

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 843 556**

51 Int. Cl.:

C12N 15/82 (2006.01)

C07K 14/195 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.08.2016** **PCT/US2016/049132**

87 Fecha y número de publicación internacional: **11.05.2017** **WO17078836**

96 Fecha de presentación y número de la solicitud europea: **26.08.2016** **E 16763143 (1)**

97 Fecha y número de publicación de la concesión europea: **14.10.2020** **EP 3371313**

54 Título: **Métodos y composiciones de transformación de plantas mejorada**

30 Prioridad:

06.11.2015 US 201562252229 P

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

19.07.2021

73 Titular/es:

PIONEER HI-BRED INTERNATIONAL, INC.
(100.0%)
7100 N.W. 62nd Avenue
Johnston, Iowa 50131-1014, US

72 Inventor/es:

ANAND, AJITH;
BASS, STEVEN HENRY;
CHO, HYEON-JE;
KLEIN, THEODORE MITCHELL;
LASSNER, MICHAEL y
MCBRIDE, KEVIN E

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 843 556 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

Métodos y composiciones de transformación de plantas mejorada

5

Campo de la divulgación

10

La presente divulgación se refiere, en general, al campo de biología molecular de plantas, incluyendo manipulación genética de plantas. Más específicamente, la presente divulgación se refiere a métodos y composiciones para transformación de plantas, que comprenden vectores que pueden usarse para generar vectores de cointegración o como un vector auxiliar ternario.

Antecedentes

15

20

Agrobacterium, un patógeno de plantas natural, se ha usado ampliamente para la transformación de plantas dicotiledóneas y, más recientemente, para transformación de plantas monocotiledóneas. La ventaja del sistema de transferencia génica mediada por *Agrobacterium* es que ofrece la posibilidad de regenerar células transgénicas a frecuencias relativamente altas sin una reducción importante en las tasas de regeneración de las plantas. Además, el proceso de transferencia de ADN al genoma vegetal está bien caracterizado con respecto a otros métodos de suministro de ADN. El ADN transferido mediante *Agrobacterium* tiene menor probabilidad de experimentar algún reordenamiento importante que si el ADN se transfiriera mediante suministro directo, y se integra en el genoma vegetal a menudo en un solo número de copias o en bajos números de copias.

25

El sistema de transferencia génica mediada por *Agrobacterium* normalmente usado es un sistema de vector de transformación binario donde la *Agrobacterium* se ha genomanipulado para que incluya un plásmido auxiliar Ti desarmado o no oncogénico, que codifica las funciones vir necesarias para transferencia del ADN, y un plásmido diferente mucho más pequeño denominado el plásmido de vector binario, que porta el ADN transferido o la región de ADN T. El ADN T se define por secuencias en cada extremo, denominadas límites de ADN T, que desempeñan una función importante en la producción de ADN T y en el proceso de transferencia.

30

35

40

Los vectores binarios son vectores en los que los genes de virulencia están colocados en un plásmido diferente al que porta la región de ADN T (Bevan, 1984, Nucl. Acids. Res. 12: 8711-8721). El desarrollo de vectores binarios de ADN T ha hecho que la transformación de células vegetales sea más fácil ya que no requieren recombinación. El hallazgo de que algunos de los genes de virulencia mostraban efectos de dosificación génica (Jin *et al.*, J. Bacteriol. (1987) 169:4417-4425) dio lugar al desarrollo de un vector superbinario, que portaba genes de virulencia adicionales (Komari, T., *et al.*, Plant Cell Rep. (1990), 9:303-306). Estos primeros vectores superbinarios portaban un fragmento "vir" grande (~14,8 kpb) del plásmido Ti de hipervirulencia, pTiBo542, que se había introducido en un vector binario convencional (*ibid*). Los vectores superbinarios provocaron una transformación de plantas considerablemente mejorada. Por ejemplo, Hiei, Y., *et al.* (Plant J. (1994) 6:271-282) describieron la transformación eficaz de arroz por *Agrobacterium*, y posteriormente hubo informes del uso de este sistema para maíz, cebada y trigo (Ishida, Y., *et al.*, Nat. Biotech. (1996) 14:745-750; Tingay, S., *et al.*, Plant J. (1997) 11:1369-1376; y Cheng, M., *et al.*, Plant Physiol. (1997) 115:971-980; véase también la patente de Estados Unidos n.º 5.591.616 de Hiei *et al.*). Ejemplos de vectores superbinarios anteriores incluyen pTOK162 (solicitud de patente japonesa (Kokai) n.º 4-222527, documento EP-A-504.869, documento EP-A-604.662 y patente de Estados Unidos n.º 5.591.616) y pTOK233 (véase Komari, T., *ibid*; e Ishida, Y., *et al.*, *ibid*).

45

50

Sin embargo, el diseño de vectores superbinarios anteriores tiene varios inconvenientes. Por ejemplo, el fragmento vir grande y la estructura del vector portan ADN no esencial cuantioso. Además, el uso del gen de resistencia a tetraciclina tetA retarda el crecimiento bacteriano y también es un mal marcador de selección para *Agrobacterium* cepa C58 ya que esta cepa ya tiene resistencia parcial al antibiótico. Además, un gran origen de replicación/región de división (derivada de RK2, con un tamaño de aproximadamente ~15 kpb) produjo un vector bastante grande que no era fácilmente susceptible a manipulación adicional y está asociado con grados variables de inestabilidad.

55

60

La limitación con respecto a manipulación adicional se exacerbó adicionalmente por el hecho de que la reconstitución del vector de ADN T superbinario requería recombinación homóloga entre el vector "superbinario" y el vector de ADN T en una cepa de *Agrobacterium* destinataria tal como LBA4404 o C58, que es receptiva a recombinación homóloga. El proceso de recombinación homóloga es relativamente ineficaz y el vector cointegrado resultante a menudo contenía eliminaciones inesperadas. Además, el cribado de muchos clones candidatos era necesario para identificar aislados de *Agrobacterium* adecuados para transformación de plantas. En cepas tales como C58 y EHA101, se sabe que se producen mutantes espontáneos resistentes a tetraciclina a alta frecuencia (Luo, Z.Q. y Farrand, S.K., (1999) J. Bacteriol. 181:618-626), lo que impedía adicionalmente la identificación de clones recombinantes verdaderos.

65

A pesar de los avances en biología molecular de plantas, particularmente la transformación de plantas y vectores útiles en la misma, sigue habiendo una necesidad en la técnica de transformación para producir plantas transgénicas eficazmente. En particular, hay una necesidad de vectores superbinarios mejorados que tengan genes vir de tamaño óptimo con secuencias no esenciales mínimas, una mezcla óptima de genes vir para virulencia mejorada, un origen de replicación más pequeño y marcadores de selección mejorados para selección de *Agrobacterium*. De forma ideal, dichos

vectores mejorados podrían usarse fácilmente para generar vectores de cointegración o usarse como un vector auxiliar ternario para transformación de plantas. Estas necesidades y otras se abordan por la presente divulgación.

Breve sumario

La presente divulgación comprende métodos y composiciones para vectores que comprenden genes vir.

La invención proporciona un vector que comprende:

- (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*;
 - (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.;
 - (c) un gen marcador de selección; y
 - (d) genes de virulencia de *Agrobacterium* spp.:
 - (i) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,
 - (ii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,
 - (iii) *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, y
 - (iv) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91,
- en el que el vector que comprende *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y *r-virG* que tiene la SEQ ID NO: 91 comprende además *r-galls* que tiene la SEQ ID NO: 101.

La invención proporciona además un vector que comprende:

- (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2;
 - (b) un origen de replicación para la propagación en *Agrobacterium* spp. que tiene la SEQ ID NO: 3;
 - (c) un gen marcador de selección que tiene la SEQ ID NO: 1; y
 - (d) genes de virulencia de *Agrobacterium* spp.:
 - (i) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,
 - (ii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,
 - (iii) *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, y
 - (iv) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91,
- en el que el vector que comprende *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y *r-virG* que tiene la SEQ ID NO: 91 comprende además *r-galls* que tiene la SEQ ID NO: 101;
- (II)(d) genes de virulencia de *Agrobacterium* spp.:
 - (i) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,
 - (ii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,
 - (iii) *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente, o *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente,
 - (iv) *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o *r-virE3* que tiene la SEQ ID NO: 100, y
 - (v) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91,
- en el que el vector que comprende *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, *r-virE3* que tiene la SEQ ID NO: 100 y *r-virG* que tiene la SEQ ID NO: 91 comprende además *r-galls* que tiene la SEQ ID NO: 101; o
- (III)(d) genes de virulencia de *Agrobacterium* spp.:
 - (i) *virA* que tiene la SEQ ID NO: 26 o *r-virA* que tiene la SEQ ID NO: 79,
 - (ii) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,
 - (iii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,
 - (iv) *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente, o *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente,
 - (v) *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o *r-virE3* que tiene la SEQ ID NO: 100,
 - (vi) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91, y
 - (vii) *virJ* que tiene la SEQ ID NO: 27,
- en el que el vector comprende *r-virA* que tiene la SEQ ID NO: 79, *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, *r-virE3* que tiene la SEQ ID NO: 100, *r-virG* que tiene la SEQ ID NO: 91 y *virJ* que tiene la SEQ ID NO: 27 comprende además *r-galls* que tiene la SEQ ID NO: 101.

La invención proporciona además también un método para la transformación de una planta, que comprende las etapas de:

- (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* o una cepa de *Ochrobactrum*, que comprende un primer vector como se describe anteriormente, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta;
- (b) cocultivar el tejido con la cepa de *Agrobacterium* o la cepa de *Ochrobactrum*; y
- (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés.

La invención proporciona además un kit que comprende:

- (a) un vector como se describe anteriormente; y
- (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium* u *Ochrobactrum*.

Se divulga además en este documento un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Agrobacterium* spp. *virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*. En un aspecto, el vector comprende además genes de virulencia de *Agrobacterium* spp. *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP* o *virQ*, o combinaciones de los mismos. En un aspecto, el vector comprende genes de virulencia de *Agrobacterium* sp. *virB1-B11* (SEQ ID NO: 4-14, respectivamente), *virC1-C2* (SEQ ID NO: 16-17, respectivamente); *virD1-D2* (SEQ ID NO: 18-19, respectivamente) y genes *virG* (SEQ ID NO: 15). En otro aspecto, el vector comprende genes de virulencia de *Agrobacterium* sp. *virA* (SEQ ID NO: 26), *virB1-B11* (SEQ ID NO: 4-14, respectivamente), *virC1-C2* (SEQ ID NO: 16-17, respectivamente); *virD1-D5* (SEQ ID NO: 18-22, respectivamente), *virE1-E3* (SEQ ID NO: 23-25), *virG* (SEQ ID NO: 15) y genes *virJ* (SEQ ID NO: 27).

En un aspecto, la presente divulgación proporciona además métodos para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Agrobacterium* spp. *virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la *Agrobacterium*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés.

En un aspecto, la presente divulgación proporciona además kits que comprenden: (a) un vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Agrobacterium* spp. *virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium*.

La presente divulgación comprende métodos y composiciones para vectores que comprenden genes *vir*. En diversos aspectos, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Rhizobiaceae* *virB1-B11* o *r-virB1-B11*, *virC1-C2* o *r-virC1-C2*, *virD1-D2* o *r-virD1-D2*, y *virG* o *r-virG*, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2*, y *r-virG* comprende además un gen de virulencia *r-galls*, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más de los genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de

virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5*, y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos del mismo. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos del mismo. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos del mismo. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos del mismo. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRI, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59, 60 o 102, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y

fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variante y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de las mismas. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En otro aspecto, la divulgación proporciona además un vector que comprende: (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (b) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (c) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (d) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp. *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

En otro aspecto, la divulgación proporciona además un vector que comprende: (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (b) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (c) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (d) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp. *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente o genes de virulencia *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100 y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D5*, *r-virE3* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

En otro aspecto, la divulgación proporciona además un vector que comprende: (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (b) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (c) un gen marcador de selección que tiene la SEQ ID NO: 1; y (d) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp., un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente o genes de virulencia *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente o un gen de virulencia *r-virE3* que tiene las SEQ ID NO: 100 y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virA*, *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D5*, *r-virE3* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

En un aspecto, la presente divulgación proporciona además un método para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* o una cepa de *Ochrobactrum* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Rhizobiaceae virB1-B11* o *r-virB1-B11*, *virC1-C2* o *r-virC1-C2*, *virD1-D2* o *r-virD1-D2*, y *virG* o *r-virG*, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2* y *r-virG* comprende además un gen de

virulencia *r-galls*, o variantes y derivados de los mismos, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la cepa de *Agrobacterium* o la cepa de *Ochrobactrum*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más de los genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un

aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*tolA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un método para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* o una cepa de *Ochrobactrum* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (ii) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (iii) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (iv) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp. genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados de los mismos, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la cepa de *Agrobacterium* o la cepa de *Ochrobactrum*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14,

respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más de los genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen la SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación

deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un método para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* o una cepa de *Ochrobactrum* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (ii) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (iii) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (iv) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp. genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente o genes de virulencia *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100 y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D5*, *r-virE3* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados de los mismos, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la cepa de *Agrobacterium* o la cepa de *Ochrobactrum*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5* o *r-virE3*, *r-virF* o variantes y derivados de los mismos.

En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomycin. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y

fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un método para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* o una cepa de *Ochrobactrum* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (ii) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (iii) un gen marcador de selección que tiene la SEQ ID NO: 1; y (iv) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp., un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente o genes de virulencia *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente o un gen de virulencia *r-virE3* que tiene las SEQ ID NO: 100 y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virA*, *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D5*, *r-virE3* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados de los mismos, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la cepa de *Agrobacterium* o la cepa de *Ochrobactrum*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más de los genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen la SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la

SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID

NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un kit que comprende: (a) un vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Rhizobiaceae* *virB1-B11* o *r-virB1-B11*, *virC1-C2* o *r-virC1-C2*, *virD1-D2* o *r-virD1-D2*, y *virG* o *r-virG*, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2*, y *r-virG* comprende además un gen de virulencia *r-galls*, o variantes y derivados del mismo; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium* u *Ochrobactrum*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen la SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la

propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *oriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un kit que comprende: (a) un vector que comprende: (i) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (ii) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (iii) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (iv) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp. genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium* u *Ochrobactrum*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de

virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más de los genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes

y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un kit que comprende: (a) un vector que comprende: (i) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (ii) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (iii) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (iv) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp. genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente o genes de virulencia *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100 y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D5*, *r-virE3* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium* u *Ochrobactrum*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y

derivados del mismo. En un aspecto, el vector comprende además uno o más de los genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen la SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* o *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto,

el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

En un aspecto, la presente divulgación proporciona además un kit que comprende: (a) un vector que comprende: (i) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (ii) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (iii) un gen marcador de selección que tiene la SEQ ID NO: 1; y (iv) genes de virulencia que comprenden genes de virulencia de *Agrobacterium* spp., un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, genes de virulencia *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, genes de virulencia *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente o genes de virulencia *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente o un gen de virulencia *r-virE3* que tiene las SEQ ID NO: 100 y un gen de virulencia *virG* que tiene la SEQ ID NO: 15 o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virA*, *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D5*, *r-virE3* y *r-virG* comprende además un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium* u *Ochrobactrum*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia de *Agrobacterium* spp. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium rhizogenes*. En un aspecto, los genes de virulencia de *Agrobacterium* spp. son genes de virulencia de *Agrobacterium tumefaciens*. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virB1-virB11* que tienen las SEQ ID NO: 4-14, respectivamente, o genes de virulencia *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o genes de virulencia *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o genes de virulencia *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virG* que tiene la SEQ ID NO: 15, o un gen de virulencia *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo. En un aspecto, el vector comprende además uno o más genes de virulencia de *Rhizobiaceae* *virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virK*, *virL*, *virM*, *virP*, *virQ*, *r-virA*, *r-virD3*, *r-virD4*, *r-virD5* o *r-virE3*, *r-virF* o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virA* que tiene la SEQ ID NO: 26 o un gen de virulencia *r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o genes de virulencia *r-virD3-D5* que tienen la SEQ ID NO: 96-98, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o un gen de virulencia *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de virulencia *virH-H1* que tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virK* que tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virL* que tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virM* que tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virP* que

tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae* es un gen de virulencia *virQ* que tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virD3-D5* y *virE1-E3* o *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el vector comprende además los genes de virulencia de *Rhizobiaceae* *virA*, *virD3-D5* y *virE1-E3*, o *r-virA*, *r-virD3-D5* y *r-virE3*, o variantes y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1. En un aspecto, el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de pSC101. En un aspecto, el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de p15A. En un aspecto, el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de R6K. En un aspecto, el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de alto número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de número de copias intermedio. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación de bajo número de copias. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es una variante del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. En un aspecto, el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. En un aspecto, el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación de pRK2 es un microorigen de replicación de pRK2. En un aspecto, el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 es un miniorigen de replicación de pRK2. En un aspecto, el mini pRK2 tiene la SEQ ID NO: 66, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*. En un aspecto, el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65, o variantes y fragmentos de las mismas. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación de pSa tiene la SEQ ID NO: 53, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma. En un aspecto, el vector comprende además una secuencia derivada del operón par DE. En un aspecto, el operón par DE tiene la SEQ ID NO: 55, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto, el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*. En un aspecto, el gen marcador de selección es *aacC1*. En un aspecto, el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aadA*. En un aspecto, el gen marcador de selección *aadA* tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt1*. En un aspecto, el gen marcador de selección *npt1* tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *npt2*. En un aspecto, el gen marcador de selección *npt2* tiene la SEQ ID NO: 41, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *hpt*. En un aspecto, el gen marcador de selección *hpt* tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *SpcN*. En un aspecto, el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto, el gen marcador de selección es *aph*. En un aspecto, el gen marcador de selección *aph* tiene la SEQ ID NO: 78, o variante y fragmentos de la misma. En un aspecto, el gen marcador de selección no proporciona resistencia a tetraciclina. En un aspecto, el gen marcador de selección no es un gen *tetAR*. En un aspecto, el gen marcador de selección es un gen marcador de contraselección. En un aspecto, el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *adhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*. En un aspecto, el vector no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma. En un aspecto, el vector no comprende una secuencia de operón *tra* o una

secuencia de operón *trb*, o variantes o fragmentos de las mismas. En un aspecto, el vector no comprende la SEQ ID NO: 63, o variantes o fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 34, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 35, o variantes y fragmentos de la misma. En un aspecto, el vector tiene la SEQ ID NO: 36, o variantes y fragmentos de la misma.

Breve descripción de los dibujos

La figura 1 muestra un mapa plasmídico del vector pSB1.

La figura 2 muestra un mapa plasmídico del vector pPHP70298, que tiene la secuencia de la SEQ ID NO: 34.

La figura 3 muestra un mapa plasmídico del vector pPHP71539, que tiene la secuencia de la SEQ ID NO: 35.

La figura 4 muestra un mapa plasmídico del vector pPHP79761, que tiene la secuencia de la SEQ ID NO: 36, con la ubicación de sitios singulares de enzimas de restricción. Estos sitios de restricción se incluyeron entre elementos funcionales para facilitar su intercambio con otros elementos funcionalmente equivalentes.

La figura 5 muestra un diagrama que muestra la relación de los genes indicados en un fragmento de 37 171 pb del plásmido de *Agrobacterium tumefaciens* Ti pTiBo542, secuencia de referencia del NCBI: NC_010929.1

La figura 6 muestra un diagrama que muestra la relación de los genes indicados en un fragmento de 2600 pb, que comprende el origen de replicación de pVS1.

La figura 7A muestra la expresión transitoria de RFP (Ds-RED) en la cultivariedad de trigo HC0456D 15 días después de la infección; cepa AGL1. La figura 7B muestra la expresión transitoria de RFP después de 15 días tras la infección; cepa LBA4404Thy-71539. Las flechas indican las zonas con fuerte expresión de Ds-RED. La barra de escala es igual a 2 mm.

Descripción detallada

Las divulgaciones en este documento se describirán más completamente en adelante en este documento con referencia a los dibujos adjuntos.

Como se usa en la memoria descriptiva y en las reivindicaciones, la expresión "que comprende" pueden incluir el aspecto de "que consiste en". Salvo que se definan de otro modo, todos los términos técnicos y científicos usados en este documento tienen el mismo significado que el normalmente entendido por un experto en la materia a la que pertenecen las composiciones y métodos divulgados. En esta memoria descriptiva y en las reivindicaciones posteriores, se hará referencia a varios términos que se definirán en este documento.

I. Métodos y composiciones para vectores mejorados que comprenden genes *vir*

La presente divulgación comprende métodos y composiciones para los vectores que comprenden genes *vir*.

Como se usa en este documento, "pPHP" se refiere al plásmido PHP, que va seguido de dígitos numéricos. Por ejemplo, pPHP70298 se refiere al plásmido PHP70298. Por ejemplo, pVIR7 se refiere al plásmido VIR7.

La tabla 23 proporciona una lista de números de identificación de secuencia (SEQ ID NO:) proporcionados en esta divulgación.

En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Rhizobiaceae virB1-B11*, *virC1-C2*, *virD1-D2* y *virG* o genes de virulencia de *Rhizobiaceae r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2*, *r-virG* y *r-galls*, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae virB1-virB11* tienen las SEQ ID NO: 4-14, respectivamente, o los genes de virulencia de *Rhizobiaceae r-virB1-B11* tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos; los genes de virulencia de *Rhizobiaceae virC1-C2* tienen las SEQ ID NO: 16-17, respectivamente, o los genes de virulencia de *Rhizobiaceae r-virC1-C2* tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos; los genes de virulencia de *Rhizobiaceae virD1-D2* tienen las SEQ ID NO: 18-19, o los genes de virulencia de *Rhizobiaceae r-virD1-D2* tienen las SEQ ID NO: 94-95, respectivamente, o variantes y derivados de los mismos; el gen de virulencia de *Rhizobiaceae virG* tiene la SEQ ID NO: 15, o el gen de virulencia de *Rhizobiaceae r-virG* tiene la SEQ ID NO: 91, respectivamente, o variantes y derivados del mismo; y el gen de virulencia de *Rhizobiaceae r-galls* tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

En aspectos, los genes de virulencia de *Rhizobiaceae* son genes de *Agrobacterium* spp., *Rhizobium* spp., *Sinorhizobium* spp., *Mesorhizobium* spp., *Phyllobacterium* spp., *Ochrobactrum* spp. o *Bradyrhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Rhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Sinorhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Mesorhizobium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Phyllobacterium* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Ochrobactrum* spp. En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Bradyrhizobium* spp.

En un aspecto, los genes de virulencia de *Rhizobiaceae* son genes de *Agrobacterium* spp. En un aspecto, los genes de *Agrobacterium* spp. son genes de *Agrobacterium albertimagni*, *Agrobacterium larrymoorei*, *Agrobacterium radiobacter*,

Agrobacterium rhizogenes, *Agrobacterium rubi*, *Agrobacterium tumefaciens* o *Agrobacterium vitis*. En un aspecto, los genes de *Agrobacterium* spp. son de *Agrobacterium rhizogenes* o *Agrobacterium tumefaciens*. En un aspecto, los genes de *Agrobacterium* spp. son de *Agrobacterium rhizogenes*. En un aspecto, los genes de *Agrobacterium* spp. son de *Agrobacterium tumefaciens*.

5

Pueden usarse varias cepas de tipo silvestre o desarmadas (no patógenas) de *Agrobacterium tumefaciens* y *Agrobacterium rhizogenes* que albergan plásmidos Ti o Ri para transferencia génica en plantas. Los genes de síntesis de fitohormonas ubicados en el ADN T de *Agrobacteria* de tipo silvestre que albergan un plásmido Ti o Ri se expresan en células vegetales después de la transformación, y causan formación de tumores o un fenotipo de raíces pilosas dependiendo de la cepa o especie de *Agrobacterium*. El ADN T de *Agrobacteria* puede genomaniplularse para remplazar muchos de sus determinantes de virulencia y patogenia (por desarme) con una o más secuencias de interés y retener la capacidad de transferir el ADN T modificado en una célula vegetal e integrarse en un genoma. Las cepas que contienen dichos plásmidos Ti desarmados se usan ampliamente para transformación de plantas.

10

En algunos aspectos, una construcción comprende un plásmido Ti (*Agrobacterium tumefaciens*) o un plásmido Ri (*Agrobacterium rhizogenes*). En algunos aspectos, la construcción comprende uno o más genes de virulencia. Los genes de virulencia pueden ser de un plásmido Ti y se representan en este documento como las SEQ ID NO: 4-27 y las SEQ ID NO: 42-49. Los genes de virulencia pueden ser de un plásmido Ri y se representan en este documento como las SEQ ID NO: 79-101. Los genes de virulencia del plásmido Ri divulgados en este documento se representan usando una "r" delante del nombre del gen *vir*. Por ejemplo, *r-virA* (SEQ ID NO: 79), *r-virB1* (SEQ ID NO: 80), *r-virB2* (SEQ ID NO: 81), *r-virB3* (SEQ ID NO: 82), *r-virB4* (SEQ ID NO: 83), *r-virB5* (SEQ ID NO: 84), *r-virB6* (SEQ ID NO: 85), *r-virB7* (SEQ ID NO: 86), *r-virB8* (SEQ ID NO: 87), *r-virB9* (SEQ ID NO: 88), *r-virB10* (SEQ ID NO: 89), *r-virB11* (SEQ ID NO: 90), *r-virG* (SEQ ID NO: 91), *r-virC1* (SEQ ID NO: 92), *r-virC2* (SEQ ID NO: 93), *r-virD1* (SEQ ID NO: 94), *r-virD2* (SEQ ID NO: 95), *r-virD3* (SEQ ID NO: 96), *r-virD4* (SEQ ID NO: 97), *r-virD5* (SEQ ID NO: 98), *r-virF* (SEQ ID NO: 99), *r-virE3* (SEQ ID NO: 100) y *r-galls* (SEQ ID NO: 101). Véase la tabla 23 en este documento. En este documento pueden usarse diferentes combinaciones de los genes de virulencia. El gen *r-galls* (SEQ ID NO: 101) es necesario para la virulencia con los genes *vir* del plásmido Ri descritos en este documento.

25

La región Vir en el plásmido Ti/Ri es una colección de genes cuya función agregada es escindir la región de ADN T del plásmido y promover su transferencia e integración en el genoma de la planta. El sistema *vir* se induce mediante señales producidas por las plantas en respuesta a formación de heridas. Los compuestos fenólicos, tales como acetosiringona, siringoaldehído o acetovainillona activan el gen *virA*, que codifica un receptor que es una proteína transmembranaria expresada constitutivamente. El gen *virA* activado actúa como una cinasa, fosforilando el gen *virG*. En su forma fosforilada, *virG* actúa como un activador transcripcional para los operones génicos *vir* restantes. El operón *virB* codifica proteínas que producen una estructura similar a poro/pelo. *VirC* se une a la secuencia de sobreexcitación. *VirD1* y *virD2* tienen actividad endonucleasa, y realizan cortes monocatenarios dentro de los límites izquierdo y derecho, y *virD4* es una proteína de acoplamiento. *VirE* se une al ADN T monocatenario, protegiéndolo durante la fase de transporte del proceso. Una vez en la célula vegetal, se sintetiza la hebra complementaria del ADN T.

30

35

Estos y otros genes *vir* funcionan en *trans*, de modo que ninguno de estos genes tiene que incluirse en los vectores de clonación. Por ejemplo, cepas de *Agrobacterium* modificadas pueden proporcionar todas las funciones Vir necesarias en los plásmidos donde la región de ADN T se ha eliminado, permitiendo que la célula proporcione las funciones *vir* para la transferencia de ADN T. En un ejemplo, hay cepas derivadas de C58 en que una parte de pBR322 se usó para remplazar la región de ADN T, y que proporciona resistencia a ampicilina.

40

45

Se proporcionan construcciones que incluyen una o más secuencias de interés para la expresión y/o inserción en un genoma celular. Las construcciones pueden contenerse dentro de un vector tales como vectores binarios, ternarios o de ADN T. Una construcción se refiere a una molécula polinucleotídica compuesta de diversos tipos de secuencias nucleotídicas que tienen diferentes funciones y/o actividades. Diversos tipos de secuencias incluyen conectores, adaptadores, regiones reguladoras, intrones, sitios de restricción, potenciadores, aislantes, marcadores cribables, marcadores de selección, promotores, casetes de expresión, polinucleótidos codificantes, polinucleótidos silenciadores, secuencias de terminación, orígenes de replicación, sitios de recombinación, casetes de escisión, recombinasas, factores de proliferación celular, elementos de retención de promotores, otros sitios que ayudan a la construcción o análisis de vectores, o cualquier combinación de los mismos. En algunos ejemplos, una construcción comprende uno o más casetes de expresión, en la que un polinucleótido está unido de forma funcional a una secuencia reguladora. Unido de forma funcional es una unión funcional entre dos o más elementos. Por ejemplo, una unión funcional entre un polinucleótido codificante y una secuencia reguladora (por ejemplo, un promotor) es una unión funcional que permite la expresión del polinucleótido codificante. Los elementos unidos de forma funcional pueden ser contiguos o no contiguos. Cuando se usa para hacer referencia a la unión de dos regiones codificantes de proteína, por unidas de forma funcional se pretende que las regiones codificantes estén en el mismo marco de lectura. Un polinucleótido codificante incluye cualquier polinucleótido que codifica un polipéptido, o que codifica un polinucleótido silenciador que reduce la expresión de genes diana. Ejemplos no limitantes de un polinucleótido silenciador incluyen un ARN interferente pequeño, microARN, ARN antisentido y una estructura de horquilla. La construcción también puede contener varios componentes genéticos para facilitar la transformación de la célula o tejido vegetal y para regular la expresión de cualquier secuencia de ácido nucleico estructural. En algunos ejemplos, los componentes genéticos se orientan de modo que expresen un ARNm, opcionalmente el ARNm se traduce en una proteína. La expresión de una secuencia codificante estructural vegetal (un gen, ADNc, ADN sintético

50

55

60

65

u otro ADN) que existe en forma bicatenaria implica la transcripción de ARN mensajero (ARNm) a partir de una hebra del ADN por la enzima ARN polimerasa y el posterior procesamiento del transcrito primario de ARNm dentro del núcleo. Este procesamiento implica una región no traducida 3' que poliadenila los extremos 3' del ARNm.

- 5 En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Agrobacterium virB1-B11*; *virC1-C2*; *virD1-D2*; y *virG*, o los genes de virulencia de *Agrobacterium r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2*, *r-virG* y *r-galls*, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Agrobacterium virB1-B11* tienen las SEQ ID NO: 4-14, respectivamente, o variantes y derivados de los mismos; los genes de virulencia de *Agrobacterium virC1-C2* tienen las SEQ ID NO: 16-17, respectivamente, o variantes y derivados de los mismos; los genes de virulencia de *Agrobacterium* spp. *virD1-D2* tienen las SEQ ID NO: 18-19, respectivamente, o variantes y derivados de los mismos; y el gen de virulencia de *Agrobacterium virG* tiene la SEQ ID NO: 15, o variantes y derivados del mismo; o los genes de virulencia de *Agrobacterium r-virB1-B11* tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos, *r-virC1-C2* tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos, *r-virD1-D2* tienen las SEQ ID NO: 95-96, respectivamente, o variantes y derivados de los mismos, *r-virG* tiene la SEQ ID NO: 91, o variantes y derivados del mismo, y *r-galls* tiene la SEQ ID NO: 101, o variantes y derivados del mismo.
- 10
- 15
- 20 En un aspecto, el vector comprende además uno o más genes de virulencia de *Rhizobiaceae virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virJ*, *virK*, *virL*, *virM*, *virP* o *virQ*, o variantes y derivados de los mismos, o uno o más genes de virulencia de *Rhizobiaceae r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF*, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* y *r-virF* comprende además un gen de virulencia *r-galls*, o variantes y derivados del mismo. Por tanto, en un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Rhizobiaceae virB1-B11*; *virC1-C2*; *virD1-D2* y *virG*, o variantes y derivados de los mismos, o los genes de virulencia de *Rhizobiaceae r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2* y *r-virG*, o variantes y derivados de los mismos y *r-galls*, o variantes y derivados del mismo; y opcionalmente uno o más genes de virulencia de *Rhizobiaceae virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virJ*, *virK*, *virL*, *virM*, *virP* o *virQ*, o variantes y derivados de los mismos o uno o más genes de virulencia de *Rhizobiaceae r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF*, o variantes y derivados de los mismos y *r-galls*, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virA* tiene la SEQ ID NO: 26 o el gen de virulencia de *Rhizobiaceae r-virA* tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae virD3-D5* tienen las SEQ ID NO: 20-22, respectivamente, o los genes de virulencia de *Rhizobiaceae r-virD3-D5* tienen las SEQ ID NO: 94-96, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae virE1-E3* tienen las SEQ ID NO: 23-25, respectivamente, o el gen de virulencia de *Rhizobiaceae r-virE3* tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae virH-H1* tienen las SEQ ID NO: 42-43 respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae virJ* tiene la SEQ ID NO: 27, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virK* tiene la SEQ ID NO: 45, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virL* tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virM* tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virP* tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virQ* tiene la SEQ ID NO: 49, o variantes y derivados del mismo.
- 25
- 30
- 35
- 40
- 45

- En un aspecto, el vector comprende además uno o más genes de virulencia de *Agrobacterium virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virJ*, *virK*, *virL*, *virM*, *virP* o *virQ*, o variantes y derivados de los mismos, o uno o más genes de virulencia de *Agrobacterium r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF*, o variantes y derivados de los mismos, y *r-galls*, o variantes y derivados del mismo. Por tanto, en un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Agrobacterium virB1-B11*; *virC1-C2*; *virD1-D2*; y *virG*, o variantes y derivados de los mismos, o los genes de virulencia de *Agrobacterium r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2*, *r-virG* y *r-galls*, o variantes y derivados de los mismos; y opcionalmente uno o más genes de virulencia de *Agrobacterium virA*, *virD3*, *virD4*, *virD5*, *virE1*, *virE2*, *virE3*, *virH*, *virH1*, *virH2*, *virJ*, *virK*, *virL*, *virM*, *virP* o *virQ*, o variantes y derivados de los mismos, o uno o más genes de virulencia de *Agrobacterium r-virA*, *r-virD3*, *r-virD4*, *r-virD5*, *r-virE3* o *r-virF* o variantes y derivados de los mismos, y *r-galls*, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virA* tiene la SEQ ID NO: 26 o el gen de virulencia de *Rhizobiaceae r-virA* tiene la SEQ ID NO: 79, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae virD3-D5* tienen las SEQ ID NO: 20-22, respectivamente, o los genes de virulencia de *Rhizobiaceae r-virD3-D5* tienen las SEQ ID NO: 94-96, respectivamente, o variantes y derivados de los mismos. En un aspecto, los genes de virulencia de *Rhizobiaceae virE1-E3* tienen las SEQ ID NO: 23-25, respectivamente, o el gen de virulencia de *Rhizobiaceae r-virE3* tiene la SEQ ID NO: 100, o variantes y derivados del mismo. En un aspecto, los genes de virulencia de *Rhizobiaceae virH-H2* tienen las SEQ ID NO: 42-43, respectivamente, o variantes y derivados de los mismos. En un aspecto, el gen de virulencia de *Rhizobiaceae virJ* tiene la SEQ ID NO: 27, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virK* tiene la SEQ ID NO: 45, o variantes y derivados del mismo.
- 50
- 55
- 60
- 65

- 5 mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virL* tiene la SEQ ID NO: 46, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virM* tiene la SEQ ID NO: 47, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virP* tiene la SEQ ID NO: 48, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae virQ* tiene la SEQ ID NO: 49, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae r-virF* tiene la SEQ ID NO: 99, o variantes y derivados del mismo. En un aspecto, el gen de virulencia de *Rhizobiaceae r-galls* tiene la SEQ ID NO: 101, o variantes y derivados del mismo.
- 10 En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Agrobacterium virB1-B11*; *virC1-C2*; *virD1-D5*; *virE1-E3*; y *virG*, o variantes y derivados de los mismos, o los genes de virulencia de *Agrobacterium r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, *r-virE3* que tiene las SEQ ID NO: 100, *r-virG* que tiene la SEQ ID NO: 91 y *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.
- 15 En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Agrobacterium virA*; *virB1-B11*; *virC1-C2*; *virD1-D5*; *virE1-E3*; y *virG*, o variantes y derivados de los mismos, o los genes de virulencia de *Agrobacterium r-virB1-B11*; *r-virC1-C2*; *r-virD1-D2*; *r-virG*; y *r-galls*, o variantes y derivados de los mismos.
- 20 En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (c) un gen marcador de selección; y (d) genes de virulencia de *Agrobacterium virA*; *virB1-B11*; *virC1-C2*; *virD1-D5*; *virE1-E3*; *virG*; y *virJ*, o variantes y derivados de los mismos, o los genes de virulencia de *Agrobacterium r-virA* que tienen la SEQ ID NO: 79, *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, *r-virE3* que tiene la SEQ ID NO: 100, *r-virG* que tiene la SEQ ID NO: 91 y *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.
- 25 La presente divulgación proporciona un vector que comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*, derivado de un origen de replicación de Col E1, pSC101, p15A o R6K, o variantes funcionales y derivados de los mismos. En un aspecto, puede usarse cualquier origen de replicación funcional en *Agrobacterium* en la construcción de los vectores divulgados. Por ejemplo, pueden seleccionarse diferentes orígenes de replicación para conseguir diferentes frecuencias y cualidades (una sola copia de ADN T, sin estructura) de acontecimientos de transformación. En un aspecto, el uno o más orígenes de replicación es un origen que es funcional en *Agrobacterium*, *E. coli* o ambos. En un aspecto, el uno o más orígenes de replicación se selecciona del grupo que consiste en pVS1, pSa, RK2, pRi, incPa, incW, Col E1, pRSF1010, pBBR1 o variantes funcionales y derivados de los mismos.
- 30 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de Col E1. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de Col E1, que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma.
- 35 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de pSC101. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de pSC101, que tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma.
- 40 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de p15A. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de p15A, que tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma.
- 45 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de R6K. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de R6K, que tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma.
- 50 En diversos aspectos, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. puede ser un origen de replicación de bajo número de copias, un origen de replicación de número de copias intermedio o un origen de replicación de alto número de copias. Se entiende que un origen de replicación de bajo número de copias proporciona aproximadamente 1-2 copias del vector por célula (por ejemplo, véase Li *et al.*, Plant Cell Report (2015)
- 55 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de Col E1. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de Col E1, que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma.
- 60 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de pSC101. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de pSC101, que tiene la SEQ ID NO: 50, o variantes y fragmentos de la misma.
- 65 En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de p15A. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de p15A, que tiene la SEQ ID NO: 51, o variantes y fragmentos de la misma.
- En un aspecto, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de R6K. En un aspecto adicional, el vector comprende un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* derivado de un origen de replicación de R6K, que tiene la SEQ ID NO: 52, o variantes y fragmentos de la misma.

- 34:745-54; y Cho y Winans Proc. Natl. Acad. Sci. USA (2005) 102:14843-848). Además, se entiende que un origen de replicación de número de copias intermedio proporciona aproximadamente 7-12 copias del vector por célula (por ejemplo, véase Oltmanns, *et al.*, Plant Physiol. (2010) 152:1158-1166). Los ejemplos ejemplares, pero no limitantes, de orígenes de replicación de número de copias intermedio incluyen los derivados de pRK2 y variantes de copia baja de pVS1. Debe
- 5 apreciarse que un origen de replicación de alto número de copias proporciona aproximadamente 15-20 y más copias del vector por célula (por ejemplo, véase Oltmanns, *et al.*, Plant Physiol. (2010) 152:1158-1166; y Li *et al.*, Plant Cell Report (2015) 34:745-54). Los ejemplos ejemplares, pero no limitantes, de orígenes de replicación de número de copias intermedio incluyen los derivados de repABC y variantes de copia alta de pVS1.
- 10 En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium*
- 15 spp. deriva del origen de replicación de pVS1. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. deriva del origen de replicación de pSa. En diversos aspectos, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. tiene la SEQ ID NO: 3, 37, 38, 57, 58, 59 o 60, o variantes y fragmentos de las mismas.
- 20 En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un origen de replicación compatible de repABC. El origen de replicación compatible de repABC puede tener las SEQ ID NO: 57, 58, 59 o 60, o variantes y fragmentos de las mismas.
- 25 En algunos aspectos, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación. Por ejemplo, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. pueden derivar de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. pueden derivar del origen de replicación de pRK2. En un
- 30 aspecto adicional, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. pueden derivar del origen de replicación de pRK2 que tiene la SEQ ID NO: 38, o variantes y fragmentos de la misma. En un aspecto, el origen de replicación deriva del origen de replicación de pSa. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. pueden derivar del origen de replicación de pSa (SEQ ID NO: 53), o variantes y fragmentos del mismo. En un aspecto adicional, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. pueden derivar del origen de replicación de pRSF1010. En un aspecto adicional, el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37, o variantes y fragmentos de la misma.
- 40 Las variantes del origen de replicación de pRK2 incluyen un mini o microorigen de replicación de pRK2. En un aspecto, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un microorigen de replicación de pRK2. En un aspecto adicional, el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es un microorigen de replicación de pRK2 y tiene la SEQ ID NO: 54, o variantes y fragmentos de la misma.
- 45 En aspectos, el vector divulgado comprende además una secuencia derivada del operón par DE. En un aspecto adicional, el vector divulgado que comprende un origen de replicación de pRK2, en particular un micro o miniorigen de replicación de pRK2, puede comprender además una secuencia derivada del operón par DE. La secuencia del operón par DE puede tener la SEQ ID NO: 55, o variantes y fragmentos de la misma.
- 50 En un aspecto, el marcador de selección proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomomicina. En un aspecto adicional, el marcador de selección es *aacC1*, *npt1*, *npt2*, *hpt*, *aadA*, *SpcN* o *aph*. En un aspecto, el marcador de selección tiene la SEQ ID NO: 1, 39, 40, 41, 67, 77 o 78, o variantes y fragmentos de las mismas, correspondientes, respectivamente, a *aacC1*, *aadA*, *npt1*, *npt2*, *hpt*, *SpcN* y *aph*. En un aspecto, el marcador de selección es *aacC1*. En un aspecto adicional, el marcador de selección es *aacC1*, y tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma. En un aspecto, el marcador de selección es *aadA*. En un aspecto adicional, el marcador de selección es *aadA*, y tiene la SEQ ID NO: 39, o variantes y fragmentos de la misma. En un aspecto, el marcador de selección es *npt1*.
- 60 En un aspecto adicional, el marcador de selección es *npt1*, y tiene la SEQ ID NO: 40, o variantes y fragmentos de la misma. En un aspecto, el marcador de selección es *npt2*. En un aspecto adicional, el marcador de selección es *npt2*, y tiene la SEQ ID NO: 41. En un aspecto, el marcador de selección es *hpt*. En un aspecto adicional, el marcador de selección es *hpt*, y tiene la SEQ ID NO: 67, o variantes y fragmentos de la misma. En un aspecto adicional, el marcador de selección es *SpcN*, y tiene la SEQ ID NO: 77, o variantes y fragmentos de la misma. En un aspecto adicional, el marcador de selección es *aph*, y tiene la SEQ ID NO: 78, o variantes y fragmentos de la misma. En diversos aspectos, el marcador de selección no es un marcador de selección de tetraciclina. En un aspecto, el marcador de selección no es *tetAR*.
- 65

En un aspecto, el marcador de selección es un marcador de contraselección o marcador de selección negativa. Como puede apreciarse, se entiende que el vector divulgado puede comprender combinaciones particulares de genes de virulencia (*vir*), orígenes de replicación y marcadores de selección. La tabla 1 a continuación proporciona combinaciones ejemplares, pero no limitantes, que comprenden genes de virulencia, uno o más marcadores de selección y orígenes de replicación. En un aspecto, el marcador de contraselección es *sacB*, *rpsL* (*strA*), *pheS*, *dhfr* (*folA*), *lacY*, *Gata-1*, *ccdB* o *thyA*.

Tabla 1.

N.º de combinación	Genes de virulencia	Origen de replicación	Marcador de selección
1	virB1-B11; virC1-C2; virD1-D2; y virG	pVS1	aacC1
2	virB1-B11; virC1-C2; virD1-D5; virE1-E3; y virG	pVS1	aacC1
3	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	pVS1	aacC1
4	virB1-B11; virC1-C2; virD1-D5; virE1-E3; y virG	pRi repABC	aacC1
5	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	pRi repABC	aacC1
6	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	pRK2	aacC1
7	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	pSa	aacC1
8	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	pSa y par DE	aacC1
9	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	pBBR1	aacC1
10	virA; virB1-B11; virC1-C2; virD1-D5; virE1-E3; virG; y virJ	RFS1010	aacC1

10

En un aspecto, la presente divulgación proporciona un vector que no comprende la SEQ ID NO: 61, o variantes o fragmentos de la misma.

15

En un aspecto, la presente divulgación proporciona un vector que no comprende la SEQ ID NO: 62, o variantes o fragmentos de la misma.

20

En un aspecto, la presente divulgación proporciona un vector que no comprende una secuencia de operón *tra* o *trb*, o variantes o fragmentos del mismo. En un aspecto adicional, la presente divulgación proporciona un vector que no comprende una secuencia de operón *tra* o *trb*, en el que la secuencia de operón *tra* o *trb* tiene la SEQ ID NO: 63, o variantes o fragmentos de la misma.

25

En un aspecto, la presente divulgación proporciona un vector que tiene las SEQ ID NO: 34, 35 o 36, o variantes y fragmentos de las mismas.

30

En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (b) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (c) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (d) genes de virulencia que comprenden genes de virulencia de *Agrobacterium virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o variantes y derivados de los mismos; *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o variantes y derivados de los mismos; *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o variantes y derivados de los mismos; y *virG* que tiene la SEQ ID NO: 15, o variantes y derivados del mismo, o los genes de virulencia de *Agrobacterium r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos, *r-virD1-D2* que tienen las SEQ ID NO: 95-96, respectivamente, o variantes y derivados de los mismos, *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo, y *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

35

En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (b) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (c) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos de la misma; y (d) genes de virulencia que comprenden genes de virulencia de *Agrobacterium virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o variantes y derivados de los mismos; *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o variantes y derivados de los mismos; *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente, o variantes y derivados de los mismos; *virE1-E3* que tienen las SEQ ID NO: 23-25, o variantes y derivados de los mismos; y *virG* que tiene la SEQ ID NO: 15, o variantes y derivados del mismo, o los genes de virulencia de *Agrobacterium r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos; *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos; *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, o variantes y derivados de los mismos; *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo; *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo; y *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

En un aspecto, la presente divulgación proporciona un vector que comprende: (a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2, o variantes y fragmentos de la misma; (b) un origen de replicación para la propagación en *Agrobacterium* spp., que tiene la SEQ ID NO: 3, o variantes y fragmentos de la misma; (c) un gen marcador de selección que tiene la SEQ ID NO: 1, o variantes y fragmentos del mismo; y (d) genes de virulencia que comprenden *virA* de *Agrobacterium* que tiene la SEQ ID NO: 26, o variantes y derivados del mismo; *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o variantes y derivados de los mismos; *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o variantes y derivados de los mismos; *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente, o variantes y derivados de los mismos; *virE1-E3* que tienen las SEQ ID NO: 23-25, o variantes y derivados de los mismos; *virG* que tiene la SEQ ID NO: 15, o variantes y derivados del mismo; y *virJ* que tiene la SEQ ID NO: 27, o variantes y derivados del mismo, o los genes de virulencia de *Agrobacterium r-virA* que tiene la SEQ ID NO: 79, o variantes y derivados del mismo; *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, o variantes y derivados de los mismos; *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, o variantes y derivados de los mismos; *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, o variantes y derivados de los mismos; *r-virE3* que tiene la SEQ ID NO: 100, o variantes y derivados del mismo; *r-virG* que tiene la SEQ ID NO: 91, o variantes y derivados del mismo; y *r-galls* que tiene la SEQ ID NO: 101, o variantes y derivados del mismo.

En un aspecto, la presente divulgación proporciona además métodos para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Rhizobiaceae virB1-B11*, *virC1-C2*, *virD1-D2* y genes *virG*, o genes de virulencia de *Rhizobiaceae r-virB1-B11*, *r-virC1-C*, *r-virD1-D2*, *r-virG* y *r-galls*, o variantes y derivados de los mismos, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la *Agrobacterium*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés.

En un aspecto, la presente divulgación proporciona además métodos para la transformación de una planta, que comprende las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Agrobacterium virB1-B11*, *virC1-C2*, *virD1-D2* y genes *virG*, o genes de virulencia de *Agrobacterium r-virB1-B11*, *r-virC1-C*, *r-virD1-D2*, *r-virG* y *r-galls*, o variantes y derivados de los mismos, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la *Agrobacterium*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés.

En un aspecto, la presente divulgación proporciona además kits que comprenden: (a) un vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Rhizobiaceae virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*, o variantes y derivados de los mismos; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium*.

En un aspecto, la presente divulgación proporciona además kits que comprenden: (a) un vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Agrobacterium virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*, o variantes y derivados de los mismos, o genes de virulencia de *Agrobacterium r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2*, *r-virG* y *r-galls*; y (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium*.

"Planta" incluye referencias a plantas completas, órganos de la planta, tejidos vegetales, semillas y células vegetales, y la descendencia de los mismos. La descendencia, las variantes y los mutantes de las plantas regeneradas también se

incluyen dentro del alcance de la presente divulgación, con la condición de que estas partes comprendan los polinucleótidos introducidos o se transformaran usando los vectores de la presente divulgación. Las células vegetales incluyen, sin limitación, células de semillas, cultivos en suspensión, embriones, regiones meristemáticas, tejido de callo, hojas, raíces, brotes, gametófitos, esporófitos, polen y microesporas.

Como se usa en este documento, "regeneración" se refiere al proceso de hacer crecer una planta a partir de una célula o células vegetales (por ejemplo, protoplasto vegetal, callo o explante).

Como se usa en este documento, el término "protoplasto" se refiere a una célula vegetal aislada sin paredes celulares, que tiene la potencia de regeneración en cultivo celular o una planta completa.

Las partes de plantas incluyen tejidos diferenciados e indiferenciados incluyendo, aunque sin limitación, los siguientes: raíces, tallos, brotes, hojas, polen, semillas, tejido tumoral y diversas formas de células y cultivo (por ejemplo, células individuales, protoplastos, embriones y tejido de callo). El tejido vegetal puede estar en una planta o en un órgano de la planta, tejido o cultivo celular.

La presente divulgación también incluye plantas obtenidas por cualquiera de los métodos divulgados o composiciones de este documento.

La presente divulgación también incluye semillas de una planta obtenida por cualquiera de los métodos o composiciones divulgadas en este documento.

Variantes y fragmentos

Por "fragmento" se pretende una parte de un polinucleótido o una parte de la secuencia de aminoácidos y, por tanto, la proteína codificada por el mismo. Los fragmentos de un polinucleótido pueden codificar fragmentos de la proteína que retienen la actividad biológica de la proteína natural. Por tanto, los fragmentos de una secuencia de nucleótidos pueden variar de al menos aproximadamente 10 nucleótidos, aproximadamente 15 nucleótidos, aproximadamente 16 nucleótidos, aproximadamente 17 nucleótidos, aproximadamente 18 nucleótidos, aproximadamente 19 nucleótidos, aproximadamente 20 nucleótidos, aproximadamente 22 nucleótidos, aproximadamente 50 nucleótidos, aproximadamente 75 nucleótidos, aproximadamente 100 nucleótidos, aproximadamente 200 nucleótidos, aproximadamente 300 nucleótidos, aproximadamente 400 nucleótidos, aproximadamente 500 nucleótidos, aproximadamente 600 nucleótidos y hasta el polinucleótido de longitud completa empleado.

Por "derivado" se pretende un polinucleótido o una parte de un polinucleótido que posee actividad que es sustancialmente similar a la actividad biológica del polinucleótido de referencia. Un derivado de un polinucleótido de gen de virulencia será funcional y retendrá la actividad génica de virulencia.

"Variante" pretende significar una secuencia sustancialmente similar. Para los polinucleótidos, una variante comprende una eliminación y/o adición y/o sustitución de uno o más nucleótidos en uno o más sitios internos dentro del polinucleótido natural y/o una sustitución de uno o más nucleótidos en uno o más sitios en el polinucleótido natural. Una variante de un polinucleótido de gen de virulencia retendrá la actividad génica de virulencia. Como se usa en este documento, un polinucleótido o polipéptido "natural" comprende una secuencia de nucleótidos o secuencia de aminoácidos de origen natural, respectivamente. Para los polinucleótidos, las variantes conservativas incluyen aquellas secuencias que, a causa de la degeneración del código genético, codifican la secuencia de aminoácidos de un polipéptido codificado por un gen de virulencia. Los polinucleótidos variantes también incluyen polinucleótidos derivados sintéticamente, tales como los generados, por ejemplo, usando mutagénesis dirigida al sitio, pero siguen reteniendo la actividad deseada. En general, las variantes de un polinucleótido divulgado particular (es decir, un gen de virulencia) tendrán al menos aproximadamente un 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más identidad de secuencia con el polinucleótido particular, que se determina por programas de alineación de secuencias y parámetros descritos en otra parte en este documento.

Las variantes de un polinucleótido divulgado particular (es decir, el polinucleótido de referencia) también pueden evaluarse por comparación del porcentaje de identidad de secuencia entre el polipéptido codificado por un polinucleótido variante y el polipéptido codificado por el polinucleótido de referencia. El porcentaje de identidad de secuencia entre dos polipéptidos cualesquiera puede calcularse usando programas de alineación de secuencias y parámetros descritos en otra parte en este documento. Cuando se evalúa cualquier par dado de polinucleótidos divulgados empleados por comparación del porcentaje de identidad de secuencia compartida por los dos polipéptidos que codifican, el porcentaje de identidad de secuencia entre los dos polipéptidos codificados es de al menos aproximadamente un 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más identidad de secuencia.

Las siguientes expresiones se usan para describir las relaciones de secuencia entre dos o más polinucleótidos o polipéptidos: (a) "secuencia de referencia", (b) "ventana de comparación". (c) "identidad de secuencia" y (d) "porcentaje de identidad de secuencia".

(a) Como se usa en este documento, "secuencia de referencia" se refiere a una secuencia definida usada como base para la comparación de secuencias. Una secuencia de referencia puede ser un subconjunto o la totalidad de una secuencia especificada; por ejemplo, un segmento de un ADNc o de una secuencia génica de longitud completa, o el ADNc o la secuencia génica de longitud completa.

(b) Como se usa en este documento, "ventana de comparación" hace referencia a un segmento contiguo y especificado de una secuencia polinucleotídica, en la que la secuencia polinucleotídica en la ventana de comparación puede comprender adiciones o eliminaciones (es decir, huecos) en comparación con la secuencia de referencia (que no comprende adiciones o eliminaciones) para la alineación óptima de los dos polinucleótidos. En general, la ventana de comparación es de al menos 20 nucleótidos contiguos de longitud y, opcionalmente, puede ser de 30, 40, 50, 100 o más larga. Los expertos en la materia comprenderán que para evitar una similitud elevada con una secuencia de referencia debido a la inclusión de huecos en la secuencia polinucleotídica, se suele introducir una penalización por huecos y se sustrae del número de coincidencias.

Salvo que se indique de otro modo, los valores de identidad/similitud de secuencia proporcionados en este documento se refieren al valor obtenido usando GAP Versión 10 usando los siguientes parámetros: % de identidad y % de similitud para una secuencia de nucleótidos usando un peso de hueco de 50 y un peso de longitud de 3, y la matriz de puntuación `nwsgapdna.cmp`; % de identidad y % de similitud para una secuencia de aminoácidos usando un peso de hueco de 8 y un peso de longitud de 2, y la matriz de puntuación `BLOSUM62`; o cualquier programa equivalente de los mismos. Por "programa equivalente" se pretende cualquier programa de comparación de secuencias que, para dos secuencias cualesquiera en cuestión, genera una alineación que tiene emparejamientos idénticos de nucleótidos o residuos aminoácidos y un porcentaje de identidad de secuencia idéntico en comparación con la alineación correspondiente generada por GAP Versión 10.

(c) Como se usa en este documento, "identidad de secuencia" o "identidad", en el contexto de dos secuencias polinucleotídicas o polipeptídicas, hace referencia a los residuos en las dos secuencias que son iguales cuando se alinean para una correspondencia máxima sobre una ventaja de comparación especificada. Cuando se usa el porcentaje de identidad de secuencia en referencia a proteínas, se reconoce que las posiciones de los residuos que no son idénticas a menudo difieren por sustituciones aminoacídicas conservativas, donde los residuos aminoacídicos están sustituidos por otros residuos aminoacídicos con propiedades químicas similares (por ejemplo, carga o hidrofobia) y, por lo tanto, no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservativas, el porcentaje de identidad de secuencia se puede ajustar al alza para corregir la naturaleza conservativa de la sustitución. Se dice que las secuencias que difieren en dichas sustituciones conservativas tienen "similitud de secuencia" o "similitud". Los medios para hacer este ajuste son bien conocidos para los expertos en la materia. Típicamente, esto implica puntuar una sustitución conservativa como un emparejamiento erróneo parcial en lugar de completo, aumentando de ese modo el porcentaje de identidad de secuencia. Por tanto, por ejemplo, cuando a un aminoácido idéntico se le da una puntuación de 1 y a una sustitución no conservativa se le da una puntuación de cero, a una sustitución conservativa se le da una puntuación entre cero y 1. La puntuación de las sustituciones conservativas se calcula, por ejemplo, como se implementa en el programa `PC/GENE` (Intelligenetics, Mountain View, California).

(d) Como se usa en este documento, "porcentaje de identidad de secuencia" significa el valor determinado al comparar dos secuencias alineadas de forma óptima sobre una ventana de comparación, en la que la parte de la secuencia polinucleotídica en la ventana de comparación puede comprender adiciones o eliminaciones (es decir, huecos) en comparación con la secuencia de referencia (que no comprende adiciones o eliminaciones) para la alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las que aparece la base del ácido nucleico o el residuo aminoacídico idéntico en ambas secuencias para obtener el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para obtener el porcentaje de identidad de secuencia.

Se proporciona además un método para identificar una variante de gen de virulencia expuesto en las SEQ ID NO: 4-27 y 42-49, o variantes y derivados de las mismas. Dichos métodos comprenden obtener un derivado candidato de una cualquiera de las SEQ ID NO: 4-27 y 42-49, o variantes y derivados de las mismas, que sea de suficiente longitud para retener la actividad génica de virulencia en cuestión; remplazar el gen de virulencia referido en un vector de control para producir un vector de la variante candidata y determinar si el derivado polinucleotídico de virulencia candidato tiene la actividad del gen de virulencia referido y proporcionar, de ese modo, la función de transformación deseada del vector como se describe en este documento. Los métodos de identificación de dichas variantes candidatas, basados en el efecto de transformación deseado, a la luz de los contenidos proporcionados en este documento, son conocidos. En diversos aspectos, debe entenderse que la expresión "... las SEQ ID NO: 4-27 y 42-49, o variantes o derivados de las mismas..." pretende significar que las secuencias divulgadas comprenden las SEQ ID NO: 4-27 y 42-49, y/o derivados de las SEQ ID NO: 4-27 y 42-49, las variantes de las SEQ ID NO: 4-27 y 42-49, y/o los derivados de las SEQ ID NO: 4-27 y 42-49, individualmente (o) o incluyendo algunas o todas las secuencias enumeradas.

II. Transformación de plantas

En aspectos, los métodos de la presente divulgación implican introducir un polinucleótido en una célula vegetal. "Introducir" pretende significar presentar a la célula vegetal el polinucleótido de tal manera que la secuencia obtenga acceso al interior de una célula vegetal. Los métodos de la presente divulgación implican introducir un polinucleótido en una célula vegetal usando métodos tales como transformación mediada por *Agrobacterium* (patente de Estados Unidos n.º 5.563.055 y patente de Estados Unidos n.º 5.981.840).

En aspectos, los vectores de la presente divulgación pueden usarse para mejorar la eficacia y velocidad de introducción de un polinucleótido en una célula vegetal.

En aspectos, los vectores de la presente divulgación son útiles para transformar una o más células de un explante. El explante, incluyendo tejido vegetal somático maduro e inmaduro, puede usarse como fuente o material de explante en la presente divulgación siempre que pueda producir material embriogénico o embriones somáticos. El tejido vegetal somático adecuado incluye tejido estaminífero (es decir, flores masculinas), pistilífero (es decir, flores femeninas), flores completas, discos de bulbo, tallos en floración y brácteas. Las flores inmaduras y discos de bulbo son las fuentes preferidas de tejido vegetal somático. En un aspecto, el explante derivado de la planta usado para la transformación incluye embriones inmaduros, embriones cigóticos de 1-5 mm y embriones de 3,5-5 mm.

El explante usado en los métodos divulgados puede derivarse de una monocotiledónea incluyendo, aunque sin limitación, cebada, maíz, mijo, avena, arroz, centeno, *Setaria* spp., sorgo, caña de azúcar, pasto varilla, triticale, césped o trigo. Como alternativa, el explante usado en los métodos divulgados puede derivarse de una dicotiledónea incluyendo, aunque sin limitación, col rizada, coliflor, brécol, panta de mostaza, col, guisante, trébol, alfalfa, judía verde, tomate, mandioca, soja, colza, alfalfa, girasol, cártamo, tabaco, *Arabidopsis* o algodón.

En un aspecto adicional, el explante usado en los métodos divulgados puede derivarse de una planta que es un miembro de la familia *Poaceae*. Ejemplos no limitantes de plantas adecuadas de las que puede derivarse un explante de la divulgación incluyen cultivos de cereales incluyendo, aunque sin limitación, cebada, maíz, avena, arroz, centeno, sorgo, trigo, mijo, triticale; cultivos de tallos y hojas incluyendo, aunque sin limitación, bambú, carrizo, poa de los prados, juncos, ballico, caña de azúcar; hierbas de pasto, hierbas ornamentales y otras hierbas tales como pasto varilla y césped.

En un aspecto adicional, el explante usado en los métodos divulgados puede derivarse de cualquier planta, incluyendo plantas superiores, por ejemplo, clases de *Angiospermae* y *Gymnospermae*. Plantas de las subclases de las *Dicotyledonae* y las *Monocotyledonae* son adecuadas. Las especies adecuadas pueden venir de la familia *Acanthaceae*, *Alliaceae*, *Alstroemeriaceae*, *Amaryllidaceae*, *Apocynaceae*, *Arecaceae*, *Asteraceae*, *Berberidaceae*, *Bixaceae*, *Brassicaceae*, *Bromeliaceae*, *Cannabaceae*, *Caryophyllaceae*, *Cephalotaxaceae*, *Chenopodiaceae*, *Colchicaceae*, *Cucurbitaceae*, *Dioscoreaceae*, *Ephedraceae*, *Erythroxylaceae*, *Euphorbiaceae*, *Fabaceae*, *Lamiaceae*, *Linaceae*, *Lycopodiaceae*, *Malvaceae*, *Melanthiaceae*, *Musaceae*, *Myrtaceae*, *Nyssaceae*, *Papaveraceae*, *Pinaceae*, *Plantaginaceae*, *Poaceae*, *Rosaceae*, *Rubiaceae*, *Salicaceae*, *Sapindaceae*, *Solanaceae*, *Taxaceae*, *Theaceae* y *Vitaceae*.

Las especies adecuadas de las que puede derivarse el explante usado en los métodos divulgados incluyen miembros del género *Abelmoschus*, *Abies*, *Acer*, *Agrostis*, *Allium*, *Alstroemeria*, *Ananas*, *Andrographis*, *Andropogon*, *Artemisia*, *Arundo*, *Atropa*, *Berberis*, *Beta*, *Bixa*, *Brassica*, *Calendula*, *Camellia*, *Camptotheca*, *Cannabis*, *Capsicum*, *Carthamus*, *Catharanthus*, *Cephalotaxus*, *Chrysanthemum*, *Cinchona*, *Citrullus*, *Coffea*, *Colchicum*, *Coleus*, *Cucumis*, *Cucurbita*, *Cynodon*, *Datura*, *Daucus*, *Dianthus*, *Digitalis*, *Dioscorea*, *Elaeis*, *Ephedra*, *Erianthus*, *Erythroxylum*, *Eucalyptus*, *Festuca*, *Fragaria*, *Galanthus*, *Glycine*, *Gossypium*, *Helianthus*, *Hevea*, *Hordeum*, *Hyoscyamus*, *Jatropha*, *Juglans*, *Lactuca*, *Lavendula*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Lycopodium*, *Manihot*, *Medicago*, *Mentha*, *Miscanthus*, *Moringa*, *Musa*, *Nicotiana*, *Oryza*, *Panicum*, *Papaver*, *Parthenium*, *Pennisetum*, *Petunia*, *Phalaris*, *Phleum*, *Pinus*, *Poa*, *Poinsettia*, *Populus*, *Rauwolfia*, *Ricinus*, *Rosa*, *Rosmarinus*, *Saccharum*, *Salix*, *Sanguinaria*, *Scopolia*, *Secale*, *Solanum*, *Sorghum*, *Spartina*, *Spinacea*, *Tanacetum*, *Taxus*, *Theobroma*, *Triticosecale*, *Triticum*, *Uniola*, *Veratrum*, *Vinca*, *Vitis* y *Zea*.

En un aspecto adicional, el explante usado en los métodos divulgados puede derivarse de una planta que es importante o interesante para agricultura, horticultura, biomasa para la producción de moléculas de combustible líquido y otros productos químicos y/o silvicultura. Ejemplos no limitantes incluyen, por ejemplo, *Panicum virgatum* (pasto varilla), *Sorghum bicolor* (sorgo, hierba del Sudán), *Miscanthus giganteus* (miscanto), *Saccharum* spp. (caña), *Populus balsamifera* (álamo), *Zea mays* (maíz), *Glycine max* (soja), *Brassica napus* (colza), *Triticum aestivum* (trigo), *Gossypium hirsutum* (algodón), *Oryza sativa* (arroz), *Helianthus annuus* (girasol), *Medicago sativa* (alfalfa), *Beta vulgaris* (remolacha azucarera), *Pennisetum glaucum* (espadaña), *Panicum* spp., *Sorghum* spp., *Miscanthus* spp., *Saccharum* spp., *Erianthus* spp., *Populus* spp., *Andropogon gerardii* (hierba tallo azul grande), *Pennisetum purpureum* (hierba de elefante), *Phalaris arundinacea* (alpiste cinta), *Cynodon dactylon* (gramilla colorada), *Festuca arundinacea* (festuca alta), *Spartina pectinata* (espartillo de la pradera), *Arundo donax* (caña común), *Secale cereale* (centeno), *Salix* spp. (sauce), *Eucalyptus* spp. (eucalipto), *Triticosecale* spp. (triticum-trigo X centeno), bambú, *Carthamus tinctorius* (cártamo), *Jatropha curcas* (jatrofa), *Ricinus communis* (ricino), *Elaeis guineensis* (palma), *Linum usitatissimum* (lino), *Brassica juncea*, *Manihot esculenta* (mandioca), *Lycopersicon esculentum* (tomate), *Lactuca sativa* (lechuga), *Musa paradisiaca* (plátano), *Solanum tuberosum* (patata), *Brassica oleracea* (brécol, coliflor, coles de Bruselas), *Camellia sinensis* (té), *Fragaria ananassa* (fresa), *Theobroma cacao* (cacao), *Coffea arabica* (café), *Vitis vinifera* (uva), *Ananas comosus* (piña), *Capsicum annum* (pimiento picante y dulce), *Allium cepa* (cebolla), *Cucumis melo* (melón), *Cucumis sativus* (pepino), *Cucurbita maxima* (calabaza), *Cucurbita moschata* (calabaza), *Spinacea oleracea* (espinaca), *Citrullus lanatus* (sandía), *Abelmoschus esculentus* (quingombó), *Solanum melongena* (berenjena), *Papaver somniferum* (adormidera), *Papaver orientale*, *Taxus baccata*, *Taxus brevifolia*, *Artemisia annua*, *Cannabis sativa*, *Camptotheca acuminata*, *Catharanthus roseus*, *Vinca rosea*, *Cinchona officinalis*, *Colchicum autumnale*, *Veratrum californica*, *Digitalis lanata*, *Digitalis purpurea*, *Dioscorea* spp., *Andrographis paniculata*, *Atropa belladonna*, *Datura stamonium*, *Berberis* spp., *Cephalotaxus* spp., *Ephedra sinica*, *Ephedra* spp., *Erythroxylum coca*, *Galanthus wornorii*, *Scopolia* spp., *Lycopodium serratum* (= *Huperzia serrata*),

5 *Lycopodium* spp., *Rauwolfia serpentina*, *Rauwolfia* spp., *Sanguinaria canadensis*, *Hyoscyamus* spp., *Calendula officinalis*, *Chrysanthemum parthenium*, *Coleus forskohlii*, *Tanacetum parthenium*, *Parthenium argentatum* (guayule), *Hevea* spp. (caucho), *Mentha spicata* (menta), *Mentha piperita* (menta), *Bixa orellana*, *Alstroemeria* spp., *Rosa* spp. (rosa), *Dianthus caryophyllus* (clavel), *Petunia* spp. (petunia), *Poinsettia pulcherrima* (poinsetia), *Nicotiana tabacum* (tabaco), *Lupinus albus* (altramuz), *Uniola paniculata* (avena), césped inclinado (*Agrostis* spp.), *Populus tremuloides* (álamo), *Pinus* spp. (pino), *Abies* spp. (abeto), *Acer* spp. (arce), *Hordeum vulgare* (cebada), *Poa pratensis* (pasto azul), *Lolium* spp. (ballico), *Phleum pratense* (fleó) y coníferas. Son de interés plantas que se cultivan para la producción de energía, los llamados cultivos energéticos, tales como cultivos energéticos basados en celulosa como *Panicum virgatum* (pasto varilla), *Sorghum bicolor* (sorgo, hierba del Sudán), *Miscanthus giganteus* (miscanto), *Saccharum* spp. (caña), *Populus balsamifera* (álamo), *Andropogon gerardii* (hierba tallo azul grande), *Pennisetum purpureum* (hierba de elefante), *Phalaris arundinacea* (alpiste cinta), *Cynodon dactylon* (gramilla colorada), *Festuca arundinacea* (festuca alta), *Spartina pectinata* (espartillo de la pradera), *Medicago sativa* (alfalfa), *Arundo donax* (caña común), *Secale cereale* (centeno), *Salix* spp. (sauce), *Eucalyptus* spp. (eucalipto), *Triticosecale* spp. (triticum-trigo X centeno) y bambú; y cultivos energéticos basados en almidón como *Zea mays* (maíz) y *Manihot esculenta* (mandioca); y cultivos energéticos basados en sacarosa como *Saccharum* spp. (caña de azúcar) y *Beta vulgaris* (remolacha azucarera); y cultivos energéticos productores de biodiésel como *Glycine max* (soja), *Brassica napus* (colza), *Helianthus annuus* (girasol), *Carthamus tinctorius* (cártamo), *Jatropha curcas* (jatrofa), *Ricinus communis* (ricino), *Elaeis guineensis* (palma), *Linum usitatissimum* (lino) y *Brassica juncea*.

20 Como se usa en este documento, una "planta fuente de energía renovable de biomasa" o "biomasa para la producción de moléculas de combustible líquido y otros productos químicos" significa una planta que tiene o produce material (sin procesar o procesado) que comprende energía solar almacenada que puede convertirse en energía eléctrica, combustibles líquidos y otros productos químicos útiles. En términos generales, dichas plantas comprenden cultivos energéticos especializados, así como plantas agrícolas y leñosas. Ejemplos de plantas fuente de energía renovable de biomasa incluyen: *Panicum virgatum* (pasto varilla), *Sorghum bicolor* (sorgo, hierba del Sudán), *Miscanthus giganteus* (miscanto), *Saccharum* spp. (caña), *Populus balsamifera* (álamo), *Andropogon gerardii* (hierba tallo azul grande), *Pennisetum purpureum* (hierba de elefante), *Phalaris arundinacea* (alpiste cinta), *Cynodon dactylon* (gramilla colorada), *Festuca arundinacea* (festuca alta), *Spartina pectinata* (espartillo de la pradera), *Medicago sativa* (alfalfa), *Arundo donax* (caña común), *Secale cereale* (centeno), *Salix* spp. (sauce), *Eucalyptus* spp. (eucalipto), *Triticosecale* spp. (triticum-trigo X centeno), bambú, *Zea mays* (maíz), *Manihot esculenta* (mandioca), *Saccharum* spp. (caña de azúcar), *Beta vulgaris* (remolacha azucarera), *Glycine max* (soja), *Brassica napus* (colza), *Helianthus annuus* (girasol), *Carthamus tinctorius* (cártamo), *Jatropha curcas* (jatrofa), *Ricinus communis* (ricino), *Elaeis guineensis* (palma), *Linum usitatissimum* (lino) y *Brassica juncea*.

35 Las células que se han transformado pueden hacerse crecer en plantas de acuerdo con maneras convencionales. Véase, por ejemplo, McCormick *et al.* (1986) Plant Cell Reports 5:81-84. Estas plantas después pueden cultivarse, y polinizarse con la misma cepa transformada o diferentes cepas, e identificarse la descendencia resultante que tiene la expresión constitutiva de la característica fenotípica deseada. Pueden cultivarse dos o más generaciones para garantizar que la expresión de la característica fenotípica deseada se mantiene de forma estable y se hereda y después pueden recogerse las semillas para garantizar que se ha conseguido la expresión de la característica fenotípica deseada. De esta manera, las composiciones y métodos descritos en este documento proporcionan semillas transformadas (también denominadas "semillas transgénicas") que comprenden un polinucleótido que se ha introducido en una planta usando un vector de la presente divulgación, incorporado de forma estable en su genoma.

45 Por tanto, los métodos y composiciones de la presente divulgación pueden usarse para la transformación de cualquier especie de planta y el desarrollo de plantas transgénicas de cualquier especie incluyendo, aunque sin limitación, monocotiledóneas y dicotiledóneas. Ejemplos de especies de plantas de interés incluyen, aunque sin limitación, maíz (*Zea mays*), *Brassica* spp. (por ejemplo, *B. napus*, *B. rapa*, *B. juncea*), particularmente aquellas especies de *Brassica* útiles como fuente de aceite de semilla, alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), mijo (por ejemplo, espadaña (*Pennisetum glaucum*), mijo común (*Panicum miliaceum*), panizo (*Setaria italica*), coracán (*Eleusine coracana*), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), patata (*Solanum tuberosum*), cacahuetes (*Arachis hypogaea*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), batata (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea* spp.), coco (*Cocos nucifera*), piña (*Ananas comosus*), árboles cítricos (*Citrus* spp.), cacao (*Theobroma cacao*), té (*Camellia sinensis*), plátano (*Musa* spp.), aguacate (*Persea americana*), higo (*Ficus casica*), guayaba (*Psidium guajava*), mango (*Mangifera indica*), oliva (*Olea europaea*), papaya (*Carica papaya*), anacardo (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almendra (*Prunus amygdalus*), remolacha azucarera (*Beta vulgaris*), caña de azúcar (*Saccharum* spp.), avena, cebada, hortalizas, plantas ornamentales y coníferas.

60 Las hortalizas incluyen tomates (*Lycopersicon esculentum*), lechuga (por ejemplo, *Lactuca sativa*), judías verdes (*Phaseolus vulgaris*), habas (*Phaseolus limensis*), guisantes (*Lathyrus* spp.) y miembros del género *Cucumis* tales como pepino (*C. sativus*), cantalupo (*C. cantalupensis*) y melón (*C. melo*). Las plantas ornamentales incluyen azalea (*Rhododendron* spp.), hortensia (*Macrophylla hydrangea*), hibisco (*Hibiscus rosasanensis*), rosas (*Rosa* spp.), tulipanes (*Tulipa* spp.), narcisos (*Narcissus* spp.), petunias (*Petunia hybrida*), claveles (*Dianthus caryophyllus*), poinsetia (*Euphorbia pulcherrima*) y crisantemo.

65 Las coníferas que pueden emplearse en la práctica de la presente divulgación incluyen, por ejemplo, pinos tales como pino taeda (*Pinus taeda*), pino del incienso (*Pinus elliotii*), pino ponderosa (*Pinus ponderosa*), pino murrayana (*Pinus contorta*), y pino de Monterrey (*Pinus radiata*); pino Oregón (*Pseudotsuga menziesii*); tsuga occidental (*Tsuga canadensis*);

- tsuga oriental (*Tsuga heterophylla*); tsuga montañosa (*Tsuga mertensiana*); alerce o pino alerce (*Larix occidentalis*); picea blanca (*Picea glauca*); secuoya (*Sequoia sempervirens*); abetos verdaderos tales como abeto plateado (*Abies amabilis*) y abeto balsámico (*Abies balsamea*); y cedros tales como cedro rojo occidental (*Thuja plicata*) y falso ciprés de Nootka (*Chamaecyparis nootkatensis*). Las especies de eucaliptos pueden emplearse en la práctica de la presente divulgación, incluyendo *E. grandis* (y sus híbridos, como "urograndis"), *E. globulus*, *E. camaldulensis*, *E. tereticornis*, *E. viminalis*, *E. nitens*, *E. saligna* y *E. urophylla*. De forma óptima, las plantas de la presente divulgación son plantas de cultivo (por ejemplo, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, cacahuete, sorgo, trigo, mijo, tabaco, etc.), de forma más óptima plantas de maíz y soja, de forma aún más óptima, plantas de maíz.
- Plantas de interés particular incluyen plantas de cereales que proporcionan semillas de interés, plantas de semillas oleaginosas y plantas leguminosas. Semillas de interés incluyen, aunque sin limitación, semillas de cereal, tales como maíz, trigo, cebada, arroz, sorgo y centeno. Las plantas de semillas oleaginosas incluyen, aunque sin limitación, algodón, soja, cártamo, girasol, *Brassica*, maíz, alfalfa, palma, y coco, las plantas leguminosas incluyen, aunque sin limitación, judías y guisantes. Las judías incluyen guar, algarrobo, fenugreco, soja, judías de jardín, caupí, soja verde, judía blanca, haba, lentejas, garbanzos.

III. Métodos para mejorar rasgos y características vegetales

- La presente divulgación proporciona composiciones y métodos novedosos para producir plantas transformadas con eficacia aumentada. Los métodos y composiciones divulgadas pueden comprender además polinucleótidos que proporcionan rasgos y características mejoradas. Por tanto, la presente divulgación proporciona además métodos para la transformación de una planta, comprendiendo el método las etapas de: (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Rhizobiaceae* *virB1-B11* o *r-virB1-B11*, *virC1-C2* o *r-virC1-C2*, *virD1-D2* o *r-virD1-D2*, y *virG* o *r-virG*, o variantes y derivados de los mismos, en el que el vector que comprende los genes de virulencia *r-virB1-B11*, *r-virC1-C2*, *r-virD1-D2* y *r-virG* comprende además un gen de virulencia *r-galls*, o variantes y derivados del mismo, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta; (b) cocultivar el tejido con la *Agrobacterium*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés; en el que la secuencia polinucleotídica proporciona un rasgo o característica mejorada.

- Como se usa en este documento, "rasgo" se refiere a una característica fisiológica, morfológica, bioquímica o física de una planta o material o célula vegetal particular. En algunos casos, esta característica es visible a simple vista, tal como tamaño de la semilla o la planta, o puede medirse por técnicas bioquímicas, tal como detectando el contenido de proteínas, almidón o aceite de la semilla u hojas, o por observación de un proceso metabólico o fisiológico, por ejemplo, midiendo la captación de dióxido de carbono, o por la observación del nivel de expresión de un gen o genes, por ejemplo, empleando análisis de Northern, RT-PCR, ensayos de expresión génica en micromatriz o sistemas de expresión de genes indicadores, o por observaciones agrícolas tales como tolerancia a las agresiones, producción o tolerancia a patógenos. Un "rasgo potenciado", como se usa al describir los aspectos de la presente divulgación incluye, por ejemplo, eficacia del uso de agua o tolerancia a las sequías mejorada o potenciada, tolerancia a agresión osmótica, tolerancia a agresión por alta salinidad, tolerancia a agresión térmica, tolerancia potenciada al frío, incluyendo tolerancia a germinación en frío, producción aumentada, eficacia del uso de nitrógeno potenciada, crecimiento y desarrollo temprano de la planta, crecimiento y desarrollo de la planta tardío, contenido potenciado de proteínas en las semillas y producción potenciada de aceite en las semillas.

- Puede usarse cualquier polinucleótido de interés en los métodos de la presente divulgación. Diversos cambios en el fenotipo son de interés incluyendo, aunque sin limitación, modificación de la composición de ácidos grasos en una planta, alteración del contenido de aminoácidos, contenido de almidón o contenido de carbohidratos de una planta, alteración del mecanismo de defensa contra patógenos de la planta, influencia en el tamaño del grano y carga de sacarosa. El gen de interés también puede estar implicado en la regulación del flujo entrante de nutrientes, y en la regulación de la expresión de genes de fitato, particularmente para reducir los niveles de fitato en la semilla. Estos resultados pueden conseguirse proporcionando la expresión de productos heterólogos o la expresión aumentada de productos endógenos en las plantas. Como alternativa, los resultados pueden conseguirse proporcionando una reducción de la expresión de uno o más productos endógenos, particularmente enzimas o cofactores en la planta. Estos cambios provocan un cambio en el fenotipo de la planta transformada.

- Los genes de interés son un reflejo de los mercados comerciales y los intereses de los implicados en el desarrollo del cultivo. Los cultivos y mercados de interés cambian, y según se abren las naciones en desarrollo a los mercados mundiales, surgirán también nuevos cultivos y tecnologías. Además, según aumenta la comprensión de los rasgos y características agronómicas tales como la producción y heterosis, la elección de genes para su transformación cambiará en consecuencia. Las categorías generales de genes de interés incluyen, por ejemplo, aquellos genes implicados en la información, tales como dedos de cinc, aquellos implicados en la comunicación, tales como cinasas y aquellos implicados en tareas constitutivas, tales como proteínas de choque térmico. Categorías más específicas de transgenes, por ejemplo, incluyen genes que codifican rasgos importantes para agronomía, resistencia a insectos, resistencia a enfermedades, resistencia a herbicidas, esterilidad, características del grano y productos comerciales. Los genes de interés incluyen, en

general, aquellos implicados en el contenido de aceite, almidón, carbohidratos o el metabolismo de los nutrientes, así como aquellos que afectan al tamaño del grano y la carga de sacarosa.

Los polinucleótidos introducidos en un tejido diana por los métodos y composiciones divulgadas pueden unirse de forma funcional a un promotor adecuado. Un tejido diana puede incluir, aunque sin limitación, un embrión somático, semillas maduras, meristemas, explante foliar, semillas, cultivos en suspensión, embriones, regiones meristemáticas, tejido de callo, hojas, raíces, brotes, gametófitos, esporófitos, polen, microesporas y otros explantes vegetales. "Promotor" significa una región de ADN que está en dirección 5' del inicio de la transcripción y está implicada en el reconocimiento y unión de la ARN polimerasa y otras proteínas para iniciar la transcripción. Un "promotor vegetal" es un promotor que puede iniciar la transcripción en células vegetales sea su origen o no de una célula vegetal. Promotores vegetales ejemplares incluyen, aunque sin limitación, los que se obtienen de plantas, virus de plantas y bacterias que comprenden genes expresados en células vegetales tales como *Agrobacterium* o *Rhizobium*. Ejemplos de promotores bajo control del desarrollo incluyen promotores que inician la transcripción preferentemente en determinados tejidos, tales como hojas, raíces o semillas. Dichos promotores se denominan "con preferencia de tejido". Los promotores que inician la transcripción solamente en determinados tejidos se denominan "específicos de tejido". Un promotor específico de "tipo celular" dirige principalmente la expresión en determinados tipos celulares en uno o más órganos, por ejemplo, células vasculares en raíces u hojas. Un promotor "inducible" o "reprimible" puede ser un promotor que está bajo control ambiental o exógeno. Ejemplos de condiciones ambientales que pueden lograr la transcripción por promotores inducibles incluyen condiciones anaeróbicas, o determinados productos químicos, o la presencia de luz. Como alternativa, el control exógeno de un promotor inducible o reprimible puede lograrse proporcionando un producto químico adecuado u otro agente que, mediante interacción con polipéptidos diana, provoca la inducción o represión del promotor. Los promotores específicos de tejido, con preferencia de tejido, específicos de tipo celular e inducibles constituyen la clase de promotores "no constitutivos". Un promotor "constitutivo" es un promotor que es activo en la mayoría de circunstancias. Como se usa en este documento, "orientación antisentido" incluye referencias a una secuencia polinucleotídica que está unida de forma funcional a un promotor en una orientación donde se transcribe la hebra de antisentido. La hebra de antisentido es suficientemente complementaria a un producto de transcripción endógeno, de modo que la traducción del producto de transcripción endógeno a menudo se inhibe. "Unido de forma funcional" se refiere a la asociación de dos o más fragmentos de ácido nucleico en un solo fragmento de ácido nucleico, de modo que la función de uno se ve afectada por el otro. Por ejemplo, un promotor está unido de forma funcional con una secuencia codificante cuando puede afectar a la expresión de esa secuencia codificante (es decir, que la secuencia codificante está bajo el control transcripcional del promotor). Las secuencias codificantes pueden unirse de forma funcional a secuencias reguladoras en orientación de sentido o de antisentido.

Los rasgos agrónomicamente importantes tales como contenido de aceite, almidón y proteínas pueden alterarse genéticamente además de usar métodos de fitomejoramiento tradicionales. Las modificaciones incluyen aumentar el contenido de ácido oleico, aceites saturados e insaturados, aumentar los niveles de lisina y azufre, proporcionar aminoácidos esenciales y también la modificación del almidón. Se describen modificaciones de la proteína hordotionina en las patentes de Estados Unidos n.º 5.703.049, 5.885.801, 5.885.802 y 5.990.389. Otro ejemplo es la proteína de semilla rica en lisina y/o azufre codificada por la albúmina 2S de soja descrita en la patente de Estados Unidos n.º 5.850.016, y el inhibidor de quimotripsina de cebada, descrito en Williamson *et al.* (1987) Eur. J. Biochem. 165:99-106.

Pueden prepararse derivados de las secuencias codificantes por mutagénesis dirigida al sitio para aumentar el nivel de aminoácidos preseleccionados en el polipéptido codificado. Por ejemplo, las proteínas vegetales ricas en metionina tales como de semillas de girasol (Lilley *et al.* (1989) Proceedings of the World Congress on Vegetable Protein Utilization in Human Foods and Animal Feedstuffs, ed. Applewhite (American Oil Chemists Society, Champaign, Ill.), pág. 497-502); maíz (Pedersen *et al.* (1986) J. Biol. Chem. 261:6279; Kiriara *et al.* (1988) Gene 71:359); y arroz (Musumura *et al.* (1989) Plant Mol. Biol. 12:123) podrían usarse. Otros genes agrónomicamente importantes codifican látex, Flourey 2, factores de crecimiento, factores de almacenamiento en semillas y factores de transcripción.

Los genes de resistencia a insectos pueden conferir resistencia a plagas tales como gusano de la raíz, gusano cortador y barrenador del maíz europeo, que provocan daños importantes a los cultivos, lo que da como resultado que sean un gran obstáculo para la producción. Dichos genes incluyen, por ejemplo, genes de proteínas tóxicas de *Bacillus thuringiensis* (patentes de Estados Unidos n.º 5.366.892; 5.747.450; 5.736.514; 5.723.756; 5.593.881; y Geiser *et al.* (1986) Gene 48:109).

Los genes que codifican rasgos de resistencia a enfermedades incluyen genes de detoxificación, tales como contra fumonósina (patente de Estados Unidos n.º 5.792.931); y genes de avirulencia (avr) y de resistencia a enfermedades (R) (Jones *et al.* (1994) Science 266:789; Martin *et al.* (1993) Science 262:1432; y Mindrinos *et al.* (1994) Cell 78:1089).

Los rasgos de resistencia a herbicidas pueden incluir genes que codifican resistencia a herbicidas que actúan inhibiendo la acción de la acetolactato sintasa (ALS), en particular los herbicidas de tipo sulfonilurea (por ejemplo, el gen de la acetolactato sintasa (ALS) que contiene mutaciones que dan lugar a dicha resistencia, en particular las mutaciones S4 y/o Hra), genes que codifican resistencia a herbicidas que actúan inhibiendo la acción de la glutamina sintasa, tal como fosfotricina o basta (por ejemplo, el gen bar), glifosato (por ejemplo, el gen EPSPS y el gen GAT; véase, por ejemplo, la publicación de Estados Unidos n.º 20040082770 y el documento WO 03/092360) u otros genes de este tipo conocidos en la técnica. El gen bar codifica resistencia al herbicida basta, el gen nptII codifica resistencia a los antibióticos kanamicina y geneticina, y los mutantes del gen ALS codifican resistencia al herbicida clorsulfurona.

También pueden codificarse genes de esterilidad en un casete de expresión y proporcionar una alternativa al descope físico. Ejemplos de genes usados de estas maneras incluyen genes con preferencia de tejido masculino y genes con fenotipos de esterilidad masculina tales como QM, descrito en la patente de Estados Unidos n.º 5.583.210. Otros genes incluyen cinasas y aquellos que codifican compuestos tóxicos para el desarrollo de gametófitos masculinos o femeninos.

La calidad del grano se refleja en rasgos tales como los niveles y tipos de aceites, saturados e insaturados, la calidad y cantidad de aminoácidos esenciales, y los niveles de celulosa. En el maíz, se describen proteínas hordotionina modificadas en la patente de Estados Unidos n.º 5.703.049, 5.885.801, 5.885.802 y 5.990.389.

También pueden codificarse rasgos comerciales en un gen o genes para una composición de rasgos mejorada, por ejemplo, almidón para producción de etanol, o expresión potenciada de proteínas. Otro uso comercial importante de las plantas transformadas es la producción de polímeros y bioplásticos, tal como se describe en la patente de Estados Unidos n.º 5.602.321. Genes tales como β -cetotiolasa, PHBasa (polihidroxibutirato sintasa), y acetoacetyl-CoA reductasa (véase Schubert *et al.* (1988) J. Bacteriol. 170:5837-5847) facilitan la expresión de polihidroxialcanoatos (PHA).

Los productos exógenos incluyen enzimas y productos vegetales, así como aquellos de otras fuentes, incluyendo procariotas y otros eucariotas. Dichos productos incluyen enzimas, cofactores y hormonas. El nivel de proteínas, particularmente proteínas modificadas que tienen distribución mejorada de aminoácidos para mejorar el valor nutritivo de la planta, puede aumentarse. Esto se consigue mediante la expresión de dichas proteínas que tienen un contenido potenciado de aminoácidos.

En un aspecto, rasgos agronómicos adicionales de interés que pueden introducirse en un tejido diana con eficacia y velocidad aumentada son rasgos tales como producción aumentada u otros rasgos que proporcionan un valor aumentado de la planta incluyendo, por ejemplo, calidad mejorada de la semilla. Son de particular interés rasgos que proporcionan eficacia del uso de agua o tolerancia a las sequías mejorada o potenciada, tolerancia a agresión osmótica, tolerancia a agresión por alta salinidad, tolerancia a agresión térmica, tolerancia potenciada al frío, incluyendo tolerancia a germinación en frío, producción aumentada, eficacia del uso de nitrógeno potenciada, crecimiento y desarrollo temprano de la planta, crecimiento y desarrollo de la planta tardío, contenido potenciado de proteínas en la planta y producción potenciada de aceite en las semillas.

Muchos rasgos agronómicos pueden afectar a la "producción" incluyendo, sin limitación, la altura de planta, el número de vainas, la posición de las vainas en la planta, el número de internudos, la incidencia de desintegración de la vaina, tamaño del grano, eficacia de nodulación y fijación de nitrógeno, eficacia de asimilación de nutrientes, resistencia a agresión biótica y abiótica, asimilación de carbono, arquitectura de la planta, resistencia a encamado, porcentaje de germinación de las semillas, vigor de las plántulas y rasgos juveniles. Otros rasgos que pueden afectar a la producción incluyen eficacia de germinación (incluyendo germinación en condiciones de agresión), tasa de crecimiento (incluyendo tasa de crecimiento en condiciones de agresión), número de espigas, número de semillas por espiga, tamaño de la semilla, composición de la semilla (almidón, aceite, proteínas) y características del llenado de la semilla. También es de interés la generación de plantas transgénicas que demuestran propiedades fenotípicas deseables que puedan conferir o no un aumento en la producción global de la planta. Dichas propiedades incluyen potenciación de la morfología de la planta, la fisiología de la planta o componentes mejorados de la semilla madura recogida de la planta transgénica.

La "producción aumentada" de una planta transgénica de la presente divulgación puede evidenciarse y medirse de varias maneras, incluyendo comprobación del peso, número de semillas por planta, peso de las semillas, número de semillas por área unitaria (es decir, semillas o peso de las semillas, por 4050 m² (acre)), fanegas por 4050 m² (acre), toneladas por 4050 m² (acre) y kilogramos por hectárea. Por ejemplo, la producción de maíz puede medirse como la producción de granos de maíz con cáscara por unidad de área de producción, por ejemplo, en fanegas por 4050 m² (acre) o toneladas métricas por hectárea, a menudo presentada en una base ajustada a la humedad, por ejemplo, a un 15,5 % de humedad. La producción aumentada puede ser el resultado de la utilización mejorada de compuestos bioquímicos clave, tales como nitrógeno, fósforo y carbohidratos, o de la tolerancia mejorada a las agresiones ambientales, tales como frío, calor, sequía, salinidad y ataque por plagas o patógenos. También puede usarse ADN recombinante potenciador de rasgos para proporcionar plantas transgénicas que tienen crecimiento y desarrollo mejorado y, finalmente, producción aumentada, como resultado de la expresión modificada de reguladores del crecimiento de la planta o la modificación del ciclo celular o las rutas de fotosíntesis.

Muchos rasgos agronómicos pueden afectar a la "producción" incluyendo, sin limitación, la altura de planta, el número de vainas, la posición de las vainas en la planta, el número de internudos, la incidencia de desintegración de la vaina, tamaño del grano, eficacia de nodulación y fijación de nitrógeno, eficacia de asimilación de nutrientes, resistencia a agresión biótica y abiótica, asimilación de carbono, arquitectura de la planta, resistencia a encamado, porcentaje de germinación de las semillas, vigor de las plántulas y rasgos juveniles. Otros rasgos que pueden afectar a la producción incluyen, aunque sin limitación, eficacia de germinación (incluyendo germinación en condiciones de agresión), tasa de crecimiento (incluyendo tasa de crecimiento en condiciones de agresión), número de espigas, número de semillas por espiga, tamaño de la semilla, composición de la semilla (almidón, aceite, proteínas) y características del llenado de la semilla. También es de interés la generación de plantas transgénicas que demuestran propiedades fenotípicas deseables que puedan conferir o no un aumento en la producción global de la planta. Dichas propiedades incluyen, aunque sin limitación,

potenciación de la morfología de la planta, la fisiología de la planta y componentes mejorados de la semilla madura recogida de la planta transgénica.

IV. Métodos para suprimir genes

En un aspecto, los métodos y composiciones divulgadas pueden usarse para introducir en una célula vegetal polinucleótidos útiles para la supresión génica de un gen diana en una planta derivada de la célula vegetal. La reducción de la actividad de genes específicos (también conocido como silenciamiento génico o supresión génica) es deseable para varios aspectos de manipulación genética en plantas. Los expertos en la materia conocen bien muchas técnicas para silenciamiento génico incluyendo, aunque sin limitación, tecnología de antisentido (véase, por ejemplo, Sheehy *et al.* (1988) Proc. Natl. Acad. Sci. USA 85:8805-8809; y patentes de Estados Unidos n.º 5.107.065; 5.453.566; y 5.759.829); cosupresión (por ejemplo, Taylor (1997) Plant Cell 9:1245; Jorgensen (1990) Trends Biotech. 8(12):340-344; Flavell (1994) Proc. Natl. Acad. Sci. USA 91:3490-3496; Finnegan *et al.* (1994) Bio/Technology 12: 883-888; y Neuhuber *et al.* (1994) Mol. Gen. Genet. 244:230-241); interferencia de ARN (Napoli *et al.* (1990) Plant Cell 2:279-289; patente de Estados Unidos n.º 5.034.323; Sharp (1999) Genes Dev. 13:139-141; Zamore *et al.* (2000) Cell 101:25-33; Javier (2003) Nature 425:257-263; y Montgomery *et al.* (1998) Proc. Natl. Acad. Sci. USA 95:15502-15507), silenciamiento génico inducido por virus (Burton, *et al.* (2000) Plant Cell 12:691-705; y Baulcombe (1999) Curr. Op. Plant Bio. 2:109-113); ribozimas específicas de ARN diana (Haseloff *et al.* (1988) Nature 334: 585-591); estructuras de horquilla (Smith *et al.* (2000) Nature 407:319-320; documentos WO 99/53050; WO 02/00904; y WO 98/53083); ribozimas (Steinecke *et al.* (1992) EMBO J. 11:1525; patente de Estados Unidos n.º 4.987.071; y Perriman *et al.* (1993) Antisense Res. Dev. 3:253); modificación dirigida mediada por oligonucleótidos (por ejemplo, documentos WO 03/076574 y WO 99/25853); moléculas dirigidas a dedos de Zn (por ejemplo, documentos WO 01/52620; WO 03/048345; y WO 00/42219); microARN artificiales (documento US8106180; Schwab *et al.* (2006) Plant Cell 18:1121-1133); y otros métodos o combinaciones de los métodos anteriores conocidos por los expertos en la materia.

V. Métodos para introducir tecnologías de edición genómica en plantas

En un aspecto, los métodos y composiciones divulgadas pueden usarse para introducir en una célula vegetal polinucleótidos útiles para abordar un sitio específico para su modificación en el genoma de una planta derivada de la célula vegetal. Las modificaciones específicas de sitio que pueden introducirse con los métodos y composiciones divulgadas incluyen las producidas usando cualquier método para introducir una modificación específica de sitio incluyendo, aunque sin limitación, mediante el uso de oligonucleótidos de reparación génica (por ejemplo, publicación de Estados Unidos 2013/0019349), o mediante el uso de tecnologías de rotura bicatenaria tales como TALEN, meganucleasas, nucleasas de dedos de cinc y CRISPR-Cas. Por ejemplo, los métodos y composiciones divulgadas pueden usarse para introducir un sistema de CRISPR-Cas en células vegetales, con fines de modificación genómica de una secuencia diana en el genoma de una célula vegetal o planta derivada de la célula vegetal, para seleccionar plantas, para eliminar una base o una secuencia, para edición génica y para insertar un polinucleótido de interés en el genoma de una planta derivada de una célula vegetal. Por tanto, los métodos y composiciones divulgadas pueden usarse conjuntamente con un sistema de CRISPR-Cas para proporcionar un sistema eficaz para modificar o alterar sitios diana y nucleótidos de interés dentro del genoma de una planta, célula vegetal o semilla.

En un aspecto, la presente divulgación comprende métodos y composiciones para su transformación, en los que el método comprende las etapas de: (a) poner en contacto una célula vegetal con una cepa de *Agrobacterium* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Agrobacterium* *virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*; un segundo vector que puede expresar un nucleótido guía; y una tercera construcción que puede expresar una endonucleasa Cas, en el que el nucleótido guía y la endonucleasa Cas pueden formar un complejo que posibilita que la endonucleasa Cas introduzca una rotura bicatenaria en el sitio diana; (b) cocultivar el tejido con la *Agrobacterium*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés.

En un aspecto, el gen de la endonucleasa Cas es una endonucleasa Cas9 optimizada para plantas, en el que la endonucleasa Cas9 optimizada para plantas puede unirse a y crear una rotura bicatenaria en una secuencia diana genómica del genoma de la planta.

La endonucleasa Cas se guía mediante el nucleótido guía para reconocer e introducir opcionalmente una rotura bicatenaria en un sitio diana específico en el genoma de una célula vegetal. El sistema de CRISPR-Cas proporciona un sistema eficaz para modificar sitios diana dentro del genoma de una planta, célula vegetal o semilla. Se proporcionan además métodos y composiciones que emplean un sistema de polinucleótido guía/endonucleasa Cas para proporcionar un sistema eficaz para modificar sitios diana dentro del genoma de una célula y para editar una secuencia de nucleótidos en el genoma de una célula. Una vez identificado el sitio diana genómico, puede emplearse una diversidad de métodos para modificar adicionalmente los sitios diana, de modo que contienen una diversidad de polinucleótidos de interés. Las composiciones y métodos divulgados pueden usarse para introducir un sistema de CRISPR-Cas para editar una secuencia de nucleótidos en el genoma de una célula. La secuencia de nucleótidos a editar (la secuencia de nucleótidos de interés) puede ubicarse dentro o fuera de un sitio diana que se reconoce por una endonucleasa Cas.

Los locus de CRISPR (repeticiones palindrómicas cortas agrupadas intercaladas de forma regular) (también conocidos como SPIDR-repeticiones directas intercaladas SPacer) constituyen una familia de locus de ADN recientemente descritos. Los locus de CRISPR consisten en repeticiones de ADN cortas y muy conservadas (típicamente de 24 a 40 pb, repetidas de 1 a 140 veces - también denominadas repeticiones CRISPR) que son parcialmente palindrómicas. Las secuencias repetidas (habitualmente específicas para una especie) están intercaladas por secuencias variables de longitud constante (típicamente de 20 a 58 dependiendo del locus CRISPR (documento WO2007/025097, publicado el 1 de marzo de 2007).

Los locus de CRISPR se reconocieron en primer lugar en *E. coli* (Ishino *et al.* (1987) J. Bacterial. 169:5429-5433; Nakata *et al.* (1989) J. Bacterial. 171:3553-3556). Se han identificado repeticiones de secuencia cortas intercaladas similares en *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena* y *Mycobacterium tuberculosis* (Groenen *et al.* (1993) Mol. Microbiol. 10:1057-1065; Hoe *et al.* (1999) Emerg. Infect. Dis. 5:254-263; Masepohl *et al.* (1996) Biochim. Biophys. Acta 1307:26-30; Mojica *et al.* (1995) Mol. Microbiol. 17:85-93). Los locus de CRISPR difieren de otras SSR en la estructura de las repeticiones, que se han denominado repeticiones espaciadas de forma regular cortas (SRSR) (Janssen *et al.* (2002) OMICS J. Integ. Biol. 6:23-33; Mojica *et al.* (2000) Mol. Microbiol. 36:244-246). Las repeticiones son elementos cortos que aparecen en grupos, que siempre están espaciados de forma regular por secuencias variables de longitud constante (Mojica *et al.* (2000) Mol. Microbiol. 36:244-246).

El gen Cas incluye un gen que está generalmente acoplado, asociado o cerca de, o en las cercanías de locus de CRISPR flanqueantes. Las expresiones "gen Cas" y "gen (Cas) asociado a CRISPR" se usan indistintamente en este documento. Se presenta una revisión extensa de la familia de la proteína Cas en Haft *et al.* (2005) Computational Biology, PLoS Comput Biol 1 (6): e60. doi:10.1371 / journal.pcbi.0010060.

Además de las cuatro familias génicas descritas inicialmente, se han descrito 41 familias génicas (Cas) asociadas a CRISPR adicionales en el documento WO/2015/026883. Esta referencia muestra que los sistemas de CRISPR pertenecen a clases diferentes, con diferentes patrones de repetición, conjuntos de genes e intervalos de especie. El número de genes Cas en un locus CRISPR dado puede variar entre especies. La endonucleasa Cas se refiere a una proteína Cas codificada por un gen Cas, en la que la proteína Cas puede introducir una rotura bicatenaria en una secuencia diana de ADN. La endonucleasa Cas se guía mediante el polinucleótido guía para reconocer e introducir opcionalmente una rotura bicatenaria en un sitio diana específico en el genoma de una célula. Como se usa en este documento, la expresión "sistema de polinucleótido guía/endonucleasa Cas" incluye un complejo de una endonucleasa Cas y un polinucleótido guía que puede introducir una rotura bicatenaria en una secuencia diana de ADN. La endonucleasa Cas desenrolla el ADN bicatenario en cercana proximidad al sitio diana genómico y escinde ambas hebras de ADN tras el reconocimiento de una secuencia diana mediante un nucleótido guía, pero solamente si el motivo adyacente de protoespaciador (PAM) correcto está orientado aproximadamente en el extremo 3' de la secuencia diana (véase la figura 2A y la figura 2B del documento WO/2015/026883, publicado el 26 de febrero de 2015).

En un aspecto, el gen de la endonucleasa Cas es una endonucleasa Cas9, tal como, aunque sin limitación, genes Cas9 enumerados y divulgados en el documento WO2007/025097, publicado el 1 de marzo de 2007. En otro aspecto, el gen de la endonucleasa Cas es endonucleasa Cas9 optimizada para plantas, maíz o soja, tal como, aunque sin limitación, las divulgadas en el documento WO/2015/026883. En otro aspecto, el gen de la endonucleasa Cas está unido de forma funcional a una señal de dirección nuclear de SV40 en dirección 5' de la región de codones de Cas y una señal de localización nuclear VirD2 bipartita (Tinland *et al.* (1992) Proc. Natl. Acad. Sci. USA 89:7442-6) en dirección 3' de la región de codones de Cas.

En un aspecto, el gen de la endonucleasa Cas es un gen de la endonucleasa Cas9, o cualquier fragmento funcional o variante del mismo, divulgado en el documento WO/2015/026883.

Las expresiones "fragmento funcional", "fragmento que es funcionalmente equivalente" y "fragmento funcionalmente equivalente" se usan indistintamente en este documento. Estas expresiones se refieren a una parte o subsecuencia de la secuencia de la endonucleasa Cas de la presente divulgación en que se retiene la capacidad de crear una rotura bicatenaria.

Las expresiones "variante funcional", "variante que es funcionalmente equivalente" y "variante funcionalmente equivalente" se usan indistintamente en este documento. Estas expresiones se refieren a una variante de la endonucleasa Cas de la presente divulgación en que se retiene la capacidad de crear una rotura bicatenaria. Los fragmentos y variantes y derivados pueden obtenerse mediante métodos tales como mutagénesis dirigida al sitio y construcción sintética.

En un aspecto, el gen de la endonucleasa Cas es un gen Cas9 de *Streptococcus pyogenes* con codones optimizados para plantas que puede reconocer cualquier secuencia genómica de la forma N(12-30)NGG, que puede abordarse.

Las endonucleasas son enzimas que escinden el enlace fosfodiéster dentro de una cadena polinucleotídica, e incluyen endonucleasas de restricción que escinden ADN en sitios específicos sin dañar las bases. Las endonucleasas de restricción incluyen endonucleasas de tipo I, tipo II, tipo III y tipo IV, que incluyen además subtipos. En los sistemas de tipo I y tipo III, están contenidas tanto la actividad metilasa como la actividad de restricción en un solo complejo. Las endonucleasas también incluyen meganucleasas, también conocidas como endonucleasas de asentamiento (HEasas), que como las endonucleasas de restricción, se unen y cortan en un sitio de reconocimiento específico, sin embargo, los sitios de reconocimiento para las meganucleasas son típicamente más largos, de aproximadamente 18 pb o más.

(Solicitud de patente PCT/US 12/30061 presentada el 22 de marzo de 2012). Las meganucleasas se han clasificado en cuatro familias basadas en los motivos de secuencia conservados, las familias son las familias de secuencia LAGLIDADG, GIY-YIG, H-N-H e His-Cys. Estos motivos participan en la coordinación de iones metálicos e hidrólisis de enlaces fosfodiéster. Las meganucleasas son destacables por sus sitios de reconocimiento largos, y por tolerar algunos polimorfismos de secuencia en sus sustratos de ADN. La convención de nomenclatura para las meganucleasas es similar a la convención para otras endonucleasas de restricción. Las meganucleasas también se caracterizan por el prefijo F-, I- o PI- para enzimas codificadas por ORF independientes, intrones e inteínas, respectivamente. Una etapa en el proceso de reconocimiento implica la escisión polinucleotídica en o cerca del sitio de reconocimiento. Esta actividad de escisión puede usarse para producir una rotura bicatenaria. Para revisiones de recombinasas específicas de sitio y sus sitios de reconocimiento, véase, Sauer (1994) *Curr Op Biotechnol* 5:521-7; y Sadowski (1993) *FASEB* 7:760-7. En algunos ejemplos, la recombinasa es de las familias integrasa o resolvasa. Las nucleasas efectoras TAL son una clase nueva de nucleasas específicas de secuencia que pueden usarse para generar roturas bicatenarias en secuencias diana específicas en el genoma de una planta u otro organismo. (Miller, *et al.* (2011) *Nature Biotechnology* 29:143-148). Las nucleasas con dedos de cinc (ZFN) son agentes que inducen roturas bicatenarias genomanipuladas que comprenden un dominio de unión al ADN de tipo dedos de cinc y un dominio de un agente inductor de rotura bicatenaria. La especificidad por el sitio de reconocimiento la confiere el dominio de dedos de cinc, que típicamente comprende dos, tres o cuatro dedos de cinc, por ejemplo, que tiene una estructura C2H2, sin embargo, se conocen otras estructuras de dedos de cinc y se han genomanipulado. Los dominios de dedos de cinc son susceptibles al diseño de polipéptidos que se unen específicamente a una secuencia de reconocimiento polinucleotídica seleccionada. Las ZFN incluyen un dominio de dedos de cinc de unión a ADN genomanipulado unido a un dominio de endonucleasa no específico, por ejemplo, dominio de nucleasa de una endonucleasa de tipo Ms tal como FokI. Pueden fusionarse funcionalidades adicionales al dominio de unión de dedos de cinc, incluyendo dominios activadores transcripcionales, dominios represores de la transcripción y metilasas. En algunos ejemplos, se requiere dimerización del dominio nucleasa para la actividad de escisión. Cada dedo de cinc reconoce tres pares de bases consecutivas en el ADN diana. Por ejemplo, un dominio de 3 dedos reconoce una secuencia de 9 nucleótidos contiguos para la unión tras la dimerización, mientras que se usan dos conjuntos de tripletes de dedos de cinc para unirse a una secuencia de reconocimiento de 18 nucleótidos.

Las bacterias y arqueobacterias han desarrollado defensas inmunitarias adaptativas llamadas sistemas asociados a repeticiones palindrómicas cortas agrupadas intercaladas de forma regular (CRISPR)/CRISPR (Cas) que usan ARN corto para dirigir la degradación de ácidos nucleicos exógenos (documento WO2007/025097, publicado el 1 de marzo de 2007). El sistema de CRISPR/Cas de tipo II de bacterias emplea un ARNcr y ARNcrtra para guiar la endonucleasa Cas a su diana de ADN. El ARNcr (ARN de CRISPR) contiene la región complementaria a una hebra de la diana de ADN bicatenario y forma pares de bases con el ARNcrtra (ARN de CRISPR de activación en *trans*) que forma un ARN bicatenario que dirige la endonucleasa Cas para escindir la diana de ADN.

Como se usa en este documento, la expresión "nucleótido guía" se refiere a una fusión sintética de dos moléculas de ARN, un ARNcr (ARN de CRISPR) que comprende un dominio de dirección variable y un ARNcrtra. En un aspecto, el nucleótido guía comprende un dominio de dirección variable de secuencias de 12 a 30 nucleótidos y un fragmento de ARN que puede interactuar con una endonucleasa Cas.

Como se usa en este documento, la expresión "polinucleótido guía", se refiere a una secuencia polinucleotídica que puede formar un complejo con una endonucleasa Cas y posibilita que la endonucleasa Cas reconozca y opcionalmente escinda un sitio diana de ADN. El polinucleótido guía puede ser una molécula sencilla o una molécula doble. La secuencia polinucleotídica guía puede ser una secuencia de ARN, una secuencia de ADN o una combinación de las mismas (una secuencia de combinación de ARN-ADN). Opcionalmente, el polinucleótido guía puede comprender al menos una modificación de nucleótido, enlace o unión fosfodiéster tal como, aunque sin limitación, ácido nucleico bloqueado (LNA), 5-metil dC, 2,6-diaminopurina, 2'-fluoro A, 2'-fluoro U, 2'-O-metil ARN, enlace fosforotioato, enlace a una molécula de colesterol, enlace a una molécula de polietilenglicol, enlace a una molécula espaciadora 18 (cadena de hexaetilenglicol) o enlace covalente 5' a 3' que provoca circularización. Un polinucleótido guía que únicamente comprende ácidos ribonucleicos también se denomina "nucleótido guía".

El polinucleótido guía puede ser una molécula doble (también denominada polinucleótido guía bicatenario) que comprende un primer dominio de secuencia de nucleótidos (denominado dominio de dirección variable o dominio VT) que es complementario a una secuencia de nucleótidos en un ADN diana y un segundo dominio de secuencia de nucleótidos (denominado dominio de reconocimiento de endonucleasa Cas o dominio CER) que interactúa con un polipéptido de endonucleasa Cas. El dominio CER del polinucleótido guía de molécula doble comprende dos moléculas diferentes que hibridan a lo largo de una región de complementariedad. Las dos moléculas diferentes pueden ser secuencias de ARN, de ADN y/o de combinación de ARN-ADN. En un aspecto, la primera molécula del polinucleótido guía bicatenario que comprende un dominio VT ligado a un dominio CER se denomina "ADNcr" (cuando está compuesto de un tramo contiguo de nucleótidos de ADN) o "ARNcr" (cuando está compuesto de un tramo contiguo de nucleótidos de ARN) o "ADN-ARNcr" (cuando está compuesto de una combinación de nucleótidos de ADN y ARN). El nucleótido cr puede comprender un fragmento del ARNcr que existe de forma natural en bacterias y arqueobacterias. En un aspecto, el tamaño del fragmento del ARNcr que existe de forma natural en bacterias y arqueobacterias que está presente en un nucleótido cr divulgado en este documento puede variar de, aunque sin limitación, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos.

En un aspecto, la segunda molécula del polinucleótido guía bicatenario que comprende un dominio CER se denomina "ARNcrtra" (cuando está compuesto de un tramo contiguo de nucleótidos de ARN) o "ADNcrtra" (cuando está compuesto de un tramo contiguo de nucleótidos de ADN) o "ADN-ARNcrtra" (cuando está compuesto de una combinación de nucleótidos de ADN y ARN). En un aspecto, el ARN que guía el complejo de endonucleasa Cas9 de ARN es un ARN bicatenario que comprende un ARNcr-ARNcrtra bicatenario.

El polinucleótido guía también puede ser una molécula sencilla que comprende un primer dominio de secuencia de nucleótidos (denominado dominio de dirección variable o dominio VT) que es complementario a una secuencia de nucleótidos en un ADN diana y un segundo dominio de nucleótidos (denominado dominio de reconocimiento de endonucleasa Cas o dominio CER) que interactúa con un polipéptido de endonucleasa Cas. Por "dominio" se entiende un tramo contiguo de nucleótidos que puede ser de secuencia de ARN, ADN y/o combinación de ARN-ADN. El dominio VT y/o el dominio CER de un polinucleótido guía sencillo puede comprender una secuencia de ARN, una secuencia de ADN o una secuencia de combinación de ARN-ADN. En un aspecto, el polinucleótido guía sencillo comprende un nucleótido (que comprende un dominio VT ligado a un dominio CER) ligado a un nucleótido (que comprende un dominio CER), en el que el enlace es una secuencia de nucleótidos que comprende una secuencia de ARN, una secuencia de ADN o una secuencia de combinación de ARN-ADN. El polinucleótido guía sencillo compuesto de secuencias del nucleótido y el nucleótido puede denominarse "nucleótido guía sencillo" (cuando está compuesto de un tramo contiguo de nucleótidos de ARN) o "ADN guía sencillo" (cuando está compuesto de un tramo contiguo de nucleótidos de ADN) o "nucleótido-ADN guía sencillo" (cuando está compuesto de una combinación de nucleótidos de ARN y ADN). En un aspecto de la divulgación, el nucleótido guía sencillo comprende un ARNc o fragmento de ARNc y un ARNcrtra o fragmento de ARNcrtra del sistema de CRISPR/Cas de tipo II que puede formar un complejo con una endonucleasa Cas de tipo II, en el que el complejo de nucleótido guía y endonucleasa Cas puede dirigir la endonucleasa Cas a un sitio diana genómico de la planta, posibilitando que la endonucleasa Cas introduzca una rotura bicatenaria en el sitio diana genómico. Un aspecto de usar un polinucleótido guía sencillo frente a un polinucleótido guía bicatenario es que tiene que prepararse solamente un casete de expresión para expresar el polinucleótido guía sencillo.

La expresión "dominio de dirección variable" o "dominio VT" se usa indistintamente en este documento e incluye una secuencia de nucleótidos que es complementaria a una hebra (secuencia de nucleótidos) de un sitio diana de ADN bicatenario. El % de complementación entre el primer dominio de secuencia de nucleótidos (dominio VT) y la secuencia diana puede ser de al menos un 50%, 51 %, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61 %, 62%, 63%, 63%, 65%, 66%, 67%, 68%, 69%, 70%, 71 %, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81 %, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%. El dominio diana variable puede ser de al menos 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 nucleótidos de longitud. En un aspecto, el dominio de dirección variable comprende un tramo contiguo de 12 a 30 nucleótidos. El dominio de dirección variable puede estar compuesto de una secuencia de ADN, una secuencia de ARN, una secuencia de ADN modificada, una secuencia de ARN modificada o cualquier combinación de las mismas.

La expresión "dominio de reconocimiento de endonucleasa Cas" o "dominio CER" de un polinucleótido guía se usa indistintamente en este documento e incluye una secuencia de nucleótidos (tal como un segundo dominio de secuencia de nucleótidos de un polinucleótido guía), que interactúa con un polipéptido de endonucleasa Cas. El dominio CER puede estar compuesto de una secuencia de ADN, una secuencia de ARN, una secuencia de ADN modificada, una secuencia de ARN modificada (véanse, por ejemplo, las modificaciones descritas en este documento), o cualquier combinación de las mismas.

La secuencia de nucleótidos que conecta el nucleótido y el nucleótido de un polinucleótido guía sencillo puede comprender una secuencia de ARN, una secuencia de ADN o una secuencia de combinación de ARN-ADN. En un aspecto, la secuencia de nucleótidos que conecta el nucleótido y el nucleótido de un polinucleótido guía sencillo puede ser de al menos 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100 nucleótidos de longitud. En otro aspecto, la secuencia de nucleótidos que conecta el nucleótido y el nucleótido de un polinucleótido guía sencillo puede comprender una secuencia de tetrabucle tal como, aunque sin limitación, una secuencia de tetrabucle GAAA.

La modificación de la secuencia de nucleótidos del polinucleótido guía, el dominio VT y/o el dominio CER puede seleccionarse de, aunque sin limitación, el grupo que consiste en un capuchón 5', una cola poliadenilada 3', una secuencia de ribointerruptor, una secuencia de control de la estabilidad, una secuencia que forma un ARNc bicatenario, una modificación o secuencia que dirige el polinucleótido guía a una localización subcelular, una modificación o secuencia que proporciona rastreo, una modificación o secuencia que proporciona un sitio de unión para proteínas, un ácido nucleico bloqueado (LNA), un nucleótido 5-metil dC, un nucleótido 2,6-diaminopurina, un nucleótido 2'-fluoro A, un nucleótido 2'-fluoro U; un nucleótido de 2'-O-metil ARN, un enlace fosforotioato, unión a una molécula de colesterol, unión a una molécula de polietilenglicol, unión a una molécula espaciadora 18, una unión covalente 5' a 3' o cualquier combinación de las mismas. Estas modificaciones pueden provocar al menos un rasgo característico beneficioso adicional, en las que el rasgo característico beneficioso adicional se selecciona del grupo de una estabilidad modificada o regulada, una dirección subcelular, rastreo, un marcador fluorescente, un sitio de unión para una proteína o complejo proteínico, afinidad de unión

modificada a secuencia diana complementaria, resistencia modificada a degradación celular y permeabilidad celular aumentada.

5 En aspectos, el nucleótido guía y la endonucleasa Cas pueden formar un complejo que posibilita que la endonucleasa Cas introduzca una rotura bicatenaria en un sitio diana de ADN.

En aspectos, el dominio diana variable es de 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 nucleótidos de longitud.

10 En aspectos, el nucleótido guía comprende un ARNc (o fragmento de ARNc) y un ARNcrtra o (fragmento de ARNcrtra) del sistema de CRISPR/Cas de tipo II que puede formar un complejo con una endonucleasa Cas de tipo II, en el que el complejo de nucleótido guía y endonucleasa Cas puede dirigir la endonucleasa Cas a un sitio diana genómico de la planta, posibilitando que la endonucleasa Cas introduzca una rotura bicatenaria en el sitio diana genómico. En un aspecto, el nucleótido guía puede introducirse en una planta o célula vegetal directamente usando cualquier método conocido en la técnica tal como, aunque sin limitación, bombardeo de partículas o aplicaciones tópicas.

15 En aspectos, el nucleótido guía puede introducirse indirectamente introduciendo una molécula de ADN recombinante que comprende la secuencia de ADN guía correspondiente unida de forma funcional a un promotor específico de plantas que puede transcribir el nucleótido guía en la célula vegetal. La expresión "ADN guía correspondiente" incluye una molécula de ADN que es idéntica a la molécula de ARN, pero tiene una "T" sustituida en el lugar de cada "U" de la molécula de ARN.

20 En aspectos, el nucleótido guía se introduce mediante bombardeo de partículas o usando los métodos y composiciones divulgadas para transformación con *Agrobacterium* de una construcción de ADN recombinante que comprende el ADN guía correspondiente unido de forma funcional a un promotor de polimerasa III U6 de plantas.

25 En aspectos, el ARN que guía el complejo de endonucleasa Cas9 de ARN es un ARN bicatenario que comprende un ARNcr-ARNcrtra bicatenario. Una ventaja de usar un nucleótido guía frente a un ARNcr-ARNcrtra bicatenario es que tiene que prepararse solamente un casete de expresión para expresar el nucleótido guía fusionado.

30 La expresiones "sitio diana", "secuencia diana", "ADN diana", "locus diana", "sitio diana genómico", "secuencia diana genómica" y "locus diana genómico" se usan indistintamente en este documento y se refieren a una secuencia polinucleotídica en el genoma (que incluye ADN de cloroplastos y mitocondrial) de una célula vegetal en que se induce una rotura bicatenaria en el genoma de la célula vegetal por una endonucleasa Cas. El sitio diana puede ser un sitio endógeno en el genoma de la planta o, como alternativa, el sitio diana puede ser heterólogo a la planta y, de ese modo, no existir de forma natural en el genoma, o el sitio diana puede encontrarse en una localización genómica heteróloga en comparación con donde existe de forma natural.

35 Como se usan en este documento, las expresiones "secuencia diana endógena" y "secuencia diana natural" se usan indistintamente en este documento para referirse a una secuencia diana que es endógena o natural al genoma de una planta y está en la posición endógena o natural de esa secuencia diana en el genoma de la planta. En un aspecto, el sitio diana puede ser similar a un sitio de reconocimiento de ADN o sitio diana que se reconoce específicamente y/o se une por un agente que induce rotura bicatenaria tal como una endonucleasa LIG3-4 (publicación de patente de Estados Unidos 2009-0133152 A1 (publicada el 21 de mayo de 2009) o una meganucleasa MS26++ (solicitud de patente de Estados Unidos 13/526912 presentada el 19 de junio de 2012).

40 Un "sitio diana artificial" o "secuencia diana artificial" se usan indistintamente en este documento y se refieren a una secuencia diana que se ha introducido en el genoma de una planta. Dicha secuencia diana artificial puede ser idéntica en secuencia a una secuencia diana endógena o natural en el genoma de una planta, pero estar localizada en una posición diferente (es decir, una posición no endógena o no natural) en el genoma de una planta.

45 Un "sitio diana alterado", "secuencia diana alterada", "sitio diana modificado" y "secuencia diana modificada" se usan indistintamente en este documento y se refieren a una secuencia diana como se divulga en este documento que comprende al menos una alteración en comparación con una secuencia diana no alterada. Dichas "alteraciones" incluyen, por ejemplo: (i) remplazo de al menos un nucleótido, (ii) una eliminación de al menos un nucleótido, (iii) una inserción de al menos un nucleótido o (iv) cualquier combinación de (i) - (iii).

55 **VI. Métodos para introducir nucleótidos para integración específica de sitio**

En un aspecto, los métodos y composiciones divulgadas pueden usarse para introducir en una célula vegetal, con eficacia y velocidad aumentadas, polinucleótidos útiles para la integración dirigida de secuencias de nucleótidos en una planta derivada de la célula vegetal. Por ejemplo, los métodos y composiciones divulgadas pueden usarse para introducir casetes de transferencia que comprenden secuencias de nucleótidos de interés flanqueadas por sitios de recombinación no idénticos para transformar una planta que comprende un sitio diana. En un aspecto, el sitio diana contiene al menos un conjunto de sitios de recombinación no idénticos correspondientes a aquellos en el casete de transferencia. El intercambio de las secuencias de nucleótidos flanqueadas por los sitios de recombinación se logra mediante una recombinasa. Por tanto, los métodos y composiciones divulgadas pueden usarse para la introducción de casetes de transferencia para la integración dirigida de secuencias de nucleótidos, en la que los casetes de transferencia que están flanqueados por sitios de recombinación no idénticos se reconocen por una recombinasa que reconoce e implementa la recombinación en los

sitios de recombinación no idénticos. Por consiguiente, los métodos y composiciones divulgadas pueden usarse para mejorar la eficacia y velocidad de desarrollo de las plantas, derivadas de células vegetales, que contienen sitios de recombinación no idénticos.

En un aspecto, la presente divulgación proporciona además métodos para la transformación de una planta, en los que el método comprende introducir un polinucleótido de interés en un sitio diana en el genoma de una célula vegetal, comprendiendo el método las etapas de: (a) poner en contacto un explante de la planta con una cepa de *Agrobacterium* que comprende un primer vector que comprende: (i) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*; (ii) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.; (iii) un gen marcador de selección; y (iv) genes de virulencia de *Agrobacterium* *virB1-B11*; *virC1-C2*; *virD1-D2*; y genes *virG*, y un segundo vector que comprenden un casete de transferencia que comprende el polinucleótido de interés flanqueado por sitios de recombinación no idénticos; (b) cocultivar el tejido con la *Agrobacterium*; y (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés; en el que el explante deriva de una planta con un genoma que comprende un sitio diana flanqueado por sitios de recombinación no idénticos que corresponden a los sitios flanqueantes del casete de transferencia. El método puede comprender además proporcionar una recombinasa que reconoce e implementa la recombinación en los sitios de recombinación no idénticos, proporcionándose la recombinasa al explante vegetal, una plántula derivada del embrión somático o una planta derivada de una plántula derivada de una célula vegetal.

Por tanto, los métodos y composiciones divulgadas pueden comprender además composiciones y métodos para la integración direccional, dirigida de nucleótidos exógenos en una planta transformada. En un aspecto, los métodos divulgados usan sitios de recombinación novedosos en un sistema de dirección génica que facilita la dirección direccional de genes deseados y secuencias de nucleótidos en sitios de recombinación correspondientes previamente introducidos en el genoma de la planta diana.

En un aspecto, se introduce una secuencia de nucleótidos flanqueada por dos sitios de recombinación no idénticos en una o más células de un explante derivado del genoma del organismo diana que establece un sitio diana para la inserción de secuencias de nucleótidos de interés. Una vez establecida una planta estable o tejido cultivado, se introduce una segunda construcción, o secuencia de nucleótidos de interés, flanqueada por sitios de recombinación correspondientes como los que flanquean el sitio diana, en la planta o tejidos transformados de forma estable en presencia de una proteína recombinasa. Este proceso provoca el intercambio de las secuencias de nucleótidos entre los sitios de recombinación no idénticos del sitio diana y el casete de transferencia.

Se reconoce que la planta transformada preparada de esta manera puede comprender múltiples sitios diana; es decir, conjuntos de sitios de recombinación no idénticos. De esta manera, están disponibles múltiples manipulaciones del sitio diana en la planta transformada. Por sitio diana en la planta transformada se pretende una secuencia de ADN que se ha insertado en el genoma de la planta transformada y comprende sitios de recombinación no idénticos.

Ejemplos de sitios de recombinación para su uso en el método divulgado son conocidos en la técnica e incluyen sitios FRT (véase, por ejemplo, Schlake y Bode (1994) *Biochemistry* 33: 12746-12751; Huang *et al.* (1991) *Nucleic Acids Research* 19: 443-448; Paul D. Sadowski (1995) *In Progress in Nucleic Acid Research and Molecular Biology* vol. 51, pág. 53-91; Michael M. Cox (1989) *In Mobile DNA*, Berg y Howe (eds) American Society of Microbiology, Washington D. C., pág. 116-670; Dixon *et al.* (1995) 18: 449-458; Umlauf y Cox (1988) *The EMBO Journal* 7: 1845-1852; Buchholz *et al.* (1996) *Nucleic Acids Research* 24: 3118-3119; Kilby *et al.* (1993) *Trends Genet.* 9: 413-421; Rossant y Geagy (1995) *Nat. Med.* 1: 592-594; Albert *et al.* (1995) *The Plant J.* 7: 649-659; Bayley *et al.* (1992) *Plant Mol. Biol.* 18: 353-361; Odell *et al.* (1990) *Mol. Gen. Genet.* 223: 369-378; y Dale y Ow (1991) *Proc. Natl. Acad. Sci. USA* 88: 10558-10562.) ; Lox (Albert *et al.* (1995) *Plant J.* 7: 649-659; Qui *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91: 1706-1710; Stuurman *et al.* (1996) *Plant Mol. Biol.* 32: 901-913; Odell *et al.* (1990) *Mol. Gen. Gevet.* 223: 369-378; Dale *et al.* (1990) *Gene* 91: 79-85; y Bayley *et al.* (1992) *Plant Mol. Biol.* 18: 353-361.) El plásmido 2 μ , encontrado en la mayoría de cepas de origen natural de *Saccharomyces cerevisiae*, codifica una recombinasa específica de sitio que promueve una inversión del ADN entre dos repeticiones invertidas. Esta inversión desempeña una función central en la amplificación del número de copias del plásmido.

La proteína, denominada proteína FLP, cataliza acontecimientos de recombinación específica de sitio. El sitio de recombinación mínimo (FRT) se ha definido y contiene dos repeticiones invertidas de 13 pares de bases (pb) que rodean un espaciador asimétrico de 8 pb. La proteína FLP escinde el sitio en las uniones de las repeticiones y el espaciador y se une covalentemente al ADN mediante un 3'fosfato. Las recombinasas específicas de sitio como FLP escinden y vuelven a ligar el ADN en secuencias diana específicas, provocando una recombinación definida de manera precisa entre dos sitios idénticos. Para funcionar, el sistema necesita los sitios de recombinación y la recombinasa. No se necesitan factores auxiliares. Por tanto, el sistema completo puede insertarse y funcionar en células vegetales. El sistema de recombinación específico de sitio de FLP/FRT de levadura ha demostrado funcionar en plantas. Hasta la fecha, el sistema se ha utilizado para la escisión de ADN indeseado. Véase, Lyznik *et al.* (1993) *Nucleic Acid Res.* 21: 969-975. En contraste, la presente divulgación utiliza FRT no idénticas para el intercambio, dirección, disposición, inserción y control de la expresión de secuencias de nucleótidos en el genoma de la planta.

En un aspecto, se necesita un organismo transformado de interés, tal como un explante de una planta, que contiene un sitio diana integrado en su genoma. El sitio diana se caracteriza por estar flanqueado por sitios de recombinación no idénticos. Se requiere adicionalmente un casete de dirección que contenga una secuencia de nucleótidos flanqueada por correspondientes sitios de recombinación no idénticos como aquellos sitios contenidos en el sitio diana del organismo transformado. Se requiere una recombinasa que reconoce los sitios de recombinación no idénticos y cataliza la recombinación específica de sitio.

Se reconoce que la recombinasa puede proporcionarse mediante cualquier medio conocido en la técnica. Es decir, puede proporcionarse en el organismo o célula vegetal transformando el organismo con un casete de expresión que pueda expresar la recombinasa en el organismo, por expresión transitoria o proporcionando ARN mensajero (ARNm) para la recombinasa o la proteína recombinasa.

Por "sitios de recombinación no idénticos" se pretende que los sitios de recombinación flanqueantes no sean idénticos en secuencia y no recombinarán o la recombinación entre los sitios será mínima. Es decir, un sitio de recombinación flanqueante puede ser un sitio FRT donde el segundo sitio de recombinación puede ser un sitio FRT mutado. Los sitios de recombinación no idénticos usados en los métodos de la presente divulgación evitan o suprimen enormemente la recombinación entre los dos sitios de recombinación flanqueantes y la escisión de la secuencia de nucleótidos contenida en los mismos. Por consiguiente, se reconoce que puede utilizarse cualquier sitio de recombinación no idéntico adecuado en la presente divulgación, incluyendo sitios FRT y FRT mutantes, sitios FRT y lox, sitios lox y lox mutantes, así como otros sitios de recombinación conocidos en la técnica.

Por sitio de recombinación no idéntico adecuado supone que en presencia de recombinasa activa, la escisión de secuencias entre dos sitios de recombinación no idénticos se produce, si acaso, con una eficacia considerablemente menor que el intercambio mediado por recombinación que dirige la disposición de las secuencias de nucleótidos en el genoma de la planta. Por tanto, sitios no idénticos adecuados para su uso en la presente divulgación incluyen aquellos sitios donde la eficacia de recombinación entre los sitios es baja; por ejemplo, donde la eficacia es de menos de aproximadamente un 30 a aproximadamente un 50 %, preferiblemente menos de aproximadamente un 10 a aproximadamente un 30 %, más preferiblemente menos de aproximadamente un 5 a aproximadamente un 10 %.

Como se indica anteriormente, los sitios de recombinación en el casete de dirección corresponden a los del sitio diana de la planta transformada. Es decir, si el sitio diana de la planta transformada contiene sitios de recombinación no idénticos flanqueantes de FRT y FRT mutante, el casete de dirección contendrá los mismos sitios de recombinación no idénticos de FRT y FRT mutante.

Además, se reconoce que la recombinasa, que se usa en los métodos divulgados, dependerá de los sitios de recombinación en el sitio diana de la planta transformada y el casete de dirección. Es decir, si se utilizan sitios FRT, se necesitará la recombinasa FLP. De la misma manera, cuando se utilizan sitios lox, se requiere la recombinasa Cre. Si los sitios de recombinación no idénticos comprenden tanto un sitio FRT como uno lox, tanto la recombinasa FLP como la Cre serán necesarias en la célula vegetal.

La recombinasa FLP es una proteína que cataliza una reacción específica de sitio que está implicada en la amplificación del número de copias del plásmido 2 μ de *S. cerevisiae* durante la replicación del ADN. La proteína FLP se ha clonado y expresado. Véase, por ejemplo, Cox (1993) Proc. Natl. Acad. Sci. U. S. A. 80: 4223-4227. La recombinasa FLP para su uso en los métodos divulgados para integración dirigida puede derivar del género *Saccharomyces*. Puede ser preferible sintetizar la recombinasa usando codones preferidos de plantas para su expresión óptima en una planta de interés. Véase, por ejemplo, la solicitud de Estados Unidos n.º de serie 08/972.258 presentada el 18 de noviembre de 1997, titulada "Novel Nucleic Acid Sequence Encoding FLP Recombinase".

La recombinasa Cre del bacteriófago cataliza la recombinación específica de sitio entre dos sitios lox. La recombinasa Cre es conocida en la técnica. Véase, por ejemplo, Guo *et al.* (1997) Nature 389: 40-46; Abremski *et al.* (1984) J. Biol. Chem. 259: 1509-1514; Chen *et al.* (1996) Somat. Cell Mol. Genet. 22: 477-488; y Shaikh *et al.* (1977) J. Biol. Chem. 272: 5695-5702. Dicha una o más secuencias Cre también pueden sintetizarse usando codones preferidos de plantas.

Cuando sea apropiado, las secuencias de nucleótidos a insertar en el genoma de la planta pueden optimizarse para expresión aumentada en la planta transformada. Cuando se usan genes de mamífero, levadura o bacterianos en la presente divulgación, pueden sintetizarse usando codones preferidos de plantas para expresión mejorada. Se reconoce que para la expresión monocotiledóneas, también pueden sintetizarse genes de dicotiledóneas usando codones preferidos de monocotiledóneas. Hay métodos disponibles en la técnica para sintetizar genes preferidos de plantas. Véanse, por ejemplo, las patentes de Estados Unidos n.º 5.380.831, 5.436.391 y Murray *et al.* (1989) Nucleic Acids Res. 17: 477-498. Los codones preferidos de plantas pueden determinarse a partir de los codones utilizados más frecuentemente en las proteínas expresadas en la planta de interés. Se reconoce que pueden construirse secuencias preferidas de monocotiledóneas o dicotiledóneas, así como secuencias preferidas de plantas para especies de plantas particulares. Véanse, por ejemplo, los documentos EPA 0359472; EPA 0385962; WO 91/16432; Perlak *et al.* (1991) Proc. Natl. Acad. Sci. USA, 88: 3324-3328; y Murray *et al.* (1989) Nucleic Acids Research, 17: 477-498. Patente de Estados Unidos n.º 5.380.831; y patente de Estados Unidos n.º 5.436.391. Se reconoce además que la totalidad o cualquier parte

de la secuencia génica puede optimizarse o ser sintética. Es decir, también pueden usarse secuencias completamente optimizadas o parcialmente optimizadas.

Se conocen modificaciones de secuencia adicionales que potencian la expresión génica en un hospedador celular y pueden usarse en la presente divulgación. Estas incluyen la eliminación de secuencias que codifican señales de poliadenilación putativas, señales sitio de corte y empalme de exones e intrones, repeticiones de tipo transposón y otras secuencias bien caracterizadas de este tipo, que pueden ser perjudiciales para la expresión génica. El contenido de G-C de la secuencia puede ajustarse a niveles promedio para un hospedador celular dado, calculado por referencia a genes conocidos expresados en la célula hospedadora. Cuando es posible, la secuencia se modifica para evitar las estructuras de MARNA secundarias en forma de horquilla previstas.

La presente divulgación también abarca sitios diana de recombinación de FLP (FRT) novedosos. El FRT se ha identificado como una secuencia mínima que comprende dos repeticiones de 13 pares de bases, separadas por un espaciador de 8 bases, de la siguiente manera: 5'-GAAGTTCCTATTC [TCTAGAAA] GTATAGGAACCTC3' en el que los nucleótidos dentro de los corchetes indican la región espaciadora. Los nucleótidos en la región espaciadora pueden remplazarse con una combinación de nucleótidos, siempre que las dos repeticiones de 13 bases estén separadas por ocho nucleótidos. Algunas sustituciones de nucleótidos en la región espaciadora pueden funcionar mejor que otras y la determinación de esas sustituciones pertenece a las habilidades de la técnica. El espaciador de ocho pares de bases está implicado en el emparejamiento de ADN-ADN durante el intercambio de hebras. La asimetría de la región determina la dirección de alineación del sitio en el acontecimiento de recombinación, que posteriormente dará lugar a inversión o escisión. Como se indica anteriormente, la mayor parte del espaciador puede mutarse sin pérdida de función. Véase, por ejemplo, Schlake y Bode (1994) *Biochemistry* 33: 12746-12751.

Pueden usarse sitios mutantes FRT novedosos en la práctica de los métodos divulgados. Dichos sitios mutantes pueden construirse por mutagénesis basada en PCR. Aunque se conocen sitios FRT mutantes (por ejemplo, véase el documento WO/1999/025821, publicado el 27 de mayo de 1999), se reconoce que pueden usarse otros sitios FRT mutantes en la práctica de la presente divulgación. En aspectos, los métodos y composiciones de la presente divulgación pueden usarse sitios de recombinación no idénticos o sitios FRT para la inserción dirigida y expresión de secuencias de nucleótidos en un genoma vegetal.

Como se analiza anteriormente, poner un ADN genómico que contiene un sitio diana con sitios de recombinación no idénticos junto con un vector que contiene un casete de transferencia con correspondientes sitios de recombinación no idénticos, en presencia de la recombinasa, provoca recombinación. La secuencia de nucleótidos del casete de transferencia localizado entre los sitios de recombinación flanqueantes se intercambia con la secuencia de nucleótidos del sitio diana localizado entre los sitios de recombinación flanqueantes. De esta manera, las secuencias de nucleótidos de interés pueden incorporarse de forma precisa en el genoma del hospedador.

Se reconoce que pueden ponerse en práctica muchas variaciones de la presente divulgación. Por ejemplo, pueden construirse sitios diana que tienen múltiples sitios de recombinación no idénticos. Por tanto, pueden apilarse u ordenarse múltiples genes o secuencias de nucleótidos en localizaciones precisas en el genoma vegetal. Asimismo, una vez se ha establecido un sitio diana dentro del genoma, pueden introducirse sitios de recombinación adicionales incorporando dichos sitios dentro de la secuencia de nucleótidos del casete de transferencia y la transferencia de los sitios a la secuencia diana. Por tanto, una vez se ha establecido un sitio diana, es posible añadir posteriormente sitios, o alterar sitios a través de recombinación.

Otra variación incluye proporcionar un promotor o región de inicio de la transcripción unida de forma funcional con el sitio diana en un organismo. Preferiblemente, el promotor estará 5' al primer sitio de recombinación. Al transformar el organismo con un casete de transferencia que comprende una región codificante, se producirá la expresión de la región codificante tras la integración del casete de transferencia en el sitio diana. Este aspecto proporciona un método para seleccionar células transformadas, particularmente células vegetales, proporcionando una secuencia marcadora de selección como secuencia codificante.

Otras ventajas del presente sistema incluyen la capacidad de reducir la complejidad de la integración de transgenes o ADN transferido en un organismo utilizando casetes de transferencia como se analizan anteriormente y seleccionando organismos con patrones de integración simples. De la misma manera, pueden identificarse sitios preferidos dentro del genoma comparando varios acontecimientos de transformación. Un sitio preferido dentro del genoma incluye uno que no altera la expresión de las secuencias esenciales y proporciona expresión adecuada de la secuencia transgénica.

Los métodos divulgados también proporcionan un medio para combinar múltiples casetes en una localización dentro del genoma. Pueden añadirse o eliminarse sitios de recombinación en sitios diana dentro del genoma.

Puede usarse cualquier medio conocido en la técnica para poner los tres componentes del sistema juntos en la presente divulgación. Por ejemplo, una planta puede transformarse de forma estable para que albergue el sitio diana en su genoma. La recombinasa puede expresarse o proporcionarse de forma transitoria. Como alternativa, puede integrarse de forma estable una secuencia de nucleótidos que pueda expresar la recombinasa en el genoma de la planta. En presencia del

sitio diana correspondiente y la recombinasa, el casete de transferencia, flanqueado por sitios de recombinación no idénticos correspondientes, se inserta en el genoma de la planta transformada.

Como alternativa, los componentes del sistema pueden ponerse juntos cruzando sexualmente plantas transformadas. En este aspecto, una planta transformada, progenitora uno, que contiene un sitio diana integrado en su genoma, puede cruzarse sexualmente con una segunda planta, progenitora dos, que se ha transformado genéticamente con un casete de transferencia que contiene sitios de recombinación no idénticos flanqueantes, que corresponden a los de la planta uno. Ya sea la planta uno o la planta dos contiene dentro de su genoma una secuencia de nucleótidos que expresa recombinasa. La recombinasa puede estar bajo el control de un promotor constitutivo o inducible.

Los promotores inducibles incluyen los descritos en este documento anteriormente, así como promotores inducibles por calor, promotores sensibles a estradiol y promotores inducibles por productos químicos. Los promotores inducibles por patógenos incluyen los de proteínas relacionadas con patogénesis (proteínas PR), que se inducen después de infección por un patógeno; por ejemplo, proteínas PR, proteínas SAR, beta-1,3-glucanasa, quitinasa, etc. Véase, por ejemplo, Redolfi, *et al.* (1983) *Neth. J. Plant Pathol.* 89: 245-254; Uknes *et al.* (1992) *The Plant Cell* 4: 645-656; y Van Loon (1985) *Plant Mol. Virol.* 4: 111-116. De esta manera, puede controlarse la expresión de la recombinasa y la posterior actividad en los sitios de recombinación.

Los promotores constitutivos para su uso en la expresión de genes en plantas son conocidos en la técnica. Dichos promotores incluyen, aunque sin limitación, el promotor 35S del virus del mosaico de la coliflor (Depicker *et al.* (1982) *Mol. Appl. Genet.* 1: 561-573; Odell *et al.* (1985) *Nature* 313: 810-812), promotor de la ubiquitina (Christensen *et al.* (1992) *Plant Mol. Biol.* 18: 675-689), promotores de genes tales como la ribulosa bifsosfato carboxilasa (De Almeida *et al.* (1989) *Mol. Gen. Genet.* 218: 78-98), actina (McElroy *et al.* (1990) *Plant J.* 2: 163-171), histona y DnaJ (Baszczynski *et al.* (1997) *Maydica* 42: 189-201).

Las composiciones y métodos divulgados son útiles en dirigir la integración de secuencias de nucleótidos transferidas a un sitio cromosómico específico. La secuencia de nucleótidos puede codificar cualquier secuencia de nucleótidos de interés. Genes de interés particulares incluyen los que proporcionan un rasgo característico funcional fácilmente analizable a la célula y/u organismo hospedador, tal como genes marcadores, así como otros genes que alteran el fenotipo de las células destinatarias. Por tanto, los genes que logran el crecimiento de la planta, la altura, la susceptibilidad a enfermedades, insectos, y el valor nutritivo pueden utilizarse en la presente divulgación. La secuencia de nucleótidos también puede codificar una secuencia de "antisentido" para inactivar o modificar la expresión génica.

Se reconoce que las secuencias de nucleótidos pueden utilizarse en una unidad o casete de expresión funcional. Por unidad o casete de expresión funcional se pretende la secuencia de nucleótidos de interés con un promotor funcional y, en la mayoría de los casos, una región de terminación. Hay diversas maneras de conseguir la unidad de expresión funcional dentro de la práctica de la presente divulgación. En un aspecto de la presente divulgación, el ácido nucleico de interés se transfiere o inserta en el genoma como una unidad de expresión funcional.

Como alternativa, la secuencia de nucleótidos puede insertarse en un sitio dentro del genoma que está 3' a una región promotora. En este último caso, la inserción de la secuencia codificante 3' a la región promotora es tal que se consigue una unidad de expresión funcional tras la integración. Por conveniencia, para la expresión en plantas, el ácido nucleico que codifica sitios diana y los casetes de transferencia, incluyendo las secuencias de nucleótidos de interés, pueden estar contenidos dentro de casetes de expresión. Los casetes de expresión comprenderán una región de inicio transcripcional o promotor, unido de forma funcional al ácido nucleico que codifica el péptido de interés. Dicho casete de expresión está provisto de una pluralidad de sitios de restricción para la inserción del gen o genes de interés para que esté bajo la regulación transcripcional de las regiones reguladoras.

La región de inicio transcripcional, el promotor, puede ser natural u homólogo o exógeno o heterólogo al hospedador, o podría ser la secuencia natural o una secuencia sintética. Por exógeno se pretende que la región de inicio transcripcional no se encuentre en el hospedador de tipo silvestre en que se introduce la región de inicio transcripcional. Puede usarse un promotor natural o heterólogo con respecto a la secuencia codificante de interés.

El casete transcripcional puede incluir, en la dirección 5'-3' de transcripción, una región de inicio transcripcional y traduccional, una secuencia de ADN de interés y una región de terminación transcripcional y traduccional funcional en plantas. La región de terminación puede ser natural con la región de inicio transcripcional, puede ser natural con la secuencia de ADN de interés, o puede derivar de otra fuente. Están disponibles regiones de terminación convenientes del gen inhibidor de la proteinasa de la patata (PinII) o secuencias del plásmido Ti de *A. tumefaciens*, tal como las regiones de terminación de la nopalina sintasa, la octopina sintasa y la opalina sintasa. Véase también, Guerinneau *et al.*, (1991) *Mol. Gen. Genet.* 262: 141-144; Proudfoot (1991) *Cell* 64: 671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5: 141-149; Mogen *et al.* (1990) *Plant Cell* 2: 1261-1272; Munroe *et al.* (1990) *Gene* 91: 151-158; Ballas *et al.* 1989) *Nucleic Acids Res.* 17: 7891-7903; Joshi *et al.* (1987) *Nucleic Acid Res.* 15: 9627-9639.

Los casetes de expresión adicionalmente pueden contener secuencias líder 5' en la construcción del casete de expresión. Dichas secuencias líder pueden actuar potenciando la traducción. Los líderes de traducción son conocidos en la técnica e incluyen: líderes de picornavirus, por ejemplo, líder de EMCV (región no codificante 5' del virus de la encefalomiocarditis)

(Elroy-Stein, O., Fuerst, T. R., y Moss, B. (1989) PNAS USA, 86: 6126-6130); líderes de potyvirus, por ejemplo, líder de TEV (virus del grabado del tabaco) (Allison *et al.* (1986); líder de MDMV (virus del mosaico enano del maíz); Virology, 154: 9-20), y proteína de unión de la cadena pesada de la inmunoglobulina humana (BiP), (Macejak, D. G., y P. Sarnow (1991) Nature, 353: 90-94); líder no traducido de la proteína de envuelta MARNA del virus del mosaico de la alfalfa (ARN 4 de AMV), (Jobling, S. A., y Gehrke, L., (1987) Nature, 325: 622-625; líder del virus del mosaico del tabaco (TMV), (Gallie *et al.* (1989) Molecular Biology of RNA, páginas 237-256, Gallie *et al.* (1987) Nucl. Acids Res. 15: 3257-3273; líder del virus del moteado clorótico del maíz (MCMV) (Lornmel, S. A. *et al.* (1991) Virology, 81: 382-385). Véase también, Della-Cioppa *et al.* (1987) Plant Physiology, 84: 965-968; y secuencias no traducidas 5' del maíz endógenas. También pueden utilizarse otros métodos conocidos para potenciar la traducción, por ejemplo, intrones.

Los casetes de expresión pueden contener uno o más de un gen o secuencia de ácido nucleico a transferir y expresar en la planta transformada. Por tanto, cada secuencia de ácido nucleico se unirá de forma funcional a secuencias reguladoras 5' y 3'. Como alternativa, pueden proporcionarse múltiples casetes de expresión.

Parte experimental

Ejemplo 1: Plásmidos vir con rasgos característicos mejorados.

Se preparó una serie de vectores superbinarios mejorados. Los números de identificación de secuencia (SEQ ID NO) descritos en este documento se enumeran en la tabla 23. Los vectores, pPHP70298, pPHP71539 y pPHP79761, se muestran como mapas plasmídicos en las figuras 2, 3 y 4, respectivamente, y tienen las secuencias expuestas en las SEQ ID NO: 34, 35 y 36, respectivamente (véase la tabla 23). Para comparaciones, un mapa plasmídico del plásmido del gen vir pSB1 (Komari, T., *et al.*, Plant Cell Rep. (1990), 9:303-306) se muestra en la figura 1. Entre las mejoras a pPHP70298 (SEQ ID NO: 34), pPHP71539 (SEQ ID NO: 35) y pPHP79761 (SEQ ID NO: 36) está la presencia de un origen de replicación más pequeño, más estable, pVS1, en lugar de un origen de replicación grande, inestable (RK2). Además, el desplazamiento del marco de lectura en el operón *virC1-2* se repara, que codifica la proteína de unión de sobreexcitación para transferencia de ADN T potenciada. Es posible que el desplazamiento del marco de lectura disminuya la traducción de *virC1-2* como segundo gen en un ARNm dicistrónico. La funcionalidad génica de virulencia mejorada también se consiguió introduciendo *virD1* y *virD2* (pPHP70298); el operón *virD* completo (es decir, *virD1-D5*) junto con los genes *virE1-3* y *virG* (pPHP71539); o el operón *virD* completo (es decir, *virD1-D5*) junto con los genes *virE1-3*, *virA*, *virG* y *virJ* (pPHP71539). La disposición de los genes vir en los vectores pPHP70298 y pPHP71539 (véanse las figuras 2 y 3) es un fragmento contiguo como está presente en el plásmido pTiBo542 original (véase la figura 4). Sin embargo, en pPHP79761, la disposición de los genes *virA* y *virJ* se dispone de una manera contigua (véase la figura 4), mientras que estos genes no son contiguos en el plásmido pTiBo542 original (véase la figura 4). En particular, la disposición de los genes *virA* y *virJ* en pPHP79761 elimina un probable elemento de ADN transponible (IS292) flanqueado por secuencias repetidas y cuatro posibles secuencias codificantes (véase CDS 51-54 en la entrada de GenBank NC 010929). Estas construcciones son más pequeñas que pSB1, a pesar de tener un número aumentado de genes de virulencia.

En comparación con pSB1, se ha eliminado el ADN no esencial y los elementos repetitivos. Por ejemplo, los vectores, pPHP70298, pPHP71539 y pPHP79761, no contienen los operones *tra* y *trb* truncados de 7,0 kpb o genes flanqueantes incluidos en el origen de replicación de RK2 de 16,2 kb encontrado en pSB1 (véase la figura 1). Se encontraron dos secuencias adicionales en pSB1 (véase la figura 1) que se eliminaron en los vectores pPHP70298, pPHP71539 y pPHP79761: (a) una secuencia palindrómica de 32 pb (SEQ ID NO: 61); y (b) una secuencia de repetición invertida de 146 pb (SEQ ID NO: 62). Se cree que estos elementos en la estructura del vector disminuyen la estabilidad global del plásmido y, por consiguiente, su eliminación de los presentes vectores debe aumentar la estabilidad global del plásmido.

Los vectores, pPHP70298, pPHP71539 y pPHP79761, también carecen del fragmento de pBR322 de 2,7 kpb encontrado en pSB1, que comprende el origen de replicación, una secuencia codificante de beta lactamasa y sitios COS lambda flanqueados por poli-G de 18 pb inestables. Los vectores, pPHP70298, pPHP71539 y pPHP79761, en su lugar usan un origen de replicación de ColE1 de 1,2 kpb que proporciona mantenimiento estable en *Escherichia coli*.

El marcador de selección es un casete de gentamicina, que proporciona estabilidad potenciada del plásmido y crecimiento más rápido de *Agrobacterium*. El gen de resistencia a tetraciclina *tetAR* en pSB1 a menudo da lugar a un crecimiento lento o a colonias sin plásmido en la cepa de *Agrobacterium* C58 o colonias de *Agrobacterium* mutantes resistentes a tetraciclina (C58 y sus derivados tales como AGL0, AGL1, se ha descrito que dan lugar a mutantes resistentes a *tet* espontáneos a alta frecuencia; véase Luo, Z.K. y Farrand, S.K. (1999) J Bacteriol. 181:618-626). El casete de gentamicina usado fue el GmR sintético *aacC1* (basado en el número de acceso de GenBank DQ530421; SEQ ID NO: 1), que confiere resistencia a gentamicina, que no da lugar a crecimiento lento o virulencia reducida en la cepa de *Agrobacterium* recombinante que alberga el plásmido.

En resumen, los vectores se construyeron de una manera similar. La construcción de pPHP79761 se describe aquí como un ejemplo de los métodos usados en la construcción de pPHP70298 y pPHP71539. Los genes de virulencia se obtuvieron del plásmido Ti de *Agrobacterium tumefaciens*, pTiBo542 (secuencia de referencia del NCBI: NC_010929.1; véase la figura 5). Un experto en la materia apreciaría que los genes vir del plásmido Ri de *Agrobacterium* también serían útiles en los vectores divulgados en este documento.

Se diseñaron cebadores de PCR para amplificar las regiones codificantes y reguladoras de *virA*, *virJ* y *virB1-virE3* a partir de una preparación de ADN genómico de *Agrobacterium tumefaciens* cepa AGL1. La secuencia *virB1-E3* de 24 kb se amplificó en 5 trozos para facilitar la PCR. Se diseñaron fragmentos/derivados con extremos solapantes de 40 pb para facilitar métodos de clonación uniformes. Se incluyeron sitios de enzimas de restricción singulares entre elementos funcionales para facilitar su intercambio con otros elementos funcionalmente equivalentes, como se muestra en la figura 4. No se incluyó un elemento genético móvil marcado como IS292 en la secuencia de referencia del NCBI pTiBo542 que está presente entre los genes *virA* y *virJ*, en pPHP79761 para reducir el tamaño y la posibilidad de su transferencia a plantas. Se sintetizó un gen de gentamicina acetiltransferasa (basado en el número de acceso de GenBank DQ530421) con sitios de restricción flanqueantes para conferir resistencia a gentamicina a sus hospedadores bacterianos. Se incluyó un origen de replicación de ColE1 correspondiente a los pares de bases 2348-3247 de pBR322 (número de acceso de GenBank J01749) para la replicación estable en *E. coli*. Se incluyó un fragmento de 2,6 kb que codifica el origen de replicación de pVS1 (véase Stanisich, V. J. Bacteriol. (1984) 129:1227-1233) para el mantenimiento estable en *Agrobacterium* (véase la figura 6).

Ejemplo 2: Acontecimientos de maíz transgénico mediante *Agrobacterium* usando el pPHP70298 auxiliar.

El vector, pPHP70298, se ensayó como un vector auxiliar para la transformación del maíz en dos cultivariedades de maíz diferentes, PHR03 y PH184C, usando transformación de embriones inmaduros mediada por *Agrobacterium* como se describe en este documento a continuación. La cepa de *Agrobacterium tumefaciens* LBA4404THY que alberga el vector ternario que contiene un casete de expresión ZmUbi: PMI: PINII y ZmUbi: ZsYellow: PINII (pPHP45981; ADN T; SEQ ID NO: 28) se ensayó. Se realizaron experimentos paralelos con pSB1 más pPHP45981 y pPHP70298 más pPHP45981 para comparar el efecto de estos dos vectores, pSB1 y pPHP70298, sobre la transformación del maíz. La transformación se evaluó en términos de frecuencia de transformación del callo (% Tx callo), frecuencia de transformación de plantas T0 (% Tx T0), frecuencia de acontecimientos de calidad (QE) (definida como el porcentaje de acontecimientos con todos los genes de interés que son una sola copia y no tienen ningún ADN de estructura del vector - % QE) y calidad de acontecimientos útiles (UE) (definida como el número de acontecimientos QE recuperados por cada 100 embriones infectados- % UE) para los dos vectores. Los experimentos de transformación se realizaron con un mínimo de 600 embriones y los datos de transformación se resumen en las tablas 2 y 3 para los genotipos PHR03 y PH184C, respectivamente. Los datos en las tablas demuestran que el vector, pPHP70298, produjo una mejora significativa en la frecuencia global de transformación del callo y plantas T0 en ambas cultivariedades de maíz ensayadas.

Tabla 2. Datos de transformación (Tx) obtenidos en PHR03.

Plásmido auxiliar	Tx callo (%)	Tx T0 (%)	QE (%)	UE (%)
pSB1	54	14	40,7	5,7
pPHP70298	67	29	29,5	8,5

Tabla 3. Datos de transformación (Tx) obtenidos en PH184C.

Plásmido auxiliar	Tx callo (%)	Tx T0 (%)	QE (%)	UE
pSB1	63	45	41,5	18,6
pPHP70298	84	61	20,2	12,3

Cultivo de *Agrobacterium* en medio sólido: Se añadieron cinco ml de medio de infección de *Agrobacterium* (medio 700; véase la tabla 21) y 5 µl de 3'-5'-dimetoxi-4'-hidroxiaacetofenona (acetosiringona) 100 mM a un tubo Falcon de 14 ml en una campana. Se suspendieron aproximadamente 3 inóculos completos de *Agrobacterium* en el tubo y el tubo entonces se mezcló con agitadora vortical para generar una suspensión uniforme. Se transfirió un ml de la suspensión a un tubo de espectrofotómetro y la DO de la suspensión se ajustó hasta 0,35 a 550 nm. La concentración de *Agrobacterium* fue de aproximadamente $0,5 \times 10^9$ ufc/ml. La suspensión final de *Agrobacterium* se dividió en alícuotas en tubos de microcentrifugadora de 2 ml, que contenían cada uno 1 ml de la suspensión. Las suspensiones se usaron entonces lo antes posible.

Cultivo de *Agrobacterium* en medio líquido: Un día antes de la infección, se preparó un matraz de 125 ml con 30 ml de 557A (véase la tabla 21) con 30 µl de espectinomicina (50 mg/ml) y 30 µl de acetosiringona (20 mg/ml). Se suspendió medio inóculo de *Agrobacterium* en los matraces y se colocaron en una agitadora de 200 rpm a 28 °C durante una noche. El cultivo de *Agrobacterium* se centrifugó a 5000 rpm durante 10 min. El sobrenadante se retiró y se añadió el medio de infección de *Agrobacterium* (medio 700; véase la tabla 21) con solución de acetosiringona. Las bacterias se resuspendieron por mezcla en agitadora vortical y se ajustó la DO de la suspensión de *Agrobacterium* hasta 0,35 a 550 nm.

Transformación de maíz: Se esterilizaron superficialmente mazorcas de una cultivariedad de maíz (*Zea mays* L.), PHR03 o PH184C, durante 15-20 min en lejía al 20 % (v/v) (hipoclorito de sodio al 5,25 %) más 1 gota de Tween 20 seguido de 3 lavados en agua estéril. Se aislaron embriones inmaduros (IE) de las mazorcas y se coloraron en 2 ml del medio de infección de *Agrobacterium* (medio 700; véase la tabla 21) con solución de acetosiringona. El tamaño óptimo de los

embriones fue 1,5-1,8 mm para PHR03, respectivamente. La solución se trasvasó y se añadió 1 ml de suspensión de *Agrobacterium* a los embriones y el tubo se mezcló con agitadora vorticial durante 5-10 segundos. El tubo de microcentrifugadora se dejó reposar durante 5 min en la campana. La suspensión de *Agrobacterium* y los embriones se vertieron en medio de cocultivo (7101; véase la tabla 21). Cualquier embrión que quedara en el tubo se transfirió a la placa usando una espátula estéril. La suspensión de *Agrobacterium* se trasvasó y los embriones se colocaron con el eje hacia abajo en el medio. La placa se selló con película Parafilm™ (plástico flexible resistente a la humedad, disponible en Bemis Company, Inc., 1 Neenah Center 4.^a planta, PO Box 669, Neenah, WI 54957) y se incubó en la oscuridad a 21 °C durante 1-3 días de cocultivo.

Los embriones se transfirieron a medio de reposo (605W; véase la tabla 21) sin selección. De tres a 7 días después, se transfirieron a medio de selección 13152T o 13152Z (véase la tabla 21) complementado con manosa (12,5 g/l) u otro agente selectivo (G418, 150 mg/l). Tres semanas después de la primera ronda de selección, los cultivos se transfirieron a 13152T o 13152Z fresco (véase la tabla 21) que contenía un agente selectivo a intervalos de 3 a 4 semanas. Una vez transformados, los tejidos se transfirieron a medio de maduración (289Q o 289M; véase la tabla 21) complementado con agente selectivo apropiado.

Ejemplo 3: Comparación de pPHP70298 y pPHP71539 con el plásmido pSB1 para transformación de maíz.

Se determinó el impacto de pPHP70298 y pPHP71539 para la transformación de maíz en el genotipo de maíz PH184C usando transformación de embriones inmaduros mediada por *Agrobacterium*. En este estudio, se utilizaron vectores de producción ternarios que albergan al menos uno o más genes de interés (GOI) unidos de forma funcional a genes marcadores, fosfomanosa-isomerasa (PMI) y fosfinotricina acetil transferasa (moPAT), para selección de plantas con el ADN T. Los datos de transformación se presentan en la tabla 4, y estos datos muestran que la frecuencia de transformación fue al menos 1,5-2 veces mayor en los experimentos en el callo y a nivel de plantas T0 para PH184C usando pPHP70298 y pPHP71539 en comparación con pSB1. La frecuencia de acontecimientos de calidad (QE) (definida como el porcentaje de acontecimientos con todos los genes de interés que son de una sola copia y que no tienen ningún ADN de estructura del vector) no fue significativamente diferente entre los tres plásmidos pSB1, pPHP70298 y pPHP71539, pero la frecuencia de acontecimientos útiles (UE) (definida como el número de acontecimientos QE recuperados por cada 100 embriones infectados) fue mayor tanto para pPHP70298 como para pPHP71539 para tres vectores de producción independientes. La frecuencia de transformación promedio, % QE y UE para los tres vectores de producción se resume en la tabla 4.

Tabla 4. Resumen de datos de transformación en el genotipo PH184C.

Construcción	Auxiliar	Tx callo (%)	Tx T0 (%)	QE (%)	UE
A	pSB1	21,9	12,9	40,8	5,3
	pPHP70298	36,9	24,1	26,3	6,3
	pPHP71539	48,5	33,2	28,4	9,4
B	pSB1	23,7	14,5	39,1	5,7
	pPHP70298	32,3	18,9	32,4	6,1
	pPHP71539	51	29	34	9,9
C	pSB1	23,1	13,7	30,8	4,2
	pPHP70298	30,8	17	29,3	5
	pPHP71539	56,7	31,1	27,4	8,5
Promedio	pSB1	22,9	13,7	36,9	5
	pPHP70298	33,3	19,9	29,1	5,8
	pPHP71539	52,3	31,1	29,9	9,3

Ejemplo 4: Transformación mejorada con pPHP71539.

Se ensayó además el impacto de pPHP71539 sobre la transformación de maíz en otra cultivariedad de maíz, PHR03, mediante transformación de embriones maduros mediada por *Agrobacterium*. En este estudio, se utilizaron vectores de producción que albergan al menos uno o más genes de interés (GOI) unidos de forma funcional a genes marcadores, fosfomanosa-isomerasa (PMI) y fosfinotricina acetil transferasa (moPAT), para selección de plantas con el ADN T. Como se muestra en la tabla 5, pPHP71539 mejoró notablemente la frecuencia de transformación sin influir negativamente en la frecuencia de acontecimientos de calidad, lo que provoca ganancias importantes en la recuperación de acontecimientos útiles en el genotipo PHR03 en comparación con pSB1. Estos datos documentan además la mejora conseguida con los plásmidos divulgados para la transformación de maíz en múltiples genotipos.

Tabla 5. Datos de transformación (Tx) usando pPHP71539 en el genotipo PHR03 con construcción indicada.

Construcción	Auxiliar	Tx callo (%)	Tx T0 (%)	QE (%)	UE
D	pSB1	23	4	27	1,2
	pPHP71539	47	22	23	5,1
E	pSB1	8	3	20	0,7
	pPHP71539	43	16	38	6
F	pSB1	28	11	33	3,5
	pPHP71539	45	20	27	5,4

Un experto en la materia apreciaría que *Agrobacteria* con plásmido auxiliar pVIR10 (SEQ ID NO: 36; véase la tabla 23) puede mejorar significativamente la frecuencia de transformación, la recuperación de acontecimientos de calidad y acontecimientos de calidad útiles en múltiple endogámicos de maíz.

Ejemplo 5: Suministro transitorio de ADN mejorado y expresión génica en plantas.

El plásmido, pPHP71539, se ensayó para el suministro transitorio de ADN T en explantes foliares usando el método de agroinfección. Para la agroinfección, se infiltraron al vacío explantes foliares jóvenes con una cepa de *Agrobacterium* que albergaba un casete de producción (es decir, un GOI de protección de la planta) y las hojas se recogieron 2-3 días después de la infiltración para las mediciones de proteínas. Las mediciones de proteínas se hicieron con ELISA convencional usando muestras replicadas y se repitieron dos veces. Los datos sobre suministro transitorio de ADN en hojas de maíz se presentan en la tabla 6. La cantidad de proteínas medida en hojas infiltradas con pPHP71539 fue significativamente mayor que las cantidades de proteínas observadas en las hojas infiltradas con el mismo casete de expresión génica con pSB1. Esto demuestra que pPHP71539 facilitó un mayor suministro de ADN T y una expresión transitoria mejorada de proteínas en maíz.

Tabla 6. Expresión génica transitoria en hojas de maíz.

Construcción	Vector auxiliar	Expresión de proteínas (ppm)	Expresión promedio (ppm)
G	pSB1	418	429
		396	
		377	
		525	
	pPHP71539	537	660
		604	
		810	
		691	

Un experto en la materia apreciará que *Agrobacteria* con plásmido auxiliar pVIR7, pVIR9 o pVIR10 puede mejorar significativamente la expresión transitoria de proteínas en diferentes explantes de plantas en múltiples endogámicos de maíz.

Ejemplo 6: Transformación mejorada de sorgo usando *Agrobacterium* que porta pPHP70298.

Para demostrar la funcionalidad mejorada de los vectores divulgados para mejorar la transformación de plantas en otras monocotiledóneas, se ensayó la frecuencia de transformación para transformación de sorgo mediada por *Agrobacterium* usando embriones inmaduros de la variedad de sorgo TX430. En el primer experimento, la cepa de *Agrobacterium tumefaciens* LBA4404THY-, que alberga el vector ternario, pPHP45981, que contiene un casete de expresión ZmUbi: PMI: PINII y ZmUbi: ZsYellow: PINII (SEQ ID NO: 28) se usó. Se realizaron experimentos paralelos con pPHP45981 más pSB1 y pPHP45981 más pPHP70298 para determinar la frecuencia de transformación del callo (% Tx callo), la frecuencia de transformación de plantas T0 (% Tx T0), la frecuencia de acontecimientos de calidad (% QE) y la calidad de acontecimientos útiles (UE). Los datos de transformación de dos experimentos independientes con un mínimo de 300 embriones inmaduros se resumen en la tabla 7. La cepa LBA4404THY, que alberga plásmido pPHP45981 más pPHP70298, mostró rendimiento mejorado a todos los estados de transformación del sorgo, incluyendo transformación del callo, transformación de plantas T0, acontecimientos de calidad y acontecimientos globales de calidad útiles. Se realizó ensayo adicional con construcciones que contenían al menos un rasgo apilado y casete de expresión PMI más pPHP71539. La frecuencia de transformación promedio con pPHP71539 (tabla 8) fue mayor que la frecuencia de transformación normal observada con los vectores ternarios que contienen pSB1 (tabla 7).

Tabla 7. Datos de transformación obtenidos usando pPHP70298 en sorgo con la construcción pPHP45981.

Auxiliar	Tx callo (%)	Tx T0 (%)	QE (%)	UE
pSB1	4,2	2,1	64,7	1,4
pPHP70298	10,5	5,7	71,8	4,1

Tabla 8. La frecuencia de transformación T0 observada usando pPHP71539 en sorgo.

Construcciones	Tx callo (%)	Tx T0 (%)
H	11	5
I	37	19
J	7	4

- 5 Un experto en la materia apreciaría que *Agrobacterium* con plásmido auxiliar pVIR7 (SEQ ID NO: 34), pVIR9 (SEQ ID NO: 35) o pVIR10 (SEQ ID NO: 36) puede mejorar significativamente la frecuencia de transformación, la recuperación de acontecimientos de calidad y acontecimientos de calidad útiles en diferentes líneas de sorgo.

Ejemplo 7: Integración específica de sitio mejorada usando pPHP71539.

10

Se evaluó pPHP71539 para integración específica de sitio ("SSI") mediada por *Agrobacterium* en una línea diana que alberga el locus diana creado en la cultivariedad de maíz, PHR03, como se describe en la patente de Estados Unidos n.º 6.187.994 y la solicitud provisional de Estados Unidos n.º 62/296639. Se usó un sitio diana unido de forma funcional a un elemento de retención de promotor para ayudar en la identificación de acontecimientos diana, e identificación de acontecimientos SSI. Se generaron líneas que comprenden un sitio diana de retención de promotor por transformación con una construcción que comprende pPHP64484 ZmProUbi-FRT1-NptII::PinII + ZmUbiPro::AmCyan::PinII-FRT87 (SEQ ID NO: 29). Los vectores binarios se generaron con un marcador de selección de retención de promotor más un promotor constitutivo (ZmUbiPro; SEQ ID NO: 30) que dirige el gen indicador (DsRed (SEQ ID NO: 31) flanqueado por pares de sitios de recombinación FRT no idénticos (por ejemplo, FRT1/FRT87 (SEQ ID NO: 32 y 33) dentro del ADN T, denominado donador. Los vectores binarios se movilizaron en la cepa de *Agrobacterium* AGL1 con y sin pPHP71539 para evaluar el efecto de pPHP71539 sobre la recuperación de acontecimientos SSI y la frecuencia de SSI útiles (acontecimientos SSI precisos). La caracterización molecular de los acontecimientos putativos se realizó usando ensayos de qPCR/PCR para determinar la escisión del gen diana (NptII), la integración de los genes donadores (PMI y DsRED), la ausencia de gen FLP (integración aleatoria de ADN T) y la presencia de la unión del par FRT (1/87). Se realizó una PCR múltiple para el análisis de la estructura del vector. Se caracterizó un acontecimiento SSI preciso (acontecimiento SSI útil) como uno que cumple los siguientes criterios: (1) una sola copia intacta de los genes donadores (PMI y DsRed); (2) ausencia del gen diana (NptII), FLP, ODP/MoCre; (3) presencia de ambas uniones FRT1 y FRT87; y (4) sin ninguna inserción de estructura de vector. La tabla 9 resume la frecuencia de transformación y la frecuencia de SSI precisa obtenida usando la construcción pPHP71518 en la cepa de *Agrobacterium* AGL1 con o sin el plásmido auxiliar, pPHP71539, en la cultivariedad de maíz PHR03. Se observaron mejoras significativas en la recuperación de acontecimientos SSI putativos y la recuperación de acontecimientos SSI T0 precisos con la cepa de *Agrobacterium* que contenía pPHP71518 más pPHP71539 en comparación con la cepa de *Agrobacterium* con pPHP71518 en solitario.

35 Tabla 9. Datos que representan la frecuencia de transformación y la frecuencia de SSI en PHR03 con y sin plásmido auxiliar pVIR9 que contiene la construcción donadora pPHP71518

Construcción	Embrión total	Callo		T0		SSI	
		N.º de acontecimientos	Frecuencia	N.º de acontecimientos	Frecuencia	N.º de acontecimientos	Frecuencia
71518	817	151	18,8%	78	9,5%	12	1,5%
71518+ pPHP71539	821	190	23,3%	96	11,7%	22	2,7%

Un experto en la materia apreciará que *Agrobacterium* con plásmido auxiliar pVIR7 o pVIR10 puede mejorar significativamente la frecuencia de transformación T0 y la frecuencia de SSI precisa en múltiples endogámicos de maíz.

Ejemplo 8: Transformación de trigo mejorada usando plásmido pVIR (transformación de alta eficacia y rápida producción de plantas T0 en trigo).

A. Métodos de transformación de trigo

Se retira un alícuota de *Agrobacterium* cepa LBA4404 THY-, que contiene el vector de interés, del almacenamiento a -80 °C y se siembra en estrías en medio LB sólido (12V) que contiene un agente selectivo espectinomicina. La *Agrobacterium* se cultivó en el 12V a 21 °C en la oscuridad durante 2-3 días, tiempo en que se selecciona una sola colonia

de la placa, se siembra en estrías en placa de medio 810D que contiene el agente selectivo y después se incuba a 28 °C en la oscuridad durante una noche. Se escogen de dos a tres colonias de *Agrobacterium* usando una espátula estéril y se suspenden en ~5 ml de medio de infección de trigo (WI4) (WI4, véase la tabla 10) con acetosiringona (AS) 400 µM. La densidad óptica (600 nm) de la suspensión se ajustó hasta aproximadamente 0,1 a 0,7 usando el mismo medio.

Se recogieron de cuatro a cinco adiciones que contenían semillas inmaduras (con embriones de 1,4-2,3 mm), y los embriones inmaduros se aislaron de las semillas inmaduras. Los granos de trigo se esterilizaron superficialmente durante 15 min en lejía al 20 % (v/v) (hipoclorito de sodio al 5,25 %) más 1 gota de Tween 20, seguido de 2-3 lavados en agua estéril. Después de la esterilización, los embriones inmaduros (IE) se aislaron de los granos de trigo y se colocaron en 1,5 ml de medio WI4. Los embriones inmaduros se transfirieron a un tubo de microcentrifugadora de 2 ml que contenía 0,25 ml de arena estéril más 1 ml de medio WI4, y después se centrifugaron a 10 000 RPM durante 30 segundos. Después de la primera centrifugación, el tubo de microcentrifugadora se mezcló en agitadora vortical a un ajuste medio durante 10 segundos, y se centrifugó de nuevo una segunda vez a 10 000 RPM durante 30 segundos. Se permitió que los embriones se asentaran en esta suspensión durante 20 minutos.

En la siguiente etapa, se decantó el medio WI4 y se inició la infección de los embriones inmaduros añadiendo 1,0 ml de suspensión de *Agrobacterium*. Se permitió que los embriones infectados se asentaran durante 20 minutos. La suspensión de *Agrobacterium* e IE se vertió en medio de cocultivo de trigo n.º 10 (véase la tabla 11). Los embriones se vertieron a la placa usando una espátula estéril, con el lado del eje hacia abajo en el medio, garantizando que los embriones estuvieran sumergidos en la solución. La placa se selló con película en cinta Parafilm® (plástico flexible resistente a la humedad, disponible en Bemis Company, Inc., 1 Neenah Center 4.^a planta, PO Box 669, Neenah, WI 54957) y se incubó en la oscuridad a 21 °C durante 3 días de cocultivo. Las tablas 10-15 describen el medio de infección de trigo líquido (WI4), el medio de cocultivo de trigo (WC n.º 10), el medio DBC4 de primera ronda, el medio DBC6 de segunda ronda, el medio MSA de regeneración y el medio MSB de regeneración.

Tabla 10. Composición del medio de infección líquido de trigo WI4.

WI4	
Agua DI	1000mL
Sal MS + vitaminas (M519)	4,43 g
Maltosa	30 g
Glucosa	10 g
MES	1,95 g
2,4-D (0,5 mg/l)	1 ml
Picloram (10 mg/ml)	200 µl
BAP (1 mg/l)	0,5 ml
Ajustar pH hasta 5,8 con KOH	
Añadir tras esterilización:	
Acetosiringona (400 µM)	400 µl

Tabla 11. Composición de medio de cocultivo de trigo WC n.º 10.

WC n.º 10	
Agua DI	1000mL
Sal MS + vitaminas (M519)	4,43 g
Maltosa	30 g
Glucosa	1 g
MES	1,95 g
2,4-D (0,5 mg/l)	1 ml
Picloram (10 mg/ml)	200 µl
BAP (1 mg/l)	0,5 ml
CuSO ₄ 50X (0,1 M)	49 µl
Ajustar pH hasta 5,8 con KOH y añadir 2,5 g/l	
de Phytigel.	

WC n.º 10	
Añadir tras esterilización:	
Acetosiringona (400 µM)	400 µl

Tabla 12. Composición de medio de cultivo de tejido verde de trigo DBC4.

DBC4	
dd H ₂ O	1000mL
Sal MS	4,3 g
Maltosa	30 g
Mioinositol	0,25 g
N-Z-Amina-A	1 g
Prolina	0,69 g
Tiamina-HCl (0,1 mg/ml)	10 ml
CuSO ₄ 50X (0,1 M)	49 µl
2,4-D (0,5 mg/ml)	2 ml
BAP	1 ml
Ajustar pH hasta 5,8 con KOH y después añadir 3,5 g/l de Phytigel.	
Añadir tras esterilización:	
Cef (100 mg/ml)	1 ml

Tabla 13. Composición del medio de inducción de tejido verde de trigo DBC6.

DBC6	
dd H ₂ O	1000mL
Sal MS	4,3 g
Maltosa	30 g
Mioinositol	0,25 g
N-Z-Amina-A	1 g
Prolina	0,69 g
Tiamina-HCl (0,1 mg/ml)	10 ml
CuSO ₄ 50X (0,1 M)	49 µl
2,4-D (0,5 mg/ml)	1 ml
BAP	2 ml
Ajustar pH hasta 5,8 con KOH y después añadir 3,5 g/l de Phytigel.	
Añadir tras esterilización:	
Cef (100 mg/ml)	1 ml

Tabla 14. Composición de medio de regeneración de trigo MSA.

MSA	
dd H ₂ O	1000mL
Sal MS + vitaminas (M519)	4,43 g
Sacarosa	20 g
Mioinositol	1 g
Ajustar pH hasta 5,8 con KOH y después añadir 3,5 g/l de Phytigel.	

MSA	
Añadir tras esterilización:	
Cef (100 mg/ml)	1 ml

Tabla 15. Composición de medio de regeneración de trigo MSB.

MSB	
dd H ₂ O	1000mL
Sal MS + vitaminas (M519)	4,43 g
Sacarosa	20 g
Mioinositol	1 g
Ajustar pH hasta 5,8 con KOH y después añadir 3,5 g/l de Phytigel.	
Añadir tras esterilización:	
Cef (100 mg/ml)	1 ml
IBA	0,5 ml

B. Selección de embriones de trigo

Los embriones inmaduros se transfieren a medio de inducción de tejido verde DBC4 (véase la tabla 12) (medio GT con 100 mg/l de cefotaxima (PhytoTechnology Lab., Shawnee Mission, KS)) sin selección, en la orientación del eje del embrión en contacto con el medio. Los embriones se incuban en este medio a 26-28 °C en luz tenue durante dos semanas, después se transfieren a medio DBC6 (véase la tabla 13) que contiene 100 mg/l de cefotaxima durante otras dos semanas. En este momento, todo el tejido que expresa la proteína fluorescente se separa de los tejidos no transformados en un microscopio de fluorescencia y se coloca en medio de inducción de GT DBC6 (véase la tabla 13) que contiene 100 mg/l de cefotaxima para la proliferación del tejido.

C. Regeneración de plántulas y transferencia al invernadero

El tejido que emite fluorescencia se transfiere a medio de regeneración MSA (véase la tabla 14) y se incuba a 26-28 °C en luz intensa durante 2-4 semanas. En este punto, el tejido se comprueba para la expresión uniforme de los genes marcadores fluorescentes en las plántulas transgénicas y para raíces sanas. Las plántulas se transfieren a tierra en macetas en el invernadero.

D. La introducción de los plásmidos superbinarios pPHP71539 y pPHP70298 en *Agrobacteria* que contienen casetes de expresión en plantas provocó suministro transitorio mejorado de ADN T y recuperación mejorada de acontecimientos T0 transgénicos en trigo

Se recogieron embriones inmaduros de la cultivariedad de trigo HC0456D y se infectaron con las cepas de *Agrobacterium* LBA4404 THY-pPHP71539 y AGL1 que contenía un plásmido binario de ADN T con la siguiente composición - RB-UBI-ZMPRO: MO-PAT: CONECTOR PROTEÍNICO:DS-RED:PINII-LB (SEQ ID NO: 68). La cepa de *Agrobacterium* LBA4404 THY- es una cepa más débil que previamente se había demostrado que provoca un suministro transitorio débil de ADN T y se ha descrito en otra parte como una mala cepa para la transformación de trigo, por lo tanto, se excluyó del estudio. Para la expresión transitoria de ADN T, se infectaron 20-30 embriones de la cultivariedad y se controló la expresión DS-RED a los 7 días y los 15 días después de la infección. En el experimento, se ensayaron dos cepas diferentes de *Agrobacterium* AGL1 y LBA4404 Thy-pPHP71539 para caracterizar el efecto de pVIR9 sobre la transformación estable (15 DPI) en trigo (figura 7). La cepa AGL1 natural mostró varias veces menos expresión transitoria de DsRed en 7DPI y siguió mostrando menos sectores de expresión de DsRed en 15DPI (figura 7), lo que representa una expresión estable de ADN T en los embriones infectados en comparación con la cepa LBA4404THY-pPHP71539.

En un experimento diferente, *Agrobacteria* que contenían el plásmido auxiliar pVIR mejoraron la transformación estable y aumentaron la recuperación de acontecimientos T0 en múltiples cultivariedades de trigo. Para este fin, se realizaron experimentos de transformación de trigo paralelos en dos cultivariedades de trigo diferentes (Fielder y HC0456D) con la cepa de *Agrobacterium* AGL1, con y sin pVIR auxiliar pPHP70298. Se recogieron embriones inmaduros de dos cultivariedades de trigo (Fielder y HC0456D) y se infectaron con *Agrobacteria* AGL1 y AGL1 más pPHP70298 que contenía dos plásmidos binarios diferentes (A y B) que contenían el siguiente casete de expresión; RB-UBI-ZMPRO: MO-PAT: CONECTOR PROTEÍNICO: DS-RED:PINII-LB (SEQ ID NO: 68) para capturar la frecuencia de transformación de plantas T0. Los datos de transformación se resumen en la tabla 16.

Tabla 16. Datos de transformación usando pPHP70298 en múltiples cultivariedades de trigo.

Cultivariedad	Construcción	AGL1 (% TX T0)	AGL1 + pPHP70298 (% TX T0)
Fielder	A	0,5% (26/5748)	1,4% (4/278)
HC0456D	A	0% (0/1918)	1,1% (4/375)
Fielder	B	0,4% (9/2294)	5,4% (47/863)
HC0456D	B	0%	5,3% (36/678)

Los embriones inmaduros de trigo transformados con AGL1 más pPHP70298 mostraron un rendimiento notablemente mejorado en todos los estados de transformación del trigo, incluyendo transformación del callo y transformación de plantas T0. La frecuencia de transformación de plantas T0 con *Agrobacteria* que contienen el plásmido pVIR7 se determinó mucho más alta que la cepa de tipo silvestre AGL1 sin el plásmido auxiliar pVIR7. Se ha usado *Ochrobactrum* que contiene los plásmidos pVIR7, pVIR9 o pVIR10 para transformar satisfactoriamente plantas (datos no mostrados) como se describe en la solicitud provisional de Estados Unidos n.º 62/211267.

Un experto en la materia apreciaría que *Agrobacteria* con plásmido auxiliar pVIR9 (SEQ ID NO: 35, véase la tabla 23) o pVIR10 puede mejorar significativamente la frecuencia de transformación, la recuperación de acontecimientos de calidad y acontecimientos de calidad útiles en múltiples líneas de trigo.

Ejemplo 9: El plásmido pVIR9 mejoró la transformación de maíz

El plásmido pVIR9 se ensayó para transformación de maíz usando el sistema de vector ternario para transformar el genotipo de maíz PH184C como se describe en la solicitud provisional de Estados Unidos n.º 62/248578. La introducción del plásmido superbinario pPHP71539 en *Agrobacteria* que contienen casetes de expresión en plantas provocó suministro transitorio mejorado de ADN T, fenotipo de embriones somáticos y recuperación mejorada de acontecimientos T0 transgénicos en maíz.

En resumen, se recogieron embriones inmaduros (2-2,5 mm de longitud) de maíz endogámico Pioneer PH184C) aproximadamente 11 días después de la polinización, y se infectaron con la cepa de *Agrobacterium* LBA4404 THY- que contiene los plásmidos de ADN T 1) pPHP80561 o 2) pPHP80559 con la siguiente composición; pPHP80561- RB + LOX P- ZM-AXIG1 1XOP-WUS2::IN2-1 TERM + ZM-PLTP PRO::ZM-ODP2::OS-T28 TERM + PINII+ GZ-W64A TERM + OLE PRO:MO-CRE EXÓN1:ST-LS1 INTRÓN2: MO-CRE EXÓN2-PINII TERM+ SB-UBI PRO: SB-UBI INTRÓN1:UBI PRO: ZS-GREEN:OS-UBI TERM+ LOXP+ SB-ALS PRO:: HRA::PINII TERM+ LB (SEQ ID NO: 69) y, pPHP80559- RB + LOX P- ZM-AXIG1 1XOP-WUS2::IN2-1 TERM + ZM-PLTP PRO::ZM-ODP2::OS-T28 TERM + PINII+ GZ-W64A TERM + LTP2 PRO:MO-CRE EXÓN1:ST-LS1 INTRÓN2: MO-CRE EXÓN2-PINII TERM+ SB-UBI PRO: SB-UBI INTRÓN1:UBI PRO: ZS-GREEN:OS-UBI TERM+ LOXP+ SB-ALS PRO:: HRA::PINII TERM+ LB (SEQ ID NO: 70). Estos plásmidos se movilizaban en dos cepas de *Agrobacterium* diferentes LBA4404THY-pSB1 y LBA4404THY-pVIR9 (pPHP71539). Las *Agrobacteria* con el plásmido se cultivaron en medio líquido hasta una densidad óptica de 0,5 (a 520 nm) y los embriones inmaduros (~600 embriones más, mazorca dividida, tres réplicas) se incubaron en la suspensión de *Agrobacterium* durante 5 minutos antes de retirarlos del líquido para colocarlo en medio 710I sólido. Después de 24 horas, los embriones se movieron a medio 605T (véase la tabla 21) para empezar la selección frente a *Agrobacterium*. Después de 6 días, se observaron numerosos embriones somáticos pequeños en la superficie de los embriones inmaduros tratados. Siete días después de la agroinfección, los embriones se transfirieron a medio de maduración (medio 289Q con 0,1 mg/l de imazapyr), usando el herbicida imidazolinona para seleccionar embriones transgénicos. Después de 14 días en el medio de maduración, los embriones maduros se movieron a medio de enraizamiento (medio 13158H; medio 13158 más 25 mg/l de cefotaxima) y se muestrearon trozos foliares para análisis de PCR. Los datos sobre la frecuencia TX, la frecuencia QE escindida y la frecuencia UE se representan en la tabla 17.

Tabla 17. Datos de transformación en PH184C endogámico en comparación con plásmido pVIR9 y pSB1 con dos vectores de expresión diferentes.

Tratamiento	pPHP	Auxiliar	Pro:CRE	N.º emb	N.º plantas T0	% TX	Acontecimientos escindidos	QE	% escisión QE	% UQE
A	pPHP80561	pVIR9	Ole	610	120	20%	21		18%	3,4%
	pPHP80561	pSB1	Ole	607	19	3%	4		21%	0,7%
B	pPHP80559	pVIR9	LTP2	610	113	19%	23		20%	3,8%
	pPHP80559	pSB1	LTP2	610	42	7%	6		14%	1,0%

Los embriones inmaduros transformados con *Agrobacteria* que contenían el plásmido pPHP71539 (pVIR9) produjeron números muy altos de plantas T0 y provocaron frecuencia de transformación de T0 mayor en comparación con *Agrobacteria* transformadas con el plásmido pSB1. Las *Agrobacteria* con el plásmido auxiliar pVIR9 también provocaron

frecuencia significativamente mayor de acontecimientos de calidad útiles en comparación con embriones inmaduros infectados con plásmido pSB1.

- 5 Un experto en la materia apreciaría que las *Agrobacteria* con plásmido auxiliar pVIR7 o pVIR10 en lugar de pVIR9 pueden mejorar significativamente la transformación de maíz, los acontecimientos de calidad y los acontecimientos de calidad útiles.

Ejemplo 10. Plásmido pVIR con diferentes orígenes de replicación para transformación de plantas.

- 10 Para caracterizar el efecto de diferentes orígenes de replicación en el plásmido auxiliar para transformación de maíz, se usó el sistema de vector ternario para transformar el genotipo de maíz HC69. Usando la estrategia de transformación de maíz como se describe anteriormente en el ejemplo 9 (solicitud provisional de Estados Unidos n.º 62/248578), se ensayaron cuatro plásmidos pVIR que contenían diferentes orígenes bacterianos de replicación, concretamente, pVS1, pSaparDE y RK2parDE (tabla 18). Se recogieron embriones inmaduros de maíz endogámico (HC69) y se infectaron con
- 15 *Agrobacterium* (cepa LBA4404 THY-) que contenía pPHP79066 (SEQ ID NO: 71) con la siguiente composición de ADN T y los plásmidos auxiliares con diferentes orígenes de replicación bacteriana ("ori"), concretamente, los siguientes: pVS1 (pVIR10; pPHP79761), pSaparDE (pVIR10; pPHP80399), RK2parDE (pVIR10; pPHP80566) y control (pVIR9; pPHP71539) como se detalla en la tabla 18. Para cada construcción, se transformaron 150 embriones en el endogámico HC69 usando el método de mazorca dividida para determinar la frecuencia de transformación. Todas las construcciones
- 20 transformaron maíz a frecuencia muy alta, lo que demuestra que las diferentes combinaciones de ori funcionaron en el plásmido pVIR. Los datos de transformación para el endogámico HC69 se presentan en la tabla 19.

Tabla 18. Ensayo de diferentes ori en plásmido pVIR.

N.º vector	Binario	pVIR	Ori
1	pPHP79066	pPHP79761	pVS1
2	pPHP79066	pPHP80399	pSA parDE
3	pPHP79066	pPHP80566	RK2micro parDE
4	pPHP79066	pPHP71539	pVS1

- 25 Tabla 19. Datos de transformación en el endogámico HC69 con plásmido pVIR que contiene diferente ori.

Combinaciones binario/pVIR	Embriones totales	Plantas totales	T0	Frecuencia de transformación
pPHP79066/ pPHP79761	183	154		84%
pPHP79066/ pPHP80399	160	116		73%
pPHP79066/ pPHP80566	176	173		98%
pPHP79066/pPHP71539	170	175		103%

Los plásmidos pVIR con diferente ori transformaron maíz y regeneraron acontecimientos T0. La frecuencia de transformación de T0 fue comparable entre todos los diferentes ori ensayados.

- 30 **Ejemplo 11. Sistema de suministro potenciado para reactivos CRISPR/CAS9 y mutaciones génicas con plásmido pVIR en maíz.**

- Para caracterizar el efecto del plásmido auxiliar en el suministro de CRISPR/Cas9 y la tasa de mutantes recuperados de los diferentes plásmidos auxiliares, se usó un sistema de vector ternario para transformar el genotipo de maíz PH2HT.
- 35 Usando el protocolo de transformación aleatoria, se transformaron dos construcciones diferentes, pPHP78147 (SEQ ID NO: 72) y pPHP78148 (SEQ ID NO: 73), que comprenden cada una dos secuencias de ARN guía (ARNg), específicas para generar eliminaciones de pérdida en genes dianas ZM-ARGOS8 y ZM-GCN2. Las secuencias de las secuencias de ARN guía y los casetes de expresión se describen en la SEQ ID NO: 72 y la SEQ ID NO: 73, respectivamente. Se recogieron embriones inmaduros de maíz endogámico PH184C y se infectaron con *Agrobacterium* (cepa LBA4404 THY-) que contenía plásmidos auxiliares pPSB1 o pVIR9 que contenían los casetes de expresión de ADN T pPHP78147 o
- 40 pPHP78148. Para cada construcción, se transformaron aproximadamente 280 embriones del endogámico PH2RT usando el método de mazorca dividida para determinar la frecuencia de transformación, las tasas de mutación, las eliminaciones de pérdida individuales y las tasas de eliminación de pérdida bialélica. Para determinar las tasas de mutación, las plantas T0 se sometieron a análisis de secuenciación profunda y se determinó la calidad del acontecimiento molecular ejecutando
- 45 ensayos de PCR y qPCR en genes diana como se describe previamente. Los datos de transformación y los tipos de edición génica provocados por cada construcción en PH2RT se presentan en la tabla 20.

Tabla 20. Tipos de edición génica provocados por cada construcción.

pPHP	Auxiliar	N.º embrión* (número)	N.º acontecimientos T0	% Txn T0	Tasa de mutación	N.º pérdida de 1 alelo	Tasa de pérdida de alelo individual	N.º pérdida bialélica	Tasa de pérdida bialélica
pPHP78147 (ARGOS8)	pSB1	285	130	61,8 %	42,3%	3	2,3%	7	5,4%
pPHP78147 (ARGOS8)	pVIR 9	285	129	80,3 %	46,5%	10	7,7%	6	4,7%
pPHP78148 (GCN2)	pSB1	280	112	47,1 %	86,6%	18	16,1 %	5	4,5%
pPHP78148 (GCN2)	pVIR 9	280	127	59,6 %	92,1%	22	17,3 %	8	6,3%

La frecuencia de transformación y las tasas de mutación recuperadas de nuevos vectores superbinarios pVIR fueron mucho mayores que las observadas con pSB1. La frecuencia de transformación varió dependiendo de las secuencias de ARNg y los tipos de eliminación. Para ARGOS8, en la que se abordó una eliminación de pérdida más grande (619 pb), la frecuencia de transformación observada con pVIR9 fue de aproximadamente un 80,3 % en comparación con un 61,8 % con pSB1. La frecuencia a la que se recuperaron pérdidas de alelo individual con la construcción ARGOS8 fue mayor con pVIR9 (7,7 %) en comparación con pSB1 (2,3 %). Con la segunda construcción dirigida a una eliminación de pérdida más pequeña en ZM-GCN2 (52 pb), los embriones transformados con *Agrobacterium* que contenía el plásmido auxiliar pVIR9 provocaron mayor generación de acontecimientos T0 y recuperación de acontecimientos mutantes (mutaciones, alelo individual y bialélicos) en comparación con acontecimientos recuperados de *Agrobacterium* con el pSB1 auxiliar (tabla 20).

Un experto en la materia apreciaría que *Agrobacteria* con el plásmido auxiliar pVIR7 o pVIR10 puede potenciar el suministro de CRISPR/nucleasas Cas9 o nucleasas relacionadas para mejorar la edición genómica y las modificaciones genómicas en múltiples endogámicos de maíz.

Ejemplo 12: Recuperación mejorada de acontecimientos de integraciones específicas de sitio (SSI) en la cepa de *Agrobacterium* LBA4404THY- que alberga plásmido superbinario.

Para determinar el efecto de los vectores superbinarios (pVIR9; pPHP71539), sobre la recuperación de acontecimientos SSI precisos, se movilizó el plásmido auxiliar pVIR9 en la cepa de *Agrobacterium* LBA4404THY-. La cepa LBA4404THY- ha demostrado provocar baja recuperación de acontecimientos SSI precisos (solicitud provisional de Estados Unidos n.º 62/296639). Para ensayar si pVIR9 en la cepa de *Agrobacterium* LBA4404THY- mejoraba la recuperación de acontecimientos SSI, un vector binario que contenía el casete donador con el siguiente casete de expresión RB-OS-ACTINA PRO: OS-ACTINA INTRÓN:: ZM-WUS2::PINII + UBI PRO::UBI1ZM INTRÓN::ZM-ODP2::PINII TERM:: UBI PRO::UBI1ZM INTRÓN::MO-FLP::PINII TERM:: CaMV35S TERM + FRT1:PMI::PINII TERM+ GEN DE RASGO+ FRT87-LB se movilizó en dos cepas de *Agrobacterium* diferentes AGL1 y LBA4404THY-pPHP71539 (pPHP79366; SEQ ID NO: 74). Los casetes donadoras se suministraron mediante transformación mediada por *Agrobacterium* en múltiples sitios diana con sitios de descarga FRT1-87 en el genotipo PH184C como se describe a continuación.

En resumen, la transformación mediada por *Agrobacterium* en líneas diana con sitios de descarga FRT1-FRT87 se realizó de la siguiente manera. Se esterilizaron superficialmente mazorcas de una cultivariedad de maíz (*Zea mays* L.), PH184C, durante 15-20 min en lejía al 20 % (v/v) (hipoclorito de sodio al 5,25 %) más 1 gota de Tween 20 seguido de 3 lavados en agua estéril. Los embriones inmaduros (IE), típicamente de 1,5-1,8 mm, se aislaron de las mazorcas y se colocaron en 2 ml del medio de infección de *Agrobacterium* con solución de acetosiringona. La solución se trasvasó y se añadió 1 ml de suspensión de *Agrobacterium* a los embriones, se mezcló con agitadora vortical durante 5-10 segundos, y después se incubó 5 min a temperatura ambiente. La suspensión de *Agrobacterium* y embriones se vertió en medio de cocultivo 7101 (véase la tabla 21). Cualquier embrión que quedara en el tubo se transfirió a la placa usando una espátula estéril. La suspensión de *Agrobacterium* se trasvasó y los embriones se colocaron con el eje hacia abajo en el medio. La placa se selló con cinta Parafilm™ (plástico flexible resistente a la humedad, disponible en Bemis Company, Inc., 1 Neenah Center 4.ª planta, PO Box 669, Neenah, WI 54957) y se incubó en la oscuridad a 21 °C durante 1-3 días de cocultivo.

Los embriones se transfirieron a medio de reposo 13265A (véase la tabla 21) sin selección. De tres a 7 días después, se transfirieron a medio de selección 13152Z (véase la tabla 21) complementado con manosa u otro agente selectivo apropiado. Tres semanas después de la primera ronda de selección, los cultivos se transfirieron a una segunda ronda de medio de selección 13152Z (véase la tabla 21) que contenía un agente selectivo a intervalos de 3 a 4 semanas. Una vez transformados, los tejidos verdes transgénicos se seleccionan y cultivan esencialmente como se describe en la patente de Estados Unidos 7102056, patente de Estados Unidos 8404930 y la publicación US20130055472.

Tabla 21. Formaciones de medios para transformación de maíz, selección y regeneración.

Componentes del medio	Unidades por litro	12V	810I	700	710 I	605J	605T	289Q	605W	289M
Mezcla de sal basal MS	G			4,3	4,3	4,3	4,3	4,3	4,3	4,3
Macronutrientes N6 10X	ml					60,0	60,0		60,0	
Nitrato de potasio	G					1,7	1,7		1,7	
Sales minoritarias B5H 1000X	ml					0,6	0,6		0,6	
NaFe EDTA para B5H 100X	ml					6,0	6,0		6,0	
Vitaminas de Eriksson 1000X	ml					0,4	0,4		0,4	
Reserva de vitaminas S&H 100X	ml					6,0	6,0		6,0	
Tiamina HCl	mg			1,0	1,0	0,2	0,2		0,2	
L-prolina	g				0,7	2,0	2,0	0,7	2,0	0,7
Hidrolizado de caseína (ácido)	g					0,3	0,3		0,3	
Sacarosa	g			68. 5	20,0	20,0	20,0	60,0	20,0	60,0
Glucosa	g	5,0		36. 0	10,0	0,6	0,6		10,0	
Maltosa	g									
2,4-D	mg			1,5	2,0	0,8	0,8		0,8	
Agar	g	15. 0	15,0		8,0	6,0	6,0	8,0		
Phytigel	g								3,5	
Dicamba	g					1,2	1,2		1,2	
Nitrato de plata	mg					3,4	3,4			
Carbenicilina Agribio	mg					100. 0			100. 0	
Timentina	mg						150. 0	150. 0		
Cefotaxima	mg						100. 0	100. 0		
Mioinositol	g			0,1	0,1			0,1		0,1
Ácido nicotínico	mg			0,5	0,5					
Piridoxina.HCl	mg			0,5	0,5					
Casaminoácidos de ensayo de vitaminas	g			1,0						
Tampón MES	g				0,5					
Acetosiringona	uM				100. 0					
Ácido ascórbico 10 mg/ml (7S)	mg				10,0					
Solución de reserva de vitaminas MS	ml							5,0		5,0
Zeatina	mg							0,5		0,5
Sulfato cúprico	mg							1,3	0,05	1,3
IAA 0,5 mg/ml (28A)	ml							2,0		2,0
ABA 0,1 mm	ml							1,0		1,0
Tiadiazurona	mg							0,1		0,1
Carbenicilina Agribio	mg							100. 0		100. 0

Componentes del medio	Unidades por litro	12V	810I	700	710 I	605J	605T	289Q	605W	289M
PPT(glufofosinato-NH ₄)	mg									
BAP	mg							1,0	0,1	
Extracto de levadura (BD Difco)	g		5,0							
Peptona	g		10,0							
Cloruro de sodio	g		5,0							
Espectinomicina	mg	50. 0	100. 0							
Sulfato ferroso.7H ₂ O	ml	2,0								
Tampón AB 20X (12D)	ml	50. 0								
Sales AB 20X (12E)	ml	50. 0								
Benomilo	mg									
pH		5,6	5,6	5,6	5,6	5,6	5,6	5,6	5,6	5,6

Tabla 21, continuación.

Componentes del medio	Unidades por litro	289R	13158 H	13224 B	13266 K	272X	272V	13158	13265 A	13152 T	13152 Z
Mezcla de sal basal MS	G	4,3	4,3		4,3	4,3	4,3	4,3	4,3	4,3	4,3
Macronutrientes N6 10X	ml			4,0	60,0				60,0		
Nitrato de potasio	G				1,7				1,7		
Sales minoritarias B5H 1000X	ml				0,6				0,6		
NaFe EDTA para B5H 100X	ml				6,0				6,0		
Vitaminas de Eriksson 1000X	ml			1,0	0,4				0,4		
Reserva de vitaminas S&H 100X	ml				6,0				6,0		
Tiamina HCl	mg			0,5	0,5				0,5	1,0	1,0
L-prolina	G	0,7	0,7	2,9	2,0				2,0	0,7	0,7
Hidrolizado de caseína (ácido)	G				0,3				0,3	1,0	1,0
Sacarosa	G	60. 0	60,0	190. 0	20,0	40. 0	40. 0	40. 0	20,0		
Glucosa	G				0,6				10,0		
Manosa	G									12,5	12,5
Maltosa	G									5,0	5,0
2,4-D	mg				1,6				1,6	1,0	1,0
Agar	G		8,0	6,4	6,0	6,0	6,0	6,0			
Phytigel	G								3,5		
Dicamba	G				1,2				1,2		
Nitrato de plata	mg			8,5	1,7						

Componentes del medio	Unidades por litro	289R	13158 H	13224 B	13266 K	272X	272V	13158	13265 A	13152 T	13152 Z
Carbenicilina Agribio	mg				2,0					100	
Timentina	mg	15 0,0	150. 0						150. 0		150. 0
Cefotaxima	mg	10 0,0	100. 0	25	25				100. 0		100. 0
Mioinositol	G	0,1	0,1			0,1	0,1	0,1		0,25	0,25
Ácido nicotínico	mg										
Piridoxina.HCl	mg										
Casaminoácidos de ensayo de vitaminas	G										
Tampón MES	G										
Acetosiringona	uM										
Ácido ascórbico 10 mg/ml (7S)	mg										
Solución de reserva de vitaminas MS	ml	5,0	5,0			5,0	5,0	5,0			
Zeatina	mg	0,5	0,5								
Sulfato cúprico	mg	1,3	1,3						0,05	1,2	1,2
IAA 0,5 mg/ml (28A)	ml	2,0	2,0								
ABA 0,1 mm	ml	1,0	1,0								
Tiadiazurona	mg	0,1	0,1								
Carbenicilina Agribio	mg										
PPT(glufosinato-NH ₄)	mg										
BAP	mg								0,1	0,5	0,5
Extracto de levadura (BD Difco)	G										
Peptona	G										
Cloruro de sodio	G										
Espectinomicina	mg										
Sulfato ferroso.7H ₂ O	ml										
Tampón AB 20X (12D)	ml										
Sales AB 20X (12E)	ml										
Benomilo	mg							100 .0			
pH		5,6	5,6	5,6	5,6	5,6	5,6	5,6	5,6	5,6	5,6

Tabla 21, continuación.

Componentes del medio	Unidades por litro	557A	810D	
Mezcla de sal basal MS	G			
Macronutrientes N6 10X	ml			
Nitrato de potasio	G			
Sales minoritarias B5H 1000X	ml			
NaFe EDTA para B5H 100X	ml			
Vitaminas de Eriksson 1000X	ml			
Reserva de vitaminas S&H 100X	ml			
Tiamina HCl	mg			
L-prolina	g			
Hidrolizado de caseína (ácido)	g			
Sacarosa	g	20,0		
Glucosa	g			
Maltosa	g			
2,4-D	mg			
Agar	g			
Phytigel	g			
Dicamba	g			
Nitrato de plata	mg			
Carbenicilina Agribio	mg			
Timentina	mg			
Cefotaxima	mg			
Mioinositol	g			
Ácido nicotínico	mg			
Piridoxina.HCl	mg			
Casaminoácidos de ensayo de vitaminas	g			
Tampón MES	g			
Acetosiringona	uM			
Ácido ascórbico 10 mg/ml (7S)	mg			
Solución de reserva de vitaminas MS	ml			
Zeatina	mg			
Sulfato cúprico	mg			
IAA 0,5 mg/ml (28A)	ml			
ABA 0,1 mm	ml			
Tiadiazurona	mg			
Carbenicilina Agribio	mg			
PPT(glufosinato-NH ₄)	mg			
BAP	mg			
Extracto de levadura (BD Difco)	g		5,0	
Peptona	g		10,0	
Cloruro de sodio	g		5,0	

Componentes del medio	Unidades por litro	557A	810D	
Espectinomicina	mg		50,0	
Sulfato ferroso.7H ₂ O	ml			
Tampón AB 20X (12D)	ml			
Sales AB 20X (12E)	ml			
Benomilo	mg			
Fosfato de potasio dibásico	g	10,5		
Fosfato de potasio monobásico anhidro	g	4,5		
Sulfato de amonio	g	1,0		
Citrato de sodio dihidrato	g	0,5		
Sulfato de magnesio 1,0 M	ml	1,0		
pH		5,6	6,8	

Los acontecimientos SSI se identificaron usando un ensayo de PCR múltiples para detectar la presencia/ausencia de ADN T de *Agrobacterium* aleatorio y análisis de estructura de vector, y qPCR para determinar la escisión del gen diana, inserción de FLP (integración aleatoria de ADN T) y la unión de par FRT (1/87). Los datos se resumen en la tabla 22. Se identificó un acontecimiento SSI preciso que tenía una sola copia del gen donador con uniones FRT intactas (1/87), menos la diana y FLP que no tiene inserción estructural. El uso del plásmido auxiliar (pVIR9) en la cepa LBA4404THY- mejoró la recuperación de acontecimientos SSI de esta cepa, en comparación con la recuperación de SSI normal usando la cepa LBA4404THY- sin el plásmido auxiliar. En ausencia del plásmido auxiliar (pVIR9), la cepa LBA4404THY- que alberga vectores ternarios que comprenden dos casetes de expresión de ADN T; consistiendo el primer ADN T en ZmUbiPro::FLP::PINII + FRT1-PMI::PINII + ZmUbiPro::DsRED::PINII-FRT87 (pPHP60577; SEQ ID NO: 75) y, comprendiendo el segundo ADN T loxP-Rab17Pro::Cre + NosPro::WUS::PINII + ZmUbiPro::BBM::PINII-loxP (pPHP44542; SEQ ID NO: 76) provocó una frecuencia de SSI muy baja (0,1 %) en el genotipo PHR03 (solicitud provisional de Estados Unidos n.º 62/296639) en comparación con la frecuencia de SSI (0,4-0,5 %) observada en múltiples RTL en PH184C, un endogámico reacto a SSI de *Agrobacterium*, con la cepa LBA4404THY-71539.

Tabla 22. Frecuencia de transformación y frecuencia de SSI en PH184C con la cepa AGL1 y LBA4404 con pVIR9

Construcción/diana	Agro	Plantas T0 (n.º)	Acontecimientos SSI (n.º)	SSI (%)
79366/ZMRTL A	AGL1	11	5	0,6%
793 66/ZMRTL B	AGL1	26	10	1,1%
793 66/ZMRTL0C	AGL1	39	17	2,3%
79366/ZMRTL A	LBA4404 +pVIR9	6	3	0,4%
793 66/ZMRTL B	LBA4404 +pVIR9	19	4	0,5%
793 66/ZMRTL C	LBA4404 +pVIR9	32	4	0,5%

Un experto en la materia apreciaría que la incorporación del plásmido auxiliar pVIR7 (SEQ ID NO: 34) o pVIR10 (SEQ ID NO: 36) en diferente cepa de *Agrobacterium* puede potenciar la virulencia de la cepa de *Agrobacterium* para mejorar la frecuencia de SSI en múltiples endogámicos de maíz.

Ejemplo 13: Mejora de SSI de *Agrobacterium* usando casetes donadores y/o plásmidos auxiliares con diferentes orígenes de replicación

Otra manera de mejorar el SSI de *Agrobacterium* es construir casetes donadores con diferentes orígenes de replicación bacterianos (Ori) que varían en el número de copias del plásmido (alto, medio, bajo; por ejemplo, RepABC, pRi, pVS1, RK2, etc.) que permite valorar (de alta a baja) la cantidad de moléculas de ADN donadoras suministradas a la célula vegetal. Como alternativa, los casetes donadores que albergan diferente Ori pueden combinarse con plásmidos auxiliares que portan genes adicionales de virulencia. Estos plásmidos auxiliares pueden tener diferentes orígenes de replicación bacterianos (RepABC, pRi, pVS1, RK2), con número de copias de plásmido variado. Este método puede mejorar la eficacia de acontecimientos SSI y recuperación.

Ejemplo 14: Número de identificación de secuencia (SEQ ID NO:)

Tabla 23.

SEQ ID	DESCRIPCIÓN
1	gen aacC1; <i>Pseudomonas aeruginosa</i>
2	ori ColE1; <i>Escherichia coli</i>
3	ori pVS1; <i>Pseudomonas aeruginosa</i>
4	VirB1; <i>Agrobacterium tumefaciens</i>
5	VirB2; <i>Agrobacterium tumefaciens</i>
6	VirB3; <i>Agrobacterium tumefaciens</i>
7	VirB4; <i>Agrobacterium tumefaciens</i>
8	VirB5; <i>Agrobacterium tumefaciens</i>
9	VirB6; <i>Agrobacterium tumefaciens</i>
10	VirB7; <i>Agrobacterium tumefaciens</i>
11	VirB8; <i>Agrobacterium tumefaciens</i>
12	VirB9; <i>Agrobacterium tumefaciens</i>
13	VirB10; <i>Agrobacterium tumefaciens</i>
14	VirB11; <i>Agrobacterium tumefaciens</i>
15	VirG; <i>Agrobacterium tumefaciens</i>
16	VirC1; <i>Agrobacterium tumefaciens</i>
17	VirC2; <i>Agrobacterium tumefaciens</i>
18	VirD1; <i>Agrobacterium tumefaciens</i>
19	VirD2; <i>Agrobacterium tumefaciens</i>
20	VirD3; <i>Agrobacterium tumefaciens</i>
21	VirD4; <i>Agrobacterium tumefaciens</i>
22	VirD5; <i>Agrobacterium tumefaciens</i>
23	VirE1; <i>Agrobacterium tumefaciens</i>
24	VirE2; <i>Agrobacterium tumefaciens</i>
25	VirE3; <i>Agrobacterium tumefaciens</i>
26	VirA; <i>Agrobacterium tumefaciens</i>
27	VirJ; <i>Agrobacterium tumefaciens</i>
28	PHP45981; secuencia artificial
29	PHP64484; secuencia artificial
30	ZmUbiPro; secuencia artificial
31	DsRED; <i>Discosoma sp.</i>
32	FRT1; <i>Saccharomyces cerevisiae</i>
33	FRT87; <i>Saccharomyces cerevisiae</i>
34	pVIR7; PHP70298; secuencia artificial
35	pVIR9; PHP71539; secuencia artificial
36	pVIR10; PHP79761; secuencia artificial
37	ori pRSF1010; <i>Escherichia coli</i>
38	ori pRK2; <i>Escherichia coli</i>
39	casete de selección aadA; <i>Escherichia coli</i>
40	casete de selección npt1; <i>Escherichia coli</i>

SEQ ID	DESCRIPCIÓN
41	casete de selección npt2; <i>Escherichia coli</i>
42	VirH; <i>Agrobacterium tumefaciens</i>
43	VirH1; <i>Agrobacterium tumefaciens</i>
44	VirH2; <i>Agrobacterium tumefaciens</i>
45	VirK; <i>Agrobacterium tumefaciens</i>
46	VirL; <i>Agrobacterium tumefaciens</i>
47	VirM; <i>Agrobacterium tumefaciens</i>
48	VirP; <i>Agrobacterium tumefaciens</i>
49	VirQ; <i>Agrobacterium tumefaciens</i>
50	ori pSC101; <i>Salmonella typhimurium</i>
51	ori p15A, <i>Escherichia coli</i>
52	ori gamma pir R6K; <i>Escherichia coli</i>
53	ori pSa repA; <i>Escherichia coli</i>
54	micro pRK2; <i>Escherichia coli</i>
55	PARDE; <i>Escherichia coli</i>
56	pSaPARDE; secuencia artificial
57	repABC-pRi1724; <i>Agrobacterium rhizogenes</i>
58	repABC-pTi-SAKURA; <i>Agrobacterium tumefaciens</i>
59	repABC; plásmido PR1b, <i>Ruegeria</i> sp.
60	repABC; pNGR234, <i>Sinorhizobium fredii</i>
61	palíndromo de 32 pb de pSB1; secuencia artificial
62	repetición invertida de 142 pb de pSB1; secuencia artificial
63	región tra-trb de pSB1; secuencia artificial
64	trfA; <i>Escherichia coli</i>
65	oriV; <i>Escherichia coli</i>
66	mini pRK2 (oriV-nptIII-trfA); <i>Escherichia coli</i>
67	casete de selección hpt; <i>Escherichia coli</i>
68	UBI-ZMPRO: MO-PAT: CONECTOR PROTEÍNICO:DSRED:PINII
69	pPHP80561
70	pPHP80559
71	pPHP79066
72	pPHP78147
73	pPHP78148
74	pPHP79366
75	pPHP60577
76	pPHP44542
77	SpcN; <i>Streptomyces spectabilis</i>
78	aph; <i>Legionella pneumophila</i>
79	AR-VIRA; <i>Agrobacterium rhizogenes</i>
80	AR-VIRB1; <i>Agrobacterium rhizogenes</i>

SEQ ID	DESCRIPCIÓN
81	AR-VIRB2; <i>Agrobacterium rhizogenes</i>
82	AR-VIRB3; <i>Agrobacterium rhizogenes</i>
83	AR-VIRB4; <i>Agrobacterium rhizogenes</i>
84	AR-VIRB5; <i>Agrobacterium rhizogenes</i>
85	AR-VIRB6; <i>Agrobacterium rhizogenes</i>
86	AR-VIRB7; <i>Agrobacterium rhizogenes</i>
87	AR-VIRB8; <i>Agrobacterium rhizogenes</i>
88	AR-VIRB9; <i>Agrobacterium rhizogenes</i>
89	AR-VIRB10; <i>Agrobacterium rhizogenes</i>
90	AR-VIRB11; <i>Agrobacterium rhizogenes</i>
91	AR-VIRG; <i>Agrobacterium rhizogenes</i>
92	AR-VIRC1; <i>Agrobacterium rhizogenes</i>
93	AR-VIRC2; <i>Agrobacterium rhizogenes</i>
94	AR-VIRD1; <i>Agrobacterium rhizogenes</i>
95	AR-VIRD2; <i>Agrobacterium rhizogenes</i>
96	AR-VIRD3; <i>Agrobacterium rhizogenes</i>
97	AR-VIRBD4; <i>Agrobacterium rhizogenes</i>
98	AR-VIRD5; <i>Agrobacterium rhizogenes</i>
99	AR-VIRF; <i>Agrobacterium rhizogenes</i>
100	AR-VIRE3; <i>Agrobacterium rhizogenes</i>
101	AR-GALLS; <i>Agrobacterium rhizogenes</i>
102	ori BBR1; <i>Bordetella bronchiseptica</i>

Como se usa en este documento, las formas singulares "uno", "una", y "el", "la" incluyen referentes plurales salvo que el contexto indique claramente lo contrario. Por tanto, por ejemplo, una referencia a "una célula" incluye una pluralidad de dichas células y una referencia a "la proteína" incluye referencias a una o más proteínas y equivalentes de las mismas conocidas por los expertos en la materia, y así sucesivamente. Todos los términos técnicos y científicos usados en este documento tienen el mismo significado que el normalmente entendido por un experto en la materia a la que pertenece esta divulgación salvo que se indique claramente lo contrario.

Todas las patentes, publicaciones y solicitudes de patente mencionadas en la memoria descriptiva son indicativas del nivel de especialización de los expertos en la materia a la que pertenece esta divulgación.

LISTA DE SECUENCIAS

5	[0226]		
	<110> Pioneer Hi-Bred International, Inc.		
	Anand, Ajith		
	Bass, Steven		
	Cho, Hyeon-Je		
10	Klein, Theodore		
	Lassner, Michael		
	McBride, Kevin		
	<120> Métodos y composiciones de transformación de plantas mejorada		
	<130> 6469		
15	<150> 62/252229		
	<151> 2015-11-06		
	<160> 102		
	<170> PatentIn versión 3.5		
	<210> 1		
20	<211> 534		
	<212> ADN		
	<213> Pseudomonas aeruginosa		
	<400> 1		
	atgttacgca gcagcaacga tgttacgcag cagggcagtc gccctaaaac aaagtttaggt	60	
	ggctcaagta tgggcatcat tcgcacatgt aggctcggcc ctgaccaagt caaatccatg	120	
	cgggctgctc ttgatctttt cggctcgtgag ttcggagacg tagccaccta ctcccaacat	180	
	cagccgggact ccgattacct cgggaacttg ctccgtagta agacattcat cgcgcttgct	240	
	gccttcgacc aagaagcggg tggttgcgct ctgcgggctt acgttctgcc caagtttgag	300	
	cagccgcgta gtgagatcta tatctatgat ctgcgagtct ccggcgagca ccggaggcag	360	
	ggcattgcc aacgcgctcat caatctcctc aagcatgagg ccaacgcgct tgggtgcttat	420	
	gtgatctacg tgcaagcaga ttacggtgac gatcccgagc tggctctcta taaaaagttg	480	
	ggcatatggg aagaagtgat gcactttgat atcgacccaa gtaccgccac ctaa	534	
25	<210> 2		
	<211> 1221		
	<212> ADN		
	<213> Escherichia coli		
	<400> 2		
	gaccaagttt actcatatat acttttagatt gattttaaac ttcattttta atttaaagg	60	
	atctagggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg	120	
	ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt	180	
	ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg gggttggttg	240	
	ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata	300	
30	ccaaatactg tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca	360	

ES 2 843 556 T3

	ccgcctacat acctcgctct gctaatacctg ttaccagtggt ctgctgccag tggcgataag	420
	tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc	480
	tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga	540
	tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagaaa ggcggacagg	600
	tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc agggggaaac	660
	gcctgggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg	720
	tgatgctcgt caggggggag gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg	780
	ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttata ccctgattct	840
	gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgccgag ccgaacgacc	900
	gagcgcagcg agtcagtgag cgaggaagcg gaagagcgcc tgatgcggta ttttctcctt	960
	acgcatctgt gcggtatttc acaccgcata tgggtgactc tcagtacaat ctgctctgat	1020
	gccgcatagt taagccagta tacactccgc tatcgctacg tgactgggtc atggctgcgc	1080
	cccgcacccc gccaacacccc gctgacgcgc cctgacgggc ttgtctgctc ccggcatccg	1140
	cttacagaca agctgtgacc gtctccggga gctgcatgtg tcagaggttt tcaccgtcat	1200
	caccgaaacg cgcgaggcag c	1221
	<210> 3	
	<211> 2559	
	<212> .ADN	
5	<213> Pseudomonas aeruginosa	
	<400> 3	
	taaaacagct tgcgtcatgc ggtcgtctgc tatatgatgc gatgagtaa taaacaaata	60
	cgcaagggga acgcatgaag gttatcgtg tacttaacca gaaaggcggg tcaggcaaga	120
	cgaccatcgc aaccatcta gcccgcgccc tgcaactcgc cggggccgat gttctgtag	180
	tcgattccga tcccagggc agtgcgcgcg attgggcggc cgtgcgggaa gatcaaccgc	240
	taaccgttgt cggcatcgac cgcccgacga ttgaccgca cgtgaaggcc atcggccggc	300
	gcgacttcgt agtgatcgac ggagcgcccc aggcggcgga cttggctgtg tccgcgatca	360
	aggcagccga cttcgtgctg attccggtgc agccaagccc ttacgacata tgggccaccg	420
	ccgacctggt ggagctggtt aagcagcgca ttgaggtcac ggatggaagg ctacaagcgg	480
	cctttgtcgt gtgcggggcg atcaaaggca cgcgcacatcg cggtgaggtt gccgaggcgc	540
	tggccgggta cgagctgccc attcttgagt cccgtatcac gcagcgcgtg agctaccag	600
	gcactgccgc cgccggcaca accgttcttg aatcagaacc cgagggcgac gctgcccgcg	660
	aggtccaggc gctggccgct gaaattaaat caaaactcat ttgagttaat gaggtaaaga	720
	gaaaatgagc aaaagcacia acacgctaag tgccggccgt ccgagcgcac gcagcagcaa	780

ES 2 843 556 T3

```

ggctgcaacg ttggccagcc tcgcagacac gccagccatg aagcgggtca actttcagtt      840
gccggcgagg gatcacacca agctgaagat gtacgcggta cgccaaggca agaccattac      900
cgagctgcta tctgaataca tcgcgcagct accagagtaa atgagcaaat gaataaatga      960
gtagatgaat tttagcggct aaaggaggcg gcatggaaaa tcaagaacaa ccaggcaccg     1020
acgccgtgga atgccccatg tgtggaggaa cgggcgggtg gccaggcgta agcggctggg     1080
ttgcctgccg gccctgcaat ggcactggaa cccccaagcc cgaggaatcg gcgtgagcgg     1140
tcgcaaacca tccggccccg taaaaatcgg cgcggcgctg ggtgatgacc tggtagagaa     1200
gttgaaggca gcgcaggccg cccagcggca acgcatcgag gcagaagcac gccccggtga     1260
atcgtggcaa gcagccgctg atcgaatccg caaagaatcc cggcaaccgc cggcagccgg     1320
tgcgcgctcg attaggaagc cgccaaggcg cgacgagcaa ccagattttt tcgttccgat     1380
gctctatgac gtgggcaccc gcgatagtcg cagcatcatg gacgtggccg ttttccgtct     1440
gtcgaagcgt gaccgacgag ctggcgaggt gatccgctac gagcttccag acgggcacgt     1500
agaggtttcc gcagggcccg caggcatggc cagtgtgtgg gattacgacc tggtagtgat     1560
ggcggtttcc catctaaccg aatccatgaa ccgataccgg gaagggaagg gagacaagcc     1620
cggccgcgctg ttccgtccac acgttgcgga cgtactcaag ttctgccggc gagccgatgg     1680
cggaaagcag aaagacgacc tggtagaaac ctgcattcgg ttaaacacca cgcacgttgc     1740
catgcagcgt acgaagaagg ccaagaacgg ccgcctggtg acggtatccg agggtgaagc     1800
cttgattagc cgctacaaga tcgtaaagag cgaaaccggg cggccggagt acatcgagat     1860
cgagctagct gattggatgt accgcgagat cacagaaggc aagaaccggg acgtgctgac     1920
ggttcacccc gattactttt tgatcgatcc cggcatcggc cgttttctct accgcctggc     1980
acgccgcgcc gcaggcaagg cagaagccag atggttggtc aagacgatct acgaacgcag     2040
tggcagcgcc ggagagttca agaagtctg tttcacctg cgcaagctga tcgggtcaaa     2100
tgacctgccg gagtacgatt tgaaggagga ggcggggcag gctggcccga tcctagtcac     2160
gcgctaccgc aacctgatcg agggcgaagc atccgccggg tcctaagtga cggagcagat     2220
gctagggcaa attgccctag caggggaaaa aggtcgaaaa ggtctctttc ctgtggatag     2280
cacgtacatt gggaaccaa agccgtacat tgggaaccgg aaccctaca ttgggaaccc     2340
aaagccgtac attgggaacc ggtcacacat gtaagtgact gatataaaag agaaaaaagg     2400
cgatttttcc gcctaaaact ctttaaaact tattaanaact cttaaaaacc gcctggcctg     2460
tgcataactg tctggccagc gcacagccga agagctgcaa aaagcgccta cccttcggtc     2520
gctgcgctcc ctacgccccg ccgcttcgcg tcggcctat      2559
<210> 4
<211> 720
<212> ADN
5 <213> Agrobacterium tumefaciens
<400> 4

```

ES 2 843 556 T3

	atgcttaaga gatcgggggtc gctttctctt gccttgatgg tctccttctg ttcgtcgagc	60
	cttgccacgc cactctcatc tgctgagttt gaccatgttg ctcgcaagtg tgcccatca	120
	gttgcgacat ctacgcttgc ggcgatagct aaggtggaga gtcgctttga tccttttagcg	180
	attcatgaca acacgaccgg cgaaacgctt cactggcaag atcacagcca agcaacccaa	240
	gtcgtcaggc accgtctcga tgcacggcat tcgctggatg ttggcctcat gcaaataaac	300
	tctcgaaatt tttctatgct cggctcgaca cctgacgggt cgctccaggc gtgcacatca	360
	ttatctgccg ctgcaaacat gctgaaaagt cgttatgcag gggcgaaac gattgacgag	420
	aagcaatttg cgcttcgtcg ggcgatctcc gttacaaca ccggtaatth catcgcggt	480
	tttgcaaacg gctacgtgcy aaaagttgaa acagctgctc aatcgctggt gcccggtta	540
	atcgagcctc caaaagacga tcacgaggcg ctaaaatccg aagagacgtg ggatgtttgg	600
	gggtcatatc agcgccgctc gcaggaggat ggcgctggcg gtttaatcgc tccgccaccg	660
	ccacaccagg acaacggcaa atccgcagac gacaatcaag tcttattcga cttatactaa	720
	<210> 5	
	<211> 365	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 5	
	atgcgatgct ttgagagata ccgtttacat ctaaatcgcc tctcgctctc gaatgcgatg	60
	atgcgcgtga tatcgagctg cgcaccaagc ttgtgcggtg caattgcatg gagcatttcc	120
	tcatccggac ccgcccgcagc gcaatctgcy ggtggcgga ctgacccgc cacaatggtt	180
	aacaatatat gcacgtttat ccttggtccg ttcggccagt cactcgctgt tctcggcatt	240
	gtcgttatcg ggatctcctg gatgttcggg cgggcttcgc ttgggctggt tgccggcgctc	300
	gtcggcgga ttgttatcat gtttggggcg agcttcctcg gccaaacgct cactggcggt	360
	agttg	365
	<210> 6	
	<211> 326	
10	<212> ADN	
	<213> Agrobacterium tumefaciens	
	<400> 6	
	atggctgacg gtttggaaga atcgaccctt tacctcgag ccacacggcc cgcattgttt	60
	cttggggtgc cactgacatt ggcagggtta ttcattgatgt tcgcccgtt tgcattcgtt	120
	atcgttcaga acccgctcta cgaagtcgtt ctgctgccgt tatggtttgc agcccggtc	180
	atcgctggagc gagactacaa tgcggcgagc gtcgtcctgc tatttttgcg gaccgggga	240
	agaagcattg atagtgcagt ttgggggggc gctactgtta gcccaaatcc aattagggtt	300
15	ccccacgag ggagaggaat ggtgtg	326
	<210> 7	
	<211> 2370	
	<212> ADN	
	<213> Agrobacterium tumefaciens	
20	<400> 7	

ES 2 843 556 T3

atgctcggcg cgagtggaac gaccgaaaga tccggtgaga tctatctccc ttatatggc	60
cacctcagcg accatatcgt ccttcttgaa gacggatcga tcatgacat tgcgagaatt	120
gatggcggtg cattcgagct tgaggaaact gaaatgcgca atgcgcgttg tcgtgcgttc	180
aacacgctgt tgcgcaatat cgctgatgat catgtgtcaa tatatgctca cctcgtacgt	240
catgccgacg tgccatcatc ggcgccgga cacttccgta gtgttttcgc cgctagcctg	300
aacgaagctt ttgaacagcg cgtgctctcc ggccaactcc tccgcaatga acacttcctt	360
acgttgattg tctaccaca ggcggttta gggaaggtaa agaggaggtt caccaagcta	420
agcggaaaaa gggaaaacga tctcacgggc cagatcagga acatggaaga tctttggcat	480
gttgctcgtg gctctcttaa agcgtatggc ctgcatcgtc ttggcatccg cgagaagcag	540
ggtgtgctct tcaccgaaat tggcgaagcg ctacggttga tcatgactgg tcggttcaca	600
ccggttccg tcgtcagcgg ctactcggc gcttcgattt ataccgacag agtcatttgc	660
ggcaagcgag gactcgagat cagaacgcca aaagacagtt acgttggatc catctattcg	720
tttcgcgaat accctgcaaa aacacggccg ggcgtgctca acgcgctgct atccctcgat	780
tttccacttg ttctcacgca gagtttttcg ttctgactc gcccgcaagc gcacgcgaaa	840
cttagcctca aatcgagcca gatgctgagt tccggcgata aagccgtgac tcaaatcggc	900
aaattatccg aggctgagga cgcacttgcg agcaacgaat tcgttatggg ctcacatcat	960
ttgagccttt gcgtctatgc agacgatctc aatagtcttg gggacagggg cgcgcgggct	1020
cggacacgaa tggcggatgc aggtgccgtg gttgtccaag aaggtattgg tatggaagcg	1080
gcctattggt cccaattgcc ggggaatttt aagtggcgca cacgccctgg cgcaatcact	1140
tcacgcaatt tcgcagggtt tgtctctttc gaaaactttc cagagggcgc cagctcaggc	1200
cactggggca acgcgattgc ccgatttcgt accaatggcg gaacgccttt cgactatata	1260
ccgcatgagc acgatgttgg catgacggca atattcgggc ctatcgggag gggtaagacg	1320
acgctcatga tgtttgttct agccatgctc gaacagagca tggtcgaccg tgcaggtacg	1380
gtcgtgttct ttgacaagga ccggggtggc gaattgctgg ttcgcgccac aggaggaaca	1440
tattttggcac ttcacagagg cacaccagc gggttggcgc cgttgctgg cctagaaaac	1500

ES 2 843 556 T3

acagcagcct cacacgattt tctgcgcgaa tggatcgtgg ctctcatcga gagtgatggg 1560
 cggggtggga tttctccgga agagaaccgc cgtctgggtcc gtggtatcca tcgtcagctc 1620
 tcgtttgatc cacaaatgcg ttcaatcgcg gggttacgtg aatttttgtt gcatgggccc 1680
 gccgaaggcg caggagcgcg gctccaacgc tgggtgccggg gccatgcgct tggctgggca 1740
 tttgacggcg aagttgacga agtaaaagta gatccgtcga ttaccggctt cgacatgacg 1800
 catcttctcg aatacgagga agtatgcgct cccgctgcag catatctcct gcatcggatt 1860
 ggagccatga tcgacggccg ccgttttgtg atgagctgcg atgagtttcg cgcctatttg 1920
 ttaaaccccta aatttttcgac tgtcgtcgac aaattcctcc tgaccgttcg aaaaaacaac 1980
 gggatgctaa tactggcaac gcagcaacca gagcatgttc tggaatcgcc gctaggagcc 2040
 agcttggttg cgcaatgtat gacgaagatt ttctatccat caccaaccgc agatcgatcg 2100
 gcttatgtcg atggactgaa atgtaccgaa aaggaatttc aggcgatccg tgaagacatg 2160
 acggtcggca gccgtaagtt tcttcttaaa cgagaaagtg gaagcgtcat ctgcgaattt 2220
 gatctgcggg atatgctga atatgtcgcc gtgctttcgg ggcggtccaa cacggtgcgc 2280
 tttgcaactc gactacgcga ggcacaagaa ggcaactcat ctggctggct cagcgaattc 2340
 atggcccgtc accacgaggc agaagattga 2370

<210> 8

<211> 663

<212> ADN

5 <213> Agrobacterium tumefaciens

<400> 8

atgaagacga cgcaacttat tgcaacagtt ttgacctgca gctttctata tattcagccc 60
 gcgcggggcg agtttggttg tagcgacccg gcaacggagg ctgagacgct cgcgactgcg 120
 ctgcgcgactg cggagaatct cactcagact atagcgatgg ttacgatgtt gacgtcggcc 180
 tacggcggtta ctggactact gacttcgctc aaccagaaaa atcagtatcc ttcgacgaag 240
 gacctagaca atgaaatggt ttgcgcgcga atgccaatgt cgaccacggc acgtgcgac 300
 accagcgata cagatcgtgc agtcgtgggt agtgatgctg aagcggacct gttgcgacg 360
 cagatcaccg gttccgcaaa cagcgtggc attgcggtg acaatctgga aacgatggac 420
 aaacgcttga cggcgaatgc tgatacgtct gctcagcttt cccgatctcg caatatcatg 480
 caggcaaccg tgaccaatgg tttgcttctc aagcagatcc atgacgcaat gattcaaaat 540
 gtacaggcga caagcctatt aacgatgact accgcgcagg ccggccttca cgaggcgga 600
 gaggcggccg ctcaacgcaa ggagcatcaa aagaccgctg tcatctttgg tgcctcccc 660
 taa 663

<210> 9

<211> 888

<212> ADN

10 <213> Agrobacterium tumefaciens

<400> 9

ES 2 843 556 T3

	atgaatttca cgattccggc gccgtttacg gccattcata cgatcttcga tgtagccttc	60
	acgacaggct tggactcgat gcttgagact atccaggagg cggtagtgat gccattgatc	120
	gcctgtgtca ctctttggat tattgttcag ggtattttag tcatacgcgg cgaagtcgat	180
	accgtagcgc gtatcactcg ggtgatcacg gtcaccatcg ttgttgctct aattgttggg	240
	caggctaact accaagacta tgtggtttcc atcttcgaaa agacggtccc aaactttgtt	300
	cagcagttta gtgtaaccgg cttgcctctg cagactgttc cggcacagtt ggatacaatg	360
	ttcgccgtga cccaggccgt ttttcagaaa atcgcacccg aaatcgggcc gatgaacgac	420
	caggacatcc ttgctttcca aggggcacag tgggtccttt acggcacgct ctggtctgcc	480
	ttcggagttt acgacgccgt tggaattctc acgaaagtgc ttctcgcgat cgggcctctg	540
	atcctcgtcg gatataatct tgatcgcacg cgggacatcg cagctaagtg gatcgggcaa	600
	cttatcacct acggtctctt gcttctctc ttaaacctcg tggcaacgat cgtcatccta	660
	accgaagcga ctgcgctcac ccttatgctt ggtgtaatca ccttcgccgg tacgaccgcg	720
	gccaaagtca ttgggtcttta cgaactcgat atgttttttc tgacagggga tgcgctcatt	780
	gtcgctttgc cggcgatcgc cggcaacatt ggaggcagtt actggagcgg cgcaacccaa	840
	tctgccagca gcttgtaccg tcgcttcgct caggttgagc gaggctag	888
	<210> 10	
	<211> 154	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 10	
	atgaagtatt gcctgctgtg cctagtgtgc gctttgagcg gctgccagac aaacgacaca	60
	ttagcgagct gcaaaggccc gatcttcccg ctgaatgtgg ggcgatggca gcctactccg	120
	tcagatcttc agctcggcaa ttcggttggga cgct	154
	<210> 11	
	<211> 710	
10	<212> ADN	
	<213> Agrobacterium tumefaciens	
	<400> 11	
	atgacggggc ctgaatatgc catgctagtg gcgcgcgaaa gccttgccga gcactataag	60
	gaagtagaag cctttcaaac cgcgcgagcg aaatcggcgc gacgtctctc caaactcatt	120
	gcagctgtcg cagctatcgc gattttggga aatgttgctc aagcgttcgc tatagccaca	180
	atggtgccgt tgagcaggct tgtgcccgta tatctatgga tacggccgga cggcaccgtt	240
	gacagcgagg tgtctgtctc gcgattgcct gcaactcaag aggaggccgt cgttaacgcc	300
	tcattgtggg agtacgttcg cctgcgcgag agttatgatg ccgacaccgc tcagtacgcc	360
	tacgacctgg tatcgaaact cagtgcacca acagtgcgcc aggattacca gcaattcttc	420
	aactatccca atcccagttc gcctcaagtc attcttggca aacgcggcag ggtggaggtc	480
	gagcacatcg cttcaaatga tgtaactcca agcacgcagc aaattcgcta taaaaggacc	540
	ctcgtcgctt acggcaaaaat gcctgtggtg agtacgtgga ccgcgacagt tcgctacgaa	600
	aaggtgacca gcttgcccgg cagattgaga ctaaccaacc cggcaggtct ggttgtcacc	660
15	tcctatcaga catcggaaga taccgtttca aacgtaggcc acagcgaacc	710
	<210> 12	

ES 2 843 556 T3

<211> 878
 <212> ADN
 <213> Agrobacterium tumefaciens
 <400> 12
 atgatcagaa aagcactttt catttttagca tgtttatttg ccgctgcgac tgggtcggag 60
 gctgaagaca ctccaatggc gggcaagcta gatccgcgca tgcgttattt ggcttacaat 120
 cccgatcaag tgggtcgcct ctcgacggcg gttggagcta ctttggtcgt aacattcgcc 180
 acgaacgaaa cggtgacagc ggttgccgtt tcaaatagca aagatctagc agccctaccg 240
 cggggaaatt atctattttt caaggcaagc caggtcctca cgcctcagcc agtaatcgtg 300
 ctaaccgcaa gcgactccgg gatgcgccgt tatgttttca gtataagttc caagactctg 360
 tcccacctcg ataaagagca gcccgatctc tattacagcg tccaattcgc ctaccccgcc 420
 gacgatgcgg cggctcggcg aagggaggca caacagaagg ctgttggtga cagactacac 480
 gcggaagcac aatatcaacg gaaagctgag aatttatttg atcagcctgt cacagccctt 540
 ggtgcggcgg acagtaattg gactacgctc gcccaaggcg atcgttcgct gttgccactc 600
 gaagtcttcg acaatggatt tacgacggta ttccactttc cgggcaatgt acgcataccc 660
 tccatctaca ccatcaatcc tgatggcaag gaagctgttg ccaactattc agttaaaggg 720
 agcgatgtcg agatttcttc ggtttccoga ggttggcgctc tgagggatgg ccacacagta 780
 ctatgtatct ggaacaccgc ttacgatccc gttggccaaa ggccgcaaac gggcacgggtg 840
 5 aggcccgatg tgaaacgcgt cctgaagggg gcgaaggg 878
 <210> 13
 <211> 1134
 <212> ADN
 <213> Agrobacterium tumefaciens
 10 <400> 13
 atgaataacg atagtcagca agcggcacat gaggttgatg catctggatc cctggtctcc 60
 gacaaacatc gccggcgtct ttcgggggtct cagaaattga tcgtcggagg tgcgttctc 120

ES 2 843 556 T3

gcgttatcat taagcctcat ttggctaggt gggcgccaaa agaaggtgaa tgagaacgca	180
tcgccgtcaa ctttgatcgc aacaaacacc aagccatttc atccagctcc gattgaggtg	240
ccgccggatc ctccagcggc tcaagaggct gttcagcctg ctgctcctct accgccgagg	300
ggcgaaccgg agcggcatga gccacggccg gaagaaacac cgatttttgc atatagcagc	360
ggcgatcaag gggtcagcaa acgcgccatt cagggcgaca cgggccgaag acaagaaggc	420
aagcgtgacg acaactcctt gccgaatggc gaagtgtccg gcgagaacga tttgtcgata	480
cgtatgaaac ccaccgagct gcagcccagc agcggcacgc tcttgccgca ccccgatttt	540
atggtaacgc aagggacaat aattccgtgc atcttgcaaa ccgcaatcga cacaaatttg	600
gcaggctatg taaagtgtgt cttgcctcag gatattcgtg gaacaacgaa caatatcgtg	660
cttcttgatc gtggcaccac cgttgtttggc gaaatacagc gtggcttgca acagggagat	720
ggcgcggttt ttgtgtttgt ggatcgcgcc gagacacctg accatgcgat gatctcgta	780
acatcgccaa gcgcggacga actcggtcgc tcaggattgc cgggctcggc cgacagccac	840
ttctggcagc gttttagcgg agctatgctc ttgagtgtt ttcaaggcgc cttccaggca	900
gctagcacct acgccggcag ctcggtggc gggatgagct tcaacagctt tcaaaataac	960
ggtgagcaga caactgagac agcccttaag gcaaccatca acataccgcc aaccctgaag	1020
aagaatcagg gtgacaccgt ttccattttc gtagcacggg acctcgattt ctttgggtgt	1080
taccagctcc gcctgactgg cggcgccacg cgggggagga accgccgctc ttaa	1134

<210> 14

<211> 1032

<212> ADN

5 <213> Agrobacterium tumefaciens

<400> 14

atggaagtgg atccgcaact acgctttctt ctgaagccga ttttggaatg gctcgatgac	60
ccgaagactg aagaaattgc gatcaatcga cctggagagg catttggtgcg ccaagccggc	120
attttttacca agatgccttt gcccgctctt tatgatgatc ttgaagatat cgctatttta	180
gcggggcgcg tgagaaagca ggatgtcggc ccacgtaacc ccctctgcgc cactgaactt	240
cctggtggtg aacgactaca aatctgtctg ccgccgaccg ttccctcggg caccgtcagc	300
ttgaccattc gacggccaag ctcgctgttt tctggtctta aagaagtctc ctcccgttat	360
gatgcttcga ggtggaacca gtggcagaca cgaaggaaac gccaaaatca ggatgatgaa	420
gctatccttc agcattttga caacggggat ttggaagcgt ttctgcacgc atgcgtcgtc	480
agccgactga cgatgttgct atgtggccct accggaagcg gcaagacaac aatgagcaag	540
accttgatca gcgccatccc ccccaggaa aggctaata ccatagaaga tacgctcgaa	600
ctcgctattc cacacgataa tcatgttaga ctactctact ccaagaacgg tgctgggctg	660

ES 2 843 556 T3

	ggcgctgtga ggcgcgagca cttgctccaa gcaagtctgc gtatgcggcc ggaccggata	720
	ttgcttggtg agatgcgcga cgatgcagca tgggcttatc tgagtgaagt cgtctcgga	780
	catccgggat cgatttcaac aatacacggc gcgaatccaa tccaaggatt caagaaactg	840
	ttttcccttg tcaaaagtag cgcccaaggt gctagcttgg aagatcgac actgattgac	900
	atgctctcta cggcgatcga tgtcatcatt ccattccgtg cctatgagga cgtttatgaa	960
	gtaggcgaga tctggctcgc ggcggacgca cgacgccggg gcgagaccat aggcgatctc	1020
	cttaataaat ag	1032
	<210> 15	
	<211> 804	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 15	
	atgattgtac atccttcacg tgaataattc tcaagcgtg tgaacaaggg ttcagatttt	60
	agattgagag gtgagccgtt gaaacacgtt cttcttatcg atgacgatgt cgctatgcgg	120
	catcttatta tcgaatacct tacgatccac gccttcaaag tgaccgcggg agccgacagc	180
	accagtttca ctagagtact ctcttcgcg acggtcgatg tcgtggttgt tgatctaaat	240
	ttaggtcgtg aagatgggct tgagatcgtt cgaaatctgg cggcaaagtc tgatattcca	300
	atcataatta tcagtggcga ccgccttgag gagacggata aagttgttgc actcgagcta	360
	ggagcaagtg attttatcgc taagccgttt agtacgagag agtttcttgc acgcattcgg	420
	gttgcccttg gcgtgcgccc caacgttgtc cgctccaaag accgacggtc tttttgtttt	480
	actgactgga cacttaatct caggcaacgt cgcttgatgt ccgaagctgg cggtgaggtg	540
	aaacttacgg caggtgagtt caatcttctc ctgcggtttt tagagaaacc ccgcgacgtt	600
	ctatcgcgcg agcaacttct cattgccagt cgagtacgcg acgaggaggt ttacgacagg	660
	agtatagatg ttctcatttt gcggctgcgc cgcaaacttg aggcggatcc gtcaagccct	720
	caactgataa aaacagcaag aggtgccggt tatttctttg acgcggacgt gcaggtttcg	780
	cacgggggga cgatggcagc ctga	804
	<210> 16	
	<211> 696	
	<212> ADN	
10	<213> Agrobacterium tumefaciens	
	<400> 16	
	atgcaacttt tgacgttttg ttctttcaaa gggggtgctg gcaaaaccac cgcaactcatg	60
	ggcctttgcg ctgctttggc aaatgacggt aaacgagtg ccctctttga tgccgacgaa	120
	aaccggcctc tgacgcgatg gagagaaaac gccttataaa gcagtacctg ggatcctcgc	180

	tgtgaagtct attccgccga cgaaatgccc cttcttgaag cagcctatga aaatgccgag	240
	ctcgaaggat ttgattatgc gttggccgat acgcgtggcg gctcgagcga gctcaacaac	300
	acaatcatcg ctagctcaaa cctgcttctg atccccacca tgctaacgcc gctcgacatc	360
	gatgaggcac tatctaccta ccgctacgtc atcgagctgc tgttgagtga aaatttggca	420
	attcctacag ctgttttgcg ccaacgcgtc ccggtcggcc gattgacaac atcgcaacgc	480
	aggatgtcag agacgctaga gaggcttcca gttgtaccgt ctcccatgca tgaaagagat	540
	gcatttgccg cgatgaaaga acgcggcatg ttgcatctta cactactaaa cacgggaact	600
	gatccgacga tgcgcctcat agagaggaat cttcggattg cgatggagga agtcgtggtc	660
	atttcgaaac tgatcagcaa aatcttggag gcttga	696
	<210> 17	
	<211> 609	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 17	
	atggcaattc gcaagcccg c attgtcggtc ggcgaagcac ggcggcttgc tgggtgctga	60
	cccagatcc accatcccaa cccgacactt gttccccaga agctggacct ccagcacttg	120
	cctgaaaaag ccgacgagaa agaccagcaa cgtgagcctc tcgtcgccga tcacatttac	180
	agtcccgatc gacaacttaa gctaactgtg gatgccctta gtccacctcc gtccccgaaa	240
	aagctccagg tttttctttc agcgcgaccg cccgcgcctc aagtgtcgaa aacatatgac	300
	aacctcgttc ggcaatacag tccctcgaag tcgctacaaa tgattttaag gcgcgcgttg	360
	gacgatttcg aaagcatgct ggcagatgga tcatttcgcg tggccccgaa aagttatccg	420
	atcccttcaa ctacagaaaa atccgttctc gttcagacct cacgcatgtt cccggttgcg	480
	ttgctcgagg tcgctcgaag tcattttgat ccgttggggg tggagaccgc tcgagctttc	540
	ggccacaagc tggctaccgc cgcgctcgcg tcattctttg ctggagagaa gccatcgagc	600
	aattggtga	609
	<210> 18	
	<211> 444	
10	<212> ADN	
	<213> Agrobacterium tumefaciens	
	<400> 18	
	atgtcaaaac acaccagagc cagctcgagt gagactacca tcaaccagca tcgatccctg	60
	aaagttgaag ggttcaaggt cgtgagtgcc cgtctgcgat cggccgagta tgaaaccttt	120
	tcctatcaag cgcgcctgct gggactttcg gatagtatgg caattcgcgt tcggtgctgc	180
	cgcacgagg gctttctcga aatagatgca gacacacgag aaaagatgga agccatactt	240
	cagtcocatc gaatactctc aagcaatgta tccatgcttc tatctgccta cgcgaagac	300
	cctcgatcgg atctggaggc tgtgcgagat gaacgtattg cttttggcga ggctttcgcc	360
	gccctcgatg gactactgcg ctccattttg tccgtatccc ggcgacggat cgacggtcgc	420
15	tcgttactga aaggtgcctt gtag	444
	<210> 19	
	<211> 1275	
	<212> ADN	
	<213> Agrobacterium tumefaciens	
20	<400> 19	

ES 2 843 556 T3

atgcccgatc gcgctcaagt aatcattcgc attgtgccag gaggtggaac caagaccctt	60
cagcagataa tcaatcagct ggagtacctg tcccgaagagg gaaagctgga actgcagcgt	120
tcagcccggc atctcgatat tcccgttccg ccggatcaaa tccgtgagct tgcccaaagc	180
tgggttacgg aggcgggat ttatgacgaa agtcagtcag acgatgacag gcaacaagac	240
ttaacaacac acattattgt aagcttcccc gcaggtaccg accaaaccgc agcttatgaa	300
gcaagccggg aatgggcagc cgagatgttt gggtcaggat acgggggtgg ccgctataac	360
tatctgacag cctaccacgt cgaccgcgat catccacatt tacatgtcgt ggtcaatcgt	420
cggaacttc tggggcagg gtggctgaaa atatccaggc gccatcccca gctgaattat	480
gacggcttac ggaaaaagat ggcagagatt tcaacttcgtc acggcatagt cctggatgcg	540
acttcgcgag cagaaagggg aatagcagag cgaccaatca catatgctga atatcgacgc	600
cttgagcggg tgcaggctca aaagattcaa ttcgaagata cagattttga tgagacctcg	660
cctgaggaag atcgtcggga cctcagtcga tcgttcgatc catttcgatc ggacgcacct	720
gccggcgaac cggaccgtgc aaccgcacat gacaaacaac cgcttgaacc gcacgcccg	780
ttccaggagc ccgccggctc cagcatcaaa gccgacgcac ggatccgcgt accattggag	840
agcgagcggg gtgcccgaacc atccgcgtcc aaaatccctg taactgggca tttcgggatt	900
gagacttcgt atgtcgctga agccagcgtg cccaaacaaa gcggcaattc cgatacttct	960
cgcccggtag ctgacgttgc catgcacaca gtcgagcgcc agcagcgatc aaaacgacgt	1020
catgacgagg aggcaggtcc gagcggagca aaccgtaaaa gattgaaggc cgcgcaagtt	1080
gattccgagg caaatgtcgg tgagcccgcac ggtcgcgatg acagcaacaa ggcggctgat	1140
ccggtgtctg cttccatccg taccgagcaa ccggaagctt ctccaacgtg tccgcgtgac	1200
cgtcacgatg gagaattggg agaacgcaaa cgtgcaagag gtaatcgtcg cgacgatggg	1260
cgcgggggga cctag	1275
<210> 20	
<211> 635	
<212> ADN	
5 <213> Agrobacterium tumefaciens	
<400> 20	
atggcaaagt gtcagttcac gatacgtctc gtcgcccgg cctccgtcgg actgacaggc	60
gaacggcgtg gagccgcac cgcctctagc tctgcaactgt ccaatgttca aagagatgtt	120
agggataggc tgattccaac tagctcacca agattaccaa atgcagccat attgcgtgat	180
tcctcgggaa gagcgtcgac tggctcggg tacatggcgg ctactcttca ttggtctgcg	240
atcgcgccat tatcgctaataaacagcaac gacctggctc cggccgctta tgactttgag	300
acgcgaaata acgcaagaaa tgtgactgcc aaagtgcgca gggcagtcgc tgttcccaag	360
caaggcgggc tcggcaaaac gctcgcaccc gtacccttta gtacacgtat atcaagggtc	420
aattccgacc gaagactgcc cgctgacgca gaagaccgcc ctgaaacgcg cgacccccag	480
aaaggacgtg gcagtcatgg tgcgacgcca accttacatg aaaagattgg aaccgcgttt	540
gctcgaagat tgcgaaagca tacgtactat attgtttgca gttgctgcca gaccgggagc	600
gcgttgacga tgggtgcaaa gatttcggtg aagtc	635
<210> 21	

ES 2 843 556 T3

<211> 1959

<212> ADN

<213> Agrobacterium tumefaciens

<400> 21

```

atgaactcca gcaagacttc gccccagcgt atgaccctga gcatcgtatg ttcgctggca      60

gccggttttt gtgcgccag ctgctatgta acgttccgcc ggggcttcaa cggcgaagcg      120

atgatgacgt tcgacgtttt cgctttttgg tatgagaccc cgctttactt gggttatgcc      180

agcacogtct tctggcgtgg tttatctgtt gtcattcttta cctcgctgat cgttctttca      240

agtcagctca tcatatcgct gcgcaatcag aagcatcatg ggacagctcg ttgggcagaa      300

attggcgaaa tggggcatgc tggttatctg cagcgttaca gtcgcatcaa ggggccgac      360

tttggaaga catgtggtcc cctttggttc ggcagttatt tgaccaatgg cgaacagcca      420

cacagtcttg tcgtcgcgcc aacgcgtgct ggcaaaggcg tcggcatcgt cattccaacg      480

ctgttgacct tcaagggctc ggtaatcgcc cttgacgtca agggagaatt gtttgaactg      540

acgtccagag cagcgaagc gagcggcgac gcagttttca agttctcccc cctagatcct      600

gagcgggaaga ctattgtta caatccggtc ctggatattg ccgcacttcc gcccgaaacg      660

cagttcactg aaacacgccg tctagctgcg aaccttatta cggctaaggg aaaggagca      720

gaaggcttta ttgacggcgc acgtgacctg ttcgtcgcgg gaatccttac ctgcattgag      780

cgtggcacac caacgattgg cgcggtatat gacctatttg cgcagcctgg cgaaaagtat      840

aagctttttg cgcaactcgc ggaggaaagc ctaaacaag aggctcagcg tatcttcgat      900

5 aatatggcgg gcaacgacac gaaaattctg acatcgtaca cctctgtgct gggcgacggt      960

ggactgaacc tgtgggctga tccgcttata aaagcagcga caagccggtc agacttttcc      1020

gtttacgata tccggaggaa gaagacctgc atttatcttt gtgtcagtc caacgatctg      1080

gaggtcttgg caccacttat gcgcctgatg tttcagcagc tcgtgtcaat cttgcagaga      1140

tcgctgccag gtgaagacga gtgccatgaa gttttatttc tcctcgacga attcaaacac      1200

ctgggcaagc ttgaggccat agagaccgag atcacaacca tcgccggtta cagaggccgc      1260

tttatgttta ttattcaaag tctttcggcc ttgtcgggca catacgatga cgcaggaaaa      1320

caaaaactttc tgagcaatac tggcgtaaca gtatttatgg ccacggctga tgacgaaact      1380

ccaacctaca tctcaaaaag tatcggcgaa tatacgttta aagcgcgttc gacctttac      1440

agtcaagcca gaatgttcga ccacaacatc cagatttctg atcaaggtgc agcccttttg      1500

cgccccgaac aagtgcgcct gctagacgat cagagtgaag tcgttctcat caaagggcga      1560

cctccactca aattacgaaa ggtgcagtat tattccgata gtacgctgaa aggccttttc      1620

gaacgccaga tgggctctct gcctgagccc gcacccttga tgctttccga ctatagcaac      1680

gatcaagttc aataccactt ggctccgata gcaaatttta atgaggatgc tgcaccgcaa      1740

aacagaactg tggccgagga ccatggaagt gttaaagtcg gtgctgatat ccctgaacgc      1800

gtgatgggaa taaatggtga cgaggaacaa gccgatgcgg gcgagatacc gccggaatcg      1860

gttgtgcctc cagaattgac gctcgtctct accgctcaac agcaattgtt ggaccagatt      1920

attgcacttc agcaaagatc gaggtccgca ccggcatag      1959

```

<210> 22

ES 2 843 556 T3

<211> 2511

<212> ADN

<213> Agrobacterium tumefaciens

<400> 22

atgacaggaa agtcgaaagt tcacataaga ggttcggctg acgcgcttcc tgacgttcct	60
ggcgggaagta ctaccgcccc ttttttaacc gaaccttctc gggatcaggt tgatgcctcg	120
tttgaggtcc aaaccgacta cagccagtct acttccgtgt cgtttaccta tgatgggtgtt	180
ggacttggtc ctgccgagcg tgcggcttac gagaactggt gcgaaccggg ccggcccact	240
tggaaagatc ttataatcaa ggcacgtgtc gatccgattg acgatgtgac ctggctccga	300
gatttagaag aggacacccc ctcaaccttc agatacgaag ggatgcctct gggcatcggg	360
gaacgacagg cctacgaaaa ttggcaagag gacgctcagc cgacatggga agaccttggt	420
gtcagcgcac gcttgacgga acttggccgt ccacacggga ttaccggcga gtatacatcc	480
ctcgcaggat cgaagaatac aagttcaatt tcattgaagc gaaagcggag caacttaatt	540
gatgatgaga attcatccgg atcgttttca tatgacggga tgaagctcgg ggaagccgag	600
5 cgttctgcat atggtgactg ggccgaggcg gagccacca cgtggaaaga tttggtattg	660

ES 2 843 556 T3

```

agggcacgcg tttcctcgat caatgactct gcttggtttt ttgattcaca aacatcttca 720
tcatcatttg aatacaacgg tgttcccttg ggcgagccgg aacggcaggc tctcagacaa 780
tggaaggag acgctcagcc tacctgggaa gatctcgttg ttaacgcgcg tatggcagaa 840
ctttgccatg ctggttgat tgaaggtcaa aaaggttgct ttgaagagcg cggggaggct 900
ctgcccgcgt cggaacgcgg ttcgcaacgc cccattggtc aacggacaga ttcctccgat 960
tcttttgtgt atgatggcac aaggctcgga gcacctgagc gaactgctta tgaacgctgg 1020
agtaagaggg aacggccgac ttgggaagat ctcatcttag atgcacacca ggcaggact 1080
gaaagtgacg ctgttacgac ccaagcgatt ggtcagtcgt cctcaccggt tttcttatat 1140
gaaggaaagt cgctcggaga cagggaacga aaggcttacg aaaaatggcg gcagccagcc 1200
caaccgcgat ggcaaaatct tgtagtgaac gctcgtctgg cagaaatcga tccctcagcc 1260
tggtattgcc atgagcgcga tccgcttgat gatagcgagc cgcttggtcg cccgtcgtac 1320
acaagcttga cggatagatc agacgtccct ttagacgac aatcaatcta tcgtcgttcc 1380
gacctagtaa gggagcaggt gccagaatcg tctcaaaggc aattcgcagc atgttcagaa 1440
tctgaaacga ggcctgtgca atggtttact gcttctgggt cagatgcaaa caatacggaa 1500
aatatcaccc ccagcgatcc cgtcgatcgc acgggtggag ttaagcggct aggtccaaa 1560
agcgacagaa ccgttacagc ttctatccat gacgtgaatt ccagcacaag gcgactgttg 1620
cttaacgaat ttggatcgga ggctccgcgc ccttcgccag aaaagactgt tcgcttaaga 1680
agcgacaata ttggcaccta tgggagccgg aaaaatgaac gagcgcggt cgcgaccgaa 1740
accggtgcgt atgagtcgga gcatatttct ggggtcaagg ctgtccacga tactgcgaga 1800
gcgacgaaag agggccggcg tctcgaaagg cccatgcccg cctaccttga ggataagggg 1860
cttcacgcc aacatatttg caccgggaga ggacggacca aacttgctcg gcgcggatgg 1920
ccggatgaca caagctatcg ctcgatcaa agggcaactc tgcggaccc cgttcgcgc 1980
tcggaaggcg cgacggcctc aaatgggtat caattgaacc aattgggtta cgcgcaccaa 2040
ctcgctagcg atggcctgca aagtgaatcg cccgatgggt ttgccttgcc aattcaagt 2100
gcaacaacga gctacaacta tacagtgagc cgcgatcctg tccttggtcc gccggataaa 2160
aacgaagccc ctcaattgct gcatcttggt ccccggtggt aaaccgaagc tgttcttgcc 2220
cgcgaaacag cattgactgg aaaatggccg actctcgagc gtgagcagca agtgtatcgc 2280
gagtttttg ccttatatga cgtaaaaaa gatcttgagg ccaaatacgt cggcgtaaga 2340
cggaaaaaaa aagaagttat ttctgcgtta gaccgaactg cgcgcttgat aagcacgtcg 2400
ccttcgaaag ctcgatccaa agcagagact gaaaaagcca ttgatgagct cgatgatcga 2460
cgagtttatg atccgcgtga tcgagctcaa gacaaagcgt ttaaagctg a 2511
<210> 23
<211> 198
<212> ADN
5 <213> Agrobacterium tumefaciens
<400> 23

```

ES 2 843 556 T3

	atggccatca tcaagccgca tgtgaacaaa aataggacaa cctcgccgat agagagaccg	60
	gagtctctca tagaggaaat gagcggcagt catccgccga gtgggttttac caacctggat	120
	ctcgctatga tcgagctgga ggactttgtc catcgggtgcc cgctcccaga agacaatctt	180
	gctggtcaga aggagtga	198
	<210> 24	
	<211> 1650	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 24	
	atggatccgt ctagcaatga gaatgtctat gtgggtcgcg gtcacaacat cgaaaatgat	60
	gatgacactg accccaggcg ttggaagaag gcgaatatca gttccaacac catctccgat	120
	attcagatga cgaatggcga agacgtacaa tcagggagcc ctaccggaac ggaagttgta	180
	agcccacgtc tggattatgg atcggtcgac tcctcctcca gcctttattc tggcagcgag	240
	cacggaaatc aagctgagat tcaaaaagag ctgtccgtct tgttctcgaa catgtctttg	300
	ccaggcaacg atcggcgccc ggacgaatac attctcgtgc atcaaacggg acaagatgct	360
	tttactggta ttgccaaaagg caacctcgac caaatgccca ccaaggcgga atttaacgcg	420
	tgctgccgtc tctacagga cggagccggt aattactacc cgccacctct cgcattcgac	480
	aagattagcg ttccagagca actggaggaa aaatggggga tgatggaggc gaaggaacgt	540
	aacaaactgc ggtttcagta caagttggac gtatggaatc atgcgcacgc tgatatgggg	600
	atcacgggca cagagatctt ttatcaaaca gataagaaca taaagctcga ccggaattat	660
	aaactaagac ctgaagaccg atacgtacaa acagaaaaat acggggcgccg ggaaattcaa	720
	aagcgatatc aacacgaact ccaggctgggt tcgctgctgc ccgatattat gatcaaaact	780
	ccccaaaatg acatccactt cgtgtacagg tttgccggcg acaattacgc caacaaacag	840
	ttcagcgagt ttgaacacac cgtcaagcgc aggtatggcg acgagactga gatcaaattg	900
	aagtcaaagt caggcattat gcatgactcg aaatatctgg aatcctggga acggggcagt	960
	gcggatatct gcttcgcgga attcgttggg gaaaatagag ctcaaatcg gcagtttcca	1020
	actgcgacag taaatatggg acagcagcca gacgggcagg gcggtttgac ccgcgaccgt	1080
	catgtgagcg ttgacttcct aatgcaaagc gcacccaatt cgccttgggc gcaagctttg	1140
	aaaaaggag aactgtggga tcgcgttcag ttgcttgctc gcgacggcaa ccgctatctg	1200
	tcgccgccca gattggaata ttctgaccct gcacatttca ccgagttgat gaaccggggt	1260
	ggtttacccg catcgatggg tcggcaaagc catgcggcta gtatcaaatt cgaaaagttt	1320
	gacgcgcagg cagcggttat tgtcttaaata ggcccagagt tacgtgacat tcatgacttg	1380
	tctcctgaaa aactgcaaaa tttgtccacc aaagatgtca tcgtcgccga tcgcaatgag	1440
	aatggtcaga gaactggcac gtacaccagc gtcgcggaat atgagcgctt gcagttaagg	1500
	ctgccacccg atgcagcggg ggtgcttggg gaagcaactg acaaatatct acgtgatttc	1560
	gttcggccag agccggcgtc gcgtccaatc agtgacagcc gcaggatata cgaaagtcga	1620
	ccgcgtagcc aaagcgtcaa cagcttttga	1650
10	<210> 25	
	<211> 2019	

<212> ADN

<213> Agrobacterium tumefaciens

<400> 25

atggtgaaca ctacaaagaa aagttttgcg aagtcgctta cggcagatat gcgccgttct	60
gctcagcgcg ttgtcgagca aatgcgaaaa gcattgatta ccgaagaaga ggcgctcaag	120
cggcaagcca gactggagag tcccgatagg aagcgaaagt atgctgctga tatggcgata	180
gtcgacaaac tcgacgtagg gtttcgaggg gaaataggct ataaaattct tggaataaac	240
cggcttcgag tagacaacca taaagaatta acgctgagc acggtagact tcgcaaaacc	300
aaaacgggtc tgaagcgtaa cccggtgacg caggaagtct atttgggttt atatgaaagg	360
aagtcctggt taagtgtcag cagccatttg tatgctgcgg acggcacact ccgcatgaag	420
cacgtgaaat acaaagacgg acgttttgag gaaaaatggg agcgcgacga aaatggcgac	480
ctgatccgca caaggtacgc caaccgtggc aggtctcttc aacctgtatc cgagaaaatg	540
ggcgcgccgt atcggagcgg ccctgacgac cggctctatc gcgatctaac ccgtcgaaac	600
ggtttcagac gggagacatt cgaacgggac gatcacggaa acctcgagcg tatcggcagc	660
aaccatgtcg gcttttccaa gatttcagtg aaggcaccga atcgtcaaac ctcccagacg	720
aagattcaaa aacttggtgg cgctttcaac aaatctttta ggtcccttct ggacaaggag	780
ggcaatgaaa tgggcccgca tattttgagc catcgacggc tctataacaa gcggtctgct	840
gtctacgatg aagctaccgg acaattgaag agtgccaagc ataccttcgg caagatctac	900
aggagcgaaa ccgaatatct cagcgcgggc ctcaagaagg ttcaaaaaa gatactcggg	960
gtgacggtct accggaaatt tgcggcgctc agcgagcgag aatccgaggc tgagagactg	1020
cgtagttttg aatccggtgc gcatcgccag atctggcagg agcgggcagc gattcccggg	1080
tcgcccctcc cgggagactga tgacattcat ttgcgacagc agtcgcacct agccaaagcc	1140
aacctgatc acgtcgaagc tgacgtcacg cgtgtgacag atcaacatgt tgatgttgct	1200
ggacaaacat catcgtctcc ccaacggaac ttggaaggat ggtagattc tcaatcacga	1260
tacaagccag caaacatgct gttgtcaaat ccagaccttc aagcgaacgg acctcgccca	1320
tacgaagggg tagctcatct caccctccgg cgcgataatg aatctgacgg gcacaaggag	1380
aatgatcagc ggctgcgaca tttctcccag ccagagccgt tgggtgttacc gcatcccggg	1440
tcgccggaaa taactaaggt gtttggtctg cggggagagc cggcacaccc gagtggaaat	1500
ctgcacacgg cggttggaga aacggcttgc gaaggaccgg tgatgtcttc atcctcggac	1560
aatcatcagc cagctccagg acagcaagaa cttttaagtt tccttcataa tgcgccagcc	1620
ccagtttctg tggcaatata tgatgatcaa gacgacttg cgggggaggc gcccgggggc	1680
tctttcagag gtagctcagg gcgaacgagt tcaatgtcgg agagtatctt cgacgaagat	1740
gtacaagggc atttggtagc ggattattcg atcaatacta ctaacggggt tattgaccgc	1800
caatcgttgt tcggtgaacc ggacttatcg agaggtccaa aatcggggcc agaaattcca	1860
tcggaagatt accatttgtc agcttcggaa caggaaaatt tgctgaatca attgcttagt	1920
gtgccactgc cggttccttc accgaagccc gaatgcgcga ggtctatgat tttcgaagg	1980
tcacgttcaa gagagcgttc cacctccaga gggttctaa	2019

5

<210> 26
 <211> 2490
 <212> ADN
 <213> Agrobacterium tumefaciens
 5 <400> 26
 atgaacggaa ggtattcacc gacgcggcag gatTTTtaaga caggcgcgaa gccttggTct 60
 atattagccc ttatcgttgc cgcaatgatt ttCgcgttca tggcggTtgc gTcctggcag 120
 gacaatggga ctaccCaggc aatCcttagc caaatacgat cgattaacgc tgacagcgcc 180
 tcactgcagc gcgatgtact ccgcgctcac acgggcaccg tggcgaaCTa ccgccccatt 240
 atctccaggc tgggagctct gcggaagaat ctggaagatt tgaagcaatt atttaggcaa 300
 tctcatattg taagtgagag caatgctgct caactgctac gccggctaga agtgtctcta 360
 aattcggctg acgcggcggt cgccgccttt ggtgcgcaaa atgtccgcct gcaagattcg 420
 ctggccagtt tCactcgtgc tTTaagcaat ctccaggaa aggcctcagg cgatcagact 480
 ttagaaaaac caacagaatt ggctagcatg atgctccaat ttcttcggca accaagcccg 540
 gctatttcat tcgagatcag ccttgaacta gagaggctcc aaaaacaacg cggTcttgat 600
 gaagctcccg tgcgcatact ttCgcgtgaa ggtccCatta tCttatcgct tttgccacag 660
 gtgaacgatc tggTgaacat gattcagacg tctgacaccg cagaaattgc ggaaatgttg 720
 cagcgtgagt gtttggaggt ctatagcttg aaaaatgtgg aggagcggag cgcacgtatc 780
 tttcttgggt ccgcttcagt gggTctttgc ctgtacatca tcaccttagt ctataggcta 840
 cgcaaaaaaa ccgattgggt agcgcggcgt ttagattacg aagagctaata caaagagatc 900

ES 2 843 556 T3

```

ggagtatgtt ttgaaggtga ggcggccact acgtcgtccg cgcaagctgc acttgggtatt      960
attcagcgct tctttgatgc cgatacgtgc gcgttagctc tagtggacca tgaccgtagg      1020
tgggctgtcg aaacattcgg tgcgaagcac ccaaaacccg tgtgggacga cagggtgcta      1080
cgcgaaatag tctctcgtac caaagcgaac gaacgggcca cggtattccg catcgtatcg      1140
acgcaaaaaa tcgtacattt gcctcccga attccaggtc tttcgatact actggctcac      1200
aatccacag ataaactaat tgcggtttgt tcaactgggtt accaaagcta tcgccctcga      1260
ccttgccaag gcgaaattca gcttcttgaa ctccgccaccg cctgcctctg tcaatatatc      1320
gatgttcggc gtaagcagac cgaatgcgac gttttggcca gacgattgga gcatgcgcaa      1380
cgccttgagg cagttggtac acttgccggc ggaatagcac atgaatttaa taacattttg      1440
gggtcaatcc tcgggcacgc agaattggcc caaaactcgg tgtctcgaac atctgtcacc      1500
cgaagataca ttgactatat catttcgtca ggcgacagag ccatgctcat tatcgatcag      1560
atcttgacgc tgagccgaaa acaagagcgc gtgatcaagc catttagtgt ctgagagctt      1620
gtgaccgaaa tcgctccctt gctacgtatg gctcttcccc caaccatcga gcttagtttc      1680
agatttgatc aaatgcagag cgtgatcgaa ggaagcccgc ttgaacttca acaggtacta      1740
attaacctct gcaagaatgc ttcccaagcc atgactgcaa atgggtcaa atcgacatcttc      1800
gtcggccaag cttattttacc agctaagaaa attctggcgc atgggtgttat gccacctggc      1860
gactatgttc tactatctgt cagcgacaat ggtggaggca tttcagaggc tgtgctaccc      1920
tacatttttg aaccttctt tacgacacgt gctcgaacg gtggaacggg tctcggcctc      1980
gcttctgtgc atggtcatat cagcgcgttt gcaggttaca tcgacgttag ttcaactgtt      2040
gggcatggga cgcgctttga catttatctc cctccgtctt ctaaggagcc cgtcaatccc      2100
gacagctttt tcggccgcaa taaggcaccg cgtggaaacg gggagattgt ggcacttggt      2160
gagcccgatg acctcctgcg ggaggcgtat gaagacaaga tcgccgctct gggatatgag      2220
ccggtcgggt ttcgtacctt tagtgaaatt cgcgattgga tttcaaaagg caatgaagcc      2280
gatctggtca tggtcgacca agcgtctctt cctgaagatc aaagtcctaa ttccgtggat      2340
ttagtgctca agaccgcctc catcatcatt ggcggaaatg atctcaaaat gcccttttca      2400
agggagaatg cgaccagga ctttatctg ccgaagccga tatcgtccag aactatggcg      2460
catgcaatcc taaccaaaat caagacgtag      2490
<210> 27
<211> 744
<212> ADN
5 <213> Agrobacterium tumefaciens
<400> 27
atggcgataa aattgggtatt gatactcgta ttcacactct ttcccgggc agacgctgca      60

```

tatgcgaatg accgcgccaa cggtttcatg tggtaaaccg ggggcgaaac tggagtgagg	120
cttcctcttc ggggtgttcaa tgccaagcca gccagaaca cggtgccgat catttattcc	180
ggagacgctg gatggcaaaa tatcgatgag gcgattggta cctatctgca gacggaaggg	240
attcctgtca ttggcgctcag ttacttccg tatttctggc cggagcggc cccaagcgaa	300
actgctaagg atcttgggtca cataatcgat gtctacacca agcatttccg tgtgcagaat	360
gttttacttg taggatattc tttcggcgcg gacgtcatgc cggcaagctt caataggctt	420
acgcttgagc aaaaaatcg ggttaagcaa atctctctct tggcattgtc acatcaagtc	480
gactatgtcg tctcatttag gggctggctc caactcgaaa cggaaggtaa gggcggcaat	540
cctctggatg atctcagatc cattgaccct gcaatcgctc aatgcagtga cggcgcgaa	600
gaccgtaata atgcttgccc atctctccga cagaccggcg cagaggtgat aggttcagc	660
ggaggccatc actttgataa tgatttcaaa aaactgtcta cgcgcgtcgt ctcaggcctc	720
gtggcacgcc taagtcatca gtaa	744

<210> 28

<211> 7743

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 28

gtttaccgc caatatatcc tgtcaaacc tcatagttta aactgaaggc gggaaacgac	60
aactctgatca tgagcggaga attaaggag tcaagttatg acccccgccg atgacgggg	120
acaagccgtt ttacgttttg aactgacaga accgcaacgt tgaaggagcc actcagcaag	180
ctgggtacgat tgtaatacga ctactatag gggaattga ggcgtgttta aacgctcttc	240
aactggaaga gcggttacgc tgtttaaacg ctcttcaact ggaagagcgg ttactaccgg	300
ttacttagct agctgctaag gttaccagag ctggtcacct ttgtccacca acttattaag	360
tatctagttg aagacacgtt cttcttcacg taagaagaca ctcatagtc ttcggccaga	420
atggcctctt gattcagcgg gcctagaagg ccgcatcact gactagctaa tttaaatcct	480
gaggatatcg ctatcaactt tgtatagaaa agttgggccc aattcgagct cggtagggcc	540
agaatggccc ggaccgggtt acccgaccg aagcttgcct gcctgcagtg cagcgtgacc	600
cggcgtgccc cctctctaga gataatgagc attgcatgtc taagttataa aaaattacca	660
catatTTTTT ttgtcacact tgtttgaagt gcagtttatc tatctttata catatattta	720
aactttactc tacgaataat ataactata gtactacaat aatatcagtg ttttagagaa	780
tcatataaat gaacagttag acatggtcta aaggacaatt gagtatTTTg acaacaggac	840
tctacagttt tatctTTTTa gtgtgcatgt gttctccttt ttttttgcaa atagcttcac	900

10

ES 2 843 556 T3

ctatataata cttcatccat tttatttagta catccattta gggtttaggg ttaatggttt	960
ttatagacta attttttttag tacatctatt ttattctatt ttagcctcta aattaagaaa	1020
actaaaactc tatttttagtt tttttattta ataatttaga tataaaatag aataaaataa	1080
agtgactaaa aattaacaac atacccttta agaaattaaa aaaactaagg aaacattttt	1140
cttgtttcga gtagataatg ccagcctgtt aaacgccgtc gacgagtcta acggacacca	1200
accagcgaac cagcagcgtc gcgtcgggcc aagcgaagca gacggcacgg catctctgtc	1260
gctgcctctg gaccctctc gagagttccg ctccaccgtt ggacttgctc cgctgtcggc	1320
atccagaaat tgcgtggcgg agcggcagac gtgagccggc acggcaggcg gcctcctcct	1380
cctctcacgg caccggcagc tacgggggat tcctttccca ccgctcctc gctttccctt	1440
cctgcgccgc cgtaataaat agacaccccc tccacacctt ctttcccaa cctcgtgttg	1500
ttcggagcgc acacacacac aaccagatct cccccaatc caccgcgcg cacctccgct	1560
tcaaggtagc ccgctcgtcc tcccccccc ccctctctac cttctctaga tcggcggtcc	1620
ggtccatgca tggttagggc ccggtagttc tacttctgtt catgtttgtg ttagatccgt	1680
gtttgtgtta gatccgtgct gtagcgttc gtacacggat gcgacctgta cgtcagacac	1740
gttctgattg ctaacttgcc agtgtttctc tttggggaat cctgggatgg ctctagccgt	1800
tccgcagacg ggatcgattt catgattttt tttgtttcgt tgcatagggt ttggtttgcc	1860
cttttccttt atttcaatat atgccgtgca cttgtttgtc gggcatcctt ttcattgctt	1920
tttttgtcct ggttgatgat atgtggtctg gttgggcggc cgttctagat cggagtagaa	1980
ttctgtttca aactacctgg tggatttatt aattttggat ctgtatgtgt gtgccataca	2040
tattcatagt tacgaattga agatgatgga tggaaatata gatctaggat aggtatacat	2100
gttgatgcgg gttttactga tgcataata gagatgcttt ttgttcgctt ggttgatgat	2160
atgtggtgtg gttgggcggc cgttcattcg ttctagatcg gagtagaata ctgtttcaa	2220
ctacctggtg tatttattaa ttttggaact gtatgtgtgt gtcataatac ttcattagta	2280
cgagtttaag atggatggaa atatcgatct aggataggta tacatgttga tgtgggtttt	2340
actgatgcat atacatgatg gcatatgcag catctattca tatgctctaa ccttgagtac	2400
ctatctatta taataaaca gtatgtttta taattatttt gatcttgata tacttggtg	2460
atggcatatg cagcagctat atgtggattt ttttagccct gccttcatac gctatttatt	2520
tgcttggtac tgtttctttt gtcgatgctc accctgttgt ttggtgttac ttctgcagg	2580
cgactctaga ggatccaccg gtcgccacca tggccacag caagcacggc ctgaaggagg	2640
agatgaccat gaagtaccac atggagggct gcgtgaacgg ccacaagtgc gtgatcaccg	2700
gcgagggcat cggctacccc ttcaagggca agcagaccat caacctgtgc gtgatcgagg	2760

ES 2 843 556 T3

gcggccccct gcccttcagc gaggacatcc tgagcgccgg cttcaagtac ggcgaccgga	2820
tcttcaccga gtacccccag gacatcgtgg actacttcaa gaacagctgc cccgccggct	2880
acacctgggg ccggagcttc ctgttcgagg acggcgccgt gtgcatctgt aacgtggaca	2940
tcaccgtgag cgtgaaggag aactgcatct accacaagag catcttcaac ggcgtgaact	3000
tccccgccga cggccccgtg atgaagaaga tgaccaccaa ctgggaggcc agctgcgaga	3060
agatcatgcc cgtgcctaag cagggcatcc tgaaggcgga cgtgagcatg tacctgctgc	3120
tgaaggacgg cggccgggtac cgggtgccagt tgcacaccgt gtacaaggcc aagagcgtgc	3180
ccagcaagat gcccgagtgg cacttcatcc agcacaagct gctgcgggag gaccggagcg	3240
acgccaaaga ccagaagtgg cagctgaccg agcagcccat cgccttcccc agcgccctgg	3300
cctgaagcgg ccgcaacctc gacttgtcca tcttctggat tggccaactt aattaatgta	3360
tgaataaaaa ggatgcacac atagtgcacat gctaataact ataagtggg catcaaagtt	3420
gtgtgttatg tgtaattact agttatctga ataaaagaga aagagatcat ccatatttct	3480
tatcctaaat gaatgtcacg tgtctttata attctttgat gaaccagatg catttcatta	3540
accaaatacca tatacatata aatattaatc atatataatt aatatcaatt gggttagcaa	3600
aacaaatcta gtctaggtgt gttttgcgaa tgcggccgcc accgcggtagg agctcgaatt	3660
ccgggtccggg tcacccggtc cgggcctaga aggcgatct cccgggcacc cagctttctt	3720
gtacaaagtg gccgttaacg gatcgcccg aatggcccg accgggttac ccggaccgaa	3780
gcttgcattg ctgcagtgc gctgacccg gtcgtgcccc tctctagaga taatgagcat	3840
tgcatgtcta agttataaaa aattaccaca tatttttttt gtcacacttg tttgaagtgc	3900
agtttatcta tctttataca tataattaaa ctttactcta cgaataatat aatctatagt	3960
actacaataa tatcagtgtt ttagagaatc atataaatga acagttagac atgggtctaaa	4020
ggacaattga gtattttgac aacaggactc tacagtttta tcttttttagt gtgcatgtgt	4080
tctccttttt ttttgcaaat agcttcacct atataatact tcatccattt tattagtaca	4140
tccatttagg gtttaggggt aatgggtttt atagactaat ttttttagta catctatttt	4200
attctatttt agcctctaaa ttaagaaaac taaaactcta ttttagtttt tttatttaat	4260
aatttagata taaaatagaa taaaataaag tgactaaaaa ttaaacaat accctttaag	4320
aaattaataa aactaaggaa acatttttct tgtttcgagt agataatgcc agcctgttaa	4380
acgccgtcga cgagtctaac ggacaccaac cagcgaacca gcagcgtgc gtccggccaa	4440
gcgaagcaga cggcacggca tctctgtcgc tgcctctgga cccctctcga gagttccgct	4500
ccaccgttgg acttgctccg ctgtcggcat ccagaaattg cgtggcggag cggcagacgt	4560
gagccggcac ggcaggcggc ctctcctcc tctcacggca ccggcagcta cgggggattc	4620
ctttccacc gtccttcgc tttcccttc tgcgccgcg taataaatag acacccctc	4680

ES 2 843 556 T3

cacaccctct tcccccaacc tcgtgttggt cggagcgcac acacacacaa ccagatctcc	4740
cccaaatacca cccgtcggca cctccgcttc aaggtaacgc gctcgtcctc ccccccccc	4800
ctctctacct tctctagatc ggcgttccg tccatgcatg gttaggggccc ggtagttcta	4860
cttctgttca tgtttgtgtt agatccgtgt ttgtgttaga tccgtgctgc tagcgttcgt	4920
acacgggatgc gacctgtacg tcagacacgt tctgattgct aacttgccag tgtttctctt	4980
tggggaatcc tgggatggct ctagccgttc cgcagacggg atcgatttca tgattttttt	5040
tgtttcgttg catagggttt ggtttgcctt tttcctttat ttcaatatat gccgtgcact	5100
tgtttgtcgg gtcactcttt catgcttttt tttgtcttgg ttgtgatgat gtggtctggt	5160
tgggcggtcg ttctagatcg gagtagaatt ctgtttcaaa ctacctggtg gatttattaa	5220
ttttggatct gtatgtgtgt gccatacata ttcatagtta cgaattgaag atgatggatg	5280
gaaatatcga tctaggatag gtatacatgt tgatgcgggt tttactgatg catatacaga	5340
gatgcttttt gttcgcttgg ttgtgatgat gtggtgtggt tgggcggtcg ttcattcggt	5400
ctagatcgga gtagaatact gtttcaaact acctggtgta tttattaatt ttggaactgt	5460
atgtgtgtgt catacatctt catagttacg agtttaagat ggatggaaat atcgatctag	5520
gatagggtata catgttgatg tgggtttttac tgatgcatat acatgatggc atatgcagca	5580
tctattcata tgctctaacc ttgagtacct atctattata ataaacaagt atgttttata	5640
attattttga tcttgatata cttggatgat ggcatatgca gcagctatat gtggattttt	5700
ttagccctgc cttcatacgc tattttatttg cttggtactg tttcttttgt cgatgctcac	5760
cctgttgttt ggtgttactt ctgcaggtcg actttaactt agcctaggat ccatgcaaaa	5820
actcattaac tcagtgcaaa actatgcctg gggcagcaaa acggcggtga ctgaacttta	5880
tggatgga aatccgtcca gccagccgat ggccgagctg tggatggcg cacaacgaa	5940
aagcagttca cgagtgcaga atgccgccg agatatcgtt tcaactgctg atgtgattga	6000
gagtataaaa tcgactctgc tcggagaggc cgttgccaaa cgctttggcg aactgccttt	6060
cctgttcaaa gtattatgcg cagcacagcc actctccatt caggttcac caaacaaaca	6120
caattctgaa atcggttttg ccaaagaaaa tgccgcaggt atcccgatgg atgccgccga	6180
gcgtaactat aaagatccta accacaagcc ggagctgggt tttgcgctga cgcctttcct	6240
tgcatgaac gcgtttcgtg aattttccga gattgtctcc ctactccagc cggtcgcagg	6300
tgcacatccg gcgattgctc actttttaca acagcctgat gccgaacgtt taagcgaact	6360
gttcgccagc ctgttgaata tgcagggtga agaaaaatcc cgcgcgctgg cgattttaaa	6420
atcgccctc gatagccagc aggtgaacc gtggcaaacg attcgtttaa tttctgaatt	6480
ttacccgga gacagcggtc tgttctcccc gctattgctg aatgtggtga aattgaacc	6540

ES 2 843 556 T3

tggcgaagcg atgttcctgt tcgctgaaac accgcacgct tacctgcaag gcgtggcgct	6600
ggaagtgatg gcaaaactccg ataactgtgt gcgtgcggtg ctgacgccta aatacattga	6660
tattccggaa ctggttgcca atgtgaaatt cgaagccaaa ccggctaacc agttgttgac	6720
ccagccgggtg aaacaagggtg cagaactgga cttcccgatt ccagtggatg attttgcctt	6780
ctcgctgcat gaccttagtg ataaagaaac caccattagc cagcagagtg ccgccatttt	6840
gttctgctgc gaaggcgatg caacgttgtg gaaagggttct cagcagttac agcttaaacc	6900
gggtgaatca gcgtttattg ccgccaacga atcacccggtg actgtcaaag gccacggcgcg	6960
tttagcgcgc gtttacaaca agctgtaaat cactagcgaa cgcgtaggta ccacatgggt	7020
aacctagact tgtccatctt ctggattggc caacttaatt aatgtatgaa ataaaaggat	7080
gcacacatag tgacatgcta atcactataa tgtgggcatc aaagttgtgt gttatgtgta	7140
attactagtt atctgaataa aagagaaaga gatcatccat atttcttata ctaaatgaat	7200
gtcacgtgtc tttataattc tttgatgaac cagatgcatt tcattaacca aatccatata	7260
catataaata ttaatcatat ataattaata tcaattgggt tagcaaaaaca aatctagtct	7320
aggtgtgttt tgogaatgcg gccgccaccg cgggtggagct cgaattccgg tccgggtcac	7380
ccggtccggg cctagaaggc cagcttcggc cgccccgggc aactttatta taaaaagttg	7440
atagatatct ggtctaacta actagtccta aggacccggc ggaccgatta aactgattcg	7500
gtccgaagct taagccatgg cccgggaatc ttagcggccg cctgcagagt taacggcgcg	7560
ccgactagct agctaaggta ccgagctoga attcattccg attaatcgtg gcctcttgct	7620
cttcaggatg aagagctatg tttaaacgtg caagcgctac tagacaattc agtacattaa	7680
aaacgtccgc aatgtgttat taagttgtct aagcgtcaat ttgtttacac cacaatatat	7740
cct	7743

<210> 29

<211> 6496

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 29

gtgcagcgtg acccggtcgt gccctctct agagataatg agcattgcat gtctaagtta	60
taaaaaatta ccacatattt tttttgtcac acttgtttga agtgcagttt atctatcttt	120
atacatatat ttaaaactta ctctacgaat aatataatct atagtactac aataatatca	180
gtgttttaga gaatcatata aatgaacagt tagacatggc cttaaaggaca attgagtatt	240
ttgacaacag gactctacag ttttatcttt ttagtgtgca tgtgttctcc tttttttttg	300
caaatagctt cacctatata atacttcata catthttatta gtacatccat ttagggttta	360

10

ES 2 843 556 T3

gggttaatgg tttttataga ctaatTTTTT tagtacatct attttattct attttagcct	420
ctaaattaag aaaactaaaa ctctatttta gtttttttat ttaataattt agatataaaa	480
tagaataaaa taaagtgact aaaaattaaa caaataccct ttaagaaatt aaaaaaacta	540
aggaaacatt tttcttggtt cgagtagata atgccagcct gttaaagcc gtcgacgagt	600
ctaacggaca ccaaccagcg aaccagcagc gtcgctcgg gccaaagcaa gcagacggca	660
cggcatctct gtcgctgcct ctggaccct ctcgagagtt ccgctccacc gttggacttg	720
ctccgctgtc ggcatccaga aattgctgg cgagcggca gacgtgagcc ggcacggcag	780
gcggcctcct cctcctctca cggcaccggc agctacggg gattccttcc ccaccgctcc	840
ttcgttttcc cttcctcgcc cgccgtaata aatagacacc ccctccacac cctctttccc	900
caacctcgtt ttgttcggag cgcacacaca cacaaccaga tctcccccac atccaccgt	960
cggcacctcc gcttcaagg acgcgctcg tcctcccccc cccccctctc taccttctct	1020
agatcggcgt tccggtccat gcatggtag ggcccggtag ttctacttct gttcatgttt	1080
gtgttagatc cgtgtttgtg ttagatccgt gctgctagcg ttctacacg gatgcgacct	1140
gtacgtcaga caggttctga ttgctaactt gccagtggtt ctctttgggg aatcctggga	1200
tggctctagc cgttccgcag acgggatcga tttcatgatt tttttgttt cgttgcatag	1260
ggtttggttt gcccttttcc tttatttcaa tatatgccgt gcacttggtt gtcgggtcat	1320
cttttcatgc tttttttgt cttggttggt atgatgtggt ctggttgggc ggtcgttcta	1380
gatcggagta gaattctgtt tcaaactacc tgggtgattt attaatTTTg gatctgtatg	1440
tgtgtgccat acatattcat agttacgaat tgaagatgat ggatggaaat atcgatctag	1500
gataggtata catgttgatg cgggttttac tgatgcatac acagagatgc tttttgttcg	1560
cttggttggt atgatgtggt gtggttgggc ggtcgttcat tcgttctaga tcggagtaga	1620
atactgtttc aaactacctg gtgtatttat taattttgga actgtatgtg tgtgtcatac	1680
atcttcatag ttacgagttt aagatggatg gaaatatcga tctaggatag gtatacatgt	1740
tgatgtgggt tttactgatg catatacatg atggcatatg cagcatctat tcatatgctc	1800
taaccttgag tacctatcta ttataataaa caagtatgtt ttataattat tttgatcttg	1860
atatacttgg atgatggcat atgcagcagc tatatgtgga tttttttagc cctgccttca	1920
tacgctattt atttgcttg tactgtttct tttgtcagtg ctcaccctgt tgtttggtgt	1980
tacttctgca ggtcgacttt aacttagcct agcgaagttc ctattccgaa gttcctattc	2040
tctagaaagt ataggaactt cagatccacc ggctagagga tccaccatgg ttgaacaaga	2100
tggattgcac gcaggttctc cgcccgcttg ggtggagagg ctattcggct atgactgggc	2160
acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc	2220

ES 2 843 556 T3

ggttcttttt	gtcaagaccg	acctgtccg	tgccctgaat	gaactgcagg	acgaggcagc	2280
gcggctatcg	tggctggcca	cgacgggct	tccttgcgca	gctgtgctcg	acgttgtcac	2340
tgaagcggga	agggactggc	tgctattggg	cgaagtgccg	gggcaggatc	tcctgtcatc	2400
tcaccttgct	cctgccgaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	2460
gcttgatccg	gctacctgcc	cattcgacca	ccaagcga	catcgcatcg	agcgagcacg	2520
tactcggatg	gaagccggtc	ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	2580
cgcgccagcc	gaactgttcg	ccaggtccaa	ggcgcgcatg	cccgcggcg	atgatctcgt	2640
cgtgacccat	ggcgatgcct	gcttgccgaa	tatcatgggtg	gaaaatggcc	gcttttctgg	2700
attcatcgac	tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cgttggctac	2760
ccgtgatatt	gctgaagagc	ttggcggcga	atgggtgac	cgcttcctcg	tgctttacgg	2820
tatcgccgct	cccgattcgc	agcgcatcgc	cttctatcgc	cttcttgacg	agttcttctg	2880
aggatccacc	atggttaacc	tagacttgct	catcttctgg	attggccaac	ttaattaatg	2940
tatgaaataa	aaggatgcac	acatagtgc	atgctaata	ctataatgtg	ggcatcaaag	3000
ttgtgtgtta	tgtgtaatta	ctagtatatc	gaataaaaga	gaaagagatc	atccatattt	3060
cttatcctaa	atgaatgtca	cgtgtcttta	taattctttg	atgaaccaga	tgcatttcat	3120
taaccaaatac	catatacata	taaatattaa	tcatatataa	ttaatatcaa	ttgggttagc	3180
aaaacaaatac	tagtctaggt	gtgttttgcg	aattgcggcc	gggtaccgag	ctcgaattcg	3240
gccaagttt	gtacaaaaaa	gcaggctccg	gccagaatgg	cccggaccga	agcttgcatg	3300
cctgcagtgc	agcgtgaccc	ggcgtgccc	ctctctagag	ataatgagca	ttgcatgtct	3360
aagttataaa	aaattaccac	atattttttt	tgtcacactt	gtttgaagtg	cagtttatct	3420
atctttatac	atatatttaa	actttactct	acgaataata	taatctatag	tactacaata	3480
atatcagtgt	tttagagaat	catataaatg	aacagttaga	catggtctaa	aggacaattg	3540
agtattttga	caacaggact	ctacagtttt	atcttttttag	tgtgcatgtg	ttctcctttt	3600
tttttgcaaa	tagcttcacc	tatataatac	ttcatccatt	ttattagtac	atccatttag	3660
ggtttaggg	taatggtttt	tatagactaa	tttttttagt	acatctattt	tattctattt	3720
tagcctctaa	attaagaaaa	ctaaaactct	attttagttt	ttttatttaa	taatttagat	3780
ataaaataga	ataaaataaa	gtgactaaaa	attaaacaaa	taccctttaa	gaaattaaaa	3840
aaactaagga	aacatttttc	ttgtttcgag	tagataatgc	cagcctgtta	aacgccgtcg	3900
acgagtctaa	cggacaccaa	ccagcgaacc	agcagcgtcg	cgtcgggcca	agcgaagcag	3960
acggcacggc	atctctgtcg	ctgcctctgg	accctctctg	agagttccgc	tccaccgttg	4020
gacttgctcc	gctgtcggca	tccagaaatt	gcgtggcgga	gcggcagacg	tgagccggca	4080
cggcaggcgg	cctcctcctc	ctctcacggc	accggcagct	acgggggatt	cctttccac	4140

ES 2 843 556 T3

cgctccttcg ctttcccttc ctgcgccgcc gtaataaata gacaccccct ccacaccctc	4200
tttccccaac ctcgtgttgt tcggagcgca cacacacaca accagatctc ccccaaatcc	4260
accggtcggc acctccgctt caaggtagcg cgctcgtcct cccccccccc cctctctacc	4320
ttctctagat cggcgttccg gtccatgcat ggtagggcc cggtagttct acttctgttc	4380
atgttttgt tagatccgtg tttgtgttag atccgtgctg ctacggttcg tacacggatg	4440
cgacctgtac gtcagacacg ttctgattgc taacttgcca gtgtttctct ttggggaatc	4500
ctgggatggc tctagccgtt ccgcagacgg gatcgatttc atgatttttt ttgtttcgtt	4560
gcataggggtt tgggttgccc ttttccttta tttcaatata tgccgtgcac ttgtttgtcg	4620
ggtcactctt tcatgctttt ttttgtcttg gttgtgatga tgtggtctgg ttgggcggtc	4680
gttctagatc ggagtagaat tctgtttcaa actacctggg ggatttatta attttgatc	4740
tgtatgtgtg tgccatacat attcatagtt acgaattgaa gatgatggat ggaaatatcg	4800
atctaggata ggtatacatg ttgatgcggg ttttactgat gcatatacag agatgctttt	4860
tgttcgcttg gttgtgatga tgtggtgtgg ttgggcggtc gttcattcgt tctagatcgg	4920
agtagaatac tgtttcaaac tacctgggtg atttattaat tttggaactg tatgtgtgtg	4980
tcatacatct tcatagttac gagtttaaga tggatggaaa tatcgatcta ggataggtat	5040
acatgttgat gtgggtttta ctgatgcata tacatgatgg catatgcagc atctattcat	5100
atgctctaac cttgagtacc tatctattat aataaacaag tatgttttat aattattttg	5160
atcttgatat acttggatga tggcatatgc agcagctata tgtggatttt tttagccctg	5220
ccttcatacg ctatttatct gcttggtact gtttcttttg tcgatgctca ccctgttgtt	5280
tgggtgtact tctgcaggtc gactctagag gatccaccgg tcgccaccat ggccctgtcc	5340
aacaagttca tcggcgacga catgaagatg acctaccaca tggacggctg cgtgaacggc	5400
cactacttca ccgtgaaggc cgagggcagc ggcaagccct acgagggcac ccagacctcc	5460
accttcaagg tgacctggc caacggcggc cccctggcct tctccttcga catcctgtcc	5520
accgtgttca tgtacggcaa ccgctgcttc accgcctacc ccaccagcat gcccgactac	5580
ttcaagcagg ccttccccga cggcatgtcc tacgagagaa ccttcaccta cgaggacggc	5640
ggcgtggcca ccgccagctg ggagatcagc ctgaaggga actgcttcga gcacaagtcc	5700
accttccacg gcgtgaactt ccccgccgac ggccccgtga tggccaagaa gaccaccggc	5760
tgggaccct ccttcgagaa gatgaccgtg tgcgacggca tcttgaaggc cgacgtgacc	5820
gccttcctga tgctgcaggc cggcggcaac tacagatgcc agttccacac ctctacaag	5880
accaagaagc ccgtgacct gcccccaac cacgtggtgg agcaccgcat cgccagaacc	5940
gacctggaca agggcggaac cagcgtgcag ctgaccgagc acgccgtggc ccacatcacc	6000

ES 2 843 556 T3

tccgtggtgc ccttctgaag cggccgcaac ctagacttgt ccatcttctg gattggccaa	6060
cttaattaat gtatgaaata aaaggatgca cacatagtga catgctaata actataatgt	6120
gggcatcaaa gttgtgtgtt atgtgtaatt actagttatc tgaataaaag agaaagagat	6180
catccatatt tcttataccta aatgaatgtc acgtgtcttt ataattcttt gatgaaccag	6240
atgcatttca ttaaccaaact ccatatacat ataaatatta atcatatata attaatatca	6300
attgggttag caaaacaaat ctagtctagg tgtgttttgc gaatgcggcc gccaccgagg	6360
tggagctcga attccggtcc gggcctagaa ggccgatctc ccgggcaccc agctttcttg	6420
tacaaagtgg ccgttaacgg atcccggtga agttcctatt ccgaagttcc tattctccag	6480
aaagtatagg aacttc	6496

<210> 30

<211> 1991

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 30

gtgcagcgtg acccggtcgt gccctctct agagataatg agcattgcat gtctaagtta	60
taaaaaatta ccacatattt tttttgtcac acttgtttga agtgcagttt atctatcttt	120
atacatatat ttaaacttta ctctacgaat aatataatct atagtactac aataatatca	180
gtgttttaga gaatcatata aatgaacagt tagacatggt cttaaaggaca attgagtatt	240
ttgacaacag gactctacag ttttatcttt ttagtgtgca tgtgttctcc tttttttttg	300
caaataagctt cacctatata atacttcatc cattttatta gtacatccat ttaggggtta	360
gggttaatgg tttttataga ctaatttttt tagtacatct attttattct attttagcct	420
ctaaattaag aaaactaaaa ctctatttta gtttttttat ttaataattt agatataaaa	480
tagaataaaa taaagtgact aaaaattaaa caaataccct ttaagaaatt aaaaaaacta	540
aggaaacatt tttcttgttt cgagtagata atgccagcct gttaaagcgc gtcgacgagt	600
ctaacggaca ccaaccagcg aaccagcagc gtcgctcgg gccaaagcga gcagacggca	660
cggcatctct gtcgctgcct ctggaccct ctcgagagtt ccgctccacc gttggacttg	720
ctccgctgtc ggcatccaga aattgcgtgg cggagcggca gacgtgagcc ggcacggcag	780
gcggcctcct cctcctctca cggcaccggc agctacgggg gattcctttc ccaccgctcc	840
ttcgttttcc cttcctcgcc cgccgtaata aatagacacc cctccacac cctctttccc	900
caacctcgtg ttgttcggag cgcacacaca cacaaccaga tctcccccga atccaccgt	960
cggcacctcc gcttcaaggt acgcccgtcg tcctccccc ccccccctctc taccttctct	1020
agatcggcgt tccggtccat gcatggttag ggcccggtag ttctacttct gttcatgttt	1080

10

ES 2 843 556 T3

	gtgttagatc cgtgtttgtg ttagatccgt gctgctagcg ttcgtacacg gatgcgacct	1140
	gtacgtcaga cacgttctga ttgctaactt gccagtgttt ctctttgggg aatcctggga	1200
	tggctctagc cgttccgcag acgggatcga tttcatgatt ttttttgttt cgttgcatag	1260
	ggtttggttt gcccttttcc tttattttcaa tatatgccgt gcacttgttt gtcgggtcat	1320
	cttttcatgc ttttttttgt cttggttgtg atgatgtggt ctggttgggc ggtcgttcta	1380
	gacgcgagta gaattctgtt tcaaaactacc tgggtggattt attaattttg gatctgtatg	1440
	tgtgtgccat acatattcat agttacgaat tgaagatgat ggatggaaat atcgatctag	1500
	gatagggtata catgttgatg cgggtttttac tgatgcatat acagagatgc tttttgttcg	1560
	cttggttgtg atgatgtggt gtggttgggc ggtcgttcat tcgttctaga tcggagtaga	1620
	atactgtttc aaactacctg gtgtatttat taattttgga actgtatgtg tgtgtcatac	1680
	atcttcatag ttacgagttt aagatggatg gaaatatcga tctaggatag gtatacatgt	1740
	tgatgtgggt tttactgatg catatacatg atggcatatg cagcatctat tcatatgctc	1800
	taaccttgag tacctatcta ttataataaa caagtatgtt ttataattat tttgatcttg	1860
	atatacttgg atgatggcat atgcagcagc tatatgtgga ttttttttagc cctgccttca	1920
	tacgtatatt atttgcttgg tactgtttct tttgtcgatg ctcaccctgt tgtttggtgt	1980
	tacttctgca g	1991
	<210> 31	
	<211> 685	
	<212> ADN	
5	<213> Desconocido	
	<220>	
	<223> Discosoma sp.	
	<400> 31	
	cgccaccatg gcctcctccg agaacgtcat caccgagttc atgcgcttca aggtgcgcat	60
	ggagggcacc gtgaacggcc acgagttcga gatcgagggc gagggcgagg gccgcccta	120
	cgaggggcac aacaccgtga agctgaaggc gacgaagggc ggccccctgc ccttcgcctg	180
	ggacatcctg tccccccagt tccagtacgg ctccaagggt tacgtgaagc accccgccga	240
	catccccgac tacaagaagc tgtccttccc cgagggcttc aagtgggagc gcgtgatgaa	300
	cttcgaggac ggcggtgtgg cgaccgtgac ccaggactcc tccctgcagg acggtgctt	360
	catctacaag gtgaagttca tcggcgtgaa cttcccctcc gacggccccg tgatgcagaa	420
	gaagaccatg ggctgggagg cctccaccga gcgcctgtac ccccgcgacg gcgtgctgaa	480
	ggcgagagacc cacaaggccc tgaagctgaa ggacggcggc cactacctgg tggagttcaa	540
	gtccatctac atggccaaga agcccgtgca gctgcccggc tactactacg tggacgcaa	600
10	gctggacatc acctcccaca acgaggacta caccatcgtg gagcagtacg agcgcaccga	660
	gggcccgcac cacctgttcc tgtag	685
	<210> 32	
	<211> 48	
	<212> ADN	
15	<213> Saccharomyces cerevisiae	
	<400> 32	
	gaagttccta ttccgaagtt cctattctct agaaagtata ggaacttc	48
	<210> 33	

<211> 48
 <212> ADN
 <213> Saccharomyces cerevisiae
 <400> 33
 5 gaagttccta ttccgaagtt cctatttctcc agaaagtata ggaacttc 48
 <210> 34
 <211> 18676
 <212> ADN
 <213> Secuencia artificial
 10 <220>
 <223> Construcción sintética
 <400> 34
 gcatggcatg ccgaatccat gtgggagttt attcttgaca cagatattta tgatataata 60
 actgagtaag cttaacataa ggaggaaaaa catatgttac gcagcagcaa cgatgttacg 120
 cagcagggca gtgccctaa aacaaagtta ggtggctcaa gtatgggcat cattcgcaca 180
 tgtaggctcg gccctgacca agtcaaatac atgcgggctg ctcttgatct ttctcggtcgt 240
 gagttcggag acgtagccac ctactcccaa catcagccgg actccgatta cctcgggaac 300
 ttgctccgta gtaagacatt catcgcgctt gctgccttcg accaagaagc ggtgttggtg 360
 gctctcgcgg cttacgttct gcccaagttt gagcagccgc gtagtgagat ctatatctat 420
 gatctcgcag tctccggcga gcaccggagg cagggcattg ccaccgcgct catcaatctc 480
 ctcaagcatg aggccaacgc gcttggtgct tatgtgatct acgtgcaagc agattacggt 540
 gacgatcccg cagtggctct ctatacaaag ttgggcatac gggaagaagt gatgcacttt 600
 gatatcgacc caagtaccgc cacctaataa tgtctaacaa ttcgttcaag ccgacgccgc 660
 ttccgcccgc ggcttaactc aagcgttaga tgcactatac gtaccaatcc aaatcggatc 720
 cccctgcagg gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt 780
 ttactcatat atactttaga ttgatttaaa acttcatttt taatttataa ggatctaggt 840
 gaagatcctt ttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg 900
 agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt 960

ES 2 843 556 T3

aatctgctgc ttgcaaaaca aaaaaccacc gctaccagcg gtggtttgtt tgccggatca	1020
agagctacca actctttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac	1080
tgtccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac	1140
atacctcgct ctgctaatac tgttaccagt ggctgctgcc agtggcgata agtcgtgtct	1200
taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg	1260
gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga gatacctaca	1320
gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt	1380
aagcggcagg gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta	1440
tctttatagt cctgtcgggt ttccgccacct ctgacttgag cgtcgatttt tgtgatgctc	1500
gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc	1560
cttttgctgg ccttttgctc acatgttctt tcctgcgtta tcccctgatt ctgtggataa	1620
ccgtattacc gcctttgagt gagctgatac cgctcgccgc agccgaacga ccgagcgag	1680
cgagtcagtg agcgaggaag cggagagcg cctgatgcgg tattttctcc ttacgcactc	1740
gtgcggtatt tcacaccgca tatggtgcac tctcagtaca atctgctctg atgcgcata	1800
gttaagccag tatacactcc gctatcgcta cgtgactggg tcatggctgc gccccgacac	1860
ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc tcccgcatc cgcttacaga	1920
caagctgtga ccgtctccgg gagctgcatg tgtcagaggt ttccaccgtc atcacgaaa	1980
cgcgcgaggc agcagatcta tcgatggtac ctatctgcag acggaaggga ttcctgttaa	2040
ttaaaacagc ttgcgtcatg cggctcgtgc gtatatgatg cgatgagtaa ataaacaaat	2100
acgcaagggg aacgcatgaa ggttatcgct gtacttaacc agaaaggcgg gtcaggcaag	2160
acgaccatcg caaccatct agcccgccgc ctgcaactcg ccggggccga tgttctgtta	2220
gtcgattccg atccccaggg cagtgccgc gattggcggg ccgtgcggga agatcaaccg	2280
ctaaccgttg tcggcatoga ccgccgacg attgaccgcg acgtgaaggc catcggccgg	2340
cgcgacttgc tagtgatoga cggagcgccc caggcgcgcg acttggtgt gtccgcgatc	2400
aaggcagccg acttcgtgct gattccggtg cagccaagcc cttacgacat atgggccacc	2460
gccgacctgg tggagctggt taagcagcg attgaggtca cggatggaag gctacaagcg	2520
gcctttgtcg tgtcgcgggc gatcaaaggc acgcgcatcg gcggtgaggt tgccgaggcg	2580
ctggccgggt acgagctgcc cattcttgag tcccgtatca cgcagcgct gagctacca	2640
ggcactgccg ccgccggcac aaccgttctt gaatcagaac ccgagggcga cgctgccgc	2700
gaggtccagg cgctggccgc tgaaattaaa tcaaaactca tttgagttaa tgaggtaaag	2760
agaaaatgag caaaagcaca aacacgctaa gtgccggccg tccgagcgca cgcagcagca	2820

ES 2 843 556 T3

aggetgcaac gttggccagc ctgcgagaca cgccagccat gaagcgggtc aactttcagt	2880
tgccggcgga ggatcacacc aagctgaaga tgtacgcggt acgccaaggc aagaccatta	2940
ccgagctgct atctgaatac atcgcgcagc taccagagta aatgagcaaa tgaataaatg	3000
agtagatgaa ttttagcggc taaaggaggc ggcattggaaa atcaagaaca accaggcacc	3060
gacgccgtgg aatgccccat gtgtggagga acggcggtt ggccaggcgt aagcggctgg	3120
gttgccctgcc ggcctgcaa tggcactgga accccaagc ccgaggaatc ggcgtgagcg	3180
gtcgcaaac atccggcccg gtacaaatcg gcgcggcgct ggggtgatgac ctgggtggaga	3240
agttgaaggc agcgcaggcc gccagcggc aacgcattcga ggcagaagca cgcgccggtg	3300
aatcgtggca agcagccgct gatcgaatcc gcaaagaatc ccggcaaccg ccggcagccg	3360
gtgcgccgtc gattaggaag ccgccaagc gcgacgagca accagatctt ttcgttccga	3420
tgtctatga cgtgggcacc cgcgatagtc gcagcatcat ggacgtggcc gttttccgtc	3480
tgtcgaagcg tgaccgacga gctggcgagg tgatccgcta cgagcttcca gacgggcacg	3540
tagaggtttc cgcaggcccg gcaggcatgg ccagtgtgtg ggattacgac ctgggtactga	3600
tggcggtttc ccatctaacc gaatccatga accgataccg ggaagggaag ggagacaagc	3660
ccggcccggt gttccgtcca cactgtgcg acgtactcaa gttctgccg cgagccgatg	3720
gcggaaagca gaaagacgac ctggtagaaa cctgcattcg gttaaaccac acgcacgttg	3780
ccatgcagcg tacgaagaag gccaaagaac gccgcctggg gacggtatcc gaggggtgaag	3840
ccttgattag ccgtacaag atcgtaaaga gcgaaaccg gcggccggag tacatcgaga	3900
tcgagctagc tgattggatg taccgcgaga tcacagaagg caagaaccg gacgtgctga	3960
cgttccacc cgattacttt ttgatcgatc ccggcatcgg ccgttttctc taccgcctgg	4020
cacgccgcgc cgcaggcaag gcagaagcca gatggttgtt caagacgac tacgaacgca	4080
gtggcagcgc cggagagttc aagaagttct gtttcaccgt gcgcaagctg atcgggtcaa	4140
atgacctgcc ggagtacgat ttgaaggagg aggcggggca ggctggccc atcctagtca	4200
tgcgctaccg caacctgac gagggcgaag catccgccg ttcctaattg acggagcaga	4260
tgctagggca aattgcccta gcaggggaaa aaggtcgaaa aggtctcttt cctgtggata	4320
gcacgtacat tgggaaccca aagccgtaca ttgggaaccg gaaccctgac attgggaacc	4380
caaagccgta cattgggaac cgttcacaca tgtaagtac tgatataaaa gagaaaaaag	4440
gcgatttttc cgcctaaaac tctttaaaac ttattaaaac tcttaaaacc cgcctggcct	4500
gtgcataact gtctggccag cgcacagccg aagagctgca aaaagcgct acccttcggt	4560
cgtcgcgctc cctacgcccc gccgcttcgc gtcggcctat gcatataatt gtggtttcaa	4620
aatcggctcc gtcgatacta tgttatagc caactctgaa aacaactttg aaaaagctgt	4680
tttctggtat ttgaccgcg aggtttctcg cttcaattga aatcataaag aagcaattga	4740

ES 2 843 556 T3

aaattttcga gtaaccgacc ctcccgataa tcttcaacat aaaacaacgc acttcttcca	4800
acgggagaggg cgggtgtagt tgcgagctaa ggagataagg tatgcttaag agatcggggg	4860
cgctttctct tgccttgatg gtctccttct gttcgtcgag ccttgccacg ccactctcat	4920
ctgctgagtt tgaccatggt gctcgcaagt gtgccccatc agttgcgaca tctacgcttg	4980
cggcgatagc taagggtggag agtcgctttg atccttttagc gattcatgac aacacgaccg	5040
gcgaaacgct tcactggcaa gatcacagcc aagcaacca agtcgtcagg caccgtctcg	5100
atgcacggca ttcgctggat gttggcctca tgcaaataaa ctctcgaaat ttttctatgc	5160
tcggtctgac acctgacggt gcgctccagg cgtgcacatc attatctgcc gctgcaaaca	5220
tgctgaaaag tcgttatgca ggccggcgaaa cgattgacga gaagcaattt gcgcttcgtc	5280
ggcgatctc cgcttacaac accggtaatt tcatcgcggt ttttgcaaac ggctacgtgc	5340
gaaaagttga aacagctgct caatcgctgg tgcccgcggt aatcgagcct ccaaagacg	5400
atcacgaggg gctaaaatcc gaagagacgt gggatgtttg ggggtcatat cagcgccgct	5460
cgcaggagga tggcgctggc ggtttaatcg ctccgccacc gccacaccag gacaacggca	5520
aatccgcaga cgacaatcaa gtcttattcg acttatacta aggaggtgcg cattgatgcg	5580
atgctttgag agataccggt tacatctaaa tcgcctctcg ctctcgaatg cgatgatgcg	5640
cgtgatatcg agctgcgcc caagcttggt cgggtgcaatt gcatggagca tttcctcatc	5700
cggacccgcc gcagcgcaat ctgcggttg cggcactgac ccgccacaa tgggtaacaa	5760
tatatgcacg tttatccttg gtccgttcgg ccagtcactc gctgttctcg gcattgtcgc	5820
tatcgggacg tccctggatgt tcgggggggc ttcgcttggg ctggttgccg gcgtcgtcgg	5880
cggcattggt atcatgtttg gggcgagctt cctcgccaa acgctcactg gcggtagttg	5940
atggctgacg gtttggaaga atcgaccctt tacctcgag ccacacggcc cgcattgttt	6000
cttgggggtg cactgacatt ggcagggtta ttcatgatgt tcgccggctt tgtcatcggt	6060
atcgttcaga acccgctcta cgaagtcgtt ctcggtccgt tatggtttgc agcccggtc	6120
atcggtggag gagactacaa tgcggcgagc gtcgtcctgc tatttttgcg gaccgcgga	6180
agaagcattg atagtgcagt ttgggggggc gctactgtta gcccaaacc aattaggggt	6240
ccccacgag ggagaggaat ggtgtgatgc tcggcgcgag tggaacgacc gaaagatccg	6300
gtgagatcta tctcccttat attggccacc tcagcgacca tatcgctcctt cttgaagacg	6360
gatcgatcat gaccattgcg agaattgatg gcgttgcatc cgagcttgag gaaactgaaa	6420
tgcgcaatgc gcgttgctgt gcgttcaaca cgctgttgcg caatatcgct gatgatcatg	6480
tgtcaatata tgctcacctc gtacgtcatg ccgacgtgcc atcatcgcg ccgcgacact	6540
tccgtagtgt tttcgccgct agcctgaacg aagcttttga acagcgcggt ctctccggcc	6600

ES 2 843 556 T3

aactcctccg caatgaacac ttccttacgt tgattgtcta cccacaggcg gcttttaggga	6660
aggtaaagag gaggttcacc aagctaagcg gaaaaaggga aaacgatctc acggggccaga	6720
tcaggaacat ggaagatctt tggcatgttg tcgctggctc tcttaaagcg tatggcctgc	6780
atcgtcttg catccgcgag aagcagggtg tgctcttcac cgaaattggc gaagcgctac	6840
ggttgatcat gactggtcgg ttcacaccgg ttccggtcgt cagcggctca ctcggcgctt	6900
cgatttatac cgacagagtc atttgccgga agcgaggact cgagatcaga acgccaaaag	6960
acagttacgt tggatccatc tattcgtttc gcgaataccc tgcaaaaaca cggccgggca	7020
tgctcaacgc gctgctatcc ctcgattttc cacttgttct cagcgagagt ttttcgttcc	7080
tgactcgccc gcaagcgcac gcgaaactta gcctcaaatc gagccagatg ctgagttccg	7140
gcgataaagc cgtgactcaa atcggcaaat tatccgaggc tgaggacgca ctgcgagca	7200
acgaattcgt tatgggctca catcatttga gcctttgcgt ctatgcagac gatctcaata	7260
gtcttgggga caggggcgcg cgggctcggg cacgaatggc ggatgcaggc gccgtggttg	7320
tccaagaagg tattggtatg gaagcggcct attggtccca attgccgggg aattttaagt	7380
ggcgcacacg ccctggcgca atcacttcac gcaatttcgc agggtttgct tctttcgaaa	7440
actttccaga gggcgccagc tcaggccact ggggcaacgc gattgccga tttcgtacca	7500
atggcggaac gcctttcgac tataatccgc atgagcacga tgttgccatg acggcaatat	7560
tcgggcctat cgggaggggt aagacgacgc tcatgatgtt tgttctagcc atgctcgac	7620
agagcatggt cgaccgtgca ggtacggtcg tgttctttga caaggaccgg ggtggcgaat	7680
tgctggttcg cgccacagga ggaacatatt tggcacttca cagaggcaca cccagcgggt	7740
tggcgccgtt gcgtggccta gaaaacacag cagcctcaca cgattttctg cgcgaaatgga	7800
tcgtggctct catcgagagt gatggtcggg gtgggatttc tccggaagag aaccgccgtc	7860
tggtcctggt tatccatcgt cagctctcgt ttgatccaca aatgcgttca atcgcggggt	7920
tacgtgaatt tttgttgcg gggcccgccg aaggcgcagg agcgcggctc caacgctggt	7980
gccggggcca tgcgcttggc tgggcatttg acggcgaagt tgacgaagta aagttagatc	8040
cgtcgattac cggttcgac atgacgcac tttctgaata cgaggaagta tgcgctcccg	8100
ctgcagcata tctcctgcat cggattggag ccatgatcga cggccgccgt tttgtgatga	8160
gctgcgatga gtttcgcgcc tatttgtaa accctaaatt ttcgactgtc gtcgacaaat	8220
tcctcctgac cgttcgaaaa aacaacggga tgctaatact ggcaacgcag caaccagagc	8280
atgttctgga atcgccgcta ggagccagct tggttgcgca atgtatgacg aagattttct	8340
atccatcacc aaccgcagat cgatcggctt atgtcgatgg actgaaatgt accgaaaagg	8400
aatttcaggc gatccgtgaa gacatgacgg tcggcagccg taagtttctt cttaaacgag	8460
aaagtggaag cgtcatctgc gaatttgatc tgcgggatat gcgtgaatat gtcgccgtgc	8520

ES 2 843 556 T3

tttcggggcg tgccaacacg gtgcgctttg caactcgact acgcgaggca caagaaggca	8580
actcatctgg ctggctcagc gaattcatgg cccgtcacca cgaggcagaa gattgataag	8640
gtaggaaacg atgaagacga cgcaacttat tgcaacagtt ttgacctgca gctttctata	8700
tattcagccc gcgcggggcg agtttggtgt tagcgacccg gcaacggagg ctgagacgct	8760
cgcgactgcg ctgcgactg cggagaatct cactcagact atagcgatgg ttacgatggt	8820
gacgtcggcc tacggcggtta ctggactact gacttcgctc aaccagaaaa atcagtatcc	8880
ttcgacgaag gacctagaca atgaaatgtt ttccgccgga atgccaatgt cgaccacggc	8940
acgtgcgac accagcgata cagatcgtgc agtcgtgggt agtgatgctg aagcggacct	9000
gttgcgatcg cagatcacccg gttccgcaaa cagcgctggc attgcggtcg acaatctgga	9060
aacgatggac aaacgcttga cggcgaatgc tgatacgtct gctcagcttt cccgatctcg	9120
caatatcatg caggcaaccg tgaccaatgg tttgcttctc aagcagatcc atgacgcaat	9180
gattcaaaat gtacaggcga caagcctatt aacgatgact accgcgcagg ccggccttca	9240
cgaggcggaa gaggcggccg ctcaacgcaa ggagcatcaa aagaccgtg tcatctttgg	9300
tgccctcccc taaggctggg cgatttggtc atccgccgc atcctcgccg aatgcgagct	9360
cattttatcc aacattatgc gacaaaccag tcaagttcag gtccaatcga tgaatttcac	9420
gattccggcg ccgtttacgg ccattcatac gatcttcgat gtagccttca cgacaggctt	9480
ggactcgatg cttgagacta tccaggaggc ggtgagtgcg ccattgatcg cctgtgtcac	9540
tctttggatt attgttcagg gtattttagt catacgccgc gaagtcgata cccgtagcgg	9600
tatcactcgg gtgatcacgg tcaccatcgt tgttgctcta attgttgggc aggctaacta	9660
ccaagactat gtggtttcca tcttcgaaaa gacggtccca aactttgttc agcagtttag	9720
tgtaaccggc ttgcctctgc agactgttcc ggcacagttg gatacaatgt tcgccgtgac	9780
ccaggccggt tttcagaaaa tcgcatccga aatcggtcg atgaacgacc aggacatcct	9840
tgctttccaa ggggcacagt gggtccttta cggcacgctc tggctctgcct tcggagttta	9900
cgcgcgcggt ggaattctca cgaaagtgtc tctcgcgac gggcctctga tcctcgtcgg	9960
atatattttt gatcgcacgc gggacatcgc agctaagtgg atcgggcaac ttatcaccta	10020
cggctctctt cttctcctct taaacctcgt ggcaacgac gtcacccctaa ccgaagcgac	10080
tgcgctcacc cttatgcttg gtgtaatcac cttcgccggt acgaccgcgg ccaagatcat	10140
tggctctttac gaactcgata tgttttttct gacaggggat gcgctcattg tcgctttgcc	10200
ggcgatcgcc ggcaacattg gaggcagtta ctggagcggc gcaaccaat ctgccagcag	10260
cttgtagcgt cgcttcgctc aggttgagcg aggttaggtc gcgcaaaaat tcgcctcaat	10320
ggagaattct atgaagtatt gcctgctgtg cctagtgtgc gctttgagcg gctgccagac	10380

ES 2 843 556 T3

aaacgacaca ttagcgagct gcaaaggccc gatcttcccg ctgaatgtgg ggcgatggca	10440
gcctactccg tcagatcttc agctcggcaa ttccgggtgga cgctatgacg gggcctgaat	10500
atgccatgct agtggcgcgc gaaagccttg ccgagcacta taaggaagta gaagcctttc	10560
aaaccgcgcg agcgaaatcg ggcgcgacgtc tctccaaact cattgcagct gtcgcagcta	10620
tcgcgatttt gggaaatgtt gctcaagcgt tcgctatagc cacaatggtg ccgttgagca	10680
ggcttgtgcc cgtatatcta tggatacggc cggacggcac cgttgacagc gaggtgtctg	10740
tctcgcgatt gcctgcaact caagaggag cgcgcgttaa cgcctcattg tgggagtagc	10800
ttcgccctgcg cgagagttat gatgccgaca ccgctcagta cgcctacgac ctggtatcga	10860
acttcagtgc cccaacagtg cgccaggatt accagcaatt cttcaactat cccaatccca	10920
gttcgcctca agtcattctt ggcaaacgcg gcaggggtgga ggtcgagcac atcgcttcaa	10980
atgatgtaac tccaagcacg cagcaaattc gctataaaag gaccctcgtc gttgacggca	11040
aaatgcctgt ggtgagtacg tggaccgcga cagttcgcta cgaaaagggtg accagcttgc	11100
ccggcagatt gagactaacc aaccggcgag gtctggttgt cacctcctat cagacatcgg	11160
aagataccgt ttcaaacgta ggccacagcg aaccatgac agaaaagcac ttttcatttt	11220
agcatgttta ttgcccgtg cgactgggtc ggagcgtgaa gacactccaa tggcgggcaa	11280
gctagatccg cgcatgcgtt atttggctta caatcccgat caagtgggtc gcctctcgac	11340
ggcgggttga gctacttttg tcgtaacatt cgccacgaac gaaacgggtg cagcggttgc	11400
cgtttcaaat agcaaagatc tagcagccct accgcgggga aattatctat ttttcaaggc	11460
aagccaggtc ctacgcctc agccagtaat cgtgctaacc gcaagcgact ccgggatgcg	11520
ccgttatgtt ttcagtataa gttccaagac tctgtcccac ctcgataaag agcagcccg	11580
tctctattac agcgtccaat tcgcctaccc cgccgacgat gcggcggctc ggcgaaggga	11640
ggcacaacag aaggctgttg tggacagact acacgcggaa gcacaatatc aacggaaagc	11700
tgagaattta ttggtacagc ctgtcacagc ccttgggtgcg gcggacagta attggcacta	11760
cgtcgcccaa ggcgatcgtt cgctgttgcc actcgaagtc ttcgacaatg gatttacgac	11820
ggtattccac tttccgggca atgtacgcat accctccatc tacaccatca atcctgatgg	11880
caaggaagct gttgccaaact attcagttaa agggagcgat gtcgagattt cttcggtttc	11940
ccgagggttg cgtctgaggg atggccacac agtactatgt atctggaaca ccgcttacga	12000
tcccgttggc caaaggccgc aaacgggcac ggtgaggccc gatgtgaaac gcgtcctgaa	12060
gggggcgaag ggatgaataa cgatagtcag caagcggcac atgaggttga tgcattctgga	12120
tccctggtct ccgacaaaca tcgccggcgt ctttcgggggt ctcagaaatt gatcgtcgga	12180
ggtgtcgttc tcgcgttatc attaaagcctc atttggttag gtgggcgcca aaagaagggtg	12240
aatgagaacg catcgccgtc aactttgatc gcaacaaaca ccaagccatt tcatccagct	12300

ES 2 843 556 T3

ccgattgagg tgccgcccga tcctccagcg gttcaagagg ctgttcagcc tgctgtcct	12360
ctaccgccga ggggcgaacc ggagcggcat gagccacggc cggaagaaac accgattttt	12420
gcataatagca gcggcgatca aggggtcagc aaacgcgccca ttcagggcga cacgggccga	12480
agacaagaag gcaagcgtga cgacaactcc ttgccgaatg gcgaagtgtc cggcgagaac	12540
gattttgtcga tacgtatgaa acccaccgag ctgcagccca gcagcgccac gctcttgccg	12600
caccccgatt ttatggtaac gcaagggaca ataattccgt gcattcttga aaccgcaatc	12660
gacacaaatt tggcaggcta tgtaaagtgt gtcttgctc aggatattcg tggaacaacg	12720
aacaatatcg tgcttcttga tcgtggcacc accgttgttgc gcgaaatata gcgtggcttg	12780
caacagggag atgggcgcgt ttttgtgttg tgggatcgcg ccgagacacc tgaccatgcg	12840
atgatctcgt taacatcgcc aagcgcggac gaactcggtc gctcaggatt gccgggctcg	12900
gtcgacagcc acttcttgga gcgttttagc ggagctatgc tcttgagtgt tgttcaaggc	12960
gccttccagg cagctagcac ctacgccggc agctcgggtg gcgggatgag cttcaacagc	13020
tttcaaaata acggtgagca gacaactgag acagccctta aggcaaccat caacataccg	13080
ccaaccctga agaagaatca gggtagacacc gtttccattt tcgtagcacg ggacctcgat	13140
ttctttggtg tttaccagct ccgcctgact ggcggcgcca cgcgggggag gaaccgccgc	13200
tcttaatgaa ttcaaatttc cgcttagaga taggatacat tgtaaattga agtggatccg	13260
caactacgct ttctttgaa gccgattttg gaatggctcg atgacccgaa gactgaagaa	13320
attgcgatca atcgacctgg agaggcattt gtgcgccaaag ccggcatttt taccaagatg	13380
cctttgcccg tctcttatga tgatcttga gatatcgcta ttttagcggg cgcgctgaga	13440
aagcaggatg tcggaccacg taacccctc tcgcccactg aacttcctgg tggatgaacga	13500
ctacaaatct gtctgccgcc gaccgttccc tcgggcaccg tcagcttgac cattcgacgg	13560
ccaagctcgc gtgtttctgg tcttaaagaa gtctcctccc gttatgatgc ttcgaggtgg	13620
aaccagtggc agacacgaag gaaacgccaa aatcaggatg atgaagctat ccttcagcat	13680
tttgacaacg gggatttgga agcgtttctg cacgcatgcg tcgtcagccg actgacgatg	13740
ttgctatgtg gccctaccgg aagcggcaag acaacaatga gcaagacctt gatcagcgcc	13800
atcccccccc aggaaaggct aatcaccata gaagatacgc tcgaactcgt cattccacac	13860
gataatcatg ttagactact ctactccaag aacgggtgctg ggctgggtgc tgtgagcgcc	13920
gagcaattgc tccaagcaag tctgcgtatg cgcccgacc ggatattgct tggcgagatg	13980
cgcgacgatg cagcatgggc ttatctgagt gaagtcgtct cgggacatcc gggatcgatt	14040
tcaacaatac acggcgcgaa tccaatccaa ggattcaaga aactgttttc ccttgtcaaa	14100
agtagcgccc aagggtgctag cttggaagat cgcacactga ttgacatgct ctctacggcg	14160

ES 2 843 556 T3

atcgatgtca tcattccatt ccgtgcctat gaggacgttt atgaagtagg cgagatctgg	14220
ctcgcggcgg acgcacgacg ccggggcgag accataggcg atctccttaa tcaatagtag	14280
ctgtaacctc gaagcgtttc acttgtaaca acgattgaga acttttgtca taaaattgaa	14340
atacttggtt cgcattttcg tcatccgcgg tcagccgcaa ttctgacgaa ctgcccattt	14400
agctggagat gattgtacat ccttcacgtg aaaattttctc aagcgctgtg aacaagggtt	14460
cagatttttag attgagaggt gagccgttga aacacgttct tcttatcgat gacgatgtcg	14520
ctatgcggca tcttattatc gaataacctta cgatccacgc cttcaaagtg accgcggtag	14580
ccgacagcac ccagttcact agagtactct cttccgcgac ggtcgatgtc gtggttgttg	14640
atctaaattt aggtcgtgaa gatgggcttg agatcgttcg aaatctggcg gcaaagtctg	14700
atattccaat cataattatc agtggcgacc gccttgagga gacggataaa gttgttgac	14760
tcgagctagg agcaagtgat tttatcgcta agccgttttag tacgagagag tttcttgac	14820
gcattcgggt tgccttgccg gtgcgcccc acgttgtccg ctccaaagac cgacggtctt	14880
tttgttttac tgactggaca cttaactctca ggcaacgtcg cttgatgtcc gaagctggcg	14940
gtgaggtgaa acttacggca ggtgagttca atcttctcct cgcgttttta gagaaacccc	15000
gcgacgttct atcgcgcgag caacttctca ttgccagtcg agtacgcgac gaggaggttt	15060
acgacaggag tatagatgtt ctcattttgc ggctgcgcg caaacttgag gcggatccgt	15120
caagccctca actgataaaa acagcaagag gtgccggtta tttctttgac gcggacgtgc	15180
aggtttcgca cggggggacg atggcagcct gagccaattg catttgatta atttaggtga	15240
ctgaggacgc ggccagcggc ctcaaaccta cactcaatat ttggtgaggg gttccgatag	15300
gtccctcttc accaattgct cgatggcttc tctccagcaa agaatacgcg gagcgcggcg	15360
gtagccagct tgtggccgaa agctcgagcg gtctccaacc ccaacggatc aaaatgactt	15420
cgagcgacct cgagcaacgc aaccgggaac atgcgtgagg tctgaacgag aacggatttt	15480
tctgtagttg aagggatcgg ataaacttttc ggggccacgc gaaatgatcc atctgccagc	15540
atgctttcga aatcgtccaa cgcgcgcctt aaaatcattt gtagcgactt cgagggactg	15600
tattgccgaa cgaggttgtc atatgttttc gacacttgag gcgcgggagg tcgcgctgaa	15660
agaaaaacct ggagcttttt cggggacgga ggtggactaa gggcatccac agttagctta	15720
agttgtcgat cgggactgta aatgtgatcg gcgacgagag gctcacgttg ctggtctttc	15780
tcgtcggctt tttcaggcaa gtgctggagg tccagcttct ggggaacaag tgtcgggttg	15840
ggatggtgga tctcgggtcg agcaccagca agccgcctg cttcgcgcgac cgacaatcg	15900
ggcttgcgaa ttgccatctt caagcctcca agattttgct gatcagtttc gaaatgacca	15960
cgacttctc catcgcaatc cgaagattcc tctctatgag gcgcacgtc ggatcagttc	16020
ccgtgttttag taatgtaaga tgcaacatgc cgcgttcttt catcgcggca aatgcacctc	16080

ES 2 843 556 T3

tttcatgcat gggagacggt acaactggaa ggctctctag cgtctctgac atcctgcgtt	16140
gcgatgttgt caatcgccg accgggacgc gttggcgcaa aacagctgta ggaattgcc	16200
aattttcaact caacagcagc tcgatgacgt agcggtaggt agatagtgcc tcatcgatgt	16260
cgagcggcgt tagcatggtg gggatcagaa gcaggtttga gctagcgatg attgtgttgt	16320
tgagctcgct cgagccgcca cgcgtatcgg ccaacgcata atcaaactcct tcgagctcgg	16380
cattttcata ggctgcttca agaaggggca tttcgtcggc ggaatagact tcacagcgag	16440
gatcccaggt actgctttgt aaggcgtttt ctctccatcg cgtcagaggc cggttttcgt	16500
cggcatcaaa gagggccact cgtttaccgt catttgccaa agcagcgcaa agggccatga	16560
gtgcggtggt tttgccagca ccccttttga aagaacaaaa cgtcaaaagt tgcataattct	16620
gatccgcct atcctgtgaa accggagtgc atttgtattt ttgttcgtat aaatgttttt	16680
gtgattatcg atgagtaaaa gcgttggttac actattttta tttcacattc gttataagac	16740
aattgcaaat gtagcaagta tattcagtat tgactgtaaa tgtactgttg atttcatatt	16800
gagcagggct agacttccat ccgtctaccc gggcacattt cgtgctggag tatccagacc	16860
ttccgctttc tttggaggaa gctatgtcaa aacacaccag agccacgtcg agtgagacta	16920
ccatcaacca gcatcgatcc ctgaaagtgt aaggggtcaa ggtcgtgagt gcccgctctgc	16980
gatcgccga gtatgaaacc ttttcctatc aagcgcgcct gctgggactt tcggatagta	17040
tggcaattcg cgttgcggtg cgcgcgatcg ggggctttct cgaaatagat gcagacacac	17100
gagaaaagat ggaagccata cttcagttca tcggaatact ctcaagcaat gtatccatgc	17160
ttctatctgc ctacgccgaa gacctcgat cggatctgga ggctgtgcga gatgaacgta	17220
ttgcttttgg cgaggctttc gccgccctcg atggactact gcgctccatt ttgtccgtat	17280
cccgccgacg gatcgacggt cgtcgtttac tgaaaggtgc cttgtagcac ttgaccacgc	17340
acctgacggg agaaaattgg atgccgatc gcgctcaagt aatcattcgc attgtgccag	17400
gaggtggaac caagaccctt cagcagataa tcaatcagct ggagtacctg tcccgaagg	17460
gaaagctgga actgcagcgt tcagcccggc atctcgatat tcccgttccg ccgcatcaaa	17520
tccgtgagct tgcccaaagc tgggttacgg aggccgggat ttatgacgaa agtcagtcag	17580
acgatgacag gcaacaagac ttaacaacac acattattgt aagcttcccc gcaggtaccg	17640
accaaaccgc agcttatgaa gcaagccggg aatgggcagc cgagatgttt gggtcaggat	17700
acgggggttg ccgctataac tatctgacag cctaccacgt cgaccgcgat catccacatt	17760
tacatgtcgt ggtcaatcgt cgggaacttc tggggcaggg gtggctgaaa atatccaggc	17820
gccatcccca gctgaattat gacggcttac ggaaaaagat ggcagagatt tcacttcgtc	17880
acggcatagt cctggatgcg acttcgcgag cagaaagggg aatagcagag cgaccaatca	17940

catatgctga atatcgacgc cttgagcgga tgcaggctca aaagattcaa ttcgaagata 18000
cagattttga tgagacctcg cctgaggaag atcgtcggga cctcagtcaa tcgttcgatac 18060
catttcgatac ggacgcatct gccggcgaac cggaccgtgc aacccgacat gacaaacaac 18120
cgcttgaacc gcacgcccgt ttccaggagc ccgcccggctc cagcatcaaa gccgacgcac 18180
ggatccgcgt accattggag agcgagcggg gtgcccacc atccgcgtcc aaaatccctg 18240
taactgggca tttcgggatt gagacttcgt atgtcgctga agccagcgtg cccaaacaaa 18300
gcggcaattc cgatacttct cggccggtga ctgacgttgc catgcacaca gtcgagcgcc 18360
agcagcgatac aaaacgacgt catgacgagg aggcaggctc gagcggagca aaccgtaaaa 18420
gattgaaggc cgcgcaagtt gattccgagg caaatgtcgg tgagcccgac ggtcgcgatg 18480
acagcaacaa ggccgctgat ccggtgtctg cttccatccg taccgagcaa ccggaagctt 18540
ctccaacgtg tccgcgtgac cgtcacgatg gagaattggg agaacgcaa cgtgcaagag 18600
gtaatcgctg cgacgatggg cgcgggggga cctagactag tgtacctcgc gaatgcatct 18660
agatccaatc caatat 18676

<210> 35

<211> 28390

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 35

gaccgcgag gtttctcgt tcaattgaaa tcataaagaa gcaattgaaa attttcgagt 60
aaccgaccct cccgataatc ttcaacataa aacaacgcac ttcttccaac gggagaggcg 120
gtgttagttg cgagctaagg agataaggta tgcttaagag atcggggctg ctttctcttg 180
ccttgatggt ctcttctgt tcgtcgagcc ttgccacgcc actctcatct gctgagtttg 240
accatgttgc tcgcaagtgt gcccacacag ttgcgacatc tacgcttgcg gcgatatgcta 300
aggtggagag tcgctttgat ctttagcga ttcatgacaa cacgaccggc gaaacgcttc 360
actggcaaga tcacagcaa gcaacccaag tcgtcaggca ccgtctcgat gcacggcatt 420
cgctggatgt tggcctcatg caaataaact ctcgaaattt ttctatgctc ggtctgacac 480
ctgacggtgc gctccaggcg tgcacatcat tatctgccgc tgcaaacatg ctgaaaagtc 540
gttatgcagg cggcgaaacg attgacgaga agcaatttgc gcttcgtcgg gcgatctccg 600
cttacaacac cggtaatttc atcggcgggt ttgcaaacgg ctacgtgcga aaagttgaaa 660
cagctgctca atcgctggtg cccgcgttaa tcgagcctcc aaaagacgat cacgaggcgc 720
taaaatccga agagacgtgg gatgtttggg ggtcatatca gcgccgctcg caggaggatg 780
gcgctggcgg tttaatcgct ccgccaccgc cacaccagga caacggcaaa tccgcagacg 840

10

ES 2 843 556 T3

acaatcaagt cttattcgac ttataactaag gaggtgcgca ttgatgcgat gctttgagag	900
ataccggttta catctaaatc gcctctcgct ctcgaaatgcg atgatgcgcg tgatatcgag	960
ctgcgccccca agcttgtgcg gtgcaattgc atggagcatt tcctcatccg gaccgcgcgc	1020
agcgcaatct gcgggtggcg gcaactgaccc cgccacaatg gttaacaata tatgcacggt	1080
tatccttgggt ccgttcggcc agtcactcgc tgttctcggc attgtcgcta tcgggatctc	1140
ctggatgttc gggcgggctt cgcttgggct ggttgccggc gtcgtcggcg gcattgttat	1200
catgtttggg gcgagcttcc tcggccaaac gctcactggc ggtagttgat ggctgatcgt	1260
ttggaagaat cgacccttta cctcgcagcc acacggcccg cattgtttct tggggtgcca	1320
ctgacattgg cagggttatt catgatgttc gccggctttg tcatcgttat cgttcagaac	1380
ccgctctacg aagtcgttct cgtgccgtta tggtttgacg cccggctcat cgtggagcga	1440
gactacaatg cggcgagcgt cgtcctgcta tttttcgga cccggggaag aagcattgat	1500
agtgcagttt gggggggcgc tactgttagc ccaaatccaa ttagggttcc cccacgaggg	1560
agaggaatgg tgtgatgtc ggcgcgagtg gaacgaccga aagatccggt gagatctatc	1620
tcccttatat tggccacctc agcgaccata tcgtccttct tgaagacgga tcgatcatga	1680
ccattgcgag aattgatggc gttgcattcg agcttgagga aactgaaatg cgcaatgcgc	1740
gttgtcgtgc gttcaacacg ctgttcgca atatcgctga tgatcatgtg tcaatatatg	1800
ctcacctcgt acgtcatgcc gacgtgccat catcggcgc gcgacacttc cgtagtgttt	1860
tcgccgctag cctgaacgaa gcttttgaac agcgcgtgct ctccggccaa ctccctccgca	1920
atgaacactt ccttacgttg attgtctacc cacaggcggc tttagggaag gtaaagagga	1980
ggttcaccaa gctaagcgga aaaagggaaa acgatctcac gggccagatc aggaacatgg	2040
aagatctttg gcatgttgc gctggctctc ttaaagcgta tggcctgcat cgtcttggca	2100
tccgcgagaa gcagggtgtg ctcttcaccg aaattggcga agcgctacgg ttgatcatga	2160
ctggtcgggt cacaccggtt ccggtcgtca gcggctcact cggcgcttcg atttataccg	2220
acagagtcatt ttgcggcaag cgaggactcg agatcagaac gccaaaagac agttacgttg	2280
gatccatcta ttcgtttcgc gaataccctg caaaaacacg gccgggcatg ctcaacgcgc	2340
tgctatccct cgattttcca cttgttctca cgcagagttt ttcgttcctg actcgccgc	2400
aagcgcacgc gaaacttagc ctcaaatcga gccagatgct gagttccggc gataaagccg	2460
tgactcaaat cggcaaatta tccgaggctg aggacgcact tgcgagcaac gaattcggtta	2520
tgggtcaca tcatttgagc ctttgcgtct atgcagacga tctcaatagt cttggggaca	2580
ggggcgcgcg ggctcggaca cgaatggcgg atgcaggtgc cgtggttgtc caagaaggta	2640
ttggtatgga agcggcctat tgggtccaat tgccggggaa ttttaagtgg cgcacacgcc	2700

ES 2 843 556 T3

ctggcgcaat cacttcacgc aatttcgcag ggtttgtctc tttcgaaaac tttccagagg	2760
gcgccagctc aggccactgg ggcaacgcga ttgcccgatt tcgtaccaat ggcggaacgc	2820
ctttcgacta tatcccgcat gagcacgatg ttggcatgac ggcaatattc gggcctatcg	2880
ggaggggtaa gacgacgctc atgatgtttg ttctagccat gctcgaacag agcatggtcg	2940
accgtgcagg tacggctcgtg ttctttgaca aggaccgggg tggcgaattg ctggttcgcg	3000
ccacaggagg aacatatttg gcacttcaca gaggcacacc cagcgggttg gcgccgttgc	3060
gtggcctaga aaacacagca gcttcacacg atttttctgc cgaatggatc gtggctctca	3120
tcgagagtga tggtcggggg gggattttctc cggaagagaa ccgccgtctg gtccgtggta	3180
tccatcgta gctctcgttt gatccacaaa tgcgttcaat cgcgggggta cgtgaatttt	3240
tgttgcatgg gcccgccgaa ggcgcaggag cgcggctcca acgctggtgc cggggccatg	3300
cgttggtgctg ggcatattgac ggcgaagttg acgaagtaaa gttagatccg tcgattaccg	3360
gcttcgacat gacgcattct ctcgaatacg aggaagtatg cgctcccgct gcagcatatc	3420
tcctgcatcg gattggagcc atgatcgacg gccgccgttt tgtgatgagc tgcgatgagt	3480
ttcgcgccta tttgttaaac cctaaatttt cgactgtcgt cgacaaattc ctctgaccg	3540
ttcgaaaaaa caacgggatg ctaatactgg caacgcagca accagagcat gttctggaat	3600
cgcgcgctagg agccagcttg gttgcgcaat gtatgacgaa gattttctat ccatcaccaa	3660
ccgcagatcg atcggttat gtcgatggac tgaaatgtac cgaaaaggaa tttcaggcga	3720
tccgtgaaga catgacggtc ggcagccgta agtttcttct taaacgagaa agtggaagcg	3780
tcactcgca atttgatctg cgggatatgc gtgaatatgt cgcctgtcct tcggggcgtg	3840
ccaacacggt gcgttttgca actcgactac gcgaggcaca agaaggcaac tcactcggct	3900
ggctcagcga attcatggcc cgtcaccacg aggcagaaga ttgataaggc aggaaacgat	3960
gaagacgacg caacttattg caacagtttt gacctgcagc tttctatata ttcagcccg	4020
gcggggcgag tttgttgta gcgacccggc aacggaggct gagacgctcg cgactgcgt	4080
cgcgactgcg gagaatctca ctcagactat agcgatgggt acgatgttga cgtcggccta	4140
cggcgttact ggactactga cttcgctcaa ccagaaaaat cagtatcctt cgacgaagga	4200
cctagacaat gaaatgtttt cgcgcgcaat gccaatgtcg accacggcac gtgcgatcac	4260
cagcgataca gatcgtgcag tcgtgggtag tgatgctgaa gcggacctgt tgcgatcgca	4320
gatcaccggt tccgcaaaca gcgctggcat tgcgctgac aatctggaaa cgatggacaa	4380
acgcttgacg gcgaatgctg atacgtctgc tcagctttcc cgatctcgca atatcatgca	4440
ggcaaccgtg accaatggtt tgcttctcaa gcagatccat gacgcaatga ttcaaatgt	4500
acaggcgaca agcctattaa cgatgactac cgcgcaggcc ggccttcacg aggcggaaga	4560
ggcggccgct caacgcgaag agcatcaaaa gaccgctgtc atctttggtg ccctccccta	4620

ES 2 843 556 T3

aggctgggcg atttgttcat ccgcccgcac cctcgccgaa tgcgagctca ttttatccaa	4680
cattatgcga caaaccagtc aagttcaggt ccaatcgatg aatttcacga ttccggcgcc	4740
gtttacggcc attcatacga tcttcgatgt agccttcacg acaggcttgg actcgatgct	4800
tgagactatc caggaggcgg tgagtgcgcc attgatcgcc tgtgtcactc tttggattat	4860
tgttcagggt attttagtca tacgcggcga agtcgatacc cgtagcggta tctactcgggt	4920
gatcacggtc accatcggtg ttgctctaata tgttgggcag gctaactacc aagactatgt	4980
ggtttccatc ttcgaaaaga cgggtccaaa ctttgttcag cagtttagtg taaccggctt	5040
gcctctgcag actgttccgg cacagttgga tacaatgttc gccgtgacct aggccgtttt	5100
tcagaaaatc gcatccgaaa tcggtccgat gaacgaccag gacatccttg ctttccaagg	5160
ggcacagtgg gtcctttacg gcacgctctg gtctgccttc ggagtttacg acgccgttgg	5220
aattctcacg aaagtgttcc tcgcgatcgg gcctctgatc ctgctcggtat atatttttga	5280
tcgcacgcgg gacatcgag ctaagtggat cgggcaactt atcacctacg gtctcttgct	5340
tctcctctta aacctcgtgg caacgatcgt catcctaacc gaagcgactg cgctcaccct	5400
tatgcttgggt gtaatcacct tcgccggtac gaccgcggcc aagatcattg gtctttacga	5460
actcgatatg ttttttctga caggggatgc gctcattgtc gctttgccgg cgatcgccgg	5520
caacattgga ggcagttact ggagcggcgc aaccaatct gccagcagct tgtaccgtcg	5580
cttcgctcag gttgagcgag gtaggtcgc gcaaaaattc gcctcaatgg agaattctat	5640
gaagtattgc ctgctgtgcc tagttgtcgc tttgagcggc tgccagacaa acgacacatt	5700
agcgagctgc aaaggcccgga tcttcccgct gaatgtgggg cgatggcagc ctactccgtc	5760
agatcttcag ctcggaatt cgggtggacg ctatgacggg gcctgaatat gccatgctag	5820
tggcgcgcgga aagccttgcc gagcactata aggaagtaga agcctttcaa accgcgcgag	5880
cgaatcggc gcgacgtctc tccaaactca ttgcagctgt cgcagctatc gcgattttgg	5940
gaaatgttgc tcaagcgttc gctatagcca caatggtgcc gttgagcagg cttgtgcccg	6000
tatatctatg gatacggccg gacggcaccg ttgacagcga ggtgtctgtc tcgcgattgc	6060
ctgcaactca agaggaggcc gtcgttaacg cctcattgtg ggagtacgtt cgctgcgcg	6120
agagttatga tgccgacacc gtcagtagc cctacgacct ggtatcgaac ttcagtgcc	6180
caacagtgcg ccaggattac cagcaattct tcaactatcc caatcccagt tcgcctcaag	6240
tcattcttgg caaacgcggc aggggtggag tcgagcacat cgcttcaaata gatgtaactc	6300
caagcacgca gcaaatcgc tataaaagga ccctcgctgt tgacggcaaa atgcctgtgg	6360
tgagtacgtg gaccgcgaca gttcgctacg aaaagggtgac cagcttgccc ggcagattga	6420
gactaaccac cccggcagggt ctggttgtca cctcctatca gacatcgga gataccgttt	6480

ES 2 843 556 T3

caaacgtagg ccacagcgaa ccatgatcag aaaagcactt ttcatttttag catgtttatt	6540
tgccgctgcg actggtgcgg aggctgaaga cactccaatg ggggcaagc tagatccg	6600
catgcggttat ttggcttaca atcccgatca agtggtgcg ctcctcgacgg cggttgagc	6660
tactttggtc gtaacattcg ccacgaacga aacggtgaca ggggttgccg tttcaaatag	6720
caaagatcta gcagccctac cgcggggaaa ttatctatctt ttcaaggcaa gccaggtcct	6780
cacgcctcag ccagtaatcg tgctaaccgc aagcgactcc gggatgcgcc gttatgtttt	6840
cagtataagt tccaagactc tgtcccacct cgataaagag cagcccgatc tctattacag	6900
cgtccaattc gcctaccccg ccgacgatgc gggcgctcgg cgaaggagg cacaacagaa	6960
ggctgttggtg gacagactac acgcggaagc acaatatcaa cggaagctg agaatttatt	7020
ggatcagcct gtcacagccc ttggtgcggc ggacagtaat tggcactacg tcgccaagg	7080
cgatcggttcg ctgttgccac tcgaagtctt cgacaatgga tttacgacgg tattccactt	7140
tccgggcaat gtacgcatac cctccatcta caccatcaat cctgatggca aggaagctgt	7200
tgccaactat tcagttaaag ggagcgatgt cgagatttct tcggtttccc gaggttgcg	7260
tctgagggat gggcacacag tactatgtat ctggaacacc gcttacgatc ccgttgcca	7320
aagggcgcaa acgggcacgg tgaggccga tgtgaaacgc gtcctgaagg gggcgaagg	7380
atgaataacg atagtcagca agcggcacat gaggttgatg catctggatc cctggtctcc	7440
gacaaacatc gccggcgctt ttcggggtct cagaaattga tcgtcggagg tgcgttctc	7500
gcgttatcat taagcctcat ttggctaggt gggcgccaaa agaagtgaa tgagaacgca	7560
tcgccgtcaa ctttgatcgc aacaaacacc aagccatttc atccagctcc gattgaggtg	7620
ccgccggatc ctccagcggc tcaagaggct gttcagcctg ctgctcctct accgccgagg	7680
ggcgaaccgg agcggcatga gccacggccg gaagaaacac cgatttttgc atatagcagc	7740
ggcgatcaag gggtcagcaa acgcgccatt cagggcgaca cgggccaag acaagaaggc	7800
aagcgtgacg acaactcctt gccgaatggc gaagtgtccg gcgagaacga tttgtcgata	7860
cgtatgaaac ccaccgagct gcagcccagc agcggcacgc tcttgccgca ccccgatctt	7920
atggtaacgc aagggacaat aattccgtgc atcttgcaaa ccgcaatcga cacaaatttg	7980
gcaggctatg taaagtgtgt cttgcctcag gatattcgtg gaacaacgaa caatatcgtg	8040
cttcttgatc gtggcaccac cgttggtggc gaaatacagc gtggcttgca acaggagat	8100
gggcgcgttt ttgtgttggt ggatcgcgcc gagacacctg accatgcgat gatctcgta	8160
acatcgccaa gcgcggacga actcggtcgc tcaggattgc cgggctcggc cgacagccac	8220
ttctggcagc gtttttagcg agctatgctc ttgagtgttg ttcaaggcgc cttccaggca	8280
gctagcacct acgcggcag ctcgggtggc gggatgagct tcaacagctt tcaaaataac	8340
ggtgagcaga caactgagac agcccttaag gcaaccatca acataccgcc aaccctgaag	8400

ES 2 843 556 T3

aagaatcagg gtgacaccgt ttccattttc gtagcacggg acctcgattt ctttggtgtt	8460
taccagctcc gcctgactgg cggcgccacg cgggggagga accgccgctc ttaatgaatt	8520
caaatttccg cttagagata ggatacattg taaatggaag tggatccgca actacgcttt	8580
cttctgaagc cgattttgga atggctcgat gaccgaaga ctgaagaaat tgcgatcaat	8640
cgacctggag aggcatattgt gcgccaagcc ggcattttta ccaagatgcc tttgccgctc	8700
tcttatgatg atcttgaaga tatcgctatt ttagcgggcg cgctgagaaa gcaggatgtc	8760
ggaccacgta acccctctg cgccactgaa cttcctgggtg gtgaacgact acaaactctgt	8820
ctgcgcgcca ccgttccctc gggcaccgtc agcttgacca ttcgacggcc aagctcgct	8880
gtttctggtc ttaaagaagt ctctcccggt tatgatgctt cgagggtgaa ccagtggcag	8940
acacgaagga aacgcaaaaa tcaggatgat gaagctatcc ttcagcattt tgacaacggg	9000
gatttggaag cgtttctgca cgcattgcgtc gtcagccgac tgacgatgtt gctatgtggc	9060
cctaccggaa gcggcaagac aacaatgagc aagacctga tcagcgccat cccccccag	9120
gaaaggctaa tcaccataga agatacgctc gaactcgta ttccacacga taatcatgtt	9180
agactactct actccaagaa cgggtgctggg ctgggtgctg tgagcgccga gcacttgctc	9240
caagcaagtc tgcgtatgcg gccggaccgg atattgcttg gcgagatgcg cgacgatgca	9300
gcatgggctt atctgagtga agtcgtctcg ggacatccgg gatcgatttc aacaatacac	9360
ggcgcgaaac caatccaagg attcaagaaa ctgttttccc ttgtcaaaag tagcgcccaa	9420
ggtgctagct tggaagatcg cactctgatt gacatgctct ctacggcgat cgatgtcatc	9480
attccattcc gtgcctatga ggacgtttat gaagtagggc agatctggct cgcggcggac	9540
gcacgacgcc gggcgagac cataggcgat ctcttaatc aatagtagct gtaacctcga	9600
agcgtttcac ttgtaacaac gattgagaac ttttgcata aaattgaaat acttggttcg	9660
cattttcgtc atccgcggtc agccgcaatt ctgacgaact gccatttag ctggagatga	9720
ttgtacatcc ttcacgtgaa aatttctcaa gcgctgtgaa caagggttca gatttttagat	9780
tgagaggtga gccgttgaaa cacgttcttc ttatcgatga cgatgtcgct atgcggcatc	9840
ttattatcga ataccttacg atccacgcct tcaaagtgc cgcggtagcc gacagcacc	9900
agttcactag agtactctct tccgcgacgg tcgatgtcgt ggttggtgat cttaaatttag	9960
gtcgtgaaga tgggcttgag atcgttcgaa atctggcggc aaagtctgat attccaatca	10020
taattatcag tggcgaccgc cttgaggaga cgataaagt tgttgactc gagctaggag	10080
caagtgattt tatcgctaag ccgttttagta cgagagagtt tcttgacgc attcggttg	10140
ccttgcgctg gcgccccaac gttgtccgct ccaaagaccg acggtctttt tgttttactg	10200
actggacact taatctcagg caacgtcgct tgatgtccga agctggcggg gaggtgaaac	10260

ES 2 843 556 T3

ttacggcagg	tgagttcaat	cttctcctcg	cgtttttaga	gaaaccccg	gacgttctat	10320
cgcgcgagca	acttctcatt	gccagtcgag	tacgcgacga	ggaggtttac	gacaggagta	10380
tagatgttct	cattttgctg	ctgcgcgcga	aacttgaggc	ggatccgtca	agccctcaac	10440
tgataaaaa	agcaagaggt	gccggttatt	tctttgacgc	ggacgtgcag	gtttcgcaag	10500
gggggacgat	ggcagcctga	gccaattgca	tttgattaat	ttaggtgact	gaggacgcgg	10560
ccagcggcct	caaacctaca	ctcaatattt	ggtgaggggt	tccgataggt	ccctcttcac	10620
caattgctcg	atggtctctc	tccagcaaa	aatgacgcga	gcgcggcggt	agccagcttg	10680
tgcccgaaag	ctcgagcggt	ctccaacccc	aacggatcaa	aatgacttcg	agcgacctcg	10740
agcaacgcaa	ccgggaacat	gcgtgaggtc	tgaacgagaa	cggatttttc	tgtagttgaa	10800
gggatcggat	aacttttcg	ggccacgcga	aatgatccat	ctgccagcat	gctttcgaaa	10860
tcgtccaacg	cgcgccctaa	aatcatattg	agcgacttcg	agggactgta	ttgccgaacg	10920
aggttgtcat	atgttttcga	cacttgaggc	gcgggcggtc	gcgctgaaag	aaaaacctgg	10980
agcttttttcg	gggacggagg	tggactaagg	gcatccacag	ttagcttaag	ttgtcgatcg	11040
ggactgtaaa	tgtgatcggc	gacgagaggc	tcacgttgct	ggtctttctc	gtcggctttt	11100
tcaggcaagt	gctggaggtc	cagcttctgg	ggaacaagtg	tcgggttggg	atggtggatc	11160
tcgggtcgag	caccagcaag	ccgccgtgct	tcgccgaccg	acaatgcggg	cttgccaatt	11220
gccatcttca	agcctccaag	attttgctga	tcagtttcga	aatgaccacg	acttcctcca	11280
tcgcaatccg	aagattcctc	tctatgaggc	gcatcgtcgg	atcagttccc	gtgttttagta	11340
atgtaagatg	caacatgccg	cgttctttca	tcgcggcaaa	tgcatctctt	tcatgcatgg	11400
gagacggtac	aactggaagg	ctctctagcg	tctctgacat	cctgcgttgc	gatgttgtca	11460
atcgggccgac	cgggacgcgt	tggcgcaaaa	cagctgtagg	aattgccaaa	ttttcactca	11520
acagcagctc	gatgacgtag	cggtaggtag	atagtgcctc	atcgatgtcg	agcggcggtta	11580
gcatgggtggg	gatcagaagc	aggtttgagc	tagcgatgat	tgtgttggtg	agctcgcctcg	11640
agccgccacg	cgtatcggcc	aacgcataat	caaatacctc	gagctcggca	ttttcatagg	11700
ctgcttcaag	aaggggcatt	tcgtcggcgg	aatagacttc	acagcgagga	tcccaggtac	11760
tgctttgtaa	ggcgttttct	ctccatcgcg	tcagaggccg	gttttcgctg	gcatcaaaga	11820
gggccactcg	tttaccgtca	tttgccaaa	cagcgcaaa	gcccatgagt	gcggtggttt	11880
tgccagcacc	ccctttgaaa	gaacaaaacg	tcaaaaagtg	catattctga	tccgcctat	11940
cctgtgaaac	cggagtgcac	ttgtattttt	gttcgtataa	atgtttttgt	gattatcgat	12000
gagtaaaagc	gttggttacac	tatttttatt	tcacattcgt	tataagacaa	ttgcaaatgt	12060
agcaagtata	ttcagtattg	actgtaaatg	tactgttgat	ttcatattga	gcagggctag	12120
acttccatcc	gtctaccggg	gcacatttcg	tgctggagta	tccagacctt	ccgctttctt	12180

ES 2 843 556 T3

tggaggaagc tatgtcaaaa cacaccagag ccacgtcgag tgagactacc atcaaccagc	12240
atcgatccct gaaagttgaa gggttcaagg tctgtagtgcc ccgtctgcga tcggccgagt	12300
atgaaacctt ttcctatcaa gcgcgcctgc tgggactttc ggatagtatg gcaattcgcg	12360
ttgcggtgcg ccgcatcggg ggctttctcg aaatagatgc agacacacga gaaaagatgg	12420
aagccatact tcagtccatc ggaatactct caagcaatgt atccatgctt ctatctgcct	12480
acgccgaaga ccctcgatcg gatctggagg ctgtgcgaga tgaacgtatt gcttttggcg	12540
aggctttcgc cgcctcgat ggactactgc gctccatttt gtccgtatcc cggcgacgga	12600
tcgacggtcg ctcggtactg aaaggtgcct tgtagcactt gaccacgcac ctgacgggag	12660
aaaattggat gcccgatcgc gctcaagtaa tcattcgcat tgtgccagga ggtggaacca	12720
agacccttca gcagataatc aatcagctgg agtacctgtc ccgaaaggga aagctggaac	12780
tgacggttc agcccgcat ctcgatattc ccgttcggcc ggatcaaatc cgtgagcttg	12840
cccaaagctg gggttacggg gccgggattt atgacgaaag tcagtcagac gatgacaggc	12900
aacaagactt aacaacacac attattgtaa gcttccccgc aggtaccgac caaaccgcag	12960
cttatgaagc aagccgggaa tgggcagccg agatgtttgg gtcaggatac gggggtggcc	13020
gctataacta tctgacagcc taccacgtcg acccgcatca tccacattta catgtcgtgg	13080
tcaatcgtcg ggaacttctg gggcaggggt ggctgaaaat atccaggcgc catccccagc	13140
tgaattatga cggcttacgg aaaaagatgg cagagatttc acttcgtcac ggcatagtcc	13200
tggatgcgac ttcgcgagca gaaaggggaa tagcagagcg accaatcaca tatgctgaat	13260
atcgacgcct tgagcggatg caggctcaaa agattcaatt cgaagataca gattttgatg	13320
agacctcgcc tgaggaagat cgtcgggacc tcagtcaatc gttcgatcca tttcgatcgg	13380
acgcactcgc cggcgaaacc gaccgtgcaa cccgacatga caaacaaccg cttgaaccgc	13440
acgcccgttt ccaggagccc gccggctcca gcatcaaagc cgacgcacgg atccgcgtac	13500
cattggagag cgagcgggggt gcccaaccat ccgcgtccaa aatccctgta actgggcatt	13560
tcgggattga gacttcgtat gtcgctgaag ccagcgtgcc caaacaagc ggcaattccg	13620
atacttctcg cccggtgact gacgttgcca tgcacacagt cgagcgccag cagcgatcaa	13680
aacgacgtca tgacgaggag gcaggtccga gcggagcaaa ccgtaaaaga ttgaaggccg	13740
cgcaagttag ttccgaggca aatgtcgtg agcccgacgg tcgcatgac agcaacaagg	13800
cggctgatcc ggtgtctgct tccatccgta ccgagcaacc ggaagcttct ccaacgtgtc	13860
cgcgtgaccg tcacgatgga gaattgggag aacgcaaacg tgcaagaggt aatcgtcgcg	13920
acgatgggcg cgggggggacc tagactagta gacaggaagg accgaataat ggcaaatggt	13980
cagttcacga tacgctctgc tcgcccggcc tccgtcggac tgacaggcga acggcgtgga	14040

ES 2 843 556 T3

gccgcacccg	cctctagctc	tgcactgtcc	aatgttcaaa	gagatgttag	ggataggtcg	14100
attccaacta	gctcaccaag	attaccaa	gcagccatat	tgcgtgattc	ctcggaaga	14160
gcgtcgactg	gtctgcggta	catggcggt	actcttcatt	ggtctcgat	cgccatta	14220
tcgctaataa	acagcaacga	cctggctccg	gccgttatg	actttgagac	gcgaaataac	14280
gcaagaaatg	tgactgccaa	agtcggcagg	gcagtcctg	ttccaagca	aggcggtc	14340
ggcaaaacgc	tcgcacccgt	acccttagt	acacgtatat	caagggtcaa	ttccgaccga	14400
agactgcccg	ctgacgcaga	agaccgccct	gaaacgcgcg	acccccagaa	aggacgtggc	14460
agtcattggtg	cgacgccaac	cttacatgaa	aagattggaa	ccgcgtttgc	tcgaagattg	14520
cgaagcata	cgtactatat	tgtttgcagt	tgctgccaga	cccgagcgc	gttgacgatg	14580
ggtgcaaaga	tttcggtgaa	gtcatgaact	ccagcaagac	ttcgccccag	cgtatgaccc	14640
tgagcatcgt	atgttcgctg	gcagccggtt	tttgtcggc	cagctgctat	gtaacgttcc	14700
gccggggctt	caacggcgaa	gcgatgatga	cgttcgacgt	tttcgctttt	tggtatgaga	14760
ccccgcttta	cttgggttat	gccagcaccg	tcttctggcg	tggtttatct	gttgtcatct	14820
ttacctcgct	gatcgttctt	tcaagtcagc	tcatcatatc	gctgcgcaat	cagaagcatc	14880
atgggacagc	tcgttgggca	gaaattggcg	aaatcgggca	tgctggttat	ctgcagcgtt	14940
acagtcgcat	caaggggccg	atctttggaa	agacatgtgg	tcccctttgg	ttcggcagtt	15000
atttgaccaa	tggcgaacag	ccacacagtc	ttgtcgtcgc	gccaacgcgt	gctggcaaag	15060
gcgtcggcat	cgtcattcca	acgctgttga	ccttcaaggg	ctcggtaatc	gcccttgacg	15120
tcaagggaga	attgtttgaa	ctgacgtcca	gagcacgcaa	agcgagcggc	gacgcagttt	15180
tcaagttctc	ccccctagat	cctgagcgga	agactcattg	ttacaatccg	gtcctggata	15240
ttgccgcact	tccgccgaa	cgccagttca	ctgaaacacg	ccgtctagct	gcgaacctta	15300
ttacggctaa	gggaaaggga	gcagaaggct	ttattgacgg	cgcacgtgac	ctgttcgtcg	15360
cgggaatcct	tacctgcatt	gagcgtggca	caccaacgat	tggcgcggta	tatgacctat	15420
ttgcgcagcc	tggcgaaaag	tataagcttt	ttgcgcaact	cgcggaggaa	agcctaaca	15480
aagaggctca	gcgtatcttc	gataatatgg	cgggcaacga	cacgaaaatt	ctgacatcgt	15540
acacctctgt	gctgggcgac	ggtggactga	acctgtgggc	tgatccgctt	atcaaagcag	15600
cgacaagccg	gtcagacttt	tccgtttacg	atctccggag	gaagaagacc	tgcatttatc	15660
tttgtgtcag	tcccaacgat	ctggaggtct	tggaaccact	tatgcgcctg	atgtttcagc	15720
agctcgtgtc	aatcttgacg	agatcgctgc	caggtgaaga	cagtgccat	gaagttttat	15780
ttctcctcga	cgaattcaaa	cacctgggca	agcttgaggc	catagagacc	gcgatcacia	15840
ccatcgcccg	ttacagaggc	cgttttatgt	ttattattca	aagtctttcg	gccttgctcg	15900
gcacatacga	tgacgcagga	aaacaaaact	ttctgagcaa	tactggcgta	caagtattta	15960

ES 2 843 556 T3

tggccacggc tgatgacgaa actccaacct acatctcaaa agctatcggc gaatatacgt	16020
ttaaagcgcg ttcgacctct tacagtcaag ccagaatgtt cgaccacaac atccagattt	16080
ctgatcaagg tgcagccctt ttgcgccccg aacaagtgcg cctgctagac gatcagagtg	16140
aaatcgttct catcaaaggg cgacctccac tcaaattacg aaagggtgcag tattattccg	16200
atcgtagcgt gaaaggcctt ttcgaacgcc agatgggctc tctgcctgag cccgcaccct	16260
tgatgctttc cgactatagc aacgatcaag ttcaatacca cttggctccg atagcaaatt	16320
ttaatgagga tgctgcaccg caaaacagaa ctgtggccga ggaccatgga agtgttaaag	16380
tcggtgctga tatccctgaa cgcgtgatgg gaataaatgg tgacgaggaa caagccgatg	16440
cgggcgagat accgccgaa tcggttgtgc ctccagaatt gacgctcgct ctgaccgctc	16500
aacagcaatt gttggaccag attattgcac ttcagcaaag atcgaggtcc gcaccggcat	16560
agcctgcgaa atgatcttaa tgggtgcaatt cgtttcaggc ggcattgtgc ggttcaacaa	16620
agtgtacgcc agatcccaat tggcccttta cgagggtgcg gtatgacagg aaagtcgaaa	16680
gttcacataa gaggttcggc tgacgcgctt cctgacgttc ctggcggaag tactaccgcc	16740
ccttttttaa ccgaaccttc tcgggatcag gttgatgcct cgtttgaggt ccaaaccgac	16800
tacagccagt ctacttcogt gtcgtttacc tatgatgggtg ttggacttgg tcctgccgag	16860
cgtgcccgtt acgagaactg gtgcgaaccg ggccggccca cttggaaaga tcttataatc	16920
aaggcacgtg tcgatccgat tgacgatgtg acctggctcc gagatttaga agaggacacc	16980
ccctcaacct tcagatacga agggatgcct ctgggcatcg gggaacgaca ggcctacgaa	17040
aattggcaag aggacgctca gccgacatgg gaagaccttg ttgtcagcgc acgcttgacg	17100
gaacttggcc gtccacacgg gattaccggc gagtatacat ccctcgcagg atcgaagaat	17160
acaagttcaa tttcattgaa gcgaaagcgg agcaacttaa ttgatgatga gaattcatcc	17220
ggatcgtttt catatgacgg gatgaagctc ggggaagccg agcgttctgc atatggtgac	17280
tgggcccagg cggagccacc cacgtggaaa gatttggtat tgagggcacg cgtttcctcg	17340
atcaatgact ctgcttggct ttttgattca caaacatctt catcatcatt tgaatacaac	17400
ggtgttccct tgggcgagcc ggaacggcag gctctcagac aatggcaagg agacgctcag	17460
cctacctggg aagatctcgt tgtaacgcg cgtatggcag aactttgcca tgctggttgg	17520
attgaaggct aaaaagggtg ctttgaagag cgcggggagg ctctgcccgc gtcggaacgc	17580
ggttcgcaac gcccattgg tcaacggaca gattcctccg attcctttgt gtatgatggc	17640
acaaggctcg gagcacctga gcgaactgct tatgaacgct ggagtaagag ggaacgcccg	17700
acttgggaag atctcatctt agatgcacac caggccagga ctgaaagtga cgctgttacg	17760
acccaagcga ttggctcagtc gtcctcaccg gttttcttat atgaaggaaa gtcgctcgga	17820

ES 2 843 556 T3

gacaggggaac gaaaggctta cgaaaaatgg cggcagccag cccaaccgcg atggcaaaat	17880
cttgtagtga acgctcgtct ggcagaaatc gatccctcag cctggattgc cgatgagcgc	17940
gatccgcttg atgatagcga cgcgcttggt cgcccgctcg acacaagctt gacggataga	18000
tcagacgtcc ctttagacga tcaatcaatc tatcgtcgtt ccgacctagt aagggagcag	18060
gtgccagaat cgtctcaaag gcaattcgca gcatgttcag aatctgaaac gaggcctgtg	18120
caatggttta ctgcttcttg gtcagatgca aacaatacgg aaaatatcac cgccagcgat	18180
cccgctcgatc gcacgggttg agttaagcgg ctaggctcca aaagcgacag aaccgttaca	18240
gcttctatcc atgacgtgaa ttccagcaca aggcgactgt tgcttaacga atttggatcg	18300
gaggctccgc gcccttcgcc agaaaagact gttcgtttaa gaagcgacaa tattggcacc	18360
tatgggagcc ggaaaaatga acgagcgcg ctcgcgaccg aaaccggtgc gtatgagtcg	18420
gagcatatct tcgggttcaa ggctgtccac gatactgcga gagcgacgaa agagggccgg	18480
cgtctcgaaa ggcccatgcc cgcctacctt gaggataagg ggcttcatcg ccaacatatt	18540
ggcaccggga gaggacggac caaacttgct gggcgcggtt ggccggatga cacaagctat	18600
cgctcggatc aaagggcaac tctgtcggac ccggttgcgc gctcgggaagg cgcgacggcc	18660
tcaaatgggt atcaattgaa ccaattgggt tacgcgcacc aactcgctag cgatggcctg	18720
caaagtgaat cgcccgatgg tgttgccctt ccaattcaag tggcaacaac gagctacaac	18780
tatacagtga gcccgcatcc tgtccttggt ccgccggata aaaacgaagc ccctcaattg	18840
ctgcatcttg gtccccgttg tcaaacgaa gctgttcttg ccgcgaaac agcattgact	18900
ggaaaatggc cgactctcga gcgtagcag caagtgtatc gcgagttttt ggccctatat	18960
gacgtaaaaa aagatcttga ggccaaatca gtcggcgtaa gacggaaaaa aaaagaagtt	19020
atttctgcgt tagaccgaac tgcgcgcttg ataagcacgt cgccttcgaa agctcgatcc	19080
aaagcagaga ctgaaaaagc cattgatgag ctcgatgatc gacgagttta tgatccgctg	19140
gatcgagctc aagacaaagc gtttaaacgc tgataagtcg ccaatatagt gatcatttgc	19200
agtattcgca tcgatcgctg gttgatattc tgccgctgggt cgaccggctg ctcgctcgcca	19260
aaaatgctca cagggatacg atggcctcgg tcaggccgcg tgcgctcctgt ctttccagtt	19320
cctccctttc agctcgattg tggcatcatt tattgcctgc tcattgcagt tgaaacgcga	19380
tatccgtttc aagaccggg tatggatggt actttggaga tatgagcact tgccagactt	19440
cccggtcggg gacttaaaaa ttccgataag aggttacgga cgacggacct tataagtggg	19500
tcgtctagtg gtggcgccga tcaaaacagt tccgccccg gctttgctca aagtagcaaa	19560
gcagctttat ctcggttgc ggaggatttt ctagaaaacc gcaattttgc aggagagaac	19620
atatggccat catcaagccg catgtgaaca aaaataggac aacctcgccg atagagagac	19680
cggagtctct catagaggaa atgagcggca gtcacccgcc gagtggtttt accaacctgg	19740

ES 2 843 556 T3

atctcgctat gatcgagctg gaggaactttg tccatcggtg cccgctccca gaagacaatc	19800
ttgctgggtca gaaggagtga gacgatggat ccgtctagca atgagaatgt ctatgtgggt	19860
cgcggtcaca acatcgaaaa tgatgatgac actgacccca ggcgttgga gaaggcgaat	19920
atcagttcca acaccatctc cgatattcag atgacgaatg gcgaagacgt acaatcaggg	19980
agccctaccc gaacggaagt tgtaagccca cgtctggatt atggatcggc cgactcctcc	20040
tccagccttt attctggcag cgagcacgga aatcaagctg agattcaaaa agagctgtcc	20100
gtcttggtct cgaacatgtc tttgccaggc aacgatcggc gcccgacga atacattctc	20160
gtgcatcaaa cgggacaaga tgcttttact ggtattgcca aaggcaacct cgaccaaag	20220
cccaccaagg cggaatttaa cgcgtgctgc cgtctctaca gggacggagc cggtaattac	20280
taccgccac ctctcgctt cgacaagatt agcgttccag agcaactgga ggaaaaatgg	20340
gggatgatgg aggcgaagga acgtaacaaa ctgcggtttc agtacaagtt ggacgtatgg	20400
aatcatgcgc acgctgatat ggggatcacg ggcacagaga tcttttatca aacagataag	20460
aacataaagc tcgaccgga ttataaacta agacctgaag accgatacgt acaaacagaa	20520
aaatacgggc gccgggaaat tcaaaagcga tatcaacacg aactccaggc tggttcgctg	20580
ctgcccagata ttatgatcaa aactcccaa aatgacatcc acttcgtgta caggtttgcc	20640
ggcgacaatt acgccaacaa acagttcagc gagtttgaa acaccgtcaa gcgcaggat	20700
ggcgacgaga ctgagatcaa attgaagtca aagtcaggca ttatgcatga ctcgaaatat	20760
ctggaatcct gggaacgggg cagtgcggat attcgcttcg cggaattcgt tggggaaaat	20820
agagctcaca atcggcagtt tccaactgcg acagtaaata tgggacagca gccagacggg	20880
cagggcgggt tgacccgga ccgtcatgtg agcgttgact tcctaatagca aagcgcaccc	20940
aattcgctt gggcgcaagc tttgaaaaag ggagaactgt gggatcgcgt tcagttgctt	21000
gctcgcgacg gcaaccgcta tctgtcgccg ccagattgg aatattctga ccctgcacat	21060
ttcaccgagt tgatgaaccg ggttggttta cccgcatcga tgggtcggca aagccatgcg	21120
gctagtatca aattcgaaaa gtttgacgcg caggcagcgg ttattgtctt aaatggcca	21180
gagttacgtg acattcatga ctgtctcct gaaaaactgc aaaatttgtc caccaaagat	21240
gtcatcgtcg ccgatcgcaa tgagaatggc cagagaactg gcacgtacac cagcgtcgcg	21300
gaatatgagc gcttgcaagt aaggctgcca cccgatgcag cgggggtgct tggatgaagca	21360
actgacaaat attcacgtga tttcgttcg ccagagccgg cgtcgctcc aatcagtgac	21420
agccgcagga tatacgaaag tcgaccgcgt agccaaagcg tcaacagctt ttgacgttcc	21480
tgctgccgcg tcaacgagga agctcgtttg acccggttt gccaatgaaa gggctcaatc	21540
atggtgaaca ctacaaagaa aagttttgcg aagtcgctta cggcagatat gcgccgttct	21600

ES 2 843 556 T3

gctcagcgcg	ttgtcgagca	aatgcgaaaa	gcattgatta	ccgaagaaga	ggcgctcaag	21660
cggcaagcca	gactggagag	tcccgatagg	aagcgaaagt	atgctgctga	tatggcgata	21720
gtcgacaaac	tcgacgtagg	gtttcgaggc	gaaataggct	ataaaattct	tggaaataac	21780
cggcttcgag	tagacaacca	taaagaatta	acgcgtgagc	acggtagact	tcgcaaaacc	21840
aaaacggttc	tgaagcgtaa	cccggtgacg	caggaagtct	at ttgggttt	atatgaaagg	21900
aagtcctggt	taagtgtcag	cagccatttg	tatgctgctg	acggcacact	ccgcatgaag	21960
cacgtgaaat	acaaagacgg	acgttttgag	gaaaaatggg	agcgcgacga	aaatggcgac	22020
ctgatccgca	caaggtacgc	caaccgtggc	aggctctttc	aacctgtatc	cgagaaaatg	22080
ggcgcgccgt	atcggagcgg	ccctgacgac	cggctctatc	gcgatctaac	ccgtcgaaac	22140
ggtttcagac	gggagacatt	cgaacgggac	gatcacggaa	acctcgagcg	tatcggcagc	22200
aaccatgtcg	gcttttccaa	gatttcagtg	aaggcaccca	atcgtcaaac	ctcccagacg	22260
aagattcaaa	aacttggtgg	cgctttcaac	aaatctttta	ggtcccttct	ggacaaggag	22320
ggcaatgaaa	tgggccgcga	tattttgagc	catcgacggc	tctataacaa	gcggtctgct	22380
gtctacgatg	aagctaccgg	acaattgaag	agtccaagc	ataccttcgg	caagatctac	22440
aggagcgaaa	ccgaatatct	cagcgcgggc	ctcaagaagg	tttcaaaaaa	gatactcggg	22500
gtgacgggtc	accggaaatt	tgcgcgctc	agcgagcgag	aatccgaggc	tgagagactg	22560
cgtagttttg	aatccgggtg	gcacgcgcag	atctggcagg	agcgggcagc	gattcccggg	22620
tcgcccctcc	cggagactga	tgacattcat	ttcgcacagc	agtcgcacct	agccaaagcc	22680
aacctgtatc	acgtcgaagc	tgacgtcacg	cgtgtgacag	atcaacatgt	tgatgttgct	22740
ggacaaacat	catcgtctcc	ccaacggaac	ttggaaggat	ggttagattc	tcaatcacga	22800
tacaagccag	caaacatgct	gttgtcaaat	ccagaccttc	aagcgaaagg	acctcgccca	22860
tacgaagggt	tagctcatct	caccctccgg	cgcgataatg	aatctgacgg	gcacaaggag	22920
aatgatcagc	ggctcgacac	tttctcccag	ccagagccgt	tgggtgttacc	gcaccccggg	22980
tcgcccggaaa	taactaagg	gtttggctcg	cggggagagc	cggcacaccc	gagtggaaat	23040
ctgcacacgg	cgggtggaga	aacggcttgc	gaaggaccgg	tgatgtcttc	atcctcggac	23100
aatcatcagc	cagctccagg	acagcaagaa	cttttaagtt	tccttcataa	tgcgccagcc	23160
ccagtttctg	tggcaataca	tgatgatcaa	gagcgacttg	cgggggaggg	gcccgccggc	23220
tctttcagag	gtagctcagg	gcgaacgagt	tcaatgtcgg	agagtatcct	cgacgaagat	23280
gtacaagggc	atttgggtacg	ggattattcg	atcaatacta	ctaacggggt	tattgaccgg	23340
caatcgttgt	tcggtgaacc	ggacttatcg	agaggtccaa	aatcggggcc	agaaattcca	23400
tcggaagatt	accatttgtc	agcttcggaa	caggaaaatt	tgctgaatca	attgcttagt	23460
gtgccactgc	cggttccttc	accgaagccc	gaatgcgcga	ggtctatgat	tttcgaagg	23520

ES 2 843 556 T3

tcacgttcaa gagagcgttc cacctccaga gggttctaag gtggaacgcg tctaggcttg	23580
cctatcatca ccacctcatg atggtcgcgg agaccttcga agagatatcc gacatcattg	23640
gagaagctgg gccggaaact agtgtacctc gcgaatgcat ctagatccaa tccaatatgc	23700
atggcatgcc gaatccatgt gggagtttat tcttgacaca gatatttatg atataataac	23760
tgagtaagct taacataagg aggaaaaaca tatgttacgc agcagcaacg atgttacgca	23820
gcagggcagt cgccctaaaa caaagtttagg tggctcaagt atgggcatca ttcgcacatg	23880
taggctcggc cctgaccaag tcaaatccat gcgggctgct cttgatcttt tcggtcgtga	23940
gttcggagac gtagccacct actcccaaca tcagccggac tccgattacc tcgggaactt	24000
gctccgtagt aagacattca tcgcgcttgc tgccttcgac caagaagcgg ttgttggcgc	24060
tctcgcggct tacgtttctgc ccaagtttga gcagccgcgt agtgagatct atatctatga	24120
tctcgcagtc tccggcgagc accggaggca gggcattgcc accgcgctca tcaatctcct	24180
caagcatgag gccaacgcgc ttggtgctta tgtgatctac gtgcaagcag attacgggtga	24240
cgatcccga gtggctctct atacaaagt gggcatacgg gaagaagtga tgcactttga	24300
tatcgaccca agtaccgcca cctaataatg tctaacaatt cgttcaagcc gacgcgctt	24360
cgcggcgcgg cttaactcaa gcgttagatg cactatacgt accaatccaa atcggatccc	24420
cctgcagggc tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt	24480
actcatatat acttttagatt gatttaaaac ttcattttta atttaaaagg atctagggtga	24540
agatcctttt tgataatctc atgacccaaaa tcccttaacg tgagttttcg ttccactgag	24600
cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa	24660
tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccggatcaag	24720
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg	24780
tccttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat	24840
acctcgctct gctaatacctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta	24900
ccgggttgga ctcaagacga tagttaccgg ataagcgca gcggtcgggc tgaacggggg	24960
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc	25020
gtgagctatg agaaagcgcc acgcttcccg aaggagaaaa ggccgacagg tatccggtaa	25080
gcggcagggc cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggtatc	25140
tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt	25200
cagggggggc gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttccctggcct	25260
tttgctggcc ttttgctcac atgttctttc ctgcgttatc ccctgattct gtggataacc	25320
gtattaccgc ctttgagtga gctgataccg ctgcgcgag ccgaacgacc gagcgcagcg	25380

ES 2 843 556 T3

agtcagt	gag	cgaggaagcg	gaagagcgcc	tgatgcggta	ttttctcctt	acgcatctgt	25440
gcggatattt	acaccgcata	tggtgcactc	tcagtacaat	ctgctctgat	gccgcatagt		25500
taagccagta	tacactccgc	tatcgctacg	tgactgggtc	atggctgcgc	cccgaacccc		25560
gccaacaccc	gctgacgcgc	cctgacgggc	ttgtctgctc	ccggcatccg	cttacagaca		25620
agctgtgacc	gtctccggga	gctgcatgtg	tcagaggttt	tcaccgtcat	caccgaaacg		25680
cgcgaggcag	cagatctatc	gatggtacct	atctgcagac	ggaagggatt	cctgttaatt		25740
aaaacagctt	gcgtcatgcg	gtcgtctgct	atatgatgcg	atgagtaaat	aaacaaatac		25800
gcaaggggaa	cgcatgaagg	ttatcgctgt	acttaaccag	aaaggcgggt	caggcaagac		25860
gaccatcgca	acccatctag	cccgcgccct	gcaactcgcc	ggggccgatg	ttctgttagt		25920
cgattccgat	cccaggggca	gtgcccgcga	ttgggcggcc	gtgcgggaag	atcaaccgct		25980
aaccgtttgc	ggcatcgacc	gcccgcacgat	tgaccgcgac	gtgaaggcca	tcggccggcg		26040
cgacttcgta	gtgatcgacg	gagcgcccca	ggcggcggac	ttggctgtgt	ccgcgatcaa		26100
ggcagccgac	ttcgtgctga	ttccggtgca	gccaaagccct	tacgacatat	gggccaccgc		26160
cgacctggtg	gagctgggta	agcagcgcat	tgaggtcacg	gatggaaggc	tacaagcggc		26220
ctttgtcgtg	tcgcggggcga	tcaaaggcac	gcgcacgggc	ggtgagggtg	ccgaggcgct		26280
ggccgggtac	gagctgcccc	ttcttgagtc	ccgtatcacg	cagcgctga	gctaccagg		26340
cactgccgcc	gccggcacaa	ccgttcttga	atcagaaccc	gagggcgacg	ctgccgcga		26400
ggtccaggcg	ctggccgctg	aaattaaatc	aaaactcatt	tgagttaatg	aggtaaagag		26460
aaaatgagca	aaagcacaaa	cacgctaagt	gccggccgtc	cgagcgacg	cagcagcaag		26520
gctgcaacgt	tgccagcct	cgcagacacg	ccagccatga	agcgggtcaa	ctttcagttg		26580
ccggcgggag	atcacaccaa	gctgaagatg	tacgcggtac	gccaaaggcaa	gaccattacc		26640
gagctgctat	ctgaatacat	cgcgcagcta	ccagagtaaa	tgagcaaatg	aataaatgag		26700
tagatgaatt	ttagcggcta	aaggaggcgg	catgaaaaat	caagaacaac	caggcaccga		26760
cgccgtggaa	tgcccatgt	gtggagggaac	gggcggttg	ccaggcgtaa	gcggctgggt		26820
tgcttgcgg	ccctgcaatg	gcactggaac	ccccaaagccc	gaggaatcgg	cgtgagcgg		26880
cgcaaaccat	ccggcccgg	acaaatcggc	gcggcgctgg	gtgatgacct	ggtggagaag		26940
ttgaaggcag	cgcaggccgc	ccagcggcaa	cgcacgagg	cagaagcacg	ccccggtgaa		27000
tcgtggcaag	cagccgctga	tcgaatccgc	aaagaatccc	ggcaaccgcc	ggcagccgg		27060
gcgcgctcga	ttaggaagcc	gcccagggc	gacgagcaac	cagatttttt	cgttccgatg		27120
ctctatgacg	tgggcaccgc	cgatagtcgc	agcatcatgg	acgtggccgt	tttccgtctg		27180
tcgaagcgtg	accgacgagc	tggcgagggt	atccgctacg	agcttccaga	cgggcacgta		27240
gaggtttccg	cagggccggc	aggcatggcc	agtggtggg	attacgacct	ggtactgatg		27300

ES 2 843 556 T3

gcggtttccc atctaaccga atccatgaac cgataccggg aaggggaagg agacaagccc 27360
ggccgcgctgt tccgtccaca cgttgcggac gtactcaagt tctgccggcg agccgatggc 27420
ggaaagcaga aagacgacct ggtagaaacc tgcattcggg taaacaccac gcacgttgcc 27480
atgcagcgta cgaagaaggc caagaacggc cgcctgggta cggatatccga gggatgaagg 27540
ttgattagcc gctacaagat cgtaaaagagc gaaaccgggc ggccggagta catcgagatc 27600
gagctagctg attggatgta ccgcgagatc acagaaggca agaaccggga cgtgctgacg 27660
gttcaccccg attacttttt gatcgatccc ggcatcggcc gttttctcta ccgcctggca 27720
cgccgcgccc caggcaaggc agaagccaga tgggtgttca agacgatcta cgaacgcagt 27780
ggcagcgccc gagagttcaa gaagtctctgt ttcaccgtgc gcaagctgat cgggtcaaat 27840
gacctgccgg agtacgattt gaaggaggag gcggggcagg ctggcccgat cctagtcattg 27900
cgctaccgca acctgatcga gggcgaagca tccgcgggtt cctaattgtac ggagcagatg 27960
ctagggcaaa ttgccctagc aggggaaaaa ggtcgaaaag gtctctttcc tgtggatagc 28020
acgtacattg ggaacccaaa gccgtacatt gggaaccgga acccgatcat tgggaaccca 28080
aagccgtaca ttgggaaccg gtcacacatg taagtgactg atataaaaga gaaaaaaggc 28140
gatttttccg cctaaaactc tttaaaactt attaaaactc ttaaaacccg cctggcctgt 28200
gcataactgt ctggccagcg cacagccgaa gagctgcaaa aagcgccctac ccttcggtcg 28260
ctgcgctccc tacgccccgc cgcttcgctg cggcctatgc atataattgt gggttcaaaa 28320
tcggctccgt cgatactatg ttatacgcca actctgaaaa caactttgaa aaagctgttt 28380
tctggtattt 28390

<210> 36

<211> 31939

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 36

cgccaacacc cgctgacgcg ccctgacggg cttgtctgct cccggcatcc gcttacagac 60
aagctgtgac cgtctccggg agctgcatgt gtcagagggt ttcaccgtca tcaccgaaac 120
gcgcgaggca gcagatctat cgatggtacc tatctgcaga cggaagggat tcctgttaat 180
taaaacagct tgcgtcatgc ggtcgtctgc tatatgatgc gatgagtaa taaacaaata 240
cgcaagggga acgcatgaag gttatcgctg tacttaacca gaaaggcggg tcaggcaaga 300
cgaccatcgc aaccatcta gcccgcccc tgcaactcgc cggggccgat gttctgttag 360
tcgattccga tcccagggc agtgcccgcg attgggcggc cgtgcgggaa gatcaaccgc 420

10

ES 2 843 556 T3

taaccgttgt cggcatcgac cgcccgacga ttgaccgcga cgtgaaggcc atcggccggc	480
gcgacttcgt agtgatcgac ggagcgcccc aggcggcgga cttggctgtg tccgcgatca	540
aggcagccga cttcgtgctg attccggtgc agccaagccc ttacgacata tgggccaccg	600
ccgacctggt ggagctggtt aagcagcgca ttgaggtcac ggatggaagg ctacaagcgg	660
cctttgtcgt gtcgcgggcg atcaaaggca cgcgcatcgg cggtagaggtt gccgagcgcg	720
tggccgggta cgagctgccc attcttgagt cccgtatcac gcagcgctg agctaccag	780
gcaactgccg cgccggcaca accgttcttg aatcagaacc cgaggcgac gctgcccgcg	840
aggtccaggc gctggccgct gaaattaaat caaaactcat ttgagttaat gaggtaaaaga	900
gaaaatgagc aaaagcacia acacgctaag tgccggccgt ccgagcgcac gcagcagcaa	960
ggctgcaacg ttggccagcc tcgcagacac gccagccatg aagcgggtca actttcagtt	1020
gccggcgag gatcacacca agctgaagat gtacgggta cgccaaggca agaccattac	1080
cgagctgcta tctgaataca tcgcgcagct accagagtaa atgagcaaat gaataaatga	1140
gtagatgaat tttagcggct aaaggaggcg gcatggaaaa tcaagaacaa ccaggcaccg	1200
acgccgtgga atgccccatg tgtggaggaa cgggcgggtt gccagcgta agcggctggg	1260
ttgcctgccg gccctgcaat ggcactggaa cccccaagcc cgaggaatcg gcgtgagcg	1320
tcgcaaacca tccggcccgg tacaatccg cgcgcgctg ggtgatgacc tggtagagaa	1380
gttgaaggca gcgcaggccg cccagcgga acgcacgag gcagaagcac gccccgtga	1440
atcgtggcaa gcagccgctg atcgaatccg caaagaatcc cggcaaccgc cggcagccgg	1500
tgcgcgctg attaggaagc cgcccaagg cgacagcaaa ccagattttt tcgttccgat	1560
gctctatgac gtgggcaccc gcgatagtcg cagcatcatg gacgtggccg ttttccgtct	1620
gtcgaagcgt gaccgacgag ctggcgaggt gatccgctac gagcttccag acgggcacgt	1680
agaggtttcc gcagggcccg caggcatggc cagtgtgtg gattacgacc tggtagat	1740
ggcggtttcc catctaaccg aatccatgaa ccgataccgg gaagggaagg gagacaagcc	1800
cggccgctg ttcgctccac acgttgcgga cgtactcaag ttctgcggc gagccgatgg	1860
cggaaagcag aaagacgacc tggtagaaac ctgcattcgg ttaaaccacca cgcacgttgc	1920
catgcagcgt acgaagaagg ccaagaacgg ccgcctggtg acggtatccg agggtagaagc	1980
cttgattagc cgctacaaga tcgtaaagag cgaaaccggg cggccggagt acatcgagat	2040
cgagctagct gattggatgt accgcgagat cacagaaggc aagaaccgg acgtgctgac	2100
ggttaccgcc gattactttt tgatcgatcc cggcatcggc cgttttctct accgcctggc	2160
acgccgcgcc gcaggcaagg cagaagccag atggttgttc aagacgatct acgaacgcag	2220
tggcagcgcc ggagagttca agaagttctg ttccaccgtg cgcaagctga tcgggtcaaa	2280
tgacctgccg gactacgatt tgaaggagga ggccgggcag gctggccga tcctagtcac	2340

ES 2 843 556 T3

gcgctaccgc aacctgatcg agggcgaagc atccgccggt tcctaattgta cggagcagat	2400
gctagggcaa attgccctag caggggaaaa aggtcgaaaa ggtctctttc ctgtggatag	2460
cacgtacatt gggaacccaa agccgtacat tgggaaccgg aaccgtaca ttgggaaccc	2520
aaagccgtac attgggaacc ggtcacacat gtaagtgact gatataaaag agaaaaaagg	2580
cgatTTTTCC gcctaaaact ctttaaaact tattaaaact cttaaaaccc gcctggcctg	2640
tgcataactg tctggccagc gcacagccga agagctgcaa aaagcgccta cccttcggtc	2700
gctgcgctcc ctacgccccg ccgcttcgcg tggccctatg catataattg tggtttcaa	2760
atcggtccg tcgatactat gttatacgcc aactctgaaa acaactttga aaaagctgtt	2820
ttctggtatt taaatcctag gaaatatcag ccgcttgtgt gttcaattct tctctgtttc	2880
acttgaaaca aattgaatag atattccgc tttcaaagcc atttacaat cctctcgtgc	2940
agcctaacgc cattggcgtg gaacaagcgt tggcacgagg aagtaagtgc gatgaacgga	3000
aggatttcac cgacgcggca ggattttaag acaggcgga agccttggtc tatattagcc	3060
cttatcgttg ccgcaatgat tttcgcttc atggcggttg cgtcctggca ggacaatggg	3120
actaccagc caatccttag ccaaatacga tcgattaacg ctgacagcgc ctcaactgcag	3180
cgcgatgtac tccgcgctca cacgggcacc gtggcgaact accgccccat tatctccagg	3240
ctgggagctc tgcggaagaa tctggaagat ttgaagcaat tatttaggca atctcatatt	3300
gtaagtgaga gcaatgctgc tcaactgcta cgccgctag aagtgtctct aaattcggct	3360
gacgggcg tgcgccctt tgggtgcgcaa aatgtccgcc tgcaagattc gctggccagt	3420
ttcaactcgt ctttaagcaa tcttccagga aaggcctcag gcgatcagac tttagaaaaa	3480
ccaacagaat tggctagcat gatgctcaa tttcttcggc aaccaagccc ggctatttca	3540
ttcgagatca gccttgaact agagaggctc caaaaacaac gcggtcttga tgaagctccc	3600
gtgcgcatac tttcgctga aggtccatt atcttatcgc ttttgccaca ggtgaacgat	3660
ctggtgaaca tgattcagac gtctgacacc gcagaaattg cggaaatgtt gcagcgtgag	3720
tgtttggagg tctatagctt gaaaaatgtg gaggagcgga gcgcacgtat ctttcttggg	3780
tccgcttcag tgggtctttg cctgtacatc atcaccttag tctataggct acgcaaaaaa	3840
accgattggt tagcgcgcg tttagattac gaagagctaa tcaaagagat cggagtatgt	3900
tttgaagggt aggcggccac tacgtcgtcc gcgcaagctg cacttggtat tattcagcgc	3960
ttctttgatg ccgatacgtg cgcgttagct ctagtggacc atgaccgtag gtgggctgtc	4020
gaaacattcg gtgcgaagca cccaaaaccc gtgtgggacg acagggtgct acgcgaaata	4080
gtctctcgta ccaaagcgaa cgaacgggcg acggtattcc gcacgtatc gacgcaaaaa	4140
atcgtagatt tgcctccga aattccaggt ctttcgatac tactggctca caaatccaca	4200

ES 2 843 556 T3

gataaactaa ttgcggtttg ttcactgggt taccaaagct atcgccctcg accttgccaa	4260
ggcgaaattc agcttcttga actcgccacc gcctgcctct gtcactatat cgatgttcgg	4320
cgtaagcaga ccgaatgcga cgttttggcc agacgattgg agcatgcgca acgccttgag	4380
gcagttggta cacttgccgg cggaatagca catgaattta ataacatttt ggggtcaatc	4440
ctcgggcacg cagaattggc ccaaaactcg gtgtctcgaa catctgtcac ccgaagatac	4500
attgactata tcatttcgtc aggcgacaga gccatgctca ttatcgatca gatcttgacg	4560
ctgagccgaa aacaagagcg cgtgatcaag ccatttagtg tctcagagct tgtgaccgaa	4620
atcgctccct tgctacgtat ggctcttccc ccaaccatcg agcttagttt cagatttgat	4680
caaatgcaga gcgtgatcga aggaagcccg cttgaacttc aacaggtact aattaacctc	4740
tgcaagaatg cttcccaagc catgactgca aatggtcaaa tcgacatctt cgtcggccaa	4800
gcttatttac cagctaagaa aattctggcg catggtgtta tgccacctgg cgactatgtt	4860
ctactatctg tcagcgacaa tgggtggaggc atttcagagg ctgtgctacc ctacattttt	4920
gaacccttct ttacgacacg tgctcgcaac ggtggaacgg gtctcggcct cgcttctgtg	4980
catggtcata tcagcgcgtt tgcaggttac atcgacgtta gttcaactgt tgggcatggg	5040
acgcgccttg acatttatct cctccgtct tctaaggagc ccgtcaatcc cgacagcttt	5100
ttcggccgca ataaggcacc gcgtggaaac ggggagattg tggcacttgt tgagcccgat	5160
gacctcctgc gggaggcgta tgaagacaag atcgccgctc tgggatatga gccggtcggt	5220
tttcgtacct ttagtgaaat tcgcgattgg atttcaaaag gcaatgaagc cgatctggtc	5280
atggctcgacc aagcgtctct tctgaagat caaagtccta attccgtgga tttagtgtc	5340
aagaccgcct ccatcatcat tggcggaat gatctcaaaa tgccccttc aagggagaat	5400
gcgaccaggg accttatct gccgaagccg atatcgcca gaactatggc gcatgcaatc	5460
ctaaccaaaa tcaagacgta gaggttgcgac gtatcaggac tgggcaatcg cggccgccgg	5520
gcttacgtgt cttttgtaac aaattgcgac gcagttgata tccacttgaa acattagtcc	5580
gaaattatcg agatctccgc tgataaggta tcgaaatggc gataaaattg gtattgatac	5640
tcgtattcac actctttccc gcggcagacg ctgcatatgc gaatgaccgc gccaacgggt	5700
tcatgtggtc aaacgggggc gaaactggag tgaggcttcc tcttcgggtg ttcaatgcca	5760
agccagccaa gaacacgggtg gcgatcattt attccggaga cgctggatgg caaaatatcg	5820
atgaggcgat tggtaacctat ctgcagacgg aagggattcc tgtcattggc gtcagttcac	5880
ttcggtatth ctggtcggag cggccccaa gcgaaactgc taaggatctt ggtcacataa	5940
tcgatgtcta caccaagcat ttcggtgtgc agaattgttt acttgtagga tattctttcg	6000
gcgcggacgt catgccggca agcttcaata ggcttacgct tgagcaaaaa aatcgggtta	6060
agcaaatctc tctcttggca ttgtcacatc aagtcgacta tgtcgtctca tttaggggct	6120

ES 2 843 556 T3

ggctccaact cgaacgga ggtaagggcg gcaatcctct ggatgatctc agatccattg	6180
accctgcaat cgtccaatgc atgtacgggc gcgaagaccg taataatgct tgcccatctc	6240
tccgacagac cggcgagag gtgataggct tcagcggagg ccatcacttt gataatgatt	6300
tcaaaaaact gtctacgcgc gtcgtctcag gcctcgtggc acgcctaagt catcagtaat	6360
cttttagttcc tgcaccgcgg gacccgcgag gtttctcgtc tcaattgaaa tcataaagaa	6420
gcaattgaaa attttcgagt aaccgaccct cccgataatc ttcaacataa aacaacgcac	6480
ttcttccaac gggagaggcg gtgttagttg cgagctaagg agataaggta tgcttaagag	6540
atcggggtcg ctttctcttg ccttgatggt ctccttctgt tcgtcgagcc ttgccacgcc	6600
actctcatct gctgagtttg accatgttgc tcgcaagtgt gccccatcag ttgcgacatc	6660
tacgcttgcg gcgatagcta aggtggagag tcgctttgat ccttttagcga ttcatgacaa	6720
cacgaccggc gaaacgcttc actggcaaga tcacagccaa gcaacccaag tcgtcaggca	6780
ccgtctcgat gcacggcatt cgctggatgt tggcctcatg caaataaaact ctcgaaattt	6840
ttctatgctc ggtctgacac ctgacgggtc gctccaggcg tgcacatcat tatctgccgc	6900
tgcaaacatg ctgaaaagtc gttatgcagc cggcgaaacg attgacgaga agcaatttgc	6960
gcttcgtcgg gcgatctccg cttacaacac cggtaatttc atcggcggtt ttgcaaacgg	7020
ctacgtgcga aaagttgaaa cagctgctca atcgctggtg cccgcgtaa tcgagcctcc	7080
aaaagacgat cacgaggcgc taaaatccga agagacgtgg gatgtttggg ggtcatatca	7140
gcgccgctcg caggaggatg gcgctggcgg tttaatcgct ccgccaccgc cacaccagga	7200
caacggcaaa tccgcagacg acaatcaagt cttattcgac ttatactaag gagtgcgca	7260
ttgatgcatg gctttgagag ataccgttta catctaaatc gcctctcgct ctgcaatgcg	7320
atgatgcgcg tgatatcgag ctgcgcccc aacctgtgctg gtgcaattgc atggagcatt	7380
tcctcatccg gaccgcgcgc agcgcaatct gcgggtggcg gcaactgacc cgcacaaatg	7440
gttaacaata tatgcacgtt tatccttggg ccgttcggcc agtcactcgc tgttctcggc	7500
attgtcgcta tcgggatctc ctggatgttc gggcgggcct cgcttgggct gggtgccggc	7560
gtcgtcggcg gcattgttat catgtttggg gcgagcttcc tcggccaaac gctcactggc	7620
ggtagttgat ggctgatcgt ttggaagaat cgacccttta cctcgagcc acacggcccg	7680
cattgtttct tggggtgcca ctgacattgg cagggttatt catgatgttc gccggctttg	7740
tcacgtttat cgttcagaac ccgctctacg aagtcgttct cgtgccgtta tggtttgag	7800
cccggctcat cgtggagcga gactacaatg cggcgagcgt cgtcctgcta tttttgcgga	7860
ccgcgggaag aagcattgat agtcagttt gggggggcgc tactgttagc ccaaaccac	7920
ttagggttcc ccacgaggg agaggaatgg tgtgatgctc ggcgcgagtg gaacgaccga	7980

ES 2 843 556 T3

aagatccggt gagatctatc tcccttatat tggccacctc agcgaccata tcgtccttct	8040
tgaagacgga tcgatcatga ccattgcgag aattgatggc gttgcattcg agcttgagga	8100
aactgaaatg cgcaatgcgc gttgtcgtgc gttcaacacg ctggttcgca atatacgtga	8160
tgatcatgtg tcaatatatg ctcacctcgt acgtcatgcc gacgtgccat catcggcgcc	8220
gcgacacttc cgtagtgttt tcgccgctag cctgaacgaa gcttttgaac agcgcggtgt	8280
ctccggccaa ctctccgca atgaacactt ccttacgttg attgtctacc cacaggcggc	8340
tttagggaag gtaaagagga ggttcaccaa gctaagcgga aaaagggaaa acgatctcac	8400
gggccagatc aggaacatgg aagatctttg gcatgttgtc gctggctctc ttaaagcgta	8460
tggcctgcat cgtcttgcca tcgcgagaa gcagggtgtg ctcttcaccg aaattggcga	8520
agcgctacgg ttgatcatga ctggtcgggt caccgcggtt ccggtcgtca gcggctcact	8580
cggcgcttcg atttataccg acagagtcac ttgcggcaag cgaggactcg agatcagaac	8640
gccaaaagac agttacgttg gatccatcta ttcgtttcgc gaataccctg caaaaacacg	8700
gccgggcatg ctcaacgcgc tgctatccct cgattttcca cttgttctca cgcagagttt	8760
ttcgttcctg actcgcccg c aagcgcacgc gaaacttagc ctcaaatcga gccagatgct	8820
gagttccggc gataaagccg tgactcaaata cggcaaatta tccgaggctg aggacgcact	8880
tcgagcaaac gaattcggtt tgggtcaca tcatttgagc ctttgcgtct atgcagacga	8940
tctcaatagt cttggggaca ggggcgcgcg ggctcggaca cgaatggcgg atgcaggtgc	9000
cgtggttgtc caagaaggta ttggtatgga agcggcctat tgggtccaat tgccggggaa	9060
ttttaagtgg cgcacacgcc ctggcgcaat cacttcacgc aatttcgcag ggtttgtctc	9120
tttcgaaaac tttccagagg gcgccagctc aggccactgg ggcaacgcga ttgcccgatt	9180
tcgtaccaat ggcggaacgc ctttcgacta tatcccgcat gagcacgatg ttggcatgac	9240
ggcaatattc gggcctatcg ggaggggtaa gacgacgctc atgatgtttg ttctagccat	9300
gctcgaacag agcatggtcg accgtgcagg tacggtcgtg ttctttgaca aggaccgggg	9360
tggcgaattg ctggttcgcg ccacaggagg aacatatttg gcacttcaca gaggcacacc	9420
cagcgggttg gcgcggttg gtggcctaga aaacacagca gcctcacacg attttctgcg	9480
cgaatggatc gtggctctca tcgagagtga tggtcggggg gggatttctc cggaagagaa	9540
ccgccgtctg gtccgtggta tccatcgtca gctctcgttt gatccacaaa tgcgttcaat	9600
cgcgggggta cgtgaatttt tgttgcatgg gcccgccgaa ggcgaggag cgcggtcca	9660
acgctggtgc cggggccatg cgcttggtg ggcatttgac ggcgaaagtt acgaagtaaa	9720
gtagatccg tcgattaccg gcttcgacat gacgcatctt ctcaatacag aggaagtatg	9780
cgtccccgt gcagcatatc tcctgcacg gattggagcc atgatcgac gccgccgttt	9840
tgtgatgagc tgcgatgagt ttcgcgccta tttgttaaac cctaaatttt cgactgtcgt	9900

ES 2 843 556 T3

cgacaaattc ctcctgaccg ttcgaaaaaa caacgggatg ctaatactgg caacgcagca	9960
accagagcat gttctggaat cgcgctagg agccagcttg gttgcgcaat gtatgacgaa	10020
gattttctat ccatcaccaa ccgcagatcg atcggttat gtogatggac tgaaatgtac	10080
cgaaaaggaa tttcaggcga tccgtgaaga catgacggtc ggcagccgta agtttcttct	10140
taaacgagaa agtggaagcg tcatctgcga atttgatctg cgggatatgc gtgaatatgt	10200
cgcctgctt tcggggcgtg ccaacacggg gcgctttgca actcgactac gcgaggcaca	10260
agaaggcaac tcatctggct ggctcagcga attcatggcc cgtcaccacg aggcagaaga	10320
ttgataaggt aggaaacgat gaagacgacg caacttattg caacagtttt gacctgcagc	10380
tttctatata ttcagcccg cggggcgagc tttgttgta gcgacccggc aacggaggct	10440
gagacgctcg cgactgcgct cgcgactgcg gagaatctca ctcagactat agcgatgggt	10500
acgatgttga cgtcggccta cggcgttact ggactactga cttcgctcaa ccagaaaaat	10560
cagtatcctt cgacgaagga cctagacaat gaaatgtttt cgccgcgaat gccaatgtcg	10620
accacggcac gtgcgatcac cagcgataca gatcgtgcag tcgtgggtag tgatgctgaa	10680
gcggacctgt tgcgatcgca gatcaccggg tccgcaaaca gcgctggcat tgcggctgac	10740
aatctggaaa cgatggacaa acgcttgacg gcgaatgctg atacgtctgc tcagctttcc	10800
cgatctcgca atatcatgca ggcaaccgtg accaatgggt tgcttctcaa gcagatccat	10860
gacgcaatga ttcaaaatgt acaggcgaca agcctattaa cgatgactac cgcgcaggcc	10920
ggccttcacg aggcggaaga ggcggccgct caacgcaagg agcatcaaaa gaccgctgtc	10980
atctttgggt ccctccccta aggetgggcg atttgttcat ccgcccgcat cctcgccgaa	11040
tgcgagctca ttttatccaa cattatgcga caaaccagtc aagttcaggt ccaatcgatg	11100
aatttcacga ttccggcgcc gtttacggcc attcatacga tcttcgatgt agccttcacg	11160
acaggccttg actcgatgct tgagactatc caggaggcgg tgagtgcgcc attgatcgcc	11220
tgtgtcactc tttggattat tgttcagggt attttagtca tacgcggcga agtcgatacc	11280
cgtagcggta tcaactcgggt gatcacggtc accatcgttg ttgctctaata tgttgggcag	11340
gctaactacc aagactatgt ggtttccatc ttcgaaaaga cgggtccaaa ctttggttcag	11400
cagtttagtg taaccggtt gcctctgcag actgttccgg cacagtggga tacaatgttc	11460
gccgtgacc aggcggtttt tcagaaaatc gcatccgaaa tcgggtccgat gaacgaccag	11520
gacatccttg ctttccaagg ggcacagtgg gtcctttacg gcacgctctg gtctgccttc	11580
ggagtttacg acgccgttg aattctcacg aaagtgttc tcgcatcgcg gcctctgatc	11640
ctcgtcggat atatttttga tcgcacgcgg gacatcgag ctaagtggat cgggcaactt	11700
atcacctacg gtctcttgct tctcctctta aacctcgtgg caacgatcgt catcctaacc	11760

ES 2 843 556 T3

gaagcgactg cgctcaccct tatgcttggg gtaatcacct tcgccggtac gaccgcggcc	11820
aagatcattg gtctttacga actcgatatg ttttttctga caggggatgc gctcattgtc	11880
gctttgccgg cgatcgccgg caacattgga ggcagttact ggagcggcgc aacccaatct	11940
gccagcagct tgtaccgtcg cttcgctcag gttgagcgag gctaggtcgc gcaaaaattc	12000
gcctcaatgg agaattctat gaagtattgc ctgctgtgcc tagttgtcgc tttgagcggc	12060
tgccagacaa acgacacatt agcgagctgc aaaggcccg tcttcccgct gaatgtgggg	12120
cgatggcagc ctactccgtc agatcttcag ctcgcaatt cgggtggacg ctatgacggg	12180
gcctgaatat gccatgctag tggcgcgcg aagccttgcc gagcactata aggaagtaga	12240
agcctttcaa accgcgcgag cgaaatcggc gcgagctctc tccaaactca ttgcagctgt	12300
cgcagctatc gcgatttttg gaaatgttgc tcaagcgttc gctatagcca caatggtgcc	12360
gttgagcagg cttgtgcccg tatactctatg gatacgcccg gacggcaccc ttgacagcga	12420
ggtgtctgtc tcgcgattgc ctgcaactca agaggaggcc gtcgttaacg cctcattgtg	12480
ggagtacgtt cgctgcgcg agagttaga tgccgacacc gctcagtacg cctacgacct	12540
ggtatcgaac ttcagtgcc caacagtgcg ccaggattac cagcaattct tcaactatcc	12600
caatcccagt tcgcctcaag tcattcttgg caaacgcggc aggggtggagg tcgagcacat	12660
cgtttcaaat gatgtaactc caagcacgca gcaaattcgc tataaaagga ccctcgtcgt	12720
tgacggcaaa atgcctgtgg tgagtacgtg gaccgcgaca gttcgtctacg aaaaggtgac	12780
cagcttgccc ggcagattga gactaacc aa cccggcaggt ctggttgtca cctcctatca	12840
gacatcggaa gataccgttt caaacgtagc ccacagcgaa ccatgatcag aaaagcactt	12900
ttcatttttag catgtttatt tgccgctgcg actggtgcgg aggctgaaga cactccaatg	12960
gcgggcaagc tagatccgcg catgcgttat ttggcttaca atcccgatca agtgggtgcg	13020
ctctcgacgg cggttggagc tactttggtc gtaacattcg ccacgaacga aacggtgaca	13080
gcggttgccg tttcaaatag caaagatcta gcagccctac cgcggggaaa ttatctattt	13140
ttcaaggcaa gccaggtcct cagcctcag ccagtaatcg tgctaaccgc aagcgactcc	13200
gggatgcgcc gttatgtttt cagtataagt tccaagactc tgtcccacct cgataaagag	13260
cagcccgatc tctattacag cgtccaattc gcctaccccg ccgacgatgc ggcggctcgg	13320
cgaagggagg cacaacagaa ggctgttgtg gacagactac acgcggaagc acaatatcaa	13380
cggaaagctg agaatttatt ggatcagcct gtcacagccc ttggtgcggc ggacagtaat	13440
tggcactacg tcgccaagc cgatcgcttcg ctgttgccac tcgaagtctt cgacaatgga	13500
tttacgacgg tattccactt tccgggcaat gtacgcatac cctccatcta caccatcaat	13560
cctgatggca aggaagctgt tgccaactat tcagttaaag ggagcgatgt cgagatttct	13620
tcggtttccc gaggttggcg tctgagggat ggccacacag tactatgtat ctggaacacc	13680

ES 2 843 556 T3

gcttacgatac	ccgttggtcca	aaggccgcaa	acgggcacgg	tgaggcccca	tgtgaaacgc	13740
gtcctgaagg	gggcgaagg	atgaataacg	atagtcagca	agcggcacat	gaggttgatg	13800
catctggatc	cctggtctcc	gacaaacatc	gccggcgtct	ttcggggtct	cagaaattga	13860
tcgtcggagg	tgtcgttctc	gcgttatcat	taagcctcat	ttggctaggt	gggcgcctaa	13920
agaaggtgaa	tgagaacgca	tcgccgtcaa	ctttgatcgc	aacaaacacc	aagccatttc	13980
atccagctcc	gattgaggtg	ccgccggatc	ctccagcgg	tcaagaggct	gttcagcctg	14040
ctgctcctct	accgccgagg	ggcgaaccgg	agcggcatga	gccacggccg	gaagaaacac	14100
cgatttttgc	atatagcagc	ggcgatcaag	gggtcagcaa	acgcgccatt	cagggcgaca	14160
cgggccgaag	acaagaaggc	aagcgtgacg	acaactcctt	gccgaatggc	gaagtgtccg	14220
gcgagaacga	tttgtcgata	cgtatgaaac	ccaccgagct	gcagcccagc	agcggcacgc	14280
tcttgccgca	ccccgatttt	atggtaacgc	aagggacaat	aattccgtgc	atcttgcaaa	14340
ccgcaatcga	cacaaatttg	gcaggctatg	taaagtgtgt	cttgccctcag	gatattcgtg	14400
gaacaacgaa	caatatcgtg	cttcttgatc	gtggcaccac	cggtgttggc	gaaatacagc	14460
gtggcttgca	acagggagat	gggcgcgttt	ttgtgttgtg	ggatcgcgcc	gagacacctg	14520
accatgcgat	gatctcgtta	acatcgccaa	gcgcggacga	actcggtcgc	tcaggattgc	14580
cgggctcgg	cgacagccac	ttctggcagc	gttttagcgg	agctatgctc	ttgagtgttg	14640
ttcaaggcgc	cttccaggca	gtagcacct	acgccggcag	ctcgggtggc	gggatgagct	14700
tcaacagctt	tcaaaataac	ggtgagcaga	caactgagac	agcccttaag	gcaaccatca	14760
acataccgcc	aaccctgaag	aagaatcagg	gtgacaccgt	ttccattttc	gtagcacggg	14820
acctcgattt	ctttggtgtt	taccagctcc	gcctgactgg	cggcgccacg	cgggggagga	14880
accgccgctc	ttaatgaatt	caaatctccg	cttagagata	ggatacattg	taaatggaag	14940
tggatccgca	actacgcttt	cttctgaagc	cgattttgga	atggctcgat	gaccggaaga	15000
ctgaagaaat	tgcgatcaat	cgacctggag	aggcatttgt	gcgccaagcc	ggcattttta	15060
ccaagatgcc	tttgcccgtc	tcttatgatg	atcttgaaga	tatcgctatt	ttagcgggcg	15120
cgctgagaaa	gcaggatgtc	ggaccacgta	acccctctct	cgccactgaa	cttcctggtg	15180
gtgaacgact	acaaatctgt	ctgccgccga	ccgttccctc	gggcaccgtc	agcttgacca	15240
ttcgacggcc	aagctcgcgt	gtttctggtc	ttaaagaagt	ctcctcccgt	tatgatgctt	15300
cgaggtggaa	ccagtggcag	acacgaagga	aacgcctaaa	tcaggatgat	gaagctatcc	15360
ttcagcattt	tgacaacggg	gatttggaag	cgtttctgca	cgcatgcgtc	gtcagccgac	15420
tgacgatgtt	gctatgtggc	cctaccggaa	gcggcaagac	aacaatgagc	aagaccttga	15480
tcagcgccat	ccccccccag	gaaaggctaa	tcaccataga	agatacgctc	gaactcgtca	15540

ES 2 843 556 T3

ttccacacga taatcatggt agactactct actccaagaa cgggtgctggg ctgggtgctg	15600
tgagcgccga gcaattgctc caagcaagtc tgcgtatgcg gccggaccgg atattgcttg	15660
gcgagatgcg cgacgatgca gcatgggctt atctgagtga agtcgtctcg ggacatccgg	15720
gatcgatttc aacaatacac ggcgcgaatc caatccaagg attcaagaaa ctgttttccc	15780
ttgtcaaaaag tagcgcccaa ggtgctagct tggaagatcg cacactgatt gacatgctct	15840
ctacggcgat cgatgtcatc attccattcc gtgcctatga ggacgtttat gaagtaggcg	15900
agatctggct cgcggcggac gcacgacgcc ggggcgagac catagcgat ctctttaatc	15960
aatagtagct gtaacctcga agcgtttcac ttgtaacaac gattgagaac ttttgcata	16020
aaattgaaat acttggttcg cattttcgtc atcccggtc agccgcaatt ctgacgaact	16080
gcccathtag ctggagatga ttgtacatcc ttcacgtgaa aatttctcaa gcgctgtgaa	16140
caagggttca gatttttagat tgagaggatga gccgttgaaa cacgttcttc ttatcgatga	16200
cgatgtcgct atgcggcatc ttattatcga ataccttacg atccacgcct tcaaagtgc	16260
cgcggtagcc gacagcacc agttcactag agtactctct tccgcgacgg tcgatgtcgt	16320
ggttggtgat ctaaatttag gtcgtgaaga tgggcttgag atcgttcgaa atctggcggc	16380
aaagtctgat attccaatca taattatcag tggcgaccgc cttgaggaga cggataaagt	16440
tgttgcactc gagctaggag caagtgattt tatcgctaag ccgttttagta cgagagagtt	16500
tcttgcacgc attcgggttg ccttgccgtg gcgccccaac gttgtccgt ccaaagaccg	16560
acggtctttt tgttttactg actggacact taatctcagg caacgtcgt tgatgtccga	16620
agctggcggg gaggtgaaac ttacggcagg tgagttcaat cttctcctcg cgtttttaga	16680
gaaacccgcg gacgttctat cgcgcgagca acttctcatt gccagtcgag tacgcgacga	16740
ggagggttac gacaggagta tagatgttct cattttgcgg ctgcgcgcga aacttgaggc	16800
ggatccgtca agccctcaac tgataaaaac agcaagagg gccggttatt tctttgacgc	16860
ggacgtgcag gtttcgcacg gggggacgat ggcagcctga gccaatgca tttgattaat	16920
ttaggtgact gaggacgcgg ccagcggcct caaacctaca ctcaatattt ggtgaggggt	16980
tccgataggt ccctcttcac caattgctcg atggcttctc tccagcaaag aatgacgcga	17040
gcgcggcggg agccagcttg tggccgaaag ctcgagcggg ctccaacccc aacggatcaa	17100
aatgacttcg agcgacctcg agcaacgcaa ccgggaacat gcgtgaggtc tgaacgagaa	17160
cggatttttc tgtagttgaa gggatcggat aacttttcgg ggccacgcga aatgatccat	17220
ctgccagcat gctttcgaaa tcgtccaacg cgcgccttaa aatcatttgt agcgacttcg	17280
agggactgta ttgccgaacg aggttgctcat atgttttcga cacttgaggc gcgggcggtc	17340
gcgctgaaag aaaaacctgg agctttttcg gggacggagg tggactaagg gcatccacag	17400
ttagcttaag ttgtcgatcg ggactgtaaa tgtgatcggc gacgagaggc tcacgttgct	17460

ES 2 843 556 T3

ggctcttctc gtcggctttt tcaggcaagt gctggaggtc cagcttctgg ggaacaagtg	17520
tcgggttggg atggtggatc tcgggtcgag caccagcaag ccgccgtgct tcgccgaccg	17580
acaatgcggg cttgcgaatt gccatcttca agcctccaag attttgctga tcagtttcga	17640
aatgaccacg acttctcca tcgcaatccg aagattctct tctatgaggc gcatcgtcgg	17700
atcagttccc gtgtttagta atgtaagatg caacatgccg cgttctttca tcgcggcaaa	17760
tgcattctct tcatgcatgg gagacggtag aactggaagg ctctctagcg tctctgacat	17820
cctgcgttgc gatgttgtca atcgccgac cgggacgcgt tggcgcaaaa cagctgtagg	17880
aattgccaaa ttttcaactca acagcagctc gatgacgtag cggtaggtag atagtgcctc	17940
atcgatgtcg agcggcggtta gcatggtggg gatcagaagc aggtttgagc tagcgatgat	18000
tgtgtgtttg agctcgctcg agccgccacg cgtatcgccc aacgcataat caaatccttc	18060
gagctcggca ttttcatagg ctgcttcaag aaggggcatt tcgtcggcgg aatagacttc	18120
acagcgagga tcccaggtag tgctttgtaa ggcgttttct ctccatcgcg tcagaggccg	18180
gttttcgtcg gcatcaaaga gggccactcg tttaccgtca tttgccaaag cagcgcaaag	18240
gcccattgagt gcggtggttt tgccagcacc ccctttgaaa gaacaaaacg tcaaaagttg	18300
catattctga tccgcctat cctgtgaaac cggagtgcatt ttgtattttt gttcgtataa	18360
atgtttttgt gattatcgat gagtaaaagc gttgttacac tatttttatt tcacattcgt	18420
tataagacaa ttgcaaatgt agcaagtata ttcagtattg actgtaaatg tactgttgat	18480
ttcatattga gcagggttag acttccatcc gtctacccgg gcacatttcg tgctggagta	18540
tccagacctt ccgctttctt tggagggaagc tatgtcaaaa cacaccagag ccacgtcgag	18600
tgagactacc atcaaccagc atcgatccct gaaagttgaa gggttcaagg tcgtgagtg	18660
ccgtctgcga tcggccgagt atgaaacctt ttcctatcaa gcgcgcctgc tgggactttc	18720
ggatagtatg gcaattcgcg ttgcggtgcg ccgcacggg ggctttctcg aaatagatgc	18780
agacacacga gaaaagatgg aagccatact tcagtccatc ggaatactct caagcaatgt	18840
atccatgctt ctatctgcct acgccgaaga ccctcgatcg gatctggagg ctgtgcgaga	18900
tgaacgtatt gcttttggcg aggctttcgc cgccctcgat ggactactgc gctccatttt	18960
gtccgtatcc cggcgacgga tcgacggtag ctctgtactg aaaggtgcct tgtagcactt	19020
gaccacgcac ctgacgggag aaaattggat gccgatcgc gctcaagtaa tcattcgcat	19080
tgtgccagga ggtggaacca agacccttca gcagataatc aatcagctgg agtacctgtc	19140
ccgaaagggg aagctggaac tgcagcgttc agcccgcat ctcgatattc ccgttccgcc	19200
ggatcaaata cgtgagcttg cccaaagctg ggttacggag gccgggattt atgacgaaag	19260
tcagtcagac gatgacaggc aacaagactt aacaacacac attattgtaa gcttccccgc	19320

ES 2 843 556 T3

aggtaccgac	caaaccgcag	cttatgaagc	aagccgggaa	tgggcagccg	agatgtttgg	19380
gtcaggatac	gggggtggcc	gctataacta	tctgacagcc	taccacgtcg	accgcgatca	19440
tccacattta	catgtcgtgg	tcaatcgtcg	ggaacttctg	gggcaggggt	ggctgaaaat	19500
atccaggcgc	catccccagc	tgaattatga	cggtttacgg	aaaaagatgg	cagagatttc	19560
acttcgtcac	ggcatagtcc	tggatgcgac	ttcgcgagca	gaaaggggaa	tagcagagcg	19620
accaatcaca	tatgctgaat	atcgacgcct	tgagcggatg	caggctcaaa	agattcaatt	19680
cgaagataca	gattttgatg	agacctcgcc	tgagggaagat	cgtcgggacc	tcagtcaatc	19740
gttcgatcca	tttcgatcgg	acgcatctgc	cggcgaaccg	gaccgtgcaa	cccgacatga	19800
caaacaaccg	cttgaaccgc	acgcccgttt	ccaggagccc	gccggctcca	gcacaaagc	19860
cgacgcacgg	atccgcgtac	cattggagag	cgagcggggg	gcccacccat	ccgcgtccaa	19920
aatccctgta	actgggcatt	tcgggattga	gacttcgtat	gtcgtggaag	ccagcgtgcc	19980
caaacaaagc	ggcaattccg	atacttctcg	cccgttgact	gacgttgcca	tgcacacagt	20040
cgagcgccag	cagcgtacaa	aacgacgtca	tgacgaggag	gcaggtcoga	gcggagcaaa	20100
ccgtaaaaga	ttgaaggccg	cgcaagttga	ttccgaggca	aatgtcgggtg	agcccgcg	20160
tcgcgatgac	agcaacaagg	cggctgatcc	ggtgtctgct	tccatccgta	ccgagcaacc	20220
ggaagcttct	ccaacgtgtc	cgcgtgaccg	tcacgatgga	gaattgggag	aacgcaaacg	20280
tgcaagaggt	aatcgtcgcg	acgatggggc	cggggggacc	tagactagta	gacaggaagg	20340
accgaataat	ggcaaatggt	cagttcacga	tacgctctgc	tcgcccggcc	tccgtcggac	20400
tgacagggca	acggcgtgga	gccgcatccg	cctctagctc	tgcaactgtc	aatgttcaaa	20460
gagatgttag	ggataggctg	attccaacta	gtccaccaag	attaccaa	gcagccatat	20520
tgctgtgattc	ctcggaaga	gcgtcgactg	gtctcggtga	catggcggct	actcttcatt	20580
ggtctgcgat	cgcgccatta	tcgctaataa	acagcaacga	cctggctccg	gccgcttatg	20640
actttgagac	gcgaaataac	gcaagaaatg	tgactgccaa	agtcggcagg	gcagtccctg	20700
ttcccaagca	aggcgggctc	ggcaaaacgc	tcgcacccgt	acccttagt	acacgtatat	20760
caagggtcaa	ttccgaccga	agactgcccg	ctgacgcaga	agaccgccct	gaaacgcgcg	20820
acccccagaa	aggacgtggc	agtcatgggtg	cgacccaac	cttacatgaa	aagattggaa	20880
ccgcgtttgc	tcgaagattg	cgaaagcata	cgtactatat	tgtttgagc	tgctgccaga	20940
cccggagcgc	gttgacgatg	ggtgcaaaga	tttcggtgaa	gtcatgaact	ccagcaagac	21000
ttcgcgccag	cgtatgacct	tgagcatcgt	atgttcgctg	gcagccgggt	tttgtgcggc	21060
cagctgctat	gtaacgttcc	gccggggctt	caacggcgaa	gcgatgatga	cgttcgacgt	21120
tttcgctttt	tggtatgaga	ccccgcttta	cttgggttat	gccagcaccg	tcttctggcg	21180
tggtttatct	gttgtcatct	ttacctcgct	gatcgttctt	tcaagtcagc	tcacatatac	21240

ES 2 843 556 T3

gctgcgcaat cagaagcatc atgggacagc tcgttgggca gaaattggcg aaatgcggca	21300
tgctggttat ctgcagcgtt acagtcgcat caaggggccg atctttggaa agacatgtgg	21360
tcccccttgg ttccggcagtt atttgaccaa tggcgaacag ccacacagtc ttgtcgtcgc	21420
gccaacgcgt gctggcaaag gcgtcggcat cgtcattcca acgctgttga cttcaaggg	21480
ctcggtaatc gcccttgacg tcaagggaga attgtttgaa ctgacgtcca gagcacgcaa	21540
agcgagcggc gacgcagttt tcaagttctc cccctagat cctgagcgga agactcattg	21600
ttacaatccg gtcctggata ttgccgact tccgcccgaa cgccagttca ctgaaacacg	21660
ccgtctagct gcgaacctta ttacggctaa gggaaagga gcagaaggct ttattgacgg	21720
cgcacgtgac ctgttcgtcg cggaatcct tacctgcatt gagcgtggca caccaacgat	21780
tggcgcggtat tatgacctat ttgcgcagcc tggcgaaaag tataagcttt ttgcgcaact	21840
cgcgaggaa agcctaaca aagaggctca gcgtatcttc gataatatgg cgggcaacga	21900
cacgaaaatt ctgacatcgt acacctctgt gctgggcgac ggtggactga acctgtgggc	21960
tgatccgctt atcaaagcag cgacaagccg gtcagacttt tccgtttacg atctccggag	22020
gaagaagacc tgcatttatc tttgtgtcag tcccaacgat ctggaggtct tggcaccact	22080
tatgcgcctg atgtttcagc agctcgtgtc aatcttgag agatcgctgc caggtgaaga	22140
cgagtgccat gaagttttat ttctcctcga cgaattcaaa cacctgggca agcttgaggc	22200
catagagacc gcgatcacia ccacgcggc ttacagaggc cgctttatgt ttattattca	22260
aagtctttcg gccttgctcg gcacatacga tgacgcagga aaacaaaact ttctgagcaa	22320
tactggcgta caagtattta tgccacggc tgatgacgaa actccaacct acatctcaaa	22380
agctatcggc gaatatacgt ttaaagcgcg ttcgacctct tacagtcaag ccagaatgtt	22440
cgaccacaac atccagatct ctgatcaagg tgcagccctt ttgcgccccg aacaagtgcg	22500
cctgctagac gatcagagtg aaatcgctct catcaaaggc cgacctccac tcaaattacg	22560
aaaggtgcag tattattccg atcgtacgct gaaaggcctt ttogaacgcc agatgggctc	22620
tctgcctgag cccgcaccct tgatgctttc cgactatagc aacgatcaag ttcaatacca	22680
cttggctccg atagcaaatt ttaatgagga tgctgcaccg caaacagaa ctgtggccga	22740
ggaccatgga agtggttaaag tcggtgctga tatccctgaa cgcgatgatg gaataaatgg	22800
tgacgaggaa caagccgatg cgggcgagat accgccgaa tcggttgtgc ctccagaatt	22860
gacgtcgtc ctgaccgtc aacagcaatt gttggaccag attattgcac ttcagcaaag	22920
atcgaggtcc gcaccggcat agcctgcgaa atgatcttaa tgggtgcaatt cgtttcaggc	22980
ggcatgttgc ggttcaacaa agtgtagcc agatcccaat tggcccttta cgaggtgtcg	23040
gtatgacagg aaagtcgaaa gttcacataa gaggttcggc tgacgcgctt cctgacgttc	23100

ES 2 843 556 T3

ctggcggaag tactaccgcc ccttttttaa ccgaaccttc tcgggatcag gttgatgcct	23160
cgtttgaggt ccaaaccgac tacagccagt ctacttccgt gtcgtttacc tatgatggtg	23220
ttggacttgg tcctgccgag cgtgcggcct acgagaactg gtgcgaaccg ggccggccca	23280
cttgaaaaga tcttataatc aaggcacgtg tcgatccgat tgacgatgtg acctggctcc	23340
gagattttaga agaggacacc ccctcaacct tcagatacga agggatgcct ctgggcatcg	23400
gggaacgaca ggcttacgaa aattggcaag aggacgctca gccgacatgg gaagaccttg	23460
ttgtcagcgc acgcttgacg gaacttggcc gtccacacgg gattaccggc gagtatacat	23520
ccctcgcagg atcgaagaat acaagttcaa tttcattgaa gcgaaagcgg agcaacttaa	23580
ttgatgatga gaattcatcc ggatcgtttt catatgacgg gatgaagctc ggggaagccg	23640
agcgttctgc atatggtgac tgggccgagg cggagccacc cacgtggaaa gatttggtat	23700
tgagggcacg cgtttcctcg atcaatgact ctgcttggct ttttgattca caaacatctt	23760
catcatcatt tgaatacaac ggtgttccct tgggcgagcc ggaacggcag gctctcagac	23820
aatggcaagg agacgctcag cctacctggg aagatctcgt tgtaaacgcg cgtatggcag	23880
aactttgcc a tgctggttgg attgaaggtc aaaaaggttg ctttgaagag cgcggggagg	23940
ctctgccgcg gtcggaacgc ggttcgcaac gcccattgg tcaacggaca gattcctccg	24000
attcttttgt gtatgatggc acaaggctcg gagcacctga gcgaactgct tatgaacgct	24060
ggagtaagag ggaacgcccg acttgggaag atctcatctt agatgcacac caggccagga	24120
ctgaaagtga cgctgttacg acccaagcga ttggtcagtc gtcctcaccg gttttcttat	24180
atgaaggaaa gtcgctcgga gacagggaac gaaaggctta cgaaaaatgg cggcagccag	24240
cccaaccgcg atggcaaaat cttgtagtga acgctcgtct ggcagaaatc gatccctcag	24300
cctggattgc cgatgagcgc gatccgcttg atgatagcga cgcgcttggc cgcccgctcg	24360
acacaagctt gacggataga tcagacgtcc ctttagacga tcaatcaatc tatcgtcgtt	24420
ccgacctagt aagggagcag gtgccagaat cgtctcaaag gcaattcgca gcatgttcag	24480
aatctgaaac gaggcctgtg caatggttta ctgcttctgg gtcagatgca aacaatacgg	24540
aaaatatcac cgccagcgat ccgctcgatc gcacgggtgg agttaagcgg ctaggctcca	24600
aaagcgacag aaccgttaca gcttctatcc atgacgtgaa ttccagcaca aggcgactgt	24660
tgcttaacga atttggatcg gaggtccgc gcccttcgcc agaaaagact gttcgtttaa	24720
gaagcgacaa tattggcacc tatgggagcc ggaaaaatga acgagcgcgg ctcgcgaccg	24780
aaaccggtgc gtatgagtcg gagcatatct tcgggttcaa ggctgtccac gatactcgga	24840
gagcgacgaa agagggcccg cgtctcgaaa ggcccatgcc cgcctacctt gaggataagg	24900
ggcttcacgc ccaacatatt ggcaccggga gaggacggac caaacttgct gggcgcggat	24960
ggccggatga cacaagctat cgctcggatc aaagggaac tctgtcggac ccggttcgcg	25020

ES 2 843 556 T3

gctcgggaagg cgcgacggcc tcaaatgggt atcaattgaa ccaattgggt tacgcgcacc	25080
aactcgctag cgatggcctg caaagtgaat cgcccgatgg tggtgccttg ccaattcaag	25140
tggcaacaac gagctacaac tatacagtga gccgcgatcc tgccttggt ccgcgggata	25200
aaaacgaagc ccctcaattg ctgcatcttg gtccccgtgg tcaaaccgaa gctgttcttg	25260
cccgcgaaac agcattgact ggaaaatggc cgactctcga gcgtgagcag caagtgtatc	25320
gcgagttttt ggccttatat gacgtaaaaa aagatcttga ggccaaatca gtcggcgtaa	25380
gacggaaaaa aaaagaagtt atttctgcgt tagaccgaac tgcgcgcttg ataagcacgt	25440
cgccttcgaa agctcgatcc aaagcagaga ctgaaaaagc cattgatgag ctcgatgac	25500
gacgagttta tgatccgctg gatcgagctc aagacaaagc gtttaaacgc tgataagtcg	25560
ccaatatagt gatcatttgc agtattcgca tcgatcgctg gttgatattc tgccgctgg	25620
cgaccggctg ctcgctcgca aaaatgctca cagggatacg atggcctcgg tcaggccg	25680
tgcgtcctgt ctttccagtt cctcccttcc agctcgattg tggcatcatt tattgcctgc	25740
tcattgcagt tgaaacgcga tatccgtttc aagaccggg tatggatgg actttggaga	25800
tatgagcact tgccagactt cccggtcgga gacttaaaaa ttccgataag aggttacgga	25860
cgacggacct tataagtgga tcgtctagtg gtggcgccga tcaaaacagt tccgccccg	25920
gctttgctca aagtagcaaa gcagctttat ctcggttgcc ggaggatttt ctagaaaacc	25980
gcaattttgc aggagagaac atatggccat catcaagccg catgtgaaca aaaataggac	26040
aacctcgccg atagagagac cggagtctct catagaggaa atgagcggca gtcacccg	26100
gagtgggtttt accaacctgg atctcgctat gatcgagctg gaggactttg tccatcggtg	26160
cccgtccca gaagacaatc ttgctggtca gaaggagtga gacgatggat cgtctagca	26220
atgagaatgt ctatgtgggt cgcggtcaca acatcgaaaa tgatgatgac actgaccca	26280
ggcgttgga gaaggcgaat atcagttcca acaccatctc cgatattcag atgacgaatg	26340
gcgaagacgt acaatcaggg agccctacc gaacggaagt tgtaagcca cgtctggatt	26400
atggatcggc cgactcctcc tccagccttt attctggcag cgagcacgga aatcaagctg	26460
agattcaaaa agagctgtcc gtcttggtct cgaacatgtc ttgccaggc aacgatcggc	26520
gcccgacga atacattctc gtgcatcaaa cgggacaaga tgcttttact ggtattgcca	26580
aaggcaacct cgaccaaag cccaccaagg cggaatttaa cgcgtgctgc cgtctctaca	26640
gggacggagc cggtaattac taccgccac ctctcgcatc cgacaagatt agcgttccag	26700
agcaactgga ggaaaaatgg gggatgatgg aggcgaagga acgtaacaaa ctgcggtttc	26760
agtacaagtt ggacgtatgg aatcatgcgc acgctgatat ggggatcacg ggcacagaga	26820
tcttttatca aacagataag aacataaagc tcgaccggaa ttataaacta agacctgaag	26880

ES 2 843 556 T3

accgatacgt	acaaacagaa	aaatacgggc	gccgggaaat	tcaaaagcga	tatcaacacg	26940
aactccaggc	tggttcgctg	ctgcccagata	ttatgatcaa	aactccccaa	aatgacatcc	27000
acttcgtgta	caggtttgcc	ggcgacaatt	acgccaaaca	acagttcagc	gagtttgaac	27060
acaccgtcaa	gcgaggtat	ggcgacgaga	ctgagatcaa	attgaagtca	aagtcaggca	27120
ttatgcatga	ctcgaaatat	ctggaatcct	gggaacgggg	cagtgcggat	attcgcttcg	27180
cggaattcgt	tggggaaaat	agagctcaca	atcggcagtt	tccaactgcg	acagtaaata	27240
tgggacagca	gccagacggg	caggggcggt	tgaccgcga	ccgtcatgtg	agcgttgact	27300
tcctaatacga	aagcgcaccc	aattcgccct	gggcgcaagc	tttgaaaaag	ggagaactgt	27360
gggatcgctg	tcagttgctt	gctcgcgacg	gcaaccgcta	tctgtcgccg	cccagattgg	27420
aatattctga	ccctgcacat	ttcaccgagt	tgatgaaccg	ggttggttta	cccgcacga	27480
tgggtcggca	aagccatgcg	gctagtatca	aattcgaaaa	gtttgacgcg	caggcagcgg	27540
ttattgtctt	aaatggccca	gagttacgtg	acattcatga	cttgtctcct	gaaaaactgc	27600
aaaatttgtc	caccaaagat	gtcatcgctg	ccgatcgcaa	tgagaatggt	cagagaactg	27660
gcacgtacac	cagcgtcgcg	gaatatgagc	gcttgcagtt	aaggctgccca	cccgatgcag	27720
cgggggtgct	tggtgaagca	actgacaaat	attcacgtga	tttcgttcgg	ccagagccgg	27780
cgtcgcgctc	aatcagtgac	agccgcagga	tatacgaaaag	tcgaccgcgt	agccaaagcg	27840
tcaacagctt	ttgacgttcc	tgctgccgcg	tcaacgagga	agctcgtttg	acccgggttt	27900
gccaatgaaa	gggctcaatc	atggtgaaca	ctacaaagaa	aagttttgcg	aagtcgctta	27960
cggcagatat	gcgcggttct	gctcagcgcg	ttgtcgagca	aatgcgaaaa	gcattgatta	28020
ccgaagaaga	ggcgtcaag	cggcaagcca	gactggagag	tcccgatagg	aagcgaaagt	28080
atgctgctga	tatggcgata	gtcgacaaac	tcgacgtagg	gtttcgaggc	gaaataggct	28140
ataaaattct	tggaataaac	cggcttcgag	tagacaacca	taaagaatta	acgcgtgagc	28200
acggtagact	tcgcaaaacc	aaaacggttc	tgaagcgtaa	cccggtgacg	caggaagtct	28260
atttgggttt	atatgaaag	aagtcctggt	taagtgtcag	cagccatttg	tatgctcgcg	28320
acggcacact	ccgcatgaag	cacgtgaaat	acaaagacgg	acgttttgag	gaaaaatggg	28380
agcgcgacga	aaatggcgac	ctgatccgca	caaggtacgc	caaccgtggc	aggctctttc	28440
aacctgtatc	cgagaaaatg	ggcgcgccgt	atcggagcgg	ccctgacgac	cggctctatc	28500
gcgatctaac	ccgtcgaaac	ggtttcagac	gggagacatt	cgaacgggac	gatcacggaa	28560
acctcgagcg	tatcggcagc	aacctgtcgc	gcttttccaa	gatttcagtg	aaggcaccca	28620
atcgtcaaac	ctcccagacg	aagattcaaa	aacttgggtg	cgctttcaac	aaatctttta	28680
ggtcccttct	ggacaaggag	ggcaatgaaa	tgggccgcga	tattttgagc	catcgacggc	28740
tctataacaa	gcggtctgct	gtctacgatg	aagctaccgg	acaattgaag	agtgccaagc	28800

ES 2 843 556 T3

ataccttcgg caagatctac aggagcgaaa ccgaatatct cagcgcgggc ctcaagaagg	28860
tttcaaaaaa gatactcggg gtgacgggtct accggaaatt tgccggcgctc agcgagcgag	28920
aatccgaggc tgagagactg cgtagttttg aatccggtgc gcatcgccag atctggcagg	28980
agcgggcagc gattcccggg tcgcccctcc cgggagactga tgacattcat ttcgcacagc	29040
agtcgcacct agccaaagcc aaccctgatc acgtcgaagc tgacgtcacg cgtgtgacag	29100
atcaacatgt tgatgttgct ggacaaacat catcgtctcc ccaacggaac ttggaaggat	29160
ggttagattc tcaatcacga tacaagccag caaacatgct gttgtcaaat ccagaccttc	29220
aagcgaacgg acctcgccca tacgaagggt tagctcatct caccctccgg cgcgataatg	29280
aatctgacgg gcacaaggag aatgatcagc ggctgcgaca tttctcccag ccagagccgt	29340
tggtgttacc gcatcccggg tcgccggaaa taactaagggt gtttggtcgc cggggagagc	29400
cggcacaccc gagtggaat ctgcacacgg cggttggaga aacggcttgc gaaggaccgg	29460
tgatgtcttc atcctcggac aatcatcagc cagctccagg acagcaagaa cttttaagtt	29520
tccttcataa tgccgcagcc ccagtttctg tggcaataca tgatgatcaa gacgacttg	29580
cgggggaggc gcccgcgggc tctttcagag gtagctcagg gcgaacgagt tcaatgtcgg	29640
agagtatctt cgacgaagat gtacaagggt atttggtacg ggattattcg atcaatacta	29700
ctaacgggtt tattgacccg caatcgttgt tcggtgaacc ggacttatcg agaggtccaa	29760
aatcggggcc agaaattcca tcggaagatt accatttgtc agcttcggaa caggaaaatt	29820
tgctgaatca attgcttagt gtgccactgc cggttccttc accgaagccc gaatgcgcga	29880
ggtctatgat tttcgaagggt tcacgttcaa gagagcgttc cacctccaga gggttctaag	29940
gtggaacgcg tctaggcttg cctatcatca ccacctcatg atggtcgcgg agaccttcga	30000
agagatatcc gacatcattg gagaagctgg gccggaaact agtgtaacct gcgaatgcat	30060
ctagatccaa tccaatatgc atggcatgcc gaatccatgt gggagtttat tcttgacaca	30120
gatatttatg atataataac tgagtaagct taacataagg aggaaaaaca tatgttacgc	30180
agcagcaacg atgttacgca gcagggcagt cggcctaaaa caaagttagg tggctcaagt	30240
atgggcatca ttcgcacatg taggctcggc cctgaccaag tcaaattccat gcgggctgct	30300
cttgatcttt tcggtcgtga gttcggagac gtagccacct actcccaaca tcagccggac	30360
tccgattacc tcgggaactt gtcocgtagt aagacattca tcgcgcttgc tgccttcgac	30420
caagaagcgg ttgttgggcg tctcgcggct tacgttctgc ccaagttaga gcagccgcgt	30480
agtgagatct atatctatga tctcgcagtc tccggcgagc accggaggca gggcattgcc	30540
accgcgctca tcaatctcct caagcatgag gccaaacgcgc ttggtgctta tgtgatctac	30600
gtgcaagcag attacggtga cgatcccgca gtggctctct atacaaagtt gggcatacgg	30660

ES 2 843 556 T3

gaagaagtga tgcactttga tatcgaccca agtaccgcca cctaataatg tctaacaatt	30720
cgttcaagcc gacgccgctt cgcggcgcgg cttaactcaa gcgttagatg cactatacgt	30780
accaatccaa atcggatccc cctgcagggc tgagataggt gcctcactga ttaagcattg	30840
gtaactgtca gaccaagttt actcatatat acttttagatt gatttaaaac ttcattttta	30900
atttaaaagg atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg	30960
tgagtttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga	31020
tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg	31080
ggtttggttg cgggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag	31140
agcgcagata ccaaatactg tccttctagt gtagccgtag ttaggccacc acttcaagaa	31200
ctctgtagca ccgcctacat acctcgctct gctaactctg ttaccagtgg ctgctgccag	31260
tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca	31320
gcggtcgggc tgaacggggg gtctgtgcac acagcccagc ttggagcgaa cgacctacac	31380
cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aagggagaaa	31440
ggcggacagc tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc	31500
agggggaaac gcctggatc tttatagtcc tgtcgggttt cggcacctct gacttgagcg	31560
tcgatttttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc	31620
ctttttacgg ttccctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttatc	31680
ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgcgcag	31740
ccgaacgacc gagcgacagc agtcagttag cgaggaagcg gaagagcgcc tgatgcggta	31800
ttttctcctt acgcatctgt gcggtatttc acaccgcata tgggtgacct tcagtacaat	31860
ctgctctgat gcgcgatagt taagccagta tacactccgc tatcgctacg tgactgggtc	31920
atggctgcgc cccgacacc	31939
<210> 37	
<211> 3671	
<212> ADN	
<213> Escherichia coli	
<400> 37	
tcagcctgcc gccttgggcc gggatgatgc gtacttgccc gccgcgaact cggttaccgt	60
ccagcccagc gcgaccagct ccggcaacgc ctgcgcgacc cgctggcggc gcttgcgcat	120
ggtcgaacca ctggcctctg acggccagac atagccgcac aaggatatcta tggaagcctt	180
gccggttttg ccggggtcga tccagccaca cagccgctgg tgcagcaggc gggcggtttc	240
gctgtccagc gcccgcacct cgtccatgct gatgcgcaca tgctggccgc caccatgac	300
ggcctgcgcg atcaaggggt tcagggccac gtacaggcgc ccgtccgcct cgtcgtggc	360

ES 2 843 556 T3

gtactccgac agcagccgaa acccctgccg cttgcggcca ttctgggcca tgatggatac	420
cttccaaagg cgctcgatgc agtcctgtat gtgcttgagc gccccaccac tatcgacctc	480
tgccccgatt tcctttgcca gcgcccgata gctacctttg accaccatgg catcagcggg	540
gacggcctcc cacttgggtt ccaggaacag ccggagctgc cgtccgcctt cggctctggg	600
ttccggggcca agcactaggc cattaggccc agccatggcc accagccctt gcaggatgcg	660
cagatcatca gcgccagcg gctccggggc gctgaactcg atccgcttgc cgtcgcctga	720
gtcatacgtc acgtccagct tgctgcgctt gcgctcgccc cgcttgaggg caccggaacag	780
gccggggggc agacagtgcg ccgggtcgtg ccggacgtgg ctgaggtgtg gcttgttctt	840
aggcttcacc acggggcacc cccttgctct tgcgctgcct ctccagcacg gcgggcttga	900
gcaccccgcc gtcattgccg ctgaaccacc gatcagcgaa cgggtgcgcca tagttgcct	960
tgctcacacc gaagcggacg aagaaccggc gctggtcgtc gtccacaccc cattcctcgg	1020
cctcggcgct ggtcatgctc gacaggtagg actgccagcg gatgttatcg accagtaccg	1080
agctgccccg gctggcctgc tgctggctgc ctgcgcccat catggccgcg cccttgctgg	1140
catggtgcag gaacacgata gagcacccg taticggcg gatggcctcc atgcgaccga	1200
tgacctgggc catggggccg ctggcgtttt cttcctcgat gtggaaccgg cgcagcgtgt	1260
ccagcaccat caggcggcg ccctcggcg cgcgcttgag gccgtcgaac cactccgggg	1320
ccatgatgtt gggcaggctg ccgatcagcg gctggatcag caggccgtca gccacggctt	1380
gccgttcctc ggcgctgagg tgcgccccaa gggcgctgcg gcggtgatga atggcggtgg	1440
gcgggtcttc ggcgggcagg tagatcaccg ggccggtggg cagttcgccc acctccagca	1500
gatccggccc gcctgcaatc tgtgcggcca gttgcagggc cagcatggat ttaccggcac	1560
caccggggca caccagcgcc ccgaccgtac cggccaccat gttgggcaaa acgtagtcca	1620
gcgggtggcg cgctgctgcg aacgcctcca gaatatgat aggcttatgg gtagccattg	1680
attgcctcct ttgcaggcag ttggtgggta ggcgctggcg gggtcactac ccccgccctg	1740
cgccgctctg agttcttcca ggcactcgcg cagcgctcgc tattcgtcgt cggtcagcca	1800
gaacttgctc tgacgcatcc ctttggcctt catgcgctcg gcatacgcg cttggcgctac	1860
agcgtcaggg ctggccagca ggtcgccggt ctgcttgtcc ttttggctct tcatatcagt	1920
caccgagaaa cttgccgggg ccgaaaggct tgtcttcgcg gaacaaggac aagggtgcagc	1980
cgtcaagggt aaggctggcc atatcagcga ctgaaaagcg gccagcctcg gccttgtttg	2040
acgtataacc aaagccaccg ggcaaccaat agccctgtc acttttgatc aggtagaccg	2100
accctgaagc gctttttctg tattccataa aaccccttc tgtgcgtgag tactcatagt	2160
ataacaggcg tgagtaccaa cgcaagcact acatgctgaa atctggcccc cccctgtcca	2220
tgctcgtctg gcgggggtgcc ggtgccgctg ccagctcggc ccgcgcaagc tggacgctgg	2280

ES 2 843 556 T3

gcagacccat gaccttgctg acggtgcgct cgatgtaatc cgcttcgtgg ccgggcttgc 2340
gctctgccag cgctgggctg gcctcggcca tggccttgcc gatttcctcg gactgcggc 2400
cccggtggc cagcttctgc gcggcgataa agtcgcactt gctgaggtca tcaccgaagc 2460
gcttgaccag ccgggccatc tcgctgcggt actcgtccag cgccgtgcgc cggtggcggc 2520
taagctgccg ctggggcagt tcgaggtgg ccagcctgcg ggccttctcc tgctgccgct 2580
gggcctgctc gatctgctgg ccagcctgct gcaccagcgc cgggccagcg gtggcggtct 2640
tgcccttgga ttacgcagc agcaccacg gctgataacc ggcgcggtg gtgtgcttgt 2700
ccttgcggtt ggtgaagccc gccaaagcgc catagtggcg gctgtcggcg ctggccgggt 2760
cggcgtcgta ctgctggcc agcgtccggg caatctgcc ccgaagtca ccgcctgcg 2820
cgtcgccac cttgacccat gcctgatagt tcttcgggct ggtttccact accagggcag 2880
gctccggcc ctgggcttct atgtcatcca ggtcaaactc gctgaggtcg tccaccagca 2940
ccagaccatg ccgctcctgc tcggcgggcc tgatatacac gtcattgcc tgggcattca 3000
tccgcttgag ccatggcgtg ttctggagca cttcggcggc tgaccattcc cggttcatca 3060
tctggccggt ggtggcgctc ctgacgcga tatcgaagcg ctcacagccc atggccttga 3120
gctgtcgcc tatggcctgc aaagtcctgt cgttcttcat cgggccacca agcgattccc 3180
acacattata cgagccggaa gcataaagt tagctagatc cgaaggatga gccgggctga 3240
atgatcgacc gagacaggcc ctgcggggct gcacacgcgc cccaccctt cgggtagggg 3300
gaaagggcgc taaagcggct aaaagcgctc cagcgtatct ctgcgggggt tgggtgtggg 3360
tttagcgggc tttgcccgcc tttccccctg ccgcgcagcg gtggggcggt gtgtagccta 3420
gcgcagcgaa tagaccagct atccggcctc tggccgggca tattgggcaa gggcagcagc 3480
gccccacaag ggcgtgata accgcgccta gtggattatt cttagataat catggtgga 3540
tttttccaac accccgccag ccccgcccc tgctgggttt gcaggtttgg gggcgtgaca 3600
gttattgcag gggttcgtga cagttattgc aggggggctg gacagttatt gcaggggttc 3660
gtgacagtta g 3671

<210> 38

<211> 15028

<212> ADN

5 <213> Escherichia coli

<400> 38

ctagcgtttg caatgcacca ggtcatcatt gaccagggcg tgttccacca ggccgctgcc 60
tcgcaactct tcgcaggctt cgccgacctg ctcgcgccac ttcttcacgc ggttggaaac 120
cgatccgcac atgaggcgga aggtttccag cttgagcggg tacggctccc ggtgcgagct 180
gaaatagtcg aacatccgtc gggccgtcgg cgacagcttg cggacttct cccatatgaa 240

ES 2 843 556 T3

tttcgtgtag tggtcgccag caaacagcac gacgatttcc tcgtcgatca ggacctggca	300
acgggacggt ttcttgccac ggtccaggac gcggaagcgg tgcagcagcg acaccgattc	360
caggtgcccc acgcggtcgg acgtgaagcc catcgccgtc gcctgtaggc gcgacaggca	420
ttcctcggcc ttcgtgtaat accggccatt gatcgaccag cccaggtcct ggcaaagctc	480
gtagaacgtg aaggtgatcg gctcgccgat aggggtgcgc ttcgcgctact ccaacacctg	540
ctgccacacc agttcgatcat cgtcggtccg cagctcgacg ccggtgtagg tgatcttcac	600
gtccttggtg acgtggaaaa tgaccttgtt ttgcagcgcc tcgcgcggga ttttcttggt	660
gcgcgtgggt aacagggcag agcggggcgt gtcgtttggc atcgctcgca tcgtgtccgg	720
ccacggcgca atatcgaaca aggaaagctg catttccttg atctgctgct tcgtgtgttt	780
cagcaacgcg gcctgcttg cctcgctgac ctgttttgcc aggtcctcgc cggcggtttt	840
tcgcttcttg gtcgtcatag ttctcgcgt gtcgatggtc atcgacttcg ccaaacctgc	900
cgctcctgt tcgagacgac gcgaacgctc cacggcgccc gatggcgagg gcagggcagg	960
gggagccagt tgcacgctgt cgcgctcgat cttggccgta gcttgctgga ccatcgagcc	1020
gacggactgg aaggtttcgc ggggcgcacg catgacgggt cggttcgca tggtttcggc	1080
atcctcggcg gaaaaccccg cgtcgatcag ttcttgccctg tatgccttcc ggtcaaactg	1140
ccgattcatt caccctcctt gcgggattgc cccgactcac gccggggcaa tgtgccctta	1200
ttcctgattt gaccgcctg gtgccttgggt gtccagataa tccaccttat cggcaatgaa	1260
gtcgggtccc tagaccgtct ggccgtcctt ctcgtacttg gtattccgaa tcttgccctg	1320
cacgaatacc agcgacccct tgcccaataa cttgccgtgg gcctcggcct gagagccaaa	1380
acacttgatg cggaagaagt cgggtgcgctc ctgcttgctg ccggcatcgt tgcgccactc	1440
ttcatatacc gctatatcga aaattgcttg cggcttgcta gaattgccat gacgtacctc	1500
ggtgtcacgg gtaagattac cgataaactg gaactgatta tggctcatat cgaaagtctc	1560
cttgagaaaag gagactctag tttagctaaa cattggttcc gctgtcaaga acttttagcg	1620
ctaaaatttt gcgggcccgc accaaagggt cgaggggcgg cttccgctgt gtacaaccag	1680
atatttttca ccaacatcct tcgtctgctc gatgagcggg gcatgacgaa acatgagctg	1740
tcggagaggg caggggtttc aatttcgttt ttatcagact taaccaacgg taaggccaac	1800
ccctcgttga aggtgatgga ggccattgcc gacgcccgtg aaactcccct acctcttctc	1860
ctggagtcca ccgacctga ccgcgaggca ctgcgcgaga ttgcgggtca tcctttcaag	1920
agcagcgtgc cggccggata cgaacgcac agtgtggttt tgccgtcaca taaggcggtt	1980
atcgtaaaga aatggggcga cgacaccgga aaaaagctgc gtggaaggct ctgacgcaa	2040
gggttagggc ttgcacttcc ttcttttagcc gctaaaacgg ccccttctct gcgggcccgc	2100

ES 2 843 556 T3

ggctcgcgca	tcatatcgac	atcctcaacg	gaagccgtgc	cggaatggc	atcgggcggg	2160
tgcgctttga	cagttgtttt	ctatcagaac	ccctacgtcg	tgcggttcga	ttagctgttt	2220
gtcttgacag	ctaaacactt	tccgtatatc	gtttgcctgt	gcgataatgt	tgctaattgat	2280
ttgttgcgta	gggttactg	aaaagtgagc	gggaaagaag	agtttcagac	catcaaggag	2340
cgggccaagc	gcaagctgga	acgcgacatg	gggtcgggacc	tggtggccgc	gctcaacgac	2400
ccgaaaaccg	ttgaagtcac	gctcaacgcg	gacggcaagg	tgtggcacga	acgccttggc	2460
gagccgatgc	ggatcacctg	cgacatgcgg	cccagccagt	cgcagcgcat	tatagaaacg	2520
gtggccggat	tccacggcaa	agaggtcacg	cggcattcgc	ccatcctgga	aggcgagttc	2580
cccttggtatg	gcagccgctt	tgccggccaa	ttgccgccgg	tcgtggccgc	gccaaccttt	2640
gcgatccgca	agcgcgcggt	cgccatcttc	acgtcggaac	agtacgtcga	ggcgggcatc	2700
atgacccgcg	agcaatacga	ggtcattaaa	agcgcgcgtc	cggcgcatcg	aaacatcctc	2760
gtcattggcg	gtactggctc	gggcaagacc	acgtcgtcga	acgcgatcat	caatgaaatg	2820
gtcgccctca	acccgtctga	gcgcgtcgct	atcatcgagg	acaccggcga	aatccagtgc	2880
gccgcagaga	acgcgcgtcca	ataccacacc	agcatcgacg	tctcgatgac	gctgctgctc	2940
aagacaacgc	tcggtatgcg	ccccgaccgc	atcctgggtc	gtgaggtacg	tggtcccgaa	3000
gcccttgatc	tggtgatggc	ctggaacacc	gggcatgaag	gaggtgccgc	caccctgcac	3060
gcaaaacaacc	ccaaagcggg	cctgagccgg	ctcgccatgc	ttatcagcat	gcacccggat	3120
tcaccgaaac	ccattgagcc	gctgattggc	gaggcggttc	atgtggtcgt	ccatatcgcc	3180
aggaccccta	gcggccgtcg	agtcaagaa	attctcgaag	ttcttggtta	cgagaacggc	3240
cagtacatca	ccaaaaccct	gtaaggagta	tttccaatga	caacggctgt	tccgttcggt	3300
ctgaccatga	atcgccgcat	tttgttctac	cttgccgtgt	tcttcgttct	cgtctctcg	3360
ttatccgcgc	atccggcgat	ggcctcggaa	ggcaccggcg	gcagcttgcc	atatgagagc	3420
tggtgacga	acctgcgcaa	ctccgtaacc	ggcccgggtg	ccttcgcgct	gtccatcatc	3480
ggcatcgctg	tcgcggcgcg	cgtgctgac	ttcgccggcg	aactcaacgc	cttcttcgga	3540
accctgatct	tcctgggtct	ggatgatggc	ctgctgggtc	gcgcgcagaa	cgtgatgagc	3600
accttctctg	gtcgtgggtc	cgaaatcgcg	gccctcggca	acggggcgct	gcaccagggtg	3660
caagtgcgcg	cggcgggatgc	cgtgcgtgcg	gtagcggctg	gacggctcgc	ctaatcatgg	3720
ctctgcgcac	gatccccatc	cgtcgcgacg	gcaaccgaga	aaacctgttc	atgggtgggtg	3780
atcgtgaaact	ggatgatgttc	tcgggcctga	tgccgtttgc	gctgattttc	agcgcccaag	3840
agctgcgggc	caccgtgggtc	ggctctgatcc	tggtgttcgg	ggcgctctat	gcgttcggaa	3900
tcattggcgaa	ggccgatccg	aagatgcggt	tcgtgtacct	gcgtcaccgc	cggtaacaagc	3960
cgtattaccc	ggcccgctcg	accccggtcc	gcgagaacac	caatagccaa	gggaagcaat	4020

ES 2 843 556 T3

accgatgatc caagcaattg cgattgcaat cgcgggcctc ggcgcgcttc tgttggtcat	4080
cctctttgcc cgcacccgcg cggtcgatgc cgaactgaaa ctgaaaaagc atcgttccaa	4140
ggacgcccgc ctggccgacg tgctcaacta cgcgcgtgtc gtogatgacg gcgtaatcgt	4200
gggcaagaac ggcagcttta tggctgcctg gctgtacaag ggcgatgaca acgcaagcag	4260
caccgaccag cagcgcgaag tagtgtccgc ccgcacaaac caggccctcg cgggcctggg	4320
aagtgggtgg atgatccatg tggacgccgt gcggcgctct gctccgaact acgcggagcg	4380
gggcctgtcg gcgttccctg accgtctgac ggcagcgatt gaagaagagc gctcggctct	4440
gccttgctcg tcgggtgatg acttcaccag ctccgcgaag tcgctcttct tgatggagcg	4500
catggggacg tgcttggaac tcacgcgcac ccccgggccg ttttagcggc taaaaaagtc	4560
atggctctgc cctcggggcg accacgcca tcatgacctt gccaaagctcgc tctgcttct	4620
cttcgatctt cgcacgacg gcgaggatcg tggcatcacc gaaccgcgcc gtgcgcgggt	4680
cgtcgggtgag ccagagtttc agcaggccgc ccaggcgccc caggctcgcca ttgatgcggg	4740
ccagctcgcg gacgtgctca tagtccacga cgcgcgtgat tttgtagccc tggccgacgg	4800
ccagcaggta ggcgcacag ctcatgccg cgcgcgccgc ctttccctca atcgctcttc	4860
gttcgtctgg aaggcagtac accttgatag gtggcgctgc cttcctggtt ggcttggttt	4920
catcagccat ccgcttgccc tcactgttta cgcgcggcgt agccggccag cctcgcagag	4980
caggattccc gttgagcacc gccaggcggc aataaggagc agtgaagaag gaacacccgc	5040
tcgcgggtgg gcctacttca cctatcctgc ccggctgacg ccgttgata caccaaggaa	5100
agtctacacg aaccctttgg caaaatcctg tatactgctc gaaaaaggat ggatataccg	5160
aaaaaatcgc tataatgacc ccgaagcagc gttatgcagc ggaaaagcgc tgcttccctg	5220
ctgtttttgt gaatatctac cgactggaaa caggcaaatg caggaaatta ctgaactgag	5280
gggacaggcg agagacgatg ccaaagagct acaccgacga gctggccgag tgggttgaat	5340
cccgcgcggc caagaagcgc cggcgtgatg aggctgcggg tgcgttcctg gcggtgaggg	5400
cggatgtcga ggcggcgcta gcgtccggct atgcgctcgt caccatttgg gagcacatgc	5460
gggaaacggg gaaggcgaag ttctcctacg agacgttccg ctgcacgcc aggcggcaca	5520
tcaaggccaa gcccgccgat gtgccgcac cgcaggccaa ggctgcggaa cccgcgcggg	5580
cacccaagac gccggagcca cggcgccga agcagggggg caaggctgaa aagccggccc	5640
ccgctgcggc cccgaccggc ttcacctca acccaacacc ggacaaaaag gatctactgt	5700
aatggcgaaa attcacatgg ttttgcaggc caagggcggg gtccggcaagt cggccatcgc	5760
cgcgatcatt gcgcagtaca agatggacaa ggggcagaca cccttgctga tcgacaccga	5820
cccggcgaac gcgacgttcg agggctacaa ggccctgaac gtccgcggcg tgaacatcat	5880

ES 2 843 556 T3

ggcggcgac	gaaattaact	cgcgcaactt	cgacaccctg	gtcgagctga	ttgcgcccac	5940
caaggatgac	gtggtgatcg	acaacgggtg	cagctcggtt	gtgcctctgt	cgcattacct	6000
catcagcaac	caggtgcccg	ctctgctgca	agaaatgggg	catgagctgg	tcatccatac	6060
cgtcgtcacc	ggcggccagg	ctctcctgga	cacggtgagc	ggcttcgccc	agctcgccag	6120
ccagttcccc	gccgaagcgc	ttttcgtggt	ctggctgaac	ccgtattggg	ggcctatcga	6180
gcatgagggc	aagagctttg	agcagatgaa	ggcgtacacg	gccaacaagg	cccgcgtgtc	6240
gtccatcatc	cagattcccg	ccctcaagga	agaaacctac	ggccgcgatt	tcagcgacat	6300
gctgcaagag	cggctgacgt	tcgaccaggc	gctggccgat	gaatcgctca	cgatcatgac	6360
gcggcaacgc	ctcaagatcg	tgccggcgcg	cctgtttgaa	cagctcgacg	cgccggccgt	6420
gctatgagcg	accagattga	agagctgac	cgggagattg	cgcccaagca	cgccatcgcc	6480
gtcggcccg	acgaccgggt	gctgacccgt	cataccatca	acgcccgggt	catggccgac	6540
agtgcggcca	agcaagagga	aatccttgcc	gcgttcaagg	aagagctgga	agggatcgcc	6600
catcgttggg	gcgaggacgc	caaggccaaa	gcggagcgga	tgctgaacgc	ggccctggcg	6660
gccagcaagg	acgcaatggc	gaaggtaatg	aaggacagcg	ccgcgcaggc	ggccgaagcg	6720
atccgcaggg	aaatcgacga	cgcccttggc	cgccagctcg	cgcccaagggt	cgccgacgcg	6780
cgccgcgtgg	cgatgatgaa	catgatcgcc	ggcgccatgg	tggtgttcgc	ggccgcccctg	6840
gtggtgtggg	cctcgttatg	aatcgagag	gcgcagatga	aaaagcccgg	cggtgcgggg	6900
ctttgttttt	gcgttagctg	ggcttgtttg	acaggcccaa	gctctgactg	cgcccgcgct	6960
cgcgctcctg	ggcctgtttc	ttctcctgct	cctgcttgcg	catcagggcc	tggtgccgtc	7020
gggctgcttc	acgcatcgaa	tcccagtcgc	cgccagctc	gggatgctcc	gcgcgcatct	7080
tgccgctcgc	cagttcctcg	atcttggggc	cgtgaatgcc	catgccttcc	ttgatttcgc	7140
gcaccatgtc	cagccgcgtg	tgccaggtct	gcaagcgggc	ttgctgttgg	gcctgctgct	7200
gctgccaggc	ggcctttgta	cgcggcaggg	acagcaagcc	gggggcattg	gactgtagct	7260
gctgcaaacg	cgcctgctga	cggctctacga	gctgttctag	gggtcctcg	atgcgctcca	7320
cctggtcattg	ctttgcctgc	acgtagagcg	caagggtctg	ctggtaggtc	tgctcgatgg	7380
gcgcggattc	taagagggcc	tgctgttcgg	tctcggcctc	ctgggcccgc	tgtagcaaat	7440
cctcgccgct	gttgccgctg	gactgcttta	ctgcggggga	ctgctgttgc	cctgctcgcg	7500
ccgtcgtcgc	agttoggctt	gccccactc	gattgactgc	ttcatttcga	gccgcagcga	7560
tgccgatctc	gattgcgtca	acggacgggg	cagcgcggag	gtgtccggct	tctccttggg	7620
tgagtcggtc	gatgccatag	ccaaagggtt	ccttccaaaa	tgcttccatt	gctggaccgt	7680
gtttctcatt	gatgcccgca	agcatcttcg	gcttgaccgc	caggtcaagc	gcgccttcat	7740
ggcggtcat	gacggacgcc	gccatgacct	tgccgcggtt	gttctcgatg	tagccgcgta	7800

ES 2 843 556 T3

atgaggcaat ggtgccgccc atcgtcagcg tgtcatcgac aacgatgtac ttctggccgg	7860
ggatcacctc cccctcgaaa gtcgggttga acgccaggcg atgatctgaa ccggctccgg	7920
ttcggggcgac cttctcccgc tgcacaatgt ccgtttcgac ctcaaggcca aggcggtcgg	7980
ccagaacgac cgccatcatg gccggaatct tgttgttccc cgccgcctcg acggcgagga	8040
ctggaacgat gcggggcttg tcgtcgccga tcagcgtctt gagctgggca acagtgtcgt	8100
ccgaaatcag gcgctcgacc aaattaagcg ccgcttcgcg gtgcgcctgc ttgcgagcct	8160
ggtattcagg ctcggttggtc aaagaaccaa ggtcgccggt gcgaaccacc ttcggaagt	8220
ctccccacgg tgcgcgctcg gctctgctgt agctgctcaa gacgcctccc tttttagccg	8280
ctaaaactct aacgagtgcg cccgcgactc aacttgacgc tttcggcact tacctgtgcc	8340
ttgccacttg cgtcataggt gatgcttttc gactcccgga tttcaggtac tttatcgaaa	8400
tctgaccggg cgtgcattac aaagtcttc cccacctggt ggtaaatgct gccgctatct	8460
gcgtggacga tgctgcgctc gtggcgctgc gacttatcgg ccttttgggc catatagatg	8520
ttgtaaatac cagggttcag ggccccggct ttatctacct tctggttcgt ccatgcgcct	8580
tggtttctcg tctggacaat tctttgcca ttcatgacca ggaggcggtg tttcattggg	8640
tgactcctga cgggtgcctc tgggtgtaaa cgtgtcctgg tcgcttgccg gctaaaaaaa	8700
agccgacctc ggcagttoga ggcgggcttt ccctagagcc gggcgcgta aggttgttcc	8760
atctatttta gtgaactgcg ttcgatttat cagttacttt cctcccgctt tgtgtttcct	8820
cccactcgtt tccgcgtcta gccgaccct caacatagcg gcctcttctt gggctgcctt	8880
tgctcttgcc cgcgcttcgt cagcgtcggc ttgcaccgtc gtaaagcgct cggcctgcct	8940
ggcgcctct tgcgccgcca acttcctttg ctctggtgg gcctcggcgt cggcctgcgc	9000
cttcgctttc accgctgcca actcgtgcg caaatctcc gcttcgcgcc tgggtggcgc	9060
gcgctcgccg cgaagcgctt gcatttcctg gttggccgcg tccagggtct tgcggctctc	9120
ttctttgaat gcgcgggctt cctggtgagc gtagtccagc tcggcgcgca gctcctgcgc	9180
tcgacgctcc acctcgtcg cccgctgcgt cgccagcgcg gcccgctgct cggctcctgc	9240
cagggcggtg cgtgcttcgg ccagggttg ccgctggcgt gcggccagct cggccgcctc	9300
ggcggcctgc tgctctagca atgtaacgcg cgcctgggct tcttcagct cgcgggctg	9360
cgcctcgaag gcgtcgcca gctccccgcg cagggcttcc aactcgttgc gctcacgac	9420
ccagccggct tgcgctgcct gcaacgattc attggcaagg gcctgggcgg cttgccagag	9480
ggcggccacg gcctggttgc cggcctgctg caccgcgtcc ggcacctgga ctgccagcg	9540
ggcggcctgc gccgtcgct ggcgtcgcca ttgcgcgatg ccggcgctgg cgtcgttcat	9600
gttgacgcgg gcggccttac gcactgcac caccgtcggg aagttctccc ggtcgccttg	9660

ES 2 843 556 T3

ctcgaacagc	tcgtccgcag	ccgcaaaaat	gcggtcgcgc	gtctctttgt	tcagttccat	9720
gttggctccg	gtaattggta	agaataataa	tactcttacc	taccttatca	gcgcaagagt	9780
ttagctgaac	agttctcgac	ttaacggcag	gttttttagc	ggctgaaggg	caggcaaaaa	9840
aagccccgca	cggtcggcgg	gggcaaaggg	tcagcgggaa	ggggattagc	gggcgtcggg	9900
cttcttcatg	cgtcggggcc	gcgcttcttg	ggatggagca	cgacgaagcg	cgcacgcgca	9960
tcgtcctcgg	ccctatcggc	ccgcgtcgcg	gtcaggaact	tgtcgcgcgc	taggtcctcc	10020
ctgggtgggca	ccaggggcat	gaactcggcc	tgtctgatgt	aggtccactc	catgaccgca	10080
tcgcagtcga	ggccgcgttc	cttcaccgtc	tcttgacaggt	cgcggtaacg	ccgctcgttg	10140
agcggctggt	aacggggccaa	ttggtcgtaa	atggctgtcg	gccatgagcg	gcctttcctg	10200
ttgagccagc	agccgacgac	gaagccggca	atgcaggccc	ctggcacaac	caggccgacg	10260
ccggggggcag	gggatggcag	cagctcgcca	accaggaacc	ccgccgcgat	gatgccgatg	10320
ccggtcaacc	agcccttgaa	actatccggc	cccgaacac	ccctgcgcac	tgccctggatg	10380
ctgcgccgga	tagcttgcaa	catcaggagc	cgtttctttt	gttcgtcagt	catggtccgc	10440
cctcaccagt	tgttcgtatc	ggtgtcggac	gaactgaaat	cgcaagagct	gccggtatcg	10500
gtccagccgc	tgctcgtgtc	gctgctgccg	aagcacggcg	aggggtccgc	gaacgccgca	10560
gacggcgat	ccggccgcag	cgcacgcgcc	agcatggccc	cggtcagcga	gccgccggcc	10620
aggtagccca	gcatggtgct	gttggtcgcc	ccggccacca	gggccgacgt	gacgaaatcg	10680
ccgtcattcc	ctctggattg	ttcgtcgctc	ggcggggcag	tgccgccgcg	cggcggcgtc	10740
gtggatggct	cgggttggtc	ggcctgcgac	ggccggcgaa	aggtgcgcag	cagctcgtaa	10800
tcgaccggt	gcggcgctcg	ggccgccgcc	ttgcgtcgcg	gtcgggtgtc	cttcttcggc	10860
tcgcgcagct	tgaacagcat	gatcgcggaa	accagcagca	acgccgcgcc	tacgcctccc	10920
gcgatgtaga	acagcatcgg	attcattctt	cggtcctcct	tgtagcggaa	ccgttgtctg	10980
tgccggcgcg	gtggcccgcg	ccgctgtctt	tggggatcag	ccctcgatga	gcgcgaccag	11040
tttcacgtcg	gcaagggttc	cctcgaaact	ctggccgctc	tcctcgtaact	tcaaccaggc	11100
atagccttcc	gccggcgggc	gacggttgag	gataaaggcg	gcagggcgct	cgtcgtgctc	11160
gacctggacg	atggcctttt	tcagcttgct	cgggtccggc	tccttcgcgc	ccttttcctt	11220
ggcgtcctta	ccgtcctggg	cgcgcctctc	gccgtcctgg	ccgtcgcggg	cctccgcgct	11280
acgctcggca	tcagtctggc	cgttgaaggc	atcgacgggt	ttgggatcgc	ggcccttctc	11340
gtccaggaac	tcgcgcagca	gcttgaccgt	gccgcgcgtg	atttcctggg	tgctcgtcgtc	11400
aagccacgcc	tcgacttcct	ccgggcgctt	cttgaaggcc	gtcaccagct	cgttcaccac	11460
ggtcacgtcg	cgcacgcggc	cgggtgtgaa	cgcacggcg	atcttctccg	gcaggtccag	11520
cagcgtgacg	tgctgggtga	tgaacgcggg	cgacttgccg	atttccttgg	cgatatcgcc	11580

ES 2 843 556 T3

tttcttcttg cccttcgcca gctcgcggcc aatgaagtcg gcaatttcgc gcggggtcag	11640
ctcgttgctg tgcaggttct cgataacctg gtcggcttcg ttgtagtcgt tgtcgatgaa	11700
cgccgggatg gacttcttgc cggcccactt cgagccacgg tagcggcggg cgccgtgatt	11760
gatgatatag cggcccggct gtccttggtt ctccgcgacc gaaatgggtg acttcacccc	11820
gcgctctttg atcgtggcac cgatttcgcg gatgctctcc ggggaaaagc cggggttgtc	11880
ggcgttcgcg ggctgatgcg gatcttcgtc gatcaggctc aggtccagct cgatagggcc	11940
ggaaccgccc tgagacgcg caggagcgtc caggaggctc gacaggctcg cgatgctatc	12000
caaccccagc cgggacggct gcgcgcgcgc tgcggcttcc tgagcggccg cagcgggtt	12060
tttcttggtg gtcttggtt gagccgcagt cattgggaaa tctccatctt cgtgaacacg	12120
taatcagcca gggcggaac ctctttcgat gccttcgcgc cggccgtttt cttgatcttc	12180
cagaccggca caccggatgc gagggcatcg gcgatgctgc tgcgcaggcc aacggtggcc	12240
ggaatcatca tcttggggtg cgcggccagc agctcggctt ggtggcgcgc gtggcgcgga	12300
ttccgcgcgt cgaccttgct gggcaccatg ccaaggaatt gcagcttggc gttcttctgg	12360
cgacggttcg caatggctgt gaccatcttc ttgatgccct ggatgctgta cgcctcaagc	12420
tcgatggggg acagcacata gtcggccgcg aagagggcgg ccgccaggcc gacgccaagg	12480
gtcggggccg tgtcgatcag gcacacgtcg aagccttggt tcgccagggc cttgatgttc	12540
gccccgaaca gtcgcggggc gtcgtccagc gacagccgtt cggcgttcgc cagtaccggg	12600
ttggactcga tgagggcgag gcgcgcggcc tggccgtcgc cggctgcggg tgcggtttcg	12660
gtccagccgc cggcagggac agcgcgaac agcttgcgtt catgcaggcc ggtagcaaag	12720
tccttgagcg tgtaggacgc attgccctgg ggggtccaggc cgatcacggc aaccgcgaag	12780
ccgcgctcga aaaagtcgaa ggcaagatgc acaagggtcg aagtcttgcc gacgccgcct	12840
ttctgggttg ccgtgaccaa agttttcatc gtttggtttc ctgttttttc ttggcgtccg	12900
cttcccactt ccggacgatg tacgcctgat gttccggcag aaccgccgtt acccgcgct	12960
accctcggg caagttcttg tcctcgaacg cggcccacac gcgatgcacc gcttgcgaca	13020
ctgcgcccct ggtcagtcgc agcgacgttg cgaacgtcgc ctgtggcttc ccatcgacta	13080
agacgccccg cgctatctcg atggtctgct gccccacttc cagcccctgg atcgctcct	13140
ggaactggct ttcggtgaag cgtttcttca tggataacac ccataatttg ctccgcgcct	13200
tggttgaaca tagcgtgac agccgcgac acatgagaga agtttagcta aacatttctc	13260
gcacgtcaac acctttagcc gctaaaactc gtccttggcg taacaaaaca aaagcccgga	13320
aaccgggctt tcgtctcttg ccgcttatgg ctctgcaccc ggctccatca ccaacaggtc	13380
gcgcacgcgc ttactcggg tgcggatcga cactgccagc ccaacaaaagc cggttgccgc	13440

ES 2 843 556 T3

```

cgccgccagg atcgcgccga tgatgccggc cacaccggcc atcgcccacc aggtcgccgc 13500
cttcgggttc cattcctgct ggtactgctt cgcaatgctg gacctcggtt caccataggc 13560
tgaccgctcg atggcgtagt ccgcttctcc ccttggcgta aaaccagcg ccgcaggcgg 13620
cattgccatg ctgcccgccg ctttcccgac cagcagcgcg gcaccaggct tgcggtccag 13680
accttcggcc acggcgagct gcgcaaggac ataatcagcc gccgacttgg ctccacgcgc 13740
ctcgatcagc tcttgctact gcgcgaaatc cttggcctcc agggccgcca tgaatcgcg 13800
acgcggcgaa ggctccgcag ggccggcgtc gtgatcgccg ccgagaatgc ccttcaccaa 13860
gttcgacgac acgaaaatca tgctgacggc tatcaccatc atgcagacgg atcgcacgaa 13920
cccgtgaat tgaacacgag cagggcaccg gcgaccacta tgccaagaat gcccaaggta 13980
aaaattgccg gccccgccat gaagtccgtg aatgccccga cggccgaagt gaagggcagg 14040
ccgccaccca ggccgcccgc ctactgccc ggcaacctgt cgctgaatgt cgatgccagc 14100
acctgoggca cgtcaatgct tccgggcgtc gcgctcgggc tgatcgccca tcccgttact 14160
gccccgatcc cggcaatggc aaggactgcc agcgtgccta tttttggggg gaggccgttc 14220
gcggccgagg ggccgcagccc ctggggggat gggaggcccg cgtagcggg ccgggagggg 14280
tcgagaaggg ggggcacccc ccttcggcgt gcgcggtcac gcgcacaggg cgcagccctg 14340
gttaaaaaca aggtttataa atattggttt aaaagcagg taaaagacag gttagcgggtg 14400
gccgaaaaac gggcggaac ctttgcaaat gctggatttt ctgcctgtgg acagcccctc 14460
aaatgtcaat aggtgcgccc ctcatctgtc agcactctgc ccctcaagt tcaaggatcg 14520
cgcccctcat ctgtcagtag tcgcgcccct caagtgtcaa taccgcaggg cacttatccc 14580
caggcttgct cacatcatct gtgggaaact cgcgtaaaat caggcgtttt cgccgatttg 14640
cgaggctggc cagctccacg tcgccggccg aaatcgagcc tgcccctcat ctgtcaacgc 14700
cgcgccgggt gagtcggccc ctcaagtgtc aacgtccgcc cctcatctgt cagtgagggc 14760
caagttttcc gcgaggtatc cacaacgccg gcggccgcgg tgtctcgcac acggcttcga 14820
cgcggtttct ggcgcgtttg cagggccata gacggccgcc agcccagcgg cgagggcaac 14880
cagcccgggt agcgtcgga aggcgctgga agccccgtag cgacgcggag aggggcgaga 14940
caagccaagg gcgcaggctc gatgcgcagc acgacatagc cggttctcgc aaggacgaga 15000
atttcctgc ggtgcccctc aagtgtca 15028
<210> 39
<211> 1011
<212> ADN
5 <213> Escherichia coli
<400> 39
atcgctcac gcaactggtc cagaaccttg accgaacgca gcggtggtaa cggcgcagtg 60

```

ES 2 843 556 T3

	gcggttttca tggcttgta tgactgtttt tttggggtac agtctatgcc tcgggcatcc	120
	aagcagcaag cgcggttacgc cgtgggtcga tgtttgatgt tatggagcag caacgatgtt	180
	acgcagcagg gcagtcgccc taaaacaaag ttaaacaatca tgagggaagc ggtgatcgcc	240
	gaagtatcga ctcaactatc agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg	300
	ttgctggccg tacatttgta cggctccgca gtggatggcg gcctgaagcc acacagtgat	360
	attgatttgc tggttacggt gacggttaagg cttgatgaaa caacgcggcg agctttgatc	420
	aacgaccttt tggaaacttc ggcttcccct ggagagagcg agattctccg cgctgtagaa	480
	gtcaccattg ttgtgcacga cgacatcatt ccgtggcggt atccagctaa gcgcgaactg	540
	caatttggag aatggcagcg caatgacatt cttgcaggta tcttcgagcc agccacgatc	600
	gacattgatc tggctatctt gctgacaaaa gcaagagaac atagcgttgc cttggtaggt	660
	ccagcggcgg aggaactctt tgatccggtt cctgaacagg atctatttga ggcgctaaat	720
	gaaaccttaa cgctatggaa ctgcgcgccc gactgggctg gcgatgagcg aaatgtagtg	780
	cttacgttgt cccgcatttg gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc	840
	gctgccgact gggcaatgga gcgcctgccg gccagtatc agcccgtcat acttgaagct	900
	agacaggctt atcttggaaga agaagaagat cgcttggcct cgcgcgcaga tcagttggaa	960
	gaatttgtcc actacgtgaa aggcgagatc accaaggtag tcggcaaata a	1011
	<210> 40	
	<211> 816	
	<212> ADN	
5	<213> Escherichia coli	
	<400> 40	
	atgagccata ttcaacggga aacgtcttgc tcgaggccgc gattaaattc caacatggat	60
	gctgatattat atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc	120
	tatcgattgt atgggaagcc cgatgcgcca gaggttgttc tgaaacatgg caaaggtagc	180
	gttgccaatg atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct	240
	cttccgacca tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg	300
	atccccggga aaacagcatt ccaggatta gaagaatatc ctgattcagg tgaaaatatt	360
	gttgatgcgc tggcagtggt cctgcgcggg ttgcattcga ttcctgtttg taattgtcct	420
	tttaacagcg atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg	480
	gttgatgcga gtgattttga tgacgagcgt aatggctggc ctggtgaaca agtctggaaa	540
	gaaatgcata agcttttggc attctcaccg gattcagtcg tcaactcatgg tgattttctca	600
	cttgataacc ttatttttga cgaggggaaa ttaatagggt gtattgatgt tggacgagtc	660
	ggaatcgcag accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct	720
	ccttcattac agaaacggct ttttcaaaaa tatggtattg ataatcctga tatgaataaa	780
	ttgcagtttc atttgatgct cgatgagttt ttctaa	816
10	<210> 41	
	<211> 795	
	<212> ADN	
	<213> Escherichia coli	
	<400> 41	

atggttgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggtatttc 60
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggtgtca 120
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180
caggacgagg cagcgcggt atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 240
ctcgacgttg tcaactgaagc gggaaggac tggttgctat tgggcgaagt gccggggcag 300
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360
cggcggtgac atacgcttga tccggtacc tgccattcg accaccaagc gaaacatcgc 420
atcgagcgag cagctactcg gatggaagcc ggtcttgctg atcaggatga tctggacgaa 480
gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac 540
ggcgatgatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 600
ggccgctttt ctggattcat cgactgtggc cggttggtg tggcggaccg ctatcaggac 660
atagcgttgg ctaccctga tattgtgaa gagcttggcg gcgaatgggc tgaccgcttc 720
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780
gacgagttct tctga 795

<210> 42

<211> 492

<212> ADN

5 <213> Agrobacterium tumefaciens

<400> 42

atgcactcaa ccagattgag aacaaacgaa gtcgcgggtca cgctaccgga agaaggcgac 60
gcacgcgttt tctttgttgg aaagatccgg accccttga aaagtcgagc tgacactcct 120
cggcagggca gtgaatccgg gccaccatgc acacttgaaa tttcagaacc gtgggcccgtg 180
gcgctaaagg gggtcgaggc gtattcgaa cttgaggttc tatattggtt gcacgagtca 240
ccacgcgata tcgtcctact aagtcggct gacgatggag agattcacgg tgctttctcc 300
ctacgagcac ctgtgaggcc caatccgatc ggaacatcga ttgtgaaagt ggagagggtg 360
tctgggaact cgattgttgt acgtggtctt gattgcctgg atggaacacc gctgttgat 420
ataaaaccgg accgctcgt atcaaagccg ctagccccgg tgaggaaaag gttctcgatt 480
cctgccgtct ga 492

<210> 43

<211> 1269

<212> ADN

10 <213> Agrobacterium tumefaciens

<400> 43

ES 2 843 556 T3

atgattgcga acagctcttc cgatgtctct atggccgacc agaagtttct aaatgtcgcg	60
aagtcaaattg aaatcgatcc cgacgccgtt cctataagca gacttgattc tgaaggtcac	120
agtatttttg cagaatggcg accgaagcgc ccgtttcttc ggagagaaga tggcgtcttt	180
ctcgtttctcc gcgcagacca tatctttctg ctgggcaccg atccacgcac ccggcaaata	240
gaaactgagc tcatgctgaa tcgaggcgtc aaagctgggtg ccgtttttga ctttataggt	300
cacagcatgt tgttttcgaa cgggtgaaacg catgggaagc ggcgctcagg tctttcaaag	360
gcgtttctcat tccgcatggt tgaagcgta cgcgccgaga ttgcgaagat aacggagcgt	420
ttgtgggacg aactacaaaa agttgacgat ttcaatttta ctgaaatgta cgcgtcgcaa	480
ttgcctgcgc tgacgatcgc aagtgtcctt ggcttgccgt ctgaggacac gccgtttttc	540
acacgacttg tttataaggt ttcccgcgtc ttgagcccgt cgtggcgaga tgaggaattc	600
gaggagattg aagcttccgc tatcgagctt caggattacg ttcggggcgt gatcgcggt	660
agtggccgtc ggatgaggga tgattttctc tcgcgctact tgaaggcggc acggaagcc	720
ggaacgcttt cgcccattga ggagatcatg caactcatgc ttatcatact cgctgggagc	780
gatacaacgc gcaactgcaat ggttatggtg acggctcttg tgctccaaa tcccgcgctc	840
tggtcttctt taactggcaa tcaatcctat gtcgcagccg ccgtggagga agggctccga	900
ttcagaccgc cagttggctc ttttccgcgg ttggccctcg aggatatcga tctggatgga	960
tacgtgttgc caaagggaag cctcctcgcg ctcaagtgtca tgtctggcct gcgagacgaa	1020
aaacactacg agaaccctca tctttttgat gttgggcgtc aacaaatgcg ttggcacctc	1080
gggtttggcg cgggagttca tcgttgccctc ggcgagacgt tggctcggat tgaattgcaa	1140
gaaggacttg gaacactttt gcgccgcgcg ccgaatctta cgggtgtagg tgactggcca	1200
cggatgatgg gtcacggagg catccggcgc gccaccggca tgacggtcaa attgagcgtc	1260
gaccggtga	1269
<210> 44	
<211> 1224	
<212> ADN	
5 <213> Agrobacterium tumefaciens	
<400> 44	
gtggaagaga ggcgcgtatc aatatcaagc atcacttgga gatttccaat gccatgtcgc	60
ccagtggacg acgtgacgac catcgatgac ctaacactcg atccctatcc gatttatcgc	120

ES 2 843 556 T3

	agaatgcgcg tgcaaaaccc ggtggtacac gtcgcctcgg tcagacggac gtttctaacg	180
	aaggcctgcg atactaagat ggtgaaggat gatccttcac gcttttagttc tgatgatccc	240
	aacacgccga tgaagccggc gcttcaggct catactttga tgcgtaagga tggagttgaa	300
	cacgctcggg aacgaatggc catgactaga gcattcgcac ccaaagcgat cgcggagcat	360
	tgggtgccga tctatcgtga catcgtgaac gagtatctag atcgccttcc gcgcggagac	420
	accgtggatc tttttgcaga gatcgttggt ccggttgctg ctgcataact ggcacatatt	480
	ctaggaatth gtgaggcatc agatgtcag atgatccgct ggtcgcaacg gcttattgac	540
	ggagccggca acttcggatg gcgacccgag ctctttgagc gatcggacga ggccaatacg	600
	gagatgaact gcctcttcaa cgacttggtc gaaaaacacc gatccgcgcc gaaccctcgc	660
	gcttttgca tcatgttgaa tgcacctgat cctatacccc taagccagat ctatgccaac	720
	atcaagatcg ccatcggcgg aggggtcaac gaggcctcgt acgctctcgg cacgatcctc	780
	tatggattgc tgaccaatcc ggagcaactc gaggaggtca agagacagca atgctggggg	840
	caggccttcg aggaagggct gcgctgggtt gctccgattc aagcgagttc gcgtctcgtg	900
	cgggagggaa cggaatttcg tggctttatc gttccgaagg gcgacattgt gatgactatt	960
	caggcctccg ccaatcgtga tgaagatgtc tttgaggatg gagaaagctt caatgtcttt	1020
	cgcccgaa atgcgcatca gtcctttggc tccggcccg c atcactgccc aggcgcgcaa	1080
	atctcgggc aaaccgtcgg cgcgatcatg ctgccatac tgttcgatcg atttccggat	1140
	atgatattgc cccaccctga attagtgcag tggcgcggtc tcggcttccg cgggcccatt	1200
	aatctacctg tgacactaag atga	1224
	<210> 45	
	<211> 438	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 45	
	atgaaacgta tcagtacaat ccttgtcggc gtatthtttg ctgcgccggt ttatgcagcc	60
	gacaatatth atacattagg gacgttgtct gaaattgaac tagcgctcac tgccggaaag	120
	ccggtgaacg tcacggttga ttttaagtttg tgtgctcccg ggggtggccga tactccggcg	180
	acgaaaacgc aaggaggcat gcgcatagat gcgtatcgaa taacaactga cggcaccctt	240
	gcgtttgcag atcagcactt caccattgac cgtgacggta aaccataac tcaatthatt	300
	cgttatcaga ttccggtcaaa cggtagcgcc gatttcacga tggttacatt caacatgcc	360
	acatatgagc gaaaaggtag cagcttgccc tacaagtgcg cgatcgacca cgggttgagt	420
	tttcgtactc cacagtga	438
	<210> 46	
	<211> 1134	
	<212> ADN	
10	<213> Agrobacterium tumefaciens	
	<400> 46	

ES 2 843 556 T3

```

atggatcatc gaaatcgccc gctggacgag agacagcagc actgggttaa cagcgtttgg      60
gagaggctct gtcgaagcgt cgcatctgcc caacatccag attccgcccg cctgccaatc      120
gcttcataatc aaagtgcacat cgcaacaagg ctgagagaaac acagtgatct cagttcgttc      180
gaaacacgtc ccgaatatat gcgagagttt tcgaacaagt tttcatggaa tttaaagcgc      240
gacattttctt ttgtgggttaa tcgcaagagt aagaacgggtg aattccgaga agatgcgttt      300
gaagagctgt tttgcgaaag gataacgtct gtccatcatt ctcttcagga ggcgatcacg      360
gcacgggttg aatccgccgt cggaaaatcg aatgacgaag gcgatgtacg ttctcgcgat      420
gtttcgacat tttacgctcg cactgctggg gagagttcta caaccaagcg attgaacatc      480
agcacaacc cagcaggcgt ggtgccgagt gctgacgtcg tcaatatctc tgaacagcag      540
gcacagtctg tcaagtatcg gcgtcctaag tttctcgttg aggtctcaaa gcacgtctat      600
agaccaatcg aacagtataa tcggatggtc ttcgatcgag tcagatttcc aaacatcact      660
cgtgaccggc gtaagttgaa tgtattgatc cgcacggaag acggcactgg ctacgaaaca      720
cctagctcgt taaaaacgcg cctgggatcg gagcagcata atctcttcac caagacaatg      780
aacgagggtc agccacccca tagtctcgcg tccacgaaga ctggtaaact gcctatagtt      840
ttggttgacg cgggaaatgg tgcaaccttt atcgccggcg agaaattccc tttgttgac      900
aacggacgcc ggataccaac gactatttct gccgatagaa ttttagtcag aaatgaaaat      960
gggactttca gcgagctttg tcagcgtttg gctggaaaca ccgaactatt gcttccacca      1020
gaagcgatag ttcaactggg actgcagcgt caacaagctg tcaatcgagc tcgtgccaga      1080
gaagcatcgg agttggagca acgaactcgc gaacattcct ctggaagagt ttga      1134

```

<210> 47

<211> 588

<212> ADN

5 <213> Agrobacterium tumefaciens

<400> 47

```

atgagagatg atccaggagt tgctcgagcg ctttcctctc cgttacaaga taagcacgac      60
attttcgaga cccagatcat tccggcaatc ggtgcgcggg ggccgtctgt tgcaaagcca      120
actttcaccg ttgtatgcgg ccagcagggt gccggaaaga gcaccttagt tcgccaatc      180
aagagtagaa cgggcggtga aagcaccacg cgaataatag ctgatgatct taatgcatac      240
attcctggta acaatgcagc cttctgcag ggctcgcatg cgttagagcg agcgaattcg      300
acggcagtga cagaatggta tcatcaatta tttgatcgca gtatcactaa tagatataat      360
attatactgg agtcttgcta tccgccaat cattacgct cgctcttaga taaggcgcgc      420
tccaatggat acagaacaga actaaatatt atcgcgacgg acaggataac gagctttaca      480
gctattcacg atcgctttga aagggcactg gcgaactcgt ttattgcctc aaccgtattg      540
ccagatgtcg aaactcacga tcattattat tcatttggcc acgtgtag      588

```

<210> 48

<211> 633

<212> ADN

<213> Agrobacterium tumefaciens

<400> 48

ES 2 843 556 T3

	atggctaaaa atgatgctta taagctaatt atatttgatt ttgatggtac attggccgac	60
	agcggcgcct ggatgatacg cgctctgagg gaaatgtccg aacgtcaccg attcgtatcg	120
	cccagcgaca ctcagattga atatctgcgt ggcccttcag tttgccaggt tctgaagtgg	180
	atgagagtac cgggtgtggcg cattcctctg atcgtgagag acttgcgtaa actggctcgc	240
	gaagccacgt tcgatatgtt ccctaggaca gaaaaggtct tgtttctcct ccaagatcaa	300
	ggggctogagc ttgcaatctt aagttcgaac agcgttgaaa atattcagcg cgtgctoggg	360
	ccattggaga gatactttgc ttatatagag ggcagctcgc caatatttgg gaagggcaaa	420
	aggatcaata agatacttcg tcggtttgaa aggtcccga atgaagtgct tcttgtcgga	480
	gatgaggtac gtgatattga ggcagcgatg agtcaaaacg ttgcctccgc gggagttacg	540
	tggggttatg ctaagagaga ggctttggcg tgcagcaatc ccacgcacat cctagaagat	600
	attgaggaac taatagactt gaaacctctc taa	633
	<210> 49	
	<211> 213	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 49	
	atgaaaacgt cgacaaaaac aatgcttgtc tacgccgcga taattttcac tccaggattt	60
	atatattttc cactatataa aattttgctc actctttata gtgaagattc aggagataca	120
	tatttttgtc agttctatac ggagttctca tttttttgga taatgggatt actggttttt	180
	gccttgctcaa cgatgcttaa aaagaagcct tag	213
	<210> 50	
	<211> 2101	
10	<212> ADN	
	<213> Salmonella typhimurium	
	<400> 50	
	gacagtaaga cgggtaagcc tgttgatgat accgctgcct tactgggtgc attagccagt	60
	ctgaatgacc tgtcacggga taatccgaag tggtcagact ggaaaatcag agggcaggaa	120
	ctgctgaaca gcaaaaagtc agatagcacc acatagcaga cccgccataa aacgccctga	180

ES 2 843 556 T3

gaagcccggtg acgggctttt cttgtattat gggtagtttc cttgcatgaa tccataaaag	240
gcgcctgtag tgccatttac cccattcac tgccagagcc gtgagcgag cgaactgaat	300
gtcacgaaaa agacagcgac tcaggtgcct gatggtcgga gacaaaagga atattcagcg	360
atttgcccga gcttgcgagg gtgctactta agcctttagg gttttaaggt ctgttttgta	420
gaggagcaaa cagcgtttgc gacatccttt tgtaatactg cggaaactgac taaagtagtg	480
agttatacac agggctggga tctattcctt ttatcttttt ttattccttc tttattctat	540
aaattataac cacttgaata taaacaaaaa aaacacacaa aggtctagcg gaatttacag	600
agggtctagc agaatttaca agttttccag caaaggctca gcagaattta cagataccca	660
caactcaaag gaaaaggact agtaattatc attgactagc ccatctcaat tggatatagt	720
attaaaatca cctagaccaa ttgagatgta tgtctgaatt agttgttttc aaagcaaatg	780
aactagcgat tagtcgctat gacttaacgg agcatgaaac caagctaatt ttatgctgtg	840
tggcactact caacccccagc attgaaaacc ctacaaggaa agaacggacg gtatcgttca	900
cttataacca atacgctcag atgatgaaca tcagtaggga aaatgcttat ggtgtattag	960
ctaaagcaac cagagagctg atgacgagaa ctgtggaaat caggaatcct ttggttaaag	1020
gctttgagat tttccagtgg acaaactatg ccaagttctc aagcgaaaaa ttagaattag	1080
tttttagtga agagatattg ccttatcttt tccagttaaa aaaattcata aaatataatc	1140
tggaacatgt taagtctttt gaaaacaaat actctatgag gatttatgag tggttattaa	1200
aagaactaac acaaaaagaa actcacaagg caaatataga gattagcctt gatgaattta	1260
agttcatggt aatgcttgaa aataactacc atgagtttaa aaggcttaac caatgggttt	1320
tgaaaccaat aagtaaagat ttaaacactt acagcaatat gaaattggtg gttgataagc	1380
gaggccgccc gactgatacg ttgattttcc aagttgaact agatagacaa atggatctcg	1440
taaccgaact tgagaacaac cagataaaaa tgaatggtga caaaatacca acaaccatta	1500
catcagattc ctacctacat aacggactaa gaaaaacact acacgatgct ttaactgcaa	1560
aaattcagct caccagtttt gaggcacaaat ttttgagtga catgcaaagt aagtatgac	1620
tcaatgggtc gttctcatgg ctacgcaaa aacaacgaac cacactagag aacatactgg	1680
ctaaatacgg aaggatctga ggttcttatg gctcttgat ctatcagtga agcatcaaga	1740
ctaacaaca aaagtagaac aactgttcac cgttacatat caaagggaaa actgtccata	1800
tgcacagatg aaaacggtgt aaaaaagata gatacatcag agctttttacg agtttttggt	1860
gcattcaaag ctgttcacca tgaacagatc gacaatgtaa cagatgaaca gcatgtaaca	1920
cctaatagaa caggtgaaac cagtaaaaca aagcaactag aacatgaaat tgaacacctg	1980
agacaacttg ttacagctca acagtcacac atagacagcc tgaaacaggc gatgctgctt	2040
atcgaatcaa agctgccgac aacacgggag ccagtgcgcg ctcccgtggg gaaaaaatca	2100
t	2101
<210> 51	
<211> 1554	
<212> ADN	
<213> Escherichia coli	
<400> 51	

ES 2 843 556 T3

```

atggaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc      60
cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca      120
ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga      180
gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc tactgatta      240
agcattggta actgtcagac caagtttact catatatact ttagattgat ttaaaacttc      300
atttttaatt taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc      360
cttaacgtga gttttcgttc cactgagcgt cagaccctt aataagatga tcttcttgag      420
atcgtttttg tctgcgcgta atctcttgct ctgaaaacga aaaaaccgcc ttgcagggcg      480
gtttttcgaa ggttctctga gctaccaact ctttgaaccg aggtaactgg cttggaggag      540
cgcagtcacc aaaacttgtc ctttcagttt agccttaacc ggcgcatgac ttcaagacta      600
actcctctaa atcaattacc agtggctgct gccagtgggtg cttttgcatg tctttccggg      660
ttggactcaa gacgatatgt accggataag gcgcagcggc cggactgaac ggggggttcg      720
tgcatacagt ccagcttggg gcgaactgcc tacccggaac tgagtgtcag gcgtggaatg      780
agacaaacgc ggccataaca gcggaatgac accggtaaac cgaaaggcag gaacaggaga      840
gcgcacgagg gagccgccag gggaaacgcc tggatatctt atagtctgtg cgggtttcgc      900
caccactgat ttgagcgtca gatttcgtga tgcttgtcag gggggcggag cctatggaaa      960
aacggctttg ccgcgccctt ctcaactccc tgtaagtat cttcctggca tcttccagga     1020
aatctccgcc ccgttcgtaa gccatttccg ctgcgcgcag tcgaacgacc gagcgtagcg     1080
agtcagtgag cgaggaagcg gaatatatcc tgtatcacat attctgctga cgcaccggtg     1140
cagccttttt tctcctgccg catgaagcac ttcaactgaca ccctcatcag tgccaacata     1200
gtaagccagt atacactccg ctacgcgtga ggtctgcctc gtgaagaagg tgttgctgac     1260
tcataccagg cctgaatcgc cccatcatcc agccagaaag tgaggagacc acggttgatg     1320
agagctttgt tgtaggtgga ccagttggtg attttgaact tttgctttgc cacggaacgg     1380
tctgcgttgt cgggaagatg cgtgatctga tccttcaact cagcaaaagt tcgatttatt     1440
caacaaagcc acgttggtgc tcaaaatctc tgatgttaca ttgcacaaga taaaaatata     1500
tcatcatgaa caataaaact gtctgcttac ataaacagta atacaagggg tggtt         1554

```

<210> 52

<211> 1576

<212> ADN

5 <213> Escherichia coli

<400> 52

ES 2 843 556 T3

gatcgctagt ttgttttgac tccatccatt agggcttcta aaacgccttc taaggccatg	60
tcagccgtta agtggtcctg tgtcactgaa aattgctttg agaggctcta agggcttctc	120
agtgcgttac atccctggct tgttgtccac aaccgttaaa ccttaaaagc tttaaaagcc	180
ttatatattc ttttttttct tataaaactt aaaaccttag aggctattta agttgctgat	240
ttatattaat tttattgttc aaacatgaga gcttagtacg tgaaacatga gagcttagta	300
cgttagccat gagagcttag tacgttagcc atgaggggtt agttcgttaa acatgagagc	360
ttagtacgtt aaacatgaga gcttagtacg tgaaacatga gagcttagta cgtactatca	420
acaggttgaa ctgctgatct tcagatctag cttaaaacag gtggcttttt aatcatcttt	480
gccaaagcatg gcgcgggttt ggggtaatat agcgactcat aaaagcgtaa aacatgagtg	540
gatagtacgt tgctaaaaca tgagataaaa attgactctc atgttattgg cgtaaagata	600
tacagaatga tgagggtttt ttatgagact caaggctcat atggacgtga acaaaaaaac	660
gaaaattcgc caccgaaacg agctaaatca caccctggct caacttcctt tgcccgcaaa	720
gcgagtgatg tatatggcgc ttgctcccat tgatagcaaa gaacctcttg aacgagggcg	780
agttttcaaa attagggctg aagacctgac agcgctcgcc aaaatcacc ccatcgcttg	840
ttatcgacaa ttaaaagagg gtggtaaatt acttggtgcc agcaaaattt cgctaagagg	900
ggatgatatc attgctttag cttaaagagct taacctgcc tttactgcta aaaactcccc	960
tgaagagtta gatcttaaca ttattgagtg gatagcttat tcaaagatg aaggatactt	1020
gtctttaaaa ttcaccagaa ccatagaacc atatatctct agccttattg ggaaaaaaa	1080
taaattcaca acgcaattgt taacggcaag cttacgctta agtagccagt attcatcttc	1140
tctttatcaa cttatcagga agcattactc taattttaag aagaaaaatt attttattat	1200
ttccgttgat gagttaaagg aagagttaac agcttatact tttgataaag atggaaatat	1260
tgagtacaaa taccctgact ttcctatttt taaaagggat gtgttaaata aagccattgc	1320
tgaattataa aagaaaacag aaatatcggt tgttggcttc actgttcattg aaaaagaagg	1380
aagaaaaatt agtaagctga agttcgaatt tgtcgttgat gaagatgaat tttctggcga	1440
taaagatgat gaagcttttt ttatgaattt atctgaagct gatgcagctt ttctcaaggt	1500
atgtgatgaa accgtacctc ccaaaaaagc taaggggtga tatatggcta aaatttacga	1560
tttccctcaa ggagcc	1576
<210> 53	
<211> 2400	
<212> ADN	
<213> Escherichia coli	
<400> 53	

ES 2 843 556 T3

cacctaaacg aaggccgggc cactcaccgc gccttttttg tacgctcata gggcagaaca	60
aaccaacgtt ttatctatac cgctacaggg tatttaattc ctatttaatc tgcgctagaa	120
tgaggcatgt ttaaccgaat ctgacgtttt ccctgcaaata gccaaaatac tatgcctatc	180
tccgggtttc gcgtgacggc caagaccgcg aaaacaaaaa atacggtttg ctggaatacg	240
cgaacgccaa aggccttcgcg ccgctacaga tcgaggaaga aattgccagc agagcaaagg	300
actggcgcaa gcgcaagctc ggagcaatca tcgaaaaggc cgagcgtggc gacgtgctac	360
tgacgcccga gattacgcgc attgccggtt ccgccctcgc cgccttgga attctcaaag	420
cggcgagcga gcgcggccta atcgtccatg tgaccaaaca gaagatcatc atggacggca	480
gcctacaaag cgacatcatg gcaaccgtgc ttggcttggc tgcacagatc gagcggcatt	540
tcattcaggc acgtaccacc gaggcgctac aagtcgccag agagcgcggc aagacgctcg	600
ggcgacccaa gggcagcaaa tcgagcgcct tgaagctgga cagccgtatt gatgaagtac	660
aggcatacgt gaaccttggc ttgccgcaaa gtcgcgcagc cgagttgtta ggcgtcagcc	720
ctcacacctt gcgcctgttc atcaaacgcc ggaacatcaa acccacaac actagaccaa	780
ccatcaccat gccggggagg gaacaacatg cctaagaaca acaaagcccc cgccatcgt	840
atcaacgaga tcatcaagac gagcctcgcg ctcgaaatgg aggatgcccg cgaagctggc	900
ttagtggct acatggcccg ttgccttgtg caagcgacca tgccccacac cgacccaag	960
accagctact ttgagcgac caatggcatc gtcaccttgt cgatcatggg caagccgagc	1020
atcggcctgc cctacggttc tatgccgcgc accttgcttg cttggatatg caccgagggc	1080
gtgcgaacga aagaccccg gttgaacctt ggccggtcgc aatcggaatt tctacaaagg	1140
ctcggaatgc acaccgatgg ccgttacacg gccacccttc gcaatcaggc gcaacgcctg	1200
ttttcatcca tgatttcgct tgccggcgag caaggcaatg acttcggcat tgagaacgctc	1260
gtcattgcca agcgcgcttt tctattctgg aatcccaagc ggccagaaga tcgggcgcta	1320
tgggatagca ccctcaccct cacaggcgat ttcttcgagg aagtcacccg ctacccggtt	1380
cctatccgaa tcgactacct gcatgccttg cggcagtcctc cgcttgcgat ggacatttac	1440
acgtggctga cctatcgctt gttcctgttg cgggccaaag gccgcccctt cgtgcaaatac	1500
ccttgggtcg ccctgcaagc gcaattcggc tcatcctatg gcagccgcgc acgcaactcg	1560
cccgaactgg acgataaggc ccgagagcgg gcagagcggg cagcactcgc cagcttcaaa	1620
tacaacttca aaaagcgcct acgcgaagtg ttgattgtct atcccaggc aagcgactgc	1680
atcgaagatg acggcgaatg cctgcgcatac aaatccacac gcctgcatgt caccgcgca	1740

ES 2 843 556 T3

	cccggaagg ggcctcgcat cggccccct ccgacttgac caggccaacg ctacgcttgg	1800
	cttggtcaag ccttcccatc caacagcccg ccgtcgagcg ggctttttta tccccggaag	1860
	cctgtggata gagggtagtt atccacgtga aaccgctaata gccccgaaa gccttgattc	1920
	acggggcttt ccggcccgct ccaaaaacta tccacgtgaa atcgctaata agggtagctg	1980
	aaatcgctaa tcggagtagc tgaaatcgct aataaggtca cgtgaaatcg ctaataaaaa	2040
	aggcacgtga gaacgctaata agccctttca gatcaacagc ttgcaaacac ccctcgctcc	2100
	ggcaagtagt tacagcaagt agtatgttca attagctttt caattatgaa tatatatatc	2160
	aattattggt cgcccttggc ttgtggacaa tgcgctacgc gcaccggctc cgcccggtga	2220
	caaccgcaag cgggttgccca ccgtcgagcg ccagcgctt tgcccaaac ccggcggccg	2280
	caacagatcg ttttataaat tttttttttt gaaaaagaaa aagcccgaaa ggcggcaacc	2340
	tctcgggctt ctggatttcc gatcaacgca ggagtcgttc ggaaagtagc tgttccagaa	2400
	<210> 54	
	<211> 599	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Construcción sintética	
	<400> 54	
	cggcggtgtg gatacctcgc ggaaaacttg gccctcactg acagatgagg ggcggacgtt	60
	gacacttgag gggccgactc acccggcgcg gcgttgacag atgaggggca ggctcgattt	120
	cggccggcga cgtggagctg gccagcctcg caaatcggcg aaaacgcctg attttacgcg	180
	agtttccac agatgatgtg gacaagcctg gggataagtg ccctgcggtg ttgacacttg	240
	aggggcgcga ctactgacag atgaggggcg cgatccttga cacttgaggg gcagagtgtc	300
	gacagatgag gggcgacact attgacattt gaggggctgt ccacaggcag aaaatccagc	360
	atttgcaagg gtttccgccc gtttttcggc caccgctaac ctgtctttta acctgctttt	420
	aaaccaatat ttataaacct tgtttttaac cagggctgcg ccctgtgcgc gtgaccgcgc	480
	acgccgaagg ggggtgcccc cccttctcga accctcccgg cccgctaacg cgggcctccc	540
	atccccccag gggctgcgcc cctcggccgc gaacggcctc accccaaaaa tggcagcgc	599
10	<210> 55	
	<211> 801	
	<212> ADN	
	<213> Escherichia coli	
	<400> 55	
	tgcgcgtccc ccttggtcaa attgggtata cccatttggg cctagtctag ccggcatggc	60
15	gcattacagc aatacgcaat ttaaatgcgc ctagcgcatt tccccgacct taatgcgcct	120

cgcgctgtag cctcacgccc acatatgtgc taatgtggtt acgtgtatTTt tatggaggtt 180
atccaatgag ccgcctgaca atcgacatga cggaccagca gcaccagagc ctgaaagccc 240
tgcccgccctt gcagggcaag accattaagc aatacgccct cgaacgtctg tcccccggtg 300
acgctgatgc cgatcaggca tggcaggaac tgaaaaccat gctggggaac cgcatacaacg 360
atgggcttgc cggcaaggtg tccaccaaga gcgtcggcga aattcttgat gaagaactca 420
gcgggggatcg cgcttgacgg cctacatcct cacggctgag gccgaagccg atctacgcgg 480
catcatccgc tacacgcgcc gggagtgggg cgcggcgag gtgcgccgct atatcgctaa 540
gctggaacag ggcatagcca ggcttgccgc cggcgaaggc ccgtttaagg acatgagcga 600
actctttccc gcgctgcgga tggcccgtg cgaacaccac tacgtttttt gcctgccgcg 660
tgccggcgaa cccgcgttgg tcgtggcgat cctgcatgag cgcattggacc tcatgacgcg 720
acttgccgac aggtctcaagg gctgatttca gccgctaaaa atcgcgccac tcacaacgtc 780
ctgatggcgt actaagctgg g 801

<210> 56

<211> 3208

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 56

cccagcttag tacgcatca ggacgttgtg agtggcgca ttttagcgg ctgaaatcag 60
cccttgagcc tgtcggcaag tcgcgtcatg aggtccatgc gctcatgcag gatcgccacg 120
accaacgcgg gttcgcccgc acgcggcagg caaaaaacgt agtgggttgc gcagcgggccc 180
atccgcagcg cgggaaagag ttcgctcatg tccttaaacg ggccttcgcc ggcggcaagc 240
ctggctatgc cctgttccag cttagcgata tagcggcgca cctgcgccgc gccccactcc 300
cggcgcgtgt agcggtatgat gccgcgtaga tcggcttcgg cctcagccgt gaggatgtag 360
gccgtcaagc gcgatccccg ctgagttcct catcaagaat ttcgccgacg ctcttggtgg 420
acaccttgcc ggcaagccca tcgttgatgc ggttccccag catgggtttc agttcctgcc 480
atgcctgatc ggcattcagc tcaccgggga acagacgttc gagggcgat tgcttaatgg 540
tcttgccctg caaggcggcc agggctttca ggctctggtg ctgctggtcc gtcattgtcg 600
ttgtcaggcg gctcattgga taacctccat aaaatacacg taaccacatt agcacatatg 660
tgggcgtag gctacagcgc gaggcgcatt aaggctcgga aaatgcgcta ggcgcattta 720
aattgcgtat tgctgtaatg cgccatgccg gctagactag gcccaaatgg gtatacccaa 780
tttgaccaag ggggacgcgc aatctagggt ctggaacagc tactttccga acgactcctg 840
cgttgatcgg aaatccagaa gcccgagagg ttgccgcctt tcgggctttt tctttttcaa 900

10

ES 2 843 556 T3

aaaaaaaaat ttataaaacg atctgttgcg gccgcgggt tgtgggcaaa ggcgctggcg	960
ctcgacggtg ggcaaccgct tgcggttgc caccggcgga gccggtgcgc gtagcgatt	1020
gtccacaagc caaggcgac caataattga tatatatatt cataattgaa aagctaattg	1080
aacatactac ttgctgtaac tacttgccg agcaggggt gtttgcaagc tgttgatctg	1140
aaagggtat tagcgctctc acgtgccttt ttgattagcg atttcacgtg accttattag	1200
cgatttcacg tactccgatt agcgatttca cgtaccctga ttagcgattt cacgtggata	1260
gtttttggag cgggccggaa agccccgtga atcaaggctt tgcggggcat tagcggtttc	1320
acgtggataa ctaccctcta tccacaggct tccggggata aaaaagcccc ctcgacggcg	1380
ggctgttga tgggaaggct tgaccaagcc aagcgtagcg ttggcctggt caagtcggag	1440
gggggccgat gcgagcgccc ttgccgggtg cgcgggtgac atgcaggcgt gtggatttga	1500
tgcgcaggca ttcccgctca tcttcgatgc agtcgcttgc ctccggatag acaatcaaca	1560
cttcgcgtag gcgctttttg aagttgtatt tgaagctggc gagtgtgcc cgctctgcc	1620
gctctcgggc cttatcgctc agttcgggcg agttgcgtgc gcggctgcca taggatgagc	1680
cgaattgcgc ttgcagggcg acccaaggga tttgcacgaa ggggcggccc ttggcccgca	1740
acaggaacac gcgataggtc agccacgtgt aaatgtccat cgcaagcgga gactgccgca	1800
aggcatgcag gtagtcgatt cggataggaa ccggtgagcg ggtgacttcc tcgaagaaat	1860
cgcctgtgag ggtgaggggt ctatcccata gcgcccgatc ttctggccgc ttgggattcc	1920
agaatagaaa agcgcgcttg gcaatgacga cgttctcaat gccgaagtca ttgccttgct	1980
cgcgggcaag cgaaatcatg gatgaaaaca ggcgttgcgc ctgattgcga agggtagggc	2040
tgtaacggcc atcgggtgtc attccgagcc tttgtagaaa ttccgattgc gaccggccaa	2100
ggttcaacac ggggtctttc gttcgcacgg cctcgggtga tatccaagca agcaagggtc	2160
gcggcataga accgtagggc agccgatgc tcggcttgcc catgatcgac aaggtagcga	2220
tgccattggt gcgctcaaag tagctggtct tggggtcggg gtggggcatg gtcgcttgca	2280
caaggcaacg ggccatgtag ccgactaagc cagcttcgcg ggcacacctc atttcgagcg	2340
cgaggctcgt cttgatgatc tcgttgatac gatggccggg ggctttgttg ttcttaggca	2400
tgttgttccc tccccggcat ggtgatggtt ggtctagtgt ttgtgggttt gatgttccgg	2460
cgtttgatga acaggcgcaa ggtgtgaggg ctgacgccta acaactcggc tgcgcgactt	2520
tgcggcaagc caaggtcac gtatgcctgt acttcatcaa tacggctgtc cagcttcaag	2580
gcgctcgatt tgctgccctt gggtcgccc agcgtcttgc cgcgctctct ggcgacttgt	2640
agcgcctcgg tggtagctgc ctgaatgaaa tgccgctcga tctgtgcagc caagccaagc	2700
acggttgcca tgatgtcgct ttgtaggctg ccgtccatga tgatcttctg tttggtcaca	2760

ES 2 843 556 T3

	tggacgatta ggccgcgctc gctcgcgct ttgagaattt ccaaggcggc gagggcgga	2820
	ccggcaatgc gcgtaatctc cggcgtcagt agcacgtcgc cacgctcggc cttttcgatg	2880
	attgctccga gcttgcgctt gcgccagtcc tttgctctgc tggcaatttc ttcctcgatc	2940
	tgtagcggcg cgaagccttt ggcgttcgcg tattcgagca aaccgtattt ttggttttcc	3000
	gggtcttggc cgtcacgcga aaccggaga taggcatagt attttgcat ttgcagggaa	3060
	aacgtcagat tcggttaaac atgcctcatt ctacgcgaga ttaaatagga attaaatacc	3120
	ctgtagcggg atagataaaa cgttggtttg ttctgcccta tgagcgtaca aaaaaggccg	3180
	ggtgagtggc ccggccttcg tttaggtg	3208
	<210> 57	
	<211> 3754	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 57	
	atgcgttata ttgagacggc aataacgcgt ctgacgctaa taaaccgtct gaaacctgag	60
	gccaagtga acgtgatcga ccggcacatt aatcgagcat caacttcgc acacatcacc	120
	cagcgcgcgg aagcgttgtc tgcgagactc agggcggttg gtgagcgcgc gtttccccct	180
	accgcacaaa agtcgcttcg atctttcact tccggcgagg tcgccgaaat tgcggcgctg	240
	tcagacggct atcttcgcca gctttcgctc gatggactcg gcccttcgcc agatttgga	300
	agcgcgggtc gccgctctta cactttggag caaatcaatc aactccgtga gtatcttgcc	360
	ggagctcgac ctaaggaagc aagtagattt tggccacgtc ggcgagtggt cgaaaaactt	420
	caagtgatca cggtcgcaaa tttcaaaggc ggctccgcga agacaacaac tgctgtttat	480
	ctcgcacaag gacttgcggt gcagggctac cgtgtcttg cggttgatct tgatcctcag	540
	gcttccctct ccgcaatgtt cggttatcag ccagagttcg acgttgccga aaacacgacc	600
	ctttatgggt cgatccgcta cgacgaccag agggtgacga tgaaggacgt tatccggcgc	660
	acctaacttta ccgggataag cattgtcccg ggtaatctcg aactgatgga gtttgaacat	720
	cagacccctc gattcatgct gcagaatcga gggcgtccg aggatctatt ttcagacgt	780
	gtcgcaggcg caatcaatca ggtcgaacag gacttcgacg tcgttgctgt cgattgcccc	840
	ccacaactgg gctttttgac catgggcgcc cttaacgcag cgagtggcat gattgtgact	900
	gtccatcccc aaatggtcga tgtcgcaccc atgagccagt tcctcttgat gacgtccgat	960
	ctcgtgtcgg taattgaaga agccggcgga cggctcgact acgactttct aaggtttctt	1020
	attacgaggc atgatccgcg cgacgtcccc gagcaggaaa tcgtggcgct gcttcgtgac	1080
	gtcttcggaa ccgacgtcat ggcggcatcc gcatggaagt cgacggcaat cgccaatgcc	1140
	ggtcttacca agcagtcgct ctatgaactc tcgcgcggcg cggtaggacg gacgacgtat	1200

ES 2 843 556 T3

gatcgagcaa	tggaatccat	caatgccgtg	aatctggaaa	ttatcgacca	gataaacaag	1260
gtctggggtc	gatgatggcg	gctcaatcgc	tgataaatca	aaaccttagt	gtgtttgcag	1320
ctgactacac	ggctttacta	gaactgttgt	ttggataggg	tcgaaacgat	gagcaaacgc	1380
acacaatcca	ttcgtcfaat	gtttgccggg	ggcaacgata	agaccgcttc	taatgacgca	1440
cagcagcctg	tgcagcgagt	tgcttcgggt	gcggtccggg	cgttgaaaga	cacgtttctca	1500
gaggtggaaa	gacattacga	agagctcaag	cagcaagttg	ctgacggagc	tgttccagtc	1560
gaaatagacc	cagaatttgt	cgacccctcg	cccttcgcgg	accggtttgc	ggaccaagac	1620
atggctgtct	tggaggtcct	taagacttcg	atcaaagata	atggacaaga	aattccgata	1680
cttgttcgac	cgcatccagc	tgagatcgga	cgatatcagg	ttgcctacgg	ccaccgcaga	1740
ctacgtgcga	ctgccgaact	cggcttgaaa	gtcaaagcct	atgtccgcga	actggatgac	1800
gatcgtcttg	tcgtcgcgca	gggcatcgag	aattccgctc	gtgaagacct	gactttcatc	1860
gaacggggcg	cgtttgcaact	caaaactggac	gaggttggtt	ttcaacggtc	gctaatacaa	1920
atggctttgt	ctgttgaccg	acacgaagct	caaaagctgg	taacagttgc	tcgggcagtt	1980
ccgcagtggt	tgattgattc	cattggccgt	gcgccgaaga	tcggtcgacc	gaggtggctg	2040
gagctcgag	aacttttgaa	agtcccgggc	gcagagaaga	aggcccgcaa	agccaccag	2100
gacaaatcct	tctcacataa	ggattctgac	aatcgtttcc	tggtcgtcct	aagagcgtcc	2160
aagatagatg	atgtgacggc	gtcgtctgta	gccccacccc	gcttggtcgc	aaaatctgct	2220
gacgggcttc	aaatcgcgac	gctcgccatc	ggcggaagc	agtgcaaaat	cgaaatgaac	2280
cgagatagag	acgagacttt	tgcaaagttt	gttatggagc	agcttcctgc	gctctacgag	2340
aacttcagaa	aggaaaatcc	gtcatctgaa	ggataaataa	ggataggaac	agcaaaagaa	2400
aggcccccga	acgtcgccgt	cgtggaagcc	tctctcaatg	gtctaagcag	cctgagaatc	2460
gcatttccat	gaatcgagc	caagagtctt	tggcacggtt	tttggtgagc	agatttcttt	2520
tgctatttga	agggtgaaag	acatgcagag	tggaaatgtg	acgacgccgt	ttgggcggcg	2580
gccgatgacg	cttgcccttag	tggcggcaca	gttcaaatcc	gccgacatca	ggcagggtaa	2640
atcgggccgac	aagtggaaga	tctatcgaga	tgcttgtagc	gcgcgatcat	tgttgggttt	2700
gcgcgatcgg	gcacttgacg	tgctcaacgc	gcttctatca	ttctaccctg	aggctgaact	2760
caaggaggaa	gcgaatctta	tcgtgtttcc	gtccaatgcc	caactcgagg	cacggggcaaa	2820
cggtatcgcc	ggcacgacgc	tacgcgagaa	cttgcccttc	ttggtcgatg	caggtctcgt	2880
ccaacgggaat	gatagcccga	acgggaaacg	ctatgcgcgc	aagagctcag	atgggtccat	2940
cgagacagcc	tttggttcca	gtctagcgcc	gcttctggcc	cgatcagaag	aacttgcggt	3000
tatggcccag	caggtcgagc	agggtaatcg	caagctcaaa	atcgtcaaaag	agcgcatcac	3060
aattgcccga	cgggatgtcc	ggaaattgat	ctctcgggca	atggaagata	acgttcccgg	3120

ES 2 843 556 T3

	ggattggggc aaggtcgagg agctttatgt tgaagccatg gcacggctgc ggcgggcaaa	3180
	aggggcgagc gaattgatcg gaattctcga cgaagtcgag atgctgaagg aagaagtagt	3240
	caatctattg gaacggcacg tcaatcttga aaaatccgac gccaatgatg acggattccg	3300
	tcagcacata cagaattcaa ataccgaatc cagcaatgaa cttgaacctt gctctgaaaa	3360
	cgagcagggc gcgaagcaga gtaaaaacgt tgaactaaga gccgaatcga taaaatcatt	3420
	cccgtttggt atgggtccttc gggcatgccc ggagatttcc atgtatggtc cggcgaggagc	3480
	aatcggcagc tggcgggacc ttatgatggc cgccgtggtc gtgcgatcaa tgctgggggt	3540
	aagcccgtcg gcctaccagg acgcctgcga agtaataggg cctgaaaacg cggccgcaat	3600
	gatggctgca atcctggagc gggcggggca catcacatcg gccggtggat atcttcgcga	3660
	tttgacgtcg agggcaaaac gcggcgagtt ttctctcgga ccaatgttga tggcgttgct	3720
	gcggggccaat ggcgggcgaac aaaggcgggc ttaa	3754
	<210> 58	
	<211> 3992	
	<212> ADN	
5	<213> Agrobacterium tumefaciens	
	<400> 58	
	gtgagcaaag ccgctgccat atcccgaat gatcgcccggt cggtagatgt taccattggt	60
	gagcatgctg agcagctcag ctctcagctt caagcgatga gcgaggcttt gtttcctccg	120
	acgtcgcaca agagcttgcg caaattcacc tcgggtgaag ccgcacgctt gatgaaaata	180
	tctgactcaa ctcttcgaaa gatgacactg gctggcgaag ggccgcaacc tgaactcgcc	240
	agcaacggac ggcgctttta caccctcggt cagataaacg aaatccggca gatgcttgcc	300
	ggctcgactc gaggacgtga aagcattgat tttgtgcctc atcgccgagg ttctgagcat	360
	ttgcaagtcg ttgctgtaac caacttcaaa ggtggctctg ggaagacgac gacgtccgct	420
	catcttgcac agtatctggc gttgcaaggt tacaggggtc tcgcagtcga tctcgatccg	480
	caggctagtc tttcagcact cctcggcgtt ctgccagaaa ctgatgtcgg tgcaaacgaa	540
	acgctctatg cggctattcg gtacgacgac acacgtcgtc cgttgcgaga tgtgatccga	600
	ccgacgtatt ttgatgtctt tcaccttgtt cctggaaatc tcgagcttat ggagttcgag	660
	cataccaccc cgaaagcatt gactgacaaa ggtacgcgcg acggattggt cttcactcgc	720
	gtggcccaag cctttgatga ggtcgcgcgac gattacgatg tcgtggtcat cgactgccct	780
	cctcagcttg ggtttttgac tctcagcggg ttgtgtgctg caacatcaat ggtaatcacc	840
	gtacatcctc agatgctgga tatcgcttcc atgagccagt ttctcctcat gacacgcgac	900
	cttctgggtg tcgtgaaaga ggcggggggc aatctccagt acgatttcat acgctatctc	960
	ttgacgcgct atgagcccca ggacgcgcg cagacgaaag tgacggcact gctgcgcaac	1020

ES 2 843 556 T3

atgttcgagg atcacgtcct tacaaatcct atgggtcaagt cggcagcggg atctgatgcc	1080
ggtttaacca agcagacgct ctatgagata gggcgagaga accttacgcg atcgacatac	1140
gaccggggcg tggaatcttt agatgcggtg aattcggaga tcgaggcttt gatcaagatg	1200
gcgtgggggc ggggtctaag aaaggctttg cgttcctcac agatctgttg ggagctccca	1260
acagacaggt gttgattcgc cccctggaca tggggcactg gagaagccgg ggtaatttga	1320
gacgacgacg cagcccatc gctaattggc cagggcgag ttgtcttgtc ttgttgggag	1380
ctcccaacca agcgcatttg caatcaaaaa tgcgacgcca cgacgccaaa cccaagaggc	1440
tgatatcatg agccgcaaag acgcaatcga tactttgttc ctcaagaagc aacctgcgac	1500
cgatagagca gcagtcgaca agtcgaccgc tcgtgttcgt accggagcga tttcggccat	1560
gggttcgtct ttgcaagaga tggctgaggg cgcgaaggct gcagctcggc tgcaggatca	1620
actggctaca ggcgaagccg tcgtgtccct ggatccatcc atgatcgacg ggtcgccgat	1680
cgcggatcgg ctgccctcag acgtggatcc gaaattcgag cagcttgagg cgagcatttc	1740
gcaggagggg cagcaggtgc cggttcttgc cagaccgcac cctgaggctg ccggtcgata	1800
tcagatcgta tatggaaggc ggcggctgcg cgcggcagta aatctgcgga gagaggtttc	1860
tgccattgtt cgaaatctca cggactgtga actggctgtg gcccagggcc gcgaaaatct	1920
taaccgtgct gacctctcgt tcattgagaa ggctctcttc gccctgcgcc tcgaagatgc	1980
gggttttgat agagccacca tcattgccgc gctatccact gacaaggccg acctcagccg	2040
ctacataact gtagcaaggc gcataaccgt gaacctcgcc acacaaatcg gcccagcgtc	2100
gaaagcgggt cgatcgcggt gggctcgact tgcggagggg cttgggaagc ctaaggcaac	2160
ggacgcaatc gaagcgatgc ttgggtcaga gcagttcaag caatctgata gcgatacccg	2220
ctttaacctc attttcaacg ccgtttcaag gccacctgcg aagactccaa aaaaggtaag	2280
ggcctggagc acgccaaagg ggaaaaaggc agcgacgacg cgacaagaaa ctggacgaac	2340
ggcgttggtt ttcgacgaga gactgggtgc aacttttggc gaatatgtcg ctgaccagtt	2400
ggacagtctg tacgccaggt tcattgaaac caacggagga ggcaagctcg accaatagtc	2460
agggtttcat ccaatttaaa gctccgctcg actgagatgg actggctctc accgcaaaag	2520
aaaaggccc ccgaaacggc gttccggaag accttctctg tagtctcgca gctaagagaa	2580
tcgcatttcc aggaatcgta gtcaagggtc ccgtaaggga aagcgtcatt tcgacgggcg	2640
gatttcaatt gcctaacaaa aggtaaaagg aatgcagac gcataatctca acgacgtcct	2700
ttgggcccgc gccgatgaca ctccggcata ttgcaagcca gatggcagca aaagcggctg	2760
catcagacac tgtcgccac aaatggcagg tcttccagca catccgtgaa tcccggggac	2820
tgatcggagc caccgaccgc tcaactctga tcctgaacgc gctgttgacg ttttaccg	2880

ES 2 843 556 T3

agaccgcctt gactggtggt gccgaactgg tcgtatggcc ttctaacgaa cagctgatgg 2940
ctcgcgccaa cggcatgccc gccacgacac tgcgccggca tcttgccata ctggttgatt 3000
gcgggctcat cattcgccgc gacagcccca atggcaagcg gttcgcccg c aagggaaggg 3060
gaggggagat tgagcaggcc tatgggttcg atctgtcgcc gatcgtcgag cgggccgagg 3120
agttccgaga tctggcccag acagtgcagg ctgaaaaaaa ggccttccgg gtggccaagg 3180
agcgttgac tcttcttcgt cgtgacattg tcaaaatgat cgaaactggc gtcgaagaga 3240
gcgttcttg aaactgggga agagttacc agacctatca ggggatcatc ggccgcctgc 3300
cacgctcgcc acctcgccag cttgtcgaga gtattgggca agagcttcag gaactttgca 3360
tcgagatccg tgacgtattg gaatctttca caaaaacgat gaatctggac gccaatgagt 3420
cccatatcgg tcgccacaaa cagaattcaa atccagactc taaatttgaa tctgaatacg 3480
gctctggaaa aaaagatgaa gcgggcggca gcgttcgga aaccgacaat gtacggagct 3540
tgccgaaacg cgagctgcct ttgggaatcg tgctggatgc ctgccccgaa atgcgggaat 3600
tgccccaggg aggtccaatt cgacattggc gcgacttgct ggccggcggt gagcttgccc 3660
ggccgatgct ggggattagt ccgagcgctt ggccgggaggc ccgcgaaacc atgggcgagc 3720
aacacgcggc gatcacgctg gcttcgatct atcagcgggc cggtcagatc aataacgctg 3780
ggggttatct gcgcagctg accgaccggg ccaaggatgg gaagttttcg acctggccga 3840
tggtcatggc gttgctccgg gcgaagctgg acgagcagaa gaatgcagtc ggcgctggaa 3900
agccgcgaac tgctgaggag gtcgaggatg acagccgcct ccacgtatcg gaatcgctgc 3960
tcaaaaacct gcgaaagccg agatcttggg ga 3992

<210> 59

<211> 865

<212> ADN

5 <213> secuencia artificial

<220>

<223> plásmido PR1b, Ruegeria sp.

<400> 59

gcgccggctg gtaatcgatg atgatgcaat ccagcgtttc gttcagcgcc tgctcgagac 60
ggacaggatc gacctttcca tctgcaacaa gtctgttggg agccgttgcc ggccgatacg 120
cctctttcca gcgatcaatc gcatcccgca gatagcgata gaaactcttc cccgcaggcg 180
cttcgcgaac caagcgggag atctgaattt cgccctcaga agtctctccg tgcgccggca 240
ccaatcgaag ccccgcccaa gaagtcttca aaaagaagct atccagcgtc tccgcagagt 300
gatccgtgta cggcgccctc tcagactgga agaggcctgc gaagtccacc atcgatggcg 360
tctcttcagt aggcattctga gacaaacct caccgcctac aaagtaaagt gtgatcgtgc 420
tttgccgatc cgcattccatc acaccaacgc gcattccgta gtgcaagctc aggtactgag 480

10

ES 2 843 556 T3

	cgaaatgtgc agcgctaagg ctcttggtg tgccgccctt ttggctggca aatgaaatga	540
	ccggaagcgg atcgccgggc ttgcccacg ctagatagtc aaaagggcgc tttgcagacg	600
	ccgccatcaa agcgcgcata tgcatacgtt cgtgaagatt gaagacacgc tcgcgtccga	660
	cgtactcacc ctttgaaaag tcgtcgttgc cattcgcaaa cttagtcaag tattggatgc	720
	tgacgttcaa aaagcgagcg cactgtcgca tcgaccatgt acgcttgata cgatggcgca	780
	acgtcaactc tttgttgacg gcgaccattg tgcgctccag acctcgtgac aggtcgggtca	840
	aaaactgttc tagacttccg ctcat	865
	<210> 60	
	<211> 3120	
	<212> ADN	
5	<213> Sinorhizobium fredii	
	<400> 60	
	atgaacgcta attcgccgcc catcgctccc gcacagccga tgcatttcga ggatctcatc	60
	ctcgaacagg gcgatctgat ttcaaagaaa ttgcatcttt taagcatgca gcagttccca	120
	ccgaatgcc aaaaactgct tcgccagttc tcgctctctg aagtggcgca gttcctcggc	180
	gtctctcaaa gtacgctgaa gaaacttcat cttgagggaa aaggcccccct ccccagaca	240
	tcgtcgtcgg gtgccggtc ttatagtgtg gacgagatgg cggaactgcg tcagtatctc	300
	gatcagcatg gacgatcgga agccaggaac tacgtgccgc accgtcggtc gggcgaaaag	360
	ctccagggtca ttgccgttgt caatttcaaa ggcgggttcg gaaagacgac gacggccgcc	420
	catctcgcgc agtacatggc gttgaccggg caccgtgtct tggcggtgga cctcgaccgc	480
	caggcgctac tttcgtcgct tcacggtttt cagcccgaac tcgacatgtc gccgtcgctc	540
	tatgagggcg tcagatatga cgatcaacgc cgatcgatca gtgagatcat ccagcctacc	600
	aatttcccg gcctcgatat cgtgcctgcg aacctcgaa ttcaggagta cgagtaagac	660
	accccgcttg cgatgtcgaa caagagctcg aacgatggca agaccttctt cacgcggatt	720
	tcgcgcgcgc tgtcggaagt caacgaccgc tatgacgttg tcgtcatcga ctgccctccc	780
	cagctcggct atctcacgat caccgcgtg actgcggcta cgagcgtcct gatcacgac	840
	catccacaga tgctcgacgt catgtcgatg ggccagttcc ttttgatgct gggcgggac	900
	ctgaagccga tccgggatgt ggggtgctgc gtcaatctcg aatggatcgt ctatctgac	960
	acccgggtac agccaacgga tgggccgcag gcgcagatgg tgggcttcac gcagacctg	1020
	ttccaccagt ttgtgctgaa gaaccagatg ctgaaatcga cagcggtttc agacgccga	1080
	atcaccaagc agaccttta cgaagtcgac aagagtcaga tgacgcggtc cacctatgag	1140
	cgcgcgatgg actccttgaa cgcagtcaac gccgagatcg tcgaactcgt ccatgcctct	1200
	tgggggagga aggttgtcag ttgagaagtg taaacttccc cgggacggcc cggtagatct	1260

ES 2 843 556 T3

caggcaaacg gagcaggcac atggcgagaa agaacctact ggcgggactg gtagacacag 1320
 cggaaatccc gcacgcagac gttgcgccc cctaccgat gcgcggggct tcgaagagca 1380
 tgggtcgctc gctcgacgag ttgtcgcgcc aggccgagaa atttctcgaa ggagaaacgg 1440
 tcgtogaact cgacctgag aactcgatg gttcgttcgt ctctgatcgc atgggcgata 1500
 gttccgagca gttcgaggaa ttgaagcagg cgattgccga gcgcgggcag gatacgccca 1560
 tccttggtgc cccccaccg tcggcagccg atcggtacca gatcgctctt ggacaccgcc 1620
 gagcccgctg tgcccgcgag ctggccgca aggttaaggc ggtcgtcaag gcgctggatg 1680
 accgcacgca cgttatcgcc cagggacagg aaaactcggc gcgcgccaac ctctccttca 1740
 ttgaacgggc gaacttcgcc tcgcacctcg agaagcttgg ctatgaccgg acaatcatag 1800
 gatcgggcgt tgctgccaat gcggccgcta tctcgaaaat gatcgccgta atagatcgta 1860
 ttctgaaga gacaatcgca agaactcggc cctgcccggc ggtcggccga gaacgctggg 1920
 tcgagttgtc attgcttggt ggtaaaaccg ccaacgaagc gaaggtgaag gcaatcgtct 1980
 ctgatccttc cttcaatgaa ctgagcactg acgatcgctt caattccctg ttttctggtc 2040
 tgaacagcgc ggcaagcct gtccggaaga cgactccaaa gatcctggag aactggcaac 2100
 cggcggacaa gaccgtctcc gctaagtatt cgaactccgc caaggcgttt gccttgctta 2160
 tgaaatcaag gaatgccggc ccctttgggc ggtatattgc agacaacctg gaccggctct 2220
 atgccgagtt cctggagcag ggtaatcgga aggaagactg acgcaaaaga aaaaggcccc 2280
 cagacggtaa ccatcgtgga agcctctctc atcgtttagc agcctgagaa tcgcatttcc 2340
 acgaatcgca gtcaagagtc tttggcaccg gaacgggtga gcggacgtct tttgcctgag 2400
 gataggtgaa agaagatgca agatggaagt gtgacgacgc ccttcgggcg gcggtcgatg 2460
 acgcttgcca tggttgcaag ccaatatatg tcgcgtgaac tcgaacctga aacatcggct 2520
 gacaagtgga agctgtttcg agcgtctctg gaagctaagc cgaagctcgg catcagcgaa 2580
 cgtgcgctct ctgttatgaa cgcacttttg agcttttatc ccgagacgac cttgtctgag 2640
 gaaaacggcc tcatcgtgtt cccctctaac atgcagttgt cgcttcgtgc gcatggcatg 2700
 gctgaagcga cgttcgctcg ccacatcgca gcgctcgtcg acgctggcct tctggcacgc 2760
 cgggatagcc caaatggcaa acgctacgcc cgcaaggacg gagacggttc gatcgatgag 2820
 gcctatgggt tttcgttggc tccgttgctg tcacgcgcac gcgagatcga gcaaatagct 2880
 gcctatgtca aaattgagcg actgcaattg cgtagacttc gcgagcgctt gacgatctgc 2940
 cggcgcgata tcggtaagct gattgaagtc gccattgagg agggagtgga cggtaattgg 3000
 gacggaattc accagcacta ccgcagccta gtagcgacca ttctcagagt ggcaacggcc 3060
 gcgaccgtcg ccccatcct cgaagagatg gagatgttgc gcgaggaaat ctccaacctt 3120
 <210> 61
 <211> 32
 <212> ADN
 5 <213> Secuencia artificial
 <220>
 <223> Construcción sintética
 <400> 61
 cgttcctcga ggcctcgagg cctcgaggaa cg 32
 10 <210> 62
 <211> 146

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Construcción sintética	
5	<400> 62	
	gagtcacgc tagatgagag ctttgttgta ggtggaccag ttggtgattt tgaacttttg	60
	ctttgccacg gaacgggtctg cgttgtcggg aagatgcgtg atctgatcct tcaactcagc	120
	aaaagttcga tttattcaac aaagcc	146
	<210> 63	
	<211> 6968	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> Construcción sintética	
	<400> 63	
	gtgtacaacc agatattttt caccaacatc cttcgtctgc tcgatgagcg gggcatgacg	60
	aaacatgagc tgtcggagag ggcaggggtt tcaatttcgt ttttatcaga cttaaccaac	120
	ggtaaggcca acccctcgtt gaaggtgatg gagggcattg ccgacgccct ggaaactccc	180
	ctacctcttc tcctggagtc caccgacctt gaccgcgagg cactcgcgga gattgcgggt	240
	catcctttca agagcagcgt gccgcccgga tacgaacgca tcagtgtggt tttgccgtca	300
	cataaggcgt ttatcgtaaa gaaatggggc gacgacaccc gaaaaaagct gcgtggaagg	360
	ctctgacgcc aagggttagg gcttgcaactt ccttcttttag ccgctaaaac ggccccttct	420
	ctgcggggcg tcggctcgcg catcatatcg acatcctcaa cggaagccgt gccgcgaatg	480
	gcatcggggc ggtgcgcttt gacagttgtt ttctatcaga acccctacgt cgtgcgggtc	540
	gattagctgt ttgtcttgca ggctaaacac tttcgggtata tcgtttgcct gtgcgataat	600
	gttgctaagt atttgttgcg taggggttac tgaaaagtga gcgggaaaga agagtttcag	660
	accatcaagg agcgggcca ggcgaagctg gaacgcgaca tgggtgcgga cctgttggcc	720
15	gcgctcaacg acccgaaaac cgttgaagtc atgctcaacg cggacggcaa ggtgtggcac	780

ES 2 843 556 T3

gaacgccttg gcgagccgat gcggtacatc tgcgacatgc ggcccagcca gtcgcaggcg	840
attatagaaa cgggtggccgg attccacggc aaagagggtca cgcggcattc gcccatcctg	900
gaaggcgagt tccccttggg tggcagccgc tttgcccggc aattgccgcc ggtcgtggcc	960
gcgccaacct ttgcgatccg caagcgcgcg gtcgccatct tcacgctgga acgtacgtc	1020
gaggcgggca tcatgacccg cgagcaatac gaggtcatta aaagcgccgt cgcggcgcat	1080
cgaacatcc tcgtcattgg cgggtactggc tcgggcaaga ccacgctcgt caacgcgatc	1140
atcaatgaaa tggtcgcctt caaccgtct gagcgctcg tcatcatcga ggacaccggc	1200
gaaatccagt gcgcccgcga gaacggcgtc caataccaca ccagcatcga cgtctcgatg	1260
acgctgctgc tcaagacaac gctgcgtatg cgcgccgacc gcacccctggg cggtgaggta	1320
cgtggccccc aagcccttga tctgttgatg gcctggaaca ccgggcatga aggaggtgcc	1380
gccaccctgc acgcaaaaa ccccaaagcg ggcctgagcc ggctcgccat gcttatcagc	1440
atgcaccggg attcacgaa acccattgag ccgctgattg gcgaggcggt tcatgtggtc	1500
gtccatatcg ccaggacccc tagcgggcgt cgagtgcgaag aaattctcga agttcttggg	1560
tacgagaacg gccagtagat caccaaaacc ctgtaaggag tatttccaat gacaacggct	1620
gttcggttcc gtctgacat gaatcgcggc attttgttct accttgccgt gttcttcgtt	1680
ctcgtctcgc cgttatccgc gcacccggcg atggcctcgg aaggcaccgg cggcagcttg	1740
ccatatgaga gctggctgac gaacctgcgc aactccgtaa ccggcccggg ggccttcgcg	1800
ctgtccatca tcggcatcgt cgtcgccggc ggcgtgctga tcttcggcgg cgaactcaac	1860
gccttcttcc gaacctgat ctctctggtt ctggtgatgg cgtgctggt cggcgcgcag	1920
aacgtgatga gcaccttctt cggctggtgt gccgaaatcg cggccctcgg caacggggcg	1980
ctgcaccagg tgcaagtcgc ggcggcggtt gccgtgcgtg cggtagcggc tggacggctc	2040
gcctaatacat ggctctgcgc acgatcccca tccgtcgcgc aggcaaccga gaaaacctgt	2100
tcatgggtgg tgatcgtgaa ctggtgatgt tctcgggcct gatggcggtt gcgctgattt	2160
tcagcgccca agagctgcgg gccaccgtgg tcggtctgat cctgtggttc gggcgctct	2220
atgcgttccg aatcatggcg aaggccgatc cgaagatgcg gttcgtgtac ctgcgtcacc	2280
gccggtacaa gccgtattac ccggcccgtt cgaccccggt ccgcgagaac accaatagcc	2340
aagggaagca ataccgatga tccaagcaat tgcgattgca atcgcgggcc tcggcgcgct	2400
tctgttggtc atcctctttg cccgcatccg cgcggtcgat gccgaactga aactgaaaaa	2460
gcacgttcc aaggacgcgg gcctggccga tctgctcaac tacgccgctg tcgtcgatga	2520
cggcgtaatc gtgggcaaga acggcagctt tatggctgcc tggctgtaca agggcgatga	2580
caacgcaagc agcaccgacc agcagcgcgga agtagtgtcc gcccgcatca accaggccct	2640

ES 2 843 556 T3

cgcgggcctg	ggaagtgggt	ggatgatcca	tgtggacgcc	gtgcggcgtc	ctgctccgaa	2700
ctacgcggag	cggggcctgt	cggcgttccc	tgaccgtctg	acggcagcga	ttgaagaaga	2760
gcgctcggtc	ttgccttgct	cgtcgggtgat	gtacttcacc	agctccgcga	agtcgctctt	2820
cttgatggag	cgcattggga	cgtgcttggc	aatcacgcgc	acccccggc	cgttttagcg	2880
gctaataaag	tcattggctct	gccctcgggc	ggaccacgcc	catcatgacc	ttgccaaagct	2940
cgtcctgctt	ctcttcgatc	ttcgcagca	gggcgaggat	cgtggcatca	ccgaaccgcg	3000
ccgtgcgcgg	gtcgtcgggt	agccagagtt	tcagcaggcc	gcccaggcgg	cccaggctgc	3060
cattgatgcg	ggccagctcg	cggacgtgct	catagtccac	gacgcccgtg	attttgtagc	3120
cctggccgac	ggccagcagg	taggccgaca	ggctcatgcc	ggccgcgcgc	gccttttctt	3180
caatcgctct	tcgttcgtct	ggaaggcagt	acaccttgat	aggtgggctg	cccttcctgg	3240
ttggcttggt	ttcatcagcc	atccgcttgc	cctcatctgt	tacgccggcg	gtagccggcc	3300
agcctcgcat	agcaggattc	ccgttgagca	ccgccagggt	cgaataaggg	acagtgaaga	3360
aggaacaccc	gctcgggggt	gggcctactt	cacctatcct	gcccggctga	cgccttgga	3420
tacaccaagg	aaagtctaca	cgaacccttt	ggcaaatcc	tgtatatcgt	gcgaaaaagg	3480
atggatatac	cgaataatc	gctataatga	ccccgaagca	gggttatgca	gcggaaaagc	3540
gctgcttccc	tgctgttttg	tggaatatct	accgactgga	aacaggcaaa	tgacgaaat	3600
tactgaactg	aggggacagg	cgagagacga	tgccaaagag	ctacaccgac	gagctggccg	3660
agtgggttga	atccgcgcg	gccaagaagc	gccggcgtga	tgaggctgcg	gttgcgttcc	3720
tgccgggtgag	ggcggtatgc	gagcgggcgt	tagcgtccgg	ctatgcgctc	gtcaccattt	3780
gggagcacat	gcgggaaacg	gggaagggtca	agttctccta	cgagacgttc	cgtcgcacg	3840
ccaggcggca	catcaaggcc	aagcccgccg	atgtgccgcg	accgcaggcc	aaggctgcgg	3900
aaccgcgcgc	ggcacccaag	acgccggagc	cacggcggcc	gaagcagggg	ggcaaggctg	3960
aaaagccggc	ccccgctgcg	gccccgaccg	gcttcacctt	caaccaaca	ccggacaaaa	4020
aggatctact	gtaatggcga	aaattcacat	ggttttgtag	ggcaaggggc	gggtcggcaa	4080
gtcggccatc	gccgcgatca	ttgcgcagta	caagatggac	aaggggcaga	cacccttggt	4140
catcgacacc	gaccgggtga	acgcgacgtt	cgagggttac	aaggccctga	acgtccgcgc	4200
gctgaacatc	atggccggcg	acgaaattaa	ctcgcgcaac	ttcgacaccc	tggtcgagct	4260
gattgcgcgc	accaaggatg	acgtggtgat	cgacaacggg	gccagctcgt	tcgtgcctct	4320
gtcgcattac	ctcatcagca	accagggtgc	ggctctgctg	caagaaatgg	ggcatgagct	4380
ggtcatccat	accgtcgtca	ccggcggcca	ggctctcctg	gacacgggtga	gcggcttcgc	4440
ccagctcgcc	agccagttcc	cggccgaagc	gcttttcgtg	gtctggctga	acccgtattg	4500
ggggcctatc	gagcatgagg	gcaagagctt	tgagcagatg	aaggcgtaca	cggccaacaa	4560

ES 2 843 556 T3

ggcccgcggtg	tcgcccatca	tccagattcc	ggccctcaag	gaagaaacct	acggccgcga	4620
tttcagcgac	atgctgcaag	agcggctgac	gttcgaccag	gcgctggccg	atgaatcgct	4680
cacgatcatg	acgcggcaac	gcctcaagat	cgtgcggcgc	ggcctgtttg	aacagctcga	4740
cgcgggcgcc	gtgctatgag	cgaccagatt	gaagagctga	tccgggagat	tgcggccaa	4800
cacggcatcg	ccgtcggccg	cgacgaccgc	gtgctgatcc	tgcataccat	caacgcccg	4860
ctcatggccg	acagtgcggc	caagcaagag	gaaatccttg	ccgcgttcaa	ggaagagctg	4920
gaagggatcg	cccatcggtg	ggcgaggac	gccaaggcca	aagcggagcg	gatgctgaac	4980
gcggccctgg	ggccagcaa	ggacgcaatg	gcgaaggtaa	tgaaggacag	cgccgcgcag	5040
gcggccgaag	cgatccgcag	ggaaatcgac	gacggccttg	gccgccagct	cgcgcccaag	5100
gtcggggacg	cgcgggcggt	ggcgatgatg	aacatgatcg	ccggcgccat	ggtgttggtc	5160
gcggccgccc	tggtggtgtg	ggcctcggtt	tgaatcgag	aggcgagat	gaaaaagccc	5220
ggcgttgccg	ggctttgttt	ttgcgttagc	tgggcttggt	tgacaggccc	aagctctgac	5280
tgcgcccgcg	ctcgcgctcc	tgggcctgtt	tcttctcctg	ctcctgcttg	cgcatcaggg	5340
cctggtgccc	tcgggctgct	tcacgcatcg	aatcccagtc	gccggccagc	tcgggatgct	5400
ccgcgcgcg	cttgcgcgct	gccagttcct	cgatcttggt	cgcggtgaatg	cccatgcctt	5460
ccttgatttc	gcgcacccatg	tccagccgcg	tgtgcaggg	ctgcaagcgg	gcttgctggt	5520
gggcctgctg	ctgctgccag	gggcctttg	tacgcggcag	ggacagcaag	ccgggggcat	5580
tggactgtag	ctgctgcaaa	cgcgccctgct	gacggtctac	gagctgttct	aggcggtcct	5640
cgatgcgctc	cacctgggtca	tgctttgcct	gcacgtagag	cgcaagggtc	tgctggtagg	5700
tctgctcgat	gggcgcggat	tctaagagg	cctgctgttc	cgtctcgccc	tctggggccg	5760
cctgtagcaa	atcctcgccg	ctgttgccgc	tggactgctt	tactgccggg	gactgctggt	5820
gccctgctcg	cgccgtcgct	gcagttcggc	ttgccccac	tcgattgact	gcttcatttc	5880
gagccgcagc	gatgcgatct	cggtatgctg	caacggacgg	ggcagcgccg	aggtgtccgg	5940
cttctccttg	ggtgagtcgg	tcgatgccat	agccaaagg	ttccttcaa	aatgcgtcca	6000
ttgctggacc	gtgtttctca	ttgatgcccg	caagcatctt	cggttgacc	gccagggtcaa	6060
gcgcgccttc	atgggcgggc	atgacggacg	ccgccatgac	cttgccgccc	ttgttctcga	6120
tgtagccgcg	taatgaggca	atggtgccgc	ccatcgctag	cgtgtcatcg	acaacgatgt	6180
acttctggcc	ggggatcacc	tccccctcga	aagtcgggtt	gaacgccagg	cgatgatctg	6240
aaccggctcc	ggttcggggc	accttctccc	gctgcacaat	gtccgtttcg	acctcaaggc	6300
caaggcggtc	ggccagaacg	accgccatca	tggccggaat	cttggtgttc	ccgcccgcct	6360
cgacggcgag	gactggaacg	atgcggggct	tgtcgtcgcc	gatcagcgtc	ttgagctggg	6420

ES 2 843 556 T3

	caacagtgtc gtccgaaatc aggcgctcga ccaaattaag cgccgcttcc gcgtcgccct	6480
	gcttcgcagc ctggtattca ggctcgttgg tcaaagaacc aaggtcgccg ttgcgaacca	6540
	ccttcgggaa gtctccccac ggtgcgcgct cggtcttgct gtagctgctc aagacgcctc	6600
	ccttttttagc cgctaaaact ctaacgagtg cgcccgcgac tcaacttgac gctttcggca	6660
	cttacctgtg ccttgccact tgcgtcatag gtgatgcttt tcgcactccc gatttcaggt	6720
	actttatcga aatctgaccg ggcgtgcatt acaaagttct tccccacctg ttggtaaatg	6780
	ctgccgctat ctgcgtggac gatgctgccg tcgtggcgct gcgacttata gcccttttgg	6840
	gccatataga tggtgtaaat gccaggtttc agggccccgg ctttatctac cttctggttc	6900
	gtccatgcgc cttggttctc ggtctggaca attctttgcc cattcatgac caggaggcgg	6960
	tgtttcat	6968
	<210> 64	
	<211> 1660	
	<212> ADN	
5	<213> Escherichia coli	
	<400> 64	
	gtggcgcaac gatgccggcg acaagcagga gcgcaccgac ttcttccgca tcaagtgttt	60
	tggctctcag gccgaggccc acggcaagta tttgggcaag gggtcgctgg tattcgtgca	120
	gggcaagatt cggaatacca agtacgagaa ggacggccag acggtctacg ggaccgactt	180
	cattgccgat aaggtggatt atctggacac caaggcacca ggcgggtcaa atcaggaata	240
	agggcacatt gccccggcgt gagtccgggc aatcccgaag ggagggtgaa tgaatcgac	300
	gtttgaccgg aaggcataca ggcaagaact gatcgacgag gggttttccg ccgaggatgc	360
	cgaaaccatc gcaagccgca ccgtcatgag tcgcgcccgcc gaaaccttcc agtccgtcgg	420
	ctcgatggtc cagcaagcta cggccaagat cgagcgcgac agcgtgcaac tggctcccc	480
	tgccctgccc gcgccatcgg ccgccgtgga gcgttcgcgt cgtctcgaac aggaggcggc	540
	aggtttggcg aagtcgatga ccatcgacac gcgaggaact atgacgacca agaagcgaaa	600
	aaccgccggc gaggacctgg caaaacaggt cagcgaggcc aagcaggccg cgttgctgaa	660
	acacacgaag cagcagatca aggaaatgca gctttccttg ttcgatattg cgcctgggcc	720
	ggacacgatg cgagcgatgc caaacgacac ggcccgctct gccctgttca ccacgcgcaa	780
	caagaaaatc ccgcgcgagg cgctgcaaaa caaggtcatt ttccacgtca acaaggacgt	840
	gaagatcacc tacaccggcg tcgagctgag ggccgacgat gacgaactgg tgtggcagca	900
	ggtgttgag tacgcgaagc gcacccctat cggcgagccg atcaccttca cgttctacga	960
	gctttgccag gacctgggct ggtcgatcaa tggccggtat tacacgaagg ccgaggaatg	1020
	cctgtcgcgc ctacaggcga cggcgatggg cttcacgtcc gaccgcgttg ggcacctgga	1080

ES 2 843 556 T3

atcgggtgtcg ctgctgcacc gcttccgctg cctggaccgt ggcaagaaaa cgtcccgttg 1140
ccaggctcctg atcgacgagg aaatcgctgt gctgtttgct ggcgaccact acacgaaatt 1200
catatgggag aagtaccgca agctgtcgcc gacggcccga cggatgttcg actatttcag 1260
ctcgcaccgg gagccgtacc cgctcaagct ggaaaccttc cgcctcatgt gcggtacgga 1320
ttccacccgc gtgaagaagt ggcgcgagca ggtcggcgaa gcctgcgaag agttgcgagg 1380
cagcggcctg gtggaacacg cctgggtcaa tgatgacctg gtgcattgca aacgctaggg 1440
ccttgtgggg tcagttccgg ctgggggttc agcagccagc gctttactgg catttcagga 1500
acaagcgggc actgctcgac gcaactgtct cgctcagtat cgctcgggac gcacggcgcg 1560
ctctacgaac tgccgataaa cagaggatta aaattgacaa ttgtgattaa ggctcagatt 1620
cgacggcttg gagcggccga cgtgcaggat ttccgcgaga 1660

<210> 65

<211> 868

<212> ADN

5 <213> Escherichia coli

<400> 65

agaacacgag cacggcaccg gcgaccacta tgccaagaat gccaaggtta aaaattgccg 60
gccccgccat gaagtccgtg aatgccccga cggccgaagt gaagggcagg ccgccacca 120
ggccgccgcc ctactgccc ggcacctggt cgctgaatgt cgatgccagc acctgcggca 180
cgtcaatgct tccgggctgc gcgctcgggc tgatcgcca tcccgttact gccccgatcc 240
cggcaatggc aaggactgcc agcgtgcca tttttgggt gaggcggttc gcggccgagg 300
ggcgcagccc ctggggggat gggaggcccg cgttagcggg cgggagggt tcgagaaggg 360
ggggcacccc ccttcggcgt gcgcggtcac gcgcacaggg cgcagccctg gttaaaaaca 420
aggtttataa atattggttt aaaagcagg taaaagacag gttagcgtg gccgaaaaac 480
ggcgcgaaac ccttgcaaat gctggatttt ctgcctgtgg acagcccctc aaatgtcaat 540
aggtgcgccc ctcatctgtc agcactctgc ccctcaagtg tcaaggatcg cggccctcat 600
ctgtcagtag tcgcgcccct caagtgtcaa taccgcaggg cacttatccc caggcttgtc 660
cacatcatct gtgggaaact cgcgtaaaat caggcgtttt cgccgatttg cgaggctggc 720
cagctccacg tcgccggccg aaatcgagcc tgcccctcat ctgtcaacgc cgcgccgggt 780
gagtcggccc ctcaagtgtc aacgtccgcc cctcatctgt cagtgagggc caagttttcc 840
gcgaggtatc cacaacgccg gcggccgc 868

<210> 66

<211> 3129

10 <212> ADN

<213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 66

ES 2 843 556 T3

cgctcaccgg gctggttgcc ctccgctg ggctggcggc cgtctatggc cctgcaaacg	60
cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt tgtggatacc	120
tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgaggggccg	180
actcaccggc cgccggcgtg acagatgagg ggcaggctcg atttcggccg gcgacgtgga	240
gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc ccacagatga	300
tgtggacaag cctggggata agtgcctg gcgtattgaca cttgaggggc gcgactactg	360
acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgaggggcgc	420
acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc aagggtttcc	480
gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa	540
acctgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg aaggggggtg	600
cccccccttc tcgaacctc ccggcccgct ctcgagttgg cagcatcacc cataattgtg	660
gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa	720
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct	780
tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata	840
aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt	900
aaaagatacg gaaggaatgt ctctgcttaa ggtatataag ctggtgggag aaaatgaaaa	960
cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga	1020
aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaaggtec tgcactttga	1080
acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc tttgctcgga	1140
agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcacag	1200
gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt	1260
agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga	1320
agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc	1380
cgaagaggaa cttgtctttt ccacggcgga cctgggagac agcaacatct ttgtgaaaga	1440
tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaaca agtggtatga	1500
cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct	1560
attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact	1620
ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca	1680
ccgacttctt ccgcatcaag tggtttggct ctcaggccga ggccacggc aagtatttgg	1740
gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg	1800

ES 2 843 556 T3

gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg	1860
caccagggcg gtcaaatacag gaataagggc acattgcccc ggctgagtc ggggcaatcc	1920
cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg	1980
acgcgggggtt ttccgcccag gatgccgaaa ccacgcgaag ccgcaccgtc atgcgtgcgc	2040
cccgcgaaac cttccagtcc gtcggctcga tggccagca agctacggcc aagatcgagc	2100
gcgacagcgt gcaactggct cccctgccc tgcccgccc atcgccccc gtggagcgtt	2160
cgctgctct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc gacacgcgag	2220
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg	2280
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt	2340
ccttgctcga tattgcgccg tggccggaca cgatgcgagc gatgccaac gacacggccc	2400
gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgtg caaaacaagg	2460
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag ctgcgggccg	2520
acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgacc cctatcggcg	2580
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggtcg atcaatggcc	2640
ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca	2700
cgtccgaccg cggtgggcac ctggaatcgg tgcgctgct gcaccgcttc cgctcctgg	2760
accgtggcaa gaaaacgtcc cggtgccagg tcctgatcga cgaggaaatc gtcgtgctgt	2820
ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg	2880
cccgacggat gttcgactat ttacgctcgc accgggagcc gtaccgctc aagctggaaa	2940
ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gagcaggtcg	3000
gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg	3060
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg ggttcagcca	3120
gcgctttac	3129
<210> 67	
<211> 1026	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> Construcción sintética	
<400> 67	
atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga aaagttcgac	60
agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt cagcttcgat	120
gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt ctacaaagat	180

ES 2 843 556 T3

cgttatgttt atcggcactt tgcacgccc gcgctccga ttccggaagt gcttgacatt 240
 ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg tgtcacgttg 300
 caagacctgc ctgaaaccga actgcccgcgt gttctgcagc cggctcgcgga ggccatggat 360
 gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg accgcaagga 420
 atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc ccatgtgtat 480
 cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgagggc tctcgatgag 540
 ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc ggatttcggc 600
 tccaacaatg tctgacgga caatggccgc ataacagcgg tcattgactg gagcgagggc 660
 atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc gtggttggt 720
 tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc aggatcgccg 780
 cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag cttggttgac 840
 ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt ccgatccgga 900
 gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg gaccgatggc 960
 tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc gagggcaaag 1020
 gaatag 1026

<210> 68

<211> 4148

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 68

gtttaccgc caatatatcc tgtcaaacac tgatagtta aactgaaggc gggaaacgac 60
 aatctgatca tgagcggaga attaaggag tcacgttatg acccccgccg atgacgggg 120
 acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagcaag 180
 ctggtacgat tgtaatacga ctactatag ggcgaattga gcgctgtta aacgtcttc 240
 aactggaaga gcggttacca gagctggtca cttttgtcca ccaagatgga actgcggcct 300
 cgagatccac tagagcttgc atgcctgcag tgcagcgtga cccggtcgtg cccctctcta 360
 gagataatga gcattgcatg tctaagttat aaaaaattac cacatatttt tttgtcaca 420
 cttgtttgaa gtgcagttta tctatcttta tacatatatt taaactttac tctacgaata 480
 atataatcta tagtactaca ataatatcag tgttttagag aatcatataa atgaacagtt 540
 agacatggtc taaaggacaa ttgagtattt tgacaacagg actctacagt tttatctttt 600
 tagtgtgcat gtgttctcct ttttttttgc aaatagcttc acctatataa tacttcatcc 660
 attttattag tacatccatt tagggtttag ggtaaatggt ttttatagac taattttttt 720

10

ES 2 843 556 T3

agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaaac tctattttag	780
tttttttatt taataattta gatataaaat agaataaaat aaagtgacta aaaattaaac	840
aaataccctt taagaaatta aaaaaactaa ggaaacattt ttcttgtttc gagtagataa	900
tgccagcctg ttaaacgccg tcgacgagtc taacggacac caaccagcga accagcagcg	960
tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg tcgctgcctc tggacccctc	1020
tcgagagtgc cgctccaccg ttggacttgc tccgctgtcg gcattccagaa attgcgtggc	1080
ggagcggcag acgtgagccg gcacggcagg cggcctcctc ctctctcac ggcaccggca	1140
gctacggggg attcctttcc caccgctcct tcgctttccc ttctcgcgcc gccgtaataa	1200
atagacaccc cctccacacc ctctttcccc aacctcgtgt tggtcggagc gcacacacac	1260
acaaccagat ctcccccata tccaccgctc ggcacctccg cttcaaggta cgcgcgtcgt	1320
cctccccccc cccctctctt accttctcta gatcggcggt ccggtccatg catggttagg	1380
gcccggtagt tctactcttg ttcatgtttg tgtagatcc gtgttttgtg tagatccgtg	1440
ctgctagcgt tcgtacacgg atgcgacctg tacgtcagac acgttctgat tgctaacttg	1500
ccagtgtttc tctttgggga atcctgggat ggctctagcc gttccgcaga cgggatcgat	1560
ttcatgattt tttttgttgc gttgcatagg gtttggtttg cccttttcct ttatttcaat	1620
atatgccgtg cacttgtttg tcgggtcatc ttttcatgct tttttttgtc ttggttgtga	1680
tgatgtggtc tggttgggag gtcgttctag atcggagtag aattctgttt caaactacct	1740
ggtggattta ttaatttttg atctgtatgt gtgtgccata catattcata gttacgaatt	1800
gaagatgatg gatggaata tcgatctagg ataggatac atgttgatgc gggttttact	1860
gatgcatata cagagatgct ttttgttcgc ttggttgtga tgatgtggtg tggttgggag	1920
gtcgttcatt cgttctagat cggagtagaa tactgtttca aactacctgg tgtatttatt	1980
aattttggaa ctgtatgtgt gtgtcataca tcttcatagt tacgagttta agatggatgg	2040
aaatatcgat ctaggatagg tatacatgtt gatgtgggtt ttactgatgc atatacatga	2100
tggcatatgc agcatctatt catatgctct aaccttgagt acctatctat tataataaac	2160
aagtatgttt tataattatt ttgatcttga tatacttggg tgatggcata tgcagcagct	2220
atatgtggat ttttttagcc ctgccttcat acgctattta tttgcttggg actgtttctt	2280
ttgtcgatgc tcaccctgtt gtttggtgtt acttctgcag gtgcacttta acttagccta	2340
ggatccacac gacaccatgt ccccgagcg ccgccccgct gagatccgcc cggccaccgc	2400
cggcgacatg gccgccgtgt gcgacatcgt gaaccactac atcgagacct ccaccgtgaa	2460
cttccgcacc gagccgcaga ccccgagga gtggatcgac gacctggagc gcctccagga	2520
ccgctacccg tggctcgtgg ccgaggtgga gggcgtggtg gccggcatcg cctacgccgg	2580

ES 2 843 556 T3

cccggtggaag gcccgcaacg cctacgactg gaccgtggag tccaccgtgt acgtgtccca	2640
ccgccaccag cgcctcggcc tcggctccac cctctacacc cacctcctca agagcatgga	2700
ggcccagggc ttcaagtccg tgggtggcgt gatcggcctc ccgaacgacc cgtccgtgcg	2760
cctccacgag gccctcggct acaccgcccg cggcaccctc cgcgcgcccg gctacaagca	2820
cggcggctgg cagcagctcg gcttctggca gcgcgacttc gagctgccgg ccccgcccg	2880
cccggtgctc ccggtgacgc agatctccgg tggaggcggc agcgggtggc gaggtccgg	2940
aggcgggtgg tccatggcct cctccgagga cgtcatcaag gagttcatgc gcttcaaggt	3000
gcgcatggag ggctccgtga acggccacga gttcgagatc gagggcgagg gcgagggccg	3060
cccctacgag ggcacccaga ccgccaagct gaaggtgacc aagggcggcc ccctgccctt	3120
cgcctgggac atcctgtccc ccagttcca gtacggctcc aaggtgtacg tgaagcacc	3180
cgccgacatc cccgactaca agaagctgtc cttccccgag ggcttcaagt gggagcgcgt	3240
gatgaacttc gaggacggcg gcgtggtgac cgtgaccag gactcctccc tgcaggacgg	3300
ctccttcac tacaaggtga agttcatcgg cgtgaacttc ccctccgacg gccccgta	3360
gcagaagaag actatgggct gggaggcctc caccgagcgc ctgtaccccc gcgacggcgt	3420
gctgaagggc gagatccaca aggccctgaa gctgaaggac ggcgccact acctggtgga	3480
gttcaagtcc atctacatgg ccaagaagcc cgtgcagctg cccggctact actacgtgga	3540
ctccaagctg gacatcacct cccacaacga ggactacacc atcgtggagc agtacgagcg	3600
cgccgagggc cgccaccacc tgttcctgta gtcaggatct gagtcgaaac ctagacttgt	3660
ccatcttctg gattggccaa cttaattaat gtatgaaata aaaggatgca cacatagtga	3720
catgctaata actataatgt gggcatcaaa gttgtgtgtt atgtgtaatt actagttatc	3780
tgaataaaag agaaagagat catccatatt tcttatccta aatgaatgtc acgtgtcttt	3840
ataattcttt gatgaaccag atgcatttca ttaaccaa atcatatacat ataaatatta	3900
atcatatata attaatatca attgggttag caaaacaaat ctagtctagg tgtgttttgc	3960
gaatgcggcc gccaccgcgg tggagctcga attcattccg attaatcgtg gcctcttgct	4020
cttcaggatg aagagctatg tttaaacgtg caagcgctac tagacaattc agtacattaa	4080
aaacgtccgc aatgtgttat taagttgtct aagcgtcaat ttgtttacac cacaatatat	4140
cctgccac	4148
<210> 69	
<211> 18762	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> Construcción sintética	
<400> 69	

ES 2 843 556 T3

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac	60
aatctgatca tgagcggaga attaaggagag tcacgttatg acccccgcgg atgacgcggg	120
acaagccggt ttacgttttg aactgacaga accgcaacgt tgaaggagcc actcagcaag	180
ctggtacgat tgtaatacga ctactatag ggcgaattga gcgctgttta aacgctcttc	240
aactggaaga gcggttacgc tgtttaaacg ctcttcaact ggaagagcgg ttactaccgg	300
ttactagct agctgctaag gttaccagag ctggtcacct ttgtccacca acttattaag	360
tatctagttg aagacacggt cttcttcacg taagaagaca ctcagtagtc ttcggccaga	420
atggcctctt gattcagcgg gcctagaagg ccggtacact gactagctaa tttaaatcct	480
gaggatctgg tcttcctaag gacccgggat atcggaccga ttaaaacttta attcgggtccg	540
ataacttcgt atagcataca ttatacgaag ttatatgcat acgcgtctta agtcgatcgc	600
tatcaacttt gtatagaaaa gttgggccga gctcggtagc gccagaatgg cccggaccgg	660
gttaccgaat tcaggcgacc catcgctgct ttgtctacat catgttcttc atcatcctcc	720
ccaggcgacg cgtgctgctg ttcttattca gactaccgtt cgagtgactg catggcgtag	780
atctttctgc atcgactttg tacggctaca tcgaacatat acacgagatg tctcgtgtga	840
atagagtcac taatgcctta agcatcgggt actccgtagg gtacattctg ttcttcttat	900
ttgtgcatat ttttattggt gtttactgat tatacagagta gttatacata catgcacata	960
catatcatca catatatcac aatatTTTTT taaattaaat taaaactaaa aatgactaaa	1020
tttctaacac caacgacatt gtaatgtttt ctccaacaac tttaacctatt ctacattggt	1080
ctattttcga tttcactcta taaacaacat agtctacaat ggaaaacagt gctttgtacg	1140
actatatacg cgatgtgtgg ctacaacata agacaatata gtcgtttgaa gattgaacct	1200
atatatcggg acggttaatc cgtctatgta cgtgggcatg acgaacaccc gtgataacga	1260
aggattaacg tgcacaatca taaatccaaa gtaggagcgg tgcgatgatga gaatcgctct	1320
cagtactcga cataatgaac cttacgaggt acaacaggca ggcaggcagg gaccaggggc	1380
cgcctttatt tcaggctcgc tggccccacg ggcgtgctgc gtgcacgaag ggcactaccc	1440
caacctctca ccgaaaaccg cgctggatcg gcaaatcaaa cgagggtggtg ccccggtgcc	1500
actctccacg tccacggcac catccctctg cagccgctca ccagccatgc cgtgtcgcgg	1560
aacggcacia ccacccccaa cccactcacg aaaccccgtc ccggccgtgc ccgtgtcggg	1620
ccgcgctcgg caacgaggcg gcccgcgctg ctgagtcgcc tggacaccgg acacctgtc	1680
ggccctttgt ttattcatcc cgaaatctca tctgccccca cggccgactg cgtgtcgcgg	1740
cccgatata tatacccatc gttatcgatc gactggggac tctatcagtg atagagtcta	1800
gaggatcgct caggaaggcc gctgagatag aggcattggc gccaatgcgg gcggcggtgg	1860

ES 2 843 556 T3

agcgggagga ggcagcggca ggcgcagcgt ggctgcgccg gcggtgtgcc gcccagcgg	1920
ctcgcggttg acgccgacgc cggagcagat caggatgctg aaggagctct actacggtg	1980
cggcatccgg tcgccagct cggagcagat ccagcgcac accgccatgc tgcggcagca	2040
cggcaagatc gagggcaaga acgtcttcta ctggttccag aaccacaagg cccgcgagcg	2100
ccagaagcgc cgcctcacca gcctcgacgt caacgtgccc gccgcggcg cggccgacgc	2160
caccaccagc caactcggcg tcctctcgct gtcgtcgccg ccgccttcag gcgcggcgcc	2220
tcctctgccc accctcggct tctacgcgc cggcaatggc ggcggatcgg ctgtgctgct	2280
ggacacgagt tccgactggg gcagcagcgg cgctgccatg gccaccgaga catgcttcct	2340
gcaggactac atgggcgtga cggacacggg cagctcgctc cagtggccac gcttctcgtc	2400
gtcggacacg ataatggcgg cggccgcggc gcgggcggcg acgacgcggg cggccgagac	2460
gtccctcttc tcccgacct gcggcgacga cggcggcagc ggtagcagca gctacttgcc	2520
gttctggggt gccgcgtcca caactgccgg cgccacttct tccgttgca tccagcagca	2580
acaccagctg caggagcagt acagctttta cagcaacagc aacagcacc agctggccgg	2640
caccggcaac caagacgtat cggcaacagc agcagcagcc gccgccctgg agctgagcct	2700
cagctcatgg tgctccctt accctgctgc agggagtatg tgagagcaac gcgagctgcc	2760
actgctcttc actggtaccg ttaacagatc attcgacaaa gcagcattag tccgttgatc	2820
ggtggaagac cactcgtcag tgttgagttg aatgtttgat caataaaata cggcaatgct	2880
gtaagggttg ttttttatgc cattgataat aactgtact gttcagttgt tgaactctat	2940
ttcttagcca tgccaagtgc tttcttatt ttgaataaca ttacagcaaa aagttgaaag	3000
acaaaaaaaa aaacccccga acagagtgtt ttgggtccca agcttcttta gactgtgttc	3060
ggcgttcccc ctaaatttct cccctatat ctactcact tgtcacatca gcgttctctt	3120
tccccctata tctccacgct ctacagcagt tccacctata tcaaacctct ataccacc	3180
acaacaatat tatatacttt catcttcaac taactcatgt acctccaat tttttctac	3240
taataattat ttacgtgcac agaaacttag caaggagaga gagagcgggg tgaccacct	3300
tgctagttgg atattacctc ttctcttcaa agtatccttg aacgctcacc ggttatcaaa	3360
tctctacact atagctctgt agtcttgcta gatagttagt tctttagctc tcggtgacca	3420
agcttggcgc gatcaagctt atcgataccg tcgacctcga agcttggta cccggtccgg	3480
gcctagaagg ccagcttcaa gttgtacaa aaaagcaggc tccggccaga atggcccgga	3540
ccgggttacc gaattcttac cctagctccc tgcggctgtt acgcggtccc ccatcaatct	3600
tctgttcttg cggttgtagc ctgtgtaaca gtgctagagt atgtatgata aataggtttt	3660
aagtctgctt acatgacatt ttttattgtg gaagagacat ataaaaatta gagagagtgg	3720
ttctcatgca acggcggacg gcccggtgct aaaagagctt caagacaaaa taatgaaaca	3780

ES 2 843 556 T3

ggaagagagt agatttatct aagagccaac tttattatat gaatgtgttt attgttggct	3840
ttagatgata tggtaaggag ttagagctaa taatagatag gctctattat tattattatt	3900
aattaaactc gctctaagga ggaaagtggg aggaaggagc gaggacgaag actactggaa	3960
gcctcgtgca tggatgatgg atgtggtgtc tcttaatgta ggtggccgga ggatgtacgt	4020
gttaattgctg cgataagcac tcagatccaa ccgcaacta cctccacact gacacactga	4080
tagagagaaa gagagacctc cgacgactgc cgccgcagat gagccacgta cgtatacgac	4140
gtctgccggc cggctcaggc tgcgccatc accctgctcg aaagtgcgct taggcggcgc	4200
cagctacata ggagtatcta gtctagccag ttagtatact actactgcgc tgatgatgaa	4260
ttaactctgc atagatactg tacttgctc cctccaacac ccaaccacct cctgctcggc	4320
tcttaataac ttggacacgg atcgatgcca tccaaggaag aacacgacga cgacgacgga	4380
acatccacca tgcaagcttg catccatacg ccgatacgcg tgcattccat catccacct	4440
tatttccatt ttccaccgat cacacgtaca caggcctatt taaggagcga catcccactg	4500
caactctcct caccactcat caccagctag ctctagcaaa gcacttgcca tctaccgacc	4560
gccgcattcc aaacagcccg acgagctagc agagcggcag gcacctccct cctcaaggaa	4620
cccatggcca ctgtgaacaa ctggctcgtc ttctccctct ccccgagga gctgccgcc	4680
tcccagacga cggactccac actcatctcg gccgccaccg ccgacctgt ctccggcgat	4740
gtctgcttca acatcccca agattggagc atgaggggat cagagcttcc ggcgctcgtc	4800
gcggagccga agctggagga ctctctcggc ggcatctcct tctccgagca gcatcacaag	4860
gccaactgca acatgatacc cagcactagc agcacagttt gctacgcgag ctgaggtgct	4920
agcaccggct accatcacca gctgtaccac cagcccacca gctcagcgct ccacttcgcg	4980
gactccgtaa tgggtggttc ctccggccgt gtccacgacg gcggtgccat gctcagcgcg	5040
gccgccgcta acggtgtcgc tggcgctgcc agtgccaacg gcggcgccat cgggctgtcc	5100
atgattaaga actggctgcg gagccaaccg gcgcccatgc agccgagggg ggcgcgcgct	5160
gagggcgcg aggggctctc tttgtccatg aacatggcgg ggacgacca aggcgctgct	5220
ggcatgccac ttctcgtcgg agagcgcgca cgggcgccc agagtgtatc gacgtcagca	5280
cagggtgag ccgtcgtcgt cacggcgccg aaggaggata gcggtggcag cgggtgtgcc	5340
ggcgctctag tagccgtgag cacggacacg ggtggcagcg gcggcgcgtc ggctgacaac	5400
acggcaagga agacggtgga cacgttcggg cagcgcacgt cgatttaccg tggcgtgaca	5460
aggcatagat ggactgggag atatgaggca catctttggg ataacagttg cagaaggga	5520
gggcaaacctc gtaagggtcg tcaagtctat ttaggtggct atgataaaga ggagaaagct	5580
gctagggctt atgatcttgc tgctctgaag tactggggtg ccacaacaac aacaaat	5640

ES 2 843 556 T3

ccagttagta	actacgaaaa	ggagctcgag	gacatgaagc	acatgacaag	gcaggagttt	5700
gtagcgtctc	tgagaaggaa	gagcagtggt	ttctccagag	gtgcatccat	ttacagggga	5760
gtgactaggc	atcaccaaca	tggaagatgg	caagcacgga	ttggacgagt	tgcagggaac	5820
aaggatcttt	acttgggcac	cttcagcacc	caggaggagg	cagcggaggc	gtacgacatc	5880
gcggcgatca	agttccgcgg	cctcaacgcc	gtcaccaact	tcgacatgag	ccgctacgac	5940
gtgaagagca	tcctggacag	cagcgccctc	cccatcggca	gcgccgcaa	gcgcctcaag	6000
gagggcgagg	ccgcagcgtc	cgcgcagcac	caccacgccg	gcgtggtgag	ctacgacgtc	6060
ggccgcacgc	cctcgcagct	cggcgacggc	ggagccctgg	cggcgcgcta	cggcgcgcac	6120
taccacggcg	ccgcctggcc	gaccatcgcg	ttccagccgg	gcgccgccag	cacaggcctg	6180
taccacccgt	acgcgcagca	gccaatgcgc	ggcgcggggt	ggtgcaagca	ggagcaggac	6240
cacgcgggtga	tcgcggccgc	gcacagcctg	caggacctcc	accacctgaa	cctgggcgcg	6300
gccggcgcg	acgacttttt	ctcggcaggg	cagcaggccg	ccgccgctgc	gatgcacggc	6360
ctgggtagca	tcgacagtgc	gtcgtctcgag	cacagcaccg	gctccaactc	cgtcgtctac	6420
aacggcgggg	tcggcgacag	caacggcgcc	agcgccgtcg	gcggcagtg	cgggtggctac	6480
atgatgccga	tgagcgctgc	cggagcaacc	actacatcgg	caatggtgag	ccacgagcag	6540
gtgcatgcac	gggcctacga	cgaagccaag	caggctgctc	agatggggta	cgagagctac	6600
ctggtgaacg	cggagaacaa	tggtggcgga	aggatgtctg	catgggggac	tgtcgtgtct	6660
gcagccgcgg	cggcagcagc	aagcagcaac	gacaacatgg	ccgccgacgt	cgggcatggc	6720
ggcgcgcagc	tcttcagtgt	ctggaacgac	acttaagcgt	acctagtgg	acctgacatc	6780
ttatagtctg	caacctctcg	tgtctgaatt	cctatcttta	tcaagtgtta	ttgcttccac	6840
gactatagga	cagctttcgt	cgaagccttt	tgctcatgtg	atctcgaagg	attcatctag	6900
tctgattttt	cgtgacttgt	atcggtttta	ttggattcat	ccaacatata	tcaataaaaa	6960
atgagttgtg	tttcctttct	tcctagtcca	gttaaaatta	tttcctcctc	gcgcttgtgc	7020
tgtaattgtc	tgtgtacctg	ttgtttgtga	ctgtgttagt	tccttggat	atgatttcgt	7080
atttgatatg	tacatggaga	tagcttagct	tcattattgg	agtatgaagt	tagtatgaca	7140
tagtcactct	cctggaaaat	tgacactgca	aaccatattt	ttattctgaa	ccacaaatcc	7200
tagtcagtcc	gctggcata	gccgtccgtt	tgctgaatcc	agaacgtggg	tttgagatg	7260
tacggctgag	atgcctctat	gcgaagggga	tttcgtggtg	aaacgagatg	ggagtagagc	7320
aacgcccgtg	gaagatgctt	caaaattcca	cacttttgag	caacgatcgg	cagtagtaag	7380
gtagacgatt	tcaagatcaa	agcatatgaa	gataaacaac	atcaacaaca	aaatttggtg	7440
gggttctata	gagagaaaca	gagctacata	catacactgt	tttgatatcta	ccatctgaga	7500
tgatgaaaag	atgaaaaact	aaagaatgcc	ccggcgccaa	cgccaggaca	cgccgcgcgc	7560

ES 2 843 556 T3

gcgtcaccgc agccatctct tgaccagcc ggcgctgtat atttacacac gttgcagcat	7620
cgatcaccac ctgttcgac gcgtcgccgt caccgggac tgagtcgaaa cctagacttg	7680
tccatcttct ggattggcca acttaattaa tgtatgaaat aaaaggatgc acacatagt	7740
acatgctaata cactataatg tgggcatcaa agttgtgtgt tatgtgtaata tactagtatt	7800
ctgaataaaa gagaaagaga tcatccatat ttcttatcct aaatgaatgt cacgtgtctt	7860
tataattctt tgatgaacca gatgcatttc attaaccaaa tccatataca tataaatatt	7920
aatcatatat aattaatatc aattgggtta gcaaaacaaa tctagtctag gtgtgttttg	7980
cgaatgcgac cttcttatgt gcttctagtc tccaaatgtg gttgatagtt attttgctct	8040
aagatcaaca gtaatgaagt ataaatcatc gttgtggtgt gctactcggg taattgagca	8100
ttaacacaca caaacatgac gaggatggtta taatctccaa aaatgtgtac tttgttaggt	8160
gggacctat agccttgatt aatgtgctat gttaggcatg cctggaaacg tgtgacgcat	8220
atgttttgtg aacctgttga tattatatgt gcttttatat taccatattt tattaataa	8280
ctaataattt ttactagtaa gatataacat tctatctagc ttaaaaacta accataaata	8340
ttccataata actagattta ccaaactaat atactaaata tacataataa atacaaaatt	8400
aacaagacaa taatcaatat ttatgagctt aatatattta gacattatgg ttggtcgacg	8460
ataatcatgc taacttttcg taattgcttg attgaaatat gcttagaata atgcctcttt	8520
gttctacatg gcaaatagg accattatgg tgtaacacc tgggaaccac aaacaccccg	8580
aaatgctact aaactacaca actaaccttc atatataaaa ttctgacagc atctcctttg	8640
aaaatttgca tagacgtgga agcaacagag tataaacaga tatcatgata agaaaacata	8700
ctagacatta ataatctgct agaaatggga agaactctaa cttgacgact gcgtaactga	8760
ctagagtcac acttagctga ccctagtcac ttacaactga cttcgtgtcc tagatcgatg	8820
ggccctggcc gaagcttggg caccgggtcc gggcctagaa ggccagcttc aagtttgtag	8880
aaaaaagcag gctccggcca gaatggcccg gaccgggtta ccgaattcga gctcggtagc	8940
ctgggatccg attgactatc tcatctctca aaccaaacac ctcaaataa tctgctatcg	9000
ggattggcat tcctgtatcc ctacgcccgt gtacccctg tttagagaac ctccaaagg	9060
ataagatggc gaagattatt gttgtcttgt ctttcatcat atatcgagtc tttccctagg	9120
atattattat tggcaatgag cattacacgg ttaatcgatt gagagaacat gcatctcacc	9180
ttcagcaaat aattacgata atccatattt tacgcttcgt aacttctcat gagtttcgat	9240
atacaaattt gttttctgga caccctacca ttcatcctct tcggagaaga gaggaagtgt	9300
cctcaattta aatatgttgt catgctgtag ttcttcacaa aatctcaaca ggtaccaagc	9360
acattgtttc cacaaattat attttagtca caataaatct atattattat taatatacta	9420

ES 2 843 556 T3

aaactatact	gacgctcaga	tgcttttact	agttcttgct	agtatgtgat	gtaggtctac	9480
gtggaccaga	aaatagttag	acacggaaga	caaaagaagt	aaaagaggcc	cggactacgg	9540
cccacatgag	attcggcccc	gccacctccg	gcaaccagcg	gccgatccaa	cggcagtgcg	9600
cgcacacaca	caacctcgta	tatatcgccg	cgcggaagcg	gcgcgaccga	ggaagccttg	9660
tcctcgacac	ccctacaca	ggtgtcgcg	tgcccccgac	acgagtcctg	catgctctcc	9720
acgcggccgc	gccagatccc	gcctccgcgc	gttgccacgc	cctctataaa	caccagctc	9780
tcctcgccc	tcctctacct	cactcgtagt	cgtagctcaa	gcacagcgg	cagcggcagc	9840
ggcaggagct	ctgggcagcg	tgccgacgtg	gggtacctag	ctcgtctctg	tagcctacct	9900
taagatatcg	gatccatgtc	caacctgctc	acggttcacc	agaaccttcc	ggctcttcca	9960
gtggacgcga	cgtccgatga	agtcaggaag	aacctcatgg	acatgttccg	cgacaggcaa	10020
gcgttcagcg	agcacacctg	gaagatgctg	ctctccgtct	gccgctcctg	ggctgcatgg	10080
tgcaagctga	acaacaggaa	gtggttcccc	gctgagcccg	aggacgtgag	ggattacctt	10140
ctgtacctgc	aagcgcgagg	tttgtttctg	cttctacctt	tgatatatat	ataataatta	10200
tcattaatta	gtagtaatat	aatatattcaa	atattttttt	caaaataaaa	gaatgtagta	10260
tatagcaatt	gcttttctgt	agtttataag	tgtgtatat	ttaatttata	acttttctaa	10320
tatatgacca	aaacatggtg	atgcctaggt	ctggcagtg	agaccatcca	gcaacacctt	10380
ggacaactga	acatgcttca	caggcgctcc	ggcctccgc	gccccagcga	ctcgaacgcc	10440
gtgagcctcg	tcctgcgccg	catcaggaag	gaaaacgtcg	atgccggcga	aagggcaaag	10500
caggccctcg	cgttcgagag	gaccgatttc	gaccaggtcc	gcagcctgat	ggagaacagc	10560
gacaggtgcc	aggacattag	gaacctggcg	ttcctcggaa	ttgcatacaa	cacgctcctc	10620
aggatcgcg	aaattgcccc	cattcgctg	aaggacatta	gccgcaccga	cggcggcagg	10680
atgcttatcc	acattggcag	gaccaagacg	ctcgtttcca	ccgcaggcgt	cgaaaaggcc	10740
ctcagcctcg	gagtgaacaa	gctcgtcgaa	cgtcggtat	ccgtgtccgg	cgtcgcggac	10800
gacccaaaca	actacctctt	ctgccgcgtc	cgcaagaacg	gggtggctgc	ccctagcgcc	10860
accagccaac	tcagcacgag	ggccttggaa	ggtatttttcg	aggccaccca	ccgcctgac	10920
tacggcgcg	aggatgacag	cgttcaacgc	tacctcgcat	ggtccgggca	ctccgcccgc	10980
gttggagctg	ctagggacat	ggcccgcgcc	ggtgtttcca	tccccgaaat	catgcaggcg	11040
ggtggatgga	cgaacgtgaa	cattgtcatg	aactacattc	gcaaccttga	cagcgagacg	11100
ggcgcaatgg	ttcgctcct	ggaagatggt	gactgaggta	cccaacctag	acttgtccat	11160
cttctggatt	ggccaactta	attaatgtat	gaaataaaaag	gatgcacaca	tagtgacatg	11220
ctaatacta	taatgtgggc	atcaaagttg	tgtgttatgt	gtaattacta	gttatctgaa	11280
taaaagagaa	agagatcatc	catatttctt	atcctaaatg	aatgtcacgt	gtctttataa	11340

ES 2 843 556 T3

ttctttgatg aaccagatgc atttcattaa ccaaattccat atacatataa atattaatca	11400
tatataatta atatcaattg ggtagcaaaa acaaattctag tctaggtgtg ttttgcgatc	11460
cgatatcgat gggccctggc cgaagcttgg tcaccgggtc cgggcctaga aggcgatct	11520
cccgggcacc cagctttctt gtacaaagtg gccgttaacg gatcgccag aatggcccgg	11580
accgggttac cgaattcctg cccttaaggc caattgttca agattcattc aacaattgaa	11640
acatctccca tgattaaatc agtataaggt tgctatggtc ttgcttgaca aagttttttt	11700
ttgagggaaat ttcaactaaa tttttgagtg aaactatcaa atactgattt taaaaaattt	11760
ttataaaaagg aagcgcagag ataaaaggcc atctatgcta caaaagtacc caaaaatgta	11820
atcctaaagt atgaattgca tttttttttt gtttggacga aaggaaagga gtattaccac	11880
aagaatgata tcatcttcat atttagatct tttttgggta aagcttgaga ttctctaaat	11940
atagagaaat cagaagaaaa aaaaaccgtg ttttgggtgt tttgatttct agcctccaca	12000
ataactttga cggcgtcgac aagtctaacg gacaccaagc agcgaaccac cagcgcgag	12060
ccaagcgaag cagacggccg agacgttgac accttcggcg cggcatctct cgagagttcc	12120
gctccggcgc tccacctoca ccgtggcggt tttcttattc cgttccgttc cgcctoctgc	12180
tctgctcctc tccacaccac acggcacgaa accgttacgg caccggcagc acccagcacg	12240
ggagagggga ttcctttccc accgttcctt ccctttccgc cccgccgcta taaatagcca	12300
gccccatccc cagctttttt ccccaatctc atctcctctc tcctgttgtt cggagcacac	12360
gcacaatccg atcgatcccc aaatcccctt cgtctctcct cgcgagcctc gtggatccca	12420
gcttcaaggt acggcgatcg atcatcccc ctctctctct ctaccttctt ttctctagac	12480
tacatcggat ggcgatccat ggtagggcc tgctagtctc ccttctgtt ttgtcgatgg	12540
ctgcgaggca caatagatct gatggcggtta tgacggctaa cttgtcatgt tgttgcgatt	12600
tatagtccct ttaggagatc agtttaattt ctcgatgggt tcgagatcgg tggatccatg	12660
ttagtaccct aagatccgcg ctgttaggggt tcgtagatgg aggcgacctg ttctgattgt	12720
taacttgtca gtacctggga aatcctggga tggttctagc tcgtccgcag atgagatcga	12780
tttcatgatc ctctgtatct tgtttcgttg cctaggttcc gtctaatacta tccgtgggat	12840
gatgtagatg ttttgatcgt gtaactacg tcttgtaaag ttaattgtca ggtcataatt	12900
tttagcatgc cttttttttt gtttggtttt gtctaattgg gctgtcgttc tagatcagag	12960
tagaagactg ttccaaaacta cctgctggat ttattgaact tggatctgta tgtgtgtcac	13020
atatcttcat aaattcatga ttaagatgga ttgaaatata ttttatcttt ttggtatgga	13080
tagttctata tgttggtgtg gctttgttag atgtatacat gcttagatac atgaagcaac	13140
gtgctgctac tgtttagtaa ttgctgttca tttgtctaat aaacagataa ggatatgtat	13200

ttatgttgct	gttggttttg	ctggtacttt	gttgataca	aatgcttcaa	tacagaaaac	13260
agcatgctgc	tacgatttac	catttatcta	atcttatcat	atgtctaata	taataaacia	13320
acatgctttt	aaattatctt	catatgcttg	gatgatggca	tacacagcgg	ctatgtgtgg	13380
ttttttaaat	accagcatc	atgggcatgc	atgacactgc	tttaatatgc	tttttatttg	13440
cttgagactg	tttcttttgt	ttatactgac	cctttagttc	ggtgactcct	ctgcaggtcg	13500
actttaactt	agcctaggat	ccatggccca	gtccaagcac	ggcctgacca	aggagatgac	13560
catgaagtac	cgcattggag	gctgcgtgga	cgccacaag	ttcgtgatca	cggcgagggg	13620
catcggttac	ccctcaagg	gcaagcaggc	catcaacctg	tgcgtggtgg	agggcgggcc	13680
cttgcccttc	gccgaggaca	tcttgtccgc	cgccttcatt	tacggcaacc	gcgtgttcac	13740
cagatcccc	caggacatcg	tcgactactt	caagaactcc	tgccccgccg	gctacacctg	13800
ggaccgctcc	ttcctgttcg	aggacggcgc	cgtgtgcata	tgcaacgccg	acatcacctg	13860
gagcgtggag	gagaactgca	tgtaccacga	gtccaagtcc	tacggcgtga	acttccccgc	13920
cgcaggcccc	gtgatgaaga	agatgaccga	caactgggag	ccctcctgcg	agaagatcat	13980
ccccgtgccc	aagcagggca	tcttgaaggg	cgcgtgagc	atgtacctgc	tgctgaagga	14040
cgttgccgcg	ttgcgtgccc	agttcgacac	cgtgtacaag	gccaagtccg	tgccccgcaa	14100
gatgcccgcg	tggcacttca	tccagcacia	gctgacctgc	gaggaccgca	gcgacgcaa	14160
gaaccagaag	tggcacctga	cgcgacgcg	catgcctccc	ggctccgcct	tgccctccgg	14220
actcagatct	cgatagggtg	ccaatggcca	gttaacagat	ccagctgctg	ctgttctagg	14280
gttcacaagt	ctgcctatct	gtcttcccca	atggagctat	ggttgtctgg	tctggtcctt	14340
ggtcgtgtcc	cgtttcattg	tgtactatct	acctgtaata	tgtatcctta	agtctggttt	14400
gatggtgtct	gaaacgtttt	gctgtggtag	agcagcatgg	aagaactata	atgaataagt	14460
gatccctaata	cattgtgtcc	aaattttgct	tctgctatac	ccttttgtgc	tgtttcttat	14520
gttttgctta	aaaatttgat	ctgacaaaac	aatttgtcta	aattatgctc	ttgttctgac	14580
tgtgtactgt	acacatttgt	attgctaccg	gtttggcaca	tagttttctc	tttgattgca	14640
gatcaaacc	ttctgattgc	agattgcttg	gctatggctg	atgtgctcac	tgttcttctt	14700
tagcttgat	attgctgatg	gtggcttgcc	tacggctgaa	aaagtgtctg	ctgcgccgac	14760
ttggcaactg	ttgctccagt	aatgctttgt	cctctcttcc	tagccgcga	cttggcactt	14820
ggcagctgtt	gctcctgtaa	tgcgttgcta	tctcttatct	cttgagatgt	actatgattt	14880
tgaggctgtt	tgacttgtgc	cagattgcca	gtgctgatgc	tatattatct	caggttttta	14940
tgctagtata	tgttttactg	tgtgctgcct	gtgtcgagtt	acgtaatacta	ttgaaccgct	15000
atgcgtgttg	atgccctgtt	gttgcttgcc	agattgatgc	ccctgttgct	tgagcttgcc	15060
aagctgtagt	gaaatcctaa	aattcgaaac	caaactctatt	ctgaactttt	gaagatctgt	15120

ES 2 843 556 T3

gatgtccgga agaaaagatt caaacagata aacgggactc tccagaatct ataaaaataaa	15180
tgatcgacaa acgctttgct ttggtaggct tcctagttag ttagcggacc gaagcttcgg	15240
tccgggccta gaaggccagc ttcggccgcc ccgggcaact ttattataca aagttgatag	15300
atcgaataac ttcgtatagc atacattata cgaagttata cctggtggcg ccgctagggg	15360
ctgcaggaat tcgatatcaa gcttatcgat accgtcgacc tcgagggggg gcccggtacc	15420
ctgggatctg cccttacttt aacgcctcta accaacaccc ctttatcttt ataaggaaca	15480
ataaacagaa tttgccccac tgttctaaat cacctaataa tatccccagc taaaaacaat	15540
aaaggtttcc tagaattaag acaagcatga ctgttcctcc aggagggttt ggaacattgt	15600
tgcagtcttg cagatacggg cgaagggtga gaaacagagc ggagggttg aggtgacctc	15660
ggtagtcgac gccggagttg agcttgacaa cgacggggcg gccctgatg gacttgagga	15720
agtccgatgg cgtcttcacc gtcccgccgg cgcccgaggc gggcctgtcg ctgccgccgc	15780
cgccgctgct catcttgccg gctgtgcccc cggcggtgtc cctgtgttgc ggatcgccgg	15840
tgggccaggt ggatgcgagg gcgaccctt tggactccgg ccggagccgc cggatccctg	15900
gtcggtgtca gtgccgttta ctctggggcc cacgtgtcag taccgtctgt agatgacaac	15960
aaccgctcgt ccacagtcac gtccaaaata tcctttcttc ttttttttcg attcggatat	16020
ctatcttcct ttttttttcc caaaaatctt cttgacgcac cagcgcgcac gtttgtggta	16080
aacgccgaca cgtcgggtcc acgtcgatag accccaccca ccagtgagta gcgtgtacgt	16140
attcgggggt gacggacgtg tcgccgtcgt cttgctagtc ccattcccat ctgagccaca	16200
catctctgaa caaaaaaag gagggaggcc tccacgcaca tccccctccg tgccaccgcg	16260
cccaaaccct cgcgccgcct ccgagacagc cgccgcaacc atggccaccg ccgccgccgc	16320
gtctaccgcg ctactcggcg ccaactaccg tgccgccaag gcgaggcgcc gggcgcacct	16380
cctggccacc cgccgcgccc tcgccgcgcc catcaggtgc tcagcggcgt caccgcctat	16440
gccgatggct ccccggccca ccccgctccg gccgtggggc cccaccgagc cccgcaaggg	16500
tgctgacatc ctcgtcgagt ccctcgagcg ctgcggcgtc cgcgacgtct tcgcctaccc	16560
cggcggcgcg tccatggaga tccaccaggc actcaccgcg tccccgtca tcgccaacca	16620
cctcttcgcg cagcagcaag gggaggcctt tgccgcctcc ggctacgcgc gtcctcggg	16680
ccgcgtcggc gtctgcatcg ccacctccgg ccccgcgccc accaacctag tctccgcgct	16740
cgccgacgcg ctgctcgatt ccgtcccat ggtcgccatc acgggacagg tggcgcgacg	16800
catgattggc accgacgcct tccaggagac gcccatcgtc gaggtcaccg gctccatcac	16860
caagcacaac tacctggtcc tcgacgtcga cgacatcccc cgcgtcgtgc aggaggcttt	16920
cttcctcgcc tcctctggtc gaccagggcc ggtgcttgtc gacatcccca aggacatcca	16980

ES 2 843 556 T3

gcagcagatg gcggtgcctg tctgggacaa gcccatgagt ctgcctgggt acattgcgcg	17040
ccttcccaag cccctgcga ctgagttgct tgagcaggtg ctgcgtcttg ttggtgaatc	17100
gcggcgccct gttctttatg tgggcggtgg ctgcgcagca tctggtgagg agttgcgacg	17160
ctttgtggag ctgactggaa tcccggtcac aactactctt atgggcctcg gcaacttccc	17220
cagcgacgac ccactgtctc tgcgcatgct aggtatgcat gggacggtgt atgcaaatta	17280
tgcagtggat aaggccgacg tgttgcttgc acttggtgtg cggtttgatg atcgcgtgac	17340
aggaagatt gaggcctttg caagcagggc taagattgtg cacgttgata ttgatccggc	17400
tgagattggc aagaacaagc agccacatgt gtccatctgt gcagatgtta agcttgcttt	17460
gcagggcatg aatgctcttc ttgaaggaag cacatcaaag aagagctttg actttggctc	17520
atggaacgat gagttggatc agcagaagag ggaattcccc cttgggtata aaacatctaa	17580
tgaggagatc cagccacaat atgctattca ggttcttgat gagctgacga aaggcgaggc	17640
catcatcggc acaggtgttg ggcagcacca gatgtgggcg gcacagtact acacttacia	17700
gcggccaagg cagtggttgt cttcagctgg tcttggggct atgggatttg gtttgccggc	17760
tgctgctggt gcttctgtgg caaaccagcgtgtgactgtt gttgacatcg atggagatgg	17820
tagctttctc atgaacgttc aggagctagc tatgatccga attgagaacc tcccagtga	17880
ggtctttgtg ctaaacaacc agcacctggg gatgggtggg cagttggagg acaggttcta	17940
taaggccaac agagcgacac catacttggg aaaccagag aatgaaagtg agatatatcc	18000
agatttcgtg acgatcgcca aagggttcaa cattccagcg gtccgtgtga caaagaagaa	18060
cgaagtccgc gcagcgataa agaagatgct cgagactcca gggccgtacc tcttgatat	18120
aatcgccca caccaggagc atgtgttgcc tatgatccct agtgggtggg ctttcaagga	18180
tatgatcctg gatggatgat gcaggactgt gtactgacta gctagtcagt taacctagac	18240
ttgtccatct tctggattgg ccaacttaat taatgtatga aataaaagga tgcacacata	18300
gtgacatgct aatcactata atgtgggcat caaagttgtg tgttatgtgt aattactagt	18360
tatctgaata aaagagaaa agatcatcca tatttcttat cctaaatgaa tgtcacgtgt	18420
ctttataatt ctttgatgaa ccagatgcat ttcattaacc aaatccatat acatataaat	18480
attaatcata tataattaat atcaattggg ttagcaaaac aaatctagtc taggtgtgtt	18540
ttgcgaatgc ggccgcctgc agagttaacg gcgcgccgac tagctagcta aggtaccgag	18600
ctcgaattca ttccgattaa tcgtggcctc ttgctcttca ggatgaagag ctatgtttaa	18660
acgtgcaagc gctactagac aattcagtac attaaaaacg tccgcaatgt gttattaagt	18720
tgtctaagcg tcaatttggt tacaccacaa tatatcctgc ca	18762
<210> 70	
<211> 18716	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> Construcción sintética	
<400> 70	

ES 2 843 556 T3

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac	60
aatctgatca tgagcggaga attaaggagag tcacgttatg acccccgcg atgacgcggg	120
acaagccggt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagcaag	180
ctggtacgat tgtaatacga ctactatag gccgaattga gcgctgttta aacgctcttc	240
aactggaaga gcggttacgc tgtttaaacg ctcttcaact ggaagagcgg ttactaccgg	300
ttcactagct agctgctaag gttaccagag ctggtcacct ttgtccacca acttattaag	360
tatctagttg aagacacgtt cttcttcacg taagaagaca ctcagtagtc ttcggccaga	420
atggcctctt gattcagcgg gcctagaagg ccggatcact gactagctaa tttaaatcct	480
gaggatctgg tcttcctaag gaccgggat atcggaccga ttaaacttta attcgggtccg	540
ataacttcgt atagcataca ttatacgaag ttatatgcat acgctctta agtcgatcgc	600
tatcaacttt gtatagaaaa gttgggccga gctcggtagc gccagaatgg cccggaccgg	660
gttaccgaat tcaggcgacc catcgctgct ttgtctacat catgttcttc atcatcctcc	720
ccaggcgacg cgtgctgctg ttcttattca gactaccgtt cgagtgactg catggcgtag	780
atctttctgc atcgactttg tacggctaca tcgaacatat acacgagatg tctcgtgtga	840
atagagtcac taatgcctta agcatcgggt actccgtagg gtacattctg ttcttcttat	900
ttgtgcatat ttttattggt gtttactgat tatacgagta gttatacata catgcacata	960
catatcatca catatatcac aatatTTTTT taaattaaat taaaactaaa aatgactaaa	1020
tttctaacac caacgacatt gtaatgtttt ctccaacaac tttacctatt ctacattggt	1080
ctatttcgaa tttcactcta taaacaacat agtctacaat ggaaaacagt gctttgtacg	1140
actatatacg cgatgtgtgg ctacaacata agacaatata gtcgtttgaa gattgaacct	1200
atatatcggg acggttaatc cgtctatgta cgtgggcatg acgaacaccc gtgataacga	1260
aggattaacg tgcacaatca taaatccaaa gtaggagcgg tgcattgatga gaatcgctct	1320
cagtactcga cataatgaac cttacgaggt acaacaggca ggcaggcagg gaccaggggc	1380
cgcctttatt tcaggctcgc tggccccacg gccgtgctgc gtgcacgaag ggcactaccc	1440
caacctctca ccgaaaaccg cgctggatcg gcaaataaaa cgaggtggtg ccccggtccc	1500
actctccacg tccacggcac catccctctg cagccgctca ccagccatgc cgtgtcgcgg	1560
aacggcacaa ccaccccaa cccactcacg aaaccccgtc ccggccgtgc ccgtgtcggg	1620
ccgcgctcgg caacgaggcg gcccgcgctg ctgagtcgcc tggacacccg acaccctgtc	1680

ES 2 843 556 T3

ggccctttgt	ttattcatcc	cgaaatctca	tctgccccca	cgcccgactg	cgctgcgccg	1740
cccgatata	tatacccatc	gttatcgatc	gactggggac	tctatcagtg	atagagtcta	1800
gaggatcgct	caggaaggcc	gctgagatag	aggcatggcg	gccaatgcgg	gcggcggtgg	1860
agcgggagga	ggcagcggca	gcggcagcgt	ggctgcgccg	gcggtgtgcc	gccccagcgg	1920
ctcgcgggtg	acgcgcgacg	cggagcagat	caggatgctg	aaggagctct	actacggtg	1980
cggcatccgg	tcgcccagct	cggagcagat	ccagcgcatc	accgccatgc	tgccgcagca	2040
cggcaagatc	gagggcaaga	acgtcttcta	ctggttccag	aaccacaagg	ccgcgcgagc	2100
ccagaagcgc	cgcctcacca	gcctcgacgt	caacgtgccc	gccgccggcg	cggccgacgc	2160
caccaccagc	caactcggcg	tcctctcgct	gtcgtcgccg	ccgccttcag	gcgcggcgcc	2220
tccctcgccc	accctcggct	tctacgccgc	cggcaatggc	ggcggatcgg	ctgtgctgct	2280
ggacacgagt	tccgactggg	gcagcagcgg	cgctgccatg	gccaccgaga	catgcttctt	2340
gcaggactac	atgggcgtga	cggacacggg	cagctcgctg	cagtggccac	gcttctcgtc	2400
gtcggacacg	ataatggcgg	cggccgcggc	gcggcgggcg	acgacgcggg	cgcccagagc	2460
gtccctctc	ttcccagact	gcggcgacga	cggcggcagc	ggtagcagca	gctacttgcc	2520
gttctggggt	gccgcgtcca	caactgccgg	cgcacttctt	tccgttgcca	tccagcagca	2580
acaccagctg	caggagcagt	acagctttta	cagcaacagc	aacagcacc	agctggccgg	2640
caccggcaac	caagacgtat	cggcaacagc	agcagcagcc	gccgccctgg	agctgagcct	2700
cagctcatgg	tgctcccctt	accctgctgc	agggagtatg	tgagagcaac	gcgagctgcc	2760
actgctcttc	actggtaccg	ttaacagatc	attcgacaaa	gcagcattag	tccgttgatc	2820
ggtggaagac	cactogtcag	tgttgagttg	aatgtttgat	caataaaata	cggcaatgct	2880
gtaagggttg	ttttttatgc	cattgataat	acactgtact	gttcagttgt	tgaactctat	2940
ttcttagcca	tgccaagtgc	ttttcttatt	ttgaataaca	ttacagcaaa	aagttgaaag	3000
acaaaaaaaa	aaacccccga	acagagtgc	ttgggtccca	agcttcttta	gactgtgttc	3060
ggcgttcccc	ctaaatttct	ccccctatat	ctcactcact	tgtcacatca	gcgttctctt	3120
tccccctata	tctccacgct	ctacagcagt	tccacctata	tcaaacctct	ataccccacc	3180
acaacaatat	tatatacttt	catcttcaac	taactcatgt	accttccaat	ttttttctac	3240
taataattat	ttacgtgcac	agaaaacttag	caaggagaga	gagagcgggg	tgaccacact	3300
tgetagttgg	atattacctc	ttctcttcaa	agtatccttg	aacgctcacc	ggttatcaaa	3360
tctctacact	atagctctgt	agtcttgcta	gatagttagt	tctttagctc	tcggtgacca	3420
agcttggcgc	gatcaagctt	atcgataccg	tcgacctcga	agcttgggtca	cccggtcggg	3480
gcctagaagg	ccagcttcaa	gtttgtacaa	aaaagcaggc	tccggccaga	atggcccggg	3540
ccgggttacc	gaattcttac	cctagctccc	tgccgctggt	acgcggtccc	ccatcaatct	3600

ES 2 843 556 T3

tctgttcttg	cggttgtagc	ctgtgtaaca	gtgctagagt	atgtatgata	aataggtttt	3660
aagtctgctt	acatgacatt	ttttattgtg	gaagagacat	ataaaaatta	gagagagtgg	3720
ttctcatgca	acggcggacg	gcccggtgct	aaaagagctt	caagacaaaa	taatgaaaca	3780
ggaagagagt	agatttatct	aagagccaac	tttattatat	gaatgtgttt	attgttggct	3840
ttagatgata	tggtaaggag	ttagagctaa	taatagatag	gctctattat	tattattatt	3900
aattaaactc	gctctaagga	ggaaagtggg	aggaagggac	gaggacgaag	actactggaa	3960
gcacgtgca	tggatgatgg	atgtggtgtc	tcttaatgta	ggtggccgga	ggatgtacgt	4020
gttaattgcg	cgataagcac	tcagatccaa	ccgcaaacta	cctccacact	gacacactga	4080
tagagagaaa	gagagacctc	cgacgactgc	cgccgcagat	gagccacgta	cgtatacgac	4140
gtctgccggc	cggtctcaggc	tgcgcccatc	accctgctcg	aaagtgcgct	taggcggcgc	4200
cagctacata	ggagtatcta	gtctagccag	ttagtatact	actactgcgc	tgatgatgaa	4260
ttaactctgc	atagatactg	tacttgccctc	cctccaacac	ccaaccacct	cctgctcggc	4320
tcttaataac	ttggacacgg	atcgatgcca	tccaagggaag	aacacgacga	cgacgacgga	4380
acatccacca	tgcaagcttg	catccatacg	ccgatacgcg	tgcatccatc	catccaccat	4440
tatttccatt	ttccaccgat	cacacgtaca	caggcctatt	taaggagcga	catcccactg	4500
caactctcct	caccactcat	caccagctag	ctctagcaaa	gcacttgcca	tctaccgacc	4560
gccgcattcc	aaacagcccg	acgagctagc	agagcggcag	gcacctccct	cctcaaggaa	4620
cccatggcca	ctgtgaacaa	ctggctcgct	ttctccctct	ccccgcagga	gctgccgccc	4680
tcccagacga	cggactccac	actcatctcg	gccgccaccg	ccgaccatgt	ctccggcgat	4740
gtctgcttca	acatccccca	agattggagc	atgaggggat	cagagctttc	ggcgtctgtc	4800
gcggagccga	agctggagga	cttcctcggc	ggcatctcct	tctccgagca	gcatcacaag	4860
gccaactgca	acatgatacc	cagcactagc	agcacagttt	gctacgcgag	ctcaggtgct	4920
agcaccggct	accatcacca	gctgtaccac	cagcccacca	gctcagcgct	ccacttcgcg	4980
gactccgtaa	tgggtggcttc	ctcggccggt	gtccacgacg	gcggtgccat	gctcagcgcg	5040
gccgccgcta	acggtgtcgc	tggcgctgcc	agtgccaacg	gcggcggcag	cgggctgtcc	5100
atgattaaga	actggctgcg	gagccaaccg	gcgcccatgc	agccgagggt	ggcggcgggt	5160
gagggcgcgc	aggggctctc	tttgtccatg	aacatggcgg	ggacgaccca	aggcgtgct	5220
ggcatgccac	ttctcgctgg	agagcgcgca	cgggcgcccg	agagtgtatc	gacgtcagca	5280
cagggtggag	ccgtcgctcg	cacggcgccg	aaggaggata	gcggtggcag	cgggtgtgcc	5340
ggcgtcttag	tagccgtgag	cacggacacg	ggtggcagcg	gcggcgcgtc	ggctgacaac	5400
acggcaagga	agacggtgga	cacgttcggg	cagcgcacgt	cgatttaccg	tggcgtgaca	5460

ES 2 843 556 T3

aggcatagat	ggactgggag	atatgaggca	catctttggg	ataacagttg	cagaagggaa	5520
gggcaaactc	gtaagggtcg	tcaagtctat	ttaggtggct	atgataaaga	ggagaaagct	5580
gctagggcct	atgatcttgc	tgctctgaag	tactggggtg	ccacaacaac	aacaaatttt	5640
ccagttagta	actacgaaaa	ggagctcgag	gacatgaagc	acatgacaag	gcaggagttt	5700
gtagcgtctc	tgagaaggaa	gagcagtggt	ttctccagag	gtgcatccat	ttacagggga	5760
gtgactaggc	atcaccaaca	tggaagatgg	caagcacgga	ttggacgagt	tgacagggaac	5820
aaggatcttt	acttgggcac	cttcagcacc	caggaggagg	cagcggaggc	gtacgacatc	5880
gcggcgatca	agttccgcgg	cctcaacgcc	gtcaccaact	tcgacatgag	ccgctacgac	5940
gtgaagagca	tcctggacag	cagcgcctc	cccacgggca	gcgccgcca	gcgcctcaag	6000
gaggccgagg	ccgcagcgtc	cgcgcagcac	caccacgccg	gcgtggtgag	ctacgacgtc	6060
ggccgcacgc	cctgcagct	cggcgacggc	ggagccctgg	cggcgcgcta	cggcgcgcac	6120
taccacggcg	ccgcctggcc	gaccatcgcg	ttccagccgg	gcgccgccag	cacaggcctg	6180
taccacccgt	acgcgcagca	gccaatgcgc	ggcgccgggt	ggtgcaagca	ggagcaggac	6240
cacgcgggtga	tcgcggccgc	gcacagcctg	caggacctcc	accacctgaa	cctgggcgcg	6300
gccggcgcg	acgacttttt	ctcggcaggg	cagcaggccg	ccgccgctgc	gatgcacggc	6360
ctgggtagca	tcgacagtgc	gtcgtctgag	cacagcaccg	gctccaaactc	cgtcgtctac	6420
aacggcgggg	tcggcgacag	caacggcgcc	agcgccgtcg	gcggcagtg	cgggtggctac	6480
atgatgccga	tgagcgctgc	cggagcaacc	actacatcgg	caatggtgag	ccacgagcag	6540
gtgcatgcac	gggcctacga	cgaagccaag	caggctgctc	agatggggta	cgagagctac	6600
ctggtgaacg	cggagaacaa	tggtggcgga	aggatgtctg	catgggggac	tgctcgtgtct	6660
gcagccgcgg	cggcagcagc	aagcagcaac	gacaacatgg	ccgccgacgt	cgggcacggc	6720
ggcgcgacgc	tcttcagtgt	ctggaacgac	acttaagcgt	acctagtgg	acctgacatc	6780
ttatagtctg	caacctctcg	tgtctgaatt	cctatcttta	tcaagtgtta	ttgcttccac	6840
gactatagga	cagctttcgt	cgaagccttt	tgctcatgtg	atctcgaagg	attcatctag	6900
tctgattttt	cgtgacttgt	atcggtttta	ttggattcat	ccaacatata	tcaataaaaa	6960
atgagttgtg	tttcctttct	tcctagtcca	gttaaaatta	tttccctcct	gcgcttggtg	7020
tgtaattgtc	tgtgtacctg	ttgtttgtga	ctgtgttagt	tccttggtat	atgatttcgt	7080
atttgatatg	tacatggaga	tagcttagct	tcattattgg	agtatgaagt	tagtatgaca	7140
tagtcactct	cctggaaaat	tgacactgca	aaccatattt	ttattctgaa	ccacaaatcc	7200
tagtcagtcc	gctggcataat	gccgtccgtt	tgctgaatcc	agaacgtggg	tttgagagtg	7260
tacggctgag	atgcctctat	gcgaagggga	tttcgtggtg	aaacgagatg	ggagtagagc	7320
aacgcccgtg	gaagatgctt	caaaactcca	cacttttgag	caacgatcgg	cagtagtaag	7380

ES 2 843 556 T3

gtagacgatt tcaagatcaa agcatatgaa gataaacaac atcaacaaca aaattttgttg	7440
gggttctata gagagaaaca gagctacata catacactgt tttgtatcta ccatctgaga	7500
tgatgaaaag atgaaaaact aaagaatgcc ccggcgccaa cgccaggaca cgccgcgcgc	7560
gcgtcaccocg agccatctct tgacccagcc ggcgctgtat atttacacac gttgcagcat	7620
cgatcaccac ctgttcgacg gcgtcgccgt caccgggacg tgagtcgaaa cctagacttg	7680
tccatcttct ggattggcca acttaattaa tgtatgaaat aaaaggatgc acacatagtg	7740
acatgctaata cactataatg tgggcatcaa agttgtgtgt tatgtgtaat tactagttat	7800
ctgaataaaa gagaaagaga tcatccatat ttcttatcct aaatgaatgt cacgtgtctt	7860
tataattctt tgatgaacca gatgcatttc attaaccaaa tccatataca tataaatatt	7920
aatcatatat aattaatatc aattgggtta gcaaaacaaa tctagtctag gtgtgttttg	7980
cgaatgcgac cttcttatgt gcttctagtc tccaaatgtg gttgatagtt attttgctct	8040
aagatcaaca gtaatgaagt ataaatcatc gttgtggtgt gctactcggt taattgagca	8100
ttaacacaca caaacatgac gaggatggta taatctccaa aaatgtgtac tttgttaggt	8160
gggaccctat agccttgatt aatgtgctat gttaggcatg cctggaaacg tgtgacgcat	8220
atgtttttgtg aacctgttga tattatatgt gcttttatat taccatattt tattaaaata	8280
ctaataattt ttactagtaa gatataacat tctatctagc ttaaaaacta accataaata	8340
ttccataata actagattta ccaaactaat atactaaata tacataataa atacaaaatt	8400
aacaagacaa taatcaatat ttatgagctt aatatattta gacattatgg ttggtcgacg	8460
ataatcatgc taacttttcg taattgcttg attgaaatat gcttagaata atgcctcttt	8520
gttctacatg gcaaataggg accattatgg tgtaacaccc tgggaaccac aaacaccccg	8580
aaatgctact aaactacaca actaaccttc atataataaa ttctgacagc atctcctttg	8640
aaaatttgca tagacgtgga agcaacagag tataaacaga tatcatgata agaaaacata	8700
ctagacatta ataatctgct agaaatggga agaactcctaa cttgacgact gcgtaactga	8760
ctagagtcac acttagctga ccctagtcac ttacaactga ctctgtgtcc tagatcgatg	8820
ggccctggcc gaagcttggc caccgggtcc gggcctagaa ggccagcttc aagtttgtac	8880
aaaaaagcag gctccggcca gaatggcccg gaccgaagct ggccgctcta gaactagtgg	8940
atctcgatgt gtagtctacg agaagggtta accgtctctt cgtgagaata accgtggcct	9000
aaaaataagc cgatgaggat aaataaaatg tgggtgtaca gtacttcaag aggtttactc	9060
atcaagagga tgcttttccg atgagctcta gtagtacatc ggacctcaca tacctccatt	9120
gtggtgaaat attttgtgct catthtagtga tgggtaaatt ttgtttatgt cactctaggt	9180
tttgacattt cagttttgcc actcttaggt tttgacaaat aatttccatt ccgcggcaaa	9240

ES 2 843 556 T3

agcaaaacaa ttttatttta cttttaccac tcttagcttt cacaatgtat cacaaatgcc	9300
actctagaaa ttctgtttat gccacagaat gtgaaaaaaa acactcactt atttgaagcc	9360
aaggtgttca tggcatggaa atgtgacata aagtaacgtt cgtgtataag aaaaaattgt	9420
actcctcgta acaagagacg gaaacatcat gagacaatcg cgtttggaag gctttgcac	9480
acctttggat gatgcgcatg aatggagtcg tctgcttgct agccttcgcc taccgccac	9540
tgagtccggg cggcaactac catcggcgaa cgaccagct gacctctacc gaccggactt	9600
gaatgcgcta ccttcgtcag cgacgatggc cgcgtacgct ggcgacgtgc ccccgcatgc	9660
atggcggcac atggcgagct cagaccgtgc gtggctggct acaaatacgt accccgtgag	9720
tgccctagct agaaacttac acctgcaact gcgagagcga gcgtgtgagt gtagccgagt	9780
agatcccccg ggctgcaggt cgactctaga ggatccgaag gagatagaac cgatccacca	9840
tgtccaacct gctcacggtt caccagaacc ttccggtctt tccagtggac gcgacgtccg	9900
atgaagtcag gaagaacctc atggacatgt tccgcgacag gcaagcgttc agcgagcaca	9960
cctggaagat gctgctctcc gtctgccgct cctgggctgc atgggtcaag ctgaacaaca	10020
ggaagtgggt ccccgctgag cccgaggacg tgagggatta ccttctgtac ctgcaagcgc	10080
gaggtttgtt tctgcttcta cctttgatat atatataata attatcatta attagtagta	10140
atataatatt tcaaatattt ttttcaaat aaaagaatgt agtatatagc aattgctttt	10200
ctgtagttta taagtgtgta tattttaatt tataactttt ctaatatatg accaaaacat	10260
ggtgatgcct aggtctggca gtgaagacca tccagcaaca ccttggacaa ctgaacatgc	10320
ttcacaggcg ctccggcctc ccgcgccccg gcgactcgaa cgcctgagc ctcgctcatgc	10380
gccgcacatc gaaggaaaac gtcgatgccg gcgaaagggc aaagcaggcc ctgcgcttcg	10440
agaggaccga tttcgaccag gtccgcagcc tgatggagaa cagcgacagg tgccaggaca	10500
ttaggaacct ggcgttcctc ggaattgcat acaacacgct cctcaggatc gcggaaattg	10560
cccgcatctg cgtgaaggac attagccgca ccgacggcgg caggatgctt atccacattg	10620
gcaggaccaa gacgctcggt tccaccgcag gcgtcgaaaa ggccctcagc ctcgagtgga	10680
ccaagctcgt cgaacgctgg atctccgtgt ccggcgtcgc ggacgacca aacaactacc	10740
tcttctgccg cgtccgcaag aacgggggtg ctgcccctag cgccaccagc caactcagca	10800
cgagggcctt ggaagggtatt ttcgaggcca cccaccgcct gatctacggc gcgaaggatg	10860
acagcggcca acgctacctc gcatggctcc ggcaactccg ccgcgttgga gctgctaggg	10920
acatggcccc cgccggtgtt tccatccccg aaatcatgca ggcggttgga tggacgaacg	10980
tgaacattgt catgaactac attcgcaacc ttgacagcga gacgggcgca atggttcgcc	11040
tcctggaaga tggtgactga gctagacca gctttcttgt acaaagtgga cgtaaacgga	11100
tccagacttg tccatcttct ggattggcca acttaattaa tgtatgaaat aaaaggatgc	11160

ES 2 843 556 T3

acacatagtg acatgctaata cactataatg tgggcatcaa agttgtgtgt tatgtgtaat	11220
tactagtatt ctgaataaaa gagaaagaga tcatccatat ttcttatcct aatgaatgt	11280
cacgtgtcct tataattcct tgatgaacca gatgcatttc attaaccaaa tccatataca	11340
tataaatatt aatcatatat aattaatatc aattgggtta gcaaaacaaa tctagtctag	11400
gtgtgttttg cgaattgagg caagcttcgg ccgcccagc ttggtcaccc ggtccggggc	11460
tagaaggccg atctccggg caccagctt tctgtataa agtggccgtt aacggatcgg	11520
ccagaatggc ccggaccggg ttaccgaatt cctgccctta aggccaattg ttcaagattc	11580
attcaacaat tgaacatct cccatgatta aatcagtata aggttgctat ggtcttgctt	11640
gacaaagttt ttttttgagg gaatttcaac taaatttttg agtgaaacta tcaaatactg	11700
attttaaaaa atttttataa aaggaagcgc agagataaaa ggccatctat gctacaaaag	11760
tacccaaaaa tgtaatccta aagtatgaat tgcatttttt ttttgtttg acgaaaggaa	11820
aggagtatta ccacaagaat gatatcatct tcatatttag atcttttttg ggtaaagctt	11880
gagattctct aaatatagag aaatcagaag aaaaaaaac cgtgttttg tggttttgat	11940
ttctagcctc cacaataact ttgacggcgt cgacaagtct aacggacacc aagcagcgaa	12000
ccaccagcgc cgagccaagc gaagcagacg gccgagacgt tgacaccttc ggcgcggcat	12060
ctctcgagag ttccgctccg gcgtccacc tccaccgctg gcggtttctt attccgttcc	12120
gttccgcctc ctgctctgct cctctccaca ccacacggca cgaaaccgtt acggcaccgg	12180
cagcaccag caccggagag gggattcctt tcccaccgtt ccttccctt ccgccccgcc	12240
gctataaata gccagcccca tccccagctt ttttcccaa tctcatctcc tctctcctgt	12300
tgttcggagc acacgcacaa tccgatcgat ccccaaatcc ccttcgtctc tctcgcgag	12360
cctcgtggat ccagcttca aggtacggcg atcgatcct cccctcctt ctctctacct	12420
tcttttctct agactacatc ggatggcgat ccatggttag ggcctgctag tttcccttcc	12480
tgttttgtcg atggctgcga ggcacaatag atctgatggc gttatgacgg ctaacttgtc	12540
atgttggtgc gatttatagt cccttttaga gatcagttta atttctcgga tggttcgaga	12600
tcggtgggcc atggttagta ccctaagatc cgcgctgtta gggttcgtag atggaggcga	12660
cctgttctga ttgttaactt gtcagtacct gggaaatcct gggatggttc tagctcgtcc	12720
gcagatgaga tcgatttcat gatcctctgt atcttgtttc gttgcctagg ttccgtctaa	12780
tctatccgtg gtatgatgta gatgttttga tctgtctaac tacgtcttgt aaagttaatt	12840
gtcagggtcat aatttttagc atgccttttt ttttgtttg ttttgcctaa ttgggctgtc	12900
gttctagatc agagtagaag actgttccaa actacctgct ggatttattg aacttgatc	12960
tgtatgtgtg tcacatatct tcataaatc atgattaaga tggattgaaa tatcttttat	13020

ES 2 843 556 T3

cttttttgta	tggatagttc	tatatgttgg	tgtggctttg	ttagatgtat	acatgcttag	13080
atacatgaag	caacgtgctg	ctactgttta	gtaattgctg	ttcatttgtc	taataaacag	13140
ataaggatat	gtatttatgt	tgctgttgg	tttgcctgga	ctttgttgg	tacaaatgct	13200
tcaatacaga	aaacagcatg	ctgctacgat	ttaccattta	tctaatttta	tcatatgtct	13260
aatctaataa	acaaacatgc	ttttaaatta	tcttcatatg	cttgatgat	ggcatacaca	13320
gcggctatgt	gtggtttttt	aaataccag	catcatgggc	atgcatgaca	ctgctttaat	13380
atgcttttta	tttgcttgag	actgtttctt	ttgtttatac	tgaccttta	gttcggtgac	13440
tcttctgcag	gtcgacttta	acttagccta	ggatccatgg	cccagtcaa	gcacggcctg	13500
accaaggaga	tgaccatgaa	gtaccgcatg	gagggtgctg	tggacggcca	caagttcgtg	13560
atcacccggc	agggcatcgg	ctaccccttc	aaggccaagc	aggccatcaa	cctgtgcgtg	13620
gtggagggcg	gccccttgcc	cttcgccgag	gacatcttgt	ccgccgcctt	catgtacggc	13680
aaccgcgtgt	tcaccgagta	ccccaggac	atcgctgact	acttcaagaa	ctcctgcccc	13740
gccggctaca	cctgggaccg	ctccttcctg	ttcgaggacg	gcgccgtgtg	catctgcaac	13800
gccgacatca	ccgtgagcgt	ggaggagaac	tgcatgtacc	acgagtcaa	gttctacggc	13860
gtgaacttcc	ccgccgacgg	ccccgtgatg	aagaagatga	ccgacaactg	ggagccctcc	13920
tgcgagaaga	tcatccccgt	gccaagcag	ggcatcttga	agggcgacgt	gagcatgtac	13980
ctgctgctga	aggacggtgg	ccgcttgccg	tgccagttcg	acaccgtgta	caaggccaag	14040
tccgtgcccc	gcaagatgcc	cgactggcac	ttcatccagc	acaagctgac	ccgcgaggac	14100
cgcagcgacg	ccaagaacca	gaagtggcac	ctgaccgagc	acgccatcgc	ctccggtccc	14160
gccttgccct	ccggactcag	atctcgatag	ggtaccaatg	gccagttaac	agatccagct	14220
gctgctgttc	tagggttcac	aagtctgcct	atttgtcttc	cccaatggag	ctatggttgt	14280
ctggtctggt	ccttggtcgt	gtcccgtttc	attgtgtact	atttacctgt	aatgtgtatc	14340
cttaagtctg	gtttgatgg	gtctgaaacg	ttttgctgtg	gtagagcagc	atggaagaac	14400
tataatgaat	aagtgatccc	taatcattgt	gtccaaattt	tgcttctgct	ataccctttt	14460
gtgctgtttc	ttatgttttg	cttaaaaatt	tgatctgaca	aacaaatttg	tctaaattat	14520
gctcttgctc	tgactgtgta	ctgtacacat	ttgtattgct	accggtttgg	cacatagttt	14580
tctctttgat	tgcatatcaa	acccttctga	ttgcagattg	cttggtatg	gctgatgtgc	14640
tcaactgtct	tctttagctt	gtatattgct	gatggtggct	tggctacggc	tgaaaaagtt	14700
gctgctgcgc	cgacttggca	actgttgctc	cagtaatgct	ttgtcctctc	ttcctagccg	14760
ccgacttggc	acttggcagc	tgttgctcct	gtaatgcgtt	gttatctctt	atctcttgag	14820
atgtactatg	attttgaggc	tgtttgactt	gtgccagatt	gccagtgcgt	atgctatatt	14880
atctcaggtt	tttatgctag	tatatgtttt	actgtgtgct	gcctgtgtcg	agttacgtaa	14940

ES 2 843 556 T3

tctattgaac cgctatgcgt gttgatgccc tgttggtgct tgccagattg atgccccctgt	15000
tgcttgacgc ttgcaagctg tagtgaaatc ctaaaattcg aaaccaaatac tattctgaac	15060
ttttgaagat ctgtgatgtc cggaagaaaa gattcaaaca gataaacggg actctocaga	15120
atctataaaa taaatgatcg acaaacgctt tgctttggta ggcttcctag ttagttagcg	15180
gaccgaagct tcggtccggg cctagaaggc cagcttcggc cggcccgggc aactttatta	15240
tacaaagttg atagatcgaa taacttcgta tagcatacat tatacgaagt tatacctggg	15300
ggcgccgcta ggggctgcag gaattcgata tcaagcttat cgataccgtc gacctcgagg	15360
ggggggcccg taccctggga tctgccctta ctttaacgcc tctaaccaac acccctttat	15420
ctttataagg aacaataaac agaatttgcc ccactgttct aaatcaccta ataatatccc	15480
cagctaaaaa caataaagggt ttcctagaat taagacaagc atgactgttc ctccaggagg	15540
gtttggaaca ttgttgacgt cttgcagata cgggcgaagg gtgagaaaca gagcggaggg	15600
ctggagggtg cctcggtagt cgacgccgga gttgagcttg acaacgacgg ggcggcccct	15660
gatggacttg aggaagtccg atggcgctct caccgtcccg ccggcgcccg aggcgggcct	15720
gtcgctgccg ccgcccgcgc tgctcatctt gcgcgctgtg ccccgccggg tgtccctgtg	15780
ttgcggatcg cgggtggggc aggtggatgc gagggcgacc cgtttgact ccggccggag	15840
ccgcccgatc cctggtcggt gtcagtgcg tttactctgg gcccacgtg tcagtaccgt	15900
ctgtagatga caacaaccg tcgtccacag tcatgtccaa aatatccttt cttctttttt	15960
ttcgattcgg atatctatct tccttttttt tttccaaaa tcttcttgac gcaccagcgc	16020
gcacgtttgt ggtaaacgcc gacacgtcgg tcccacgtcg atagacccca cccaccagt	16080
agtagcgtgt acgtattcgg gggtagcga cgtgtcgccg tcgtcttgct agtcccattc	16140
ccatctgagc cacacatctc tgaacaaaaa aaaggaggga ggccctccacg cacatcccc	16200
tccgtgccac ccgccccaaa ccctcgcgcc gcctccgaga cagccgccgc aaccatggcc	16260
accgccgccg ccgctcttac cgcgctcact ggcgccacta ccgctgcgcc caaggcgagg	16320
cgccggggcg acctcctggc caccgcgccg gccctcgccg cgcccatcag gtgctcagcg	16380
gcgtcaccgc ccatgccgat ggctcccccg gccaccccg cccggccgtg gggccccacc	16440
gagccccgca aggggtgctga catcctcgtc gagtcctcg agcgctgcgg cgtccgcgac	16500
gtcttcgcct acccggcgg cgcgtccatg gagatccacc aggcactcac ccgctcccc	16560
gtcatcgcca accacctctt ccgccacgag caaggggagg cctttgccgc ctccggctac	16620
gcgcgctcct cgggcccgcgt cggcgtctgc atcgccacct ccggcccccg cgccaccaac	16680
ctagtctccg cgctcgccga cgcgctgctc gattccgtcc ccatggctgc catcacggga	16740
cagggtggcg gacgcatgat tggcaccgac gccttcagg agacgcccac cgtcgaggtc	16800

acccgctcca	tcaccaagca	caactacctg	gtcctcgacg	tcgacgacat	ccccgcgctc	16860
gtgcaggagg	ctttcttcct	cgcctcctct	ggtcgaccag	ggccggtgct	tgtcgacatc	16920
cccaaggaca	tccagcagca	gatggcgggt	cctgtctggg	acaagcccat	gagtctgcct	16980
gggtacattg	cgcgccttcc	caagccccct	gcgactgagt	tgcttgagca	ggtgctgcgt	17040
cttggtgggt	aatcgcgggc	ccctgttctt	tatgtggggc	gtggctgcgc	agcatctggt	17100
gaggagttgc	gacgctttgt	ggagctgact	ggaatcccgg	tcacaactac	tcttatgggc	17160
ctcggcaact	tccccagcga	cgaccactg	tctctgcgca	tgctaggtat	gcattgggacg	17220
gtgtatgcaa	attatgcagt	ggataaggcc	gatctgttgc	ttgcacttgg	tgtgcggttt	17280
gatgatcgcg	tgacagggaa	gattgaggct	tttgcaagca	gggctaagat	tgtgcacggt	17340
gatattgatc	cggctgagat	tggcaagaac	aagcagccac	atgtgtccat	ctgtgcagat	17400
gttaagcttg	ctttgcaggg	catgaatgct	cttcttgaag	gaagcacatc	aaagaagagc	17460
tttgactttg	gctcatggaa	cgatgagttg	gatcagcaga	agagggaatt	cccccttggg	17520
tataaaacat	ctaataagga	gatccagcca	caatatgcta	ttcaggttct	tgatgagctg	17580
acgaaaggcg	aggccatcat	cggcacaggt	gttgggcagc	accagatgtg	ggcggcacag	17640
tactacactt	acaagcggcc	aaggcagtgg	ttgtcttcag	ctgggtcttg	ggctatggga	17700
tttggtttgc	cggctgctgc	tgggtgcttct	gtggcaaacc	cagggtgtcac	tgttgttgac	17760
atcgatggag	atggtagctt	tctcatgaac	gttcaggagc	tagctatgat	ccgaattgag	17820
aacctcccag	tgaaggctct	tgtgctaaac	aaccagcacc	tggggatggg	ggtgcagttg	17880
gaggacaggt	tctataaggc	caacagagcg	cacacatact	tgggaaaccc	agagaatgaa	17940
agtgagatat	atccagattt	cgtgacgatc	gccaaagggt	tcaacattcc	agcggccggt	18000
gtgacaaaga	agaacgaagt	ccgcgcagcg	ataaagaaga	tgctcgagac	tccagggccg	18060
tacctcttgg	atataatcgt	cccacaccag	gagcatgtgt	tgccatgat	ccctagtggg	18120
ggggctttca	aggatatgat	cctggatggg	gatggcagga	ctgtgtactg	actagctagt	18180
cagttaacct	agacttgtcc	atcttctgga	ttggccaact	taattaatgt	atgaaataaa	18240
aggatgcaca	catagtgcac	tgctaatac	tataatgtgg	gcataaaagt	tgtgtgttat	18300
gtgtaattac	tagttatctg	aataaaagag	aaagagatca	tccatatttc	ttatcctaaa	18360
tgaatgtcac	gtgtctttat	aattctttga	tgaaccagat	gcatttcatt	aaccaaacc	18420
atatacatat	aatattaat	catatataat	taatatcaat	tgggttagca	aaacaaatct	18480
agtctaggtg	tgttttgcga	atgcggcccg	ctgcagagtt	aacggcgcgc	cgactagcta	18540
gctaaggtag	cgagctcgaa	ttcattccga	ttaatcgtgg	cctcttgctc	ttcaggatga	18600
agagctatgt	ttaaactgtc	aagcgtact	agacaattca	gtacattaaa	aacgtccgca	18660
atgtgttatt	aagttgtcta	agcgtcaatt	tgtttacacc	acaatatatc	ctgccca	18716
<210>	71					
<211>	13825					
<212>	ADN					
<213>	Secuencia artificial					
<220>						
<223>	Construcción sintética					
<400>	71					

ES 2 843 556 T3

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac	60
aatctgatca tgagcggaga attaaggag tgacgttatg acccccgcg atgacgcggg	120
acaagccggt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagcaag	180
ctggtacgat tgtaatacga ctactatag ggcgaattga gcgctgttta aacgtcttc	240
aactggaaga gcggttacgc tgtttaaacg ctcttcaact ggaagagcgg ttactaccgg	300
ttcactagct agctgctaata cgagctagtt accctatgag gtgacatgaa gcgctcacgg	360
ttactatgac ggtagcttc acgactgttg gtggcagtag cgtacgactt agctatagtt	420
ccggacttac ccttaagcga tttaaatcct gaggatatcg ctatcaactt tgtatagaaa	480
agttgggccc agctcggtag gcccagaatg gcccgaccg ggttaccgaa ttcaggcgac	540
ccatcgctgc tttgtctaca tcatgttctt catcatcctc ccaggcgac gcgtgctgct	600
gttcttattc agactaccgt tcgagtgaact gcatggcgta catctttctg catcgacttt	660
gtacggctac atogaacata tacacgagat gtctcgtgtg aatagagtca ctaatgcctt	720
aagcatcggg tactccgtag ggtacattct gttcttctta tttgtgcata tttttattgt	780
tgtttactga ttatacgagt agttatacat acatgcacat acatatcatc acatatatca	840
caatatTTTT ctaaatTaaa ttaaaactaa aaatgactaa atttctaaca ccaacgacat	900
tgtaatgttt tctccaacaa ctttacctat tctacattgt totatttoga atttcaactc	960
ataaacaaca tagtctacaa tggaaaacag tgctttgtac gactatatac gcgatgtgtg	1020
gctacaacat aagacaatat agtcgtttga agattgaacc tatatatcgg tacggttaat	1080
ccgtctatgt acgtgggcat gacgaacacc cgtgataacg aaggattaac gtgcacaatc	1140
ataaatccaa agtaggagcg gtgcatgatg agaatcgctc tcagtactcg acataatgaa	1200
ccttacgagg tacaacaggc aggcaggcag ggaccagggg ccgcctttat ttcaggctcg	1260
ctggccccac gggcggtgctg cgtgcacgaa gggcactacc ccaacctctc accgaaaacc	1320
gcgctggatc ggcaaatcaa acgaggtggt gcccogtgcc cactctccac gtccacggca	1380
ccatccctct gcagccgctc accagccatg ccgtgtcgcg gaacggcaca accaccccca	1440
acccactcac gaaaccccggt cccggccgtg cccgtgtcgcg tccgcgctcg gcaacgaggg	1500
ggcccgcgct gctgagtcct ctggacaccc gacaccctgt cggccctttg tttattcatc	1560

ES 2 843 556 T3

ccgaaatctc	atctgcccc	acggccgact	gcgctgcgcc	gcccggatat	atatacccat	1620
cgttatcgat	cgactgggga	ctctatcagt	gatagagtct	agaggatcgc	tcaggaaggc	1680
cgtgagata	gaggcatggc	ggccaatgcg	ggcggcgggtg	gagcgggagg	aggcagcggc	1740
agcggcagcg	tggtgcgcc	ggcgggtgtc	cgccccagcg	gctcgcgggtg	gacgccgacg	1800
ccggagcaga	tcaggatgct	gaaggagctc	tactacggct	gcggcatccg	gtcgcccagc	1860
tcggagcaga	tccagcgcat	caccgccatg	ctgcggcagc	acggcaagat	cgagggcaag	1920
aacgtcttct	actggttcca	gaaccacaag	gcccgcgagc	gccagaagcg	ccgcctcacc	1980
agcctcgacg	tcaacgtgcc	cgccgccggc	gcggccgacg	ccaccaccag	ccaactcggc	2040
gtcctctcgc	tgctgcgcc	gccgccttca	ggcgcggcgc	ctccctcgcc	caccctcggc	2100
ttctacgccg	ccggcaatgg	cggcgggatcg	gctgtgctgc	tggacacgag	ttccgactgg	2160
ggcagcagcg	gcgctgccat	ggccaccgag	acatgcttcc	tgcaggacta	catgggcgtg	2220
acggacacgg	gcagctcgtc	gcagtggcca	cgcttctcgt	cgtcggacac	gataatggcg	2280
gcggccgcgg	cgcgggcggc	gacgacgcgg	gcgcccagag	cgctccctct	cttcccgaac	2340
tgcggcgacg	acggcggcag	cggtagcagc	agctacttgc	cgttctgggg	tgccgcgtcc	2400
acaactgccg	gcgccacttc	ttccgttgcg	atccagcagc	aacaccagct	gcaggagcag	2460
tacagctttt	acagcaacag	caacagcacc	cagctggccg	gcaccggcaa	ccaagacgta	2520
tcggcaacag	cagcagcagc	cgccgccctg	gagctgagcc	tcagctcatg	gtgctcccct	2580
taccctgctg	cagggagtat	gtgagagcaa	cgcgagctgc	cactgctcct	cactggtacc	2640
gttaacagat	cattcgacaa	agcagcatta	gtccgttgat	cggtggaaga	ccactcgtca	2700
gtgttgagtt	gaatgtttga	tcaataaaat	acggcaatgc	tgtaagggtt	gttttttatg	2760
ccattgataa	tacactgtac	tgttcagttg	ttgaaactta	tttcttagcc	atgccaaagt	2820
cttttcttat	tttgaataac	attacagcaa	aaagttgaaa	gacaaaaaaa	aaaacccccg	2880
aacagagtgc	tttgggtccc	aagcttcttt	agactgtggt	cggcgttccc	cctaaatttc	2940
tccccctata	tctcactcac	ttgtcacatc	agcgttctct	ttccccctat	atctccacgc	3000
tctacagcag	ttccacctat	atcaaacctc	tataccccac	cacaacaata	ttatataactt	3060
tcatcttcaa	ctaactcatg	taccttccaa	tttttttcta	ctaataatta	tttacgtgca	3120
cagaaactta	gcaaggagag	agagagcggg	gtgaccacac	ttgctagttg	gatattacct	3180
cttctcttca	aagtatcctt	gaacgctcac	cggttatcaa	atctctacac	tatagctctg	3240
tagtcttgct	agatagttag	ttcttttagct	ctcggtgacc	aagcttggcg	cgatcaagct	3300
tatcgatacc	gtcgacctcg	aagcttggtc	acccggtccg	ggcctagaag	gccagcttca	3360
agtttgatca	aaaaagcagg	ctccggccag	aatggcccg	accgggttac	cgaattctta	3420
ccetagtctc	ctgcggctgt	tacgcgggtcc	cccatcaatc	ttctgttctt	gcggttgtag	3480

ES 2 843 556 T3

cctgtgtaac agtgctagag tatgtatgat aaataggttt taagtctgct tacatgacat	3540
tttttattgt ggaagagaca tataaaaatt agagagagtg gttctcatgc aacggcggac	3600
ggcccggtgc taaaagagct tcaagacaaa ataatgaaac aggaagagag tagatttatc	3660
taagagccaa ctttattata tgaatgtgtt tattgttggc tttagatgat atggtaagga	3720
gtagagcta ataatagata ggctctatta ttattattat taattaaact cgctctaagg	3780
aggaaagtgg gaggaaggga cgaggacgaa gactactgga agcatcgtgc atggatgatg	3840
gatgtggtgt ctcttaatgt aggtggccgg aggatgtacg tgtaattgc gcgataagca	3900
ctcagatcca accgcaaact acctccacac tgacacactg atagagagaa agagagacct	3960
ccgacgactg ccgccgcaga tgagccacgt acgtatacga cgtctgccgg ccggctcagg	4020
ctgccgccat caccctgctc gaaagtcgcg ttagcgggcg ccagctacat aggagtatct	4080
agtctagcca gtagtatac tactactgcg ctgatgatga attaactctg catagatact	4140
gtacttgcct ccctccaaca cccaaccacc tcctgctcgg ctcttaataa cttggacacg	4200
gatcgatgcc atccaaggaa gaacacgacg acgacgacgg aacatccacc atgcaagctt	4260
gcatacatac gccgatacgc gtgcatccat ccatccacca ttatttccat tttccaccga	4320
tcacacgtac acaggcctat ttaaggagcg acatcccact gcaactctcc tcaccactca	4380
tcaccagcta gctctagcaa agcacttgcc atctaccgac cgccgcatte caaacagccc	4440
gacgagctag cagagcggca ggcacctccc tcctcaagga acccatggcc actgtgaaca	4500
actggctcgc tttctccctc tccccgcagg agctgccgcc ctcccagacg acggactcca	4560
cactcatctc ggccgccacc gccgaccatg tctccggcga tgtctgcttc aacatcccc	4620
aagattggag catgagggga tcagagcttt cggcgctcgt cgcgagccg aagctggagg	4680
acttctcgg cggtcatctc ttctccgagc agcatcacia ggccaactgc aacatgatac	4740
ccagcactag cagcacagtt tgctacgcga gctcaggtgc tagcaccggc taccatcacc	4800
agctgtacca ccagcccacc agctcagcgc tccacttcgc ggactccgta atggtggctt	4860
cctcggccgg tgtccacgac ggcggtgcca tgctcagcgc ggccgccgct aacggtgtcg	4920
ctggcgctgc cagtgccaac ggcggcgga tcgggctgtc catgattaag aactggctgc	4980
ggagccaacc ggcgccatg cagccgaggg tggcgggggc tgagggcgcg caggggctct	5040
ctttgtccat gaacatggcg gggacgacct aaggcgctgc tggcatgcca cttctcgctg	5100
gagagcgcgc acgggcgccc gagagtgtat cgacgtcagc acagggtgga gccgtcgtcg	5160
tcacggcgcc gaaggaggat agcgggtggca gcggtgttgc cggcgctcta gtagccgtga	5220
gcacggacac ggggtggcagc ggcggcgcggt cggctgacaa cacggcaagg aagacgggtg	5280
acacgttcgg gcagcgcacg tcgatttacc gtggcgtgac aaggcataga tggactggga	5340

ES 2 843 556 T3

gatatgaggc acatcttttg gataacagtt gcagaaggga agggcaaaact cgtaagggtc	5400
gtcaagtcta tttaggtggc tatgataaag aggagaaagc tgctagggct tatgatcttg	5460
ctgctctgaa gtactggggt gccacaacaa caacaaattht tccagtgaht aactacgaaa	5520
aggagctcga ggacatgaag cacatgacaa ggcaggagtt tgtagcgtct ctgagaagga	5580
agagcagtggt tttctccaga ggtgcatcca tttacagggg agtgactagg catcaccaac	5640
atggaagatg gcaagcacgg attggacgag ttgcaggga caaggatctt tacttgggca	5700
ccttcagcac ccaggaggag gcagcggagg cgtacgacat cgcggcgatc aagttccgcg	5760
gcctcaacgc cgtcaccaac ttcgacatga gccgctacga cgtgaagagc atcctggaca	5820
gcagcgcctt ccccatcggc agcgcggcca agcgctcaa ggaggccgag gccgcagcgt	5880
ccgcgcagca ccaccacgcc ggcgtggtga gctacgacgt cggccgcac cgcctcgagc	5940
tcggcgacgg cggagccctg gcggcggcgt acggcgcgca ctaccacggc gccgcctggc	6000
cgaccatcgc gttccagccg ggcgcggcca gcacaggcct gtaccacccg tacgcgcagc	6060
agccaatgcg cggcggcggg tgggtcaagc aggagcagga ccacgcggtg atcgcggccg	6120
cgcacagcct gcaggaacct caccacctga acctgggcgc ggccggcgcg cacgacttht	6180
tctcggcagg gcagcaggcc gccgcgctg cgatgcacgg cctgggtagc atcgacagtg	6240
cgtcgctcga gcacagcacc ggctccaaact ccgtcgtcta caacggcggg gtccggcgaca	6300
gcaacggcgc cagcgcctg gccgcgagtg gcggtggtta catgatgccg atgagcgctg	6360
ccggagcaac cactacatcg gcaatggtga gccacgagca ggtgcatgca cgggcctacg	6420
acgaagccaa gcaggctgct cagatggggt acgagagcta cctggtgaac gcggagaaca	6480
atggtggcgg aaggatgtct gcatggggga ctgtcgtgtc tgcagccgcg gcggcagcag	6540
caagcagcaa cgacaacatg gccgcgacg tcgggcatgg cggcgcgag ctcttcagtg	6600
tctggaacga cacttaagcg tacctagtgg tacctgacat cttatagtct gcaacctctc	6660
gtgtctgaat tcctatcttht atcaagtgtt attgcttcca cgactatagg acagcttht	6720
tcgaaagctt ttgctcatgt gatctcgaag gattcatcta gtctgatttht tcgtgacttg	6780
tatcggttht attggattca tccaacatat atcaataaaa aatgagttgt gtttccttht	6840
ttcctagttht agtthaaatt atttccctcc tgcgcttggt ctgtaattgt ctgtgtacct	6900
gttgtthtgt actgtgttag ttccttgga tatgatttht tatttgatat gtacatggag	6960
atagcttagc ttcattattg gagtatgaag ttagtatgac atagtcactc tcctggaaaa	7020
ttgacactgc aaacctatt tttattctga accacaaatc ctagtacgtc cgctggcata	7080
tgccgtccgt ttgctgaatc cagaacgtgg gtttgagat gtacggctga gatgcctcta	7140
tgcgaagggg atttcgtggt gaaacgagat gggagtagag caacgcccg ggaagatgct	7200
tcaaaacttht acactthtga gcaacgatcg gcagtagtaa ggtagacgat ttcaagatca	7260

ES 2 843 556 T3

aagcatatga agataaacia catcaaciaa aaaatttgtt ggggttctat agagagaaac	7320
agagctacat acatacactg ttttgtatct accatctgag atgatgaaaa gatgaaaaac	7380
taaagaatgc cccggcgcca acgccaggac acgccgcgcg cgcgtcaccc gagccatctc	7440
ttgaccacgc cggcgctgta tatttacaca cgttcgagca tcgatcacca cctgttcgat	7500
cgcgtcgccg tcaccggtac cgagctcgaa ttccgggtccg ggcctagaag gccgatctcc	7560
cgggcaccca gctttcttgt acaaagtggc cgttaacgga tcggccagaa tggcccgga	7620
cgggttaccg aattcgagct cgggtaccctg ggatctgccc ttactttaac gcctctaacc	7680
aacacccctt tatctttata aggaacaata aacagaattt gcccactgt tctaaatcac	7740
ctaataatat cccagctaa aaacaataaa ggtttcctag aattaagaca agcatgactg	7800
ttcctccagg agggtttga acattgttgc agtcttgag atacgggcga agggtgagaa	7860
acagagcgga gggctggagg tgacctcgt agtcgacgcc ggagttgagc ttgacaacga	7920
cggggcgccc cctgatggac ttgaggaagt ccgatggcgt cttcacccgc ccgccggcgc	7980
ccgaggcggg cctgtcgtg ccgcccgccg cgtcgtcat cttgcgcgt gtgcccccg	8040
cgggtgccct gtgttgcgga tcgcgggtgg gccaggtgga tgcgaggcg acccgtttg	8100
actccggccg gagccgcgg atccctggtc ggtgtcagtg ccgtttactc tgggccccac	8160
gtgtcagtag cgtctgtaga tgacaacia ccgtcgtcca cagtcagtgc caaaatatcc	8220
tttcttctt ttttctgatt cggatatcta tcttcctttt ttttttcaa aaatcttctt	8280
gacgcaccag cgcgcacgtt tgtggtaaac gccgacacgt cgggtccacg tcgatagacc	8340
ccacccacca gtgagtagcg tgtacgtatt cgggggtgac ggacgtgtcg ccgtcgtctt	8400
gctagtccca ttcccatctg agccacacat ctctgaacia aaaaaaggag ggaggcctcc	8460
acgcacatcc ccctccgtgc caccgcgcc aaaccctcgc gccgcctccg agacagccgc	8520
cgcaaccatg gccaccgcg ccgccgcgtc taccgcgtc actggcgcca ctaccgctgc	8580
gccaaggcg aggcgcggg cgcacctct ggccacccgc cgcgcctcg ccgcgccat	8640
caggtgctca gcggcgctac ccgccatgcc gatggctccc ccggccaccc cgtccggcc	8700
gtggggcccc accgagcccc gcaagggtgc tgacatctc gtogagtccc tcgagcgctg	8760
cggcgtccgc gacgtcttcg cctaccccg cggcgcgtcc atggagatcc accaggcact	8820
caccgcgtcc ccgctcatcg ccaaccacct cttccgccac gagcaaggag aggcctttgc	8880
cgcctccgc tacgcgcgt cctcggggcg cgtcggcgtc tgcacgcca cctccggccc	8940
cggcgccacc aacctagtct ccgcgctcgc cgacgcgtg ctcgattccg tccccatggt	9000
cgccatcacg ggacaggtgg cgcgacgcat gattggcacc gacgccttc aggagacgcc	9060
catcgtcgag gtcaccgcgt ccatcaccaa gcacaactac ctggtcctcg acgtcgacga	9120

ES 2 843 556 T3

catccccgc gtcgtgcagg aggttttctt cctcgcctcc tctggtcgac cagggccggt	9180
gcttgtcgac atccccaagg acatccagca gcagatggcg gtgcctgtct gggacaagcc	9240
catgagtctg cctgggtaca ttgcgcgcct tcccaagccc cctgcgactg agttgcttga	9300
gcagggtgctg cgtcttgttg gtgaatcgcg gcgcctgtt ctttatgtgg gcgggtggctg	9360
cgcagcatct ggtgaggagt tgcgacgctt tgtggagctg actggaatcc cggtcacaac	9420
tactcttatg ggctcggca acttccccag cgacgacca ctgtctctgc gcatgctagg	9480
tatgcatggg acggtgtatg caaattatgc agtggataag gccgatctgt tgcttgcact	9540
tggtgtgcgg tttgatgac gcgtgacagg gaagattgag gcttttgcaa gcagggctaa	9600
gattgtgcac gttgatattg atccggctga gattggcaag aacaagcagc cacatgtgtc	9660
catctgtgca gatgttaagc ttgctttgca gggcatgaat gctcttcttg aaggaagcac	9720
atcaaagaag agctttgact ttggctcatg gaacgatgag ttggatcagc agaagaggga	9780
attccccctt ggtataaaaa catctaata ggagatccag ccacaatatg ctattcaggt	9840
tcttgatgag ctgacgaaag gcgaggccat catcggcaca ggtgttgggc agcaccagat	9900
gtgggcgcca cagtactaca cttacaagcg gccaaaggcag tggttgtctt cagctgtctt	9960
tggggctatg ggatttgggt tgccggctgc tgcgtgtgct tctgtggcaa acccaggtgt	10020
cactgttgtt gacatcgatg gagatggtag ctttctcatg aacgttcagg agctagctat	10080
gatccgaatt gagaacctcc cagtgaaggc ctttgtgcta aacaaccagc acctggggat	10140
ggtggtgcag ttggaggaca ggttctataa ggccaacaga gcgcacacat acttgggaaa	10200
cccagagaat gaaagtgaga tataatccaga tttcgtgacg atcgccaaag ggttcaacat	10260
tccagcggtc cgtgtgacaa agaagaacga agtccgcgca gcgataaaga agatgctcga	10320
gactccaggg ccgtacctct tggatataat cgtcccacac caggagcatg tgttgcctat	10380
gatccctagt ggtggggctt tcaaggatat gatcctggat ggtgatggca ggactgtgta	10440
ctgactagct agtcagttaa cagatctgcc agatcctcgg tgtacaaata acccgtctta	10500
tcctatgaga cgggcggcg tcagtgtgtt ctggaggaat ttttatgtca gagccttttt	10560
ctttgtgtgc ttgatgtaga tgccaaggga agcttattgg ctggtgaagc ttgatgcaaa	10620
ataaattatg gaactctgtt ttttgtttat ctaataataa ctagcaaata tgcttccatt	10680
gcattgaaac taacagcctt ttgtgtttcc aagttttatt ttgtgacaat gtcacttatt	10740
tcaattagtt gtggaatcgg aaacttgca gactaacttg gaaactccaa tccctcagca	10800
tcctggactt tttcctggtg taatccatgt agataattatt ttaatcatca ttttagttct	10860
ggagggtttt ccatctccgg ttttgcctcc ctttcttcaa aaaaaaaaaa aaaatgccgt	10920
aggcgcgcca acgcccacct gttgttcaaa ctcatgggca cgagtggctc gaagatttta	10980
tacaacaatt gctgtagttt caccgttgct ggtgaagaag cattttttta aaaaaatata	11040

ES 2 843 556 T3

gtggtattca ttttaattag tttagttgtg cagcgagcaa aatttgaca tgctctgctc	11100
ggaatctgat cgacctagac acagattagc agcagtagct ttgtcatctg ttccaagagt	11160
tgcgatctga tagaagaaaa aaaaaccttc ccttcaatgt aaaaccgaaa ctaacaagaa	11220
agaagcacag tgccgttttag gcaagtatgg agtacgtatt aagcatgtag aaggccatgc	11280
atgaaccact aacaagaaag aagcgagtg ccattcaggc aagcataagc atgtagattg	11340
gcatgcatga acaactaaca gtagatcgtc tctggtctga ttagaagttt tttgggaagc	11400
caagaaatca tgtacaactg gttccatctc aaattccgtg gcaaaaagag gcctaagcaa	11460
cataaagctt cgggccgggc ctagaaggcc attgggtcat cggatcccgg gcaactttat	11520
tatacaaagt tgatagatat ctggtctaac taactagtcc taaggaccgc gcggaccgaa	11580
gctggccgct ctagaactag tggatctcga tgtgtagtct acgagaaggg ttaaccgtct	11640
cttcgtgaga ataaccgtgg cctaaaaata agccgatgag gataaataaa atgtggtggt	11700
acagtacttc aagaggttta ctcatcaaga ggatgctttt ccgatgagct ctagtagtac	11760
atcggacctc acatacctcc attgtggtga aatattttgt gctcatttag tgatgggtaa	11820
attttgttta tgtcactcta ggttttgaca tttcagtttt gccactctta ggttttgaca	11880
aataatttcc attccgcggc aaaagcaaaa caattttatt ttacttttac cactcttagc	11940
tttcacaatg tatcacaat gccactctag aaattctgtt tatgccacag aatgtgaaaa	12000
aaaacactca cttatttgaa gccaaagggt tcatggcatg gaaatgtgac ataaagtaac	12060
gttcgtgtat aagaaaaaat tgtactcctc gtaacaagag acggaaacat catgagacaa	12120
tcgcgttttg aaggttttgc atcacctttg gatgatgcgc atgaatggag tcgtctgctt	12180
gctagccttc gcctaccgcc cactgagtcc gggcggaac taccatcggc gaacgacca	12240
gctgacctct accgaccgga cttgaatgcg ctacctctgt cagcgacgat ggccgcgtac	12300
gctggcgacg tgcccccgca tgcattggcg cacatggcga gctcagaccg tgcgtggctg	12360
gctacaaata cgtaccccgat gaggcccta gctagaaact tacacctgca actgcgagag	12420
cgagcgtgtg agtgtagccg agtagatccc ccgggctgca ggtcgactct agaggatcca	12480
ccggtcgcca ccatggccca cagcaagcac ggcctgaagg aggagatgac catgaagtac	12540
cacatggagg gctgcgtgaa cggccacaag ttcgtgatca ccggcgaggg catcggctac	12600
cccttcaagg gcaagcagac catcaacctg tgcgtgatcg agggcgccc cctgcccttc	12660
agcgaggaca tcctgagcgc cggttcaag tacggcgacc ggatcttcac cgagtacccc	12720
caggacatcg tggactactt caagaacagc tgccccgccg gctacacctg gggccggagc	12780
ttcctgttcg aggacggcgc cgtgtgcatc tgtaacctgg acatcaccgt gagcgtgaag	12840
gagaactgca tctaccacaa gagcatcttc aacggcgtga acttccccgc cgacggcccc	12900

ES 2 843 556 T3

gtgatgaaga agatgaccac caactgggag gccagctgcg agaagatcat gcccgtagcct 12960
aagcagggca tcctgaaggc cgacgtgagc atgtacctgc tgctgaagga cggcgccggc 13020
taccgggtgcc agttcgacac cgtgtacaag gccaaagagcg tgcccagcaa gatgcccagag 13080
tggcacttca tccagcacia gctgtgctgg gaggaccgga ggcagcacia gaaccagaag 13140
tggcagctga ccgagcacgc catcgccctc cccagcgccc tggcctgaag cggcccatgg 13200
atattcgaac gcgtaggtac cacatgggta acctagactt gtccatcttc tggattggcc 13260
aacttaatta atgtatgaaa taaaaggatg cacacatagt gacatgctaa tcactataat 13320
gtgggcatca aagttgtgtg ttatgtgtaa ttactagtta tctgaataaa agagaaagag 13380
atcatccata tttcttatcc taaatgaatg tcacgtgtct ttataattct ttgatgaacc 13440
agatgcattt cattaaccaa atccatatac atataaatat taatcatata taattaatat 13500
caattgggtt agcaaaacaa atctagtcta ggtgtgtttt gcgaatgcgg ccgccaccgc 13560
ggtggagctc gaattccggt ccgaagctta agccatggcc cgggaatctt agcggccgcc 13620
tgacaggtta acggcgcgcc gactagctag ctaaggtagc gagctcgaat tcattccgat 13680
taatcgtggc ctcttgctct tcaggatgaa gagctatgtt taaacgtgca agcgtacta 13740
gacaattcag tacattaaaa acgtccgcaa tgtgttatta agttgtctaa gcgtcaattt 13800
gtttacacca caatatatcc tgcca 13825

<210> 72

<211> 20570

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 72

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac 60
aatctgatca tgagcggaga attaaaggag tcacgttatg acccccgcgc atgacgcggg 120
acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagcaag 180
ctggtacgat tgtaatacga ctactatag ggcgaattga gcgctgttta aacgctcttc 240
aactggaaga gcggttacgc tgtttaaacg ctcttcaact ggaagagcgg ttactaccgg 300
ttcactagct agctgctaata cgagctagtt accctatgag gtgacatgaa gcgctcacgg 360
ttactatgac ggtagcttc acgactgttg gtggcagtag cgtacgactt agctatagtt 420
ccggacttac ccttaagata acttcgtata gcatacatta tacgaagtta tgggcccacg 480
cgtgaagagc ggtagaccga gctggtcacc tttgtccacc aagatggaac tgtcttcggc 540
cagaatggcc tccggaattt atcgctatca actttgtata gaaaagtgg gccgaattcg 600
agctcggtag ggccagaatg gcccgaccg ggttaccgaa ttcgagctcg gtaccctggg 660

10

ES 2 843 556 T3

atccggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc	720
gattaagttg ggtaacgccg gggttttccc agtcacgacg ttgtaaaacg acggccagtg	780
ccaagctcag atcagcttgc atgcctgcag tgcagcgtga cccggtcgtg cccctctcta	840
gagataatga gcattgcatg tctaagttat aaaaaattac cacatatattt ttttgtcaca	900
cttgtttgaa gtgcagttta tctatcttta tacatatatt taaactttac tctacgaata	960
atataatcta tagtactaca ataatatcag tgttttagag aatcatataa atgaacagtt	1020
agacatggtc taaaggacaa ttgagtattt tgacaacagg actctacagt tttatctttt	1080
tagtgtgcat gtgttctcct ttttttttgc aaatagcttc acctatataa tacttcatcc	1140
attttattag tacatccatt tagggtttag ggttaatggt ttttatagac taattttttt	1200
agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaaac tctattttag	1260
tttttttatt taataattta gatataaaat agaataaaat aaagtgacta aaaattaaac	1320
aaataccctt taagaaatta aaaaaactaa ggaaacattt ttcttgtttc gagtagataa	1380
tgccagcctg ttaaacgccg tcgacgagtc taacggacac caaccagcga accagcagcg	1440
tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg tcgctgcctc tggacccctc	1500
tcgagagttc cgctccaccg ttggacttgc tccgctgtcg gcatccagaa attgcgtggc	1560
ggagcggcag acgtgagccg gcacggcagg cggcctcctc ctctctcac ggcacgggca	1620
gctacggggg attcctttcc cacgcctcct tcgctttccc ttctctgccc gccgtaataa	1680
atagacaccc cctccacacc ctctttcccc aacctcgtgt tgttcggagc gcacacacac	1740
acaaccagat ctccccaaaa tccaccctgc ggcacctccg cttcaaggta cgcctcgt	1800
cctccccccc cccctctct accttctcta gatcggcggt ccggtccatg catggttagg	1860
gcccggtagt tctactctg ttcattgttg tgtagatcc gtgtttgtgt tagatccgtg	1920
ctgctagcgt tcgtacacgg atgcgacctg tacgtcagac acgttctgat tgctaacttg	1980
ccagtgtttc tctttgggga atcctgggat ggctctagcc gttccgcaga cgggatcgat	2040
ttcatgattt ttttgtttc gttgcatagg gtttggttg cccttttcct ttatttcaat	2100
atatgccgtg cacttgtttg tcgggtcatc ttttcatgct tttttttgtc ttggttgtga	2160
tgatgtggtc tggttgggcg gtcgttctag atcggagtag aattctgttt caaactacct	2220
ggtggattta ttaattttgg atctgtatgt gtgtgccata catattcata gttacgaatt	2280
gaagatgatg gatggaata tcgatctagg ataggtatac atgttgatgc gggttttact	2340
gatgcatata cagagatgct ttttgttcgc ttggttgtga tgatgtggtg tggttgggcg	2400
gtcgttcatt cgttctagat cggagtagaa tactgtttca aactacctgg tgtatttatt	2460
aattttggaa ctgtatgtgt gtgtcataca tcttcatagt tacgagttta agatggatgg	2520

ES 2 843 556 T3

aaatatcgat ctaggatagg tatacatggt gatgtgggtt ttactgatgc atatacatga	2580
tggcatatgc agcatctatt catatgctct aaccttgagt acctatctat tataataaac	2640
aagtatgttt tataattatt ttgatcttga tatacttggga tgatggcata tgcagcagct	2700
atatgtggat ttttttagcc ctgccttcat acgctattta tttgcttggg actgtttcct	2760
ttgtcgatgc tcaccctggt gtttgggtgtt acttctgcag gtcgactcta gaggatccac	2820
cgttcgccac catggcaccg aagaagaagc gcaaggatgat ggacaagaag tacagcatcg	2880
gcctcgacat cggcaccac tcggtgggct gggccgtcat caccgacgaa tataaggtcc	2940
cgtcgaagaa gttcaaggtc ctccgcaata cagaccgcca cagcatcaag aaaaacttga	3000
tcggcgccct cctgttcgat agcggcgaga ccgcgagggc gaccaggctc aagaggaccg	3060
ccaggagacg gtacactagg cgcaagaaca ggatctgcta cctgcaggag atcttcagca	3120
acgagatggc gaaggtggac gactccttct tccaccgcct ggaggaaatca ttcctggtgg	3180
aggaggacaa gaagcatgag cggcacccaa tcttcggcaa catcgtcgac gaggtaagtt	3240
tctgcttcta cctttgatat atatataata attatcatta attagtagta atataatatt	3300
tcaaataattt ttttcaaat aaaagaatgt agtatatagc aattgctttt ctgtagtta	3360
taagtgtgta tattttaatt tataactttt ctaatatatg accaaaacat ggtgatgtgc	3420
aggtggccta ccacgagaag tacccgacaa tctaccacct ccggaagaaa ctggtggaca	3480
gcacagacaa ggcggacctc cggctcatct accttgccct cgcgcatatg atcaagttcc	3540
gcggccactt cctcatcgag ggcgacctga acccgacaa ctccgacgtg gacaagctgt	3600
tcatccagct cgtcgacagc tacaatcaac tggttcgagga gaacccata aacgctagcg	3660
gcgtggacgc caaggccatc ctctcggcca ggctctcgaa atcaagaagg ctggagaacc	3720
ttatcgcgca gttgccaggc gaaaagaaga acggcctctt cggcaacctt attgcgctca	3780
gcctcggcct gacgccgaac ttcaaatcaa acttcgacct cgcggaggac gccaaactcc	3840
agctctcaaa ggacacctac gacgacgacc tcgacaacct cctggcccag ataggagacc	3900
agtacgcgga cctcttcctc gccgccaaaga acctctccga cgctatcctg ctacgcgaca	3960
tccttcgggt caacaccgaa attaccaagg caccgctgtc cgccagcatg attaaacgct	4020
acgacgagca ccatcaggac ctacgctgc tcaaggcact cgtccgccag cagctccccg	4080
agaagtacaa ggagatcttc ttcgaccaat caaaaaacgg ctacgcggga tatatcgacg	4140
gcggtgccag ccaggaagag ttctacaagt tcatcaaacc aatcctggag aagatggacg	4200
gcaccgagga gttgctggtc aagctcaaca gggaggacct cctcaggaag cagaggacct	4260
tcgacaacgg ctccatcccg catcagatcc acctgggcga actgcatgcc atcctgcggc	4320
gccaggagga cttctaccgg ttctgaagg ataaccggga gaagatcgag aagatcttga	4380
cgttcgcat ccatactac gtgggccgcg tggctcgcgg caactccggg ttgcctgga	4440

ES 2 843 556 T3

tgacccggaa gtcggaggag accatcacac cctggaactt tgaggaggtg gtcgataaag	4500
gcgctagcgc tcagagcttc atcgagcgca tgaccaactt cgataaaaac ctgccaatg	4560
aaaaagtcct ccccaagcac tcgctgctct acgagtactt caccgtgtac aacgagctca	4620
ccaaggtcaa atacgtcacc gagggcatgc ggaagccggc gttcctgagc ggcgagcaga	4680
agaaggcgat agtggacctc ctcttcaaga ccaacaggaa ggtgaccgtg aagcaattaa	4740
aagaggacta cttcaagaaa atagagtgtc tcgactccgt ggagatctcg ggcgtggagg	4800
atcggttcaa cgcctcactc ggcacgtatc acgacctcct caagatcatt aaagacaagg	4860
acttctcga caacgaggag aacgaggaca tcctcgagga catcgtctc accctgaccc	4920
tgttcgagga ccgcgaaatg atcgaggaga ggctgaagac ctacgcgcac ctgttcgacg	4980
acaaggtcat gaaacagctc aagagggcgc gctacactgg ttggggaagg ctgtcccgca	5040
agctcattaa tggcatcagg gacaagcaga gcggcaagac catcctggac ttctcaagt	5100
ccgacgggtt cgccaaccgc aacttcatgc agctcattca cgacgactcg ctcacgttca	5160
aggaagacat ccagaaggca cagggtgagcg ggcaggggtga ctccctccac gaacacatcg	5220
ccaacctggc cggctcgccg gccattaaaa agggcatcct gcagacggtc aaggtcgtcg	5280
acgagctcgt gaaggtgatg ggccggcaca agcccgaaaa tatcgtcata gagatggcca	5340
gggagaacca gaccaccaa aaagggcaga agaactcgcg cgagcggatg aaacggatcg	5400
aggagggcat taaagagctc ggggtccaga tcctgaagga gcaccccggtg gaaaataccc	5460
agctccagaa tgaaaagctc tacctctact acctgcagaa cggcccgac atgtacgtgg	5520
accaggagct ggacattaat cggtatcgg actacgacgt cgaccacatc gtgccgcagt	5580
cgttcctcaa ggacgatagc atcgacaaca aggtgctcac ccggtcggat aaaaatcggg	5640
gcaagagcga caacgtgcc agcgaggagg tcgtgaagaa gatgaaaaac tactggcgcc	5700
agctcctcaa cgcgaaactg atcaccagc gcaagttcga caacctgacg aaggcggaac	5760
gcggtggctt gagcgaactc gataaggcgg gcttcataaa aaggcagctg gtcgagacgc	5820
gccagatcac gaagcatgtc gccagatcc tggacagccg catgaatact aagtacgatg	5880
aaaacgacaa gctgatccgg gaggtgaagg tgatcacgct gaagtccaag ctcggtcgg	5940
acttccgcaa ggacttccag ttctacaagg tccgcgagat caacaactac caccacgcc	6000
acgacgccta cctgaatcg gtggtcggga ccgccctgat caagaagtac ccgaagctgg	6060
agtcggagtt cgtgtacggc gactacaagg tctacgacgt gcgcaaatg atcgccaagt	6120
ccgagcagga gatcggcaag gccacggcaa aatacttctt ctactcgaac atcatgaact	6180
tcttcaagac cgagatcacc ctgcgaacg gcgagatccg caagcgcccg ctcatcgaaa	6240
ccaacggcga gacgggcgag atcgtctggg ataagggccg ggatttcgcg acggtccgca	6300

ES 2 843 556 T3

aggtgctctc catgccgcaa gtcaatatcg tgaaaaagac ggaggtccag acgggcgggt	6360
tcagcaagga gtccatcctc ccgaagcgca actccgacaa gctcatcgcg aggaagaagg	6420
attgggaccc gaaaaaatat ggcggcttcg acagcccgac cgtcgcatac agcgtcctcg	6480
tcgtggcgaa ggtggagaag ggcaagtcaa agaagctcaa gtccgtgaag gagctgctcg	6540
ggatcacgat tatggagcgg tcctccttcg agaagaacct gatcgacttc ctagaggcca	6600
agggatataa ggaggtcaag aaggacctga ttattaaact gccgaagtac tcgctcttcg	6660
agctggaaaa cggccgcaag aggatgctcg cctcccgagg cgagttgcag aagggcacg	6720
agctcgccct cccgagcaaa tacgtcaatt tcctgtacct cgctagccac tatgaaaagc	6780
tcaagggcag cccggaggac aacgagcaga agcagctctt cgtggagcag cacaagcatt	6840
acctggacga gatcatcgag cagatcagcg agttctcgaa gcgggtgac ctcgccgacg	6900
cgaacctgga caaggtgctg tcggcatata acaagcaccg cgacaaacca atacgcgagc	6960
aggccgaaaa tatcatccac ctcttcaccc tcaccaacct cggcgctccg gcagccttca	7020
agtacttcga caccacgatt gaccggaagc ggtacacgag cacgaaggag gtgctcgatg	7080
cgacgctgat ccaccagagc atcacagggc tctatgaaac acgcacgac ctgagccagc	7140
tgggcgga caagagacca cgggaccgcc acgatggcga gctgggaggc cgcaagcggg	7200
caaggtaggt accgttaacc tagacttgct catctctctg attggccaac ttaattaatg	7260
tatgaaataa aaggatgcac acatagtgc atgctaata ctataatgtg ggcataaag	7320
ttgtgtgtta tgtgtaatta ctagtatat gaataaaaga gaaagagatc atccatattt	7380
cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga tgcatttcat	7440
taaccaaata catatacata taaatattaa tcatatataa ttaatatcaa ttgggttagc	7500
aaaacaaata tagtctaggt gtgttttgcg aattcgtaat catgtcatag ctgtttcctg	7560
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgt	7620
aagcctgggg tgccataatga gtgagctaac tcacattaat tgcgttgccg tcaactgccg	7680
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga	7740
gagggcggtt gcgtattggg cgctcttcgg ctgatccgat atcgatgggc cctggccgaa	7800
gcttggtcac ccggtccggg cctagaaggc cagcttcaag tttgtacaaa aaagcaggct	7860
ccggccagaa tggcccgac cgggttaccg aattcgagct cggtagccct ggatccgata	7920
tcgatgggcc ctggccgaag ctttgagagt acaatgatga acctagatta atcaatgcca	7980
aagtctgaaa aatgcaccct cagtctatga tccagaaaat caagattgct tgaggccctg	8040
ttcggttgtt ccgattaga gccccgatt aattcctagc cggattactt ctctaattta	8100
tatagatttt gatgagctgg aatgaatcct ggcttattcc ggtacaaccg aacaggccct	8160
gaaggatacc agtaatcgct gagctaaatt ggcattgctgt cagagtgtca gtattgcagc	8220

ES 2 843 556 T3

aaggtagtga gataaccggc atcatggtgc cagtttgatg gcaccattag ggtagagat	8280
ggtggccatg ggcgcagtgc ctggccaact ttgtatgata tatggcaggg tgaataggaa	8340
agtaaaattg tattgtaaaa agggatttct tctgtttgtt agcgcatgta caaggaatgc	8400
aagttttgag cgagggggca tcaaagatct ggctgtgttt ccagctgttt ttgttagccc	8460
catcgaatcc ttgacataat gatcccgtt aaataagcaa cctcgcttgt atagtccctt	8520
gtgctctaac acacgatgat gataagtcgt aaaatagtgg tgtccaaaga atttccaggc	8580
ccagttgtaa aagctaaaat gctattcgaa tttctactag cagtaagtcg tgtttagaaa	8640
ttatTTTTTT atataccttt ttctcttcta tgtacagtag gacacagtgt cagcgccg	8700
ttgacggaga atatttgcaa aaaagtaaaa gagaaagtca tagcggcgta tgtgccaaaa	8760
acttcgtcac agagagggcc ataagaaaca tggcccacgg cccaatacga agcaccgca	8820
cgaagcccaa acagcagtcg gtaggtggag caaagcgctg ggtaatacgc aaacgtttt	8880
tcccaccttg actaatcaca agagtggagc gtaccttata aaccgagccg caagcaccga	8940
attgcatcat atctatccgt agccgtttta gagctagaaa tagcaagtta aaataaggct	9000
agtccgttat caacttgaaa aagtggcacc gagtcggtgc tttttttttg cggccgcgaa	9060
ttcctgcagg gccctcttgt cggaccagt tcccaccacg ttggtgaggg gcaccagct	9120
ttcttgtaga aagtggcgt taacggatcg gccagaatgg cccggaccgg gttaccgaat	9180
tcgagctcgg taccctggga tccgatatcg atgggcctg gccgaagctt tgagagtaca	9240
atgatgaacc tagattaatc aatgccaaag tctgaaaaat gcaccctcag tctatgatcc	9300
agaaaatcaa gattgcttga ggccctgttc ggttgttccg gattagagcc ccgattaat	9360
tcctagccgg attacttctc taatttataat agattttgat gagctggaat gaatcctggc	9420
ttattccggt acaaccgaac aggcctgaa ggataccagt aatcgctgag ctaaattggc	9480
atgctgtcag agtgtcagta ttgcagcaag gtagtgagat aaccggcatc atggtgccag	9540
tttgatggca ccattagggg tagagatggg ggccatgggc gcatgtcctg gccaaacttg	9600
tatgatatat ggcaggttga ataggaaagt aaaattgtat tgtaaaaagg gatttcttct	9660
gtttgttagc gcatgtacaa ggaatgcaag ttttgagcga gggggcatca aagatctggc	9720
tgtgtttcca gctgtttttg ttagcccat cgaatccttg acataatgat cccgcttaa	9780
taagcaacct cgcttgata gtcccttggt ctctaacaca cgatgatgat aagtcgtaaa	9840
atagtgtgt ccaaagaatt tccaggcca gttgtaaaag ctaaaatgct attcgaattt	9900
ctactagcag taagtcgtgt ttagaaatta tttttttata tacctttttt ccttctatgt	9960
acagtaggac acagtgtcag cgccgcttg acggagaata tttgcaaaaa agtaaaagag	10020
aaagtcatag cggcgatatg gccaaaaact tcgtcacaga gagggccata agaaacatgg	10080

ES 2 843 556 T3

ccacgcccc	aatacgaagc	accgcgacga	agcccaaaca	gcagtccgta	ggtggagcaa	10140
agcgcgtgggt	aatacgcaaa	cgttttgtcc	caccttgact	aatcacaaga	gtggagcgta	10200
ccttataaac	cgagccgcaa	gcaccgaatt	gttacttctc	taagcacggc	gttttagagc	10260
tagaaatagc	aagttaaaat	aaggctagtc	cgttatcaac	ttgaaaaagt	ggcaccgagt	10320
cgtgtccttt	tttttgcggc	cgcgaattcc	tgcaggggcc	tcttgtcgga	ccagttgcc	10380
accacgttgg	tgaggggcaa	ctttattata	caaagttgat	agataatacg	cgtctatagt	10440
attttaaaat	tgcattaaca	aacatgtcct	aattggtact	cctgagatac	tataccctcc	10500
tgttttaaaa	tagttggcat	tatcgaatta	tcattttact	ttttaatggt	ttctcttctt	10560
ttaatatatt	ttatgaattt	taatgtattt	taaaatgtta	tgagttcg	tctggacttt	10620
tctgctgctg	ctacacttgg	gtgtactggg	cctaaattca	gcctgaccga	ccgcctgcat	10680
tgaataatgg	atgagcaccg	gtaaaatccg	cgtacccaac	tttcgagaag	aaccgagacg	10740
tgccggggccg	ggccaccgac	gcacggcacc	agcgactgca	cacgtcccgc	cggcgtagct	10800
gtacgtgctg	ttccctcact	ggccgcccaa	tccactcatg	catgccacg	tacaccctg	10860
ccgtggcgcg	cccagatcct	aatcctttcg	ccgttctgca	cttctgctgc	ctataaatgg	10920
cggcatcgac	cgtcacctgc	ttcaccaccg	gcgagccaca	tcgagaacac	gatcgagcac	10980
acaagcacga	agactcgttt	aggagaaacc	acaaaccacc	aagccgtgca	agcaccaggc	11040
ttgggcaccc	gctccgggct	tagaaggcca	gcttcaagtt	tgtacaaaaa	agcaggcttc	11100
gaaggagata	gaaccgatcc	accatgtcca	acctgctcac	ggttcaccag	aaccttccgg	11160
ctcttccagt	ggacgcgacg	tccgatgaag	tcaggaagaa	cctcatggac	atgttccgcg	11220
acaggcaagc	gttcagcgag	cacacctgga	agatgctgct	ctccgtctgc	cgtcctggg	11280
ctgcatggtg	caagctgaac	aacaggaagt	ggttccccgc	tgagcccgag	gacgtgaggg	11340
attaccttct	gtacctgcaa	gcgcgaggtt	tgtttctgct	tctacctttg	atatatatat	11400
aataattatc	attaattagt	agtaataata	tatttcaa	atttttttca	aaataaaaga	11460
atgtagtata	tagcaattgc	ttttctgtag	tttataagtg	tgtatatttt	aatttataac	11520
ttttctaata	tatgaccaa	acatggtgat	gcctaggtct	ggcagtgaag	accatccagc	11580
aacaccttgg	acaactgaac	atgcttcaca	ggcgctccg	cctccgcgc	cccagcgact	11640
cgaacgccgt	gagcctcgtc	atgcgccgca	tcaggaagga	aaacgtcgat	gccggcgaaa	11700
gggcaaagca	ggccctcgcg	ttcgagagga	ccgatttcga	ccaggtccgc	agcctgatgg	11760
agaacagcga	caggtgccag	gacattagga	acctggcggt	cctcggaatt	gcatacaaca	11820
cgtcctcag	gatcgcgga	attgcccgca	ttcgctgaa	ggacattagc	cgcaccgacg	11880
gcggcaggat	gcttatccac	attggcagga	ccaagacgct	cgtttccacc	gcaggcgctg	11940
aaaaggccct	cagcctcgga	gtgaccaagc	tcgtcgaacg	ctggatctcc	gtgtccggcg	12000

ES 2 843 556 T3

tcgcggacga cccaaacaac tacctcttct gccgcgtccg caagaacggg gtggctgccc	12060
ctagcgccac cagccaactc agcacgaggg ccttggaagg tattttcgag gccaccacc	12120
gcctgatcta cggcgcgaag gatgacagcg gtcaacgcta cctcgcatgg tccgggcact	12180
ccgcccgcgt tggagctgct agggacatgg cccgcgccgg tgtttccatc cccgaaatca	12240
tgcaggcggg tggatggacg aacgtgaaca ttgtcatgaa ctacattcgc aaccttgaca	12300
gcgagacggg cgcaatggtt cgctcctgg aagatggtga ctgagctaga ccagctttc	12360
ttgtacaaag tggccgttaa cggatccaga cttgtccatc ttctggattg gccaaactaa	12420
ttaatgtatg aaataaaagg atgcacacat agtgacatgc taatcactat aatgtgggca	12480
tcaaagttgt gtgttatgtg taattactag ttatctgaat aaaagagaaa gagatcatcc	12540
atatcttcta tcctaaatga atgtcacgtg tctttataat tctttgatga accagatgca	12600
tttcattaac caaatccata tacatataaa tattaatcat atataattaa tatcaattgg	12660
gtagcaaaa caaatctagt ctagggtgtg tttgcgaatg cggccgccac cgcggcctaa	12720
ggggtcaccg gcctagaagg ccatttaa at cagatatctg gtccctaagg acccggcgga	12780
ccgattaaac tgattcggtc cgacgcgtta cctaataact tcgtatagca tacattatac	12840
gaagttatta gctaactaac tagtaagctt aagccatggg tcgacggatc cgttaactca	12900
cttaagatgt actcgacaat ggtgccctca taccggcatg tgtttcctag aaataatcaa	12960
tatattgatt gagatttatc tcgatatat tctgaactat gttcatcata taaataactg	13020
aaaacatcaa atcgtaattt taaactcatg ctttgtcaat acatagataa tacaatatta	13080
cttcacatc ccaatgatgt cctagcacaa cctattgaat gttaatgttt ggttgtgtgg	13140
gggtgtgttt ataacataga tgtgattatt tgtgcttttt gttgagtata tacatatatg	13200
gtatgttgat ttgatatagt gatggacaca tgctttggcc ttggatatcc aaatcacttg	13260
tacttgacg aagcaaaa taatataagt ttagaagtaa acttgtaact gtgtccaaac	13320
atgctcacac aaagtcatat cgcattatat ttttttggt aatattcaac acatgtattt	13380
tttacaagaa cccaaatttt acagacaaat gcagcattgt agacatgtag aattctttga	13440
agcatgtgaa cttaacaaca ctaatgtcat taaatcaact ggaccctatg agtaacaatt	13500
tcgatattgc aaacaccaa ttatggaact tatttgctga aaaaattatg atcaatgtga	13560
agtttaaatt attataccat aaatatatca aagatttttt ttgaggaagg taaaaattgc	13620
atggaatggg ctgccaacg tgatagctca cttttatgct aggtagcatt accaaagatg	13680
ggaatgttct gatgaacacc aaaccactc aaataatatt tatatttggg ttgtttagtt	13740
gtaaaagtga agaccaaga ttaaagtacc aattgtcgat gccatacact agaagtgaag	13800
aatgttgtaa gcgttaggag tagatgtgtc aatggggacc gatactcgtt atcgatgggc	13860

ES 2 843 556 T3

cgacgtaag cttgcatgcc tgcagtgcag cgtgacccgg tcgtgccct ctctagagat	13920
aatgagcatt gcatgtctaa gttataaaaa attaccacat attttttttg tcacacttgt	13980
ttgaagtcca gtttatctat ctttatacat atatttaaac tttactctac gaataatata	14040
atctatagta ctacaataat atcagtgttt tagagaatca tataaatgaa cagttagaca	14100
tgttctaaag gacaattgag tattttgaca acaggactct acagttttat ctttttagtg	14160
tgcattgtgt ctcttttttt tttgcaaata gcttcaccta tataatactt catccatttt	14220
attagtacat ccatttaggg tttaggggta atggttttta tagactaatt ttttttagtac	14280
atctatttta ttctatttta gcctctaaat taagaaaact aaaactctat tttagttttt	14340
ttattttaata atttagatat aaaatagaat aaaataaagt gactaaaaat taaacaaata	14400
ccctttaaga aattaaaaaa actaaggaaa catttttctt gtttcgagta gataatgcca	14460
gcctgttaaa cgccgtcgac gagtctaacg gacaccaacc agcgaaccag cagcgtcgcg	14520
tcggggccaag cgaagcagac ggcacggcat ctctgtcgct gcctctggac ccctctcgag	14580
agttccgctc caccgttggg cttgctccgc tgtcggcatc cagaaattgc gtggcgagac	14640
ggcagacgtg agccggcacg gcaggcgcc tcctcctcct ctcacggcac cggcagctac	14700
gggggattcc tttcccaccg ctctctcgct ttcccttctt cggccgcctt aataaataga	14760
cacccctcc acaccctctt tccccaacct cgtgttgttc ggagcgcaca cacacacaac	14820
cagatctccc ccaaattccac ccgtcggcac ctccgcttca aggtacgccg ctctcctcc	14880
ccccccccc tctctacctt ctctagatcg gcgttcgggt ccatgcatgg ttagggcccg	14940
gtagtcttac ttctgttcat gtttgtgtta gatccgtgtt tgtgttagat ccgtgctgct	15000
agcgttcgta cacggatgcg acctgtacgt cagacacgtt ctgattgcta acctgccagt	15060
gtttctcttt ggggaatcct gggatggctc tagccgttcc gcagacggga tcgatttcat	15120
gatttttttt gtttcgttgc atagggtttg gtttgccctt ttcttttatt tcaatatatg	15180
ccgtgcactt gtttgcggg tcactctttc atgctttttt ttgtcttggg tgtgatgatg	15240
tgttctggtt gggcggtcgt tctagatcgg agtagaattc tgtttcaaac tacctggtgg	15300
atttattaat tttggatctg tatgtgtgtg ccatacatat tcatagttag gaattgaaga	15360
tgatggatgg aaatatcgat ctaggatagg tatacatgtt gatgcgggtt ttactgatgc	15420
atatacagag atgctttttg ttccgttggg tgtgatgatg tgggtgtggtt gggcggtcgt	15480
tcattcgttc tagatcggag tagaatactg tttcaaacta cctggtgtat ttattaattt	15540
tggaactgta tgtgtgtgtc atacatcttc atagttagca gtttaagatg gatggaaata	15600
tcgatctagg ataggatata atgttgatgt gggttttact gatgcatata catgatggca	15660
tatgcagcat ctattcataat gctctaacct tgagtaccta tctattataa taaacaagta	15720
tgttttataa ttattttgat cttgatatac ttggatgatg gcatatgcag cagctatatg	15780

ES 2 843 556 T3

tggatTTTTT tagccctgcc ttcatacgct atttatttgc ttggtactgt ttcttttgtc	15840
gatgctcacc ctgttgtttg gtgttacttc tgcaggtcga ctctagagga tccaccatgg	15900
ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct	15960
atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc	16020
aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg	16080
acgaggcagc gcggctatcg ttgctggcca cgacgggcgt tccttgcgca gctgtgctcg	16140
acgttgctac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc	16200
tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc	16260
ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg	16320
agcgagcacg tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc	16380
atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg	16440
atgatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc	16500
gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag	16560
cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttctcgc	16620
tgctttacgg tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg	16680
agtcttctg aggatccacc atggttaacc tagacttgct catcttctgg attggccaac	16740
ttaattaatg tatgaaataa aaggatgcac acatagtgc atgctaata ctataatgtg	16800
ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct gaataaaaaga gaaagagatc	16860
atccatattt cttatccata atgaatgtca cgtgtcttta taattctttg atgaaccaga	16920
tgcatttcat taaccaaaac catatacata taaatattaa tcatatataa ttaatataca	16980
ttgggttagc aaaacaaaac tagtctaggt gtgttttgcg aatgcgacct tcttatgtgc	17040
ttctagtctc caaatgtggt tgatagtatt tttgctctaa gatcaacagt aatgaagtat	17100
aaatcatcgt tgtggtgtgc tactcggtta attgagcatt aacacacaca aacatgacga	17160
ggatgggtata atctccaaaa atgtgtactt tgttaggtgg gaccctatag ccttgattaa	17220
tgtgctatgt taggcatgcc tggaaacgtg tgacgcataat gttttgtgaa cctgttgata	17280
ttatatgtgc ttttatatta ccatatttta ttaaaataact aatattttatt actagtaaga	17340
tataacattc tatctagctt aaaaactaac cataaatatt ccataataac tagatttacc	17400
aaactaatat actaaatata cataataaat acaaaattaa caagacaata atcaatattt	17460
atgagcttaa tatatttaga cattatggtt ggtcgacgat aatcatgcta acttttcgta	17520
attgcttgat tgaaatatgc ttagaataat gcctctttgt tctacatggc aaatagggac	17580
cattatggtg taacaccctg ggaaccacaa acaccccgaa atgctactaa actacacaac	17640

ES 2 843 556 T3

taaccttcat	atataaaatt	tcgacagcat	ctcctttgaa	aatttgcata	gacgtggaag	17700
caacagagta	taaacagata	tcatgataag	aaaacatact	agacattaat	aatctgctag	17760
aaatgggaag	aatcctaact	tgacgactgc	gtaactgact	agagtacac	ttagctgacc	17820
ctagtcaact	acaactgact	tcgtgtccta	gggcgatccg	atatcgatgg	gccctggccg	17880
aagcttggtc	acccggtccg	ggcctagaag	gccagcttcg	gccgccccaa	ttcccatgga	17940
gtcaaagatt	caaatagagg	acctaacaga	actcgccgta	aagactggcg	aacagttcat	18000
acagagtctc	ttacgactca	atgacaagaa	gaaaatcttc	gtcaacatgg	tggagcacga	18060
cacgcttgtc	tactccaaaa	atatcaaaga	tacagtctca	gaagacccaa	gggcaattga	18120
gacttttcaa	caaagggtaa	tatccggaaa	cctcctcgga	ttccattgcc	cagctatctg	18180
tcactttatt	gtgaagatag	tggaaaagga	aggtggctcc	tacaaatgcc	atcattgcga	18240
taaaggaaag	gccatcggtg	aagatgcctc	tgccgacagt	ggtcccaaag	atggaccccc	18300
acccacgagg	agcatcggtg	aaaaagaaga	cgttccaacc	acgtcttcaa	agcaagtgga	18360
ttgatgtgat	atctccactg	acgtaaggga	tgacgcacaa	tcccactaag	ctgaccgaag	18420
ctggccgctc	tagaactagt	ggatctcgat	gtgtagtcta	cgagaagggt	taaccgtctc	18480
ttcgtgagaa	taaccgtggc	ctaaaaataa	gccgatgagg	ataaataaaa	tgtgggtgta	18540
cagtacttca	agaggtttac	tcatcaagag	gatgcttttc	cgatgagctc	tagtagtaca	18600
tcggacctca	catacctcca	ttgtggtgaa	atattttgtg	ctcatttagt	gatgggtaaa	18660
ttttgtttat	gtcactctag	gttttgacat	ttcagttttg	ccactcttag	gttttgacaa	18720
ataattttcca	ttccgcgga	aaagcaaaac	aattttat	tacttttacc	actcttagct	18780
ttcacaatgt	atcacaatg	ccactctaga	aattctgttt	atgccacaga	atgtgaaaaa	18840
aaacactcac	ttatttgaa	ccaaggtgtt	catggcatgg	aatgtgaca	taaagtaacg	18900
ttcgtgtata	agaaaaaatt	gtactcctcg	taacaagaga	cggaaacatc	atgagacaat	18960
cgcgttttga	aggcttttga	tcaccttttg	atgatgcgca	tgaatggagt	cgtctgcttg	19020
ctagccttcg	cctaccgccc	actgagtcgg	ggcggcaact	accatcgggc	aacgaccag	19080
ctgacctcta	ccgaccggac	ttgaatgcgc	taccttcgtc	agcgacgatg	gccgcgtacg	19140
ctggcgacgt	gccccgcgat	gcatggcggc	acatggcgag	ctcagaccgt	gcgtggctgg	19200
ctacaaatac	gtaccccggtg	agtgccctag	ctagaaactt	acacctgcaa	ctgcgagagc	19260
gagcgtgtga	gtgtagccga	gtagatcccc	cggtgtgcag	gtcgactcta	gaggatccac	19320
cgtgcgccac	catggccctg	tccaacaagt	tcatcggcga	cgacatgaag	atgacctacc	19380
acatggacgg	ctgcgtgaac	ggccactact	tcaccgtgaa	gggcgagggc	agcggcaagc	19440
cctacgaggg	caccagacc	tccaccttca	aggtgacct	ggccaacggc	ggccccctgg	19500
ccttctcctt	cgacatcctg	tccaccgtgt	tcatgtacgg	caaccgctgc	ttcaccgcct	19560

ES 2 843 556 T3

accccaccag catgcccagac tacttcaagc aggccttccc cgacggcatg tcctacgaga 19620
 gaaccttcac ctacgaggac ggcggcgtgg ccaccgccag ctgggagatc agcctgaagg 19680
 gcaactgctt cgagcacaag tccaccttcc acggcgtgaa cttccccgcc gacggccccg 19740
 tgatggccaa gaagaccacc ggctgggacc cctccttcga gaagatgacc gtgtgcgacg 19800
 gcatcttgaa gggcgacgtg accgccttcc tgatgctgca gggcggcggc aactacagat 19860
 gccagttcca cacctcctac aagaccaaga agcccgtagc catgcccccc aaccacgtgg 19920
 tggagcaccg catcgccaga accgacctgg acaagggcgg caacagcgtg cagctgaccg 19980
 agcacgccgt ggcccacatc acctccgtgg tgcccttctg aagcggccca tggatatctg 20040
 aacgcgtcgc tgaaatcacc agtctctctc tacaaatcta tctctctcta taataatgtg 20100
 tgagtagttc ccagataagg gaattagggt tcttataggg ttctgctcat gtgttgagca 20160
 tataagaaac ccttagtatg tatttgtatt tgtaaaatac ttctatcaat aaaatttcta 20220
 attcctaaaa ccaaatcca gtggcgagct cgagccaggg tgcttcctga gctgattccg 20280
 atgacttcgt aggttcctag ctcaagccgc tcgtgtccaa gcgtcactta cgattagcta 20340
 atgattacgg catctaggac cgactagcta actaactagt acgccgacta gctagctaag 20400
 gtaccgagct cgaattcatt ccgattaatc gtggcctctt gctcttcagg atgaagagct 20460
 atgtttaaac gtgcaagcgc tactagacaa ttcagtacat taaaaacgtc cgcaatgtgt 20520
 tattaagttg tctaagcgtc aatttgttta caccacaata tctctgcca 20570

<210> 73

<211> 20572

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 73

gtttaccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac 60
 aatctgatca tgagcggaga attaaggagac tcacgttatg acccccgccg atgacgcggg 120
 acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagcaag 180
 ctggtacgat tgtaatacga ctactatag gcggaattga gcgctgttta aacgtcttc 240
 aactggaaga gcggttacgc tgtttaaacg ctcttcaact ggaagagcgg ttactaccgg 300
 ttactagct agctgctaata cgagctagtt accctatgag gtgacatgaa gcgctcacgg 360
 ttactatgac ggtagcttc acgactgttg gtggcagtag cgtacgactt agctatagtt 420
 ccggacttac ccttaagata acttcgtata gcatacatc tacgaagtta tgggcccacg 480
 cgtgaagagc ggttaccaga gctggtcacc tttgtccacc aagatggaac tgtcttcggc 540

10

ES 2 843 556 T3

cagaatggcc tccggaatth atcgtatca actttgtata gaaaagtgg gccgaattcg	600
agctcggtag gcccagaatg gcccgaccg ggttaccgaa ttcgagctcg gtaccctggg	660
atccggtgag gccctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc	720
gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg	780
ccaagctcag atcagcttgc atgcctgcag tgcagcgtga cccggtcgtg cccctctcta	840
gagataatga gcattgcatg tctaagttat aaaaaattac cacatattht tttgtcaca	900
cttgtttgaa gtgcagttta tctatcttta tacatatatt taaactttac tctacgaata	960
atataatcta tagtactaca ataatatcag tgttttagag aatcatataa atgaacagtt	1020
agacatggtc taaaggacaa ttgagtatth tgacaacagg actctacagt tttatcttht	1080
tagtgtgcat gtgttctcct ttttttttgc aaatagcttc acctatataa tacttcatcc	1140
atthttattag tacatccatt tagggtttag ggthaatggg ttttatagac taatthttth	1200
agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaaac tctatthtag	1260
tttttttatt taataatthta gatataaaat agaataaaat aaagtgacta aaaatthaac	1320
aaataccctt taagaaatta aaaaaactaa ggaaacatth ttcttgthtc gagtagataa	1380
tgccagcctg ttaaacgccg tcgacgagtc taacggacac caaccagcga accagcagcg	1440
tcgcgtcggg ccaagcgaag cagacggcac ggcctctctg tcgctgcctc tggaccctc	1500
tcgagagttc cgctccaccg ttggacttgc tccgctgtcg gcctccagaa attgcgtggc	1560
ggagcggcag acgtgagccg gcacggcagg cggcctcctc ctctctcac ggcaccggca	1620
gtacggggg attcctthtc caccgctcct tcgctthccc ttctcgccc gccgtaataa	1680
atagacaccc cctccacacc ctctthcccc aacctcgtgt tgttcggagc gcacacacac	1740
acaaccagat ctccccaaa tccaccgctc ggcacctccg cttcaaggta cgcgctcgt	1800
cctccccccc cccctctct acctctctcta gatcggcgtt ccggtccatg catggttagg	1860
gcccggtagt tctacttctg ttcattgttg tgttagatcc gtgtttgtgt tagatccgtg	1920
ctgctagcgt tcgtacacgg atcgacctg tacgtcagac acgttctgat tgctaacttg	1980
ccagtgtthc tctttgggga atcctgggat ggctctagcc gttccgcaga cgggatcgat	2040
ttcatgattt tttttgtthc gttgcatagg gtttggttg ccctthtctt ttattthcaat	2100
atatgccgtg cacttgthtg tcgggtcatc ttttcatgct tttttttgtc ttggttgtga	2160
tgatgtggtc tggttgggag gtcgttctag atcgagtag aattctgtth caaactacct	2220
ggtggattta ttaattthtg atctgtatgt gtgtgccata catattcata gttacgaatt	2280
gaagatgatg gatggaaata tcgatctagg ataggtatac atgttgatgc gggthttact	2340
gatgcataa cagagatgct tttgttcgc ttggttgtga tgatgtggtg tggttgggag	2400
gtcgttcatt cgttctagat cggagtagaa tactgtthca aactacctgg tgtatttht	2460

ES 2 843 556 T3

aatTTTggaa ctgtatgtgt gtgtcataca tcttcatagt tacgagttta agatggatgg	2520
aaatatcgat ctaggatagg tatacatgtt gatgtggggt ttactgatgc atatacatga	2580
tggcatatgc agcatctatt catatgctct aaccttgagt acctatctat tataataaac	2640
aagtatgttt tataattatt ttgatcttga tataacttga tgatggcata tgcagcagct	2700
atatgtggat ttttttagcc ctgccttcat acgctattta tttgcttggg actgtttctt	2760
ttgtcgatgc tcacctgtt gtttggtgtt acttctgcag gtgcactcta gaggatccac	2820
cggtcgccac catggcacccg aagaagaagc gcaaggtgat ggacaagaag tacagcatcg	2880
gcctcgacat cggcaccaac tcggtgggct gggccgtcat cacggacgaa tataaggtcc	2940
cgtcgaagaa gttcaaggtc ctcggaata cagaccgcca cagcatcaag aaaaacttga	3000
tcggcgccct cctgttcgat agcggcgaga ccgcgagggc gaccaggctc aagaggaccg	3060
ccaggagacg gtacactagg cgcaagaaca ggatctgcta cctgcaggag atcttcagca	3120
acgagatggc gaaggtggac gactccttct tccaccgctt ggaggaatca ttcctgggtg	3180
aggaggacaa gaagcatgag cggcacccaa tcttcggcaa catcgtcgac gaggtaagtt	3240
tctgcttcta cctttgatat atatataata attatcatta attagtagta atataatatt	3300
tcaaatatTT ttttcaaat aaaagaatgt agtatatagc aattgctttt ctgtagttta	3360
taagtgtgta tattttaatt tataactttt ctaatatatg accaaaacat ggtgatgtgc	3420
aggtggccta ccacgagaag tacccgacaa tctaccacct ccggaagaaa ctggtggaca	3480
gcacagacaa ggcggacctc cggtcatctt accttgcctt cgcgcatatg atcaagttcc	3540
gcggccactt cctcatcgag ggcgacctga acccgacaa ctccgacgtg gacaagctgt	3600
tcatccagct cgtgcagacg tacaatcaac tgttcgagga gaaccata aacgctagcg	3660
gcgtggacgc caaggccatc ctctcgcca ggctctcgaa atcaagaagg ctggagaacc	3720
ttatcgcgca gttgccaggc gaaaagaaga acggcctctt cggcaacctt attgcgctca	3780
gcctcgccct gacgccgaac ttcaaatcaa acttcgacct cgcggaggac gccaaagctcc	3840
agctctcaaa ggacacctac gacgacgacc tcgacaacct cctggcccag ataggagacc	3900
agtacgcgga cctcttcctc gccgccaaga acctctccga cgctatcctg ctacgcgaca	3960
tccttcgggt caacaccgaa attaccaagg caccgctgtc cgccagcatg attaaacgct	4020
acgacgagca ccatcaggac ctacgctgc tcaaggcact cgtccgcccag cagctccccg	4080
agaagtacaa ggagatcttc ttcgaccaat caaaaaacgg ctacgcggga tatatcgacg	4140
gcggtgccag ccaggaagag ttctacaagt tcatcaaacc aatcctggag aagatggacg	4200
gcaccgagga gttgctggtc aagctcaaca gggaggacct cctcaggaag cagaggacct	4260
tcgacaacgg ctccatcccg catcagatcc acctggcgga actgcatgcc atcctgcggc	4320

ES 2 843 556 T3

gccaggagga cttctacccg ttcctgaagg ataaccggga gaagatcgag aagatcttga	4380
cgttccgcat cccatactac gtgggcccg cggctcgcg caactcccg ttgcctgga	4440
tgaccggaa gtccgaggag accatcacac cctggaactt tgaggaggtg gtcgataagg	4500
gcgctagcgc tcagagcttc atcgagcgca tgaccaactt cgataaaaac ctgcccaatg	4560
aaaaagtcc cccaagcac tcgctgctct acgagtactt caccgtgtac aacgagctca	4620
ccaaggtcaa atacgtcacc gagggcatgc ggaagccggc gttcctgagc ggcgagcaga	4680
agaaggcgat agtggacctc ctcttcaaga ccaacaggaa ggtgacctg aagcaattaa	4740
aagaggacta cttcaagaaa atagagtgtc tcgactccgt ggagatctcg ggcgtggagg	4800
atcggttcaa cgcctcactc ggcacgtatc acgacctcct caagatcatt aaagacaagg	4860
acttcctcga caacgaggag aacgaggaca tcctcgagga catcgtcctc accctgaccc	4920
tgttcgagga ccgcgaaatg atcgaggaga ggctgaagac ctacgcgcac ctgttcgacg	4980
acaaggtcat gaaacagctc aagaggcgcc gctacactgg ttggggaagg ctgtcccgca	5040
agctcattaa tggcatcagg gacaagcaga gcggcaagac catcctggac ttcctcaagt	5100
ccgacgggtt cgccaaccgc aacttcatgc agctcattca cgacgactcg ctcacgttca	5160
aggaagacat ccagaaggca caggtgagcg ggcagggtga ctccctccac gaacacatcg	5220
ccaacctggc cggctcgccg gccattaaaa agggcatcct gcagacggtc aaggtcgtcg	5280
acgagctcgt gaaggtgatg ggccggcaca agcccgaata tatcgtcata gagatggcca	5340
gggagaacca gaccacccaa aaagggcaga agaactcgcg cgagcggatg aaacggatcg	5400
aggagggcat taaagagctc ggggtcccaga tcctgaagga gcaccccggtg gaaaataccc	5460
agctccagaa tgaaaagctc tacctctact acctgcagaa cggccgcgac atgtacgtgg	5520
accaggagct ggacattaat cggctatcgg actacgacgt cgaccacatc gtgccgcagt	5580
cgttcctcaa ggacgatagc atcgacaaca aggtgctcac ccggtcggat aaaaatcggg	5640
gcaagagcga caacgtgcc agcgaggagg tcgtgaagaa gatgaaaaac tactggcgcc	5700
agtcctcaa cgcgaaactg atcaccacgc gcaagttcga caacctgacg aaggcggaac	5760
gcggtggcct gagcgaactc gataaggcgg gcttcataaa aaggcagctg gtcgagacgc	5820
gccagatcac gaagcatgtc gccagatcc tggacagccg catgaatact aagtacgatg	5880
aaaacgacaa gctgatccgg gaggtgaagg tgatcacgct gaagtccaag ctcggtcgg	5940
acttcgcaa ggacttccag ttctacaagg tccgcgagat caacaactac caccacgcc	6000
acgacgccta cctgaatgcg gtggtcggga ccgccctgat caagaagtac ccgaagctgg	6060
agtcggagtt cgtgtacggc gactacaagg tctacgacgt gcgcaaatg atcgccaagt	6120
ccgagcagga gatcggcaag gccacggcaa aatacttctt ctactcgaac atcatgaact	6180
tcttcaagac cgagatcacc ctgcggaacg gcgagatccg caagcgcccg ctcatcgaaa	6240

ES 2 843 556 T3

ccaacggcga gacgggcgag atcgtctggg ataagggccg ggatttcgcg acggtccgca	6300
aggtgctctc catgccgcaa gtcaatatcg tgaaaaagac ggaggtccag acgggcgggt	6360
tcagcaagga gtccatcctc ccgaagcgca actccgacaa gctcatcgcg aggaagaagg	6420
attgggaccc gaaaaaatat ggcggcttcg acagcccgac cgctcgcatc agcgtcctcg	6480
tcgtggcgaa ggtggagaag ggcaagtcaa agaagctcaa gtccgtgaag gagctgctcg	6540
ggatcacgat tatggagcgg tcctccttcg agaagaacct gatcgacttc ctagaggcca	6600
agggatataa ggaggtcaag aaggacctga ttattaaact gccgaagtac tcgctcttcg	6660
agctggaaaa cggccgcaag aggatgctcg cctccgcagg cgagttgcag aagggcaacg	6720
agctcgccct cccgagcaaa tacgtcaatt tcctgtacct cgctagccac tatgaaaagc	6780
tcaagggcag cccggaggac aacgagcaga agcagctctt cgtggagcag cacaagcatt	6840
acctggacga gatcatcgag cagatcagcg agttctcgaa gcgggtgatc ctgccgcagc	6900
cgaacctgga caaggtgctg tcggcatata acaagcaccg cgacaaacca atacgcgagc	6960
aggccgaaaa tatcatccac ctcttcaccc tcaccaacct cggcgctccg gcagccttca	7020
agtacttcga caccacgatt gaccggaagc ggtacacgag cacgaaggag gtgctcgatg	7080
cgacgctgat ccaccagagc atcacagggc tctatgaaac acgcatcgac ctgagccagc	7140
tgggcccgaga caagagacca cgggaccgcc acgatggcga gctgggaggc cgcaagcggg	7200
caaggtaggt accgttaacc tagacttgct catcttctgg attggccaac ttaattaatg	7260
tatgaaataa aaggatgcac acatagtac atgctaata ctataatgtg ggcatcaaag	7320
ttgtgtgtta tgtgtaatta ctagtatatc gaataaaaga gaaagagatc atccatattt	7380
cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga tgcatttcat	7440
taaccaaadc catatacata taaatattaa tcatatataa ttaatataa ttgggttagc	7500
aaaacaaatc tagtctaggt gtgttttgcg aattcgtaat catgtcatag ctgtttcctg	7560
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgtg	7620
aagcctgggg tgcctaataa gtgagctaac tcacattaat tgcgttgccg tctactgccc	7680
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcgcccaa cgcgcgggga	7740
gaggcgggtt gcgtattggg cgctcttcg ctgatccgat atcgatgggc cctggccgaa	7800
gcttggtcac ccggtccggg cctagaaggc cagcttcaag tttgtacaaa aaagcaggct	7860
ccggccagaa tggcccggac cgggttaccg aattcgagct cggtagccctg ggatccgata	7920
tcgatgggcc ctggccgaag ctttgagagt acaatgatga acctagatta atcaatgcca	7980
aagtctgaaa aatgcacctc cagtctatga tccagaaaat caagattgct tgaggccctg	8040
ttcggttggt ccggattaga gccccggatt aattcctagc cggattactt ctctaattta	8100

ES 2 843 556 T3

tatagatttt gatgagctgg aatgaatcct ggcttattcc ggtacaaccg aacaggccct	8160
gaaggatacc agtaatcgct gagctaaatt ggcattgctgt cagagtgtca gtattgcagc	8220
aaggtagtga gataaccggc atcatgggtgc cagtttgatg gcaccattag ggtagagat	8280
ggtggccatg ggcgcattgc ctggccaact ttgtatgata tatggcaggg tgaataggaa	8340
agtaaaattg tattgtaaaa agggatttct tctgtttgtt agcgcattga caaggatgc	8400
aagttttgag cgagggggca tcaaagatct ggctgtgttt ccagctgttt ttgttagccc	8460
catcgaatcc ttgacataat gatcccgctt aaataagcaa cctcgcttgt atagttcctt	8520
gtgctctaac acacgatgat gataagtcgt aaaatagtggt tgtccaaaga atttccaggc	8580
ccagttgtaa aagctaaaat gctattcgaa tttctactag cagtaagtcg tgtttagaaa	8640
ttatTTTTTT atataccttt tttccttcta tgtacagtag gacacagtgt cagcgccgcy	8700
ttgacggaga atatttgcaa aaaagtaaaa gagaaagtca tagcggcgta tgtgccaaaa	8760
acttcgtcac agagagggcc ataagaaaca tggcccacgg cccaatacga agcaccgcga	8820
cgaagcccaa acagcagtcg gtagggtggag caaagcgctg ggtaatacgc aaacgttttg	8880
tcccaccttg actaatcaca agagtggagc gtaccttata aaccgagccg caagcaccga	8940
attgctccag ctgctctaatt cgatggtttt agagctagaa atagcaagtt aaaataaggc	9000
tagtccgtta tcaacttgaa aaagtggcac cgagtcggtg cttttttttt gcggccgcga	9060
attcctgcag ggccctcttg tcggaccagt tgcccaccac gttggtgagg ggcaccagc	9120
tttcttgtag aaagtggccg ttaacggatc ggccagaatg gcccgagccg ggttaccgaa	9180
ttcgagctcg gtaccctggg atccgatatc gatgggccct ggccgaagct ttgagagtac	9240
aatgatgaac ctagattaat caatgccaaa gtctgaaaaa tgcaccctca gtctatgatc	9300
cagaaaatca agattgcttg aggcctgtt cgggtgttcc ggattagagc cccggattaa	9360
ttcctagccg gattacttct ctaatttata tagattttga tgagctggaa tgaatcctgg	9420
cttattccgg tacaaccgaa caggccctga aggataccag taatcgctga gctaaattgg	9480
catgctgtca gagtgtcagt attgcagcaa ggtagtgaga taaccggcat catggtgcca	9540
gtttgatggc accattaggg ttagagatgg tggccatggg cgcattgcct ggccaacttt	9600
gtatgatata tggcaggggtg aataggaaag taaaattgta ttgtaaaaag ggatttcttc	9660
tgtttgttag cgcattgaca aggaatgcaa gttttgagcg agggggcatc aaagatctgg	9720
ctgtgtttcc agctgttttt gttagcccca tcgaatcctt gacataatga tcccgcttaa	9780
ataagcaacc tcgcttgat agttccttgt gctctaacac acgatgatga taagtcgtaa	9840
aatagtgggtg tccaaagaat ttccaggccc agttgtaaaa gctaaaaatgc tattcgaatt	9900
tctactagca gtaagtcgtg tttagaaatt atttttttat atacctttt tcttctatg	9960
tacagtagga cacagtgtca gcgccgcgtt gacggagaat atttgcaaaa aagtaaaaga	10020

ES 2 843 556 T3

gaaagtcata gcggcgtatg tgccaaaaac ttcgtcacag agagggccat aagaaacatg	10080
gcccacggcc caatacgaag caccgcgacg aagcccaaac agcagtccgt aggtggagca	10140
aagcgctggg taatacgcga acgttttgtc ccaccttgac taatcacaag agtggagcgt	10200
accttataaa ccgagccgca agcaccgaat tgcgtggtcc ttggctgcct tggttttaga	10260
gctagaaata gcaagttaaa ataaggctag tccgttatca acttgaaaaa gtggcaccga	10320
gtcggtgctt tttttttgcg gccgcgaatt cctgcagggc cctcttgctg gaccagttgc	10380
ccaccacgtt ggtgaggggc aactttatta taaaaagttg atagataata cgcgtctata	10440
gtatttttaa attgcattaa caaacatgtc ctaattggta ctcttgagat actataccct	10500
cctgttttaa aatagttggc attatcgaat tatcatttta ctttttaatg ttttctcttc	10560
ttttaatata ttttatgaat tttaatgtat tttaaaatgt tatgcagttc gctctggact	10620
tttctgctgc gcctacactt ggggtgtactg ggcctaaatt cagcctgacc gaccgcctgc	10680
attgaataat ggatgagcac cggtaaaatc cgcgtaccca actttcgaga agaaccgaga	10740
cgtggcgggc cgggccaccg acgcacggca ccagcgactg cacacgtccc gccggcgtag	10800
gtgtacgtgc tgttccctca ctggccgccc aatccactca tgcagtccca cgtacacccc	10860
tgccgtggcg cgcccagatc ctaatccttt cgccgttctg cacttctgct gcctataaat	10920
ggcggcatcg accgtcacct gcttcaccac cggcgagcca catcgagaac acgatcgagc	10980
acacaagcac gaagactcgt ttaggagaaa ccacaaacca ccaagccgtg caagcaccag	11040
gcttgggcac ccgctccggg cttagaaggc cagcttcaag tttgtacaaa aaagcaggct	11100
tcgaaggaga tagaacgat ccaccatgtc caacctgctc acggttcacc agaacccttc	11160
ggctcttcca gtggacgca cgtccgatga agtcaggaag aacctcatgg acatgttccg	11220
cgacaggcaa gcgttcagcg agcacacctg gaagatgctg ctctccgtct gccgtcctg	11280
ggctgcatgg tgcaagctga acaacaggaa gtggttcccc gctgagcccg aggacgtgag	11340
ggattacctt ctgtacctgc aagcgcgagg tttgtttctg cttctacctt tgatatatat	11400
ataataatta tcattaatta gtagtaatat aatatttcaa atattttttt caaaataaaa	11460
gaatgtagta tatagcaatt gcttttctgt agtttataag tgtgtatatt ttaatttata	11520
acttttctaa tatatgacca aaacatggtg atgcctaggt ctggcagtga agaccatcca	11580
gcaacacctt ggacaactga acatgcttca caggcgctcc ggctcccgc gccccagcga	11640
ctcgaacgcc gtgagcctcg tcatgcgccg catcaggaag gaaaacgtcg atgccggcga	11700
aagggcaaag caggccctcg cgttcgagag gaccgatttc gaccaggtcc gcagcctgat	11760
ggagaacagc gacaggtgcc aggacattag gaacctggcg ttctcggaa ttgcatacaa	11820
cacgctcctc aggatcgcg aaattgcccg cattcgcggtg aaggacatta gccgcaccga	11880

ES 2 843 556 T3

cgcgggcagg	atgcttatcc	acattggcag	gaccaagacg	ctcgtttcca	ccgcaggcgt	11940
cgaaaaggcc	ctcagcctcg	gagtgaacaa	gctcgtcgaa	cgctggatct	ccgtgtccgg	12000
cgtcgcgga	gacccaaaca	actacctctt	ctgccgcgtc	cgcaagaacg	gggtggctgc	12060
ccctagcgcc	accagccaac	tcagcacgag	ggccttggaa	ggtattttctg	aggccacca	12120
ccgcctgatc	tacggcgcg	aggatgacag	cggtcaacgc	tacctcgcat	ggtccgggca	12180
ctccgcccgc	gttgagctg	ctagggacat	ggcccgcgcc	ggtgtttcca	tccccgaaat	12240
catgcaggcg	gggtgatgga	cgaacgtgaa	cattgtcatg	aactacattc	gcaaccttga	12300
cagcgagacg	ggcgcaatgg	ttcgctcct	ggaagatgg	gactgagcta	gaccagctt	12360
tcttgtacaa	agtggccgtt	aacggatcca	gacttgtcca	tcttctggat	tggccaactt	12420
aattaatgta	tgaataaaaa	ggatgcacac	atagtacat	gctaatact	ataatgtggg	12480
catcaaagtt	gtgtgttatg	tgtattact	agttatctga	ataaaagaga	aagagatcat	12540
ccatatttct	tatcctaaat	gaatgtcacg	tgtctttata	attctttgat	gaaccagatg	12600
catttcatta	accaaatacca	tatacatata	aatattaatc	atatataatt	aatatcaatt	12660
gggttagcaa	aacaaatcta	gtctaggtgt	gttttgcgaa	tgcggccgcc	accgcggcct	12720
aaggggtcac	cggcctagaa	ggccatttaa	atcagatata	tggttcctaa	ggaccggcg	12780
gaccgattaa	actgattcgg	tccgacgcgt	tacctaatata	cttcgtatag	catacattat	12840
acgaagttaa	tagctaacta	actagtaagc	ttaagccatg	ggtcgacgga	tccgttaact	12900
cacttaagat	gtactcgaca	atggtgccct	cataccggca	tgtgtttcct	agaaataatc	12960
aatatattga	ttgagattta	tctcgatata	tttctgaact	atgttcatca	tataaataac	13020
tgaaaacatc	aaatcgtaat	tttaaaactca	tgcttggtca	atacatagat	aatacaatat	13080
tacttcatca	tcccaatgat	gtcctagcac	aacctattga	atgttaatgt	ttggttgtgt	13140
gggggtgtgt	ttataacata	gatgtgatta	tttgtgcttt	ttgttgagta	tatacatata	13200
tggtatgttg	atttgatata	gtgatggaca	catgcttttg	ccttggatat	tcaaatact	13260
tgtacttgca	cgaagcaaaa	cataatataa	gtttagaagt	aaacttgtaa	ctgtgtccaa	13320
acatgctcac	acaaagtcac	atcgcatat	atttttttgg	taaatattca	acacatgtat	13380
tttttacaag	aacccaaatt	ttacagacaa	atgcagcatt	gtagacatgt	agaattcttt	13440
gaagcatgtg	aacttaacaa	cactaatgtc	attaaatcaa	ctggacccta	tgagtaacaa	13500
tttcgatatt	gcaaacacca	aattatggaa	cttatttgct	gaaaaaatta	tgatcaatgt	13560
gaagttttaa	ttattatacc	ataaatatat	caaagatttt	ttttgaggaa	ggtaaaaatt	13620
gcatggaatg	ggctgcccaa	cgtgatagct	cacttttatg	ctaggtagca	ttaccaaga	13680
tggaatgtt	ctgatgaaca	ccaaaccac	tcaaataata	tttatatttg	ggttgtttag	13740
ttgtaaaagt	gaagacccaa	gattaaagta	ccaattgtcg	atgccatata	ctagaagtga	13800

ES 2 843 556 T3

agaatgttgt tagcgttagg agtagatgtg tcaatgggga ccgataactcg ttatcgatgg	13860
gccgcacgta agcttgcatg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag	13920
ataatgagca ttgcatgtct aagttataaa aaattaccac atattttttt tgtcacactt	13980
gtttgaagtg cagtttatct atctttatac atatatatta actttactct acgaataata	14040
taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga	14100
catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atcttttttag	14160
tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcattccatt	14220
ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa tttttttagt	14280
acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt	14340
ttttatttaa taatttagat ataaaataga ataaaataaa gtgactaaaa attaaacaaa	14400
taccctttta gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc	14460
cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg	14520
cgtcggggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg	14580
agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga	14640
gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct	14700
acgggggatt cctttccac cgctccttcg ctttcccttc ctgcccgc gtaataaata	14760
gacacccct ccacaccctc tttcccaac ctctgttgtg tcggagcgca cacacacaca	14820
accagatctc ccccaaacc acccgtcggc acctccgctt caaggtacgc cgctcgtcct	14880
ccccccccc cctctctacc ttctctagat cggcgttcg gtccatgcat ggttagggcc	14940
cggtagttct acttctgttc atgttttgtg tagatccgtg tttgtgttag atccgtgctg	15000
ctagcgttcg tacacggatg cgacctgtac gtcagacacg ttctgattgc taacttgcca	15060
gtgtttctct ttggggaatc ctgggatggc tctagccgtt ccgcagacgg gatcgatttc	15120
atgatttttt ttgtttcgtt gcatagggtt tggtttgccc ttttccttta tttcaatata	15180
tgccgtgcac ttgtttgtcg ggtcatcttt tcatgctttt ttttgtcttg gttgtgatga	15240
tgtggtctgg ttgggcggtc gttctagatc ggagtagaat tctgtttcaa actacctggt	15300
ggatttatta attttggatc tgtatgtgtg tgccatacat attcatagtt acgaattgaa	15360
gatgatggat ggaaatatcg atctaggata ggtatacatg ttgatgcggg ttttactgat	15420
gcatatacag agatgctttt tgttcgcttg gttgtgatga tgtggtgtgg ttgggcggtc	15480
gttcattcgt tctagatcgg agtagaatac tgtttcaaac tacctggtgt atttattaat	15540
tttggaactg tatgtgtgtg tcatacatct tcatagttac gagtttaaga tggatggaaa	15600
tatcgatcta ggataggat acatgttgat gtgggtttta ctgatgcata tacatgatgg	15660

ES 2 843 556 T3

catatgcagc atctattcat atgctctaac cttgagtacc tatctattat aataaacaag	15720
tatgttttat aattattttg atcttgatat acttggatga tggcatatgc agcagctata	15780
tgtggatttt tttagccctg ccttcatacg ctatttattt gcttggtact gtttcttttg	15840
tcgatgctca ccctgttggt tgggtgttact tctgcaggtc gactctagag gatccaccat	15900
ggttgaacaa gatggattgc acgcagggtc tccggccgct tgggtggaga ggctattcgg	15960
ctatgactgg gcacaacaga caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc	16020
gcagggggcg ccggttcttt ttgtcaagac cgacctgtcc ggtgccctga atgaactgca	16080
ggacgaggca gcgcggctat cgtggctggc cacgcagggc gttccttgcg cagctgtgct	16140
cgacgttgtc actgaagcgg gaagggactg gctgctattg ggcaagtgc cggggcagga	16200
tctcctgtca tctcaccttg ctctgccga gaaagtatcc atcatggctg atgcaatgcg	16260
gcggctgcat acgcttgatc cggctacctg cccattcgac caccaagcga aacatcgcat	16320
cgagcgagca cgtactcgga tggaaagccg tcttgctgat caggatgatc tggacgaaga	16380
gcatcagggg ctgcgccag ccgaactggt cgccaggctc aaggcgcgca tgcccgcagg	16440
cgatgatctc gtctgaccc atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg	16500
ccgcttttct ggattcatcg actgtggccg gctgggtgtg gcggaccgct atcaggacat	16560
agcgttggt acccgtgata ttgctgaaga gcttggcggc gaatgggctg accgcttct	16620
cgtgctttac ggtatcgccg ctcccgattc gcagcgcatc gccttctatc gccttcttga	16680
cgagttcttc tgaggatcca ccatggttaa cctagacttg tccatcttct ggattggcca	16740
acttaattaa tgtatgaaat aaaaggatgc acacatagtg acatgctaata cactataatg	16800
tgggcatcaa agttgtgtgt tatgtgtaat tactagtatt ctgaataaaa gagaaagaga	16860
tcatccatat ttcttatcct aaatgaatgt cactgtctct tataattctt tgatgaacca	16920
gatgcatttc attaaccaaa tccatataca tataaatatt aatcatatat aattaatatc	16980
aattgggtta gcaaaacaaa tctagtctag gtgtgttttg cgaatcgac cttcttatgt	17040
gcttctagtc tccaaatgtg gttgatagtt attttgcctc aagatcaaca gtaatgaagt	17100
ataaatcatc gttgtgggtg gctactcggc taattgagca ttaacacaca caaacatgac	17160
gaggatggta taatctccaa aatgtgtac tttgttaggt gggaccctat agccttgatt	17220
aatgtgctat gttaggcatg cctggaaacg tgtgacgcat atgttttgtg aacctgttga	17280
tattatatgt gcttttatat taccatattt tattaataa ctaatattta ttactagtaa	17340
gatataacat tctatctagc ttaaaaacta accataaata ttccataata actagattta	17400
ccaaaactaat atactaaata tacataataa atacaaaatt aacaagacaa taatcaatat	17460
ttatgagctt aatatattta gacattatgg ttggtcgacg ataatcatgc taacttttcg	17520
taattgcttg attgaaatat gcttagaata atgcctcttt gttctacatg gcaaataggg	17580

ES 2 843 556 T3

accattatgg tgtaacaccc tgggaaccac aaacaccccg aaatgctact aaactacaca	17640
actaaccttc atatataaaa ttogacagc atctcctttg aaaatttgca tagacgtgga	17700
agcaacagag tataaacaga tatcatgata agaaaacata ctagacatta ataactctgct	17760
agaaatggga agaatcctaa cttgacgact gcgtaactga ctagagtcac acttagctga	17820
ccctagtcac ttacaactga cttcgtgtcc tagggcgatc cgatatcgat gggccctggc	17880
cgaagcttgg tcacccggtc cgggcctaga aggccagctt cggccgcccc aattcccatg	17940
gagtcaaaga ttcaaataga ggacctaaca gaactcgccg taaagactgg cgaacagttc	18000
atacagagtc tcttacgact caatgacaag aagaaaatct tcgtcaacat ggtggagcac	18060
gacacgcttg tctactccaa aaatatcaaa gatacagtct cagaagacca aagggcaatt	18120
gagacttttc aacaaagggt aatatccgga aacctcctcg gattccattg cccagctatc	18180
tgctacttta ttgtgaagat agtggaagga gaaggtggct cctacaaatg ccatcattgc	18240
gataaaggaa aggccatcgt tgaagatgcc tctgccgaca gtggtcccaa agatggaccc	18300
cccccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg	18360
gattgatgtg atatctccac tgacgtaagg gatgacgcac aatccacta agctgaccga	18420
agctggccgc tctagaacta gtggatctcg atgtgtagtc tacgagaagg gttaaccgtc	18480
tcttcgtgag aataaccgtg gcctaaaaat aagccgatga ggataaataa aatgtggtgg	18540
tacagtactt caagaggttt actcatcaag aggatgcttt tccgatgagc tctagtagta	18600
catcggacct cacataacct cattgtgtgt aaatatattt tgctcattta gtgatgggta	18660
aattttgttt atgtcactct aggttttgac atttcagttt tgccactctt aggttttgac	18720
aaataatttc cattccgagg caaaagcaaa acaattttat tttactttta ccactcttag	18780
ctttcacaaat gtatcacaaa tgccactcta gaaattctgt ttatgccaca gaatgtgaaa	18840
aaaaacactc acttatttga agccaagggtg ttcattggcat ggaaatgtga cataaagtaa	18900
cgctcgtgta taagaaaaaa ttgtactcct cgtaacaaga gacggaaaaca tcatgagaca	18960
atcgcgtttg gaaggctttg catcaccttt ggatgatgcg catgaatgga gtcgtctgct	19020
tgctagcctt cgcctaccgc ccactgagtc cgggcggcaa ctaccatcgg cgaacgaccc	19080
agctgacctc taccgaccgg acttgaatgc gctaccttcg tcagcgacga tggccgcgta	19140
cgctggcgac gtgccccgc atgcatggcg gcacatggcg agctcagacc gtgcgtggct	19200
ggctacaaat acgtaccccg tgagtgcctt agctagaaac ttacacctgc aactgcgaga	19260
gcgagcgtgt gagtgtagcc gagtagatcc cccgggctgc aggtcgactc tagaggatcc	19320
accggtcgcc accatggccc tgtccaacaa gttcatcggc gacgacatga agatgacctc	19380
ccacatggac ggctgcgtga acggccacta cttcacctgt aagggcgagg gcagcggcaa	19440

ES 2 843 556 T3

gccctacgag ggcacccaga cctccacctt caaggtgacc atggccaacg gcggccccct 19500
 ggccttctcc ttogacatcc tgtccaccgt gttcatgtac ggcaaccgct gcttcaccgc 19560
 ctaccccacc agcatgcccg actacttcaa gcaggccttc cccgacggca tgcctacga 19620
 gagaaccttc acctacgagg acggcggcgt ggccaccgcc agctgggaga tcagcctgaa 19680
 gggcaactgc ttogagcaca agtccacctt ccacggcgtg aacttccccg ccgacggccc 19740
 cgtgatggcc aagaagacca cggcgtggga cccctccttc gagaagatga ccgtgtgcga 19800
 cggcatcttg aagggcgacg tgaccgcctt cctgatgctg cagggcggcg gcaactacag 19860
 atgccagttc cacacctcct acaagaccaa gaagcccgtg accatgcccc ccaaccacgt 19920
 ggtggagcac cgcacgcaca gaaccgacct ggacaagggc ggcaacacgcg tgcagctgac 19980
 cgagcacgcc gtggccaca tcacctcgtt ggtgcccttc tgaagcgcc catggatatt 20040
 cgaacgcgtc gctgaaatca ccagtctctc tctacaaatc tatctctctc tataataatg 20100
 tgtgagtagt tcccagataa gggcaattagg gttcttatag ggtttcgctc atgtgttgag 20160
 catataagaa acccttagta tgtatttgta tttgtaaaat acttctatca ataaaatttc 20220
 taattcctaa aacaaaaatc cagtggcgag ctcgagccag ggtgcttcct gagctgattc 20280
 cgatgacttc gtaggttcct agctcaagcc gctcgtgtcc aagcgtcact tacgattagc 20340
 taatgattac ggcacatagc accgactagc taactaacta gtacgccgac tagctagcta 20400
 aggtaccgag ctgcaattca ttccgattaa tcgtggcctc ttgctcttca ggatgaagag 20460
 ctatgtttaa acgtgcaagc gctactagac aattcagtac attaaaaacg tccgcaatgt 20520
 gttattaagt tgtctaagcg tcaatttggt tacaccacaa tatatcctgc ca 20572

<210> 74

<211> 26187

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 74

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac 60
 aatctgatca tgagcctagt tagttagggg cagctaacgt cccctagcta actagttgct 120
 attaaccttt agcctagttt agtgcggtcg cgctagtga ctagtcta at gggcgcctag 180
 cttagattag atccccgttc ggcttagact aggttagtcg gatagggcta gtctagtggc 240
 ttagcttgga actgttaagc taggctgtta gtgcgggact tctaggctaa cctattgacc 300
 cctaacgcct agaacctaag aaacgctaag gacacctagc taactttggc gacgtccaga 360
 ggtccgattc cggctaacta actagcttaa gcttgatata gaattcctgc agcccatccc 420
 tcagccgcct ttcactatct tttttgcccg agtcattgtc atgtgaacct tggcatgtat 480

10

ES 2 843 556 T3

aatcgggtgaa ttgcgtcgat tttcctctta taggtgggcc aatgaatccg tgtgatcgcg	540
tctgattggc tagagatatg tttcttcctt gttggatgta ttttcataca taatcatatg	600
catacaaata tttcattaca ctttatagaa atggtcagta ataaacccta tcactatgtc	660
tggtgtttca ttttatttgc ttttaaacga aaattgactt cctgattcaa tatttaagga	720
tcgtcaacgg tgtgcagtta ctaaattctg gttttagga actatagtaa actattcaag	780
tcttcactta ttgtgactc acctctcgcc acatcaccac agatgttatt cacgtcttaa	840
atttgaacta cacatcatat tgacacaata ttttttttaa ataagcgatt aaaacctagc	900
ctctatgtca acaatggtgt acataaccag cgaagttag ggagtaaaaa acatcgctt	960
acacaaagt cgttttaaaa aataaagagt aaattttact ttggaccacc cttcaaccaa	1020
tgtttcactt tagaacgagt aattttatta ttgtcacttt ggaccaccct caaatctttt	1080
ttccatctac atccaattta tcatgtcaaa gaaatggtct acatacagct aaggagattt	1140
atcgacgaat agtagctagc atactcgagg tcattcatat gcttgagaag agagtcggga	1200
tagtccaaaa taaaacaaag gtaagattac ctggtcaaaa gtgaaaacat cagttaaaag	1260
gtggtataaa gtaaaatata ggtaataaaa ggtggcccaa agtgaaattt actcttttct	1320
actattataa aaattgagga tgtttttgtc ggtactttga tacgtcatth ttgtatgaat	1380
tggtttttta gtttatcgcc ttttggaaat gcatactgt atttgagtcg ggttttaagt	1440
tcgtttgctt ttgtaaatac agagggattt gtataagaaa tatcttttaa aaaaccata	1500
tgctaatttg acataatttt tgaaaaaat atatatcag gcgaattctc acaatgaaca	1560
ataataagat taaaatagct ttccccggt gcagcgcag ggtatttttt ctagtaaaaa	1620
taaaagataa acttagactc aaaacattta caaaaacaac ccctaaagt cctaaagccc	1680
aaagtgcata ccacgatcca tagcaagccc agcccaaccc aaccaacccc	1740
agtccagcca actggacaat agtctccaca cccccccact atcaccgtga gttgtccgca	1800
cgacccgcac gtctcgcagc caaaaaaaaa aaaagaaaga aaaaaaagaa aaagaaaaaa	1860
cagcaggtgg gtccgggtcg tggggggccg aaacgcgagg aggatcgca gccagcgacg	1920
aggcgggccc tccctccgct tccaaagaaa cgccccccat cgccactata tacatacccc	1980
cccctctcct cccatcccc caaccctacc accaccacca ccaccacctc cacctcctcc	2040
cccctcgctg ccggacgacg agctcctccc ccctccccct ccgcccgcgc cgcgccggtg	2100
accaccccgc ccctctcctc tttctttctc cgtttttttt ttccgtcacg gtctcgatct	2160
ttggccttgg tagtttggtt gggcgagagg cggcttcgtg cgcgcccaga tcggtgcgcg	2220
ggagggggcg gatctcgcg ctggggctct cgccggcggt gatcaggccc ggatctcgcg	2280
gggaatgggg ctctcggtg tagatctgcg atccgccgtt gttgggggag atgatggggg	2340

ES 2 843 556 T3

gtttaaaatt tccgccatgc taaacaagat caggaagagg ggaaaagggc actatggttt	2400
atatttttat atattttctgc tgcttcgtca ggcttagatg tgctagatct ttctttcttc	2460
tttttggtgg tagaatttga atccctcagc attgttcacg ggtagttttt cttttcatga	2520
tttgtgacaa atgcagcctc gtgcggagct tttttgtagg tagaaggatc catggcggcc	2580
aatgcgggcg gcggtggagc gggaggaggc agcggcagcg gcagcgtggc tgcgccggcg	2640
gtgtgccgcc ccagcggctc gcggtggacg ccgacgccgg agcagatcag gatgctgaag	2700
gagctctact acggtgcggy catccggtcg cccagctcgg agcagatcca gcgcateacc	2760
gccatgctgc ggcagcacgg caagatcgag ggcaagaacg tcttctactg gttccagaac	2820
cacaaggccc gcgagcgcca gaagcgccgc ctcaccagcc tcgacgtcaa cgtgcccgcc	2880
gccggcgcggy ccgacgccac caccagccaa ctcggcgtcc tctcgtctgc gtccgcccg	2940
ccttcaggcg cgggcgctcc ctcgccacc ctcggttct acgcccggy caatggcggy	3000
ggatcggctg tgctgctgga cacgagttcc gactggggca gcagcggcg tgctatggcc	3060
accgagacat gcttctctga ggactacatg ggcgtgacgg acacgggcag ctgctcgag	3120
tgccacgct tctcgtctgc ggacacgata atggcgcggy ccgcgcgcggy ggcggcgag	3180
acgcggggcg ccgagacgct cctctcttc ccgacctgcg gcgacgacgg cggcagcggt	3240
agcagcagct acttgccgtt ctggggtgcc gcgtccacaa ctgccggcg cacttcttc	3300
gttgcatcc aacagcaaca ccagctgcag gagcagtaca gcttttacag caacagcaac	3360
agcaccacgc tggccggcac cggcaaccaa gacgtatcgg caacagcagc agcagccgcc	3420
gccctggagc tgagcctcag ctcatggtgc tccccttacc ctgctgcagg gagtatgtga	3480
acctagactt gtccatcttc tggattggcc aacttaatta atgtatgaaa taaaaggatg	3540
cacacatagt gacatgctaa tcactataat gtgggcatca aagttgtgtg ttatgtgtaa	3600
ttactagtta tctgaataaa agagaaagag atcatccata tttcttatcc taaatgaatg	3660
tcacgtgtct ttataattct ttgatgaacc agatgcattt cattaaccaa atccatatac	3720
atataaatat taatcatata taattaatat caattgggtt agcaaaacaa atctagtcta	3780
ggtgtgtttt gcggcgatcg cgggtaccatt taaattgcgc ccgccacggc cgtggaggtc	3840
gtattccggt cagcttgcat ccctgcagtg cagcgtgacc cggctcgtgcc cctctctaga	3900
gataatgagc attgcatgtc taagttataa aaaattacca catatTTTTT ttgtcacact	3960
tgtttgaagt gcagtttatc tatctttata catatatTTA aactttactc tacgaataat	4020
ataatctata gtactacaat aatatcagtg ttttagagaa tcatataaat gaacagttag	4080
acatggtcta aaggacaatt gagtatTTTg acaacaggac tctacagttt tatctTTTTA	4140
gtgtgcatgt gttctccttt ttttttgcaa atagcttcac ctatataata cttcatccat	4200
tttattagta catccattta gggtttaggg ttaatggttt ttatagacta atTTTTTTtag	4260

ES 2 843 556 T3

tacatctatt ttattctatt ttagcctcta aattaagaaa actaaaactc tatttttagtt	4320
tttttattta ataatttaga tataaaatag aataaaataa agtgactaaa aattaaacia	4380
atacccttta agaaattaaa aaaactaag aaacattttt cttgtttcga gtagataatg	4440
ccagcctggt aaacgcccgtc gacgagctta acggacacca accagcgaac cagcagcgtc	4500
gcgtcggggc aagcgaagca gacggcacgg catctctgtc gctgcctctg gacccctctc	4560
gagagttccg ctccaccgtt ggacttgctc cgctgtcggc atccagaaat tgcgtggcgg	4620
agcggcagac gtgagccggc acggcaggcg gcctcctcct cctctcacgg caccggcagc	4680
tacgggggat tcctttccca ccgtccttc gctttccctt cctcgccgcg cgtaataaat	4740
agacaccccc tccacacct ctttcccaa cctcgtgttg ttgggagcgc acacacacac	4800
aaccagatct ccccaaatc caccgcgtc caccctcgtc tcaaggtagc ccgtcgtcc	4860
tcccccccc cctctctac cttctctaga tcggcggttc ggtccatgca tggtagggc	4920
ccggtagtct tacttctgt catgtttgtg ttagatccgt gtttgtgtta gatccgtgct	4980
gctagcgttc gtacacggat gcgacctga cgtcagacac gttctgattg ctaacttgcc	5040
agtgtttctc tttggggaat cctgggatgg ctctagccgt tccgcagacg ggatcgattt	5100
catgattttt tttgtttcgt tgcataagggt ttggtttgcc cttttccttt atttcaatat	5160
atgcctgca cttgtttgtc gggcatcctt ttcattgcttt tttttgtctt ggttgtgatg	5220
atgtggtctg gttggcggt cgttctagat cggagtagaa ttctgtttca aactacctgg	5280
tggatttatt aattttggat ctgtatgtgt gtgccataca tattcatagt tacgaattga	5340
agatgatgga tggaaatc gatctaggat aggtatacat gttgatgcgg gttttactga	5400
tgcatacata gagatgcttt ttgttcgctt ggttgtgatg atgtggtgtg gttggcggt	5460
cgttcattcg ttctagatcg gagtagaata ctgtttcaaa ctacctggtg tatttattaa	5520
ttttggaact gtatgtgtgt gtcatacatc ttcatagtta cgagtttaag atggatggaa	5580
atatcgatct aggataggta tacatgttga tgtgggtttt actgatgcat atacatgatg	5640
gcatatgcag catctattca tatgctctaa ccttgagtag ctatctatta taataaacia	5700
gtatgtttta taattatttt gatcttgata tacttggtatg atggcatatg cagcagctat	5760
atgtggattt ttttagccct gccttcatac gctatttatt tgcttggtac tgtttctttt	5820
gtcgatgctc accctgttgt ttggtgttac ttctgcaggt cgactctaga ggatccatgg	5880
ccactgtgaa caactggctc gctttctccc tctcccgca ggagctgccg ccctcccaga	5940
cgacggactc cacactcatc tcggccgcca ccgccagca tgtctccggc gatgtctgct	6000
tcaacatccc ccaagattgg agcatgaggg gatcagagct ttccggcgctc gtcgcgagc	6060
cgaagctgga ggacttcctc ggcggcatct ccttctccga gcagcatcac aaggccaact	6120

ES 2 843 556 T3

gcaacatgat acccagcact agcagcacag tttgctacgc gagctcaggt gctagcacccg	6180
gctaccatca ccagctgtac caccagccca ccagctcagc gctccacttc gcggactccg	6240
taatggtggc ttccctcggcc ggtgtccacg acggcgggtgc catgctcagc gcggccgcgcg	6300
ctaacgggtgt cgctggcgct gccagtgcc aacggcgccg catcgggctg tccatgatta	6360
agaactggct gcggagccaa ccggcgccca tgcagccgag ggtggcggcg gctgagggcg	6420
cgcaggggct ctctttgtcc atgaacatgg cggggacgac ccaaggcgct gctggcatgc	6480
cacttctcgc tggagagcgc gcacggggcg ccgagagtgt atcgacgtca gcacagggtg	6540
gagccgtcgt cgtcacggcg ccgaaggagg atagcgggtg cagcgggtgt gccggcgctc	6600
tagtagccgt gagcacggac acgggtggca gcggcgccgc gtcggctgac aacacggcaa	6660
ggaagacggt ggacacgttc gggcagcgca cgtcgattta ccgtggcgctg acaaggcata	6720
gatggactgg gagatatgag gcacatcttt gggataacag ttgcagaagg gaagggcaaa	6780
ctcgtaaagg tcgtcaagtc tatttaggtg gctatgataa agaggagaaa gctgctaggg	6840
cttatgatct tgctgctctg aagtactggg gtgccacaac aacaacaaat tttccagtga	6900
gtaactacga aaaggagctc gaggacatga agcacatgac aaggcaggag tttgtagcgt	6960
ctctgagaag gaagagcagt ggtttctcca gaggtgcac catttacagg ggagtgacta	7020
ggcatcacca acatggaaga tggcaagcac ggattggacg agttgcaggg aacaaggatc	7080
tttacttggg caccttcagc acccaggagg aggcagcggg ggcgtacgac atcgcgccga	7140
tcaagttccg cggcctcaac gccgtcacca acttcgacat gagccgctac gacgtgaaga	7200
gcatectgga cagcagcgcc ctccccatcg gcagcgccgc caagcgctc aaggagccg	7260
aggccgcagc gtccgcgcag caccaccacg ccggcgtggt gagctacgac gtcggccgca	7320
tcgcctcgca gctcggcgac ggcggagccc tggcgccggc gtacggcgcg cactaccacg	7380
gcgcgcctg gccgaccatc gcgttcacg cgggcgcgc cagcacaggc ctgtaccacc	7440
cgtacgcgca gcagccaatg cgcggcgccg ggtggtgcaa gcaggagcag gaccacgcgg	7500
tgatcgcggc cgcgcacagc ctgcaggacc tccaccacct gaacctgggc gcggccggcg	7560
cgcacgactt tttctcggca gggcagcagg ccgccgccgc tgcgatgcac ggcctgggta	7620
gcatcgacag tgcgtcgctc gagcacagca ccggctccaa ctccgtcgct tacaacggcg	7680
gggtcggcga cagcaacggc gccagcgccg tcggcgccag tggcggtggc tacatgatgc	7740
cgatgagcgc tgccggagca accactacat cggcaatggt gagccacgag caggtgcatg	7800
cacgggccta cgacgaagcc aagcaggctg ctcatatggg gtacgagagc tacctggtga	7860
acgcggagaa caatggtggc ggaaggatgt ctgcatgggg gactgtcgtg tctgcagccg	7920
cggcgccagc agcaagcagc aacgacaaca tggccgccga cgtcgggcat ggcggcgccg	7980
agctcttcag tgtctggaac gacacttaag gcgtacgtgc cggcctggct ctccgaaagg	8040

ES 2 843 556 T3

gcgtattcca	gcacactggc	ggccgttact	agacccaacc	tagacttgtc	catcttctgg	8100
attggccaac	ttaattaatg	tatgaaataa	aaggatgcac	acatagtgac	atgctaataca	8160
ctataatgtg	ggcatcaaag	ttgtgtgtta	tgtgtaatta	ctagttatct	gaataaaaga	8220
gaaagagatc	atccatattt	cttatcctaa	atgaatgtca	cgtgtcctta	taattctttg	8280
atgaaccaga	tgcatttcat	taaccaaadc	catatacata	taaatattaa	tcatatataa	8340
ttaatatcaa	ttgggttagc	aaaacaaadc	tagtctaggt	gtgttttgcg	aatgcggcct	8400
ccggattctt	atgtgcttct	agtctccaaa	tgtggttgat	agttattttg	ctctaagatc	8460
aacagtaatg	aagtataaat	catcgttggt	gtgtgctact	cgggttaattg	agcattaaca	8520
cacacaaaca	tgacgaggat	ggtataatct	ccaaaaatgt	gtactttggt	aggtgggacc	8580
ctatagcctt	gattaatgtg	ctatggttagg	catgcctgga	aacgtgtgac	gcatatgttt	8640
tgtgaacctg	ttgatattat	atgtgctttt	atattaccat	attttattaa	aataactaata	8700
tttattacta	gtaagatata	acattctatc	tagcttaaaa	actaaccata	aatattccat	8760
aataactaga	tttaccaaac	taatatacta	aatatacata	ataaatacaa	aattaacaag	8820
acaataatca	atattttatga	gcttaatata	tttagacatt	atggttggtc	gacgataatc	8880
atgctaactt	ttcgtaatgt	cttgattgaa	atatgcttag	aataatgcct	ctttgttcta	8940
catggcaaat	agggaccatt	atggtgtaac	accctgggaa	ccacaaacac	cccgaatgc	9000
tactaaacta	cacaactaac	cttcatatat	aaaatttcga	cagcatctcc	tttgaaaatt	9060
tgcatagacg	tggaagcaac	agagtataaa	cagatatcat	gataagaaaa	catactagac	9120
attaataatc	tgctagaaat	gggaagaatc	acgcgtgaagc	ttgcatgcct	gcagtgcagc	9180
gtgacccggt	cgtgcccctc	tctagagata	atgagcattg	catgtctaag	ttataaaaaa	9240
ttaccacata	ttttttttgt	cacacttggt	tgaagtgcag	tttatctatc	tttatacata	9300
tattttaact	ttactctacg	aataatataa	tctatagtac	tacaataata	tcagtgtttt	9360
agagaatcat	ataaatgaac	agttagacat	ggtctaaagg	acaattgagt	attttgacaa	9420
caggactcta	cagttttatc	tttttagtgt	gcatgtgttc	tccttttttt	ttgcaaatag	9480
cttcacctat	ataaacttcc	atccatttta	ttagtacatc	catttagggg	ttaggggttaa	9540
tggtttttat	agactaattt	tttttagtaca	tctattttat	tctatttttag	cctctaaatt	9600
aagaaaacta	aaactctatt	ttagtttttt	tatttaataa	tttagatata	aaatagaata	9660
aaataaagtg	actaaaaatt	aaacaaatac	cctttaagaa	attaaaaaaa	ctaaggaaac	9720
atttttcttg	tttcgagtag	ataatgccag	cctgttaaac	gccgtcgacg	agtctaacgg	9780
acaccaacca	gcgaaccagc	agcgtcgcg	cgggccaagc	gaagcagacg	gcacggcatc	9840
tctgtcgctg	cctctggacc	cctctcgaga	gttccgctcc	accgttggac	ttgtccgct	9900

ES 2 843 556 T3

gtcggcatcc	agaaattgcg	tggcggagcg	gcagacgtga	gccggcacgg	caggcggcct	9960
cctcctcctc	tcacggcacc	ggcagctacg	ggggattcct	ttcccacgcg	tccttcgctt	10020
tcctctcctc	gcccgccgta	ataaatagac	acccctccca	cacctctttt	ccccaacctc	10080
gtgttggtcg	gagcgcacac	acacacaacc	agatctcccc	caaatccacc	cgtcggcacc	10140
tcctgttcaa	ggtagcgccg	tcgtcctccc	ccccccccct	ctctaccttc	tctagatcgg	10200
cgttcgggtc	catgcatggt	tagggcccg	tagttctact	tctgttcatg	tttgtgttag	10260
atccgtgttt	gtgttagatc	cgtgctgcta	gcgttcgtac	acggatgcga	cctgtacgtc	10320
agacacgttc	tgattgctaa	cttgccagtg	tttctctttg	gggaatcctg	ggatggctct	10380
agccgttccg	cagacgggat	cgatttcatg	atTTTTTTTT	tttcgttgca	tagggtttgg	10440
tttgcccttt	tcctttattt	caatatatgc	cgtgcacttg	tttgtcgggt	catcttttca	10500
tgcTTTTTTT	tgtcttggtt	gtgatgatgt	ggtctgggtg	ggcggtcggt	ctagatcgga	10560
gtagaattct	gtttcaaact	acctgggtga	tttattaatt	ttggatctgt	atgtgtgtgc	10620
catacatatt	catagttacg	aattgaagat	gatggatgga	aatatcgatc	taggataggt	10680
atacatgttg	atgcgggttt	tactgatgca	tatacagaga	tgctttttgt	tcgcttggtt	10740
gtgatgatgt	ggtgtgggtg	ggcggtcggt	cattcgttct	agatcggagt	agaatactgt	10800
ttcaaactac	ctggtgtatt	tattaatTTT	ggaactgtat	gtgtgtgtca	tacatcttca	10860
tagttacgag	tttaagatgg	atggaaatat	cgatctagga	taggtataca	tgttgatgtg	10920
ggttttactg	atgcatatac	atgatggcat	atgcagcatc	tattcatatg	ctctaacctt	10980
gagtacctat	ctattataat	aaacaagtat	gttttataat	tattttgatc	ttgatatact	11040
tggtgatgtg	catatgcagc	agctatatgt	ggattTTTTT	agccctgcct	tcatacgcta	11100
tttatttgct	tggtactgtt	tcttttgtcg	atgctcacc	tgttgtttgg	tgttacttct	11160
gcaggtcgac	tttaacttag	cctaggatcc	aacaatgcc	cagttcgaca	tcctctgcaa	11220
gacccccccc	aagggtgctg	tgaggcagtt	cgtggagagg	ttcgagaggc	cctccggcga	11280
gaagatcgcc	ctctgcgccg	ccgagctcac	ctacctctgc	tggtgatgca	cccacaacgg	11340
caccgccatt	aagagggcca	ccttcatgtc	atacaacacc	atcatctcca	actccctctc	11400
cttcgacatc	gtgaacaagt	ccctccagtt	caaatacaag	accagaagg	ccaccatcct	11460
cgaggcctcc	ctcaagaagc	tcatccccgc	ctgggagttc	accatcatcc	cctactacgg	11520
ccagaagcac	cagtccgaca	tcaccgacat	cgtgtcatcc	ctccagcttc	agttcgagtc	11580
ctccgaggag	gctgacaagg	gcaactccca	ctccaagaag	atgctgaagg	ccctcctctc	11640
cgagggcgag	tccatctggg	agatcaccga	gaagatcctc	aactccttcg	agtacacctc	11700
caggttcact	aagaccaaga	ccctctacca	gttcctcttc	ctcgccacct	tcatacaactg	11760
cggcagggtc	tcagacatca	agaacgtgga	ccccaaagtcc	ttcaagctcg	tgcagaacaa	11820

ES 2 843 556 T3

gtacctaggt ttgtttctgc ttctaccttt gatatatata taataattat cattaattag	11880
tagtaatata atatttcaaa tatttttttc aaaataaaaag aatgtagtat atagcaattg	11940
cttttctgta gtttataagt gtgtatatatt taatttataa cttttctaata atatgaccaa	12000
aacatgggtga tgcctaggtg tcatcatcca gtgcctcgtg accgagacca agacctccgt	12060
gtccaggcac atctacttct tctccgctcg cggcaggatc gacccctcg tgtacctcga	12120
cgagttcctc aggaactcag agcccgctgct caagaggggtg aacaggaccg gcaactcctc	12180
ctccaacaag caggagtacc agctcctcaa ggacaacctc gtgaggtcct acaacaaggc	12240
cctcaagaag aacgccccct actccatctt cgccatcaag aacggcccca agtcccacat	12300
cgttaggcac ctcatgacct ccttcctctc aatgaagggc ctccaccgagc tcaccaacgt	12360
ggtgggcaac tgggtccgaca agagggcctc cgccgtggcc aggaccacct acaccacca	12420
gatcaccgcc atccccgacc actacttcgc cctcgtgtca aggtactacg cctacgaccc	12480
catctccaag gagatgatcg ccctcaagga cgagactaac cccatcgagg agtggcagca	12540
catcgagcag ctcaagggct ccgcccagggt ctccatcagg taccctgcct ggaacggcat	12600
catctccag gaggtgctcg actacctctc ctccctacatc aacaggagga tctgagttaa	12660
cctagacttg tccatcttct ggattggcca acttaattaa tgtatgaaat aaaaggatgc	12720
acacatagtg acatgctaata cactataatg tgggcatcaa agttgtgtgt tatgtgtaat	12780
tactagttat ctgaataaaa gagaaagaga tcatccatat ttcttatcct aaatgaatgt	12840
cacgtgtctt tataattctt tgatgaacca gatgcatttc attaaccaa tccatataca	12900
tataaatatt aatcatatat aattaatatc aattgggtta gcaaaacaaa tctagtctag	12960
gtgtgttttg cggtcacacc ggttaaaacc aaaatccagt ggcgagctct cgagtcgatc	13020
gctatcaact ttgtatagaa aagttgggccc gaattcgagc tcggtacggc cagaatggcc	13080
cggaccgggt taccgaattc gagctcggta ccctgggatc agcttcgctg aaatcaccag	13140
tctctctcta caaatctatc tctctctata ataatgtgtg agtagttccc agataaggga	13200
attaggggtc ttataggggt tcgctcatgt gttgagcata taagaaaccc ttagtatgta	13260
tttgatatttg taaaatactt ctatcaataa aatttctaata tcctaaaacc aaaatccagt	13320
ggcgagctgc tagcgaagtt cctattccga agttcctatt ctctagaaag tataggaact	13380
tcagatccac cgggatcccc gatcatgcaa aaactcatta actcagtgca aaactatgcc	13440
tggggcagca aaacggcgtt gactgaactt tatggtatgg aaaatccgtc cagccagccg	13500
atggccgagc tgtggatggg cgcacatccg aaaagcagtt cacgagtgca gaatgccgcc	13560
ggagatatcg tttcactcg tgatgtgatt gagagtgata aatcgactct gctcggagag	13620
gccgttgcca aacgctttgg cgaactgcct ttctgttca aagtattatg cgcagcacag	13680

ES 2 843 556 T3

ccactctcca	ttcagggttca	tccaaacaaa	cacaattctg	aatcggttt	tgccaaagaa	13740
aatgccgcag	gtatcccgat	ggatgccgcc	gagcgtaact	ataaagatcc	taaccacaag	13800
ccggagctgg	tttttgcgct	gacgcctttc	cttgcgatga	acgcgtttcg	tgaattttcc	13860
gagattgtct	ccctactcca	gccggtcgca	ggtgcacatc	cggcgattgc	tcaactttta	13920
caacagcctg	atgccgaacg	tttaagcgaa	ctgttcgccca	gcctgttgaa	tatgcagggt	13980
gaagaaaaat	cccgcgcgct	ggcgatttta	aaatcggtcc	tcgatagcca	gcagggtgaa	14040
ccgtggcaaa	cgattcggtt	aattttctgaa	ttttaccggg	aagacagcgg	tctgtttctc	14100
ccgctattgc	tgaatgtggt	gaaattgaac	cctggcggaag	cgatgttctt	gttcgctgaa	14160
acaccgcacg	cttacctgca	aggcgtggcg	ctggaagtga	tggcaaaactc	cgataacgtg	14220
ctgcgtgcgg	gtctgacgcc	taaatacatt	gatattccgg	aactggttgc	caatgtgaaa	14280
ttcgaagcca	aaccggctaa	ccagttgttg	accagccgg	tgaacaagg	tgcaagaactg	14340
gacttccoga	ttccagtgga	tgattttgcc	ttctcgctgc	atgaccttag	tgataaagaa	14400
accaccatta	gccagcagag	tgccgccatt	ttgttctgcg	tcgaaggcga	tgcaacgttg	14460
tggaaagggt	ctcagcagtt	acagcttaaa	ccgggtgaat	cagcgtttat	tgccgccaac	14520
gaatcaccgg	tgactgtcaa	aggccacggc	cgtttagcgc	gtgtttacaa	caagctgtaa	14580
gagcttactg	aaaaaattaa	catctcttgc	taagctgggg	gtggaacctc	gacttgtcca	14640
tcttctggat	tggccaactt	aattaatgta	tgaataaaaa	ggatgcacac	atagtacat	14700
gctaatact	ataatgtggg	catcaaagtt	gtgtgttatg	tgtaattact	agttatctga	14760
ataaaagaga	aagagatcat	ccatatttct	tatcctaaat	gaatgtcacg	tgtctttata	14820
attctttgat	gaaccagatg	catttcatta	accaaatacca	tatacatata	aatattaatc	14880
atatataatt	aatatcaatt	gggttagcaa	aacaaatcta	gtctagggtg	gttttgcgaa	14940
tgcgaccttc	ttatgtgctt	ctagtctcca	aatgtgggtg	atagttatct	tgctctaaga	15000
tcaacagtaa	tgaagtataa	atcatcggtt	tgggtgtgcta	ctcggttaat	tgagcattaa	15060
cacacacaaa	catgacgagg	atggtataat	ctccaaaaat	gtgtactttg	ttaggtggga	15120
ccctatagcc	ttgattaatg	tgctatgtta	ggcatgcctg	gaaacgtgtg	acgcatatgt	15180
tttgtgaacc	tggtgatatt	atatgtgctt	ttatattacc	atattttatt	aaaatactaa	15240
tattttattac	tagtaagata	taacattcta	tctagcttaa	aaactaacca	taaatattcc	15300
ataataacta	gatttaccaa	actaatatac	taaatataca	taataaatac	aaaattaaca	15360
agacaataat	caatatttat	gagcttaata	tatttagaca	ttatggttgg	tcgacgataa	15420
tcatagctaac	ttttcgtaat	tgcttgattg	aaatatgctt	agaataatgc	ctctttgttc	15480
tacatggcaa	atagggacca	ttatggtgta	acaccctggg	aaccacaaac	accccgaaat	15540
gctactaaac	tacacaacta	accttcatat	ataaaaatttc	gacagcatct	cctttgaaaa	15600

ES 2 843 556 T3

tttgcataga cgtggaagca acagagtata aacagatatc atgataagaa aacatactag	15660
acattaataa tctgctagaa atgggaagaa tcctaacttg acgactgcgt aactgactag	15720
agtcacactt agctgacctt agtcacttac aactgacttc gtgtcctagg cttaggctac	15780
tgctagtccg cgggtgatcc gtgatcgagt tggcgccaga cggaatctgt tctccatcgc	15840
tgacatcctc gagtagatca cattcaagct tgatatcgaa ttcttgacgc ccatccctca	15900
gccgcctttc actatctttt ttgcccgagt cattgtcatg tgaaccttgg catgtataat	15960
cggatgaattg cgtcgatttt cctcttatag gtgggccaat gaatccgtgt gatcgcgtct	16020
gattggctag agatatgttt ctctcttggt ggatgtattt tcatacataa tcatatgcat	16080
acaaatattt cattacactt tatagaaatg gtcagtaata aacctatca ctatgtctgg	16140
tgtttcattt tatttgcttt taaacgaaaa ttgacttcct gattcaatat ttaaggatcg	16200
tcaacggtgt gcagttacta aattctggtt tgtaggaact atagtaaact attcaagtct	16260
tcacttattg tgcaactcacc tctcgccaca tcaccacaga tgttattcac gtcttaaatt	16320
tgaactacac atcatattga cacaatattt tttttaaata agcgattaaa acctagcctc	16380
tatgtcaaca atggtgtaca taaccagcga agtttaggga gtaaaaaaca tcgccttaca	16440
caaagtctgc tttaaaaaat aaagagtaaa ttttactttg gaccaccctt caaccaatgt	16500
ttcacttttag aacgagtaat tttattattg tcactttgga ccaccctcaa atcttttttc	16560
catctacatc caatttatca tgtcaaagaa atggtctaca tacagctaag gagatttatc	16620
gacgaatagt agctagcata ctcgaggtca ttcatatgct tgagaagaga gtcgggtag	16680
tccaaaataa aacaaaggtg agattacctg gtcaaaagtg aaaacatcag ttaaaaggtg	16740
gtataaagta aaatatcggg aataaaaggt ggcccaaagt gaaatttact cttttctact	16800
attataaaaa ttgaggatgt ttttgcgggt actttgatac gtcatttttg tatgaattgg	16860
tttttaagtt tattcgcttt tggaaatgca tatctgtatt tgagtcgggt ttttaagttcg	16920
tttgcttttg taaatacaga gggatttgta taagaaatat ctttaaaaaa acctatatgc	16980
taatttgaca taatttttga gaaaaatata tattcaggcg aattctcaca atgaacaata	17040
ataagattaa aatagctttc cccggttgca gcgcatgggt attttttcta gtaaaaaata	17100
aagataaact tagactcaaa acatttacaa aaacaacccc taaagttcct aaagcccaa	17160
gtgctatcca cgatccatag caagccagc ccaacccaac ccaacccaac ccacccagc	17220
ccagccaact ggacaatagt ctccacaccc cccactatc accgtgagtt gtccgcacgc	17280
accgcacgtc tcgcagccaa aaaaaaaaaa agaaagaaaa aaaagaaaaa gaaaaaacag	17340
caggtgggtc cgggtcgtgg gggccggaaa cgcgaggagg atcgcgagcc agcgacgagg	17400
ccggccctcc ctccgcttc aaagaaacgc ccccatcgc cactatatac ataccccccc	17460

ctctcctccc atccccccaa ccctaccacc accaccacca ccacctccac ctctcctccc	17520
ctcgctgccg gacgacgagc tcctccccc tccccctccg ccgccgcgc gccggtaacc	17580
accccgcccc tctcctcttt cttctctcgt tttttttttc cgtcacggtc tcgatctttg	17640
gccttggtag tttgggtggg cgagaggcgg cttcgtgcgc gccagatcg gtgcgcggga	17700
ggggcgggat ctgcgggctg gggctctcgc cggcgtggat caggcccgga tctcgcgggg	17760
aatggggctc tcgatgtag atctgcgatc cggcgttgtt gggggagatg atgggggggt	17820
taaaatttcc gccatgctaa acaagatcag gaagagggga aaagggcact atggtttata	17880
tttttatata tttctgctgc ttcgtcaggc ttagatgtgc tagatctttc tttcttcttt	17940
tttggggtag aatttgaatc cctcagcatt gttcatcggg agtttttctt ttcattgattt	18000
gtgacaaatg cagcctcgtg cggagctttt ttgtaggtag aaggatccac acgacaccat	18060
gtcccccgag cgcgcgcccg tcgagatccg cccgccacc gccgcgcgaca tggccgcggt	18120
gtgcgacatc gtgaaccact acatcgagac ctccaccgtg aacttccgca ccgagccgca	18180
gaccccgag gagtggatcg acgacctgga gcgcctccag gaccgctacc cgtggctcgt	18240
ggccgaggtg gagggcgtgg tggccggcat cgcctacgcc ggcccggtga aggcccgcaa	18300
cgcctacgac tggaccgtgg agtccaccgt gtacgtgtcc caccgccacc agcgcctcgg	18360
cctcggctcc accctctaca cccacctcct caagagcatg gagggccagg gcttcaagtc	18420
cgtggtggcc gtgatcggcc tcccgaacga cccgtccgtg cgcctccacg aggccctcgg	18480
ctacaccgcc cgcggcacc tgcgcgcgc cggctacaag cacggcggct ggcacgacgt	18540
cggcttctgg cagcgcgact tcgagctgcc ggccccgcg cgcgcggtgc gcccggtgac	18600
gcagatctga gtgcacctgc aggcattgcc ctgaaatcac cagtctctct ctacaaatct	18660
atctctctct ataataatgt gtgagtagtt ccagataag ggaattaggg ttcttatagg	18720
gtttcgctca tgtgttgagc atataagaaa cccttagtat gtatttgtat ttgtaaaata	18780
cttctatcaa taaaatttct aattcctaaa accaaaatcc agtggcgagc taatgcggcc	18840
cgaataactt cgtatagcat acattatacg aagttatacc tgggtggcgc gctaggggct	18900
gcaggaattc ctgcagcccg ggggatccac tagttctaga gcggccgacc tcgacagatc	18960
taagcttact agtgccgtgg gtcgtttaag ctgccgctgt acctgtgtcg tctggtgcct	19020
tctggtgtac ctgggaggtt gtcgtctatc aagtatctgt ggttggtgtc atgagtcagt	19080
gagtcccaat actgttcgtg tcctgtgtgc attataccca aaactgttat gggcaaatca	19140
tgaataagct tgatgttcga acttaaaagt ctctgctcaa tatggtatta tggttgtttt	19200
tgttcgtctc ctaatatattg cctgggatca aattttattg gctggtgttc atttgacctc	19260
catgttcttg ctaggctcca ttttttactc tacagccata atatgtttga ttgtttggtt	19320
tgttctttgt tgtacacctg gttctgtcga gcttagtttt cgacactggc ttacagetta	19380

ES 2 843 556 T3

acatgttgct attttattgg gttctgattg ctattttatt gggttctgat tgctagtttt	19440
tgctgaatcc aaaaaccatg ttattttattt aagcgatcca ggttattatt atgatgggtg	19500
ctaagttttt ttttttccaa gggtaaattt tctggattct ccagtgtttc tgtggccgaa	19560
ttcactagtg attcagatct gatatcgatg ggccactaa ctatctatac tgtaataatg	19620
ttgtatagcc gccgtagatg tagctagttt agtcattcag cggcgatggg taataataaa	19680
gtgtcatcca tccatcacca tgggtggcaa cgtgagcaat gacctgattg aacaaattga	19740
aatgaaaaga agaaatatgt tatatgtcaa cgagatttcc tcataatgcc actgacgacg	19800
tgtgtccaag aaatgtatca gtgatacgta tattcacaat ttttttatga cttatactca	19860
caatttgttt ttttactact tatactcaca atttgttgtg ggtaccataa caatttcgat	19920
cgaatatata tcagaaagt gacgaaagta agctcactca aaaagttaaa tgggctgctg	19980
aagctgctgc aggcccaagt tttggctatt ctatccggta tccacgattt tgatggctga	20040
gggacatatg ttcgcttaag cttggtcacc cggctcgggc ctagaaggcc agcttcaagt	20100
ttgtacaaaa aagcaggctc cggccagaat ggcccgacc gggtgaccgg cgcgccaagc	20160
ttggtgacgg cgacgcgatc gaacagggtg tgatcgatgc tgcaacgtgt gtaaatatac	20220
agcgccggct gggcaagag atggctcggg tgacgcgcgc gcggcgtgct ctggcgttg	20280
cgcgggggca ttctttagtt tttcatcttt tcatcatctc agatggtaga tacaaaacag	20340
tgtatgtatg tagctctgtt tctctctata gaacccaac aaattttgtt gttgatgttg	20400
tttatcttca tatgctttga tcttgaaatc gtctacctta ctactgccga tcgttgctca	20460
aaagtgtgga agtttgaagc atcttcacg ggcgttgctc tactcccatc tcgtttcacc	20520
acgaaatccc cttcgcatag aggcattctca gccgtacatc tccaaacca cgttctggat	20580
tcagcaaacg gacggcatat gccagcggac tgactaggat ttgtggttca gaataaaaat	20640
atggtttgca gtgtcaattt tccaggagag tgactatgtc atactaactt catactccaa	20700
taatgaagct aagctatctc catgtacata tcaaatacga aatcatatcc aagggaacta	20760
acacagtcac aaacaacagg tacacagaca attacagcac aagcgcagga gggaaataat	20820
tttaactgaa ctaggaagaa aggaaacaca actcattttt tattgatata tgttggatga	20880
atccaataaa accgatataa gtcacgaaaa atcagactag atgaatcctt cgagatcaca	20940
tgagcaaac ctttcgacga aagctgtcct atagtcgtgg aagcaataac acttgataaa	21000
gataggaatt cagacacgag aggttgaga ctataagatg tcaggtacct acgcgttcga	21060
atcaatgaga tgatgcagta gaggtagagc gaagagggaa ggctgagatc gacgcattcac	21120
tgccatgtcc ctgatgatga tgaggtagca atggagatct acagagcggg caattgaagt	21180
tgagatgttg aagccagcct tccaaacact gcttatggaa gacgtgtctg caatctagct	21240

ES 2 843 556 T3

tcctcacttc ttctccggtc ttgagtttag acagacacac gatgcagtca gaggctgcgt	21300
tgctggagta gcggtacgag aagaggcgat ttaggttgag ctggtcggcg aggacgctga	21360
ggtttgaagt aacaacaaca acgggggcag aagaaggga gaggagaaga gaccggacgt	21420
gtctgaagaa agttgcgagg agagccagta gcatcagtgg gatcgaatct gatgacacgt	21480
cggagagctg accttgtagt cccatggatc ctaggctaag ttaaagtcga cctgcagaag	21540
taacaccaaa caacaggggtg agcatcgaca aaagaaacag taccaagcaa ataaatagcg	21600
tatgaaggca gggctaaaaa aatccacata tagctgctgc atatgccatc atccaagtat	21660
atcaagatca aaataattat aaaacatact tgtttattat aatagatagg tactcaaggt	21720
tagagcatat gaatagatgc tgcatatgcc atcatgtata tgcatacagta aaaccacat	21780
caacatgtat acctatccta gatcgatatt tccatccatc ttaaactcgt aactatgaag	21840
atgtatgaca cacacataca gttccaaaat taataaatac accaggtagt ttgaaacagt	21900
attctactcc gatctagaac gaatgaacga ccgccaacc acaccacatc atcacaacca	21960
agcgaacaaa aagcatctct gtatatgcat cagtaaaacc cgcataca tgtataccta	22020
tcctagatcg atatttccat ccatcatctt caattcgtaa ctatgaatat gtatggcaca	22080
cacatacaga tccaaaatta ataaatccac caggtagttt gaaacagaat tctactccga	22140
tctagaacga ccgccaacc agaccacatc atcacaacca agacaaaaa aagcatgaaa	22200
agatgacccg acaaacaagt gcacggcata tattgaaata aaggaaaagg gcaaaccaaa	22260
ccctatgcaa cgaacaaaaa aaaatcatga aatcgatccc gtctgcggaa cggctagagc	22320
catcccagga ttcccaaaag agaaacactg gcaagttagc aatcagaacg tgtctgacgt	22380
acaggtcgca tccgtgtacg aacgctagca gcacggatct aacacaaaca cggatctaac	22440
acaaacatga acagaagtag aactaccggg ccctaaccat gcatggaccg gaacgccgat	22500
ctagagaagg tagagagggg ggggggggag gacgagcggc gtaccttgaa gcggaggtgc	22560
cgacgggtgg atttggggga gatcctcgct ttctggagg aggcgtggag cgcacgcggt	22620
tggaggcgaa ggcggatgaa aaacgagagg ggctgacggg ctgtcttctg tctggcttgc	22680
gccctccga cgggttggtt ggctttatag gcgtgtggtg catataatta ggtgcgcgaa	22740
ccgccgttcc gatggaaaaa actggctatg gaggtctctg attcgtgccg gcctatcgag	22800
acggagcttt ttttattatt tttttcatc tcgcatgtat gttgtttgtt cgctgatgga	22860
atttctgttt gttcgcgctg aaaacattgg acagggttcg tttatgctgt gccagtgag	22920
gaaccataaa aggaacgctg tgcatatgga gtcacgtact tagtaaaaaa tccttcgctc	22980
tggattaaac taggttgatt atactgttgc acgtagtatc tatagagaca aaaatattat	23040
aaaaatagta gaaaccgata caaattaaag ttaagtaaac atgatgttta tctgatactc	23100
taagtcaagc aagtcatttc tagtgttttc ttctgtgcta ctagcagagc aagagattgg	23160

gacccgaggc	ggggggccatg	cagacatagc	gggatgctcc	ttaggagaga	gataaaaata	23220
aataatagaa	tatatgatta	ttggtataga	gtacagattt	aaggtaaata	ttgcttcaaa	23280
agattgtgga	atagagtaga	gaatctcgtt	gacaaaaatg	aaatattact	tttagaataa	23340
aaatttagag	ttgcataagt	atagatatca	ttagtttgga	ttgagggcta	gagctattcc	23400
agctatttag	gctctgttta	taatcctaca	gccataaatt	accagaagct	tcggtccggg	23460
cctagaaggc	cgatctcccg	ggcaccacgc	tttcttgtac	aaagtggccg	ttaacggatc	23520
ggccagaatg	gcccggaccg	ggttaccgaa	ttgctaacta	actaggagct	ccctttaatc	23580
tggcgcttga	tctgcacccg	cggttgcaa	agataaatgg	cacatttagt	gtgttatttt	23640
gcaatacctt	tcatagtaga	tatccttaaa	tgcagtttta	ggcatgtttg	ggtaattaaa	23700
taacattttt	aggaggagtt	ttagatttac	ctttctttcg	tgatgactga	tgacagacgt	23760
ggggaattca	aatgcaactc	tagcgaaagt	tcatatatat	ttcataaata	gctgaggctg	23820
gggtaattat	tttttgtaga	aaaatagaat	aggtggaatg	gtttggggaa	ggcgtaggcg	23880
ctcgtggacg	acgccccata	aaagacaaga	ggcggaattg	ccatgaattc	gaggtagcta	23940
agtaaggcgc	atatatatgc	caaaaattct	actgtcactt	tccaatttca	atgcgctgcc	24000
aaacaagccg	ccgacccttg	gatactgact	tgaattcagc	ccaattctgt	agatccaaac	24060
agggccggcg	tcagtgcctc	aggtgagaga	gcagcagacg	atgcaaagag	ccaaaagtgg	24120
aagcagacgc	agccgaagcc	gaagcccaag	cccaaaactg	ttttgtcttt	gcccagaacc	24180
gcgacgagcc	taaactgccg	cttcctccta	tctacaagtc	cctggcacat	cacgcatagt	24240
ccaaccatcg	cgcgccaggc	ataaggcgcg	ccacggggac	gcgacatgtg	gtggcgggacg	24300
cgatcaggat	agggccaggc	tggccggggc	cggccacggg	agaacggtgg	ccactcgtcc	24360
cacatccgct	tcgtcctgtc	ctgtactgcg	tcctgcccc	aacgagagcc	ggagccggcc	24420
atcccgtcgc	acactctccc	cctctatata	tgccgtcggg	gtgggggagc	ctactacagg	24480
acgacccaag	caagcaagca	agcagcgagt	acatacatac	taggcagcca	ggcagccatg	24540
atgctgcact	gcacatttgc	tatatctgag	gctcctgcgc	gcgccttggc	ccttggccag	24600
gtgtctgtca	tgcgggcgat	gccgcaggaa	gaagaagccg	cgggtggctac	gacgaccatg	24660
gccgggggca	aggtggcgcc	gctgctggcc	acggcgggcc	cgtgctgct	gctgctcccg	24720
ctggcgctgc	cgccgctgcc	gccgccgccc	acgcagctgt	tgctcgtccc	cgtggtcttg	24780
ctgctcctcg	tggcgctcct	cgcgttctgc	cccgcgcga	cctcctcgcc	gtcgccgatg	24840
catgccgccg	accacgggtc	gttcgggacc	actggatcac	cgcacctatg	ttgaggtacc	24900
cggggatcct	ctagtgcctg	gggtcgttta	agctgccgct	gtacctgtgt	cgtctggtgc	24960
cttctggtgt	acctgggagg	ttgtcgtcta	tcaagtatct	gtggttggtg	tcagagtc	25020

gtgagtccca atactgttcg tgcctgtgt gcattatacc caaaactggt atgggcaaatt 25080
catgaataag cttgatgttc gaacttaaaa gtctctgctc aatatggtat tatggttggt 25140
tttgttcgtc tcctaataatt tgcctgggat caaatTTTTat tggctggtgt tcatttgacc 25200
tccatgttct tgctaggctc cattttttac tctacagcca taatatgttt gattgtttgg 25260
tttgttcttt gttgtacacc tggttctgtc gagcttagtt ttcgacactg gcttacagct 25320
taacatgttg ctatttttatt gggttctgat tgctatttta ttgggttctg attgctagtt 25380
tttgctgaat ccaaaaacca tgttatttat ttaagcgatc caggttatta ttatgatggt 25440
ggctaagttt ttttttttcc aagggtaaat tttctggatt ctccagtgtt tctgtggccg 25500
aattcaatca ctagagtcga cctgcaggca tgcccgcggt ggagctcgaa ttccggtccg 25560
ggcctagaag gccagcttcg gccgccccgg gcaactttat tatacaaagt tgatagatcg 25620
agcgactaat taactagctg tcccacggcc taactagcac ttaatccctc agcctaacct 25680
aagagcgcta atctaggcta gtggtcactt agggctttaa ggctagcgta tacgaagttc 25740
ctattccgaa gttcctattc tccagaaagt ataggaactt ctgtacacct gagcctaact 25800
aactaggacg tcccagagtc cgattccggg ctaattaaca cctagctcgc cactcgacta 25860
attagggact aagcctagcg cttagccgtt aaaacctagc acaccctaag cacccttagt 25920
taggttcccc tcttaattaa gccctagtga gccctaagt taaggggacg ctaagagccc 25980
cctaacctag tattcggcta gagcggaact aggctaaaca cctaagcgca cctctttaag 26040
ctagatcgct agggggctag gctagagctt cgctagatta gtctaagggc agctaactaa 26100
ctaggtttta actacattaa aaacgtccgc aatgtgttat taagttgtct aagcgtcaat 26160
ttgtttacac cacaatatat cctgcc 26187

<210> 75

<211> 9982

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 75

gtttacccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac 60
aatctgatca tgagcggaga attaaggag tgacgttatg acccccgccg atgacgctgg 120
acaagccgtt ttacgttttg aactgacaga accgcaacgt tgaaggagcc actcagcaag 180
ctggtacgat tgtaatacga ctactatag ggcgaattga ggcgtgttta aacgtctctc 240
aactggaaga gcggttacta ccggttcact agctagctgc taaggttacc agagctggtc 300
acctttgtcc accaacttat taagtatcta gttgaagaca cgttcttctt cacgtaagaa 360
gacactcagt agtcttcggc cagaatggcc tcttgattca gcgggcctag aaggccggat 420

10

ES 2 843 556 T3

cactgactag ctaattttaa tcctgaggat atcgctatca actttgtata gaaaagttga	480
agcttcgctg aaatcaccag tctctctcta caaatctatc tctctctata ataagtgtg	540
agtagttccc agataaggga attaggggtt ttatagggtt tcgctcatgt gttgagcata	600
taagaaaccc ttagtatgta tttgtatttg taaaatactt ctatcaataa aatttcta	660
tcctaaaacc aaaatccagt ggcgagctgc tagcgaagtt cctattccga agttcctatt	720
ctctagaaag tataggaact tcagatccta gaggatcgat ccccgatcat gcaaaaactc	780
attaactcag tgcaaaacta tgccctggggc agcaaaacgg cgttgactga actttatggt	840
atggaaaatc cgtccagcca gccgatggcc gagctgtgga tgggcgacaca tccgaaaagc	900
agttcacgag tgcagaatgc cgccggagat atcgtttcac tgcgtgatgt gattgagagt	960
gataaatcga ctctgctcgg agaggccgtt gccaaacgct ttggcgaaact gcctttcctg	1020
ttcaaagtat tatgcgcagc acagccactc tccattcagg ttcatccaaa caaacacaat	1080
tctgaaatcg gttttgccaa agaaaatgcc gcaggtatcc cgatggatgc cgccgagcgt	1140
aactataaag atcctaacca caagccggag ctggtttttg cgctgacgcc tttccttgcg	1200
atgaacgcgt ttcgtgaatt ttccgagatt gtctccctac tccagccggt cgcaggtgca	1260
catccggcga ttgctcactt tttacaacag cctgatgccg aacgtttaag cgaactgttc	1320
gccagcctgt tgaatatgca gggatgaaga aaatcccgcg cgctggcgat tttaaaatcg	1380
gccctcgata gccagcaggg tgaaccgtgg caaacgattc gtttaatttc tgaattttac	1440
ccggaagaca gcggtctgtt ctccccgcta ttgctgaatg tggatgaaatt gaaccctggc	1500
gaagcgatgt tcctgttcgc tgaaacaccg cacgcttacc tgcaaggcgt ggcgctggaa	1560
gtgatggcaa actccgataa cgtgctgcgt gcgggtctga cgcttaaata cattgatatt	1620
ccggaactgg ttgccaatgt gaaattcgaa gccaaacggg ctaaccagtt gttgaccag	1680
ccggtgaaac aaggtgcaga actggacttc ccgattccag tggatgattt tgccttctcg	1740
ctgcatgacc ttagtgataa agaaaccacc attagccagc agagtgccgc cattttgttc	1800
tgcgtcgaag gcgatgcaac gttgtggaaa ggttctcagc agttacagct taaaccgggt	1860
gaatcagcgt ttattgccc caacgaatca ccggtgactg tcaaaggcca cgcccgttta	1920
gcgcgtgttt acaacaagct gtaagagctt actgaaaaa ttaacatctc ttgctaagat	1980
ccatgcatgg atattcgaac gcgtaggtac cacatggtta acctagactt gtccatcttc	2040
tggattggcc aacttaatta atgtatgaaa taaaaggatg cacacatagt gacatgctaa	2100
tcactataat gtgggcatca aagttgtgtg ttatgtgtaa ttactagtta tctgaataaa	2160
agagaaagag atcatccata tttcttatcc taaatgaatg tcacgtgtct ttataattct	2220
ttgatgaacc agatgcattt cattaaccaa atccatatac atataaatat taatcatata	2280

ES 2 843 556 T3

taattaatat	caattgggtt	agcaaaacaa	atctagtcta	ggtgtgtttt	gcgaattgcg	2340
gccgggtacc	gagctcgaat	tcggcccaag	tttgtacaaa	aaagcaggct	ccggccagag	2400
ttacccgga	cgaagcttgc	atgcctgcag	tgcagcgtga	cccggctcgtg	cccctctcta	2460
gagataatga	gcattgcatg	tctaagttat	aaaaaattac	cacataatctt	ttttgtcaca	2520
cttgtttgaa	gtgcagttta	tctatcttta	tacatatatt	taaactttac	tctacgaata	2580
atataatcta	tagtactaca	ataatatcag	tgttttagag	aatcatataa	atgaacagtt	2640
agacatggtc	taaaggacaa	ttgagtattt	tgacaacagg	actctacagt	tttatctttt	2700
tagtgtgcat	gtgttctcct	ttttttttgc	aaatagcttc	acctatataa	tacttcatcc	2760
attttattag	tacatccatt	tagggtttag	ggttaatggg	ttttatagac	taattttttt	2820
agtacatcta	ttttattcta	ttttagcctc	taaattaaga	aaactaaaac	tctatttttag	2880
tttttttatt	taataattta	gatataaaat	agaataaaat	aaagtgacta	aaaattaaac	2940
aaataccctt	taagaaatta	aaaaaactaa	ggaaacattt	ttcttgtttc	gagtagataa	3000
tgccagcctg	ttaaagcccg	tcgacgagtc	taacggacac	caaccagcga	accagcagcg	3060
tcgcgtcggg	ccaagcgaag	cagacggcac	ggcatctctg	tcgctgcctc	tggacccttc	3120
tcgagagttc	cgtccaccg	ttggacttgc	tccgctgtcg	gcacccagaa	attgcgtggc	3180
ggagcggcag	acgtgagccg	gcacggcagg	cgccctcctc	ctcctctcac	ggcaccggca	3240
gctacggggg	attcctttcc	caccgctcct	tcgctttccc	ttcctcgccc	gccgtaataa	3300
atagacaccc	cctccacacc	ctctttcccc	aacctcgtgt	tgttcggagc	gcacacacac	3360
acaaccagat	ctcccccaaa	tccaccgctc	ggcacctccg	cttcaaggta	cgcgcgtcgt	3420
cctccccccc	ccccctctct	accttctcta	gatcggcggt	ccgggtccatg	catggttagg	3480
gcccggtagt	tctacttctg	ttcatgtttg	tgttagatcc	gtgtttgtgt	tagatccgtg	3540
ctgctagcgt	tcgtacacgg	atcgacacctg	tacgtcagac	acgttctgat	tgctaacttg	3600
ccagtgtttc	tctttgggga	atcctgggat	ggctctagcc	gttccgcaga	cgggatcgat	3660
ttcatgattt	ttttgtttc	gttgcatagg	gtttggtttg	cccttttccct	ttatttcaat	3720
atatgccgtg	cacttgtttg	tcgggtcatc	ttttcatgct	tttttttgtc	ttggttgtga	3780
tgatgtggtc	tggttgggcg	gtcgttctag	atcgagtag	aattctgttt	caaactacct	3840
ggtggattta	ttaatttttg	atctgtatgt	gtgtgccata	catattcata	gttacgaatt	3900
gaagatgatg	gatggaaata	tcgatctagg	ataggtatac	atgttgatgc	gggttttact	3960
gatgcatata	cagagatgct	ttttgttcgc	ttggttgtga	tgatgtgggtg	tggttgggcg	4020
gtcgttcatt	cgttctagat	cggagtagaa	tactgtttca	aactacctgg	tgtatttatt	4080
aattttgga	ctgtatgtgt	gtgtcataca	tcttcatagt	tacgagttta	agatggatgg	4140
aaatatcgat	ctaggatagg	tatacatggt	gatgtggggt	ttactgatgc	atatacatga	4200

ES 2 843 556 T3

tggcatatgc agcatctatt catatgctct aaccttgagt acctatctat tataataaac	4260
aagtatgttt tataattatt ttgatcttga tatacttgga tgatggcata tgcagcagct	4320
atatgtggat ttttttagcc ctgccttcat acgctattta tttgcttggc actgtttctt	4380
ttgtogatgc tcacctgtt gtttgggtgtt acttctgcag gtogactcta gaggatccac	4440
cggtcgccac catggcctcc tccgagaacg tcatcaccga gttcatgcgc ttcaagggtgc	4500
gcatggaggg caccgtgaac ggccacgagt tcgagatcga gggcgagggc gagggccgcc	4560
cctacgaggg ccacaacacc gtgaagctga aggtgaccaa gggcggtccc ctgcccttcg	4620
cctgggacat cctgtccccc cagttccagt acgggtccaa ggtgtactg aagcaccgcc	4680
ccgacatccc cgactacaag aagctgtcct tccccgaggg cttcaagtgg gagcgcgtga	4740
tgaacttcga ggacggcggc gtggcgaccg tgaccaggga ctccctccctg caggacggct	4800
gcttcatcta caaggtaagt ttctgcttct acctttgata tatatataat aattatcatt	4860
aattagtagt aatataatat ttcaaattt tttttcaaaa taaaagaatg tagtatatag	4920
caattgcttt tctgtagttt ataagtgtgt atattttaat ttataacttt tctaataat	4980
gacaaaaaca tggtagatgtg cagggtgaagt tcatcggcgt gaacttcccc tccgacggcc	5040
ccgtgatgca gaagaagacc atgggctggg aggcctccac cgagcgcctg taccgccgcg	5100
acggcgtgct gaaggcgag acccacaagg ccctgaagct gaaggacggc ggccactacc	5160
tggtaggatt caagtcctac tacatggcca agaagccgt gcagctgccc ggctactact	5220
acgtggacgc caagctggac atcacctccc acaacgagga ctacaccatc gtggagcagt	5280
acgagcgcac cgaggggccc caccacctgt tcctgtagcg gccgaagcta acctagactt	5340
gtccatcttc tggattggcc aacttaatta atgtatgaaa taaaaggatg cacacatagt	5400
gacatgctaa tcactataat gtgggcatca aagttgtgtg ttatgtgtaa ttactagtta	5460
tctgaataaa agagaaagag atcatccata tttcttatcc taaatgaatg tcacgtgtct	5520
ttataattct ttgatgaacc agatgcattt cattaaccaa atccatatac atataaatat	5580
taatcatata taattaatat caattgggtt agcaaaaaca atctagtcta ggtgtgtttt	5640
gcgaatgcgg ccgccaccgc ggtggagctc gaattccggt ccgggtcacc cggtcggggc	5700
ctagaaggcc gatctcccg gcacccagct ttcttgtaca aagtggccgt taacggatcc	5760
cggatgaagt cctattccga agttcctatt ctccagaaag tataggaaact tcactagagc	5820
ttgcggccgc cccggagctt gcatgcctgc agtgcagcgt gaccgggtcg tgccctctc	5880
tagagataat gagcattgca tgtctaagtt ataaaaaatt accacatatt ttttttgtca	5940
cacttgtttg aagtgcagtt tatctatctt tatacatata tttaaacttt actctacgaa	6000
taatataatc tatagtacta caataatatc agtgttttag agaatcatat aatgaacag	6060

ttagacatgg tctaaaggac aattgagtat tttgacaaca ggactctaca gttttatcct	6120
tttagtgtgc atgtgttctc cttttttttt gcaaatagct tcacctatat aatacttcat	6180
ccatttttatt agtacatcca tttagggttt aggggttaatg gtttttatag actaattttt	6240
ttagtacatc tattttattc tatttttagcc tctaaattaa gaaaactaaa actctatttt	6300
agtttttttta ttttaataatt tagatataaa atagaataaa ataaagtgc taaaaattaa	6360
acaaataccc ttttaagaaat taaaaaaact aaggaaacat ttttcttggt tcgagtagat	6420
aatgccagcc tgttaaagc cgtcgacgag tctaacggac accaaccagc gaaccagcag	6480
cgtcgcgtcg ggccaagcga agcagacggc acggcatctc tgtcgcgtgcc tctggacccc	6540
tctcgagagt tccgctccac cgttggaact gctccgctgt cggcatccag aaattgcgtg	6600
gcggagcggc agacgtgagc cggcacggca ggcgccctcc tcctcctctc acggcacggc	6660
agctacgggg gattcctttc ccaccgctcc ttcgctttcc ctccctcgcc cgccgtaata	6720
aatagacacc ccctccacac cctctttccc caacctcgtg ttgttcggag cgcacacaca	6780
cacaaccaga tctcccccac atccaccgct cggcacctcc gcttcaaggc acgccgctcg	6840
tcctccccc cccccctct ctaccttctc tagatcggcg ttccggtcca tggttagggc	6900
ccggtagttc tacttctgtt catgtttgtg ttagatccgt gtttgtgtta gatccgtgct	6960
gctagcgttc gtacacggat gcgacctgta cgtcagacac gttctgattg ctaacttgcc	7020
agtgtttctc tttggggaat cctgggatgg ctctagccgt tccgcagacg ggatcgattt	7080
catgattttt tttgtttcgt tgcatagggt ttggtttgcc cttttccttt atttcaatat	7140
atgccgtgca cttgtttgtc gggcatctct ttcatgcttt tttttgtctt ggttgtgatg	7200
atgtggtctg gttgggcggc cgttctagat cggagtagaa ttctgtttca aactacctgg	7260
tggattttatt aattttggat ctgtatgtgt gtgccataca tattcatagt tacgaattga	7320
agatgatgga tggaaatata gatctaggat aggtatacat gttgatcggg gttttactga	7380
tgcatacata gagatgcttt ttgttcgctt ggttgtgatg atgtggtgtg gttgggcggc	7440
cgttcattcg ttctagatcg gagtagaata ctgtttcaaa ctacctggtg tatttattaa	7500
ttttggaact gtatgtgtgt gtcatacatc ttcatagtta cgagtttaag atggatggaa	7560
atatcgatct aggataggta tacatgttga tgtgggtttt actgatgcat atacatgatg	7620
gcatatgcag catctattca tatgctctaa ccttgagtac ctatctatta taataaacia	7680
gtatgtttta taattatttt gatcttgata tacttggtat atggcatatg cagcagctat	7740
atgtggattt ttttagccct gccttcatac gctatttatt tgcttggtac tgtttctttt	7800
gtcgtatgctc accctgttgt ttggtgttac ttctgcaggc cgactctaga ggatccaaca	7860
atgccccagt tcgacatcct ctgcaagacc ccccccaagg tgctcgtgag gcagttcgtg	7920
gagagggttc agaggccctc cggcgagaag atcgccctct gcgccgcga gctcacctac	7980

ES 2 843 556 T3

ctctgctgga tgatcaccca caacggcacc gccattaaga gggccacctt catgtcatac	8040
aacaccatca tctccaactc cctctccttc gacatcgtga acaagtcctt ccagttcaaa	8100
tacaagaccc agaaggccac catcctcgag gcctccctca agaagtcctt ccccgctgg	8160
gagttcacca tcatcccta ctacggccag aagcaccagt ccgacatcac cgacatcgtg	8220
tcatccctcc agcttcagtt cgagtcctcc gaggaggctg acaagggcaa ctccactcc	8280
aagaagatgc tgaaggcctt cctctccgag ggcgagtcga tctgggagat caccgagaag	8340
atcctcaact ccttcgagta cactccaggt ttactaaga ccaagaccct ctaccagttc	8400
ctcttcctcg ccaccttcat caactgcggc aggtttctcag acatcaagaa cgtggacccc	8460
aagtccttca agctcgtgca gaacaagtac ctaggtttgt ttctgcttct acctttgata	8520
tatatataat aattatcatt aattagtagt aatataatat ttcaaattt tttttcaaaa	8580
taaaagaatg tagtatatag caattgcttt tctgtagttt ataagtgtgt atattttaat	8640
ttataacttt tctaataatat gacaaaaaca tgggtgatgcc taggtgtcat catccagtc	8700
ctcgtgaccg agaccaagac ctccgtgtcc aggcacatct acttcttctc cgctcgcggc	8760
aggatcgacc ccctcgtgta cctcgacgag ttctcagga actcagagcc cgtgctcaag	8820
agggtgaaca ggaccggcaa ctctcctcc aacaagcagg agtaccagct cctcaaggac	8880
aacctcgtga ggtcctacaa caaggccctc aagaagaacg cccctactc catcttcgcc	8940
atcaagaacg gcccgaagtc ccacatcggg aggcacctca tgacctcctt cctctcaatg	9000
aagggcctca ccgagctcac caacgtggtg ggcaactggt ccgacaagag ggcctccgcc	9060
gtggccagga ccacctacac ccaccagatc accgccatcc ccgaccacta ctctgcctc	9120
gtgtcaaggt actacgccta cgaccccatc tccaaggaga tgatcgccct caaggacgag	9180
actaacccca tcgaggagtgc gcagcacatc gagcagctca agggctccgc cgagggtcc	9240
atcaggtacc ccgcctggaa cggcatcatc tcccaggagg tgctcgacta cctctcctcc	9300
tacatcaaca ggaggatctg agttaacctg gacttgtcca tcttctggat tggccaactt	9360
aattaatgta tgaaataaaa ggatgcacac atagtgcacat gctaactact ataagtggg	9420
catcaaagtt gtgtgttatg tgtaattact agttatctga ataaaagaga aagagatcat	9480
ccatatttct tatcctaaat gaatgtcacg tgtctttata attctttgat gaaccagatg	9540
catttcatta accaaatcca tatacatata aatattaatc atatataatt aatatcaatt	9600
gggttagcaa aacaaatcta gtctaggtgt gttttgcgaa ttgcggcccc gggcaacttt	9660
attatacaaa gttgatagat atctggtcta actaactagt cctaaggacc cggcggaccg	9720
attaaactga ttcggtccga agcttaagcc atggcccggt aatcttagcg gccgcctgca	9780
gagttaacgg cgcgccgact agctagctaa ggtaccgagc tcgaattcat tccgattaat	9840
cgtggcctct tgctcttcag gatgaagagc tatgtttaaa cgtgcaagcg ctactagaca	9900
attcagtaca ttaaaaacgt ccgcaatgtg ttattaagtt gtctaagcgt caatttgttt	9960
acaccacaat atatcctgcc ac	9982

<210> 76

<211> 10955

<212> ADN

<213> Secuencia artificial

<220>

<223> Construcción sintética

<400> 76

gtttaccgc caatatatcc tgtcaaacac tgatagttta aactgaaggc gggaaacgac	60
aactctgatca tgagcggaga attaaggagac tcacgttatg acccccgccg atgacgcggg	120
acaagccgtt ttacgtttgg aactgacaga accgcaacgt tgaaggagcc actcagcaag	180
ctggtacgat tgtaatacga ctactatag gccgaattga gcgctgttta aacgctcttc	240
aactggaaga gcggttacca gagctggtca cctttgtcca ccaagatgga actgcggccg	300
taagtgacta gggtcacgtg accctagtca cttatcgagc tagttaccct atgaggtgac	360
atgaagcgct caggttact atgacggtta gcttcacgac tgttggtggc agtagcgtac	420
gacttagcta tagttccgga cttaccgggc ccaccggtg taccgagctc gtttaaaccg	480
tcttcaactg gaagagcggc taccagagct ggtcaccttt gtccaccaag atggaactgg	540
cgcgcctcat taattaagtc agcggccgct ctagttagaag acacgttcat gtcttcatcg	600
taagaagaca ctacgtatgc ttccggccaga atggccatct ggattcagca ggcctagaag	660
gccatttaaa tcttgaggat ctggtcttcc taaggaccgc ggatatacga ccgattaaac	720
tttaattcgg tccgacctgg tggcgccgct agcgtatacg aagttcctat tccgaagttc	780
ctattctcca gaaagtatag gaacttctgt acaataactt cgtatagcat acattatacg	840
aagttatatg catacgcgtc ttaagtcgat cgctatcaac tttgtataga aaagtgggc	900
cgaattcgag ctcggtacgg ccagaatggc ccggaccggg ttaccgaatt cgagctcggc	960
accctgggat ccgatatcga tgggccctgg ccgaagcttg gtcaccggc ccgggcctag	1020
aaggccagct tcaagtttgt acaaaaaagc aggtccggc cagaatggc ccgaccggaa	1080
ttcgagctcg gatccactag taacggccgc cagtgtgctg gaattcggcc ttgacggccc	1140
gggctggtat ttcaaaacta tagtatttta aaattgcatt aacaaacatg tcctaattgg	1200
tactcctgag atactatacc ctctgtttt aaaatagttg gcattatcga attatcattt	1260
tactttttaa tgttttctct tcttttaata tattttatga attttaaatg attttaaaat	1320
gttatgcagt tcgctctgga cttttctgct gcgcctacac ttgggtgtac tgggcctaaa	1380
ttcagcctga ccgaccgcct gcattgaata atggatgagc accggtaaaa tccgcgtacc	1440

ES 2 843 556 T3

caacttttga gaagaaccga gacgtggcgg gccgggccac cgacgcacgg caccagcgac	1500
tgcacacgtc ccgccggcgt acgtgtacgt gctgttccct cactggccgc ccaatccact	1560
catgcatgcc cacgtacacc cctgccgtgg cgcgcccaga tcctaatacct ttcgccgttc	1620
tgcacttctg ctgcctataa atggcgcat cgaccgtcac ctgcttcacc accggcgagc	1680
cacatcgaga acacgatcga gcacacaagc acgaagactc gtttaggaga aaccacaaac	1740
caccaagccg tgcaagcacc aagcttggtc acccggccg gccctagaag gccagcttca	1800
agtttgtaca aaaaagcagg cttcgaagga gatagaaccg atccaccatg tccaacctgc	1860
tcacggttca ccagaacctt ccggctcttc cagtggacgc gacgtccgat gaagtcagga	1920
agaacctcat ggacatgttc cgcgacaggc aagcgttcag cgagcacacc tggaagatgc	1980
tgctctccgt ctgccgtcc tgggctgcat ggtgcaagct gaacaacagg aagtggttcc	2040
ccgctgagcc cgaggacgtg agggattacc ttctgtacct gcaagcgcga ggtttgtttc	2100
tgcttctacc tttgatatat atataataat tatcattaat tagtagtaat ataataattc	2160
aaatattttt ttcaaaataa aagaatgtag tatatagcaa ttgcttttct gtagtttata	2220
agtgtgtata ttttaattta taacttttct aatatatgac caaaacatgg tgatgcctag	2280
gtctggcagt gaagaccatc cagcaacacc ttggacaact gaacatgctt cacaggcgct	2340
ccggcctccc gcgcccagc gactcgaacg ccgtgagcct cgtcatgcgc cgcacagga	2400
aggaaaacgt cgatgccggc gaaagggcaa agcaggccct cgcgttcgag aggaccgatt	2460
tcgaccaggt ccgcagcctg atggagaaca gcgacagggt ccaggacatt aggaacctgg	2520
cgttctcctg aattgcatac aacacgctcc tcaggatcgc ggaaattgcc cgcattcgcg	2580
tgaaggacat tagccgcacc gacggcggca ggatgcttat ccacattggc aggaccaaga	2640
cgtcgtttc caccgcaggc gtcgaaaagg ccctcagcct cggagtgacc aagctcgtcg	2700
aacgctggat ctccgtgtcc ggcgtcgcgg acgacccaaa caactacctc ttctgccgcg	2760
tccgcaagaa cggggtggct gccctagcgc ccaccagcca actcagcacg agggccttgg	2820
aaggtatttt cgaggccacc caccgcctga tctacggcgc gaaggatgac agcgttcaac	2880
gctacctcgc atgggtccgg cactccgccc gcgttgagc tgctagggac atggcccgcg	2940
ccggtgtttc catccccgaa atcatgcagg cgggtggatg gacgaacgtg aacattgtca	3000
tgaactacat tcgcaacctt gacagcgaga cgggcgcaat ggttcgcctc ctggaagatg	3060
gtgactgagc tagaccagc tttctgttac aaagtggcgc ttaacggatc cagacttgtc	3120
catcttctgg attggccaac ttaattaatg tatgaaataa aaggatgcac acatagtgc	3180
atgctaatac ctataatgtg ggcataaaag ttgtgtgtta tgtgtaatta ctagttatct	3240
gaataaaaaga gaaagagatc atccatattt cttatcctaa atgaatgtca cgtgtcttta	3300

taattctttg atgaaccaga tgcatttcat taaccaaadc catatacata taaatattaa	3360
tcatatataa ttaatatcaa ttgggttagc aaaacaaadc tagtctaggt gtgttttgcg	3420
aattgctgca agcttgctgc cgtcccgact tgggtcaccg gtccgggctt agaaggccga	3480
tctcccggtc acccagcttt cttgtacaaa gtggccgtta acggatcggt cagaatggcc	3540
cggaccgggt taccgaattc gagctcggtt ccctgggac gaccgaagct gaccgaagct	3600
tgcggccgca cactgatagt ttaaactgaa ggccgggaaac gacaatctga tcatgagcgg	3660
agaattaagg gagtcacgtt atgaccccg ccatgacgc gggacaagcc gttttacgtt	3720
tggaactgac agaaccgcaa cgattgaagg agccactcag ccgctgggtt ctggagtta	3780
atgagctaag cacatacgtc agaaaccatt attgcgcgtt caaaagtcgc ctaaggtcac	3840
tatcagctag caaatatttc ttgtcaaaaa tgctccactg acgttccata aattcccctc	3900
ggtatccaat tagagtctca tattcactct cccgggggat ctgactcta gaggatcgct	3960
caggaaggcc gctgagatag aggcattggc gccaatgcgg gcggcggtgg agcgggagga	4020
ggcagcggca ggcgcagcgt ggctgcgcgc gcggtgtgcc gcccagcgg ctgcgcgtgg	4080
acgccgacgc cggagcagat caggatgctg aaggagctct actacggctg cggcatccgg	4140
tcgcccgact cggagcagat ccagcgcac accgccatgc tgcggcagca cggcaagac	4200
gagggcaaga acgttttcta ctggttccag aaccacaagg ccgcgcagcg ccagaagcgc	4260
cgctcacca gcctcagct caactgccc gccgcggcg cggccgacgc caccaccagc	4320
caactcggcg tcctctcgt gtcgtcgccg ccgccttcag gcgcggcgcc tcctcggcc	4380
accctcggct tctacgccgc cggcaatggc ggccgacgg ctgtgctgct ggacacgagt	4440
tcggactggg gcagcagcgg cgtgcctat gccaccgaga catgcttct gcaggactac	4500
atgggcgtga cggacacggg cagctcgtgc cagtggccac gcttctcgtc gtcggacacg	4560
ataatggcgg cggccgcggc gcgggcggcg acgacgcggg cggccgagac gctccctctc	4620
ttcccgacct gcggcgacga cggcggcagc ggtagcagca gctacttgc gttctgggt	4680
gccgcgtcca caactgccgg cgcacttct tcggttgca tcagcagca acaccagctg	4740
caggagcagt acagctttta cagcaacagc aacagcacc agctggcgg caccggcaac	4800
caagacgtat cggcaacagc agcagcagcc gccgccctgg agctgagcct cagctcatgg	4860
tgtccctct accctgctgc agggagtat tgagagcaac gcgagctgcc actgctcttc	4920
actgatgtct ctggaatgga aggaggagga agtgagcata gcgttggtgc gttgctgtca	4980
agggcgcaatt gtaccacatg gttaacctag acttgccat cttctggatt ggccaactta	5040
attaatgtat gaaataaaa gatgcacaca tagtgacatg ctaatcacta taatgtgggc	5100
atcaaagttg tgtgttatgt gtaattacta gttatctgaa taaaagagaa agagatcatc	5160
catatttctt atcctaaatg aatgtcacgt gtctttataa ttctttgatg aaccagatgc	5220

ES 2 843 556 T3

athttcattaa ccaaattccat ataatataaa atattaatca tatataatta atatcaattg	5280
ggttagcaaa acaaattctag tctaggtgtg ttttgcgaat gcggccgcga ctctagatca	5340
taaatcagcca taccacattc gaatgtgagt tgatccccgg cgggtgtcccc cactgaagaa	5400
actatgtgct gtagtatagc cgctggctag ctagctagtt gagtcattta gcggcgatga	5460
ttgagtaata atgtgtcacg catcaccatg catgggtggc agtctcagtg tgagcaatga	5520
cctgaatgaa caattgaaat gaaaagaaaa aagtattgtt ccaaattaaa cgttttaacc	5580
ttttaatagg ttatataaat aattgatata tgttttctgt atatgtctaa tttgttatca	5640
tccatttaga tatagacgaa aaaaaatcta agaactaaaa caaatgctaa tttgaaatga	5700
agggagtata tattgggata atgtcgatga gatccctcgt aatatcaccg acatcacacg	5760
tgtccagtta atgtatcagt gatacgtgta ttcacatttg ttgcgcgtag gcgtacccaa	5820
caattttgat cgactatcag aaagtcaacg gaagcgagtc gacctcgagg gggggcccgg	5880
taccaagata tcaaccggcg aaagatctaa gcatgcaagg gcccaagtcg acctgcagaa	5940
gcttcgggtcc gggcctagaa ggccgatctc ccgggcaccc agctttcttg taaaaagtgg	6000
ccgttaacgg atcggccaga atggcccggg ccgggttacc gaattcgagc tcggtaccct	6060
gggatcgacc gaagctgacc gaagcttgca tgcctgcagt gcagcgtgac ccggtcgtgc	6120
ccctctctag agataatgag cattgcatgt ctaagttata aaaaattacc acatattttt	6180
tttgtcacac ttgtttgaag tgcagtttat ctatctttat acatatattt aaactttact	6240
ctacgaataa tataatctat agtactacaa taatatcagt gtttttagaga atcatataaa	6300
tgaacagtta gacatgggtc aaaggacaat tgagtatttt gacaacagga ctctacagtt	6360
ttatcttttt agtgtgcatg tgttctcctt tttttttgca aatagcttca cctatataat	6420
acttcatcca ttttattagt acatccattt aggggttagg gttaatggtt tttatagact	6480
aattttttta gtacatctat tttattctat tttagcctct aaattaagaa aactaaaact	6540
ctattttagt ttttttattt aataatttag atataaaaata gaataaaaata aagtgactaa	6600
aaattaataa aataaccctt aagaaattaa aaaaactaag gaaacatttt tcttgtttcg	6660
agtagataat gccagcctgt taaacgccgt cgacgagtc aacggacacc aaccagcgaa	6720
ccagcagcgt cgcgtcgggc caagcgaagc agacggcacg gcatctctgt cgctgcctct	6780
ggacccctct cgagagttcc gctccaccgt tggacttgct ccgctgtcgg catccagaaa	6840
ttgcgtggcg gagcggcaga cgtgagccgg cacggcaggc ggctcctctc tcctctcacg	6900
gcacccggcag ctacggggga ttcccttccc accgctcctt cgctttccct tcctcgcccg	6960
ccgtaataaa tagacacccc ctccacaccc tctttcccca acctcgtgtt gttcggagcg	7020
cacacacaca caaccagatc tcccccaaat ccaccgctcg gcacctccgc ttcaagggtac	7080

ES 2 843 556 T3

gccgctcgtc	ctcccccccc	ccctctctcta	ccttctcttag	atcggcgttc	cggtccatgc	7140
atggttaggg	cccggtagtt	ctacttctgt	tcatgtttgt	gtagatccg	tgtttgtgtt	7200
agatccgtgc	tgctagcgtt	cgtacacgga	tgcgacctgt	acgtcagaca	cgttctgatt	7260
gctaacttgc	cagtgtttct	ctttggggaa	tcctgggatg	gctctagccg	ttccgcagac	7320
gggatcgatt	tcatgatttt	ttttgtttcg	ttgcataggg	tttggtttgc	ccttttcctt	7380
tatttcaata	tatgccgtgc	acttgtttgt	cggttcatct	tttcatgctt	ttttttgtct	7440
tggttgtgat	gatgtggtct	ggttgggcgg	tcgttctaga	tcggagtaga	attctgtttc	7500
aaactacctg	gtggatttat	taattttgga	tctgtatgtg	tgtgccatac	atattcatag	7560
ttacgaattg	aagatgatgg	atggaaatat	cgatctagga	taggtataca	tgttgatgcg	7620
ggttttactg	atgcatatac	agagatgctt	tttgttcgct	tggttgtgat	gatgtggtgt	7680
ggttgggcgg	tcgttcattc	gttctagatc	ggagtagaat	actgtttcaa	actacctggt	7740
gtatttatta	attttggaac	tgtatgtgtg	tgtcatacat	cttcatagtt	acgagtttaa	7800
gatggatgga	aatatcgatc	taggataggt	atacatgttg	atgtgggttt	tactgatgca	7860
tatacatgat	ggcatatgca	gcactctattc	atatgctcta	accttgagta	cctatctatt	7920
ataataaaca	agtatgtttt	ataattattt	tgatcttgat	atacttggat	gatggcatat	7980
gcagcagcta	tatgtggatt	tttttagccc	tgccctcata	cgctatttat	ttgcttggtg	8040
ctgtttcttt	tgctgatgct	cacctgttg	tttgggtgta	cttctgcagg	tcgactctag	8100
aggatccatg	gccactgtga	acaactggct	cgctttctcc	ctctccccgc	aggagctgcc	8160
gccctcccag	acgacggact	ccacactcat	ctcgcccgcc	accgccgacc	atgtctccgg	8220
cgatgtctgc	ttcaacatcc	cccaagattg	gagcatgagg	ggatcagagc	tttcggcgct	8280
cgtcgcgag	ccgaagctgg	aggacttcct	cggcggcatc	tccttctccg	agcagcatca	8340
caaggccaac	tgcaacatga	taccagcac	tagcagcaca	gtttgctacg	cgagctcagg	8400
tgctagcacc	ggctaccatc	accagctgta	ccaccagccc	accagctcag	cgctccactt	8460
cgcggactcc	gtaatggtgg	cctcctcggc	cggtgtccac	gacggcgggtg	ccatgctcag	8520
cgcggccgcc	gctaacggtg	tcgctggcgc	tgccagtgcc	aacggcggcg	gcacgggct	8580
gtccatgatt	aagaactggc	tgcgagacca	accggcgccc	atgcagccga	gggtggcggc	8640
ggctgagggc	gcgcaggggc	tctctttgtc	catgaacatg	gcggggacga	cccaaggcgc	8700
tgctggcatg	ccacttctcg	ctggagagcg	cgcacgggcg	cccagagagt	tatcgacgtc	8760
agcacagggt	ggagccgtcg	tcgtcacggc	gccgaaggag	gatagcgggtg	gcagcggtgt	8820
tgccggcgct	ctagtagccg	tgagcacgga	cacgggtggc	agcggcggcg	cgtcggctga	8880
caacacggca	aggaagacgg	tggacacggt	cgggcagcgc	acgtcgattt	accgtggcgt	8940
gacaaggcat	agatggactg	ggagatatga	ggcacatctt	tgggataaca	gttgcagaag	9000

ES 2 843 556 T3

ggaagggcaa actcgtaagg gtcgtcaagt ctatttaggt ggctatgata aagaggagaa 9060
agctgctagg gcttatgata ttgctgctct gaagtactgg ggtgccacaa caacaacaaa 9120
ttttccagtg agtaactacg aaaaggagct cgaggacatg aagcacatga caaggcagga 9180
gtttgtagcg tctctgagaa ggaagagcag tggtttctcc agaggtgcat ccatttacag 9240
gggagtgact aggcatacacc aacatggaag atggcaagca cggattggac gagtgcagc 9300
gaacaaggat ctttacttgg gcaccttcag caccaggag gaggcagcgg aggcgtacga 9360
catcgccggc atcaagttcc gcggcctcaa cgcgctcacc aacttcgaca tgagccgcta 9420
cgacgtgaag agcatcctgg acagcagcgc cctccccatc ggcagcgcgg ccaagcgcct 9480
caaggaggcc gaggccgag cgtccgcgca gcaccaccac gccggcgtgg tgagctacga 9540
cgtcgcccg c atcgccctgc agctcggcga cggcggagcc ctggcggcgg cgtacggcgc 9600
gcactaccac ggcgcgcgcct ggccgaccat cgcgttccag ccgggcgcgg ccagcacagg 9660
cctgtaccac ccgtacgcgc agcagccaat gcgcggcgcc ggggtgtgca agcaggagca 9720
ggaccacgcg gtgatcgcgg ccgcgcacag cctgcaggac ctccaccacc tgaacctggg 9780
cgcggccggc gcgcacgact ttttctcggc agggcagcag gccgcgcgg ctgcgatgca 9840
cggcctgggt agcatcgaca gtgcgtcgtc cgagcacagc accggctcca actccgtcgt 9900
ctacaacggc ggggtcggcg acagcaacgg cgcacgcgc gtcggcgga gtggcgggtg 9960
ctacatgatg ccgatgagcg ctgccggagc aaccactaca tcggcaatgg tgagccacga 10020
gcaggtgcat gcacgggcct acgacgaagc caagcaggct gctcagatgg ggtacgagag 10080
ctacctgggt aacgcggaga acaatgggtg cgggaaggatg tctgcatggg ggactgtcgt 10140
gtctgcagcc gcggcggcag cagcaagcag caacgacaac atggccgcgg acgtcggcca 10200
tggcggcgcg cagctcttca gtgtctgga cgcacttaa gcgtacgtgc cggcctggct 10260
ctccgaaagg gcgaattcca gcacactggc ggccgttact agaccaacc tagacttgtc 10320
catcttcttg attggccaac ttaattaatg tatgaaataa aaggatgcac acatagtac 10380
atgctaata ctataatgtg ggcatcaaag ttgtgtgtta tgtgtaatta ctagttatct 10440
gaataaaaga gaaagagatc atccatattt cttatcctaa atgaatgtca cgtgtcttta 10500
taattctttg atgaaccaga tgcatttcac taaccaaacc catatacata taaatattaa 10560
tcatatataa ttaatatcaa ttgggttagc aaaacaaatc tagtctaggt gtgttttgcg 10620
aattagcttg gtcacccggc ccgggcctag aaggccagct tcggccgccc cgggcaactt 10680
tattatacaa agttgataga tcgaataact tcgtatagca tacattatac gaagttatcc 10740
tgagctgatt ccgatgactt cgtaggttcc tagctcaagc cgctcgtgtc caagcgtcac 10800
ttacgattag ctaatgatta cggcatctag gaccgactag ctaactaact agtacaacg 10860
tgcaagcgct actagacaat tcagtacatt aaaaacgtcc gcaatgtgtt attaagttgt 10920
ctaagcgtca atttgtttac accacaatat atcct 10955

5

<210> 77
<211> 965
<212> ADN
<213> Streptomyces spectabilis
<400> 77

ES 2 843 556 T3

atgcagaatc ttctgaaaa gtctgatgaa acacgcctat gggaagcttt gaggggattt	60
ggtgtgtccc cgtcgcgtgt cacctacgcc cccgtcggat tcggggacta ccactggacg	120
gtgaccgacg aggacggccg gccgtggttc gccaccgtct ccgacctgga gcacaaggag	180
cactgcgggc agggggcaca ggccgcgctg aaggggctcc gacaggccat ggacacggcg	240
ctggccctgc gcgaccgtga cggactgcgg ttctgtcgtg caccgtggc cgccaccgac	300
ggcggcggtc cagtactgcc cctggacgcc cggtagcgcc tcacggtgtt ccccatgtc	360
ccgggccgca caggggagtt cgggcagcgc ctgacggagg ccgagcggga ccggctgctc	420
gccctgctcg cggagctgca cggccggacg ccccgaggga ccacaccgcc cgccgacatg	480
gagcccccg gccttcccg cgtcgcgcgc gccctggccg agtcggaagg gccctggtcg	540
ggcggcccg tgcgcgagcc cgcgcggctg ctgctgcgcg agcacgaggc gacgtccac	600
gcgcgcctcg ccgagttcga ggccctcgtc gcgcgcgtga agggccgggg cgcccgctc	660
gtcgtcacgc acggcgagcc gcatccgggc aacctcatcc tgggtgagga gggctatctc	720
ctggtggact gggacacggt gggcctcgcc ccggccgagc gcgacctctc cctgatctcg	780
gacgacccg cggacctcgc ccgctacgcc gagctgaccg gccgcacccc ggaccggac	840
gccctcgcgc tctaccggtg gagggttcc gcgccgagca ccagcgacg caggacaccg	900
agtccgctg gaagggcttc acggacacgc tggggcaact ggcgcgggc ggatccgag	960
cctag	965

<210> 78

<211> 993

<212> ADN

5 <213> Legionella pneumophila

<400> 78

atgctaaaac aaccaattca agctcaacaa cttatcgaac ttttgaaagt gcattatgga	60
attgatattc atacagcaca attcatccag ggtggtgctg atacgaatgc atttgcata	120
caagcagatt cagaatcaa gtcttatttc ataaagctaa aatacggcta tcatgatgaa	180
attaatttat cgataatccg tcttttacat gattctggaa taaaagaaat tatttttcct	240
atccatacac ttgaagcaaa attattccag caactaaagc atttttaa aattgcgtat	300
ccatttattc atgcgccc aa tggtttcacc caaaatttaa caggaaaaca gtggaaacag	360

	cttgaaaaag tattaagaca aattcatgaa acatcagttc ccatctcgat tcaacaacaa	420
	ttaagaaaag aaatatactc ccctaaatgg cgtgaaatag tcagatcctt ttataatcaa	480
	attgaatttg ataattcaga tgataagctc acggctgcct ttaaactctt ttttaaccaa	540
	aatagtgtcg caattcatcg attagtgtat acttcagaaa aactatctaa aaaaattcaa	600
	cctgatttag ataaatacgt actatgtcat tctgatatac atgcgggcaa tgtgttagtc	660
	ggtaatgaag agtcgattta cattattgat tgggatgagc ctatgttagc tccaaaagaa	720
	cgtgatttga tgttcatagg tgggtggcgtt ggtaatgtat ggaataaacc ccatgaaatc	780
	caatatTTTT atgaaggTTa tggTgaaata aatgtcgata aaacaatttt gtcttattac	840
	aggcatgaac gaattgtcga agatatcgca gtatacgggc aagacttgct ttcacgtaat	900
	caaaacaatc agtccagact tgaaagTTTT aaatatttta aagaaatgtt tgatccaaac	960
	aacgttggtg aaatagcttt tgctacagag cag	993
	<210> 79	
	<211> 2556	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 79	
	atgatttggg atcatttatc accttcccga agggttctgc ttgatatcaa atcgtgggtc	60
	gtattggtac tggtaattgc ctcgatatct tttgcgggtc tggccatagg atcttggcaa	120
	gacaacgaaa gcaacgaagc gatcctgaca gagctgcaat cgatcgacgt ggattgcgcg	180
	atgctgcagc ggaatgtgtt gcgtgctcac gccggccttc ttcggaatta ccatcctctt	240
	atcgtccctc tggggcgact gcgcacgagc gtcgcagatt tacagcagct ttttaggcgg	300
	gcgggcctcg acggggccgg tgaattttct gaattgctgg cgcggtgaa gagctcagta	360
	gatgcgaccg acgcagccgt tgcagcgttc ggtacgcaaa atgtaatgtt ccaggattcg	420
	ttggcgacct ttaaccagtc gatagcttcg cttccgaggt ctttggtac gagggatttg	480
	aactcgctag aagtcgctga gttgggctat ctgatgctcc agttttcgtc tcagccaaat	540
	ttcgagtttg cacgggaaat caaccaacgt ctcgaccggc ttcaggtttc ggcgaacggg	600
	gataaggcag gcgtcaagga gatcgttcgc agtgggagga ttatcctgac tctactgccg	660
	cgattgaaag acaccgttgg agtgatccaa acttctgaaa ccatcgacaa cacgaagaag	720
	ttgcagcgag cataccttga agcttatagc ttggccagcg cgggtgaaca gcgggtacgg	780
	gtcttctctg gcgcagtttc ggttttcttt tgcttttgca taatttccct agtccatagg	840
	ttacatctgc gaaccaaggt gttgacgaga cgactggact tcgaagaagt gatcaaaaag	900
	attgggcttt gcttcgaaga tgcttcagaa gcaaaaccgt cgttggaatc ctcaacgaat	960
	gccgcgttgg gaattatcca attattcttt gaagctcatc aatgcgcgct gggactggtc	1020

ES 2 843 556 T3

	aacatcaatg agaatgacgt tggcgtcacc ttttttggaa gcactcctgc gccagagtgg	1080
	aacgaacaac ggggtgcgcga aattgtttcc acggtagggg ccgaggacgg cgggcggatt	1140
	ttccgagcct atcctcaacg aaaagcgagc tgcttcggtg aggactctcc atttctgtgg	1200
	gtgctgcttg cgttcaaggt atccgatcga atagtcgctg tctttggcct gggatttgat	1260
	cgagagcatc tacagacgcc gaccgcacgt gaaatgcagc tgatggaact agcggctgca	1320
	tgtgttagcc actatgcggt catttcggcg c aagcagtcctc agcgagatat tctggaacgt	1380
	cgtttgaaac atgcggaacg gcttgaagca gttggtacgc ttgccggcgg cattgcgcac	1440
	gaatttaaca acattttggg tgcagttctg ggttacgcgg aaatggcaca caacgttctg	1500
	cgccgtcgta ctcatgctcg agactacatc gatcacatca tttctgaagg caatcgagcg	1560
	cgggtgattg tcaaccagat ccttgccttt agtcgaaagc gggatcgaac aacaaaacca	1620
	ttcgatcttt cggagttagt gaccgtgatc gcgcgctcgt tacgtgttgc gttgcctccg	1680
	ggcgtcgagc tcgatttcaa atatcagagc gctccaatgg ttatcgaagg aaactcactt	1740
	gaggtcgagc aaatactgat gaatttgtgc aaaaactctg ctgaagcatg ccgagagtta	1800
	gggcgtgtgg aggtcagcgt taggaggtct gtcgtacgga aacacaagggt tcttaciaaac	1860
	ggtactattc ccaccggcga gtatatcctt ctttcagttg aggataacgg cgcaggcata	1920
	acagagggcg tattatcgca ctttttcgag ctttttttta ctaccgcgc ccgcagcgggt	1980
	ggcacggggc tgggattatc cacggttcac ggtcacgtta gcgcgatggc tggttacatc	2040
	gacgtggtat caactgtcgg cagaggtacg cgcttcgacg tctacttgcc accttcgtcg	2100
	aagcagccgg taaattcaga gagttttttt gaccctggaa gaatcccgtt tgggaacggc	2160
	gagattgtcg cgggtggtcga acccgactta gccacgctgg aaatgtacga ggaaaagatc	2220
	gcggcgctgg gctatgagcc agtcggcttt aacacgttgg atggtttaat cgactgggtg	2280
	ttggaaggaa aagaacctga tcttgttcta attgatcaac cgtctttgct ggatggagac	2340
	ggtgcagggc ccttaatctc aaagctggca aaagcgcca tcatcattat tggcgaaaac	2400
	cagaaaaatg ttcccctttc tgcggatcgc gagggatctg cccgtttttt ggaaaagccg	2460
	atttcagcaa agactctcgc ctatgtcgtt cgtgcgaata tcagaactga acgaatggtc	2520
	ggctcgaata gcagcggaat cgtgattgcg tcataa	2556
	<210> 80	
	<211> 720	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 80	
	atgcggaaga cggcaggatc tttcgggatc atttttctga tggccagcc aagcaaagct	60
	gcgccgctct cgttcgctga atttaaccaa ctgcgtcgcg aatgcgcacc ttcagtcgct	120

	ccatcgacgc ttggagcgat tgcgaagggt gaaagtcagt ttgatccgct cgtgctgcat	180
	gacaataacca ccggcgaaac ttacatggc aagaatccgg cggatgctac gcaaagcgta	240
	aacaatcgcg tcgcagcagg gcattcagtc gatgtaggcc taatgcaagt gaactcgaaa	300
	aatttcgcga acttgggctt gactgccggc aatgcaatca acccttgcgt gtcctgtcg	360
	gcggcgccgg acctgctcgt acggcattat agtggcgccg acacggtaga gtcagaacaa	420
	cttgccatcc ggcgagcaat ttcggttac aataccggca acccgacacg cggcttcgcg	480
	aacggctatg tgcggagggt cgaagtggcg gctcagctac tcgtccccc gctggctcag	540
	tcggggaaaa aaggcgaccg tgatgagcgg agtcacagaag aaccctggaa cgtctgggg	600
	tcatacgacc gctccactc agcagtcggg tcgtccgcgc cgccagaggc gcagcagccg	660
	aatgagcgc aatccccga acaagatcaa gttttcgaac cgaatgatgg agatgcacca	720
	<210> 81	
	<211> 360	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 81	
	atggcatggc ttgaaggata ccgcgcaccg tcgaaattcc gtggcctctg gcacagagct	60
	gtgcgcctgc tggctccgca cgtgccgagc gtcaccggcg cgatcggtg gagcttggtt	120
	ttctgcgagc cagctgccgc gcaggccgct gggggcactg accccgccac catggtcaac	180
	aatatctgca cgtttattct cggctccgtt ggtcaatcgc tcgccgttct tggatcgta	240
	gccatcgggg tatcgtggat gttcggccga gcttccttg gtctggtagc aggtgttgtt	300
	ggcgggatcg ttatcatggt tggagccagc tttctcgcc aaacactgac agggggcggg	360
	<210> 82	
	<211> 324	
10	<212> ADN	
	<213> Agrobacterium rhizogenes	
	<400> 82	
	atgaatgccg atcttgaaga agccacactc tacttggcgg caacgagacc tgcattgttc	60
	tttggggtag cgctgacgct tgctggagtt ttcattgatg tggcgggggt tctcatcggt	120
	gttgtgcaga accccctcta tgaaattggt ctcttgccgc tgtggcttgg agcgcggctc	180
	atcgtagagc gggactataa cgcggcaagc gtggtcttgc tgtttctcca gacggcaggg	240
	agaagcgttg acggccatgc gtggggcggc gcgagcgtca gtcccaacct catcagagt	300
	gcgtcgcgag gaagaggaat gacg	324
	<210> 83	
15	<211> 2370	
	<212> ADN	
	<213> Agrobacterium rhizogenes	
	<400> 83	

ES 2 843 556 T3

atgttcggag caagcggacg gaccgagaga tctggggaaa tctacctgcc ttacgtaggg	60
cacctcagcg accatgtcgt tcttctagag gacggttcca tcttgaccat ggcgcatatc	120
agtggccttc ccttcgaact ggaggaagta gaggtccgaa atgcgcgctg ccgtgccttt	180
aacacgctct ttcgcaatat tgccgatgat aacgtctcag tatatgcca tctcgtccgt	240
cacaacgatg tgccagtacc gccggtacgc cattttcgca gcaccttcgc gtcaagccta	300
agcgaaacgt tcgagcggcg tgttctgtct ggtaaactgt ttcgaaacga ccacttcctc	360
acgctcattg tctctcctcg caccgctctc ggtaaagtgg gaagcagatt caccaggcgc	420
tatgggaaga acccaggcga tcttcgcat cagatccgac atctggagga tctttggaac	480
gttgctcgtg gcgcgcttga tgggtatgga ctctcgcac ttggcatccg agaaaaaat	540
caggtgcttt tcacggaaat tgggtgaagcc ctctgcctga taatgacttg tcgtttttta	600
ccggtgccgg tcgttagcgg atccctcggc gcctcagttt ataccgaccg ggtcatctgc	660
ggcaagcgtg ccctcgagat tagagcgccg aaagatcgtt acgttggatc tatcttttca	720
ttccgcgagt acccggcaaa gacgcggccc ggaatgctta acacactcct gtcatcaagt	780
tttccgctcg ttttgagcca gagtttctct ttcctcacac gtgcacaggc tcacgccaag	840
ctcagcctca aatcgaacca gatgactagc tccggtgaca aggcggtcac ccagatcggg	900
gaactcgccc aagcagaaga ttcatgtggc agtagcgaat tcgtcatggg ttcgcatcat	960
ctcagccttt gtgtctatgg cgatgatctc gatacccttg ccgactatgg tgcgcgagcc	1020
cgcacgagct tgtcggacgc gggcgcggtt atcgtgcagg aaagcattgg aatggaggct	1080
gcatactggt ctcagcttcc aggaaaccac agatggcgta cccgtcccgg agcgatcact	1140
tcacgcaact ttgccggctt ggtttcattc gagaacttcc ctatcggaag gaagtccggc	1200
cattggggag gtgcggttgc tcgtttccgt acaaacggcg ggacgccctt cgattatatt	1260
ccgcacgaaa acgacgtggg catgaccgcg atcttcgggc cgattggcag gggaaaaacg	1320
acacttatga ccttcatcct ggcgatgctt gagcagagcg tggttgaccg cagcgggacg	1380
atcgtttttt tcgataagga tcgtggcggg gaactgcttg tgcgcgccac cgggggcacg	1440
tacctcacgt tacgacgagg tgtcccaagc gcccttgcac ctttgcgcgg cctagaggat	1500
acggcagcgg cgccggactt tctccgcgag tggatcgtgg cgctgattga aagcgacggc	1560
agggggggaa tttcgccgga ggaaaatcgc cgcttagagc gcggtatcca acggcaactt	1620
tcgtttgagc ctaacatgag atcccttgcg ggcttaaggg agttcttatt gcacggggcc	1680
tccgaaggtg cgggagctag attgcagcgg tggtgccgag gcaacgctct tggctgggcg	1740
tttgacggcg aatcggacga agtgaagtta gatccttcga ttacgggctt cgatatgacc	1800

ES 2 843 556 T3

	catctcctcg aatatgagga ggtatgcgcg ccggccgcag catacctttt gcatcgaata	1860
	ggagcgatgg tagatggccg acgcttcgtc atgagctgcg atgagtttcg ttcctacctg	1920
	cttaatcaga aattcgcagc agtcgtcgac aagttcctgc tgaccgtccg gaagaacaac	1980
	ggaatgttag ttcttgcgac gcagcaaccc gagcacgtac ttgattcacc attgggcgca	2040
	agcctcgtag cacaatgtat gacaaaaatc ttctacccat caccactgc agaccgtctt	2100
	gcttacatag acgggttgaa gtgcaccgag cgagaatttc aggctattcg cgaggagatg	2160
	gcgattggca gccgcaaatt tctgctgaag cgcgagagcg gaagtgttgt gtgcgaattc	2220
	gatctgcgcg acatgccgga atacgtcgcc gtactatccg ggcgtgccaa tacggtgcgt	2280
	tttcccagc agttgcgcga gacgcattga gaggaaccct ctgcctggct ggaaaaattc	2340
	atgacgcgct accatgaggc acaggattga	2370
	<210> 84	
	<211> 669	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 84	
	atgaatatta cgaagctcgt aataaacgca cttttcatat gccttgttct gtcggggact	60
	gccaaaggcg agtttgttgt cagcgaccct gcaacggagg ccgaaacgct gacgacagcg	120
	atcaatactg cggcgaacct cgagcaattg atcacgatgg tgacaatggt gacctcgccc	180
	tttggcgtca ccggcatggt atcagcgatc gaccagaaaa accaatatcc ctctgccggc	240
	caactcgaca aggaaatggt ttccgccgag atgcctgctt cgacaactgc gcgtgcaatt	300
	accttagatg ccgatcgcg agtcgtgggc gacgacgctg aaggaaatct tttgcgacag	360
	cagattgcag gcgccgcaaa tgctgccggt gtcgcggctg acaatctgga tgcaatggac	420
	aagcgccttg cggcaaattc tgagacatcg ggccagctct cccgctcacg caatatcatg	480
	caggccaccg tcaccaacgg tttgcttctc aaacagatcc atgacgcaat tattcaaaac	540
	attcaggcga ccagcctggt gacgatgacc accgcacagg cggggctgca cgaggcggag	600
	gaggcggcga cccaacgaaa ggaacatcag gcaactgcgc ttatatattg gcccgctcaa	660
	ctacactga	669
	<210> 85	
	<211> 888	
	<212> ADN	
10	<213> Agrobacterium rhizogenes	
	<400> 85	
	atgaatttta gcattcctgc gccttttacg gcgatacata cgatattcga cctcgccttc	60
	accgtcgggc tggacacggt gcttggggac attcaacggg cagtcagcgc tcctctcgtg	120
	gcctgcgtga cgctatggat cattgtccag ggcacccctg taatgcgtgg ggaaatggac	180

ES 2 843 556 T3

	gcacggggcg gaatcacccg ggtaatcatg gtatcggtag ttgttgctct ttcggttgag	240
	caggcggaat atcatgatta cgttgtctcg gttttcgagg atacgattcc gaatttcatt	300
	caacagttcg gtatcagtgg gcttcctttg cagacgattc ctgctcagct cgacacgatg	360
	ttttcactca cgcaagtagc ctttcagaag attgcttctg aaattgggcc gatgaacgat	420
	caagacatcc tcgcctttca aggcgcacaa tggattttct atggaacgct ttggactgcg	480
	tttggaatct acgatgccgt cggaatcctc accaagggtgc tattggcgat cgggccattg	540
	atgctcggtg gctatctctt cgatcgact agagacatgg ccgcgaaatg gatcgccag	600
	ctcgtcactt atggcatact tctgcttctt ctaaacaatag tggcaactat cgtggttctg	660
	acggaagcga cggcactcgt gctaagtctc ggggttatca cttctgctgg cacaacggcg	720
	gccaagataa taggcctcta tgaactcgac atgttctttc tgaccggcga cgccttgatc	780
	gtcgtctctc cggaatcgc cgggaatatt ggaggagct attggagcgg tgcgacgcaa	840
	acggccggta gcttaaatcg ccatttcgcc cggacaatcc gccgttag	888
	<210> 86	
	<211> 153	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 86	
	atgaagtact ttctattgtt ttgatcatc ggcttggcgt cttgtcagac aagcgatcag	60
	ttggcgactt gtaaggggcc tatttttccg ctgaacgtcg ggcgatggca gcctgctcag	120
	tcggaccttc agcctaccaa cgcaggagag gct	153
	<210> 87	
	<211> 708	
10	<212> ADN	
	<213> Agrobacterium rhizogenes	
	<400> 87	
	atgaacggct ctgaatacgc gctgttagta gagcgggaag cattggcgga ccactataag	60
	gaagtagaag catttcagtc tgcacgtgct agatcagctc ggcaaatctc tagagccttg	120
	gctgcttttg ctgtcattgc agtcgcagga aacgtggcgc aggccttcgc ttcgcccgc	180
	atgcttccgc tgaacaaact tgttccagta tatctgtggg tgcgcccaga cggtagcgtt	240
	gacagtgagg tatctatttc gcggttgccg gcgacgcagg agcaagcggc cgtgaatgag	300
	tctctatggg agtatgttcg tctgcgggag agctactcgg cggatacagc ccaatacgc	360
	tatgatctgg tctcgagctt cagtgcacca acggtgcgtc aagattatca gcagttcttc	420
	aattacccca gtcctagctc ccccaaaacc atcattggta agcgcggaaa gctggaagcc	480
	gagcacatcg gctcaaacga acttatgact ggcgtccagc agatccgcta caaacgcact	540
	ctcatcatgg aagggcaagc tccaatagta accacctgga ccgcaacggc acattatgaa	600
	acggtaacta acttgcccgg ccgattgagg ctgacaaatc caggcggctt aatcgttacc	660
15	tcctatcaaa cttcggaaga taccgtttcg aacacgacgc ggagccag	708
	<210> 88	
	<211> 876	
	<212> ADN	
	<213> Agrobacterium rhizogenes	
20	<400> 88	

ES 2 843 556 T3

atgatcaaga atttgtttct gggcctggtt tgtatcctct ttgtaaccag cggcgcgaac	60
gcggaagaca cgccagcggc aggcaaactg gatccacgta tgcgctatct cgcctacaat	120
cccgatgagg ttgtgcacct ttcaaccgct gttggagcca cattggtcgt aacatttggg	180
tccaatgaag cgggtgacagc cgtcgcggtt tccaacagca aagatctcgc tgcgctccca	240
cgcggcaatt atctgttttt caaggccagc aaagtgctgc agccgcagcc tgtgatcgtc	300
cttaccgcaa gcgacgcagg aatgcggcgc tatgttttca gcctcgcgac cagaacgatg	360
tctcggctcg ataaagagca gcctgacctc tactacagcg tacagttcac ttatcctgcc	420
gacgttgccg ctgctcggcg aaaagaagcg gagcagagag atcttgcgga tcggatgcgg	480
gcgcaagcac aatatcagcg tcgagccgag gacttgctcg agcgtccgcc agcagggtggc	540
agcacagatg cgaaaaattg gagctatgtc gcgcaaggag accgctcgtc attacgctc	600
gaggttttcg acaacgggta ttccacgaca ttctcgtttc ccggaacgt gcgtgtacct	660
tcaatctacg tgattaaccc agatggcaag gaagccacgg ccaactattc agttaaggga	720
gattacgttg aagtagcatc agtttccagg gaatggcgtc tgcgggacgg ccatacagta	780
ctttgcatct ggaataaggc atatgacgcc gtcggacgga agccgggcac tggcacagtc	840
agaccgatg tagtgcgct gctgaaggag acgagg	876

<210> 89

<211> 1131

<212> ADN

5 <213> Agrobacterium rhizogenes

<400> 89

atggaggacg tgaatgcccc atccagagaa ggtatcgacg cgcccgatc cctcgtcacc	60
gatcctcatg gccggcgccct ctcggggtcg caaaagctcc ttgttgcggg tttagtcctg	120
gtactgtcgt tgagcctcat ttggctcggg gcgcgttcaa agaaaaggac cgagccgtct	180
ccaccaaaaca caatgatcga tgccaacacg aagccctttc ggccggcccc gattgatatt	240
cccgccaaac cggcaatgac accgccagct gcggaagcaa cggttttgcc gtcggatcaa	300
catcagcgag aacacaacga actgaggccc gaagaaacac cgatctttgc ctacagcgga	360
ggtgatcaga atggtgtcaa atccgcccc cagctgata ttcagaacgg tggccaagac	420

ES 2 843 556 T3

	aatagaaacg ccaactccct agccgcgctg gaggattctg ctgagaacga tctctctgtg	480
	cgcttgaaac cgacagtgtt gcagcctagc ctcgctatac ttttgccgca cccagatttc	540
	actgttacgc aaggaacgat cattccctgc atactgcaaa ccgccattga taaaaatctg	600
	gcagggtatg tgaaatgtgt acttcctcaa gacatacgcg gagcaaccgg aatgtttgtg	660
	cttctggatc gcggaacgac agttgtcggc gaaatccagc gtggactgca gcagggagac	720
	gcccgcgttt tcgttctctg gaatcgcgcg gagacgccct cacacgccgt cgtctctctt	780
	tcctccccgg gggccgacga actcgccgct tccggactgc cgggtacggt cgacaatcat	840
	ttctggaaac gcttttagcg ggccatgctc ctgagtgtcg tccaggcgcg agtccaggcg	900
	gcaagtagct acgcaggga ttcgaccggc gggaccagct tcaatagttt ccaaaacaat	960
	ggcgaacaag cggccgatac ggcctaagg gcggccataa acattcccc aaccttgaaa	1020
	aaaaatcagg gtgacacggt ctgatcttc gtcgcgcggtg atttggttt ctcaggaatc	1080
	taccagcttc atatgactgg tggatcggtg aaacgtcggc atctgcgcta a	1131
	<210> 90	
	<211> 1035	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 90	
	atggaagtaa atccgcaatt gcgcgccctt cttaaccag tcttgcaatg gctcgtgac	60
	ccgaggactg aagaagttgc cataaacgga ccgggagaag cctttgtacg ccaaagtggc	120
	gtcttcacga ggttcgccgc accgttctct tatgatgatc tcgaagacat cgccattcta	180
	gcaggggcat tgcaaaaaca agatgtcggc ccccgcaacc ctctttgcgc taccgagctc	240
	ccagggtggcg agcggatgca aatctgcttg ccgccaacgg taccctctgg caccgtgagc	300
	ttgacaattc gacgaccgag caatcgtgtc tctgaattgg gggaagtctc ggctcgctac	360
	gatgttcgct gatggaatca gtggcaaatc cggacgcagc gacgagatca actagacgaa	420
	gcgattctgc gcgactacga caatggcgat ctggaatcat ttttacgtgc atgtgttacc	480
	ggtaagcgga cgatgttgct ttgtgggccc accggaagcg gcaagaccac gatgagcaag	540
	accttgatca gtgctattcc gcgggaagaa aggtgataa ccattgaaga tacgctcgaa	600
	ctcgtcattc cacatgagaa ccacgtcaga ctgctctact ccaagagtgg ggcagggcta	660
	ggtgcagtga ctgcggaaca gctgctccag gcaagccttc ggatgcggcc tgaccgaata	720
	ttgctcgggt agatgcgcga cgacgcggcg tgggcgtacc tgagtgaggt ggtctcaggc	780
	caccccggt cgatttcaac catacatggg gcgaatccc tccaaggctt caaaaagctg	840
	ttctcacttg tgaagagcag tgtgcagggg gcatgcttgg aagatcgac actgattgac	900
	atgttggcaa cggcaatcga tgtcattgtt cccttcctg cttacggcga tgtttacgaa	960
	gtgggcgaag tatggctcgc cgccgatgcc cgccgacgcg gtgagacgat aggggatctt	1020
	ctaaaccagc aatag	1035
10	<210> 91	
	<211> 690	
	<212> ADN	
	<213> Agrobacterium rhizogenes	
	<400> 91	

ES 2 843 556 T3

	atgcgccatc ttattaccga gtatttaacg attcatgcct taagggtcac cgcggtatca	60
	gatagccagc aattcaaccg ggtacttgca tcagagacgg tcgatgtcgt cgtcgtggat	120
	ctcaatttag ggcgtgaaga cggacttgag atcgttcgca gtctcgccac aaaatccgat	180
	gttccaatta taatcatcag cggatgatgc cttgaggagg cggataaagt ggtggcgctc	240
	gagctgggag caacggattt tatcgcgaaa cctttcggaa cgcgcgagtt cctagcacgc	300
	atccgtgtcg cattacgcgt gcgaccgaat gtcgatgcga ccaaagaccg gcgttcattt	360
	tgtttcgccg gctggacact tagtctcaga cagcggcgat tggatatccg acaaaatggc	420
	gaggtgaagc ttacggcagg ggaattcaat ctgctcgtcg ctttcctgga aaagccacgc	480
	gacgtgcttt cccgagaaca gtcctcatc gcgagccgag tgcgtgagga ggaagtatat	540
	gacagaagca ttgatgttct ctttttgagg ctccgccgaa agcttgagca ggatgcagcg	600
	agcccaaact tgatcaaaac tgccaggggt gccggctatt tcttcgacgc tgatgtggat	660
	gtttcctatg gaggtctgat ggcggcctga	690
	<210> 92	
	<211> 696	
	<212> ADN	
5	<213> Agrobacterium rhizogenes	
	<400> 92	
	atgaaacttc tgacattttg ctcatthaaa ggaggcgccg gcaaaaccac ggcgctcatg	60
	ggccttttgc gtgcgtttgc aagagacggc aaaagagtgg ctctcttcga tgcggacgag	120
	aaccggccgt taacgcgatg gaaagaaaac gcgatgcgta gcaacacctg ggatacatcc	180
	tgcgagggtg acgctgcaga agaaatgtcg cttcttgaag cggcttatga ggacgccgag	240
	cttcaggaat ttgattatgc gctggccgat acgcatggtg gttcgagcga gctcaataac	300
	acaatcatcg ctagctcaaa cctgcttctg atcccgacaa tgtaaacgcc gctggatatc	360
	gatgaggcac tttcgacata tcgttatgtt gtcgagctgc tgctgagcga gaacttagtg	420
	attccgaccg ccgtgttacg ccagcgagtc ccggttgcc gattaaccac gtcacaacgc	480
	gcgatgtcag acatgctggt gaatctgct gtcgtacaat ctccaatgca tgaacgagac	540
	gcgttcgctg caatgaaaga gcgtggcatg ttgcatctca cattgctgaa cacgacaaat	600
	gatccgacga tgcgcctcct tgagcggaac ctccagaattg cgatggagga actcgtaca	660
	atttcagaat tgattggcga taccttgggg aattga	696
10	<210> 93	
	<211> 606	
	<212> ADN	
	<213> Agrobacterium rhizogenes	
	<400> 93	

ES 2 843 556 T3

```

atgggggatcc gcaaacctgc tttgtctgtc ggtgaagcaa gacggcttgc tgccgcgcga      60
ccgatcggtc acccaatccc agctctttcc tcgcaaaact tggcgcttc gcaattacct      120
gaaagagccg gaaaggaaaa acgaccagcg cctcctatcg ccgccaaacg tcgtgacaac      180
ttcgatcggc aatcaatgct aacggcggac gccttgagtt cgactgcgcc gccggaaaaa      240
gtccaggtct ttctttcggc acgccccccc gctcctgagg tatcgaagat atatgacaac      300
ttgattctgc agtacagcgc taccaaggcg ctgcagatga tcttgccgcg tgcgctcgcc      360
gattttgaga acatgttggc ggacggatca tttagcacgg ggctcaaag ctacccgatc      420
tcaaaagtag tcgaaaaacc tatcgttgtc ctgacttctc gcatgttccc ggtatcgctg      480
ttagaagtcg cacgcaatca ctttgatcca ttgagattgg agaccgccag ggcttttggc      540
cacaaactgg ctacggccgc acttgcatct ttctttgccg aagagaagat aagaaaagaa      600
cgctaa                                           606

```

<210> 94

<211> 447

<212> ADN

5 <213> Agrobacterium rhizogenes

<400> 94

```

atgtcgcaaa gcgaaaaatc cacttcaagt gacattgtcg accggcgtga aggcccgaa      60
attgaaggtt tcaaggctgt tagcactcgt ttgcgatcgt cggaatatga gagcttctct      120
cgccaagccc gtctgctggg gctctccgac agcatggcca taagagtgc ggttcgccgc      180
attgccggct ttctcgaaat cgatgcagag actcgtcaga tgatggaagc cattcttcat      240
tcaataggag cactctcaaa caacattgcc gcgctgctgt gtgcctatgc tgaaaaatccg      300
acaacggatt tgggcgctct gcaagctgaa cgaaacgctt tcggtaaata gttcgctgat      360
ctcgacgggtt tgctccgttc cattttgtcc gtatcacggc gacggatcga cggttgctcc      420
atgctggcgg acgccttgca gcattaa                                           447

```

<210> 95

<211> 1176

<212> ADN

10 <213> Agrobacterium rhizogenes

<400> 95

ES 2 843 556 T3

atgcccgatc ggtctcaagt catcatccgc attgtgccgg gaggtgggac gaagaccctc 60
caacagatta tcaaccagtt ggagtatcta tctcggaagg gcaagttgga gctgcagcgt 120
tcggcccgcac atctcgatac tttcgtacca ccggatgaaa tccgcgaact cgccctaagc 180
tgggttcaag aaaccgggac ttatcacgaa agtcggccag acgaggaaag gcaacaggag 240
ttgacgaccc acattattgt aagcttccct gccggtacaa gtcaggtagc agcttatgca 300
gcgagccggg agtgggcagc cgagatgttt gggtcaggcg cagggggagg ccgatacaac 360
tatcttacag cctttcacat cgatcgcgat cccccacatc tgcattgtgt cgttaatcgt 420
cgcaactttt tggggcacgg ctggctgaaa atatcccggc gccaccccca actgaattat 480
gacgccttgc gcattacaat ggccgagatt tcaacttcgtc atggcgctcgt cctcgaggcg 540
acgagccgag ctgagcgcg cgatctcgag cggccgatga cttttgcca gtatcgacgc 600
ctggagcggc agcaggccaa tcaaattcgt ttcgaggata ttgatttcga agaattctcg 660
cctgggggaa acgacggaga accggatcaa tcttttaatt cttcgtacgg ttcgctgcct 720
caaaacgtgt cagaaaactt gcgacgaaac ggcggtctgc aatcggtgtc tgaagtagt 780
tcccgggcac tgatcggtac gaacggcaac ggtgtcacgc ccgatcgtgt atcatctgga 840
aacgattccc gttctcaatc cggcagcgat aaagcctttg tggacgatgg caacttgagg 900
aacggccctc tcaccatcgc cggagacggc caggatcttg aaggccgttc tggcatgcat 960
cgtttgcaa ccgaaccctg caccacaca acaagcgaag atgatgttcg gcaacggcct 1020
catcgaaaac ggctcgcgg tgatgaggaa gagcagagcg gcgcaaaact gaccagggtg 1080
gacggaattc ggacgggagt gacaatcagc gccgtaccgg ttgccaggga cgatccgatt 1140
acatcgccga tccagccccc tggatcaaatt ccgttg 1176

<210> 96

<211> 3024

<212> ADN

5 <213> Agrobacterium rhizogenes

<400> 96

atgacccggg gcagaccaac atcgctacgg acgcattgcc tccgacagcc gatcgccggc 60
agcagggaga gccaaattct aagcgtcctc gtgatgatga cgccgagccg agcattcgta 120
agcgggtcaag ggacggggcg agccaagggg atgagggaaa cagaaggtag aaaaggatcg 180
atcatggcag atgaagaatt tcgtcgagac ttcagccgct ccgctacgtt gaactcgaat 240
aacgaaggcg ctacaggcgc accaatcccg gatacatcgg tgtcgactgc cttccgacaa 300
aaccgcagag aacacttaac gaccagttcc ttacggccat taattactgg tggggccggc 360
gaccttccag gcacttcccc cgaaaatcaa cggccgacaa cccatgcccc gaatacagca 420
gaaatactcc cggtgagatc agccgaagct ccggtggcc aatcagcggt ctttccgagg 480

ES 2 843 556 T3

agcaaggggt	attttcgtac	ggccatgcaa	tacctgctgg	agattgaaat	gcagtccact	540
gccagggcgg	accacgaggc	gacgtcgtcg	tcacgttttg	gcaaggggga	aagaaagcag	600
ctgcgcgaac	acgagcaaac	cggatgatcaa	agtgaatcaa	tggcgaagcg	gagtaccagc	660
tcgggaatgc	aggtgggaac	tattgccgcg	tctttggcta	tgaaagatag	ctcgaccgcg	720
acgccggcct	cgaaggggcc	gcaccggtct	gggatggagc	tgataaatat	cgaagggact	780
ggagagagcc	gtacgcaatc	tcagcacgaa	attgcgacac	aacctattgc	cacaagcgct	840
tttggttcag	gtcatgatgt	gatcatctct	ggagagggcc	agcctgaaga	aacattcgac	900
ggcgcaccat	attggatggt	agaaagcggg	gatagcgaga	ccaatagcct	tccgtcgatc	960
caacaaggct	tatcttggcc	agggttagat	gaccttgatt	tgctcgccga	cagcagttca	1020
ccgctcagta	ctcaggttgt	tcctcctgac	gtagccgcta	cgaatttaag	caactgaggca	1080
gggtccgccca	gcgtcgaaaag	cacccaaatc	ctcgccgaca	gtgcctattc	cgacgacgcg	1140
accgaaaccg	gtggccaaac	agcttcagga	tcgatccatg	aaggttcagt	tcgttcgcctt	1200
ggctcaccgg	aggaacgcct	cgacgccttt	cccagttga	ccgacgagga	cttggcgcag	1260
attgacgcat	tctcgcgagc	ccgctctcct	tcgagcgagc	atgctgcgca	cgctattcat	1320
gcggtcagct	tggtgatgac	cttctcacgc	aaggatcttg	ccgcggggcca	cgacggcccg	1380
cttacgaatc	ctgcgttccg	tcaggaggct	tctctgcagg	aatcgagccc	atcagaaacg	1440
aagtccgacg	acttcccgga	attaaccgac	gaggacctgg	cacggatcga	cgactctctc	1500
gaagcccgcct	caacttcgag	gctggatggt	gcacgcacta	ttcatgcggt	caggttgagg	1560
gcctctcggc	aggacccgag	cccacagaa	atgaagtccg	acgaactccc	cgaattgacc	1620
aaagaggact	tagcccgcat	cgacgcactc	tctggacccc	gcccgaactac	ggggcgggaa	1680
ggtgcgaaca	ctattcagcc	agtcagggtt	gacgatgctt	tctcagccga	agattatacc	1740
gtgggcccgc	gcggccggca	tacgaatcct	gcgttccatg	aggaagcttc	tcggcaggac	1800
ccgagcccat	cagagatgaa	gttcgatggc	ttgccggagt	tgaccgacga	ggatctggta	1860
aggattgacg	caatctctga	accgcgctcg	aagagatctc	gctccgagca	ggttgtcggg	1920
ggaaaatctc	ataggctcaa	actgaacgag	gctatcgagg	aggcgggctc	gccactacc	1980
gtttcgggaa	caagcgccgc	caggcgggca	cctatcacgc	tcgaaagcgg	agtacctagg	2040
ttgccgcagg	ctaccatac	ggaacttcct	ggctcgacg	ttgcgggtcc	actgcggatc	2100
gggaaggcgt	cgtgtgccga	tatacggcga	gatgtttcgg	ataaagggaa	acgaaaagtc	2160
tcgtcagatt	atcagattca	gcaaggtcct	ggcgcgtcct	atgcttacca	gccgcgacgc	2220
atgccgacgg	aagcgatttt	cgatgctgaa	cggcttgggt	acgaagtgct	catcggaat	2280
gagccagttc	cgcggtggcc	ttaccctctg	gagcacgttg	ttcgttttga	aacaggcaag	2340
cagctcttcg	aaagattgca	gaagtggccg	actgcggagg	agttggaaaa	cagtgtgatc	2400

ES 2 843 556 T3

ctcgtcaaag cggatagcgc gcaacgatcc tatatgcccc gaaatcagct gctcttgcac 2460
ctcggatccg aaaagtatta cgaacgaatg aaagcgctcg gtctatccga taataccttc 2520
agcggattga agtccgatag caaatatattt tatgatacac ctgctggacc acgggtaagt 2580
gaacatgatt atgagttccc gcggcaggca ggcgtgcata atccagaggt gatggttcag 2640
catcatagcg gaaccttctt cgcgttttgg ccaaaacttc tcaggctgga aaagcttgcg 2700
acagtaacga atacgtatgg cgcggaaaat gtttggctca aatcgccagc caatgcttac 2760
atgacgccgg aaacctataa ttccaagggt gcgaacggcg ccgaaaacct ttcatcgag 2820
ttcctacgtc ccgctcctgt ccacgaccac ctgcgggaac ctggaacttc aaacgcaaaa 2880
actgttgacg gacgaaacga agatcgcgcg caacaggccg accggttagc agatcagtcg 2940
atgggccgga acggggacgg ccgtgacgcg cagggtggta agagaacctt ggaaccgcg 3000
gtgcgcgaaa ggtatgggct gtag 3024
<210> 97
<211> 1983

<212> ADN

5 <213> *Agrobacterium rhizogenes*

<400> 97

atgaattcag gcaagtacac gccaataggc ctggctgcga gcatagcatg ttcattggcc 60
gtgggggttt gtgcggccag tctctatgtc acatttcgcc atggctttac gggcgaaacg 120
atgatgactt tcaacgtttt tgcgttcctg tacgagacac cgccttattt gggatacgca 180
agtctacgt tctatcgcg actagccatc atcgttgcga cgtcaacgct cgtgttgcta 240
tgccaactac tggtatcgat gcgcgagcgc gaacatcacg gcactgctcg ctgggccggc 300
tccggcgaga tgcggcacgc caggtacctc cggcgctaca gtcacgtcac gggcccacac 360
ttcggcaaga catgtggacc acgttggttt ggcagctacc taagtaacgg agaacaacct 420
cacagccttg ttgtcgcgcc aactcgcgct ggtaaggggt tggcgctcgt tatcccgcag 480
ttgctaacct tcaagggctc agtaattgcc ttggatgtga aaggcgagct atttgaactc 540
acatcgaggg cgcgtaaatc aagtggcgac gcggttttca agttctcgcc cttagatcca 600
gagcggcgaa cgcattgtta caatcccgtg ctggatattg ccgcgttgcc gcccgaaacgg 660
cgctttacag aaacgcgccc cctcgctgcg aatctcatca cagccaaagg caaggagacc 720
gaaggtttca tcgacggggc acgggatctt ttcgtagcag gcatattgag ttgcattgag 780
cgcggcacac caacgatcgg tgccgtttac gatctgtttg ctgagccggg tgaaggtac 840
aaactatattg cgcaactcgc ggaggaaaca caaaacaaag aggctcagcg tatcttcgac 900
aatatggcgg ggaatgacac caaaatcctg acatcctaca catcggtgct cggcgacggc 960
ggactcaatc tttgggccga tccactcgtt aaagcggcta caagcacctc ggatttttcc 1020

ES 2 843 556 T3

```

gtctacgatac tgcggcggaa gaggacttgt atttatcttt gtgtcagccc gaacgatctc 1080
gaagtcatag cgcccttgat gcgtcttctc ttccagcagg ttgtatcaat cctgcagcga 1140
tcgctgccac ttggagatga acggcacgaa gttctgtttc tcctcgatga gttcaagcac 1200
ttgggcaagc tggaagctgt ggagacggca attacgacga tcgccggcta caagggccgc 1260
ttcatgttta ttattcagag tctttcggct ttgacaggaa cttatgatga agctggtaag 1320
caaaattttc tcagcaatac ggggtgtcag gtattcatgg cgactgctga tgacgagacg 1380
cccaactaca tttccaaagc tatcggcgac tacacattcc aggctcgctc aacgtcgtaac 1440
agccaagcgc gcatgtttga tcataacatc caaatctccg accaaggagc accgcttttg 1500
cgcgccgaac aagtccgcct actcgacgac gattatgaaa tagtccttat caaaggccaa 1560
cctcccccta aactgagaaa agtgcgatat tattcggatt tcatattgaa gcgaattttc 1620
gaaagccaac acggttccct tcccagagcc gcatctttga tgttgccggg agacacaaac 1680
ctcgttgaag gcaagctcga ccagggaaca gccgatacaa cttcgcaaaa accggtgcag 1740
attgaagaga acgaccatgg ggaagtcgtt tcccccaaaa acagaactgt ggctgatgga 1800
gtaatgccaa tagaagttcg cccccgttca aacgaagttg atgacgagcg ggaataccaa 1860
ggcgcggacg ccgcaacctc tgagaagctt ccggaaattg ctccggctct tttggcgag 1920
cgcgaaactc tcgaccagat catttcgctc cagctacgag gtagaacagc ctccggcact 1980
tga 1983

```

<210> 98

<211> 2508

<212> ADN

5 <213> Agrobacterium rhizogenes

<400> 98

```

atgaaaccgt cgggaagctc aaaaactgga tacgccgggt cggctagatc ctgcgccag 60
gttcgcgcag gggttacgcc tgtccttcat ccaaagagc cttggaaccg ttttgcttgc 120
tccccatctg acggccaaat ggatcacccg gaaaatttat ctgcgcaatt cgcttacgat 180
ggaatgagac tgggcgcagc cgagcggctc gcgtacgaga catgggatag agcggaccgg 240
ccgagttgga aagacctgat actgagcgct cgcctgaacg cgattgacag tttcgcttgg 300
aatgtcgatg tcggagaaaag tacctcttca acttttgttt atgatggtgt tcctctggcg 360
gaaggggaac ggcattgccta cgaggaatgg gtggagccgg gacagcccag ttggcaacaa 420
ctcgttgtga acgcgcgtat tgaagagcta aacacctctg ctgcaattcc gaatgagtgt 480
agtccccctc aagaggaatt ccgatcggag gcgcctaagc gtaagcggac aagctccatt 540
ggtcaaaaga acggtttccc tgaaccattc gaatttgatg ggatgagact cggctcgcct 600
gagcgcgaag catatgagaa ttggagtaaa ccgcaaccgc catcctggaa agacctgata 660

```

ES 2 843 556 T3

```

gttgacgccc gtcttgacgc aatcgacacc tccacctggc tcaacgaatc aaatgacacc 720
tcggtcttcg agtacgaagg cgtcccgtt ggggaggggg aacgcctagc atatgaaaaa 780
tggctcgagc ccgcgcaacc aggatgggaa gacctggtt tggacgcacg cgttgcgga 840
ttgcaccagt ctgctccgac gtggtgcgaa caccaaccag gcaacgacgt ttcttccaac 900
gagtccgggc gcatctcggg cgtcccaatt aacgccgagc aggtgtcgcg tgcgttcttc 960
gtgtacgatg gagtggcgct cggagcagcc gagcgcgctg cgcgatgatc ttggagcagg 1020
ccggaccggc ccacctggga agatcttatc atagtcgcgc gccaaactgc tatcgaaggc 1080
ggtgccgttt cgaatgggat gatcgggaag acatcttctc cagtcttttt atacgagggg 1140
atgttgcttg gggatgcgga gcgtcaggcg tacggacggg ggaggcagct agcccagccg 1200
cgttggcaaa atctggtggt gaacgcgcgc ctggcagagc tcgaccggc ggcttgatt 1260
cccgatgaac atgatccgtc tgaggatggc ggggcgactg gtcttctgtc gcaagcaagc 1320
acgaccaata agtccgcct cagtttaggt gatcaaccgg aagcgcgctc gcctagtctc 1380
gcacgtgagc cagcacaaca gccgactcac gtgcaaaacc tgacgtgcgc acaattggaa 1440
gcaagacgtg ctctatatct cgggccctct gggagggatg cagaccaaac cgccagcatc 1500
gccgacagta atcgctcoga cgaggtaagc aaagttaaac ggctgggtgc caagagccgt 1560
cgaggcgcta aagcaacggc ctatgacgta atttcaagtg cggaaagact gtcgtctcac 1620
gagggttggt tgacggttca gtccactcag ccagaaaaag ccgcttgctc gagaagcgac 1680
aatatcggca cttatggaag tcgaaaaaac gaacgagctc ggcttgcgac cgagaccggg 1740
aaatacaggt cggagcatat ttctgggttc aaggttgctc acgataccct gcgtctgacc 1800
aaagagggcc ggcggtcgga aaggcctatg cccgcctatc tcgaatgcaa ggaacttcac 1860
cgacaacatg ttggtacggg aagggggcgg acccgactgg tggggcgcgg ctggccagat 1920
gacgcaagtt accgctcgga tcaaagggca actctatcgg accctgttgc gtccacggaa 1980
ggtgcaacgg cgtcaaacgg gtatcaatta aaccagctag gctacgcaca ccagctcgct 2040
aacgacgggc tgcagagcga aacgccagat ggggttatca tgccgcttca ggtcgcgact 2100
actagctaca actatacggg aagccgcgat ccggtacttt ctccgccag taaagagcaa 2160
gctccggaat tgttgacact tgggccacgt ggccagacag aagctgtact cgctcgcgaa 2220
acagcattga caggaagatg gccaacacgc gagcgtgagc agcaagtcta tcgagagttt 2280
ttagcccttt atgacgtcaa gaaggacctg gaggccaaga cacttgggtt gcggaagaaa 2340
aaagctgcgc ttgtttccgc gttgaaccgg actgccgcct caataggcac ttcacccttg 2400
aaagcccaat cgtcgagcgc agaagttgaa aaagcaactc acgagtttga tgaacgacgg 2460
gtttatgatc cgcgcgatcg cgtcggggac aaagcattac aacggtga 2508
<210> 99
<211> 930
<212> ADN
5 <213> Agrobacterium rhizogenes
<400> 99

```

ES 2 843 556 T3

atgcaggaac gcgagctttc ggtcctcggt ggtgataaga ccgaaagtgc agccacctta	60
ccgaatgaca ttttgggtga agttgccaaa catctgccga ccgacgatcc agtcgaaacg	120
gcagccaacc tcacaagctt taaacttgcg agcccgtcgg ttcgggcatc ggtcgatcag	180
agcgatgtcg gaacgttcca ccgaagcgta aatcgactgg gtgcctcaag caaggcattg	240
tatgacctgg ctgtaccccg gaacggcttt gccgagtatc cgaactatcc ggaaccggac	300
cctgctggcg agtttgcgct ggcgagccag cgcacagga caattggtcc aactttgaag	360
ttccaatcgc ccgccaggaa aacagcgatc gtcaatcaca ttctcaatat gtcggaaggc	420
ggtgagcagg cggaggcgat tatgtcgatg acgtccatt tgagtgcct cgagagggcc	480
gacaagagac gccttatcga cagagcaatt gaatacttca aggtagacgg gcctatcaac	540
tacgacagtc gccattatgc agcctacgag attgcagcgg cgcacaacca gttggaactc	600
gaacacaagt cacaaatctt cgatgcaatg gcgaagagac cccagcttgc gagactctac	660
gccgacgaac gcatccatgc gcaggagcat ttggggcttg cgtcgacatc cgaagggtccg	720
cacgagcgga gcaataaaca actggccacg gacattcttc agatcgaaca aaaaatccgc	780
actgagctcc ttcccgaac tctcagtgtc cacgaccaga tggtaaaagc ggaaggaatc	840
gctggatcga ttgagaaggc ttaccgcgat gctcgcgca acttaacccg ggcaccgcga	900
gggaggtcaa attctgatct cagccggtag	930

<210> 100

<211> 1356

<212> ADN

5 <213> Agrobacterium rhizogenes

<400> 100

atggtcagta taaataaaaa agccccgcag aaatcttcta cagcgaccac gcgcaattct	60
gatgaacgcc tggttgatca gtttagggcag gcgtaccttg tggaacgcga ggcgaggaa	120
aggcgcgcg cacttgaagc gccgccggcg ttgaaaaaat atgcctcgaa actgcaactc	180
ctcaataagc tcgatgcaga tttccgtggt gtgatcgccc acaagccggt tcgcagcgag	240
caactgcggg tcgatagcaa tggagaattg acccatagtc ggggtctcat aaaaaagaag	300
aagaaggttt tcattcgaga tcaacaact ggctcccttc ggcttattca ctacgagagg	360
tctgactgga atagtgtccg tcgctatgac tacgaccaga acggcggtgt gagcgagaaa	420
catatcaaaa gcaaagccgg tgccttcgaa gaaaaatggg agcgggacga aaacggtgaa	480
ctgatgcgca cccgatacat cgatcggcgc agacttactg gacgggcggt gcatccatt	540

ES 2 843 556 T3

tccgaggaaa tcggacgtcc ttatgagagt ggcctggaga agaggcttta tcgctgttta	600
acgcgtcggg agggctccca acaaaagact ttcgaacgcg atgacaaggg taatctggag	660
ctcattgcaa gcaaacggat gggctattcg atgtcttcac aaaaggcacc ggatcgacaa	720
acatctagaa caagtatccg caagcttggc ggagctttca gtgaatcata tagatcccgg	780
ctggataagg acggccacga gttaggtcgg gatgtctccg cacatcgaag cctcctcaac	840
aagcgttcag ccgtctacga tgacgtctaca ggggagctaa agagcagcaa gcacacgttt	900
gggaagatat acaaagcggg agcgacatac ctgaacgccg aaatcaaaga ggtctccaaa	960
aaaatcctcg gagttacggt cggtaaagaa ttgaagacat taagtggacg cgaactcgag	1020
gctcaagtat tacgtgccgc ggaacgtgcc ctccataaac aggcctggca gcacccttca	1080
gctaaccctt cgcggtcgca gcgggaaaat accaacagtc atcttggaag cgaacttgat	1140
gggccgctcg aatcggatat gaccgtcgtc catcaagccg gatccgggtt cgtcggcgac	1200
gatcgccacg ttcattgatg cgggtggatcc aggttgaatg aattgcacga atcgcgatcg	1260
tcgaatgata gggaaaagct tagttcgcgg agagccgaaa tttgcgaccg ctcgatgaat	1320
cgggatcgca ccggctcagg tactttgtcg cgatga	1356
<210> 101	
<211> 5310	
<212> ADN	
5 <213> Agrobacterium rhizogenes	
<400> 101	
atgccaaccg acgacattgt aatgtccgat cccggaatgg ctgctgttga cacgtctgtc	60
cctacgcgct tccagacaga tcttcgccag ttcagtaacc tgctggacga gagcaatatt	120
gtggaatggg tcgtaaacca cgccgcaa atcgcaatgcgt gcttcggaca agaccaactt	180
aagatttcgg cggtcggggc actcacgcac tacaaacggc cggtacaaga aatcgacctg	240
ctattccaac aggtcgtcgc aagcgacaaa ctgacccggc ttgattctcc agatggccaa	300
cctctctaca cctctacaga gctgagagct gcggaacaga atactgtacg tcatgtcacc	360
aatctttccg cagaagaaag cttcggcgtt ccgggaacca ttgttgatat ggtcgccgat	420
agaggcattt tcacaccgga gctcagagga gcattacatt atctgtctag gcgaaatcga	480
gcttccacta tgggtggcgt agcgggttcc gcaaaaacaa gtgtacttgg agccctgaac	540
gaagcgggtg acaggttcaa cgaggaagcg ctcccagacc agaaaataaa gttgattgcc	600
ttcgtaccca ccaaccgggc ggcccaggag cttcgtgaga aaggtcttcg cgacgtcgcg	660
acgactttca aggcgaaaga ccgcaccatc ggaaaaata ccatcgtcgt gatcgacgaa	720
atgtcgatgg ccaagacgca ggagatcgcg caactaatcg agacagttgc aaatgccact	780
tcctatcacc caaccgaaag gccgaaactg atatgtgtgg gagacgatcg gcaactgccc	840

ES 2 843 556 T3

cctgttggac ctggtgatct gcttccgatt attatggagc aagccggctg ctatgaactc	900
gtccagccgc ttcgacagct cgacgcgcgg tcgcgaatag aaacagaaaa gctcggacgt	960
gacatacgac gcgatgccaa gagcgccgtc gggcactatc tcacagcgct cgaacaaatg	1020
ggcctcgtcc atttcgtata tgctcctcct agcggctcaa acgtcagggc gtcagaccag	1080
atcactgacc gcatgctgga ggagatgaat agagtttgcg accgatatcc ggtgtatgat	1140
cgtcttgta tggcacactc caacaggagc gttaatcgcg ccaatatcgc tcttcatcag	1200
aggtttggtg acaaatccgg attgcaagac cagcaaata gtgtgtggtc aagcgtata	1260
gatagtggaa gttcggaatc tggaatagat gagggttcgt cgagacgcgt tcagcaaagg	1320
attaatctcg tcttgggaga caggattata ttcagcgaaa accgtctgca agcggacatc	1380
cgcaatggta cttttgcaac tgtagtcgga cttaatagca cagaagcttc ccgagctgag	1440
aatggcacat ctgctggggt tcgcattact gcaaggttag acggatctgg acgaaccgtc	1500
acttggaccg atagcgagtt taagggtttt acctacggct atgctgctac aatttataag	1560
agtcaggggg ccacagttga tcacgccatt ctctctctgtg atggtgctct ttcagataag	1620
ctgacctatg tcggcctgac gaggcacgc agtcgtctcg acgtcattgc ttctccctc	1680
gtggcaaaca atgttgaggc tctcagtgc aagctcactc agcaaagcgc gccaaataat	1740
tccattcaat tcccagcgct cgaacctgct gccttccaag agcgctcgtc ggtttcgaa	1800
gcgcacttgc taaggccgct cgaccagatg gaaatctctc caacgacgcc gctttccttc	1860
caaggcagac aagaaagctc cttcggggag aaccagctgc acgaaccatc cacatcggat	1920
gtggaaatgc cggaaagcgc ggacctttac ccgagcgcca ttgacttcga agttttccag	1980
tcttctgtca tcacgactga agagcaaata caaccggaat tctgcggtgc ggaatcgacg	2040
gctaacacct cagccacga aagccaccaa gtcacgatcg caacggtaac aaacctgggg	2100
atagcgatcg agaaacggaa acgagcggcg gcaaaggaag aaatcgattc tcgcaaaaaa	2160
atggcccgct actcagttag ttctgccgag gaagttgaag atatcacttc cctgttcgaa	2220
cgattgaaac tggcaacctc ggtcaccaac acagttcaac cgacgtatgc ggacgtcgat	2280
catatcatcg gtactgatgg ccgagtgggt ccaataccgc accgggaaga tgatccggga	2340
tcaaattggg aattgtcgtc aacgaacgcc ccctcaaacg tctcagagaa acgatcgacc	2400
aaagactcag gagaaaaaaa tatggcttcg aatgaccaga aaatgcaagg cgctgactct	2460
gcaggcagcg agcgtcattt ttcgcagact tcattgtgg aagccctgca ggctggcgat	2520
cacaagaagc tcgtggaaac gctcgccgca accggctccc ttgtcgagaa ggcggacgcc	2580
gggtcggcgc tgaaagcaat ggcggatgct tacgaaaggg atatcaaggc agatcccgaa	2640
aagggtccga tgctctatgc caatcggcgc gttgatgtag atgcgctggc caaggagatc	2700

ES 2 843 556 T3

gaaaaacagg gccttgcatc cggtcgcctg agtggcgacc ctgtcagctt caaggcgttc	2760
aacaacaagg atatccgttt cagggtcggc gacaggctca gcgtcaccag cagctacagg	2820
aatctcattg ccggtgaagc cggccggatc gaaaagatcg aaggcaaatc cgttaccttc	2880
aaacccgatg gcagcagtca gagcaagacc ttccaggcca gcaccggcga gccgggtgac	2940
agggggcgta agggctacaa aatcgaaggc ccgatgacgg tcatcggctc cgcctatgaa	3000
cttacgcgac atgatgacgg caaacgtgtt gaccaaactc atctcttgaa cgacgctgcc	3060
catggcatca acgccaatag cgctttgctg agcaaggccg acgagcggac gcagcttttc	3120
acttcgctca aggacaccgc cgacagggac atcttgacca ggcaactggg aaagccgggc	3180
agcgagctct accaggacaa attcccagcc ttgccggact acacagccgc catcgatcgg	3240
attgcaaaac gccaggaagt tgtcgcggca cgtggcgagg agaaaacttc cgccccgata	3300
gacctggcgc cccaaacgcc gcccgacgct gccgatgcgg cagtcaataa ccagaaaaga	3360
atcgctcagg ccttgacagg aggcgaccac aagaagcttg tggaaacgct agccgcaacc	3420
ggctcccttg ttgagaacac aaatctcgta tcaacgttga aagcgatagc gcagacttat	3480
gaaagagaca ttagtgcgga ccccgagaaa gtccgcgctc tctacgcaa caaacgtgat	3540
catgtagatt tgcttgccga agagatcgaa aagcagggcc ttgcatccgg tcgcctgagt	3600
ggcgaccctg tcagcttcaa ggcgtttaac aacaaggata tccgtttcag ggtcggtgac	3660
aggctcagcg tcaactagcag ctataggaat ctcatcgccg gtgaagccgg ccggatcgaa	3720
aagatcgaa gcaaatccgt taccttcaaa cccgatggca gcagtcaag caagaccttc	3780
gaggccagca ccggcgagcc gggtgacagg ggtgtcaagg gctataagat tgaggggccc	3840
acgacaatca tcggctccgc ctacgaaatc agaaatcacg acgatggcag gcgcgtcgat	3900
caggtttatg tcttgacag tgccattttc ggctcaaatg ccgcgagcgc cctgttgagc	3960
aaggccgacg aacgaaccca ggttttcacc gtgcaaagcg acacggctga cagacacatc	4020
ttggccagcc aactcagcaa gcctggcagc gagctctatc aggacaaatt ccagctctcg	4080
ccggactacg cagccgcgat cgaccggatc gcgaaacgcg aggaacttag ggccggacgt	4140
ggcgaggaaa gtcctactgt tcaaatggag ccggtgaacc aaacgcccta cgaggcggga	4200
cattggacgc ctgacagtca aaaatctgtg gtggaagccc tgcaggctgg cgatcacaag	4260
aagctcgtgg aaacgctcgc cgcaaccggc tcccttgctg agaaagcgga cgccgggtcg	4320
gcgctgaaag caatggcgga tgcttacgaa agggatatca aggcagatcc cgaaaaggtc	4380
cgcatgctct atgccaatcg gcgcgttgat gtagatgcgc tggccaagga gatcgaaaaa	4440
cagggccttg catccggctg cctgagtggc gacctgtca gcttcaaggc gttcaacaac	4500
aaggatatcc gtttcagggt cggcgacagg ctacagctca ccagcagcta caggaatctc	4560
attgccggtg aagccggccg gatcgaaaag atcgaaggca aatccgttac cttcaaacc	4620

ES 2 843 556 T3

	gatggcagca gtcagagcaa gaccttcgag gccagcaccg gcgagccggg tgacaggggc	4680
	gtcaagggct acaaaatcga aggcccgatg acggtcatcg gctccgccta tgaacttacg	4740
	cgacatgatg acggcaaacg tgttgaccaa atctatctct tgaacgacgc tgcccatggc	4800
	atcaacgcca atagcgcttt gctgagcaag gccgacgagc ggacgcagct ttctacttcg	4860
	ctcaaggaca ccgccgacag ggacatcttg gccaggcaac tgggaaagcc gggcagcgag	4920
	ctctaccagg acaaaattccc agccttgccg gactacacag ccgccatcga tcggattgca	4980
	aaacgccagg aagttgtcgc ggcacgtggc gaggtcaagc cgattgagcg caccgaagcg	5040
	cccgaatac cccagggtca agctactatc gttggcgaga cctcgatcat tcggttgccc	5100
	gatcggataa aaaccaatcg cgagaagccc caagaacttg atgtagtctc gaaagaaggg	5160
	cacgaaacta caggtagtct gaggcagagt gtccaagaga gtaatgcagc accaaaacag	5220
	acttcgccga aggcggcaaa tgacgtggat cggttgacac gggattttga cgagcgcac	5280
	cgcgctccgcg gggatggacg tggactctaa	5310
	<210> 102	
	<211> 1522	
	<212> ADN	
5	<213> Bordetella bronchiseptica	
	<400> 102	
	ctaccggcgc ggcagcgta cccgtgtcgg cggtccaac ggctcgccat cgtccagaaa	60
	acacggctca tcgggcatcg gcagcgctg ctgccgcgc cgttcccat cctccgtttc	120
	ggtcaaggct ggcaggtctg gttccatgcc cggaatgcc ggctggctgg gcggtcctc	180
	gccggggccg gtcggtagtt gctgctgcc cgatacagg gtcgggatgc ggcgcaggtc	240
	gcatgcccc aacagcgatt cgtcctggtc gtcgtgatca accaccacgg cggcactgaa	300
	caccgacagg cgcaactggt cgcggggctg gccccacgcc acgcggtcat tgaccacgta	360
	ggccgacacg gtgccggggc cgttgagctt cacgacggag atccagcgct cggccaccaa	420
	gtccttgact gcgtattgga ccgtccgcaa agaacgtccg atgagcttgg aaagtgtctt	480
	ctggctgacc accacggcgt tctggtggcc catctgcgcc acgaggtgat gcagcagcat	540
	tgccgcctg ggtttcctcg caataagccc ggcacacgcc tcatgcgctt tgcgttcctg	600
	ttgaccccag tgaccgggct tgttcttggc ttgaatgccg atttctctgg actgcgtggc	660
	catgcttata tccatgcggt aggggtgccg cacggttgcg gcacatgcg caatcagctg	720
	caacttttcg gcagcgcgac aacaattatg cgttgcgtaa aagtggcagt caattacaga	780
	ttttcttta cctacgcaat gagctattgc ggggggtgcc gcaatgagct gttgcgtacc	840
	cccctttttt aagttgttga ttttaagtc ttctgcattt cgccctatat ctagtctttt	900
	ggtgccccaa gaagggcacc cctgcggggt tccccacgc cttcggcgcg gctccccctc	960

ES 2 843 556 T3

cggcaaaaag tggccctcc ggggcttgtt gatcgactgc ggggccttcg gccttgccca	1020
aggtggcgct gcccccttgg aacccccgca ctcgccgccg tgaggctcgg ggggcaggcg	1080
ggcgggcttc gcccttcgac tgccccact cgcataaggct tgggtcgttc caggcgcgtc	1140
aaggccaagc cgctgcgcgg tcgctgcgcg agccttgacc cgccttcac ttggtgtcca	1200
accggcaagc gaagcgcgca gggcgaggc cggaggcttt tcccagaga aaattaaaaa	1260
aattgatggg gcaaggccgc agggcgcgca gttggagccg gtgggtatgt ggtcgaaggc	1320
tgggtagccg gtgggcaatc cctgttgtca agctcgtggg caggcgcagc ctgtccatca	1380
gcttgtccag cagggttgtc cacgggcca gcaagcgag ccagccggtg gccgctcgcg	1440
gccatcgcc acatatccac gggctggcaa gggagcgag cgaccgcga gggcgaagcc	1500
cggagagcaa gcccgtaggg gg	1522

REIVINDICACIONES

1. Un vector que comprende:
 - (a) un origen de replicación para la propagación y mantenimiento estable en *Escherichia coli*;
 - (b) un origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp.;
 - (c) un gen marcador de selección; y
 - (d) genes de virulencia de *Agrobacterium* spp.:
 - (i) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,
 - (ii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,
 - (iii) *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, y
 - (iv) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91,
 en el que el vector que comprende *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y *r-virG* que tiene la SEQ ID NO: 91 comprende además *r-galls* que tiene la SEQ ID NO: 101.
2. El vector de la reivindicación 1, que comprende además uno o más genes de virulencia de *Agrobacterium* spp.:
 - (a) *virA* que tiene la SEQ ID NO: 26 o *r-virA* que tiene la SEQ ID NO: 79;
 - (b) *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, o *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente;
 - (c) *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o *r-virE3* que tiene la SEQ ID NO: 100;
 - (d) *virH-H2* que tienen las SEQ ID NO: 42-44, respectivamente;
 - (e) *virJ* que tiene la SEQ ID NO: 27;
 - (f) *virK* que tiene la SEQ ID NO: 45;
 - (g) *virL* que tiene la SEQ ID NO: 46;
 - (h) *virM* que tiene la SEQ ID NO: 47;
 - (i) *virP* que tiene la SEQ ID NO: 48;
 - (j) *virQ* que tiene la SEQ ID NO: 49;
 - (k) *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, y *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente, y *r-virE3* que tiene la SEQ ID NO: 100;
 - (l) *virA* que tiene la SEQ ID NO: 26, *virD3-D5* que tienen las SEQ ID NO: 20-22, respectivamente, *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, y *virJ* que tiene la SEQ ID NO: 27, o *r-virA* que tiene la SEQ ID NO: 79, *r-virD3-D5* que tienen las SEQ ID NO: 96-98, respectivamente, y *r-virE3* que tiene la SEQ ID NO: 100.
3. El vector de la reivindicación 1 o 2, en el que el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* deriva de un origen de replicación de Col E1, pSC101, p15A o R6K; preferiblemente en el que:
 - (a) el origen de replicación derivado del origen de replicación de ColE1 tiene la SEQ ID NO: 2;
 - (b) el origen de replicación derivado del origen de replicación de pSC101 tiene la SEQ ID NO: 50;
 - (c) el origen de replicación derivado del origen de replicación de p15A tiene la SEQ ID NO: 51; o
 - (d) en el que el origen de replicación derivado del origen de replicación de R6K tiene la SEQ ID NO: 52.
4. El vector de una cualquiera de las reivindicaciones 1-3, en el que el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es:
 - (A) un origen de replicación de alto número de copias;
 - (B) un origen de replicación de número de copias intermedio;
 - (C) un origen de replicación de bajo número de copias;
 - (D) derivado de un origen de replicación de pRi, pVS1, pRSF1010, pRK2, pSa o pBBR1; preferiblemente en el que el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. es:
 - (a) una variante del origen de replicación de pRK2;
 - (b) derivado del origen de replicación de pRSF1010;
 - (c) derivado del origen de replicación de pVS1;
 - (d) derivado del origen de replicación de pSa; o
 - (e) un origen de replicación que tiene una cualquiera de las SEQ ID NO: 3, 37, 38, 53, 57, 58, 59 o 60; o
 - (E) un origen de replicación compatible de repABC; preferiblemente en el que el origen de replicación compatible de repABC tiene una cualquiera de las SEQ ID NO: 57, 58, 59 o 60.
5. El vector de la reivindicación 1 o 2, en el que el origen de replicación para la propagación y mantenimiento estable en *Escherichia coli* y el origen de replicación para la propagación y mantenimiento estable en *Agrobacterium* spp. son el mismo origen de replicación; preferiblemente en el que el origen de replicación deriva de un origen de replicación de pRK2, de un origen de replicación de pSa o un origen de replicación de pRSF1010.
6. El vector de la reivindicación 5, en el que el origen de replicación de pRK2 tiene la SEQ ID NO: 38;

y/o en el que el origen de replicación de pRK2 es:

(a) un microorigen de replicación de pRK2; preferiblemente en el que el microorigen de replicación de pRK2 tiene la SEQ ID NO: 54; o

(b) un miniorigen de replicación de pRK2; preferiblemente en el que el mini pRK2 tiene la SEQ ID NO: 66;

5 y/o en el que el origen de replicación de pRK2 comprende las secuencias *trfA* y *OriV*; preferiblemente en el que el origen de replicación de pRK2 comprende las SEQ ID NO: 64 y 65.

7. El vector de la reivindicación 5, en el que:

(a) el origen de replicación de pSa tiene la SEQ ID NO: 53; o

10 (b) el origen de replicación de pRSF1010 tiene la SEQ ID NO: 37.

8. El vector de una cualquiera de las reivindicaciones 5-7, que comprende además una secuencia derivada del operón par DE;

preferiblemente en el que el operón par DE tiene la SEQ ID NO: 55.

15

9. El vector de una cualquiera de las reivindicaciones 1-8, en el que el gen marcador de selección:

(A) proporciona resistencia a gentamicina, neomicina/kanamicina, higromicina o espectinomicina; preferiblemente en el que el gen marcador de selección es un gen *aacC1*, un gen *npt1*, un gen *npt2*, un gen *hpt*, un gen *aadA*, un gen *SpcN* o un gen *aph*;

20 más preferiblemente en el que:

(a) el gen marcador de selección *aacC1* tiene la SEQ ID NO: 1;

(b) el gen marcador de selección *aadA* tiene la SEQ ID NO: 39;

(c) el gen marcador de selección *npt1* tiene la SEQ ID NO: 40;

(d) el gen marcador de selección *npt2* tiene la SEQ ID NO: 41;

25 (e) el gen marcador de selección *hpt* tiene la SEQ ID NO: 67;

(f) el gen marcador de selección *SpcN* tiene la SEQ ID NO: 77; o

(g) el gen marcador de selección *aph* tiene la SEQ ID NO: 78;

(B) no proporciona resistencia a tetraciclina;

(C) no es un gen *tetAR*; o

30 (D) es un gen marcador de contraselección; preferiblemente en el que:

(a) el gen marcador de contraselección es un gen *sacB*, un gen *rpsL* (*strA*), un gen *pheS*, un gen *dhfr* (*folA*), un gen *lacY*, un gen *Gata-1*, un gen *ccdB* o un gen *thyA*;

(b) el vector no comprende la SEQ ID NO: 61;

y/o en el que el vector no comprende la SEQ ID NO: 62;

35 y/o en el que el vector no comprende una secuencia del operón *tra* o una secuencia del operón *trb*; preferiblemente en el que el vector no comprende la SEQ ID NO: 63.

10. El vector de la reivindicación 1, en el que el vector tiene:

(a) la SEQ ID NO: 34;

40 (b) la SEQ ID NO: 35; o

(c) la SEQ ID NO: 36.

11. Un vector que comprende:

(a) un origen de replicación para la propagación en *Escherichia coli* que tiene la SEQ ID NO: 2;

45 (b) un origen de replicación para la propagación en *Agrobacterium* spp. que tiene la SEQ ID NO: 3;

(c) un gen marcador de selección que tiene la SEQ ID NO: 1; y:

(I)(d) genes de virulencia de *Agrobacterium* spp.:

(i) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,

50 (ii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,

(iii) *virD1-D2* que tienen las SEQ ID NO: 18-19, respectivamente, o *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente, y

(iv) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91,

55 en el que el vector que comprende *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D2* que tienen las SEQ ID NO: 94-95, respectivamente y *r-virG* que tiene la SEQ ID NO: 91 comprende además *r-galls* que tiene la SEQ ID NO: 101;

(II)(d) genes de virulencia de *Agrobacterium* spp.:

(i) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,

60 (ii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,

(iii) *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente, o *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente,

65 (iv) *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o *r-virE3* que tiene la SEQ ID NO: 100, y

(v) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91,

- en el que el vector que comprende *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, *r-virE3* que tiene la SEQ ID NO: 100 y *r-virG* que tiene la SEQ ID NO: 91 comprende además *r-galls* que tiene la SEQ ID NO: 101; o
- (III)(d) genes de virulencia de *Agrobacterium* spp.:
- 5 (i) *virA* que tiene la SEQ ID NO: 26 o *r-virA* que tiene la SEQ ID NO: 79,
 - (ii) *virB1-B11* que tienen las SEQ ID NO: 4-14, respectivamente, o *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente,
 - (iii) *virC1-C2* que tienen las SEQ ID NO: 16-17, respectivamente, o *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente,
 - 10 (iv) *virD1-D5* que tienen las SEQ ID NO: 18-22, respectivamente, o *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente,
 - (v) *virE1-E3* que tienen las SEQ ID NO: 23-25, respectivamente, o *r-virE3* que tiene la SEQ ID NO: 100,
 - (vi) *virG* que tiene la SEQ ID NO: 15 o *r-virG* que tiene la SEQ ID NO: 91, y
 - (vii) *virJ* que tiene la SEQ ID NO: 27,
 - 15 en el que el vector comprende *r-virA* que tiene la SEQ ID NO: 79, *r-virB1-B11* que tienen las SEQ ID NO: 80-90, respectivamente, *r-virC1-C2* que tienen las SEQ ID NO: 92-93, respectivamente, *r-virD1-D5* que tienen las SEQ ID NO: 94-98, respectivamente, *r-virE3* que tiene la SEQ ID NO: 100, *r-virG* que tiene la SEQ ID NO: 91 y *virJ* que tiene la SEQ ID NO: 27 comprende además *r-galls* que tiene la SEQ ID NO: 101.
- 20 12. Un método para la transformación de una planta, que comprende las etapas de:
 - (a) poner en contacto un tejido de la planta con una cepa de *Agrobacterium* o una cepa de *Ochrobactrum*, que comprende un primer vector de una cualquiera de las reivindicaciones 1-11, y un segundo vector que comprende límites de ADN T y una secuencia polinucleotídica de interés para su transferencia a la planta;
 - (b) cocultivar el tejido con la cepa de *Agrobacterium* o la cepa de *Ochrobactrum*; y
 - 25 (c) regenerar una planta transformada a partir del tejido que expresa la secuencia polinucleotídica de interés.
 13. Un kit que comprende:
 - (a) un vector de una cualquiera de las reivindicaciones 1-11; y
 - (b) instrucciones para su uso en la transformación de una planta usando *Agrobacterium* u *Ochrobactrum*.
 - 30

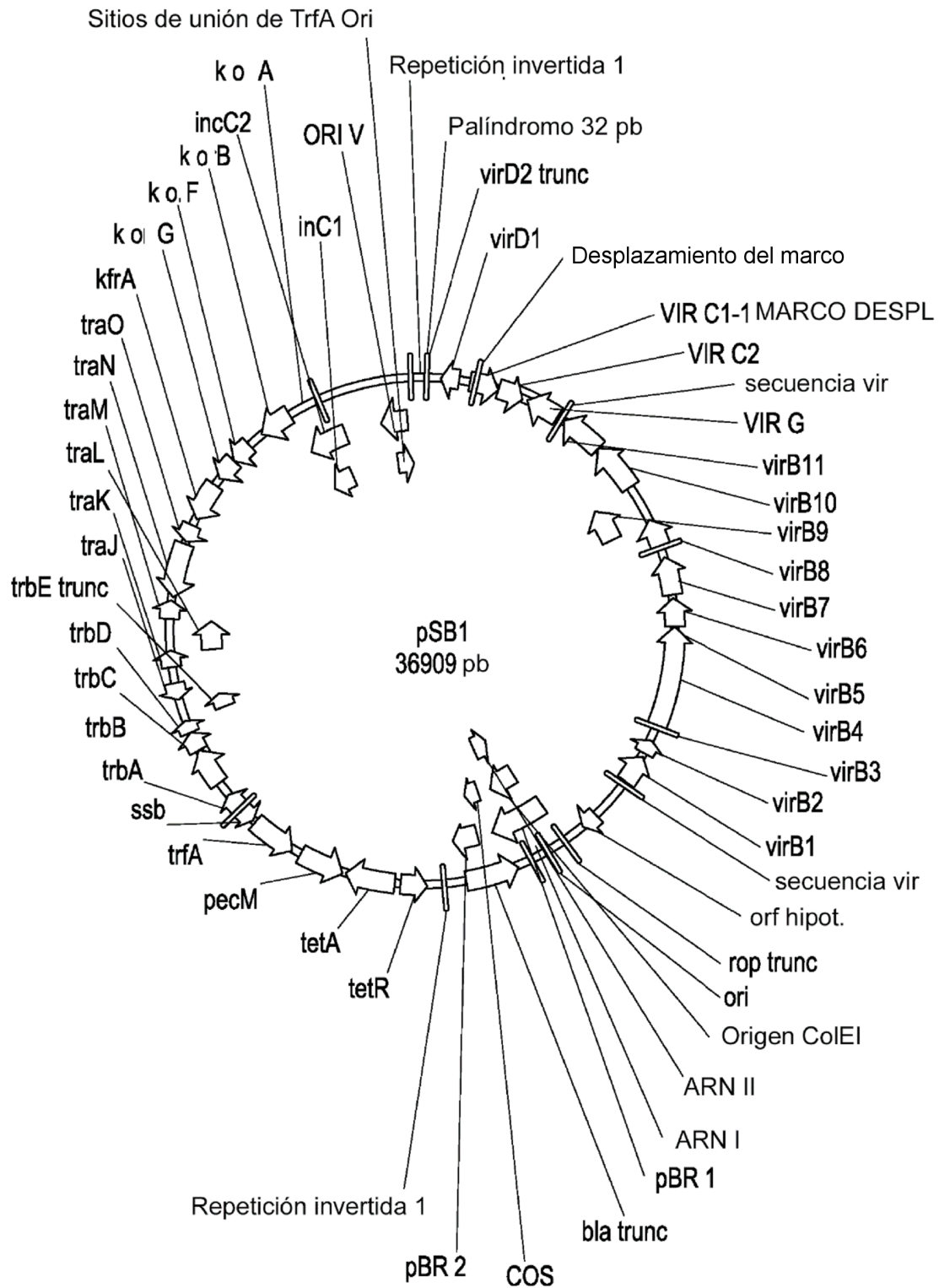


FIG. 1

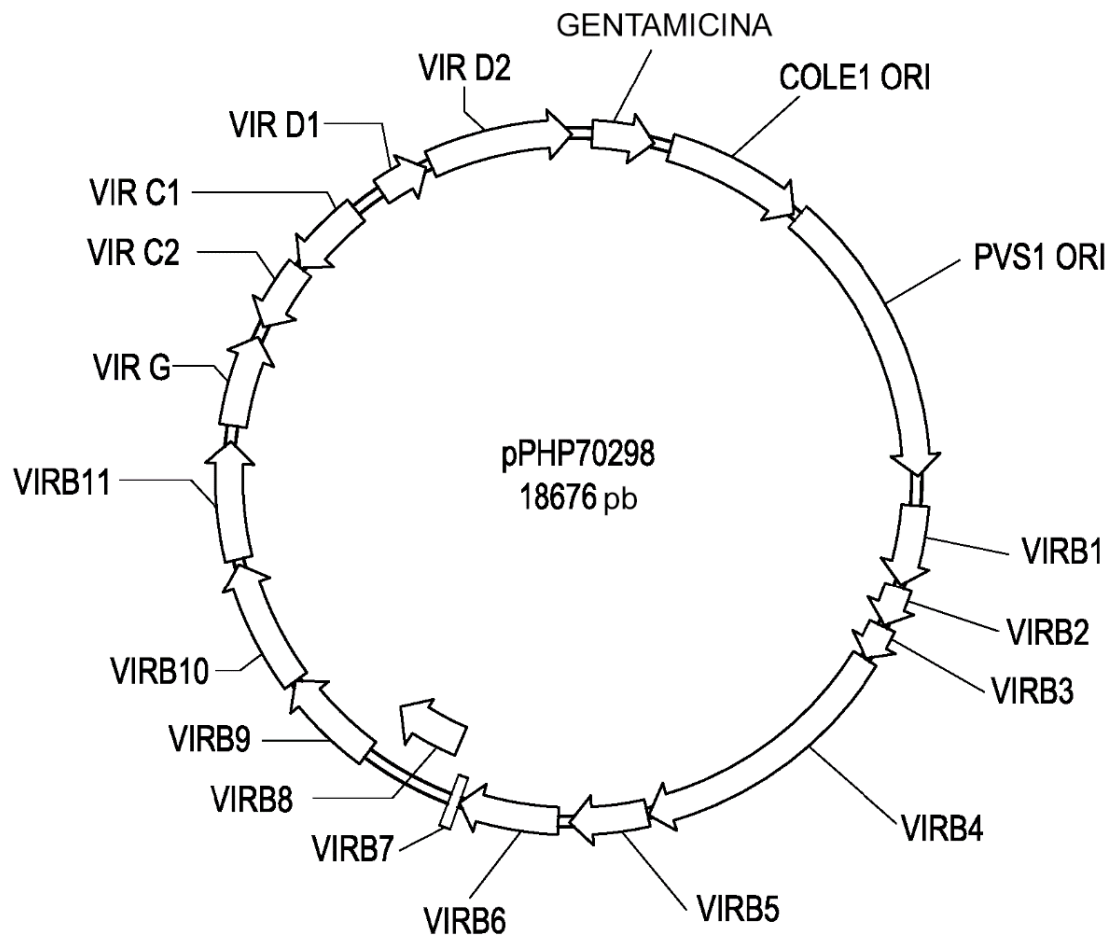


FIG. 2

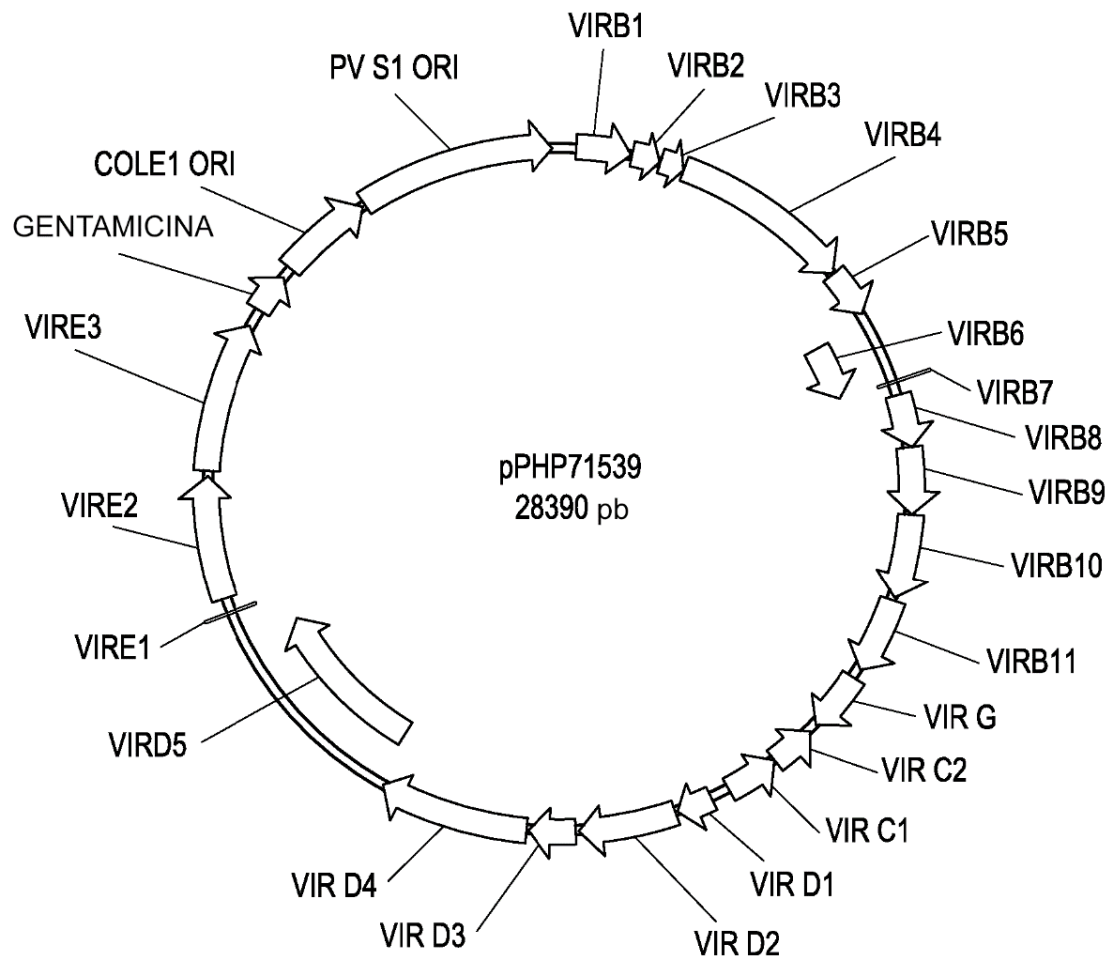


FIG. 3

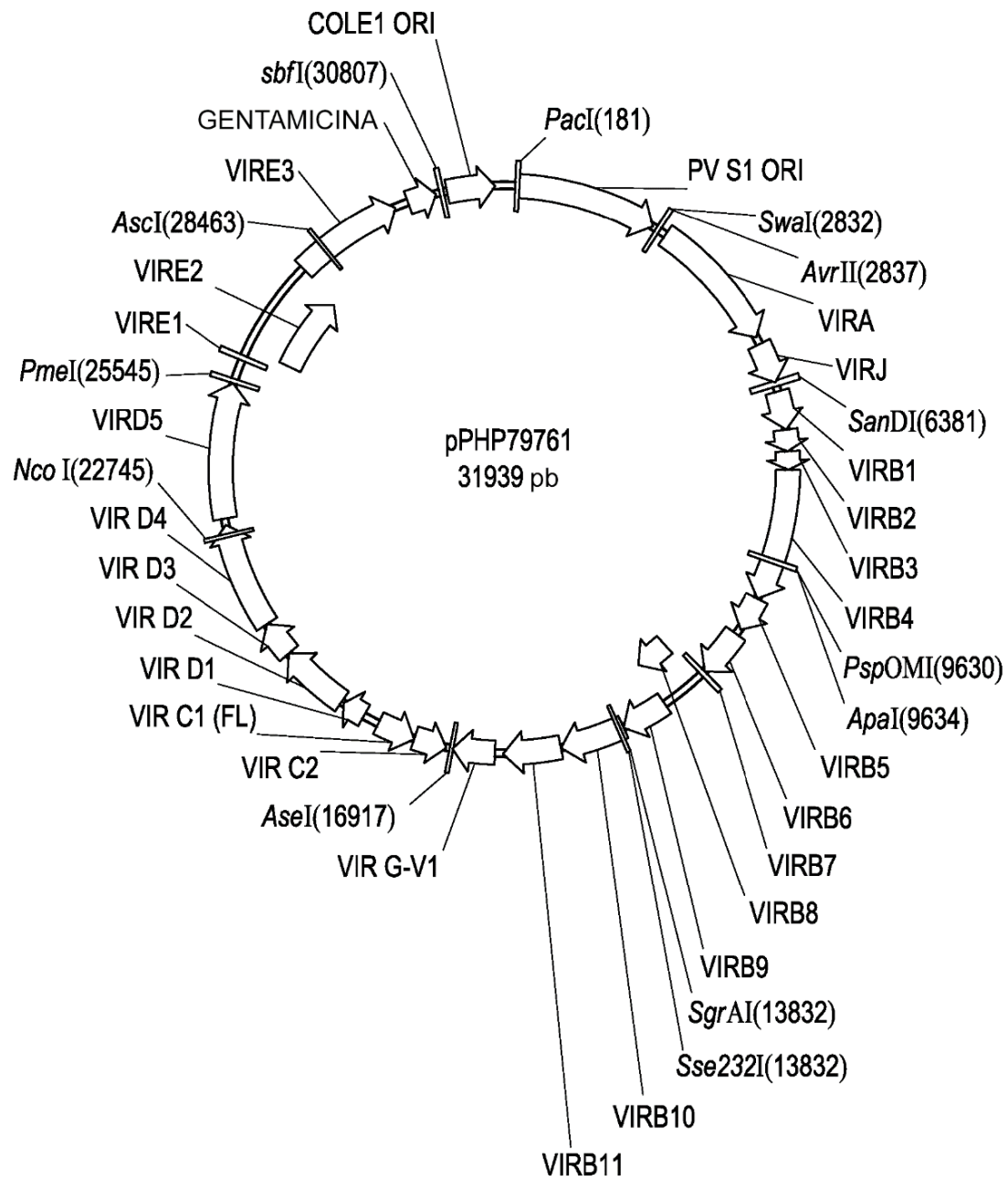


FIG. 4

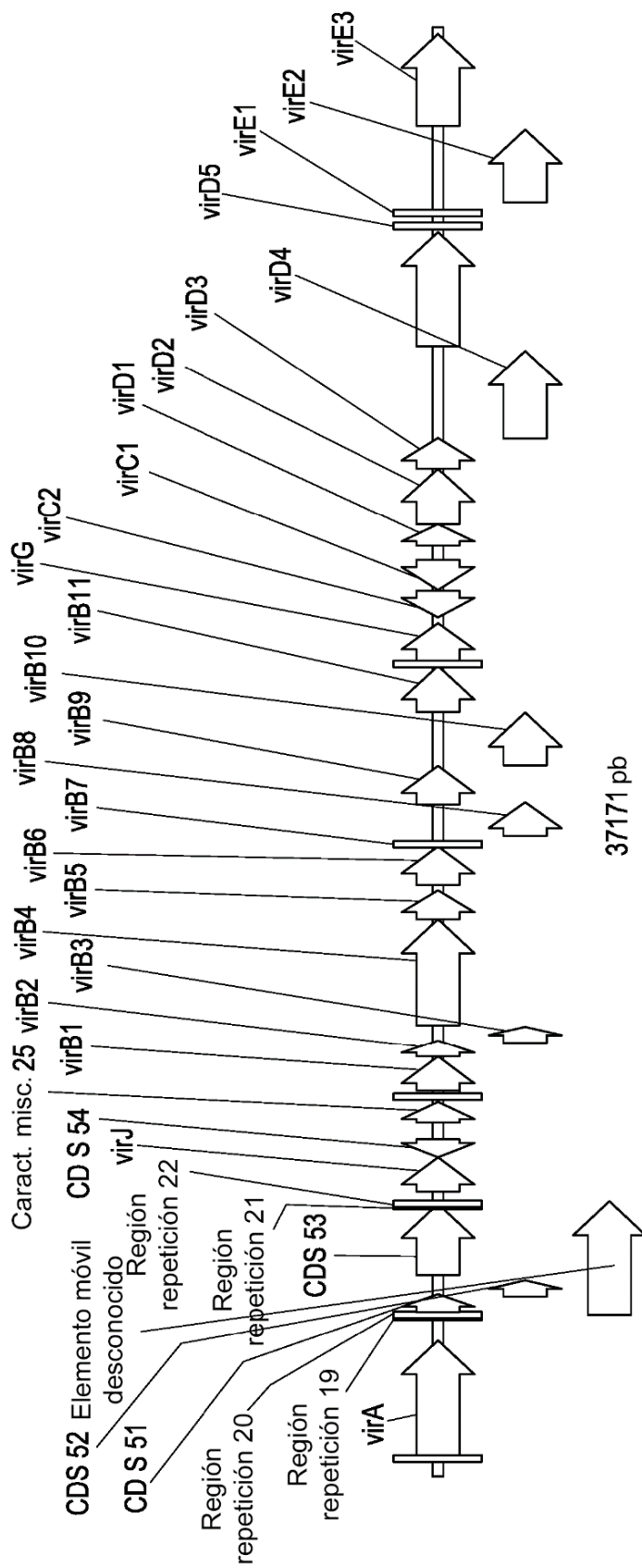


FIG. 5

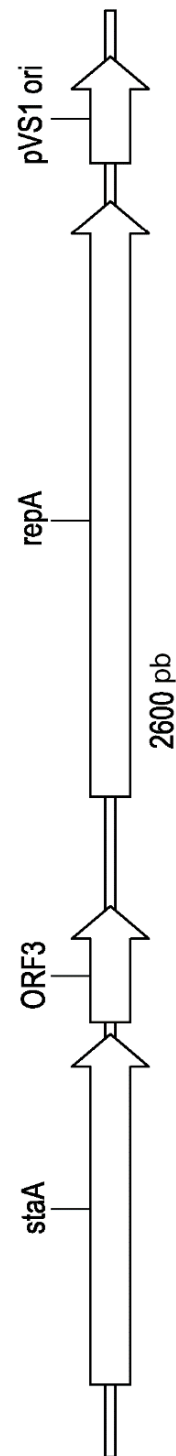


FIG. 6

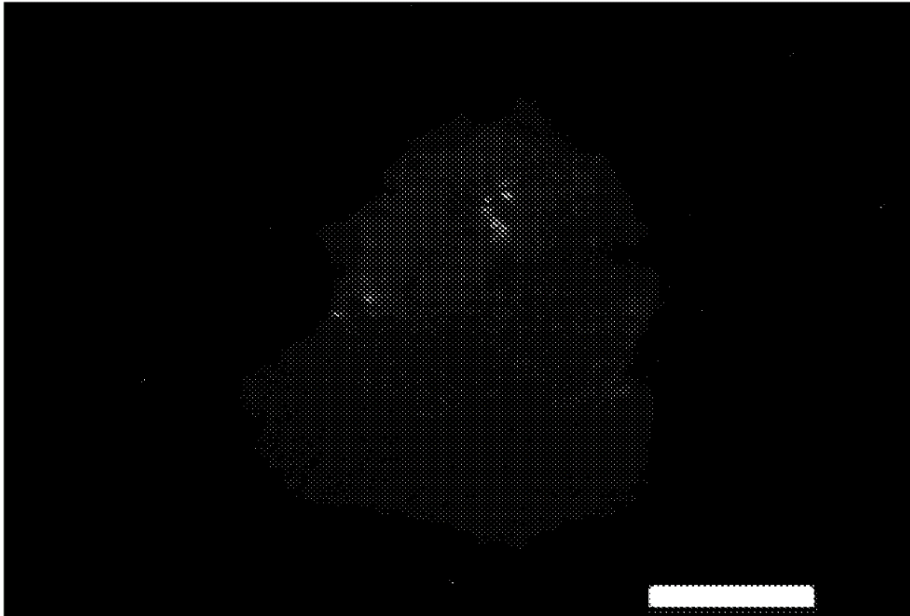


FIG. 7A

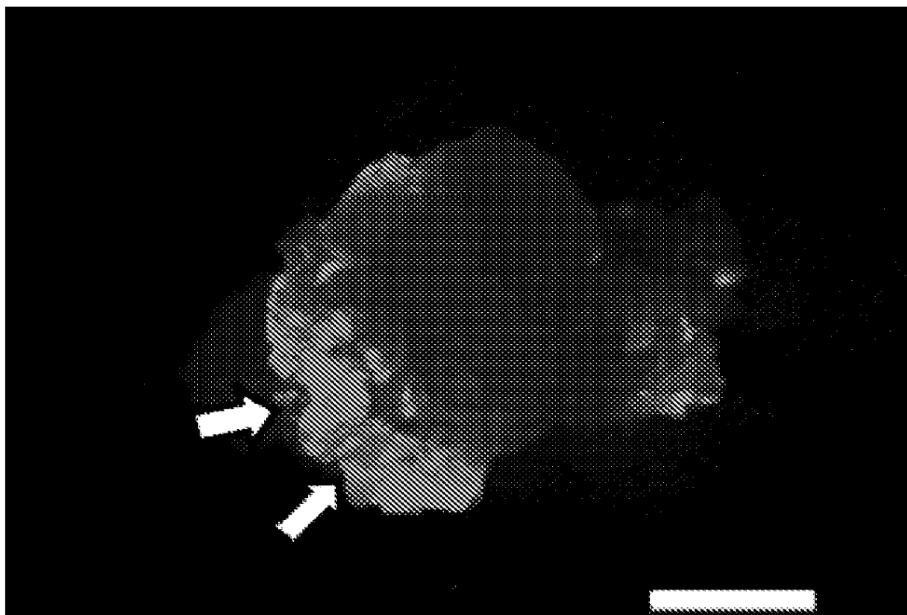


FIG. 7B