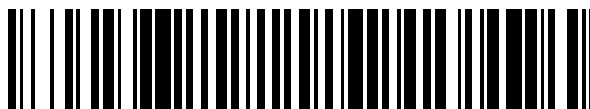


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 701 243**

51 Int. Cl.:

**C12N 9/10** (2006.01)

**C12N 15/82** (2006.01)

**C07K 14/415** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.10.2014** **PCT/EP2014/072678**

87 Fecha y número de publicación internacional: **30.04.2015** **WO15059205**

96 Fecha de presentación y número de la solicitud europea: **22.10.2014** **E 14787180 (0)**

97 Fecha y número de publicación de la concesión europea: **26.09.2018** **EP 3060576**

54 Título: **Nuevo promotor preferente en fibras en algodón**

30 Prioridad:

**24.10.2013 EP 13189991**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**21.02.2019**

73 Titular/es:

**BASF AGRICULTURAL SOLUTIONS SEED US  
LLC (50.0%)  
100 Park Avenue  
Florham Park, NJ 07932, US y  
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL  
RESEARCH ORGANISATION (50.0%)**

72 Inventor/es:

**MEULEWAETER, FRANK y  
LLEWELLYN, DANNY J.**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 701 243 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

Nuevo promotor preferente en fibras en algodón

## Campo de la invención

- 5 La presente invención se refiere a materiales y métodos para la expresión de un gen de interés preferentemente o selectivamente en fibras de plantas, tales como plantas de algodón. En particular, la invención proporciona promotores, regiones promotoras y casetes de expresión para regular la expresión preferente en fibras o selectiva en fibras en plantas de algodón.

## Antecedentes de la invención

- 10 La fibra de algodón es el material textil simple más importante en el mundo. Aproximadamente 32 millones de hectáreas (80 millones de acres) de algodón son recogidos anualmente en todo el mundo. El algodón es el quinto cultivo más grande en los EE. UU. en cuanto a producción en hectáreas, con una media de 4,2 millones de hectáreas (10,3 millones de acres) plantados en los años 2006 a 2008. Aproximadamente 90% del algodón cultivado en el mundo es *Gossypium hirsutum* L., mientras que *Gossypium barbadense* responde de aproximadamente 8%. Por consiguiente, la modificación de las características de las fibras de algodón para ajustarse mejor a los requisitos de la industria y el consumidor es un esfuerzo importante en la reproducción bien mediante métodos clásicos o bien al alterar genéticamente el genoma de las plantas de algodón. Los objetivos que se han de alcanzar incluyen un incremento de la longitud, la resistencia, la capacidad de tinción de las fibras de borra, una disminución de la producción de fibras apelmazadas, la relación de madurez de las fibras, el contenido de fibras inmaduras, la uniformidad de las fibras y el micronaire.

- 20 El desarrollo de las fibras de algodón es un procedimiento multifásico bajo la regulación de un gran número de genes, muchos de los cuales son regulados al alza o altamente expresados en el desarrollo de las células de las fibras (Li, C.H. y cols., 2002; Ruan y cols., 2003; Wang, S. y cols., 2004; Li y cols., 2005; Luo y cols., 2007).

- 25 Cada fibra de algodón es una sola célula epidérmica diferenciada del óvulo. Se producen aproximadamente medio millón de fibras por cápsula de algodón, formando algunas pelusa y formando algunas borra. La iniciación de una célula epidérmica en una fibra requiere un cambio en el destino celular, que es un proceso biológico fundamental que implica "conmutaciones" genéticas, fisiológicas y de desarrollo. Las mutaciones genéticas, la poliploidía, la polinización/fertilización y la regulación hormonal pueden afectar al número de células que se desarrollan en fibras o alterar las propiedades celulares de las fibras (pelusa frente a borra). Sin embargo, no está claro cómo estos factores controlan los cambios en la expresión génica que orquestan el patrón y el ritmo en fases tempranas del desarrollo de la fibra.

- 35 En contraste, el desarrollo morfológico de fibras de algodón está bien documentado en la especialidad. Las fibras de algodón sufren cuatro fases de desarrollo solapadas: iniciación de las células de las fibras, elongación, biosíntesis de la pared secundaria y maduración. La iniciación de las fibras es un proceso rápido. Las fibras pilosas blancas empiezan a desarrollarse inmediatamente después de la antesis y continúan hasta 3 días después de la antesis (DPA), que está seguida por elongación de las células de las fibras (hasta 20 DPA). La biosíntesis de la pared secundaria se inicia alrededor de 15 dpa y continúa hasta 35 DPA, seguido por un proceso de maduración hasta 45-60 DPA. Las fibras de algodón se derivan de células epidérmicas ovulares (tejidos maternos). Sin embargo, solo ~25-30% de las células epidérmicas se diferencian en las fibras de borra comercialmente importantes. La mayoría de las células no se diferencian en fibras o se desarrollan en fibras cortas o pelusa. Para las células comprometidas con el desarrollo de fibras, la iniciación y la elongación de las células son casi sincrónicas en cada óvulo, indicando que los cambios en la expresión génica están orquestados durante la diferenciación y el desarrollo de las fibras a través de mecanismos de señalización y/o temporización intercelulares.

Se han descrito diversos promotores que conducen la expresión de genes en la semilla de algodón. También se conocen diversos promotores que conducen la expresión preferente en fibras en el algodón.

- 50 E6 fue el primer gen de fibra de algodón identificado, y el promotor de E6 se ha usado para manipular la calidad de las fibras de algodón (John y Keller, 1996). GhRDL1, un gen altamente expresado en células de fibras de algodón en la fase de elongación, codifica una proteína que contiene el dominio BURP (Li, C.H. y cols., 2002), y el promotor de GhRDL1 exhibía una actividad específica en tricomas en plantas de *Arabidopsis* transgénicas (Wang, S. y cols., 2004). Los transcritos de GhTUB1 se acumulan preferentemente a altos niveles en la fibra, según esto, el gen de fusión pGhTUB1::GUS se expresaba en un alto nivel en fibra pero a niveles muy inferiores en otros tejidos (Li, X.B. y cols., 2002). Los promotores de tres genes de proteínas de transferencia de lípidos de algodón, LTP3, LTP6 y FSltp4, eran capaces de dirigir la expresión del gen GUS en tricomas secretores glandulares (GSTs) de hojas y tallo en plantas de tabaco transgénicas (Hsu y cols., 1999; Liu y cols., 2000; Delaney y cols., 2007), sin embargo, no exhibían una clara especificidad en tejidos. Por ejemplo, en plantas de tabaco transgénicas pFSltp4::GUS, se podía

detectar una fuerte actividad de GUS en todos los tipos de tricomas; además, la expresión de GUS también era visible en el margen foliar, el tejido vascular, los óvulos y las puntas de las raíces (Delaney y cols., 2007).

Se ha encontrado que el factor GaMYB2 de transcripción de algodón R2R3 MYB es un homólogo funcional de GLABRA1 (GL1) de Arabidopsis, un regulador clave de la formación de tricomas de Arabidopsis. GaMYB2 se expresa en células de fibras de algodón en las fases de desarrollo tempranas (Wang, S. y cols., 2004). Su promotor conduce la expresión específica de tricomas también en Arabidopsis y en la expresión específica de espigas de GST en tabaco (Shangguan y cols., 2008).

La patente de EE. UU. 7.626.081 divulga un promotor específico de semillas de algodón encontrado en el gen de alfa globulina. El promotor *Gh-sp* se deriva de un gen de proteína de semilla y es activo solamente en semillas de algodón en maduración (Song y cols., 2000).

La solicitud de patente de EE. UU. 2003/0106089 divulga un gen expresado de un modo específico en fibras y su promotor.

La patente de EE. UU. 6.211.430 divulga un promotor que dirige la expresión durante el desarrollo de fibras tardío.

La solicitud de patente de EE. UU. 20130081154 divulga otro promotor que dirige la expresión durante el desarrollo de fibras tardío.

A pesar del hecho de que están disponibles en la especialidad muchos promotores que se sabe que conducen la expresión preferente en las fibras o preferente en las semillas en plantas de algodón, sigue habiendo una necesidad de promotores alternativos capaces de expresión preferente o selectiva en fibras, p. ej. para la expresión independiente de varias secuencias de ADN extrañas de interés sin la posibilidad de inactivación postranscripcional debido al uso del mismo promotor. Además, los promotores preferentes en fibras o selectivos en fibras conocidos dirigen cada uno un patrón de expresión temporal, espacial y/o de desarrollo particular, que no siempre se ajusta a los objetivos particulares. Así, sigue habiendo una necesidad de nuevos promotores preferentes en fibras o selectivos en fibras con la capacidad de controlar la transcripción en el desarrollo de células de fibras, preferiblemente de un modo más selectivo.

Esta invención proporciona una solución a estos problemas como se describe posteriormente en el sumario, las realizaciones detalladas, los ejemplos, los dibujos y las reivindicaciones.

## Sumario de la invención

La presente solicitud divulga un ácido nucleico aislado que comprende (a) al menos aproximadamente 700 nucleótidos consecutivos de SEQ ID NO: 1, (b) un ácido nucleico que tiene una secuencia nucleotídica que comprende la secuencia nucleotídica de SEQ ID No 1 desde la posición 1 a la posición 922, o (c) una secuencia nucleotídica con al menos 95% de identidad de secuencia con las secuencias nucleotídicas de (a) o (b), teniendo dicho ácido nucleico aislado una actividad promotora preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) y no expresándose en fibras de 10 dpa y no siendo más largo de 959 nucleótidos. También se divulga en la presente un gen quimérico que comprende una región promotora como la descrita en la presente, ligada operativamente a un ácido nucleico que codifica un producto de expresión de interés, y opcionalmente una secuencia de terminación de la transcripción y poliadenilación. También se divulga en la presente una célula de planta transgénica, una planta transgénica y una semilla como las caracterizadas en las reivindicaciones. Los métodos divulgados en la presente se refieren a la producción de una planta transgénica, que efectúa una expresión preferente en fibras de un producto en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa y a la alteración de las propiedades de las fibras en una planta de algodón según se caracteriza en las reivindicaciones.

## Breve descripción de las figuras

Figura 1: Expresión relativa de genes de tipo Fb8 en diferentes tejidos en comparación con fibras de 35 dpa usando el gen de ubiquitina de algodón como referencia. Los cebadores eran para un consenso de un número de ESTs tipo Fb8 muy similares. Los valores son la media de cuatro réplicas técnicas.

Figura 2: Alineamiento de secuencias proteínicas deducidas de la proteína de tipo Fb8 codificada por BAC8K12 en comparación con la proteína codificada por ADNc procedente de proteína de tipo cFB8 nº 1 y la de tipo FB8 codificada por BAC150F2 (descrito en la solicitud de patente de EE. UU. 20130081154) alineadas usando ClustalW. Los aminoácidos idénticos se encierran en cajas en negro y los aminoácidos similares en gris.

Figura 3: Una representación esquemática del vector que comprende la región promotora de tipo Fb8(8K12) ligada operativamente a la región codificante GUS, que se ha usado para transformar algodón.

Figura 4: Secuencia nucleotídica de la región promotora del gen tipo Fb8(8K12) en *Gossypium hirsutum* (SEQ ID No. 1).

Figura 5: tinción de GUS de fibra de 28 dpa de cultivos de óvulos de algodón bombardeados con tipo Fb8(8K12)-GUS.

Figura 6: patrones de tinción de GUS en diferentes tejidos de plantas T0 de la línea de algodón T447-22-2 transformada con la construcción tipo Fb8(8K12)-promotor GUS. Se observó una tinción intensa en 20 dpa, pero no en fibras de 10 dpa. La tinción se observó en la envuelta de las semillas de semillas tanto de 10 como de 20 dpa. Había alguna tinción en glándulas pequeñas en los peciolo foliares (flechas blancas). Nótese la ausencia de tinción de tricomas.

Figura 7: tinción de GUS alrededor de los bordes cortados de cotiledones de una plántula T1 de T447-22-2 que expresa la construcción tipo Fb8(8K12)-promotor GUS, pero no a lo largo de bordes que no se han cortado. La tinción también es evidente en los cortes cercanos a la vasculatura y también tricomas glandulares pequeños situados por encima de la vasculatura.

Figura 8: Expresión de GUS en óvulos, semillas y fibras de algodón de diferentes edades para la planta T1 23 transgénica de la línea T447-22-2. Nótese que la tinción de GUS es la más intensa en las fibras más viejas. También hay alguna tinción de la superficie del óvulo y la envuelta de la semilla en las semillas más viejas.

Figura 9: Expresión de GUS en diferentes tejidos de algodón procedentes de la planta T1 23 de la línea T447-22-2 que expresa la construcción tipo Fb8(8K12)-promotor GUS. Panel A) Tinción en venas que radian desde zonas lesionadas en la base de una hoja madura. Panel B) Tinción de los dos extremos cortados de una sección del peciolo foliar. Panel C) Tinción débil de la superficie del tallo y las glándulas subepidérmicas (el recuadro muestra una sección transversal del tallo). Panel D) Tinción de GUS de los bordes cortados alrededor del tabique en una cápsula de 10 dpa de edad. Panel E) tinción de una yema de algodón de 0 dpa seccionada con tinción en la base en la que se han retirado los pétalos y las brácteas y alrededor de los óvulos de algodón. Panel F) Ausencia de tinción en una raíz con raíces laterales en desarrollo. Panel G) tinción de GUS del estilo en el que se han retirado los pétalos y las brácteas y alrededor de la punta del estigma. Las paredes de las anteras no se tiñen y los filamentos se tiñen ligeramente. Nótese la tinción intensa en los granos de polen en segregación (el contenido de la célula se ha perdido durante la clarificación) (recuadro). Panel H) Tinción en la vasculatura del extremo distal del pétalo.

Figura 10: Agotamiento de Acid Orange 7 después de la tinción a partir de una planta de algodón transgénica que comprende el ADN T de pTIB359. Barras gris claro: fibras procedentes de plantas de control no transgénicas; Barras gris oscuro: fibras transgénicas procedentes de líneas de algodón transgénicas.

## Descripción detallada de la invención

### Comentarios generales

El listado de secuencias que está contenido en el archivo llamado BCS13-2022\_ST25.txt, que tiene 36 kilobytes (medidos en el sistema operativo MS windows), comprende las secuencias 1 a 10 y fue creado el 21 de octubre de 2013, está archivado con la presente.

Según se usa en la presente, se ha de interpretar que el término "que comprende" especifica la presencia de las características, los números enteros, las etapas o los componentes indicados a los que se hace referencia, pero no excluye la presencia o la adición de una o más características, números enteros, etapas o componentes, o grupos de los mismos. Así, p. ej., un ácido nucleico que comprende una secuencia de nucleótidos puede comprender más nucleótidos que los realmente citados, es decir, estar incluido en un ácido nucleico mayor. Un gen quimérico como se describirá más adelante que comprende un ácido nucleico que esta funcionalmente o estructuralmente definido puede comprender ácidos nucleicos adicionales, etc. Sin embargo, en el contexto de la presente divulgación, el término "que comprende" también incluye "que consiste en". En otras palabras, la terminología que se refiere a un ácido nucleico "que comprende" una cierta secuencia nucleotídica, según se usa a lo largo del texto, se refiere a un ácido nucleico o una proteína que incluye o que contiene al menos la secuencia descrita, de modo que otras secuencias nucleotídicas o de aminoácidos puedan estar incluidas en el extremo 5' (o N-terminal) y/o 3' (o C-terminal), p. ej. (la secuencia nucleotídica de) una proteína marcadora seleccionable, (la secuencia nucleotídica de) un péptido transitorio y/o una secuencia líder 5' o una secuencia de remolque 3'.

### Ácidos nucleicos

En otro aspecto, la presente solicitud divulga un ácido nucleico aislado que comprende (a) al menos 700 nucleótidos consecutivos de SEQ ID NO: 1, preferiblemente al menos 700 nucleótidos consecutivos aguas arriba de la posición del nucleótido 959 en SEQ ID No 1; (b) una secuencia nucleotídica con al menos 95% de identidad de secuencia con la secuencia nucleotídica de (a); (c) una secuencia de ácido nucleico que se hibrida bajo condiciones rigurosas a



la secuencia nucleotídica de (a) o (b); y (d) una secuencia de ácido nucleico complementaria a la secuencia nucleotídica de una cualquiera de (a) a (c), en donde dicho ácido nucleico aislado tiene actividad promotora preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) y no se expresa en fibras de 10 dpa, y no tiene una longitud mayor de 959 nucleótidos.

En un aspecto particular, la presente solicitud divulga el ácido nucleico de SEQ ID No. 1 y secuencias nucleotídicas con al menos 95% de identidad de secuencia con la secuencia nucleotídica completa de SEQ ID No. 1 que tienen actividad promotora preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) y no se expresa en fibras de 10 dpa, y no tiene una longitud mayor de 959 nucleótidos.

Se ha observado que la secuencia nucleotídica del promotor y la región promotora preferente en las fibras de esta invención comparte un alto grado de identidad de secuencia nucleotídica con el promotor preferente en fibras descrito en la solicitud de patente de EE. UU. 20130081154, particularmente en la región justo aguas arriba del codón ATG de iniciación, entre las posiciones nucleotídicas 427 y 959 de SEQ ID No. 1. Otra región de alta identidad de secuencia nucleotídica entre estos dos promotores preferentes en fibras se puede encontrar más aguas arriba (posición 1 a 361) separada por una región de poca identidad de secuencia. Por otra parte, la transcripción se puede iniciar en la posición 922 o la posición 926 de SEQ ID No. 1.

También se describe un ADN promotor preferente en fibras que comprende una secuencia nucleotídica que tiene al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% de identidad de secuencia con o que es idéntica a la secuencia nucleotídica de SEQ ID No. 1 entre la posición nucleotídica 427 y 922 o que tiene al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% de identidad de secuencia con o que es idéntica a la secuencia nucleotídica de SEQ ID No. 1 entre la posición nucleotídica 427 y 926, o un promotor preferente en las fibras que tiene al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% de identidad de secuencia con o que es idéntica a la secuencia nucleotídica de SEQ ID No. 1 entre la posición nucleotídica 427 y 959.

Los ácidos nucleicos aislados de este aspecto también se denominan posteriormente en la presente "promotor" o "región promotora".

A menos que se indique otra cosa, las realizaciones descritas posteriormente para la secuencia promotora divulgada en la presente también son aplicables a realizaciones respectivas de otros aspectos divulgados en la presente.

Un "ácido nucleico aislado" o "secuencia de ácido nucleico aislado", según se usa en la presente solicitud, se refiere a un ácido nucleico que se define anteriormente que no está presente en la naturaleza, tal como, por ejemplo, un ácido nucleico artificial o sintético con una secuencia nucleotídica diferente que el ácido nucleico presente en la naturaleza o un ácido nucleico que es más corto que uno presente en la naturaleza, o que no es más largo en el ambiente natural en el que originalmente estaba presente, p. ej., una secuencia codificante de ácido nucleico asociada con un elemento regulador heterólogo tal como una secuencia codificante bacteriana ligada operativamente a un promotor expresable en plantas, o en un gen quimérico o un ácido nucleico transferido a otra célula hospedadora, tal como una célula de planta transgénica.

La longitud de la secuencia del ácido nucleico aislado de la invención, tal como, p. ej., un fragmento de SEQ ID NO: 1 según se divulga en la presente, y su posición dentro de SEQ ID NO: 1 se ha de elegir de modo que sea suficientemente largo, p. ej. que comprenda todos los elementos necesarios y suficientes, y situarse de modo que sea capaz de inducir expresión preferente en fibras. Las regiones promotoras se pueden elegir para comprender al menos 700 nucleótidos consecutivos agua arriba de la posición nucleotídica 959 en la SEQ ID No 1.

Son conocidos para el experto métodos para evaluar si una secuencia de ácido nucleico como la descrita anteriormente, que en la presente solicitud representa una secuencia promotora, es capaz de inducir la expresión de un gen quimérico en el que está comprendida o, en particular, de una secuencia de ácido nucleico ligada operativamente a la misma, de un modo preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa.

Por ejemplo, se pueden realizar estudios en genes indicadores a fin de evaluar la función inductora de la expresión de una secuencia de ácido nucleico. Esto incluye ligar operativamente la secuencia de ácido nucleico de la invención a un gen indicador tal como GUS, introducir la construcción de ácido nucleico resultante en una planta o una célula de planta, tal como en una planta de algodón, y evaluar la inducción de la expresión de dicho gen informador en diferentes tejidos de dicha planta, como también se describirá con más detalle más adelante.

El fragmento de la secuencia de ácido nucleico descrita en la presente que tiene actividad promotora preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa puede comprender según esto al menos 700, al menos 800, al menos 900, al menos 950 de SEQ ID NO: 1. En otro ejemplo, dicho ácido nucleico comprende la secuencia de ácido nucleico de SEQ ID NO: 1. En otro ejemplo, dicha secuencia promotora consiste en la secuencia de ácido nucleico de SEQ ID NO: 1.

Sin embargo, estará claro que también se pueden usar con el mismo efecto variantes de los ácidos nucleicos descritos en la presente, incluyendo inserciones, eliminaciones y sustituciones de los mismos. Generalmente, estas variantes tienen al menos 95%, al menos 96%, al menos 97%, al menos 98% o incluso al menos 99% de identidad de secuencia con SEQ ID NO: 1. Estas variantes deben retener su actividad promotora preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa, y no deben tener una longitud mayor de 959 nucleótidos.

Según se usa en la presente, el término "promotor" indica cualquier secuencia de ácido nucleico, tal como una secuencia de ADN, que sea reconocida y se una (directamente o indirectamente) mediante una ARN polimerasa dependiente de ADN durante la iniciación de la transcripción, dando como resultado la generación de una molécula de ARN que es complementaria al ADN transcrito. Esta región también se puede denominar una "región reguladora 5'". Los promotores están situados habitualmente aguas arriba de la región no traducida (UTR) 5' que precede a la secuencia codificante de proteína que se va a transcribir y tiene regiones que actúan como sitios de unión para ARN polimerasa II y otras proteínas tales como factores de transcripción para iniciar la transcripción de una secuencia ligada operativamente. Los promotores pueden contener ellos mismos subelementos (es decir, motivos promotores) tales como elementos cis o dominios mejoradores que regulan la transcripción de genes ligados operativamente. El promotor y una UTR 5' conectada también se denomina "región promotora".

Un promotor o una región promotora "preferente en fibras" en el contexto de la presente invención significa que la transcripción de una secuencia de ácido nucleico controlada por un promotor es al menos 2 veces superior, al menos 5 veces superior, al menos 10 veces superior o incluso 20 veces superior en una célula de fibra que en células de otro tejido de planta tales como, por ejemplo, tejido foliar. El término no tiene en cuenta la expresión inducida por lesiones potencial. En otras palabras, la expresión preferente en fibras puede estar todavía presente aunque el promotor o las regiones promotoras según la invención sean inducidas al lesionar una planta.

El promotor de la presente invención también puede ser específico en fibras. Un promotor "específico en fibras" significa que la transcripción de un ácido nucleico controlado por dicho promotor es más de 20 veces superior, tal como al menos 30 veces superior, al menos 40 veces superior, al menos 50 veces superior o incluso 100 veces superior en una célula de fibra que de media en las células de otros tejidos de planta en el momento del desarrollo de la fibra, tal como de 1 dpa hasta 60 dpa, tal como de 10 a 50, de 20 a 50 dpa, de 10 a 35, de 10 a 30 o de 15 a 30 dpa o de 15 a 35 dpa. El término no tiene en cuenta la expresión inducida por lesiones potencial. En otras palabras, la expresión específica en fibras todavía puede estar presente aunque los promotores descritos en la presente sean inducidos al lesionar una planta.

La confirmación de la actividad promotora para una secuencia promotora o un fragmento promotor funcional puede ser determinada por los expertos en la especialidad, por ejemplo usando una construcción promotor-indicador que comprende la secuencia promotora ligada operativamente a un marcador fácilmente puntuable según se explica adicionalmente en la presente. La capacidad de expresión preferente en fibras de los fragmentos o las variantes identificados o generados del promotor descrito en la presente se puede probar convenientemente al ligar operativamente estas secuencias de ácido nucleico a una secuencia nucleotídica que codifica un marcador fácilmente calificable, p. ej. un gen de betaglucuronidasa, introducir este gen quimérico en una planta y analizar el patrón de expresión del marcador en células de fibras en comparación con el patrón de expresión del marcador en otras partes de la planta. Otros candidatos para un marcador (o un gen indicador) son cloranfenicol acetil transferasa (CAT), betagalactosidasa (beta-GAL) y proteínas con propiedades fluorescentes o fosforescentes, tales como proteína fluorescente verde (GFP) procedente de *Aequora victoria* o luciferasa. Para confirmar la función promotora, una secuencia de ácido nucleico que representa el promotor se liga operativamente a la secuencia codificante de un gen marcador (indicador) mediante técnicas de ADN recombinante muy conocidas en la especialidad. El gen indicador se liga operativamente aguas abajo del promotor, de modo que los transcritos que se inician en el promotor avancen a través del gen indicador. El casete de expresión que contiene el gen indicador bajo el control del promotor se puede introducir en un tipo de célula apropiado mediante técnicas de transformación muy conocidas en la especialidad y descritas en otras partes en esta solicitud. Para ensayar la proteína indicadora, se preparan lisados celulares y se realizan ensayos apropiados, que son muy conocidos en la especialidad, para la proteína indicadora. Por ejemplo, si CAT era el gen indicador de elección, los lisados procedentes de células transfectadas con construcciones que contienen CAT bajo el control de un promotor bajo estudio se mezclan con cloranfenicol marcado isotópicamente y acetil-coenzima A (acetil-CoA). La enzima CAT transfiere el grupo acetilo desde la acetil-CoA a la posición 2 o 3 del cloranfenicol. La reacción se comprueba mediante cromatografía en capa fina, que separa el cloranfenicol acetilado del material sin reaccionar.

A continuación, los productos de reacción se visualizan mediante autorradiografía. El nivel de actividad enzimática corresponde a la cantidad de enzima que se elaboraba, que a su vez revela el nivel de expresión de la funcionalidad preferente en fibras del promotor o el fragmento o la variante del mismo. El nivel de expresión también se puede comparar con otros promotores para determinar la fuerza relativa del promotor bajo estudio. Una vez que se confirma la actividad y la funcionalidad, se pueden emplear análisis de mutación y/o eliminación adicionales para determinar, p. ej., una región promotora mínima y/o secuencias requeridas para iniciar la transcripción. Así, las secuencias se pueden someter a eliminación en el extremo 5' de la región promotora y/o en el extremo 3' de la región promotora, o dentro de la secuencia promotora, y/o se pueden introducir sustituciones de nucleótidos. A

continuación, estas construcciones se introducen de nuevo en células y se determinan su actividad y/o funcionalidad.

En lugar de medir la actividad de una enzima indicadora, la actividad (y la funcionalidad) del promotor de la transcripción también se puede determinar al medir el nivel de ARN que se produce. Este nivel de ARN, tal como ARNm, se puede medir bien en un solo punto temporal o bien en múltiples puntos temporales y como tal el incremento en número de veces puede ser un incremento en número de veces medio o un valor extrapolado derivado de valores medidos experimentalmente. Como es una comparación de niveles, se puede usar cualquier método que mida los niveles de ARNm. En un ejemplo, el tejido o los órganos comparados son una semilla o un tejido de semilla tal como fibras con una hoja o un tejido foliar. En otro ejemplo, se comparan múltiples tejidos u órganos. Un ejemplo para múltiples comparaciones son células de fibras comparadas con 2, 3, 4 o más tejidos u órganos seleccionados del grupo que consiste en tejido floral, ápice floral, polen, hoja, embrión, vástago, primordios foliares, ápice de vástago, raíz, punta de raíz, tejido vascular y cotiledón. Según se usan en la presente, ejemplos de órganos de plantas son semilla, hoja, raíz, etc. y un ejemplo de tejidos son primordios foliares, ápice de vástago, tejido vascular, etc. La actividad o la intensidad de un promotor se puede medir en cuanto a la cantidad de acumulación de ARNm o proteína que produce específicamente, con relación a la cantidad total de ARNm o proteína. El promotor expresa una secuencia de ácido nucleico ligada operativamente, por ejemplo a un nivel mayor de aproximadamente 0,1%, aproximadamente 0,2%, mayor de aproximadamente 0,5, 0,6, 0,7, 0,8 o aproximadamente 0,9%, mayor de aproximadamente 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8% o aproximadamente 9%, o mayor de aproximadamente 10% del ARNm total de la célula en la que está contenido. Alternativamente, la actividad o la intensidad de un promotor se puede expresar con relación a un promotor bien caracterizado para el que se evaluaba previamente la actividad de transcripción.

En otro aspecto, se proporcionan promotores preferentes en fibras que comprenden una secuencia de ácido nucleico que tiene al menos 95%, 96%, 97%, 98% o al menos 99% de identidad de secuencia con SEQ ID NO: 1 o un fragmento de la misma según se define anteriormente (o son idénticos a la misma), tales como, por ejemplo, un ácido nucleico que comprende al menos 700, al menos 800, al menos 900 o al menos 950 nucleótidos consecutivos de la secuencia nucleotídica indicada como SEQ ID NO: 1. Variantes presentes en la naturaleza del promotor divulgado en la presente se pueden identificar con el uso de técnicas de biología molecular bien conocidas, tales como, por ejemplo, con reacción en cadena de la polimerasa (PCR) y técnicas de hibridación como la esbozadas anteriormente en la presente. Estas secuencias de ácido nucleico también incluyen secuencias de ácido nucleico derivadas sintéticamente, tales como las generadas, por ejemplo, al usar mutagénesis dirigida al sitio de SEQ ID NO: 1 o un fragmento de la misma. Generalmente, las variantes de secuencia de ácido nucleico de la invención tendrán al menos 95%, al menos 96%, al menos 97%, al menos 98% o al menos 99% de identidad de secuencia con la secuencia de ácido nucleico de SEQ ID NO: 1 o un fragmento de la misma según se define anteriormente, tal como, por ejemplo, un ácido nucleico que comprende al menos 700, al menos 800, al menos 900 o al menos 950 nucleótidos consecutivos de la secuencia nucleotídica indicada como SEQ ID NO: 1. Derivados de las secuencias de ácido nucleico divulgadas en la presente pueden incluir, pero no se limitan a, eliminaciones de secuencia, mutaciones puntuales simples o múltiples, alteraciones en un sitio de reconocimiento de enzima de restricción particular, adición de elementos funcionales, u otros medios de modificación molecular que pueden mejorar o alterar de otro modo la expresión del promotor. Técnicas para obtener estos derivados son muy conocidas en la especialidad (véase, por ejemplo, J. F. Sambrook, D. W. Russell y N. Irwin (2000) *Molecular Cloning: A Laboratory Manual*). Por ejemplo, un experto en la especialidad puede delimitar los elementos funcionales dentro de los promotores divulgados en la presente y eliminar cualesquiera elementos no esenciales. Los elementos funcionales se pueden modificar o combinar para incrementar la utilidad o la expresión de las secuencias divulgadas en la presente para cualquier aplicación particular. Los expertos en la especialidad están familiarizados con los materiales didácticos estándar que describen condiciones y procedimientos específicos para la construcción, manipulación y aislamiento de macromoléculas (p. ej., moléculas de ADN, plásmidos, etc.), así como la generación de organismos recombinantes y el cribado y el aislamiento de moléculas de ADN.

La secuencia promotora de SEQ ID NO: 1 y sus fragmentos funcionales y variantes se pueden alterar, por ejemplo, para contener "ADN mejorador" para ayudar a elevar la expresión génica. Como se sabe bien en la especialidad, se pueden usar ciertos elementos de ADN para mejorar la transcripción de ADN como parte de o en asociación con un promotor. Estos mejoradores a menudo se encuentran aguas arriba (5') del inicio de la transcripción en un promotor que funciona en células eucarióticas, pero se pueden insertar aguas abajo (3') de la secuencia codificante. En algunos casos, estos elementos de ADN mejorador están dentro o son parte de intrones. Entre los intrones que son útiles como ADN mejorador están los intrones 5' desde el gen de actina 1 de arroz (véase el documento US5641876), el gen de actina 2 de arroz, el gen de alcohol deshidrogenasa de maíz, el gen de proteína de choque térmico 70 de maíz (véase el documento US5593874), el gen Shrunken 1 de maíz, el gen de fotosensibilidad 1 de *Solanum tuberosum* y el gen de la proteína de choque térmico 70 de *Petunia hybrida* (véase el documento US 5659122). Así, según se contempla en la presente, un promotor o una región promotora incluye variaciones de promotores derivados al insertar o eliminar regiones reguladoras, someter el promotor a mutagénesis aleatoria o dirigida al sitio, etc. La actividad o la intensidad de un promotor se puede medir en cuanto a las cantidades de ARN que produce, o la cantidad de acumulación de proteína en una célula o un tejido, con relación a un promotor cuya actividad de transcripción se ha evaluado previamente, según se describe anteriormente.

Según se usa en la presente, el término "porcentaje de identidad de secuencia" se refiere al porcentaje de nucleótidos idénticos entre dos segmentos de un intervalo de ADN óptimamente alineado. El alineamiento óptimo de secuencias para alinear un intervalo de comparación es muy conocido para los expertos en la especialidad y se puede efectuar mediante herramientas tales como el algoritmo de homología local de Smith y Waterman (Waterman, M. S., Chapman & Hall. Londres, 1995), el algoritmo de alineamiento de homología de Needleman y Wunsch (1970), la búsqueda del método de similitud de Pearson y Lipman (1988) y preferiblemente mediante ejecuciones informatizadas de estos algoritmos tales como GAP, BESTFIT, FASTA y TFASTA disponibles como parte de GCG (marca comercial registrada), Wisconsin Package (marca comercial registrada de Accelrys Inc., San Diego, Calif.). Una "fracción de identidad" para segmentos alineados de una secuencia de prueba y una secuencia de referencia es el número de componentes idénticos que son compartidos por las dos secuencias alineadas dividido por el número total de componentes del segmento de la secuencia de referencia, es decir, toda la secuencia de referencia o una parte definida menor de la secuencia de referencia. El porcentaje de identidad de secuencia se representa como la fracción de identidad por 100. La comparación de una o más secuencias de ADN puede ser con una secuencia de ADN de longitud completa o una porción de la misma, o con una secuencia de ADN más larga.

El término "hibridación" se refiere a la capacidad de una primera hebra de ácido nucleico para conectarse con una segunda hebra a través de apareamiento de bases por enlace de hidrógeno cuando las dos hebras de ácido nucleico tienen suficiente identidad de secuencia. La hibridación se produce cuando las dos moléculas de ácido nucleico se renaturalizan entre sí bajo condiciones apropiadas. La hibridación de ácidos nucleicos es una técnica muy conocida por los expertos en la especialidad de la manipulación de ADN. La propiedad de hibridación de un par de ácidos nucleicos dado es una indicación de su similitud o identidad. Otra indicación de que dos secuencias de ácido nucleico son en gran parte idénticas es que las dos moléculas se hibridan entre sí bajo condiciones rigurosas. Las "condiciones de hibridación rigurosas" y "condiciones de lavado de hibridación rigurosas" en el contexto de los experimentos de hibridación de ácidos nucleicos hibridación Southern y Northern dependen de la secuencia, y son diferentes bajo diferentes parámetros ambientales. Un ejemplo de condiciones de lavado muy rigurosas es 0,1 X SSC, 5X solución de Denhardt, 0,5% de SDS a 65°C durante, p. ej., aproximadamente 15 minutos. Un ejemplo de condiciones de lavado apropiadas es un lavado con 2 X SSC, 0,1% de SDS a 65°C durante, p. ej., aproximadamente 15 minutos. A menudo, un lavado muy riguroso está precedido por un lavado poco riguroso para retirar la señal de sonda de fondo. Las condiciones rigurosas también se pueden conseguir con la adición de agentes desestabilizadores tales como formamida. En general, una relación de señal a ruido de 2 X (o superior) la observada para una sonda no relacionada en el ensayo de hibridación particular indica la detección de una hibridación específica. Las condiciones muy rigurosas se seleccionan para ser iguales a la T<sub>m</sub> para una sonda particular.

En el transcurso de la presente invención, un promotor preferente en fibras, junto con una UTR 5' posterior (juntos también denominados "región promotora"), se ha identificado en *Gossypium hirsutum*. Dicho promotor controla un gen tipo Fb8 en algodón.

El promotor es altamente expresado en fibras. Se ha mostrado que la expresión es particularmente intensa durante la fase de engrosamiento celular del desarrollo de las fibras; la expresión del producto génico se detectó de 20 a 35 dpa.

Se espera que el presente promotor sea adecuado para la expresión preferente en fibras al menos durante las fases tardías de fibras en desarrollo de transgenes en algodón. El presente promotor se puede usar para expresar genes que modifican las propiedades de fibras de algodón, especialmente las propiedades de la pared celular de las fibras de algodón.

A menos que se indique otra cosa, las definiciones específicas o las características específicas de ciertos ejemplos divulgados en la presente solicitud en relación con un aspecto se pueden introducir en cualquier otro aspecto divulgado en la presente.

#### Genes quiméricos

En otro aspecto, la presente solicitud divulga un gen quimérico que comprende los ácidos nucleicos aislados descritos en la presente, es decir el promotor y las regiones promotoras y los fragmentos funcionales, ligados operativamente a una secuencia de ácido nucleico que codifica un producto de expresión de interés, y una secuencia de terminación de la transcripción y poliadenilación funcional en células vegetales.

Un gen quimérico es un gen artificial construido al ligar operativamente fragmentos de genes no relacionados u otras secuencias de ácido nucleico. En otras palabras, "gen quimérico" indica un gen que no se encuentra normalmente en una especie vegetal o se refiere a cualquier gen en el que el promotor, las partes colindantes del promotor o una o más de otras regiones reguladoras del gen no están asociados en la naturaleza con una parte o la totalidad del ácido nucleico transcrito, es decir, son heterólogos con respecto al ácido nucleico transcrito. Más particularmente, un gen quimérico es un gen artificial, es decir no presente en la naturaleza, producido por una ligazón operativa de la secuencia de ácido nucleico de la invención, tal como, p. ej., el ácido nucleico de SEQ ID NO: 1, un fragmento de la misma o una secuencia de ácido nucleico que tiene al menos 95% de identidad de secuencia con la misma, todas capaces de dirigir la expresión preferente en fibras de un producto de expresión de interés según se describe

anteriormente, con una segunda secuencia de ácido nucleico que codifica dicho producto de expresión de interés que naturalmente no está ligado operativamente a dicha secuencia de ácido nucleico.

El término "heterólogo" se refiere a la relación entre dos o más secuencias de ácido nucleico o proteína que se derivan de diferentes fuentes. Por ejemplo, un promotor es heterólogo con respecto a una secuencia de ácido nucleico ligada operativamente, tal como una secuencia codificante, si esta combinación no se encuentra normalmente en la naturaleza. Además, una secuencia particular puede ser "heteróloga" con respecto a una célula u organismo en el que se inserta (es decir no está presente naturalmente en esa célula u organismo particular). Por ejemplo, el gen quimérico divulgado en la presente es un ácido nucleico heterólogo.

El término "ligado operativamente" se refiere a la disposición espacial funcional de dos o más regiones de ácido nucleico o secuencias de ácido nucleico. Por ejemplo, una región promotora puede estar situada con relación a una secuencia de ácido nucleico que codifica un producto de expresión de interés de modo que la transcripción de dicha secuencia de ácido nucleico sea dirigida por la región promotora. Así, una región promotora está "ligada operativamente" a la secuencia de ácido nucleico.

El promotor, el fragmento o la variante del mismo que se describe anteriormente puede estar ligado operativamente a una secuencia de ácido nucleico que codifica un producto de expresión de interés que es heterólogo con respecto al promotor. La secuencia de ácido nucleico puede ser generalmente cualquier secuencia de ácido nucleico para la que se desee un nivel alterado tal como un nivel incrementado de transcripción. La secuencia de ácido nucleico puede codificar, por ejemplo, un polipéptido que es capaz de modificar las propiedades de la fibra en el algodón.

Secuencias heterólogas de ácido nucleico adecuadas para modificar las propiedades de fibras de algodón incluyen, sin limitación, las divulgadas en el documento WO02/45485, por las que la calidad de las fibras en plantas productoras de fibras, tales como algodón, se modifica al modular la actividad y/o la expresión de sacarosa sintasa en estas plantas, los ácidos nucleicos útiles para mejorar la resistencia de las fibras o para introducir cargas positivas en la pared celular de la planta según se describe en los documentos WO2006/136351 o WO2011/089021.

Un "producto de expresión" indica un producto intermedio o producto final que surge de la transcripción y opcionalmente la traducción del ácido nucleico, tal como ADN o ARN, que codifica este producto. Durante el proceso de transcripción, una secuencia de ADN bajo el control de regiones reguladoras, particularmente la secuencia promotora divulgada en la presente, se transcribe en una molécula de ARN. Una molécula de ARN puede formar por sí misma un producto de expresión y, entonces, por ejemplo, es capaz de interactuar con otro ácido nucleico u otra proteína. Alternativamente, una molécula de ARN puede ser un producto intermedio cuando es capaz de ser traducida en un péptido o una proteína. Se dice que un gen codifica una molécula de ARN como producto de expresión cuando el ARN como el producto final de la expresión del gen es capaz de interactuar con otro ácido nucleico o proteína. Ejemplos de productos de expresión de ARN incluyen ARNs inhibidores tales como, p. ej., ARN de sentido, ARN antisentido, ARN de horquilla, ribozimas, miARN o ARNsi, ARNm, ARNr y ARNt. Se dice que un gen codifica una proteína o un péptido como producto de expresión cuando el producto final de la expresión del gen es una proteína o un péptido.

Productos de expresión de interés ejemplares incluyen proteínas implicadas en la síntesis de la pared celular y la formación de fibras como las divulgadas en el documento WO2005/017157, en particular un gen que codifica una proteína de  $\beta$ -1,3 glucano sintasa, o en el documento WO2006/136351, en particular una N-acetilglucosamina transferasa que puede ser dirigida a las membranas del aparato de Golgi, tal como una N-acetilglucosamina transferasa del tipo NodC, o una quitina sintasa y las descritas anteriormente en relación con ácidos nucleicos heterólogos adecuados.

Dentro del alcance de la presente divulgación, también se puede hacer uso, en combinación con el gen quimérico descrito anteriormente, de otras secuencias reguladoras, que están situadas entre dicha secuencia de ácido nucleico que comprende un promotor y dicha secuencia de ácido nucleico que comprende la secuencia codificante del producto de expresión. Ejemplos no limitativos de estas secuencias reguladoras incluyen activadores ("mejoradores") de la traducción, por ejemplo el activador de la traducción del virus del mosaico del tabaco (TMV) descrito en la Solicitud WO 87/07644, o del virus del grabado del tabaco (TEV) descrito por Carrington & Freed 1990, J. Virol. 64: 1590-1597, o intrones tales como el intrón de histona 3 de *Arabidopsis* (Chaubet y cols., 1992).

Otras secuencias reguladoras adecuadas incluyen UTRs 5'. Según se usa en la presente, una UTR 5', también denominada una secuencia líder, es una región particular de un ARN mensajero (ARNm) situada entre el sitio de inicio de la transcripción y el codón de inicio de la región codificante. Está implicada en la estabilidad del ARNm y la eficacia de la traducción. Por ejemplo, el líder no traducido 5' de un gen de proteína de unión a/b de clorofila de petunia (cab22L) aguas abajo del sitio de inicio de la transcripción 35S se puede utilizar para aumentar los niveles en estado estacionario de expresión del gen indicador (Harpster y cols., 1988, Mol Gen Genet. 212(1):182-90). El documento WO95/006742 describe el uso de secuencias líder no traducidas 5' derivadas de genes que codifican proteínas de choque térmico para incrementar la expresión del transgén.

Estas UTRs son particularmente útiles si la secuencia de ácido nucleico que constituye el promotor de la presente invención no comprende los primeros 44 o 48 nucleótidos 5' del inicio de la traducción (es decir, los nucleótidos 926 a 959 de SEQ ID NO: 1 o los nucleótidos 922 a 959 de SEQ ID No 1). En otras palabras, estas UTRs se pueden usar ventajosamente si el promotor de la presente invención no alcanza más allá del nucleótido 922 o el nucleótido 926 de SEQ ID NO: 1. En estos casos, la región promotora resultante todavía es capaz de dirigir la expresión preferente en fibras o específica en fibras del gen ligado operativamente a la misma en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa.

El gen quimérico también comprende una secuencia de terminación de la transcripción o de poliadenilación operativa en una célula de planta, particularmente una célula de planta de algodón. Como una secuencia de terminación de la transcripción o de poliadenilación, se puede hacer uso de cualquier secuencia correspondiente de origen bacteriano, tal como, por ejemplo, el terminador nos de *Agrobacterium tumefaciens*, de origen viral, tal como, por ejemplo, el terminador 35S de CaMV, o de origen vegetal, tal como, por ejemplo, un terminador de histona como el descrito en la Solicitud de Patente publicada EP 0 633 317 A1.

En un ejemplo del gen quimérico descrito en la presente, dicho producto de expresión de interés es una proteína, un péptido o una molécula de ARN, dicha molécula de ARN capaz de modular la expresión de un gen endógeno a dicha planta. En un ejemplo, dicha proteína, péptido o molécula de ARN es capaz de modular una propiedad de las fibras. En otro ejemplo, dicha proteína, péptido o molécula de ARN esta implicado en la biosíntesis de auxina.

El término "proteína", según se usa en la presente, describe un grupo de moléculas que consisten 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 u oligómeros superiores, es decir que consisten en más de una molécula de (poli)péptido. Las moléculas de proteína o péptido que forman estos dímeros, trímeros, etc. pueden ser idénticas o no idénticas. Las estructuras de orden superior correspondientes se denominan, por consiguiente, homo- o heterodímeros, homo- o heterotrímeros, etc. Los términos "proteína" y "péptido" también se refieren a proteínas o péptidos modificados naturalmente en los que la modificación se efectúa, p. ej., mediante glicosilación, acetilación, fosforilación y similares. Estas modificaciones son muy conocidas en la especialidad.

Proteínas ejemplares adecuadas como productos de expresión incluyen proteínas implicadas en la modificación de propiedades de las fibras, tales como las que median en el incremento en la longitud de las fibras, una alteración en la resistencia de las fibras o una alteración en las propiedades de la pared celular que dé como resultado, p. ej., una carga alterada de dichas paredes celulares, una alteración en la capacidad de tinción, una disminución en la producción de fibras apelmazadas, una alteración en la relación de madurez de las fibras, una disminución en el contenido de fibras inmaduras o un incremento en la uniformidad de las fibras y el micronaire.

Dicho producto de expresión de interés también puede ser una molécula de ARN capaz de modular la expresión de un gen endógeno a dicha planta de algodón.

Ejemplos de genes diana adecuados para productos de expresión de ARN en relación con esto incluyen los implicados en la modificación de las propiedades de las fibras, tales como los que median en un incremento en la longitud de las fibras, una alteración en la resistencia de las fibras o una alteración en las propiedades de la pared celular que dé como resultado, p. ej., una carga alterada de dichas paredes celulares, una alteración en la capacidad de tinción, una disminución en la producción de fibras apelmazadas, una alteración en la relación de madurez de las fibras, una disminución en el contenido de fibras inmaduras o un incremento en la uniformidad de las fibras y el micronaire.

Para el caso de las moléculas de ARN, estará claro que si las secuencias de ácido nucleico de moléculas de DNA se definen mediante referencia a secuencias de ácido nucleico de moléculas de ARN correspondientes, la timina (T) en la secuencia de ácido nucleico se debe reemplazar por uracilo (U). Si se hace referencia a moléculas de ARN o ADN, estará claro a partir del contexto de la solicitud.

El término "capaz de modular la expresión de un gen" se refiere a la acción de una molécula de ARN, tal como una molécula de ARN inhibidora según se describe en la presente, para influir de diferentes modos en el nivel de expresión de genes diana. Esto se puede efectuar, p. ej., al inhibir la expresión de un gen diana al interactuar directamente con componentes que conducen la expresión tales como el propio gen o el ARNm transcrito que da como resultado una disminución de la expresión, o al inhibir otro gen implicado en la activación de la expresión de un gen diana aboliendo de ese modo dicha activación, o inhibir otro gen implicado en la inhibición de la expresión de un gen diana que da como resultado un incremento de expresión. La inhibición de un gen implicado en la inhibición de la expresión de un gen diana usando ARN inhibidor puede, por el contrario, dar como resultado una activación de la expresión de dicho gen diana.

Los moléculas de ARN inhibidoras disminuyen los niveles de ARNms de sus proteínas diana disponibles para la traducción en dicha proteína diana. De este modo, se puede inhibir la expresión de proteínas implicadas en respuestas no deseadas a condiciones de estrés. Esto se puede conseguir a través de técnicas bien establecidas

incluyendo cosupresión (supresión de ARN de sentido), ARN antisentido, ARN bicatenario (ARNds), ARNsi o microARN (miARN).

Una molécula de ARN como producto de expresión según se divulga en la presente comprende una parte de una secuencia de ácido nucleico que codifica una proteína diana o una secuencia homóloga para regular a la baja la expresión de dicha proteína diana. Otro ejemplo para una molécula de ARN como producto de expresión para el uso en la regulación a la baja de la expresión son moléculas de ARN antisentido que comprenden una secuencia de ácido nucleico complementaria a al menos una parte de una secuencia de ácido nucleico que codifica una proteína de interés o una secuencia homóloga. Aquí, la regulación a la baja se puede efectuar, p. ej., al introducir este ARN antisentido o un ADN quimérico que codifica esta molécula de ARN. En otro ejemplo más, la expresión de una proteína de interés se regula a la baja al introducir una molécula de ARN bicatenario que comprende una región de ARN de sentido y una de antisentido correspondientes a y respectivamente complementarias a al menos parte de una secuencia génica que codifica dicha proteína de interés, región de ARN de sentido y antisentido que son capaces de formar una región de ARN bicatenario entre sí. Esta molécula de ARN bicatenario puede ser codificada por moléculas tanto de sentido como antisentido según se describe anteriormente y por una molécula monocatenaria que es procesada para formar ARNsi o miARN.

En un ejemplo, la expresión de una proteína diana se puede regular a la baja al introducir una construcción de ADN quimérica que da una molécula de ARN de sentido capaz de regular a la baja la expresión mediante cosupresión. La región de ADN transcrita dará durante la transcripción la llamada molécula de ARN de sentido capaz de reducir la expresión de un gen que codifica una proteína diana en la planta diana o una célula de planta de un modo transcripcional o postranscripcional. La región de ADN transcrita (y la molécula de ARN resultante) comprende al menos 20 nucleótidos consecutivos que tienen al menos 95% de identidad de secuencia con la porción correspondiente de la secuencia de ácido nucleico que codifica la proteína diana presente en la célula de planta o la planta.

Alternativamente, un producto de expresión para la regulación a la baja de la expresión de una proteína diana es una molécula de ARN antisentido. La regulación a la baja o la reducción de la expresión de una proteína de interés en la planta o la célula de planta de algodón diana se efectúa de nuevo de un modo transcripcional o postranscripcional. La región de ADN transcrita (y la molécula de ARN resultante) comprende al menos 20 nucleótidos consecutivos que tienen al menos 95% de identidad de secuencia con el complemento de la porción correspondiente de la secuencia de ácido nucleico que codifica dicha proteína diana presente en la célula de planta o la planta.

Sin embargo, la secuencia de ácido nucleico mínima de la región de ARN antisentido o de sentido de aproximadamente 20 nt (nucleótidos) de la secuencia de ácido nucleico que codifica una proteína diana puede estar comprendida dentro de una molécula de ARN mayor, que varía en tamaño de 20 nt a una longitud igual al tamaño del gen diana. Las regiones de ácido nucleico de antisentido o sentido mencionadas pueden así tener una longitud de aproximadamente 21 nt a aproximadamente 5000 nt, tal como 21 nt, 40 nt, 50 nt, 100 nt, 200 nt, 300 nt, 500 nt, 1000 nt, 2000 nt o incluso aproximadamente 5000 nt o una longitud mayor. Por otra parte, no se requiere para el propósito de la invención que la secuencia de ácido nucleico de la molécula de ARN inhibidora usada o la región codificante del transgén sea completamente idéntica o complementaria al gen endógeno que codifica la proteína diana cuya expresión está destinada a ser reducida en la célula de planta. Cuanto más larga sea la secuencia, menos riguroso será el requisito para la identidad de secuencia global. Así, las regiones de sentido o antisentido pueden tener una identidad de secuencia global de aproximadamente 40% o 50% o 60% o 70% u 80% o 90% o 100% con la secuencia de ácido nucleico de un gen endógeno o el complemento del mismo. Sin embargo, según se menciona, las regiones antisentido o de sentido deben comprender una secuencia de ácido nucleico de 20 nucleótidos consecutivos que tiene de aproximadamente 95 a aproximadamente 100% de identidad de secuencia con la secuencia de ácido nucleico del gen endógeno que codifica el gen diana. El tramo de aproximadamente 95 a aproximadamente 100% de identidad de secuencia puede ser de aproximadamente 50, 75 o 100 nt.

La eficacia de los susodichos genes quiméricos para la regulación a la baja del nivel de expresión génica mediada por ARN antisentido o ARN de sentido se puede mejorar adicionalmente mediante la inclusión de elementos de ADN que dan como resultado la expresión de moléculas de ARN inhibidoras no poliadeniladas aberrantes. Uno de estos elementos de ADN adecuado para ese propósito es una región de ADN que codifica una ribozima de autoempalme. La eficacia también se puede mejorar al proveer a las moléculas de ARNs generadas de señales de localización o retención nuclear.

Además, un producto de expresión como el descrito en la presente puede ser una secuencia de ácido nucleico que da una molécula de ARN bicatenario capaz de regular a la baja la expresión de un gen que codifica una proteína diana. Durante la transcripción de la región de ADN, el ARN es capaz de formar una molécula de ARNds a través de apareamiento de bases convencional entre una región de sentido y antisentido, en donde la región de sentido y antisentido son secuencias de ácido nucleico como las descritas anteriormente en la presente. Los productos de expresión que son ARNds según la invención pueden comprender además un intrón, tal como un intrón heterólogo, situado, p. ej., en el espaciador entre las regiones de ARN de sentido y antisentido según la divulgación del

documento WO 99/53050. Para conseguir la construcción de este transgén, se puede hacer uso de los vectores descritos en el documento WO 02/059294 A1.

En un ejemplo, dicha molécula de ARN comprende una región de ARN primera y segunda en la que 1. dicha primera región de ARN comprende una secuencia de ácido nucleico de al menos 19 nucleótidos consecutivos que tiene al menos aproximadamente 94% de identidad de secuencia con la secuencia de ácido nucleico de dicho gen endógeno; 2. dicha segunda región de ARN comprende una secuencia de ácido nucleico complementaria a dichos 19 nucleótidos consecutivos de dicha primera región de ARN; 3. dichas primera y segunda región de ARN son capaces de apareamiento de bases para formar una molécula de ARN bicatenaria entre al menos dichos 19 nucleótidos consecutivos de dichas primera y segunda región. En otros ejemplos, se aplican las mismas consideraciones que se describen anteriormente para ARN de sentido y antisentido.

Otro producto de expresión ejemplar es una molécula de microARN (miARN, que se puede procesar a partir de una molécula de pre-microARN) capaz de guiar la escisión de ARNm transcrito a partir del ADN que codifica la proteína diana que se va a traducir en dicha proteína diana. Las moléculas de miARN se pueden introducir convenientemente en células de planta a través de la expresión a partir de un gen quimérico como el descrito en la presente que comprende una (segunda) secuencia de ácido nucleico que codifica como producto de expresión de interés este transcrito de miARN, pre-miARN o miARN primario.

Los miARNs son ARNs endógenos pequeños que regulan la expresión génica en plantas, pero también en otros eucariotas. Según se usa en la presente, un "miARN" es una molécula de ARN de aproximadamente 20 a 30 nucleótidos (Siomi y Siomi, 2009) de longitud que se puede cargar en un complejo de RISC y dirigir la escisión de una molécula de ARN diana, en donde la molécula de ARN diana comprende una secuencia de ácido nucleico esencialmente complementaria a la secuencia de ácido nucleico de la molécula de miARN. En miARNs ejemplares, se pueden producir uno o más de los siguientes desajustes en el miARN esencialmente complementario al ARN diana:

- Un desajuste entre el nucleótido en el extremo 5' de dicho miARN y la secuencia de ácido nucleico correspondiente en la molécula de ARN diana;

- Un desajuste entre uno cualquiera de los nucleótidos en la posición 1 a la posición 9 de dicho miARN y la correspondiente secuencia de ácido nucleico en la molécula de ARN diana;

- Tres desajustes entre uno cualquiera de los nucleótidos en la posición 12 a la posición 21 de dicho miARN y la correspondiente secuencia de ácido nucleico en la molécula de ARN diana con la condición de que no haya más de dos desajustes consecutivos;

- No está permitido desajuste en las posiciones 10 y 11 del miARN (todas las posiciones del miARN están indicadas partiendo del extremo 5' de la molécula de miARN).

Según se usa en la presente, una molécula de "pre-miARN" es una molécula de ARN de aproximadamente 100 a aproximadamente 200 nucleótidos, preferiblemente de aproximadamente 100 a aproximadamente 130 nucleótidos que pueden adoptar una estructura secundaria que comprende un tallo de ARNs y un bucle de ARN monocatenario y que comprende además la secuencia de ácido nucleico del miARN y su secuencia complementaria del miARN\* en el tallo de ARN bicatenario. Preferiblemente, el miARN y su complemento están situados de aproximadamente 10 a aproximadamente 20 nucleótidos desde los extremos libres del tallo de miARN ARNs. La longitud y la secuencia de la región de bucle monocatenario no son críticas y pueden variar considerablemente, p. ej. entre 30 y 50 nt de longitud. Preferiblemente, la diferencia en energía libre entre la estructura de ARN desapareada y apareada está entre -20 y -60 kcal/mol, por ejemplo alrededor de -40 kcal/mol. La complementariedad entre el miARN y el miARN\* no necesita ser perfecta y se pueden tolerar de aproximadamente 1 a 3 protuberancias de nucleótidos desapareados. La estructura secundaria adoptada por una molécula de ARN se puede predecir mediante algoritmos informáticos convencionales en la especialidad tales como mFold, UNAFold y RNAFold. La hebra particular del tallo de ARNs procedente del pre-miARN que se libera por la actividad de DCL y se carga sobre el complejo de RISC se determina mediante el grado de complementariedad en el extremo 5', con lo que la hebra que en su extremo 5' es la menor implicada en la unión por enlace de hidrógeno entre los nucleótidos de las diferentes hebras del tallo de ARNs escindido se carga en el complejo de RISC y determinará la especificidad de secuencia de la degradación de la molécula de ARN diana. Sin embargo, si empíricamente la molécula de miARN procedente de una molécula de pre-miARN sintético particular no es funcional debido a que la hebra "errónea" está cargada en el complejo de RISC, será inmediatamente evidente que este problema se puede resolver al intercambiar la posición de la molécula de miARN y su complemento en las hebras respectivas del tallo de ARNs de la molécula de pre-miARN. Como se sabe en la especialidad, la unión entre A y U que implica dos enlaces de hidrógeno o G y U que implica dos enlaces de hidrógeno es menos fuerte que entre G y C que implica tres enlaces de hidrógeno.



Las moléculas de miARN pueden estar comprendidas dentro de sus moléculas de pre-miARN presentes en la naturaleza, pero también se pueden introducir en andamiajes de la molécula de pre-miARN existentes al intercambiar la secuencia de ácido nucleico de la molécula de miARN normalmente procesada a partir de esta molécula de pre-miARN existente para la secuencia de ácido nucleico de otro miARN de interés. El andamiaje del pre-miARN también puede ser completamente sintético. Asimismo, las moléculas de miARN sintéticas pueden estar comprendidas dentro de, y ser procesadas a partir de, andamiajes de molécula de pre-miARN existentes o andamiajes de pre-miARN sintéticos.

Los productos de expresión ejemplares también pueden ser ribozimas que catalizan bien su propia escisión o bien la escisión de otros ARNs.

En un ejemplo del gen quimérico divulgado en la presente, modular la expresión es incrementar la expresión y dicha secuencia de ácido nucleico que codifica un producto de expresión de interés codifica un ARN, que cuando se transcribe da una molécula de ARN capaz de incrementar la expresión de un gen endógeno a dicha planta de algodón. Estos genes se podrían correlacionar positivamente con la longitud de las fibras, la resistencia de las fibras o una alteración deseada en las propiedades de la pared celular, que dé como resultado, p. ej., una carga alterada de dichas paredes celulares, una alteración en la capacidad de tinción, una disminución en la producción de fibras apelmazadas, una alteración en la relación de madurez de las fibras, una disminución en el contenido de fibras inmaduras o un incremento en la uniformidad de las fibras y el micronaire o 2. da una molécula de ARN capaz de disminuir la expresión de un gen endógeno a dicha planta de algodón, en donde dicho gen puede estar correlacionado negativamente con la longitud de las fibras, la resistencia de las fibras o una alteración deseada en las propiedades de la pared celular que da como resultado, p. ej., una carga alterada de dichas paredes celulares, una alteración en la capacidad de tinción, una disminución en la producción de fibras apelmazadas, una alteración en la relación de madurez de las fibras, una disminución en el contenido de fibras inmaduras o un incremento en la uniformidad de las fibras y el micronaire

En otro ejemplo del gen quimérico descrito en la presente, dicho producto de expresión es un gen indicador o un gen específico para fibras según se describe en otras partes en esta solicitud.

En un ejemplo del gen quimérico descrito en la presente, dicha molécula de ARN comprende una primera y una segunda región de ARN en las que 1. dicha primera región de ARN comprende una secuencia de ácido nucleico de al menos 19 nucleótidos consecutivos que tiene al menos aproximadamente 94% de identidad de secuencia con la secuencia de ácido nucleico de dicho gen endógeno; 2. dicha segunda región de ARN comprende una secuencia de ácido nucleico complementaria a dichos 19 nucleótidos consecutivos de dicha primera región de ARN; y 3. dichas primera y segunda región de ARN son capaces de apareamiento de bases para formar una molécula de ARN bicatenaria entre al menos dichos 19 nucleótidos consecutivos de dichas primera y segunda región.

#### Células de planta y plantas

La presente invención también se dirige a células de planta de algodón transgénicas y plantas de algodón transgénicas que comprenden una secuencia de ácido nucleico como la descrita anteriormente, es decir la secuencia promotora divulgada en la presente, ligada operativamente a una secuencia de ácido nucleico heteróloga que codifica un producto de expresión de interés. Alternativamente, dichas células de planta o plantas transgénicas comprenden en gen quimérico divulgado en la presente. Se describen anteriormente secuencias promotoras y productos de expresión de interés ejemplares u otros elementos reguladores.

Una célula de planta de algodón o una planta de algodón transgénica se puede producir al introducir la secuencia o las secuencias de ácido nucleico que se describen anteriormente en plantas de algodón o células de planta de algodón. "Introducir" en relación con la presente solicitud se refiere a la colocación de información genética en una célula de planta o una planta por medios artificiales. Esto se puede efectuar mediante cualquier método conocido en la especialidad para introducir ARN o ADN en células de planta, protoplastos, callos, raíces, tubérculos, semillas, tallos, hojas, plántulas, embriones, polen y microesporas, otros tejido de planta, o plantas enteras. Más particularmente, "introducir" significa integrar establemente en el genoma de la planta.

Las plantas que contienen al menos una secuencia de ácido nucleico transformada se denominan "plantas transgénicas". Transgénico y recombinante se refieren a un organismo hospedador tal como una planta en la que se ha introducido una molécula de ácido nucleico heteróloga (p. ej., la secuencia de ácido nucleico, el gen quimérico según se describen en la presente). El ácido nucleico se puede integrar establemente en el genoma de la planta. Métodos específicos para la introducción se describen en relación con los métodos divulgados en la presente.

"Algodón" o "planta de algodón", según se usa en la presente, puede ser cualquier especie del género *Gossypium* útil para cultivar y cosechar fibras de algodón. Las especies de algodón más comúnmente usadas son *Gossypium barbadense*, *G. hirsutum*, *G. arboreum* y *G. herbaceum*. Especies adicionales incluyen *G. africanum* y *G. raimondii*. También se incluyen la progenie de cruces de cualquiera de las especies anteriores con otras especies o cruces entre estas especies.

Una célula de planta de algodón puede ser cualquier célula que comprenda esencialmente la información genética necesaria para definir una planta de algodón, que puede, aparte del gen quimérico divulgado en la presente, estar complementada mediante uno o más transgenes adicionales. Las células se pueden derivar de los diversos órganos y/o tejidos que forman una planta de algodón, incluyendo, pero no limitados a, frutos, semillas, embriones, tejido reproductor, regiones meristemáticas, tejido calloso, hojas, raíces, vástagos, flores, tejido vascular, gametofitos, esporofitos, polen y microesporas.

La presente solicitud también divulga una planta transgénica que consiste en las células de planta de algodón transgénica descritas anteriormente en la presente, o que comprende el gen quimérico descrito en la presente establemente integrado en el genoma de la planta. Esto se puede efectuar mediante protocolos de transformación descritos en cualquier parte de esta solicitud.

En otra realización, la presente invención se refiere a una semilla generada a partir de una planta transgénica descrita en la presente, en donde dicha semilla comprende el gen quimérico descrito en la presente.

La semilla es formada por una planta embrionaria encerrada con nutrientes almacenados por una envuelta seminal. Es el producto del óvulo madurado de plantas gimnospermas y angiospermas, a las últimas de las cuales pertenece el algodón, que se produce después de la fertilización y hasta un cierto punto el crecimiento dentro de la planta madre.

También se divulgan en la presente fibras de algodón y aceite de semillas de algodón obtenibles de las plantas divulgadas en la presente. Las fibras de algodón divulgadas en la presente se pueden distinguir de otras fibras al aplicar el método de detección divulgado en el documento WO2010/015423 y verificar la presencia del ácido nucleico de (a) o un gen quimérico de (b) en las fibras. Según esto, el ácido nucleico de (a) también se puede usar para rastrear paredes celulares, en particular fibras de algodón según la invención.

#### Métodos y usos

Por otra parte, la presente invención se refiere a un método para producir una planta de algodón transgénica (a) proporcionando un gen quimérico descrito en la presente; y (b) introducir dicho gen quimérico en una planta.

Está disponible un número de métodos para introducir ADN en células de planta o plantas, bien mediante transformación o bien mediante introgresión. La transformación de algodón mediada por *Agrobacterium* se ha descrito, p. ej., en la patente de EE. UU. 5.004.863, en la patente de EE. UU. 6.483.013 y el documento WO 2000/71733.

Las plantas también se pueden transformar mediante bombardeo de partículas: Partículas de oro o volframio se revisten con ADN y a continuación se disparan a células de planta jóvenes o embriones de planta. Este método también permite la transformación de plástidos de planta. La transformación de algodón mediante bombardeo de partículas se presenta, p. ej., en el documento WO 92/15675.

La transformación (transducción) viral se puede usar para la expresión transitoria o estable de un gen, dependiendo de la naturaleza del genoma viral. El material genético deseado se empaqueta en un virus de planta adecuado y se deja que el virus modificado infecte la planta. La progenie de las plantas infectadas está libre de virus y también está libre del gen insertado. Métodos adecuados para la transformación viral se describen o se detallan adicionalmente, p. ej., en los documentos WO 90/12107, WO 03/052108 o WO 2005/098004.

"Introgresión" significa la integración de un gen en un genoma de planta por medios naturales, es decir al cruzar una planta que comprende el gen quimérico descrito en la presente con una planta que no comprende dicho gen quimérico. Los descendientes se puede seleccionar con respecto a los que comprenden el gen quimérico.

Protocolos de transformación e introgresión adicionales también se pueden encontrar en la patente de EE. UU. 7.172.881.

El gen quimérico se puede introducir, p. ej., mediante transformación, en plantas de algodón a partir de las cuales se puede derivar el callo embriogénico, tales como Coker 312, Coker310, 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, non Acala "picker" Siokra, la 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 los mismos.

5 En un aspecto adicional, la presente solicitud divulga un método para cultivar algodón que comprende (a1) proporcionar la planta transgénica descrita en la presente o producida mediante el método descrito en la presente; o (a2) introducir un gen quimérico descrito en la presente en una planta; (b) cultivar la planta de (a1) o (a2); y (c) recoger el algodón producido por dicha planta.

10 "Cultivar" se refiere a crear el ambiente para que las plantas crezcan, se multipliquen y/o envejezcan. Condiciones de cultivo adecuadas para plantas específicas son muy conocidas en la especialidad.

En otro aspecto, la presente solicitud divulga un método para producir una semilla que comprende el gen quimérico divulgado en la presente que comprende (a) cultivar una planta transgénica que comprende el gen quimérico descrito en la presente o el vector descrito en la presente, una planta transgénica descrita en la presente o una planta transgénica obtenida mediante el método descrito en la presente, en donde dicha planta transgénica produce dicha semilla y dicho gen quimérico está comprendido en dicha semilla, y (b) aislar dicha semilla de dicha planta transgénica.

20 En un ejemplo del método para producir una planta transgénica o el método para producir una semilla, la planta es una planta de algodón como la descrita en cualquier parte en esta solicitud.

En otro aspecto, la presente solicitud divulga un método para efectuar una expresión preferente en fibras de un producto en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa, que comprende introducir el gen quimérico divulgado en la presente en el genoma de una planta de algodón; o proporcionar la planta de algodón transgénica divulgada en la presente.

En un aspecto adicional, la presente solicitud divulga un método para alterar las propiedades de la fibras en una planta de algodón, que comprende introducir el gen quimérico divulgado en la presente en el genoma de una planta de algodón; o proporcionar la planta de algodón transgénica divulgada en la presente.

En un ejemplo, el método comprende además cultivar dicha planta hasta que se generan semillas.

En otro ejemplo basado en la etapa adicional anterior, el método es para incrementar el rendimiento de algodón de una planta de algodón y comprende además recoger el algodón producido por dicha planta de algodón. En otras palabras, se divulga en la presente un método para incrementar el rendimiento de algodón de una planta de algodón, que comprende introducir el gen quimérico divulgado en la presente en el genoma de una planta de algodón; o proporcionar la planta transgénica divulgada en la presente; cultivar dicha planta hasta que se generan semillas; y recoger el algodón producido por dicha planta de algodón.

El término "incrementar el rendimiento" en relación con la presente solicitud se refiere a un incremento en la producción de fibras de algodón que se puede conseguir, p. ej., al incrementar el número de fibras producidas en una semilla de algodón, la longitud de las fibras o la resistencia de las fibras. Se han descrito anteriormente genes y productos de expresión de los mismos implicados en conferir estas propiedades.

Características ejemplares para una modificación de las propiedades de las fibras son la longitud de las fibras, la resistencia de las fibras, una alteración en las propiedades de la pared celular que dé como resultado, p. ej., una carga alterada de las paredes celulares, la capacidad de tinción, una disminución en la producción o el contenido de fibras apelmazadas, la relación de madurez de las fibras, el contenido de fibras inmaduras, la uniformidad de las fibras y el micronaire.

En otro aspecto, la presente solicitud divulga el uso del gen quimérico divulgado en la presente o la planta o la célula de planta de algodón transgénica divulgada en la presente para la expresión preferente en fibras de un producto en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa, para alterar las propiedades de las fibras en algodón o para incrementar el rendimiento de algodón. Las definiciones y los ejemplos adicionales descritos anteriormente para otros aspectos divulgados en la presente se aplican igualmente al presente aspecto.

Las células de planta o las plantas transformadas descritas en la presente tales como las obtenidas mediante los métodos descritos en la presente se pueden usar además en procedimientos de reproducción muy conocidos en la especialidad, tales como cruzamiento, autofecundación y retrocruzamiento. Los programas de reproducción pueden implicar el cruzamiento para generar una generación F1 (primer filial), seguido por varias generaciones de autofecundación (que generan F2, F3, etc.). El programa de reproducción también puede implicar etapas de retrocruzamiento (BC), por las que la descendencia se retrocruza con una de las líneas parentales, denominada el progenitor recurrente.

Según esto, también se divulga en la presente un método para producir plantas que comprenden el gen quimérico divulgado en la presente, que comprende la etapa de cruzar la planta de algodón divulgada en la presente con otra planta o consigo misma y seleccionar la descendencia que comprende dicho gen quimérico.

5 Las células de planta o las plantas transgénicas obtenidas mediante los métodos divulgados en la presente también se pueden usar adicionalmente en procedimientos de transformación posteriores, p. ej. para introducir un gen quimérico adicional.

10 Las células de planta y las plantas transgénicas que comprenden un gen recombinante según la invención se pueden usar para producir alimentos, piensos (tales como aceite, harina, grano, almidón, harina fina o proteína) o un producto industrial tal como biocombustible, fibra, productos químicos industriales, un producto farmacéutico o un producto nutracéutico, particularmente para la producción de fibras de algodón.

15 Las plantas o las semillas de algodón que comprenden el gen quimérico divulgado en la presente o obtenidas mediante los métodos divulgados en la presente se pueden tratar adicionalmente con herbicidas para algodón tales como diurona, fluometurona, MSMA, oxifluorfen, prometrina, trifluralina, carfentrazona, cletodim, fluazifop-butilo, glifosato, norflurazona, pendimetalina, piritiobaco-sodio, trifloxisulfurona, tepraloxidim, glufosinato, flumioxazina, tidiazurona; insecticidas para algodón tales como acefato, aldicarb, clorpirifós, cipermetrina, deltametrina, abamectina, acetamiprida, benzoato de emamectina, imidacloprida, indoxacarb, lambda-cihalotrina, espinosad, 20 tiodicarb, gamma-cihalotrina, espiromesifeno, piridialilo, flonicamida, flubendiamida, triflumurona, rinaxipir, beta-ciflutrina, espirotetramato, clotianidina, tiametoxam, tiacloprida, dinetofurano, flubendiamida, ciazipir, espinosad, espinotoram, gamma-cihalotrina, 4-[[[(6-clorpiridin-3-il)metil](2,2-difluoretil)amino]furan-2(5h)-ona, tiodicarb, avermectina, flonicamida, piridialilo, espiromesifeno, sulfoxaflor; y fungicidas para algodón tales como azoxistrobina, bixafeno, boscalida, carbendazim, clorotalonilo, cobre, ciproconazol, difenoconazol, dimoxistrobina, epoxiconazol, 25 fenamidona, fluazinam, fluopiram, fluoxastrobina, fluxapiraxad, iprodiona, isopirazam, isotianilo, mancozeb, maneb, metominostrobin, pentiopirad, picoxistrobina, propineb, prothioconazol, piraclostrobina, quintozeno, tebuconazol, tetraconazol, tiofanato-metilo, trifloxistrobina. Para un tratamiento con herbicidas para algodón, dicha plantas o semillas de algodón comprenden además preferiblemente un rasgo que confiere una tolerancia al herbicida respectivo o son naturalmente tolerantes a un herbicida.

30 Los siguientes Ejemplos no limitativos describen el aislamiento de un promotor selectivo para fibras, y la construcción de genes quiméricos para la expresión selectiva en el desarrollo de células de fibra. A menos que se indique otra cosa en los Ejemplos, todas las técnicas de ADN recombinante se llevan a cabo según protocolos estándar como los descritos en Sambrook y cols. (1989) Molecular Cloning: A Laboratory Manual, Segunda Edición, Cold Spring Harbor Laboratory Press, NY y en los Volúmenes 1 y 2 de Ausubel y cols. (1994) Current Protocols in Molecular Biology, Current Protocols, USA. Materiales y métodos estándar para el trabajo molecular en plantas se describen en Plant Molecular Biology Labfax (1993) de R.D.D. Croy, publicados conjuntamente por BIOS Scientific Publications Ltd (Reino Unido) y Blackwell Scientific Publications, Reino Unido.

40 A lo largo de la descripción y los Ejemplos, se hace referencia a las siguientes secuencias representadas en el listado de secuencias:

SEQ ID No 1: secuencia nucleotídica de la región promotora selectiva en fibras de un gen tipo FB8 procedente de *Gossypium hirsutum* (BAC8K12)

45 SEQ ID No. 2: secuencia nucleotídica del cebador oligonucleotídico tipo FB8\_F

SEQ ID No. 3: secuencia nucleotídica del cebador oligonucleotídico tipo FB8\_R

50 SEQ ID No. 4: secuencia nucleotídica del cebador oligonucleotídico tipo QFB8\_F

SEQ ID No. 5: secuencia nucleotídica del cebador oligonucleotídico tipo QFB8\_R

55 SEQ ID No. 6: secuencia nucleotídica del cebador oligonucleotídico UbiQ\_F

SEQ ID No. 7: secuencia nucleotídica del cebador oligonucleotídico UbiQ\_R

SEQ ID No. 8: secuencia nucleotídica del cebador oligonucleotídico tipo gwFb8Pro(8K12)\_F

60 SEQ ID No. 9: secuencia nucleotídica del cebador oligonucleotídico tipo gwFb8Pro(8K12)\_R

SEQ ID No. 10: secuencia nucleotídica del gen tipo FB8 8K12 (región promotora y región codificante)

SEQ ID No. 11: secuencia de ácido nucleico que codifica el gen de quitina sintasa 2 de *Neurospora crassa*

65 SEQ ID No. 12: señal que se dirige al aparato de Golgi procedente de *Arabidopsis thaliana*

SEQ ID No. 13: secuencia codificante del gen gfa (glutamina:fructosa-6-fosfato amidotransferasa) de *E. coli*

SEQ ID No. 14: secuencia nucleotídica de pTIB259

## Ejemplos

### 5 Ejemplo 1: Identificación de posibles genes expresados preferentemente en fibras

Una lista de aproximadamente 450 genes expresados en fibras potenciales se generó a partir de un análisis de la bibliografía y datos públicos de micromatrices y esta se redujo a través de un procedimiento de examen de la frecuencia de presencia de ESTs homólogas a esos genes en bases de datos públicas (como un sustituto de la abundancia de transcritos) y a continuación seguimiento con RT-PCR cuantitativa a través de una gama de tipos de tejidos y edades de las fibras. Se decidió considerar particularmente el gen F203F9 (DR176749, tipo Fb8). Este es un gen originalmente identificado por Liu y cols. (2006) mediante hibridación sustractiva por supresión a partir de las fibras de *G. hirsutum* cultivar TM-1. Liu y cols., 2006 apuntaron que el clon F203F9 (presentado como el registro del Genbank DR176749) se expresaba predominantemente en fibras de 20 dpa, pero no se expresaba en fibras de 10 dpa. En esa publicación, la EST se anotó como muy similar al ADNc de *G. barbadense* tipo Fb8 (AF531368), pero también está estrechamente relacionada con el ADNc de *G. barbadense* FbLate-2 (U34401) cDNA, ambos miembros de una familia multigénica en algodón.

Una búsqueda por blastN de las ESTs del algodón del Genbank con DR176749 identificaba varias ESTs muy similares, incluyendo AF531368, confirmando que este gen era un miembro de una familia multigénica. DR176749 era idéntico a lo largo de la mayor parte de su longitud a la secuencia EST más larga de AF531368 de modo que esta se usó para diseñar cebadores (tipo Fb8\_F: 5'-attcttttcttctatttggttaacct-3' (SEQ ID NO: 2) y tipo Fb8\_R: 5'-tagtgctcgagccagactga-3' (SEQ ID NO: 3)) que flanqueaban la región codificante de EST tipo Fb8 AF531368 para amplificar el gen o los genes procedentes de ADNc y ADN genómico de fibras de algodón Upland para examinar la diversidad dentro de esta familia génica.

Tanto los clones de ADNc (procedentes de una mezcla de ARNm de fibras Coker315 de 10 y 20 dpa) como tipo FB8 genómicos (procedentes de ADN de Coker315) se clonaron y secuenciaron y estos variaban tanto en longitud como en secuencia, confirmando la naturaleza multigénica de esta familia de genes tipo Fb8. Tres clones de longitud similar pero diferente, por ejemplo, se recuperaron de ADNc de fibras, siendo el clon tipo cFB8 #1 el más cercano en identidad a AF531368, el gen tipo FB8 de *G. barbadense*.

Este ADNc codificaba una proteína predicha de 276 aminoácidos considerablemente más larga que los 165 residuos de aminoácido de la proteína tipo Fb8 de *G. barbadense* pero muy similar en secuencia.

La proteína no ha reconocido dominios conservados (según se confirmaba al usar la plataforma de predicción proporcionada en <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>), pero tiene un péptido de señal predicho para secreción (según se confirmaba usando la plataforma de predicción proporcionada en <http://www.cbs.dtu.dk/services/SignalP/>) y tiene una estructura de secuencia repetida interesante de tres dominios ricos en glutamato-lisina según se predice por el cbersitio RADAR (<http://www.ebi.ac.uk/Tools/Radar/>). La proteína tipo FB8 de *G. barbadense* tiene solo la primera de estas tres repeticiones.

Se diseñaron cebadores de Q-PCR hasta casi el extremo C-terminal de una secuencia tipo FB8 de consenso y se usaron para perfilar la expresión relativa de genes tipo Fb8 en diferentes tejidos del cultivar de *G. hirsutum* Coker 315 usando el gen de ubiquitina como una referencia. La preparación de ARN se llevó a cabo según Wan y Wilkins (1994) con hasta 2 g de fibras trituradas u otros tejidos. Todas las preparaciones de ARN se trataron con ADNsa y se purificaron mediante extracción orgánica. Los experimentos de PCR en tiempo real usaban 5 µg de ARN total para la producción de ADNc según las recomendaciones del fabricante (Superscript II, BRL Life Technologies, Gaithersburg, MD, EE. UU. de A.) y se cebaron con 0,5 µg de oligo-dT (dT18).

Los cebadores usados se diseñaron con el Cebador 3 ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)) y eran tipo QFb8\_F: 5'-gagtgccaagagtcacacga-3' (SEQ ID NO: 4) y tipo QFb8\_R: 5'-ttggatcccaccctaaaca-3' (SEQ ID NO: 5) que amplifican un producto de 113 bp. Los cebadores del gen de ubiquitina de algodón eran UbiQ\_F: 5'-aatccatttgacactggtc-3' (SEQ ID NO: 6) y UbiQ\_R: 5'-ctagtggccagctcacatca-3' (SEQ ID NO: 7) que amplifican un producto de 131 bp (Al Ghazi y cols., 2009). Los experimentos de PCR en tiempo real cuantitativa se llevaron a cabo en un sistema de PCR en tiempo real rápido Applied Biosystems 7900HT (CA. EE. UU. de A.) según el siguiente procedimiento. Se pipetearon en placas de 96 pocillos 15 µl de una mezcla madre que consistía en 10 µl de 2X SYBR Green JumpStart Taq Ready Mix (Sigma), 0,5 µl de cada uno de 20 µM de oligonucleótidos directo e inverso correspondientes a un gen diana dado y 4 µl de agua de calidad para PCR. Las plantillas (5 µl de una dilución 1/500 de la reacción de ADNc) se añadieron a continuación a las mezclas madre y se transfirieron al ciclador térmico. Las condiciones de ciclación eran 5 min de desnaturalización a 95°C seguido por 40 ciclos de desnaturalización a 95°C durante 15 segundos, renaturalización a 60°C durante 15 segundos y elongación a 72°C durante 20 segundos. Después de la amplificación, se llevó a cabo una fase de disociación para detectar cualesquiera productos

complejos. El análisis de datos se realizó con el software RQmanager V1.2 (Applied Biosystems) y la abundancia relativa de transcritos se comparó con fibras de 35 dpa determinada usando el gen de ubiquitina de algodón como un gen de referencia para corregir la entrada de ADNc usando el método  $\Delta\Delta Ct$ .

- 5 Según los datos de Q-PCR (Figura 1), la familia tipo Fb8 se expresa altamente principalmente en la fase de engrosamiento de la pared celular de desarrollo de fibras (de 20 a 35 dpa), aunque es la más alta en fibras de aproximadamente 35 dpa, con poca o ninguna expresión en otras partes.

#### Ejemplo 2: Aislamiento de un promotor tipo Fb8

- 10 La inserción del clon de ADNc tipo cFB8#1 se cortó, se purificó en gel y se usó como una sonda para cribar una biblioteca de BAC de inserción grande procedente de *G. hirsutum* cv Acala Maxxa (GH\_MBb, Clemson Genomics Institute) según el protocolo recomendado por los proveedores (<http://www.genome.clemson.edu/resources/protocols>). Se obtuvieron varios BACs de hibridación fuerte, se preparó ADN y se agrupó por el tipo de patrón de restricción usando transferencia Southern. La secuenciación con cebadores en la región codificante de los clones con patrones de restricción únicos identificaba el clon BAC8K12. 15 BAC8K12 codificaba una proteína casi idéntica al tipo cFb8#1 (Figura 2).

- 20 La secuenciación aguas arriba en el promotor y abajo en la región codificante se llevó a cabo de un modo por etapas a partir de ADN de BAC. El gen codificado no parecía contener un intrón y codificaba una proteína de 277 aminoácidos. La secuencia del promotor y el gen para BAC8K12 se muestra en SEQ ID NO: 10 (región codificante en letras mayúsculas; la región promotora usada para la construcción de expresión de indicador está subrayada. Las regiones contenidas dentro de los cebadores Gateway usados para amplificar el promotor procedente de ADN de BAC se indican en negrita):

aagcttaaaatcacttatcaatttcaaaaacagaggtagccgaatgctaagagcttaaaatggctcttttgtttcttttc  
ttgcaaacggtggagagaagagggaatgaagattgaccatattttttattatgttttaacatataatattaataatttaacat  
aattatacttttggtgaatgtgacagtgggatatacgtaaagtataacattatacttttgaagcagttggctggtctaccca  
agagtgatcaaatgtgagctgcctcaatgagccaattttgcctataatggataaaggcactttgtctagtccaactgctcaca  
gaataatgttaaaatgaaattaaaacaaggtgggtggtcacacacaaaaataaaatattaatgtggtgtttggttagtcgaa  
tttatcataattttatagaattgagactaaaaacattaagagaataaattctataacaagacaatttagtaaaatgtcctttta  
ggtaatttaagtactcttaacaaaaataaaatcaaatcaaatcaacaaataagataatataacatacggaacatcccact  
tataatcttcatcccgtaattcttattgaaaagtaattcttattactcgaatcaaatgctctcacaaactattatctaataaa  
gaaaaacacttaattttataacattttttcatattttgaaagattattttgtattttacgtaaaaatattgacatagattgagc  
accttttaacataattccaccataagtcaattatgtagatgagaattgggtacaacaacgtggagccaaatcccacaaacc  
atctctcatctctctataaaaaggcttgcacacatagacaacaatccacacaaaaatacacttaaaattcttttcttctatt  
tggttaaccATGGCTCATAACTTTTCGTCATCCTTTCTTCCTTTTCCAACCTTTTACT  
CATTACTGTCTACTAATGATCGGTAGCCACACCGTCTCGTCAGCGGCTCG  
ACATTTATTCCACACACAAACAACCTCATCAGAGCTGCCACAATTGGCTTC  
AAAATACGAAAAGCACGAAGAGTCTGAATACAAACAGCCAAAATATCAT  
GAAGAGTACCCAAAACATGAGAAGCCTGAAATGTACAAGGAGGAAAAAC  
AAAAACCCGCAAACATCATGAAGAGTACCACGAGTCACGCGAATCGAAG  
GAGCACGAAGAGTACGATAAAGAAAAACCCGATTTCCCAAATGGGAAA  
AGCCTAAAGAGCACGAGAAACACGAAGTCGAATATCCGAAAATACCCGA  
GTACAAGGACAAACAAGATGAGGATAAGGAACATAAAAAATGAAGAGTAC  
CACGAATCACGTGAATCGAAGGAGCACGAAGAATACGAGAAAGAAAAAC  
CCGAGTTCCCAAACGGGAAAAGCCTAAAGAGCACGAGAAACACGAAGT  
CGAATATTTCGAAATACCCGAGTACAAGGAAAAGGAAGATAAGAGTAAG  
AAACATAAAGATGAAGAGTGCCAGGAGTCACACGAATCGAAAGAGCACG  
AAGAGTACGAGAAAGAAAAACCCGATTTCCCAAATGGGAAAAGCCTAA  
AGGGCACGAGAAACATAAAGCCGAATATCCGAAAATACCTGAGTGCAAG  
GAAAACTAGATGAGGATAAGGAACATAAACATGAGTTCCCAAAGCATG  
AAAAAGAAGAGGAGAAGAAACCTGAGAAAGGCATAGTACCCTGAggtgggta  
aaatgcctgaatggccgaagtccatgtttactcagctggtcgagcactaagccttaagccatgacactgggtgcatgtgc  
catcatcatgcagtaatttcatgggatatcgtaattatattgtaataaaaagatgtgagtg

Ejemplo 3: Producción de una construcción de expresión que comprende tipo pFb8 ligado operativamente a GUS

Un fragmento de 959 bp de la región promotora aguas arriba del codón de inicio ATG se amplificó a partir de ADN de BAC de BAC8K12 usando los cebadores Gateway tipo gwFb8Pro(8K12)\_F (SEQ ID No. 8):

5 5'-ggggacaagttgtacaaaaagcaggctgaagcttaaaatcacttatcaatttcaaaaacaga-3' y tipo gwFb8Pro(8K12)\_R (SEQ ID No. 9):

10 5'-ggggaccactttgtacagaagaagctgggtggttaaccaaatagaaagaaagaattttaagt-3' (las secuencias específicas del gen están en negrita) que contienen los sitios de recombinación attB1 y attB2 para la recombinación *in vitro* usando el sistema de clonación Gateway (Invitrogen). La amplificación por PCR a partir de ADN de BAC seguida por depuración con PEG era según las instrucciones del fabricante y la recombinación *in vitro* usando la reacción BP con el vector intermedio pDONR201 (Invitrogen) y la transformación en células DH5alfa daban como resultado la producción de pDONR/ tipo gwFb8(8K12). El vector intermedio se secuenció con los cebadores pDONR-F y pDONR-R (Invitrogen) y a continuación se recombinó en un vector de destino del gen indicador a GUS usando la reacción LR según se recomendaba. Este vector contiene los límites izquierdo y derecho del ADN T, un gen GUS sin

promotor aguas abajo de un casete de recombinación Gateway attR, un terminador procedente de enzima málica de *Flavaria bidentis* y un gen de resistencia a kanamicina expresable en plantas con un promotor procedente del ARN subgenómico del segmento 1 y un terminador del segmento 3 del virus del raquitismo del trébol subterráneo según se describe en Schünmann y cols., (2003). El plásmido resultante tipo pFb8(812K)-GUS (Figura 3) se secuenció con un cebador en el gen indicador de GUS para confirmar la recombinación correcta del promotor y a continuación se transfirió a la cepa de *Agrobacterium* AGL1 (Lazo y cols., 1991) mediante electroporación.

La secuencia del promotor en la dirección 5'-3' (Fig 4) corresponde a SEQ ID NO: 1

Ejemplo 4: Expresión de una construcción que comprende la región promotora de tipo pFb8 ligada operativamente a GUS

La funcionalidad del promotor se probó mediante bombardeo de vector tipo pFb8(8K12)-GUS en óvulos de algodón cultivados y tinción para la actividad de la enzima GUS (Jefferson y cols., 1987). La actividad de Gus se ensayó en fibras de 7 dpa, 21 dpa y 28 dpa (véase la Tabla 1 y la Figura 5).

| Promotor       | 7 dpa      | 16 dpa | 29 dpa     |
|----------------|------------|--------|------------|
| 35S            | 40/10      | 1/18   | no probado |
| tipo Fb8(8K12) | No probado | 1/40   | 5/48       |

Tabla 1: número de manchas de GUS/número de óvulos analizados. El promotor tipo Fb8(8K12) era capaz de conducir la expresión de GUS entre 16 dpa y 29 dpa.

La transformación de algodón con la cepa de *Agrobacterium* que soporta el vector tipo pFb8(8K12)-GUS se llevó a cabo esencialmente como se describe en Murray y cols., (1999) usando los segmentos de cotiledón de plántula del cultivar de algodón Coker315-11 (una selección del cultivar Coker 315 regenerable) y selección con respecto a la resistencia a sulfato de kanamicina.

Se generó un total de 22 transformantes primarios que representaban 11 episodios de transformación independientes. Las plantas se transfirieron al suelo y se dejó que florecieran. Las flores se siguieron en la antesis y las cápsulas tanto de 10 como de 20 dpa se recogieron y semillas enteras con fibras se tiñeron con respecto a la enzima GUS usando tinciones histoquímicas de X-gluc según se describe por Jefferson y cols., 1987. También se tiñeron con respecto a GUS hojas inmaduras, segmentos de peciolo, raíces y secciones de tallo de ramas tanto jóvenes como más viejas. Los patrones de tinción para cuatro líneas independientes diferentes se resumen en la Tabla 2 e imágenes representativas se proporcionan en la Figura 6.

| Línea de transformante | Hoja                                                                      | Peciolo | Tallo                                                                                               | Raíz | fibra de 10 dpa | fibra de 20 dpa |
|------------------------|---------------------------------------------------------------------------|---------|-----------------------------------------------------------------------------------------------------|------|-----------------|-----------------|
| T447-1-1               | - (sin tinción de tricomas)                                               | -       | - (sin tinción de tricomas)                                                                         | -    | -               | +++             |
| T447-21-6              | - (sin tinción de tricomas)                                               | -       | - (sin tinción de tricomas)                                                                         | -    | -               | +++             |
| T447-22-2              | - (sin tinción de tricomas)                                               | -       | - (sin tinción de tricomas)                                                                         | -    | -               | +++             |
| T447-56                | + (sin tinción de tricomas, algo de tinción cerca de los bordes cortados) | -       | + algo de tinción de glándulas superficiales cerca de los bordes cortados, sin tinción de tricomas) | -    | -               | +++             |

Tabla 2: Tejidos en los que se detectaba actividad de la enzima GUS en cuatro plantas de algodón T<sub>0</sub> diferentes transformadas con la construcción tipo pFb8(8K12)-GUS. +, expresión débil o localizada, ++ expresión media, +++ expresión muy fuerte.

El promotor tipo Fb8 (8K12) también parece ser inducible por lesiones según se indica por tinción de GUS alrededor de los bordes cortados de las hojas. Un ejemplo representativo se muestra en la Figura 7.

El promotor tipo Fb8(8K12) parece ser altamente preferente en fibras aunque hay algo de expresión en tricomas glandulares pequeños en otras partes de la planta posibles durante las fases tardías de desarrollo de esas glándulas. El promotor también es inducible por lesiones y esto está de acuerdo con el gen tipo Fb8 que está implicado en el crecimiento y el engrosamiento rápidos de la pared celular implicados en el procedimiento diverso de expansión y maduración de fibras, maduración de glándulas secretoras y reparación de la pared celular después de la lesión.



Ejemplo 5: Expresión de una construcción que comprende tipo pFb8 ligado operativamente a GUS en plantas T1

Semillas T1 procedentes de transformantes primarios independientes con alta expresión en fibra de 20 dpa se recogieron y se plantaron en el invernadero. Estas líneas eran: T447-1-1, T447-21-6, T447-22-2 y T447-56. Se observaban patrones de tinción similares en las cuatro líneas. Se tiñeron secciones de cotiledones con X-Gluc según se describe previamente. Se observó actividad de Gus alrededor de los bordes cortados y a lo largo de la vasculatura que conduce más allá de los bordes cortados. También se observó tinción en algunos tricomas glandulares pequeños que están distribuidos a través del cotiledón, particularmente a lo largo de la vasculatura. Estas reacciones son típicas de una respuesta a lesiones y también se han observado en otros tejidos de la planta T1 incluyendo las hojas y los peciolo. Un ejemplo representativo se muestra en la Figura 7.

Dos plantas T1 procedentes de cada línea se cultivaron hasta la madurez y las flores se siguieron para recoger cápsulas de diferentes edades (0, 10, 20 y 30 dpa) y se recogieron varios otros tejidos incluyendo hoja, peciolo, tallo, raíz, envuelta de la cápsula (a 10 y 20 dpa), capullo pequeño y diversas partes de las flores a 0 dpa (incluyendo bráctea floral, pétalo, estambres y estigma). Estos se tiñeron con respecto a la actividad de GUS. La tinción GUS era en solución X-Gluc durante la noche a 37°C según se describe previamente. Los tejidos se aclararon en una serie de etanol hasta que no se extraía más clorofila y se fotografiaron. Todas las líneas tenían patrones de tinción similares, pero el patrón de tinción para la línea T-447-22-2 se resume en la Tabla 3, mientras que las figuras 8 y 9 son fotografías de fibras de diferentes edades y de diferentes tejidos de algodón en plantas T1 de T-447-22-2.

Tabla 3: Patrón de tinción en diferentes tejidos de algodón de plantas T1 de la línea T-447-22-2

| Tejido                               | Comentario                                                                                                                                               |
|--------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| Óvulo de 0 dpa                       | Superficie del óvulo solo muy débilmente teñida                                                                                                          |
| Fibra y semilla de 10 dpa            | Fibras no teñidas, envuelta seminal moderadamente teñida                                                                                                 |
| Fibra y semilla de 20 dpa            | Fibras moderadamente teñidas, envuelta seminal moderadamente teñida                                                                                      |
| Fibra de 30 dpa                      | Fibras intensamente teñidas, envuelta seminal teñida                                                                                                     |
| Cotiledón                            | Algunos tricomas glandulares pequeños teñidos, tinción más fuerte de los bordes cortados principalmente después de la vasculatura. Sin tricomas teñidos. |
| Hoja                                 | Tinción en venas cerca de zonas lesionadas, sin tricomas pilosos teñidos sobre las hojas.                                                                |
| Peciolo foliar                       | Tinción cerca de los extremos cortados, tricomas pilosos no teñidos                                                                                      |
| Tallo                                | Tinción débil en el tallo y las glándulas subepidérmicas. Sin tricomas pilosos teñidos                                                                   |
| Raíz                                 | Sin tinción                                                                                                                                              |
| Envuelta de cápsula (0, 10 y 20 dpa) | Tinción moderada del tabique de la envuelta de la cápsula que separa los lóculos y en algunos bordes cortados                                            |
| Capullo                              | Tinción alrededor de las zonas lesionadas cerca de las brácteas y otros sitios lesionados y en y alrededor de los óvulos.                                |
| Bráctea floral                       | Tinción alrededor de algunos bordes cortados cerca de la base de la bráctea, pero no los tricomas pilosos.                                               |
| Pétalo                               | Algo de tinción de la vasculatura cerca del extremo distal del pétalo. Sin tinción cerca de la base del pétalo o en los tricomas alrededor de la base.   |
| Estambre/Antera/Polen                | Sin tinción en la pared de la antera, pero tinción débil en el filamento. Tinción fuerte en el polen                                                     |
| Estigma/Estilo                       | Tinción alrededor de la punta del estigma y en la base del estilo en la que se retiran los pétalos y las brácteas.                                       |

En general, estos datos apoyan que el promotor tipo Fb8 8K12 es preferente en fibras y se expresa a lo largo del desarrollo de las fibras predominantemente en fibras más viejas que sufren deposición de la pared celular secundaria. Se expresa en cualquier parte de la planta, pero habitualmente en un número relativamente pequeño de células o tejidos especializados, incluyendo tricomas glandulares pequeños o glándulas subepidérmicas sobre cotiledones, tallos y peciolo. El promotor parece ser sensible a lesiones en una variedad de tejidos y es regulado al alza en la vasculatura después de una lesión de modo que puede ser sensible a una señal hormonal transmitida a través de la vasculatura.

## Ejemplo 6: Expresión de quitina sintasa bajo el control del promotor 8K12 en algodón

Las quitina sintasas se pueden expresar en plantas de algodón para incrementar las cargas positivas en la fibra de algodón al introducir polímeros de quitina en la pared celular de las fibras. Para esto, la expresión preferente en fibras o específica en fibras es importante ya que las plantas transformadas con un gen de quitina sintasa apenas muestran un fenotipo apreciable si el promotor que controla la expresión de la quitina sintasa está conduciendo la expresión en muchos otros tejidos o tipos de células distintos a las células de fibras.

Se generó el siguiente vector de ADN T que comprende genes quiméricos según la invención: vector de ADN T (pTIB359) que comprende una secuencia de ácido nucleico que codifica el gen de quitina sintasa 2 de *Neurospora crassa* (SEQ ID NO: 11) que comprende una señal que se dirige al aparato de Golgi procedente de *Arabidopsis thaliana* (Pagny y cols., 2003, SEQ ID NO: 12) bajo el control del promotor 8K12 y que comprende además un gen epsps recombinante como gen marcador seleccionable, así como un gen recombinante que comprende la región codificante gfa (glutamina:fructosa-6-fosfato amidotransferasa) procedente de *E. coli* (Frohberg y Essigmann, 2006; SEQ ID NO: 13) bajo el control del promotor 8K12. La secuencia nucleotídica de este vector se representa en SEQ ID No. 14. Los elementos genéticos del vector se representan en la Tabla 4.

Tabla 4: Elementos de pTIB359

| Inicio | Fin   | Nombre       | Región                                                                                                                                                                                          |
|--------|-------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1      | 25    | RB           | Repetición del límite derecho procedente del ADN T de <i>Agrobacterium tumefaciens</i>                                                                                                          |
| 31     | 60    | MCS          | Sitios de clonación policonectores                                                                                                                                                              |
| 61     | 1018  | Pfb8-lke-2   | secuencia que incluye la región promotora del gen tipo fb8 (8K12) de <i>Gossypium hirsutum</i>                                                                                                  |
| 1020   | 1124  | RPxylTat     | secuencia codificante para el péptido de retención del aparato de Golgi del gen de beta-1,2-xilosiltransferasa de <i>Arabidopsis thaliana</i> (Pagny y cols., 2003)                             |
| 1125   | 3956  | chs2Nc-1Pc   | secuencia codificante del gen de quitina sintasa 2 de <i>Neurospora crassa</i>                                                                                                                  |
| 3969   | 4205  | 3'35S        | secuencia que incluye la región no traducida 3' del transcrito 35S del virus del mosaico de la coliflor                                                                                         |
| 4207   | 4242  | MCS          | Sitios de clonación policonectores                                                                                                                                                              |
| 4243   | 5201  | Pfb8-lke-2   | secuencia que incluye la región promotora del gen tipo fb8 de <i>Gossypium hirsutum</i> (algodón) (8K12)                                                                                        |
| 5202   | 7031  | gfaEc-1Pb    | región codificante del gen de glutamina:fructosa-6-fosfato amidotransferasa de <i>Escherichia coli</i> (Frohberg y Essigmann, 2006), adaptada al uso de codones de la planta                    |
| 7032   | 7045  | MCS          | Sitios de clonación policonectores                                                                                                                                                              |
| 7049   | 7709  | 3'histonAt   | secuencia que incluye la región no traducida 3' del gen de histona H4 de <i>Arabidopsis thaliana</i>                                                                                            |
| 7760   | 8674  | Ph4a748 ABC  | secuencia que incluye la región promotora del gen de histona H4 de <i>Arabidopsis thaliana</i>                                                                                                  |
| 8725   | 9187  | intrónI h3At | primer intrón del gen II de la variante de histona H3.III de <i>Arabidopsis thaliana</i>                                                                                                        |
| 9194   | 9565  | TPotp C      | secuencia codificante del péptido transitorio optimizado, que contiene la secuencia de los genes de la subunidad pequeña RuBisCO de <i>Zea mays</i> (maíz) y <i>Helianthus annuus</i> (girasol) |
| 9566   | 10903 | 2mepsps      | secuencia codificante del gen de 5-enol-piruvilsiquimato-3-fosfato sintasa doblemente mutante de <i>Zea mays</i> (maíz)                                                                         |
| 10927  | 11587 | 3'histon At  | secuencia que incluye la región no traducida 3' del gen de histona H4 de <i>Arabidopsis thaliana</i>                                                                                            |
| 11686  | 11724 | LB           | Repetición del límite izquierdo procedente del ADN T de <i>Agrobacterium tumefaciens</i>                                                                                                        |
| 13585  | 12031 | aadA C       | Sm/Sp R - la secuencia codificante del gen de aminoglicósido adeniltransferasa (aadA) de <i>Escherichia coli</i> .                                                                              |

13815 16604 ORI pVS1 Origen de replicación plasmídico procedente de pVS1 para el mantenimiento estable en *Agrobacterium*.

16605 17668 ORI ColE1 fragmento que incluye el origen de replicación del plásmido pBR322 para la replicación en *Escherichia coli*

Este vector se transfirió en una cepa de *Agrobacterium* apropiada que se usó para transformar la variedad de algodón Coker312-17. Se obtuvieron 17 líneas T0 que desarrollaban semillas. Los experimentos de campo con la progenie de al menos una de estas líneas transgénicas demostraban que se podían obtener plantas transgénicas con un fenotipo casi idéntico al control silvestre.

5

Cuando fibras de algodón recogidas de plantas cultivadas en invernadero de la al menos una de las susodichas líneas transgénicas se teñían con Acid Orange 7, las líneas transgénicas daban como resultado un agotamiento mayor del colorante remanente (13,9 para fibras procedentes de plantas transgénicas frente a 10,9 para fibras procedentes de controles no transgénicos; es decir un incremento de 27% de agotamiento). Estos resultados se representan en la Figura 10.

10

## REFERENCIAS

- Al-Ghazi Y, Bourot S, Arioli T, Dennis ES, Llewellyn DJ (2009) Transcript profiling during fiber development identifies pathways in secondary metabolism and cell wall structure that may contribute to cotton fiber quality. *Plant and Cell Physiology* 50: 1364-1381.
- Carrington & Freed 1990. Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. *J. Virology* 64, p. 1590-1597.
- Chaubet y cols. (1992). Genes encoding a histone H3.3-like variant in *Arabidopsis* contain intervening sequences. *J Molecular Biology* 225(2):569-74.
- Cornelissen y Vandewiele (1989). Both RNA level and translation efficiency are reduced by anti-sense RNA in transgenic tobacco. *Nucleic Acids Research* 17(3):833-43.
- Delaney y cols. (2007). The fiber specificity of the cotton FSIt4 gene promoter is regulated by an AT-rich promoter region and the AT-hook transcription factor GhAT1. *Plant and Cell Physiology* 48, 1426-1437.
- Eisner, T. y cols. (1998). When defence backfires: detrimental effect of a plant's protective trichomes on an insect beneficial to the plant. *Proceeding of the National Academy of Sciences, USA* 95, p. 4410-4414.
- Frohberg y Essigmann (2006) WO2006/032538.
- Hsu y cols. (1999). Analysis of promoter activity of cotton lipid transfer protein gene LTP6 in transgenic tobacco plants. *Plant Science* 143, p. 63-70.
- Jefferson, R.A., Kavanagh, T.A. y Bevan, M.W. (1987) GUS fusions:  $\beta$  - glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6: 3901-3907.
- John y Keller (1996). Metabolic pathway engineering in cotton: biosynthesis of polyhydroxybutyrate in fiber cell. *Proceeding of the National Academy of Sciences, USA* 93, 12678-12773.
- Kim y Triplett, 82001). Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. *Plant Physiology* 127, p. 1361-1366.
- Lazo, G.R., Stein, P.A. y Ludwig, R.A. (1991) A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Biotechnology* 9: 963-967.
- Li y cols. (2005). The cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation. *The Plant Cell* 17, p. 859-875.
- Lindbo, J.A. (2007). High-efficiency protein expression in plants from agroinfection-compatible Tobacco mosaic virus expression vectors. *BMC Biotechnology* 27 de agosto de 2007;7:52.
- Liu y cols. (2000). Cloning and promoter analysis of the cotton lipid transfer protein gene Ltp3. *Biochimica et Biophysica Acta* 1487, p. 106-111.
- Liu, D., Zhang, X., Tu, L., Zhu, L., Guo X. (2006) Isolation by suppression-subtractive hybridization of genes preferentially expressed during early and late fiber development stages in cotton. *Molecular Biology* 40: 741-749.
- Luo y cols. (2007). GhDET2, a steroid 5 $\alpha$ -reductase, plays an important role in cotton fiber cell initiation and elongation. *The Plant Journal* 51, p. 419-430.
- Medeiros y Tingey, (2006). Glandular trichomes of *Solanum berthaultii* and its hybrids with *Solanum tuberosum* affect nymphal emergence, development, and survival of *Empoasca fabae* (Homoptera: Cicadellidae). *Journal of Economic Entomology* 99, p. 1483-1489.
- Murray, F., Llewellyn, D., McFadden, H., Last, D., Dennis, E.S. y Peacock, W.J. (1999). Expression of the *Talaromyces flavus* glucose oxidase gene in cotton and tobacco reduces fungal infection, but is also phytotoxic. *Molecular Breeding* 5: 219-232.
- Needleman y Wunsch (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Molecular Biology*, 48, p.443-453.

- S. Pagny, F. Bouissonie, M. Sarkar, M.L. Follet-Gueye, A. Driouich, H. Schachter, L. Faye y V. Gomord (2003) Structural requirements for Arabidopsis  $\beta$ 1,2-xylosyltransferase activity and targeting to the Golgi. *The Plant Journal*. Vol 33: 189-203
- 5 Pearson y Lipman (1988). Improved tools for biological sequence comparison. *Proceeding of the National Academy of Sciences, USA* 85, p.2444-48.
- Ranger y Hower, 2001. Role of the glandular trichomes in resistance of perennial alfalfa to the potato leafhopper (Homoptera: Cicadellidae). *Journal of Economic Entomology* 94, p. 950-957.
- 10 Ruan y cols. (2003). Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. *The Plant Cell* 15, p. 952-964.
- 15 Sambrook, J.F., Russell, D.W. e Irwin, N. (2000). *Molecular Cloning: A Laboratory Manual*, 3ª edición Volúmenes 1, 2, and 3. Cold Spring Harbor Laboratory Press.
- Schünmann PHD, Llewellyn D, Surin B, Boevink P, Defeyter RC, Waterhouse PM (2003) A suite of novel promoters and terminators for plant biotechnology. *Functional Plant Biology* 30: 443-452.
- 20 Shangguan y cols. (2008). Promoter of a cotton fiber MYB gene functional in trichomes of Arabidopsis and glandular trichomes of tobacco. *Journal of Experimental Botany* 59(13), p. 3533-3542.
- Siomi, H. y Siomi, M. (2009). On the road to reading the RNA-interference code. *Nature* 457, 396-404.
- 25 Song y cols. (2000). Expression of two tissue-specific promoters in transgenic cotton plants. *Journal of Cotton Science* 4, p. 217-223.
- Szymanski y cols. (2000). Progress in the molecular genetic analysis of trichome initiation and morphogenesis in Arabidopsis. *Trends in Plant Science* 5, p. 214-219.
- 30 Wagner y cols. (2004). New approaches for studying and exploiting an old protuberance, the plant trichome. *Annals of Botany* 93, p. 3-11.
- 35 Wan CY, Wilkins TA (1994) A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Analytical Biochemistry* 223: 7-12.
- Wang, S. y cols. (2004). Control of plant trichome development by a cotton fiber MYB gene. *The Plant Cell* 16, p. 2323-2334.
- 40 Waterman, M. S. (1995). *Introduction to Computational Biology: Maps, sequences and genomes*. Chapman & Hall. London.
- Werker, E. (2000). Trichome diversity and development. *Advances in Botanical Research* 31, p. 1-35.

## LISTADO DE SECUENCIAS

|    |                                                                    |     |
|----|--------------------------------------------------------------------|-----|
|    | <110> Bayer CropScience NV                                         |     |
|    | Commonwealth Scientific and Industrial Research Organisation       |     |
| 5  | <120> Nuevo promotor preferente en fibras en algodón               |     |
|    | <130> BCS 13-2022                                                  |     |
| 10 | <160> 14                                                           |     |
|    | <170> PatentIn versión 3.5                                         |     |
|    | <210> 1                                                            |     |
| 15 | <211> 959                                                          |     |
|    | <212> ADN                                                          |     |
| 20 | <213> Secuencia Artificial                                         |     |
|    | <220>                                                              |     |
|    | <223> región promotora selectiva en fibras procedente de BAC8K12   |     |
| 25 | <400> 1                                                            |     |
|    | aagcttaaaa tcacttatca atttcaaaaa cagaggtag ccgaatgcta agagcttaaa   | 60  |
|    | aatggcttct tttgtttctt tttcttgcaa acggtggaga gaagaggga atgaagattg   | 120 |
|    | accatatctt tttttattat gttttaacat ataattataa taatttaatc ataattatac  | 180 |
|    | tttgggtgaat gtgacagtgg ggatatacgt aaagtatata acattatact ttttgcaagc | 240 |
|    | agttggctgg tctacccaag agtgatcaaa ttttgagctg ccttcaatga gccaattttt  | 300 |
|    | gcctataatg gataaaggca ctttgtctag ttcaactgct cacagaataa tgttaaaatg  | 360 |
|    | aaattaaaac aaggtgggct ggtcacacca aaaaaataaa atattaatgt ggtgtttggt  | 420 |
|    | tagtcgaatt tatcataatt ttatagaaat tgagactaaa aaacattaag agaataaatt  | 480 |
|    | ctataacaaa gacaatttag taaaaatgtc cttttaggta attttaagta ctcttaacca  | 540 |
|    | aaataaaaaa ttcaaatcaa attaaccaaa taagataata taacatacgg aacatccac   | 600 |
|    | ttataatctt acatccccgt aatcttatta tgaaaagtaa tcttatatta ctggaatcaa  | 660 |
|    | atgctctcac aaactattat ctaaataaag aaaaacactt aatttttata acattttttt  | 720 |
|    | tcatatattt gaaagattat attttgtatt ttacgtaaa aatatttgac atagattgag   | 780 |
|    | caccttttta acataattcc accataagtc aattatgtag atgagaaatt ggtacaaaca  | 840 |
|    | acgtggagcc aaatcccacc aaaccatctc tcacctctc ctataaaag cttgctacac    | 900 |
|    | atagacaaca atccacacaa aaatacactt aaaattcttt tctttctatt tggttaacc   | 959 |
| 30 | <210> 2                                                            |     |
|    | <211> 28                                                           |     |
|    | <212> ADN                                                          |     |
| 35 | <213> Secuencia Artificial                                         |     |

<220>

<223> cebador oligonucleotídico tipo Fb8\_F

5 <400> 2

attcttttct ttctatttgg ttaacat 28

<210> 3

10 <211> 20

<212> ADN

15 <213> Secuencia Artificial

<220>

<223> cebador oligonucleotídico tipo FB8\_R

20 <400> 3

tagtgctcga gccagactga 20

25 <210> 4

<211> 20

<212> ADN

30 <213> Secuencia Artificial

<220>

35 <223> cebador oligonucleotídico QFb81ike\_F

<400> 4

gagtgccaag agtcacacga 20

40 <210> 5

<211> 20

45 <212> ADN

<213> Secuencia Artificial

<220>

50 <223> cebador oligonucleotídico tipo QFb8\_R

<400> 5

55 ttggatccca cccttaaaca 20

<210> 6

<211> 20

60 <212> ADN

<213> Secuencia Artificial

65 <220>

<223> cebador oligonucleotídico UbiQ\_F  
 <400> 6  
 5 aatccacttt gcacctgggc 20  
 <210> 7  
 <211> 20  
 10 <212> ADN  
 <213> Secuencia Artificial  
 15 <220>  
 <223> cebador oligonucleotídico UbiQ-R  
 <400> 7  
 20 ctagtgcca gctcacatca 20  
 <210> 8  
 25 <211> 63  
 <212> ADN  
 <213> Secuencia Artificial  
 30 <220>  
 <223> secuencia oligonucleotídica tipo gwFb8Pro(8K12)\_F  
 35 <400> 8  
 ggggacaagt ttgtacaaaa aagcaggcta agcttaaaat cacttatcaa tttcaaaaac 60  
 aga 63  
 <210> 9  
 40 <211> 62  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> cebador oligonucleotídico tipo gwFb8Pro(8K12)\_R  
 50 <400> 9  
 ggggaccact ttgtacaaga aagctgggtg gttaacaaaa tagaaagaaa agaattttta 60  
 gt 62  
 55 <210> 10  
 <211> 1943  
 <212> ADN  
 60 <213> Secuencia Artificial



<220>

<223> promotor y gen para BAC8K12

5 <400> 10

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aagcttaaaa tcacttatca atttcaaaaa cagaggtag ccgaatgcta agagcttaaa   | 60   |
| aatggcttct tttgtttctt tttcttgcaa acggtggaga gaagaggaa atgaagattg   | 120  |
| accatatctt tttttattat gttttaacat ataattataa taatttaatc ataattatac  | 180  |
| tttggtgaat gtgacagtgg ggatatacgt aaagtatata acattatact ttttgcaagc  | 240  |
| agttggctgg tctacccaag agtgatcaaa ttttgagctg ccttcaatga gccaatTTTT  | 300  |
| gcctataatg gataaaggca ctttgtctag ttcaactgct cacagaataa tgttaaaatg  | 360  |
| aaattaaac aaggtgggct ggtcacacca aaaaaataa atattaatgt ggtgtttggt    | 420  |
| tagtcgaatt tatcataatt ttatagaaat tgagactaaa aaacattaag agaataaatt  | 480  |
| ctataacaaa gacaatttag taaaaatgtc cttttaggta attttaagta ctcttaacca  | 540  |
| aaataaaaaa ttcaaataca attaaccaaa taagataata taacatacgg aacatcccac  | 600  |
| ttataatctt acatccccgt aatcttatta tgaaaagtaa tcttatatta ctgcaatcaa  | 660  |
| atgctctcac aaactattat ctaaataaag aaaaacactt aatttttata acattttttt  | 720  |
| tcatatattt gaaagattat attttgtatt ttacgtaaa aatatttgac atagattgag   | 780  |
| caccttttta acataattcc accataagtc aattatgtag atgagaaatt ggtacaaaca  | 840  |
| acgtggagcc aaatcccacc aaaccatctc tcatcctctc ctataaaagg cttgctacac  | 900  |
| atagacaaca atccacacaa aaatacactt aaaattcttt tctttctatt tggttaacca  | 960  |
| tggctcataa ctttgcgtcat cctttcttcc ttttccaact tttactcatt actgtctcac | 1020 |
| taatgatcgg tagccacacc gtctcgtag cggctcgaca tttattccac acacaaacaa   | 1080 |
| cctcatcaga gctgccacaa ttggcttcaa aatacgaaaa gcacgaagag tctgaataca  | 1140 |
| aacagccaaa atatcatgaa gagtacccaa aacatgagaa gcctgaaatg tacaaggagg  | 1200 |
| aaaaacaaaa accctgcaaa catcatgaag agtaccacga gtcacgcgaa tcgaaggagc  | 1260 |
| acgaagagta cgataaagaa aaacccgatt tccccaaatg ggaaaagcct aaagagcacg  | 1320 |
| agaaacacga agtcgaatat ccgaaaatac ccgagtacaa ggacaaacaa gatgaggata  | 1380 |
| aggaacataa aaatgaagag taccacgaat cacgtgaatc gaaggagcac gaagaatacg  | 1440 |
| agaaagaaaa acccgagttc cccaaacggg aaaagcctaa agagcacgag aaacacgaag  | 1500 |
| tcgaatatcc ggaaataccc gagtacaagg aaaaggaaga taagagtaag aaacataaag  | 1560 |
| atgaagagtg ccaggagtgca cacgaatcga aagagcacga agagtacgag aaagaaaaac | 1620 |
| ccgatttccc caaatgggaa aagcctaaag ggcacgagaa acataaagcc gaatatccga  | 1680 |
| aaatacctga gtgcaaggaa aaactagatg aggataagga acataaacat gagttcccaa  | 1740 |
| agcatgaaaa agaagaggag aagaaacctg agaaaggcat agtaccctga gtgggttaaa  | 1800 |
| atgcctgaat ggccgaagtc catgtttact cagtctggct cgagcactaa gccttaagcc  | 1860 |
| atatgacact ggtgcatgtg ccatcatcat gcagtaattt catgggatat cgtaattata  | 1920 |
| ttgttaataa aaagatgtga gtg                                          | 1943 |

# ES 2 701 243 T3

<210> 11

<211> 2831

5 <212> ADN

<213> Neurospora crassa

<400> 11

10 agtccagaat cagcaaccgg ttatcgagtt ccgccacaag gacggtagca gccttcagaa 60  
 atcgatgtca tgccaggcca gggacaccgg gatcgagtta cggaaatgcg aggcgaccgc 120  
 ttccctcggc accagcgcct ttacactaca atagcccaag tcgcgcagcg agtcattatc 180  
 cacggtagca tggaggttat gcgacgacg tgacagttag catgggaccg gacgacgatc 240  
 gtacagatat ctttggcccc gaaaccgatc tcagcgaaac gcgccacctc aacgacgcat 300  
 acgggtttcg gtcattcccag atcaccctca gcgaagatcc ccacggcacc cacgcgcgtt 360  
 cccggtacga cgacgaagac gatgtgagca ccacttattc ctccaacacg ggcaccagcg 420

# ES 2 701 243 T3

|             |            |            |            |            |            |      |
|-------------|------------|------------|------------|------------|------------|------|
| cttcagggtgt | cgacaagttc | gagcattacg | gtcccattcc | ggaggaaggc | aagcacgagc | 480  |
| ggcgcgggcgt | gcgaccacca | cagatgtcga | ggaaggaagt | ccagctcatc | aacggcgaac | 540  |
| tcgttctcga  | gtgcaagatt | ccgactatat | tgtattcgtt | tttggccagg | agagacgaag | 600  |
| tggagtttac  | gcacatgcgg | tacacagccg | tcacttgtga | ccctgatgac | tttgttgcca | 660  |
| gggggttacia | gttgcgccag | aatatcggtc | gtaccgccag | ggagacggag | ctgttcacat | 720  |
| gcgtgaccat  | gtacaacgag | gacgagttcg | gattcacacg | gactatgcac | gcagtgatga | 780  |
| agaacatttc  | gcatttttgt | tcccgaatac | agagtaggac | gtggggagcg | gatgggtggc | 840  |
| agaagattgt  | ggtctgtgtg | gtttcggatg | gacgagagat | cattcacccc | cggaccttgg | 900  |
| acgccctcgc  | agccatgggc | gtttaccagc | acggatcgcg | caagaacttt | gtcaaccaga | 960  |
| aggcggtgca  | ggcccacgtt | tacgagtaca | cgacacaagt | gtctctggac | agcgacctca | 1020 |
| agttcaaggg  | cgccgagaag | ggcatcgtgc | cctgccagat | gattttttgc | ttgaaggaga | 1080 |
| agaacaaaaa  | gaaactcaac | tgcataagat | ggttcttcaa | cgcctttggc | aaagccttga | 1140 |
| acccgaatgt  | gtgtatcctc | ctagacgtcg | gcaccgcccc | cggcggcaca | agtctctacc | 1200 |
| atctctggaa  | agccttcgac | acggattcca | acgtggcggg | ggcctgcggg | gaaatcaaag | 1260 |
| cgtgaagggt  | gcgggtttgg | gggaatttgc | tcaaccctct | ggtggctagt | cagaactttg | 1320 |
| agtacaagat  | gagcaatatt | ctggacaaa  | cgttgagatc | ggtgtttggg | tacatcacgg | 1380 |
| tgttgccggg  | cgccttgtcg | gcgtatcggt | accatgcgct | gcagaacgat | gagacgggcc | 1440 |
| atgggcccgt  | gagtcagtat | ttcaaggggc | agacgctcca | tgggcagcac | gcggatgtgt | 1500 |
| ttacggcgaa  | catgtacttg | gccgaggacc | gaattctgtg | ttgggagttg | gtggccaaga | 1560 |
| ggggtgagag  | gtgggtgttg | aagtatgtga | aggggtgtac | gggtgagacg | gatgtgcctg | 1620 |
| acaccgtccc  | ggaattcgtc | tgcacaacgc | gtcgttggct | caacggtgcc | ttcttcgccg | 1680 |
| ccgtctactc  | cctcgtccac | tttcgacaaa | tctggaaaac | cgaccacacc | tttatgcgca | 1740 |
| aagcccttct  | ccacgtcgaa | ttcctctacc | acctcctgca | actcctcttc | acctacttct | 1800 |
| ccctggccaa  | cttctacctc | gccttctact | ttatcgccgg | cggactcgcc | gatccccacg | 1860 |
| tcgacccttt  | taactcggac | ggccacgtcg | cgcgcacatc | cttcaacatc | ctccgctacg | 1920 |
| tctgcgtcct  | gctgatctgc | acacaattca | tcttgtccct | cggcaaccgt | ccgcagggtg | 1980 |
| ccaaaagaat  | gtatctcgca | tccatgatca | tctacgccgt | catcatggtg | tacaccacct | 2040 |
| tcgccaccat  | cttcacgtgc | gtgcgacaaa | tccaaccctc | tcaaaaatcc | gacgacaagc | 2100 |
| ccgacctcga  | actcggcaac | aacgtcttca | ccaacctgat | cgtctccgtg | gctagtaccc | 2160 |
| tcgggctcta  | cttcgtcatg | tcctttctct | atctcgaccc | ctggcacatg | ttcacctcgg | 2220 |
| ccatccagta  | ctttgtcctg | ctgccttctc | acatctgcac | gctccagatc | tacgcctttt | 2280 |
| gcaacaccca  | cgacgtcaca | tggggcacca | aaggcgacaa | cgtgatgcgc | accgatctcg | 2340 |

# ES 2 701 243 T3

|    |                                                                     |      |
|----|---------------------------------------------------------------------|------|
|    | gagggcgccat tggcaagggg agcaccgtcg aactggaaat gccttcggac caactcgaca  | 2400 |
|    | tcgactcggg atacgacgaa tgtctacgaa atctccggga tcgctcatg gtccctgccg    | 2460 |
|    | ttcccgtgtc cgaggaccag ctgcagcagg attactacaa gtcggtgcgc acgtacatgg   | 2520 |
|    | tggtgtcgtg gatggtggcc aacgcgacgc tggccatggc ggtgtcgga gcgtatggcg    | 2580 |
|    | attcgaaaat tggggataat ttttacttgc ggtttatcct gtgggcgggtg gcggccctgg  | 2640 |
|    | cgctgttttag agcgttgggg tcgacgacgt ttgcggcgat taatctggtg agtgctctcg  | 2700 |
|    | tggagggcag ggtcaggctg aggttgaata tgaaagggtt taggtggatt aaggagaagt   | 2760 |
|    | ggggggatgc ggatgtgaag ggcaagtttg aggggttggg ggatcgggcg agggggttgg   | 2820 |
|    | cgaggcgggtg a                                                       | 2831 |
|    | <210> 12                                                            |      |
| 5  | <211> 105                                                           |      |
|    | <212> ADN                                                           |      |
|    | <213> Arabidopsis thaliana                                          |      |
| 10 | <400> 12                                                            |      |
|    | atgagtaaac ggaatccgaa gattctgaag atttttctgt atatgttact tctcaactct   | 60   |
|    | ctctttctca tcatctactt cgtttttcac tcatcgtcgt tttca                   | 105  |
| 15 | <210> 13                                                            |      |
|    | <211> 1830                                                          |      |
|    | <212> ADN                                                           |      |
| 20 | <213> Escherichia coli                                              |      |
|    | <400> 13                                                            |      |
|    | atgtgcggaa ttgttggtgc tatcgcccaa agagacgttg ctgagatttt gttagagggt   | 60   |
|    | ctgcgaaggc tagagtatag aggatatgac tccgctggtc tggctgtcgt tgatgctgag   | 120  |
|    | ggtcatatga caaggctaag aaggtttaga aaggttcaga tgcttgctca ggcagctgag   | 180  |
|    | gaacatccat tgcattggagg tactggtatt gcacatacca ggtgggctac tcatggggag  | 240  |
|    | ccatcagaag ttaatgctca tccacatgtg agtgagcata tcgttgtagt tcacaatggg   | 300  |
|    | ataattgaaa accacgaacc attgaggga gagttaagg caaggagata tacttttgtg     | 360  |
|    | agtgagaactg aacttgaggt tattgcacat ttagtgaact gggaaactcaa acaggggggc | 420  |
|    | acattgcgtg aggctgtggt aagagctatt cctcaactta gaggtgcata cggtagctgt   | 480  |
|    | attatggatt caagacaccc agatactctc cttgcagcta gatcaggtag tcccttggtc   | 540  |
|    | ataggacttg gaatgggtga aaattttatc gctagcgacc aattggcctt attgccagtt   | 600  |
|    | acaagacgat ttattttcct tgaagagggc gatattgctg agattactag aaggctctgtg  | 660  |
|    | aacatctttg ataagactgg cgctgagggt aaacgtcagg atatcgagtc taaccttcaa   | 720  |
| 25 | tacgatgctg gtgataaagg aatttacagg cattatatgc aaaaggaaat ttatgaacaa   | 780  |

# ES 2 701 243 T3

```

ccaaatgcta tcaaaaacac acttactggc cgtattttctc atggacaggt cgatttaagc      840
gagcttggtc ctaatgcaga cgaactgcta tcaaaagtgt agcacatata gatactggca      900
tgcggaacta gttataattc aggaatgggt tctagatact ggttcgaaag cttggcaggt      960
ataccttggt atgtagagat cgcttctgag tttaggtata gaaagtctgc tgtgcgtaga     1020
aattcattaa tgattacatt atctcaatcc ggagaaacag cagatacact ggctggattg     1080
aggctttcta aggaactcgg atatctgggt tcacttgcta tttgtaatgt accaggttcc     1140
tcattggttc gtgaatcaga tctagcactt atgacaaatg caggaactga aataggtgtg     1200
gcaagtacca aggctttcac aacccaactg accgtacttt taatgttggg agcaaaaactc     1260
agtcgattaa aggggctaga tgcactctatc gaacatgata ttgttcacgg gcttcaagct     1320
ctcccttcaa gaattgaaca aatgctttca caagataaga gaatagaggc attggctgaa     1380
gatttttccg acaaacatca cgcattgttt cttggacgtg gcgatcaata tccaattgca     1440
ttggaaggag ctttgaagtt gaaagaaata agttacattc acgcagaagc atatgcagct     1500
ggagaactca agcatggtcc tttggcactc atcgacgtg acatgcccggt gatcgtagtg     1560
gctcctaata acgaactgct cgaaaagctt aaatcaaata tcgaagaggt tcgagctaga     1620
ggaggtcagc tttacgtttt cgctgaacaa gatgctggat tcgtgtcaag cgataaatatg     1680
catataattg aaatgcctca cggtgaagaa gtgattgcac ctatatttta tacagtccca     1740
ttgcaacttc tagcttacca tgttgcaact attaaaggaa ctgatgttga tcagcctaga     1800
aacctagcaa aatctgtaac agtcgaataa                                     1830

```

<210> 14

5 <211> 17967

<212> ADN

<213> Secuencia Artificial

10

<220>

<223> pTIB359

15 <400> 14

```

aattacaacg gtatatatcc tgccagtact gggccccctc gagggcgatc gcgcggccgc      60
aagcttaaaa tcacttatca atttcaaaaa cagagggttag ccgaatgcta agagcttaaa     120
aatggcttct tttgtttctt tttcttgcaa acggtggaga gaagaggga atgaagattg     180
accatatctt tttttattat gttttaacat ataattataa taatttaatc ataattatac     240
tttgggtgaat gtgacagtgg ggatatacgt aaagtatata acattatact ttttgcaagc     300
agttggctgg tctacccaag agtgatcaaa ttttgagctg ccttcaatga gccaatTTTT     360
gcctataatg gataaaggca ctttgtctag ttcaactgct cacagaataa tgttaaaatg     420
aaattaaaac aaggtgggct ggtcacacca aaaaaataaa atattaatgt ggtgtttggg     480
tagtcgaatt tatcataatt ttatagaaat tgagactaaa aaacattaag agaataaatt     540

```

# ES 2 701 243 T3

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ctataacaaa gacaatttag taaaaatgtc ctttttaggta attttaagta ctcttaacca  | 600  |
| aaataaaaaa ttcaaataca attaaccaaa taagataata taacatacgg aacatcccac   | 660  |
| ttataatctt acatccccgt aatcttatta tgaaaagtaa tcttatatta ctcgaaatcaa  | 720  |
| atgctctcac aaactattat ctaaataaag aaaaacactt aatttttata acattttttt   | 780  |
| tcataatatt gaaagattat attttgtatt ttacgtaaa aatatttgac atagattgag    | 840  |
| caccttttta acataattcc accataagtc aattatgtag atgagaaatt ggtacaaaca   | 900  |
| acgtggagcc aaatcccacc aaaccatctc tcacctctc ctataaaagg cttgctacac    | 960  |
| atagacaaca atccacacaa aaatacactt aaaattcttt tctttctatt tggttaacca   | 1020 |
| tgagtaaacg gaatccgaag attctgaaga tttttctgta tatgttactt ctcaactctc   | 1080 |
| tctttctcat catctacttc gtttttctact catcgtcgtt ttcagagtcc agaatacagca | 1140 |
| accggttatt gagttccgcc acaaggacgg tacgagcctt cagaaatcga tgtcatgcca   | 1200 |
| ggccagggac accgggatcg agttacggaa atgcgagggc accgcttccc tcggcaccag   | 1260 |
| cgcctttaca ctacaatagc ccaagtcgcg cagcgagtca ttatccacgg taccatggag   | 1320 |
| gttatgcgga cgacgtgaca gttagcatgg gaccggacga cgatcgtaca gatatctttg   | 1380 |
| gccccgaaac cgatctcagc gaaacgcgcc acctcaacga cgcatacggg tttcggtcac   | 1440 |
| cccagatcac cctcagcgaa gatccccacg gcacccacgc gcgttcccgg tacgacgacg   | 1500 |
| aagacgatgt gagcaccact tattcctcca acacgggcac cagcgcttca ggtgtcgaca   | 1560 |
| agttcgagca ttacgggtccc attccggagg aaggcaagca cgagcggcgc ggcgtgcgac  | 1620 |
| caccacagat gtcgaggaag gaagtccagc tcacccaacgg cgaactcgtt ctcgagtcca  | 1680 |
| agattccgac tatattgtat tcgtttttgc ccaggagaga cgaagtggag ttacgcaca    | 1740 |
| tgcggtacac agccgtcact tgtgacctg atgactttgt tgccaggggt tacaagttgc    | 1800 |
| gccagaatat cggtcgtacc gccagggaga cggagctgtt catctgcgtg accatgtaca   | 1860 |
| acgaggacga gttcggattc acacggacta tgcacgcagt gatgaagaac atttcgcatt   | 1920 |
| tttgttcccg aaacaagagt aggacgtggg gagcggatgg gtggcagaag attgtgtct    | 1980 |
| gtgtggtttc ggatggacga gagatcattc acccccggac cttggacgcc ctgcagcca    | 2040 |
| tgggcgttta ccagcacggt atcgccaaga actttgtcaa ccagaaggcg gtgcaggccc   | 2100 |
| acgtttacga gtacacgaca caagtgtctc tggacagcga cctcaagttc aaggcgccg    | 2160 |
| agaagggcat cgtgccctgc cagatgattt ttgcttgaa ggagaagaac caaaagaaac    | 2220 |
| tcaactcgca tagatggttc ttcaacgcct ttggcaaagc cttgaaccgg aatgtgtgta   | 2280 |
| tcctcctaga cgtcggcacc cgccccggcg gcacaagtct ctaccatctc tggaaagcct   | 2340 |
| tcgacacgga ttccaacgtg gcgggggcct gcggggaaat caaagcgatg aaggggcggt   | 2400 |

# ES 2 701 243 T3

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ttggcgggaa tttgctcaac cctctggtgg ctagtcagaa ctttgagtac aagatgagca  | 2460 |
| atattcttga caaacctgtt gagtcggtgt ttgggtacat cacggtgttg ccgggcgcct  | 2520 |
| tgtcggcgta tcggtaccat gcgctgcaga acgatgagac gggccatggg ccgttgagtc  | 2580 |
| agtatttcaa gggcgagacg ctccatgggc agcacgcgga tgtgtttacg gcgaacatgt  | 2640 |
| acttgccga ggaccgaatt ctgtgttggg agttggtggc caagaggggt gagaggtggg   | 2700 |
| tgttgaagta tgtgaagggg tgtacgggtg agacggatgt gcctgacacc gtcccgaat   | 2760 |
| tcgtctcgca acgtcgtcgt tggctcaacg gtgccttctt cgccgcctc tactccctcg   | 2820 |
| tccactttcg acaaactctg aaaaccgacc acacctttat gcgcaaagcc cttctccacg  | 2880 |
| tcgaattcct ctaccacctc ctgcaactcc tcttcaccta cttctccctg gccaaattct  | 2940 |
| acctcgctt ctactttatc gccggcggac tcgccgatcc ccacgtcgac ctttttaact   | 3000 |
| cggacggcca cgtcgcgcgc atcatcttca acatcctcgc ctacgtctgc gtctgtctga  | 3060 |
| tctgcacaca attcatcttg tccctcggca accgtccgca gggcgccaaa agaattgata  | 3120 |
| tcgcatccat gatcatctac gccgtcatca tgggtgtacac caccttcgcc accatcttca | 3180 |
| tcgtcgtgcg acaaattcaa cctctcaaa aatccgacga caagcccgac ctcgaaactcg  | 3240 |
| gcaacaacgt cttcaccaac ctgatcgtct ccgtggctag taccctcggg ctctacttcg  | 3300 |
| tcattgtcctt tctctatctc gacctctggc acatgttcac ctgggccatc cagtactttg | 3360 |
| tcctgtctgc ttctacatc tgcacgctcc agatctacgc cttttgcaac acccacgacg   | 3420 |
| tcacatgggg caccaaaggc gacaacgtga tgcgcaccga tctcggaggc gccattggca  | 3480 |
| agggaaagcac cgtcgaactg gaaatgcctt cggaccaact cgacatcgac tcgggatacg | 3540 |
| acgaatgtct acgaaatctc cgggatcgcg tcatggctcc tgcggttccc gtgtccgagg  | 3600 |
| accagctgca gcaggattac tacaagtccg tgcgcacgta catgggtgtg tcgtggatgg  | 3660 |
| tggccaacgc gacgctggcc atggcggtgt cggaagcgta tggcgattcg gaaattgggg  | 3720 |
| ataattttta cttgcggttt atcctgtggg cgggtggcggc cctggcgctg tttagagcgt | 3780 |
| tggggtcgac gacgtttgcg gcgattaatc tgggtgagtgc tctcgtggag ggcagggtca | 3840 |
| ggctgagggt gaatatgaaa ggggttaggt ggattaaagga gaagtggggg gatgcggatg | 3900 |
| tgaagggcaa gtttgagggg ttgggggatc gggcgagggg gttggcgagg cggtgagcta  | 3960 |
| gcaagcttgg acacgctgaa atcaccagtc tctctctaca aatctatctc tctctatttt  | 4020 |
| ctccataata atgtgtgagt agttcccaga taagggaatt agggttccta tagggtttcg  | 4080 |
| ctcatgtgtt gagcatataa gaaaccctta gtatgtatct gtatttgtaa aatacttcta  | 4140 |
| tcaataaaat ttctaattcc taaaaccaa atccagtact aaaatccaga tcatgcatgg   | 4200 |
| tacagcacgc gtctctcagg ccgggttaa ttaagcggcc gcaagcttaa aatcacttat   | 4260 |
| caatttcaaa aacagagggt agccgaatgc taagagctta aaaatggctt cttttgtttc  | 4320 |

# ES 2 701 243 T3

|                                                                     |      |
|---------------------------------------------------------------------|------|
| tttttcttgc aaacggtgga gagaagaggg aaatgaagat tgaccatatac tttttttatt  | 4380 |
| atgttttaac atataatatt aataatttaa tcataattat actttggtga atgtgacagt   | 4440 |
| ggggatatac gtaaagtata taacattata ctttttgcaa gcagttggct ggtctaccca   | 4500 |
| agagtgatca aattttgagc tgccttcaat gagccaattt ttgcctataa tggataaagg   | 4560 |
| cactttgtct agttcaactg ctacacagaat aatgttaaaa tgaaattaaa acaagggtggg | 4620 |
| ctggtcacac caaaaaata aaatattaat gtggtgtttg gttagtcgaa tttatcataa    | 4680 |
| ttttatagaa attgagacta aaaaacatta agagaataaa ttctataaca aagacaattt   | 4740 |
| agtaaaaatg tccttttagg taattttaag tactcttaac caaaataaaa aattcaaatac  | 4800 |
| aaattaacca aataagataa tataacatac ggaacatccc acttataatc ttacatcccc   | 4860 |
| gtaatcttat tatgaaaagt aatcttatat tactcgaatc aaatgctctc acaaactatt   | 4920 |
| atctaataa agaaaaacac ttaattttta taacattttt tttcatatat ttgaaagatt    | 4980 |
| atattttgta tttttacgta aaaatatattg acatagattg agcacctttt taacataatt  | 5040 |
| ccaccataag tcaattatgt agatgagaaa ttggtacaaa caacgtggag ccaaattcca   | 5100 |
| ccaaaccatc tctcatcctc tcctataaaa ggcttgctac acatagacaa caatccacac   | 5160 |
| aaaaatacac ttaaaattct tttctttcta tttggttaac catgtgcgga attggtggtg   | 5220 |
| ctatcgccca aagagacgtt gctgagattt tgtagaggg tctgcgaagg ctagagtata    | 5280 |
| gaggatatga ctccgctggt ctggctgtcg ttgatgtga gggcatatg acaaggctaa     | 5340 |
| gaaggtagg aaagggttcag atgcttgctc aggcagctga ggaacatcca ttgcatggag   | 5400 |
| gtactggtat tgcacatacc aggtgggcta ctcatgggga gccatcagaa gttaatgctc   | 5460 |
| atccacatgt gagtgagcat atcgttgtag ttcacaatgg gataattgaa aaccacgaac   | 5520 |
| cattgagggg agagttaaag gcaagaggat atacttttgt gagtgagact gacactgagg   | 5580 |
| ttattgcaca tttagtgaac tgggaactca aacagggggg cacattgcgt gaggctgtgt   | 5640 |
| taagagctat tcctcaactt agaggtgcat acggtactgt tattatggat tcaagacacc   | 5700 |
| cagatactct ccttgacgct agatcaggta gtcccttggc cataggactt ggaatgggtg   | 5760 |
| aaaattttat cgctagcgac caattggcct tattgccagt tacaagacga tttattttcc   | 5820 |
| ttgaagaggg cgatattgct gagattacta gaaggctgtg gaacatcttt gataagactg   | 5880 |
| gcgctgaggt taaacgtcag gatatcgagt ctaaccttca atacgatgct ggtgataaag   | 5940 |
| gaatttacag gcattatatg caaaaggaaa tttatgaaca accaaatgct atcaaaaaca   | 6000 |
| cacttactgg ccgtatttct catggacagg tcgatttaag cgagcttggc cctaatagcag  | 6060 |
| acgaactgct atcaaaagtt gagcacatac agatactggc atgcggaact agttataatt   | 6120 |
| caggaatggt gtctagatac tggttcgaaa gcttggcagg tataccttgt gatgtagaga   | 6180 |



# ES 2 701 243 T3

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tcgcttctga gtttaggtat agaaagtctg ctgtgcgtag aaattcatta atgattacat  | 6240 |
| tatctcaatc cggagaaaca gcagatacac tggctggatt gaggctttct aaggaactcg  | 6300 |
| gatatctggg ttcacttgct atttgtaatg taccaggttc ctcatgtgtt cgtgaatcag  | 6360 |
| atctagcact tatgacaaat gcaggaactg aaataggtgt ggcaagtacc aaggctttca  | 6420 |
| caacccaact gaccgtactt ttaatgttgg tagcaaaact cagtcgatta aaggggctag  | 6480 |
| atgcatctat cgaacatgat attgttcacg ggcttcaagc tctcccttca agaattgaac  | 6540 |
| aaatgctttc acaagataag agaatagagg cattggctga agatttttcc gacaaacatc  | 6600 |
| acgcattgtt tcttggacgt ggcgatcaat atccaattgc attggaagga gctttgaagt  | 6660 |
| tgaagaaat aagttacatt cacgcagaag catatgcagc tggagaactc aagcatggtc   | 6720 |
| ctttggcact catcgacgct gacatgccg tgatcgtagt ggctcctaataacgaactgc    | 6780 |
| tcgaaaagct taaatcaaatacgaagagg ttcgagctag aggaggtcag ctttacgttt    | 6840 |
| tcgctgaaca agatgctgga ttcgtgtcaa gcgataatat gcatataatt gaaatgcctc  | 6900 |
| acgttgaaga agtgattgca cctatatattt atacagctcc attgcaactt ctagcttacc | 6960 |
| atgttgcaact tattaagga actgatgttg atcagcctag aaacctagca aaatctgtaa  | 7020 |
| cagtcgaata aacgcgtggc gcgccccga tccgcgtttg tgttttctgg gtttctcact   | 7080 |
| taagcgtctg cgttttactt ttgtattggg tttggcgttt agtagtttgc ggtagcgttc  | 7140 |
| ttgttatgtg taattacgct ttttcttctt gcttcagcag tttcggttga aatataaatc  | 7200 |
| gaatcaagtt tcactttatc agcgttgttt taaattttgg cattaaattg gtgaaaattg  | 7260 |
| cttcaatttt gtatctaaat agaagagaca acatgaaatt cgacttttga cctcaaactc  | 7320 |
| tcgaacattt atttcctgat ttcacgatgg atgaggataa cgaaaggcg gttcctatgt   | 7380 |
| ccgggaaagt tcccgtagaa gacaatgagc aaagctactg aaacgcggac acgacgtcgc  | 7440 |
| attggtacgg atatgagtta aaccgactca attcctttat taagacataa accgattttg  | 7500 |
| gttaaagtgt aacagtgagc tgatataaaa ccgaaacaaa ccggtacaag tttgattgag  | 7560 |
| caacttgatg acaaaacttca gaattttggt tattgaatga aaatcatagt ctaatcgtaa | 7620 |
| aaaatgtaca gaagaaaagc tagagcagaa caaagattct atattctggt tccaatttat  | 7680 |
| catcgcttta acgtccctca gatttgatcg gggaattcga tatcattacc ctgttatccc  | 7740 |
| taaagcttat taatgtttgt cgaggagaaa tatgagtcga ggcatggata cactaagttc  | 7800 |
| ccctgaagtg agcatgatct ttgatgctga gatgattccc agagcaagat agtttgtgct  | 7860 |
| gcaagtgaca caattgtaat gaaaccacca ctcaacgaat ttacttgtgg ctttgacatg  | 7920 |
| tcgtgtgctc tgtttgtatt tgtgagtgcc ggttggtaat tatttttgtt aatgtgattt  | 7980 |
| taaaacctct tatgtaaaata gttactttat ctattgaagt gtgttcttgt ggtctatagt | 8040 |
| ttctcaaagg gaaattaaaa tgttgacatc ccatttaciaa ttgataactt ggtatacaca | 8100 |

# ES 2 701 243 T3

|                                                                   |      |
|-------------------------------------------------------------------|------|
| aactttgtaa atttggtgat atttatggtc gaaagaaggc aatacccatt gtatgttcca | 8160 |
| atatcaatat caatacgata acttgataat actaacatat gattgtcatt gtttttccag | 8220 |
| tatcaatata cattaagcta ctacaaaatt agtataaatc actatattat aaatcttttt | 8280 |
| cggttgtaac ttgtaattcg tgggttttta aaataaaagc atgtgaaaat tttcaaataa | 8340 |
| tgtgatggcg caattttatt ttccgagttc caaaatattg ccgcttcatt accctaattt | 8400 |
| gtggcgccac atgtaaaaca aaagacgatt cttagtggct atcactgcc aacgcggat   | 8460 |
| cactaatatg aaccgtcgat taaaacagat cgacggttta tacatcattt tattgtacac | 8520 |
| acggatcgat atctcagccg ttagatttaa tatgcgatct gattgctcaa aaaatagact | 8580 |
| ctccgtcttt gcctataaaa acaatttcac atctttctca cccaaatcta ctcttaaccg | 8640 |
| ttcttcttct tctacagaca tcaatttctc tcgactctag aggatccaag cttatcgatt | 8700 |
| tcgaaccct caggcgaaga acaggtatga tttgtttgta attagatcag gggtttaggt  | 8760 |
| ctttccatta ctttttaatg ttttttctgt tactgtctcc gcgatctgat tttacgacaa | 8820 |
| tagagtttcg ggttttgtcc cattccagtt tgaaaataaa ggtccgtctt ttaagtttgc | 8880 |
| tggatcgata aacctgtgaa gattgagtct agtcgattta ttggatgatc cattcttcat | 8940 |
| cgtttttttc ttgcttcgaa gttctgtata accagatttg tctgtgtgcg attgtcatta | 9000 |
| cctagccgtg tatcgagaac tagggttttc gagtcaattt tgccctttt gggtatatct  | 9060 |
| ggttcgataa cgattcatct ggattagggt ttttaagtgt gacgtttagt attccaattt | 9120 |
| cttcaaaatt tagttatgga taatgaaaat cccaattga ctgttcaatt tcttgttaaa  | 9180 |
| tgccgagatc acaatggctt cgatctctc ctcagtcgag accgttagcc ggaccgcccc  | 9240 |
| tgctcaggcc aacatggtgg ctccgttcac cggccttaag tccaacgccg ccttccccac | 9300 |
| caccaagaag gctaacgact tctccaccct tcccagcaac ggtggaagag ttcaatgtat | 9360 |
| gcaggtgtgg ccggcctacg gcaacaagaa gttcgagacg ctgtcgtacc tgccgcccgt | 9420 |
| gtctatggcg cccaccgtga tgatggcctc gtcggccacc gccgtcgctc cgttccaggg | 9480 |
| gctcaagtcc accgccagcc tcccgcgtcg ccgcccgtcc tccagaagcc tcggcaacgt | 9540 |
| cagcaacggc ggaaggatcc ggtgcatggc cggcgccgag gagatcgtgc tgcagcccat | 9600 |
| caaggagatc tccggcaccg tcaagctgcc ggggtccaag tcgctttcca accggtacct | 9660 |
| cctactcgcc gcctgtccg aggggacaac agtggttgat aacctgctga acagtgagga  | 9720 |
| tgtccactac atgctcgggg ccttgaggac tcttggtctc tctgtcgaag cggacaaagc | 9780 |
| tgccaaaaga gctgtagttg ttggctgtgg tggaaagttc ccagttgagg atgctaaaga | 9840 |
| ggaagtgcag ctcttcttgg ggaatgctgg aatcgcaatg cggtccttga cagcagctgt | 9900 |
| tactgctgct ggtggaaatg caacttacgt gcttgatgga gtaccaagaa tgagggagag | 9960 |

# ES 2 701 243 T3

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| accattggc gacttggtg tcggattgaa gcagcttggg gcagatgttg attgtttcct    | 10020 |
| tggcactgac tgcccacctg ttcgtgtcaa tggaatcgga gggctacctg gtggcaaggt  | 10080 |
| caagctgtct ggctccatca gcagtcagta cttgagtgcc ttgctgatgg ctgctccttt  | 10140 |
| ggctcttggg gatgtggaga ttgaaatcat tgataaatta atctccattc cgtacgtcga  | 10200 |
| aatgacattg agattgatgg agcgttttgg tgtgaaagca gagcattctg atagctggga  | 10260 |
| cagattctac attaaggag gtcaaaaata caagtccct aaaaatgcct atgttgaagg    | 10320 |
| tgatgcctca agcgcaagct atttcttggc tgggtgctga attactggag ggactgtgac  | 10380 |
| tgtggaaggt tgtggcacca ccagtttgca gggatgatgg aagtttgctg aggtactgga  | 10440 |
| gatgatggga gcgaagggtta catggaccga gactagcgta actgttactg gccaccgcg  | 10500 |
| ggagccattt gggaggaaac acctcaaggc gattgatgtc aacatgaaca agatgcctga  | 10560 |
| tgtcgccatg actcttgctg tgggtgccct ctttgccgat ggcccgacag ccatcagaga  | 10620 |
| cgtggcttcc tggagagtaa aggagaccga gaggatgggt gcgatccgga cggagctaac  | 10680 |
| caagctggga gcatctgttg aggaagggcc ggactactgc atcatcacgc cgcggagaa   | 10740 |
| gctgaacgtg acggcgatcg acacgtacga cgaccacagg atggcgatgg ctttctccct  | 10800 |
| tgccgcctgt gccgaggtcc ccgtcaccat ccgggaccct ggggtgcaccc ggaagacctt | 10860 |
| ccccgactac ttcgatgtgc tgagcacttt cgtcaagaat taagctctag aactagtga   | 10920 |
| tccccgatc cgcgtttgtg ttttctgggt ttctcactta agcgtctgcg ttttactttt   | 10980 |
| gtattgggtt tggcgtttag tagtttgcgg tagcgttctt gttatgtgta attacgcttt  | 11040 |
| ttcttcttgc ttcagcagtt tcggttgaaa tataaatcga atcaagtttc actttatcag  | 11100 |
| cgttgtttta aattttggca ttaaattgggt gaaaattgct tcaattttgt atctaaatag | 11160 |
| aagagacaac atgaaattcg acttttgacc tcaaatcttc gaacatttat ttcctgattt  | 11220 |
| cacgatggat gaggataacg aaaggcggt tcctatgtcc gggaaagtcc ccgtagaaga   | 11280 |
| caatgagcaa agctactgaa acgcggacac gacgtcgcat tggtaggat atgagttaaa   | 11340 |
| ccgactcaat tcctttatta agacataaac cgattttgggt taaagtgtaa cagtgcgtg  | 11400 |
| atataaaacc gaaacaaacc ggtacaagtt tgattgagca acttgatgac aaacttcaga  | 11460 |
| attttggtta ttgaatgaaa atcatagtct aatcgtaaaa aatgtacaga agaaaagcta  | 11520 |
| gagcagaaca aagattctat attctgggtc caatttatca tcgctttaac gtccctcaga  | 11580 |
| tttgatcggg aaaccaaaac gtcgtgagac agtttggtta actataacgg tcctaaggta  | 11640 |
| gcgatcgagg cattacggca ttacggcact cgcgagggtc cgaattcgag catggagcca  | 11700 |
| tttacaattg aatatatcct gccgccgctg ccgctttgca ccgggtggag cttgcatgtt  | 11760 |
| ggtttctacg cagaactgag ccggttaggc agataatttc cattgagaac tgagccatgt  | 11820 |
| gcaccttccc cccaacacgg tgagcgacgg ggcaacggag tgatccacat gggactttta  | 11880 |

# ES 2 701 243 T3

|             |            |            |            |            |            |       |
|-------------|------------|------------|------------|------------|------------|-------|
| aacatcatcc  | gtcggatggc | gttgcgagag | aagcagtcga | tccgtgagat | cagccgacgc | 11940 |
| accgggcagg  | cgcgcaacac | gatcgcaaag | tatttgaacg | caggtacaat | cgagccgacg | 12000 |
| ttcacggtac  | cggaacgacc | aagcaagcta | gcttagtaaa | gccctcgcta | gattttaatg | 12060 |
| cggatgttgc  | gattacttcg | ccaactattg | cgataacaag | aaaaagccag | cctttcatga | 12120 |
| tatatctccc  | aatttgtgta | gggcttatta | tgcacgctta | aaaataataa | aagcagactt | 12180 |
| gacctgatag  | tttggctgtg | agcaattatg | tgcttagtgc | atctaacgct | tgagttaagc | 12240 |
| cgcgccgcga  | agcggcgctg | gcttgaacga | attgttagac | attatttgcc | gactaccttg | 12300 |
| gtgatctcgc  | ctttcacgta | gtggacaaat | tcttccaact | gatctgcgcg | cgaggccaag | 12360 |
| cgatcttctt  | cttgtccaag | ataagcctgt | ctagcttcaa | gtatgacggg | ctgatactgg | 12420 |
| gccggcaggc  | gctccattgc | ccagtcggca | gcgacatcct | tcggcgcgat | tttgccggtt | 12480 |
| actgcgctgt  | accaaagcgc | ggacaacgta | agcactacat | ttcgctcatc | gccagcccag | 12540 |
| tcgggcggcg  | agttccatag | cgttaagggt | tcatttagcg | cctcaaatag | atcctgttca | 12600 |
| ggaaccggt   | caaagagttc | ctccgccgct | ggacctacca | aggcaacgct | atgttctctt | 12660 |
| gcttttgtca  | gcaagatagc | cagatcaatg | tcgatcgtgg | ctggctcgaa | gatacctgca | 12720 |
| agaatgtcat  | tcgcgtgcc  | ttctccaaat | tgcagttcgc | gcttagctgg | ataacgccac | 12780 |
| ggaatgatgt  | cgctcgtcac | aacaatggtg | acttctacag | cgcggagaat | ctcgctctct | 12840 |
| ccagggaag   | ccgaagtttc | caaaaggctg | ttgatcaaag | ctcgccgctg | tgtttcatca | 12900 |
| agccttacgg  | tcaccgtaac | cagcaaatca | atatcactgt | gtggcttcag | gccgccatcc | 12960 |
| actgcggagc  | cgtacaaatg | tacggccagc | aacgtcgggt | cgagatggcg | ctcgatgacg | 13020 |
| ccaactacct  | ctgatagtgt | agtcgatact | tcggcgatca | ccgcttcctt | catgatgttt | 13080 |
| aactttgttt  | tagggcgact | gccctgctgc | gtaacatcgt | tgctgctcca | taacatcaaa | 13140 |
| catcgaccca  | cggcgtaacg | cgcttgctgc | ttggatgccc | gaggcataga | ctgtaccca  | 13200 |
| aaaaaacagt  | cataacaagc | catgaaaacc | gccactgcgc | cgttaccacc | gctgcgttcg | 13260 |
| gtcaagggtc  | tggaaccagt | gcgtgagcgc | atacgctact | tgcatcacag | cttacgaacc | 13320 |
| gaacaggctt  | atgtccactg | ggttcgtgcc | ttcatccggt | tccacgggtg | gcgtcacccg | 13380 |
| gcaaccttgg  | gcagcagcga | agtcgaggca | tttctgtcct | ggctggcgaa | cgagcgcaag | 13440 |
| gtttcggtct  | ccacgcatcg | tcaggcattg | gcggccttgc | tgctcttcta | cggcaagtgc | 13500 |
| tgtgcacgga  | tctgccctgg | cttcaggaga | tcggaagacc | tcggccgtcc | ggcgcttgc  | 13560 |
| cgggtggtgct | gaccccggtg | gaagtctcta | gagctctaga | gggttcgcat | cctcggtttt | 13620 |
| ctggaaggcg  | agcatcggtt | gttcgcccag | cttctgtatg | gaacgggcat | gcggatcagt | 13680 |
| gagggtttgc  | aactgcgggt | caaggatctg | gatttcgata | acggcacgat | catcgtgcgg | 13740 |

# ES 2 701 243 T3

|                                                                   |       |
|-------------------------------------------------------------------|-------|
| gagggcaagg gctccaagga tggggccttg atgttaccog agagcttggc acccagcctg | 13800 |
| cgcgagcagg gatcgatacc gtgcggtgc atgaaatcct ggccggtttg tctgatgcca  | 13860 |
| agctggcggc ctggccggcc agcttgccg ctgaagaaac cgagcgccgc cgtctaaaaa  | 13920 |
| ggtgatgtgt atttgagtaa aacagcttgc gtcacgcggt cgctgcgtat atgatgcgat | 13980 |
| gagtaataa acaaatagc aaggggaacg catgaagggt atcgctgtac ttaaccagaa   | 14040 |
| aggcgggtca ggcaagacga ccatcgcaac ccatctagcc cgcgccctgc aactcgccgg | 14100 |
| ggccgatgtt ctgttagtcg attccgatcc ccagggcagt gcccgcgatt gggcgccgt  | 14160 |
| gcggaagat caaccgctaa ccgttgctgg catcgaccgc ccgacgattg acccgacgt   | 14220 |
| gaaggccatc ggccggcgcg acttcgtagt gatcgacgga gcgcccagg cgcgggactt  | 14280 |
| ggctgtgtcc gcgatcaagg cagccgactt cgtgctgatt ccggtgcagc caagccctta | 14340 |
| cgacatatgg gccaccgccc acctggtgga gctgggtaag cagcgcatth aggtcacgga | 14400 |
| tggaggcta caagcgccct ttgtcgtgtc gggggcgatc aaaggcacgc gcatcgccgg  | 14460 |
| tgaggttgcc gagcgctgg ccgggtacga gctgccatt cttgagtccc gtatcacgca   | 14520 |
| gcgctgagc taccaggca ctgcccgcc cggcacaaac gttcttgaat cagaaccgga    | 14580 |
| ggcgacgct gcccgcgagg tccaggcgct ggccgctgaa attaaatcaa aactcatttg  | 14640 |
| agttaatgag gtaaagagaa aatgagcaaa agcacaaaac cgctaagtgc cggccgtccg | 14700 |
| agcgacgca gcagcaaggc tgcaacgttg gccagcctgg cagacacgcc agccatgaag  | 14760 |
| cgggtcaact ttcagttgcc ggcgaggat cacaccaagc tgaagatgta cgcggtacgc  | 14820 |
| caaggcaaga ccattaccga gctgctatct gaatacatcg cgagctacc agagtaaatg  | 14880 |
| agcaaatgaa taaatgagta gatgaatth agcggtctaa ggaggcgga tggaaatca    | 14940 |
| agaacaacca ggcaccgacg ccgtggaatg ccccatgtgt ggaggaaagg gcggttgcc  | 15000 |
| aggcgtaagc ggctgggttg tctgccggcc ctgcaatggc actggaaccc ccaagccga  | 15060 |
| ggaatcggcg tgacggtcgc aaaccatccg gcccggtaca aatcggcgcg gcgctgggtg | 15120 |
| atgacctggt ggagaagttg aaggccgcgc aggcgcgcca gcggcaacgc atcgaggcag | 15180 |
| aagcacgccc cgggtgaatcg tggcaagcgg ccgctgatcg aatccgcaa gaatcccgcc | 15240 |
| aaccgcccgc agccggtgcg ccgtcgatta ggaagccgcc caagggcgac gagcaaccag | 15300 |
| attttttcgt tccgatgctc tatgacgtgg gcaccgcca tagtcgcagc atcatggacg  | 15360 |
| tggccgtttt ccgtctgtcg aagcgtgacc gacgagctgg cgaggtgac cgtacgagc   | 15420 |
| ttccagacgg gcacgtagag gtttccgcag ggccggccgg catggccagt gtgtgggatt | 15480 |
| acgacctggt actgatggcg gtttcccatc taaccgaatc catgaaccga taccgggaag | 15540 |
| ggaagggaga caagcccgcc cgctgttcc gtccacacgt tgggacgta ctcaagttct   | 15600 |
| gccggcgagc cgatggcgga aagcagaaag acgacctggt agaaacctgc attcggttaa | 15660 |

# ES 2 701 243 T3

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| acaccacgca cgttgccatg cagcgtacga agaaggccaa gaacggccgc ctggtgacgg  | 15720 |
| tatccgaggg tgaagccttg attagccgct acaagatcgt aaagagcgaa accgggcggc  | 15780 |
| cggagtacat cgagatcgag ctagctgatt ggatgtaccg cgagatcaca gaaggcaaga  | 15840 |
| acccggacgt gctgacgggt caccocgatt actttttgat cgatcccggc atcggccggt  | 15900 |
| ttctctaccg cctggcacgc cgcgccgag gcaaggcaga agccagatgg ttgttcaaga   | 15960 |
| cgatctacga acgcagtggc agcgccggag agttcaagaa gttctgtttc accgtgcgca  | 16020 |
| agctgatcgg gtcaaattgac ctgccggagt acgatttgaa ggaggaggcg gggcaggctg | 16080 |
| gcccgatcct agtcatgcgc taccgcaacc tgatcgaggg cgaagcatcc gccggttcct  | 16140 |
| aatgtacgga gcagatgcta gggcaaattg ccctagcagg ggaaaaaggc cgaaaaggtc  | 16200 |
| tctttcctgt ggatagcacg tacattggga acccaaagcc gtacattggg aaccggaacc  | 16260 |
| cgtacattgg gaacccaaag ccgtacattg ggaaccggtc acacatgtaa gtgactgata  | 16320 |
| taaaagagaa aaaaggcgat tttccgcct aaaactcttt aaaacttatt aaaactctta   | 16380 |
| aaaccgcctt ggcctgtgca taactgtctg gccagcgcac agccgaagag ctgcaaaaag  | 16440 |
| cgcctaccct tcggtcgctg cgctccctac gcccgcgcgc ttcgcgtcgg cctatcgcg   | 16500 |
| ccgctggccg ctcaaaaatg gctggcctac ggccaggcaa tctaccaggg cgcggacaag  | 16560 |
| ccgcgcgctc gccactcgac cgcggcgcc cacatcaagg caccctgcct cgcgcgtttc   | 16620 |
| ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg agacgggtcac agcttgtctg | 16680 |
| taagcggatg ccgggagcag acaagcccgt cagggcgcggt cagcgggtgt tggcgggtgt | 16740 |
| cggggcgcgag ccatgaccca gtacagtagc gatagcggag tgtatactgg cttaactatg | 16800 |
| cggcatcaga gcagattgta ctgagagtgc accatatgcg gtgtgaaata ccgcacagat  | 16860 |
| gcgtaaggag aaaataccgc atcaggcgct cttccgcttc ctgcgtcact gactcgctgc  | 16920 |
| gctcggctgt tcggctgcgg cgagcgggat cagctcactc aaaggcggtg atacggttat  | 16980 |
| ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca  | 17040 |
| ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc  | 17100 |
| atcacaaaaa tcgacgctca agtcagagggt ggcgaaaccc gacaggacta taaagatacc | 17160 |
| aggcgtttcc ccttggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg  | 17220 |
| gatacctgtc cgcttttctc cttcgggaa gcgtggcgct ttctcatagc tcacgctgta   | 17280 |
| ggtatctcag ttccgtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg  | 17340 |
| ttcagccccg ccgctgcgcc ttatccggta actatcgtct tgagtcacac ccggttaagc  | 17400 |
| acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag  | 17460 |
| gcggtgctac agagttcttg aagtgggtggc ctaactacgg ctacactaga aggacagtat | 17520 |

# ES 2 701 243 T3

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat  | 17580 |
| ccggcaaaca aaccaccgct ggtagcgggtg gtttttttgt ttgcaagcag cagattacgc | 17640 |
| gcagaaaaaa aggatctcaa gaagatccgg aaaacgcaag cgcaaagaga aagcaggtag  | 17700 |
| cttgcagtgg gcttacatgg cgatagctag actgggcggt tttatggaca gcaagcgaac  | 17760 |
| cggaattgcc agattcgaag ctcggtcccg tgggtgttct gtcgtctcgt tgtacaacga  | 17820 |
| aatccattcc cattccgcgc tcaagatggc ttcccctcgg cagttcatca gggctaaatc  | 17880 |
| aatctagccg acttgtccgg tgaaatgggc tgcactcaa cagaaacaat caaacaaca    | 17940 |
| tacacagcga cttattcaca cgcgaca                                      | 17967 |

# REIVINDICACIONES

1. Un ácido nucleico que comprende una secuencia nucleotídica seleccionada del grupo de
  - (a) al menos 700 nucleótidos consecutivos de la secuencia nucleotídica indicada como SEQ ID NO: 1;
  - (b) la secuencia nucleotídica de SEQ ID No. 1 desde el nucleótido en la posición 1 hasta el nucleótido en la posición 922;
  - (c) una secuencia nucleotídica que tiene al menos 95% de identidad de secuencia con la secuencia de ácido nucleico de (a) o (b); yen donde el ácido nucleico tiene actividad promotora preferente en fibras en algodón entre 15 y 35 días después de la antesis (dpa) y no se expresa en fibras de 10 dpa, y en donde dicho ácido nucleico que tiene actividad promotora preferente en fibras no tiene más de 959 nucleótidos de longitud.
2. Un gen recombinante que comprende
  - (a) el ácido nucleico según la reivindicación 1 ligado operativamente a
  - (b) un ácido nucleico heterólogo que codifica un producto de expresión de interés, y opcionalmente
  - (c) una región de terminación de la transcripción y poliadenilación.
3. El gen recombinante según la reivindicación 2, en el que dicho producto de expresión de interés es:
  - (a) una proteína o una molécula de ARN capaz de modular la expresión de un gen endógeno a dicha planta; o
  - (b) un gen indicador o un gen específico en fibras.
4. El gen recombinante según la reivindicación 3, en el que dicha molécula de ARN comprende una primera y una segunda región de ARN en donde
  - (a) dicha primera región de ARN comprende una secuencia nucleotídica de al menos 19 nucleótidos consecutivos que tiene al menos aproximadamente 94% de identidad de secuencia con la secuencia nucleotídica de dicho gen endógeno;
  - (b) dicha segunda región de ARN comprende una secuencia nucleotídica complementaria a dichos 19 nucleótidos consecutivos de dicha primera región de ARN; y
  - (c) dichas primera y segunda región de ARN son capaces de apareamiento de bases para formar una molécula de ARN bicatenaria entre al menos dichos 19 nucleótidos consecutivos de dichas primera y segunda región.
5. El gen recombinante según la reivindicación 2, en el que dicho ácido nucleico heterólogo que codifica un producto de expresión de interés codifica quitina sintasa o glutamina:fructosa-6-fosfato amidotransferasa.
6. Una célula de planta de algodón transgénica que comprende un gen recombinante según una cualquiera de las reivindicaciones 2 a 5.
7. Una planta de algodón transgénica que comprende el gen recombinante según una cualquiera de las reivindicaciones 2 a 5 integrado establemente en su genoma o que consiste en las células de planta de algodón transgénica según la reivindicación 6.
8. La planta transgénica según la reivindicación 7, que es *G. hirsutum*, *G. barbadense*, *G. arboreum* o *G. herbaceum*.
9. Una semilla generada a partir de una planta transgénica según la reivindicación 7, en donde la semilla comprende un gen recombinante según la reivindicación 2.
10. Fibras de algodón obtenibles de la planta transgénica según la reivindicación 7.
11. Un método para producir una planta de algodón transgénica, que comprende
  - (a) proporcionar un gen recombinante según la reivindicación 2; y
  - (b) introducir dicho gen recombinante en una planta.



12. Un método para efectuar una expresión preferente en fibras de un producto en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa, que comprende

- 5 (a) introducir el gen recombinante según la reivindicación 2 en el genoma de una planta de algodón; o  
(b) proporcionar la planta transgénica según la reivindicación 7.

13. Un método para alterar las propiedades de las fibras en una planta de algodón, que comprende

- 10 (a) introducir el gen recombinante según la reivindicación 2 en el genoma de una planta de algodón; o  
(b) proporcionar la planta transgénica según la reivindicación 7.

- 15 14. Uso del gen recombinante según la reivindicación 2 o la planta transgénica según la reivindicación 7 para la expresión preferente en fibras de un producto en algodón entre 15 y 35 días después de la antesis (dpa) pero no en fibras de 10 dpa, para alterar las propiedades de las fibras en algodón o para incrementar el rendimiento del algodón.

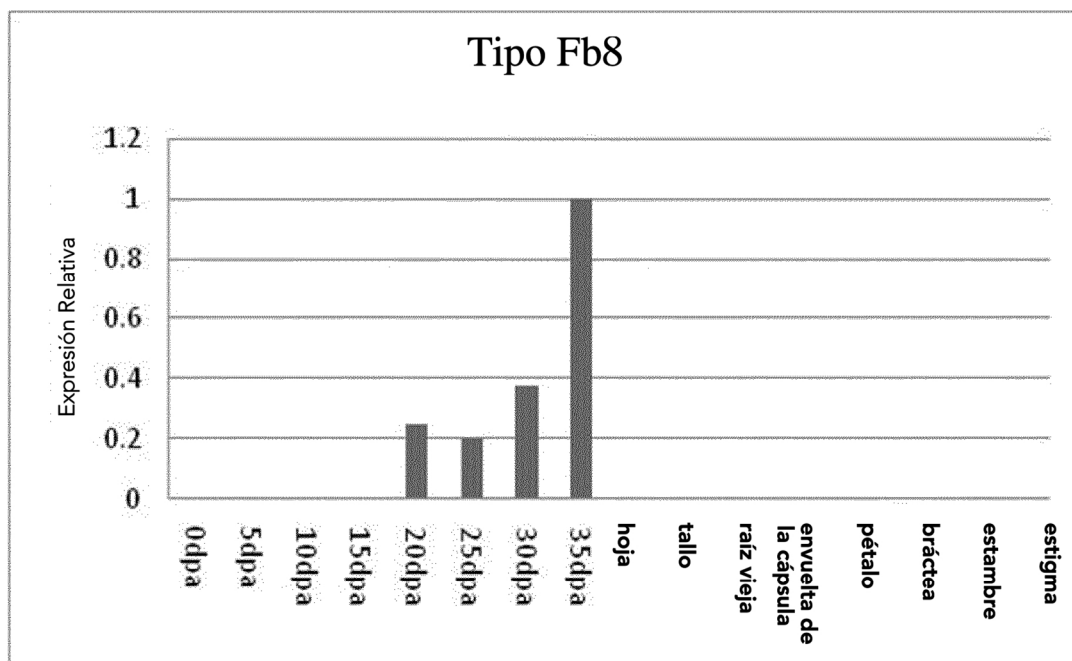


Figura 1

|            |                                                       |     |
|------------|-------------------------------------------------------|-----|
| 150F2      | MAHSLURHPFFLFQLLLTISVSLVIGSHTVSSAARHLFHTQTTSSELPQLAS  | 50  |
| 8K12       | MAHNFRHPFFLFQLLLTITVSLMIGSHTVSSAARHLFHTQTTSSELPQLAS   | 50  |
| tipo cFB8* | MAHNFRHPFFLFQLLLTITVSLMIGSHTVSSAARHLFHTQTTSSELPQLAS   | 50  |
| 150F2      | KYEKHEES EYEKPEYKQPKYHEEYSKLEKPEMOKKEEKQKPKCKHHEEYHES | 100 |
| 8K12       | KYEKHEES EYEKPEYKQPKYHEEYKHEKPEMYKEEKQKPKCKHHEEYHES   | 95  |
| tipo cFB8* | KYEKHEES EYEKPEYKQPKYHEEYKHEKPEMYKEEKQKPKCKHHEEYHES   | 95  |
| 150F2      | HESKEQKEYEKENL EFPKWEKPKHEHEKHEVEYPKIPEYKMKQDEGKEHKH  | 150 |
| 8K12       | RESKEHEEYDKEKPD FPKWEKPKHEHEKHEVEYPKIPEYKQKQDEDEKHEKN | 145 |
| tipo cFB8* | RESKEHEEYDKEKPD FPKWEKPKHEHEKHEVEYPKIPEYKDKQDEDEKHEKN | 145 |
| 150F2      | EYHESRESKEHEEYKKEKPEFPKLEKPKHEHEKHEVEYHETPEYKKKQDE    | 200 |
| 8K12       | EYHESRESKEHEEYKKEKPEFPKREKPKHEHEKHEVEYSEIPEYKEKEDK    | 195 |
| tipo cFB8* | EYHESRESKEHEEYKKEKPEFPKREKPKHEHEKHEVEYSEIPEYKEREDK    | 195 |
| 150F2      | GNCHK HEPKHEHEEKKPL ER KA EY                          | 225 |
| 8K12       | SKKHKDEECQESHESKEHEEYEKEKPD FPKWEKPKGHEKHKA EYPKIPEQ  | 245 |
| tipo cFB8* | SKKHKDEECQESHESKEHEEYEKEKPD FPKWEKPKGHEKHKA EYPKIPEQ  | 245 |
| 150F2      | LTWAK MPTWSKSMETJPGSATKP                              | 248 |
| 8K12       | KEKLDDEKHEHKHEFPKHEKEEEKKPEKGIVP                      | 276 |
| tipo cFB8* | KEKLDDEKHEHKHEFPKHEKEEEKKPEKGIVP                      | 276 |

Figura 2

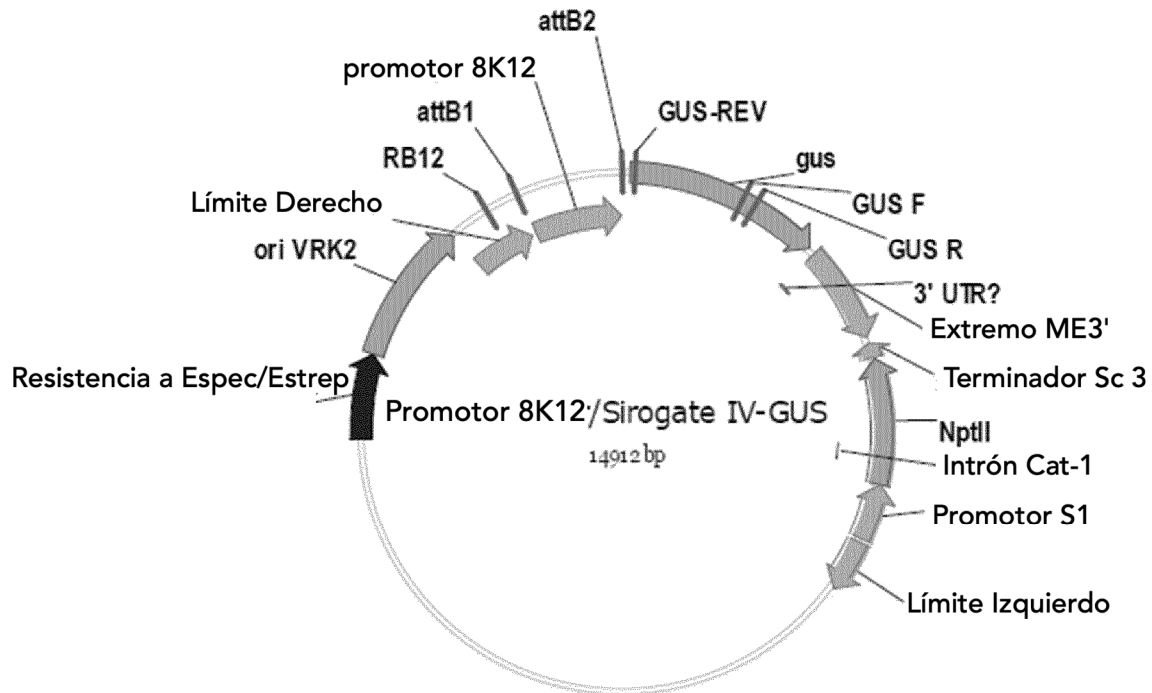


Figura 3

aagcttaaaatcacttatcaatttcaaaaacagaggtagccgaatgctaagagcttaaaaatggcttctttgtttct  
 tttcttgcaaacggtggagagaagagggaatgaagattgaccatatctttttattatgttttaacataataatta  
 ataatttaatacataattatactttggtgaatgtgacagtggggatatacgtaaagtataacattatacttttgcaa  
 gcagttggctggtctaccaagagtgatcaaattttgagctgccttcaatgagccaattttgcctataatggataaa  
 ggcactttgtctagttcaactgctcacagaataatgttaaaatgaaattaaaacaaggtgggctggtcacaccaa  
 aaaataaaatattaatgtggtgtttggttagtcgaatttatcataattttatagaaattgagactaaaaaacattaag  
 agaataaattctataacaaagacaatttagtaaaatgtccttttaggtaattttaagtactcttaaccaaataaaa  
 aattcaaatcaaattaaccaataagataatataacatacgaacatcccacttataatcttacatccccgtaatctt  
 attatgaaaagtaattcttatattactgaatcaaagtcttcacaaactattatctaaataaagaaaaacacttaatt  
 ttataacattttttcatatatttgaaagattatattgtatttttacgtaaaaatatttgacatagattgagcaccttt  
 ttaacataattccaccataagtcaattatgtagatgagaaattggtacaaacaacgtggagccaaatcccacaaa  
 ccatctctcatcctctctataaaaggcttgctacacatagacaacaatccacacaaaaatacacttaaaattctttt  
 ctttctatttggttaacc

Figura 4



Figura 5

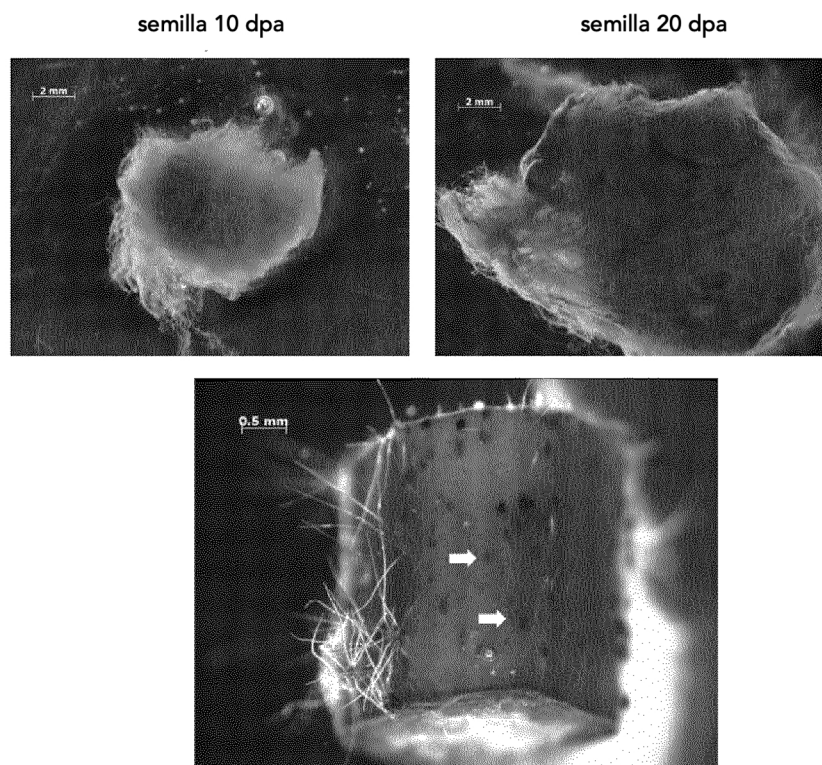


Figura 6

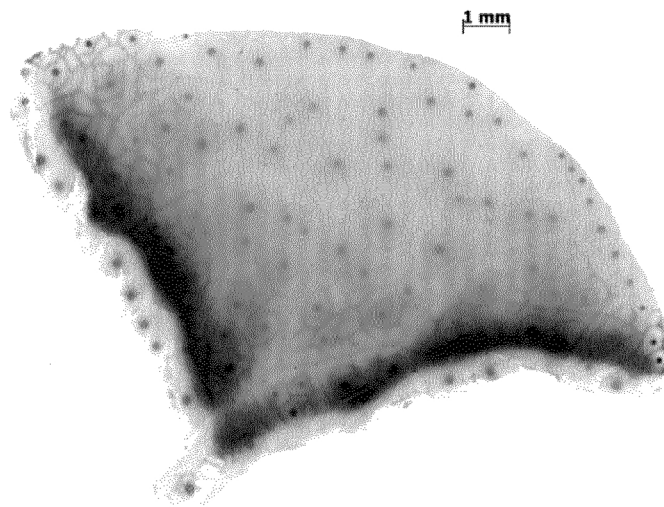


Figura 7

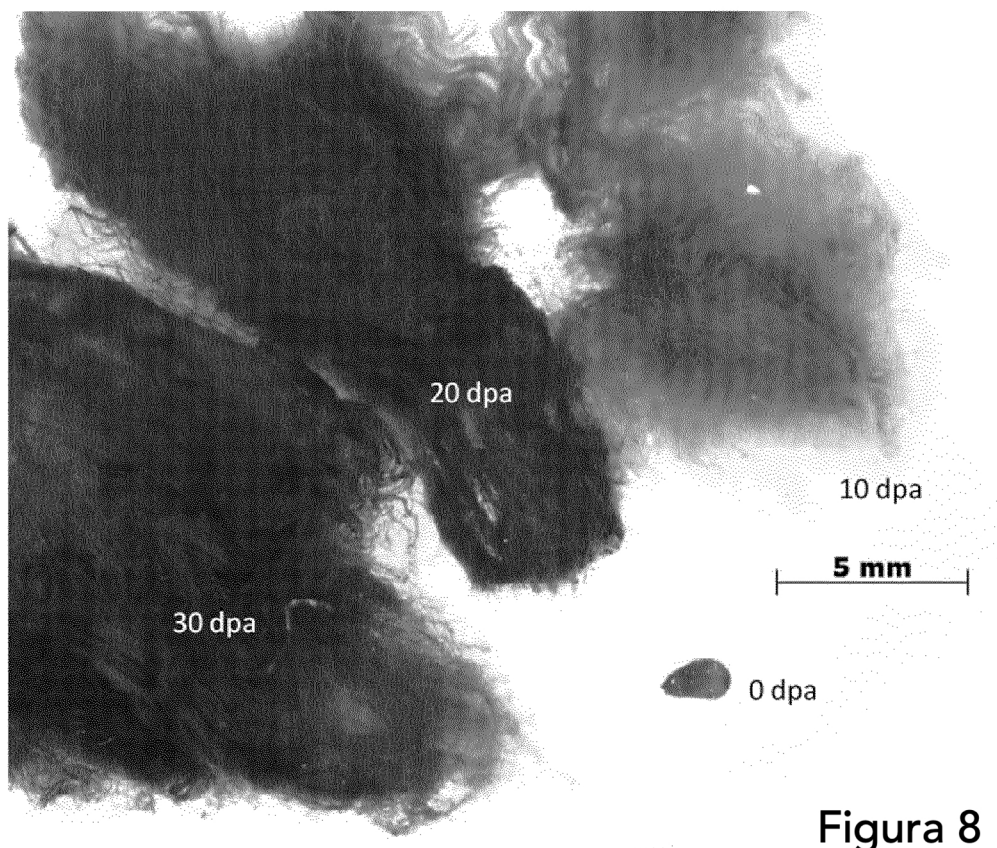


Figura 8

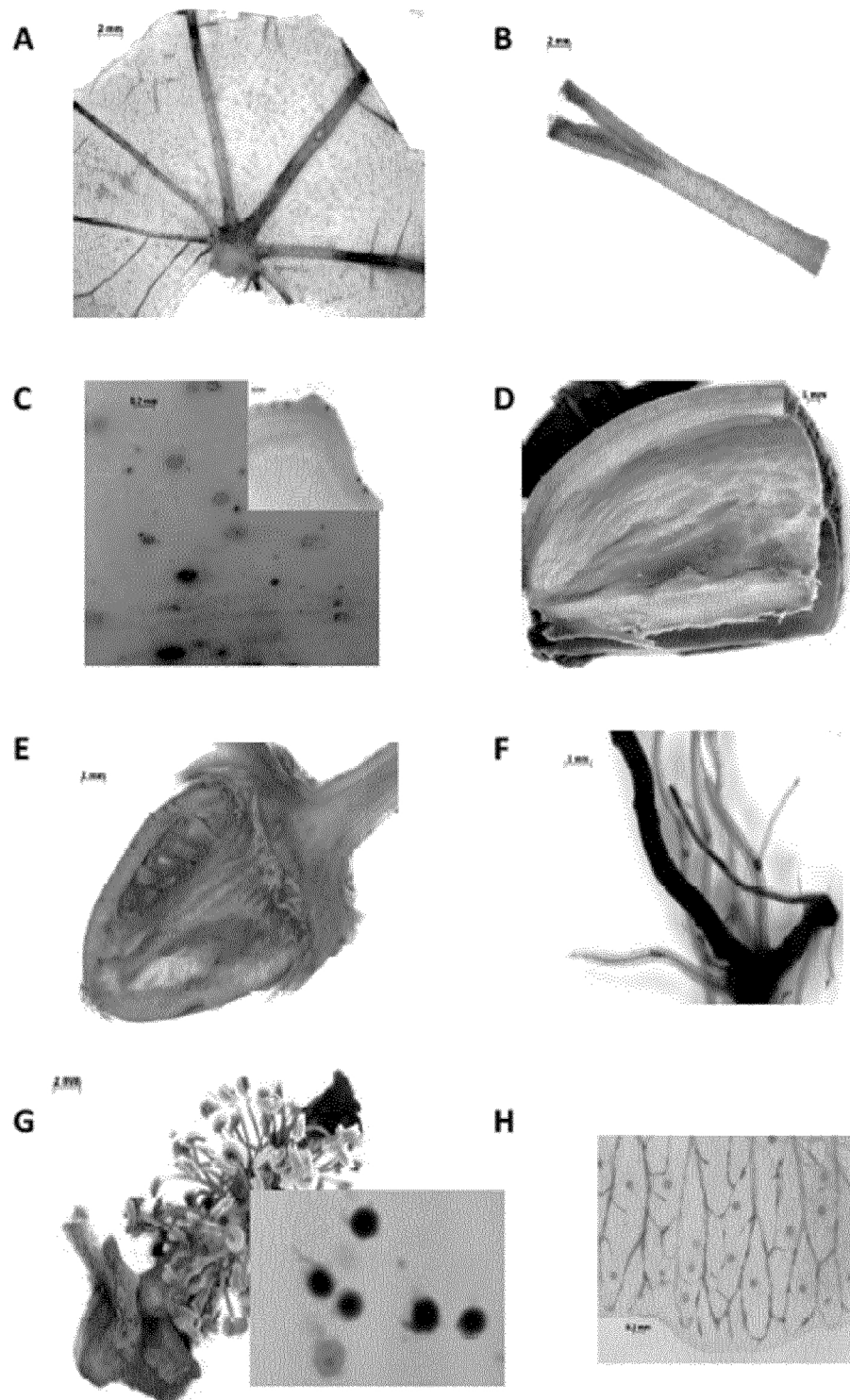


Figura 9

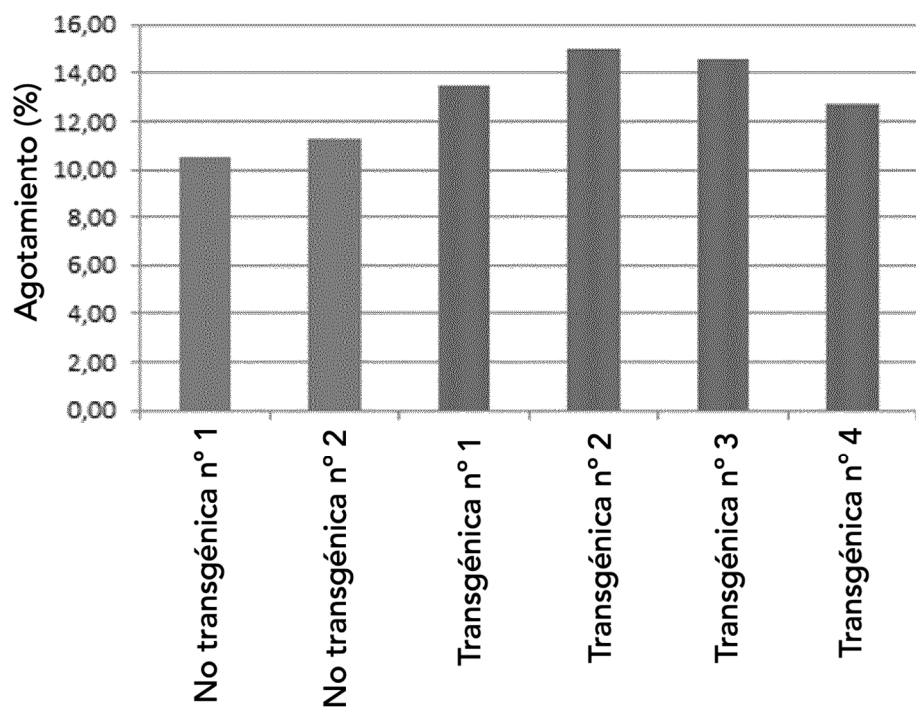


Figura 10