

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 657 825**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12N 5/04 (2006.01)

C12N 5/10 (2006.01)

C12N 9/16 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.05.2012 PCT/EP2012/060181**

87 Fecha y número de publicación internacional: **13.12.2012 WO12168124**

96 Fecha de presentación y número de la solicitud europea: **30.05.2012 E 12729391 (8)**

97 Fecha y número de publicación de la concesión europea: **29.11.2017 EP 2718443**

54 Título: **Métodos y medios para modificar el genoma de una planta en un sitio preseleccionado**

30 Prioridad:

06.06.2011 US 201161493579 P
06.06.2011 EP 11004570

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
07.03.2018

73 Titular/es:

BAYER CROPSCIENCE NV (100.0%)
J.E. Mommaertsiaan 14
1831 Diegem, BE

72 Inventor/es:

D'HALLUIN, KATELIJN

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 657 825 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 medios para modificar el genoma de una planta en un sitio preseleccionado

CAMPO DE LA INVENCÓN

5 [1] La invención se refiere al campo de la agronomía. Más particularmente, la invención proporciona métodos y medios para introducir una modificación dirigida, que incluye inserción, delección o sustitución, en una secuencia de nucleótidos localizada con precisión en el genoma de una planta de algodón utilizando callo embriogénico.

ANTECEDENTES DE LA TÉCNICA

10 [2] La necesidad de introducir modificaciones dirigidas en los genomas de las plantas, incluido el control sobre la ubicación de la integración del ADN extraño en las plantas, se ha vuelto cada vez más importante, y se han desarrollado varios métodos en un esfuerzo de satisfacer esta necesidad (para una revisión, véase Kumar y Fladung, 2001, *Trends in Plant Science*, 6, págs. 155-159). Estos métodos se basan principalmente en la introducción inicial de una rotura de ADN de doble cadena en la ubicación fijada como objetivo a través de la expresión de una enzima inductora de la rotura de la doble cadena (DSBI).

15 [3] Se ha demostrado que la activación del locus diana y/o la reparación o ADN del donante mediante la inducción de roturas de ADN de doble cadena (DSB) a través de endonucleasas de corte raro, tal como I-SceI, aumenta la frecuencia de la recombinación homóloga en varios órdenes de magnitud (Puchta *et al.*, 1996, *Proc. Natl. Acad. Sci. U.S.A.*, 93, págs. 5055-5060; Chilton y Que, *Plant Physiol.*, 2003; D'Halluin *et al.* 2008 *Plant Biotechnol. J.* 6, 93-102).

20 [4] El documento WO96/14408 describe un ADN aislado que codifica la enzima I-SceI. Esta secuencia de ADN puede incorporarse en vectores de clonación y expresión, líneas celulares transformadas y animales transgénicos. Los vectores son útiles en el mapeo de genes y la inserción de genes dirigida al sitio.

25 [5] El documento WO00/46386 describe métodos para modificar, reparar, atenuar e inactivar un gen u otro ADN cromosómico en una célula a través de una rotura de doble cadena inducida por I-SceI. También se describen métodos de tratamiento o profilaxis de una enfermedad genética en un individuo que lo necesite. Se describen, además, endonucleasas de restricción quiméricas.

[6] El documento WO 2005/049842 describe métodos y medios para mejorar la inserción de ADN fijado como objetivo en plantas utilizando enzimas inductoras de la "rotura de la doble cadena" (DSBI) de corte raro, así como secuencias de nucleótidos que codifican I-SceI mejoradas.

30 [7] El documento WO2006/105946 describe un método para el intercambio exacto en células vegetales y plantas de una secuencia de ADN diana para una secuencia de ADN de interés a través de la recombinación homóloga, con lo que el marcador seleccionable o rastreado utilizado durante la fase de recombinación homóloga para la selección temporal de los eventos de reemplazo de genes pueden separarse posteriormente sin dejar huella y sin recurrir al cultivo in vitro durante la etapa de separación, empleando el método descrito en el mismo para la separación de un ADN seleccionado por expresión específica de microesporas de una endonucleasa DSBI de corte raro.

35 [8] El documento WO2008/037436 describe variantes de los métodos y medios del documento WO2006/105946 en los que la etapa de separación de un fragmento de ADN seleccionado, inducida por una endonucleasa de escisión rara, inductora de una rotura de la doble cadena, está bajo el control de un promotor específico de la línea germinal. Otras realizaciones del método se basaron en la unión en los extremos no homóloga en un extremo del ADN reparador y la recombinación homóloga en el otro extremo. El documento WO08/148559 describe variantes de los métodos del documento WO2008/037436, es decir, métodos para el intercambio exacto en células eucariotas tales como células vegetales, de una secuencia de ADN diana para una secuencia de ADN de interés a través de recombinación homóloga, en donde el marcador seleccionable o rastreado utilizado durante la fase de recombinación homóloga para la selección temporal de los eventos de reemplazo génicos se puede separar subsiguientemente sin dejar huella, empleando un método para la separación de un ADN seleccionado, flanqueado por dos secuencias de nucleótidos en repeticiones directas.

40

45

[9] El documento WO 2003/004659 describe sistemas de recombinación y un método para separar secuencias de ácidos nucleicos del ADN cromosómico de organismos eucarióticos. La invención también se refiere a organismos transgénicos (preferiblemente plantas) que contienen dichos sistemas o producidos por dicho método.

50 [10] El documento WO 2006/032426 describe sistemas y métodos de recombinación mejorados para eliminar secuencias marcadoras del genoma de las plantas. Particularmente, la invención se basa en el uso de un casete de expresión que comprende el promotor de ubiquitina de perejil, y operativamente unida a él una secuencia de ácidos nucleicos que codifica una ADN-endonucleasa específica para la secuencia.

55 [11] El documento WO 2009/006297 describe métodos y composiciones para alterar el genoma de una célula vegetal monocotiledónea y una planta monocotiledónea, que implica el uso de un agente inductor de la rotura de la doble cadena para alterar una planta monocotiledónea o secuencia genómica de una célula vegetal que comprende una secuencia de reconocimiento para el agente inductor de la rotura de doble cadena.

[12] El documento WO 2004/006667 describe métodos mejorados de regeneración y transformación de algodón mediada por *Agrobacterium* mediante embriogénesis somática.

5 [13] El documento WO 2005/103271 describe métodos para la transformación de plantas de alta eficacia a través de la conjugación de ADN-T mediada por *Agrobacterium* a células cultivadas en suspensión o callos, empleando membranas de filtros como soporte sólido poroso para el co-cultivo de donantes y receptores de ADN-T.

[14] El documento WO 2008/112633 se refiere a la escisión de material de explante que comprende tejido meristemático a partir de semillas de algodón. Se describen métodos para la preparación, el almacenamiento, la transformación y la selección de tejidos o identificación de plantas transformadas, así como tejidos de meristemas transformables y plantas producidos por métodos de este tipo, y aparatos para la preparación de tejidos.

10 [15] Métodos de transformación del genoma utilizando una enzima DSBI se han aplicado en plantas tales como el tabaco (véase, p. ej., Puchta et al., 1996, Townsend et al., Nature 459:442-445, 2009) y maíz (véase, p. ej., el documento WO 2009/006297, Shukla et al., Nature 459:437-441, 2009). Sin embargo, todavía existe la necesidad de desarrollar métodos para la modificación dirigida del genoma de plantas que sean más recalcitrantes en el cultivo y la transformación de tejidos, tales como el algodón. La presente invención proporciona una contribución frente a la técnica al proporcionar este tipo de métodos, utilizando callo embriogénico de algodón para la introducción de una endonucleasa sola o en combinación con un ADN reparador que ha de utilizarse como un molde para la reparación de la rotura de ADN de doble cadena.

SUMARIO DE LA INVENCION

20 [16] En una primera realización, la invención proporciona un método para modificar el genoma de una célula de planta de algodón en un sitio predefinido, que comprende las etapas de

a. inducir una rotura de ADN de doble cadena en la vecindad de o en dicho sitio predefinido, induciéndose dicha rotura de la doble cadena mediante la introducción en dicha célula de una enzima endonucleasa de escisión rara que reconoce una secuencia de reconocimiento en la vecindad de o en dicho sitio predefinido;

25 b. seleccionar una célula vegetal en la que dicha rotura de ADN de doble cadena se ha reparado dando como resultado una modificación en el genoma en dicho sitio preseleccionado, en el que dicha modificación se selecciona de

i. un reemplazo de al menos un nucleótido;

ii. una delección de al menos un nucleótido;

30 iii. una inserción de al menos un nucleótido; o

iv. cualquier combinación de i.-iii.;

caracterizado porque dicha célula está comprendida dentro del callo embriogénico quebradizo, en donde dicho callo embriogénico quebradizo se ha mantenido en condiciones de luz tenue y en un medio que comprende carbono activo.

35 [17] En una realización, la enzima endonucleasa puede introducirse en dicha célula mediante la administración a dicha célula de una molécula de ADN que codifica dicha enzima endonucleasa.

[18] En otra realización, antes de la etapa b. se suministra una molécula de ADN reparador extraña a dicha célula, utilizándose dicha molécula de ADN reparador extraña como un molde para la reparación de dicha rotura de ADN de doble cadena.

40 [19] En una realización específica, el callo embriogénico se induce a partir de explantes de hipocotilo. El callo embriogénico puede ser inducido en un medio que comprende carbono activo. Antes, durante y después de dicha introducción de dicha enzima endonucleasa, el callo embriogénico puede incubarse en medio sin hormonas. El callo también puede incubarse en un medio sólido antes y después de dicha introducción de dicha enzima endonucleasa. Además, durante o después de dicha introducción de la enzima endonucleasa, el callo embriogénico puede incubarse durante 1 a 4 días en un medio no selectivo.

45 [20] En una realización, la enzima endonucleasa que codifica ADN y/o el ADN reparador extraño se administra a las células del callo embriogénico mediante bombardeo de partículas. El bombardeo puede realizarse con aproximadamente 0,5 pmol de ADN reparador extraño y/o aproximadamente 0,5 pmol de ADN que codifica la endonucleasa. Antes del bombardeo, el callo embriogénico puede incubarse en un medio que comprende manitol 0,2 M y sorbitol 0,2 M durante aproximadamente 2 a aproximadamente 20 horas. Además, durante o después de dicha introducción de dicha enzima endonucleasa, el callo embriogénico puede incubarse durante 1 a 4 días en un medio no selectivo que comprende manitol 0,2 M.

55 [21] En otra realización, el ADN que codifica la enzima endonucleasa y/o el ADN reparador extraño se administra dentro de la célula del callo embriogénico utilizando la transformación mediada por *Agrobacterium*. La transferencia de ADN mediada por *Agrobacterium* se puede realizar co-cultivando el callo con la o las cepas de *Agrobacterium*

que comprenden la o las moléculas de ADN durante aproximadamente tres días en un medio que comprende 100 µM de acetosiringona y/o 100 mg/l de L-cisteína. Después de la transformación, los callos pueden incubarse en un medio que comprende 250 mg/L o 125 mg/L de triacilina.

5 [22] En una realización particular, el ADN reparador extraño comprende al menos una secuencia de nucleótidos
 10 flanqueante que tiene suficiente homología con la región de ADN aguas arriba o aguas abajo de dicho sitio
 predefinido para permitir la recombinación con dicha región de ADN aguas arriba o aguas abajo. El ADN reparador
 extraño puede comprender dos secuencias de nucleótidos flanqueantes situadas en extremos opuestos de dicho
 ADN extraño, teniendo una de dichas secuencias de nucleótidos flanqueantes suficiente homología con la región de
 15 ADN aguas arriba de dicho sitio predefinido, teniendo la otra secuencia de nucleótidos flanqueante suficiente
 homología con la secuencia aguas abajo de dicho sitio predefinido para permitir la recombinación entre dichas
 secuencias de nucleótidos flanqueantes y dichas regiones de ADN aguas arriba y aguas abajo. El ADN reparador
 extraño también puede comprender un gen marcador seleccionable y/o un gen de interés expresable en plantas. El
 gen de interés expresable en plantas se puede seleccionar del grupo de un gen de tolerancia a herbicidas, un gen
 de resistencia a insectos, un gen de resistencia a enfermedades, un gen de resistencia al estrés abiótico, un gen que
 20 codifica una enzima implicada en la biosíntesis de aceite o la biosíntesis de hidratos de carbono, un gen que codifica
 una enzima implicada en la resistencia de las fibras o la longitud de las fibras, un gen que codifica una enzima
 implicada en la biosíntesis de metabolitos secundarios.

20 [23] En otra realización, el ADN extraño consiste en dos secuencias de nucleótidos flanqueantes, teniendo una de
 dichas secuencias de nucleótidos flanqueantes suficiente homología con la región de ADN aguas arriba de dicho
 sitio predefinido, teniendo la otra secuencia de nucleótidos flanqueante suficiente homología con la secuencia aguas
 abajo de dicha sitio predefinido para permitir la recombinación entre dichas secuencias de nucleótidos flanqueantes
 y dichas regiones de ADN aguas arriba y aguas abajo.

[24] En aún otra realización, el sitio preseleccionado está flanqueado por dos regiones con suficiente homología para
 la recombinación entre sí.

25 [25] La invención proporciona, además, métodos para modificar el genoma de una célula de planta de algodón en un
 sitio predefinido utilizando callo embriogénico tal como se describió arriba, en los que dicha célula de planta de
 algodón se regenera adicionalmente en una planta de algodón. La planta de algodón, así generada, puede cruzarse
 adicionalmente con otra planta.

30 [26] También se proporciona una célula de planta de algodón que comprende una modificación en un sitio
 predefinido del genoma, obtenida mediante cualquiera de los métodos descritos arriba, así como una planta de
 algodón o fibra o semilla o material de propagación de esa planta, que comprende una modificación en un sitio
 predefinido del genoma.

35 [27] La invención se refiere, además, a un método para cultivar una planta de algodón tal como se describe arriba,
 que comprende, además, la etapa de aplicar un producto químico a dicha planta o sustrato, en el que se cultiva
 dicha planta, así como un método para producir una planta que comprende una modificación en un sitio predefinido
 del genoma, que comprende la etapa de cruzar una planta que consiste esencialmente en las células vegetales tal
 como se describe arriba o una planta tal como se describe arriba con otra planta o consigo misma y, opcionalmente,
 cosechar semillas.

LEYENDAS DE LAS FIGURAS

40 [28] Figura 1. Perspectiva esquemática de la inserción/reemplazo dirigido a través de al menos recombinación
 homóloga unilateral. Las tijeras indican sitios de reconocimiento para las enzimas DSBI (I-SceI), las flechas bloque
 representan promotores, las flechas P3 y P4 representan cebadores, las cruces y las cruces punteadas representan
 la recombinación homóloga potencial entre la construcción diana pTCV117 y el ADN reparador pTIB323. Después
 45 de seleccionar la resistencia a la higromicina (hygR) y la sensibilidad al glifosato (glyphS), pueden ocurrir dos grupos
 de eventos; (I) o bien (II) el ADN reparador que confiere resistencia a la higromicina se integra en el genoma en un
 sitio aleatorio y el gen EPSPS en el locus diana se retira del control bidireccional del promotor 35S mediante
 inducción de la rotura de ADN de doble cadena en el primer (izquierda) sitio I-SceI, en cuyo caso la amplificación por
 PCR con los cebadores P3 x P4 no resulta en un producto de la PCR, o (I) el gen EPSPS en el locus diana está
 50 truncado a través de recombinación homóloga con el fragmento 3'EPSPS del ADN reparador, con lo que el gen de
 resistencia a higromicina se incorpora en el locus diana, y en donde en el otro extremo son posibles diversos
 escenarios, dependiendo de cuál de los sitios I-SceI se utilice o utilicen (indicado por el diseño rayado), permitiendo
 en cada caso amplificar la región entre los cebadores P3 y P4.

55 [29] Figura 2: Perspectiva esquemática de la inserción/reemplazo dirigido a través de la unión de los extremos no
 homóloga. La selección para la resistencia a la higromicina puede dar como resultado una integración aleatoria del
 ADN reparador pTIB236 en el genoma, con lo que la amplificación por PCR con los cebadores P3 X P4 y P1 x P2 no
 da como resultado un producto de PCR, o (I-III) dependiendo de qué sitio I-SceI se utilice, inserción del ADN
 reparador en el sitio de DSBI (I y III) o reemplazo del gen BAR por el ADN reparador (II), dando como resultado
 productos de PCR con cebadores P3xP4 y P1xP2 de diversas longitudes.

DESCRIPCIÓN DETALLADA DE DIFERENTES REALIZACIONES DE LA INVENCION

[30] Los autores de la invención han encontrado que la modificación dirigida del genoma en algodón mediante la inducción de rotura de ADN de doble cadena puede realizarse de manera eficiente mediante el uso de callo embriogénico quebradizo para la introducción de una enzima inductora de la rotura de la doble cadena (DSBI), p. ej., introduciendo una molécula de ADN que codifica una enzima de este tipo, y opcionalmente una molécula de ADN que funciona como un molde para la reparación de la rotura de ADN de doble cadena, p. ej., a través de métodos de transferencia directa de ADN (bombardeo de partículas) o mediante la administración de ADN mediado por *Agrobacterium*. Utilizando estos procedimientos de administración en callo embriogénico quebradizo, se pueden realizar inserciones, reemplazos y deleciones dirigidos en el genoma nuclear de una planta de algodón en la vecindad del sitio de inducción de rotura de la doble cadena.

[31] Tal como se utiliza en esta memoria, una endonucleasa es una enzima capaz de inducir una rotura del ADN en una secuencia de nucleótidos particular, denominada el "sitio de reconocimiento" o la "secuencia de reconocimiento", es decir, son endonucleasas específicas para el sitio. Son "de escisión rara" en el sentido de que, debido a sus sitios de reconocimiento específicos, habitualmente largos (alrededor de 12-40 nt), tienen una frecuencia de escisión muy baja, incluso en los genomas de plantas grandes, p. ej., cortan menos de 20 veces, menos de 10 veces, menos de 5 veces o solo una vez en el genoma diana. Las enzimas inductoras de la rotura del ADN de doble cadena (DSBI) de escisión rara son endonucleasas de escisión rara que inducen una rotura del ADN de doble cadena (DSB) en su sitio de reconocimiento. Las endonucleasas de migración, a veces también denominadas meganucleasas, constituyen una familia de endonucleasas de escisión rara de este tipo. Pueden ser codificadas por intrones, genes independientes o secuencias intermedias, y presentan sorprendentes propiedades estructurales y funcionales que las distinguen de las enzimas de restricción más clásicas, habitualmente de los sistemas de restricción/modificación bacteriana tipo II. Sus sitios de reconocimiento tienen una asimetría general que contrasta con la simetría de diadas característica de la mayoría de los sitios de reconocimiento de enzimas de restricción. Se ha demostrado que varias endonucleasas de migración codificadas por intrones o inteínas fomentan la migración de sus respectivos elementos genéticos en sitios alélicos sin intrones o sin inteínas. Al hacer una rotura de la doble cadena específica del sitio en los alelos sin intrones o sin inteínas, estas nucleasas crean extremos recombinogénicos, que participan en un proceso de conversión génica que duplica la secuencia codificadora y conduce a la inserción de un intrón o de una secuencia intermedia al nivel de ADN.

[32] Una endonucleasa de migración bien caracterizada es I-SceI. I-SceI es una endonucleasa específica para el sitio, responsable de la movilidad del intrón en las mitocondrias en *Saccharomyces cerevisiae*. La enzima es codificada por el intrón opcional Sc LSU.1 del gen 21S de ARNr e inicia una rotura del ADN de doble cadena en el sitio de inserción del intrón generando un corte escalonado de 4 pb con colgantes 3'OH. El sitio de reconocimiento de la endonucleasa I-SceI se extiende a lo largo de una secuencia no simétrica de 18 pb (Colleaux y col., 1988 *Proc. Natl. Acad. Sci. USA* 85: 6022-6026). La secuencia de aminoácidos para I-SceI y un código universal equivalente del gen I-SceI mitocondrial han sido proporcionados, p. ej., por el documento WO 96/14408. El documento WO 96/14408 describe, además, varias variantes de proteína I-SceI que todavía son funcionales.

[33] La solicitud PCT PCT/EP04/013122 (incorporada en esta memoria como referencia) proporciona variantes de secuencias de nucleótidos sintéticas de I-SceI que se han optimizado para la expresión en plantas. La secuencia de nucleótidos de dichas regiones codificadoras de I-SceI sintéticas se recoge en la SEQ ID N° 1 en el código UIPAC. Los símbolos del código UIPAC tienen su significado habitual, es decir, N = A o C o G o T; R = A o G; Y = C o T; B = C o G o T (no A); V = A o C o G (no T); D = A o G o T (no C); H = A o C o T (no G); K = G o T; M = A o C; S = G o C; W = A o T.

[34] En la Tabla I del documento WO 03/004659 (páginas 17 a 20) (incorporado aquí como referencia) se proporciona una lista de otras enzimas DSBI de escisión rara y sus sitios de reconocimiento respectivos. Estos incluyen I-Sce I, I-Chu I, I-Dmo I, I-Cre I, I-Csm I, PI-Fli I, Pt-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, PI-Civ I, PI-Ctr I, PI-Aae I, PI-BSU I, PI-Dhal, PI-Dra I, PI-Mav I, PI-Mch I, PI-Mfu I, PI-Mfl I, PI-Mga I, PI-Mgo I, PI-Min I, PI-Mka I, PI-Mle I, PI-Mma I, PI-Msh I, PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu I, PI-Rma I, PI-Spb I, PI-Ssp I, PI-Fac I, PI-Mja I, PI-Pho I, PI-Tag I, PI-Thy I, PI-Tko I o PI-Tsp I.

[35] Además, hay métodos disponibles para diseñar endonucleasas de escisión rara personalizadas que reconocen básicamente cualquier secuencia de nucleótidos diana de elección. En síntesis, las enzimas de restricción quiméricas se pueden preparar utilizando híbridos entre un dominio dedo de zinc diseñado para reconocer una secuencia de nucleótidos específica y el dominio de escisión de ADN no específico de una enzima de restricción natural, tal como FokI. A enzimas de este tipo se las alude generalmente como endonucleasas de dedo de zinc (ZFE). Métodos de este tipo se han descrito, p. ej., en los documentos WO 03/080809, WO94/18313 o WO95/09233 y en Isalan et al., 2001, *Nature Biotechnology* 19, 656-660; Liu et al. 1997, *Proc. Natl. Acad. Sci. USA* 94, 5525-5530). En el documento WO2004/067736 se describe otra forma de producir endonucleasas o meganucleasas de escisión rara hechas a medida, mediante selección a partir de una colección de variantes. Las meganucleasas hechas a medida con especificidad de secuencia alterada y afinidad de unión al ADN también pueden obtenerse mediante un diseño racional tal como se describe en el documento WO2007/047859. Otro ejemplo de endonucleasas de escisión rara diseñadas a medida incluye las denominadas nucleasas TALE, que se basan en efectores de tipo activador de la transcripción (TALE) del género bacteriano *Xanthomonas* condensadas al dominio catalítico de, p. ej., FOKI. La especificidad de unión al ADN de estas TALE se define por di-residuos de repetición variable (RVDs) de unidades repetitivas de 34/35 aminoácidos dispuestas en tándem, que pueden modificarse para

reconocer secuencias diana específicas (Christian et al., 2010, Genetics 186: 757-761, documentos WO11/072246, WO10/079430 y W011/146121. A dichas endonucleasas de escisión rara diseñadas a medida también se las alude como endonucleasas que se producen de forma no natural.

[36] Dado que las meganucleasas re-diseñadas se derivan de endonucleasas que se producen de forma natural, los sitios de reconocimiento del potencial disponibles no son completamente aleatorios, pero parecen tener un cierto grado de semejanza con la secuencia de nucleótidos originalmente reconocida por la endonucleasa que se producen de forma natural sobre la que se basa la meganucleasa re-diseñada. Como afirma Gao et al (2010, *The Plant Journal*, págs. 1-11), el método de diseño de proteínas basado en la estructura para modificar las características de unión a ADN de I-Crel se basa en la inspección visual de la co- estructura cristalina de co-I-Crel-ADN que conduce a una predicción de un gran número de sustituciones de aminoácidos que cambian la preferencia de base de I-Crel en posiciones particulares en su sitio de reconocimiento. Las sustituciones de aminoácidos individuales se evaluaron experimentalmente, y las que confirieron el cambio deseado en la preferencia base se añadieron a una base de datos de mutaciones que pueden "mezclarse y emparejarse" para generar derivados de I-Crel que reconocen sitios de ADN altamente divergentes. En teoría, la diversidad combinatoria disponible utilizando la base de datos de mutaciones actual es suficiente para fijar como objetivo una endonucleasa modificada aproximadamente cada 1000 pb en una secuencia de ADN aleatoria.

[37] Las meganucleasas re-diseñadas se pueden basar en la meganucleasa I-Crel que se produce de forma natural para su uso como armazón. I-Crel es una endonucleasa de migración encontrada en los cloroplastos de *Chlamydomonas reinhardtii* (Thompson et al., 1992, *Gene* 119, 247-251). Esta endonucleasa es un homodímero que reconoce un sitio de ADN pseudo-palindrómico de 22 pb en el gen 23SrRNA y crea un corte de ADN de doble cadena que se utiliza para la introducción de un intrón. I-Crel es un miembro de un grupo de endonucleasas que portan un solo motivo LAGLIDADG. Las enzimas LAGLIDADG contienen una o dos copias del motivo de consenso. Las enzimas de un solo motivo, tales como I-Crel, funcionan como homodímeros, mientras que las enzimas de doble motivo son monómeros con dos dominios separados. Por consiguiente, al re-diseñar las meganucleasas derivadas de un armazón I-Crel para reconocer una secuencia de nucleótidos de 22 pb de interés, se diseñan dos unidades monoméricas, cada una reconociendo una parte del sitio de reconocimiento de 22 pb, que se necesitan en concierto para inducir una rotura de la doble cadena en el sitio de reconocimiento de 22 pb (documento WO2007/047859). Se puede lograr una acción concertada enlazando las dos unidades monoméricas en una meganucleasa de una sola cadena, o también se puede lograr fomentando la formación de heterodímeros tal como se describe, p. ej., en el documento WO2007/047859. Ejemplos de tales meganucleasas diseñadas específicamente se describen, p. ej., en los documentos EP10005926.0 y EP10005941.9 (sin publicar).

[38] Por lo tanto, para la acción concertada de tales endonucleasas dimericas, las subunidades necesitan ser dimerizadas con el fin de poder inducir una rotura de la doble cadena en el sitio preseleccionado en el genoma. Se puede lograr una acción concertada potenciada enlazando las dos unidades monoméricas en una meganucleasa de una sola cadena, o también se puede lograr fomentando la formación de heterodímeros tal como se describe, p. ej., en el documento WO2007/047859.

[39] En la técnica se conocen diversos métodos para la administración de ADN a las células/tejidos, e incluyen la electroporación tal como se ilustra en la patente de EE.UU. N° 5.384.253; bombardeo con microproyectiles (biolística) tal como se ilustra en las Patentes de EE.UU. N°s 5.015.580; 5.550.318; 5.538.880; 6.160.208; 6.399.861; y 6.403.865; transformación mediada por *Agrobacterium* tal como se ilustra en las Patentes de EE.UU. N°s 5.635.055; 5,824,877; 5.591.616; 5.981.840; y 6.384.301; transformación de protoplastos tal como se ilustra en la Patente de EE.UU. N° 5.508.184, electroporación, transformación asistida químicamente, transformación mediada por liposomas (véase, p. ej., A. Deshayes et al. (1985) *EMBO J.* 4:2731-7), fibra de carbono, fibra de carburo de silicio o fibra de borato de aluminio (generalmente denominadas bigotes) (véase, p. ej., J. Brisibe, *Exp. Bot.* 51 (343):187-196 (2000); Dunwell (1999) *Methods Mol. Biol.* 1 11:375-82; y Patente de EE.UU. N° 5.464.765), microinyección (véase, p. ej., T.J. Reich, et al. (1986) *Biotechnology* 4: 1001-1004) y transformación mediada por virus (véase, p. ej., S.B. Gelvin, (2005) *Nat Biotechnol.* 23:684-5), bombardeo de células vegetales con ADN extraño heterólogo adherido a partículas, aceleración de partículas balísticas, transformación del haz de aerosol (Solicitud de Patente de EE.UU. N° 20010026941; Patente de EE.UU. N° 4.945.050, Publicación Internacional N° WO 91/00915; Solicitud de Patente de EE.UU. N° 2002015066, documento WO 01/038514; transformación de Lec1, transformación de PEG y diversos otros métodos no mediados por partículas directas para transferir ADN.

[40] La modificación dirigida del genoma de células de algodón para la modificación dirigida del genoma de acuerdo con la invención se realiza en callo embriogénico, preferiblemente callo quebradizo. El término "callo" o la expresión "callo embriogénico" se refiere a una masa desorganizada de células y racimos de células principalmente embriogénicos producidos como consecuencia del cultivo de tejidos vegetales. Callo quebradizo se refiere al callo con una textura quebradiza con el potencial de formar brotes y raíces y finalmente regenerarse en plantas enteras. Un callo de este tipo puede distinguirse adicionalmente por una masa de células de color verde loro/cremosas, fácilmente dispersas en medio líquido, y una forma nodular. Por lo tanto, tal como se utiliza en esta memoria, "una célula vegetal comprendida dentro de un callo embriogénico" se refiere a que la célula es una célula callosa en sí misma, es decir, esa célula es parte del tejido del callo.

[41] El callo puede regenerarse/inducirse a partir de diversos explantes de tejido tal como hipocótilo, cotiledón, embriones zigóticos inmaduros, hojas, anteras, pétalos, óvulos, raíces y meristemas, células madre y pecíolos. En

una realización de la presente invención, el explante se toma del hipocotilo o cotiledón. En una realización, la inducción del callo embriogénico se realiza incubando los explantes en un medio que comprende carbono activo durante aproximadamente 2 a 4 meses, preferiblemente 4 meses, o al menos hasta que se haya formado callo embriogénico en condiciones de luz tenue.

5 [42] En una realización, los callos se mantienen en medio sin hormonas durante todo el proceso de regeneración del callo, transferencia de ADN y posterior selección y regeneración. Hormonas, tal como se utilizan en esta memoria, se refieren a hormonas vegetales tales como auxinas, p. ej., 2,4-D y citoquininas (p. ej., Kin). En una realización, las células se mantienen en medio sólido durante todo el proceso.

10 [43] En otra realización, después de la administración de ADN, los callos se mantienen durante 1-4 días, preferiblemente 3 días, en un medio no selectivo, es decir, un medio que no contiene un compuesto de selección. El medio no selectivo puede comprender los componentes del sustrato M100. Después de 1-4 días en un medio no selectivo, los callos pueden transferirse a un medio que puede comprender los componentes del sustrato M100 y un compuesto de selección. El utilizar un compuesto de selección después de la transformación permite el enriquecimiento de los eventos de recombinación dirigida en caso de que se co-administre un ADN reparador que comprende un gen marcador seleccionable que confiere tolerancia al compuesto de selección. Después de la selección del callo embriogénico, la inducción del embrión y la germinación del embrión pueden tener lugar en un medio selectivo que puede comprender los componentes del sustrato M104 y carbono activo. El desarrollo adicional del embrión puede tener lugar en un sustrato no selectivo que puede comprender los componentes del sustrato M702 y la regeneración de plantas puede tener lugar en medio que comprende los componentes del sustrato M700. Se describen a continuación componentes de los diversos sustratos.

20 [44] Tal como se utiliza en esta memoria, la administración de ADN se refiere a la introducción de una o más moléculas de ADN en una célula. Esto se refiere tanto a la transfección estable, en la que la molécula de ADN introducida se integra de forma estable en el genoma de la célula huésped, como a la presencia transitoria de esa o esas moléculas en la célula. Quedará claro que para realizar los métodos de la invención, no se requiere que las células se transformen establemente con el ADN que codifica la endonucleasa, sino que la expresión transitoria de la endonucleasa puede ser ya suficiente para inducir la rotura de la doble cadena del ADN.

30 [45] En la técnica se conocen diversos métodos para la administración de ADN a células/tejidos, e incluyen la electroporación, tal como se ilustra en la patente de EE.UU. N° 5.384.253; bombardeo con microproyectiles (biolística) tal como se ilustra en las Patentes de EE.UU. N°s 5.015.580; 5.550.318; 5.538.880; 6.160.208; 6.399.861; y 6.403.865; transformación mediada por *Agrobacterium* tal como se ilustra en las Patentes de EE.UU. N°s 5.635.055; 5.824.877; 5.591.616; 5.981.840; y 6.384.301; transformación de protoplastos tal como se ilustra en la Patente de EE.UU. N° 5.508.184, electroporación, transformación asistida químicamente, transformación mediada por liposomas (véase, p. ej., A. Deshayes et al. (1985) *EMBO J.* 4:2731-7), fibra de carbono, fibra de carburo de silicio o fibra de borato de aluminio (generalmente denominadas bigotes) (véase, p. ej., J. Brisibe, *Exp. Bot.* 51 (343):187-196 (2000); Dunwell (1999) *Methods Mol. Biol.* 1 11:375-82; y Patente de EE.UU. N° 5.464.765), microinyección (véase, p. ej., T.J. Reich, et al. (1986) *Biotechnology* 4: 1001-1004) y transformación mediada por virus (véase, p. ej., S.B. Gelvin, (2005) *Nat Biotechnol.* 23:684-5), bombardeo de células vegetales con ADN extraño heterólogo adherido a partículas, aceleración de partículas balísticas, transformación del haz de aerosol (Solicitud de Patente de EE.UU. N° 20010026941; Patente de EE.UU. N° 4.945.050, Publicación Internacional N° WO 91/00915; Solicitud de Patente de EE.UU. N° 2002015066, documento WO 01/038514; transformación de Lec1, transformación de PEG y diversos otros métodos no mediados por partículas directas para transferir ADN.

40 [46] En una realización específica, la administración de ADN en callo que comprende las células de algodón de acuerdo con la invención se realiza mediante métodos de transferencia directa de ADN, tales como bombardeo con partículas.

45 [47] En una realización, antes del bombardeo con partículas, los callos se pre-plasmolizan en un medio que comprende manitol y sorbitol durante aproximadamente 2 a aproximadamente 20 horas, preferiblemente aproximadamente 2 a 4 horas.

50 [48] En otra realización específica, la administración de ADN a células de algodón de acuerdo con la invención se realiza por transformación mediada por *Agrobacterium*. En una realización, los callos embriogénicos se ponen en contacto con una cepa de *Agrobacterium* que contiene el ADN a ser introducido en las células de algodón, después de lo cual los callos se co-cultivan con la cepa de *Agrobacterium* en medio que comprende acetosiringona y L-cisteína durante aproximadamente 3 días en la oscuridad.

[49] En otra realización, después del co-cultivo, los callos embriogénicos transformados se seleccionan en un medio de selección (es decir, que comprende uno o más compuestos de selección) que comprende además triacilina.

55 [50] Se entenderá que el ADN que codifica endonucleasas y ADN reparador extraño se puede co-administrar a la célula o tejido (p. ej. callo) de forma secuencial (o secuencialmente inversa), utilizando los mismos o diferentes métodos de administración, o se puede co-administrar simultáneamente, p. ej., en donde el ADN reparador extraño y el ADN que codifica endonucleasas están comprendidos dentro de la misma mezcla o incluso en la misma molécula.

[51] La enzima endonucleasa puede, pero no necesita comprender una señal de localización nuclear (NLS) (Raikhel, *Plant Physiol.* 100: 1627-1632 (1992) y referencias en el mismo) tal como el NLS del antígeno T grande de SV40 (Kalderon *et al. Cell* 39: 499-509, 1984). La señal de localización nuclear puede estar ubicada en cualquier parte de la proteína, pero está ubicada convenientemente en el extremo N-terminal de la proteína. La señal de localización nuclear puede reemplazar a uno o más de los aminoácidos de la enzima inductora de la rotura de la doble cadena.

[52] La inducción de una rotura de la doble cadena en un sitio preseleccionado permite varias aplicaciones potenciales. Si no se introduce ADN reparador extraño, la región de ADN cerca del sitio de reconocimiento de endonucleasa puede alterarse por delección, reemplazo o inserción de uno o varios a muchos nucleótidos. De ese modo, la formación de delecciones o inserciones pequeñas o más grandes en el sitio preseleccionado puede, por ejemplo, inactivar el gen que comprende la secuencia de nucleótidos del sitio preseleccionado/sitio de reconocimiento. Si las regiones de ADN genómico localizadas aguas arriba y aguas abajo del sitio preseleccionado o del sitio de reconocimiento tienen suficiente homología entre sí para permitir la recombinación entre la región de ADN aguas arriba y aguas abajo, la región de ADN intermedia, es decir, la región de ADN entre las dos regiones homólogas aguas arriba y aguas abajo de ADN puede ser eliminada (en bucle). Esto se puede utilizar, por ejemplo, para separar secuencias previamente introducidas tal como genes marcadores tal como se describe, p. ej., en el documento WO 06/105946.

[53] Si la inducción de la rotura de ADN de doble cadena va acompañada de la introducción de una molécula de ADN reparador extraño que se utiliza como molde, la reparación de la rotura de la doble cadena puede producirse básicamente de tres maneras. El ADN reparador puede integrarse en el ADN genómico en el sitio DSB por unión no homóloga en ambos extremos, o si una o dos regiones flanqueantes con homología con las regiones de aguas arriba y/o aguas abajo del sitio preseleccionado están presentes en el ADN reparador, la integración del ADN reparador también puede producirse (parcialmente) a través de la recombinación homóloga. Como tal, la rotura de la doble cadena en el sitio preseleccionado facilitará también el reemplazo de una región de ADN en la vecindad de ese sitio por una región de ADN de interés, p. ej., tal como se describe en los documentos WO 06/105946, WO08/037436 o WO08/148559.

[54] Para insertar un ADN extraño mediante recombinación homóloga en el sitio preseleccionado, el ADN extraño puede comprender al menos una región de ADN flanqueante que tiene una secuencia de nucleótidos que es similar a la secuencia de nucleótidos de la región de ADN de aguas arriba o aguas abajo del sitio preseleccionado. El ADN extraño también puede comprender dos regiones flanqueantes de ADN, ubicadas en extremos opuestos de la molécula y que tienen suficiente homología con la secuencia de nucleótidos de la región de ADN de aguas arriba y aguas abajo del sitio preseleccionado, respectivamente, para permitir la recombinación entre dichas regiones flanqueantes y dicha región de aguas arriba y aguas abajo.

Como se utiliza en esta memoria, "un sitio preseleccionado" o "un sitio predefinido" indica una secuencia de nucleótidos particular en el genoma nuclear de la planta, ubicada en o cerca de la secuencia de ADN diana, en la que se desea insertar el ADN extraño o intercambiar la secuencia de ADN diana. Una persona experta en la técnica podría elegir una enzima inductora de la rotura de ADN de doble cadena ("DSBI") que reconozca la secuencia de nucleótidos diana seleccionada o modificar una endonucleasa DSBI de ese tipo. Alternativamente, se puede introducir un sitio de reconocimiento de endonucleasa DSBI en el genoma de la planta utilizando cualquier método de transformación convencional o mediante reproducción convencional utilizando una línea de planta que tenga un sitio de reconocimiento de endonucleasa DSBI en su genoma, y cualquier ADN extraño deseado puede ser introducido después de ello en el sitio diana preseleccionado, previamente introducido.

[55] Tal como se utiliza en esta memoria, "ubicado en la vecindad" se refiere al sitio de inducción de la rotura de la doble cadena del ADN, es decir, el sitio de reconocimiento de la endonucleasa, que se encuentra a una distancia de entre 500 pb, 1 kpb, 2 kpb, 3 kpb, 4 kpb, 5 kpb a 10 kpb del sitio predefinido, es decir, el sitio en el ADN genómico que se va a modificar (el sitio diana).

[56] Tal como se utiliza en esta memoria, "una región de ADN flanqueante" es un ADN con una secuencia de nucleótidos que tiene homología con las regiones de ADN, respectivamente aguas arriba y/o aguas abajo de la secuencia de ADN diana o del sitio preseleccionado. Esto permite controlar mejor la precisión de la modificación prevista. De hecho, la integración mediante recombinación homóloga permitirá una unión precisa del fragmento de ADN extraño al genoma nuclear de la planta hasta el nivel de nucleótidos.

[57] Para tener suficiente homología para la recombinación, las regiones de ADN flanqueantes del ADN reparador pueden variar en longitud, y deberían ser de al menos aproximadamente 10 nucleótidos de longitud. Sin embargo, la región flanqueante puede ser tan larga como sea prácticamente posible (p. ej., hasta aproximadamente 100-150 kb tal como cromosomas artificiales bacterianos completos (BACs)). Preferiblemente, la región flanqueante será de aproximadamente 50 pb a aproximadamente 2000 pb. Además, las regiones que flanquean el ADN extraño de interés no necesitan ser idénticas a las regiones de ADN que flanquean el sitio preseleccionado y pueden tener entre aproximadamente 80% y aproximadamente 100% de identidad de la secuencia, de preferencia aproximadamente 95% a aproximadamente 100% de identidad de la secuencia con las regiones de ADN que flanquean al sitio preseleccionado. Cuanto más larga sea la región flanqueante, menos estricto será el requisito de homología. Además, se prefiere que la identidad de secuencia sea tan alta como sea posible en la práctica en la vecindad del DSB. Además, para lograr el intercambio de la secuencia de ADN diana sin cambiar la secuencia de ADN de las

secuencias de ADN adyacentes, las secuencias de ADN flanqueantes deberían ser preferiblemente idénticas a las regiones de ADN aguas arriba y aguas abajo que flanquean el sitio preseleccionado o la secuencia de ADN diana destinada a ser intercambiada. Se aplican los mismos criterios para la recombinación entre la región aguas arriba y aguas abajo que tienen homología entre sí para separar las secuencias de ADN intermedias.

5 [58] Además, las regiones que flanquean el ADN extraño de interés no necesitan tener homología con las regiones que flanquean inmediatamente el sitio preseleccionado, sino que pueden tener homología con una región de ADN del genoma más alejada de ese sitio preseleccionado. La recombinación homóloga entre el ADN genómico y el ADN reparador dará como resultado una separación del ADN diana entre el sitio de inserción preseleccionado y la región de ADN de homología. En otras palabras, el ADN diana ubicado entre las regiones de homología será sustituido por el ADN extraño entre las regiones flanqueantes. Cuando el ADN reparador consiste sólo en las dos secuencias flanqueantes, es decir, carece de cualquier secuencia intermedia, este enfoque puede utilizarse para eliminar específicamente la región genómica ubicada entre las dos regiones de homología.

15 [59] Resultará claro que, en el caso en el que las regiones de homología estén presentes en el ADN reparador extraño, también se pueden utilizar recombinasas específicas para el sitio para llevar a cabo los métodos de la invención. Las recombinasas específicas para el sitio requieren dos sitios de reconocimiento, que pueden ubicarse en la misma molécula de ADN, pero también en dos moléculas de ADN diferentes, entre las cuales se produce la recombinación. De este modo, un ADN reparador que comprende al menos un sitio de reconocimiento de este tipo puede dirigirse a un locus genómico que también comprende al menos uno de dichos sitios. Ejemplos de recombinasas específicas para el sitio son bien conocidas en la técnica e incluyen, por ejemplo, el sistema Cre-Lox del bacteriófago P1 (Austin et al., 1981, Cell, 25:729-736), el sistema Flp-Frt de *Saccharomyces cerevisiae* (Broach et al., 1982, Cell, 29:227-234), el sistema R-RS de *Zygosaccharomyces rouxii* (Araki et al., 1985, J. Mol. Biol., 182: 191-203) y la integrasa del fago PhiC31 de *Streptomyces* (Thorpe y Smith, 1998, Proc. Natl. Acad. Sci., 95: 5505-5510; Groth et al., 2000, Proc. Natl. Acad. Sci., 97: 5995-6000).

25 [60] El ADN extraño también puede comprender un marcador seleccionable o rastreable, que puede o no separarse después de la inserción tal como se describe, p. ej., en los documentos WO06/105946, WO08/037436 y WO08/148559.

30 [61] "Marcadores seleccionables o rastreables", tal como se utilizan en esta memoria, tienen el significado habitual en la técnica e incluyen, pero no se limitan a fosfotricina acetiltransferasa, neomicina fosfotransferasa, glifosato oxidasa, enzima EPSP tolerante al glifosato, gen nitrilasa, acetolactato sintasa mutante o gen acetohidroxiácido sintasa, β -glucuronidasa (GUS), genes de locus R, proteína fluorescente verde y similares. Marcadores seleccionables pueden proporcionar tolerancia o resistencia a compuestos de selección tales como fosfotricina, neomicina, glifosato, higromicina, herbicidas inhibidores de ALS (p. ej., sulfonilurea y similares) o pueden proporcionar otros medios para seleccionar o enriquecer las células en las que ha tenido lugar la modificación deseada, p. ej., por medios visuales (tinción de GUS, fluorescencia).

35 [62] La selección de la célula vegetal o planta en la que el marcador seleccionable o rastreable y el resto de la molécula de ADN extraña se ha introducido por recombinación homóloga a través de las regiones de ADN flanqueantes puede lograrse, p. ej., seleccionando la ausencia de secuencias presentes en el ADN transformante, pero ubicadas fuera de las regiones de ADN flanqueantes. De hecho, la presencia de secuencias del ADN transformante fuera de las regiones de ADN flanqueantes indicaría que la formación de las células vegetales transformadas se realiza por inserción aleatoria de ADN. Con este fin, se pueden incluir marcadores seleccionables o rastreables en la molécula de ADN transformante fuera de las regiones de ADN flanqueantes, que luego se pueden utilizar para identificar aquellas células vegetales que no tienen los marcadores seleccionables o rastreables localizados fuera del ADN transformante y que pueden haber surgido por recombinación homóloga a través de las regiones de ADN flanqueantes. Alternativamente, la molécula de ADN transformante puede contener marcadores seleccionables fuera de las regiones de ADN flanqueantes que permiten la selección para la ausencia de tales genes (genes marcadores de selección negativos).

40 [63] Resultará claro que los métodos de acuerdo con la invención permiten la inserción de cualquier ADN de interés, incluyendo ADN que comprenda una secuencia de nucleótidos con una firma de secuencia de nucleótidos particular, p. ej., para la identificación posterior. El ADN de interés también puede ser uno o más genes expresables en plantas, incluyendo pero no limitados a un gen de tolerancia a herbicidas, un gen de resistencia a insectos, un gen de resistencia a enfermedades, un gen de resistencia al estrés abiótico, un gen que codifica una enzima implicada en la biosíntesis de aceite o la biosíntesis de hidratos de carbono, un gen que codifica una enzima implicada en la resistencia de las fibras y/o longitud de las fibras, un gen que codifica una enzima implicada en la biosíntesis de metabolitos secundarios.

55 [64] Genes de tolerancia a herbicidas incluyen un gen que codifica la enzima 5-enolpiruvil shikimato-3-fosfato sintasa (EPSPS). Ejemplos de genes EPSPS de este tipo son el gen AroA (CT7 mutante) de la bacteria *Salmonella typhimurium* (Comai et al., 1983, Science 221, 370-371), el gen CP4 de la bacteria *Agrobacterium* sp. (Barry et al., 1992, Curr. Topics Plant Physiol. 7, 139-145), los genes que codifican un EPSPS de *Petunia* (Shah et al., 1986, Science 233, 478-481), un EPSPS de Tomate (Gasser et al., 1988, J. Biol. Chem. 263, 4280-4289) o un EPSPS de Eleusina (documento WO 01/66704). También puede ser un EPSPS mutado tal como se describe, por ejemplo, en los documentos EP 0837944, WO 00/66746, WO 00/66747 o WO 02/26995. Plantas tolerantes al glifosato también

pueden obtenerse expresando un gen que codifica una enzima glifosato oxido-reductasa tal como se describe en las patentes de EE.UU. N°s 5.776.760 y 5.463.175. Plantas tolerantes al glifosato también pueden obtenerse expresando un gen que codifica una enzima glifosato acetil transferasa tal como se describe, por ejemplo, en los documentos WO 02/36782, WO 03/092360, WO 05/012515 y WO 07/024782. Plantas tolerantes al glifosato también se pueden obtener seleccionando plantas que contienen mutaciones que se producen de forma natural de los genes arriba mencionados tal como se describe, por ejemplo, en los documentos WO 01/024615 o WO 03/013226. Los genes EPSPS que confieren tolerancia al glifosato se describen, p. ej., en las solicitudes de patente de EE.UU. N°s 11/517.991, 10/739.610, 12/139.408, 12/352.532, 11/312.866, 11/315.678, 12/421.292, 11/400.598, 11/651.752, 11/681.285, 11/605.824, 12/468.205, 11/760.570, 11/762.526, 11/769.327, 11/769.255, 11/943801 o 12/362.774. Otros genes que confieren tolerancia al glifosato, tales como genes descarboxilasa, se describen, p. ej., en las solicitudes de patente de EE.UU. 11/588.811, 11/185.342, 12/364.724, 11/185.560 o 12/423.926.

[65] Otros genes de tolerancia a herbicidas pueden codificar una enzima que destoxifica el herbicida o una enzima glutamina sintasa mutante que es resistente a la inhibición, p. ej., descrita en la Solicitud de Patente de EE.UU. N° 11/760.602. Una de estas enzimas desintoxicantes eficientes es una enzima que codifica una fosfinotricina acetiltransferasa (tal como la proteína bar o pat de especies de *Streptomyces*). Las fosfinotricina acetiltransferasas se describen, p. ej., en las patentes de EE.UU. N°s 5.561.236; 5.648.477; 5.646.024; 5.273.894; 5.637.489; 5.276.268; 5.739.082; 5.908.810 y 7.112.665.

[66] Los genes de tolerancia a herbicidas también pueden conferir tolerancia a los herbicidas que inhiben la enzima hidroxifenilpiruvato dioxigenasa (HPPD). Hidroxifenilpiruvato dioxigenasas son enzimas que catalizan la reacción en la que el para-hidroxifenilpiruvato (HPP) se transforma en homogentisato. Plantas tolerantes a inhibidores de HPPD pueden transformarse con un gen que codifica una enzima HPPD resistente que se produce de forma natural, o un gen que codifica una enzima HPPD mutada o química tal como se describe en los documentos WO 96/38567, WO 99/24585 y WO 99/24586, WO 2009/144079, WO 2002/046387 o US 6.768.044. La tolerancia a inhibidores de HPPD también puede obtenerse mediante la transformación de plantas con genes que codifican determinadas enzimas que permiten la formación de homogentisato a pesar de la inhibición de la enzima HPPD nativa por parte del inhibidor de HPPD. Dichas plantas y genes se describen en los documentos WO 99/34008 y WO 02/36787. La tolerancia de plantas a inhibidores de HPPD también se puede mejorar transformando plantas con un gen que codifica una enzima que tiene actividad de prefenato deshidrogenasa (PDH) además de un gen que codifica una enzima tolerante a HPPD, tal como se describe en el documento WO 2004/024928. Además, las plantas pueden hacerse más tolerantes a los herbicidas inhibidores de HPPD añadiendo a su genoma un gen que codifica una enzima capaz de metabolizar o degradar inhibidores de HPPD tales como las enzimas CYP450 mostradas en los documentos WO 2007/103567 y WO 2008/150473.

[67] Todavía otros genes de tolerancia a herbicidas codifican enzimas ALS variantes (también conocidas como acetohidroxiácido sintasa, AHAS) tal como se describe, por ejemplo, en Tranel y Wright (2002, *Weed Science* 50:700-712), pero también, en las patentes de EE.UU. N°s 5.605.011, 5.378.824, 5.141.870 y 5.013.659. La producción de plantas tolerantes a sulfonilurea y plantas tolerantes a imidazolinona se describe en las Patentes de EE.UU. N°s 5.605.011; 5.013.659; 5.141.870; 5.767.361; 5.731.180; 5.304.732; 4.761.373; 5.331.107; 5.928.937; y 5.378.824; y publicación internacional WO 96/33270. Otros genes de tolerancia a imidazolinona también se describen, por ejemplo, en los documentos WO 2004/040012, WO 2004/106529, WO 2005/020673, WO 2005/093093, WO 2006/007373, WO 2006/015376, WO 2006/024351 y WO 2006/060634. Otros genes de tolerancia a sulfonilurea e imidazolinona se describen, por ejemplo, en el documento WO 07/024782 y la solicitud de patente de EE.UU. N° 61/288958.

[68] El gen de resistencia a insectos puede comprender una secuencia codificante que codifica:

1) una proteína cristal insecticida de *Bacillus thuringiensis* o una porción insecticida de la misma, tal como las proteínas cristal insecticidas enumeradas por Crickmore et al. (1998, *Microbiology and Molecular Biology Reviews*, 62: 807-813), actualizado por Crickmore et al. (2005) en la nomenclatura de la toxina de *Bacillus thuringiensis*, en línea en: http://www.lifesci.sussex.ac.uk/Home/Neil_Crickmore/Bt/, o porciones insecticidas de la misma, p. ej., proteínas de las clases de proteínas Cry Cry1Ab, Cry1Ac, Cry1B, Cry1C, Cry1D, Cry1F, Cry2Ab, Cry3Aa o Cry3Bb o porciones insecticidas de las mismas (p. ej., documentos EP 1999141 y WO 2007/107302), o proteínas codificadas por genes sintéticos tal como se describen p. ej., en la Solicitud de Patente de EE.UU. N° 12/249.016; o

2) una proteína cristal de *Bacillus thuringiensis* o una porción de la misma que es insecticida en presencia de una segunda otra proteína cristal de *Bacillus thuringiensis* o una porción de la misma, tal como la toxina binaria compuesta por las proteínas cristal Cry34 y Cry35 (Moellenbeck et al. 2001, *Nat. Biotechnol.* 19: 668-72; Schnepf et al., 2006, *Applied Environm. Microbiol.* 71, 1765-1774) o la toxina binaria compuesta por las proteínas Cry1A o Cry1F y las proteínas Cry2Aa o Cry2Ab o Cry2Ae (Solicitud de Patente de EE.UU. N° 12/214.022 y documento EP 08010791.5); o

3) una proteína insecticida híbrida que comprende partes de diferentes proteínas cristal insecticidas de *Bacillus thuringiensis* tal como un híbrido de las proteínas de 1) anteriores o un híbrido de las proteínas de 2) anteriores, p. ej., la proteína Cry1A.105 producida por un evento de maíz MON89034 (documento WO 2007/027777); o

4) una proteína de uno cualquiera de 1) a 3) anterior, en donde algunos, particularmente 1 a 10 aminoácidos han sido reemplazados por otro aminoácido para obtener una mayor actividad insecticida para una especie de insecto

diana, y/o para expandir la gama de especies de insectos diana afectadas, y/o debido a cambios introducidos en el ADN codificante durante la clonación o transformación, tal como la proteína Cry3Bb1 en los eventos de maíz MON863 o MON88017, o la proteína Cry3A en el evento de maíz MIR604; o

5) una proteína insecticida secretada de *Bacillus thuringiensis* o *Bacillus cereus*, o una porción insecticida de la misma, tal como las proteínas insecticidas vegetativas (VIP) enumeradas en: http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Btip.html, p. ej., proteínas de la clase de proteínas VIP3Aa; o

6) una proteína secretada de *Bacillus thuringiensis* o *Bacillus cereus* que es insecticida en presencia de una segunda proteína secretada de *Bacillus thuringiensis* o *B. cereus*, tal como la toxina binaria compuesta por las proteínas VIP1A y VIP2A (documento WO 94/21795); o

7) una proteína insecticida híbrida que comprende partes de diferentes proteínas secretadas de *Bacillus thuringiensis* o *Bacillus cereus*, tal como un híbrido de las proteínas en 1) anteriores o un híbrido de las proteínas en 2) anteriores; u

8) una proteína de uno cualquiera de 5) a 7) anteriores, en donde algunos, particularmente 1 a 10 aminoácidos han sido reemplazados por otro aminoácido para obtener una mayor actividad insecticida para una especie de insecto diana, y/o para expandir la gama de especies de insectos diana afectados, y/o debido a cambios introducidos en el ADN codificante durante la clonación o transformación (aunque todavía codifican una proteína insecticida) tal como la proteína VIP3Aa en el evento de algodón COT102; o

9) una proteína secretada de *Bacillus thuringiensis* o *Bacillus cereus* que es insecticida en presencia de una proteína cristal de *Bacillus thuringiensis* tal como la toxina binaria compuesta por VIP3 y Cry1A o Cry1F (Solicitudes de Patente de EE.UU. N°s 61/126083 y 61/195019), o la toxina binaria compuesta por la proteína VIP3 y las proteínas Cry2Aa o Cry2Ab o Cry2Ae (Solicitud de Patente de EE.UU. N° 12/214.022 y documento EP 08010791.5);

10) una proteína de 9) anterior, en donde algunos, particularmente 1 a 10 aminoácidos han sido reemplazados por otro aminoácido para obtener una mayor actividad insecticida para una especie de insecto diana y/o para expandir la gama de especies de insectos diana afectados, y/o debido a cambios introducidos en el ADN codificante durante la clonación o la transformación (aunque todavía codifican una proteína insecticida).

[69] Un "gen resistente a insectos" tal como se utiliza en esta memoria, incluye, además, transgenes que comprenden una secuencia que produce, tras la expresión, un ARN de doble cadena que tras la ingestión por una plaga de insectos de plantas inhibe el crecimiento de esta plaga de insectos tal como se describe, p. ej., en los documentos WO 2007/080126, WO 2006/129204, WO 2007/074405, WO 2007/080127 y WO 2007/035650.

[70] Genes de tolerancia al estrés abiótico incluyen

1) un transgén capaz de reducir la expresión y/o la actividad del gen de la poli(ADP-ribosa) polimerasa (PARP) en las células vegetales o plantas tal como se describe en los documentos WO 00/04173, WO/2006/045633, EP 04077984.5 o EP 06009836.5.

2) un transgén capaz de reducir la expresión y/o la actividad de los genes que codifican PARG de las células vegetales o plantas tal como se describe, p. ej., en el documento WO 2004/090140.

3) un transgen que codifica una enzima funcional de la planta de la vía de síntesis de salvado del dinucleótido de nicotinamida adenina incluyendo nicotinamidasa, nicotinato fosforribosiltransferasa, mononucleótido adenil transferasa de ácido nicotínico, nicotinamida adenina dinucleótido sintetasa o nicotinamida fosforilsiltransferasa tal como se describe, p. ej., en los documentos EP 04077624.7, WO 2006/133827, PCT/EP07/002433, EP 1999263 o WO 2007/107326.

[71] Enzimas implicadas en la biosíntesis de hidratos de carbono incluyen las descritas, p. ej., en los documentos EP 0571427, WO 95/04826, EP 0719338, WO 96/15248, WO 96/19581, WO 96/27674, WO 97/11188, WO 97/26362, WO 97/32985, WO 97/42328, WO 97/44472, WO 97/45545, WO 98/27212, WO 98/40503, WO 99/58688, WO 99/58690, WO 99/58654, WO 00/08184, WO 00/08185, WO 00/08175, WO 00/28052, WO 00/77229, WO 01/12782, WO 01/12826, WO 02/101059, WO 03/071860, WO 2004/056999, WO 2005/030942, WO 2005/030941, WO 2005/095632, WO 2005/095617, WO 2005/095619, WO 2005/095618, WO 2005/123927, WO 2006/018319, WO 2006/103107, WO 2006/108702, WO 2007/009823, WO 00/22140, WO 2006/063862, WO 2006/072603, WO 02/034923, EP 06090134.5, EP 06090228.5, EP 06090227.7, EP 07090007.1, EP 07090009.7, WO 01/14569, WO 02/79410, WO 03/33540, WO 2004/078983, WO 01/19975, WO 95/26407, WO 96/34968, WO 98/20145, WO 99/12950, WO 99/66050, WO 99/53072, US 6.734.341, WO 00/11192, WO 98/22604, WO 98/32326, WO 01/98509, WO 01/98509, WO 2005/002359, US 5.824.790, US 6.013.861, WO 94/04693, WO 94/09144, WO 94/11520, WO 95/35026 o WO 97/20936 o enzimas implicadas en la producción de polifruktosa, especialmente del tipo inulina y levan, tal como se describe en los documentos EP 0663956, WO 96/01904, WO 96/21023, WO 98/39460 y WO 99/24593, la producción de alfa-1,4-glucanos tal como se describe en los documentos WO 95/31553, US 2002031826, US 6.284.479, US 5.712.107, WO 97/47806, WO 97/47807, WO 97/47808 y WO 00/14249, la producción de alfa-1,6 alfa ramificado-1,4-glucanos tal como se describe en el documento WO 00/73422, la producción de alternano tal como se describe, p. ej., en los documentos WO 00/47727, WO 00/73422, EP 06077301.7, US 5.908.975 y EP 0728213, la producción de hialuronano tal como se describe, p. ej., en los

documentos WO 2006/032538, WO 2007/039314, WO 2007/039315, WO 2007/039316, JP 2006304779 y WO 2005/012529.

[72] La persona experta en la técnica apreciará que, además del genoma nuclear, los métodos de la invención también se pueden aplicar para modificar, p. ej., el genoma del cloroplasto o el genoma mitocondrial, por lo que la inducción de DSB en el sitio predefinido puede ser adicionalmente potenciado al proporcionar la señal de fijación de objetivo correcta a la enzima endonucleasa.

[73] También se describen células de plantas de algodón y plantas generadas de acuerdo con los métodos de la invención. Plantas de este tipo pueden contener una secuencia de ADN heterólogo o extraño insertada en o en lugar de una secuencia diana o pueden contener una delección, y sólo serán diferentes de sus plantas progenitoras por la presencia de la modificación particular.

[74] Tal como se describe, las células vegetales de la invención, es decir, una célula vegetal que comprende la combinación de ADN-T así como células vegetales generadas de acuerdo con los métodos descritos en esta memoria que comprenden la modificación genómica deseada, pueden ser células no propagantes.

[75] Las plantas de algodón obtenidas mediante los métodos descritos en esta memoria pueden cruzarse adicionalmente mediante técnicas de reproducción tradicionales con otras plantas para obtener plantas de progenie que comprenden la modificación fijada como objetivo.

[76] Las plantas y semillas de algodón descritas en esta memoria pueden tratarse adicionalmente con un compuesto químico, tal como un compuesto químico seleccionado de las siguientes listas:

- Herbicidas: Diuron, Fluometuron, MSMA, Oxifluorfen, Prometrin, Trifluralin, Carfentrazone, Cletodim, Fluazifop-butilo, Glifosato, Norflurazon, Pendimetalina, Piritiobac-sódico, Trifloxisulfuron, Tepraloxidim, Glufosinato, Flumioxazin, Tiazuron

- Insecticidas: acefato, aldicarb, clorpirifos, cipermetrina, deltametrina, abamectina, acetamiprid, emamectina benzoato, imidacloprid, indoxacarb, lambda-cihalotrina, spinosad, tiodicarb, gamma-cialotrina, espiromesifeno, piridilil, flonicamid, flubendiamida, triflumuron, Rynaxypyr, beta-ciflutrina, Espirotetramato, Clotianidina, Tiametoxam, Tiacloprid, Dinotofuran, Flubendiamida, Cyazypyr, Spinosad, Spintoram, gamma Cihalotrina, 4 -[[[(6-Clorpiridin-3-il)metil]-(2,2-difluoretil)amino]furan-2(5H)-ona, Tiodicarb, Avermectina, Flonicamid, Piridilil, Spiromesifen, Sulfoxaflor

- Fungicidas: Azoxistrobina, Bixafen, Boscalid, Carbendazim, Clorotalonil, Cobre, Ciproconazol, Difenconazol, Dimoxistrobina, Epoxiconazol, Fenamidona, Fluazinam, Fluopiram, Fluoxastrobina, Fluxaproxad, Iprodiona, Isopirazam, Isotianilo, Mancozeb, Maneb, Metominostrobin, Penthiopirad, Picoxistrobina, Propineb, Protiociconazol, piraclostrobina, quintozeno, tebuconazol, tetraconazol, tiofanato-metilo, trifloxistrobina

[77] Algodón, tal como se utiliza en esta memoria, se refiere a cualquier variedad de algodón existente. Por ejemplo, la célula de la planta de algodón puede ser de una variedad útil para el cultivo de algodón. Las variedades de algodón más utilizadas son *Gossypium barbadense*, *G. hirsutum*, *G. arboreum* y *G. herbaceum*. Otras variedades incluyen *G. africanum* y *G. raimondii*.

[78] Ejemplos de plantas de algodón descritas en esta memoria incluyen aquellas de las que se puede derivar callo embriogénico, tales como Coker 312, Coker310, Coker 5Acala SJ-5, GSC25110, FIBERMAX 819, Siokra 1-3, T25, GSA75, Acala SJ2, Acala SJ4, Acala SJ5, Acala SJ-C1, Acala B1644, Acala B1654-26, Acala B1654-43, Acala B3991, Acala GC356, Acala GC510, Acala GAM1, Acala C1, Acala Royale, Acala Maxxa, Acala Prema, Acala B638, Acala B1810, Acala B2724, Acala B4894, Acala B5002, no Acala "picker" Siokra, variedad "stripper" FC2017, Coker 315, STONEVILLE 506, STONEVILLE 825, DP50, DP61, DP90, DP77, DES119, McN235, HBX87, HBX191, HBX107, FC 3027, CHEMBRED A1, CHEMBRED A2, CHEMBRED A3, CHEMBRED A4, CHEMBRED B1, CHEMBRED B2, CHEMBRED B3, CHEMBRED C1, CHEMBRED C2, CHEMBRED C3, CHEMBRED C4, PAYMASTER 145, HS26, HS46, SICALA, PIMA S6 ORO BLANCO PIMA, FIBERMAX FM5013, FIBERMAX FM5015, FIBERMAX FM5017, FIBERMAX FM989, FIBERMAX FM832, FIBERMAX FM966, FIBERMAX FM958, FIBERMAX FM989, FIBERMAX FM958, FIBERMAX FM832, FIBERMAX FM991, FIBERMAX FM819, FIBERMAX FM800, FIBERMAX FM960, FIBERMAX FM966, FIBERMAX FM981, FIBERMAX FM5035, FIBERMAX FM5044, FIBERMAX FM5045, FIBERMAX FM5013, FIBERMAX FM5015, FIBERMAX FM5017 o FIBERMAX FM5024 y plantas con genotipos derivados de las mismas. Estos son adecuados para aplicar los métodos arriba descritos.

[79] Tal como se utiliza en esta memoria, "que comprende(n)" debe interpretarse como que especifica la presencia de las características, los números enteros, las etapas o los componentes mencionados, pero no excluye la presencia o adición de una o más características, números enteros, etapas o componentes o grupos de los mismos. Por lo tanto, p. ej., un ácido nucleico o una proteína que comprende una secuencia de nucleótidos o aminoácidos, puede comprender más nucleótidos o aminoácidos que los realmente citados, es decir, puede estar embebida en un ácido nucleico o una proteína más grande. Un gen quimérico que comprende una región de ADN que se define funcional o estructuralmente puede comprender regiones de ADN adicionales, etc.

[80] El término "planta" también incluye la progenie de plantas que conservan las características distintivas de los parentales tales como la semilla obtenida por autofecundación o cruzamiento, p. ej., semilla híbrida, plantas híbridas y partes de plantas derivadas de la misma.

[81] Tal como se utiliza en esta memoria, "parte de planta" incluye cualquier órgano de la planta o tejido de la planta, incluyendo, pero no limitado a frutos, semillas, embriones, fibras, regiones meristemáticas, tejido de callo, hojas, raíces, brotes, flores, gametofitos, esporofitos, polen y microesporas.

5 [82] Los términos "proteína" o "polipéptido", tal como se utilizan en esta memoria, describen un grupo de moléculas que consiste en más de 30 aminoácidos, mientras que el término "péptido" describe moléculas que consisten en hasta 30 aminoácidos. Las proteínas y los péptidos pueden formar, además, dímeros, trímeros y oligómeros superiores, es decir, consisten en más de una molécula de (poli)péptido. Las moléculas de proteínas o péptidos que forman tales dímeros, trímeros, etc. pueden ser idénticas o no idénticas. Las estructuras de orden superior correspondientes se denominan, por consiguiente, homodímeros o heterodímeros, homotrímeros o heterotrímeros, etc. Los términos "proteína" y "péptido" también se refieren a proteínas o péptidos modificados de forma natural, en los que la modificación se realiza, p. ej., mediante glicosilación, acetilación, fosforilación y similares. Modificaciones de este tipo son bien conocidas en la técnica.

15 [83] Para el propósito de esta invención, la "identidad de secuencia" de dos secuencias de nucleótidos o aminoácidos relacionadas, expresada como un porcentaje, se refiere al número de posiciones en las dos secuencias óptimamente alineadas que tienen residuos idénticos (x100), dividido por el número de posiciones comparadas. Un hueco, es decir, una posición en un alineamiento en donde un residuo está presente en una secuencia pero no en la otra, se considera como una posición con residuos no idénticos. El alineamiento de las dos secuencias se realiza mediante el algoritmo de Needleman y Wunsch (Needleman y Wunsch, 1970). El alineamiento de secuencia asistido por ordenador anterior, puede realizarse convenientemente utilizando un programa de software estándar tal como GAP que es parte del Paquete Wisconsin Versión 10.1 (Genetics Computer Group, Madison, Wisconsin, EE.UU), utilizando la matriz de puntuación por defecto con una penalidad de creación del hueco de 50 y una penalidad por extensión del hueco de 3.

25 [84] La lista de secuencias contenida en el archivo denominado "BCS11-2008_WO_ST25", que es de 59 kilobytes (tamaño medido en Microsoft Windows®), contiene 4 secuencias SEQ ID NO: 1 a SEQ ID NO: 4, se archiva mediante presentación electrónica y se incorpora en esta memoria como referencia.

[85] Los siguientes Ejemplos no limitantes describen métodos para modificar el genoma de una célula de planta de algodón utilizando una enzima DSBI y métodos de administración tanto mediada por *Agrobacterium* como administración directa en callo embriogénico.

30 [86] A menos que se indique lo contrario en los Ejemplos, todas las técnicas de ADN recombinante se llevan a cabo de acuerdo con protocolos estándares tal como se describe en Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor Laboratory Press, NY y en los volúmenes 1 y 2 de Ausubel *et al.* (1994) *Current Protocols in Molecular Biology*, Current Protocols, USA. Los materiales y métodos estándares para el trabajo molecular de plantas se describen en *Plant Molecular Biology Labfax* (1993) por R.D.D. Cray, publicado conjuntamente por BIOS Scientific Publications Ltd (Reino Unido) y Blackwell Scientific Publications, Reino Unido. Otras referencias para técnicas de biología molecular estándar incluyen Sambrook y Russell (2001) *Molecular Cloning: A Laboratory Manual*, tercera edición, Cold Spring Harbor Laboratory Press, NY, Volúmenes I y II de Brown (1998) *Molecular Biology LabFax*, segunda edición, Academic Press (Reino Unido). Los materiales y métodos estándares para las reacciones en cadena de la polimerasa se pueden encontrar en Dieffenbach y Dveksler (1995) *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, y en McPherson *et al.* (2000) *PCR – Basics: From Background to Bench*, primera edición, Springer Verlag, Alemania.

[87]

EJEMPLOS

Ejemplo 1: construcción del vector

45 [88] Utilizando técnicas de ADN recombinante estándares se construyeron los siguientes vectores de ADN que comprenden los siguientes elementos enlazados operativamente (representados esquemáticamente en las figuras 1 y 2):

[89] Vector de ADN diana pTCV117 (SEC ID NO 1)

- LB: límite izquierdo del ADN-T (nt 12540-12564)
- Sitio I-SceI (nt 31-14)
- 50 • 3'nos: secuencia que incluye la región 3' no traducida del gen de la nopalina sintasa del ADN-T de pTiT37 (Depicker *et al.*, 1982) (nt 32-220, complemento inverso)
- Barra: región codificadora del gen BAR de *Streptomyces hygrosopicus* (nt 240-791, complemento inverso)
- 5'cab22L: secuencia que incluye la secuencia conductora del gen de la proteína de unión a clorofila a/b de *Petunia hybrida* (Harpster *et al.*, 1988) (nt 801-857, complemento inverso)
- 55 • P35S2: promotor 35S (nt 858-1405, complemento inverso)

- 3'g7: terminación del extremo 3' y región de poliadenilación del gen 7 de *Agrobacterium tumefaciens* octopina tipo ADN-T (nt 1612-1409, complemento inverso)
- Sitio de reconocimiento de I-SceI (nt 1643 - 1617)
- 5 • intron 1 h3At: secuencia que incluye el primer intrón del gen II de la variante de histona H3.III de *Arabidopsis thaliana* (Chaubet *et al.*, 1992) (nt 1662-2127)
- TPotpC: secuencia codificadora del péptido de tránsito optimizado, que contiene la secuencia de los genes de la subunidad pequeña RuBisCO de *Zea mays* (maíz) y *Helianthus annuus* (girasol), tal como se describe por Lebrun *et al.* (1996), (nt 2145 - 2510)
- 2mepsps: región codificadora de EPSPS doble mutante del maíz (nt 2517-3852)
- 10 • 3'histonAt: secuencia que incluye la región 3' no traducida del gen de histona H4 de *Arabidopsis thaliana* (Chaboute *et al.*, 1987) (nt 3871-4537)
- RB: límite derecho de ADN-T (nt 4594-4618)
- [90] Vector de ADN reparador pTIB232 (SED ID NO 2):
- LB: límite izquierdo del ADN-T (nt 25-1)
- 15 • bar (del1-403): fragmento 5' del gen BAR de *Streptomyces hygrosopicus* (nt 46-448, complemento inverso)
- 5'cab22L: secuencia que incluye la secuencia conductora del gen de la proteína de unión a clorofila a/b de *Petunia hybrida* (Harpster *et al.*, 1988) (nt 458-514, complemento inverso)
- P35S2: promotor 35S (nt 515-1062, complemento inverso)
- 20 • 3'g7: Terminación del extremo 3' y región de poliadenilación del gen 7 de *Agrobacterium tumefaciens* octopina tipo ADN-T (nt1409-1612)
- Pcsvmv XYZ: fragmento del promotor CsVMV (1285-1724)
- 5'cvmv: conductor del promotor CsVMV (nt 1725-1797)
- hyg-1Pa: gen de resistencia a higromicina de *E. coli* (nt 1804-2829)
- 3'35S: terminación de la transcripción de 3' 35S y región de poliadenilación (nt 2841-3065)
- 25 • 2mepsps (5'del): fragmento 3' del gen EPSPS doble mutante del maíz (nt 3096-3501)
- 3'histonAt: región del extremo 3' del gen de la histona 4 de *Arabidopsis* (nt 3520-4186)
- RB: límite derecho de ADN-T (nt 4267-4243)
- [91] ADN reparador y vector de expresión de endonucleasa pTIB236 (SEC ID NO 3):
- LB: límite izquierdo del ADN-T (nt 1-25)
- 30 • 3'35S: terminación de la transcripción de 35S y región de poliadenilación (nt 104-328, complemento inverso)
- hyg-1 Pa: gen de resistencia a higromicina de *E. coli* (nt 340-1365, complemento inverso)
- 5'cvmv: conductor del promotor CsVMV (nt 1372-1444, complemento inverso)
- Pcsvmv XYZ: fragmento del promotor CsVMV (1445-1884, complemento inverso)
- 3'35S: región de poliadenilación de 35S (nt 1963-2097, complemento inverso)
- 35 • I-SceI: región codificadora de SceI de código universal (nt 2185-2919, complemento inverso)
- NLSsv40: señal de localización nuclear de SV40 (nt-2887-2910-, complemento inverso)
- 5'ats1 b: secuencia conductora del gen rbcS ATS1A de *Arabidopsis thaliana* (nt 2936-2976, complemento inverso)
- P35S2: promotor 35S (nt 2976-3496, complemento inverso)
- RB: límite derecho de ADN-T (nt 4592-5655)
- 40 [92] Vector de expresión de endonucleasa ptrr26 (SEQ ID NO 4) que comprende la región codificadora de I-SceI de código universal (documento WO 2006/074956):
- RB: límite derecho de ADN-T
- P35S2: promotor 35S
- 5'ats1b: secuencia conductora del gen rbcS ATS1A de *Arabidopsis thaliana*

- NLSsv40: señal de localización nuclear de SV40
- I-Scel: región codificadora de I-Scel de código universal
- 3³⁵S
- RB: límite izquierdo de ADN-T

5 Ejemplo 2: medios y tampones

[93] Medios y tampones utilizados durante la generación y transformación del callo embrionario tal como se describe a continuación en los ejemplos 3, 4 y 5:

Sustrato de co-cultivo: M100 con ½ concentración de sales MS pH = 5,2, + 100 µM AS + 100 mg/L de L-cisteína (L-cisteína siempre debe estar recién preparada y añadirse después de someterla a autoclave)

10 **Sustrato M100:** sales MS, vitaminas B5, MES 0.5 g/L, MgCl₂·6H₂O 0,94 g/L, gelrita 2g/L, glucosa 30 g/L, pH 5,8
Sustrato M104: = sustrato M100 + 1 g/L de KNO₃, pH 5,8

Sustrato M700: sales de Stewart + vitaminas, MgCl₂·6H₂O 0,47 g/L, gelrita 1g/L, agar de plantas 2,25 g/L, sacarosa 20 g/L, pH 6,8

15 **Sustrato M702:** sales de Stewart + vitaminas, MgCl₂·6H₂O 0,71 g/L, gelrite 1,5 g/L, agar de plantas 5 g/L, sacarosa 5 g/L, pH 6,8

AC: carbono activo 2 g/L

AS: acetosiringona 100 µM en DMSO

Ejemplo 3: Generación de callos embriogénicos quebradizos

20 [94] Semillas de algodón de Coker 312 se germinaron en medio de germinación sólido M100 sin hormonas durante 7-10 días en la oscuridad a 28°C. A continuación, la inducción del callo embriogénico se realizó incubando explantes de hipocotilo de las plántulas en medio M100 sólido (sin hormonas). Después de aproximadamente 2 meses cuando el callo de la herida en la superficie cortada del hipocotilo comienza a mostrar una proliferación rápida, el subcultivo adicional para el enriquecimiento y mantenimiento del callo embriogénico se realiza en medio M100 sólido con carbono activo (2 g/L). La inducción y el mantenimiento del callo embriogénico se produce en condiciones de luz tenue (intensidad: 1 a 7 µmol m⁻² s⁻¹, fotoperíodo: 16 H de luz/8 H de oscuridad) a 28°C.

25

Ejemplo 4: Transformación de callo embriogénico de algodón mediante bombardeo de partículas

[95] Se siguió el siguiente procedimiento para transformar callos embriogénicos de algodón utilizando bombardeo de partículas:

30 • Callo embriogénico de algodón (EC) quebradizo del Ejemplo 3 de una línea diana en la cual se ha de introducir una modificación fijada como objetivo inducida por DSB, se recoge 2 a 3 semanas después del subcultivo y se siembra en una capa delgada mediante un filtro Buchner encima de un papel de filtro en sustrato M100 con manitol 0,2 M y sorbitol 0,2 M durante ~ 2 a ~ 20 horas antes del bombardeo.

35 • Después de la preplasmolisis en sustrato M100 con manitol 0,2 M y sorbitol 0,2 M durante ~ 2 a ~ 20 horas, el EC se bombardea con el ADN de endonucleasa (~ 0,5 pmol), opcionalmente en presencia de un ADN reparador (~0,5 pmol)

• Condiciones de bombardeo:

o diámetro de las partículas de oro: 0.3-3 µm

o disco de rotura: 1100-1350 psi

o distancia al tejido diana: 9 cm

40 o vacío de cámara ~ 27 (en Hg)

o sistema de administración de partículas biolísticas BioRAD PPS_1000/He

• Después del bombardeo, los filtros se transfieren al sustrato M100 con manitol 0,2 M o al sustrato M100 sin agente selectivo.

45 • Después de 1 a 4 días en sustrato no selectivo bajo condiciones de luz tenue a 28°C, los filtros se transfieren al sustrato selectivo M100 con 50 mg/L de higromicina o glifosato 1 mM o 5 mg/L de PPT

• Después de aproximadamente 2 a 3 semanas, los callos proliferantes se selecciona de los filtros y se subcultivan adicionalmente como pequeñas pilas en el sustrato selectivo M100 con 50 mg/L de higromicina o glifosato 1 mM o 5 mg/L de PPT. Después de un período de subcultivo de ~ 6 semanas con ~ 3 intervalos de subcultivo semanales, en sustrato M100 selectivo bajo condiciones de luz tenue a 28°C, se pueden seleccionar embriones EC/somáticos transformados.

50

- Se realiza un rastreo molecular para la identificación de eventos de modificación fijados como objetivo al nivel de embriones EC/somáticos transformados.

5 • La regeneración de la planta se inicia a partir de los eventos de modificación específicos sembrando embriones EC/somáticos en M104 con carbono activo (AC) y el agente selectivo correspondiente en condiciones de luz (intensidad: 40 a 70 $\mu\text{mol m}^{-2} \text{s}^{-1}$; fotoperiodo: 16H de luz/8 H de oscuridad) a 28°C.

- Después de aproximadamente un mes, se transfieren embriones individuales de aproximadamente 0,5-1 cm sobre un papel de filtro en M104 con AC y el agente selectivo correspondiente.

- Los embriones que germinan bien adicionales se transfieren a un sustrato de germinación no selectivo M702 en condiciones de luz (intensidad: 40 a 70 $\mu\text{mol m}^{-2} \text{s}^{-1}$; fotoperiodo: 16H de luz/8H de oscuridad) a 28°C.

10 • Después de uno a dos meses, los embriones en desarrollo adicionales se transfieren al sustrato M700 en condiciones de luz (intensidad: 40 a 70 $\mu\text{mol m}^{-2} \text{s}^{-1}$; fotoperiodo: 16H de luz/8 H de oscuridad) a 28°C para desarrollar pequeñas plántulas.

Ejemplo 5: Transformación de callo embriogénico de algodón mediante administración de ADN mediada por Agrobacterium

15 [96] Se siguió el siguiente procedimiento para transformar callos embriogénicos de algodón utilizando Agrobacterium:

20 • Callo embriogénico de algodón (EC) quebradizo del Ejemplo 3 de una línea diana en la cual se ha de introducir una modificación fijada como objetivo inducida por DSB, se recoge 2 a 3 semanas después de subcultivo en sustrato 100 y se sumerge durante 20' en una suspensión de Agrobacterium de 5×10^8 células/ml en sustrato M100 pH 5,2, con 100 μM de acetosiringona (AS). La cepa de Agrobacterium porta un vector que contiene el ADN reparador y el gen que codifica la endonucleasa.

25 • Después de 3 días de co-cultivo en la oscuridad a 24°C en M100 con $\frac{1}{2}$ concentración de sales MS de pH 5,2, con 100 μM de AS y 100 mg/L de L-cisteína, el EC es sembrado por encima de un filtro papel por medio de un filtro Buchner o es transferido como pequeñas pilas sobre sustrato M100 pH 5,8, 250 mg/L de triacilina y un agente selectivo (higromicina 50 mg/L o PPT 5 mg/L o glifosato 1 a 1,5 mM) y se incuban en luz tenue (intensidad: 1 a 7 $\mu\text{mol m}^{-2} \text{s}^{-1}$; 16 H de luz/8 H de oscuridad) a 28°C.

30 • Después de 2 a 3 semanas, los callos se subcultivan adicionalmente como pequeñas pilas sobre sustrato M100 selectivo con 50 mg/L de higromicina o glifosato 1 mM o 5 mg/L de PPT. Después de un período de subcultivo de ~ 6 semanas con ~ 3 intervalos de subcultivo semanales, en sustrato M100 selectivo bajo condiciones de luz tenue a 28°C, se pueden seleccionar embriones EC/somáticos transformados.

- Se realiza un rastreo molecular para la identificación de eventos de modificación fijados como objetivo a nivel de embriones EC/somáticos transformados.

35 • La regeneración de plantas se inicia a partir de los eventos de modificación fijados como objetivo sembrando embriones CE/somáticos en M104 con carbono activo y el agente selectivo correspondiente bajo condiciones de luz (intensidad: 40 a 70 $\mu\text{mol m}^{-2} \text{s}^{-1}$; fotoperiodo 16 H de luz/8 H de oscuridad) a 28°C. A partir de este paso, los sustratos ya no contienen antibióticos.

- Después de aproximadamente un mes, se transfieren embriones individuales de aproximadamente 0,5-1 cm sobre un papel de filtro en M104 con carbono activo (AC) y el sustrato selectivo correspondiente.

40 • Los embriones que germinan adicionales se transfieren a un sustrato de germinación no selectivo M702 en condiciones de luz (intensidad: 40 a 70 $\mu\text{mol m}^{-2} \text{s}^{-1}$; fotoperiodo: (16H de luz/8H de oscuridad) a 28°C.

- Después de uno a dos meses, los embriones en desarrollo adicionales se transfieren al sustrato M700 en condiciones de luz (intensidad: 40 a 70 $\mu\text{mol m}^{-2} \text{s}^{-1}$; fotoperiodo: (16H de luz/8 H de oscuridad) a 28°C para desarrollar pequeñas plántulas.

Ejemplo 6: Generación de plantas diana para la evaluación de la modificación del genoma seleccionado

45 [97] Plantas de algodón transgénico que comprenden el vector de ADN diana pTCV117 se generaron mediante Agrobacterium tal como se describe en el ejemplo 5. El vector pTCV117 comprende un gen de barra funcional impulsado por 35S, ubicado entre dos sitios de reconocimiento de I-SceI y un gen EPSPS sin promotor (véase la Figura 1). Para evaluar la recombinación fijada como objetivo, originalmente se pretendía transformar estas plantas con un casete de expresión 35S-I-SceI y un ADN reparador con regiones de homología para el ADN diana para la restauración del gen epsps sin promotor mediante la inserción de un promotor de histona, resultando con ello la adquisición de la tolerancia al glifosato. Sin embargo, estas plantas pTCV117 parecían tener ya altos niveles de tolerancia al glifosato, probablemente debido a la actividad transcripcional bidireccional del promotor 35S, haciendo con ello inutilizable el ensayo.

50

Ejemplo 7: Inserción o reemplazo fijado como objetivo a través de recombinación homóloga

[98] Por lo tanto, se construyó un nuevo ADN reparador pTIB232 con regiones de homología con el ADN diana, que comprende un gen de resistencia a la higromicina impulsado por CSVMV flanqueado en un extremo por un fragmento 3' del gen epsps (permitiendo la recombinación homóloga con el gen EPSPS en el locus diana, reemplazando así la parte 5' del gen EPSPS por el gen higromicina) y en el otro extremo flanqueado por un promotor 35S enlazado a un fragmento génico de 5' bar), tal como se indica esquemáticamente en la Figura 1. Las plantas diana pTCV117 se transformaron con el vector pTIB232 y el vector ptrr26 que comprende una región codificadora de I-SceI de código universal operativamente enlazada a un promotor 35S utilizando bombardeo de partículas tal como se describe en el Ejemplo 4.

[99] En primer lugar, se rastrearon transformantes en cuanto a la tolerancia a la higromicina, ya que esto indica la inserción del ADN reparador de pTIB232. Los eventos HygR se evaluaron posteriormente para la pérdida de resistencia a glifosato, que es indicativa de recombinación en el locus diana. Los 36 recombinantes hygR y GlyS así obtenidos se caracterizaron adicionalmente por análisis de PCR utilizando el par de cebadores P3 xP4 que reconoce la región genómica aguas arriba del gen EPSPS y el gen de higromicina (Figura 1). Esto dio como resultado la identificación de 8 eventos potenciales de selección de genes correctos (reemplazo de la parte 5' del gen epsps por el gen de higromicina por recombinación homóloga al menos unilateral: configuración I en la Figura 1), que posteriormente se confirmó que era de hecho corregir los eventos de fijación como objetivo génico por análisis de la secuencia del producto de PCR obtenido por el conjunto de cebadores P3, P4).

Ejemplo 8: Inserción fijado como objetivo a través de unión en los extremos no homóloga

[100] Plantas pTCV117 se transformaron con ADN reparador y vector de expresión de endonucleasa pTIB236, que comprende un gen higR funcional y un gen codificador de I-SceI de código universal funcional, sin homología con el sitio diana (véase la Figura 2), utilizando transferencia de ADN mediada por Agrobacterium tal como se describe en el Ejemplo 5.

[101] Los 200 eventos hygR que se obtuvieron de este modo se analizaron mediante PCR utilizando 2 pares de cebadores P1xP2 y P3xP4; el par de cebadores P1xP2 que reconoce el gen hygR y la región genómica aguas abajo del gen bar y el otro par que reconoce la región genómica aguas arriba del gen EPSPS y el gen de higromicina (Figura 2). Esto resultó en la identificación de 17 eventos potenciales de inserción/reemplazo fijados como objetivo. El análisis de secuencia realizado en 4 de estos eventos demostró que en realidad eran eventos de inserción/reemplazo fijados como objetivo (configuración I, II en la Figura 2).

Ejemplo 9: Inserción o reemplazo fijado como objetivo a través de unión no homóloga de los extremos utilizando transformación mediada por Agrobacterium

[102] La línea diana G4GH198-01901 se transformó tal como se describe en el Ejemplo 5 con la cepa de Agrobacterium A5280=EHA101 (pTIB236) que comprende el ADN reparador y el vector de expresión de endonucleasa pTIB236.

[103] Se seleccionaron eventos resistentes a la higromicina (HygR), de los cuales 200 se analizaron por PCR para la inserción del gen hyg en el o los sitios de I-SceI diana utilizando los pares de cebadores P1xP2 y P3xP4. De estos 200 eventos, se encontró que aproximadamente el 10% eran eventos específicos. El análisis de la secuencia de algunos de estos eventos específicos reveló que el gen bar había sido reemplazado por el gen hyg o que el gen hyg había sido insertado junto al gen bar (evento apilado).

Ejemplo 10: Inducción de la rotura de la doble cadena fijada como objetivo utilizando una meganucleasa diseñada a medida

[104] Callos embriogénicos de plantas de algodón resistentes a PPT que contienen un gen quimérico que comprende el gen bar bajo el control del promotor CSVMV se transformaron mediante bombardeo de partículas con un vector que comprende un gen quimérico que codifica una meganucleasa modificada a medida que reconoce una secuencia diana en el gen bar, como una cadena sencilla (pCV170: SEQ ID NO 5) o como un heterodímero (pCV177: SEQ ID NO 6). Se realizó una co-administración del vector de meganucleasa con un vector que contiene el gen 2 mEPSPS bajo el control de un promotor expresable en plantas que confiere tolerancia a glifosato como un gen marcador seleccionable. Se obtuvieron aproximadamente 3000 callos resistentes al glifosato, de los cuales 85 eventos parecían sensibles a PPT, lo cual indica una alteración del gen bar. De estos, se caracterizaron 79 eventos para el genotipo mediante PCR utilizando cebadores que flanquean el sitio diana y la secuenciación posterior del producto de PCR (tabla 1).

Tabla 1: Caracterización de eventos de transformación resistentes a glifosato sensibles a PPT. La ausencia de un producto de PCR es indicativa de una deleción grande alrededor del sitio diana.

pCV170 (sc) Producto PCR obtenible		nº eventos	cambio en sitio diana	nº eventos
no		11	deleción grande	11
sí		8	sin mutación	6
			reemplazo/inserción	1
			deleción	1
pCV177 (hd) Producto PCR obtenible		nº eventos	cambio en sitio diana	nº eventos
no		33	deleción grande	33
sí		27	sin mutación	19
			Inserción	4
			deleción	4

5 [105] Por lo tanto, estos resultados demuestran que es posible inducir una rotura de ADN de doble cadena fijada como objetivo en la posición deseada con una meganucleasa de cadena sencilla así como una meganucleasa heterodímica diseñada específicamente y que pueden obtenerse deleciones fijadas como objetivo, reemplazos e inserciones. usando estas meganucleasas en algodón.

LISTADO DE SECUENCIAS

- <110> Bayer Cropscience N.V.
D'Halluin, Katelijjn
- 5 <120> MÉTODOS Y MEDIOS PARA MODIFICAR EL GENOMA DE UNA PLANTA EN UN SITIO
PRESELECCIONADO
- <130> BCS-11-2008
- <160> 4
- <170> PatenIn versión 3.3
- 10 <210> 1
- <211> 12641
- <212> ADN
- <213> Artificial
- <220>
- 15 <223> vector

```

<400> 1
agcttgccgat cgctagggat aacagggtaa tcgcgtatta aatgtataat tgcgggactc      60
taatcataaa aacccatctc ataaataacg tcatgcatta catgttaatt attacatgct      120
taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca ggattcaatc      180
ttaagaaact ttattgcca atgtttgaac gatctgcttc ggatcctaga cgcgtagat      240
cagatctcgg tgacgggcag gaccggacgg ggcggtagcg gcaggctgaa gtccagctgc      300
cagaaaccca cgatcatgcca gttcccgctc ttgaagccgg ccgcccgcag catgccgcgg      360
ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag cccgatgaca      420
gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtaggt gtagagcgtg      480
gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc ggccgtccag      540
tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac ctccgctcc      600
acctcggcga cgagccaggg atagcgcctc cgcagacgga cgaggtcgct cgtccactcc      660
tgcggttctt cgcgctcggg acggaagttg accggtgctt tctcgatgta gtggttgacg      720
atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc cgggcgtcgt      780
tctgggtcca tggttttggt ttaataagaa gagaaaagag ttcttttgtt atggctgaag      840
taatagagaa atgagctcga gtcctctcca aatgaaatga acttccttat atagaggaag      900
ggtcttgoga aggatagtgg gattgtgcgt catcccttac gtcagtggag atatcacatc      960
aatccacttg ctttgaagac gtggttgaa cgtcttcttt ttccacgatg ctccctcgtg      1020
gtgggggtcc atctttggga cactgtcgg cagaggcatc ttgaacgata gcctttcctt      1080
tatcgcaatg atggcatttg taggtgccac ctcccttttc tactgtcctt ttgatgaagt      1140
gacagatagc tgggcaatgg aatccgagga ggtttccga tattaccctt tgttgaaaag      1200
tctcaatagc cctttggtct tctgagactg tatctttgat attcttgag tagacgagag      1260
    
```

ES 2 657 825 T3

tgtcgtgctc caccatgttg acgaagattt tcttcttgctc attgagtcgt aaaagactct 1320
 gtatgaactg ttcgccagtc ttcacggcga gttctgtag atcctcgatc tgaatttttg 1380
 actccatgta tgggtgcatat ggcgcgccta agctagctag atcatcaatt tatgtattac 1440
 acataatata gcactcagtc tttcatctac ggcaatgtac cagctgatat aatcagttat 1500
 tgaaatattt ctgaatttaa acttgcatca ataaatttat gtttttgctt ggactataat 1560
 acctgacttg ttattttatc aataaatatt taaactatat ttctttcaag atacgtatta 1620
 ccctgttatc cctaaagctt atcgatttgc aaccctcag gcgaagaaca ggtatgattt 1680
 gtttgtaatt agatcagggg tttaggtctt tccattactt tttaatgttt tttctgttac 1740
 tgtctccgcg atctgatttt acgacaatag agtttcgggt tttgtccat tccagtttga 1800
 aaataaaggt ccgctctttta agtttgctgg atcgataaac ctgtgaagat tgagtctagt 1860
 cgatttattg gatgatccat tcttcatcgt tttttcttg cttcgaagtt ctgtataacc 1920
 agatttgtct gtgtgcgatt gtcattacct agccgtgat cgagaactag ggttttcgag 1980
 tcaattttgc cccttttggg tataatctggg tcgataacga ttcattctgga ttagggtttt 2040
 aagtgggac gtttagtatt ccaatttctt caaaatttag ttatggataa tgaaaatccc 2100
 caattgactg ttcaatttct tgttaaatgc gcagatccc atggcttoga tctcctcctc 2160
 agtcgcgacc gtttagccga ccgccctgc tcaggccaac atgggtggctc cgttcaccgg 2220
 ccttaagtcc aacgccgcct tccccaccac caagaaggct aacgacttct ccacccttcc 2280
 cagcaacggg ggaagagttc aatgatgca ggtgtggccg gcctacggca acaagaagtt 2340
 cgagacgctg tcgtacctgc cgccgctgc tatggcggcc accgtgatga tggcctcgtc 2400
 ggccaccgcc gtcgctccgt tccaggggct caagtccacc gccagcctcc ccgtcgcccg 2460
 ccgctcctcc agaagcctcg gcaacgtcag caacggcga aggatccggt gcatggccgg 2520
 cgccgaggag atcgtgctgc agcccatcaa ggagatctcc ggcaaccgtca agctgccggg 2580
 gtccaagtgc ctttccaacc ggatcctcct actcgccgcc ctgtccgagg ggacaacagt 2640
 ggttgataac ctgctgaaca gtgaggatgt cactacatg ctcggggctc tgaggactct 2700
 tggctctctc gtcgaagcgg acaaagctgc caaaagagct gtagttgttg gctgtggtgg 2760
 aaagttccca gttgaggatg ctaaagagga agtgcagctc ttcttgggga atgctggaat 2820
 cgcaatgcgg tccttgacag cagctgttac tgctgctggg ggaaatgcaa cttacgtgct 2880
 tgatggagta ccaagaatga gggagagacc cattggcgac ttggttgcg gattgaagca 2940
 gcttggtgca gatgttgatt gtttccttgg cactgactgc ccacctgttc gtgtcaatgg 3000
 aatcggaggg ctacctggtg gcaaggtcaa gctgtctggc tccatcagca gtcagtactt 3060
 gagtgccttg ctgatggctg ctcctttggc tcttggggat gtggagattg aaatcattga 3120

ES 2 657 825 T3

taaattaatc tccattccgt acgtcgaaat gacattgaga ttgatggagc gttttggtgt 3180
 gaaagcagag cattctgata gctgggacag attctacatt aagggaggtc aaaaatacaa 3240
 gtcccctaaa aatgcctatg ttgaaggtga tgcctcaagc gcaagctatt tcttggtctgg 3300
 tgctgcaatt actggagggg ctgtgactgt ggaaggttgt ggcaccacca gtttgacagg 3360
 tgatgtgaag tttgctgagg tactggagat gatgggagcg aaggttacat ggaccgagac 3420
 tagcgtaact gttactggcc caccgcggga gccatttggg aggaaacacc tcaaggcgat 3480
 tgatgtcaac atgaacaaga tgctgatgt cgccatgact cttgctgtgg ttgccctctt 3540
 tgccgatggc ccgacagcca tcagagacgt ggcttcctgg agagtaaag agaccgagag 3600
 gatggttcg atccggacgg agctaaccaa gctgggagca tctgttgagg aagggccgga 3660
 ctactgcatc atcacccgc cggagaagct gaacgtgacg gcgatcgaca cgtacgacga 3720
 ccacaggatg gcgatggctt tctoccttgc cgcctgtgcc gaggtccccg tcaccatccg 3780
 ggaccctggg tgcaccggga agacctccc cgactacttc gatgtgctga gcactttcgt 3840
 caagaattaa gctctagaac tagtggatcc cccgatccgc gtttgtgtt tctgggtttc 3900
 tcacttaagc gtctgcgtt tacttttcta ttgggtttgg cgtttagtag tttgcgtag 3960
 cgttcttgtt atgtgtaatt acgcttttcc ttcttgcttc agcagtttcg gttgaaatat 4020
 aaatcgaatc aagtttcaact ttatcagcgt tgttttaaat tttggcatta aattggtgaa 4080
 aattgcttca attttgcctc taaatagaag agacaacatg aaattcgact tttgacctca 4140
 aatcttcgaa catttatttc ctgatttcac gatggatgag gataacgaaa gggcggttcc 4200
 tatgtccggg aaagtcccc tagaagacaa tgagcaaagc tactgaaacg cggacacgac 4260
 gtgcatttg tacggatatg agttaaaccg actcaattcc tttattaaga cataaacgga 4320
 ttttggttaa agtgtaacag tgagctgata taaaaccgaa acaaaccggt acaagtttga 4380
 ttgagcaact tgatgacaaa cttcagaatt ttggttattg aatgaaaatc atagtctaat 4440
 cgtaaaaaat gtacagaaga aaagctagag cagaacaaag attctatatt ctggttccaa 4500
 tttatcatcg ctttaacgtc cctcagattt gatcgggctg caggaattaa acgcccgggc 4560
 acgtgggatc ctctagagtc gacggccgag tactggcagg atatataccg ttgtaatttg 4620
 tcgctgtgga ataagtcgct gtgtatgttt gtttgattgt ttctgttga gtgcagccca 4680
 tttcaccgga caagtcggct agattgattt agccctgatg aactgccgag gggaagccat 4740
 cttgagcgcg gaatgggaat ggatttcgtt gtacaacgag acgacagAAC acccagggga 4800
 ccgagcttcg atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac 4860
 catatTTTTG aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata 4920
 ggatggcaag atcctggtat cggctgcga ttccgactcg tccaacatca atacaaccta 4980
 ttaatttccc ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg 5040

ES 2 657 825 T3

aatccggtga gaatggcaaa agtttatgca tttctttcca gacttgttca acaggccagc 5100
cattacgctc gtcatcaaaa tcaactcgcac caaccaaacc gttattcatt cgtgattgcg 5160
cctgagcgag acgaaatacg ccgctgttaa aaggacaatt acaaacagga atcgaatgca 5220
accggcgcag gaacactgcc agcgcaccaa caatattttc acctgaatca ggatattctt 5280
ctaatacctg gaatgctggt tttccgggga tcgcagtggg gagtaaccat gcatcatcag 5340
gagtacggat aaaatgcttg atggtcggaa gaggcataaa ttccgtcagc cagtttagtc 5400
tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc agaaacaact 5460
ctggcgcacg gggcttccca tacaatcgat agattgtcgc acctgattgc ccgacattat 5520
ccgaatctgg caattccggt tcgcttgctg tccataaaac cgcccagtct agctatcgcc 5580
atgtaagccc actgcaagct acctgctttc tctttgcgct tgcgttttcc ggatcttctt 5640
gagatccttt tttctgcgc gtaatctgct gcttgcaaac aaaaaacca ccgctaccag 5700
cggtggtttg tttgccggat caagagctac caactctttt tccgaagta actggcttca 5760
gcagagcgca gataccaaat actgtccttc tagttagacc gtagttaggc caccacttca 5820
agaactctgt agcaccgcct acatacctcg ctctgctaact cctgttacca gtggctgctg 5880
ccagtggcga taagtctgtt cttaccgggt tggactcaag acgatagtta ccggataagg 5940
cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttgagc cgaacgacct 6000
acaccgaact gagataccta cagcgtgagc tatgagaaaag cgccacgctt cccgaagga 6060
gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc 6120
ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac ctctgacttg 6180
agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg 6240
cggccttttt acggttctcg gccttttgct ggccttttgc tcacatgttc tttcctgcgt 6300
tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat accgctcgcc 6360
gcagccgaac gaccgagcgc agcagtcagc tgagcgagga agcgggaagag cgcctgatgc 6420
ggtattttct ccttacgcat ctgtgcggta tttcacaccg catatggtgc actctcagta 6480
caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatcgc tacgtgactg 6540
ggtcatggct gcgccccgac acccgccaac acccgctgac gcgcccctgac gggcttgtct 6600
gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag 6660
gttttcaccg tcatcaccga aacgcgcgag gcaggggtgcc ttgatgtggg ccgccggcgt 6720
cgagtggcga cggcgcggct tgtccgcgcc ctggtagatt gcctggccgt aggccagcca 6780
tttttgagcg gccagcggcc gcgataggcc gacgcgaagc ggcggggcgt agggagcgca 6840
gcgaccgaag ggtaggcgct ttttgcagct ctccggctgt gcgctggcca gacagttatg 6900

ES 2 657 825 T3

cacagccag gcgggtttta agagtttta taagtttta agagttttag gcggaaaaat 6960
cgcctttttt ctcttttata tcagtcactt acatgtgtga ccggttccca atgtacggct 7020
ttgggttccc aatgtacggg ttccggttcc caatgtacgg ctttgggttc ccaatgtacg 7080
tgctatccac aggaaagaga ccttttcgac ctttttcccc tgctagggca atttgcccta 7140
gcatctgctc cgtacattag gaaccggcgg atgcttcgcc ctcgatcagg ttgcggtagc 7200
gcatgactag gatcgggcca gcctgccccg cctcctcctt caaatcgtac tccggcaggt 7260
catttgacc cgtacagcttg cgcacgggtga aacagaactt cttgaactct ccggcgctgc 7320
cactgcgttc gtagatcgtc ttgaacaacc atctggcttc tgccttgctt gcggcgcggc 7380
gtgccaggcg gtagagaaaa cggccgatgc cgggatcgat caaaaagtaa tccgggtgaa 7440
ccgtcagcac gtccgggttc ttgccttctg tgatctcggg gtacatccaa tcagctagct 7500
cgatctcgat gtactccggc cgcgccggtt cgctctttac gatctttag cggtaatca 7560
aggcttcacc ctggatacc gtcaccaggc ggccgttctt ggccttcttc gtacgctgca 7620
tggcaacgtg cgtggtggtt aaccgaatgc aggtttctac caggtcgtct ttctgctttc 7680
cgccatcggc tcgccggcag aacttgagta cgtccgcaac gtgtggacgg aacacgcggc 7740
cgggcttgtc tcccttcctt tcccggatc ggttcatgga ttcggttaga tgggaaaccg 7800
ccatcagtac caggtcgtaa tcccacacac tggccatgcc ggccggccct gcggaaacct 7860
ctacgtgccc gtctggaagc tcgtagcggg tcacctcgcc agctcgtcgg tcacgcttcg 7920
acagacggaa aacggccacg tccatgatgc tgcgactatc gcgggtgccc acgtcataga 7980
gcatcggaa gaaaaaatct ggttgctcgt cgccttggg cggcttcta atcgacggcg 8040
caccggctgc cggcgggttc cgggattctt tgcggattcg atcagcggcc gcttgccacg 8100
attcaccggg gcgtgcttct gcctcgatgc gttgccgctg ggccgctgc gcggccttca 8160
acttctccac caggtcatca cccagcggc cgcgatttg taccgggccc gatggtttg 8220
gaccgtcacg ccgattcctc gggcttgggg gttccagtgc cattgcaggg ccggcagaca 8280
accagccgc ttacgcctgg ccaaccgccc gttcctccac acatggggca ttccacggcg 8340
tcgggtgctg gttgttctt attttccatg ccgcctcctt tagccgctaa aattcatcta 8400
ctcatttatt catttgctca ttactctgg tagctcggg atgtattcag atagcagctc 8460
ggtaatggtc ttgccttggc gtaccgctga catcttcagc ttggtgtgat cctccgcccg 8520
caactgaaag ttgaccgct tcatggctgg cgtgtctgcc aggtggcca acggtgcagc 8580
cttgctgctg cgtgcgctcg gacggccggc acttagcgtg tttgtgcttt tgctcatttt 8640
ctctttacct cattaactca aatgagttt gatttaattt cagcggccag cgctggacc 8700
tcgcccggc cgtgcgccctc gggttctgat tcaagaacgg ttgtccggc ggcggcagtg 8760
cctgggtagc tcacgcgctg cgtgatacgg gactcaagaa tgggcagctc gtaccggcc 8820

ES 2 657 825 T3

agcgcctcgg caacctcacc gccgatgcgc gtgcctttga tcgcccgcga caccgacaaag 8880
 gccgcttgta gccttccatc cgtgacctca atgcgctgct taaccagctc caccaggctcg 8940
 gcggtggccc atatgtcgtc agggcttggc tgcaccggaa tcagcacgaa gtcggctgcc 9000
 ttgatcgcgg acacagccaa gtccgccgcc tggggcgctc cgtcgatcac tacgaagtcg 9060
 cgccggccga tggccttcac gtcgcggtca atcgtcgggc ggtcgatgcc gacaacggtt 9120
 agcggttgat cttcccgcac ggccgcccaa tcgcgggcac tgccctgggg atcggaatcg 9180
 actaacagaa catcggcccc ggcgagttgc agggcgggg ctagatgggt tgcgatggtc 9240
 gtcttgccctg acccgccttt ctggttaagt acagcgataa ccttcacgctg tccccttgc 9300
 gtatttgttt atttactcat cgcacatcat acgcagcgac cgcacgacgc aagctgtttt 9360
 actcaaatac acatcacctt tttagacggc ggcgctcggg ttcttcagcg gccaaagctgg 9420
 ccggccaggc cgccagcttg gcatcagaca aaccggccag gatttcacgc agccgcacgg 9480
 ttgagacgtg cggggcgggc tcgaacacgt acccggccgc gatcatctcc gcctcgatct 9540
 cttcggtaat gaaaaacggt tcgtcctggc cgtcctgggt cggtttcatg cttgttcctc 9600
 ttggcgttca ttctcggcgg ccgccagggc gtcggcctcg gtcaatgcgt cctcacggaa 9660
 ggcaccgcgc cgcctggcct cgggtggcgt cacttcctcg ctgcgctcaa gtgcgcggta 9720
 cagggtcgag cgatgcacgc caagcagtc agccgcctct ttcaaggctg gcccttcctg 9780
 gtcgatcagc tcgcgggcgt ggcgatctg tgcgggggtg agggtagggc gggggccaaa 9840
 cttcacgcct cgggccttgg cggcctcgcg cccgctcgg gtgcggctca tgattagga 9900
 acgctcgaac toggcaatgc cggcgaacac ggtcaacacc atgcggccgg ccggcgtggt 9960
 ggtgtcggcc cacggctctg ccaggctacg caggcccgcg ccggcctcct ggatgcgctc 10020
 ggcaatgtcc agtaggtcgc ggggtcctgc gccagggcg tctagcctgg tcaactgtcac 10080
 aacgtcgcga gggcgtaggt ggtcaagcat cctggccagc tccggggcgt cgcgcctggt 10140
 gccggtgatc ttctcggaaa acagcttggg gcagccggcc gcgtgcagtt cggcccgttg 10200
 gttggtcaag tcctggctgt cgggtcctgac gcgggcatag cccagcaggc cagcggcggc 10260
 gctcttgctt atggcgtaat gtctccggtt ctagtgcgaa gtattctact ttatgcgact 10320
 aaaacacgcg acaagaaaac gccaggaaa gggcagggcg gcagcctgtc gcgtaactta 10380
 ggacttgctg gacatgtcgt tttcagaaga cggctgcact gaacgtcaga agccgactgc 10440
 actatagcag cggaggggtt ggatcgatcc ctgctcgcgc aggctgggtg ccaagctctc 10500
 ggtaacatc aaggccgat ccttgagacc cttgccctcc cgcacgatga tcgtgccgtg 10560
 atcgaaatcc agatccttga cccgcagttg caaacctca ctgatccgca tgcccgttcc 10620
 atacagaagc tgggcgaaca aacgatgctc gccttcaga aaaccgagga tgccaaccac 10680

ES 2 657 825 T3

ttcatccggg gtcagcacca cgggcaagcg cccggacggc cgaggtcttc cgatctcctg 10740
 aagccagggc agatccgtgc acagcacttg ccgtagaaga acagcaaggc cgccaatgcc 10800
 tgacgatgcg tggagaccga aaccttgccg tcgttcgcca gccaggacag aaatgcctcg 10860
 acttcgctgc tgcccaaggt tgccgggtga cgcacaccgt ggaaacggat gaaggcacga 10920
 acccagtgga cataagcctg ttcggttcgt aagctgtaat gcaagtagcg tatgcgctca 10980
 cgcaactggt ccagaacctt gaccgaacgc agcggtggtg acggcgcagt ggcggttttc 11040
 atggcttght atgactgttt ttttggggta cagtctatgc ctccggcatc caagcagcaa 11100
 gcgctgtacg ccgtgggtcg atgtttgatg ttatggagca gcaacgatgt tacgcagcag 11160
 ggcagtcgcc ctaaaacaaa gttaaacatc atgagggaaag cggtgatcgc cgaagtatcg 11220
 actcaactat cagaggtagt tggcgtcatc gagcgccatc tcgaaccgac gttgctggcc 11280
 gtacatttgt acggctccgc agtggatggc ggcctgaagc cacacagtga tattgatttg 11340
 ctggttacgg tgaccgtaag gcttgatgaa acaacgcggc gagctttgat caacgacctt 11400
 ttggaaactt cggcttcccc tggagagagc gagattctcc gcgctgtaga agtcaccatt 11460
 gttgtgcacg acgacatcat tccgtggcgt tatccagcta agcgcgaact gcaatttggg 11520
 gaatggcagc gcaatgacat tcttgacaggt atcttcgagc cagccacgat cgacattgat 11580
 ctggctatct tgctgacaaa agcaagagaa catagcgttg ccttggtagg tccagcggcg 11640
 gaggaactct ttgatccggt tcctgaacag gatctatttg aggcgctaaa tgaaacctta 11700
 acgctatgga actcgccgcc cgactgggct ggcgatgagc gaaatgtagt gcttacgttg 11760
 tcccgcatth ggtacagcgc agtaaccggc aaaatcgcgc cgaaggatgt cgctgccgac 11820
 tgggcaatgg agcgcctgcc ggcccagtat cagcccgtca tacttgaagc tagacaggct 11880
 tatcttgac aagaagaaga tcgcttgccc tcgcgcgagc atcagttgga agaatttgtc 11940
 cactacgtga aaggcgagat caccaagta gtcggcaaat aatgtctaac aattcgttca 12000
 agccgacgcc gcttcgcggc gcggcttaac tcaagcgtta gatgcactaa gcacataatt 12060
 gctcacagcc aaactatcag gtcaagtctg cttttattat ttttaagcgt gcataataag 12120
 ccctacacaa attgggagat atatcatgaa aggctggcct tttcttgta tcgcaatagt 12180
 tggcgaagta atcgcaacat ccgcattaaa atctagcgag ggctttacta agctagcttg 12240
 cttggtcgtt ccggtaccgt gaacgtcggc tcgattgtac ctgcttcaa atactttgcg 12300
 atcgtgttgc gcgcctgccc ggtgcgtcgg ctgatctcac ggatcgactg cttctctcgc 12360
 aacgccatcc gacggatgat gtttaaaagt cccatgtgga tcaactccgtt gccccgtcgc 12420
 tcaccgtgtt ggggggaag tgcacatggc tcagttctca atggaaatta tctgcctaac 12480
 cggctcagtt ctgcgtagaa accaacatgc aagctccacc gggtgcaaag cggcagcggc 12540
 ggcaggatat attcaattgt aaatggctcc atggcgatcg ctacgtatct agaattcctg 12600

ES 2 657 825 T3

caggctcgagt cgcgacgtac gttcgaacaa ttggttttaa a 12641

<210> 2
 <211> 12188
 <212> ADN
 <213> Artificial

<220>
 <223> vector

<400> 2
 cggcaggata tattcaattg taaatggctc catggcgatc gctacgcacg ctcgggctcgt 60
 tgggcagccc gatgacagcg accacgctct tgaagccctg tgcctccagg gacttcagca 120
 ggtgggtgta gagcgtggag cccagtcctc tccgctgggtg gcggggggag acgtacacgg 180
 tcgactcggc cgtccagctg taggcgttgc gtgccttcca ggggcccgcg taggcgatgc 240
 cggcgacctc gccgtccacc tcggcgacga gccagggata gcgctccgc agacggacga 300
 ggtcgtccgt ccaactcctgc ggttcctgcg gctcggtagc gaagttgacc gtgcttctct 360
 cgatgtagtg gttgacgatg gtgcagaccg ccggcatgtc cgcctcggtg gcacggcgga 420
 tgtcggccgg gcgctcgtct ggtccatgg ttttggtta ataagaagag aaaagagttc 480
 ttttgttatg gctgaagtaa tagagaaatg agctcagatc ctctccaaat gaaatgaact 540
 tccttatata gaggaagggc cttgcgaagg atagtgggat tgtgcgtcat cccttacgctc 600
 agtggagata tcacatcaat ccaacttgctt tgaagacgtg gttggaacgt cttcttttctc 660
 cacgatgctc ctcgtgggtg ggggtccatc tttgggacca ctgtcggcag aggcactctg 720
 aacgatagcc tttcctttat cgcaatgatg gcatttgtag gtgccacctt ccttttctac 780
 tgtccttttg atgaagtgac agatagctgg gcaatggaat ccgaggaggt ttcccgatat 840
 taccctttgt tgaaaagtct caatagccct ttggtcttct gagactgtat ctttgatatt 900
 cttggagtag acgagagtgt cgtgctccac catggtgacg aagatthtct tcttgtcatt 960
 gagtcgtaaa agactctgta tgaactgttc gccagtctc acggcgagtt ctgttagatc 1020
 ctcgatctga atttttgact ccatgtatgg tgcatatggc gcgcctaagc tagctagatc 1080
 atcaatttat gtattacaca taatatcgca ctcagtctt catctacggc aatgtaccag 1140
 ctgatataat cagttattga aatattctg aatttaaact tgcacataa aatttatggt 1200
 tttgcttggc ctataatacc tgacttgta ttttatcaat aaatattta actatatttc 1260
 tttcaagata cgtatctaga attcgaaggt aattatccaa gatgtagcat caagaatcca 1320
 atgtttacgg gaaaaactat ggaagtatta tgtgagctca gcaagaagca gatcaatatg 1380
 cggcacatat gcaacctatg ttcaaaaatg aagaatgtac agatacaaga tcctatactg 1440
 ccagaatagc aagaagaata cgtagaatg gaaaaagaag aaccaggcga agaaaagaat 1500

ES 2 657 825 T3

cttgaagacg taagcactga cgacaacaat gaaaagaaga agataaggtc ggtgattgtg 1560
 aaagagacat agaggacaca tgtaagggtg aaaatgtaag ggcggaagt aaccttatca 1620
 caaaggaatc ttatcccca ctacttatcc ttttatattt ttccgtgtca tttttgccct 1680
 tgagttttcc tatataagga accaagttcg gcatttgtga aaacaagaaa aaatttgggtg 1740
 taagctatth tctttgaagt actgaggata caacttcaga gaaatttcta agtttctctc 1800
 gagatgaaaa agcctgaact caccgcgacg tctgtcgaga agtttctgat cgaaaagttc 1860
 gacagcgtct ccgacctgat gcagctctcg gagggcgaag aatctcgtgc tttcagcttc 1920
 gatgtaggag ggcgtggata tgcctgcgg gtaaatagct gcgccgatgg tttctacaaa 1980
 gatcgttatg tttatcggca ctttgcacg gccgcgctcc cgattccgga agtgcttgac 2040
 attggggagt tcagcgagag cctgacctat tgcactccc gccgtgcaca ggggtgcacg 2100
 ttgcaagacc tgccgaaac cgaactgcc gctgttctgc agccggtcgc ggaggccatg 2160
 gatgctatcg ctgcccga tcttagccag acgagcgggt tcggcccatt cggaccgcaa 2220
 ggaatcggtc aatacactac atggcgtgat tcatatgcg cgattgctga tccccatgtg 2280
 tatcactggc aaactgtgat ggacgacacc gtcagtgcgt ccgtcgcgca ggctctcgat 2340
 gagctgatgc tttggcccga ggactgcccc gaagtccggc acctcgtgca cgcggatttc 2400
 ggctccaaca atgtcctgac ggacaatggc cgcataacag cggtcattga ctggagcgag 2460
 gcgatgttcg gggattccca atacgaggtc gccaacatct tcttctggag gccgtggtt 2520
 gcttgatgg agcagcagac gcgctacttc gagcggaggc atccggagct tgcaggatcg 2580
 ccgcgcctcc gggcgtatat gctccgcat ggtcttgacc aactctatca gagcttggt 2640
 gacggcaatt tcgatgatgc agcttggcgc cagggtcgat gcgacgcaat cgtccgatcc 2700
 ggagccggga ctgtcggcgc tacacaaatc gcccgagaa gcgcggccgt ctggaccgat 2760
 ggctgtgtag aagtactcgc cgatagtga aaccgacgcc ccagcactcg tccgagggca 2820
 aaggaatagg atatcaagct tggacacgct gaaatcacca gtctctctct acaaatctat 2880
 ctctctctat tttctcata ataatgtgtg agtagttccc agataaggga attagggttc 2940
 ctatagggtt tcgctcatgt gttgagcata taagaaacc ttagtatgta tttgtatttg 3000
 taaaatactt ctatcaataa aatttctaat tcctaaaacc aaaatccagt actaaaatcc 3060
 agatctaact ataacggtcc taaggtagcg accgcgggag ccatttggga ggaaacacct 3120
 caaggcgatt gatgtcaaca tgaacaagat gcctgatgtc gccatgactc ttgctgtggt 3180
 tgcctcttt gccgatggc cgacagccat cagagacgtg gcttcctgga gagtaaagga 3240
 gaccgagagg atggttgca tccggacgga gtaaccaag ctgggagcat ctggtgagga 3300
 agggccggac tactgcatca tcacgccgcc ggagaagctg aacgtgacgg cgatcgacac 3360
 gtacgacgac cacaggatgg cgatggcttt ctcccttgcc gcctgtgcc aggtccccgt 3420

ES 2 657 825 T3

caccatccgg gaccctgggt gcacccggaa gaccttcccc gactacttcg atgtgctgag 3480
 cactttcgtc aagaattaag ctctagaact agtggatccc cggatccgcg tttgtgtttt 3540
 ctgggtttct cacttaagcg tctgcgtttt acttttgtat tgggtttggc gtttagtagt 3600
 ttgcggtagc gttcttgta tgtgtaatta cgctttttct tcttgcttca gcagtttcgg 3660
 ttgaaatata aatcgaatca agtttcaact tatcagcgtt gttttaaatt ttggcattaa 3720
 attggtgaaa attgcttcaa ttttgtatct aaatagaaga gacaacatga aattcgactt 3780
 ttgacctcaa atcttcgaac atttatttcc tgatttcacg atggatgagg ataacgaaag 3840
 ggcggttcct atgtccggga aagttcccgt agaagacaat gagcaaagct actgaaacgc 3900
 ggacacgacg tcgcattggt acggatatga gttaaaccga ctcaattcct ttattaagac 3960
 ataaaccgat tttggttaaa gtgtaacagt gagctgatat aaaaccgaaa caaaccgta 4020
 caagtttgat tgagcaactt gatgacaaac ttcagaattt tggttattga atgaaaatca 4080
 tagtctaadc gtaaaaaatg tacagaagaa aagctagagc agaacaaaga ttctatattc 4140
 tggttccaat ttatcatcgc tttaacgtcc ctgagatttg atcgggctgc aggaattaaa 4200
 cgcccgggca cgtgggatcc tctagagtcg acggccgagt actggcagga tatataccgt 4260
 tgaatttgt cgcgtgtgaa taagtcgctg tgtatgtttg tttgattggt tctggttgag 4320
 tgcagcccat ttcaccggac aagtcggcta gattgattta gccctgatga actgccgagg 4380
 ggaagccatc ttgagcgcgg aatgggaatg gatttcgttg tacaacgaga cgacagaaca 4440
 cccacgggac cgagcttcga tcgagcatca aatgaaactg caatttattc atatcaggat 4500
 tatcaatacc atatttttga aaaagccgtt tctgtaatga aggagaaaac tcaccgaggc 4560
 agttccatag gatggcaaga tcctggtatc ggtctgcgat tccgactcgt ccaacatcaa 4620
 tacaacctat taatttcccc tcgtcaaaaa taaggttatc aagtgagaaa tcaccatgag 4680
 tgacgactga atccggtgag aatggcaaaa gtttatgcat ttctttccag acttgttcaa 4740
 caggccagcc attacgctcg tcatcaaaat cactcgcatac aaccaaacg ttattcattc 4800
 gtgattgctc ctgagcgaga cgaaatacgc cgctgttaaa aggacaatta caaacaggaa 4860
 tcgaatgcaa cggcgcgagg aacactgcc aacactgcc aacactgcc aatattttca cctgaatcag 4920
 gatattcttc taatacctgg aatgctgttt ttccggggat cgcagtggtg agtaaccatg 4980
 catcatcagg agtacggata aaatgcttga tggctggaag aggcataaat tccgtcagcc 5040
 agtttagtct gaccatctca tctgtaacat cattggcaac gctacctttg ccatgtttca 5100
 gaaacaactc tggcgcacgc ggcttcccat acaatcgata gattgtcgca cctgattgcc 5160
 cgacattatc cgaatctggc aattccgggt cgcttgctgt ccataaaacc gccagtcta 5220
 gctatcgcca tgtaagccca ctgcaagcta cctgctttct ctttgcgctt gcgttttccg 5280

ES 2 657 825 T3

gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 5340
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 5400
ctggcttcag cagagcgagc ataccaaata ctgtccttct agtgtagccg tagttaggcc 5460
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaata ctgttaccag 5520
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 5580
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttgagc 5640
gaacgacctc caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 5700
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaca ggagagcgca 5760
cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc 5820
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 5880
ccagcaacgc ggccttttta cggctcctgg ccttttgctg gccttttgct cacatgttct 5940
ttcctgcggt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata 6000
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 6060
gcctgatgcg gtattttctc cttacgcacg tgtgcggtat ttcacacgcg atatggtgca 6120
ctctcagtac aatctgctct gatgccgat agttaagcca gtatacactc cgctatcgct 6180
acgtgactgg gtcattggtg cgcctcgaca cccgccaca cccgctgacg cgccctgacg 6240
ggcttgtctg ctcccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat 6300
gtgtcagagg ttttcaccgt catcaccgaa acgcgcgagg cagggtgcct tgatgtgggc 6360
gccgcggtc gagtggcgac ggcgcggctt gtccgcgccc tggtagattg cctggccgta 6420
ggccagccat ttttgagcgg ccagcggccg cgataggccg acgcgaagcg gcggggcgta 6480
gggagcgag cgaccgaagg gtaggcgctt tttgcagctc ttcggctgtg cgctggccag 6540
acagttatgc acaggccag cgggttttaa gagttttaat aagttttaa gagttttag 6600
cggaaaaatc gccttttttc tcttttatat cagtcactta catgtgtgac cggttcccaa 6660
tgtacggctt tgggttccca atgtacgggt tccggttccc aatgtacggc tttgggttcc 6720
caatgtacgt gctatccaca ggaaagagac cttttcgacc tttttcccct gctagggcaa 6780
tttgccctag catctgctcc gtacattagg aaccggcgga tgcttcgccc tcgatcaggt 6840
tgcggtagcg catgactagg atcgggccag cctgccccgc ctcctccttc aaatcgact 6900
ccgcaggtc atttgaccg atcagcttgc gcacggtgaa acagaacttc ttgaacttc 6960
cggcgtgcc actgcgttcg tagatcgtct tgaacaacca tctggcttct gccttgctg 7020
cggcgggcg tgccaggcg tagagaaac gcccgatgcc gggatcgatc aaaaagtaat 7080
cggggtgaac cgtcagcacg tccgggttct tgcttctgt gatctcgcg tacatccaat 7140
cagctagctc gatctcgatg tactccggcc gcccggttc gctctttacg atctttagc 7200

ES 2 657 825 T3

ggctaatacaa ggcttcaccc tcggataaccg tcaccaggcg gccgttcttg gccttcttcg 7260
 tacgctgcat ggcaacgtgc gtggtgttta accgaatgca ggtttctacc aggtcgtctt 7320
 tctgctttcc gccatcggct cgccggcaga acttgagtac gtccgcaacg tgtggacgga 7380
 acacgcggcc gggcttgtct cccttccctt cccggtatcg gttcatggat tcggttagat 7440
 gggaaaccgc catcagtacc aggtcgtaat cccacacact ggccatgccg gccggccctg 7500
 cggaaacctc tacgtgcccg tctggaagct cgtagcggat cacctcgcca gctcgtcggc 7560
 cacgcttcga cagacgaaa acggccacgt ccatgatgct gcgactatcg cgggtgccca 7620
 cgtcatagag catcggaacg aaaaaatctg gttgctcgtc gcccttgggc ggcttcctaa 7680
 tcgacggcgc accggtgcc ggcggttgc gggattcttt gcggattcga tcagcggccg 7740
 cttgccacga ttcaccgggg cgtgcttctg cctcagatgcg ttgccgctgg gcggcctgcg 7800
 cggccttcaa cttctccacc aggtcatcac ccagcggcg gccgatttgt accgggcccg 7860
 atggtttgcg accgtcacgc cgattcctcg ggcttggggg ttccagtgcc attgcagggc 7920
 cggcagacaa cccagccgct tacgcctggc caaccgcccg ttccctccaca catggggcat 7980
 tccacggcgt cggcgcctgg ttgttcttga tttccatgc cgcctccttt agccgctaaa 8040
 atcctctac tcatttattc atttgctcat ttactctggt agctgcgcga tgtattcaga 8100
 tagcagctcg gtaatggtct tgccttggcg taccgcgtac atcttcagct tgggtgatc 8160
 ctccgccggc aactgaaagt tgaccgcctt catggctggc gtgtctgcca ggctggccaa 8220
 cgttcagacc ttgctgctgc gtgcgctcgg acggccggca cttagcgtgt ttgtgctttt 8280
 gctcattttc tctttacctc attaaactcaa atgagttttg atttaatttc agcggccagc 8340
 gcctggacct cggcggcagc gtcgccctcg ggttctgatt caagaacggt tgtgccggcg 8400
 gcggcagtc ctgggtagct cacgcgctgc gtgatacggg actcaagaat gggcagctcg 8460
 taccgggcca gcgcctcggc aacctcaccg ccgatgcgcg tgcctttgat cggccgcgac 8520
 acgacaaagg ccgctttag ccttccatcc gtgacctcaa tgcgctgctt aaccagctcc 8580
 accaggtcgg cggcggccca tatgtcgtaa gggcttggct gcaccggaat cagcacgaag 8640
 tcggtgcct tgatcgcgga cacagccaag tccgcgcct ggggcgctcc gtcgatcact 8700
 acgaagtcgc gccggccgat ggcttcacg tcgcggtcaa tcgtcgggcg gtcgatgccg 8760
 acaacggtta gcggttgatc ttcccgcacg gccgcccaat cgcgggact gccctgggga 8820
 tcggaatcga ctaacagaac atcggccccg gcgagttgca gggcgcgggc tagatgggtt 8880
 gcgatggtcg tcttgctga cccgcctttc tggttaagta cagcgataac cttcatgctg 8940
 tccccttgcg tatttgttta ttactcatc gcatcatata cgcagcgacc gcatgacgca 9000
 agctgtttta ctcaaataca catcaccttt ttagacggcg gcgctcgggt tcttcagcgg 9060

ES 2 657 825 T3

ccaagctggc cggccaggcc gccagcttgg catcagacaa accggccagg atttcatgca 9120
 gccgcacggc tgagacgtgc gggggcggct cgaacacgta cccggccgcg atcatctccg 9180
 cctcgatctc ttcggtaatg aaaaacggtt cgtcctggcc gtctgtgtgc ggtttcatgc 9240
 ttgttcctct tggcgttcat tctcggcggc cggcaggcg tcggcctcgg tcaatcgctc 9300
 ctcacggaag gcaccgcgcc gcctggcctc ggtggcgctc acttcctcgc tgcgctcaag 9360
 tgcgcggtac agggtcgagc gatgcacgcc aagcagtgca gccgcctctt tcacggtgcg 9420
 gccttcctgg tcgatcagct cgcggcgctg cgcgatctgt gccggggtga gggtagggcg 9480
 ggggcaaac ttcacgcctc gggccttggc ggctcgcgc ccgctccggg tgcggtcgat 9540
 gattagggaa cgctcgaact cggcaatgcc ggcaaacacg gtcaacacca tgcggccggc 9600
 cggcgtggtg gtgtcggccc acggctctgc caggctacgc aggcccgcg cggcctcctg 9660
 gatgcgctcg gcaatgtcca gtaggtcggc ggtgctgagg gccaggcggc cttagcctggt 9720
 cactgtcaca acgtcgccag ggcgtagggtg gtcaagcatc ctggccagct ccgggcggtc 9780
 gcgcctggtg ccggtgatct tctcgaaaa cagcttgggtg cagccggccg cgtgcagttc 9840
 ggcccgttgg ttggtcaagt cctggtcgtc ggtgctgacg cgggcatagc ccagcaggcc 9900
 agcggcggcg ctcttgttca tggcgtaatg tctccggttc tagtcgcaag tattctactt 9960
 tatgcgacta aaacacgcga caagaaaacg ccagaaaag ggcaggcggc cagcctgtcg 10020
 cgtaacttag gacttgtgcg acatgtcgtt ttcagaagac ggctgactg aacgtcagaa 10080
 gccgactgca ctatagcagc ggagggggtg gatcgatccc tgctcgcgca ggctgggtgc 10140
 caagctctcg ggtaacatca aggccgatc cttggagccc ttgccctccc gcacgatgat 10200
 cgtgccgtga tcgaaatcca gatccttgac ccgcagttgc aaaccctcac tgatccgcat 10260
 gcccgctcca tacagaagct gggcgaacaa acgatgctcg ccttcagaa aaccgaggat 10320
 gcgaaccact tcatccgggg tcagcaccac cggcaagcgc ccggacggcc gaggtcttcc 10380
 gatctcctga agccagggca gatccgtgca cagcacttgc cgtagaagaa cagcaaggcc 10440
 gccaatgcct gacgatgctg ggagaccgaa accttgcgct cgttcgccag ccaggacaga 10500
 aatgcctcga ctctcgtgct gcccaagggt gccgggtgac gcacaccgtg gaaacggatg 10560
 aaggcacgaa cccagtggac ataagcctgt tcggttcgta agctgtaatg caagtagcgt 10620
 atgcgctcac gcaactggtc cagaacctg accgaacgca gcggtggtaa cggcgcagtg 10680
 gcggttttca tggcttggtt tgactgtttt tttgggttac agtctatgcc tcgggcatcc 10740
 aagcagcaag cgcgttacgc cgtgggtcga tgtttgatgt tatggagcag caacgatggt 10800
 acgcagcagg gcagtcgccc taaaacaaag ttaaacatca tgaggggaagc ggtgatcgcc 10860
 gaagtatcga ctcaactatc agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg 10920
 ttgctggccg tacatttgta cggctccgca gtggatggcg gcctgaagcc acacagtgat 10980

ES 2 657 825 T3

attgatttgc tggttacggt gaccgtaagg cttgatgaaa caacgcggcg agctttgatc 11040
aacgaccttt tggaaacttc ggcttcccct ggagagagcg agattctccg cgctgtagaa 11100
gtcaccattg ttgtgcacga cgacatcatt ccgtggcggt atccagctaa gcgcgaactg 11160
caatttggag aatggcagcg caatgacatt cttgcaggta tcttcgagcc agccacgatc 11220
gacattgatc tggctatctt gctgacaaaa gcaagagaac atagcgttgc cttggtaggt 11280
ccagcggcgg aggaactctt tgatccgggt cctgaacagg atctatttga ggcgctaaat 11340
gaaaccttaa cgctatggaa ctgcgcgcc gactgggctg gcgatgagcg aaatgtagtg 11400
cttacgttgt cccgcatttg gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc 11460
gctgcgcgact gggcaatgga gcgcctgccc gcccagatc agcccgtcat acttgaagct 11520
agacaggctt atcttggaca agaagaagat cgcttggcct cgcgcgcaga tcagttggaa 11580
gaatttgtcc actacgtgaa aggcgagatc accaaggtag tcggcaaata atgtctaaca 11640
attcgttcaa gcgcagccg cttcgcggcg cggcttaact caagcgttag atgcactaag 11700
cacataattg ctcacagcca aactatcagg tcaagtctgc ttttattatt ttttaagcgtg 11760
cataataagc cctacacaaa ttgggagata tatcatgaaa ggctggcttt ttcttgttat 11820
cgcaatagtt ggcgaagtaa tcgcaacatc cgcattaaaa tctagcgagg gctttactaa 11880
gctagcttgc ttggtcgttc cggtagcgtg aacgtcggct cgattgtacc tgcgttcaaa 11940
tactttgcga tcgtgttgcg cgcctgcccg gtgcgtcggc tgatctcacg gatcgactgc 12000
ttctctcgca acgccatccg acggatgatg tttaaaagtc ccatgtggat cactccggtg 12060
ccccgtcgtc caccgtgttg gggggaaggc gcacatggct cagttctcaa tggaaattat 12120
ctgcctaacc ggctcagttc tgcgtagaaa ccaacatgca agctccaccg ggtgcaaagc 12180
ggcagcgg 12188

<210> 3
<211> 11500
<212> ADN
<213> Artificial

<220>
<223> vector

<400> 3
cggcaggata tattcaattg taaatggctc catggcgatc gctaccctct gcagtcgacg 60
ggcccgttgt ccgcgggtcg ctaccttagg accggtatag ttagatctgg attttagtac 120
tggattttgg ttttaggaat tagaaatttt attgatagaa gtattttaca aatacaata 180
catactaagg gtttcttata tgctcaacac atgagcgaaa ccctatagga accctaattc 240
ccttatctgg gaactactca cacattatta tggagaaaat agagagagat agattttag 300

ES 2 657 825 T3

agagagactg gtgatttcag cgtgtccaag cttgatatcc tattucltttg ccctcggacg 360
 agtgctgggg cgtcggtttc cactatcggc gactacttct acacagccat cggccagac 420
 ggccgcgctt ctgcgggcga tttgtgtacg cccgacagtc ccggctccgg atcggacgat 480
 tgcgtcgcac cgaccctcgg cccaagctgc atcatogaaa ttgcogtcaa ccaagctctg 540
 atagagttgg tcaagaccaa tgcggagcat atacgcccg aggcgcggcg atcctgcaag 600
 ctccggatgc ctccgctcga agtagcgcgt ctgctgctcc atacaagcca accacggcct 660
 ccagaagaag atgttggcga cctcgtattg ggaatccccg aacatcgcct cgctccagtc 720
 aatgaccgct gttatgcggc cattgtcogt caggacattg ttggagccga aatccgcgtg 780
 cacgaggtgc cggacttcgg ggcagtcctc ggcccaaagc atcagctcat cgagagcctg 840
 cgcgacggac gactgacgg tgcctccat cacagtttgc cagtgataca catggggatc 900
 agcaatcgcg catatgaaat cacgccatgt agtgattga ccgattcctt gcggtccgaa 960
 tgggccgaac ccgctcgtct ggctaagatc ggccgcagcg atagcatcca tggcctccgc 1020
 gaccggctgc agaacagcgg gcagttcggg ttcaggcagg tcttgcaacg tgacaccctg 1080
 tgcacggcgg gagatgcaat aggtcaggct ctgcctgaac tccccaatgt caagcacttc 1140
 cggaatcggg agcgcggccg atgcaaagtg ccgataaaca taacgatctt tgtagaacc 1200
 atcggcgcag ctatttacc gcaggacata tccacgccct cctacatcga agctgaaagc 1260
 acgagattct tcgccctccg agagctgcat caggctggag acgctgtcga acttttcgat 1320
 cagaaacttc tcgacagacg tcgcggtgag ttcaggcttt tcatctcga gacaaactta 1380
 caaatttctc tgaagtgtga tcctcagtac ttcaaagaaa atagcttaca ccaaattttt 1440
 tcttgttttc acaaatgcgg aacttggttc cttatatagg aaaactcaag ggcaaaaatg 1500
 acacggaaaa atataaaagg ataagtagtg ggggataaga ttcctttgtg ataaggttac 1560
 tttccgccct tacattttcc accttacctg tgcctctat gtctctttca caatcaccga 1620
 ccttatcttc ttcttttcat tgttgcgtc agtgcttacg tcttcaagat tcttttcttc 1680
 gcctggttct tctttttcaa tttctacgta tttctcttcg tattctggca gtataggatc 1740
 ttgtatctgt acattcttca tttttgaaca taggttgcac atgtgccgca tattgatctg 1800
 cttcttgctg agctcacata atacttccat agtttttccc gtaaacattg gattcttgat 1860
 gctacatctt ggataattac cttcgaattc gtcttgggga tccgatatct agagtatacg 1920
 cgttaacgcg gccgctgtac catgcatgat ctggatttta gtactggatt ttggttttag 1980
 gaattagaaa ttttattgat agaagtattt tacaataca aatacactact aagggtttct 2040
 tatatgctca acacatgagc gaaaccctat aggaacccta attcccttat ctgggaacta 2100
 ctcacacatt attatggaga aaatagagag agatagattt gtagagagag actggtgatt 2160
 tcagcgtgtc caagcttgc agccctattt caggaaagtt tcggaggaga tagtgttcgg 2220

ES 2 657 825 T3

cagtttgtag atcatctgcg ggatcaggta cggtttgatc aggttgtaga agatcaggta 2280
 agacatagaa tcgatgtaga tgatcggttt gtttttggtg atttttacgt aacagttcag 2340
 ttggaatttg ttacgcagac ccttaaccag gtattctact tcttcgaaag tgaaagactg 2400
 ggtgttcagt acgatcgatt tgttggtaga gtttttggtg taatcccatt taccaccatc 2460
 atccatgaac cagtatgcca gagacatcgg ggtcaggtag ttttcaacca ggtgttcgg 2520
 gatggttttt ttgttgtaa cgatgaacag gttagccagt ttggtgaaag cttggtgttt 2580
 gaaagtctgg gcgccccagg tgattaccag gttaccaggg tggttaaacac gttctttttt 2640
 gtgcggcggg gacagtacc actgatcgtc cagcagacat acgtggtcca tgtatgcttt 2700
 gtttttccac tcgaactgca tacagtaggt tttacctca tcacgagaac ggatgtaagc 2760
 atcaccagg atcagaccga tacctgcttc gaactgttcg atgttcagtt cgatcagctg 2820
 ggatttgat tctttcagca gtttagagtt cggaccagg ttcattacct ggtttttttt 2880
 gatgtaacc ttgcgcttct tcttggggg tttagccatg gttttggact tcttcttctt 2940
 cttcttttgc ttaattctcg agtctctcc aatgaaatg aacttcctta tatagaggaa 3000
 gggcttgcg aaggatagtg ggattgtgc tcatccctta cgtcagtga gatcacat 3060
 caatccactt gctttgaaga cgtggttga acgtcttctt tttccacgat gctcctogtg 3120
 ggtgggggtc catctttggg accactgtcg gcagaggcat cttgaacgat agcctttcct 3180
 ttatcgcaat gatggcattt gtaggtgcca ccttctttt ctactgtcct tttgatgaag 3240
 tgacagatag ctgggcaatg gaatccgagg aggtttcccg atattacct ttgttgaaaa 3300
 gtctcaatag ccctttggtc ttctgagact gtatcttga tattcttga gtagacgaga 3360
 gtgtcgtgct ccaccatggt gacgaagatt ttcttcttgt cattgagtcg taaaagactc 3420
 tgtatgaact gttcgccagt cttcacggcg agttctgta gatcctcgat ctgaattttt 3480
 gactccatgt atggtgcata tggcgcgcca tgcatacgt ggtaccaatt gcctgcagg 3540
 cgacggccga gtactggcag gatataacc gttgtaattt gtcgctgtg aataagtcgc 3600
 tgtgtatggt tgtttgattg tttctgttgg agtgcagccc atttcaccgg acaagtggc 3660
 tagattgatt tagccctgat gaactgccga ggggaagcca tcttgagcgc ggaatgggaa 3720
 tggatttctg tgtacaacga gacgacagaa caccacggg accgagcttc gatcgagcat 3780
 caaatgaaac tgcaattht tcatatcag attatcaata ccatattht gaaaaagccg 3840
 tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa gatcctggta 3900
 tcggtctgcg attccgactc gtccaacatc aatacaacct attaatthc cctctgcaaa 3960
 aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg agaatggcaa 4020
 aagtttatgc atttctttcc agacttgctc aacaggccag ccattacgct cgtcatcaaa 4080

ES 2 657 825 T3

atcactcgc tcaaccaaac cgttattcat tcgtgattgc gcctgagcga gacgaaatac 4140
 gccgctgtta aaaggacaat tacaacacag aatcgaatgc aaccggcgca ggaacactgc 4200
 cagcgcacatca acaatatttt cacctgaatc aggatattct tctaatacct ggaatgctgt 4260
 ttttccgggg atcgcagtg tgagtaacca tgcacatca ggagtacgga taaaatgctt 4320
 gatggtcggga agaggcataa attccgtcag ccagtttagt ctgaccatct catctgtaac 4380
 atcattggca acgctacctt tgccatgttt cagaaacaac tctggcgcgcat cgggcttccc 4440
 atacaatcga tagattgtcg cacctgattg cccgacatta tccgaatctg gcaattccgg 4500
 ttcgcttgct gtccataaaa ccgcccagtc tagctatcgc catgtaagcc cactgcaagc 4560
 tacctgcttt ctctttgctg ttgctgtttc cggatcttct tgagatcctt tttttctgctg 4620
 cgtaatctgc tgcttgcaaa caaaaaaac accgctacca gcggtgggtt gtttgccgga 4680
 tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agatacacia 4740
 tactgtcctt ctagtgtagc cgtagtttag ccaccacttc aagaactctg tagcaccgcc 4800
 tacatacctc gctctgctaa tccctgttacc agtggctgct gccagtggcg ataagtctgtg 4860
 tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac 4920
 ggggggttgc tgcacacagc ccagcttggga gcgaacgacc tacaccgaac tgagatacct 4980
 acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc 5040
 ggtaagcggc agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg 5100
 gtatctttat agtctctgct ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg 5160
 ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct 5220
 ggccttttgc tggccttttg ctccatgtt ctttctgctg ttatcccctg attctgtgga 5280
 taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg 5340
 cagcagatca gtgagcagag aagcgggaaga gcgcctgatg cggatatttc tccttacgca 5400
 tctgtcgggt atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc 5460
 atagttaagc cagtatacac tccgctatcg ctacgtgact gggcatggc tgcgccccga 5520
 caccgcgcaa caccgcgctga cgcgcctga cgggcttgtc tgctcccggc atccgcttac 5580
 agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcacaccg 5640
 aaacgcgcga ggcaggggtgc cttgatgtgg gcgccggcgg tcgagtgggc acggcgcggc 5700
 ttgtccgcgc cctggtagat tgccctggcc taggccagcc atttttgagc ggccagcggc 5760
 cgcgataggc cgacgcgaag cggcggggcg tagggagcgc agcgaccgaa gggtaggcgc 5820
 tttttgcagc tcttcggctg tgcgctggcc agacagttat gcacaggcca ggccgggtttt 5880
 aagagtttta ataagtttta aagagtttta ggcgaaaaa tcgccttttt tctcttttat 5940
 atcagtcact tacatgtgtg accggttccc aatgtacggc tttgggttcc caatgtacgg 6000

ES 2 657 825 T3

gttccggttc ccaatgtacg gctttggggtt cccaatgtac gtgctatcca caggaaagag 6060
 accttttcga cctttttccc ctgctagggc aatttgccct agcatctgct ccgtacatta 6120
 ggaaccggcg gatgcttcgc cctcgatcag gttgcggtag cgcgatgacta ggatcgggcc 6180
 agcctgcccc gcctcctcct tcaaactcgtc ctccggcagg tcatttgacc cgatcagctt 6240
 gcgcacggtg aaacagaact tcttgaactc tccggcgctg ccaactgcgtt cgtagatcgt 6300
 cttgaacaac catctggctt ctgccttgcc tgcggcgcggtg cgtgccaggc ggtagagaaa 6360
 acggccgatg ccgggatcga tcaaaaagta atcgggggtga accgtcagca cgtccggggtt 6420
 cttgccttct gtgatctcgc ggtacatcca atcagctagc tcgatctcga tgtactccgg 6480
 ccgcccgggtt tcgctcttta cgatcttgta gcggctaatac aaggcttcac cctcggatac 6540
 cgtcaccagg cggccgttct tggccttctt cgtacgctgc atggcaacgt gcgtggtggtt 6600
 taaccgaatg caggtttcta ccaggtcgtc tttctgcttt ccgccatcgg ctccgccggca 6660
 gaacttgagt acgtccgcaa cgtgtggacg gaacacgcgg ccgggcttgt ctcccttccc 6720
 ttcccggtat cggttcatgg attcgggttag atgggaaacc gccatcagta ccaggtcgta 6780
 atcccacaca ctggccatgc cggccggccc tgcggaaacc tctacgtgcc cgtctggaag 6840
 ctctagcgg atcacctcgc cagctcgtcg gtcacgcttc gacagacgga aaacggccac 6900
 gtccatgatg ctgcgactat cgcgggtgcc cacgtcatag agcatcggaa cgaaaaaatc 6960
 tggttgctcg tcgcccttgg gcggcttctt aatcgacggc gcaccggctg ccggcgggtt 7020
 ccgggattct ttgcggattc gatcagcggc cgcttgccac gattcaccgg ggcgtgcttc 7080
 tgcctcgatg cgttgccgct gggcggcctg cgcggccttc aacttctcca ccaggtcac 7140
 acccagcgc gcgccgattt gtaccgggcc ggatggtttg cgaccgtcac gccgattcct 7200
 cgggcttggg ggttccagtg ccattgcagg gccggcagac aaccagccg cttacgcctg 7260
 gccaacgcc cgttctcca cacatggggc attccacggc gtccggtgcct ggttgttctt 7320
 gattttccat gccgcctcct ttagccgcta aaattcatct actcatttat tcatttgctc 7380
 atttactctg gtagctcgc gatgtattca gatagcagct cggtaatggt cttgccttgg 7440
 cgtaccgctg acatcttcag cttggtgtga tcctccgccg gcaactgaaa gttgaccgcg 7500
 ttcatggctg gcgtgtctgc caggctggcc aacgttgcag ccttgctgct gcgtgcgctc 7560
 ggacggccgg cacttagcgt gtttgtgctt ttgctcattt tctctttacc tcattaactc 7620
 aatgagttt tgatttaatt tcagcggcca gcgcctggac ctccgggca gcgtcgcctt 7680
 cgggttctga ttcaagaacg gttgtgccgg cggcggcagc gcctgggtag ctcacgcgct 7740
 gcgtgatacg ggactcaaga atgggcagct cgtaccggc cagcgcctcg gcaacctcac 7800
 cgccgatgcg cgtgcctttg atcggccgcg acacgacaaa ggccgcttgt agccttccat 7860

ES 2 657 825 T3

ccgtgacctc aatgcgctgc ttaaccagct ccaccaggtc ggcggtggcc catatgtcgt 7920
 aagggcttgg ctgcaccgga atcagcacga agtcggctgc cttgatcgcg gacacagcca 7980
 agtccgccgc ctggggcgct ccgtcgatca ctacgaagtc gcgccggccg atggccttca 8040
 cgtcgcggtc aatcgtcggg cggtcgatgc cgacaacggt tagcggttga tcttcccgca 8100
 cggccgccca atcgcgggca ctgccctggg gatcgggaatc gactaacaga acatcggccc 8160
 cggcgagtgt cagggcgcgg gctagatggg ttgcatgggt cgtcttgccct gaccgcctt 8220
 tctggttaag tacagcgata accttcatgc gttccccttg cgtatttgtt tatttactca 8280
 tcgcatcata tacgcagcga ccgcatgacg caagctgttt tactcaaata cacatcacct 8340
 ttttagacgg cggcgctcgg tttcttcagc ggccaagctg gccggccagg ccgccagctt 8400
 ggcacagac aaaccggcca ggatttcatg cagccgcacg gttgagacgt gcgcgggcg 8460
 ctogaacacg taccgggccg cgatcatctc gcctcgcac tcttcggtaa tgaaaaacgg 8520
 ttcgtcctgg ccgtcctggt gcggtttcat gcttgctcct cttggcgctt attctcggcg 8580
 gccgccaggc cgtcggcctc ggtcaatgcg tcctcacgga aggcaccgcg ccgcctggcc 8640
 tcggtgggcg tcacttcctc gctgcgctca agtgcgcggt acagggtcga gcgatgcacg 8700
 ccaagcagtg cagccgcctc tttcacggtg cggccttctt ggtcgatcag ctccggggcg 8760
 tgcgcgatct gtgcgggggt gaggttaggg cgggggcca aacttcacgcc tcgggccttg 8820
 gcggcctcgc gcccgctccg ggtgcggtcg atgattaggg aacgctcga ctcggcaatg 8880
 ccggcgaaca cggtaaacac catgcggccg gccggcgctg tgggtgcggc ccacggctct 8940
 gccaggctac gcaggcccg cccggcctcc tggatgcgct cggcaatgct cagtaggtcg 9000
 cgggtgctgc gggccaggcg gtctagcctg gtcactgtca caacgctgcc agggcgtagg 9060
 tggtaagca tcctggccag ctccgggcg tgcgcctgg tgcgggtgat cttctcggaa 9120
 aacagcttgg tgcagccggc cgcgtgcagt tcggcccgtt ggttgggtcaa gtcctggtcg 9180
 tcggtgctga cgcgggcata gccagcagg ccagcggcg cgctcttgtt catggcgtaa 9240
 tgtctccggt tctagtcgca agtattctac tttatgcgac taaaacacgc gacaagaaaa 9300
 cgcagggaaa agggcagggc ggcagcctgt cgcgtaactt aggacttgtg cgacatgtcg 9360
 ttttcagaag acggctgcac tgaacgtcag aagccgactg cactatagca gcggaggggt 9420
 tggatcgatc cctgctcgcg caggctgggt gccaaactct cgggtaacat caaggcccga 9480
 tccttgagc ccttgccctc ccgcacgatg atcgtgccgt gatcgaatc cagatccttg 9540
 acccgagtt gcaaaccctc actgatccgc atgcccgttc catacagaag ctgggcgaac 9600
 aaacgatgct cgccttcag aaaaccgagg atgcgaacca cttcatccgg ggtcagcacc 9660
 accggcaagc gcccggacgg ccgaggtctt ccgatctcct gaagccaggg cagatccgtg 9720
 cacagcactt gccgtagaag aacagcaagg ccgccaatgc ctgacgatgc gtggagaccg 9780

ES 2 657 825 T3

aaaccttgcg ctcgttcgcc agccaggaca gaaatgcctc gacttcgctg ctgccaag 9840
 ttgcccgggtg acgcacaccg tggaaacgga tgaaggcacg aaccagtg acataagcct 9900
 gttcggttcg taagctgtaa tgcaagtagc gtatgcgctc acgcaactgg tccagaacct 9960
 tgaccgaacg cagcgggtgt aacggcgcag tggcggtttt catggcttgt tatgactgtt 10020
 tttttgggtg acagtctatg cctcgggcat ccaagcagca agcgcgttac gccgtgggtc 10080
 gatgtttgat gttatggagc agcaacgatg ttacgcagca gggcagtcgc cctaaaacaa 10140
 agttaaacat catgagggaa gcggtgatcg ccgaagtatc gactcaacta tcagaggtag 10200
 ttggcgtcat cgagcgcctat ctgaaaccga cgttgctggc cgtacatttg tacggctccg 10260
 cagtggatgg cggcctgaag ccacacagtg atattgattt gctggttaag gtgaccgtaa 10320
 ggcttgatga aacaacgcgg cgagcttga tcaacgacct tttggaaact tcggcttccc 10380
 ctggagagag cgagattctc cgcgctgtag aagtcacat tgttgtgcac gacgacatca 10440
 ttccgtggcg ttatccagct aagcgcgaac tgcaatttgg agaatggcag cgcaatgaca 10500
 ttcttgacag tatcttcgag ccagccacga tcgacattga tctggctatc ttgctgacaa 10560
 aagcaagaga acatagcgtt gccttggtag gtccagcggc ggaggaactc tttgatccgg 10620
 ttctgaaca ggatctattt gaggcgctaa atgaaacctt aacgctatgg aactcgccgc 10680
 ccgactgggc tggcgatgag cgaaatgtag tgcttacgtt gtcccgcatt tggtagcagc 10740
 cagtaaccgg caaaatcgcg ccgaaggatg tcgctgcgca ctgggcaatg gagcgcctgc 10800
 cggcccagta tcagcccgtc atacttgaag ctagacaggc ttatcttga caagaagaag 10860
 atcgtttggc ctgcgcgca gatcagttgg aagaatttgt ccactacgtg aaaggcgaga 10920
 tcaccaaggt agtcggcaaa taatgtctaa caattcgttc aagccgacgc cgcttcgcg 10980
 cgcggcttaa ctcaagcgtt agatgcacta agcacataat tgctcacagc caaactatca 11040
 ggtcaagtct gcttttatta tttttaagcg tgcataataa gccctacaca aattgggaga 11100
 tataatcatga aaggctggct ttttcttgtt atcgcaatag ttggcgaagt aatcgcaaca 11160
 tcgcattaa aatctagcga gggctttact aagctagctt gcttggctgt tcgggtaccg 11220
 tgaacgtcgg ctcgattgta cctgcgttca aatactttgc gatcgtgttg cgcgcctgcc 11280
 cggcgcgtcg gctgatctca cggatcgact gcttctctcg caacgccatc cgacggatga 11340
 tgtttaaag tcccatgtgg atcaactcgt tgcccgcgtc ctcaccgtgt tggggggaag 11400
 gtgcacatgg ctcagttctc aatggaaatt atctgcctaa ccggctcagt tctgcgtaga 11460
 aaccaacatg caagctccac cgggtgcaaa gcggcagcgg 11500

<210> 4
 <211> 10443
 <212> ADN

ES 2 657 825 T3

<213> Artificial

<220>

<223> vector

<400> 4

```

agattcgaag ctcggtcccg tgggtgttct gtcgtctcgt tgtacaacga aatccattcc      60
cattccgcgc tcaagatggc ttcccctcgg cagttcatca gggctaaatc aatctagccg     120
acttgtccgg tgaaatgggc tgcaactcaa cagaacaatc caaacaaca tacacagcga     180
cttattcaca cgcgacaaat tacaacggta tatatcctgc cagtactcgg ccgtcgcacct    240
gcaggcaatt ggtacctacg tatgcatggc gcgccatatg caccatacat ggagtcaaaa     300
atcagatcgc aggatctaac agaactcggc gtgaagactg gcgaacagtt catacagagt     360
cttttacgac tcaatgacaa gaagaaaatc ttcgtcaaca tgggtggagca cgacactctc     420
gtctactcca agaatatcaa agatacagtc tcagaagacc aaagggctat tgagactttt     480
caacaaaggg taatatcggg aaacctcctc ggattccatt gccagctat ctgtcacttc     540
atcaaaagga cagtagaaaa ggaaggtggc acctacaaat gccatcattg cgataaagga     600
aaggttatcg ttcaagatgc ctctgccgac agtgggtcca aagatggacc cccaccacg     660
aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt ggattgatgt     720
gatatctcca ctgacgtaag ggatgacgca caatcccact atccttcgca agacccttc     780
tctatataag gaagttcatt tcatttggag aggactcgag aattaagcaa aagaagaaga     840
agaagaagtc caaaccatg gctaaacccc ccaagaagaa gcgcaaggtt aacatcaaaa     900
aaaaccaggt aatgaacctg ggtccgaact ctaaactgct gaaagaatac aatcccagc     960
tgatcgaact gaacatcgaa cagttcgaag caggtatcgg tctgatcctg ggtgatgctt    1020
acatccgttc tcgtgatgaa ggtaaaacct actgtatgca gttcagatgg aaaaacaaag    1080
catacatgga ccacgtatgt ctgctgtacg atcagtggtt actgtccccg ccgcacaaaa    1140
aagaacgtgt taaccacctg ggtaacctgg taatcacctg gggcgcccag actttcaaac    1200
accaagcttt caacaaactg gctaacctgt tcatogttaa caacaaaaaa accatccoga    1260
acaacctggt tgaaaactac ctgacccoga tgtctctggc atactggttc atggatgatg    1320
gtggtaaatg ggattacaac aaaaactcta ccaacaatc gatcgtactg aacaccagt     1380
ctttcacttt cgaagaagta gaatacctgg ttaaggtctc gcgtaacaaa ttccaactga    1440
actgttacgt aaaaatcaac aaaaacaaac cgatcatcta catcgattct atgtcttacc    1500
tgatcttcta caacctgatc aaacctgacc tgatcccgca gatgatgtac aaactgccga    1560
acactatctc ctccgaaact ttctgaaat agggctagca agcttgaca cgctgaaatc     1620
accagtctct ctctacaaat ctatctctct ctatcttctc cataataatg tgtgagtagt    1680
tcccagataa gggaattagg gttcctatag ggtttcgtc atgtgttgag catataagaa    1740

```

ES 2 657 825 T3

acccttagta tgtatttgta tttgtaaaat acttctatca ataaaatttc taattcctaa 1800
 aacccaaaatc cagtactaaa atccagatca tgcattggtac agcggccgcg ttaacgcgta 1860
 tactctagag cgatcgccat ggagccattt acaattgaat atatcctgcc gccgctgccg 1920
 ctttgacccc ggtggagcct gcatgttggt ttctacgcag aactgagccg gttaggcaga 1980
 taatttccat tgagaactga gccatgtgca ccttcccccc aacacgggtga gcgacggggc 2040
 aacggagtga tccacatggg acttttaaac atcatccgtc ggatggcggt gcgagagaag 2100
 cagtcgatcc gtgagatcag ccgacgcacc gggcaggcgc gcaacacgat cgcaaagtat 2160
 ttgaacgcag gtacaatcga gccgacgttc acggtaccgg aacgaccaag caagctagct 2220
 tagtaaagcc ctgcctagat tttaatgcgg atgttgcat tacttcgcca actattgcca 2280
 taacaagaaa aagccagcct ttcattgat atctccaat ttgtgtaggg cttattatgc 2340
 acgcttaaaa ataataaaag cagacttgac ctgatagttt ggctgtgagc aattatgtgc 2400
 ttagtgcatt taacgcttga gtttaagccgc gccgcgaagc ggctgcggct tgaacgaatt 2460
 gttagacatt atttgccgac taccttggtg atctcgcctt tcacgtagtg gacaaattct 2520
 tccaactgat ctgcgcgca ggccaagcga tcttcttctt gtccaagata agcctgtcta 2580
 gcttcaagta tgacgggctg atactgggcc ggcaggcgcct ccattgccca gtcggcagcg 2640
 acatccttcg gcgcatgctt gccggttact gcgctgtacc aaatgcccga caacgtaagc 2700
 actacatttc gctcatcgcc agcccagtcg ggcggcaggt tccatagcgt taaggtttca 2760
 tttagcgcct caaatagatc ctgttcagga accggatcaa agagttcctc cgccgctgga 2820
 cctaccaagg caacgctatg ttctcttgct tttgtcagca agatagccag atcaatgtcg 2880
 atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc 2940
 agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact 3000
 tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttcaa aaggtcgttg 3060
 atcaaagctc gccgcgctgt ttcattcaagc cttacgggtca ccgtaaccag caaatcaata 3120
 tcaactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac 3180
 gtcggttcga gatggcgtc gatgacgcca actacctctg atagttgagt cgatacttcg 3240
 gcgatcaccg cttccctcat gatgtttaac tttgttttag ggcgactgcc ctgctgcgta 3300
 acatcgcttc tgctccataa catcaaacat cgaccacggt cgtaacgcgc ttgctgcttg 3360
 gatgcccag gcatagactg taccclaaaa aacagtcatt aacaagccat gaaaaccgcc 3420
 actgcgccgt taccaccgct gcgttcggtc aaggttctgg accagttgctg tgagcgcata 3480
 cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc 3540
 atccgcttcc acggtgtgct tcaccggca acctgggca gcagcgaagt cgaggcattt 3600

ES 2 657 825 T3

ctgtcctggc tggcgaacga gcgcaaggtt tcggtctcca cgcatcgtca ggcatcggcg 3660
 gccttgctgt tcttctacgg caagtgcgtg gcacggatct gccctggctt caggagatcg 3720
 gaagacctcg gccgtccggg cgcttgccgg tggtgctgac cccggatgaa gtggttcgca 3780
 tcctcgggtt tctggaaggc gagcatcgtt tgttcgcca gcttctgtat ggaacgggtc 3840
 gacgcgttta atgaccagca cagtcgtgat ggcaaggtca gaatagcgtc gaggtctgcc 3900
 tcgtgaagaa ggtggtgctg actcatacca ggcctgaatc gcccacatcat ccagccagaa 3960
 agtgagggag ccacggttga tgagagcttt gttgtaggtg gaccagttgg tgattttgaa 4020
 cttttgcttt gccacggaac ggtctgcgtt gtcggaaga tgcgtgatct gatccttcaa 4080
 ctcagcaaaa gttcgattta ttcaacaaag ccgccgtccc gtcaagtcag cgtaatgctc 4140
 tgccagtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga 4200
 aactgcaatt tattcatatc aggattatca ataccatatt ttgaaaaag ccgtttctgt 4260
 aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggctc 4320
 gcgattccga ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg 4380
 ttatcaagtg agaaatcacc atgagtgcg actgaatccg gtgagaatgg caaaagtta 4440
 tgcatctctt tccagacttg ttcaacaggc cagccattac gctcgtcatc aaaatcactc 4500
 gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa tacgcggctg 4560
 ttaaaggac aattacaaac aggaatcga tgcaaccggc gcaggaacac tgccagcgca 4620
 tcaacaatat ttccacctga atcaggatat tcttctaata cctggaatgc tgtttttccg 4680
 gggatcgcag tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc 4740
 ggaagaggca taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg 4800
 gcaacgctac ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat 4860
 cgatagattg tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa 4920
 tcagcatcca tgttgaatt taatcgcggc ctcgagcaag acgtttcccg ttgaatatgg 4980
 ctcataaacac cccttgatt actgtttatg taagcagaca gttttattgt tcatgatgat 5040
 atatttttat cttgtgcaat gtaacatcag agatttgag acacaacgtg gctttgttga 5100
 ataaatcgaa cttttgctga gttgaaggat cagatcacgc atcttcccga caacgcagac 5160
 cgttccgtgg caaagcaaaa gttcaaaatc accaactggt ccacctaaa caaagctctc 5220
 atcaaccgtg gctccctcac tttctggctg gatgatgggg cgattcaggc ctggatgag 5280
 tcagcaaacac cttcttcacg aggcagacct cagcgtatatt ctgaccttgc catcacgact 5340
 gtgctggca ttaaaccgct cgacggatca gtgagggttt gcaactgagg gtcaaggatc 5400
 tggatttcga tcacggcacg atcatcgtgc gggagggcaa gggctccaag gatcgggcct 5460
 tgatgttacc cgagagcttg gcaccagcc tgcgcgagca gggatcgatc caaccctcc 5520

ES 2 657 825 T3

gctgctatag tgcagtcggc ttctgacgtt cagtgcagcc gtcttctgaa aacgacatgt 5580
cgcacaagtc ctaagttacg cgacaggctg ccgccctgcc cttttcctgg cgttttcttg 5640
tcgctggttt tagtcgcata aagtagaata cttgcgacta gaaccggaga cattacgcca 5700
tgaacaagag cgccgccgct ggcctgctgg gctatgcccg cgtcagcacc gacgaccagg 5760
acttgaccaa ccaacggggc gaactgcacg cggccggctg caccaagctg ttttccgaga 5820
agatcaccgg caccagggcg gaccgcccg agctggccag gatgcttgac cacctacgcc 5880
ctggcgacgt tgtgacagtg accaggctag accgcctggc ccgcagcacc cgcgacctac 5940
tggacattgc cgagcgcata caggaggccg gcgcgggctt gcgtagcctg gcagagccgt 6000
gggcccacac caccacgccg gccggccgca tgggtgtgac cgtgttcgcc ggcattgccg 6060
agttcgagcg ttcctaatac atcgaccgca cccggagcgg gcgcgaggcc gccaggccc 6120
gaggcgtgaa gtttggtccc cgccctacc tcacccggc acagatcgc cacgcccgcg 6180
agctgatcga ccaggaaggc cgcaccgtga aagaggcggc tgcactgctt ggcgtgcatc 6240
gctcgaccct gtaccgcgca cttgagcgc gcgaggaagt gacgccacc gaggccaggc 6300
ggcgcggtgc cttccgtgag gacgcattga ccgaggccga cccctggcg gccgccgaga 6360
atgaaccca agaggaacaa gcatgaaacc gcaccaggac ggccaggacg aaccgttttt 6420
cattaccgaa gagatcgagg cggagatgat ccgcccggg tacgtgttcg agccgcccg 6480
gcacgtctca accgtgcggc tgcattgaaat cctggccggt ttgtctgatg ccaagctggc 6540
ggcctggccg gccagcttg ccgctgaaga aaccgagcgc cgcctctaa aaaggatg 6600
tgtatttgag taaaacagct tgcgtcatgc ggtcgtgctg tatatgatgc gatgagtaaa 6660
taaacaaata cgcaagggga acgcatgaag gttatcgtg tacttaacca gaaaggcggg 6720
tcaggcaaga cgaccatcgc aaccatcta gccgcgccc tgcaactcgc cggggccgat 6780
gttctgtag tcgattccga tccccaggc agtgcccggc attggggggc cgtgcccggaa 6840
gatcaaccgc taaccgttgt cggcatcgac cgcgccgacga ttgaccgca cgtgaaggcc 6900
atcgccggc gcgacttcgt agtgatcgac ggagcgcccc aggcggcgga cttggctgtg 6960
tccgcgatca aggcagccga cttcgtgctg attccggtgc agccaagccc ttacgacata 7020
tgggccaccg ccgacctggt ggagctggtt aagcagcgc ttgaggtcac ggatggaagg 7080
ctacaagcgg cttttgtcgt gtcgccccg atcaaaggca cgcgcatcgg cggtgaggtt 7140
gccgagggc tggccgggta cgagctgcc attcttgagt cccgtatcac gcagcgcgtg 7200
agctaccag gactgccgc cgcggcaca accgttctg aatcagaacc cgagggcgac 7260
gctgcccgcg aggtccaggc gctggccgct gaaattaaat caaaactcat ttgagttaat 7320
gaggtaaaga gaaaatgagc aaaagcaca acacgctaag tgccggccgt ccgagcgac 7380

ES 2 657 825 T3

gcagcagcaa ggctgcaacg ttggccagcc tggcagacac gccagccatg aagcgggtca 7440
actttcagtt gccggcggag gatcacacca agctgaagat gtacgoggta cgccaaggca 7500
agaccattac cgagctgcta tctgaataca tcgcgagct accagagtaa atgagcaaat 7560
gaataaatga gtagatgaat tttagcggct aaaggaggcg gcatggaaaa tcaagaacaa 7620
ccaggcaccg acgcoctgga atgccccatg tgtggaggaa cgggcgggtg gccaggcgta 7680
agcggctggg ttgtctgccc gccctgcaat ggcactggaa cccccaagcc cgaggaatcg 7740
gcgtgacggg cgcaaaccat ccggcccggg acaaatcggc gcggcgctgg gtgatgacct 7800
ggtggagaag ttgaaggccg cgcaggccgc ccagcggcaa cgcacgagg cagaagcacg 7860
ccccggtgaa tcgtggcaag cggccgctga tcgaatccgc aaagaatccc ggcaaccgcc 7920
ggcagccggg gcgccgtcga ttaggaagcc gcccaagggc gacgagcaac cagatTTTTT 7980
cgttccgatg ctctatgacg tgggcacccg cgatagtcgc agcatcatgg acgtggccgt 8040
tttccgtctg tcgaagcgtg accgacgagc tggcgaggtg atccgctacg agcttccaga 8100
cgggcacgta gaggtttccg cagggccggc cggcatggcc agtgtgtggg attacgacct 8160
ggtactgatg gcggtttccc atctaaccga atccatgaac cgataaccggg aagggaaggg 8220
agacaagccc ggccgcgtgt tccgtccaca cgttgccggc gtactcaagt tctgccggcg 8280
agccgatggc ggaaagcaga aagacgacct ggtagaaacc tgcattcggg taaacaccac 8340
gcacgttgcc atgcagcgtg cgaagaaggc caagaacggc cgctcgggta cggtatccga 8400
gggtgaagcc ttgattagcc gctacaagat cgtaaagagc gaaaccgggc ggccggagta 8460
catcgagatc gagctagctg attggatgta ccgcgagatc acagaaggca agaaccggga 8520
cgtgctgacg gttcaccocg attactTTTT gatcgatccc ggcatcggcc gttttctcta 8580
ccgcctggca cgccgcgccc caggcaaggc agaagccaga tggttgttca agacgatcta 8640
cgaacgcagt ggcagcggcg gagagttcaa gaagttctgt ttcaccgtgc gcaagctgat 8700
cgggtcaaat gacctgccgg agtacgattt gaaggaggag gcggggcagg ctggcccgat 8760
cctagtcatg cgctaccgca acctgatcga gggcgaagca tccgccgggt cctaatgtac 8820
ggagcagatg ctagggcaaa ttgccttagc aggggaaaaa ggtcgaaaag gtctctttcc 8880
tgtggatagc acgtacattg ggaacccaaa gccgtacatt gggaaccgga acccgatcat 8940
tgggaaccca aagccgtaca ttgggaaccg gtcacacatg taagtgactg atataaaaga 9000
gaaaaaaggc gatTTTTTccg cctaaaactc tttaaaactt attaaaactc ttaaaaccg 9060
cctggcctgt gcataactgt ctggccagcg cacagccgaa gagctgcaaa aagcgcctac 9120
ccttcggctg ctgcgctccc tacgccccgc cgcttcgctg cggcctatcg cggccgctgg 9180
ccgctcaaaa atggctggcc tacggccagg caatctacca gggcggggac aagccgcgcc 9240
gtcgccactc gaccgcggc gccacatca aggcaccctg cctcgcgctg ttcgggtgatg 9300

ES 2 657 825 T3

acggtgaaaa cctctgacac atgcagctcc cggagacggt cacagcttgt ctgtaagcgg 9360
 atgccgggag cagacaagcc cgtcagggcg cgtcagcggg tgttggcggg tgtcggggcg 9420
 cagccatgac ccagtcacgt agcgatagcg gagtgtatac tggcttaact atgcggcac 9480
 agagcagatt gtactgagag tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag 9540
 gagaaaatac cgcacaggc gctcttcgc ttctctcgc actgactcgc tgcgctcgtt 9600
 cgttcggctg cggcgagcgg taccagctca ctcaaaggcg gtaatacggg tatccacaga 9660
 atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 9720
 taaaaggcc gcggtgctgg cgtttttcca taggctccgc cccctgacg agcatcacia 9780
 aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 9840
 tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 9900
 gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 9960
 cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccggtcagcc 10020
 cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt 10080
 atcgccactg gcagcagcca ctggaacag gattagcaga gcgaggtatg taggcggtgc 10140
 tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggtat 10200
 ctgctctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 10260
 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 10320
 aaaaggatct caagaagatc cggaaaacgc aagcgcaaag agaaagcagg tagcttgacg 10380
 tggccttaca tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt 10440
 gcc 10443

REIVINDICACIONES

1. Un método para modificar el genoma de una célula de planta de algodón en un sitio predefinido, que comprende las etapas de
- 5 a. inducir una rotura de ADN de doble cadena en la vecindad de o en dicho sitio predefinido, induciéndose dicha rotura de la doble cadena mediante la introducción en dicha célula de una enzima endonucleasa de escisión rara que reconoce una secuencia de reconocimiento en la vecindad de o en dicho sitio predefinido;
- 10 b. seleccionar una célula vegetal en la que dicha rotura de ADN de doble cadena se ha reparado dando como resultado una modificación en el genoma en dicho sitio preseleccionado, en el que dicha modificación se selecciona de
- i. un reemplazo de al menos un nucleótido;
- ii. una delección de al menos un nucleótido;
- iii. una inserción de al menos un nucleótido; o
- iv. cualquier combinación de i.-iii.;
- 15 caracterizado por que dicha célula está comprendida dentro del callo embriogénico quebradizo, en donde dicho callo embriogénico quebradizo se ha mantenido en condiciones de luz tenue y en un medio que comprende carbono activo, en donde dichas condiciones de luz tenue implican una intensidad de aproximadamente $1 \text{ a } 7 \mu\text{mol m}^2 \text{ s}^{-1}$ con un fotoperiodo de aproximadamente 16 horas de luz/8 horas de oscuridad.
- 20 2. El método de la reivindicación 1, en el que dicho callo embriogénico quebradizo se ha mantenido bajo condiciones de luz tenue y en un medio que comprende carbono activo durante aproximadamente 2 a 4 meses.
3. El método de la reivindicación 1 o 2, en el que dicha enzima endonucleasa se introduce en dicha célula mediante la administración a dicha célula de una molécula de ADN que codifica dicha enzima endonucleasa.
4. El método de una cualquiera de las reivindicaciones 1-3, en el que antes de la etapa b. se suministra una molécula de ADN reparador extraño a dicha célula, utilizándose dicha molécula de ADN reparador extraño como un molde para la reparación de dicha rotura de ADN de doble cadena.
- 25 5. El método de una cualquiera de las reivindicaciones 1-4, en el que dicho callo embriogénico se incuba en medio sin hormonas.
6. El método de una cualquiera de las reivindicaciones 1-5, en el que dicho callo embriogénico se incuba durante 1 a 4 días en un medio no selectivo durante o después de dicha introducción de dicha enzima endonucleasa.
- 30 7. El método de una cualquiera de las reivindicaciones 3-6, en el que dicha administración de ADN se realiza mediante bombardeo con partículas o utilizando *Agrobacterium*.
8. El método de acuerdo con una cualquiera de las reivindicaciones 3-7, en el que dicho ADN reparador extraño comprende al menos una secuencia de nucleótidos flanqueante que tiene suficiente homología con la región de ADN aguas arriba o aguas abajo de dicho sitio predefinido para permitir la recombinación con dicha región de ADN aguas arriba o aguas abajo, en el que dicho ADN reparador extraño comprende dos secuencias de nucleótidos flanqueantes situadas en extremos opuestos de dicho ADN extraño, teniendo una de dichas secuencias de nucleótidos flanqueantes suficiente homología con la región de ADN aguas arriba de dicho sitio predefinido, teniendo la otra secuencia de nucleótidos flanqueante suficiente homología con la secuencia aguas abajo de dicho sitio predefinido para permitir la recombinación entre dichas secuencias de nucleótidos flanqueantes y dichas regiones de
- 35 ADN aguas arriba y aguas abajo.
9. El método de acuerdo con una cualquiera de las reivindicaciones 4-8, en el que dicho ADN reparador extraño comprende un gen marcador seleccionable.
10. El método de acuerdo con una cualquiera de las reivindicaciones 4-9, en el que dicho ADN extraño comprende un gen de interés expresable en plantas.
- 45 11. El método de acuerdo con la reivindicación 10, en el que dicho gen expresable en plantas se selecciona del grupo de un gen de tolerancia a herbicidas, un gen de resistencia a insectos, un gen de resistencia a enfermedades, un gen de resistencia al estrés abiótico, un gen que codifica una enzima implicada en la biosíntesis de aceite o la biosíntesis de hidratos de carbono, un gen que codifica una enzima implicada en la resistencia de las fibras o la longitud de las fibras, un gen que codifica una enzima implicada en la biosíntesis de metabolitos secundarios.
- 50 12. El método de acuerdo con una cualquiera de las reivindicaciones 4-7, en el que dicho ADN extraño consiste en dos secuencias de nucleótidos flanqueantes, teniendo una de dichas secuencias de nucleótidos flanqueantes suficiente homología con la región de ADN aguas arriba de dicho sitio predefinido, teniendo la otra secuencia de nucleótidos flanqueante suficiente homología con la secuencia aguas abajo de dicho sitio predefinido para permitir la

recombinación entre dichas secuencias de nucleótidos flanqueantes y dichas regiones de ADN aguas arriba y aguas abajo.

13. El método de una cualquiera de las reivindicaciones 1-3 o 5-7, en el que dicho sitio preseleccionado está flanqueado por dos regiones con suficiente homología para la recombinación entre sí.

- 5 14. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 13, en el que dicha célula de la planta de algodón se regenera adicionalmente en una planta de algodón.

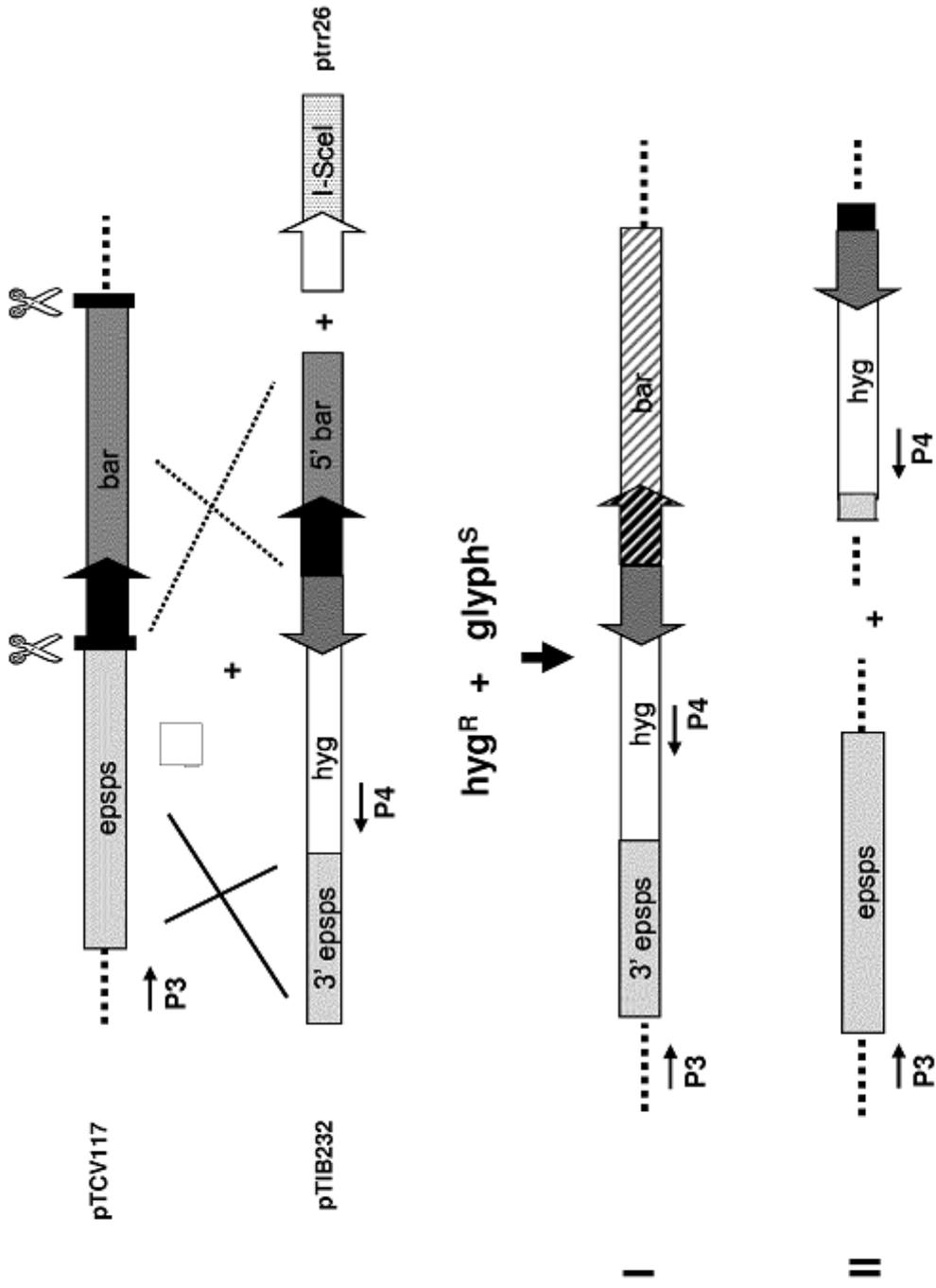


Figure 1

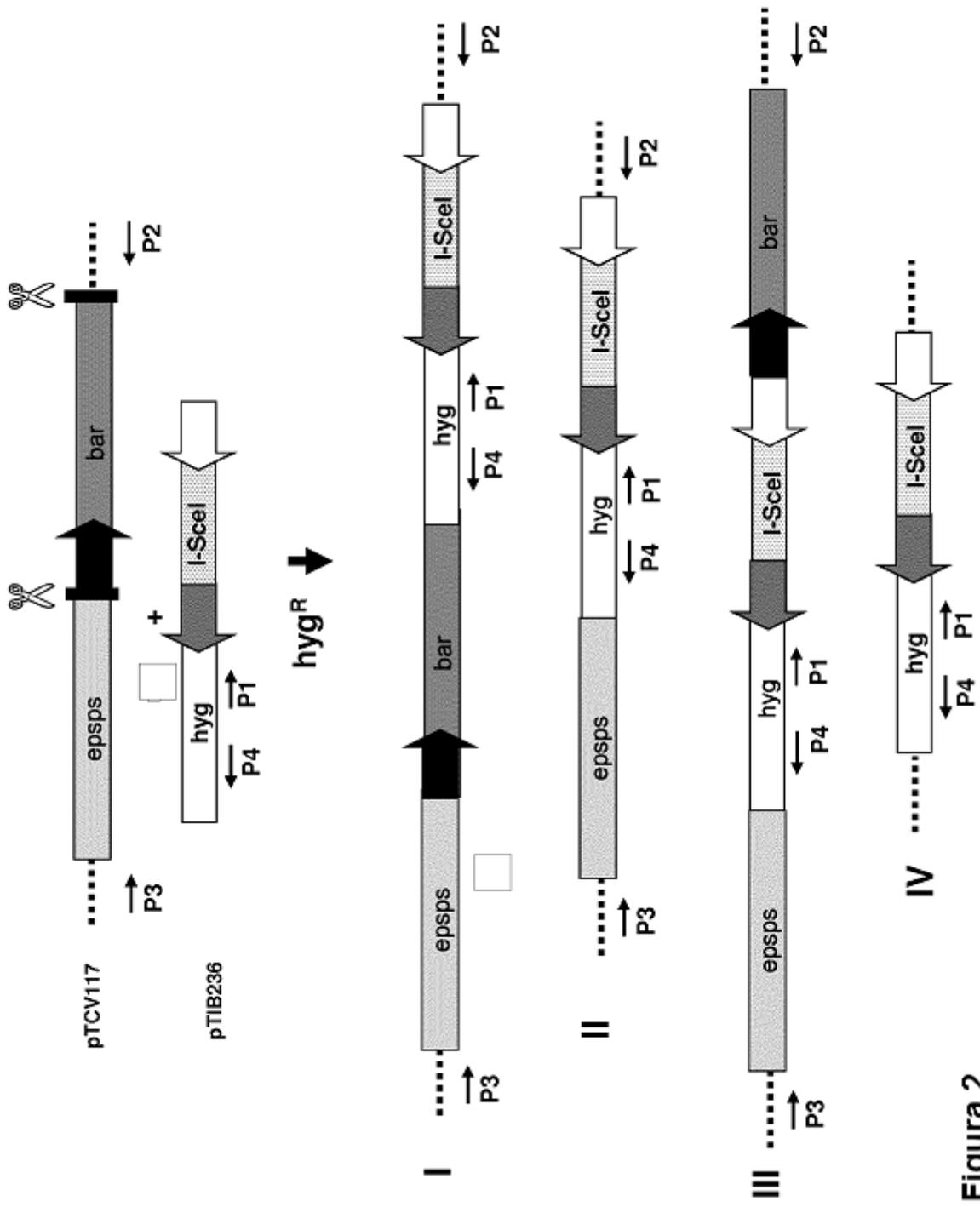


Figura 2