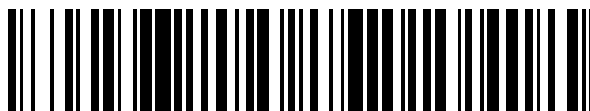


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 656 079**

51 Int. Cl.:

**C12N 15/09** (2006.01)

**C12N 5/10** (2006.01)

**C12P 21/02** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.12.2011 PCT/JP2011/078938**

87 Fecha y número de publicación internacional: **21.06.2012 WO12081629**

96 Fecha de presentación y número de la solicitud europea: **14.12.2011 E 11848634 (9)**

97 Fecha y número de publicación de la concesión europea: **08.11.2017 EP 2653541**

54 Título: **Procedimiento para producir proteínas**

30 Prioridad:

**15.12.2010 JP 2010279850**

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

**23.02.2018**

73 Titular/es:

**KYOWA HAKKO KIRIN CO., LTD. (100.0%)  
1-6-1, Ohtemachi, Chiyoda-ku  
Tokyo , JP**

72 Inventor/es:

**KUROKAWA MEGUMI;  
HAYASHI YOKO y  
TSUKAHARA MASAYOSHI**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

ES 2 656 079 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

Procedimiento para producir proteínas.

5 **Campo técnico**

La presente invención se refiere a un procedimiento para la producción de una proteína de interés, que comprende introducir un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable y que comprende además un par de secuencias de transposón en ambos extremos del fragmento génico, en una célula de mamífero en suspensión; integrar el fragmento génico insertado entre el par de secuencias de transposón en un cromosoma de la célula de mamífero; obtener una célula de mamífero en suspensión que produce la proteína de interés, y cultivar en suspensión la célula de mamífero, y una célula de mamífero en suspensión que expresa la proteína de interés mediante el procedimiento.

15 **Antecedentes de la técnica**

La producción de proteínas exógenas mediante técnicas de ADN recombinante se utiliza en diversas industrias, tales como la industria farmacéutica y la industria alimentaria. En la mayoría de casos, la producción de proteínas recombinantes se lleva a cabo mediante la introducción de un vector de expresión que comprende una secuencia de nucleótidos codificante de una proteína de interés en un hospedador, tal como *Escherichia coli*, levadura, célula de insecto, célula vegetal y célula animal, seleccionar un transformante en el que se integra el vector de expresión en el cromosoma y cultivar además la línea de células transformadas bajo condiciones de cultivo apropiadas.

Sin embargo, con el fin de desarrollar un hospedador que puede producir una proteína exógena eficientemente, resulta necesario seleccionar una célula hospedadora con buena productividad de cada proteína de interés, de manera que se desea una innovación técnica adicional de las técnicas de producción de proteínas exógenas para cada hospedador.

En los sistemas bacterianos, tales como *Escherichia coli*, o en los sistemas de levadura, que son diferentes de las células animales, las modificaciones postraduccionales, tales como la modificación con cadenas sacáridas, resultan difíciles de llevar a cabo en muchos casos y, de esta manera, provocan un problema para la producción de proteína que presente su actividad.

Debido a que la proteína producida se somete a una modificación postraducciona, tal como la fosforilación y la adición de cadenas sacáridas en el sistema de insecto, este sistema presenta el mérito de que puede expresarse una proteína que presenta su actividad fisiológica original. Sin embargo, debido a que la estructura de la cadena sacárida de la proteína secretada es diferente de la de las células derivadas de mamífero, la antigenicidad y similares se convierten en un problema al aplicar la proteína al uso farmacéutico.

Además, debido a que se utiliza un virus recombinante en el sistema de células de insecto al introducir un gen exógeno, existe el problema de que, desde el punto de vista de la seguridad, resulta necesaria la inactivación y contención del virus.

En el sistema de células animales, pueden llevarse a cabo modificaciones postraduccionales, tales como la fosforilación, la adición de cadenas sacáridas y el plegamiento, en las proteínas derivadas de animales superiores, incluyendo el ser humano, de una manera más similar a la que se produce en el cuerpo in vivo. Resultan necesarias estas modificaciones postraduccionales exactas para recrear la actividad fisiológica que presenta una proteína en su forma de proteína recombinante, y habitualmente se aplica un sistema de producción de proteínas en el que se utiliza una célula de mamífero como hospedadora a los productos farmacéuticos y similares que requieren dicha actividad fisiológica.

Sin embargo, un sistema de expresión de proteínas en el que se utilice una célula de mamífero como hospedadora generalmente presenta una productividad baja y también conlleva en muchos casos un problema de estabilidad de los genes introducidos. La mejora de la productividad de una proteína utilizando una célula de mamífero en cultivo como hospedadora no sólo resulta muy importante para producir medicamentos para el tratamiento, agentes diagnósticos y similares, sino que también contribuye en gran medida a la investigación y desarrollo de los mismos. De esta manera, resulta urgente desarrollar un sistema de expresión génica que posibilite de manera sencilla la obtención de una línea celular de elevada productividad utilizando una célula de mamífero en cultivo, en particular una célula de ovario de hámster chino (célula CHO), como hospedadora.

Un transposón es un elemento genético transponible que puede desplazarse de un locus a otro en el cromosoma. Un transposón es una herramienta potente para el estudio de la biología y genética moleculares y se utiliza para un propósito, tal como la mutagénesis, la captura génica y la preparación de individuos transgénicos, en insectos o nemátodos (por ejemplo, *Drosophila melanogaster* o *Caenorhabditis elegans*) y

plantas. Sin embargo, el desarrollo de dicha técnica se ha retrasado para los animales vertebrados, incluyendo las células de mamífero.

Sin embargo, en los últimos años, se ha informado de transposones que presentan actividades también en animales vertebrados, y algunos de ellos han demostrado una actividad en células de mamífero, tales como las células derivadas de ratones y seres humanos. Entre los ejemplos típicos se incluyen los transposones Tol1 (referencia de patente nº 1) y Tol2 (referencia no de patente nº 1), que se clonan a partir del pez-arroz japonés ("medaka"); "Sleeping Beauty", reconstruido a partir de un transposón no autónomo existente en el genoma de los peces del género *Onchorhynchus* (referencia no de patente nº 2), el transposón artificial "Frog prince" (referencia de no de patente nº 3), que se deriva de una rana, y el transposón piggyBac (referencia no de patente nº 4), que se deriva de un insecto.

Estos transposones de ADN han sido utilizados para mutagénesis, captura génica, preparación de individuos transgénicos, expresión de proteínas resistentes a fármacos y similares, como herramienta de introducción de genes para producir un nuevo fenotipo en un genoma de una célula de mamífero (referencias no de patente nº 5 a nº 12).

En el caso de los insectos, se ha estudiado un procedimiento en el que se introduce un gen exógeno en el cromosoma del gusano de la seda utilizando el transposón piggyBac derivado de un insecto lepidóptero para expresar la proteína codificada por dicho gen exógeno y se ha dado a conocer un procedimiento de producción de proteínas utilizando las técnicas anteriormente indicadas (referencia de patente nº 2).

Sin embargo, debido a que la proteína de interés no se expresa a niveles suficientes y se produce en todo el cuerpo del gusano de la seda, provoca un problema económico debido a la necesidad de una técnica de purificación avanzada para recuperar la proteína exógena expresada en una forma altamente purificada a partir del líquido corporal que incluye una gran cantidad de proteínas contaminantes.

Además, se conoce un ejemplo en el que una proteína relacionada con la resistencia a G418 se expresa en una célula de mamífero utilizando el transposón Tol2 derivado de medaka (referencia no de patente nº 12).

Como procedimiento para cribar eficientemente células de nivel elevado de expresión, es conocida la atenuación de un gen marcador seleccionable. Como procedimiento para la atenuación, es conocida la modificación de aminoácidos en un gen de resistencia a la neomicina (referencias no de patente nº 13 y nº 14) y la unión de una secuencia de desestabilización al gen *dhfr* (referencia no de patente nº 15). Alternativamente, se ha demostrado que pueden obtenerse células de elevado nivel de expresión mediante la utilización de un gen marcador seleccionable atenuado.

Por otra parte, también se demuestra que el número de células resistentes a fármacos se reduce drásticamente mediante la atenuación y que, como resultado, existe la posibilidad de que no se obtenga ninguna célula resistente a fármaco. De esta manera, todavía se desea la creación de un procedimiento para cribar eficientemente las células de alto nivel de expresión.

Es conocido que en los genes codificantes de proteína, existe un sesgo en el uso de los codones según la especie y que la expresión de la eritropoyetina humana en una célula de CHO se mejora mediante la optimización de este sesgo de los codones (referencia no de patente nº 16).

Se ha descrito un procedimiento para la generación rápida de líneas celulares de alta expresión estables para la producción de proteínas recombinantes con una elevada eficacia de la integración estable utilizando simultáneamente una presión selectiva baja durante sólo un periodo de tiempo corto, que utiliza una transposasa de piggyBac expresada transitoriamente para mediar en la integración estable de un transgén flanqueado por extremos del transposón PB (literatura de patentes nº 3).

Se ha descrito un rendimiento mejorado de proteínas recombinante utilizando el marcador seleccionable DHFR de codones desoptimizados en un plásmido de expresión CHEF I (literatura no de patentes nº 17).

[Listado de referencias]

[Literatura de patentes]

- [Literatura de patentes nº 1] Documento nº WO 2008/072540
- [Literatura de patentes nº 2] solicitud no examinada publicada de patente japonesa nº 2001-532188
- [Literatura de patentes nº 3] solicitud publicada de patente US nº US 2010/0311116.

[Literatura no de patentes]

- [Literatura no de patentes nº 1] Nature 383:30, 1996.

- [Literatura no de patentes nº 2] Cell 91:501-510, 1997.  
 [Literatura no de patentes nº 3] Nucleic Acids Res. 31:6873-6881, 2003.  
 [Literatura no de patentes nº 4] Insect Mol. Biol. 5:141-151, 1996.  
 [Literatura no de patentes nº 5] Genetics 166:895-899, 2004.  
 5 [Literatura no de patentes nº 6] PLoS Genet. 2:e169, 2006.  
 [Literatura no de patentes nº 7] Proc. Natl. Acad. Sci. USA 95:10769-10773, 1998.  
 [Literatura no de patentes nº 8] Proc. Natl. Acad. Sci. USA 98:6759-6764, 2001.  
 [Literatura no de patentes nº 9] Nature 436:221-226, 2005.  
 [Literatura no de patentes nº 10] Nucleic Acids Res. 31:6873-6881, 2003.  
 10 [Literatura no de patentes nº 11] Nucleic Acids Res. 35:e87, 2007.  
 [Literatura no de patentes nº 12] Proc. Natl. Acad. Sci. USA 103:15008-15013, 2006.  
 [Literatura no de patentes nº 13] Biotech. Bioeng. 89:530-538, 2005.  
 [Literatura no de patentes nº 14] Journal of Immunological Methods 295:49-56, 2004.  
 [Literatura no de patentes nº 15] Metabolic Engineering 9:304-316, 2007.  
 15 [Literatura no de patentes nº 16] Gene 199:293-301, 1997.  
 [Literatura no de patentes nº 17] Biotech. Prog. 26(6):1558-1566, 2010.

### Exposición de la invención

#### 20 Problemas que debe resolver la invención

Con el fin de producir y analizar una proteína de interés, resulta necesario seleccionar una línea celular que exprese establemente y a nivel elevado una proteína de interés, utilizando una célula de cultivo derivada de un mamífero. Sin embargo, la preparación y el cultivo de la célula que produce la proteína de interés requiere un esfuerzo y tiempo considerables.

Además, aunque es conocido que las proteínas se expresan en la célula de mamífero utilizando una secuencia de transposón, no se conoce la preparación de una célula con un nivel elevado de expresión de una proteína y que, de esta manera, pueda utilizarse como sistema de producción de la proteína mediante la utilización de una secuencia de transposón, ni una célula de producción elevada que comprenda una secuencia de transposón ni un procedimiento de producción de una proteína utilizando la célula. Además, no se conoce ningún ejemplo de obtención de una célula de nivel elevado de expresión mediante la modificación de un codón para suprimir la expresión (traducción) de un gen de resistencia a fármaco.

Tal como se ha indicado anteriormente, se ha necesitado la expresión de una proteína de interés en una gran cantidad mediante el establecimiento de un sistema de producción de proteínas que pueda producir a nivel elevado una proteína de interés utilizando una célula de mamífero en cultivo de manera eficiente y en un corto periodo de tiempo.

De esta manera, los objetivos de la presente invención son proporcionar una célula capaz de expresar a nivel elevado una proteína de interés que pueda establecerse eficientemente y un procedimiento para producir la proteína de interés utilizando la célula.

#### 45 Medios para resolver los problemas

Con el fin de resolver los problemas anteriormente indicados, los presentes inventores han llevado a cabo estudios intensivos y como consecuencia han encontrado que puede producirse eficientemente una célula de producción que exprese a nivel elevado una proteína de interés mediante la introducción de un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable atenuado y que también comprende un par de secuencias de transposón en ambos extremos del fragmento génico, en una célula de mamífero en suspensión, y la integración del fragmento génico insertado entre el par de secuencias de transposón en un cromosoma de la célula de mamífero en suspensión. Además, se ha encontrado que puede reducirse drásticamente el tiempo para preparar una línea celular de alta expresión de la proteína de interés y, de esta manera, se ha llevado a cabo la invención. Por lo tanto, el objetivo de la presente invención es proporcionar un nuevo procedimiento de preparación de una célula de producción que puede preparar eficientemente la célula de producción que expresa a nivel elevado un gen exógeno y un procedimiento de producción de una proteína recombinante.

Específicamente, la presente invención se refiere a lo siguiente:

1. Un procedimiento para producir una proteína de interés, que comprende: introducir un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable atenuado y además comprende un par de secuencias de transposón en ambos extremos del fragmento génico, en una célula de CHO en suspensión que es capaz de sobrevivir y proliferar en un medio sin suero, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador

- seleccionable antes de la modificación y para comprender codones utilizados a una frecuencia baja en la célula de CHO, de manera que el nivel de expresión en la célula de CHO se ha reducido, y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2; integrar el fragmento génico que comprende el ADN codificante de la proteína de interés insertado entre el par de secuencias de transposón en un cromosoma de la célula de CHO; obtener una célula de CHO que expresa la proteína de interés y cultivar en suspensión la célula de mamífero.
- 5
2. Un procedimiento para producir una proteína de interés, que comprende las etapas (A) y (B) a continuación:
- 10
- (A) una etapa de introducción simultánea de los vectores de expresión (a) y (b) siguientes en una célula de CHO en suspensión que es capaz de sobrevivir y proliferar en un medio sin suero; integrar un fragmento génico insertado entre un par de secuencias de transposón en un cromosoma de la célula de CHO mediante una transposasa de expresión transitoria; y obtener una célula de CHO en suspensión que expresa la proteína de interés:
- 15
- (a) un vector de expresión que comprende el fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable atenuado y que comprende además el par de secuencias de transposón en ambos extremos del fragmento génico, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender codones utilizados a una frecuencia baja en la célula de CHO de manera que el nivel de expresión en la célula de CHO se reduce y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2,
- 20
- (b) un vector de expresión que comprende un ADN codificante de la transposasa que reconoce las secuencias de transposón y presenta actividad de transferencia del fragmento génico insertado entre el par de secuencias de transposón en el cromosoma,
- 25
- (B) una etapa de cultivo en suspensión de la célula de CHO en suspensión que expresa la proteína de interés con el fin de producir la proteína de interés.
- 30
3. El procedimiento descrito en el ítem 1 o 2, anteriormente, en el que la célula de CHO en suspensión es cualquiera de las células seleccionada de entre CHO-K1, CHO-K1SV, DUKXB11, CHO/DG44, Pro-3 y CHO-2.
- 35
4. El procedimiento descrito en cualquiera de los ítems 1 a 3, anteriormente, en el que el gen marcador seleccionable se modifica en 10% o más de la secuencia de nucleótidos codificante del gen marcador seleccionable antes de la modificación.
- 40
5. El procedimiento descrito en cualquiera de los ítems 1 a 4, anteriormente, en el que el gen marcador seleccionable es modificado de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA de entre los codones correspondientes a residuo de leucina incluidos en el gen.
- 45
6. El procedimiento descrito en cualquiera de los ítems 1 a 5, anteriormente, en el que el gen marcador seleccionable es modificado de manera que 70% o más de los codones correspondientes al residuo de alanina son GCG de entre los codones correspondientes a residuo de alanina incluidos en el gen.
- 50
7. El procedimiento descrito en cualquiera de los ítems 1 a 6, anteriormente, en el que el gen marcador seleccionable es modificado de manera que todos los codones correspondientes al residuo de leucina incluidos en el gen son TTA o todos los codones correspondientes a residuo de alanina incluidos en el gen son GCG.
- 55
8. El procedimiento descrito en cualquiera de los ítems 1 a 7, anteriormente, en el que el gen marcador seleccionable es un gen marcador seleccionable seleccionado de entre el grupo que consiste en gen de resistencia a neomicina, gen de resistencia a puromicina, gen de resistencia a higromicina, un gen de resistencia a zeocina y un gen de resistencia a blasticidina.
- 60
9. El procedimiento descrito en cualquiera de los ítems 1 a 8, anteriormente, en el que las secuencias de nucleótidos derivadas de un par de transposones Tol2 son la secuencia de nucleótidos mostrada en SEC ID nº 2 y la secuencia de nucleótidos mostrada en SEC ID nº 3.
- 65
10. El procedimiento descrito en cualquiera de los ítems 1 a 8, anteriormente, en el que las secuencias de nucleótidos derivadas de un par de transposones Tol1 son las secuencias de nucleótidos mostradas en SEC ID nº 35 y la secuencia de nucleótidos mostrada en SEC ID nº 36.
11. Una célula de CHO en suspensión que es capaz de sobrevivir y proliferar en un medio sin suero, en la que se introduce un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de

- una proteína de interés y un gen marcador seleccionable atenuado y que comprende además un par de secuencias de transposón en ambos extremos del fragmento génico, en la que el gen marcador seleccionable atenuado es un gen marcador seleccionado modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender codones utilizados a una frecuencia baja en la célula de CHO de manera que se reduce el nivel de expresión en la célula de CHO y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2, y en el que el fragmento génico insertado entre el par de secuencias de transposón se integra en un cromosoma de la célula de CHO en suspensión y la célula de mamífero en suspensión produce la proteína de interés.
12. Una célula de CHO en suspensión que es capaz de sobrevivir y proliferar en un medio sin suero, que presenta un cromosoma en el que se ha integrado un fragmento génico insertado entre un par de transposones y que produce una proteína de interés obtenible mediante la introducción simultánea de los vectores (a) y (b) a continuación:
- (a) un vector de expresión de proteínas que comprende un fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable atenuado y que comprende además el par de secuencias de transposón en ambos extremos del fragmento génico, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender codones utilizados a una frecuencia baja en la célula de CHO de manera que se reduce el nivel de expresión en la célula de CHO y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposiciones Tol1 o Tol2,
- (b) un vector de expresión que comprende un ADN codificante de una transposasa que reconoce las secuencias de transposón y presenta actividad de transferencia del fragmento génico insertado entre el par de secuencias de transposón al interior del cromosoma.
13. La célula de CHO descrita en el ítem 11 o 12, anteriormente, que es cualquiera de las células seleccionada de entre CHO-K1, CHO-K1SV, DUKXB11, CHO/DG44, Pro-3 y CHO-S.
14. La célula de CHO descrita en cualquiera de los ítems 11 a 13, anteriormente, en la que el gen marcador seleccionable se modifica en 10% o más de la secuencia de nucleótidos codificante del gen marcador seleccionable antes de la modificación.
15. La célula de CHO descrita en cualquiera de los ítems 11 a 14, anteriormente, en la que el gen marcador seleccionable se modifica de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA de entre los codones correspondientes al residuo leucina incluidos en el gen.
16. La célula de CHO indicada en cualquiera de los ítems 11 a 15, anteriormente, en la que el gen marcador seleccionable se modifica de manera que 70% o más de los codones correspondientes al residuo de alanina son GCG de entre los codones correspondientes al residuo alanina incluidos en el gen.
17. La célula de mamífero indicada en cualquiera de los ítems 11 a 16, anteriormente, en la que el gen marcador seleccionable se modifica de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA o todos los codones correspondientes al residuo de alanina incluidos en el gen son GCG.
18. La célula de CHO indicada en cualquiera de los ítems 11 a 17, anteriormente, en la que el gen marcador seleccionable es un gen marcador seleccionable seleccionado de entre el grupo que consiste en un gen de resistencia a la neomicina, un gen de resistencia a la puromicina, un gen de resistencia a la higromicina, un gen de resistencia a la zeocina y un gen de resistencia a la blasticidina.
19. La célula de CHO indicada en cualquiera de los ítems 11 a 17, anteriormente, en la que las secuencias de nucleótidos derivadas de un par de transposones Tol2 son la secuencia de nucleótidos mostrada en SEC ID nº 2 y la secuencia de nucleótidos mostrada en SEC ID nº 3.
20. La célula de CHO indicada en cualquiera de los ítems 11 a 17, anteriormente, en la que las secuencias de nucleótidos derivadas de un par de transposones Tol2 son las secuencias de nucleótidos mostradas en SEC ID nº 35 y la secuencia de nucleótidos mostrada en SEC ID nº 36.
21. Un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de una proteína de interés y un marcador seleccionable atenuado, y comprende además un par de secuencias de transposón en ambos extremos del fragmento génico, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos como gen marcador seleccionable antes de la modificación y para comprender codones utilizados a baja frecuencia en la célula de CHO de manera que el nivel de expresión en la célula de CHO se reduzca y en la que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2.

22. El vector de expresión indicado en el ítem 21, anteriormente, en el que el par de secuencias de transposón derivadas de un par de transposones Tol2 son la secuencia de nucleótidos mostrada en la SEC ID nº 2 y la secuencia de nucleótidos mostrada en la SEC ID nº 3 o el par de secuencias de transposón derivadas de un par de transposones Tol1 son la secuencia de nucleótidos mostrada en la SEC ID nº 35 y la secuencia de nucleótidos mostrada en la SECID nº 36.

23. El vector indicado en el ítem 21 o 22, anteriormente, en el que el gen marcador seleccionable se modifica en 10% o más de la secuencia de nucleótidos codificante del gen marcador seleccionable antes de la modificación.

24. El vector indicado en cualquiera de los ítems 21 a 23, anteriormente, en el que el marcador seleccionable se modifica de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA de entre los codones correspondientes al residuo leucina incluidos en el gen.

25. El vector indicado en cualquiera de los ítems 21 a 24, anteriormente, en el que el gen marcador seleccionable se modifica de manera que 70% o más de los codones correspondientes al residuo de alanina son GCG de entre los codones correspondientes al residuo alanina incluidos en el gen.

26. El vector indicado en cualquiera de los ítems 21 a 25, anteriormente, en el que el gen marcador seleccionable modificado de manera que el nivel de expresión en la célula de mamífero ha sido reducido se modifica de manera que todos los codones correspondientes al residuo de leucina incluidos en el gen son TTA o la totalidad de los codones correspondientes al residuo de alanina incluidos en el gen son GCG.

27. El vector indicado en cualquiera de los ítems 21 a 26, anteriormente, en el que el gen marcador seleccionable es un gen marcador seleccionable seleccionado de entre el grupo que consiste en un gen de resistencia a neomicina, un gen de resistencia a puromicina, un gen de resistencia a higromicina, un gen de resistencia a zeocina y un gen de resistencia a blasticidina.

### Efecto de la invención

Según el procedimiento de producción de proteínas de la presente invención, puede producirse eficientemente una proteína de interés mediante la utilización de una célula de mamífero. La célula de la presente invención puede utilizarse como célula de producción de proteínas para la producción de una proteína recombinante.

### Breve descripción de los dibujos

[Figura 1] La figura 1 muestra la estructura del vector de expresión A de anticuerpos. En la figura 1, Tol2-L representa un fragmento de ADN que comprende la secuencia de Tol2-L (SEC ID nº 2) y Tol2-R representa un fragmento de ADN que comprende la secuencia de Tol2-R (SEC ID nº 3), CMV representa un promotor del CMV, poli A representa un sitio de poliadenilación, Hc representa un gen de cadena pesada del anticuerpo de CD98, Lc representa un gen de cadena ligera de anticuerpo anti-CD98 humano, SO representa un promotor de SV40, SV representa un sitio de poliadenilación de SV40 y Neo-r representa un gen de resistencia a neomicina.

### Formas de realización para poner en práctica la invención

La presente invención se refiere a un procedimiento para producir una proteína de interés, que comprende introducir un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable y que comprende además un par (dos) secuencias de transposón en ambos extremos del fragmento génico, en una célula de mamífero en suspensión; integrar el fragmento génico insertado entre el par de secuencias de transposón en un cromosoma de la célula de mamífero; obtener una célula de mamífero en suspensión que produce la proteína de interés y cultivar en suspensión la célula de mamífero, y una célula de mamífero en suspensión que expresa la proteína de interés mediante el procedimiento.

Entre los ejemplos de la célula productora de una proteína de interés de la presente invención se incluyen una célula de mamífero en suspensión, en la que se ha introducido un vector de expresión que comprende un fragmento génico que comprende un ADN codificante de una proteína de interés y un gen marcador seleccionable y que comprende además un par de secuencias de transposón en ambos extremos del fragmento génico, el fragmento génico insertado entre el par de secuencias de transposón se integra en un cromosoma y la célula de mamífero en suspensión produce la proteína de interés.

Además, entre los ejemplos de la célula productora de una proteína de interés de la presente invención se incluye una célula de mamífero en suspensión que presenta un cromosoma en el que se integra un fragmento génico insertado entre un par de transposones y que produce la proteína de interés obtenible mediante la introducción simultánea de los vectores (a) y (b), a continuación:

(a) un vector de expresión que comprende el fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable y que comprende además el par de secuencias de transposón en ambos extremos del fragmento génico,

5

(b) un vector de expresión que comprende un ADN codificante de una transposasa que reconoce las secuencias de transposón y presenta actividad de transferencia del fragmento génico insertado entre el par de secuencias de transposón al interior del cromosoma.

10 Entre los ejemplos del procedimiento para producir una proteína de interés de la presente invención se incluyen un procedimiento para producir una proteína de interés, que comprende las etapas (A) y (B) a continuación:

15 (A) una etapa de introducción simultánea de los vectores (a) y (b) proporcionados a continuación, en una célula de mamífero en suspensión y obtención de una célula de mamífero en suspensión que expresa la proteína de interés mediante la integración de un fragmento génico insertado entre un par de secuencias de transposón en un cromosoma de la célula de mamífero mediante una transposasa expresada transitoriamente:

20 (a) un vector de expresión que comprende el fragmento génico que comprende un ADN codificante de la proteína de interés y un gen marcador seleccionable y que comprende además el par de secuencias de transposón en ambos extremos del fragmento génico,

25 (b) un vector de expresión que comprende un ADN codificante de la transposasa que reconoce las secuencias de transposón y presenta actividad de transferencia del fragmento génico insertado entre el par de secuencias de transposón al interior del cromosoma, y

(B) una etapa de cultivo en suspensión de la célula de mamífero en suspensión que expresa la proteína de interés para producir la proteína de interés.

30 Entre los términos utilizados en la presente memoria se incluyen las definiciones siguientes.

El transposón es un elemento genético transponible y el término transposón se refiere a una unidad génica que se desplaza en un cromosoma o de un cromosoma a otro cromosoma (transposición) manteniendo simultáneamente una determinada estructura.

35

El transposón presenta secuencias de transposón repetidas (también denominadas secuencia repetida invertida (secuencia RI) o secuencia repetida invertida terminal (secuencia RIT)) que se sitúa en la misma dirección o en la dirección contraria en ambos extremos de una unidad génica y una secuencia de nucleótidos codificante de una transposasa que reconoce la secuencia de transposón para introducir un gen existente entre las secuencias de transposón.

40

La transposasa traducida a partir del transposón puede introducir un ADN mediante el reconocimiento de las secuencias de transposón de ambos extremos del transposón, extrayendo el fragmento de ADN insertado entre el par de secuencias de transposón e insertando el fragmento en el sitio que debe introducirse.

45

La expresión secuencia de transposón se refiere a la secuencia de nucleótidos de un transposón reconocido por una transposasa y presenta el mismo significado que la secuencia RI o secuencia RIT. La secuencia puede comprender una fracción repetida imperfecta con la condición de que pueda introducirse (insertarse en otra posición del genoma) por la actividad de una transposasa y exista una secuencia de transposón específica de una transposasa.

50

La secuencia de transposón que debe ser utilizada en la presente invención puede ser cualquier secuencia con la condición de que sea una secuencia de nucleótidos derivada de los transposones Tol1 y Tol2 que pueda ser reconocida por una transposasa y ser transpuesta en células de mamífero. Entre los ejemplos de la misma se incluyen los transposones Tol1 y Tol2 derivados del pez medaka.

55

En particular, de entre ellas, resultan preferentes las secuencias de nucleótidos derivadas del transposón Tol2 derivado del pez medaka que comprenden la secuencia de nucleótidos mostrada en SEC ID nº 6. Entre los ejemplos de secuencia de nucleótidos derivada de un par de transposones Tol2 se incluye la secuencia de nucleótidos que comprende una secuencia de nucleótidos en las posiciones 1 a 2.229 y la secuencia de nucleótidos en las posiciones 4.148 a 4.682 en la secuencia de nucleótidos del transposón Tol2 mostrada en SEC ID nº 6 del Listado de secuencias.

60

Como secuencia de nucleótidos derivada de un par de transposones Tol2, resultan más preferentes la secuencia de nucleótidos en las posiciones 1 a 200 (SEC ID nº 2) (en lo sucesivo denominada "secuencia de Tol2-L") y la secuencia de nucleótidos en las posiciones 2.285 a 2.788 (SEC ID nº 3) (en lo sucesivo denominada "secuencia

65



de Tol2-R") en la secuencia de nucleótidos de transposón Tol2 mostrada en SEC ID nº 1 del Listado de secuencias.

5 Como secuencia de transposón de la presente invención, puede utilizarse la secuencia de nucleótidos derivada del transposón Tol1 derivado del pez medaka que consiste en la secuencia de nucleótidos mostrada en SEC ID nº 37 del Listado de secuencias. Entre los ejemplos de secuencia de nucleótidos derivada de un par de transposones Tol1 se incluye una secuencia de nucleótidos en las posiciones 1 a 157 y una secuencia de nucleótidos en las posiciones 1.748 a 1.855 en la secuencia de nucleótidos derivada del transposón Tol1 derivado del pez medaka que consiste en la secuencia de nucleótidos mostrada en SEC ID nº 37 del Listado de secuencias.

15 Como secuencia de nucleótidos derivada de un par de transposones Tol1, resulta más preferente la región en las posiciones 1 a 200 (SEC ID nº 35) (en lo sucesivo denominada "secuencia de Tol1-L") y la región en las posiciones 1.351 a 1.855 (SEC ID nº 36) (en adelante denominada "secuencia de Tol1-R") en la secuencia de nucleótidos derivada del transposón Tol1 que consiste en la secuencia de nucleótidos mostrada en SEC ID nº 37 del Listado de secuencias.

20 Entre los ejemplos de la secuencia de transposón de la presente invención se incluyen secuencias de transposón las reacciones de transposón de las cuales se controlan mediante la utilización de una secuencia parcial de una secuencia de transposón específica para el transposón anteriormente indicado, mediante el ajuste de la longitud de la secuencia de nucleótidos y mediante la modificación de la secuencia de nucleótidos por adición, deleción o sustitución. Con respecto al control de la reacción de transposición de un transposón, la reacción de transposición puede acelerarse o suprimirse mediante la elevación o reducción del reconocimiento de la secuencia del transposón por una transposasa, respectivamente.

25 El término transposasa se refiere a un enzima que reconoce secuencias de nucleótidos que presentan secuencias de transposón y transfieren un fragmento génico existente entre las secuencias de nucleótidos en un cromosoma o del cromosoma a otro cromosoma.

30 Entre los ejemplos de la transposasa se incluyen enzimas derivados de Tol1 y Tol2 que se derivan del pez medaka.

35 Como transposasa, puede utilizarse un enzima nativo y puede utilizarse cualquier transposasa en la que una parte de sus aminoácidos ha sido sustituida, delecionada, insertada y/o añadida, con la condición de que se mantenga la misma actividad de transposición que la de la transposasa. Mediante el control de la actividad enzimática de la transposasa, puede controlarse la reacción de transposición del ADN existente entre las secuencias de transposón.

40 Con el fin de analizar si presenta o no una actividad de transposición similar a la de la transposasa, puede medirse mediante el sistema de análisis de 2 componentes dado a conocer en la solicitud no examinada y publicada de patente japonesa nº 2003-235575. En particular, puede analizarse si un elemento Tol2 no autónomo puede transferirse e insertarse o no en un cromosoma de una célula de mamífero mediante la actividad de una transposasa mediante la utilización separada de un plásmido que comprende un transposón Tol2 de transposasa Tol2 delecionada (transposón no autónomo derivado de Tol2) y un plásmido que comprende transposasa Tol2.

50 La expresión transposón no autónoma en la presente invención se refiere a un transposón que ha perdido una transposasa existente dentro del transposón y que, por lo tanto, no puede llevarse a cabo su transposición autónomamente. El transposón no autónomo puede transferir el ADN insertado entre las secuencias de transposón del transposón no autónoma al interior del cromosoma de la célula hospedadora, al permitir que una proteína transposasa, un ARNm codificante de la proteína transposasa o un ADN codificante de la proteína transposasa se encuentren presentes simultáneamente en la célula.

55 El gen de transposasa se refiere a un gen codificante de una transposasa. Con el fin de mejorar su eficiencia de expresión en una célula de mamífero, puede conectarse una secuencia que ajusta el espacio entre la secuencia de consenso de Kozak (Kozak M., Nucleic Acids Res. 12:857-872, 1984) o una secuencia de unión ribosómica, secuencia de Shine-Dalgarno y el codón de inicio, a una distancia apropiada (por ejemplo de 6 a 18 bases), con un sitio situado cadena arriba del codón de inicio de traducción ATG del gen.

60 Según la presente invención, con el fin de integrar un vector de expresión en el cromosoma de una célula hospedadora, se permite que actúe una transposasa sobre el vector de expresión. Con el fin de permitir que una transposasa actúe sobre una célula, puede inyectarse el enzima transposasa en la misma, o puede introducirse un gen con el ADN codificante de la transposasa en un vector de expresión deseado y éste transfectarse en la célula. Además, mediante la transfección de la célula con un ARN codificante de un gen de transposasa, la transposasa puede expresarse dentro de la célula.

65

El vector de expresión que puede utilizarse en la presente memoria no se encuentra particularmente limitado. Puede utilizarse cualquier vector de expresión mediante la selección opcional de entre los vectores de expresión conocidos por el experto en la materia, dependiendo de la célula hospedadora en la que se introduce el vector de expresión que comprende un gen de transposasa, el uso y similares.

5

En el caso de que la proteína de interés que comprende dos o más polipéptidos se produzca mediante el procedimiento de la presente invención, el vector de expresión puede integrarse en el cromosoma de una célula hospedadora mediante la inserción del ADN codificante de cada uno de los dos o más polipéptidos en el mismo vector o en vectores de expresión diferentes. Específicamente, puede insertarse una cadena pesada y una

10

cadena ligera de anticuerpo en diferentes vectores de expresión y el vector de expresión puede integrarse en un cromosoma de una célula hospedadora.

La transposasa puede insertarse en un vector de expresión para la expresión junto con la proteína de interés o puede insertarse en un vector diferente del vector de expresión. Puede permitirse que la transposasa actúe transitoriamente o puede permitirse que actúe continuamente, aunque resulta preferente permitir que la transposasa actúe transitoriamente con el fin de preparar una célula para la producción estable.

15

Con el fin de permitir que la transposasa actúe transitoriamente, puede insertarse, por ejemplo, un gen de transposasa en un plásmido de expresión que sea diferente del vector de expresión que presenta una proteína de interés y puede transfectarse una célula con ellos.

20

La expresión "vector de expresión" se refiere a un vector de expresión para la utilización en la introducción en una célula de mamífero. El vector de expresión utilizado en la presente invención presenta una estructura en la que se encuentra presente por lo menos un par de secuencias de transposón en ambos lados de un casete de expresión.

25

La expresión "casete de expresión" se refiere a una secuencia de nucleótidos que presenta una región de control de la expresión génica necesaria para expresar una proteína de interés y una secuencia codificante de la proteína de interés. Entre los ejemplos de la región de control de la expresión génica se incluyen un intensificador, un promotor y un terminador. El casete de expresión puede incluir un gen marcador seleccionable.

30

Puede utilizarse cualquier promotor, con la condición de que pueda funcionar en una célula animal. Entre los ejemplos se incluye un promotor de gen IT (inmediato temprano) del citomegalovirus (CMV), un promotor temprano del SV40, un promotor de retrovirus, un promotor metalotioneína, un promotor de choque térmico, un promotor SR $\alpha$ , virus de la leucemia murina de Moloney, un intensificador y similares. Además, puede utilizarse el intensificador del gen IT del CMV humano junto con el promotor.

35

El gen marcador seleccionable se refiere a un gen marcador opcional que puede utilizarse para distinguir una célula en la que se introduce un vector plásmido de una célula que no presenta el vector. Entre los ejemplos del gen marcador seleccionable se incluyen un gen de resistencia a fármaco (tal como un gen de resistencia a la neomicina, un gen DHFR, un gen de resistencia a puomicina, un gen de resistencia a blasticidina, un gen de resistencia a zeocina y un gen de resistencia a higromicina), genes marcadores de fluorescencia y bioluminiscencia (tal como la proteína fluorescente verde, GFP por sus siglas en inglés, Green Fluorescent Protein) y similares.

45

Un gen marcador seleccionable atenuado es un gen marcador seleccionable que se modifica de manera que se reduce la actividad de la proteína codificada por el gen marcador seleccionable en el interior de la célula.

50

Entre los ejemplos del gen marcador seleccionable que se modifica de manera que la actividad en la célula sea baja se incluyen: (A) un gen marcador seleccionable en el que se modifica una secuencia de aminoácidos de una proteína codificada por un gen marcador seleccionable de manera que se reduce la actividad de la proteína en la célula, o (B) un gen marcador seleccionable en el que una secuencia de nucleótidos que controla la expresión de un gen marcador seleccionable se modifica o una secuencia de nucleótidos dentro de un ORF (por sus siglas en inglés, Open Reading Frame, marco de lectura abierto) se modifica de manera que se reduce la expresión del gen marcador seleccionable.

55

Entre los ejemplos del gen marcador seleccionable en el que se modifica la secuencia de aminoácidos de la proteína codificada por el gen marcador seleccionable de manera que se reduce la actividad de la proteína en la célula se incluyen el gen de resistencia a neomicina descrito por Sauter et al. [Biotech. Bioeng. 89:530-538, 2005] o Chen et al. [Journal of Immunological Methods 295:49-56, 2004].

60

Entre los ejemplos del procedimiento para reducir el nivel de expresión de una proteína en la célula mediante la modificación de una secuencia de nucleótidos que controla la expresión del gen marcador seleccionable se incluye un procedimiento para modificar la secuencia del promotor, la secuencia del terminador, la secuencia intensificadora, la secuencia de consenso de Kozak o la secuencia de Shine-Dalgarno, que controla la expresión del gen marcador seleccionable.

65

Más específicamente, entre los ejemplos se incluye un procedimiento en el que una secuencia promotora que controla la expresión de un gen marcador seleccionable se sustituye por una secuencia promotora más débil.

5 Entre los ejemplos del procedimiento de reducción del nivel de expresión de la proteína en la célula mediante la modificación de una secuencia de nucleótidos en el ORF de un gen marcador seleccionable se incluye un procedimiento en el que se sustituye un codón en el ORF por un codón sinónimo que presenta una frecuencia todavía más baja de uso de codones en la célula.

10 Entre los ejemplos del gen marcador seleccionable atenuado de la presente invención se incluye un marcador seleccionable en el que el codón anteriormente indicado en el ORF del gen se sustituye por un codón sinónimo que presenta una frecuencia todavía más baja de uso del codón en la célula.

15 En las células de diversas especies biológicas, el codón sinónimo que presenta una frecuencia todavía más baja de uso de entre cada codón sinónimo puede seleccionarse basándose en la literatura conocida, bases de datos y similares.

20 Entre los ejemplos de dicha sustitución por un codón sinónimo con una frecuencia más baja de uso, específicamente en el caso de la célula CO, se incluye la sustitución del codón de leucina por TTA, la sustitución del codón de arginina por CGA o CGT, la sustitución del codón de alanina por GCG, la sustitución del codón de valina por GTA, la sustitución del codón de serina por TCG, la sustitución del codón de isoleucina por ATA, la sustitución del codón de treonina por ACG, la sustitución del codón de prolina por CCG, la sustitución del codón de ácido glutámico por GAA, la sustitución del codón de tirosina por TAT, la sustitución del codón de lisina por AAA, la sustitución del codón de fenilalanina por TTT, la sustitución del codón de histidina por CAT, la sustitución del codón de glutamina por CAA, la sustitución del codón de asparagina por AAT, la sustitución del codón de ácido aspártico por GAT, la sustitución del codón de cisteína por TGT y la sustitución del codón de glicina por GGT.

30 En un gen marcador seleccionable atenuado, el número de codones que debe colocarse en comparación con el gen marcador seleccionable antes de la modificación no se encuentra particularmente limitado con la condición de que la célula productora de proteínas pueda obtenerse eficientemente, aunque resulta preferente sustituir codones correspondientes a 20 o más residuos aminoácidos.

35 En un gen marcador seleccionable atenuado, el número de bases que debe modificarse en comparación con el gen marcador seleccionable antes de la modificación no se encuentra particularmente limitado, aunque resulta preferente modificar 10% o más de la secuencia de nucleótidos codificante del gen marcador seleccionable.

40 Además, en un gen marcador seleccionable atenuado, los residuos aminoácidos codificados por los codones que deben sustituirse no se encuentran particularmente limitados, aunque entre los ejemplos preferentes se incluye leucina, alanina, serina y valina.

45 En el caso de un gen marcador seleccionable atenuado, en el que en que los codones correspondientes a la leucina son sustituidos, no se encuentra particularmente limitado pero resulta preferente sustituir los codones correspondientes a 70% o más de los residuos de leucina de entre los codones correspondientes al total de residuos de leucina contenidos en el gen marcador seleccionable. Además, en el caso de un gen marcador seleccionable atenuado, en el caso de que los codones correspondientes a alanina se sustituyan de manera no particularmente limitada, resulta preferente sustituir los codones correspondientes a 70% o más de los residuos de alanina de entre los codones correspondientes al total de residuos de alanina contenidos en el gen marcador seleccionable.

50 Entre los ejemplos específicos del gen marcador seleccionable atenuado obtenido mediante dicha modificación en la que los codones se sustituyen por codones sinónimos con una frecuencia de uso más baja se incluyen un gen de resistencia a neomicina que comprende la secuencia de nucleótidos representada por SEC ID n° 9, n° 11 o n° 13, un gen de resistencia a puromicina que comprende la secuencia de nucleótidos representada por SEC ID n° 21, n° 23 o n° 25, un gen de resistencia a zeocina que consiste en la secuencia de nucleótidos representada por SEC ID n° 27 o n° 29 y un gen de resistencia a higromicina que comprende la secuencia de nucleótidos representada por SEC ID n° 31 o n° 33.

60 Además, resulta posible atenuar un gen marcador seleccionable también mediante el incremento considerable de la concentración de un fármaco en comparación con la concentración utilizada convencionalmente al seleccionar una célula resistente a fármaco en la preparación de una célula productora de anticuerpo o llevando a cabo una administración adicional antes de que el gen de resistencia a fármaco metabolice y degrade el fármaco.

65 El procedimiento para introducir el vector de expresión anteriormente indicado que comprende una secuencia de transposón, un vector plásmido para expresar una transposasa o ARN no se encuentra particularmente limitado. Entre los ejemplos se incluyen transfección con fosfato de calcio, electroporación, un procedimiento de

liposomas, un procedimiento de pistola génica, lipofección y similares. Entre los ejemplos del procedimiento para introducir directamente una transposasa en forma de una proteína se incluyen una técnica de microinyección o la introducción en la célula mediante endocitosis. La introducción génica puede llevarse a cabo mediante el procedimiento descrito en Shin Idenshi Kogaku Handbook (New Genetic Engineering Handbook), editado por Masami Muramatsu y Tadashi Yamamoto, publicado por Yodo-sha, ISBN 9784897063737.

La célula hospedadora puede ser cualquier célula de CHO con la condición de que pueda subcultivarse y expresar establemente una proteína de interés.

Entre los ejemplos de la célula hospedadora particular se incluyen la célula de ovario de hámster chino, CHO (Journal of Experimental Medicine 108:945, 1958; Proc. Natl. Acad. Sci. USA 60:1275, 1968; Genetics 55:513, 1968; Chromosom 41:129, 1973; Methods in Cell Science 18:115, 1996; Radiation Research 148:260, 1997; Proc. Natl. Acad. Sci. USA. 77:4216, 1980; Proc. Natl. Acad. Sci. 60:1275, 1968; Cell 6:121, 1975; Molecular Cell Genetics, Apéndices I y II (páginas 883 a 900)), CHO/DG44 (ATCC n° CRL-9096), CHO-K1 (ATCC n° CCL-61), DUKXB11 (ATCC n° CCL-9096), Pro-5 (ATCC n° CCL-1781), CHO-S (Life Technologies, n° de cat. 11619), Pro-3 y sublíneas de la línea celular de células de CHO.

Además, la célula hospedadora anteriormente indicado también puede utilizarse en el procedimiento de producción de proteínas de la presente invención mediante la modificación de la célula de manera que resulte adecuada para la producción de proteínas, debido a la modificación del ADN cromosómico, la introducción de un gen exógeno y similares.

Además, en la presente invención, con el fin de controlar la estructura de cadena sacárida unida a una proteína de interés que debe producirse, también puede utilizarse Lec13 que ha adquirido resistencia a lectina [Somatic Cell and Molecular Genetics 12:55, 1986] o una célula de CHO de la que se ha delecionado el gen de  $\alpha$ 1,6-fucosiltransferasa (documentos n° WO2005/35586 y n° WO2002/31140) como célula hospedadora para producir una proteína de interés de la presente invención.

La proteína de interés producida en la presente invención puede ser cualquier proteína con la condición de que pueda expresarse mediante el procedimiento de producción de una proteína utilizando un transposón no autónomo de la presente invención. En particular, entre los ejemplos de la proteína de interés se incluyen una proteína sérica humana, una hormona peptídica, un factor de crecimiento, una citoquina, un factor de coagulación sanguínea, una proteína fibrinolítica, un anticuerpo, fragmentos parciales de diversas proteínas y similares.

Entre los ejemplos de la proteína de interés producida en la presente invención se incluyen preferentemente un anticuerpo monoclonal, tal como un anticuerpo quimérico, un anticuerpo humanizado y un anticuerpo humano, una proteína de fusión de Fc, una proteína de unión a albúmina y un fragmento parcial de los mismos.

La actividad efectora de un anticuerpo monoclonal producido en la presente invención puede controlarse mediante diversos procedimientos. Entre los ejemplos de los procedimientos conocidos se incluyen un procedimiento para controlar la cantidad de fucosa (en lo sucesivo también denominada "fucosa nuclear") que se encuentra unida a N-acetilglucosamina (GlcNAc) mediante enlace  $\alpha$ -1,6 en un extremo reductor de una cadena sacárida unida mediante N de tipo complejo que se encuentra unida a la asparagina (Asn) en la posición 297 de una región Fc de un anticuerpo (documentos n° WO 2005/035586, n° WO 2002/31140 y n° WO 00/61739), un procedimiento para controlar una actividad efectora mediante la modificación del residuo o residuos aminoácidos de una región Fc del anticuerpo o similar. La actividad efectora del anticuerpo monoclonal producido en la presente invención puede controlarse mediante la utilización de cualquiera de los procedimientos.

La actividad efectora se refiere a una actividad dependiente de anticuerpo que resulta inducida por una región Fc de un anticuerpo. Como actividad efectora se conoce la citotoxicidad celular dependiente de anticuerpos (actividad ADCC), la citotoxicidad dependiente del complemento (actividad CDC), la fagocitosis dependiente de anticuerpos (actividad ADP) por células fagocíticas, tales como macrófagos o células dendríticas, y similares.

Además, mediante el control del contenido de fucosa nuclear de una cadena sacárida unida mediante N de tipo complejo de Fc de un anticuerpo monoclonal que se produce en la presente invención, puede incrementarse o reducirse la actividad efectora del anticuerpo. Como procedimiento para reducir el contenido de fucosa que se encuentra unida a una cadena sacárida unida mediante N de tipo complejo unida a Fc del anticuerpo, puede obtenerse un anticuerpo al que no se encuentra unido fucosa, mediante la expresión de un anticuerpo utilizando una célula de CHO que es deficiente en un gen codificante de  $\alpha$ 1,6-fucosiltransferasa.

El anticuerpo al que la fucosa no se encuentra unida presenta una elevada actividad ADCC. Por otra parte, como procedimiento para incrementar el contenido de fucosa que se encuentra unida a una cadena sacárida unida mediante N de tipo complejo unida a Fc de un anticuerpo, puede obtenerse un anticuerpo al que se encuentra unida fucosa, mediante la expresión de un anticuerpo utilizando una célula hospedadora en la que se introduce un gen codificante de  $\alpha$ 1,6-fucosiltransferasa. El anticuerpo al que se encuentra unido la fucosa presenta una

actividad ADCC más baja que el anticuerpo al que no se encuentra unida fucosa.

Además, mediante la modificación de uno o más residuos aminoácidos en la región Fc de un anticuerpo, puede incrementarse o reducirse la actividad ADCC o la actividad CDC. Por ejemplo, la actividad CDC de un anticuerpo puede incrementarse mediante la utilización de la secuencia de aminoácidos de la región Fc indicada en el documento nº US 2007/0148165. Además, la actividad ADCC o la actividad CDC puede incrementarse o reducirse llevando a cabo la modificación de aminoácidos indicada en las patentes US nº 6.737.056, nº 7.297.775 o nº 7.317.091.

La expresión "célula de mamífero en suspensión" en la presente invención se refiere a una célula que no se adhiere a un anclaje de cultivo celular de recubrimiento para facilitar la adhesión de las células en cultivo, tal como micropelotas, un recipiente de cultivo para el cultivo de tejidos (también denominado recipiente de cultivo de tejidos o de cultivo en adhesión y similares) y similares, y puede sobrevivir y crecer en suspensión en la solución de cultivo. Con la condición de que la célula no se adhiera al anclaje de cultivo celular, la célula puede sobrevivir y crecer en estado de células individuales en la solución de cultivo o sobrevivir y crecer en un estado de masa de células formada mediante la aglutinación de dos o más células.

Además, la célula de mamífero en suspensión que debe utilizarse en la presente invención es una célula que puede sobrevivir y crecer en un medio sin suero que no contiene suero de feto bovino (en lo sucesivo denominado FCS) y similares, en suspensión en la solución de cultivo sin adherirse al anclaje de cultivo celular. Resulta más preferente una célula de mamífero que puede sobrevivir y crecer en suspensión en un medio sin proteínas que no contiene proteínas.

El recipiente de cultivo para el cultivo de tejidos puede ser cualquiera, tal como un matraz, una placa Petri y similares, con la condición de que se haya aplicado al mismo un recubrimiento para el cultivo en adhesión. En particular, puede confirmarse que es una célula de mamífero en suspensión utilizando un matraz de cultivo de tejidos disponible comercialmente (fabricado por Greiner), un matraz de cultivo en adhesión (fabricado por Sumitomo Bakelite) y similares.

Como célula de mamífero en suspensión para la utilización en la presente invención, puede seleccionarse una célula preparada mediante la adaptación adicional de una célula que originalmente presentaba una propiedad en suspensión al cultivo en suspensión, o una célula de mamífero en suspensión preparada mediante la adaptación de una célula de mamífero adhesiva a condiciones de cultivo en suspensión. Entre los ejemplos de la célula que originalmente presentaba una propiedad en suspensión se incluyen la célula de CHO-S (fabricada por Invitrogen) y similares.

La célula de mamífero en suspensión preparada mediante la adaptación de una célula de mamífero adhesiva a condiciones de cultivo en suspensión en la presente invención puede prepararse mediante el procedimiento descrito en Mol. Biotechnol. 15(3):249-57, 2000, o mediante el procedimiento mostrado a continuación, y puede prepararse mediante el establecimiento de una célula que muestra una propiedad de proliferación y una propiedad de supervivencia similares a las presentes antes de adaptar el cultivo en suspensión o superiores a las presentes antes de la adaptación al cultivo en suspensión (J. Biotechnol. 130(3):282-90, 2007).

La expresión "similar a las condiciones antes de la adaptación al cultivo en suspensión" se refiere a que la proporción de supervivencia, la tasa de proliferación (tiempo de duplicación) y similares de la célula adaptada al cultivo en suspensión son sustancialmente iguales a las de la célula antes de la adaptación al cultivo en suspensión.

En la presente invención, entre los ejemplos del procedimiento de adaptación de una célula de mamífero adhesiva a las condiciones de cultivo en suspensión se incluye el procedimiento siguiente. El contenido en suero de un medio que contenía suero se redujo a 1/10 y se repitió el subcultivo a una densidad relativamente elevada de células. Cuando la célula de mamífero consigue sobrevivir y proliferar, el contenido en suero se reduce adicionalmente y se repite el subcultivo. Mediante este procedimiento puede prepararse una célula de mamífero en suspensión que puede sobrevivir y proliferar bajo condiciones sin suero.

Además, también puede prepararse una célula de mamífero en suspensión mediante el cultivo con la adición de un surfactante no iónico apropiado, tal como Pluronic-F68 o similar en la solución de cultivo. Un ejemplo de la célula de mamífero en suspensión en la que se adapta la célula de mamífero adhesiva a condiciones de cultivo en suspensión es una célula de CHO.

En la presente invención, la célula de CHO en suspensión preferentemente presenta una propiedad por la que, al llevar a cabo el cultivo en suspensión con  $2 \times 10^5$  células/ml, la densidad celular después del cultivo durante 3 o 4 días preferentemente es de  $5 \times 10^5$  células/ml o superior, más preferentemente de  $8 \times 10^5$  células/ml o superior, particularmente preferentemente de  $1 \times 10^6$  células/ml o superior, más preferentemente de  $1,5 \times 10^6$  células/ml o superior. Además, el tiempo de duplicación de la célula de CHO en suspensión de la presente invención preferentemente es de 48 horas o inferior, más preferentemente de 24 horas o inferior, particularmente

preferentemente de 18 horas o inferior, más preferentemente de 11 horas o inferior.

Entre los ejemplos del medio para el cultivo en suspensión se incluye medio disponible comercialmente, tal como medio CD OptiCHO (Invitrogen), medio EX-CELL 325-PF (SAFC Biosciences), medio SFM4CHO (HyClone) y similares. Además, también puede obtenerse mediante la mezcla de sacáridos, aminoácidos, vitaminas, sales metálicas y similares, los cuales resultan necesarios para el cultivo de las células de CHO.

El cultivo en suspensión puede llevarse a cabo mediante la utilización de un recipiente de cultivo que puede utilizarse para el cultivo en suspensión bajo una condición de cultivo capaz de cultivo en suspensión. Entre los ejemplos de recipiente de cultivo se incluyen una placa de 96 pocillos para el cultivo celular en suspensión (fabricada por Corning), un matraz T (fabricado por Becton Dickinson), un matraz cónico (fabricado por Corning) y similares.

Con respecto a las condiciones de cultivo, por ejemplo, puede cultivarse estáticamente en una atmósfera con 5% de CO<sub>2</sub> a una temperatura de cultivo de 37°C. También puede utilizarse un equipo de cultivo bajo agitación, tal como un equipo de cultivo para el uso exclusivo en cultivo en suspensión, por ejemplo Wave Bioreactor (fabricado por GE Healthcare Bioscience).

Con respecto a las condiciones de cultivo en suspensión para una célula de CHO utilizando el equipo Wave Bioreactor, la célula puede cultivarse mediante el procedimiento descrito en la página de Internet principal de GE Healthcare Bioscience: [http://www.gelifesciences.co.jp/tech\\_support/manual/pdf/cellcult/wave\\_03\\_16.pdf](http://www.gelifesciences.co.jp/tech_support/manual/pdf/cellcult/wave_03_16.pdf).

Además del cultivo bajo agitación, también puede utilizarse el cultivo utilizando un equipo de agitación mediante rotación, tal como un biorreactor. El cultivo utilizando un biorreactor puede llevarse a cabo mediante el procedimiento descrito en Cytotechnology 52:199-207, 2006, y similares.

La purificación de la proteína producida por la célula en cultivo se lleva a cabo mediante la separación de la proteína de interés de impurezas diferentes de la proteína de interés en una solución de cultivo u homogeneizado celular que contiene la proteína. Entre los ejemplos del procedimiento de separación se incluyen la centrifugación, la diálisis, la precipitación con sulfato amónico, la cromatografía de columna, la filtración o similares. La separación puede llevarse a cabo basándose en diferencias en las propiedades físicoquímicas de la proteína de interés y de las impurezas o en diferencias en su avidéz para el portador de columna mismo.

Como procedimiento para purificar la proteína, la purificación se lleva a cabo mediante el procedimiento descrito en Protein Experimentation Note (el primer volumen) - Extraction, Separation and Expression of Recombinant Protein (traducción de un libro de texto escrito en japonés) (editado por Masato Okada y Kaori Miyazaki, publicado por Yodo-sha, ISBN 9784897069180) y similares.

La presente invención ha sido descrita anteriormente mostrando formas de realización preferentes de la misma en aras de la fácil comprensión. En adelante, la presente invención se describe basándose en ejemplos, aunque las explicaciones anteriormente proporcionadas y los ejemplos, posteriormente, se proporcionan meramente con fines de ejemplificación y no con fines limitativos de la invención. De acuerdo con ello, el alcance de la presente invención no se encuentra limitado a las formas de realización y ejemplos que se describen específicamente en la presente memoria, sino que se encuentra limitada según las reivindicaciones únicamente.

En adelante, se muestran ejemplos para describir adicionalmente la presente memoria específicamente aunque la presente invención no se encuentra limitada a la descripción de dichos ejemplos.

Se llevaron a cabo diversas técnicas experimentales relacionadas con la recombinación descrita a continuación, tales como la clonación y similares, de acuerdo con las técnicas de ingeniería genética descritas en Molecular Cloning, 2a edición, editada por J. Sambrook, E.F. Fritschy y T. Maniatis, Current Protocols in Molecular Biology, editado por Frederick M. Ausubel et al., publicado por Current Protocols, y similares.

## Ejemplos

### **[Ejemplo 1] Preparación de un vector de transposón que expresa gen de resistencia a la neomicina y anticuerpo anti-CD98 humano**

#### (1) Preparación de un vector de transposón que expresa gen de resistencia a neomicina de tipo salvaje y anticuerpo anti-CD98 humano.

Un plásmido que comprendía un casete de expresión génica para la utilización con células de mamífero que comprendía un gen de anticuerpo humano arbitrario y un gen marcador de resistencia a fármaco insertado entre un par de secuencias de nucleótidos derivadas de Tol2 se utilizó como vector plásmido para la expresión de proteínas.

- 5 El ADN del gen que debía utilizarse se obtuvo llevando a cabo síntesis química de manera artificial basándose en una secuencia de nucleótidos convencionalmente conocida o mediante la preparación de cebadores de las secuencias de ambos extremos y llevando a cabo de esta manera una PCR utilizando una fuente de ADN apropiada. Con el fin de realizar estas últimas manipulaciones génicas, se añadió un sitio de digestión con enzima de restricción al extremo del cebador. En la secuencia de nucleótidos del transposón Tol2 no autónomo (SEC ID nº 1) dado a conocer en el documento nº JP-A-2003-235575, se utilizó una secuencia de nucleótidos en las posiciones 1 a 200 (secuencia de Tol2-L) (SEC ID nº 2) y una secuencia de nucleótidos en las posiciones 2.285 a 2.788 (secuencia de Tol2-R) (SEC ID nº 3) como las secuencias de transposón.
- 10 Se sintetizó un fragmento de ADN que comprendía la secuencia de Tol2-R y la secuencia de Tol2-L.
- 15 Se preparó un fragmento de ADN que incluía una secuencia de nucleótidos (SEC ID nº 15) que codificaba la cadena H de anticuerpo bajo el control del promotor del CMV, amplificado basándose en el vector del anticuerpo anti-CD98 humano N5KG1-Val C2IgG1NS/117L (patente japonesa nº 4324637) como casete génico de cadena pesada de anticuerpo y un fragmento de ADN que comprendía una secuencia de nucleótidos (SEC ID nº 17) que codificaba cadena ligera de anticuerpo bajo el control del promotor del SV40, amplificado basándose en el vector de anticuerpo anti-CD98 humano N5KG1-Val C2IgG1NS/117L, como casete génico de cadena ligera de anticuerpo.
- 20 Como casete génico de resistencia a la neomicina, se preparó un fragmento de ADN que comprendía un ADN que comprendía una secuencia de nucleótidos codificante de un gen de resistencia a la neomicina bajo el control de promotor del SV40 (un ADN que codifica una fosfotransferasa de neomicina que consiste en la secuencia de nucleótidos representada por SEC ID nº 7 y GenBank nº de acceso U47120.2).
- 25 Se preparó un vector de expresión A de anticuerpo anti-CD98 mediante la conexión del casete de expresión génica de cadena pesada de anticuerpo anteriormente indicado, el casete de expresión génica de cadena ligera de anticuerpo y el casete de expresión génica de resistencia a la neomicina y conectando adicionalmente sus dos extremos con un fragmento de ADN que comprendía una secuencia de Tol2-R y un fragmento de ADN que comprendía una secuencia de Tol2-L (figura 1).
- 30 (2) Preparación de vector de transposón de expresión de anticuerpo anti-CD98 humano que comprendía un gen 1 de resistencia a la neomicina de tipo modificado.
- 35 Se preparó un vector B de transposón de expresión de anticuerpo anti-CD98 humano en el que el gen de resistencia a neomicina del vector A de transposón de expresión de anticuerpo anti-CD98 humano obtenido en (1) que comprendía un gen de resistencia a neomicina de tipo salvaje se sustituyó por el gen 1 de resistencia a neomicina de tipo modificado que comprendía la secuencia de nucleótidos representada por SEC ID nº 9.
- 40 El gen 1 de resistencia a neomicina de tipo modificado codifica una secuencia de aminoácidos idéntica a la del gen de resistencia a neomicina de tipo salvaje y fue modificada para presentar una secuencia de nucleótidos en la que 167 bases correspondientes a 22% de la secuencia entera habían sido modificados. Específicamente, de entre el total de 32 residuos de leucina, los codones correspondientes a 25 residuos de leucina habían sido modificados de manera que fuesen TTA.
- 45 (3) Preparación de vector de transposón de expresión de anticuerpo anti-CD98 humano que comprendía un gen 2 de resistencia a neomicina de tipo modificado.
- 50 Un vector C de transposón de expresión de anticuerpo anti-CD98 humano en el que el gen de resistencia a neomicina del vector A de transposón de expresión de anticuerpo anti-CD98 humano obtenido en (1) que comprendía un gen de resistencia a neomicina de tipo salvaje se sustituyó por un gen 2 de resistencia a neomicina de tipo modificado que comprendía la secuencia de nucleótidos representada por SEC ID nº 11.
- 55 El gen 2 de resistencia a neomicina de tipo modificada codificaba la secuencia de aminoácidos idéntica al gen de resistencia a neomicina de tipo salvaje y presentaba una secuencia de nucleótidos en la que las 180 bases correspondientes a 23% de la secuencia entera habían sido modificadas. Específicamente, de entre el total de 32 residuos de leucina, los codones correspondientes a 28 residuos de leucina habían sido modificados de manera que fuesen TTA.
- 60 (4) Preparación de vector de transposón de expresión de anticuerpo anti-CD98 humano que presentaba un gen 3 de resistencia a neomicina de tipo modificado.
- 65 Se preparó un vector D de transposón de expresión de anticuerpo anti-CD98 humano en el que el gen de resistencia a neomicina del vector A de transposón de expresión de anticuerpo anti-CD98 humano obtenido en (1) que comprendía un gen de resistencia a neomicina de tipo salvaje se sustituyó por un gen 3 de resistencia a neomicina de tipo modificado que comprendía la secuencia de nucleótidos representada por la SEC ID nº 13.

El gen 3 de resistencia a neomicina de tipo modificado codificaba una secuencia de aminoácidos idéntica a la del gen de resistencia a neomicina de tipo salvaje y presentaba una secuencia de nucleótidos en la que 203 bases correspondientes a 26% de la secuencia entera habían sido modificados. Específicamente, de entre el total de 32 residuos de leucina, los codones correspondientes a 30 residuos de leucina se habían modificado de manera que fuesen TTA.

**[Ejemplo 2] Producción de anticuerpos por células de CHO productoras de anticuerpos que expresan gen de resistencia a neomicina de tipo modificado.**

Se prepararon células A a D productoras de anticuerpos mediante la introducción de cada uno de los vectores A a D de transposón de expresión de anti-CD98 humano preparados en los Ejemplos 1(1) a (4) en las células de CHO-K1 en suspensión junto con un vector pCAGGS-T2TP que expresa una transposasa Tol2 que comprende la secuencia de aminoácidos representada por SEC ID nº 5 [Kawakami K. y Noda T., Genetics 166:895-899, 2004].

La introducción de vectores en las células de CHO en suspensión se llevó a cabo mediante la suspensión de células de CHO ( $4 \times 10^6$  células) en 400 µl de tampón PBS y la cotransfección del vector de transposón de expresión de anticuerpo anti-CD98 humano (10 µg) y del vector de expresión de transposasa Tol2 llamado pCAGGS-T2TP (20 µg) directamente en forma de ADN circular mediante electroporación.

En este caso, el vector de expresión de transposasa Tol2 también se introdujo directamente como ADN circular con el fin de expresar transitoriamente la transposasa Tol2.

Además, como control que no utilizaba transposasa Tol2, el vector D de transposón de expresión de anticuerpo anti-CD98 humano (10 µg) del Ejemplo 1(4) se linealizó utilizando el enzima de restricción *PciI* (TAKARA BIO INC.) y después se introdujo en células de CHO-K1 en suspensión mediante electroporación.

La electroporación se llevó a cabo utilizando un electroporador [sistema Gene Pulser Xcell (fabricado por Bio-Rad)] bajo las condiciones de 300 V de voltaje y 500 µF de capacidad electrostática y a temperatura ambiente utilizando una cubeta de 4 mm de hueco (fabricada por Bio-Rad).

Tras la introducción génica mediante electroporación, se inocularon las células en cada cubeta en una placa de 96 pocillos y se cultivaron durante 3 días en un incubador de CO<sub>2</sub> utilizando un medio CD OptiCHO (Invitrogen) complementado con 5% de hidrolizado de soja.

A continuación, a partir del intercambio de medio tras 4 días de la introducción génica, se llevó a cabo el cultivo en presencia de G418 (Geneticin®, Invitrogen) mediante la adición de G418, proporcionando una concentración final de 500 µg/ml y se llevó a cabo el cultivo durante 3 semanas mientras se cambiaba el medio a intervalos de una semana.

Tras el cultivo, se determinó la expresión del anticuerpo utilizando el ensayo LANCE® (Perkin-Elmer Corp.) mediante un procedimiento de tipo sándwich al que se aplicó FRET (Transferencia por resonancia de energía fluorescente). Se muestran los resultados en la Tabla 1.

[Tabla 1]

	Células productoras de anticuerpos				Células de control
	A (tipo salvaje)	B (tipo modificado 1)	C (tipo modificado 2)	D (tipo modificado 3)	
Nivel de expresión de anticuerpos (mg/l) de células que mostraban la expresión máxima	0,5	2,0	1,6	5,1	-
Nivel de expresión medio de anticuerpos (mg/l) de las 10 células superiores	0,5	0,7	0,7	1,7	-

Tal como se muestra en la Tabla 1, los niveles de expresión de anticuerpo anti-CD98 humano de las células B a D que expresaban los genes de resistencia a neomicina de tipo modificado eran más elevados que el de la célula A que expresaba el gen de resistencia a neomicina de tipo salvaje.

En particular, en el caso de la célula D productora de anticuerpo anti-CD98 humano que expresaba el gen 3 de resistencia a neomicina de tipo modificado, se obtuvo una línea celular que mostraba un nivel de expresión 10 veces más elevado que el de la célula A productora de anticuerpo anti-CD98 humano que expresaba el gen de resistencia a neomicina de tipo salvaje.



Además, incluso al utilizar el gen 3 de resistencia a neomicina de tipo modificado, no se pudo obtener una célula que expresase el anticuerpo anti-CD98 humano a partir de una célula de control en la que no se había cotransfectado el vector de expresión de transposasa Tol2, a pesar de linearizar el vector.

5 **[Ejemplo 3] Preparación de vector de transposón que expresaba gen de resistencia a puomicina y anticuerpo anti-CD98 humano.**

10 (1) Preparación de vector de transposón de expresión de anticuerpo anti-CD98 humano que comprendía el gen 1 de resistencia a puomicina de tipo modificado.

Se preparó el vector E de transposón de expresión de anticuerpo anti-CD98 humano en el que se había sustituido el gen de resistencia a neomicina del vector A de transposón de expresión de anticuerpo anti-CD98 humano obtenido en el ejemplo 1(1), que comprendía el gen de resistencia a neomicina de tipo salvaje, por un gen 1 de resistencia a puomicina de tipo modificado que consistía en la secuencia de nucleótidos representada por SEC ID nº 21.

El gen 1 de resistencia a puomicina de tipo modificado codificaba una secuencia de aminoácidos idéntica a la del gen de resistencia a puomicina de tipo salvaje que consistía en la secuencia de nucleótidos representada por la SEC ID nº 19 (un gen de puomicin-N-acetiltransferasa, que consiste en la secuencia de nucleótidos dada a conocer en GenBank nº de acceso U07648.1) y presentaba una secuencia de nucleótidos en la que se habían modificado 17 bases, que corresponde a 3% del total de bases.

Específicamente, de entre el total de 28 residuos de alanina contenidos en el gen de resistencia a puomicina, los codones correspondientes a los 17 residuos de alanina se modificaron por GCG mediante modificación y, junto con los codones que ya eran GCG en el tipo salvaje, los codones que correspondían a todos los residuos de alanina fueron modificados a GCG.

30 (2) Preparación de vector de transposón de expresión de anticuerpo anti-CD98 humano que comprendía el gen 2 de resistencia a puomicina de tipo modificado.

Se preparó el vector F de transposón de expresión de anticuerpo anti-CD98 humano, en el que el gen de resistencia a neomicina del vector A de transposón de expresión de anticuerpo anti-CD98 humano obtenido en el Ejemplo 1(1) que comprendía el gen de resistencia a neomicina de tipo salvaje se sustituyó por el gen 2 de resistencia a puomicina de tipo modificado que comprendía la secuencia de nucleótidos representada por la SEC ID nº 23.

El gen 2 de resistencia a puomicina de tipo modificado codificaba una secuencia de aminoácidos idéntica a la del gen de resistencia a puomicina de tipo salvaje y presentaba una secuencia de nucleótidos en la que 79 bases, correspondiente a 14% del total de bases, habían sido modificadas. Específicamente, además de la modificación de los codones que correspondían a los residuos de alanina del gen 1 de resistencia a puomicina de tipo modificado, se modificaron los codones correspondientes a los residuos de alanina de manera que fuesen TTA, y los codones correspondientes a los residuos de valina se modificaron de manera que fuesen GTA y el codón de serina se modificó para que fuese TCG.

45 **[Ejemplo 4] Producción de anticuerpos por células de CHO productoras de anticuerpos que expresan gen 1 de resistencia a puomicina de tipo modificado**

Se prepararon las células E y F productoras de anticuerpos mediante la introducción del vector E de transposón de expresión de anticuerpo anti-CD98 humano del Ejemplo 3(1), que comprendía el gen 1 de resistencia a puomicina de tipo modificado, el vector F de transposón de expresión de anticuerpo anti-CD98 humano del Ejemplo 3(2) que comprendía el gen 2 de resistencia a puomicina de tipo modificado y el vector de expresión de transposasa Tol2 llamado pCAGGS-T2TP en células de CHO-K1 en suspensión.

La introducción de los vectores en las células en suspensión se llevó a cabo mediante la suspensión de las células de CHO en suspensión ( $4 \times 10^6$  células) en 400 µl de tampón PBS y cotransfectando el vector de transposón de expresión de anticuerpo anti-CD98 humano que comprendía el gen de resistencia a puomicina de tipo modificado en forma de ADN circular (10 µg) y pCAGGS-T2TP (20 µg) directamente mediante electroporación.

En este caso, el vector de expresión de transposasa Tol2 llamado pCAGGS-T2TP también se introdujo directamente en forma de ADN circular con el fin de expresar transitoriamente la transposasa Tol2.

La electroporación se llevó a cabo utilizando un electroporador (sistema Gene Pulser Xcell, fabricado por Bio-Rad) bajo las condiciones de 300 V de voltaje y 500 µF de capacidad electrostática y a temperatura ambiente, utilizando una cubeta de 4 mm de camino óptico (fabricada por Bio-Rad).

Tras la introducción génica mediante electroporación, las células en cada cubeta fueron inoculadas en una placa de 96 pocillos y cultivadas durante 3 días en un incubador de CO<sub>2</sub> utilizando un medio CD OptiCHO (Invitrogen) complementado con 5% de hidrolizado de soja.

5 A continuación, tras el cambio de medio 2 días después de la introducción génica, se llevó a cabo el cultivo durante 4 semanas añadiendo simultáneamente puromicina (P9620, Sigma-Aldrich), proporcionando una concentración final de 5 µg/ml y realizando el cambio de medio a medio que contenía puromicina a intervalos de una semana.

10 Tras el cultivo, se determinó el nivel de expresión del anticuerpo utilizando el ensayo LANCE® (Perkin-Elmer Corp.) mediante un procedimiento de tipo sándwich al que se aplicó FRET (transferencia de energía por resonancia de fluorescencia). Se muestran los resultados en la Tabla 2.

[Tabla 2]

	Células productoras de anticuerpos	
	E (Modificación 1)	F (Modificación 2)
Nivel de expresión de anticuerpo (mg/l) de células que muestran la expresión máxima	1,0	2,2
Nivel promedio de expresión de anticuerpo (mg/l) de las 10 células superiores	0,7	1,6

Tal como se muestra en la Tabla 2, la célula F productora de anticuerpos que expresaba el gen 2 de resistencia a puromicina de tipo modificado mostraba una productividad de anticuerpos dos o más veces superior a la de la célula E productora de anticuerpos, que expresaba el gen 1 de resistencia a puromicina de tipo modificado.

20 **[Ejemplo 5] Producción de anticuerpos por células de CHO productoras de anticuerpos que expresaban gen 2 de resistencia a puromicina de tipo modificado.**

25 La célula F productora de anticuerpos obtenida en el Ejemplo 4 que expresaba el gen 2 de resistencia a puromicina de tipo modificado se cultivó en un matraz cónico para producir anticuerpo anti-CD98 humano.

30 Específicamente, la célula F productora de anticuerpos se cultivó para la expansión utilizando una placa de 96 pocillos, una placa de 24 pocillos y un aplaca de 6 pocillos, en este orden. Se seleccionaron dos líneas celulares de la célula F productora de anticuerpos en la que el número de células se había incrementado en grado suficiente (líneas celulares 1 y 2) y se suspendieron en 35 ml de medio CD OptiCHO (Invitrogen) complementado con 5% de hidrolizado de soja de manera que se alcanzase una densidad celular de  $2 \times 10^5$  células/ml y se cultivaron durante 1 semana en un agitador utilizando un matraz cónico de 125 ml de capacidad (con una tapa doblada, Corning Glassworks) en una atmósfera con 5% de CO<sub>2</sub> a 37°C, produciendo de esta manera el anticuerpo anti-CD98 humano.

35 Se determinó la cantidad del anticuerpo en el medio después del cultivo mediante HPCL (Waters Associates, Inc.). Se muestran los resultados en la Tabla 3.

[Tabla 3]

	Línea celular 1	Línea celular 2
Nivel de expresión de anticuerpo (mg/l)	15,6	14,8

45 Los resultados anteriormente proporcionados demuestran que en la suspensión, las células de CHO, el gen de anticuerpo insertado entre un par de secuencias de transposón y el gen de resistencia a fármaco de tipo modificado, son introducidos eficientemente en el cromosoma del hospedador y también resultan eficaces para la selección de una célula de expresión elevada. Además, se encontró que la célula obtenida de esta manera podía cultivarse para la expansión y que resultaba posible la producción de la proteína de interés bajo condiciones de cultivo en suspensión.

50 **[Ejemplo de referencia] (1) Preparación de células de CHO en suspensión.**

55 Se desprendieron células de CHO-K1 adhesivas EC85051005 (Colección Europea de Cultivos Celulares) que habían sido cultivadas en medio  $\alpha$ -MEM (Invitrogen) que contenía 10% de suero (FCS), mediante un tratamiento de tripsina y después fueron recuperadas, seguido de un cultivo bajo agitación a 37°C en un incubador con 5% de CO<sub>2</sub> utilizando medio  $\alpha$ -MEM fresco que contenía 10% de FCS. Varios días después se confirmó el crecimiento de estas células y seguidamente se llevó a cabo un cultivo bajo agitación mediante la inoculación de las mismas en medio  $\alpha$ -MEM que contenía 5% de FCS a una densidad de  $2 \times 10^5$  células/ml, seguido del cultivo bajo agitación.

5 Varios días más después, se llevó a cabo la inoculación de manera similar utilizando medio  $\alpha$ -MEM que contenía 5% de FCS. Finalmente, se prepararon células adaptadas al cultivo en suspensión mediante la repetición del subcultivo y con el cultivo bajo agitación con el medio  $\alpha$ -MEM sin suero, confirmando que las células presentaban la misma capacidad de crecimiento que el cultivo de las mismas en presencia de suero.

La presente solicitud se basa en la solicitud de patente japonesa (n° 2010-279850), presentada el 15 de diciembre de 2010.

10 **Aplicabilidad industrial**

Según el procedimiento de producción de la proteína de la presente invención, puede producirse eficientemente una proteína de interés utilizando una célula de mamífero. La célula de la presente invención puede utilizarse como célula productora de proteínas para la producción de una proteína recombinante.

15 [Listado de secuencias]

SEC ID n° 1 - Descripción de secuencia artificial: secuencia de nucleótidos de transposón Tol2 no autónomo.

20 SEC ID n° 2 - Descripción de secuencia artificial: secuencia de Tol2-L.

SEC ID n° 3 - Descripción de secuencia artificial: secuencia de Tol2-R.

25 SEC ID n° 7 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a neomicina de tipo salvaje.

SEC ID n° 8 - Descripción de secuencia artificial: secuencia de aminoácidos codificada por el gen de resistencia a neomicina de tipo salvaje.

30 SEC ID n° 9 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a neomicina de tipo modificado.

SEC ID n° 10 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

35 SEC ID n° 11 - Descripción de secuencia artificial: secuencia de nucleótidos de gen de resistencia a neomicina de tipo modificado.

SEC ID n° 12 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

40 SEC ID n° 13 - Descripción de secuencia artificial: secuencia de nucleótidos de gen de resistencia a neomicina de tipo modificado.

SEC ID n° 14 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

45 SEC ID n° 15 - Descripción de secuencia artificial: secuencia de nucleótidos codificante de región variable de cadena pesada de anticuerpo anti-CD98 humano.

SEC ID n° 16 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

50 SEC ID n° 17 - Descripción de secuencia artificial: secuencia de nucleótidos codificante de región variable de cadena ligera de anticuerpo anti-CD98 humano.

SEC ID n° 18 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

55 SEC ID n° 19 - Descripción de secuencia artificial: secuencia de nucleótidos de gen de resistencia a puromicina de tipo salvaje.

SEC ID n° 20 - Descripción de secuencia artificial: secuencia de aminoácidos codificada por el gen de resistencia a puromicina de tipo salvaje.

60 SEC ID n° 21 - Descripción de secuencia artificial: secuencia de nucleótidos de gen de resistencia a puromicina de tipo modificado.

SEC ID n° 22 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

65 SEC ID n° 23 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a

puromicina de tipo modificado.

SEC ID nº 24 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

5 SEC ID nº 25 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a puromicina de tipo modificado.

SEC ID nº 26 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

10 SEC ID nº 27 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a zeocina de tipo modificado.

SEC ID nº 28 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

15 SEC ID nº 29 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a zeocina de tipo modificado.

SEC ID nº 30 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

20 SEC ID nº 31 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a higromicina de tipo modificado.

SEC ID nº 32 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

25 SEC ID nº 33 - Descripción de secuencia artificial: secuencia de nucleótidos del gen de resistencia a higromicina de tipo modificado.

SEC ID nº 34 - Descripción de secuencia artificial: secuencia de aminoácidos de constructo sintético.

30 **Listado de secuencias**

<110> Inter-University Research Institute Corporation Research Organization of Information and Systems  
Kyowa Hakko Kirin Co., Ltd.

35 <120> Procedimiento de producción de proteínas mediante la utilización de un gen de resistencia a fármaco atenuado

<130> W503695

40 <150> JP2010-279850

<151> 2010-12-15

<160> 37

45 <170> PatentIn versión 3.3

<210> 1

<211> 2788

50 <212> ADN

<213> Artificial

<220>

55 <223> Descripción de secuencia artificial: transposón Tol2 no autólogo

<400> 1

ES 2 656 079 T3

cagaggtgta aagtacttga gtaatthttac ttgattactg tacttaagta ttatthtttg 60  
 ggatthtttac ttactttgag tacaattaaa aatcaatact ttactthtta cttaattaca 120  
 thttthttaga aaaaaagta cthtttactc cttacaatth tatttacagt caaaaagtac 180  
 ttatthtttg gagatcactt cattctatth tcccttgcta ttaccaaacc aattgaattg 240  
 cgctgatgcc cagthtaatt taaatgttat ttattctgcc tatgaaaatc gthttcacat 300  
 tatatgaaat tggtcagaca tgttcattgg tccthtgga gtgacgtcat gtcacatcta 360  
 ttaccacaat gcacagcacc ttgacctgga aattagggaa attataacag tcaatcagtg 420  
 gaagaaaatg gaggaagtat gtgattcatc agcagctgcg agcagcacag tccaaaatca 480  
 gccacaggat caagagcacc cgtggccgta tcttcgcaa thctthttct taagtgggtg 540  
 aaataaagat tcattcaaga tgaaatgtgt cctctgtctc ccgcttaata aagaaatac 600  
 ggccttcaaa agttcgccat caaacctaag gaagcatatt gaggtaagta cattaagtat 660  
 thtgthttac tgatagthtt thttthttth thttthttth thtttggtg tgcathttth 720  
 gacgttgatg gcgcgcctth tatatgtgta gtaggcctat thtcaactaat gcatgcgatt 780  
 gacaatataa ggctcacgta ataaaatgct aaaatgcatt tgtaattggt aacgttaggt 840  
 ccacgggaaa thtggcgcct attgcagctt tgaataatca ttatcattcc gtgctctcat 900  
 tgtgttgaa thcatgcaaa acacaagaaa accaagcgag aaathttth ccaaacatgt 960  
 tgtattgtca aaacggtaac actthacaat gaggttgatt agthcatgta ttaactaaca 1020  
 ttaaataacc atgagcaata catttgthac tgtatctgtt aatctthgtt aacgttagtt 1080

ES 2 656 079 T3

aatagaaata cagatgttca ttgtttgttc atgttagttc acagtgcatt aactaatgtt 1140  
aacaagatat aaagtattag taaatgttga aattaacatg tatacgtgca gttcattatt 1200  
agttcatggt aactaatgta gttaactaac gaaccttatt gtaaaagtgt taccatcaaa 1260  
actaatgtaa tgaatcaat tcaccctgtc atgtcagcct tacagtcctg tgtttttgtc 1320  
aatataatca gaaataaaat taatgtttga ttgtcactaa atgctactgt atttctaaaa 1380  
tcaacaagta ttaacatta taaagtgtgc aattggctgc aaatgtcagt tttattaaag 1440  
ggttagttca cccaaaaatg aaaataatgt cattaatgac tcgccctcat gtcgttccaa 1500  
gcccgtaaga cctccgttca tcttcagaac acagtttaag atattttaga tttagtccga 1560  
gagctttctg tgcctccatt gagaatgtat gtacggtata ctgtccatgt ccagaagggt 1620  
aataaaaaca tcaaagtagt ccatgtgaca tcagtgggtt agttagaatt ttttgaagca 1680  
tcgaatacat tttggtccaa aaataacaaa acctacgact ttattcggca ttgtattctc 1740  
ttccgggtct gttgtcaatc cgcgttcacg acttcgcagt gacgctacaa tgctgaataa 1800  
agtcgtaggt tttgttattt ttggacccaa atgtattttc gatgcttcaa ataattctac 1860  
ctaaccctact gatgtcacat ggactacttt gatgttttta ttaccttctt ggacatggac 1920  
agtataccgt acatacattt tcagtggagg gacagaaagc tctcggacta aatctaaaat 1980  
atcttaaact gtgttccgaa gatgaacgga ggtgttacgg gcttggaacg acatgaggggt 2040  
gagtcattaa tgacatcttt tcatttttgg gtgaactaac cctttaatgc tgtaatcaga 2100  
gagtgtatgt gtaattgtta catttattgc atacaatata aatatttatt tgttgttttt 2160  
acagagaatg cacccaaatt acctcaaaaa ctactctaaa ttgacagcac agaagagaaa 2220  
gatcgggaca gatctcatat gctcggggc ccatctggcc tgtgtttcag acaccagggg 2280  
gtctctgctc acgtttcctg ctatttgcag cctctctatc aagactaata cacctcttcc 2340  
cgcatcggct gcctgtgaga ggcttttcag cactgcagga ttgcttttca gccccaaaag 2400  
agctaggctt gacactaaca attttgagaa tcagcttcta ctgaagtaa atctgaggtt 2460  
ttacaacttt gagtagcgtg tactggcatt agattgtctg tcttatagtt tgataattaa 2520  
atacaaacag ttctaaagca ggataaaacc ttgtatgcat ttcatttaat gttttttgag 2580  
attaaaagct taacaagaa tctctagttt tctttcttgc ttttactttt acttccttaa 2640  
tactcaagta caattttaat ggagtacttt tttactttta ctcaagtaag attctagcca 2700  
gatactttta cttttaattg agtaaaattt tccctaagta cttgtacttt cacttgagta 2760  
aaatttttga gtacttttta cacctctg 2788

<210> 2

<211> 200

5 <212> ADN

<213> Artificial

<220>

<223> Descripción de secuencia artificial: secuencia del transposón To12-L

10 <400> 2

ES 2 656 079 T3

cagaggtgta aagtacttga gtaatTTTtac ttgattactg tacttaagta ttatTTTTgg 60

ggatTTTTtac ttactttgag tacaattaaa aatcaatact tttactTTta cttaattaca 120

TTTTTTtaga aaaaaagta cTTTTtactc cttacaattt tatttacagt caaaaagtac 180

ttatTTTTtg gagatcactt 200

<210> 3  
 <211> 504  
 5 <212> ADN  
 <213> Artificial

<220>  
 10 <223> Descripción de secuencia artificial: secuencia del transposón To12-R

<400> 3

ctgctcaagT ttctgtctat ttgcagcctc tctatcaaga ctaatacacc tcttcccgca 60

tgggtgctgct gtgagaggct tttcagcact gcaggattgc ttttcagccc caaaagagct 120

aggcttgaca ctaacaattt tgagaatcag cttctactga agttaaactt gaggtTTTtac 180

aactttgagT agcgtgtact ggcattagat tgtctgtctt atagTTtgat aattaaatac 240

aaacagTtct aaagcaggat aaaaccttgt atgcatttca tTtaatgTtt tttgagatta 300

aaagctTaaa caagaatctc tagTtttctt tcttgctTtt actTTtactt ccttaatact 360

caagtacaat tTtaatggag tactTTTTTta cTTTtactca agtaagattc tagccagata 420

cTTTtactTt taattgagta aaatTTTccc taagtacttg tactTTtact tgagTaaaat 480

TTTTgagTac TTTTtacacc tctg 504

<210> 4  
 15 <211> 2156  
 <212> ADN  
 <213> Oryzias latipes

<220>  
 20 <221> CDS  
 <222> (85)..(2034)

<400> 4

acgtcatgTc acatctatta ccacaatgca cagcaccttg acctggaaat tagggaaatt 60

ataacagTca atcagTggaa gaaa atg gag gaa gta tgt gat tca tca gca 111

Met Glu Glu Val Cys Asp Ser Ser Ala

1 5

gct gcg agc agc aca gTc caa aat cag cca cag gat caa gag cac ccg 159

Ala Ala Ser Ser Thr Val Gln Asn Gln Pro Gln Asp Gln Glu His Pro

10 15 20 25

ES 2 656 079 T3

tgg ccg tat ctt cgc gaa ttc ttt tct tta agt ggt gta aat aaa gat	207
Trp Pro Tyr Leu Arg Glu Phe Phe Ser Leu Ser Gly Val Asn Lys Asp	
30 35 40	
tca ttc aag atg aaa tgt gtc ctc tgt ctc ccg ctt aat aaa gaa ata	255
Ser Phe Lys Met Lys Cys Val Leu Cys Leu Pro Leu Asn Lys Glu Ile	
45 50 55	
tcg gcc ttc aaa agt tcg cca tca aac cta agg aag cat att gag aga	303
Ser Ala Phe Lys Ser Ser Pro Ser Asn Leu Arg Lys His Ile Glu Arg	
60 65 70	
atg cac cca aat tac ctc aaa aac tac tct aaa ttg aca gca cag aag	351
Met His Pro Asn Tyr Leu Lys Asn Tyr Ser Lys Leu Thr Ala Gln Lys	
75 80 85	
aga aag atc ggg acc tcc acc cat gct tcc agc agt aag caa ctg aaa	399
Arg Lys Ile Gly Thr Ser Thr His Ala Ser Ser Ser Lys Gln Leu Lys	
90 95 100 105	
gtt gac tca gtt ttc cca gtc aaa cat gtg tct cca gtc act gtg aac	447
Val Asp Ser Val Phe Pro Val Lys His Val Ser Pro Val Thr Val Asn	
110 115 120	
aaa gct ata tta agg tac atc att caa gga ctt cat cct ttc agc act	495
Lys Ala Ile Leu Arg Tyr Ile Ile Gln Gly Leu His Pro Phe Ser Thr	
125 130 135	
gtt gat ctg cca tca ttt aaa gag ctg att agt aca ctg cag cct ggc	543
Val Asp Leu Pro Ser Phe Lys Glu Leu Ile Ser Thr Leu Gln Pro Gly	
140 145 150	
att tct gtc att aca agg cct act tta cgc tcc aag ata gct gaa gct	591
Ile Ser Val Ile Thr Arg Pro Thr Leu Arg Ser Lys Ile Ala Glu Ala	
155 160 165	
gct ctg atc atg aaa cag aaa gtg act gct gcc atg agt gaa gtt gaa	639
Ala Leu Ile Met Lys Gln Lys Val Thr Ala Ala Met Ser Glu Val Glu	
170 175 180 185	
tgg att gca acc aca acg gat tgt tgg act gca cgt aga aag tca ttc	687
Trp Ile Ala Thr Thr Thr Asp Cys Trp Thr Ala Arg Arg Lys Ser Phe	
190 195 200	
att ggt gta act gct cac tgg atc aac cct gga agt ctt gaa aga cat	735
Ile Gly Val Thr Ala His Trp Ile Asn Pro Gly Ser Leu Glu Arg His	
205 210 215	
tcc gct gca ctt gcc tgc aaa aga tta atg ggc tct cat act ttt gag	783
Ser Ala Ala Leu Ala Cys Lys Arg Leu Met Gly Ser His Thr Phe Glu	
220 225 230	
gta ctg gcc agt gcc atg aat gat atc cac tca gag tat gaa ata cgt	831
Val Leu Ala Ser Ala Met Asn Asp Ile His Ser Glu Tyr Glu Ile Arg	
235 240 245	
gac aag gtt gtt tgc aca acc aca gac agt ggt tcc aac ttt atg aag	879
Asp Lys Val Val Cys Thr Thr Thr Asp Ser Gly Ser Asn Phe Met Lys	
250 255 260 265	
gct ttc aga gtt ttt ggt gtg gaa aac aat gat atc gag act gag gca	927
Ala Phe Arg Val Phe Gly Val Glu Asn Asn Asp Ile Glu Thr Glu Ala	
270 275 280	



ES 2 656 079 T3

aga agg tgt gaa agt gat gac act gat tct gaa ggc tgt ggt gag gga	975
Arg Arg Cys Glu Ser Asp Asp Thr Asp Ser Glu Gly Cys Gly Glu Gly	
285 290 295	
agt gat ggt gtg gaa ttc caa gat gcc tca cga gtc ctg gac caa gac	1023
Ser Asp Gly Val Glu Phe Gln Asp Ala Ser Arg Val Leu Asp Gln Asp	
300 305 310	
gat ggc ttc gaa ttc cag cta cca aaa cat caa aag tgt gcc tgt cac	1071
Asp Gly Phe Glu Phe Gln Leu Pro Lys His Gln Lys Cys Ala Cys His	
315 320 325	
tta ctt aac cta gtc tca agc gtt gat gcc caa aaa gct ctc tca aat	1119
Leu Leu Asn Leu Val Ser Ser Val Asp Ala Gln Lys Ala Leu Ser Asn	
330 335 340 345	
gaa cac tac aag aaa ctc tac aga tct gtc ttt ggc aaa tgc caa gct	1167
Glu His Tyr Lys Lys Leu Tyr Arg Ser Val Phe Gly Lys Cys Gln Ala	
350 355 360	
tta tgg aat aaa agc agc cga tgg gct cta gca gct gaa gct gtt gaa	1215
Leu Trp Asn Lys Ser Ser Arg Ser Ala Leu Ala Ala Glu Ala Val Glu	
365 370 375	
tca gaa agc cgg ctt cag ctt tta agg cca aac caa acg cgg tgg aat	1263
Ser Glu Ser Arg Leu Gln Leu Leu Arg Pro Asn Gln Thr Arg Trp Asn	
380 385 390	
tca act ttt atg gct gtt gac aga att ctt caa att tgc aaa gaa gca	1311
Ser Thr Phe Met Ala Val Asp Arg Ile Leu Gln Ile Cys Lys Glu Ala	
395 400 405	
gga gaa ggc gca ctt cgg aat ata tgc acc tct ctt gag gtt cca atg	1359
Gly Glu Gly Ala Leu Arg Asn Ile Cys Thr Ser Leu Glu Val Pro Met	
410 415 420 425	
ttt aat cca gca gaa atg ctg ttc ttg aca gag tgg gcc aac aca atg	1407
Phe Asn Pro Ala Glu Met Leu Phe Leu Thr Glu Trp Ala Asn Thr Met	
430 435 440	
cgt cca gtt gca aaa gta ctc gac atc ttg caa gcg gaa acg aat aca	1455
Arg Pro Val Ala Lys Val Leu Asp Ile Leu Gln Ala Glu Thr Asn Thr	
445 450 455	
cag ctg ggg tgg ctg ctg cct agt gtc cat cag tta agc ttg aaa ctt	1503
Gln Leu Gly Trp Leu Leu Pro Ser Val His Gln Leu Ser Leu Lys Leu	
460 465 470	
cag cga ctc cac cat tct ctc agg tac tgt gac cca ctt gtg gat gcc	1551
Gln Arg Leu His His Ser Leu Arg Tyr Cys Asp Pro Leu Val Asp Ala	
475 480 485	
cta caa caa gga atc caa aca cga ttc aag cat atg ttt gaa gat cct	1599
Leu Gln Gln Gly Ile Gln Thr Arg Phe Lys His Met Phe Glu Asp Pro	
490 495 500 505	
gag atc ata gca gct gcc atc ctt ctc cct aaa ttt cgg acc tct tgg	1647
Glu Ile Ile Ala Ala Ile Leu Leu Pro Lys Phe Arg Thr Ser Trp	
510 515 520	
aca aat gat gaa acc atc ata aaa cga ggc atg gac tac atc aga gtg	1695
Thr Asn Asp Glu Thr Ile Ile Lys Arg Gly Met Asp Tyr Ile Arg Val	

ES 2 656 079 T3

	525		530		535	
	cat ctg gag cct ttg gac cac aag aag gaa ttg gcc aac agt tca tct					1743
	His Leu Glu Pro Leu Asp His Lys Lys Glu Leu Ala Asn Ser Ser Ser					
	540		545		550	
	gat gat gaa gat ttt ttc gct tct ttg aaa ccg aca aca cat gaa gcc					1791
	Asp Asp Glu Asp Phe Phe Ala Ser Leu Lys Pro Thr Thr His Glu Ala					
	555		560		565	
	agc aaa gag ttg gat gga tat ctg gcc tgt gtt tca gac acc agg gag					1839
	Ser Lys Glu Leu Asp Gly Tyr Leu Ala Cys Val Ser Asp Thr Arg Glu					
	570		575		580	585
	tct ctg ctc acg ttt cct gct att tgc agc ctc tct atc aag act aat					1887
	Ser Leu Leu Thr Phe Pro Ala Ile Cys Ser Leu Ser Ile Lys Thr Asn					
		590		595		600
	aca cct ctt ccc gca tgc gct gcc tgt gag agg ctt ttc agc act gca					1935
	Thr Pro Leu Pro Ala Ser Ala Ala Cys Glu Arg Leu Phe Ser Thr Ala					
		605		610		615
	gga ttg ctt ttc agc ccc aaa aga gct agg ctt gac act aac aat ttt					1983
	Gly Leu Leu Phe Ser Pro Lys Arg Ala Arg Leu Asp Thr Asn Asn Phe					
		620		625		630
	gag aat cag ctt cta ctg aag tta aat ctg agg ttt tac aac ttt gag					2031
	Glu Asn Gln Leu Leu Leu Lys Leu Asn Leu Arg Phe Tyr Asn Phe Glu					
		635		640		645
	tag cgtgtactgg cattagattg tctgtcttat agtttgataa ttaaatacaa					2084
	acagttctaa agcaggataa aaccttgtat gcatttcatt taatgttttt tgagattaaa					2144
	agcttaaaca ag					2156
	<210> 5					
	<211> 649					
5	<212> PRT					
	<213> <i>Oryzias latipes</i>					
	<400> 5					
	Met Glu Glu Val Cys Asp Ser Ser Ala Ala Ala Ser Ser Thr Val Gln					
	1		5		10	15
	Asn Gln Pro Gln Asp Gln Glu His Pro Trp Pro Tyr Leu Arg Glu Phe					
		20		25		30
	Phe Ser Leu Ser Gly Val Asn Lys Asp Ser Phe Lys Met Lys Cys Val					
		35		40		45
	Leu Cys Leu Pro Leu Asn Lys Glu Ile Ser Ala Phe Lys Ser Ser Pro					
		50		55		60
	Ser Asn Leu Arg Lys His Ile Glu Arg Met His Pro Asn Tyr Leu Lys					
		65		70		75
						80

ES 2 656 079 T3

Asn Tyr Ser Lys Leu Thr Ala Gln Lys Arg Lys Ile Gly Thr Ser Thr  
 85 90 95  
 His Ala Ser Ser Ser Lys Gln Leu Lys Val Asp Ser Val Phe Pro Val  
 100 105 110  
 Lys His Val Ser Pro Val Thr Val Asn Lys Ala Ile Leu Arg Tyr Ile  
 115 120 125  
 Ile Gln Gly Leu His Pro Phe Ser Thr Val Asp Leu Pro Ser Phe Lys  
 130 135 140  
 Glu Leu Ile Ser Thr Leu Gln Pro Gly Ile Ser Val Ile Thr Arg Pro  
 145 150 155 160  
 Thr Leu Arg Ser Lys Ile Ala Glu Ala Ala Leu Ile Met Lys Gln Lys  
 165 170 175  
 Val Thr Ala Ala Met Ser Glu Val Glu Trp Ile Ala Thr Thr Thr Asp  
 180 185 190  
 Cys Trp Thr Ala Arg Arg Lys Ser Phe Ile Gly Val Thr Ala His Trp  
 195 200 205  
 Ile Asn Pro Gly Ser Leu Glu Arg His Ser Ala Ala Leu Ala Cys Lys  
 210 215 220  
 Arg Leu Met Gly Ser His Thr Phe Glu Val Leu Ala Ser Ala Met Asn  
 225 230 235 240  
 Asp Ile His Ser Glu Tyr Glu Ile Arg Asp Lys Val Val Cys Thr Thr  
 245 250 255  
 Thr Asp Ser Gly Ser Asn Phe Met Lys Ala Phe Arg Val Phe Gly Val  
 260 265 270  
 Glu Asn Asn Asp Ile Glu Thr Glu Ala Arg Arg Cys Glu Ser Asp Asp  
 275 280 285  
 Thr Asp Ser Glu Gly Cys Gly Glu Gly Ser Asp Gly Val Glu Phe Gln  
 290 295 300  
 Asp Ala Ser Arg Val Leu Asp Gln Asp Asp Gly Phe Glu Phe Gln Leu  
 305 310 315 320  
 Pro Lys His Gln Lys Cys Ala Cys His Leu Leu Asn Leu Val Ser Ser  
 325 330 335

ES 2 656 079 T3

Val Asp Ala Gln Lys Ala Leu Ser Asn Glu His Tyr Lys Lys Leu Tyr  
 340 345 350

Arg Ser Val Phe Gly Lys Cys Gln Ala Leu Trp Asn Lys Ser Ser Arg  
 355 360 365

Ser Ala Leu Ala Ala Glu Ala Val Glu Ser Glu Ser Arg Leu Gln Leu  
 370 375 380

Leu Arg Pro Asn Gln Thr Arg Trp Asn Ser Thr Phe Met Ala Val Asp  
 385 390 395 400

Arg Ile Leu Gln Ile Cys Lys Glu Ala Gly Glu Gly Ala Leu Arg Asn  
 405 410 415

Ile Cys Thr Ser Leu Glu Val Pro Met Phe Asn Pro Ala Glu Met Leu  
 420 425 430

Phe Leu Thr Glu Trp Ala Asn Thr Met Arg Pro Val Ala Lys Val Leu  
 435 440 445

Asp Ile Leu Gln Ala Glu Thr Asn Thr Gln Leu Gly Trp Leu Leu Pro  
 450 455 460

Ser Val His Gln Leu Ser Leu Lys Leu Gln Arg Leu His His Ser Leu  
 465 470 475 480

Arg Tyr Cys Asp Pro Leu Val Asp Ala Leu Gln Gln Gly Ile Gln Thr  
 485 490 495

Arg Phe Lys His Met Phe Glu Asp Pro Glu Ile Ile Ala Ala Ala Ile  
 500 505 510

Leu Leu Pro Lys Phe Arg Thr Ser Trp Thr Asn Asp Glu Thr Ile Ile  
 515 520 525

Lys Arg Gly Met Asp Tyr Ile Arg Val His Leu Glu Pro Leu Asp His  
 530 535 540

Lys Lys Glu Leu Ala Asn Ser Ser Ser Asp Asp Glu Asp Phe Phe Ala  
 545 550 555 560

Ser Leu Lys Pro Thr Thr His Glu Ala Ser Lys Glu Leu Asp Gly Tyr  
 565 570 575

Leu Ala Cys Val Ser Asp Thr Arg Glu Ser Leu Leu Thr Phe Pro Ala

ES 2 656 079 T3

580

585

590

Ile Cys Ser Leu Ser Ile Lys Thr Asn Thr Pro Leu Pro Ala Ser Ala  
595 600 605

Ala Cys Glu Arg Leu Phe Ser Thr Ala Gly Leu Leu Phe Ser Pro Lys  
610 615 620

Arg Ala Arg Leu Asp Thr Asn Asn Phe Glu Asn Gln Leu Leu Leu Lys  
625 630 635 640

Leu Asn Leu Arg Phe Tyr Asn Phe Glu  
645

<210> 6

<211> 4682

<212> ADN

<213> *Oryzias latipes*

5

<400> 6

cagaggtgta aagtacttga gtaattttac ttgattactg tacttaagta ttatTTTTGG 60  
 ggatttttac tttacttgag tacaattaaa aatcaactact tttactttta cttaattaca 120  
 tttttttaga aaaaaaagta ctttttactc cttacaattt tatttacagt caaaaagtac 180  
 ttatTTTTtg gagatcactt cattctattt tcccttgcta ttaccaaacc aattgaattg 240  
 cgctgatgcc cagtttaatt taaatgttat ttattctgcc tatgaaaatc gttttcacat 300  
 tatatgaaat tggtcagaca tgttcattgg tcctttggaa gtgacgtcat gtcacatcta 360  
 ttaccacaat gcacagcacc ttgacctgga aattagggaa attataacag tcaatcagtg 420  
 gaagaaaatg gaggaagtat gtgattcatc agcagctgcg agcagcacag tccaaaatca 480  
 gccacaggat caagagcacc cgtggccgta tcttcgcgaa ttcttttctt taagtgggtg 540  
 aaataaagat tcattcaaga tgaaatgtgt cctctgtctc cgccttaata aagaaatc 600  
 ggccttcaaa agttcgccat caaacctaag gaagcatatt gaggtaagta cattaagtat 660  
 tttgTTTTac tgatagtttt tttttttttt tttttttttt tttttgggtg tgcattgttt 720  
 gacgttgatg gcgcgccttt tatatgtgta gtaggcctat tttcactaat gcatgcgatt 780  
 gacaatataa ggctcacgta ataaaatgct aaaatgcatt tgtaattggt aacgttaggt 840  
 ccacgggaaa tttggcgcct attgcagctt tgaataatca ttatcattcc gtgctctcat 900  
 tgtgTTtgaa ttcattgcaaa acacaagaaa accaagcgag aaatTTTTtt ccaaactgt 960  
 tgtattgtca aaacggtaac actttacaat gaggttgatt agttcatgta ttaactaaca 1020  
 ttaaataacc atgagcaata cttttgttac tgtatctggt aatctttggt aacgttaggt 1080  
 aatagaataa cagatgttca ttgTTTgttc atgttagttc acagtgcatt aactaatggt 1140

ES 2 656 079 T3

aacaagatat aaagtattag taaatgttga aattaacatg tatacgtgca gttcattatt 1200  
agttcatggt aactaatgta gttaactaac gaaccttatt gtaaaagtgt taccatcaaa 1260  
actaatgtaa tgaaatcaat tcacctgtc atgtcagcct tacagtccctg tgtttttgtc 1320  
aatataatca gaaataaaaat taatgtttga ttgtcactaa atgctactgt atttctaaaa 1380  
tcaacaagta tttaacatta taaagtgtgc aattggctgc aaatgtcagt tttattaaag 1440  
ggttagtcca cccaaaaatg aaaataatgt cattaatgac tcgccctcat gtcgttccaa 1500  
gcccgtaaga cctccgttca tcttcagaac acagtttaag atattttaga tttagtccga 1560  
gagctttctg tgcctccatt gagaatgtat gtacggata ctgtccatgt ccagaaaggt 1620  
aataaaaaca tcaaatgagt ccatgtgaca tcagtgggtt agttagaatt tttgaagca 1680  
tcgaatacat tttggtccaa aaataacaaa acctacgact ttattcggca ttgtattctc 1740  
ttccgggtct gttgtcaatc cgcgttcacg acttcgcagt gacgctacaa tgetgaataa 1800  
agtcgtaggt tttgttattt ttggaccaa atgtattttc gatgcttcaa ataattctac 1860  
ctaaccact gatgtcacat ggactacttt gatgttttta ttacctttct ggacatggac 1920  
agtataccgt acatacattt tcagtggagg gacagaaagc tctcggacta aatctaaaat 1980  
atcttaaact gtgttccgaa gatgaacgga ggtgttacgg gottggaacg acatgagggt 2040  
gagtcattaa tgacatcttt tcatttttgg gtgaactaac ctttaatgc tgtaatcaga 2100  
gagtgatgt gtaattgtta cttttattgc atacaatata aatatttatt tgttgttttt 2160  
acagagaatg cacccaaatt acctcaaaaa ctactctaaa ttgacagcac agaagagaaa 2220  
gatcgggacc tccaccatg ctccagcag taagcaactg aaagttgact cagttttccc 2280  
agtcaaacat gtgtctccag tcaactgtgaa caaagctata ttaaggtaca tcattcaagg 2340  
acttcactct ttcagcactg ttgatctgcc atcatttaaa gagctgatta gtacactgca 2400  
gcctggcatt tctgtcatta caaggcctac tttacgctcc aagatagctg aagctgctct 2460  
gatcatgaaa cagaaaagtga ctgctgccat gagtgaagtt gaatggattg caaccacaac 2520  
ggattgttg actgcacgta gaaagtcatt cattgggtga actgctcact ggatcaacc 2580  
tggaagtctt gaaagacatt ccgctgcact tgccctcaaa agattaatgg gctctcatac 2640  
ttttgaggta ctggccagtg ccatgaatga tatccactca gagtatgaaa tacgtgacaa 2700  
ggttgtttgc acaaccacag acagtgttc caactttatg aaggcttca gagtttttg 2760  
tgtgaaaac aatgatatcg agactgaggc aagaaggtgt gaaagtgatg aactgattc 2820  
tgaaggctgt ggtgagggaa gtgatgtgtt ggaattccaa gatgcctcac ggtcctgga 2880  
ccaagacgat ggcttcgaat tccagctacc aaaacatcaa aagtgtgcct gtcacttact 2940  
taacctagtc tcaagcgttg atgccccaaa agctctctca aatgaacct acaagaaact 3000  
ctacagatct gtctttggca aatgccaaagc tttatggaat aaaagcagcc gatcggctct 3060

ES 2 656 079 T3

agcagctgaa gctggtgaat cagaaagccg gcttcagctt ttaaggccaa accaaaacgcg 3120  
 gtggaattca acttttatgg ctggtgacag aattcttcaa atttgcaaag aagcaggaga 3180  
 aggcgcactt cggaaatatat gcacctctct tgaggttcca atgtaagtgt ttttccctc 3240  
 tatcgatgta aacaaatgtg ggttggtttt gttaataact ctttgattat gctgatttct 3300  
 cctgtaggtt taatccagca gaaatgctgt tcttgacaga gtgggccaac acaatgcgctc 3360  
 cagttgcaaa agtactcgac atcttgcaag cggaaacgaa tacacagctg gggtaggtgc 3420  
 tgcctagtgt ccatcagtta agcttgaaac ttcagcgact ccaccattct ctcaggtact 3480  
 gtgacccact tgtggatgcc ctacaacaag gaatccaaac acgattcaag catatgtttg 3540  
 aagatcctga gatcatagca gctgccatcc ttctccctaa atttcggacc tcttgacaa 3600  
 atgatgaaac catcataaaa cgaggtaaat gaatgcaagc aacatacact tgacgaattc 3660  
 taatctgggc aacctttgag ccataccaaa attattcttt tatttattta tttttgact 3720  
 ttttaggaat gttatatccc atctttggct gtgatctca tatgaatatt gatgtaaagt 3780  
 attcttgacag caggtttag ttatccctca gtgttcttg aaaccaaact catatgtatc 3840  
 atatgtggtt tggaaatgca gttagatttt atgctaaaat aagggatttg catgatttta 3900  
 gatgtagatg actgcacgta aatgtagtta atgacaaaat ccataaaatt tgttcccagt 3960  
 cagaagcccc tcaaccaaac ttttcttctg gtctgctcac tgtgcttcta ggcatggact 4020  
 acatcagagt gcatctggag cctttggacc acaagaagga attggccaac agttcatctg 4080  
 atgatgaaga tttttctgct tctttgaaac cgacaacaca tgaagccagc aaagagttgg 4140  
 atggatatct ggcctgtggt tcagacacca gggagtctct gctcacgttt cctgctatct 4200  
 gcagcctctc tatcaagact aatacacctc ttcccgcatc ggctgcctgt gagaggcttt 4260  
 tcagcactgc aggattgctt ttcagcccca aaagagctag gcttgacact aacaattttg 4320  
 agaatcagct tctactgaag ttaaactctga ggttttacia ctttgagtag cgtgtactgg 4380  
 cattagattg tctgtcttat agtttgataa ttaaatacaa acagttctaa agcaggataa 4440  
 aacctgtat gcatttcatt taatgttttt tgagattaaa agcttaaca agaatctcta 4500  
 gttttctttc ttgcttttac ttttacttcc ttaatactca agtacaattt taatggagta 4560  
 ctttttact tttactcaag taagattcta gccagatact tttactttta attgagtaa 4620  
 atttcccta agtacttcta ctttcaactg agtaaaattt ttgagtactt tttacacctc 4680  
 tg 4682

- <210> 7
- <211> 795
- 5 <212> ADN
- <213> desconocido
- <220>
- <223> gen de resistencia a neomicina de tipo salvaje
- 10 <220>
- <221> CDS
- <222> (1)..(795)
- <400> 7

ES 2 656 079 T3

atg att gaa caa gat gga ttg cac gca ggt tct ccg gcc gct tgg gtg 48  
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val  
1 5 10 15

gag agg cta ttc ggc tat gac tgg gca caa cag aca atc ggc tgc tct 96  
Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser  
20 25 30

gat gcc gcc gtg ttc cgg ctg tca gcg cag ggg cgc ccg gtt ctt ttt 144  
Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe  
35 40 45

gtc aag acc gac ctg tcc ggt gcc ctg aat gaa ctg cag gac gag gca 192  
Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala  
50 55 60

gcg cgg cta tcg tgg ctg gcc acg acg ggc gtt cct tgc gca gct gtg 240  
Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val  
65 70 75 80

ctc gac gtt gtc act gaa gcg gga agg gac tgg ctg cta ttg ggc gaa 288  
Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu  
85 90 95

gtg ccg ggg cag gat ctc ctg tca tct cac ctt gct cct gcc gag aaa 336  
Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys  
100 105 110

gta tcc atc atg gct gat gca atg cgg cgg ctg cat acg ctt gat ccg 384  
Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro  
115 120 125

gct acc tgc cca ttc gac cac caa gcg aaa cat cgc atc gag cga gca 432  
Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala  
130 135 140

cgt act cgg atg gaa gcc ggt ctt gtc gat cag gat gat ctg gac gaa 480  
Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu  
145 150 155 160

gag cat cag ggg ctc gcg cca gcc gaa ctg ttc gcc agg ctc aag gcg 528  
Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
165 170 175

cgc atg ccc gac ggc gag gat ctc gtc gtg acc cat ggc gat gcc tgc 576  
Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
180 185 190

ttg ccg aat atc atg gtg gaa aat ggc cgc ttt tct gga ttc atc gac 624  
Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
195 200 205

tgt gcc cgg ctg ggt gtg gcg gac cgc tat cag gac ata gcg ttg gct 672  
Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
210 215 220

acc cgt gat att gct gaa gag ctt ggc ggc gaa tgg gct gac cgc ttc 720  
Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
225 230 235 240

ctc gtg ctt tac ggt atc gcc gct ccc gat tcg cag cgc atc gcc ttc 768  
Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
245 250 255

tat cgc ctt ctt gac gag ttc ttc tga 795  
Tyr Arg Leu Leu Asp Glu Phe Phe  
260

5 <210> 8  
<211> 264



ES 2 656 079 T3

<212> PRT  
 <213> desconocido

<220>  
 5 <223> Constructo sintético

<400> 8

```

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1           5           10           15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
          20           25           30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
          35           40           45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
          50           55           60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65           70           75           80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
          85           90           95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
          100           105           110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
          115           120           125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
          130           135           140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145           150           155           160
  
```

ES 2 656 079 T3

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
 165 170 175

Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
 180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
 195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
 210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
 225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
 245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe  
 260

<210> 9  
 <211> 795  
 5 <212> ADN  
 <213> Artificial

<220>  
 10 <223> gen de resistencia a neomicina modificado

<220>  
 <221> CDS  
 <222> (1)..(795)

15 <400> 9  
 atg att gaa caa gat gga ttg cac gca ggt tct ccg gcc gct tgg gtg 48  
 Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val  
 1 5 10 15  
 gag agg cta ttc ggc tat gac tgg gca caa cag aca atc ggc tgc tct 96  
 Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser  
 20 25 30  
 gat gcc gcc gtg ttc cgg ctg tca gcg cag ggg cgc ccg gtt ctt ttt 144  
 Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe  
 35 40 45  
 gtc aag acc gac ctg tcc ggt gcc ctg aat gaa ctg caa gat gaa gcg 192  
 Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala  
 50 55 60  
 gcg cga tta tcg tgg tta gcg acg acg ggg gta ccg tgt gcg gcg gta 240  
 Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val  
 65 70 75 80

ES 2 656 079 T3

tta gat gta gta acg gaa gcg ggg cga gat tgg tta tta tta ggg gaa 288  
 Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu  
 85 90 95  
  
 gta ccg ggg caa gat tta tta tcg tcg cat tta gcg ccg gcg gaa aaa 336  
 Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys  
 100 105 110  
  
 gta tcg ata atg gcg gat gcg atg cga cga tta cat acg tta gat ccg 384  
 Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro  
 115 120 125  
  
 gcg acg tgt ccg ttt gat cat caa gcg aaa cat cga ata gaa cga gcg 432  
 Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala  
 130 135 140  
  
 cga acg cga atg gaa gcg ggg tta gta gat caa gat gat tta gat gaa 480  
 Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu  
 145 150 155 160  
  
 gaa cat caa ggg tta gcg ccg gcg gaa tta ttt gcg cga tta aaa gcg 528  
 Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
 165 170 175  
  
 cga atg ccg gat ggg gaa gat tta gta gta acg cat ggg gat gcg tgt 576  
 Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
 180 185 190  
  
 tta ccg aat ata atg gta gaa aat ggg cga ttt tcg ggg ttt ata gat 624  
 Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
 195 200 205  
  
 tgt ggg cga tta ggg gta gcg gat cgt tat caa gat ata gcg tta gcg 672  
 Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
 210 215 220  
  
 acg cga gat ata gcg gaa gaa tta ggg ggg gaa tgg gcg gat cga ttt 720  
 Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
 225 230 235 240  
  
 tta gta tta tat ggg ata gcg gcg ccg gat tcg caa cga ata gcg ttt 768  
 Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
 245 250 255  
  
 tat cga tta tta gat gaa ttt ttt tga 795  
 Tyr Arg Leu Leu Asp Glu Phe Phe  
 260

<210> 10  
 <211> 264  
 5 <212> PRT  
 <213> Artificial

<220>  
 <223> Constructo sintético

<400> 10

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val  
 1 5 10 15

ES 2 656 079 T3

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser  
 20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe  
 35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala  
 50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val  
 65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu  
 85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys  
 100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro  
 115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala  
 130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu  
 145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
 165 170 175

Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
 180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
 195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
 210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
 225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
 245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe  
 260

<210> 11  
 <211> 795  
 5 <212> ADN  
 <213> artificial

<220>  
 <223> gen de resistencia a neomicina modificado

10 <220>  
 <221> CDS

ES 2 656 079 T3

<222> (1)..(795)

<400> 11

atg att gaa caa gat gga ttg cac gca ggt tct ccg gcc gct tgg gtg	48
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val	
1 5 10 15	
gag agg cta ttc ggc tat gac tgg gca caa cag aca atc ggc tgc tct	96
Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser	
20 25 30	
gat gcc gcc gtg ttc cgg ctg tca gcg cag ggg cgc ccg gtt ctt ttt	144
Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe	
35 40 45	
gta aaa acg gat tta tcg ggg gcg tta aat gaa tta caa gat gaa gcg	192
Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala	
50 55 60	
gcg cga tta tcg tgg tta gcg acg acg ggg gta ccg tgt gcg gcg gta	240
Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val	
65 70 75 80	
tta gat gta gta acg gaa gcg ggg cga gat tgg tta tta tta ggg gaa	288
Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu	
85 90 95	
gta ccg ggg caa gat tta tta tcg tcg cat tta gcg ccg gcg gaa aaa	336
Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys	
100 105 110	
gta tcg ata atg gcg gat gcg atg cga cga tta cat acg tta gat ccg	384
Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro	
115 120 125	
gcg acg tgt ccg ttt gat cat caa gcg aaa cat cga ata gaa cga gcg	432
Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala	
130 135 140	
cga acg cga atg gaa gcg ggg tta gta gat caa gat gat tta gat gaa	480
Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu	
145 150 155 160	
gaa cat caa ggg tta gcg ccg gcg gaa tta ttt gcg cga tta aaa gcg	528
Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala	
165 170 175	
cga atg ccg gat ggg gaa gat tta gta gta acg cat ggg gat gcg tgt	576
Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys	
180 185 190	
tta ccg aat ata atg gta gaa aat ggg cga ttt tcg ggg ttt ata gat	624
Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp	
195 200 205	
tgt ggg cga tta ggg gta gcg gat cgt tat caa gat ata gcg tta gcg	672
Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala	
210 215 220	
acg cga gat ata gcg gaa gaa tta ggg ggg gaa tgg gcg gat cga ttt	720
Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe	
225 230 235 240	
tta gta tta tat ggg ata gcg gcg ccg gat tcg caa cga ata gcg ttt	768
Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe	
245 250 255	
tat cga tta tta gat gaa ttt ttt tga	795
Tyr Arg Leu Leu Asp Glu Phe Phe	
260	

ES 2 656 079 T3

<210> 12  
<211> 264  
<212> PRT  
<213> artificial

5

<220>  
<223> Constructo sintético

<400> 12  
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val  
1 5 10 15  
  
Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser  
20 25 30  
  
Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe  
35 40 45  
  
Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala  
50 55 60  
  
Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val  
65 70 75 80  
  
Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu  
85 90 95  
  
Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys  
100 105 110  
  
Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro  
115 120 125

ES 2 656 079 T3

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala  
130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu  
145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
165 170 175

Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe  
260

<210> 13  
<211> 795  
5 <212> ADN  
<213> artificial

<220>  
10 <223> gen de resistencia a neomicina modificado

<220>  
<221> CDS  
<222> (1)..(795)

15 <400> 13  
atg att gaa caa gat gga ttg cac gca ggt tct ccg gcc gct tgg gtg 48  
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val  
1 5 10 15  
gag agg cta ttt ggg tat gat tgg gcg caa caa acg ata ggg tgt tcg 96  
Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser  
20 25 30  
gat gcg gcg gta ttt cga tta tcg gcg caa ggg cga ccg gta tta ttt 144  
Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe  
35 40 45

ES 2 656 079 T3

gta aaa acg gat tta tcg ggg gcg tta aat gaa tta caa gat gaa gcg 192  
Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala  
50 55 60

gcg cga tta tcg tgg tta gcg acg acg ggg gta ccg tgt gcg gcg gta 240  
Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val  
65 70 75 80

tta gat gta gta acg gaa gcg ggg cga gat tgg tta tta tta ggg gaa 288  
Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu  
85 90 95

gta ccg ggg caa gat tta tta tcg tcg cat tta gcg ccg gcg gaa aaa 336  
Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys  
100 105 110

gta tcg ata atg gcg gat gcg atg cga cga tta cat acg tta gat ccg 384  
Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro  
115 120 125

gcg acg tgt ccg ttt gat cat caa gcg aaa cat cga ata gaa cga gcg 432  
Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala  
130 135 140

cga acg cga atg gaa gcg ggg tta gta gat caa gat gat tta gat gaa 480  
Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu  
145 150 155 160

gaa cat caa ggg tta gcg ccg gcg gaa tta ttt gcg cga tta aaa gcg 528  
Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
165 170 175

cga atg ccg gat ggg gaa gat tta gta gta acg cat ggg gat gcg tgt 576  
Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
180 185 190

tta ccg aat ata atg gta gaa aat ggg cga ttt tcg ggg ttt ata gat 624  
Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
195 200 205

tgt ggg cga tta ggg gta gcg gat cgt tat caa gat ata gcg tta gcg 672  
Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
210 215 220

acg cga gat ata gcg gaa gaa tta ggg ggg gaa tgg gcg gat cga ttt 720  
Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
225 230 235 240

tta gta tta tat ggg ata gcg gcg ccg gat tcg caa cga ata gcg ttt 768  
Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
245 250 255

tat cga tta tta gat gaa ttt ttt tga 795  
Tyr Arg Leu Leu Asp Glu Phe Phe  
260

<210> 14  
<211> 264  
5 <212> PRT  
<213> artificial

<220>  
10 <223> Constructo sintético

<400> 14



ES 2 656 079 T3

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val  
 1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser  
 20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe  
 35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala  
 50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val  
 65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu  
 85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys  
 100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro  
 115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala  
 130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu  
 145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala  
 165 170 175

Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys  
 180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp  
 195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala  
 210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe  
 225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe  
 245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe  
 260

- 5 <210> 15
- <211> 417
- <212> ADN
- <213> artificial

ES 2 656 079 T3

<220>

<223> cadena pesada de anticuerpo anti-CD98 humano

<220>

5 <221> CDS

<222> (1)..(417)

<400> 15

atg aag cac ctg tgg ttc ttc ctc ctg ctg gtg gcg gct ccc aga tgg 48  
Met Lys His Leu Trp Phe Phe Leu Leu Leu Val Ala Ala Pro Arg Trp  
1 5 10 15

gtc ctg tcc cag ctg cag ctg cag gag tcg ggc cca gga ctg gtg aag 96  
Val Leu Ser Gln Leu Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys  
20 25 30

cct tcg gag acc ctg tcc ctc acc tgc act gtc tct ggt ggc tcc atc 144  
Pro Ser Glu Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile  
35 40 45

agc agt agt agt tac tac tgg ggc tgg atc cgc cag ccc cca ggg aag 192  
Ser Ser Ser Ser Tyr Tyr Trp Gly Trp Ile Arg Gln Pro Pro Gly Lys  
50 55 60

ggg ctg gag tgg att ggg agt atc tat tat agt ggg agt acc tac tac 240  
Gly Leu Glu Trp Ile Gly Ser Ile Tyr Tyr Ser Gly Ser Thr Tyr Tyr  
65 70 75 80

aac ccg tcc ctc aag agt cga gtc acc ata tcc gta gac acg tcc aag 288  
Asn Pro Ser Leu Lys Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys  
85 90 95

aac cag ttc tcc ctg aag ctg agc tct gtg acc gcc gca gac acg gct 336  
Asn Gln Phe Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala  
100 105 110

gtg tat tac tgt gcg aga caa ggg acg ggg ctc gcc cta ttt gac tac 384  
Val Tyr Tyr Cys Ala Arg Gln Gly Thr Gly Leu Ala Leu Phe Asp Tyr  
115 120 125

tgg ggc cag gga acc ctg gtc acc gtc tcc tca 417  
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
130 135

10

<210> 16

<211> 139

<212> PRT

<213> artificial

15

<220>

<223> Constructo sintético

<400> 16

ES 2 656 079 T3

Met Lys His Leu Trp Phe Phe Leu Leu Leu Val Ala Ala Pro Arg Trp  
1 5 10 15

Val Leu Ser Gln Leu Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys  
20 25 30

Pro Ser Glu Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile  
35 40 45

Ser Ser Ser Ser Tyr Tyr Trp Gly Trp Ile Arg Gln Pro Pro Gly Lys  
50 55 60

Gly Leu Glu Trp Ile Gly Ser Ile Tyr Tyr Ser Gly Ser Thr Tyr Tyr  
65 70 75 80

Asn Pro Ser Leu Lys Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys  
85 90 95

Asn Gln Phe Ser Leu Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala  
100 105 110

Val Tyr Tyr Cys Ala Arg Gln Gly Thr Gly Leu Ala Leu Phe Asp Tyr  
115 120 125

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
130 135

<210> 17

<211> 387

5 <212> ADN

<213> artificial

<220>

10 <223> cadena ligera de anticuerpo anti-CD98 humano

<220>

<221> CDS

<222> (1)..(387)

15 <400> 17

atg gaa acc cca gcg cag ctt ctc ttc ctc ctg cta ctc tgg ctc cca 48  
Met Glu Thr Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu Pro  
1 5 10 15

gat acc acc gga gaa att gtg ttg acg cag tct cca ggc acc ctg tct 96  
Asp Thr Thr Gly Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser

ES 2 656 079 T3

	20		25		30	
	ttg tct cca ggg gaa aga gcc acc ctc tcc tgc agg gcc agt cag agt					144
	Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser					
	35		40		45	
	ggt agc agc agc ttc tta gcc tgg tac cag cag aaa cct ggc cag gct					192
	Val Ser Ser Ser Phe Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala					
	50		55		60	
	ccc agg ctc ctc atc tat ggt gca tcc agc agg gcc act ggc atc cca					240
	Pro Arg Leu Leu Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Ile Pro					
	65		70		75	80
	gac agg ttc agt ggc agt ggg tct ggg aca gac ttc act ctc acc atc					288
	Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile					
		85		90		95
	agc aga ctg gag cct gaa gat ttc gca gtg tat tac tgt cag cag tat					336
	Ser Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr					
		100		105		110
	ggt agc tca cct cta ttc act ttc ggc cct ggg acc aaa gtg gat atc					384
	Gly Ser Ser Pro Leu Phe Thr Phe Gly Pro Gly Thr Lys Val Asp Ile					
		115		120		125
	aaa					387
	Lys					
	<210> 18					
	<211> 129					
5	<212> PRT					
	<213> artificial					
	<220>					
10	<223> Constructo sintético					
	<400> 18					
	Met Glu Thr Pro Ala Gln Leu Leu Phe Leu Leu Leu Trp Leu Pro					
	1		5		10	15
	Asp Thr Thr Gly Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser					
		20		25		30
	Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser					
		35		40		45
	Val Ser Ser Ser Phe Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala					
		50		55		60
	Pro Arg Leu Leu Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Ile Pro					
		65		70		75
	Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile					
		85		90		95
	Ser Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr					
		100		105		110
	Gly Ser Ser Pro Leu Phe Thr Phe Gly Pro Gly Thr Lys Val Asp Ile					
		115		120		125
	Lys					

ES 2 656 079 T3

<210> 19  
 <211> 600  
 <212> ADN  
 <213> desconocido  
 5  
 <220>  
 <223> gen de resistencia a puromicina de tipo salvaje  
 <220>  
 10 <221> CDS  
 <222> (1)..(600)  
 <400> 19  
 atg acc gag tac aag ccc acg gtg cgc ctc gcc acc cgc gac gac gtc 48  
 Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15  
 ccc cgc gcc gta cgc acc ctc gcc gcc gcg ttc gcc gac tac ccc gcc 96  
 Pro Arg Ala Val Arg Thr Leu Ala Ala Phe Ala Asp Tyr Pro Ala  
 20 25 30  
 acg cgc cac acc gtc gac ccg gac cgc cac atc gag cgg gtc acc gag 144  
 Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45  
 ctg caa gaa ctc ttc ctc acg cgc gtc ggg ctc gac atc ggc aag gtg 192  
 Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60  
 tgg gtc gcg gac gac gcc gcc gcg gtg gcg gtc tgg acc acg ccg gag 240  
 Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80  
 agc gtc gaa gcg ggg gcg gtg ttc gcc gag atc gcc ccg cgc atg gcc 288  
 Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95  
 gag ttg agc ggt tcc cgg ctg gcc gcg cag caa cag atg gaa ggc ctc 336  
 Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 100 105 110  
 ctg gcg ccg cac cgg ccc aag gag ccc gcg tgg ttc ctg gcc acc gtc 384  
 Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125  
 ggc gtc tcg ccc gac cac cag gcc aag ggt ctg gcc agc gcc gtc gtg 432  
 Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140  
 ctc ccc gga gtg gag gcg gcc gag cgc gcc ggg gtg ccc gcc ttc ctg 480  
 Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160  
 gag acc tcc gcg ccc cgc aac ctc ccc ttc tac gag cgg ctc gcc ttc 528  
 Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175  
 acc gtc acc gcc gac gtc gag gtg ccc gaa gga ccg cgc acc tgg tgc 576  
 Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
 180 185 190  
 atg acc cgc aag ccc ggt gcc tga 600  
 Met Thr Arg Lys Pro Gly Ala  
 195  
 <210> 20  
 <211> 199  
 <212> PRT  
 <213> desconocido  
 20

ES 2 656 079 T3

<220>

<223> Constructo sintético

<400> 20

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
1 5 10 15

Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
20 25 30

Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
35 40 45

Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
50 55 60

Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
65 70 75 80

Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
85 90 95

Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
100 105 110

Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
115 120 125

Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
130 135 140

Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
145 150 155 160

Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
165 170 175

Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
180 185 190

Met Thr Arg Lys Pro Gly Ala  
195

- 10 <210> 21
- <211> 600
- <212> ADN
- <213> artificial

- 15 <220>
- <223> gen de resistencia a puromicina modificado

- <220>
- <221> CDS
- <222> (1)..(600)

- 20 <400> 21

ES 2 656 079 T3

atg acc gag tac aag ccc acg gtg cgc ctc gcg acc cgc gac gac gtc 48  
 Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15

ccc cgg gcg gta cgc acc ctc gcg gcg gcg ttc gcg gac tac ccc gcg 96  
 Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
 20 25 30

acg cgc cac acc gtc gac ccg gac cgc cac atc gag cgg gtc acc gag 144  
 Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45

ctg caa gaa ctc ttc ctc acg cgc gtc ggg ctc gac atc ggc aag gtg 192  
 Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60

tgg gtc gcg gac gac ggc gcg gcg gtg gcg gtc tgg acc acg ccg gag 240  
 Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80

agc gtc gaa gcg ggg gcg gtg ttc gcg gag atc ggc ccg cgc atg gcg 288  
 Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95

gag ttg agc ggt tcc cgg ctg gcg gcg cag caa cag atg gaa ggc ctc 336  
 Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 100 105 110

ctg gcg ccg cac cgg ccc aag gag ccc gcg tgg ttc ctg gcg acc gtc 384  
 Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125

ggc gtc tcg ccc gac cac cag ggc aag ggt ctg ggc agc gcg gtc gtg 432  
 Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140

ctc ccc gga gtg gag gcg gcg gag cgc gcg ggg gtg ccc gcg ttc ctg 480  
 Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160

gag acc tcc gcg ccc cgc aac ctc ccc ttc tac gag cgg ctc ggc ttc 528  
 Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175

acc gtc acc gcg gac gtc gag gtg ccc gaa gga ccg cgc acc tgg tgc 576  
 Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
 180 185 190

atg acc cgc aag ccc ggt gcg tga 600  
 Met Thr Arg Lys Pro Gly Ala  
 195

5 <210> 22  
 <211> 199  
 <212> PRT  
 <213> artificial

10 <220>  
 <223> Constructo sintético

<400> 22

ES 2 656 079 T3

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15

Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
 20 25 30

Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45

Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60

Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80

Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95

Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 100 105 110

Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125

Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140

Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160

Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175

Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
 180 185 190

Met Thr Arg Lys Pro Gly Ala  
 195

- 5 <210> 23
- <211> 600
- <212> ADN
- <213> artificial
- 10 <220>
- <223> gen de resistencia a puromicina modificado
- <220>
- <221> CDS
- 15 <222> (1)..(600)
- <400> 23



ES 2 656 079 T3

atg acc gag tac aag ccc acg gta cgc tta gcg acc cgc gac gac gta 48  
 Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15

ccc cgg gcg gta cgc acc tta gcg gcg gcg ttc gcg gac tac ccc gcg 96  
 Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
 20 25 30

acg cgc cac acc gta gac ccg gac cgc cac atc gag cgg gta acc gag 144  
 Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45

tta caa gaa tta ttc tta acg cgc gta ggg tta gac atc ggc aag gta 192  
 Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60

tgg gta gcg gac gac ggc gcg gcg gta gcg gta tgg acc acg ccg gag 240  
 Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80

tcg gta gaa gcg ggg gcg gta ttc gcg gag atc ggc ccg cgc atg gcg 288  
 Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95

gag tta tcg ggt tcg cgg tta gcg gcg cag caa cag atg gaa ggc tta 336  
 Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 100 105 110

tta gcg ccg cac cgg ccc aag gag ccc gcg tgg ttc tta gcg acc gta 384  
 Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125

ggc gta tcg ccc gac cac cag ggc aag ggt tta ggc tcg gcg gta gta 432  
 Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140

tta ccc gga gta gag gcg gcg gag cgc gcg ggg gta ccc gcg ttc tta 480  
 Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160

gag acc tcg gcg ccc cgc aac tta ccc ttc tac gag cgg tta ggc ttc 528  
 Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175

acc gta acc gcg gac gta gag gta ccc gaa gga ccg cgc acc tgg tgc 576  
 Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
 180 185 190

atg acc cgc aag ccc ggt gcg tga 600  
 Met Thr Arg Lys Pro Gly Ala  
 195

5 <210> 24  
 <211> 199  
 <212> PRT  
 <213> artificial

10 <220>  
 <223> Constructo sintético

<400> 24

ES 2 656 079 T3

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15

Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
 20 25 30

Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45

Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60

Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80

Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95

Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 100 105 110

Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125

Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140

Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160

Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175

Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
 180 185 190

Met Thr Arg Lys Pro Gly Ala  
 195

5 <210> 25  
 <211> 600  
 <212> ADN  
 <213> artificial

10 <220>  
 <223> gen de resistencia a puromicina modificado

<220>  
 <221> CDS

15 <222> (1)..(600)

<400> 25

ES 2 656 079 T3

atg acg gaa tat aaa ccg acg gta cgt tta gcg acg cgt gat gat gta 48  
 Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
 1 5 10 15

ccg cgt gcg gta cgt acg tta gcg gcg gcg ttt gcg gat tat ccg gcg 96  
 Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
 20 25 30

acg cgt cat acg gta gat ccg gat cgt cat ata gaa cgt gta acg gaa 144  
 Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
 35 40 45

tta caa gaa tta ttt tta acg cgt gta ggt tta gat ata ggt aaa gta 192  
 Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
 50 55 60

tgg gta gcg gat gat ggt gcg gcg gta gcg gta tgg acg acg ccg gaa 240  
 Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
 65 70 75 80

tcg gta gaa gcg ggt gcg gta ttt gcg gaa ata ggt ccg cgt atg gcg 288  
 Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
 85 90 95

gaa tta tcg ggt tcg cgt tta gcg gcg caa caa caa atg gaa ggt tta 336  
 Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
 100 105 110

tta gcg ccg cat cgt ccg aaa gaa ccg gcg tgg ttt tta gcg acg gta 384  
 Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
 115 120 125

ggt gta tcg ccg gat cat caa ggt aaa ggt tta ggt tcg gcg gta gta 432  
 Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
 130 135 140

tta ccg ggt gta gaa gcg gcg gaa cgt gcg ggt gta ccg gcg ttt tta 480  
 Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
 145 150 155 160

gaa acg tcg gcg ccg cgt aat tta ccg ttt tat gaa cgt tta ggt ttt 528  
 Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
 165 170 175

acg gta acg gcg gat gta gaa gta ccg gaa ggt ccg cgt acg tgg tgt 576  
 Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
 180 185 190

atg acg cgt aaa ccg ggt gcg tga 600  
 Met Thr Arg Lys Pro Gly Ala  
 195

5 <210> 26  
 <211> 199  
 <212> PRT  
 <213> artificial

10 <220>  
 <223> Constructo sintético

<400> 26

ES 2 656 079 T3

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val  
1 5 10 15

Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala  
20 25 30

Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu  
35 40 45

Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val  
50 55 60

Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val Trp Thr Thr Pro Glu  
65 70 75 80

Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala  
85 90 95

Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu  
100 105 110

Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val  
115 120 125

Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val  
130 135 140

Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu  
145 150 155 160

Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe  
165 170 175

Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys  
180 185 190

Met Thr Arg Lys Pro Gly Ala  
195

- 5 <210> 27
- <211> 375
- <212> ADN
- <213> artificial

- 10 <220>
- <223> gen de resistencia a zeocina modificado

- <220>
- <221> CDS
- 15 <222> (1)..(375)

<400> 27

ES 2 656 079 T3

atg gcg aag tta acc tcg gcg gtt ccg gta tta acc gcg cgc gac gtc 48  
 Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val  
 1 5 10 15

gcg gga gcg gtc gag ttc tgg acc gac cgg tta ggg ttc tcg cgg gac 96  
 Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp  
 20 25 30

ttc gta gag gac gac ttc gcg ggt gta gtc cgg gac gac gta acc tta 144  
 Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu  
 35 40 45

ttc atc tcg gcg gtc cag gac cag gta gta ccg gac aac acc tta gcg 192  
 Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala  
 50 55 60

tgg gta tgg gta cgc ggc tta gac gag tta tac gcg gag tgg tcg gag 240  
 Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu  
 65 70 75 80

gtc gta tcg acg aac ttc cgg gac gcc tcg ggg ccg gcg atg acc gag 288  
 Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu  
 85 90 95

atc gcc gag cag ccg tgg ggg cgg gag ttc gcg tta cgc gac ccg gcg 336  
 Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala  
 100 105 110

ggc aac tgc gta cac ttc gta gcg gag gag cag gac tga 375  
 Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp  
 115 120

5 <210> 28  
 <211> 124  
 <212> PRT  
 <213> artificial

10 <220>  
 <223> Constructo sintético

<400> 28  
 Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val  
 1 5 10 15

Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp  
 20 25 30

Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu  
 35 40 45

Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala  
 50 55 60

Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu  
 65 70 75 80

Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu  
 85 90 95

Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala  
 100 105 110

Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp  
 115 120

ES 2 656 079 T3

<210> 29  
 <211> 375  
 <212> ADN  
 5 <213> artificial  
  
 <220>  
 <223> gen de resistencia a zeocina modificado  
  
 10 <220>  
 <221> CDS  
 <222> (1)..(375)  
  
 <400> 29  
 atg gcg aaa tta acg tcg gcg gta ccg gta tta acg gcg cgt gat gta 48  
 Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val  
 1 5 10 15  
  
 gcg ggt gcg gta gaa ttt tgg acg gat cgt tta ggt ttt tcg cgt gat 96  
 Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp  
 20 25 30  
  
 ttt gta gaa gat gat ttt gcg ggt gta gta cgt gat gat gta acg tta 144  
 Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu  
 35 40 45  
  
 ttt ata tcg gcg gta caa gat caa gta gta ccg gat aat acg tta gcg 192  
 Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala  
 50 55 60  
  
 tgg gta tgg gta cgt ggt tta gat gaa tta tat gcg gaa tgg tcg gaa 240  
 Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu  
 65 70 75 80  
  
 gta gta tcg acg aat ttt cgt gat gcg tcg ggt ccg gcg atg acg gaa 288  
 Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu  
 85 90 95  
  
 ata ggt gaa caa ccg tgg ggt cgt gaa ttt gcg tta cgt gat ccg gcg 336  
 Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala  
 100 105 110  
  
 ggt aat tgt gta cat ttt gta gcg gaa gaa caa gat tga 375  
 Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp  
 115 120  
 15  
  
 <210> 30  
 <211> 124  
 <212> PRT  
 20 <213> artificial  
  
 <220>  
 <223> Constructo sintético  
  
 25 <400> 30

ES 2 656 079 T3

Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val  
1 5 10 15

Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp  
20 25 30

Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu  
35 40 45

Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala  
50 55 60

Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu  
65 70 75 80

Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu  
85 90 95

Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala  
100 105 110

Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp  
115 120

- 5 <210> 31
- <211> 1026
- <212> ADN
- <213> artificial

- 10 <220>
- <223> gen de resistencia a higromicina modificado

- <220>
- <221> CDS
- 15 <222> (1)..(1026)

<400> 31

ES 2 656 079 T3

atg aaa aag cct gaa tta acc gcg acg tcg gta gag aag ttt tta atc	48
Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile	
1 5 10 15	
gaa aag ttc gac tcg gta tcg gac tta atg cag tta tcg gag ggc gaa	96
Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu	
20 25 30	
gaa tcg cgt gcg ttc tcg ttc gat gta gga ggg cgt gga tat gta tta	144
Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu	
35 40 45	
cgt gta aat tcg tgc gcg gat ggt ttc tac aaa gat cgt tat gta tat	192
Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr	
50 55 60	
cgt cac ttt gcg tcg gcg gcg tta ccg att ccg gaa gta tta gac att	240
Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile	
65 70 75 80	
ggg gaa ttc tcg gag tcg tta acc tat tgc atc tcg cgc cgt gcg cag	288
Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln	
85 90 95	
ggt gta acg ttg caa gac tta cct gaa acc gaa tta ccc gcg gta tta	336
Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu	
100 105 110	
cag ccg gta gcg gag gcg atg gat gcg atc gcg gcg gcg gat tta tcg	384
Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser	
115 120 125	
cag acg tcg ggg ttc gcc cca ttc gga ccg caa gga atc ggt caa tac	432
Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr	



ES 2 656 079 T3

130	135	140	
act aca tgg cgt gat ttc ata tgc gcg att gcg gat ccc cat gta tat			480
Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr			
145	150	155	160
cac tgg caa act gta atg gac gac acc gta tgc gcg tgc gta gcg cag			528
His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln			
	165	170	175
gcg tta gat gag tta atg tta tgg gcg gag gac tgc ccc gaa gta cgt			576
Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg			
	180	185	190
cac tta gta cac gcg gat ttc ggc tgc aac aat gta tta acg gac aat			624
His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn			
	195	200	205
ggc cgc ata aca gcg gta att gac tgg tgc gag gcg atg ttc ggg gat			672
Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu Ala Met Phe Gly Asp			
	210	215	220
tcg caa tac gag gta gcg aac atc ttc ttc tgg cgt ccg tgg ttg gcg			720
Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp Arg Pro Trp Leu Ala			
	225	230	235
tgt atg gag cag cag acg cgc tac ttc gag cgt cgt cat ccg gag tta			768
Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu			
	245	250	255
gcg gga tgc ccg cgt tta cgt gcg tat atg tta cgc att ggt ctt gac			816
Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp			
	260	265	270
caa tta tat cag tgc ttg gta gac ggc aat ttc gat gat gcg gcg tgg			864
Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp			
	275	280	285
gcg cag ggt cga tgc gac gcg atc gta cga tgc gga gcg ggg act gta			912
Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val			
	290	295	300
ggg cgt aca caa atc gcg cgc cgt tgc gcg gcg gta tgg acc gat ggc			960
Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly			
	305	310	315
tgt gta gaa gta tta gcg gat tgc gga aac cga cgc ccc tgc act cgt			1008
Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg			
	325	330	335
ccg cgt gcg aag gaa tag			1026
Pro Arg Ala Lys Glu			
	340		

<210> 32  
 <211> 341  
 <212> PRT  
 <213> artificial

5

<220>  
 <223> Constructo sintético

10

<400> 32

ES 2 656 079 T3

Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile  
1 5 10 15

Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu  
20 25 30

Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu  
35 40 45

Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr  
50 55 60

Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile  
65 70 75 80

Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln  
85 90 95

Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu  
100 105 110

Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser  
115 120 125

Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr  
130 135 140

Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr  
145 150 155 160

His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln  
165 170 175

Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg  
180 185 190

His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn  
195 200 205

Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu Ala Met Phe Gly Asp  
210 215 220

Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp Arg Pro Trp Leu Ala  
225 230 235 240

Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu

ES 2 656 079 T3

	245		250		255	
	Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp					
		260		265		270
	Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp					
		275		280		285
	Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val					
		290		295		300
	Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly					
		305		310		315
	Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg					
			325		330	
	Pro Arg Ala Lys Glu					
		340				
	<210> 33					
	<211> 1026					
5	<212> ADN					
	<213> artificial					
	<220>					
10	<223> gen de resistencia a higromicina modificado					
	<220>					
	<221> CDS					
	<222> (1)..(1026)					
15	<400> 33					
	atg aaa aaa ccg gaa tta acg gcg acg tcg gta gaa aaa ttt tta ata					48
	Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile					
	1	5		10		15
	gaa aaa ttt gat tcg gta tcg gat tta atg caa tta tcg gaa ggt gaa					96
	Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu					
		20		25		30
	gaa tcg cgt gcg ttt tcg ttt gat gta ggt ggt cgt ggt tat gta tta					144
	Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu					
		35		40		45
	cgt gta aat tcg tgt gcg gat ggt ttt tat aaa gat cgt tat gta tat					192
	Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr					
		50		55		60
	cgt cat ttt gcg tcg gcg gcg tta ccg ata ccg gaa gta tta gat ata					240
	Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile					
		65		70		75
	ggt gaa ttt tcg gaa tcg tta acg tat tgt ata tcg cgt cgt gcg caa					288
	Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln					

ES 2 656 079 T3

85	90	95	
ggt gta acg tta caa gat tta ccg gaa acg gaa tta ccg gcg gta tta			336
Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu			
100	105	110	
caa ccg gta gcg gaa gcg atg gat gcg ata gcg gcg gcg gat tta tcg			384
Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Asp Leu Ser			
115	120	125	
caa acg tcg ggt ttt ggt ccg ttt ggt ccg caa ggt ata ggt caa tat			432
Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr			
130	135	140	
acg acg tgg cgt gat ttt ata tgt gcg ata gcg gat ccg cat gta tat			480
Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr			
145	150	155	160
cat tgg caa acg gta atg gat gat acg gta tcg gcg tcg gta gcg caa			528
His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln			
165	170	175	
gcg tta gat gaa tta atg tta tgg gcg gaa gat tgt ccg gaa gta cgt			576
Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg			
180	185	190	
cat tta gta cat gcg gat ttt ggt tcg aat aat gta tta acg gat aat			624
His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn			
195	200	205	
ggt cgt ata acg gcg gta ata gat tgg tcg gaa gcg atg ttt ggt gat			672
Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu Ala Met Phe Gly Asp			
210	215	220	
tcg caa tat gaa gta gcg aat ata ttt ttt tgg cgt ccg tgg tta gcg			720
Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp Arg Pro Trp Leu Ala			
225	230	235	240
tgt atg gaa caa caa acg cgt tat ttt gaa cgt cgt cat ccg gaa tta			768
Cys Met Glu Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu			
245	250	255	
gcg ggt tcg ccg cgt tta cgt gcg tat atg tta cgt ata ggt tta gat			816
Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp			
260	265	270	
caa tta tat caa tcg tta gta gat ggt aat ttt gat gat gcg gcg tgg			864
Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Ala Ala Trp			
275	280	285	
gcg caa ggt cgt tgt gat gcg ata gta cgt tcg ggt gcg ggt acg gta			912
Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val			
290	295	300	
ggt cgt acg caa ata gcg cgt cgt tcg gcg gcg gta tgg acg gat ggt			960
Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly			
305	310	315	320
tgt gta gaa gta tta gcg gat tcg ggt aat cgt cgt ccg tcg acg cgt			1008
Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg			
325	330	335	
ccg cgt gcg aaa gaa tga			1026
Pro Arg Ala Lys Glu			
340			

5 <210> 34  
 <211> 341  
 <212> PRT  
 <213> artificial

10 <220>

ES 2 656 079 T3

<223> Constructo sintético

<400> 34

Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile  
 1 5 10 15

Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu  
 20 25 30

Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu  
 35 40 45

Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr  
 50 55 60

Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile  
 65 70 75 80

Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln  
 85 90 95

Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu  
 100 105 110

Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser  
 115 120 125

Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr  
 130 135 140

Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr  
 145 150 155 160

His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln  
 165 170 175

Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg  
 180 185 190

His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn

ES 2 656 079 T3

195	200	205																			
Gly	Arg	Ile	Thr	Ala	Val	Ile	Asp	Trp	Ser	Glu	Ala	Met	Phe	Gly	Asp						
210						215					220										
Ser	Gln	Tyr	Glu	Val	Ala	Asn	Ile	Phe	Phe	Trp	Arg	Pro	Trp	Leu	Ala						
225					230					235					240						
Cys	Met	Glu	Gln	Gln	Thr	Arg	Tyr	Phe	Glu	Arg	Arg	His	Pro	Glu	Leu						
				245					250					255							
Ala	Gly	Ser	Pro	Arg	Leu	Arg	Ala	Tyr	Met	Leu	Arg	Ile	Gly	Leu	Asp						
			260					265					270								
Gln	Leu	Tyr	Gln	Ser	Leu	Val	Asp	Gly	Asn	Phe	Asp	Asp	Ala	Ala	Trp						
		275					280					285									
Ala	Gln	Gly	Arg	Cys	Asp	Ala	Ile	Val	Arg	Ser	Gly	Ala	Gly	Thr	Val						
	290					295					300										
Gly	Arg	Thr	Gln	Ile	Ala	Arg	Arg	Ser	Ala	Ala	Val	Trp	Thr	Asp	Gly						
305					310						315				320						
Cys	Val	Glu	Val	Leu	Ala	Asp	Ser	Gly	Asn	Arg	Arg	Pro	Ser	Thr	Arg						
				325					330					335							
Pro	Arg	Ala	Lys	Glu																	
			340																		

5 <210> 35  
 <211> 200  
 <212> ADN  
 <213> artificial

10 <220>  
 <223> secuencia del transposón Tol1-L

<400> 35  
 cagtagcggg tctaggcacg ggccgtccgg gcggtggcct ggggcgaaa actgaagggg 60  
 ggcggcaccg gcggctcagc cctttgtaat atattaatat gcaccactat tggtttactt 120  
 atgtcacagt ttgtaagttt gtaacagcct gaacctggcc gcgccgccgc cctcgccccg 180  
 cagctgcgct ctctgtctt 200

15 <210> 36  
 <211> 505  
 <212> ADN  
 <213> artificial

20 <220>  
 <223> secuencia del transposón Tol1-R

<400> 36

ES 2 656 079 T3

atatttttag ccaatagaat ttccataaat ctgtaggtag ttttaaaaat gaatatttac 60  
 catttactgc aactctatgg ggacaaaaca taatgtaaca ggtcataact aaaaatgtgc 120  
 caatcaaagg attgaagacg gaaaacatga gttaatTTTT cttctctgaa gtagagatcg 180  
 atatagaaca tgacaattta aatttccaat tcataaatgt ttttaaaata tttattttat 240  
 attatttatt taacattgag tttgattcaa tattttctta gctaactgta tttttgccat 300  
 gcttatggtc ttttattttt tgtgttctga taacttttat aatgcttttc agaattttga 360  
 catcttttgt atccacttct taatttcaat gacaataaaa catttcagtt gacgaagaca 420  
 aacaaagttc tgttgtgact atgggggggg ggggcgcctg gggatggtct cgccccggga 480  
 gtaattcagg gtagaaccgc cactg 505

<210> 37  
 <211> 1855  
 <212> ADN  
 <213> artificial

5

<220>  
 <223> transposón Tol1 no autólogo

10

<400> 37  
 cagtagcggg tctaggcacg ggccgtccgg gcggtggcct ggggcggaaa actgaagggg 60  
 ggcggcaccg gcggtcagc cctttgtaat atattaatat gcaccactat tggtttactt 120  
 atgtcacagt ttgtaagttt gtaacagcct gaacctggcc gcgccgccgc cctcgccccg 180  
 cagctgcgct ctctgtctt tgagaagtag acacaaatgt gtgtgaagaa ggagaagggg 240  
 gggggcgcgg ggtgagcacg gagcgtcgcc gcgtttgcgc atgcgcaaaa cctggctggc 300  
 tcatctttca ggggaggcga cggtcgcggg cttgatgaaa aaaataaaaag taaaaactgc 360  
 gactgcgccg tcatgtagcg aatcagcgcc cctggctgta gctgcacgcg ctctgctgg 420  
 aatgtgtga agaggggggg gggggggggg gctgcgggga atcagttcaa ttgtgggacg 480  
 cttccaaatt aagtggctag gtggggacaa ggcggggggt ttgaatctac ttcataaaac 540  
 cttttatat tataagtcag tcataaggtg acattctata acctacattt taataaaggt 600  
 ataaaaata tattctgctt tttttgggtt aattttgtgt gaaatgtcca aataaaaaaa 660  
 atggcaacac aaaacaatgc tgtcactaag gtgacagttg gttcagtcga cggacttgat 720  
 gccttcttcg tgacgtgagg acatttatgc caacaaacg ccaataaaca tctaaaatat 780  
 ggaaaagaaa aggtcaaagc catctggtgc ccaatttaga aagaaaagaa aagaagaaga 840  
 ggagaaaaga gataaagaaa agggtaagtc ctcacagctt gatgcatgtt ttttctaaat 900  
 tctaattgcta cctgocctac aacaacgttg ccgatgaaaa ctttattttg gtcgatgacc 960

# ES 2 656 079 T3

```

aacactgaat taggccc aaa  tgttgcaaat agcgtcattt tttttttttt ttttagattt 1020
tattcttaaa aatttgctct gccttaactt gtaacattag ttatgattca tgtgtctgtc 1080
tgctctgctg taacacaaaag gttttgttgg gttttgctgt tgtatactag ctcataatgt 1140
taaaaaagct gtgatggtta cacagcatgc tggctgtgcc ataagatgct aatggggcaa 1200
ataatttgag attggtcatt aatttaataa tcatttgtgg cagcctaaac gttttcacia 1260
tgtttttttg acatttaact ggggatttag gggttaattt tgagcctgca tatgaagttt 1320
atTTTTtatt tgttttacia atgtgggatt atTTTTttag ccaatagaat ttccataaat 1380
ctgtaggtag ttttaaaaat gaatatTTac catttactgc aactctatgg ggacaaaaca 1440
taatgtaaca ggtcataact aaaaatgtgc caatcaaagg attgaagacg gaaaacatga 1500
gtaatttttt ctctctgaa gtagagatcg atatagaaca tgacaattta aatttccaat 1560
tcataaatgt ttttaaaata tttattttat attatttatt taacattgag tttgattcaa 1620
tattttctta gctaactgta tttttgcat gcttatggtc ttttattttt tgtgttctga 1680
taacttttat aatgcttttc agaattttga catcttttgt atccacttct taatttcaat 1740
gacaataaaa catttcagtt gacgaagaca aacaaagttc tgttgtgact atgggggggg 1800
ggggcgctg gggatggctc cgcccgggga gtaattcagg gtagaaccgc cactg 1855

```



**REIVINDICACIONES**

1. Procedimiento para producir una proteína de interés, que comprende:

- 5 (i) introducir un vector de expresión que comprende un fragmento génico que comprende un ADN que codifica la proteína de interés y un gen marcador seleccionable atenuado y comprende asimismo un par de secuencias de transposón en ambos extremos del fragmento génico, en una célula de CHO en suspensión que puede sobrevivir y proliferar en un medio sin suero, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender unos codones utilizados a una frecuencia baja en la célula de CHO de manera que el nivel de expresión en la célula de CHO se reduce y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2;
- 10
- 15 (ii) integrar el fragmento génico que comprende el ADN que codifica la proteína de interés insertado entre el par de secuencias de transposón en un cromosoma de la célula de CHO;
- (iii) obtener una célula de CHO que expresa la proteína de interés; y
- 20 (iv) cultivar en suspensión la célula de CHO.

2. Procedimiento para producir una proteína de interés, que comprende las etapas (A) y (B) siguientes:

- 25 (A) una etapa de introducir simultáneamente los vectores de expresión (a) y (b) siguientes en una célula de CHO en suspensión que puede sobrevivir y proliferar en un medio sin suero; integrar un fragmento génico insertado entre un par de secuencias de transposón en un cromosoma de la célula de CHO mediante una transposasa expresada de manera transitoria; y obtener una célula de CHO en suspensión que expresa la proteína de interés:
- 30 (a) un vector de expresión que comprende el fragmento génico que comprende un ADN que codifica la proteína de interés y un gen marcador seleccionable atenuado y comprende asimismo el par de secuencias de transposón en ambos extremos del fragmento génico, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender unos codones utilizados a una frecuencia baja en la célula de CHO de manera que se reduce el nivel de expresión en la célula de CHO y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2,
- 35
- 40 (b) un vector de expresión que comprende un ADN que codifica la transposasa que reconoce las secuencias de transposón y presenta una actividad de transferencia del fragmento génico insertado entre el par de secuencias de transposón en el cromosoma,
- (B) una etapa de cultivar en suspensión la célula de CHO en suspensión que expresa la proteína de interés para producir la proteína de interés.
- 45

3. Procedimiento según la reivindicación 1 o 2, en el que la célula de CHO en suspensión es cualquiera de las células seleccionadas de entre CHO-K1, CHO-K1SV, DUKXB11, CHO/DG44, Pro-3 y CHO-S.

4. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el gen marcador seleccionable es:

- 50 (a) modificado en 10% o más de la secuencia de nucleótidos que codifica el gen marcador seleccionable antes de la modificación; y/o
- 55 (b) modificado de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA de entre los codones correspondientes al residuo de leucina incluidos en el gen; y/o
- (c) modificado de manera que 70% o más de los codones correspondientes al residuo de alanina son GCG de entre los codones correspondientes al residuo de alanina incluidos en el gen; y/o
- 60 (d) modificado de manera que todos los codones correspondientes al residuo de leucina incluidos en el gen son TTA o todos los codones correspondientes al residuo de alanina incluidos en el gen son GCG.

5. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el gen marcador seleccionable es un gen marcador seleccionable seleccionado de entre el grupo que consiste en un gen de resistencia a neomicina, un gen de resistencia a puromicina, un gen de resistencia a higromicina, un gen de resistencia a

zeocina y un gen de resistencia a blasticidina.

6. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que:

- 5 (a) las secuencias de nucleótidos derivadas de un par de transposones Tol2 son la secuencia de nucleótidos representada en SEC ID n°: 2 y la secuencia de nucleótidos representada en SEC ID n°: 3; o
- (b) las secuencias de nucleótidos derivadas de un par de transposones Tol1 son la secuencia de nucleótidos representada en SEC ID n°: 35 y la secuencia de nucleótidos representada en SEC ID n°: 36.

10

7. Célula de CHO en suspensión que puede sobrevivir y proliferar en un medio sin suero, en la que se introduce un vector de expresión que comprende un fragmento génico que comprende un ADN que codifica una proteína de interés y un gen marcador seleccionable atenuado y comprende asimismo un par de secuencias de transposón en ambos extremos del fragmento génico, en la que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender unos codones utilizados a una frecuencia baja en la célula de CHO de manera que se reduce el nivel de expresión en la célula de CHO y en la que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2, y en la que el fragmento génico insertado entre el par de secuencias de transposón se integra en un cromosoma de la célula de CHO en suspensión y la célula de mamífero en suspensión produce la proteína de interés.

15

20

8. Célula de CHO en suspensión que puede sobrevivir y proliferar en un medio sin suero, que presenta un cromosoma en el que se integra un fragmento génico insertado entre un par de transposones y que produce una proteína de interés y que puede obtenerse introduciendo simultáneamente los vectores (a) y (b) siguientes:

25

(a) un vector de expresión de proteínas que comprende un fragmento génico que comprende un ADN que codifica la proteína de interés y un gen marcador seleccionable atenuado y comprende asimismo el par de secuencias de transposón en ambos extremos del fragmento génico, en la que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender unos codones utilizados a una frecuencia baja en la célula de CHO de manera que el nivel de expresión en la célula de CHO se reduce y en la que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2,

30

35

(b) un vector de expresión que comprende un ADN que codifica una transposasa que reconoce las secuencias de transposón y presenta una actividad de transferencia del fragmento génico insertado entre el par de secuencias de transposón en el cromosoma.

40

9. Célula de CHO según la reivindicación 7 u 8, que es cualquiera de las células seleccionadas de entre CHO-K1, CHO-K1SV, DUKXB11, CHO/DG44, Pro-3 o CHO-S.

10. Célula de CHO según cualquiera de las reivindicaciones 7 a 9, en la que el gen marcador seleccionable es:

45

(a) modificado en 10% o más de la secuencia de nucleótidos que codifica el gen marcador seleccionable antes de la modificación; y/o

(b) modificado de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA de entre los codones correspondientes al residuo de leucina incluidos en el gen; y/o

50

(c) modificado de manera que 70% o más de los codones correspondientes al residuo de alanina son GCG de entre los codones correspondientes al residuo de alanina incluidos en el gen; y/o

(d) modificado de manera que todos los codones correspondientes al residuo de leucina incluidos en el gen son TTA o todos los codones correspondientes al residuo de alanina incluidos en el gen son GCG.

55

11. Célula de CHO según cualquiera de las reivindicaciones 7 a 10, en la que:

60

(a) el gen marcador seleccionable es un gen marcador seleccionable seleccionado de entre el grupo que consiste en un gen de resistencia a neomicina, un gen de resistencia a puromicina, un gen de resistencia a higromicina, un gen de resistencia a zeocina y un gen de resistencia a blasticidina; y/o

(b)

65

(i) las secuencias de nucleótidos derivadas de un par de transposones Tol2 son la secuencia de nucleótidos representada en la SEC ID n°: 2 y la secuencia de nucleótidos representada en la SEC ID n°: 3; o

- (ii) las secuencias de nucleótidos derivadas de un par de transposones Tol1 son la secuencia de nucleótidos representada en la SEC ID n°: 35 y la secuencia de nucleótidos representada en la SEC ID n°: 36.

5

12. Vector de expresión, que comprende un fragmento génico que comprende un ADN que codifica una proteína de interés y un marcador seleccionable atenuado, y comprende asimismo un par de secuencias de transposón en ambos extremos del fragmento génico, en el que el gen marcador seleccionable atenuado es un gen marcador seleccionable modificado para codificar la misma secuencia de aminoácidos que el gen marcador seleccionable antes de la modificación y para comprender codones utilizados a una frecuencia baja en la célula de CHO de manera que el nivel de expresión en la célula de CHO se reduce y en el que el par de secuencias de transposón son secuencias de nucleótidos derivadas de un par de transposones Tol1 o Tol2.

10

13. Vector de expresión según la reivindicación 12, en el que:

15

(a) el par de secuencias de transposón derivadas de un par de transposones Tol2 son la secuencia de nucleótidos representada en la SEC ID n°: 2 y la secuencia de nucleótidos representada en la SEC ID n°: 3; o

20

(b) el par de secuencias de transposón derivadas de un par de transposones Tol1 son la secuencia de nucleótidos representada en la SEC ID n°: 35 y la secuencia de nucleótidos representada en la SEC ID n°: 36.

14. Vector según la reivindicación 12 o 13, en el que el gen marcador seleccionable es:

25

(a) modificado en 10% o más de la secuencia de nucleótidos que codifica el gen marcador seleccionable antes de la modificación; y/o

30

(b) modificado de manera que 70% o más de los codones correspondientes al residuo de leucina son TTA de entre los codones correspondientes al residuo de leucina incluidos en el gen; y/o

(c) modificado de manera que 70% o más de los codones correspondientes al residuo de alanina son GCG de entre los codones correspondientes al residuo de alanina incluidos en el gen; y/o

35

(d) modificado de manera que todos los codones correspondientes al residuo de leucina incluidos en el gen son TTA o todos los codones correspondientes al residuo de alanina incluidos en el gen son GCG.

15. Vector según cualquiera de las reivindicaciones 12 a 14, en el que el gen marcador seleccionable es un gen marcador seleccionable seleccionado de entre el grupo que consiste en un gen de resistencia a neomicina, un gen de resistencia a puromicina, un gen de resistencia a higromicina, un gen de resistencia a zeocina y un gen de resistencia a blasticidina.

40

*Fig. 1*

