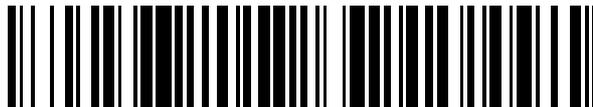


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 548 763**

51 Int. Cl.:

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.01.2008 E 08701076 (5)**

97 Fecha y número de publicación de la concesión europea: **22.07.2015 EP 2099917**

54 Título: **Expresión diferencial de alelos específicos de subgenoma en algodón y usos de los mismos**

30 Prioridad:

11.01.2007 EP 07000550
11.01.2007 US 884564 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.10.2015

73 Titular/es:

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

72 Inventor/es:

ENGELN, STEVEN y
ARIOLI, ANTONIO

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 548 763 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

Expresión diferencial de alelos específicos de subgenoma en algodón y usos de los mismos

5 **Campo de la invención**

La presente invención se refiere al campo de la agricultura, más específicamente al uso de técnicas de biología molecular para modificar plantas productoras de fibras, en particular, plantas de algodón, y/o acelerar el cultivo de tales plantas que contienen fibras. Se proporcionan promotores vegetales con un perfil de expresión preferente de fibra. Además, los promotores están sometidos a regulación temporal.

Técnica anterior

El algodón proporciona gran parte de la fibra de alta calidad para la industria textil. La modificación de las características de las fibras de algodón para adaptarse mejor a los requisitos de la industria es un esfuerzo considerable en el cultivo por procedimientos clásicos o mediante la modificación genética del genoma de las plantas de algodón.

Aproximadamente el 90 % del algodón que se cultiva en el mundo es *Gossypium hirsutum* L., mientras que *Gossypium barbadense* supone aproximadamente el 8 %. Como en la mayoría de las plantas con flor, se cree que los genomas del algodón han incurrido en uno o más eventos de poliploidización y han evolucionado por la unión de varios genomas divergentes en un núcleo común. El mercado del algodón está dominado por las formas mejoradas de dos especies tetraploides "AD", *Gossypium hirsutum* L. y *Gossypium barbadense* L. Se cree que los algodones tetraploides se formaron hace aproximadamente 1-2 millones de años, en el Nuevo Mundo, por hibridación entre un taxón materno de genoma "A" del Viejo Mundo parecido a *Gossypium herbaceum* y un taxón paterno de genoma "D" del nuevo mundo parecido a *Gossypium raimondii* o *Gossypium gossypoides*. Los taxones silvestres de *Gossypium* tetraploide AD y diploide de genoma A producen fibras hilables. En Asia se sigue cultivando y desarrollando una especie diploide de genoma A, *Gossypium arboreum*. Su pariente cercano y posible progenitor, la especie diploide de genoma A *G. herbaceum* también produce fibra hilable. Aunque las semillas de los diploides de genoma D son puberales, ninguna produce fibras hilables. No se han domesticado los taxones de ninguno de los demás genomas diploides de *Gossypium* reconocidos (B, C, E, F y G). La selección dirigida intensa llevada a cabo por el ser humano ha producido se forma consistente algodones tetraploides AD que tienen un rendimiento mejor y/o mejores características de calidad en comparación con las variedades de cultivo diploides de genoma A. En el cultivo selectivo de *G. hirsutum* (AADD) se ha enfatizado el rendimiento máximo, mientras que en el de *G. barbadense* (AADD) se valoran sus fibras de mayor longitud, resistencia y finura (Jiang et al. 1998 - Proc Natl Acad Sci USA. 14 de abril de 1998; 95(8): 4419-4424).

Una fibra de algodón es una sola célula que se inicia en la epidermis del tegumento exterior de los óvulos en la anthesis o justo antes. Posteriormente, las fibras se elongan rápidamente durante aproximadamente 3 semanas antes de cambiar a la síntesis intensiva de celulosa de la pared celular secundaria. Las células fibrosas se interconectan únicamente con la capa subyacente de la semilla en los extremos basales y la entrada de solutos, agua y otras moléculas se produce a través de los plasmodesmas o la membrana plasmática. Ruan et al. 2001 (Plant Cell 13: 47-63) demostraron un cierre transitorio de los plasmodesmas durante la elongación de las fibras. Ruan et al. 2004 (Plant Physiology - Vol. 136: págs. 4104-4113) compararon la duración del cierre de los plasmodesmas entre distintos genotipos de algodón de diferentes longitudes de fibra y descubrieron una correlación positiva entre la duración del cierre de los plasmodesmas y la longitud de las fibras. Además, se presentaron pruebas microscópicas que demostraban la deposición de callosa y la degradación en la base de la fibra, que se correlacionan con la cronología del cierre y la reapertura de los plasmodesmas. Además, la expresión de un gen de β -1,3-endoglucanasa (GhGluc1) en las fibras, que permite degradar la callosa, presentaba una correlación con la reapertura de los plasmodesmas en la base de la fibra.

En el documento WO2005/017157 se describen procedimientos y medios para modular la longitud de las fibras en plantas productoras de fibras tales como el algodón mediante la modificación de la fase de elongación de las fibras como se describe en Ruan et al. 2001. La fase de elongación de las fibras se puede aumentar o disminuir al interferir con la deposición de callosa en los plasmodesmas en la base de las células fibrosas.

Además, sería interesante para la modificación de las fibras por ingeniería genética la presencia de promotores que se expresan de forma preferente o específica sólo en células fibrosas y/o que se expresan únicamente a partir de una fase del desarrollo de la fibra en particular.

El documento WO2004/018620 se refiere a una molécula de ácido nucleico aislada que codifica una quitinasa de algodón endógena y su promotor, que se expresan preferentemente en las fibras durante la deposición de la pared secundaria. También se desvelan el polipéptido codificado por la molécula de ácido nucleico, una construcción de ADN que enlaza la molécula de ácido nucleico aislada con un promotor, la construcción de ADN incorporada en un sistema de expresión, una célula huésped, una planta o una semilla vegetal. El documento también se refiere a una construcción de ADN que enlaza el promotor aislado con un segundo ADN, así como sistemas de expresión, células

huésped, plantas o semillas vegetales que contienen la construcción de ADN. También se desvelan procedimientos para conferir resistencia a insectos y hongos, regular el contenido de las fibras en celulosa y procedimientos para expresar un gen de forma preferente en las fibras durante la deposición de la pared secundaria.

- 5 Sería útil contar con promotores alternativos que dirigieran la expresión génica de forma preferente y/o intensa en las fibras durante la deposición de la pared secundaria, es decir, de forma intensa y continua (por ejemplo, a >50 % de su actividad máxima) por ejemplo, desde el inicio de la deposición de la pared secundaria hasta su finalización o, por ejemplo, de la fase de maduración en adelante. El inicio de la deposición de la pared secundaria se define como el momento en que comienza a aumentar el peso en seco por unidad de longitud de una fibra de algodón o en el que comienza a aumentar el peso seco por unidad de área de superficie de cualquier célula a través de la síntesis de un nuevo material de pared que contiene más del 40 % (p/p) de celulosa. En el caso de la fibra de algodón de *G. hirsutum* L., se espera que esto ocurra entre 14 y 17 DPA cuando se cultivan las plantas de algodón en condiciones típicas en el invernadero o en el campo (temperatura diurna de 26-34 °C, temperatura nocturna de 20-26 °C, intensidad lumínica mayor o igual a 1000 einstein/m²/s, con agua y nutrientes minerales adecuados). El final de la formación de la pared celular secundaria y el comienzo de la fase de maduración se suelen producir alrededor de 35 DPA en el caso de la fibra de algodón de *G. hirsutum* L.

Además, sería útil contar con promotores alternativos que dirigieran la expresión génica únicamente o de forma preferente en las fibras al mismo tiempo que excluyen o minimizan la expresión en otros tejidos.

Las invenciones descritas de aquí en adelante en el presente documento en los diferentes modos de realización, ejemplos, figuras y reivindicaciones proporcionan promotores o regiones promotoras específicos(as) de fibras y/o preferentes de fibras.

25 Sumario de la invención

En un modo de realización de la invención, se proporciona un promotor preferente de células fibrosas que comprende una secuencia de nucleótidos seleccionada del siguiente grupo de secuencias de nucleótidos:

- 30 a) una secuencia de nucleótidos que comprende la secuencia de nucleótidos de SEC ID NO 9 desde el nucleótido e la posición 465 hasta el nucleótido de la posición 2307;
 b) una secuencia de nucleótidos que comprende la secuencia de nucleótidos que tiene una identidad de secuencia de al menos el 95 % con la secuencia de nucleótidos mencionada en a) en la que dicho promotor preferente de células fibrosas que comprende dicha secuencia de nucleótidos dirige la expresión génica más allá de 30 DPA en *Gossypium hirsutum*.

En otro modo de realización más de la invención, se proporciona un gen quimérico que comprende las siguientes regiones de ADN enlazadas de forma funcional

- 40 a) un promotor preferente de células fibrosas como se describe en el presente documento;
 b) una región heteróloga de ADN que codifica un ARN biológicamente activo de interés; y
 c) una señal de terminación de la transcripción y de poliadenilación.

El ARN biológicamente activo puede codificar una proteína de interés, tal como una N-acetilglucosamina transferasa, preferentemente una N-acetilglucosamina transferasa de tipo NodC, una celulosa sintasa, una sacarosa sintasa, preferentemente una sacarosa sintasa de tipo C, una sacarosa fosfato sintasa o una β-1,3-endoglucanasa. El ARN biológicamente activo también puede ser un ARN inhibidor o un ARN silenciador, tal como una ribozima, microARN, una horquilla de ARN bicatenario, dirigida en particular a la regulación por disminución de la expresión de un gen endógeno del algodón, tal como β-1,3-endoglucanasa xilano sintasa (cs1F), xiloglucano sintasa (cs1D) o 1,3-1,4 glucano sintasa (cs1C).

También se proporcionan células vegetales y plantas, tales como plantas de algodón, y semillas de las mismas, que comprenden tales genes quiméricos.

55 La invención proporciona además un procedimiento para expresar un ARN biológicamente activo de forma preferente en una célula fibrosa o una planta productora de fibras, tal como una planta de algodón, procedimiento que comprende las etapas de dotar a las células de las plantas de un gen quimérico según la invención y cultivar las plantas.

60 En otro modo de realización más, la invención se refiere al uso de un promotor específico de fibra y/o preferente de fibra según la invención para la expresión preferente de un ARN biológicamente activo en células fibrosas de una planta productora de fibras tal como una planta de algodón.

Breve descripción de las figuras

65 **Figura 1:** alineación de regiones promotoras del gen de β-1,3-endoglucanasa del subgenoma A y del subgenoma D

de *Gossypium hirsutum*. La secuencia de nucleótidos del promotor del subgenoma A (secuencia superior) corresponde a la secuencia de nucleótidos de SEC ID NO 9 desde la posición 1460 hasta la posición 2408. La secuencia de nucleótidos del promotor del subgenoma D (secuencia inferior) corresponde a la secuencia de nucleótidos de SEC ID NO 10 desde la posición 2372 hasta la posición 3372. Las diferencias en la secuencia de nucleótidos se indican con recuadros grises. Los nucleótidos que no tienen un nucleótido correspondiente en la otra región promotora se indican con guiones en la secuencia de nucleótidos que carece de los nucleótidos. La homología total entre las dos regiones promotoras es de aproximadamente el 71 % de identidad de secuencia.

Figura 2: alineación de regiones codificantes del gen de β -1,3-endoglucanasa (sin intrón) del subgenoma A y del subgenoma D de *Gossypium hirsutum*. La secuencia de nucleótidos de la región codificante del alelo de subgenoma A (secuencia superior) corresponde a la secuencia de nucleótidos de SEC ID NO 9 desde la posición 2410 hasta la posición 2443 y desde la posición 2556 hasta la 3499. La secuencia de nucleótidos de la región codificante del alelo de subgenoma D (secuencia inferior) corresponde a la secuencia de nucleótidos de SEC ID NO 10 desde la posición 3373 hasta la posición 3406 y desde la posición 3501 hasta la 4444. Las diferencias en la secuencia de nucleótidos se indican con recuadros grises. Los nucleótidos que no tienen un nucleótido correspondiente en la otra región promotora se indican con guiones en la secuencia de nucleótidos que carece de los nucleótidos. La homología total entre las dos regiones promotoras es de aproximadamente el 97% de identidad de secuencia.

Figura 3: alineación de la secuencia intrónica del gen de β -1,3-endoglucanasa del subgenoma A y del subgenoma D de *Gossypium hirsutum*. La secuencia superior corresponde a la secuencia de nucleótidos del intrón del alelo de subgenoma A, mientras que la secuencia inferior corresponda a la secuencia de nucleótidos del intrón del alelo de subgenoma D.

Figura 4: alineación de la secuencia de aminoácidos de la β -1,3-endoglucanasa codificada por el alelo de subgenoma A (secuencia superior correspondiente a la SEC ID NO 11) y el alelo de subgenoma D (secuencia inferior correspondiente a la SEC ID NO 12).

Figura 5: alineación de la secuencia de nucleótidos de SEC ID 1 a 4 desde la posición de nucleótido 201 a la 211. El sitio de reconocimiento de la enzima de restricción *Alw1* presente en la secuencia de nucleótidos del alelo de subgenoma A sólo del gen de β -1,3-endoglucanasa está subrayado y se representa en letra cursiva.

Figura 6: correlación entre la curva de crecimiento de la fibra (panel A) y la expresión del gen de β -1,3-endoglucanasa (*GhGluc1*) (panel B). Se extrajo y se igualó ADN de una colección de ADNc de fibras (en desarrollo) en *Gossypium hirsutum*. Los fragmentos de PCR se amplificaron con los cebadores oligonucleotídicos SE002 y SE003 (SEC ID NO 5 y 6) y se digirieron con *Alw1*. Un producto amplificado por PCR para la variante de genoma A proporciona 3 fragmentos (479 pb + 118 pb + 59 pb), mientras que para la variante de genoma D sólo proporciona 2 fragmentos (538 pb + 118 pb).

Figura 7: análisis de la expresión de GhGluc1 en hojas, raíces y tallos de algodón por RT-PCR. Se sometió ARN extraído de hojas, raíces y tallos de algodón a un análisis por RT-PCR para detectar la expresión de glucanasa 1 o proteína fosfatasa 2A (pp2A). La síntesis de la primera hebra de ADNc se realizó con el sistema SuperScript First-Strand Synthesis de Invitrogen. Carril 1: escalera de 100 pb; carril 2: muestra de ARN de hoja de algodón + cebadores específicos de Gluc1 + transcriptasa inversa; carril 3: muestra de ARN de hoja de algodón + cebadores específicos de Gluc1 - transcriptasa inversa; carril 4: muestra de ARN de raíz de algodón + cebadores específicos de Gluc1 + transcriptasa inversa; carril 5: muestra de ARN de raíz de algodón + cebadores específicos de Gluc1 - transcriptasa inversa; carril 6: muestra de ARN de tallo de algodón + cebadores específicos de Gluc1 + transcriptasa inversa; carril 7: muestra de ARN de tallo de algodón + cebadores específicos de Gluc1 - transcriptasa inversa; carril 8: ADN genómico (Fibermax 400 ng) + cebadores específicos de Gluc1; carril 9: escalera de 100 pb; carril 10: muestra de ARN de hoja de algodón + cebadores específicos de pp2A + transcriptasa inversa; carril 11: muestra de ARN de hoja de algodón + cebadores específicos de pp2A - transcriptasa inversa; carril 12: muestra de ARN de raíz de algodón + cebadores específicos de pp2A + transcriptasa inversa; carril 13: muestra de ARN de raíz de algodón + cebadores específicos de pp2A - transcriptasa inversa; carril 14: muestra de ARN de tallo de algodón + cebadores específicos de pp2A + transcriptasa inversa; carril 15: muestra de ARN de tallo de algodón + cebadores específicos de pp2A - transcriptasa inversa; carril 16: escalera de 100 pb; carril 17: ARN de control positivo (suministrado con el kit).

Modos de realización detallados

La presente invención se basa en el descubrimiento inesperado de que la cronología de la expresión de los alelos específicos de subgenoma A y D del gen de β -1,3-endoglucanasa específico de fibra (*GhGluc1*) se diferente al menos en *Gossypium hirsutum*. Mientras el inicio de la expresión del alelo específico de subgenoma D se correlaciona con la finalización de la fase de elongación rápida (aproximadamente de 14 a 17 días postantesis, en lo sucesivo "DPA"), el inicio de la expresión del subgenoma A se retrasa hasta el comienzo de la fase tardía de maduración de la fibra (aproximadamente 35-40 DPA), en función de las condiciones de cultivo.

En consecuencia, en este documento se describe la identificación de promotores específicos de fibra novedosos que

se expresan de forma preferente en células fibrosas desde el final de la fase de elongación rápida o desde el inicio de la fase de maduración de la fibra. Estos promotores tienen utilidad, por ejemplo, en la expresión de genes quiméricos productores de biomoléculas novedosas en fases tardías del desarrollo de las fibras, por ejemplo, al codificar XET potenciado, el dominio de unión a celulosa (CBD), los genes mutantes de raíces hinchadas (RSW), el color, la plasticidad, la capacidad de digestión de la celulosa, propiedades/compuestos de pared celular novedosos. Los promotores también se podrían usar para dirigir la expresión génica -30, 40 DPA en las fibras; por ejemplo, para la producción de quitina en fases tardías del desarrollo de las fibras de algodón (30, 40 DPA) o para la expresión génica 40 DPA en las fibras, por ejemplo, para la producción de quitina en fases muy tardías del desarrollo de las fibras de algodón (40 DPA). Los promotores según la invención también se podrían usar para dirigir la expresión de genes Cesa 40 DPA para potenciar la biosíntesis de celulosa y potenciar las propiedades de las fibras o para producir ARN biológicamente activo destinado a la silenciación específica de fase de genes endógenos del algodón, tal como la silenciación de β -1,3-endoglucanasa para cultivar fibras más largas. Los promotores también se podrían aplicar en la expresión específica de fase de genes quiméricos que codifican, por ejemplo, callosa sintasa para modificar las propiedades y la calidad de las fibras.

En este documento se describe un fragmento promotor preferente de fibra que se puede aislar del algodón, en particular de la especie *Gossypium hirsutum*, y que se sitúa aguas arriba de las secuencias de ácido nucleico que codifican una β -1,3-endoglucanasa específica de fibra que tiene una secuencia de aminoácidos que comprende la secuencia de aminoácidos de SEC ID 11, 12, 7 u 8 o variantes de la misma.

Tal como se usa en el presente documento, el término "promotor" denota cualquier ADN reconocido y unido (directa o indirectamente) por una polimerasa de ARN dependiente de ADN durante la iniciación de la transcripción. Un promotor incluye el sitio de iniciación de la transcripción y sitios de unión para los factores de iniciación de la transcripción y la polimerasa de ARN, y puede comprender diversos sitios más (por ejemplo, potenciadores) a los que se pueden unir proteínas reguladoras de la expresión génica.

Tal como se usa en el presente documento, el término "región reguladora" significa cualquier ADN que participa en la dirección de la transcripción y el control (es decir, la regulación) de la cronología y el grado de transcripción de una secuencia de ADN dada, tal como un ADN que codifica una proteína o polipéptido. Por ejemplo, una región reguladora (o "región promotora") en 5' es una secuencia de ADN situada aguas arriba (es decir, en 5') de una secuencia codificante y que comprende el promotor y la secuencia líder no traducida en 5'. Una secuencia reguladora en 3' es una secuencia de ADN situada corriente abajo (es decir, en 3') de la secuencia codificante y que comprende señales de formación (y/o regulación) de extremos en 3' en la transcripción adecuados, incluidas una o más señales de poliadenilación.

El término "gen" significa cualquier fragmento de ADN que comprende una región de ADN (la "región de ADN transcrito") que se transcribe en una molécula de ARN (por ejemplo, un ARNm) en una célula bajo el control de regiones reguladoras adecuadas, por ejemplo, una región promotora que se puede expresar en plantas. Por tanto, un gen puede comprender varios fragmentos de ADN enlazados de forma funcional, tales como un promotor, una secuencia líder no traducida en 5', una región codificante y una región no traducida en 3' que comprende un sitio de poliadenilación. Un gen vegetal endógeno es un gen que se encuentra de forma natural en una especie vegetal. Una gen quimérico es cualquier gen que no se encuentra normalmente en una especie vegetal o, de forma alternativa, cualquier gen en el que el promotor no está asociado en la naturaleza con parte o con la totalidad de la región de ADN transcrito (una región de ADN "heteróloga") o con al menos otra región reguladora del gen.

El término "expresión de un gen" se refiere al proceso en el que una región de ADN bajo el control de regiones reguladoras, en particular el promotor, se transcribe en un ARN que es biológicamente activo, es decir, que puede interactuar con otro ácido nucleico o que se puede traducir en un polipéptido o una proteína biológicamente activos. Se dice que un gen codifica un ARN cuando el producto final de la expresión del gen es ARN biológicamente activo, tal como un ARN antisentido o una ribozima. Se dice que un gen codifica una proteína cuando el producto final de la expresión del gen es una proteína o un polipéptido funcional o biológicamente activo.

El término "selectivo de fibra" o "selectivo de célula fibrosa" o "específico de fibra" o "específico de célula fibrosa" respecto a la expresión de un ADN según la presente invención, se refiere, con fines prácticos, a la expresión altamente específica de un ADN en células fibrosas de plantas, tales como plantas de algodón ("selectiva de células fibrosas"). En otras palabras, las concentraciones de transcrito de un ADN en tejidos diferentes de las células fibrosas están por debajo del límite de detección o son muy bajas (menos de aproximadamente 0,2 picogramos por microgramo de de ARN total).

El término "preferente de fibra" o "preferente de célula fibrosa" respecto a la expresión de un ADN según la presente invención se refiere a un patrón de expresión por el que el ADN se expresa predominantemente en células fibrosas o fibras, pero se puede identificar la expresión en otros tejidos de la planta. Preferentemente, la expresión en células fibrosas es de aproximadamente 2 a aproximadamente 10 veces mayor en las células fibrosas que en otros tejidos.

Las secuencias de nucleótidos de SEC ID NO 9 y SEC ID NO 10 representan las secuencias de nucleótidos de las moléculas de ADN genómico que codifican respectivamente una β -1,3 endoglucanasa selectiva de fibra de

subgenoma A y de subgenoma D. El sitio de iniciación de la transcripción de SEC ID NO 9 se ha determinado en la posición 2308, mientras que el sitio de iniciación de la transcripción de SEC ID NO 10 se ha determinado en la posición 3270. Por tanto, en un modo de realización de la invención se proporciona un promotor selectivo de fibra que tiene la secuencia de nucleótidos de SEC ID NO 9 desde la posición de nucleótido 1 hasta la posición de nucleótido 2307 o 2308; también se describe un promotor selectivo de fibra que tiene la secuencia de nucleótidos de SEC ID NO 10 desde la posición de nucleótido 1 hasta la posición de nucleótido 3269 o 3270.

El sitio de iniciación de la traducción de la secuencia de nucleótidos de SEC ID NO 9 se ha determinado como el codón ATG situado en las posiciones 2410 a 2412. La región no traducida en 5' del promotor selectivo de fibra del alelo subgenómico A que codifica la β -1,3 endoglucanasa en consecuencia tiene una secuencia de nucleótidos de SEC ID NO 9 desde la posición de nucleótido 2308 hasta la posición de nucleótido 2409. En consecuencia, una región promotora selectiva de fibra del alelo subgenómico A es una región promotora que comprende los nucleótidos 1 a 2409 de la SEC ID NO 9.

Del mismo modo, el sitio de iniciación de la traducción de la secuencia de nucleótidos de SEC ID NO 10 se ha determinado como el codón ATG situado en las posiciones 3373 a 3375. La región no traducida en 5' del promotor selectivo de fibra del alelo subgenómico D que codifica la β -1,3 endoglucanasa en consecuencia tiene una secuencia de nucleótidos de SEC ID NO 10 desde la posición de nucleótido 3270 hasta la posición de nucleótido 3372. En consecuencia, una región promotora selectiva de fibra del alelo subgenómico D es una región promotora que comprende los nucleótidos 1 a 3372 de la SEC ID NO 10.

La alineación de las regiones promotoras selectivas de fibra de SEC ID NO 9 y SEC ID NO 10 (figura 1) reveló una homología de secuencia general en las secuencias de nucleótidos de aproximadamente 1000 nucleótidos aguas arriba del codón de inicio de la traducción ATG, en particular en la región aproximadamente 150 nucleótidos aguas arriba del codón de inicio de la traducción ATG.

No es necesario decir que el promotor o la región promotora pueden estar contenidos en secuencias de nucleótidos mayores. Por tanto, se puede determinar una región promotora selectiva de célula fibrosa como la región aguas arriba (es decir, situada en 5') del codón que codifica el primer aminoácido de la proteína codificada por el ARNm indicado en la SEC ID NO 9 o 10. Una región promotora de este tipo puede tener al menos de aproximadamente 400 a 500 pb, al menos aproximadamente 1000 pb, al menos 1200 pb, al menos aproximadamente 1300 pb o al menos de aproximadamente 1500 a 2000 pb aguas arriba del codón de inicio. Por comodidad, es preferente que esta región promotora no se extienda más de aproximadamente 3000 a 5000 pb aguas arriba del codón de inicio. El tamaño del fragmento se puede determinar parcialmente por la presencia de sitios de restricción adecuados. Por ejemplo, para la región promotora selectiva de fibra del alelo de subgenoma A que codifica la β -1,3 endoglucanasa, se puede clonar con facilidad un fragmento de aproximadamente 1945 pb (desde la posición de nucleótido 465 hasta la posición de nucleótido 2409 de la SEC ID NO 9). La presencia de enzimas de restricción adecuadas permite generar con facilidad regiones promotoras de aproximadamente 1036 pb (desde el nucleótido 1374 hasta la posición de nucleótido 2409 de la SEC ID NO 9) o de aproximadamente 879 pb (desde el nucleótido 1531 hasta la posición de nucleótido 2409 de la SEC ID NO 9). Para la región promotora selectiva de fibra del alelo de subgenoma D que codifica la β -1,3 endoglucanasa, se puede clonar con facilidad un fragmento de aproximadamente 1976 pb (desde la posición de nucleótido 1397 hasta la posición de nucleótido 3372 de la SEC ID NO 10). La presencia de enzimas de restricción adecuadas permite generar con facilidad regiones promotoras de aproximadamente 1002 pb (desde el nucleótido 1397 hasta la posición de nucleótido 3372 de la SEC ID NO 10) o de aproximadamente 655 pb (desde el nucleótido 2371 hasta la posición de nucleótido 3372 de la SEC ID NO 10).

Además, resultará evidente que se pueden aislar promotores o regiones promotoras selectivos de fibras equivalentes de otras plantas de algodón o plantas progenitoras de algodón. Con este fin, se pueden aislar fragmentos promotores de otras plantas de algodón, tales como *G. barbadense*, que incluyen las llamadas variedades PIMA, o de otras variedades mediante un fragmento promotor como se describe en el presente documento como sonda e identificar secuencias de nucleótidos de estas otras plantas que hibridan en condiciones de hibridación rigurosas.

Tal como se usa en el presente documento, "condiciones de hibridación rigurosas" significa que, por lo general, se producirá la hibridación si hay una identidad de secuencia de al menos el 95 % y, preferentemente, de al menos el 97 % entre la sonda y la secuencia diana. Son ejemplos de condiciones de hibridación rigurosas, la incubación durante una noche en una solución que comprende formamida al 50 %, 5 x SSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato de sodio 50 mM (pH 7,6), solución de Denhardt 5x, sulfato de dextrano al 10 % y 20 μ g/ml de ADN de vehículo desnaturalizado cortado, tal como ADN de esperma de salmón, seguido de lavado del soporte de hibridación en 0,1 x SSC a aproximadamente 65 °C. Otras condiciones de hibridación y lavado son bien conocidas y se ejemplifican en Sambrook et al., Molecular Cloning: A Laboratory Manual, segunda edición, Cold Spring Harbor, NY (1989), en particular en el capítulo 11.

Estos promotores selectivos de fibra equivalentes aislados de otras plantas de algodón o plantas progenitoras de algodón pueden representar promotores variantes en los que se han reemplazado uno o más nucleótidos por sustitución, delección o inserción. También se puede generar un promotor selectivo de fibra variante sintéticamente.

La figura 1 representa la alineación de promotores selectivos de fibra del subgenoma A y D de *G. hirsutum* y las posiciones variantes, donde la secuencia de nucleótidos se puede cambiar aparentemente sin un efecto significativo sobre la capacidad de iniciación de la transcripción selectiva de fibra, se han indicado con recuadros grises. A partir de la figura 1 y de su leyenda también es evidente que la identidad de secuencia total entre los dos fragmentos promotores selectivos de fibra ejemplificados puede ser de tan sólo aproximadamente el 71 % de identidad de secuencia.

En consecuencia, en otro modo de realización de la presente invención, se describen promotores y regiones promotoras selectivos de fibra que comprenden una secuencia de nucleótidos que tiene una identidad de secuencia de al menos el 95 % con los promotores y las regiones promotoras selectivos de fibra descritos en el presente documento.

Para el fin de la presente invención, la "identidad de secuencia" de dos secuencias de nucleótidos o aminoácidos relacionadas, expresada como un porcentaje, se refiere al número de posiciones de las dos secuencias alineadas óptimamente que tienen residuos idénticos ($\times 100$) dividido entre el número de posiciones comparadas. Un hueco, es decir, una posición de una alineación en la que un residuo está presente en una secuencia pero no en la otra, se considera una posición con residuos no idénticos. La alineación de las dos secuencias se realiza por el algoritmo de Needleman y Wunsch (Needleman y Wunsch (1970). La alineación de secuencias asistida por ordenador anterior se puede realizar con facilidad usando un programa informático estándar tal como GAP, que forma parte del paquete informático Wisconsin versión 10.1 (Genetics Computer Group, Madison, Wisconsin, EE. UU.) con la matriz de puntuación predeterminada con una penalización por creación de hueco de 50 y una penalización por extensión de hueco de 3.

Resultará evidente que, cuando se definen secuencias de nucleótidos de moléculas de ARN por referencia a la secuencia de nucleótidos de moléculas de ADN correspondientes, se debe reemplazar la timina (T) de la secuencia de nucleótidos con uracilo (U). Si se hace referencia a moléculas de ARN o ADN quedará claro por el contexto de la solicitud.

La capacidad de expresión selectiva de fibra del promotor o las regiones promotoras identificados o generados se puede probar fácilmente enlazando de forma funcional de estas moléculas de ADN a una región codificante que codifica un marcador fácilmente puntuable, por ejemplo, un gen de β -glucuronidasa, introduciendo este gen quimérico en una planta productora de fibras y analizando el patrón de expresión del marcador en células fibrosas (preferentemente durante el desarrollo de las fibras) en comparación con el patrón de expresión del marcador en otras partes de las plantas.

No es necesario decir que los promotores y las regiones promotoras de la invención también pueden comprender elementos adicionales conocidos por mejorar la eficacia de la transcripción, tales como potenciadores, intrones, etc.

Además, las secuencias UTR en 5' ejemplificadas de las regiones promotoras selectivas de fibra ejemplificadas son bastante similares en la secuencia de nucleótidos y se espera que las secuencias UTR en 5' sean intercambiables.

Además, la invención incluye moléculas de ADN que comprenden los promotores o las regiones promotoras selectivos de fibra de la invención, enlazados de forma funcional a una o más regiones heterólogas que codifican un ARN, una proteína o un péptido biológicamente activo. Los promotores de la invención se pueden usar para expresar cualquier región codificante heteróloga deseada.

Los ejemplos de otras moléculas de ADN que codifican proteínas que se podrían expresar con el promotor de la presente invención incluyen, pero sin limitación, celulosa sintasas homólogas y heterólogas (genes *CesA*), tanto en forma normal como mutada (Arioli et al., "Molecular Analysis of Cellulose Biosynthesis in Arabidopsis", *Science*, 279: 717-720 (1998); Holland et al., "A Comparative Analysis of the Plant Cellulose Synthase (*CesA*) Gene Family", *Plant Physiol.*, 123: 1313-1324 (2000)); genes que pueden modular el reparto del carbono en la celulosa (Delmer, "Cellulose Biosynthesis in Developing Cotton Fibers" en: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 85-112 (1999)) como la sacarosa sintasa (Amor et al., "A Membrane-Associated Form of Sucrose Synthase and Its Potential Role Synthesis of Cellulose and Callose in Plants", *Proc. Natl. Acad. Sci. USA*, 92: 9353-9357 (1995), sacarosa fosfato sintasa (Haigler et al., "Transgenic Cotton Over-Expressing Sucrose Phosphate Synthase Produces Higher Quality Fibers with Increased Cellulose Content and Has Enhanced Seed Cotton Yield" Resumen 477. En: *Proceedings of Plant Biology 2000*, 15-19 de julio, San Diego, CA. American Society of Plant Physiologists, Rockville, MD, (2000), UDPG-pirofosforilasa (Waffler y Meier, "Enzyme Activities in Developing Cotton Fibers" *Plant Physiol. Biochem.* 32: 697-702 (1994), pirofosfatasa inorgánica (Geigenberger et al., "Overexpression of Pyrophosphatase Leads to Increased Sucrose Degradation and Starch Synthesis, Increased Activities of Enzymes for Sucrose-Starch Interconversions, and Increased Levels of Nucleotides in Growing Potato Tubers", *Planta*, 205:428-437 (1998)), hexocinasas (Smeekens, "Sugar Regulation of Gene Expression" *Curr. Op. Plant Biol.*, 1: 230-234 (1998) e invertasas (Sturm y Tang, "The Sucrose-Cleaving Enzymes of Plants are Crucial for Development, Growth, and Carbon Partitioning", *Trends Plant Sci.*, 4: 401-407 (1999)); genes que podrían afectar a las propiedades moleculares y biofísicas de la celulosa, incluido en grado de polimerización, el grado de cristalinidad, el tamaño de los cristales y la orientación de

las microfibrillas (es decir, genes codificantes de proteínas, incluidos polímeros proteínicos que cocristalizan o dominios de unión a celulosa y enzimas sintetizadoras de polisacáridos y de otros tipos) (Delmer, "Cellulose Biosynthesis: Exciting Times for a Difficult Field of Study", *Ann. Rev. Plant Physiol. Mol. Biol.* 50: 245-276 (1999); Delmer, "Cellulose Biosynthesis in Developing Cotton Fibers. En: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 85-112 (1999); Hsieh, "Structural Development of Cotton Fibers. En: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 137-166 (1999)); factores de transcripción tales como genes MYB que podrían prolongar el crecimiento por elongación y/o cambiar la cronología o el grado de deposición de la pared secundaria (Wilkins y Jernstedt, "Molecular Genetics of Developing Cotton Fibers. En: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 231-270 (1999)); genes para efectuar la síntesis de hormonas vegetales y cambiar las propiedades de las fibras (John, "Genetic Engineering Strategies for Cotton Fiber Modification. En: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 271-289 (1999)); genes para elementos del citoesqueleto o proteínas asociadas al citoesqueleto que podrían afectar a las propiedades de las fibras (Seagull, "Cytoskeletal Involvement in Cotton Fiber Growth and Development", *Micron*, 24: 643-660 (1993)); genes para sintetizar lípidos o modificar enzimas que podrían cambiar las propiedades de la membrana y, de este modo, mejorar la calidad de las fibras, incluido en condiciones de estrés ambiental (Haigler, "The Crystallinity of Cotton Cellulose in Relation to Cotton Improvement", *Proc. Cotton Fiber Cellulose: Structure, Function and Utilization Conference*, National Cotton Council of America: Memphis, TN., págs. 211-225 (1992)); enzimas tales como xiloglucano endotransferasa, peroxidasa, expansina o ATPasa vacuolar que, a través del aumento o la disminución de su actividad, podrían prolongar o aumentar el crecimiento de la extensión durante la deposición de la pared secundaria (Wilkins y Jernstedt, "Molecular Genetics of Developing Cotton Fibers. En: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 231-270 (1999)); genes para polímeros proteínicos o plásticos que se podrían retener en la luz de la fibra o integrarse en la pared celular para aumentar la resistencia de la fibra o cambiar sus propiedades textiles (John, "Genetic Engineering Strategies for Cotton Fiber Modification", en: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 271-289 (1999); Guda et al., "Hyperexpression of an Environmentally Friendly Synthetic Polymer Gene", *Biotechnology Letters*, 17: 745-750 (1995)); genes para enzimas biosintéticas de la matriz de la pared celular vegetal o sus proteínas reguladoras, para que se puedan integrar otros hidratos de carbono en la pared celular y cambiar las propiedades de la fibra (Haigler, "The Relationship Between Polymerization and Crystallization in Cellulose Biogenesis", en C. H. Haigler y P. Weimer, eds., *Biosynthesis and Biodegradation of Cellulose*, Nueva York: Marcel Dekker, págs. 99-124 (1991); Andrawis et al., "Cotton Fiber Annexins: A Potential Role in the Regulation of Callose Synthase", *Plant J.*, 3: 763-772 (1993); genes para moléculas tales como taninos, suberinas o tintes que podrían conferir un color valioso a las fibras (Ryser, "Cotton Fiber Initiation and Histodifferentiation", en: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 1-46 (1999); genes para moléculas tales como cutina, suberina o cera que podrían cambiar la capacidad de absorción y la resistencia de las fibras de algodón (mayo "Genetic Variation in Fiber Quality", en: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 183-230 (1999); Ryser, "Cotton Fiber Initiation and Histodifferentiation", en: A. S. Basra (ed.), *Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing*, The Haworth Press, Nueva York, págs. 1-46 (1999); y genes para moléculas de transducción de señales, tales como Rac, que pueden regular los cambios entre las fases del desarrollo de las fibras (Delmer et al., "Genes Encoding Small GTP-Binding Proteins Analogous to Mammalian rac are Preferentially Expressed in Developing Cotton Fibers" *Mol. Gen. Genet.*, 248: 43-51 (1995).

50 Son regiones codificantes de proteínas particularmente preferentes las regiones codificantes de N-acetilglucosamina que se describe, por ejemplo, en el documento WO2006/136351 o los genes de sacarosa sintasa que se describen en los documentos WO2002/45485 o EP06015433.3.

El ARN biológicamente activo también puede codificar los llamados ARN antisentido, ARN sentido, ARN bicatenario como se describe en el documento WO99/53053 o moléculas de micro ARN diseñadas según normas bien conocidas en la técnica para regular por disminución la expresión de otros genes comprendidos en la célula de las plantas productoras de fibras o incluso de genes comprendidos en un patógeno o una plaga que se alimenta de la planta productora de fibras.

También se proporcionan procedimientos para expresar una proteína o ARN biológicamente activo específicamente en células fibrosas de una planta productora de fibras que comprenden la etapa de introducir en las células fibrosas de plantas productoras de fibras una molécula de ADN que comprende un promotor o una región promotora selectivos de fibra como se describe en el presente documento enlazados de forma funcional a una región de ADN transcrito que codifica la molécula de ARN biológicamente activo.

La invención también engloba los genes quiméricos descritos en el presente documento, así como plantas, semillas, tejidos que comprenden estos genes quiméricos.

65 Los procedimientos para transformar plantas son bien conocidos en la técnica y son de importancia menor para la

presente invención. Los procedimientos para transformar plantas de algodón también son bien conocidos en la técnica. Se describe la transformación del algodón mediada por *Agrobacterium*, por ejemplo, en la patente de EE. UU. 5.004.863 o en la patente de EE. UU. 6.483.013 y se informa de la transformación del algodón por bombardeo de partículas, por ejemplo, en el documento WO92/15675.

5 Los genes quiméricos según la invención se pueden introducir en plantas en manera estable o de manera transitoria usando procedimientos bien conocidos en la técnica. Los genes quiméricos se pueden introducir en plantas o se pueden generar dentro de la célula vegetal, como se describe, por ejemplo, en el documento EP 1339859.

10 Los genes quiméricos se pueden introducir por transformación en plantas de algodón de las que se pueden obtener callos embrionarios, tales como Coker 312, Coker310, Coker 5Acala SJ-5, GSC25110, FIBERMAX 819, Siokra 1-3, T25, GSA75, Acala SJ2, Acala SJ4, Acala SJ5, Acala SJ-C1, Acala B1644, Acala B1654-26, Acala B1654-43, Acala B3991, Acala GC356, Acala GC510, Acala GAM1, Acala CI, Acala Royale, Acala Maxxa, Acala Prema, Acala B638, Acala B1810, Acala B2724, Acala B4894, Acala B5002, "picker" Siokra no Acala, variedad "stripper" FC2017, Coker
15 315, STONEVILLE 506, STONEVILLE 825, DP50, DP61, DP90, DP77, DES 119, 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.

25 Tal como se usa en el presente documento, "algodón" incluye *Gossypium hirsutum* o *Gossypium barbadense*. Las "plantas progenitoras de algodón" incluyen *Gossypium arboreum*, *Gossypium herbaceum* y *Gossypium raimondii* y *Gossypium longicalyx*.

30 Los procedimientos y los medios de la presente invención también se pueden emplear para otras especies vegetales tales como el cáñamo, el yute, el lino y las plantas leñosas, incluidas, pero sin limitación, *Pinus spp.*, *Populus spp.*, *Picea spp.*, *Eucalyptus spp.* etc.

35 La planta transformada obtenida se puede usar en un esquema de cultivo convencional para producir más plantas transformadas con las mismas características o para introducir el gen quimérico según la invención en otras variedades de la misma planta o de una especie relacionada o en plantas híbridas. Las semillas obtenidas de las plantas transformadas contienen los genes quiméricos de la invención como inserto genómico estable y también se engloban en la invención.

40 Tal como se usa en el presente documento, "que comprende" debe interpretarse como la especificación de la presencia de los rasgos, números enteros, etapas o componentes indicados a los que se hace referencia, pero no excluye la presencia o la adición de uno o más rasgos, números enteros, etapas o componentes o grupos de los mismos. Así pues, por ejemplo, un ácido nucleico o una proteína que comprende una secuencia de nucleótidos o aminoácidos, puede comprender más nucleótidos o aminoácidos que los que se mencionan realmente, es decir, estar incluidos en un ácido nucleico o una proteína mayores. Un gen quimérico que comprende una región de ADN, que está definida funcional o estructuralmente, puede comprender regiones adicionales de ADN, etc.

45 En los siguientes ejemplos se describe la identificación de dos β -1,3-endoglucanasas de subtipo A y D de algodón, así como sus regiones promotoras, y el análisis de la cronología de la expresión y la implicación en la síntesis de la pared celular vegetal secundaria. También se describen genes quiméricos para la expresión preferente de fibra en plantas productoras de fibras, tales como el algodón, y usos de los mismos. A menos que se indique lo contrario en los ejemplos, todas las técnicas de ADN recombinante se llevan a cabo según protocolos estándar como los descritos en Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor Laboratory Press, NY y en los volúmenes 1 y 2 de Ausubel et al. (1994) *Current Protocols in Molecular Biology*, Current Protocols, EE. UU. Se describen materiales y procedimientos estándar para el trabajo molecular con plantas en *Plant Molecular Biology Labfax* (1993) por R.D.D. Croy, publicado conjuntamente por BIOS Scientific Publications Ltd (Reino Unido) y Blackwell Scientific Publications, Reino Unido.

A lo largo de la descripción y los ejemplos se hace referencia a las siguientes secuencias representadas en la lista de secuencias:

60 SEC ID NO 1: fragmento genómico amplificado de β -1,3-endoglucanasa de *Gossypium hirsutum* Fiber Max966, subtipo A
SEC ID NO 2: fragmento genómico amplificado de β -1,3-endoglucanasa de *Gossypium hirsutum* Fiber Max966, subtipo D
65 SEC ID NO 3: fragmento genómico amplificado de β -1,3-endoglucanasa de *Gossypium arboreum*
SEC ID NO 4: fragmento genómico amplificado de β -1,3-endoglucanasa de *Gossypium raimondii*

- SEC ID NO 5: cebador inverso (SE002) para la amplificación del fragmento genómico de β -1,3-endoglucanasa
- SEC ID NO 6: cebador directo (SE003) para la amplificación del fragmento genómico de β -1,3-endoglucanasa
- 5 SEC ID NO 7: secuencia de aminoácidos codificada por el fragmento genómico amplificado de la β -1,3-endoglucanasa de *Gossypium hirsutum* Fiber Max966, subtipo A
- SEC ID NO 8: secuencia de aminoácidos codificada por el fragmento genómico amplificado de la β -1,3-endoglucanasa de *Gossypium hirsutum* Fiber Max966, subtipo D
- SEC ID NO 9: secuencia de nucleótidos del clon genómico completo de la β -1,3-endoglucanasa de *Gossypium hirsutum* Fiber Max966, subtipo A, incluida la secuencia promotora
- 10 SEC ID NO 10: secuencia de nucleótidos del clon genómico completo de la β -1,3-endoglucanasa de *Gossypium hirsutum* Fiber Max966, subtipo D, incluida la secuencia promotora
- SEC ID NO 11: secuencia de aminoácidos de la β -1,3-endoglucanasa completa de *Gossypium hirsutum* Fiber Max966, subtipo A
- 15 SEC ID NO 12: secuencia de aminoácidos de la β -1,3-endoglucanasa completa de *Gossypium hirsutum* Fiber Max966, subtipo D

Ejemplo 1: identificación de los alelos específicos de subgenoma A y D que codifican una β -1,3-glucanasa específica de fibra en plantas de algodón y plantas progenitoras de algodón

20 Se ha propuesto que la callosa participa en el proceso de mantenimiento de la turgencia en las células fibrosas en crecimiento. La retirada de la callosa de la base de las estructuras puente entre células de algodón fibrosas y no fibrosas, conocidas como plasmodesmas, disipa la presión de la turgencia. Esto tiene como consecuencia la finalización de la elongación rápida de la fibra. Por lo tanto, al retrasar la retirada de la callosa debería aumentar la longitud de la fibra.

25 La enzima β -1,3-glucanasa cataliza la hidrólisis de los enlaces β -1,3-D-glucosídicos del β -1,3-D-glucano (callosa). El gen de una β -1,3-glucanasa específica de fibra codificada por *GhGluc1* era indetectable cuando la callosa se depositaba en la base de la fibra, pero se hacía evidente en el momento de la degradación de la callosa.

30 En resumen, parece que el cierre de los plasmodesmas desempeña un papel importante en la elongación de las fibras de algodón. La deposición y la degradación de la callosa deberían estar implicadas en el cierre y la reapertura de los plasmodesmas, respectivamente. La expresión de *GhGluc1* podría desempeñar un papel en este proceso al degradar la callosa y, por tanto, abrir los plasmodesmas. (Ruan et al., 2004, Plant Physiology - 136: 4104-4113).

35 Con base en la secuencia de nucleótidos de *GhGluc1* (número de acceso de EMBL D88416) descrita en Ruan et al., 2004, Plant Physiology - 136: 4104-4113, se diseñaron 2 cebadores (SE002: ggccgaagccgatcttatctagg (cebador inverso; SEC ID NO 5) y SE003: cggcaacaatcttccatctccag (cebador directo, SEC ID NO 6)) para amplificar fragmentos de ADN genómico para *G. hirsutum* (genoma AD), *G. arboreum* (genoma A) y *G. raimondii* (genoma D). Estos fragmentos se han secuenciado (véanse las SEC ID NO 1-4). Para *G. hirsutum* se obtuvieron 2 secuencias de consenso, para *G. arboreum* se obtuvo 1 secuencia de consenso y para *G. raimondii* se obtuvo 1 secuencia de consenso.

Resumen de los polimorfismos entre las dos secuencias de *G. hirsutum* y las dos secuencias diploides:

ID	taxón	genoma	75	120	135	167	170	147	201	202	256	265	270	271	247	279	280	299	307	327	355	360	358	376	492	494	495	569	571
GaGluc1	<i>G. arboreum</i>	A	A	-	G	A	A	C	C	G	C	A	G	G	C	T	C	C	G	G	C	C	C	G	T	C	G	A	C
GhGluc1-SGA	<i>G. hirsutum</i>	AD	T	C	C	A	A	C	C	G	A	A	G	G	C	T	C	C	G	G	G	C	C	G	T	C	G	A	C
GhGluc1-SGD	<i>G. hirsutum</i>	AD	A	C	C	G	C	G	A	A	C	G	A	G	G	C	T	T	A	A	C	T	T	G	C	G	T	G	T
GrGluc1	<i>G. raimondii</i>	D	A	C	C	G	C	G	A	A	C	G	A	A	G	C	T	T	A	A	C	C	T	A	C	G	T	G	T

45 En consecuencia, un tipo de secuencia corresponde a la secuencia de genoma A y el otro tipo corresponde al genoma D.

50 Con la digestión con *Alw1* (sitio de reconocimiento = GGATC) en el fragmento amplificado por PCR, se pueden distinguir ambas variantes subgenómicas dentro del *hirsutum* o el genoma (véase la figura 5). Por tanto, se puede usar el siguiente protocolo para distinguir los alelos de subgenoma A y D de GhGluc1. Los cebadores usados son:

- a. SE002: GGCCGAAGCCGATCTTATCTAGG (SEC ID NO 5)
- b. SE003: CGGCAACAATCTTCCATCTCCAG (SEC ID NO 6)

55 La longitud prevista para el producto de PCR es de 655 pb.

Condiciones de PCR:

	1 x
ADN molde (200 ng/μl)	1 μl
Tampón GreenGoTaq 5x	5 μl
SE002 (10 μM)	0,75 μl
SE003 (10 μM)	0,75 μl
dNTP (20 mM)	0,5 μl
Polimerasa GoTaq	0,25 μl
Agua MQ	16,75 μl
	25 μl

Perfil de PCR

- 5
- | | |
|----------------|--------|
| 95 °C - 5 min | } 5 x |
| 95 °C - 1 min | |
| 58 °C - 1 min | |
| 72 °C - 2 min | } 25 x |
| 93 °C - 30 s | |
| 58 °C - 30 s | |
| 72 °C - 1 min | |
| 72 °C - 10 min | |

Después de la amplificación por PCR, se digiere el fragmento de PCR con digestión por AlwI (3 h de incubación a 37 °C) con 10 μl de molde; 1 μl de enzima AlwI; 2 μl de tampón de restricción NEB 4; 7 μl de agua MQ. Los fragmentos resultantes se analizan en gel de TAE al 1,5 % teñido con EtBr.

- 10 Los tamaños previstos de las bandas para el fragmento de PCR específico del alelo de subgenoma A son: 479 + 118 + 59 pb. Los tamaños previstos de las bandas para el fragmento de PCR específico del alelo de subgenoma D son: 538 + 118 pb.

15 **Ejemplo 2: expresión diferencial de los alelos específicos de subgenoma A y D que codifican una β-1,3-glucanasa específica de fibra y correlación con las fases de desarrollo de la fibra en *Gossypium hirsutum***

Se analizó la expresión específica alélica en todo el perfil de crecimiento/ desarrollo de las fibras para *G. hirsutum*. Se extrajo ADN de las colecciones de ADNc creadas a partir de células fibrosas y la semilla 5 DPA y a partir de células fibrosas 10, 15, 20 y 40 DPA, se igualó la concentración y se realizó la amplificación por PCR mencionada anteriormente. Las diferencias en las intensidades de banda corresponden a las diferencias relativas de la expresión (figura 6, carriles 2, 4, 6, 8 y 10. Se incluyó un control positivo (ADN genómico) (figura 6, carril 12). Se realizó una digestión con AlwI como se describe en el ejemplo 1 para distinguir entre los 2 alelos subgenómicos de los genes GhGluc1 (figura 6, carriles 3, 5, 7, 9, 11 y 13).

25 El perfil de expresión se puede resumir de la siguiente manera:

5 dpa	10 dpa	15 dpa	20 dpa	30 dpa	40 dpa
/	/	D	D	A y D	A y D

30 Durante la fase de elongación rápida en el desarrollo de la fibra (5 y 10 dpa) no hay expresión de *GhGluc1*. Por lo tanto, no se degrada el tapón de callosa, no se libera la turgencia de la célula y continúa la elongación. Se cree que la banda débil 5 dpa indica algo de expresión en la semilla (no en la fibra). En la transición entre la fase de elongación y la fase de formación de la pared celular secundaria (15 DPA), se expresa *GhGluc1*. Esto libera la turgencia de la célula y se detiene la elongación rápida. Sólo se expresa la variante de tipo genoma D del gen *GhGluc1*. Se observa más expresión 20 DPA que 15 DPA (banda más intensa). La variante de tipo genoma A no se expresa ni 15 ni 20 DPA. La variante de tipo genoma A, así como la variante de tipo genoma D, se expresa 30 y 40 DPA y podría desempeñar un papel en la maduración o en otras propiedades de las fibras.

40 **Ejemplo 3: aislamiento e identificación de la región promotora de los alelos específicos de subgenoma A y D que codifican una β-1,3-glucanasa selectiva de fibra**

Se usaron fragmentos de PCR que comprendían la secuencia de nucleótidos de los alelos específicos subgenómicos A y D de GhGluc1 para cribar una colección BAC que contenía clones de ADN genómico de la variedad *Gossypium hirsutum*. Se identificaron 4 clones diferentes, 2 para cada variante subgenómica. Se identificó la secuencia de nucleótidos de los fragmentos genómicos para cada una de las variantes alélicas y se representa en la SEC ID NO 9 (genoma A) y la SEC ID NO 10 (genoma D).

Para la variante de genoma A se pudo identificar una caja TATA en las posiciones de 2278 a 2281 y un sitio de iniciación de la transcripción en la posición 2308. La secuencia líder no traducida en 5' se extiende desde el nucleótido 2308 hasta el 2409; el codón de inicio de la traducción se sitúa en las posiciones de 2410 a 2412. La secuencia codificante consiste en dos exones (nt 2410 a 2443 y nt 2556 a 3499) separados por una secuencia intrónica (2444 a 2555). El codón de detención de la traducción se sitúa en las posiciones de 3497 a 3499 y el sitio de poliadenilación se sitúa en la posición 3624.

Para la variante de genoma D se pudo identificar una caja TATA en las posiciones de 3242 a 3245 y un sitio de iniciación de la transcripción en la posición 3270. La secuencia líder no traducida en 5' se extiende desde el nucleótido 3270 hasta el 3372; el codón de inicio de la traducción se sitúa en las posiciones de 3373 a 3375. La secuencia codificante consiste en dos exones (nt 3373 a 3406 y nt 3501 a 4444) separados por una secuencia intrónica (nt 3407 a 3500). El codón de detención de la traducción se sitúa en las posiciones de 4442 a 4444 y el sitio de poliadenilación se sitúa en la posición 4566.

Se ha alineado la secuencia de nucleótidos de las regiones codificantes fusionadas (figura 2) y se indican las diferencias; de forma similar, se han alineado las secuencias de nucleótidos de los intrones (figura 3). La figura 4 muestra una alineación de las proteínas codificadas, mientras que la figura 1 muestra una alineación de las secuencias de nucleótidos situadas aguas arriba de la región codificante.

Ejemplo 4: construcciones de genes quiméricos que comprenden diferentes regiones promotoras selectivas de fibra enlazadas de forma funcional a un gen marcador

Se enlazaron de forma funcional los siguientes fragmentos de ADN con técnicas recombinantes estándar:

- a. La secuencia de nucleótidos del promotor específico del alelo A que comprende la secuencia de nucleótidos de SEC ID NO 9 desde el nucleótido 465 hasta el nucleótido 2409.
- b. Una región codificante de β -glucuronidasa (GUS)
- c. Un fragmento que comprende una señal de terminación de la transcripción y de poliadenilación (CaMV) en 3' 35S (indicado en lo sucesivo como Gluc1-SGA (A1.9))
- y
- d. La secuencia de nucleótidos del promotor específico del alelo A que comprende la secuencia de nucleótidos de SEC ID NO 9 desde el nucleótido 1374 hasta el nucleótido 2409.
- e. Una región codificante de β -glucuronidasa (GUS)
- f. Un fragmento que comprende una señal de terminación de la transcripción y de poliadenilación (CaMV) en 3' 35S.
- y
- g. La secuencia de nucleótidos del promotor específico del alelo A que comprende la secuencia de nucleótidos de SEC ID NO 9 desde el nucleótido 1531 hasta el nucleótido 2409.
- h. Una región codificante de β -glucuronidasa (GUS)
- i. Un fragmento que comprende una señal de terminación de la transcripción y de poliadenilación (CaMV) en 3' 35S.
- y
- j. La secuencia de nucleótidos del promotor específico del alelo D que comprende la secuencia de nucleótidos de SEC ID NO 10 desde el nucleótido 1397 hasta el nucleótido 3372 (indicada en lo sucesivo como Gluc1-SGD (D2.0))
- k. Una región codificante de β -glucuronidasa (GUS)
- l. Un fragmento que comprende una señal de terminación de la transcripción y de poliadenilación (CaMV) en 3' 35S
- y
- m. La secuencia de nucleótidos del promotor específico del alelo D que comprende la secuencia de nucleótidos de SEC ID NO 10 desde el nucleótido 2371 hasta el nucleótido 3372.
- n. Una región codificante de β -glucuronidasa (GUS)
- Un fragmento que comprende una señal de terminación de la transcripción y de poliadenilación (CaMV) en 3' 35S
- o
- o. La secuencia de nucleótidos del promotor específico del alelo D que comprende la secuencia de nucleótidos de SEC ID NO 10 desde el nucleótido 2718 hasta el nucleótido 3372.
- p. Una región codificante de β -glucuronidasa (GUS)
- q. Un fragmento que comprende una señal de terminación de la transcripción y de poliadenilación (CaMV) en 3' 35S

Se clonaron los genes quiméricos anteriores por separado en un vector de ADN-T en presencia de un gen marcador seleccionable quimérico y se introdujeron en una cepa de *Agrobacterium tumefaciens* que comprendía un plásmido colaborador Ti desarmado. Las cepas de *Agrobacterium* se usaron para generar plantas de algodón transgénicas según el procedimiento descrito en el documento WO00/71733. Las células fibrosas y otros tejidos de estas plantas se someten a tinción histoquímica.

La expresión de las diferentes construcciones químéricas se observa predominantemente en las células fibrosas y las fibras.

Ejemplo 5: construcciones de genes químéricos que comprenden diferentes regiones promotoras selectivas de fibra enlazadas de forma funcional a la región codificante de la N-acetilglucosamina transferasa

Se ensamblaron construcciones similares a las construcciones del ejemplo 4, pero en las que se ha intercambiado la región codificante Gus por una región codificante NodC (véase el documento WO2006/136531, específicamente las SEC ID 1 a 9) precedida por una secuencia líder no traducida Cab22, con técnicas de ADN recombinante actuales.

Se clonan estos genes químéricos por separado en un vector de ADN-T en presencia de un gen marcador seleccionable químérico y se introducen en una cepa de *Agrobacterium tumefaciens* que comprende un plásmido colaborador Ti desarmado. Las cepas de *Agrobacterium* se usan para generar plantas de algodón transgénicas según el procedimiento descrito en el documento WO00/71733. Se analizan las células fibrosas y otros tejidos de estas plantas para determinar la presencia de oligosacáridos con carga positiva, como se describe en el documento WO2006/136531.

Ejemplo 7: análisis de la especificidad de los promotores de gluc1 por las fibras

Se aisló ARN de hojas, raíces y tallos de algodón y se analizó para determinar la expresión de genes gluc1 y del gen de la proteína fosfatasa 2A (una serina/treonina fosfatasa ubicua y conservada con amplia especificidad de sustrato y diversas funciones celulares). La síntesis de la primera hebra de ADNc se realizó con el sistema SuperScript First Strand Synthesis para RT-PCR (Invitrogen), después de lo cual se añadieron los cebadores específicos de gluc1 o pp2A y se realizó la reacción de RT-PCR. Los resultados se visualizan en la figura 7. El resultado positivo del carril 17 indicaba que se había completado con éxito la síntesis de la primera hebra. Los resultados positivos de los carriles 10, 12, 14 con cebadores específicos de pp2A en muestras de ARN de hoja, raíz y tallo indicaba específicamente que la síntesis de la primera hebra en la muestra de hoja, raíz y tallo había funcionado. Los carriles 11, 13, 15 son controles negativos en los que se omite la transcriptasa inversa de la reacción. El carril 8 es un control positivo que consiste en ADN genómico para comprobar que la reacción de PCR se ha realizado correctamente. En las líneas 2, 4, 6 no se puede detectar ninguna señal que indique la ausencia de expresión de Gluc1 en hojas, raíces y tallos de algodón. Los carriles 3, 5, 7 son controles negativos en los que se omite la transcriptasa inversa.

En conclusión, no se pudo detectar expresión de Gluc1 en el tejido de hoja, raíz y tallo de algodón.

Ejemplo 8: transformación de óvulos de algodón con cepas de Agrobacterium que contienen el

Se aislaron óvulos de flores de algodón 0 DPA, 1 DPA y 2 DPA y se usaron para experimentos de cultivos de óvulos (como se describe, por ejemplo, por Feng y Brown, 2000, *In vitro Cellular & Developmental Biology*, volumen 36, número 4, páginas 293-299) después de la transformación con *Agrobacterium tumefaciens* portadoras de los diferentes genes químéricos de fusión del promotor de gluc1::GUS indicados en el ejemplo 4 como Gluc1-SGA (A1.9) y Gluc1-SGD (D2.0). Se sometieron las células fibrosas/células iniciales a un análisis histoquímico para determinar la expresión de GUS después de 1 semana, 2 semanas, 3 semanas, 4 semanas o 5 semanas. Como control positivo se incluyeron óvulos cultivados que se habían infectado con una cepa de *Agrobacterium* que albergaba un gen químérico de fusión del promotor CaMV35::Gus. Obsérvese que la región codificante GUS usada contenía una secuencia intrónica para evitar la expresión accidental de GUS en las agrobacterias infectantes. Los resultados se resumen en la tabla 1.

Tabla 1: análisis histoquímico de Gus en los óvulos de algodón cultivados y transfectados

	7 DPA	14 DPA	21 DPA	28 DPA	35 DPA
CaMV35S-gus	+	+	+	+	+
Gluc1-SGA-gus	-	-	-	-	+
Gluc1-SGD-gus	-	+	+	+	-

Estos datos corroboran los perfiles de expresión deducidos para los promotores de los diferentes alelos subgenómicos para GhGluc1 a partir de los datos del ejemplo 2.

LISTADO DE SECUENCIAS

- <110> Bayer CropScience NV Engelen, Steven Arioli, Tony
- <120> Expresión diferencial de alelos específicos de subgenoma en algodón y usos de los mismos.
- <130> BCS 07-2001
- <150> EP07000550.9

<151> 11-01-2007

<150> US60/884.564

<151> 11-01-2007

5 <160> 12

<170> PatentIn versión 3.5

10 <210> 1
<211> 689
<212> ADN
<213> *Gossypium hirsutum*

15 <400> 1

tggccgtgat attggtggtt gctatgggtt gaacggcaac aatcttccat ctccaggaga	60
tgttattaat cttttcaaaa ctagtggcat aaacaatatc aggctctacc agccttacct	120
tgaagtgctc gaagcagcaa ggggatcggg aatatccctc tcgatgagta cgacaaacga	180
ggacatacaa agcctcgcga cggatcaaag tgcagccgat gcatgggtta acaccaacat	240
cgtcccttat aaggaagatg ttcaattcag gttcatcadc attgggaatg aagccattcc	300
aggacagtca agctcttaca ttcttgggtc catgaacaac ataatgaact cgctggcctc	360
atltgggcta ggcacgacga aggttacgac cgtgggtccc atgaatgccc taagtacctc	420
gtaccctcct tcagacggcg cttttggaag cgatataaca tcgatcatga ctagtatcat	480
ggcattctg gttcgcagc attcgcctc cctgatcaat gtgtaccctt attttgccta	540
tgcctcagac cccactcata tttccctcaa ctacgccttg ttcacctcga ccgcaccggt	600
ggtggtcgac caaggcttgg aatactacaa cctctttgac ggcattggtc atgctttcaa	660
tgccgcccta gataagatcg gcttcggcc	689

<210> 2

20 <211> 689
<212> ADN
<213> *Gossypium hirsutum*

25 <400> 2

tggccgtgat attggtggtt gctatgggtt gaacggcaac aatcttccat ctccaggaga	60
tgttattaat ctttacaaaa ctagtggcat aaacaatatc aggctctacc agccttacct	120
tgaagtgctc gaagcagcaa ggggatcggg aatatccctc tcgatgggtc cgagaaacga	180
ggacatacaa agcctcgcga aagatcaaag tgcagccgat gcatgggtta acaccaacat	240
cgtcccttat aaggacgatg ttcagttcaa gttgatcact attgggaatg aagccatttc	300
aggacaatca agctcttaca ttcttggatc catgaacaac ataatgaact cgctcgcctt	360
atltgggtta ggcacgacga aggttacgac cgtgggtccc atgaatgccc taagtacctc	420
gtaccctcct tcagacggcg cttttggaag cgatataaca tcgatcatga ctagtatcat	480
ggcattctg gctgtacagg attcgcctc cctgatcaat gtgtaccctt attttgccta	540
tgcctcagac cccactcata tttccctcga ttacgccttg ttcacctcga ccgcaccggt	600
ggtggtcgac caaggcttgg aatactacaa cctctttgac ggcattggtc atgctttcaa	660
tgccgcccta gataagatcg gcttcggcc	689

ES 2 548 763 T3

	<210> 3		
	<211> 655		
	<212> ADN		
	<213> <i>Gossypium arboreum</i>		
5	<400> 3		
	cggcaacaat cttccatctc caggagatgt tattaatctt tacaaaacta gtggcataaa	60	
	caatatcagg ctctaccagc cttacctgaa gtgctcgaag gagcaagggg atcggaata	120	
	tccctctcga tgagtacgac aaacgaggac atacaaagcc tcgcaacgga tcaaagtgca	180	
	gccgatgcat gggttaacac caacatcgtc cttataagg acgatgttca attcaggttc	240	
	atcatcattg ggaatgaagc cattccagga cagtcaagct cttacattcc tggtgccatg	300	
	aacaacataa tgaactcgct cgcctcattt gggctaggca cgacgaaggt tacgaccgtg	360	
	gtcccgatga atgccctaag tacctcgtac cctcctcag acggcgcttt tggaagcgat	420	
	ataacatcga tcatgactag tatcatggcc attctggttc gacaggattc gccctcctg	480	
	atcaatgtgt acccttattt tgcctatgcc tcagaccca ctcattttc cctcaactac	540	
	gccttgttca cctcgaccgc accggtggtg gtcgaccaag gcttggaata ctacaacctc	600	
	tttgacggca tggtcgatgc tttcaatgcc gccctagata agatcggctt cggcc	655	
10	<210> 4		
	<211> 656		
	<212> ADN		
	<213> <i>Gossypium raimondii</i>		
15	<400> 4		
	cggcaacaat cttccatctc caggagatgt tattaatctt tacaaaacta gtggcataaa	60	
	caatatcagg ctctaccagc cttaccctga agtgctcga gacgcaaggg gatcgggaat	120	
	atccctctcg atgggtccga gaaacgagga catacaaagc ctcgcaaaag atcaaagtgc	180	
	agccgatgca tgggttaaca ccaacatcgt cccttataag gacgatgttc agttcaaatt	240	
	gatcactatt gggaatgaag ccatttcagg acaatcaagc tcttacattc ctgatgccat	300	
	gaacaacata atgaactcgc tcgcctcatt tgggttaggc acaacgaagg ttacgaccgt	360	
	ggtcccgatg aatgccctaa gtacctcgt cctccttca gacggcgctt ttggaagcga	420	
	tataacatcg atcatgacta gtatcatggc cattctggct gtacaggatt cgccccctc	480	
	gatcaatgtg tacccttatt ttgcctatgc ctcagacccc actcatattt ccctcgatta	540	
	cgcttggtc acctcgaccg caccggtggt ggtcgaccaa ggcttggaat actacaacct	600	
	ctttgacggc atggtcgatg ctttcaatgc cgcctagat aagatcggct tcggcc	656	
20	<210> 5		
	<211> 23		
	<212> ADN		
	<213> Artificial		
25	<220>		
	<223> cebador oligonucleótido para la amplificación de fragmentos genómicos de endoglucanasas		
	<400> 5		
30	ggccgaagcc gatcttatct agg	23	

ES 2 548 763 T3

<210> 6
 <211> 23
 <212> ADN
 <213> Artificial
 5
 <220>
 <223> cebador oligonucleótido para la amplificación de fragmentos genómicos de endoglucanasas
 <400> 6
 10 cggaacaat cttccatctc cag 23
 <210> 7
 <211> 229
 <212> PRT
 15 <213> *Gossypium hirsutum*
 <400> 7
 Gly Arg Asp Ile Gly Val Cys Tyr Gly Leu Asn Gly Asn Asn Leu Pro
 1 5 10 15
 Ser Pro Gly Asp Val Ile Asn Leu Phe Lys Thr Ser Gly Ile Asn Asn
 20 25 30
 Ile Arg Leu Tyr Gln Pro Tyr Pro Glu Val Leu Glu Ala Ala Arg Gly
 35 40 45
 Ser Gly Ile Ser Leu Ser Met Ser Thr Thr Asn Glu Asp Ile Gln Ser
 50 55 60
 Leu Ala Thr Asp Gln Ser Ala Ala Asp Ala Trp Val Asn Thr Asn Ile
 65 70 75 80
 Val Pro Tyr Lys Glu Asp Val Gln Phe Arg Phe Ile Ile Ile Gly Asn
 85 90 95
 Glu Ala Ile Pro Gly Gln Ser Ser Ser Tyr Ile Pro Gly Ala Met Asn
 100 105 110
 Asn Ile Met Asn Ser Leu Ala Ser Phe Gly Leu Gly Thr Thr Lys Val
 115 120 125
 Thr Thr Val Val Pro Met Asn Ala Leu Ser Thr Ser Tyr Pro Pro Ser

ES 2 548 763 T3

130 135 140
 Asp Gly Ala Phe Gly Ser Asp Ile Thr Ser Ile Met Thr Ser Ile Met
 145 150 155 160
 Ala Ile Leu Val Arg Gln Asp Ser Pro Leu Leu Ile Asn Val Tyr Pro
 165 170 175
 Tyr Phe Ala Tyr Ala Ser Asp Pro Thr His Ile Ser Leu Asn Tyr Ala
 180 185 190
 Leu Phe Thr Ser Thr Ala Pro Val Val Val Asp Gln Gly Leu Glu Tyr
 195 200 205
 Tyr Asn Leu Phe Asp Gly Met Val Asp Ala Phe Asn Ala Ala Leu Asp
 210 215 220
 Lys Ile Gly Phe Gly
 225

<210> 8
 <211> 229
 <212> PRT
 <213> *Gossypium hirsutum*

<400> 8

Gly Arg Asp Ile Gly Val Cys Tyr Gly Leu Asn Gly Asn Asn Leu Pro
 1 5 10 15
 Ser Pro Gly Asp Val Ile Asn Leu Tyr Lys Thr Ser Gly Ile Asn Asn
 20 25 30
 Ile Arg Leu Tyr Gln Pro Tyr Pro Glu Val Leu Glu Ala Ala Arg Gly
 35 40 45
 Ser Gly Ile Ser Leu Ser Met Gly Pro Arg Asn Glu Asp Ile Gln Ser
 50 55 60
 Leu Ala Lys Asp Gln Ser Ala Ala Asp Ala Trp Val Asn Thr Asn Ile
 65 70 75 80
 Val Pro Tyr Lys Asp Asp Val Gln Phe Lys Leu Ile Thr Ile Gly Asn
 85 90 95
 Glu Ala Ile Ser Gly Gln Ser Ser Ser Tyr Ile Pro Asp Ala Met Asn
 100 105 110
 Asn Ile Met Asn Ser Leu Ala Leu Phe Gly Leu Gly Thr Thr Lys Val
 115 120 125
 Thr Thr Val Val Pro Met Asn Ala Leu Ser Thr Ser Tyr Pro Pro Ser
 130 135 140

5

10

ES 2 548 763 T3

Asp Gly Ala Phe Gly Ser Asp Ile Thr Ser Ile Met Thr Ser Ile Met
145 150 155 160

Ala Ile Leu Ala Val Gln Asp Ser Pro Leu Leu Ile Asn Val Tyr Pro
165 170 175

Tyr Phe Ala Tyr Ala Ser Asp Pro Thr His Ile Ser Leu Asp Tyr Ala
180 185 190

Leu Phe Thr Ser Thr Ala Pro Val Val Val Asp Gln Gly Leu Glu Tyr
195 200 205

Tyr Asn Leu Phe Asp Gly Met Val Asp Ala Phe Asn Ala Ala Leu Asp
210 215 220

Lys Ile Gly Phe Gly
225

- 5 <210> 9
<211> 6009
<212> ADN
<213> *Gossypium hirsutum*
- 10 <220>
<221> misc_feature
<222> (2278)..(2281)
<223> caja TATA
- 15 <220>
<221> misc_feature
<222> (2308)..(2308)
<223> sitio de inicio de la transcripción
- 20 <220>
<221> misc_feature
<222> (2308)..(2409)
<223> secuencia líder no traducida 5'
- 25 <220>
<221> misc_feature
<222> (2410)..(2412)
<223> codón de inicio de la traducción (AGT.)
- 30 <220>
<221> misc_feature
<222> (2410)... (2443)
<223> secuencia codificante parte I
- 35 <220>
<221> misc_feature
<222> (2444)... (2555)
<223> secuencia intrónica
- 40 <220>
<221> misc_feature
<222> (2556)..(3499)
<223> secuencia codificante parte II
- 45 <220>
<221> misc_feature
<222> (3497)..(3499)

ES 2 548 763 T3

<223> codón de terminación de la traducción

<220>

<221> misc_feature

5 <222> (3624)..(3624)

<223> sitio de poliadenilación

<400> 9

```

cgcgccatat aatTTTTatgt gtgtaatttg ttgggttaat tacttaaaat agtatatTTT      60
taattgctgt aattaatgta agataatTTT tattatTTTga atcattgcac aaaattaaaa      120
tagaataatt tattaaaca ttcaaataata ataataatcc aaattataat tatagtatTT      180
ttacaatatt caatatacaa tatagTTTTa cttcatacaa ttaatataaa aaaatattat      240
tcaaaataat aactaataaa cataattacc atatattaat tattTTTgata tttcgaacat      300
aacgctaata aaaaatttcc taatcattat taaatcattt gtataaacta taaagaaatt      360
gatatattgt aaattaaact ttattcattt tttttcttaa tactcaataa attaatcata      420
ataactcata aataatatat aattaaaata atcataacat ggtagattat ataaataggg      480
ggcgaatcta gggagctggc atgacccta aaatagaatt ttctatTTTg acctatcaaa      540
atTTTTaaaa ttttaaatta gtaaaggtaa atTTTgtactt tgacctctta aaatgataaa      600
atTTTacttt aatcTTTTaa aatttacatt ttactatca taaaaattac aatttgattt      660
tgcccctaaa atTTTTTTct agcttagccc tgtatataaa tatattatTT ataatTTTTa      720
tatttaaaat ataaagTTTT taattataca aataattaa atctgatatt taaaactaaa      780
gtaatttctt ttttctTTTT actTTTTTTT aattgcaaca taatggTTTa aatatctata      840
taacgtatga agtaattTga tataaatTTT atTTTaatTT attattatat aaattcattt      900
agtaaaaact tttaatagaa tcaaaatTTT tattTgTaaa ttcgataact tttcttatca      960
agtatattTg tgagaaccaa atatttagta aaattaatat tcttattTat aaatatgata      1020
aatcttataa aaaaatattt aaaaTgaaa aaattgtaca aatattataa aaaaatattt      1080
aaaatgaaa acattgtaca aaggctatat aagaagtTca aaagtttctt cgaccatgta      1140
ctcttataga gattatagat agattataaa actatatgta gtttctctta actTTTaaat      1200
aagaggataa atgtatTTTt atgtactcaa acttatatat ttttatattg acaataatat      1260
caatatcaac ctaattaaga ttcatTctaa catTaatgTt gaagatTTTT aataaaaagaa      1320
aaggTtaata aattaattag aacacaaaca aacacaaatt taagtggTat gtaaggTcct      1380
tgaccCAAag gaaaaattTg ttacgtcgat taaattataa attaatTTaa agtaaaaatta      1440
cattTTaacc taaaaaaaga gaaaagtata tctaatttct tcgaaaatgg aaagaaaatt      1500
ataaatTTat ggcattTcta aaaaatttct gaattcgcta ctaaaagatg aattataaaa      1560
atccgaagca ttaccagaag atggatcacc aaatcacaaa caatcaatga aaagTaatga      1620
taattaattg aaagtgagca tTTaatTTTg atagccatat acttctgct gaatttatag      1680
gttctcatta atgcaattaa attatattcg acacctTTTg aatgaaataa aatgacacaa      1740
gaggaaagac ggtTcatcta tttttctTTT caatcgccca tcaaaatacc aaaaatgTaa      1800
ctacatgcaa aaaatcaaat atgaaaaata tTcatatTTT gatattTTTaa tatattgtgt      1860

```

10

ES 2 548 763 T3

gttcaaaacg taaatgtatt gaaaaattat gatggtgttg ttgctgtatg tccataaaat 1920
 tcaatgtact cacatttatc aaatgtatac tttgagagaa gttatTTTTga taataactcaa 1980
 gtttttttta tagatgggaa aatttttttaa attatTTTTt gattttgatg aaatgtatat 2040
 ataaatttta attcgataca tataaatata tatgtaaatt ttaaatttaa atttaataat 2100
 atacaattaa gaaaataatt tataaatatt ttccgattaa aaataaatct ggaaagaaga 2160
 aatgtcaaca ctttttcatt aaatacaatt aggatgggac acgatacctt catgcattga 2220
 tatctcaggt ggtccaaaaa ctcggaatcc tttttgaaa aaaacttcca gagagagtat 2280
 ataaatccag cagtaggcac aagaaacgag caccagtat tgactttcct ttgtaaaaaa 2340
 aaaaagtgct gagatcaaga aatatagtga aatatgggtc caagattttc tgggttttta 2400
 atctaagcaa tgctgttttt aactcaactc ctctctctaa caggtaaaac aaacttctct 2460
 acagtgattt tacagtaa atggctttga aaaatataca acaaaacatt tatcttcaat 2520
 ccattttaat tactgatcta ctatatatgt tgcagatggc cgtgatattg gtgtttgcta 2580
 tggtttgaac ggcaacaatc ttccatctcc aggagatggt attaatcttt tcaaaactag 2640
 tggcataaac aatcaggc tctaccagcc ttaccctgaa gtgctcgaag cagcaagggg 2700
 atcgggaata tccctctcga tgagtacgac aaacgaggac atacaaagcc tcgcaacgga 2760
 tcaaagtgca gccgatgcat gggttaacac caacatcgtc ccttataagg aagatgttca 2820
 attcaggttc atcatcattg ggaatgaagc cattccagga cagtcaagct cttacattcc 2880
 tgggtgccatg aacaacataa tgaactcgtt ggcctcattt gggctaggca cgacgaaggt 2940
 tacgaccgtg gtcccgatga atgccctaag tacctcgtac cctccttcag acggcgcttt 3000
 tgggaagc gatatacga tcatgactag tatcatggcc attctgggtc gacaggattc 3060
 gccctcctg atcaatgtgt acccttattt tgcctatgcc tcagacccca ctcatatttc 3120
 cctcaactac gccttgttca cctcgaccgc accgggtggtg gtcgaccaag gcttgggaata 3180
 ctacaacctc tttgacggca tggtcgatgc tttcaatgcc gccctagata agatcggctt 3240
 cggccaaatt actctcattg tagccgaaac tggatggccg accgcccgtta acgagcctta 3300
 cacgagtgtc gcgaacgctc aaacttataa caagaacttg ttgaatcatg tgacgcagaa 3360
 agggactccg aaaagacctg aatatataat gccgacgttt ttcttcgaga tgttcaacga 3420
 gaacttgaag caaccacag ttgagcagaa tttcggattc ttcttcccca atatgaacct 3480
 tgtttatcca ttttggtgaa cttgaaatgt tattgttggc tatttaaate ttttgccaga 3540
 gacgcttcat atagtctctg catatTTTga aagtggaaaa tcaatctaaa tataaataag 3600
 ttttatttgt tgttttttaa ttaaataaaa ttttaaataat tttaaaaaca tctttattgg 3660
 taattaaata ttaaataaaa agtttaatat tcaaatttta tcaattcaaa aataaaataa 3720
 aaatatatta aatttatttt tacgaataaa ttgattttct attaatgcag attttaaata 3780
 atttgatata aattttcaat tcaacaatag taattttgat cacatcaaag gagaaagggg 3840
 aagatttaac ttttaattggt gacctaatat aacacgttga aaacggagtt cccaataagg 3900

ES 2 548 763 T3

caaaatgact tgtaatgacg aaagagatgt ccaagtgaaa tctgctttaa agtgaaagaa 3960
gcataaaagg ataactaaat aactcatgat ctaaattgaa gttctataaa atgcaacttt 4020
catctagaaa caaggatgt cttaaagatg gttttatgaa tttgtcttaa ttgggtttta 4080
tgcaatgaat tcatggatag cacatctcta attatacgtt gctggtttat atgagagtgg 4140
tgcagaagtt aattgtgctt taaatacttg cttagtgttt atgaaatttg aaaagtgtta 4200
tatacttata ataaaaataa ttcgattcgg aatccaattc agggttcgcac tcaatataat 4260
aaaattttac agatatcttg aaggggatct tcttcttctc tacttctcga gcagtgttat 4320
atatttaca taaagataac tcaattcgag atccgacctc atataataaa attctacaga 4380
catatcaaag aggggatct tcttcttccc tacatcttga ccttcttgat caaaatgacc 4440
ttccttatat ttttacatac gttgattata tgaatcaaaa gaaagatacc aaaaagtttt 4500
taaaaataaa caacgggggtt cttatgtaga gatgcttatg ggccgggccc gactcaacta 4560
aaaatttagg cacattcatt gggcccaggt cgggcctaac caaaaatgg gcctaaaatt 4620
ttgcccaagc ttgactcaaa taaaaatgct aaaattcggg cctgaccccg tattaatttt 4680
atattatttt atataacttt taaatatata taatatataa aaaatactaa aaaaattaaa 4740
ataaatattt cccaactaaa ctaaaattat taagaaaaat aattcatatt agcgtataaa 4800
ttggaaattg accaaaatta aaattattgt atagttaatc tatattaaaa ggacatgtaa 4860
ttaaaaacca ttaaaactat tatacaataa attaaatctt cattgtatac atagaaaggc 4920
attaataatt aaaaaactat attaagatat aaactaaatt caaaattatt aaaaacaaga 4980
actaaataaa aaagcaattg aaaattacga attaatgtta aatcaaatg ttaaaatcaa 5040
gggacttaaa taaaaatata ccaaaatata aaacattagc ttcctttccc atccacgtga 5100
atgcaaagtt tacatgggtg ttctagtgt ttgtgcgact ccaacctttt atttacctct 5160
tttttcttt atttgaacaa ttatttgata atgattagaa ttttgggatt gttgctcctc 5220
gtacgtgcaa cacttaaaat cactatgatt tttcataatt tatataacct atacgtttt 5280
ggaaattaat tttatttttt atattatttt aataaaaata ccatctacct tttttaattt 5340
atgatccctt tcatatttaa aaattcaaat tgacaattgt ctaactaac accgtcacac 5400
tccaataaga ttgtaatttc ctccatcttg atattacact caaaagcatg ttgccaaaca 5460
acaaatcaac tagccttttt ctaccactat tcatcatctt cttagagtg tgtttatgtc 5520
atgtgccgag attttaggta tggtcacgtt gtggctttaa actcaaatct attgcccatg 5580
agtctaagtt agcctccgat ctcactaaa gagaggcttg gcacacttta cctagccaag 5640
tacacaagga atagagctat tagaaagcat taaagagtta ggagaatgtg gaagtgtttt 5700
tattactcaa agctaacttg gatacaataa aaggagggag cctctccttt aggcaagctt 5760
cttttgatct gatggttaca attaatctcg aataggaggg gtcaaacttc tcaactcagtt 5820
tcatattatc tcttgggtgct tggttggcct ccgcttgag acaactttag ataacaccta 5880
gtcttaacac ttttagcttc acattgtacg catccttcat tactcaaatg ccacaaagcc 5940
tccttactta aggcctttgg tcgctcccac taccttcggc tttagactca tctaagatct 6000
tcccaatcg 6009

<210> 10
 <211> 6877
 <212> ADN
 <213> *Gossypium hirsutum*
 5
 <220>
 <221> misc_feature
 <222> (3242)..(3245)
 <223> caja TATA
 10
 <220>
 <221> misc_feature
 <222> (3270)..(3270)
 <223> sitio de inicio de la transcripción
 15
 <220>
 <221> misc_feature
 <222> (3270)..(3372)
 <223> secuencia líder no traducida 5'
 20
 <220>
 <221> misc_feature
 <222> (3373)..(3375)
 <223> codón de inicio de la traducción (AGT.)
 25
 <220>
 <221> misc_feature
 <222> (3373)..(3406)
 <223> secuencia codificante parte I
 30
 <220>
 <221> misc_feature
 <222> (3407)..(3500)
 <223> secuencia intrónica
 35
 <220>
 <221> misc_feature
 <222> (3501)..(4444)
 <223> secuencia codificante parte II
 40
 <220>
 <221> misc_feature
 <222> (4442)..(4444)
 <223> codón de terminación
 45
 <220>
 <221> misc_feature
 <222> (4566)..(4566)
 <223> sitio de poliadenilación
 50
 <400> 10
 ttcaaactta ctcgcttgca caaaaataat tttataaaag tattttaaatt ataaaatttt 60
 tatgttgata atatTTTTat atacatttta tattttaaga aataataatt ttttaggaat 120
 tagaaaaaaa atgtagaata atatcattga tattttaatt tttcaaaaaa ttaaaaataa 180
 gttcacgtag tctaatttta tctattttta tttttatact ttcaaattga gataaatatc 240
 aaagaacttt tggttcaata tgcaatttga tacttaaatt ttaatttgat gtaattatta 300
 catgaaactt ggcttgggt ttatacgtat acatgaaatt tttattttga ttcaattgta 360
 cgcatttaaa gaaatgaaaa tggtttctaat tcaataatat tattagtgat ttgtgaaatt 420

ES 2 548 763 T3

taaaactttt atgcattaaa ccacacaaaa tcagagttta tgtatgatat tgcacattgg	480
actatagttc atgcatattt tttatatattt atccatgtca aattttgaaa tttcattcctt	540
aacttatatg atagcagtta aatttgttaa gtcaaactct agtattagtt atatactata	600
cataacttgt agagtttagt ttaagttcac taatttgatt attttttattc tgtttatttt	660
ttcaatttca agatttaagt ttaagctta acttaacaa tagtcattaa atttattaac	720
taaaatgtcc tggggttttt tgtaagtatt ataatatgtt tgccacgtga gattttggta	780
aaagtagagt ttaacttaac aaatttaatg gctactactt agtaaggatt agaatttcaa	840
aattaaaaaa aaaattatag aggctaaaga tgatcaaatt agaggtttaa attaagtcaa	900
attaanaatag ttctggatat taactattta aattaattaa tgtcattata aaattagagg	960
tctaaattat gtaaaattaa aatataaaaa ctaaactctcg aatgtgagta tagtataagg	1020
atcaaaaagtg attttgggtca ttttctttta tttacaaata ttcttagaga tgctctttta	1080
tatataatgg tttctaattg gatatgcgcy gcatataatt ttatgtgttt aatttatitt	1140
attaattatt taaataatat atttttaatt actataatta atgtaaaata ttttttatta	1200
tttgaatcac tgcacaaaat taaaatatac taacttaatt aacaattcaa atataataat	1260
aatctaaatt ataattaaag catttttaca atattcagta tataatatag ttttacttta	1320
tataattaat acaaagaaat attattcaaa ataataacta atcaacataa ttactatata	1380
tcaattatit t gatatttctg aacataatgc taataaaaaa tttcctaattc attattaaat	1440
catttgata aactataaag aaattgatat attgtaaatt aaacttttaa ctattcaatt	1500
ttttctaat agtcaataaa ttaatcataa taattcataa ttaatataa attaacataa	1560
ccataacata gaatttttta ttttggccca ttaaaatttt taaaatttta aattagtaaa	1620
ggaaaaatta cactttgacc ctttaaaaat gataaaattt tattttaatc ctttaaaatt	1680
gacattttta ctattgtaa aattacaatt taattttgcc cccttaaaaa atttttctag	1740
cttcgccctt gtgtataaat atattaatta caatttttat atttgaatta tataaataat	1800
taaattttga tattttaaac taaagtaatc tctttttttt ttactttttt ttaattgaaa	1860
cataatggtt taaatatcta tattacgtat gaagtaattt aatataaatt ttattttaat	1920
ttattattat ataaattcat ttagtaaaaa cttttaatag aatcaaaaatt tttatttgta	1980
aattcgataa cttttcttat caagtaaatt tgyttgaatta aatatttagt aaaattaata	2040
ttttatttta taaatatgat aatcttata aaaaataaaa aatattttaa aatgaaaaac	2100
attgtacaaa ggctatataa gaagttcaaa agtttcttcg accctgtact ctaatagaga	2160
ttatagatag attatagaac tattcatagt ttctcttaac ctttaataa gaattttagt	2220
gtactcaaac ttacatattt ttatattgat aataatgtca ataccagccg agttaagatt	2280
cactcgacat taatgttgaa aatttttaat aaaagaaaat gttgataagt taattagaac	2340
acaagcaagc acaaatttaa gtgtaagta aggtccttga ccctaattgga aaaattgtta	2400
tgttgattaa attataaatt aatttaaggt aaaattatat tttgacctaa aaaaatgaaa	2460

ES 2 548 763 T3

aaaatatatc tagtttcttc gaaaatgaaa agaaaataat aaattgatac attataaaat 2520
 ttatggcatt tctaaaaaaa ttctgaattt gatgaaatta taataaaaaa aaagtttaaa 2580
 aacatataga tttcaagaat agtgggaaaa ttatatattga acaacactga agaaatccaa 2640
 agcattagca gaaaatggat caccaaatca caaacaatca gtgaaaagta atgataatta 2700
 attgaaagtg agcatttaaa tttgatagcc atatacttcc tgctgaattt ataggttctc 2760
 attaatgcaa ttaaattata tttgtcactt tttgaatgaa ataaatgaca cagttcatct 2820
 attttttttc tttcaatcgc ccatcaaaat accgaaaatg taactacatt aaaaaagatc 2880
 gaaaaatatt catatatttga tattttaata gatttgtgtg tcaaggcgta atgtactaaa 2940
 aaattatgat ggtggtgtcg ctgtatgtcc ataaaattca atgtattcgc atgtatcaaa 3000
 tgtaaathtt gacacaagtt attctaataa taatcaagtt atttttatac atgagataca 3060
 tctcaaaatt atttttatat atccgaaaa tcataacgta cgatcaaact agaaagagga 3120
 agtgtcaaaa cctattcatt atatgcaaat atgatgggac acgataccct catgcattga 3180
 tatctcatat tgtccaaaaa ctcagaatcc tttttgaaaa aaaaaaattc cagagagagt 3240
 gtataaatcc agcagtggtc acaagaaacg agcaccagtt attgacattc ctttgtaaaa 3300
 aaaaaaagaa gctgagatca agaaatatag tgaatatatg gtccaacatt ttctgggttt 3360
 ttaatctcag caatgggtgtt ttttaactcaa ctctctctc taacaggtaa aacaaacttc 3420
 tctacagtga ttttacggta agtatggctt tgaaaaatat acaacaaaac atttatactg 3480
 atctaccata tatgttgcag atggccgtga tattggtgtt tgctatggtt tgaacggcaa 3540
 caatcttcca tctccaggag atgttattaa tctttacaaa actagtggca taacaatat 3600
 caggctctac cagccttacc ctgaagtget cgaagcagca aggggatcgg gaatatccct 3660
 ctgatgggt ccgagaaacg aggacataca aagcctcgca aaagatcaaa gtgcagccga 3720
 tgcattgggtt aacaccaaca tcgtccctta taaggacgat gttcagttca agttgatcac 3780
 tattgggaat gaagccattt caggacaatc aagctcttac attcctgatg ccatgaacaa 3840
 cataatgaac tcgctgcct tatttgggtt aggcacgacg aaggttacga ccgtggtccc 3900
 gatgaatgcc ctaagtacct cgtaccctcc ttcagacggc gcttttgtaa gcgatataac 3960
 atcgatcatg actagtatca tggccattct ggctgtacag gattcgcccc tctgatcaa 4020
 tgtgtaccct tattttgcct atgcctcaga cccactcat atttccctcg attacgcctt 4080
 gttcacctcg accgcaccgg tgggtggtcga ccaaggcttg gaatactaca acctcttga 4140
 cgcatggtc gatgctttca atgccgccct agataagatc ggcttcggcc aaattactct 4200
 cattgtagcc gaaactggat ggccgaccgc cggtaacgag ccttacacga gtgtcgcgaa 4260
 cgctcaaact tataacaaga acttgttaaa tcatgtgacg cagaagggga ctccgaaaag 4320
 acctgaatat ataatgccga cgtttttctt cgagatgttc aacgaggatt tgaagcaacc 4380
 cacagttgag cagaatttcg gattcttctt ccccaatatg aaccctgttt atccattttg 4440
 gtgaagtga aatggtgttg gctatttaaa tcttttgcca gagacgcttc atatagtttc 4500
 tgcatathtt gaaagtggaa aatcaatcta aatattaata agttttatgt gttgtttttt 4560

ES 2 548 763 T3

aattaaataa aattttaaat attttaaaaa tatctttatt ggtaattaa tattaaataa 4620
aaagtttaat attcaaattt tatcaattca aaaataaaat aaaaatatat taaatttatt 4680
tttacgaata aattgatttt ctattaatac agattttgaa taatttgata taaattttaa 4740
attcaacaat agtaattttg atcacatcaa aggagaaagg gaaagattta actttaattg 4800
gtgacctaat ataacacggt gaaaacggag ctcccaggaa ggcaaatga cttgtaatga 4860
cgaaagagat gtccaagtag aatctgcatt aaagtgaaaa aagcataaaa ggataagtaa 4920
actcatgatc tgacataaat tgaagtctta taaaatgcaa ctttcatcta gaaacaagg 4980
atgtcttaaa tgatgtttta tgaatttgc ttaactgggt tttatgcaat gaattcatgg 5040
atagcacctc actaattata cgttgctggt ttatatgaga gtgggtgcaga agttaattgt 5100
gctttaaata cttgcttagt gttcaagaaa tttgaaaagt attatatatt tataataaaa 5160
ataattcaga tccgactcaa tctagtaaaa ttttacaac attctaaagg ggatcttctt 5220
ttttctctac ttattgatca gtgttatata cttataataa agacaacctg atttgagatc 5280
cggcctaata taataaaatt ctacagacat ctcaagggag agatcttctt cttccctaca 5340
tcttgacctt tttgatcaaa atttctccc ctctatttcc acattgggtg atcatatgaa 5400
tcaacagaaa ggtaccaaaa agtttttaa aataaaciaa ggggttctta tgaattcat 5460
atgatataat gggctcaatt attagaatca attttaagtt taaacaaatt taaaattcaa 5520
aactcaattc catttttgtt tgaacggaaa gttactaatt gttaagaaaa ataattcata 5580
ttagcgtata aattggaaat tgaccaaaaac taaaattatt gtatagttaa tctatattaa 5640
aaggacatgt aattaaaac cattaaaact attatagaat aaattaaatc ttattctat 5700
acatacaaag tcattaataa ttaaaaaact atattaagat ataaactata ttcaaaaaat 5760
attaaaaaca ataactaat aaaaaaaca attgaaaatt acgaattaat gttaaaatca 5820
agggacttaa ataaaaatat cccaaaatac aaaacattag cttccttcc catccacgtg 5880
attgcaaagt ttacatgggt tttcctagtg cttgtgagac tccaacctt tatttacttt 5940
tttcttttct ttatttgaac aattatttga taatgattag aattttggga ttgttgctca 6000
tcgtacgtgc aacacttaa atcactatga tttttcataa tttatatac ctatatcgtt 6060
ttgaaatta atgttattat ttatattggt ttaataaaaa taccatctac ctctttaat 6120
ttatgatcca tttcttattt gaaaattcaa attgacagtt gtctaactaa acaccatcgc 6180
actccaataa aattgtaatt ttttctatcg tgaatagtag actcaaaagt atgttgtaa 6240
caaacaaatc aattagcctt tttctacctc tattcatcat cttcttaata gcgtgtttat 6300
gtcacgtggt gagattttag ttccggtcac gtgtggcctt aaaccgaat ttcttacgca 6360
tgagtctaag ttagcctctg atcctcgcta tggagatgct tggcacagtt tacctaggta 6420
agtaaacaag gaatagagct attagaaagc atcagagagt taggagaatg tggaaagtgt 6480
tctattactc aaagctaact tggatacaaa taaaagaggg agcctctcct ttaggcaagc 6540
ctattttgat ctgacgggtg caattaatct cgaataggag gggtcgaact tctcactcag 6600

ES 2 548 763 T3

tttcacatta tctcttggtg cttagttggc ctccgccttg āgacacattc aaataacacc 6660
 tagtcttaac acttttggct tcttattgtg cgtatccttc attactcaaa tgccacaaag 6720
 cctcattact taagactctc ggtcgctccc actaccttcg acttttagact catctaagat 6780
 cttcccaatc gtagacaact tggccttggg ggggaaatct tgcaccctac ggggccttac 6840
 ataagaagca attaaatggc tttctctcac ccacctt 6877

<210> 11
 <211> 325
 <212> PRT
 <213> *Gossypium hirsutum*

5

<400> 11

Met Leu Phe Leu Thr Gln Leu Leu Ser Leu Thr Asp Gly Arg Asp Ile
 1 5 10 15
 Gly Val Cys Tyr Gly Leu Asn Gly Asn Asn Leu Pro Ser Pro Gly Asp
 20 25 30
 Val Ile Asn Leu Phe Lys Thr Ser Gly Ile Asn Asn Ile Arg Leu Tyr
 35 40 45
 Gln Pro Tyr Pro Glu Val Leu Glu Ala Ala Arg Gly Ser Gly Ile Ser
 50 55 60
 Leu Ser Met Ser Thr Thr Asn Glu Asp Ile Gln Ser Leu Ala Thr Asp
 65 70 75 80
 Gln Ser Ala Ala Asp Ala Trp Val Asn Thr Asn Ile Val Pro Tyr Lys
 85 90 95
 Glu Asp Val Gln Phe Arg Phe Ile Ile Ile Gly Asn Glu Ala Ile Pro
 100 105 110
 Gly Gln Ser Ser Ser Tyr Ile Pro Gly Ala Met Asn Asn Ile Met Asn
 115 120 125
 Ser Leu Ala Ser Phe Gly Leu Gly Thr Thr Lys Val Thr Thr Val Val
 130 135 140
 Pro Met Asn Ala Leu Ser Thr Ser Tyr Pro Pro Ser Asp Gly Ala Phe
 145 150 155 160
 Gly Ser Asp Ile Thr Ser Ile Met Thr Ser Ile Met Ala Ile Leu Val
 165 170 175
 Arg Gln Asp Ser Pro Leu Leu Ile Asn Val Tyr Pro Tyr Phe Ala Tyr
 180 185 190
 Ala Ser Asp Pro Thr His Ile Ser Leu Asn Tyr Ala Leu Phe Thr Ser
 195 200 205

10

ES 2 548 763 T3

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

<210> 12
 <211> 325
 <212> PRT
 <213> *Gossypium hirsutum*

<400> 12

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

5

10

REIVINDICACIONES

1. Un promotor preferente de células fibrosas que comprende una secuencia de nucleótidos seleccionada del siguiente grupo de secuencias de nucleótidos:
- 5
- a) una secuencia de nucleótidos que comprende la secuencia de nucleótidos de SEC ID NO 9 desde el nucleótido de la posición 465 hasta el nucleótido de la posición 2307;
 - b) una secuencia de nucleótidos que comprende la secuencia de nucleótidos que tiene una identidad de secuencia de al menos el 95 % con dicha secuencia de nucleótidos mencionada en a), en la que dicho promotor
- 10
- preferente de células fibrosas que comprende dicha secuencia de nucleótidos dirige la expresión génica más allá de 30 DPA en *Gossypium hirsutum*.
2. El promotor preferente de células fibrosas según la reivindicación 1, que comprende la secuencia de nucleótidos de SEC ID NO 9 desde la posición 1374 hasta la posición 2307.
- 15
3. Un gen quimérico que comprende las siguientes regiones de ADN enlazadas de forma funcional
- a) un promotor preferente de células fibrosas según una cualquiera de las reivindicaciones 1 o 2;
 - b) una región heteróloga de ADN que codifica un ARN biológicamente activo de interés; y
 - c) una señal de terminación de la transcripción y de poliadenilación.
- 20
4. El gen quimérico según la reivindicación 3, en el que dicha región de ADN heterólogo codifica una molécula seleccionada de entre una N-acetilglucosamina transferasa, preferentemente una N-acetilglucosamina transferasa de tipo NodC, una celulosa sintasa, una sacarosa sintasa, preferentemente una sacarosa sintasa de tipo C; una sacarosa fosfato sintasa; una β -1,3-endoglucanasa, csIF, csID y una 1,3-1,4 glucano sintasa csIC.
- 25
5. El gen quimérico según la reivindicación 3, en el que dicho ARN biológicamente activo es una ribozima, un microARN, una horquilla de ARN bicatenario.
- 30
6. Una célula vegetal que comprende un gen quimérico según una cualquiera de las reivindicaciones 3 a 5.
7. Una planta que comprende en sus células un gen quimérico según una cualquiera de las reivindicaciones 3 a 5.
8. La planta según la reivindicación 7, que es una planta de algodón.
- 35
9. Una semilla de una planta que comprende en sus células un gen quimérico según una cualquiera de las reivindicaciones 3 a 5.
10. Un procedimiento para expresar un ARN biológicamente activo de forma preferente en una célula fibrosa de una planta productora de fibras, tal como una planta de algodón, procedimiento que comprende
- 40
- a) dotar a las células de dichas plantas de un gen quimérico según una cualquiera de las reivindicaciones 3 a 5; y
 - b) cultivar dichas plantas.
- 45
11. Uso de un promotor preferente de células fibrosas según una cualquiera de las reivindicaciones 1 o 2 para la expresión preferente de un ARN biológicamente activo en células fibrosas de una planta productora de fibras.

```

1 -----ttaaattataaataaattaaagtaaaattacattttaaccctaaaaaa-gagaaaagtatatcta
1 gtccttgaccctaattggaataattggtatgattaaattataaataaattaaaggtaaaattatattttgacctaaaaaaatgaaaaaaatatatcta
66 atttctcgaaaaatggaagaaaaattataaatt-----tatggcatttctaaaaaaattctgaattcgct-----actaaaagatg
101 gtttctcgaaaaatggaagaaaaataaattgatacattalaaaaattatggcatttctaaaaaaattctgaatttgaatgaaattalaataaaaaaaa
142 a-----aattata-----aaatccgaagcattccagaaagatggatcaccacccaaatccaca
201 agtttaaaaacatatagattccaagaatagtggaataattatattgacaacacgtgaagaaatccaagcattagcagaaaaaggatcaccacaaatccaca
190 acaatacaatgaaaaagtaattgataaattaaattgaaagtgagcatttaatttggatagccatatacttctctgctgaatttataggttctctcaatgcaatt
301 acaatacaatgaaaaagtaattgataaattaaattgaaagtgagcatttaatttggatagccatatacttctctgctgaatttataggttctctcaatgcaatt
290 aaattatattcgacaccttttgaatgaaataaaatgacacagaggaaagacgggttcattctatttttt--ctttcaatcgcccatcaaaaataccaaaaat
401 aaattatattgtcacttttgaatgaaata-aatgacaca-----gttcactattttttttcttcaatcgcccatcaaaaataccgaaaaat
388 gtaactacatgcaaaaa-accaaaatgaaaaaataatcctatatttggatattttaaataatttggatgtaaaacgtaaaatgattgaaaaaattatgatgg
488 gtaactacatgcaaaaaagatc-----gaaaaatattcctatatttggatattttaaataatttggatgtaaaacgtaaaatgattgaaaaaattatgatgg
487 tgttgtgctgtatgtccataaaaatccaatgtactcaccatttatcaaatgtatactttgaggaaggttattttgataaatactcaagtttttttttatagat
581 tgttgtcgtgtatgtccataaaaatccaatgtatcgaatgtataatcttgacacgaagttatcttaataataatacaagttatcttttttatacat
587 gggaaaatttttaaattatttttggattttgataaattttaaattcgatacacaataaataataatgtaaaatttttaaaatttaaaattta
681 gagatacatcctcaaaaattattttt-----atataatccgaaaa-----
687 ataatacaaatgaagaaaaataattataaataatttccgattaaaaataaatactggaaaagaaatgtcaacacttttcttaataacaatcagaataggat
718 -----atcaataacgta-----cgatcaaa-----ctagaaaagggaaagtgtcaaaaactctattcattatatacgcaaatatgat
787 gggacacgataaccttcatgcattgatctcaggggtccaaaaactcggaaatcctttttgaaaaaaaac--ttccagagagagatatataaaatccagcag
784 gggacacgataacctcctcatgcattgatatactcatatgtccaaaaactcagaatcctttttgaaaaaaaataatccagagagagtgatataaaatccagcag
885 taggcacaagaacagcaccagttattgactttcctttgtaaaaaaataaaggt-gctgagatcaagaaataatgtagaaataatgggtcccaagattttctg
884 tgtgcacaagaacagcaccagttattgacattcctttgtaaaaaaataaaggtgagatcaagaaataatgtagaaataatgggtcccaacattttctg
984 ggttttttaacttaagca
984 ggttttttaactcagca

```

Figura 1

1 atgctgtttttaaactcaactcctctcttaacagatggccgtgatatggttgctatggtttgaaacggcaacaactctcccatctccaggagatgtta
 1 atgggtgtttttaaactcaactcctctcttaacagatggccgtgatatggttgctatggtttgaaacggcaacaactctcccatctccaggagatgtta

 101 ttaactctttcaaaaactagtggtgataaaacataatcaggctctaccagccttacctgaagtgtcgaagcagcaaggggatcgggaataatccctctcgat
 101 ttaactctttcaaaaactagtggtgataaaacataatcaggctctaccagccttacctgaagtgtcgaagcagcaaggggatcgggaataatccctctcgat

 201 gagtacgacaaaacgaggacatacaaaaagcctcgcgaacggatcaaaagtgcagccgatgcatgggttaacaccaacaacatcgtcccttataagggaagatgttcaa
 201 gggtcggagaaaacgaggacatacaaaaagcctcgcgaacggatcaaaagtgcagccgatgcatgggttaacaccaacaacatcgtcccttataagggaagatgttcaa

 301 ttcagggttcacatcatctgggaatgaaagcattccaggacagtcgaagctcttacatccctgggtgcgatgaacaacataatgaaactcgtggcctcattt
 301 ttcagggttgatcactatctgggaatgaaagcattccaggacaaatcaagctcttacatccctgggtgcgatgaacaacataatgaaactcgtggcctcattt

 401 ggttaggcacgacgaaaggttacgaccgtgggtcccgatgaatgcccctaagctacccctcctcagacggcgttttggaagcgatataaacatcgat
 401 ggttaggcacgacgaaaggttacgaccgtgggtcccgatgaatgcccctaagctacccctcctcagacggcgttttggaagcgatataaacatcgat

 501 catgactagtatcatggccattctggttcgacaggattcgccccctcctgatcaaatgtgtaccttattttgcctatgcctcagacccccactcatattcc
 501 catgactagtatcatggccattctggttcgacaggattcgccccctcctgatcaaatgtgtaccttattttgcctatgcctcagacccccactcatattcc

 601 ctcaactacgcttggttcacctcgaccgaccgggtgggtggtcgaccaggcttggaaatactacaacctcttgacggcatgggtcggatgcttcaatgccc
 601 ctcgatcacgcttggttcacctcgaccgaccgggtgggtggtcgaccaggcttggaaatactacaacctcttgacggcatgggtcggatgcttcaatgccc

 701 ccctagataagatcgggttcggccaaattactctcatctgtagccgaaactggatggccgaccgctggtaacgagccttacacgagtgctcgcgaaacgctca
 701 ccctagataagatcgggttcggccaaattactctcatctgtagccgaaactggatggccgaccgctggtaacgagccttacacgagtgctcgcgaaacgctca

 801 aacttataacaagaacttggtgaaatcattgacgcagaaaggactccgaaaaagacctgaatataaatgccgacctttttcttcgagatgttcaacgag
 801 aacttataacaagaacttggtgaaatcattgacgcagaaaggactccgaaaaagacctgaatataaatgccgacctttttcttcgagatgttcaacgag

 901 aacttgaagcaacccacagttgagcagaattcggattcttcttcccccaatgaacccctgtttatccattttgggtga
 901 gatctgaagcaacccacagttgagcagaattcggattcttcttcccccaatgaacccctgtttatccattttgggtga

Figura 2

```
1 gtaaaacaaacttctctacagtgatcttacagtaaatggccttgaaaaatatacaacaaacatttatcttcaatccattttaattactgatctacta  
1 gtaaaacaaacttctctacagtgatcttacagtaaatggccttgaaaaatatacaacaaacatttta-----tactgatctacca  
101 tataatggttcag  
83 tataatggttcag
```

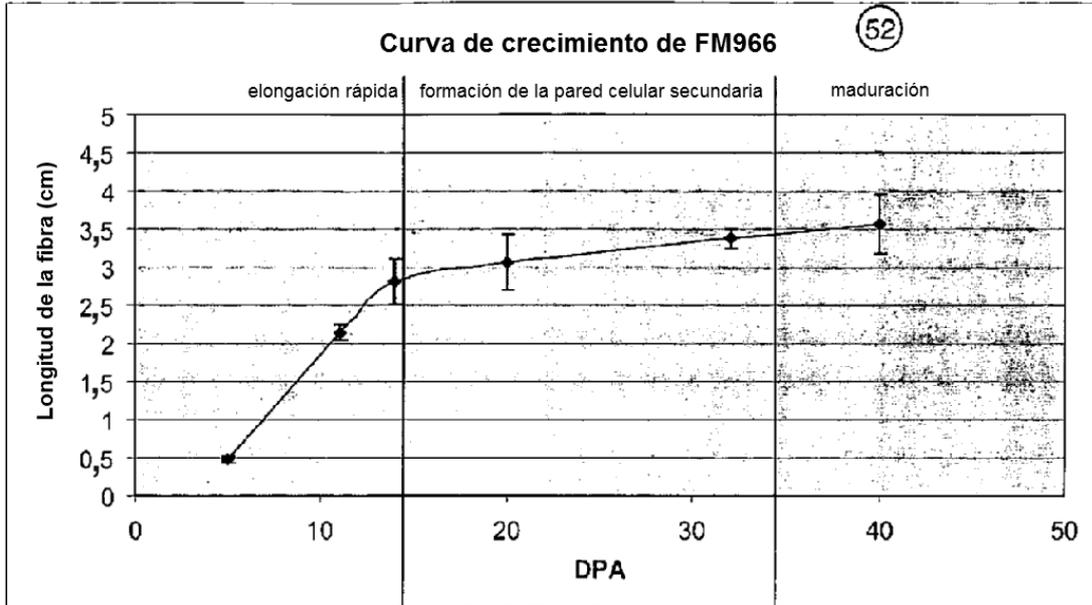
Figura 3

1 mlflltqlsltdgrdigvcyglngnlnpspgdvinlktsginnirlyqpypevleaargsgislsmgmediqslakdqsaadawvntni vpykédvq
 1 mvfltlqlsltdgrdigvcyglngnlnpspgdvinlktsginnirlyqpypevleaargsgislsmgmediqslakdqsaadawvntni vpykédvq
 101 fffliigneaipggsssyipgamnnimnslafglgttkvttvvpmmalstsyppsdgafgsditsimtsimailvqdspllinvpyfayasdpthis
 101 fkliligneaigqsssyipgamnnimnslafglgttkvttvvpmmalstsyppsdgafgsditsimtsimailvqdspllinvpyfayasdpthis
 201 lnyalftstapvvvdqgleyynlfdgmvdafnaaldkigfgqitli vaetgwptagnepytsvanaqtnknllnhvtqkgtprpeyimpdffemfne
 201 lnyalftstapvvvdqgleyynlfdgmvdafnaaldkigfgqitli vaetgwptagnepytsvanaqtnknllnhvtqkgtprpeyimpdffemfne
 301 dlkqptveqnfghffpnmnpvypfw
 301 dlkqptveqnfghffpnmnpvypfw

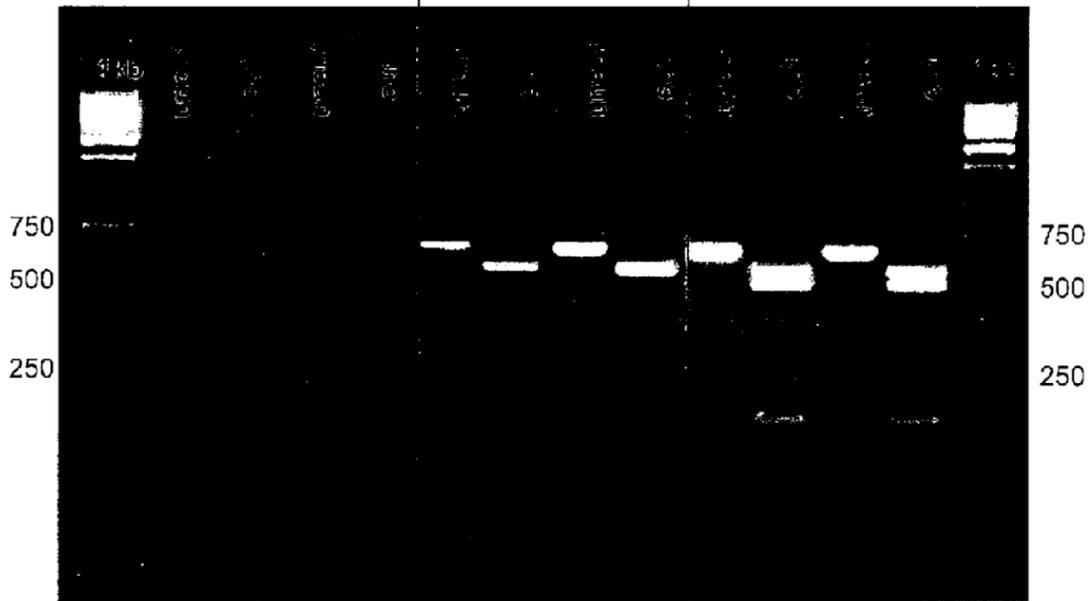
Figura 4

GaGluc1#-#Garb#leaves#gDNA CCGATCAAAGT
 GhGluc1SGA#FiberMax966#Ghir#leaves#gDNA CCGATCAAAGT
 GhGluc1SGD#FiberMax966#Ghir#leaves#gDNA AAGATCAAAGT
 GrGluc1#-#Grai#leaves#gDNA AAGATCAAAGT

Figura 5



Fibra + semilla 5 dpa Fibra 10 dpa Fibra 15 dpa Fibra 20 dpa Fibra 40 dpa ADN genómico FM966



Tipo genoma A = 476 + 118 + 59
 Tipo genoma D = 538 + 118

Figura 6

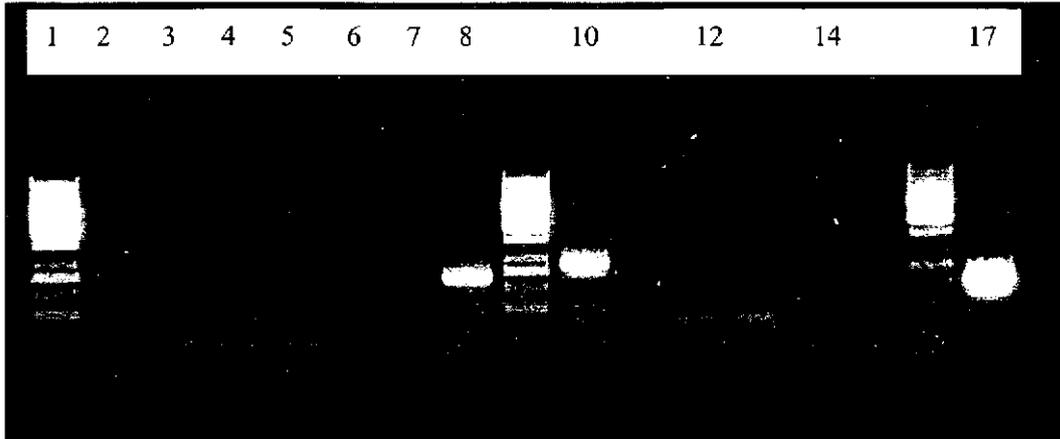


Figura 7