctcaaac ttaatagccg gattccctga 3720
 tagataatct cactaaatga attctcataa aactctcact aagatttaga gaaggcttc 3780
 agggttgaat tctgaacat taagaacagc atgittttta aaagttaac ttggtgattg 3840
 gaccaggact tcatctaggc taigaatgct cagaatgga ggtccittac caaacagctt 3900
 gagtttgtgt ataaagtgt ctcatctct taagagtcag agaaacagaa ccaagcgact 3960

tcactataat ttgatctgag gaagtttctt actcacaata ggtaaatgaa ggcacatact 4020
 aaccagcaat ataaacaaca atatcaagtg tcattcacac atgcaaaaaa cagacaaaat 4080
 cccaaactct gtgttctaac aaatcgcaaa aacctcacta acaataaatt gaaatgacca 4140
 aatgtttgga ctgaaaagca atgccttggg agcctagcca tgcctaactc aaataacaga 4200
 accatctcga tgttaaaatc ctcacagatc aagctgtgta tgtctcgggt caagacttgc 4260
 ccaaaaagca gtgagcacac acttaagagg gaaaaaatct acctcagcct cctaaatgca 4320
 atcatctcta cagagttgc aggcccaag cttcaacgtg ttctgctgga caacgcagta 4380
 gaaagctgac aagcaggtgg ccttcccaca ctgactgaac caccctcatg cccaigtcca 4440
 ttcatittct tccccacccc atgtgctata acagacctcc tggctcaggg cactctttcc 4500
 ttccctgactc ccttcactta atgactttgt acttttaggt gcaaaaatta tctgcagaaa 4560
 tccacactga aaaccaagct tgagaaaggc agcaataacc aacatttita caagaagaac 4620
 aaggtcaata tcaagcccat cagattcaaa tagcaagcat ggatgaaaat gaaagattga 4680
 aaggctttag tgccttctta atgtattaaa tatccattta atttacaatt aagctcactg 4740
 tgcctcactgg ccttttaate agctttccag gtcctgctca gacttgccta ggacatgga 4800
 atgaaagaac ctatacattt atggaccaat ctaccttaac taacttgta agtgitcctg 4860
 catcaagcag aagaaacatc agtgaactg atacaggaat taacccttg itaatccata 4920
 aaacttaag gagcgggac caatcttctg gcttcccctg gccacgctgg aagaagaatt 4980
 gtcttgcgc acacataaaa tacacgaaca ctaataatag ctgctaagct ttaaaaaaat 5040
 tgcaaaaaag gaaaatctca taatttttg ttgtttgta ggtggagcct cactctgta 5100
 cccagcccg agtgcagtg caccatcttg gctcactgca acctctgct cctgggttca 5160
 agccattctc ctgctcagc ctcccagta gctgggatga taggcgtgtg ccaccatgcc 5220
 cagctaattt tctattttt agtagagacg gggtttcacc atgttgcca gctgctctc 5280
 aaactcctga cctcaggtga tccaccacc tggcctccc aaagtgtgtg gattacaggt 5340
 gtgagccacc gtccccgcc aatgtttta gaacgtttac gaatttgiat tgggcccat 5400
 tcaaagcctt cacagctgc atgcagcctg cagcccgcg ttggacaagc ttggattaga 5460
 gaaatctaca gagacaaact agtgaactag tagcctctg atagctcatg atttcaaga 5520
 aacttaggat gactatgtgt aaagaccaca aacatcaatt taactgaatg gttcccacca 5580
 cactggaatg aggaagctga gcaaaactcag aggactctaa gaaaggctg atgcatctg 5640
 aactgttcgg aaltataaac tctctaac atgtttcaaa gccagaactt gtaggagtg 5700
 ttctgataca cggattaaa gagggatgac aaagtgtctg tccccacac tggtaaaagg 5760
 gacaggtcat tglatgctg gcaatgcagg ctgctgaaaa gaatgtatct gtcaaaagta 5820
 atcaaagtaa tgacccaga agctccaga aacagactgg taaattcagg ttgctttcag 5880
 acttccaca tcttggcaca caagggaaa gacaaaacta acalttacag agcattatat 5940
 ttgatattac atttaatccc cattaaaaag atactattc ccgtttcact agtgaaaaag 6000
 ttgatctttc aaaggttaa ttatttaaca ccaaggtaa aggtaagtt ggagagacca 6060
 gattcaaacc cagtctgaca ttaaaacatg tgltttccc ccacatctc tctgctaatt 6120
 aacctcaat ctaaaaactg acttgccta cacttgagc cccatctac aaactctccc 6180
 tgacgttatt aaitcagctg tcactgtgca cctacaacgt gccagacacc atactctca 6240
 acactctgta ggcacagaag gaacagataa aaatccctac ctcatagat attattctag 6300

ggtaacaca gtaaaataa acattaaaat agttttcaca tagtagcaaa ttccataaag 6360
 caaaataaaa cagaagaagg aatagcaaat gagggagaig ccctcttaaa catggtgctg 6420
 agggaaagcc tccttgagaa agataicatt tacccecaaaa ataaaaaagc aagtaataga 6480
 aaaaacaggt aaaaggtgtt ctagacactt aaacctgcca cattgagaac tcagggttct 6540
 gatgcaaaac ctgcctgcat agaatgcatt aacttatttt tatacattta acaaaacaaa 6600
 cttactttaa gaactgtgtt ctaaaggaag gagcatatta caggaaggca atttttggtc 6660
 agagtagaca cacttaaaaa ctaaacctat tgaagacca agaacaactg aaagtctttg 6720
 ctttgcaga ttttgacca aaaggaaaat taaagaaaca caccgtgcc atccaatgat 6780
 ttcaccaagg aattttaaga gagaaaatcc tacttcttcc tcaccagta gccagtgaaa 6840
 tgactgagca aattcacaag ttcactgggg ctgctttcat gtaacacagg gacaacacat 6900
 gacagacaca gtggaacctt acaggttgc tagtattga aagactgtga agaggaggag 6960
 atgtcaaat tcaaagtctt aatgatgta gttttaagta tgttcagcaa ttcaccact 7020
 cagtagtaa gccagctaca gttgaaagca atcagaaatt tgagggtgt gaaataagca 7080
 gaagcacaga agttaaggat ttgtattctt cccacatitt ccactttatt ttatactgt 7140
 gagaaaaaac aaatttaata gttttctgct gtataagaga gacacattca ctttatgtca 7200
 cagtaagagt cactcaatit taatacaact atctcaatgt ataaattaac attctcccc 7260
 ctgccacac atagtaagtc tcttatgatg ttgctgatta gagaagcaaa agttgccgt 7320
 acaattctct tctgcatit taatataaac aatcatcagt ctttcttca tagagtgcag 7380
 tgtggcact atcatcagaa tgtaccagca ctgggtgtgc aaagtttaca aagattagca 7440
 agagcaaaaag tgttgagatt ttigaattc atgctgctgc aaagaagat gtaaaaactc 7500
 actcaccata gaggaccaca cagaaactca ggcatgaagt tatatggctg tgtgagtgt 7560
 ttgggagaag gaacggaaag cacttcacc aacctatag cctgagcaaa ttaatgcaaa 7620
 acctcagaag ctacaaaaaa gtttatctac ctaaataaa attggtgtcc acagcagtag 7680
 ccagcaaat gcctgcgaag cgcaaaagtg taaatatttt agggctgtg ggtcatatgg 7740
 tctctgttaa acaatatgta aatgaatggg tgtggctgtg ttccaataaa acttcattta 7800
 taaaaagagg cagcatggt catccagtca gcaagctata atgtaccaac ccccggtcta 7860
 aactaacca aataccttt aataagccaa agaaactgtg tctcttagg ccggaagcgg 7920
 tggctcacac ctataatccc agcattttgg gagccgagg cgggagatc acctgagtc 7980
 aggagttga gaccatcctg gccaacatgg tgaacccta ttttactaa aaatacaaaa 8040
 attagccagg ctgctgccc ggccctgta atgccaacta ctggggagc tgaagcacga 8100
 gaatcgcttg aaccagagc gcagaggtg cagcgagcct agatcacgc attgcactcc 8160
 agcctgggca acaagagaga aactccgtct caaaaaaaaa aaaggaaata aaagtataca 8220
 aagtgaaac aaagaaatta aactgccctt atttgccagt gacattactg tctatgcaca 8280
 aaattccaaa aatctacaaa aaagcttcta gtactaaaa tgagtttagc aaggtttag 8340
 aatccaaggt cagcatataa cataaatca ccttctata tactagcaat caccaactgg 8400
 aaattgagaa gtatcattca caacagtacc acaaacatga aataaatgig 8450

<210> 2
 <211> 8420
 <212> ADN

<213> *Homo sapiens*

<400> 2

ES 2 565 428 T3

tcttagtatg gtaaaccctt tgaagtagat tcaaatgaga atgggaagag agaaaaagga 60
 gagaagcaac ataagaaatc tctttaagg aattttatat agagagaaac agaggaaica 120
 gttgatagtt gaaattatt ttaaagaaaa tgggttattt taaagaaaa aggtattaca 180
 acatgtttgc actattgigg gaataatcaa gttgagacag aaaattattt ttaaggaag 240
 agtctaattg ctgaagtga agagaatgaa tgagaccctg tgcataagtg tgatcagata 300
 ggagcatgta cagctcaagt aagaacagga agaaagagac aataaacatg tacagatagg 360
 atggectggt cgatgtggg gtgaaaagac atgagagta ttactgatta ctctatttc 420
 cccagtgaaa taggaagcca gttcataaa ccaaaatgaa gaggagcgag gcagtattgg 480
 aagttcagga aaagtaatag gtgtaaaaat atgtaaagta gaattaccag ggagtatgaa 540
 gatacatttc caattaagga tgaagaattt aaagtgaggc cagccaatac ccctgcttgg 600
 cttcagctac atcagctgca taggttcagg cacagaatac atggaacatt gtatttaaat 660
 agggcctgga ttttcaaaa gtaacacaaat gaagaagaga gatgcaaggc tatttgaggg 720
 tgtttgtggg agagattgta aatatttagc taagtaagaa ggggactgca aattttagtg 780
 gtataaagga atgaggaaaa gtgtaaatc atgggggta aagaatgttt ggagccaagg 840
 cactagaggc aattagctga aatgtaggt gattattggt gagtgacatg gtttaaatga 900
 aaagtataga aggtacaat tatccatcat gaaaagtct aggttacaac taagatcga 960
 gtagctgaag tagaatgaaa gtagaatgga ctttccata tccagccagg ttcagtgaca 1020
 gaaggttagg aaacaaatta taaaccactt gagagaacat atcccctaag ttgttttgc 1080
 tattttctt tcagcatata ttgttggaa tgccaactat gttcagttca attaatatgg 1140
 gcttcttaa taaggctcc agcactggat aatcctgcca tttattttga tacattccat 1200
 cctgctgctc agatctattg gcactctacag gatgtctttt gagaagatgg gcattcacat 1260
 ccctatgtcc tagcaaattt ccaactcaga aaaccacatt aggcttctct atatatcttc 1320
 caactatttc aatggaaaa acaattctct gatttcttcc tatgatattt atcaaagaga 1380
 atggtgctg ccagttctag ggtgggggaa ctcaatacaa atcaccaacc tttagatgac 1440
 acctgtctt caaagtctt tcaaagtctg gcagaaaaaa agtaccaggt ggctataaga 1500
 ccaccagga gttcagtcat gcattctaag tagcagatca ctggaatgta atggctagt 1560
 gagttcatt tactctctc tcttggta catgttaccg ccttgtacc ctgcacgttc 1620
 tcttccag acttacaag catgttctt tgaattcgtt ctcttttaa attcacacag 1680
 tcttaatgat tcttcttca caagagtctt tcaactctac aattcagttc aagtcacca 1740
 catgcttatt atgagcaagg gtctgggact taggggaaaa ggaataaaa agatgaatga 1800
 aatgatgcc ctgcagcca agagcttgc gtgaaaagg aagtttgct tacattgcct 1860
 ccctaatecc ttggctaggc cagaacagaa tattgtctaa aacctctca cgtcagcagt 1920
 cctctgggt ggtgactgga agtagaattt aaacaaaaat ataattgaca cataataatt 1980
 gtgcatact atagggtaca atctgatgtt tccatattgt ttaaatggg tgcattgtgt 2040
 aatgatcaa ttgaggtaat ttatccacca cctgaagag agattttca atattctcat 2100
 tgcgaagaag caggaatttt tagcagacaa ctgagatgct tcttgttcc actaagtcat 2160
 tctgacgat gatttacata acttgttgt tttttgtgt gtgtgtttt gagacagagt 2220

ctactttgt cgactaggct gaagtgcagt ggcacaatct cggctcactg caacctccac 2280
 ctccccgggt caaacgattc tcctgcctca gcctcctgag tagctgggat tacaggtgca 2340
 tgcaactagg cctggctaatt ttttatattt ttaatacaga tgggattica ccatgttggc 2400
 cctgctggtg tcaaatcctg ggcctcaagt gatctaccag ctgcccctc ccaaagtcca 2460
 gggattacag gtgtgagaca ccaagcctgg tacatttaca tttcttatct ggatctttcc 2520
 tttagtaagt gctaaggaat cctacttccc ccaatatttt ttctatttc aatgttttag 2580
 catgtatcat gtactactt tgcagacatt tgattttccc ctttgtttac tgaagatgat 2640
 atttttatag cttttgtaat agaagtattc taaaatctgc ctgcaacctc tctttctgac 2700
 tctgcatttt aggaataat tctctgttgt ggaatgaaaa aaaaaacaga gcctgtggag 2760
 tcagagatct catttcaaat tatagttatc cctaggaata aatctgagtg acaggtagta 2820
 tagtataata ataagtataa agctatggtt aaggaaaact caacaacctt atctgtaaat 2880
 tgggatgaca acagcctacg tcaaaaaaat gtgaaggtaa atgagataat gtaaggctga 2940
 tacttagtaa gcaattfaaa aacacccaaa aaactatgac catgattact ctacttactc 3000
 taitttctta tgcctcaggc aatgaacta ctaatgacct aggggtcctt ccccattctc 3060
 ttcttcacaa ggaatattc tctctctgtg tctgtttat taaaatctac tgcctctttt 3120
 agaagccttt ccagatcatc ccatggccaa gaacgatcgc tgcctcctct tctttacata 3180
 cagatgtttt tctcctgctt gacaattatt tttgtgcaat taitttcctt ttgattgtgt 3240
 ttifaatgtc cccccaccc cacaattttc cagactgttt gctccacgag agaggagacc 3300
 atcctctctg tgcctaccgt tgtatgacca gtaicctgag gagtggctgt tacataatta 3360
 catcaggcac tcaataaaaa ttgatgaat aaacactgga ttttaaggca ggtatcatat 3420
 cttacatagc atatcatatc ttacatttta tgcctctcac ataatacca cagagtgaag 3480
 fatatgacag ataaggtcat ttctcttgat aaglacatag tccagtctga aacagatag 3540
 ccaaaaaaaaa acaaacctgg agtaaacaaag atgaattgtt ttaatagagg cactgtatta 3600
 gtttcttagg actgccagaa caaatcacct caaacttagt ggcctgaaaac acaaaaaatt 3660
 taitgtctca cagttataga tgttagaagt ataaaattaa ggtctcagtg ggattggctc 3720
 ctctggggg ctgtggaaga gaatctgtcc caagccttca cacigttaag tacagtactg 3780
 gagggatagg acttcaactt gctctatctc agatagagag gagccatttg ttgtgaattg 3840
 agaagagggg tatgttgaat ccataataag cacataaaaa cttggctggt tcataggaga 3900
 agtaacatgt ttccagctct agtaaaaaac aaattgaagt ggcctataaa aaggtacaga 3960
 gtacgacaga atgaaaaata aatgaacaag aatacagaga ggatgtggtg aattatcatg 4020
 ttccctaat atgttattgg acactaaatg gtattagaat tatttatcaa taataattct 4080
 aaactgttgc aattgaaaga atataattaag tgggtttata tgagaagtg caggccattc 4140
 tcatttctgt ccaatgggag aaacattttc gtttagagacc tccgtgaata atacagtctt 4200
 itagttagga gagctgcat ttgagtggtg caggcagaat ggcgatctct caccacaca 4260
 aacactaaga tagagagaga cagagacaga gacagagaca gcagagagag acagagaaag 4320
 gaagtacagg tactcagata gagataagcc atttcttgac attaagaaat aaagtagaat 4380
 ccattggagg gaaataaaac tgcctcagga acagagttaa ttcacatata catgcaggta 4440
 aacacacact gcttgatact tactgtggac tttgaaaatt atgaatgtgt gtgtgtgtgt 4500
 gtgtgtacat tcagccctcc atatccatgg attitgcatt cacagattca accaaccatg 4560

aattaaaaac attiggaat aacaacatt aaaatataac aatacaaca taaaaataat 4620
acaaataaaa aatatagtt aacaactgt tacatagcat gtagttgta ttaagtagta 4680
taaatctaga gattacttaa tgtataccag aggatgcata ggctatatgc aaatactatg 4740
ccactttaa ctgataaga cagalactaa acttcatctt agccaaaagt cagagaaca 4800
atataactat gccatttiac ataaggact tgagctgagc atcctcagat ttcagtatct 4860
ttggagtcc tggaaacaat tccttgttt atatatatat atgtgtgtgt atatatatat 4920
atatatatac acacatatat atatatatat atatatgata gctactgagt gacaggtgat 4980
attataccat accactgtc acicagtagc tgtatagca tatgtatata tatacatata 5040
catatagtg tgtatgtgta tgtgtgtgtg tgtgtgtgtg tgtgtgtaig ctgtctttcc 5100
tcggtatcac aggaattgg agatatatat atcttttca gtacaaaaa aatigaacac 5160
agatgggtat ggtaccagaa cagaaggtaa agacacatga aaaaaattg caacaacatg 5220
aatggaactg gagatcatta ittgaggaga aataatccag gcacagaaaa acaagcattt 5280
tattatitta ggtgaaagc aaacatttta ttttagtga aataatccag gcacagaaag 5340
acaaacattg catgttcca ittatitgtg gtagtataaa atcaaaaca tagaacgtat 5400
ggaggtagac agcagaagga tagttacca aggctgcaa gggtagtga ggctttgagg 5460
gtgagtgagg gatggttatt ggtacaaaa aatagttaga aagaataat aatatctagt 5520
attaatagc acaacaggtt gactatagtc aaaataacat aattgtaca ttaaataatg 5580
aaatfaata tatatacaag actagaacac caagtgaat gactccagct tgcgaaccc 5640
acattgatca ccatgcttc cccaaggaa gcigtacaat gtctggctcg tccagaaccc 5700
catcattat cactagcaat ctattgtcca taatcatgtt taaattaata gcattttaa 5760
ggtacaaata tttttaaaa aacaataat tatttaattc gccttttaa agcttttaa 5820
aaagttttt aaaaacttt itaaagtcct gaggactatt tctttaaag tctcagtta 5880
cagagctcca tatattggc tatgatagcc ttacctgatt cttgccaaga atctagtcc 5940
cagaaaatgc aaatacaag taagcaactg aaaaataaac aaataagttg gaggtatgct 6000
acctgtgaa atatgacct gcgcaaacac ctatgccact tgcctatgaa atcatatagg 6060
tttctggtt gcagtttga ctgaatgag gattttacgc tggaccaca gggggccct 6120
ctgtcaata cgtactccat ttgtgtatia agtcaaaaat gaaatggaag agaaaagaaa 6180
catcgtatgc ccaagtctc ttaattgaa tggagtaaa agggaaaca cgaatgagaa 6240
aagtactctg cccttttaag aatcttgcct tcacattcct gatgaagtt ttttctctc 6300
tctactgat tccatttca ctctattaca tagcaccgtg tccccagga gctcctgat 6360
gaaggacatc actcagctgt gtttaagtac tggacaata aatatactag tttcaatgtc 6420
taggctatgg gtttcttt ttactgaagg tatgacata agctgccag gccigactaa 6480
attaatagta ataataatta ataattgcaa atttttattc tattaagtt cttggcttga 6540
ctgttagaaa tagcaacatt catctgaaat gccccctct acacttatgt ctaaggacaa 6600
atcccacata caccacagat aacttcatt tacatgttt attctgttac caaactaaat 6660
ttttatcata tagtctgtg ctactgaac tcttcagtaa tctcaacat accatgtaa 6720
gcattaagca cagttcaac acagagcaa tagcaataa ctgttagta ttataacatt 6780
attatgtgtt ttcagtcat taaaccactg gtctgatacc tagccaaca tctatttaa 6840

ES 2 565 428 T3

ccacataatc cagttgaata atatatgata atataataaa atggcgataa gtgctaaata 6900
 tccagataga aacacagatg gaacacagaca gctttcccaa gaaatagaga aaatagtaga 6960
 taggcgatct aggcctaagc actctaagca gaagctaagt taccacagga tatcttggca 7020
 atctgtggca cgtgaaccct tttcttctgg agctctggaac tatgttgcaa ctctcacttt 7080
 ctccctatct agagactcag tttgttccct tgtgattatc agcagttgag aaatccttag 7140
 accttctgaa aggactactt tttaaattta tatatataat attfaaaata catatcttta 7200
 tatataatat atatttaaat atataatatt taaattaata tatatttaaa tatataatat 7260
 ttaaattaat atatatatta ataaataaat ttatatatta atatataata attaaatat 7320
 attttaagt aacagagagt aaaggattat ttgaagaga aactcctggt tcccacttaa 7380
 aatcctttct tgtttccaag ttttcaaat ggagccctct ttaccagctt gccccctcag 7440
 agataagctg ttcccctact taccagatc tgagatctga aaacattcct tttcctgta 7500
 gttcagctag gacaaagatg gagctttttg ataaaatttg gcaaacacat tttttaaga 7560
 tgaaaatttt taaaattga aaaaaaaca ttatagaaa gagacttcta atccaaattt 7620
 aacttctcaa actatgtttt gaccggctag cataatgttt cagtctttct ggagaatgcc 7680
 ccttgaact gttttctct acacaacttc ctccittcct ttgactttcc tgctctggaa 7740
 gggagaaca ggaagaggac agatcaaat actcaagagg aaggacaaga aataaggac 7800
 caaattatca acaattggag aaagaaagct gatgtcagta tcatttcata tatgattatg 7860
 tcagagctag gtggataagc caatcctgtt gaatagcata cttttcctgc tactcctgaa 7920
 gggtaagag gtctttctct tacaagccg tcctagctag taatcttaca ggtgcaaaaa 7980
 gcttgtttc atgttatitc ttagtaactc aaaatacctc taaagttata catattatga 8040
 aagtactaca gtcacagtc tgagaaaagg agtaaataag acaatgtata taaaacact 8100
 tggctcagcc cctggctctg tggttgataa atattaagtt agtattcatt attattataa 8160
 tttcaaaga gtccattaaa agatatagaa gaaggaggc agcaataaca ctaagagaaa 8220
 attcattat ctccaactat ttatcctcta gcccataata attgccatta gaaagagcaa 8280
 cttaacaaa aattttaagt tgcaatagat gttcaacttt aaatccatcc cagaaaaatt 8340
 tctaacaaa ggagcataga agatttgatc ttattttcta agtagtatag acttaattgt 8400
 gagaacaaaa taaaacttg 8420

- <210> 3
- <211> 8475
- <212> ADN
- <213> *Homo sapiens*
- <400> 3

ES 2 565 428 T3

gcataacttg taagaaatgg agtgaggctc cagttcaaac tggcttctgt atgacttcaa 60
agccaaagtc agcaacttag aaggcaaaaa ttataattta gttggcaaat acgagaaaag 120
gtcagaaaca catgaaatga agctcaatag gaacacttac agggtagcag ggtagtagcc 180
tagggaaaaa agtcagacac taaaatggtt taaatagta agttcaaggg acaggtaaag 240
accttagtgg gtaagaagcc aatcagcaga cgaactgcaa gcaagcactg tctctcttc 300
ccttctgtct cctctttag taactgacca caattaagcc tccctagggg aataatgaag 360
taatcctctt attatcagca atggtctgat ccagtgccag gcaccacaga caacttgggt 420
ttcagagaag atccttcaag atgaacaaag ggtcaaaata aaaaattcta gaagagagaa 480

gactgatcac aatttaatgt aagccttsga aggaactgat ctctaccctc cttaacatct 540
caagaacttc ctacagatca ttgatgttg agtgtgtgtg agtctagtag aaaaatgaat 600
tttigtittc taacttggat atgtgattag gatgttaata attaatcttg ggctaatatt 660
gaaggatcti taigtatggc ttcttaaagc attgatcaca aagactgcat gttcataaac 720
tgagctgcac ttgttaggat tctagatgtt tgaattttct tgtgttattt tggctcaga 780
ttctagaca aattttctca aattcctatt tcactttttg acatatcatg agtgactcaa 840
atgtttgcc ttgagtcgga aaacaccag catlaggaat aggcacataa acataatact 900
tcaagcttca gatttaagct caattataaa gtgtttaag gctgtgctga tagttcttct 960
gagtagaatt cctacaacta tgggtttgtc tataataaaa tgttcactct atattgaacg 1020
ccttatttaa aactcgaat gtgtaagtag taataaagaa aatagtcct cctgtaacca 1080
aagctaggac cgattacatg ttcacttgc tgacagatac aatcacctat attaggagca 1140
atcagcactt ccttacaac taacaactg agatgtatg tcccatgg ctatgaagat 1200
ttctttatt tactcagaat agtctgtagg atctgccagc tggccctgat tataccagct 1260
gcaccaatg atcacagtga acattatit acattctaaa taactgtgc aaggtagcc 1320
atggtttct gagtttcta tcaccttgt gttcagtc ctcaaatgt aatttgtaaa 1380
gctgtgttt caggcaaac taacaaat agcatctaat caataacct actatgtcca 1440
cccatacct ataacacaga agtagggaa gagtgagaaa ggtggaagt gagaataga 1500
ggcccaaaa gaaagtitta tcacaggaat atctagatgt ctctgggat tctctgtta 1560
agagcttga cactcatata aatgcagaat tactctctt ctctcttgt ggttagaagg 1620
ccaaggtgc caiggttaata ctaccaaca tatacaaaag ctggcagga aaaaigtac 1680
cttcagaaat ttataatct gatataaat aggtcaagaa atataataa actagtttct 1740
ttggttct tagaaacct gaaaactta aattagaac ttagaagct ttaaatcaga 1800
ctttgtagt aaaaaaggaa atttagt cttccagcat tagaattccg tgattctctg 1860
actctgagc tggattaaat ctaccagc tgagtggaaa cttaagtaac tagctggtg 1920
cctttagtga tcttccact taiggctct tccgctaag aagttcatca tctgactta 1980
ctttcttgg gcaaaagtc tgaactact tcttgggc aaagtggaa agcagagtc 2040
aaagtcaat gaaatggga caaactcact tctactcct tggtaagg gccatttca 2100
gtagccctt ttaagatta gttcattca agattgata agctgtttg acttactat 2160
agatcttati atccatgca gtttaagtt tgccttccact aatctatct gaattcaaaa 2220
ggtaaaaagc taatgctcag tctatcaga tttatctat ttatfaatag aatgtgatt 2280
tttttaagca tatacaata atagtaatga taggaccata aatgtggatg gctctttaca 2340
agtcactaac attacataa ttctcaaca acacactctg aggcataac aaacttttag 2400
aaataacaca atggctacg gaactccagc catctagctt catggctcc cactttaatt 2460
tcaaaaacac agaactgtc acattcatt acatgattag ggcagagctt aactgtatct 2520
catgtagcac ctacatcatt cticagaca acttattgcc tttacagac aagaaaactg 2580
ggctcaaaa aaggactgc ttataactgg ctaataaaga ggaactctg gttcaagtg 2640
agtccaatc tttctccac ccacagctc tctaaagtc attacagaaa tcatagagc 2700
agttctcca cgttattgct taggttcta aagagcagtg acctaataca acatgctcta 2760
taatttatta ctgattaac tatttcta aggattcact ttaactttt aacttgtaaa 2820

taigtctaat aaacaccact gaatatgcaa cctctttctt catggccttg tggttgtaaa 2880
 gcaagctagt aatatatgic tgtggatttg tgctaataaa gttctataca cctcattaat 2940
 tccacaaatc ctactgggta ttctttatct gccagatcct acgctaggta ctggatacac 3000
 agtactgaac aaaatgggta caaatgagcc tcacagagct tgtttcattg aaaagcagag 3060
 agatacacac taalcaacaa attaatagta acacactacg atgtgttttg aaggaaaatt 3120
 agagcatcaa agagacggtg ttagcaggtg gaggggagct cttttagatg gagaatgaga 3180
 atgcctccct aaagacatgg gaataaattg agatcacaaa aatgagaaa tagccagcct 3240
 tgagaagagc agaaggaaga acattcaaag gaaaagaaag tgcatactgg aaagcctgaa 3300
 cactagagtt tggigtatgt aaggagctga gcaatggta cttgtgtgat aagatgtgtg 3360
 gatgtggsgt gggggcagg ggtgagtcct acgcagctct taagtgtgtc ctgagactcc 3420
 tgtgttttc atcagccaca acctgaataa ctgtgtgta atccaaaaat gattacagat 3480
 taaacatata aaaatcat tacaccata gtacctaacg caaggacaca gtattctatc 3540
 ttttcaatga agatctgcat gaagtaaaat tattatata aattttaggt attgatatag 3600
 atacatcagt ggalagatat agatatgtg ctctgtgata gaaaaaagt ttaaaggat 3660
 attaaaagt cttatcttc agggttgaag attgtggcaa ctttcatttc tttttaatt 3720
 taagaaaaa gtgtattat ggggattag catgtttgtg ggtatatgta ttttttaat 3780
 taaaaataa acaacaaaat gaaaacgttt ttcttctatg aaagccta at aagaagaat 3840
 ttcagctgtt ttaacttagg gagctaaaaa catcaaatcc aagaatgttc tctggaactg 3900
 agctcaatc atttttattt gagtaagaat tggatacatt tccatcccct tgggctcca 3960
 gtctgtcaat attttacttt tcagcgataa aaagacacat gtagataatc acagtacct 4020
 cagtaacttt ccttctctta ttaagtita ttttattct atcgtagttt tccctgttaa 4080
 agatttttc tttttgctta catatataat ttagagaat aacaatgcac acacaaaaa 4140
 ttctcttgt tctgctagac ctggactttt tctctaata atatctccat tttttgctt 4200
 tttcagagc tattttgaa gcaaaggaga gaattgctat atagctgact tccctctctc 4260
 atcaacagtg ttttaacagt ttttaagcaa aagtcagctt tgtttatcta agatttttt 4320
 tgcctggcatt taacctacc ctgcctccc ttcccgaat ccacttcagc caacctctca 4380
 ttgcacaggt accacctct aacataactg aaataatgc taccattact gcatctgtct 4440
 agcaaagaat ctcaaattt cccacttgg tglaaattat tttgtaatct ctagtgttta 4500
 aggtgcgctt gtctatcta atcccctcc tggcaggaca ccttacagaa cctaccctt 4560
 aactagtca ttaagcacca tcagggacgg atggctgtgt cactggtctg tttggtattc 4620
 cctactgac ctaccatgig gtgattatct atgacttccc taatccctgg ctgccttagc 4680
 tggactggc tgacatgctt ctgagttgc cgtggcttt acagtccttt actgcccattg 4740
 ccactttgga gataggcagg gctagtactt ttctatata gcccccacac ttgactttgt 4800
 gtttcacagt aggtgaaaaa gtgggtctc tttctttta cttttcttc cacaagaiga 4860
 taaagctagg ggaagcctgt gacatggtt taittctgca actgcaalga ttgattggtg 4920
 ctctctgct cttacttct aaactttgt ctgagtgca gatccctagc agtttctatc 4980
 cctgtctcg ctaaaaaga atggatgtg acctcaggg cctagttctt ttaattaaa 5040
 ttgtatttt gttatcatta ttattattat tttttgaga tgggtctta ctctgtgcc 5100

caggctgaag tgcagtggtg caatcacagc tcactgtttt agcctcctga gtagctggga 5160
 ctacaagcgt catgccacca tgcctctttt taatttttta aaatggtttt ctgccittca 5220
 ttctaagcac ttctcaattg taaccaagag ataactcttt ttatgaattc ttaaagttat 5280
 caacagatac tcaaagtttt agcaaagtct aatgatatt aagcttctcc ttattgcca 5340
 agtgacttca atgactattt gtaattgca accaagggtc attttttaa tgaatata 5400
 ttattattat atataataa ttaaggtcct caaalaccta aaagtttagc aaaatctaaa 5460
 taatatttg ccatattctt tattactgta ttatctcctt tcatgttgc tgataaagac 5520
 ataccaaga ctgggcaatt taaaaaaga agaggttcac tggactcaca gttccacgtg 5580
 gctggggagg cctcacaatc acggcagctt acgggattgt tgagaaatga cacttctcaa 5640
 gctggggcct aactatctct gtggtagttg ttctgattca agtatfgaat tggtttttt 5700
 tgttttttt gagatggagt ttcttcttg ttgccaggc tggagigcaa tggcacgatc 5760
 tcagctcacc gcaacctctg cctcccgggt tcaagigatt ctctgcttc agcctccaa 5820
 gtagctggga ctacaggcat gagccaccac acccagctaa tttgtattt ttatagaga 5880
 catgtttct ccatgttgg caggctggc tcaaacctcc aacctcagg gatccacctg 5940
 ccttggcctc ctaaagtct gggattacag gcataagcca cctgcccgg ctggagcatt 6000
 ggtatataaa agctgcctag gtaactctaa cctttggccc catacatctg aaggatacct 6060
 acaatgcacc tgaaaaatgc aactgaaca gtagttccct gggaccacac actcagaaag 6120
 ggggtgatc agagatcta gggaccagga gggtggaaga cctaaggcag cactacagat 6180
 gatggagaaa aaccacttg ggggggca tctaacctt gagaatcact gagatcatc 6240
 agaagtatt gatcctacag cattaatatt gtattgtatt gtattagat atatatatag 6300
 tgtatataa tagtattagt atatatattg tattgtatta gcataatata actaattgta 6360
 ttgtattgta ttatataa tagtattgta ttatataa tatacagat atatgtatat 6420
 atactaatac aatgtactaa tacaatacaa taccatata atataacta acacaatata 6480
 attagtata atatatata atataacta acaatacaat actatataa tactaataca 6540
 atatatat atatactac caagacatag tagtggctg atgtctggc gccacactca 6600
 tctctacct tcagctctg tctaccaat atcatttct tctgggatct ttgcagtcca 6660
 aggaacttca tcttgatag cccaccctt actaacttt ttttttttt tttttttga 6720
 gacggagtct cgtctgtca cccaggctg agtcagtggt tttgatctc gctcactca 6780
 agctccacct cctgggatca caccattct ctgctcagc ctccaagta gctgggacta 6840
 cagggtccc caccacacc agcctaagt ttaccctgt tagcaaggat ggtctcgatc 6900
 tctgacctc atgatccatc cgccttggc tctaaagtg ctgggattac aggcataagc 6960
 caccgaccc gccaccctt tactaattt tagtaacct caaggattaa aggaaattg 7020
 ccttacctat ttaacaggaa tcaacagggt taatctcact cctttctaa aataattta 7080
 taaacattg agacaatct atctatccct gctaaactg tgggaatta ctgccattta 7140
 atgtaatcag tctactcatt tagtttgcct aaggatttt tgaaaaaaca gttaatgaa 7200
 tgacttaag gaataaccag gaagtgaag tctccaatag taagaatgaa ctcttctct 7260
 ctggataatc aaatgggtc ttctctctc aggtagatca tgcatttcc tcacttacac 7320
 tgaacaggta aacaacataa ttactgact caacttctag ttaattcctt cttttatcac 7380
 tgagtatct ttggctggga gttttgttg ctatgtctcc atttttctt gttatcacag 7440

ES 2 565 428 T3

```

tcctataaca taccaatcct tcaatataac tcatctttaa attgtggtt taccttctca 7500
agaagttatt aattatgcca gtgctaaatc ttctaaaaig attgttgact tgttgaitag 7560
cccccaigca attccccct cccgtccctc agcacgtaag gaatggccct ttgcctactt 7620
ccacagatcc ttaaatctac cagttagaag ctaatagcct acctctctac caggaaggaa 7680
ctgtgggctg gaacataata caigttgact tataatttct tagaaaattg tggagaaac 7740
atcaaactcc tgattccagg atatgcaaaa gacacatcat taaaaagcaa aacaaaacaa 7800
aacaacctc attgacggt gctagtagtg gcatatttca tcaagatcag ctcaaatata 7860
tagaagttag atttcacac aaattagact gtagtgcttt ttttttaac ttatctttac 7920
catatgattt ttaacggtaa aaaaaatcgt ttgagatatt agatgtataa tatttatcat 7980
ccaattactt cattagtcca atcttttttc aatggcgctc ctgcatctga gaataaggtc 8040
agaaaatttc atgttctgat ttcatgctga tttcagaag aaaaatgta 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 gcaaacctgc attagaaagt tagttttcca 8340
tcacaaaaac agtaacaaaa ggaatataaa gtaagttact ttaataatat aagaagaggg 8400
gcaggccggg cgcagtggtc cagcctgta atcccagcac tttgggagcc tgaggcgggt 8460
ggatcacctg aggtc 8475

```

- 5
- <210> 4
 - <211> 8401
 - <212> ADN
 - <213> *Homo sapiens*
 - <400> 4

ES 2 565 428 T3

tttcatctaa gactacattt ctattgtttt atataatcag cccccctaag atcaacatgt	60
ccacatfttt tggcaaagac aaagcctact gatttcagga tcattatftt cctftttcaa	120
aagcacaaac ccaaactgag aaataaatca agagaaattc tcctfttttc taigctaatt	180
tagaagtaga gtctttatft ctttcaaac ccaaagagaa tcagacatac aatatgaatt	240
tatctactft cgcttgctca gactgagagg aaagattaat atttfcaggc tgttagtcaa	300
aactgttcat tcaaatatta ttaataaaa tccaagaacc agctaaaaag tcgcttaagc	360
taagaaacct tcaccagcct catgggaat tegtacagt tttccactag aatagcctat	420
aatgccttac tgaaaatgc taagttcata tcttgtaac taacattfta attcaatctg	480
cagaataata taigcttctt tagtgctaag atatgaatat tagaggcatt ctttcttaa	540
atttctatft agttatactt tcacaaata ctatataata ttaaaattct gcatgtggca	600
taaaacat ttaaatggag aaggtaatgt gtagggagtt tatttctgtt tgctattaga	660
acttgtgttt attcttggtt aaaaaactg cagattaca catagaaaa aacaaaagta	720
tgttgtatat ctcttacagt agaagataaa gagtagttct aaatttagaa aggaaaaata	780
aatatacaca gtgaaaatat gtgfcagta gatgttaatc aaagatcaac tattgctgag	840
accagcaata ttaaatccct gcacaattac tcatattata atgagaattt taaaagaaa	900
atatgaacac ataacataat gaaggcagaa gtcactctca tccttcatct ttgtattccc	960
aattcaggaa gctggtatag tatcttcatt ataattacta ttcaacaaac atttftaaaa	1020

tgaatgaata aggaatgaat gatgagaaaa atgataaaca tctccctctg tctcctggga 1080
 gttactgca ctactttctt ttaaatttaa ttaatcctca atgtccttgt aaaatagcca 1140
 aaggaaaaat gtattfacat tactctaaat atlgatgcaa tctacaaaaa gtgttaaaca 1200
 acttctcaaa agtaaataaa atgttcacaa tccagctagg ataaaaggat ttaaatcatt 1260
 tcttaggtag agggctttca attagagccc ctgctgcatt aacctggga actcatctca 1320
 ctctcttcat gatggagccc tgagtgttc tgctaactcg tactctacca ttctaagtct 1380
 ttaaggctc ctttcagcc cttctctctc gtaatccaca aatactgaga ccaaggcatt 1440
 tttgggtca gtctaatit caagcattct atcttgcctt ccccaaatga actcacatt 1500
 attagaccat atgttctat attagttcag gaaggggaa aaaatgttaa tcacacttgt 1560
 atataagaga tcatagaaaa acagtttact aacctgtgaa aataccattc attctctgtt 1620
 tactcttggc ccacagctaa gcaatcagca ggatataaat gtaccctatg ttactattc 1680
 agtattcata agtatactac ttatgaattg gaaatctgac acaacattta catgacctaa 1740
 tttgaaaaa ttaaaatagt gtaaggcccc taggcttaat tttacagggg aaagattaaa 1800
 gggacacaag caaacatata ttctctctct gtctgtggg acactggtaa tttttgact 1860
 taaaatattt gatacttaa atgccaaact tctacattc tgcagtaaca aggcagttat 1920
 catattgaat accatttctt tctctccagt aagtagagt aatattagca catgaactga 1980
 aatatlaag tgattataaa aacgtccaaa taattcatt aaaatttagc ttggcaaaat 2040
 gttagttca tgttcttggc agaagtcctt ttatattat attcaaatga aatgaacaat 2100
 ttacaagcaa aggaaatggc atcaaatatt tcacacctg cctccaagg tttattgatt 2160
 catgttttt gctcagatct aggtttctc actcagaaa agaggagaat gtaccatac 2220
 ttgggaaaac aagtttccga tggcacagct ttgatcaaac agcaaaattc tatcatcta 2280
 tttattgcca tctgacagta tgacaaatgg tccatgtgc gatattcaca ctgcatgca 2340
 gtcaaacctg taagtcaag gatatgaaat aalagtaact atacattaag cacagaagaa 2400
 aatgaaacaa acaaaaagg tttaaaccaa ccaaaaataa gtcttattt ggatgttcta 2460
 tatgttctta cattctctca gctctttgt gtctattga acacaattct aacaagctg 2520
 attatttat ttccattcac atattacag caacaagctg aaaaagtaga acgggtgta 2580
 gagagacagg acaagtaca gattaggct tgaagtccc ctgaccagtc gacagcaacc 2640
 acatggaata atgactcatg tgcattaatg atcacactaa atgatattg ttttttacc 2700
 tagtcttca actgacagct taaagaactt caggttgtc tgattcttga gcctctcta 2760
 cagcttcaga gaggacttc atttatttt ggatcaaatg ctccacaact agttgaaact 2820
 ggaattaaat tttatatgaa gttcctagat gatttaaagc tgaagaaga agaataatga 2880
 atcataagaa aacttgcctg tacagatac aaaaaggaa gttaccatcc ctcatgctaa 2940
 tcttttcat ttaaaataa caggatctaa aaaaataat gctgggaagt cctaaccaca 3000
 tcaagaatgc ctgagatcag tgaccaggg aaccttccag aatggatgaa atagacccaa 3060
 agctgaattc acctaatit agggccaaaa acccaaaaaa caaaacaaga ccaaaaaaat 3120
 cttcagatac tgggagaaca aatctcaat gcicaattgt atcttatgaa aacaatttt 3180
 caaaataaaa caagagat ttaagattca ttaagttctt gtcatttcaa attttaagaa 3240
 aaatatttc taatggaatt acataatit atatgattct tctagttata tccatggtaa 3300

taaatactct ttccagttgg aaataaaacc catttgtgct atattattag ggaaaatatac 3360
 tacataaatt agtitttaatt ttaactaaag tctatctttt gaaltcataa gcataaaatt 3420
 ttaaccactt gcaaaattta taacacactt aaggtagtca gatgcctigt caagtagtitt 3480
 acaaaaagtg attttcacct gtttgtttta ataacagtgc atcgatttta tgaaaatcag 3540
 gcatgccctc gggtcctaac aaagtatacg aagctgaatg gatctatgcc aaatagcca 3600
 gattttactt tctgagtctg attttatact tctgtcctct ttcttaccac atggcttcca 3660
 glatcactta cagactaacc cttcaaaagg agaaggctaa gttactaaca ttgggaaggc 3720
 ttatgaaagt gaagcatagt tatgagccag caatgttttt atttaggaa tegtgtcaaa 3780
 ccatacactt aagcaagctc tggggaatga gagtggggg gaatcaactc ttttatttgc 3840
 taattggat ttccittaaa agatagagtt ctccagatt itaactgtgt taatagttag 3900
 tctagaaaaa ttggagattt gtgtgcatat attttatgtt gtaaacagac acataccag 3960
 agacactgag agagacagac agacagtaaa cagaggagca ctaaccacaa acggtttaca 4020
 aatgacctct gtctcattc acctgtctgt tccccactt gccttttata gcaactatag 4080
 caacagccat gagagtcatt gtggaagaa ataaaaataa attaaaaat cctggaagct 4140
 tglaaagaat gtgagcaaa gggaggaagt tglaaaaaa atgaataaag ggcaccgatc 4200
 cagagtattg aagaaggcag agtggagagc ctagtatga gtaactgtga cccagatc 4260
 ctctcccaca gaatctgtac agctctccgt ttagacagt itaaacttaa tttaaattat 4320
 caaacagaca ctttctcaa acatataaat gatgaggcag ttcattcagg ctgtatgtat 4380
 aaagtgttc cagccactt tttctaaigg ctctctata tctttacat ggagacaatg 4440
 agagatttgc ttaggacaat ttgactgtaa tttagaagta ggaatggga agtatttga 4500
 tcttctttgc ctaactcaca ttagtactc aagtaagcat tcttccgtt attgcatttt 4560
 cctgattaca agtittatgt tttctctaaa acacataca aagaaatgt cctaagcact 4620
 atcgaggggg aagccatgac atttatccac cactgtcagc aaaaacatga acttagccct 4680
 caacagaata tttcacttca tctagtgtc acctctcgt cacctgcact ggagtcacca 4740
 ctggccigt ttgtaagacc aggatgcacc gctgaaataa aaaggggtca gacaataca 4800
 gaaaagccag tagaaatgc caaatgtatc agaatacaca cagcctttct aaggatateg 4860
 cccaagagga aggcctctaga gccccacctg aacaggatt ttgacttca cagataaatt 4920
 atttaattt caatacaca attcaattaa agaaagggaa atacaaggct aaacaataa 4980
 gaaatgaaga caaaaacca acctttcaa tctaaagaaa ataactgttt itaaagacac 5040
 agatgaagat caggaacca aacagaaga aaggaaggc aattaacgct ggcaictgat 5100
 aacaacgaaa agtatggagt ctggagaatc gctagactct aaaaattata aaggittaga 5160
 ctggacttt gtacactgaa gaaaagaaaa ctgcatgcat ttatactgac caatgtacac 5220
 tattgtctgt ttttaacttt tgtgtatag taggtagat tttttttaa gtgaaagcaa 5280
 gcttattaag aaagtaaaag aataaaaagg tggcttctcc ataggcagaa aactagcgt 5340
 gttttttat tagaaattgt tattcaataa tagtacaigt tacaataaa taccatttta 5400
 aactgaaaaa attgtagact ttcaaatcag ttagggtgt caccctaaaa aaggcattt 5460
 tttcccctta gtcctctgt tcatgttct cacaacaaga aatgggctaa tgcctatgat 5520
 aataataca aacactgct tctgtcaggc cctgtctga ataccgtctg catatgtata 5580
 ggaaagggtt aactcagcag gtcctgtttg cccagactct gtacatttcc aagaaagtc 5640

tgcctttagg actggtcctt gcccagctcc tggagaatga gctctcagct tttagaaaat 5700
 tctatctgct aagaatagtt ttgcatgctc caggtcttgg gccacaaaat atcagtttaa 5760
 tcagatggtt tatgttaaca agtatgattt atggcaaaaca tagatctcta atctccattt 5820
 ctctctcata tatctatatt tatctatcca tatatatgia cctatatata tcaaataatga 5880
 agatatgttt atagcaattg catataaata gagagatagt atgtagtagg aagagagaca 5940
 tagatattat tcttcatttt agaattgltat cttggtagt ttaaaggaa aaacttaaga 6000
 tgtgttcaa ttgcagtatg agtttcaggt atgtacatgt tatgtgtgtg tgtgagagac 6060
 acacacaaac acatttcaa catgtttttat gtttaagctc aatattcaa cacagaaata 6120
 taacatctat tcttaatatg ttttatgtaa gtacagcagc agcattatta aatactgtat 6180
 ttctatggtg attgaaaatt agtaggcaga gaattttgt aatggttctt aataattttt 6240
 gtaatagtaa atgattactt ttgttttagt atagttttat aatctataca tgaataaagt 6300
 ggatatttct attcatatag aaatgtgatt tactctcatg tacttatcta caigctaaaa 6360
 ccataagtia tcaattttag ttctgtcca aggcactttt actgaataaa aataatcagc 6420
 taattttata ttttctgat tcaattttat atgcccgtgt aatgttccgg gtttttttt 6480
 ttaatttct gtaaatcaga atattcagat gttgaaaag tctttgccit cagatttaaa 6540
 agataccttt gaaatgtagc atatcccaa atgcaacca gaggctggca atgtcaacat 6600
 ttttctgttt taaaaaacct cttatgaaa ctattgccat actaaatttt ttacttgctg 6660
 atgacttaca gctgaaaagg attctgtaca tataagacat caaatattga ggatactgga 6720
 acttttaaat taatggcaa gaaagtcaac aaaggaagtt catatgaaat caaactagta 6780
 atatgattac aaaaaaaaaa gtttaaaatt ttcttggcc ccagctttat catttctgag 6840
 ccaaatacaa ttctatcga atcacctgaa actgaaatca ccattctagg ctggttttcc 6900
 cataaagatg gactgctcca aaaagaggaa tcaagaaaga atttggctca cagtgaatta 6960
 ttcaatttgt cttagttaag taaaaataa atctgactgt taactacaga aatcatttca 7020
 aattctgtgg tgataataa gtaatgacca cttttcagct ggagggacta acttcttttt 7080
 tttttttgct gcatatalag ctgtgttaca ttttaatgtg aaatgatgac tgcatacagct 7140
 tatatccatg gaggagattt tagcattcag cttgggtctc ccagtcaata tctacagctc 7200
 tcttcttaag gagatcgatg acacagatac atacagacta acaaatgtga taccaataat 7260
 caagaattca ctcagttaag attttgcca ctgatttcca cacaagaaac ctagaattta 7320
 ctgattcttt gtgctgttga ggctccactc atttccctga atcacaaaag ctacagagta 7380
 tttagataga aatataccta ctcttaacat gaaccatttt aaatataatgt attactgtgt 7440
 ccacaggagt acactttaaa gcagggactt cactcttcaa tctctccat cacgtgttac 7500
 ctaaagtggc atgiggttcc ctaaagctta ataactgaca ttgctttaa aaaggggttt 7560
 gcttcccgac taatgtgaa aaagtctgaa aaatgatttt aaatctttca ctaaatttct 7620
 catttggta cgtggaggaa aatgatttca ccaaatagat actctcatta attttttaat 7680
 gtaatttata aaagaaatga aatatttga taaattccag atttcccca ccatgagctt 7740
 ctccgaaagt atactccatc acagactgct cactaagaag cttacttga gtcaaagtga 7800
 ccgaatttaa gggacataa tgactactt tgctacacag aaacattatc catctctaac 7860
 acttccctat gagatggaag acggacttct aatcaggta cagagaggc tctgccaact 7920

ES 2 565 428 T3

tcaggccttt gatgaataag aatggtgag agcgctcatc ataaatgaat tcagtataac 7980
tgagtgagaa agtgagagaa ccagagaaat aaatcctcat gtagaaaatt taggggatg 8040
aaatgccaaa tgccagttaa ccaaagcttt cttgtcata aagcaacttc tataaaaatt 8100
gctgaaaata aattcttcat ggctcaatgt gaatcagtaa tttccatctc tattacactg 8160
tigitlacc aaaactatt tttatgact aagactcaga gtttccaga gtgtttcca 8220
caaaacaact gtttgagat actccagatc tgtaatcaag taagtctgaa aaacccaaa 8280
tacctcactc acctcttga tatgcataaa gcacactaat atataacgtt ctaaaagcc 8340
aatcattaaa accgttttat atgtttaag cattcctag acatatttg ctacaaatct 8400
a 8401

5 <210> 5
<211> 8427
<212> ADN
<213> *Homo sapiens*

<400> 5

ES 2 565 428 T3

gcacctgcca ccacgcccag ctaattttct attttcagta gagatgaggt ttgccatgt 60
 tggccaggct ggtctcgaac tcttgacctc aggtgatcca cccgcctcag cctaccaaaag 120
 agctgggatt acaggcgtga gccaccgccc ctggccatat taacaaattt taatcacaa 180
 ctatgtgggg ggggaggcta gtattattac agcagattgg tttgctatat aaacaagtac 240
 tttaaaaat atttcttggg ccagcgttgg tggctcagc ctgtaatccc agcactttgg 300
 gaggccgagg tggcagatc acttgaggcc aagagttaag agaccagcct ggccaacatg 360
 gtgaaacccc atcttacta aaaatataac aattagccag gcatggaggt gcatgcctgt 420
 aattccagct gcctcagagg ctgaggcatg agaactgctg gatcctggga ggcagagggt 480
 gcagtgagct gatattgccc cactgcactc catatccagc ctgggcaaca tggcaagact 540
 ccgtctcaa taataaata aataaataa taaaactaaa ggcagagttt tcttaaataa 600
 acatggtagc cctcagcaac aatattgtaa gaactcctcg caagagaaaa agctggaata 660
 agatactggc taagcaagta agaaaggcac tgccttgcct ctgcatacat tcaaactaag 720
 acatatacat tgcagcttac acttacattt tccaatatcc ccaggcatcc ctttcccttc 780
 tcaaacagcc aaaaggaacc agccatgcaa ataaaaatac aagtcaaga gcctaaaaga 840
 agtcagctc ctaaaagaga aaattaatgt aaagaattaa gatttttga aactacactt 900
 tctttctggg gctgittact ggcctccaat acatcaatcc tghtaacctg tgaactacag 960
 tgatagattg gtacatgctt ctaaacacaa cagaattttt ccaaggttac atacactgta 1020
 acaaaagggg cattttgag catcttattt tcttaatca actagtttgg atattctaac 1080
 agtcaaaca ttgtaacaa taaatttca ttacctttg aactttctga agtcaaccaa 1140
 aggcttggg tatggatgca atgagtacta gacaggcaga gctgaatact agtcaaaata 1200
 ttcagttact ggtgtgatag tcttttggg ggcatacacc acttagggag aaactgaggt 1260
 gcaaggacat ttacacaca gcaaaaacat tctcaggaat ttgtcacacc attaccataa 1320
 gccaaaaac tcaaggctt agaacagcct gagcttctga tcaaattata ttgtaaaaag 1380
 agaggaaaa aatgtgaagc gtgctatfff ttaaaataac agtaactact actactgctg 1440
 ctgctgctaa ttctaaactt ttactgagcc ctattatgt gccaaagacc gtgctaggta 1500
 cggatcagaa ttttaacaat taatccctgt acaacccctc tgatattagt taataaaatt 1560

aaagtagaat cctcaccaaa aaaattttaa ctttccaaat aaaaatataa ataaattatt 1620
aaagacattt cacctcttc tctgccctcag actacatttt caagtattaa atttacacta 1680
aaaccacatt tattttcagg aattccagtt aaagcgtaca gatattcaag atgttgacaa 1740
ttattacaga agaatcacag aactctgaaa ttaaatactg gcacagaaaa ccttccatcc 1800
aaccttacgg aacaactatc cccattttaa aaaaaaagga acagcatata taccaggctt 1860
gataataaga ggcttctcat gccacacta gcaatgaatg atgccataat tataaagaga 1920
cctgtatcgc cacatgcata aaaataatth acatctgcta agtcaagttt tcaatatatt 1980
atthttgtgt taaaccttat agtagctgat aaaaaatata ataaactaat ctaaggtaaa 2040
ctaaaacact agttgtttc tgaagactca ctttagaatt tgagcagcat aataatcata 2100
atattagtaa tcaactact tagcagaaag ttcttagagg gctgggaagc lgtgtataat 2160
aaaaaggagc agacaagaag gaagggtttt ccttactgtt taaatcaact acaggctcca 2220
gcatgcagtg ctctaactc aagttaagca aaaactgcaa tgcatactgg gacttgtagt 2280
aagtaacca cgttatcaca gcaagttca agaaagtctg aactatctag cacaattga 2340
ctatatctta ttatcagagt ctaatcaaat ttaaatacaa ttgtatgtt ctctgatgtg 2400
gcacacagtt tctctagcac ataccgaaa aaglatcaat atttagacca acatthtcac 2460
attgaaaaa tcttacgtag gagaagcaca gaaaaaatg ctgaaaaagc aaaaaactt 2520
gatgaataaa aatatataat ttgaaatag thttttaaag ttgaaatgga tccatttcaa 2580
cattctctaa tcttccccca caaaaagttt aatgttttg gccgggcgcg gtggctcagc 2640
ccgtaatcc caacactta ggagctgag gcgggtgaat taccagatca agagatcagc 2700
accatcctgg ccaacatggt gaaacctct ctactaaaaa taaaaaat agttggcct 2760
ggggcgcac gcctgtagtc ccagctactc aggagctga gacaggagaa ttgcttgaac 2820
ctggagggtg gaggtctcag tgagctaata tgcaccact gcactccagc ctggtgacag 2880
tgtgagattc attctcaaaa aaaaaaaaaa aaaaagtta atgttttaa cagttgtctt 2940
tttaacaatt attcaagatg taitttataa ataattttc ttgaagaaaa ttctcagaag 3000
caaacattcc ccatattcta atattgccca ccaggaaata atthttttag taatccgac 3060
acaccccatc acaaaaaaca acaaaaaaca ctgaagtctt gcttttgica agtcttact 3120
caatatttat gccctccatt cctcacctct aatccctac acacacacac acacacgac 3180
acatcccac acacacacgc ttctacaaag aacacttaga aaaacagtat tccaactaca 3240
agcccacttc tctcatccac tgaccttc tgaaaacaca aaagatttt taagctatca 3300
gtaacacgtc caaacacaag ctgataagtt tgagctagaa ttacatata tacagttgct 3360
acacaccctc ctatthctg caagtctgtg gaaggaggct gggaaagaac taagtcaat 3420
ctgcatcagg aggcctaaca caggtgtgtg gttatthca ggcaacagca ccttcacaaa 3480
catgtthtg aatatagtc aagaaattcc taacaaggaa agataagctg gcacacaaat 3540
ttaacgcaat ccagctaaaa atcatctgca acacatgcta ctacattca ccataaaagt 3600
gacggctac tataaaggat ttgaagcttc gtcaatataa cactactgccc ataaggccag 3660
agatagcagt tgccatggtt actataccca ctthttatcag gaaattactg tcatcccc 3720
aaagthttg gtacttatti aaaattttaa aaaaacacac acaatttagg gttctgactg 3780
ttaattgagt gaaataatca actactgtt gatttgaag taigtgctt tggagatgca 3840
catggttaac aatacttga tctgcagcag aaaaaaatc aatthcttc tctgctctc 3900

tctcctcaag tactgacagt ttgtattctc aatgcagcca aaacaataaa acaaaaccca 3960
tctttttggc ttctgtgttt aagttatfff tcccctaggc ccacaacag agtcaaaata 4020
aagcctagat catcaacctg ttaggcctca tccccttctt atcccctcca tactggttca 4080
ctttcttgac tacttagaaa aggcagaaaa catttctgta actgattcca aagtatagaa 4140
aagaatagtt gccitcaact gagatffff caccaaagtc tttttatff acitfffff 4200
taaggcagg agaggggaga gacttgcagg gtactgaaag ggagaagtg aggagtattc 4260
aaattgccac acaagtctag tgaagaaag ttgctttaga agagtccaaa ggaaggctga 4320
acctcacata taatttctaa aagctttgga agagttcacc ataatttta gactgaattg 4380
agggacaagt aatagaaaag ttattcataa agtctacttc aacatffff caaaagataa 4440
ctattcaaaa atttaacaca catataagaa ttatacgaag ccctacaaaa tagtatggcc 4500
acatatacac acaaacatac aaagtagaaa acataagcta ttaagaaat aattatctac 4560
aataaattca atgcaatgit aacataatc ctctffffa aaaaatcgca aagcagcaaa 4620
aacatacacc tgagaaaat aatgtgatca aaacgttaa gaattcttag gcctataaaa 4680
aaagcccatg taaaaagct cctgagaagt caacataat cattaatatt tcccagcaca 4740
aaataatag aaaaatcaaa catgtttcaa gaaatcagtt ctagatag atataaaaga 4800
attcattaa aggtcagaga cctaaaacti taattccttc ctttctctgt ttgaatagta 4860
attaatata aaagccttca gcaataaaat actaaggata caaaattta aagcacatta 4920
atataagctt aacttcaglia tgccttca gaaagctta ctattcactg tctgtaggat 4980
gaaaagtta ataacacct gagagtttc attttatct aacagttaa gtgttttct 5040
cacgcttca agaagcaagt ttctatatt acitctaaa ggggcaatt tcaaaagaat 5100
agtcacttct aaaatttaag atactatacc tttgatagg ctcataaaca cagggttctt 5160
aattatctat atttactit aaaatgctt tattccaat ttgtgagcag agtttataag 5220
aaagctgaaa ctcaaggct taacttttg ggttatttt acacaaaaat attcagtc 5280
actcctctag atttgagtag tcaattcctt gtgcatcct ctaaaataga aaaacaaaa 5340
tgatatacc atataacct aatactaaca catacagata tacatcttt tcaactgtga 5400
acaagcttga aagctttagg cagtaagaat tttcagaaa gttagcagag tcagtcaaaa 5460
cattcaaac ttgaacctg acatctgta ctctgcaat aagagctat agaagaatca 5520
gggaacttac atactacta aaatcaacta ctatcacatc acatcaatgg agaaatgaag 5580
aaaactgta ataggggaca tacaattcac aggatcttca aaaggaaaa tgatctttt 5640
tttttttta aattatgaga aactgactag gcagcatfff ttcaaaagca gcttcaaac 5700
tatacaaaag acatffffg taaccacagc agtatftaaa aaacaaaaat ttaggccggg 5760
cgtggtgct cacgcctata atcccagcac ttgggagcc caaggcaggt gcatcacctg 5820
agtcaggagt tcaagaccag cctgaccaac atggtgatac cccgtctta ctcaaaatc 5880
aaaacttagc cggcctagt ggcggacacc tclataatca cagctactca ggaagctgag 5940
aggcaggaga atccttgaa cctgggagcc agaggttga gtgagccgag atcacgccgt 6000
tgactccag cctgggaaac agagcgagac tccgtctcaa aaaataaaaa aataaaaaa 6060
ctatagtgc cagggtgcac tttaaatgta ttactttctc aactgatatg gaaaagtta 6120
gcatttaag acagaagcti ctgtccatgt attaattagt tacctatctc acaacttaa 6180

tatctgcatg ctttcttacc atttatgaag aacttttata tgtattatct catttggct 6240
 tactgagaaa acagtatltt gcctacaaaa tagacaaaat tcaaagcaga tttatcaaac 6300
 tttctagcat ccccaaattt ttaaaacttc gacacaaaac tttacaagca accacagtgg 6360
 catgatattt tcagtataa tcaattcacc taacactaac agagtttcaa aggaccatgt 6420
 gclataaatg ctatgaaact gttaaagtag ctatattcat ctttatgcag ttactgttac 6480
 atcaacaatg acctaccact gatacaactt gacttacagt tcaagaatct cagtctttgc 6540
 aggctaactt aaglacatca accatatgta ttataaaagc cgagtgccta aaaatigatc 6600
 talattagaa tcatagtctg taaatccgag gggaaaaaac tacaagaagt ctaaaatttt 6660
 ttcaacacac tatacccttt tccaaaatct caactactct atatcttatt tgtattaata 6720
 ttatagggat gataacaagg cttaaagccc taaatcatac caactacttt tgtttataac 6780
 aaitacaat aattttttaa aatacatgct caacatccca ctcatcaaca caagactaat 6840
 tccccitca aataaaataa ttctaaacag tgctctgtac caagggccag aatccttata 6900
 ctatccgcaa tgcacatct actttgtaca gtcaaagact tcactttcaa gtagcaaca 6960
 ttatttatga atggaatttt taaatggact tactcaaat ctttctgaa ctttaagggtg 7020
 ttaatctgt tgcttagctg aagctaagca gagctgtaat aagtagcaag acctcaaaa 7080
 ttcaaaaatt tccittatct tgctgtagca cctcctgctg gatagcattt agagatcttc 7140
 atgtaagcag aagaagagta tticagaggc agctccttcc agaagactga ataggaaaa 7200
 ggatggacc ttcaagcta aaagaatag gccccatcca tcacttatac cttctaaaa 7260
 tacaatttag cccaggtagg tgccttttc atctattact actccagttc cacaaagact 7320
 tgcctcagtc caaaatacaa catgcttaa taaagcctgc aaaatgtct aaaaactaag 7380
 ttaaaaagca ttcaatagca cccaagcaaa acactttatt atgggcagcc aagcaatgtc 7440
 agtcaaaact taaatactat tatgtacca aaagcaaaag tctgatgta aaaaaaaaa 7500
 aaaaaagcc cctggaatat tctgaacatg ttagccagat gtttgtgtt tgagaacttt 7560
 gtgcactatt actatgctct tcacttaagg atagttgtac atctacaac gttttaagta 7620
 cagaaatttt ttataaaca ttagcataac tglacacaaa atttctctt tgcctatgaa 7680
 agataggctc tgggatttga aatgtattt ttacagacatt ttaaatgacc ccctaaaata 7740
 aactagtttt aagcccacaa caccgattcc ataaacaagt aaagacagaa gaagagaata 7800
 agaaggaact taccaaaatt aaaaatgaata atagtattc cagtaaaaat gtagtaacag 7860
 ttccaacaa tgcgtaaac caaataaatt gtgaaactta aaaaaggaag gaggggcca 7920
 gtcttcaaag accaaaagca aagctgacct atttatttct attgcttaga gtgaacacca 7980
 gatgtaaaca aatatataa acactgaaaa gtacgcttac atggttagc ctcaatttca 8040
 gtacccttac cagccctca ataaagctac agatgttggg gagaactcgc tcaaaaagga 8100
 gataattcca gcccctcgc ttaagaatc cctatcaagt gaacctgtga aaagacttcc 8160
 ttcccagagt gcacaactgc ttfaaaaaa aaaaactttc atcagcccaa ataatctga 8220
 ttctaalt caactatcca ttatttatat ataaatgttc ttccctctct aactttcca 8280
 gctcagcat ctacattcct gacaccgact attagcaaaa atgcacaact ccttccccag 8340
 ctatggggca aatctttgaa atctgaaaca cagccacaaa gttcactgtc aaggccaggt 8400
 gatgaggccc acacatgccc ggacctt 8427

ES 2 565 428 T3

<210> 6
 <211> 1704
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Fragmento A

10

```

    ggtaccacc aagctggcta ggtaagcttg ctaccgccac catggtgctg cagaccagg 60
    tgttcattc cctgctgctg tggatctccg gcscatattg cgatatcgtg atgattaaac 120
    gtaccggtgc cccccctcc gtgttcattc tccccctc cgacgagcag ctgaagtcgg 180
    gcaccgcctc cgtgggtgtc ctgctgaata acttctaccc cagagaggcc aaggtgcagt 240
    ggaagggtga caacgcccgt cagtcgggga actcccagga gagcgtgacc gagcaggaca 300
    gcaaggacag cacctacagc ctgagcagca cctgaccct gagcaaagcc gactacgaga 360
    agcacaaggt gtaccctgc gaggtgacc accaggcct gagctcccc gtaccaaga 420
    gcttcaacag ggggagtggt tagggcccg tttaaacggg tggcatccct gtgaccctc 480
    cccagtgctt ctctggccc tggaaattgc cactccagtg cccaccagcc ttgtcctaat 540
    aaaattaagt tgcatcattt tgtctgacta ggtgtccttc tataatatta tggggtggag 600
    gggggtgta tggagcaagg ggcaagttgg gaagacaacc tgtaggcct gggggtcta 660
    ttgggaacca agctggagtg cagtggcaca atcttggctc actgcaatct cccctcctg 720
    ggttcaagc attctctgc ctcagcctcc cgagttgtg ggattccagg catgcatgac 780
    caggctcacc taattttgt tttttgta gagacgggt tccaccatg tggccaggct 840
    ggttccaac tcctaactc aggtgatcta cccaccctgg cctcccaat tctggtgatt 900
    acaggcgtga accactgctc cagcgcctt gtaccggcgc attaagcgc gggggtggt 960
    tggttaccg cagcgtgacc gctacacttg ccagcgcct agcggcgcct cttttcctt 1020
    tcttccctc ctttctgcc acgttcgcc gctttcccg tcaagctcta aatcggggc 1080
    tccctttag gtccgattt agtgccttac ggcacctga ccccaaaaa ctgattagg 1140
    gtgatggtc acgtagttg ccatgcctt gatagcggg ttttcgcct ttgacttgg 1200
    agtccaggt ctttaatagt ggactctgt tccaaactgg aacaacactc aaccctatct 1260
    cggctcttc tttgatita taaggattt tccgatttc ggcctattg ttaaaaaatg 1320
    agctgattt acaaaaattt aacgcgaatt aattctgtg aatgtgtgc agttagggtg 1380
    tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 1440
    agcaaccagg tgggaaagt cccaggctc cccaggcgc agaagtatgc aaagcatgca 1500
    tctcaattg tcagcaacca tagtccgcc cctaactcc cccatcccgc ccctaactc 1560
    gcccagttc gcccattctc cggccatgg ctgactaatt tttttattt atgcagaggc 1620
    cgaggccgc tctgcctctg agctattcca gaagtagta ggagccttt ttggaggcct 1680
    aggcctttgc aaaaagctcc cggg 1704
    
```

<210> 7
 <211> 1120
 <212> ADN
 <213> Secuencia artificial

15

ES 2 565 428 T3

<220>

<223> señal de IgG1 humana + constante de IgG1 humana

<400> 7

5

```

tgctagcggc accatgaaac acctgtggtt ctctctctg ctggtggcag ctcccagatg      60
ggctctgagc cagggtgcaat tgtgcaggcg gttagctcag cctccaccaa ggcccgaagc     120
gtcttcccc tggcacccct ctccaagagc acctctggcg gcacagccgc cctgggctgc     180
ctggtcaagg actacttccc cgaacccgtg accgtgagct ggaactcagg cgccttgacc     240
agcggcgtgc acaccttccc cgtctctctg cagtcctcag gactctactc cctcagcagc     300
gtggtgaccg tgccttcag cagcttgggc acccagacct acatctgcaa cgtgaatcac     360
aagcccagca acaccaaggt ggacaagaga gttgagccca aatcttgtga caaaactcac     420
acatgcccac cctgcccagc acctgaactc ctggggggac cctcagtctt cctcttcccc     480
ccaaaaccca aggacaccct catgatctcc cggaccctg aggtcacatg cgtggtggtg     540
gacgtgagcc acgaagacc ttaggtcaag ttcaactggt acgtggacgg cgtggaggtg     600
cataatgcca agacaaagcc ccggaggag cagtacaaca gcacgtaccg ggtggtcagc     660
gtctcaccg tctgcacca ggactggctg aatggcaagg agtacaagtg caaggtctcc     720
aacaagccc tcccagccc catcgagaaa accatctcca aagccaaagg ccagccccgg     780
gaaccacagg tgtaccct gccccatcc cgggaggaga tgaccaagaa ccaggtcagc     840
ctgacctgcc tggtaaaagg cttctatccc agcgacatcg ccgtggagtg ggagagcaat     900
ggccagccc agaacaacta caagaccacc ctctccgtgc tggactccga cggctccttc     960
ttctctaca gcaagctcac cgtggacaag agcaggtggc agcagggcaa cgtcttctca    1020
tgctccgtga tgcatgaggc tctgcacaac cactacacc agaagagcct ctccctgtct    1080
cccggcaaat gagatatcg gcccgttta acgggtggca                               1120

```

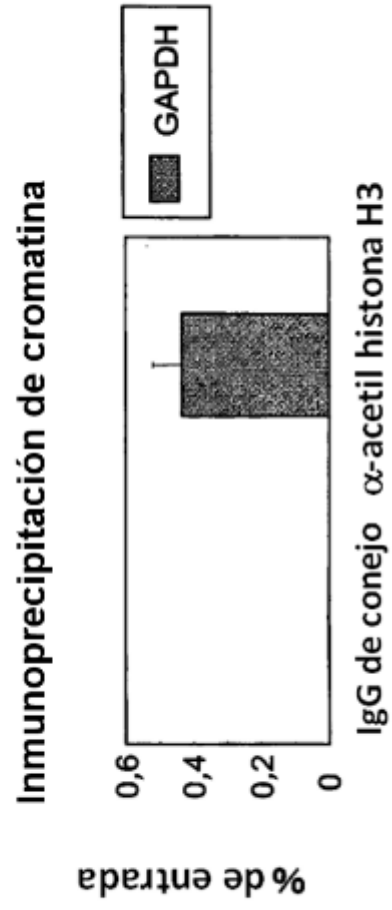

REIVINDICACIONES

1. Un polinucleótido que consiste en una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias.
- 5 2. Un polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias, en el que el polinucleótido es un fragmento parcial de la SEC ID N°: 2 y tiene una actividad de potenciar la expresión de un gen exógeno.
3. Un polinucleótido que consiste en una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más respecto de la secuencia de polinucleótido del polinucleótido de acuerdo con las reivindicaciones 1 o 2 y que tiene una actividad de potenciar la expresión de un gen exógeno.
- 10 4. Un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótido del polinucleótido de acuerdo con una cualquiera de las reivindicaciones 1 a 3 y que tiene una actividad de potenciar la expresión de un gen exógeno.
5. Un polinucleótido que consiste en una secuencia de polinucleótido que contiene:
 - 15 - al menos una de las secuencias de polinucleótido del polinucleótido de acuerdo con una cualquiera de las reivindicaciones 1 a 3; y
 - al menos uno de
 - (a) una secuencia de polinucleótido representada por la SEC ID N°: 1 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos;
 - 20 (b) una secuencia de polinucleótido representada por la SEC ID N°: 3 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos;
 - (c) una secuencia de polinucleótido representada por la SEC ID N°: 4 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos; y/o
 - 25 (d) una secuencia de polinucleótido representada por la SEC ID N°: 5 en el listado de secuencias, secuencia que tiene una homología del 95 % o más respecto de la misma o un fragmento parcial de cualquiera de las dos;

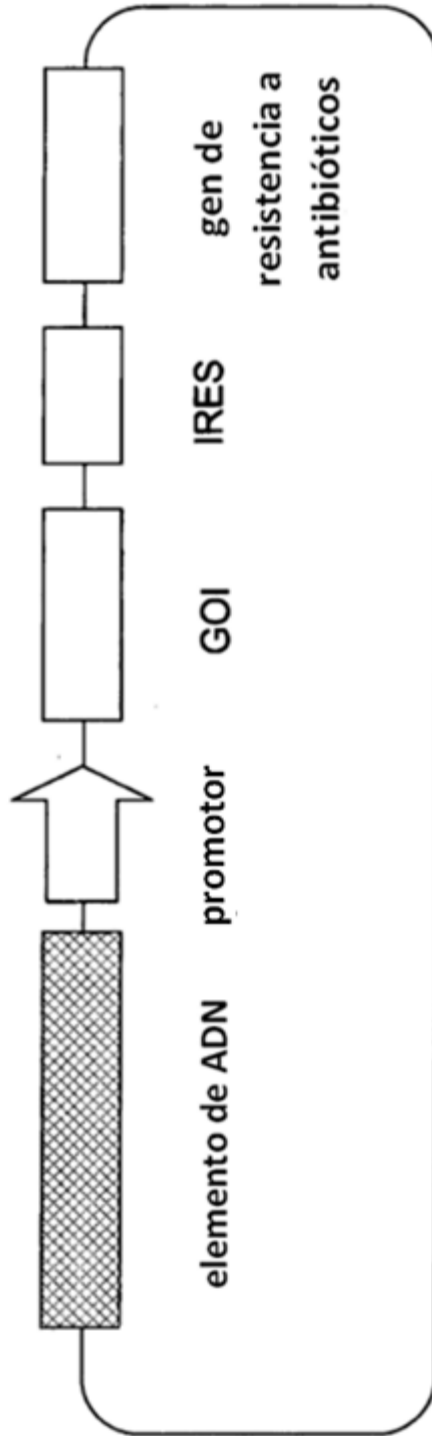
y que tiene una actividad de potenciar la expresión de un gen exógeno.
- 30 6. Un vector que comprende un elemento de ADN que tiene la secuencia de polinucleótido de:
 - (i) un polinucleótido de acuerdo con una cualquiera de las reivindicaciones 1 a 5;
 - (ii) un polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
 - 35 (iii) un polinucleótido que comprende una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más con la secuencia de polinucleótido del polinucleótido que consiste en al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno; o
 - 40 (iv) un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótido del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno.
7. Un vector de acuerdo con la reivindicación 6, en el que el vector es un vector de expresión de un gen exógeno que comprende además un gen exógeno, y en el que el elemento de ADN se posiciona con respecto al gen exógeno de una manera funcionalmente eficaz.
- 45 8. El vector de expresión de un gen exógeno de acuerdo con la reivindicación 7, en el que la proteína codificada por el gen exógeno es una proteína multimérica o una proteína hetero-multimérica.
9. El vector de expresión de un gen exógeno de acuerdo con la reivindicación 8, en el que la proteína codificada por el gen exógeno es una proteína hetero-multimérica y es un anticuerpo o un fragmento funcional del mismo.
- 50 10. Una célula transformada en la que:
 - a) se ha introducido el vector de acuerdo con la reivindicación 6 y un gen exógeno, o
 - b) se ha introducido el vector de expresión de un gen exógeno de acuerdo con una cualquiera de las reivindicaciones 7 a 9.

11. La célula transformada de acuerdo con la reivindicación 10, en la que la célula es una célula cultivada procedente de un mamífero preferentemente en la que la célula cultivada es una célula que se selecciona de entre el grupo que consiste en células COS-1, células 293, y células CHO.
- 5 12. La célula transformada de acuerdo con la reivindicación 10 u 11, en la que la proteína codificada por el gen exógeno es una proteína multimérica o una proteína hetero-multimérica.
13. La célula transformada de acuerdo con la reivindicación 12, en la que la proteína codificada por el gen exógeno es una proteína hetero-multimérica y es un anticuerpo o un fragmento funcional del mismo.
- 10 14. Un procedimiento de producción de una proteína **caracterizado porque** comprende cultivar la célula transformada de acuerdo con una cualquiera de las reivindicaciones 10 a 13 y obtener la proteína codificada por el gen exógeno a partir del producto de cultivo resultante.
- 15 15. Un procedimiento para potenciar la expresión de un gen exógeno en una célula transformada en la que se ha introducido un gen exógeno o un vector de expresión de un gen exógeno, **caracterizado porque** usa:
- (i) un polinucleótido de acuerdo con una cualquiera de las reivindicaciones 1 a 5;
 - (ii) un polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
 - (iii) un polinucleótido que comprende una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más respecto de la secuencia de polinucleótido del polinucleótido que consiste en al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
 - (iv) un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótido del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno; o
 - (v) un vector de acuerdo con la reivindicación 6 o un vector de expresión de un gen exógeno de acuerdo con una cualquiera de las reivindicaciones 7 a 9.
16. El uso de:
- (i) el polinucleótido de acuerdo con una cualquiera de las reivindicaciones 1 a 5;
 - (ii) un polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
 - (iii) un polinucleótido que consiste en una secuencia de polinucleótido que tiene una homología del 95 % o más o del 99 % o más respecto de la secuencia de polinucleótido del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno;
 - (iv) un polinucleótido que consiste en una secuencia de polinucleótido que contiene dos o más de las secuencias de polinucleótido del polinucleótido que comprende al menos 3000, 2000 o 1500 nucleótidos consecutivos de una secuencia de polinucleótido representada por la SEC ID N°: 2 en el listado de secuencias y que tiene una actividad de potenciar la expresión de un gen exógeno; o
 - (v) un vector de acuerdo con la reivindicación 6, para potenciar la expresión de un gen exógeno en una célula transformada.

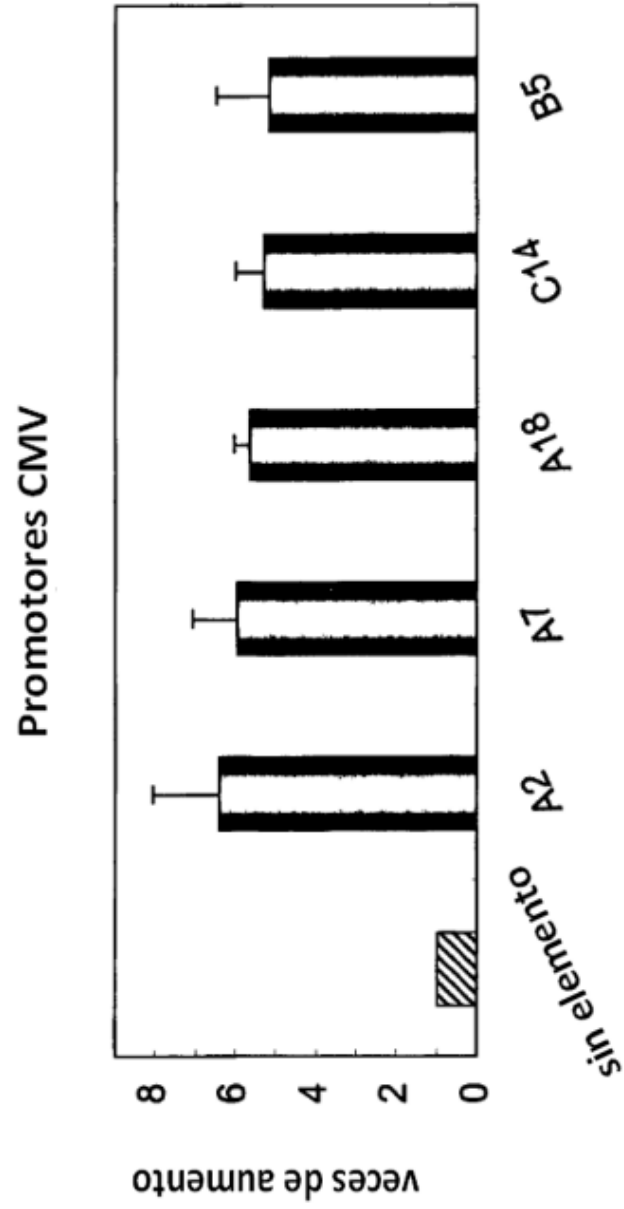
[Fig.1]



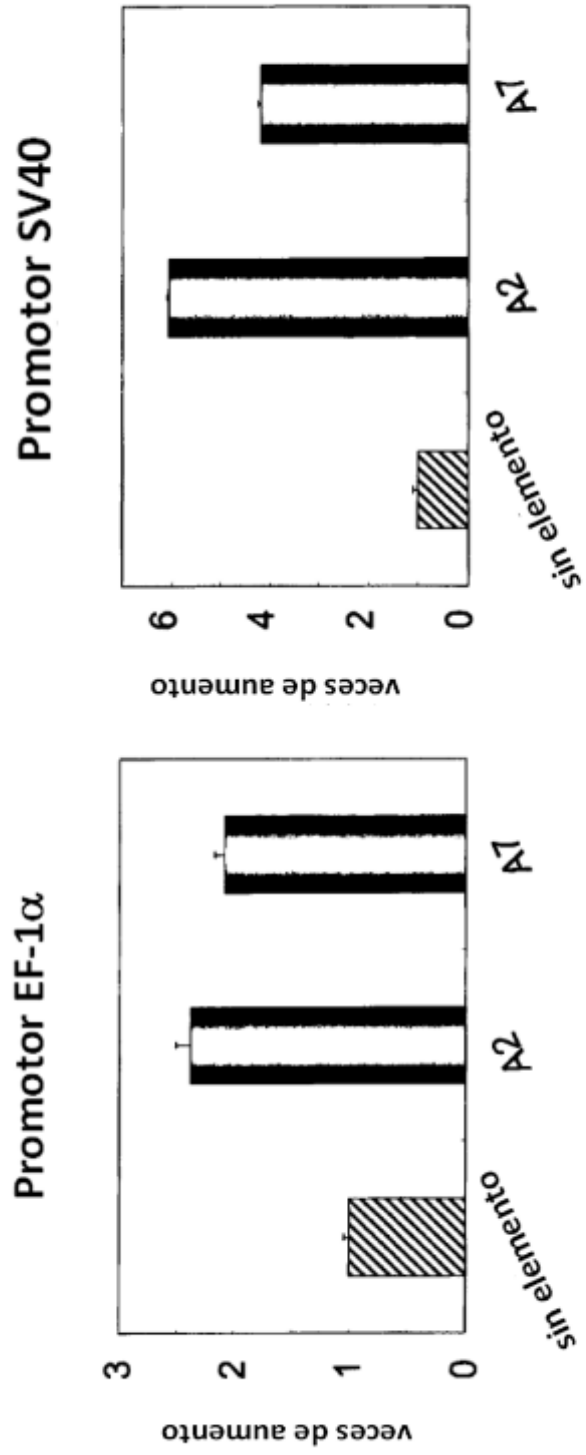
[Fig.2]



GOI : Gen de interés (SEAP)

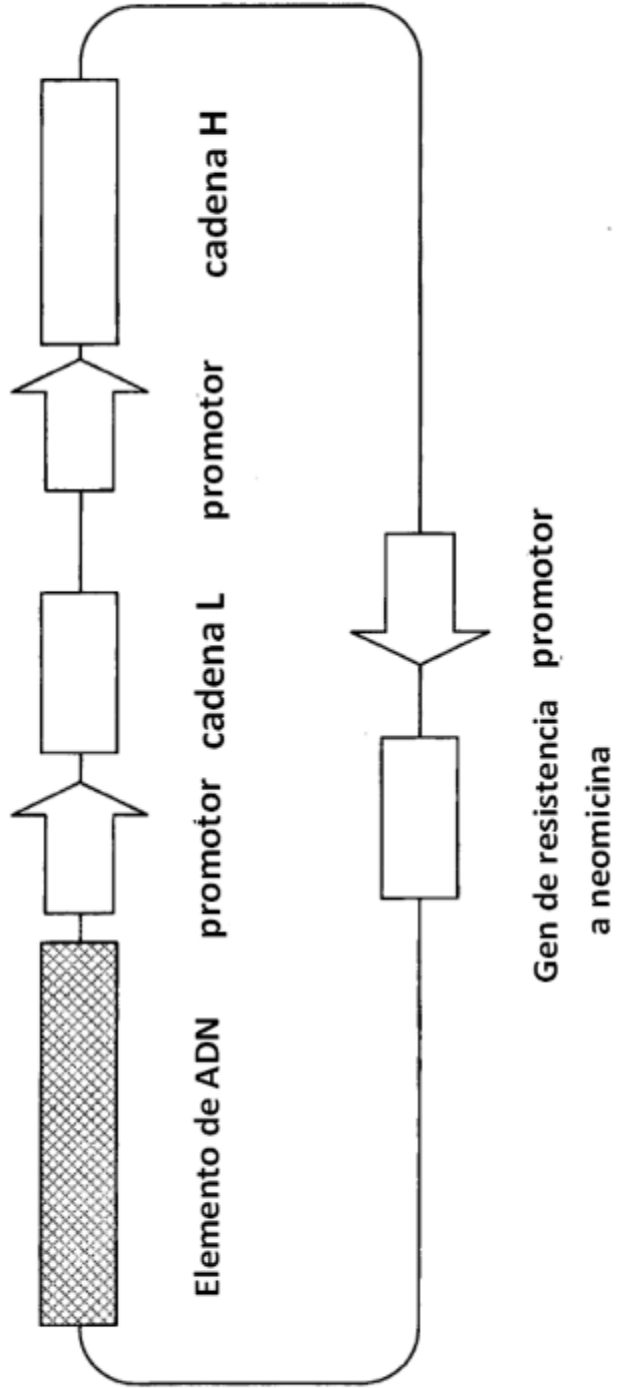


[Fig.3]

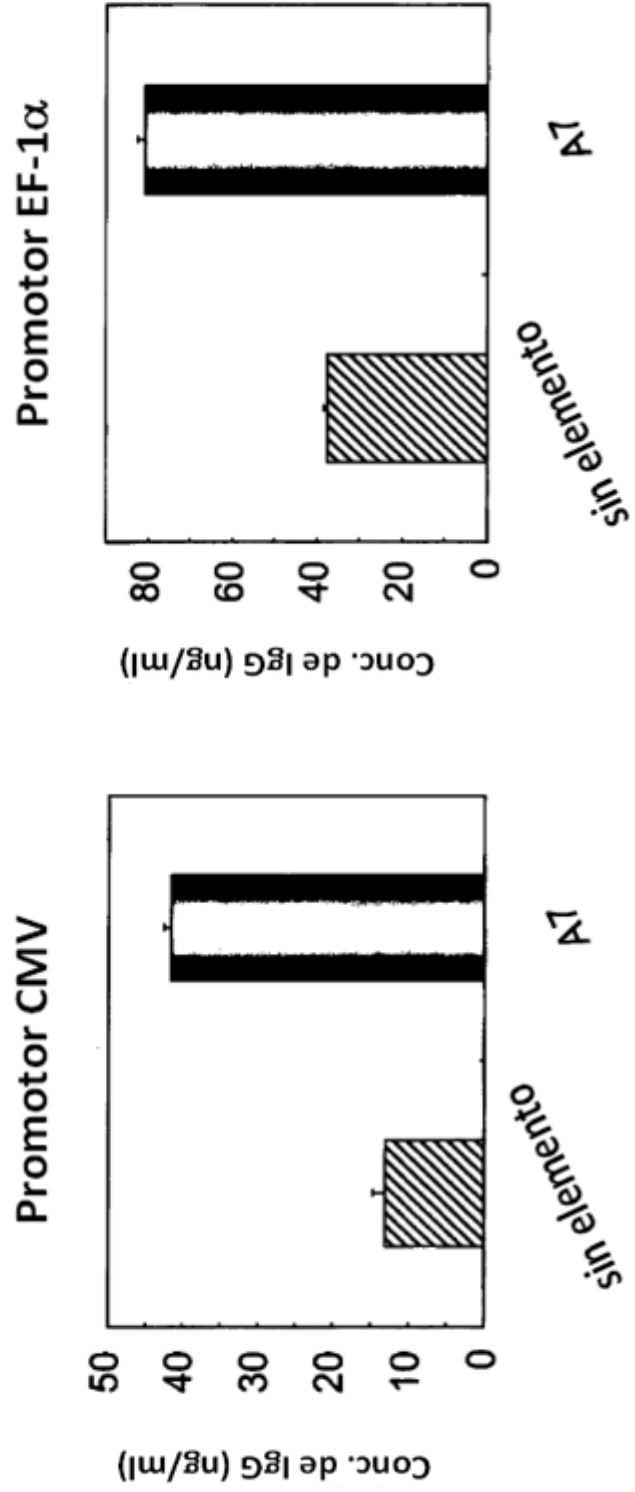


[Fig.4]

[Fig.5]



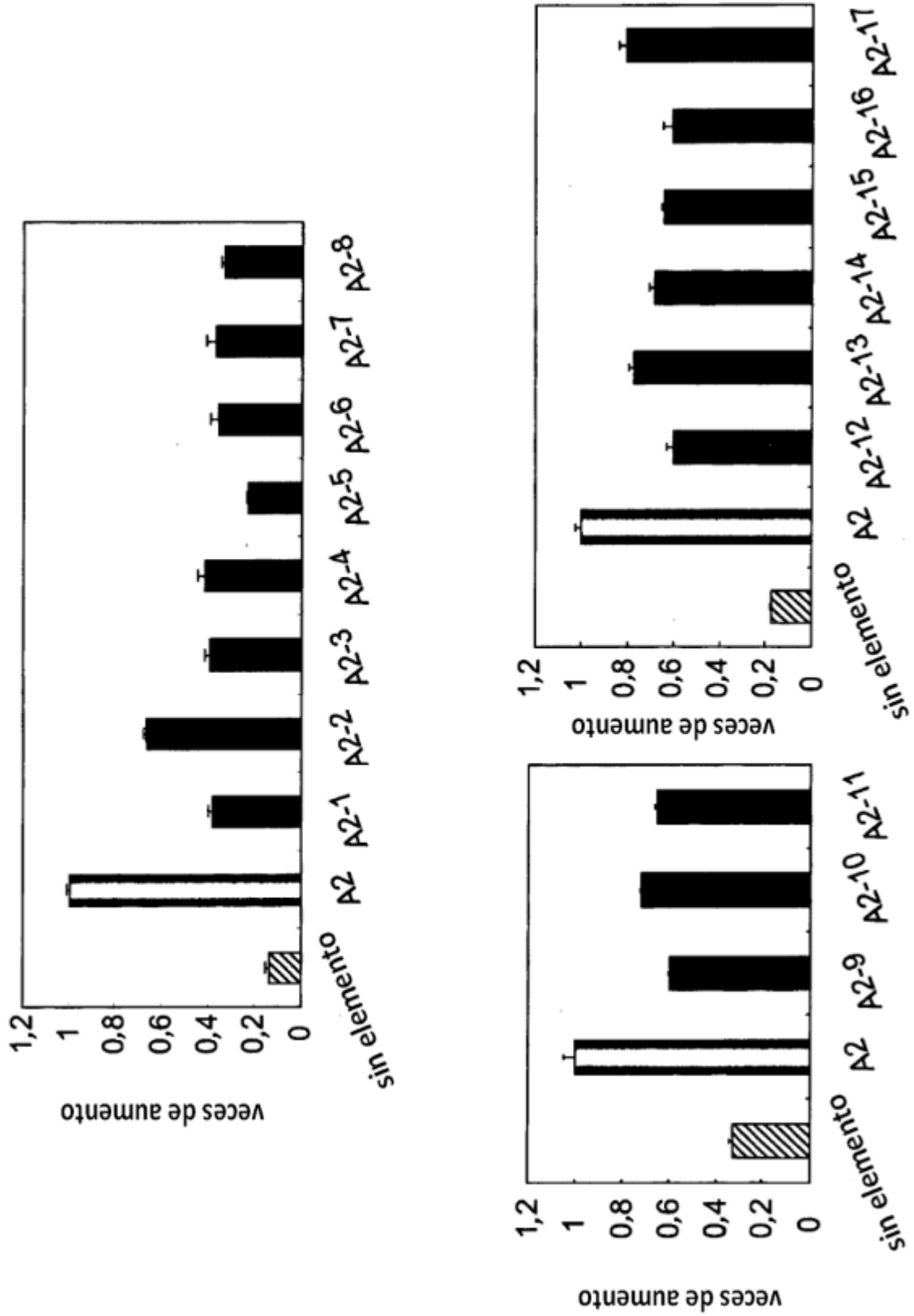
[Fig.6]



[Fig.7]

ID	Localización de inicio	Localización final	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

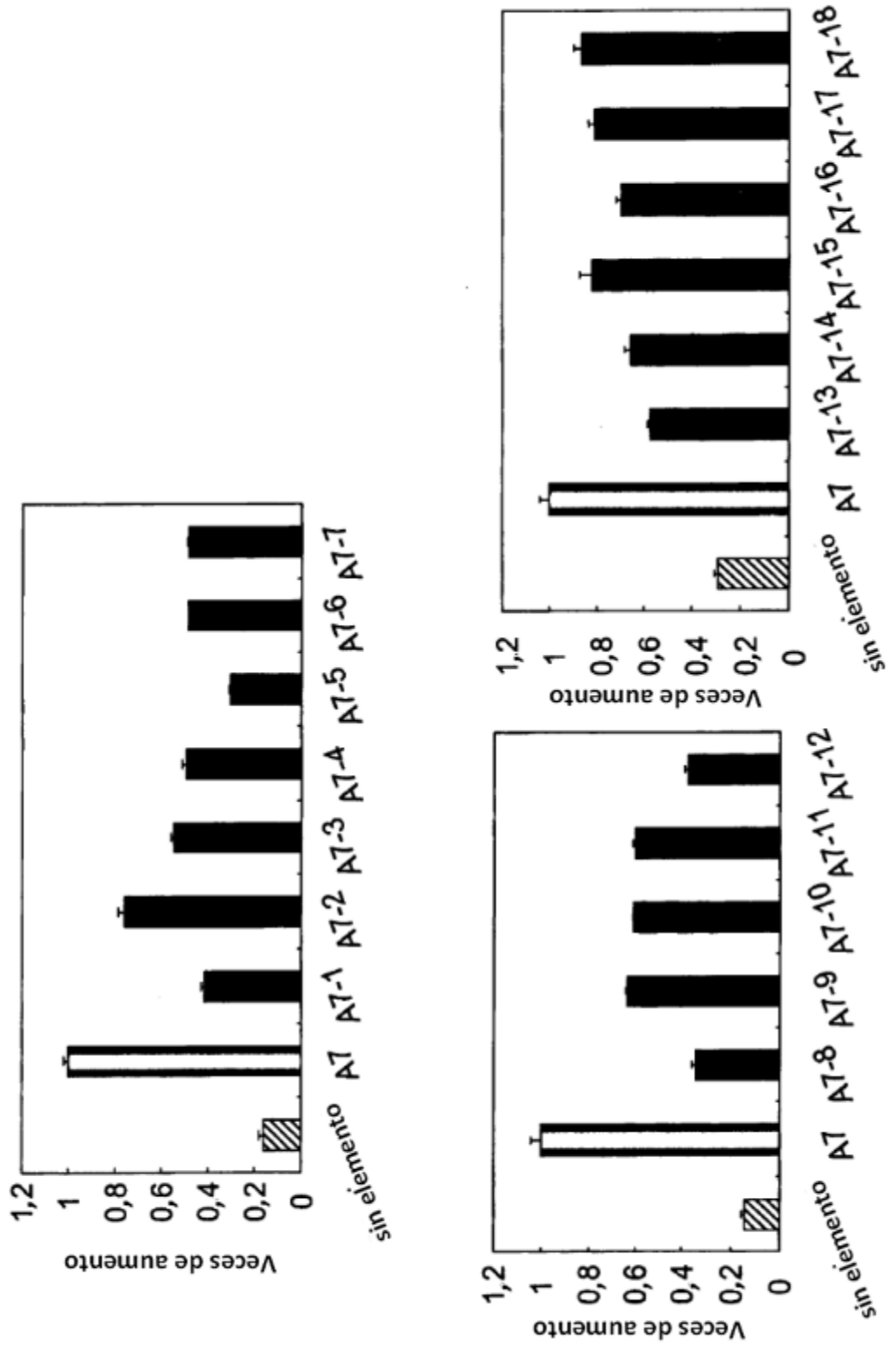
[Fig.8]



[Fig.9]

ID	Localización de inicio	Localización final	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

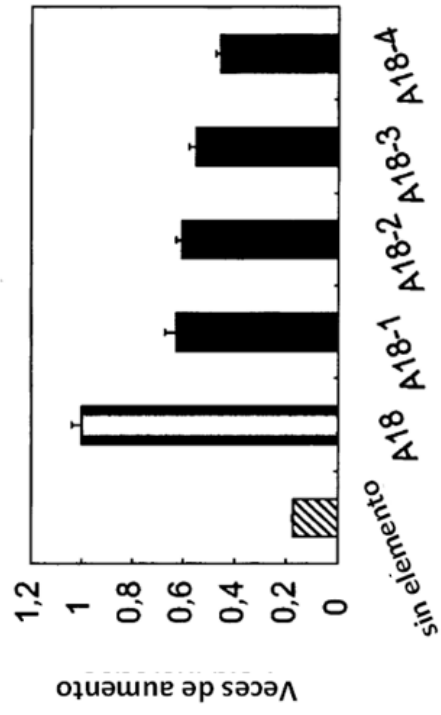
[Fig.10]



[Fig.11]

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

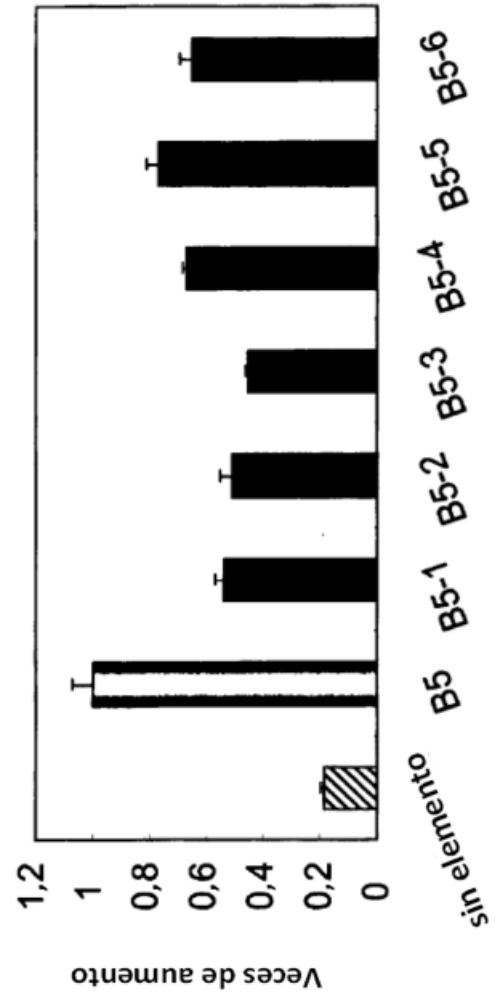
[Fig.12]



[Fig.13]

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

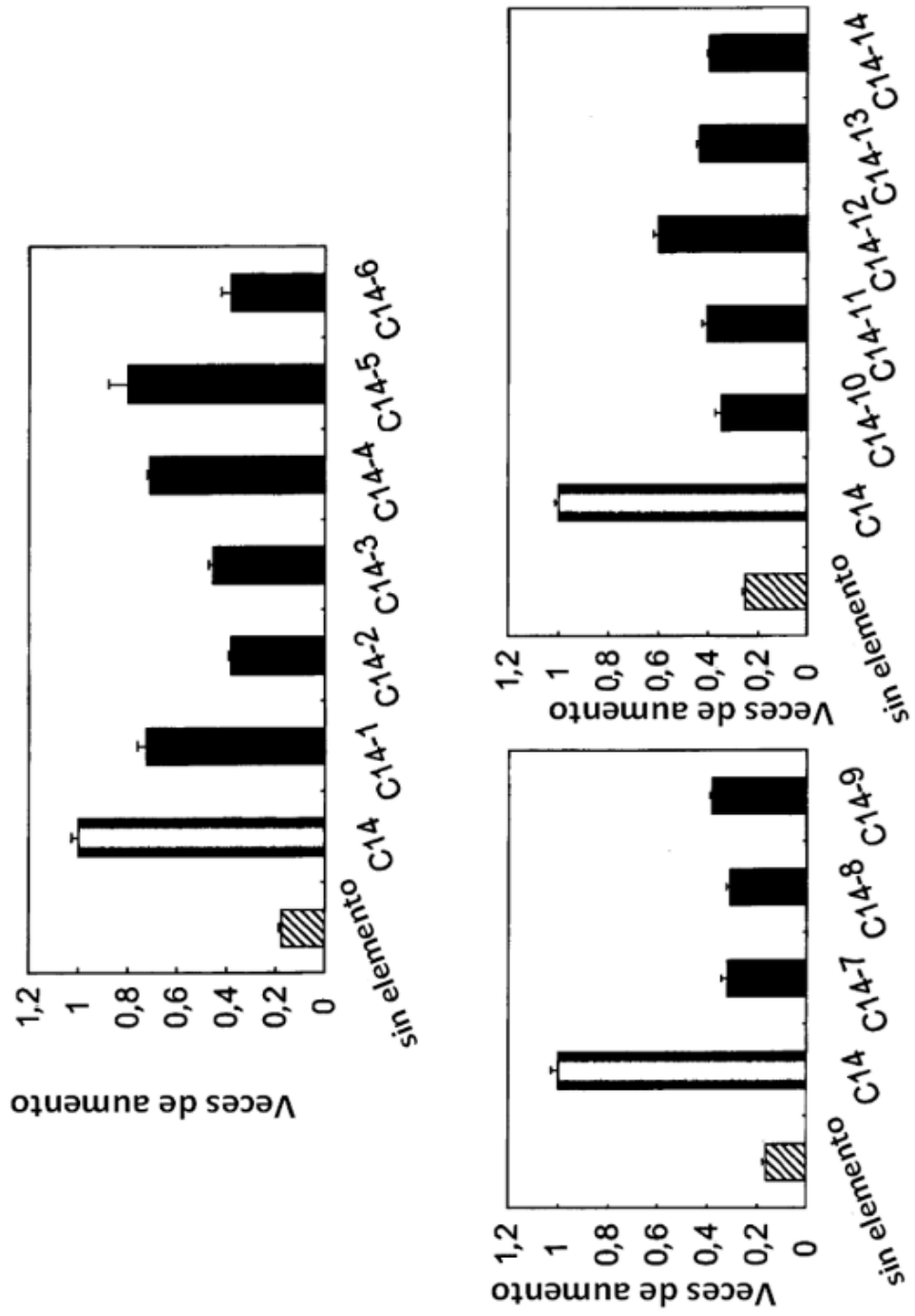
[Fig.14]

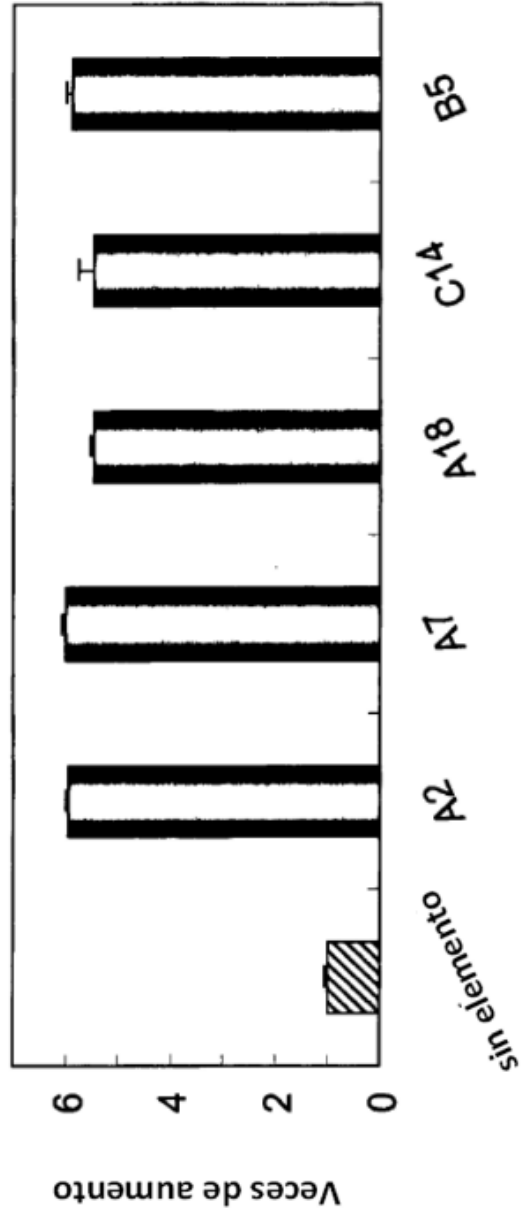


[Fig.15]

ID	Localización de inicio	Localización final	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. 16]





[Fig.17]

[Fig. 18]

Los puntos de inicio y final basándose en la secuencia de longitud completa del elemento de ADN		Punto de inicio		Punto de final	
A2		A7		A18	
A2	1	8450	A7	1	8420
A2-1	1	3000	A7-1	601	3600
A2-2	2801	5800	A7-2	3601	8420
A2-3	5401	8450	A7-3	5401	8420
A2-4	701	2700	A7-4	3401	6400
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
			A18	1	8475
			A18-1	1	5040
			A18-2	1001	6002
			A18-3	2001	7000
			A18-4	3000	7000

[Fig.19]

Los puntos de inicio y final basándose en la secuencia de longitud completa del elemento de ADN
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

Los puntos de inicio y final basándose en la secuencia de longitud completa del elemento de ADN
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