

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 672 349**

51 Int. Cl.:

C12N 15/82 (2006.01)

C07K 14/415 (2006.01)

A01H 5/10 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.11.2014 PCT/US2014/066599**

87 Fecha y número de publicación internacional: **28.05.2015 WO15077447**

96 Fecha de presentación y número de la solicitud europea: **20.11.2014 E 14859333 (8)**

97 Fecha y número de publicación de la concesión europea: **16.05.2018 EP 2922961**

54 Título: **Secuencia de nucleótidos que codifica la proteína homeobox4 relacionada con wuschel (WOX4) de corchorus olitorius y corchorus capsularis y métodos de uso de los mismos**

30 Prioridad:

22.11.2013 US 201361907617 P

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

14.06.2018

73 Titular/es:

**BANGLADESH JUTE RESEARCH INSTITUTE
(100.0%)
Manik Mia Avenue
Dhaka 1207, BD**

72 Inventor/es:

**ALAM, MAQSUDUL;
ISLAM, MOHAMMED, SHAHIDUL;
AHMED, BORHAN;
HAQUE, MOHAMMED, SAMIUL y
ALAM, MOHAMMED, MONJURUL**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 672 349 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

Secuencia de nucleótidos que codifica la proteína homeobox4 relacionada con wuschel (WOX4) de corchorus olitorius y corchorus capsularis y métodos de uso de los mismos

Campo de la invención

- 5 La presente invención se refiere en general al campo de la biología molecular y se refiere a un método para potenciar los rasgos relacionados con el rendimiento de fibras al modular la expresión en una planta de un ácido nucleico que codifica el polipéptido de proteína homeobox4 relacionada con WUSCHEL. Más particularmente, la presente invención proporciona los homólogos de homeobox4 (WOX4) relacionada con WUSCHEL aislados de dos especies de plantas de yute, a saber, *C. olitorius* y *C. capsularis*, y su aplicación en la producción de fibra de estopa
10 en la planta de yute, así como también una planta de yute transgénica de las mismas.

Antecedentes de la invención

- 15 El yute es una fibra natural ecológica y biodegradable. Es un recurso renovable con una alta producción de biomasa por unidad de área de tierra. Más de 100 especies de yute (que incluyen parientes silvestres) producen fibra de líber natural. El yute se produce a partir de plantas del género *Corchorus*, que alguna vez se clasificaron con la familia Tiliaceae, más recientemente con Malvaceae. Sin embargo, solo dos de estas especies, *C. olitorius* y *C. capsularis*, producen fibra de alta calidad adecuada para su uso en propósitos industriales. Siendo una fibra natural, el yute se puede utilizar de diversas maneras, complementando o reemplazando sintéticos, y ha recibido una atención creciente por parte de la industria. Como existe una creciente demanda mundial de fibra de yute, es necesario seguir mejorando la producción de fibra de yute. Por lo tanto, existe un interés significativo en estudiar la biosíntesis de la
20 fibra y explorar la biología molecular implicada en la biosíntesis de esta fibra.

- 25 La fibra de yute es una fibra extraaxilar que está compuesta y/o comprende dos tipos de fibra: (i) fibra de floema primaria que se desarrolla a partir del procámbium en la región del protofloema a través de la división y modificación celular y (ii) fibra de floema secundaria que se desarrolla a partir del cámbium por la actividad de las iniciales fusiformes y radiales (Maiti and Mitra, 1972. Bull Bot Soc Bengal 26:79-85). Estos tejidos de procámbium y cámbium se comunican célula a célula mediados por el factor inhibidor de la diferenciación del elemento traqueal (TDIF) y la proteína quinasa de membrana del receptor TDIF (TDR), que promueve la proliferación de células de procámbium y suprime su diferenciación del xilema (Figura 1; Hirakawa et al., 2010. Célula vegetal, 22:2618-2629; Etchells et al., 2013. Development 140, 2224-2234).

- 30 La familia de genes HOMEBOX relacionada con WUSCHEL (WOX) realiza funciones relacionadas durante la iniciación y/o mantenimiento de diversas células embrionarias, meristemáticas e iniciales de órganos (Haecker et al, 2004). Entre las proteínas de la familia de genes HOMEBOX relacionadas con WUSCHEL (WOX), la WOX4 actúa como un regulador clave de la ruta de señalización de TDIF (Hirakawa et al. 2010) y se expresa preferiblemente en el procámbium y el cámbium (Schradler et al., 2004; Ji et al., 2010 e Hirakawa et al. 2010). Por ejemplo, el TDIF-TDR induce la transcripción del factor de transcripción principal HOMEBOX4 relacionado con WUSCHEL (WOX4) que promueve el mantenimiento de las células madre de procámbium/cámbium en *Arabidopsis* y en tomate. El polipéptido de HOMEBOX4 relacionado con (WOX4) WUSCHEL cataliza el inicio de la fibra de líber en la planta. Sin embargo, no existen muchos informes de caracterización o tecnologías existentes proporcionadas en la técnica anterior que se relacionen con este polipéptido. La patente de los Estados Unidos No. 2011/0283420 A1 ha divulgado un polipéptido similar a la homeobox 1 relacionada con wuschel (similar a WOX1) para rasgos
35 relacionados con el rendimiento mejorado en plantas. En otra Patente E.P. No. 1451301 B1 se divulga el uso del gen wuschel en la promoción de la embriogénesis somática en plantas. Recientemente, se divulgaron algunos homólogos del gen wuschel en la patente de Estados Unidos No. 2010/0100981 A1.

- 45 En vista del hecho de que la proteína homeobox4 relacionada con WUSCHEL podría desempeñar una función importante en la ruta de biosíntesis de fibra de yute, es deseable que la industria proporcione un enfoque genético relacionado con la biosíntesis de la fibra en la planta al explorar y utilizar la biología molecular y la información genética de homeobox4 relacionada con WUSCHEL (WOX4). Además, debido a que la ruta de biosíntesis de fibra y la composición genética de cada especie de planta normalmente varía, también es preferible un enfoque específico a especie para optimizar el rendimiento de fibra de las plantas de yute y obtener resultados compatibles para permitir su uso en la industria.

- 50 Resumen de la invención

Un objeto de la presente invención, entre otros, es proporcionar un gen que codifica una proteína, derivada de *C. olitorius* y *C. capsularis*, que está involucrada en catalizar el inicio de la formación de fibra de líber, y tiene una secuencia similar de homeobox4 relacionada con WUSCHEL (WOX4). Más específicamente, la presente invención

proporciona un gen que tiene la capacidad de inducir fibra de floema, proporcionando de esta manera una proteína codificada de ese modo, y usos de la misma.

5 Otro objeto de la presente invención es proporcionar la biología molecular e información genética de homeobox4 relacionada con WUSCHEL (WOX4) para ser explotada/utilizada para mejorar la producción, crecimiento, resistencia y rendimiento de fibra en las plantas de *C. olitorius* y *C. capsularis*, así como en otras plantas productoras de fibras de líber.

Aún otro objeto de la presente invención es obtener una planta transgénica de *C. olitorius* y *C. capsularis* con una mayor producción de fibra regulando la biosíntesis de la homeobox4 relacionada con WUSCHEL (WOX4) en la planta.

10 Todavía otro objeto de la presente invención es proporcionar polinucleótidos aislados que tengan secuencias de nucleótidos específicas, que puedan facilitar la realización del método divulgado, y proporcionar acceso a plantas transgénicas de *C. olitorius* y *C. capsularis*.

Un objeto adicional de la presente invención es proporcionar una potencial forma comercialmente factible de aumentar la producción de fibra para productos a base de fibra de yute.

15 En un primer aspecto la presente invención proporciona un polinucleótido aislado que codifica un polipéptido homeobox4 relacionado con WUSCHEL derivado de la planta *C. olitorius* o *C. capsularis*, que comprende una molécula de ácido nucleico seleccionada del grupo que consiste de: (a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma; y (b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos
20 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma.

En un segundo aspecto la invención proporciona un polipéptido homeobox4 relacionado con WUSCHEL aislado que comprende una secuencia de aminoácidos seleccionada del grupo que consiste de: (a) una secuencia de aminoácidos establecida en la SEQ ID NO: 3 o SEQ ID NO: 6, o fragmento biológicamente activo de la misma; y (b)
25 una secuencia de aminoácidos que comprende una secuencia de aminoácidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 3 o 6. El polipéptido comprende una función seleccionada del grupo que consiste de la catalización de la iniciación, formación, mejora, y variación, para de esta manera modificar la composición de fibra de floema en la planta de *C. olitorius* o *C. capsularis*.

La presente invención proporciona un gen aislado de *C. olitorius* que codifica una proteína que tiene la secuencia de aminoácidos como se establece en la SEQ ID NO: 3, que está involucrada en catalizar el inicio de la formación de fibra de floema y tiene una secuencia similar a homeobox4 relacionada con WUSCHEL (WOX4). La presente invención proporciona adicionalmente un gen que codifica una proteína que tiene una secuencia de aminoácidos modificada por la adición o supresión de uno o una pluralidad de aminoácidos y/o reemplazo con otros aminoácidos
30 en la secuencia de aminoácidos como se establece en la SEQ ID NO: 3, que está involucrada en catalizar el inicio de formación de fibra de floema y tiene una secuencia similar a homeobox4 relacionada con WUSCHEL (WOX4). La presente invención también proporciona un gen aislado de *C. capsularis* que codifica una proteína que tiene la secuencia de aminoácidos como se establece en la SEQ ID NO: 6, que está involucrada en catalizar el inicio de formación de fibra de floema y tiene una secuencia similar a homeobox4 relacionada con WUSCHEL (WOX4). La presente invención proporciona adicionalmente un gen que codifica una proteína que tiene una secuencia de aminoácidos modificada por la adición o supresión de uno o una pluralidad de aminoácidos y/o reemplazo con otros aminoácidos en la secuencia de aminoácidos como se establece en la SEQ ID NO: 6, que está involucrada en catalizar el inicio de formación de fibra de floema y tiene una secuencia similar a homeobox4 relacionada con WUSCHEL (WOX4).
35 40

De acuerdo con una de las realizaciones preferidas de la presente invención, la planta de *C. olitorius* utilizada es la variedad O-4, y la planta de *C. capsularis* utilizada es la variedad CVL-1.
45

Un tercer aspecto de la presente invención divulga una construcción de gen recombinante que comprende un polinucleótido aislado del primer aspecto, en el que el polinucleótido se puede expresar en una célula anfitriona para producir un homólogo de homeobox4 relacionada con WUSCHEL (WOX4) en la planta de *C. olitorius* y *C. capsularis*, respectivamente.

50 Una realización adicional de la presente invención es un transformante que comprende una construcción de gen recombinante del tercer aspecto capaz de expresar un polinucleótido para producir un homólogo de proteína de homeobox4 relacionada con WUSCHEL (WOX4).

Otra realización proporciona una semilla de un transformante de la invención, en la que dicha semilla comprende un polinucleótido aislado del aspecto.

Otro aspecto proporciona un método para producir una planta o planta transgénica que tiene rendimiento de fibra mejorado o potenciado en relación con plantas de control, que comprende:

5 (a) introducir en una célula vegetal una construcción de gen recombinante que comprende un polinucleótido del tercer aspecto; y

(b) cultivar la célula vegetal bajo condiciones para promover el crecimiento y desarrollo de la planta; y

(c) expresar un polipéptido seleccionado del grupo que consiste en:

10 (i) un polipéptido que comprende una secuencia de aminoácidos establecida en la SEQ ID NO: 3 o SEQ ID NO: 6, o un fragmento biológicamente activo de la misma; y

(ii) un polipéptido que tiene por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 3 o SEQ ID NO: 6.

15 Un aspecto adicional de la invención proporciona un método para inducir, iniciar, mejorar, o potenciar el crecimiento de la planta, altura de la planta, fibra y rendimiento de semilla en una planta de yute, que comprende incorporar en una planta de yute la construcción de gen recombinante del tercer aspecto, en la que la construcción de gen recombinante comprende adicionalmente una región promotora unida de forma operable a una molécula de ácido nucleico establecida en (a) o (b), en la que dicho promotor mejora la transcripción o expresión de la molécula de ácido nucleico.

20 Un experto en la técnica apreciará fácilmente que la presente invención está bien adaptada para llevar a cabo los objetos y obtener los fines y ventajas mencionados, así como también aquellos inherentes a los mismos.

Estas y otras características, aspectos y ventajas de la presente invención se entenderán mejor con referencia a la siguiente descripción y reivindicaciones.

Breve descripción de los dibujos

25 La Figura 1 muestra la ruta de señalización TDIF-TDR. La señalización de TDIF-TDR diverge en dos rutas, una es la WOX4 implicada en la proliferación de células de floema, que promueve el inicio de células de fibra y la otra es un factor hipotético X que inhibe la diferenciación de xilema de las células madre vasculares.

La Figura 2 muestra el árbol filogenético que compara la SEQ ID NO. 3 de *C. olitorius* y la SEQ ID NO. 6 de *C. capsularis* junto con otras secuencias de aminoácidos, que producen la proteína de homeobox4 relacionada con WUSCHEL(WOX4).

30 Descripción detallada de la invención

La invención se puede comprender más completamente a partir de la siguiente descripción detallada y los dibujos acompañantes que forman parte de esta solicitud.

35 Las definiciones y/o métodos proporcionados en este documento definen la presente invención y guían a aquellos expertos en la técnica en la práctica de la presente invención. Salvo que se establezca lo contrario, los términos se deben entender de acuerdo con el uso convencional por parte de los expertos en la técnica pertinente. En la medida en que se encuentre que cualquiera de las definiciones y/o métodos son inconsistentes con cualquiera de las definiciones y/o métodos proporcionados en cualquier referencia de patente o diferente a patente, se entiende que dicha definición y/o método que se ha proporcionado/adoptado expresamente en esta solicitud se utilizará en este documento. Los términos singulares “un”, “una” y “el” incluyen referentes plurales a menos que el contexto indique claramente lo contrario. De manera similar, la palabra “o” tiene la intención de incluir “y” a menos que el contexto indique claramente lo contrario. Por lo tanto, “que comprende A o B” significa que incluye A, o B, o A y B. Se debe entender adicionalmente que todos los tamaños de bases o aminoácidos, y todos los valores de peso molecular o de masa molecular, dados para ácidos nucleicos o polipéptidos son aproximados, y se proporcionan para la descripción. Aunque se pueden utilizar métodos y materiales similares o equivalentes a aquellos descritos en este documento en la práctica o prueba de la presente divulgación, métodos y materiales adecuados se describen a continuación.

40

45

La presente invención se refiere a polinucleótidos aislados que codifican la proteína homeobox4 relacionada con WUSCHEL extraída de *C. olitorius* y *C. capsularis*, y los correspondientes polipéptidos derivados de la misma. Más particularmente, la presente invención proporciona homólogos de homeobox4 relacionada con WUSCHEL, y su aplicación en rendimiento de fibra mejorado en *C. olitorius* y *C. capsularis*, ambas especies de plantas de yute, así como plantas transgénicas de *C. olitorius* y *C. capsularis* relacionadas. Las secuencias genómicas de la invención que codifican las enzimas se identificaron principalmente mediante comparación de secuencias de nucleótidos de ADN genómico de *C. olitorius* y *C. capsularis* y las secuencias de nucleótidos de genes enzimáticos conocidos de otras plantas. Antes de esta invención, no se conocían secuencias de nucleótidos de estos genes de *C. olitorius* y *C. capsularis*, los marcos de lectura, las posiciones de los exones e intrones, la estructura de las enzimas y su posible utilidad en el desarrollo de plantas de yute con alto rendimiento de fibra.

El análisis de la secuencia del genoma de especies de yute cultivadas comercialmente, *C. olitorius* y *C. capsularis*, revela que ambas especies tienen genes únicos que codifican enzimas que poseen actividades catalíticas preferidas para potenciar la producción de fibra. Las secuencias de nucleótidos se anotaron inicialmente mediante programas de software, tales como Augustus, Analizador de Ácidos Nucleicos Basado en Semi-HMM (SNAP), Geneid (Genome Bioinformatics Research Lab), que puede identificar regiones de codificación putativas, intrones y uniones de corte y empalme. Se realizó una curación automatizada y manual adicional de las secuencias de nucleótidos para refinar y establecer una caracterización precisa de las regiones de codificación y otras características génicas.

Se secuenciaron parcial o completamente más de 30.096 ADNc de *C. olitorius* y 37.031 ADNc de *C. capsularis*. De ellos se desarrolló un único ADNc de cada una de las especies, que codifican nuevas enzimas, con funciones putativas, actividades catalíticas preferidas para mejorar la producción de fibra en yute.

Los marcos de lectura abiertos (ORF) se analizan después de la secuenciación total o parcial de clones de colecciones de ADNc derivadas de ARNm de *C. olitorius* y *C. capsularis* y se analizan adicionalmente utilizando un software de análisis de secuencia y al determinar la homología con secuencias conocidas en bases de datos (públicas/privadas).

En el contexto de esta divulgación, una serie de términos utilizados a lo largo de la especificación tienen los significados indicados a menos que se indique expresamente que tienen un significado diferente.

Como se utiliza en este documento, un "polinucleótido" es una secuencia de nucleótidos tal como un fragmento de ácido nucleico. Un polinucleótido puede ser un polímero de ARN o ADN de cadena doble o sencilla, que opcionalmente contiene bases de nucleótidos sintéticas, no naturales o alteradas. Un polinucleótido en forma de un polímero de ADN puede comprender y/o consistir de uno o más segmentos de ADNc, ADN genómico, ADN sintético o mezclas/combinaciones de los mismos. Un polinucleótido aislado de la presente invención puede incluir por lo menos uno de los 150 nucleótidos contiguos (tanto corriente arriba como corriente abajo) derivados de la SEQ ID No. 1 y, SEQ ID No. 4, o el complemento de dichas secuencias.

"Polipéptido" como se utiliza en el presente documento, es una cadena lineal única de aminoácidos unida entre sí por enlaces peptídicos, y que tiene usualmente una secuencia de más de 100 aminoácidos de longitud.

"Aislado" significa alterado "por la mano del hombre" del estado natural. Si una composición o sustancia es de origen natural, se ha "aislado" si se ha cambiado o eliminado de su entorno original, o ambos. Por ejemplo, un polinucleótido o un polipéptido presente de manera natural en una planta o animal vivo no está "aislado", pero el mismo polinucleótido o polipéptido separado de los materiales coexistentes de su estado natural está "aislado", tal como se emplea en este documento el término.

El término "gen", como se utiliza en el presente documento, se define como las secuencias genómicas de la planta *C. olitorius* y *C. capsularis*, particularmente la secuencia polinucleotídica que codifica la secuencia polipeptídica de las enzimas homeobox4 relacionadas con WUSCHEL implicadas en las actividades catalíticas preferidas para mejorar la producción de fibra en yute. El término puede incluir adicionalmente las moléculas de ácido nucleico que comprenden las secuencias de nucleótidos de cadena arriba, cadena abajo y/o de intrón.

Una "secuencia de codificación" o "región de codificación" se refiere a una molécula de ácido nucleico que tiene información de secuencia necesaria para producir un producto génico, tal como un aminoácido o polipéptido, cuando se expresa la secuencia. La secuencia de codificación puede comprender y/o consistir de secuencias no traducidas (que incluyen intrones o regiones 5' o 3' no traducidas) dentro de regiones traducidas, o puede carecer de dichas secuencias interpuestas no traducidas (por ejemplo, como en ADNc).

El término "oligonucleótido", como se utiliza en el presente documento, es un polinucleótido corto o una porción de polinucleótido, que puede comprender preferiblemente 10-1000, lo más preferiblemente 12 a 50 nucleótidos de longitud. Con respecto a la realización de la presente invención, los nucleótidos contenidos dentro de los oligonucleótidos pueden ser análogos o derivados de nucleótidos de origen natural.

El término “cebador”, como se utiliza en el presente documento, es un oligonucleótido capaz de unirse a una secuencia de ácido nucleico objetivo y cebar la síntesis de ácido nucleico. Un oligonucleótido de amplificación como se define en este documento puede tener preferiblemente de 10 a 50, más preferiblemente de 15 a 25, nucleótidos de longitud. Adicionalmente, los oligonucleótidos de amplificación de la presente invención se pueden sintetizar químicamente y dichos oligonucleótidos no son ácidos nucleicos naturales.

La abreviatura utilizada en toda la especificación para referirse a ácidos nucleicos que comprenden secuencias de nucleótidos son las abreviaturas convencionales de una letra. Por lo tanto, cuando se incluyen en un ácido nucleico, los nucleótidos de codificación que se producen de forma natural se abrevian de la siguiente manera: adenina (A), guanina (G), citosina (C), timina (T) y uracilo (U). Adicionalmente, a menos que se especifique lo contrario, las secuencias de ácido nucleico presentadas en este documento tienen la dirección 5' → 3'.

Como se utiliza en este documento, el término “complementario” y sus derivados se utilizan en referencia al emparejamiento de ácidos nucleicos por las reglas bien conocidas de que A se empareja con T o U y C se empareja con G. El complemento puede ser “parcial” o “completo”. En complemento parcial, solo algunas de las bases de ácido nucleico se emparejan de acuerdo con las reglas de emparejamiento de bases; mientras que, en complemento completo o total, todas las bases se emparejan de acuerdo con la regla de emparejamiento. El grado de complemento entre las cadenas de ácido nucleico puede tener efectos significativos sobre la eficacia y la fuerza de la hibridación entre las cadenas de ácidos nucleicos, como es bien conocido en la técnica. La eficacia y la fuerza de dicha hibridación dependen del método de detección.

El término “célula anfitriona”, como se utiliza en el presente documento, incluye cualquier tipo de célula que sea susceptible a transformación, transfección, transducción, expresión y similares con una construcción de ácido nucleico o vector de expresión que comprende y/o consiste en un polinucleótido de la presente invención. La célula anfitriona adecuada incluye células de hongos y/o plantas, especialmente células de plantas que producen fibras de líber.

El término “ligado operativamente” generalmente indica aquí una configuración en la que una secuencia de control se coloca en una posición apropiada con respecto a la secuencia de codificación de la secuencia de polinucleótidos de tal manera que la secuencia de control dirija la expresión de la secuencia de codificación de un polipéptido. Por ejemplo, un promotor se puede ligar operativamente con una secuencia de codificación cuando afecta a la expresión de esa secuencia de codificación, es decir, la secuencia de codificación está bajo el control transcripcional del promotor.

Un “vector” generalmente se refiere a un replicón, tal como un plásmido, fago, cósmido, levadura o virus, a los que pueden insertar operativamente otro segmento de ácido nucleico para provocar la replicación o expresión del segmento. El término “vector” también pretende referirse a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha ligado. Un tipo de vector es un “plásmido”, que se refiere a un bucle de ADN de doble cadena circular en el que se pueden ligar segmentos de ADN adicionales. Otro tipo de vector es un vector viral, donde se pueden ligar segmentos de ADN adicionales en el genoma viral. Ciertos vectores son capaces de replicación autónoma en una célula anfitriona en la que se introducen (por ejemplo, vectores bacterianos que tienen un origen bacteriano de replicación y vectores de mamífero episómicos). Otros vectores se pueden integrar en el genoma de una célula anfitriona después de la introducción en la célula anfitriona, y de ese modo se replican junto con el genoma anfitrión. Más aún, ciertos vectores son capaces de dirigir la expresión de los genes a los que están ligados operativamente. Dichos vectores se denominan en la presente memoria “vectores de expresión recombinantes” (o simplemente, “vectores de expresión”). En general, los vectores de expresión de utilidad en las técnicas de ADN recombinante a menudo están en forma de plásmidos. En la presente especificación, “plásmido” y “vector” se pueden utilizar indistintamente ya que el plásmido es la forma de vector utilizada más comúnmente. Sin embargo, la invención pretende incluir dichas otras formas de vectores de expresión, tales como vectores víricos (por ejemplo, retrovirus con replicación defectuosa, adenovirus y virus asociados a adeno), que cumplen funciones equivalentes.

El término “construcción de ácido nucleico” o “construcción de ADN” se utiliza a veces para referirse a una secuencia de codificación o secuencias ligadas operativamente a secuencias reguladoras apropiadas e insertadas en un vector para transformar una célula. Este término se puede utilizar indistintamente con el término “ADN transformante” o “transgén”.

El término “promotor”, como se utiliza en el presente documento, se refiere a una secuencia de ácidos nucleicos que funciona para dirigir la transcripción de un gen corriente abajo. El promotor generalmente será apropiado para la célula anfitriona en la que se expresa el gen objetivo. El promotor junto con otras secuencias de ácido nucleico reguladoras de la transcripción y la traducción (también denominadas “secuencias de control”) es necesario para expresar un gen dado. En general, las secuencias reguladoras transcripcionales y traduccionales incluyen, pero no se limitan a, secuencias promotoras, sitios de unión ribosómica, secuencias de inicio y parada de transcripción, secuencias de inicio y parada de traducción y secuencias potenciadoras o activadoras.

5 Como se utiliza en el presente documento, "polipéptido WUSCHEL" o "polipéptido WUS" significa un polipéptido que tiene actividad wuschel, es decir, implicado en la iniciación y el mantenimiento de las células madre en las plantas. La actividad de Wuschel estimula el crecimiento celular, incluyendo las células madre. Wuschel es una proteína de homeodominio de plantas, que comprende un motivo de homeodominio hélice-bucle-hélice-giro-hélice "atípico" (en comparación con el motivo de homeodominio de animales) que comprende residuos de aminoácidos adicionales en el bucle y/o giro del dominio. Las proteínas Wuschel pueden comprender y/o consistir adicionalmente e otros motivos conservados, tales como los dos dominios de terminal C de wuschel conservados, los dominios (E/R)TLPLFP y A(A/S)LEL(S/T)L. El término también incluye fragmentos, variantes y homólogos con cualquiera de las funciones mencionadas anteriormente.

10 El término "homólogos", como se utiliza en el presente documento, se refiere a una proteína que abarca péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones, supresiones y/o inserciones de aminoácidos en relación con la proteína no modificada en cuestión y que tienen actividad biológica y funcional similar a la Proteína no modificada de la que se derivan.

Una supresión se refiere a la eliminación de uno o más aminoácidos de una proteína.

15 Una inserción se refiere a uno o más residuos de aminoácidos que se introducen en un sitio predeterminado en una proteína. Las inserciones pueden comprender fusiones de terminal N y/o terminal C, así como inserciones intra-secuencia de aminoácidos únicos o múltiples. Generalmente, las inserciones dentro de la secuencia de aminoácidos serán más pequeñas que las fusiones de terminal N o C, del orden de aproximadamente 1 a 10 residuos.

20 Una sustitución se refiere al reemplazo de aminoácidos de la proteína con otros aminoácidos que tienen propiedades similares (tales como hidrofobicidad similar, hidrofilia, antigenicidad, propensión a formar o romper estructuras α -helicoidales o estructuras de lámina β), las sustituciones de aminoácidos son normalmente de residuos únicos, pero pueden estar agrupados dependiendo de las restricciones funcionales que se colocan sobre el polipéptido y pueden variar desde 1 hasta 10 aminoácidos; las inserciones usualmente serán del orden de aproximadamente 1 a 10 residuos de aminoácidos. Las sustituciones de aminoácidos son preferiblemente sustituciones de aminoácidos conservadoras. Las tablas de sustitución conservadora son bien conocidas en la técnica (véase, por ejemplo, Creighton (1984) Proteins. W.H. Freeman and Company (Eds) y la Tabla 1 a continuación),

Tabla 1: Ejemplos de sustituciones de aminoácidos conservadoras

Residuo	Sustituciones conservadoras	Residuo	Sustituciones conservadoras
Ala	Ser	Leu	Ile; Val
Arg	Lys	Lys	Arg; Gln
Asn	Gln; His	Met	Leu; Ile
Asp	Glu	Phe	Met; Leu; Tyr
Gln	Asn	Ser	Thr; Gly
Cys	Ser	Thr	Ser; Val
Glu	Asp	Trp	Tyr
Gly	Pro	Tyr	Trp; Phe
His	Asn; Gln	Val	Ile; Leu
Ile	Leu, Val		

30 Las sustituciones, supresiones y/o inserciones de aminoácidos se pueden realizar fácilmente utilizando técnicas de síntesis de péptidos bien conocidas en la técnica, tales como síntesis de péptidos en fase sólida y/o cualquier otra técnica sintética, o mediante manipulación de ADN recombinante. Los métodos para la manipulación de secuencias

de ADN para producir variantes de sustitución, inserción o supresión de una proteína son bien conocidos en la técnica. Por ejemplo, las técnicas para realizar mutaciones de sustitución en sitios predeterminados en ADN son bien conocidas por aquellos expertos en la técnica e incluyen la mutagénesis de M13, mutagénesis in vitro T7-Gen. (USB, Cleveland, OH), mutagénesis dirigida al sitio de cambio rápido (Stratagene, San Diego, CA), mutagénesis dirigida a sitio mediada por PCR u otros protocolos de mutagénesis dirigida a sitio.

Como se utiliza en este documento, una "porción biológicamente activa" se puede referir a un fragmento de proteína de homeobox4 relacionada con WUSCHEL que tiene una actividad biológica para catalizar la iniciación, formación, mejora o variación en la composición de fibra de floema en la planta de *C. olitorius* o *C. capsularis*. Las porciones biológicamente activas de una proteína homeobox4 relacionada con WUSCHEL incluyen péptidos o polipéptidos que comprenden secuencias de aminoácidos suficientemente idénticas o derivadas de la secuencia de aminoácidos de la proteína de homeobox4 relacionada con WUSCHEL, por ejemplo, la secuencia de aminoácidos como se establece en la SEQ ID NO: 3 o SEQ ID NO: 6, que incluye pocos aminoácidos diferentes a las proteínas de homeobox4 relacionada con WUSCHEL de longitud completa, y exhibe por lo menos una actividad de una proteína de homeobox4 relacionada con WUSCHEL. Normalmente, las porciones biológicamente activas comprenden un dominio o motivo con por lo menos una actividad de la proteína de homeobox4 relacionada con WUSCHEL. Una porción biológicamente activa de una proteína de homeobox4 relacionada con WUSCHEL puede ser un polipéptido que tiene, por ejemplo, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, o 223 aminoácidos en longitud. La proteína de homeobox4 relacionada con WUSCHEL puede tener una secuencia de aminoácidos establecida en la SEQ ID NO: 3 o SEQ ID NO: 6. En otras realizaciones, la proteína de homeobox4 relacionada con WUSCHEL es sustancialmente idéntica a la SEQ ID NO: 3 o SEQ ID NO: 6, y retiene la actividad funcional de la proteína de la SEQ ID NO: 3 o SEQ ID NO: 6, todavía difiere en la secuencia de aminoácidos debido a la variación alélica natural o mutagénesis. En otra realización, la proteína de homeobox4 relacionada con WUSCHEL comprende una secuencia de aminoácidos de por lo menos aproximadamente 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% o más idéntica a la SEQ ID NO: 3 o SEQ ID NO: 6.

El término "dominio", como se utiliza en el presente documento, se refiere a un conjunto de aminoácidos conservados en posiciones específicas a lo largo de un alineamiento de secuencias de proteínas relacionadas evolutivamente. Mientras que los aminoácidos en otras posiciones pueden variar entre homólogos, los aminoácidos que están altamente conservados en posiciones específicas indican aminoácidos que probablemente sean esenciales en la estructura, estabilidad o función de una proteína. Identificados por su alto grado de conservación en secuencias alineadas de una familia de homólogos proteicos, se pueden utilizar como identificadores para determinar si algún polipéptido en cuestión pertenece a una familia polipeptídica previamente identificada.

El término "motivo", como se utiliza en el presente documento, se refiere a una región corta conservada en la secuencia de proteínas relacionadas evolutivamente. Los motivos son con frecuencia partes de dominios altamente conservados, pero también pueden incluir solo una parte del dominio, o estar localizados fuera del dominio conservado (si todos los aminoácidos del motivo quedan fuera de un dominio definido).

Existen bases de datos especializadas para la identificación de dominios, por ejemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Ácidos nucleicos Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), una sintaxis de perfil generalizada para motivos de secuencias biomoleculares y su función en la interpretación automática de secuencias. (En) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)), or Pfam (Bateman et al., Ácidos nucleicos Research 30(1): 276-280 (2002)). Un conjunto de herramientas para el análisis in silico de secuencias de proteínas está disponible en el servidor de proteómica ExPASy (Instituto Suizo de Bioinformática (Gasteiger et al., ExPASy: The Proteomics Server for Indepth Protein Knowledge and Analysis, Ácidos nucleicos Res. 31:3784-3788(2003)). También se pueden identificar dominios o motivos utilizando técnicas de rutina, tales como mediante alineación de secuencia.

Para los propósitos de la invención, "transgénico", "transgén" o "recombinante" significa con respecto a, por ejemplo, una secuencia de ácidos nucleicos, un casete de expresión, construcción génica o un vector que comprende y/o que consiste en la secuencia de ácidos nucleicos o un organismo transformado con las secuencias de ácidos nucleicos, casetes de expresión o vectores de acuerdo con la invención, todas esas construcciones producidas por métodos recombinantes en los que: (a) las secuencias de ácidos nucleicos que codifican proteínas útiles en los métodos de la invención, o b) secuencia de control genético que está operativamente ligada a la secuencia de ácidos nucleicos de acuerdo con la invención, por ejemplo un promotor, o (c) a) y b) no se encuentran en su entorno genético natural o han sido modificadas por métodos recombinantes. La modificación puede tomar la forma de, por ejemplo, una sustitución, adición, supresión, inversión o inserción de uno o más residuos de nucleótidos. El entorno genético natural se entiende como el locus genómico o cromosómico natural en la planta original o la presencia en una colección genómica. En el caso de una colección genómica, el entorno genético natural de la secuencia de ácidos nucleicos se retiene preferiblemente, por lo menos en parte. El entorno flanquea la secuencia de ácidos nucleicos por lo menos en un lado y tiene una longitud de secuencia de aproximadamente 50 pb, preferiblemente de aproximadamente 500 pb. Un casete de expresión de origen natural - por ejemplo la combinación de origen natural

del promotor natural de las secuencias de ácidos nucleicos con la secuencia de ácidos nucleicos correspondiente que codifica un polipéptido útil en los métodos de la presente invención, como se definió anteriormente - se convierte en un casete de expresión transgénica cuando este casete de expresión se modifica por métodos no naturales, sintéticos ("artificiales") tales como, por ejemplo, el tratamiento mutagénico. Se describen métodos adecuados, por ejemplo, en el documento US 5,565,350 o documento WO 00/15815.

Por lo tanto, se entiende que una planta transgénica para el propósito de la invención que incluye aquellas plantas en las que los ácidos nucleicos utilizados en el método de la invención no están en su locus natural en el genoma de dicha planta, y por lo tanto es posible que los ácidos nucleicos ácidos que se expresan de manera homóloga o heteróloga. Sin embargo, como se mencionó, transgénico también significa que, mientras que los ácidos nucleicos de acuerdo con la invención o utilizados en el método de la invención están en su posición natural en el genoma de una planta, la secuencia se ha modificado con respecto a la secuencia natural, y/o que se han modificado las secuencias reguladoras de las secuencias naturales. Por transgénico se entiende preferiblemente la expresión de los ácidos nucleicos de acuerdo con la invención en un locus no natural en el genoma, en el que tiene lugar la expresión homóloga o, preferiblemente, heteróloga de los ácidos nucleicos.

El término "introducción" o "transformación", como se utiliza en el presente documento, abarca la transferencia de un polinucleótido exógeno a una célula anfitriona, independientemente del método utilizado para la transferencia. El tejido vegetal capaz de propagación clonal posterior ya sea por organogénesis o embriogénesis, se puede transformar con una construcción genética de la presente invención y una planta completa regenerada a partir de la misma. El tejido particular elegido variará dependiendo de los sistemas de propagación clonal disponibles, y más adecuados para la especie particular que se va a transformar. Los objetivos tisulares de ejemplo incluyen discos foliares, polen, embriones, cotiledones, hipocótilos, megagametofitos, tejido calizo, tejido meristemático existente (por ejemplo, meristema apical, yemas axilares y meristema radicular) y tejido de meristema inducido (por ejemplo, Meristema cotiledónico y meristema hipocótilo). El polinucleótido se puede introducir transitoria o establemente en una célula anfitriona y se puede mantener en forma no integrada, por ejemplo, como un plásmido. Alternativamente, se puede integrar en el genoma del anfitrión. La célula vegetal transformada resultante se puede utilizar entonces para regenerar una planta transformada de una manera conocida por los expertos en la técnica.

La transferencia de genes extraños al genoma de una planta se llama transformación.

La transformación de especies de plantas es ahora una técnica bastante rutinaria. Ventajosamente, se puede utilizar cualquiera de diversos métodos de transformación para introducir el gen de interés en una célula antecesora adecuada. Los métodos descritos para la transformación y regeneración de plantas a partir de tejidos vegetales o células de plantas se pueden utilizar para la transformación transitoria o estable. Los métodos de transformación incluyen el uso de liposomas, electroporación y productos químicos que aumentan la absorción libre de ADN, la inyección del ADN directamente en la planta, el bombardeo con pistolas de partículas, la transformación utilizando virus o polen y la microproyección. Las plantas transgénicas, que incluyen plantas de cultivo transgénicas, se producen preferiblemente mediante transformación mediada por *Agrobacterium*. Un método de transformación ventajoso es la transformación en planta ((Sajib et. al. *Célula vegetal Tiss. Organ Cult.* (2008) 95, 333-34).

En general después de la transformación, las células vegetales o agrupamientos celulares se seleccionan por la presencia de uno o más marcadores que están codificados por genes expresables en plantas co-transferidos con el gen de interés, después de lo cual el material transformado se regenera en una planta completa. Para seleccionar las plantas transformadas, el material vegetal obtenido en la transformación está, por regla general, sometido a condiciones selectivas para que las plantas transformadas se puedan distinguir de las plantas no transformadas. Por ejemplo, las semillas obtenidas de la manera descrita anteriormente se pueden plantar y, después de un período de crecimiento inicial, someterse a una selección adecuada mediante pulverización. Una posibilidad adicional consiste en hacer crecer las semillas, si es apropiado después de la esterilización, en placas de agar utilizando un agente de selección adecuado de modo que solo las semillas transformadas puedan crecer en las plantas. Alternativamente, las plantas transformadas se criban para detectar la presencia de un marcador seleccionable tal como los descritos anteriormente.

Después de la transferencia y regeneración del ADN, también se pueden evaluar las plantas putativamente transformadas, por ejemplo, utilizando análisis de Southern, para la presencia del gen de interés, número de copias y/o organización genómica. Alternativa o adicionalmente, los niveles de expresión del ADN recientemente introducido se pueden monitorizar utilizando análisis Northern y/o Western, ambas técnicas son bien conocidas por las personas con experiencia normal en la técnica.

Las plantas transformadas generadas se pueden propagar por una variedad de medios, tales como por propagación clonal o técnicas clásicas de reproducción. Por ejemplo, puede seleccionarse una planta transformada de primera generación (o T1) y pueden seleccionarse transformantes homocigóticos de segunda generación (o T2), y las plantas T2 luego se pueden propagar adicionalmente mediante técnicas de reproducción clásicas. Los organismos transformados generados pueden tomar una variedad de formas. Por ejemplo, pueden ser quimeras de células transformadas y células no transformadas; transformantes clonales (por ejemplo, todas las células transformadas

para contener el casete de expresión); injertos de tejidos transformados y no transformados (por ejemplo, en plantas, un rizoma transformado injertado en un vástago no transformado).

El término “expresión aumentada” o “sobreexpresión” como se utiliza en el presente documento se refiere a cualquier forma de expresión que es adicional al nivel de expresión de tipo salvaje original.

5 Los métodos para incrementar la expresión de genes o productos génicos están bien documentados en la técnica e incluyen, por ejemplo, la sobreexpresión impulsada por promotores apropiados, el uso de potenciadores de la transcripción o potenciadores de la traducción. Los ácidos nucleicos aislados que sirven como elementos promotores o potenciadores se pueden introducir en una posición apropiada (normalmente corriente arriba) de una forma no heteróloga de un polinucleótido para regular positivamente la expresión de un ácido nucleico que codifica el polipéptido de interés. Por ejemplo, los promotores endógenos se pueden alterar in vivo mediante mutación, supresión y/o sustitución (véase, Kmiec, US 5,565,350; Zarlring et al., WO9322443) o se pueden introducir promotores aislados en una célula vegetal en la orientación y distancia adecuada de un gen de la presente invención con el fin controlar la expresión del gen.

15 Si se desea la expresión del polipéptido, generalmente es deseable incluir una región de poliadenilación en el extremo 3' de una región de codificación de un polinucleótido. La región de poliadenilación se puede derivar del gen natural, de una variedad de otros genes vegetales, o del T-ADN. La secuencia del extremo 3' que se va a agregar se puede derivar, por ejemplo, de los genes de nopalina sintasa u octopina sintasa, o alternativamente de otro gen de plantas, o menos preferiblemente de cualquier otro gen eucariota.

20 También se puede agregar una secuencia de intrón a la región 5' no traducida (UTR) o a la secuencia de codificación de la secuencia de codificación parcial para aumentar la cantidad del mensaje maduro que se acumula en el citosol. Se ha mostrado que la inclusión de un intrón capaz de corte y empalme en la unidad de transcripción tanto en construcciones de expresión de plantas como de animales aumenta la expresión génica tanto a nivel de ARNm como de proteínas hasta 1000 veces (Buchman and Berg (1988) Mol. Cell Biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Dicha mejora del intrón de expresión génica es normalmente mayor cuando se coloca cerca del extremo 5' de la unidad de transcripción. El uso de los intrones de maíz Adh1-S intrón 1, 2 y 6, el intrón Bronze-1 es conocido en la técnica. Para información general, véase: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

30 Para determinar el porcentaje de identidad de dos secuencias de aminoácidos o de dos secuencias de ácidos nucleicos, las secuencias se pueden alinear para propósitos de comparación óptimos (por ejemplo, se pueden introducir espacios en una o ambas secuencias de una primera o segunda secuencia de aminoácidos o ácidos nucleicos para alineación óptima y las secuencias no idénticas pueden descartarse para propósitos de comparación). La longitud de una secuencia de referencia alineada para propósitos de comparación puede ser por lo menos el 95% de la longitud de la secuencia de referencia. Luego se pueden comparar los residuos de aminoácidos o nucleótidos en las correspondientes posiciones de aminoácidos o posiciones de nucleótidos. Cuando una posición en la primera secuencia está ocupada por el mismo residuo de aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición (como se utiliza en este documento “identidad” de aminoácido o ácido nucleico es equivalente a “homología” de aminoácido o ácido nucleico). El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de espacios, y la longitud de cada espacio, que se deben introducir para la alineación óptima de las dos secuencias.

45 La comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias se pueden lograr utilizando un algoritmo matemático. En una realización, el porcentaje de identidad entre dos secuencias de aminoácidos se puede determinar utilizando algoritmo de Needleman y Wunsch (J. Mol. Biol. 48: 444-453 (1970)) que se ha incorporado al programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), utilizando ya sea una matriz Blosum 62 o una matriz PAM250, y un peso de espacio de 16, 14, 12, 10, 8, 6 o 4 y un peso de longitud de 1, 2, 3, 4, 5 o 6. En aún otra realización preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se pueden determinar utilizando el programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), utilizando una matriz NWSgapADN. CMP y un peso de espacio de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 o 6. En otra realización, el porcentaje de identidad entre dos secuencias de aminoácidos o nucleótidos se puede determinar utilizando el algoritmo de E. Meyers and W. Miller (Comput. Appl. Biosci. 4:11-17 (1988)) que se ha incorporado al programa ALIGN (versión 2.0 o 2.0U), utilizando una tabla de residuos de peso PAM120, una penalización de longitud de espacio de 12 y una penalización de espacio de 4.

55 Los programas informáticos de ejemplo que se pueden utilizar para determinar la identidad entre dos secuencias incluyen, entre otros, el conjunto de programas BLAST, por ejemplo, BLASTN, BLASTX y TBLASTX, BLASTP y TBLASTN, de acceso público en www.ncbi.nlm.nih.gov/BLAST.

Las búsquedas de secuencia se llevan a cabo normalmente utilizando el programa BLASTN, cuando se evalúa una secuencia de ácidos nucleicos dada con respecto a secuencias de ácidos nucleicos en las secuencias de ADN de

GenBank y otras bases de datos públicas. Se prefiere el programa BLASTX para buscar secuencias de ácidos nucleicos que se han traducido en todos los marcos de lectura frente a secuencias de aminoácidos en las Secuencias de Proteínas GenBank y otras bases de datos públicas.

5 Se realiza una alineación preferida de las secuencias seleccionadas para determinar el “% de identidad” entre dos o más secuencias utilizando, por ejemplo, el programa CLUSTAL-W.

10 Como se establece, una realización de la presente invención son polinucleótidos aislados que codifican el polipéptido de homeobox4 relacionada con WUSCHEL encontrado en las plantas *C. olitorius* y *C. capsularis* que comprende y/o que consiste de secuencia de nucleótidos como se establece en la SEQ ID NO 2 y SEQ ID NO 5, respectivamente. De forma correspondiente, los polipéptidos de homeobox4 relacionada con WUSCHEL respectivos codificados por las secuencias de nucleótidos poseen las secuencias de aminoácidos establecidas en la SEQ ID NO 3 y SEQ ID NO 6. De acuerdo con una realización de la presente invención, la SEQ ID NO 3 se refiere a la secuencia de polipéptidos del homólogo de homeobox4 relacionada con WUSCHEL (WOX4) derivado de *C. olitorius*, y la SEQ ID NO 6 se refiere a la secuencia de polipéptidos de homólogo de homeobox4 relacionada con WUSCHEL (WOX4) derivado de *C. capsularis*. Ambas de estas enzimas están presentes en la ruta de biosíntesis de fibra en las plantas de *C. olitorius* y *C. capsularis* para catalizar las moléculas iniciadoras para cebar la biosíntesis de célula de fibra en la planta.

La presente invención también proporciona una secuencia génica que codifica los homólogos de homeobox4 relacionada con WUSCHEL (WOX4) a partir de las plantas *C. olitorius* y *C. capsularis*.

20 En una variante, el polinucleótido de 1250 pb de longitud ilustrado en la SEQ ID No. 1 es el gen de longitud completa aislado de *C. olitorius*. Esta secuencia génica incluye por lo menos 150 nucleótidos contiguos tanto corriente arriba como corriente abajo del gen. Esto también proporciona la secuencia intrónica del gen.

En otra variante, el polinucleótido de 1237 pb de longitud ilustrado en la SEQ ID No. 4 es el gen de longitud completa aislado de *C. capsularis*. Esta secuencia génica incluye por lo menos 150 nucleótidos contiguos tanto de la corriente arriba como corriente abajo del gen. Esto también proporciona la secuencia intrónica del gen.

25 En una realización de la presente invención, se proporciona un polinucleótido aislado de codificación un polipéptido que comprende secuencia de nucleótidos establecida en la SEQ ID NO 2 y/o SEQ ID NO 5. La SEQ ID NO 2 se refiere a la secuencia de polinucleótidos de la secuencia de homólogos de homeobox4 relacionada con WUSCHEL (WOX4) derivada de *C. olitorius* y SEQ ID NO 5 se refiere a la secuencia de polinucleótidos de la secuencia de homólogos de homeobox4 relacionada con WUSCHEL (WOX4) derivada de *C. capsularis*.

30 En aún otra realización, una molécula de ácido nucleico aislada que es capaz de codificar un polipéptido de homeobox4 relacionada con WUSCHEL, o fragmento biológicamente activo del mismo, comprende una secuencia de nucleótidos que tiene por lo menos aproximadamente 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% o más idéntica a la longitud completa de la secuencia de nucleótidos establecida en la SEQ ID NO: 2, SEQ ID NO: 5, o cualquier complemento de la misma.

35 En una realización, el polinucleótido de 669 bp de longitud ilustrado en la SEQ ID No. 2 es el clon de ADNc de longitud completa que codifica la proteína de homeobox4 relacionada con WUSCHEL (WOX4) que exhibe un marco de lectura abierto que codifica un polipéptido de 222 aminoácidos, como en la SEQ ID No. 3, con una masa molecular calculada de aproximadamente 25.54 kD. A través del análisis SMART de la SEQ ID No. 3, se revela la presencia del dominio homeobox en la secuencia. Esto es factor de unión de ADN, que está involucrado en la regulación transcripcional de los procesos clave de desarrollo de la planta. Esta es la proteína de homeobox4 relacionada con WUSCHEL (WOX4) involucrada en la proliferación de células vasculares de la planta.

40 En una realización, el polinucleótido de 672 bp de longitud ilustrado en la SEQ ID No. 5 es el clon de ADNc de longitud completa que codifica la proteína de homeobox4 relacionada con WUSCHEL (WOX4) que exhibe un marco de lectura abierto que codifica un polipéptido de 223 aminoácidos, como en la SEQ ID No. 6, con una masa molecular calculada de aproximadamente 25.67 kD. A través del análisis SMART de la SEQ ID No. 6, se revela la presencia del dominio homeobox en la secuencia. Esto es factor de unión de ADN, que está involucrada en la regulación transcripcional de los procesos clave de desarrollo de la planta. Esta es la proteína de homeobox4 relacionada con WUSCHEL (WOX4) involucrada en la proliferación de células vasculares de la planta.

45 De acuerdo con la realización preferida de la presente invención, el polinucleótido aislado ilustrado en la SEQ ID No. 2 se puede obtener mediante amplificación por PCR de la región conservada del gen utilizando ARN total aislado de la planta de *C. olitorius* y la SEQ ID No. 5 se puede obtener mediante amplificación por PCR de la región conservada de este gen utilizando ARN total aislado de la planta de *C. capsularis*. Como se establece en la descripción precedente, la planta de *C. olitorius* aplicada es la variedad 0-4 y la *C. capsularis* aplicada es la variedad CVL-1.

En otra realización de la presente invención, se divulga una construcción de gen recombinante que comprende un polinucleótido que tiene la secuencia de nucleótidos establecida en la SEQ ID NO 2 y/o SEQ ID NO 5, en la que el polinucleótido se puede expresar en una célula anfitriona, y es trasladable para producir el homólogo de proteína de homeobox4 relacionada con WUSCHEL (WOX4) en las plantas de *C. olitorius* y *C. capsularis*. El procedimiento para amplificar, clonar y secuenciar el la homeobox4 relacionada con WUSCHEL (WOX4) a partir de las plantas de *C. olitorius* y *C. capsularis* se detalla adicionalmente en el Ejemplo 2. Preferiblemente, la construcción de gen recombinante comprende adicionalmente una región promotora unida de forma operable para mejorar la expresión de la plantilla de polinucleótido. Bajo el control de transcripción del promotor específico, la expresión de la región de codificación dentro de la construcción de genes recombinantes que contienen el polinucleótido de la SEQ ID NO 2 y/o SEQ ID NO 5 luego se puede mejorar, lo que conduce a un mayor rendimiento de la proteína de homeobox4 relacionada con WUSCHEL (WOX4).

De acuerdo con una realización de la invención, la expresión modulada es la expresión o actividad aumentada, por ejemplo sobreexpresión de un polipéptido homeobox4 relacionado con WUSCHEL que codifica la molécula de ácido nucleico, por ejemplo, de una molécula de ácido nucleico que codifica la SEQ ID NO 2 y SEQ ID NO 5. Los métodos para aumentar la expresión de ácidos nucleicos o genes, o productos génicos, están bien documentados en la técnica y se proporcionan ejemplos en la sección de definiciones.

La invención también proporciona un método para la producción de plantas transgénicas que tienen rendimiento de fibra mejorado en relación con las plantas de control, que comprenden introducción y expresión en una planta de cualquier ácido nucleico que codifica un polipéptido homeobox4 relacionado con WUSCHEL como se definió en este documento anteriormente.

Más específicamente, la presente invención proporciona un método para la producción de plantas transgénicas que tienen rendimiento de fibra mejorado en comparación con las plantas de control nulas, cuyo método comprende:

(i) introducir y expresar en una planta o célula vegetal un polipéptido homeobox4 relacionado con WUSCHEL que codifica una construcción de ácidos nucleicos o genética que comprende y/o que consiste de un polipéptido homeobox4 relacionado con WUSCHEL que codifica ácido nucleico; y

(ii) cultivar la célula vegetal bajo condiciones que promueve el crecimiento y desarrollo de la célula de fibra.

Otro aspecto de la invención se refiere a polinucleótido aislado que codifica un polipéptido homeobox4 relacionado con WUSCHEL, y se deriva de la planta *C. olitorius*, que comprende una molécula de ácido nucleico seleccionada del grupo que consiste de:

a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida in de la SEQ ID NO 2, o un complemento de la misma; y

b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2, o un complemento de la misma.

En ciertas realizaciones, la planta de *C. olitorius* es la variedad O-4.

Otro aspecto de la invención se refiere a un polinucleótido aislado que codifica un polipéptido homeobox4 relacionado con WUSCHEL y se deriva de la planta *C. capsularis*, que comprende una molécula de ácido nucleico seleccionada del grupo que consiste de:

a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO: 5, o un complemento de la misma; y

b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 5, o un complemento de la misma.

En ciertas realizaciones, la planta de *C. capsularis* es la variedad CVL-1.

Otro aspecto de la invención se refiere a un polipéptido homeobox4 relacionado con WUSCHEL aislado que comprende una secuencia de aminoácidos establecida en la SEQ ID: NO: 3, o fragmento biológicamente activo de la misma, dicho polipéptido comprende una función seleccionada del grupo que consiste de la catalización de la iniciación, formación, mejora, y variación, para de esta manera modificar la composición de fibra de floema en la planta de *C. olitorius*.

En ciertas realizaciones, la planta de *C. olitorius* es la variedad O-4.

En ciertas realizaciones, dicho polipéptido comprende por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 3.

- 5 Otro aspecto de la invención se refiere a un polipéptido homebox4 relacionado con WUSCHEL aislado que comprende una secuencia de aminoácidos como se establece en la SEQ ID NO: 6, o fragmento biológicamente activo de la misma, dicho polipéptido comprende una o más funciones seleccionadas del grupo que consiste de la catalización de la iniciación, formación, mejora, y variación, para de esta manera modificar la composición de fibra de floema en la planta de *C. capsularis*.

En ciertas realizaciones, la planta de *C. capsularis* es la variedad CVL-1.

- 10 En ciertas realizaciones, dicho polipéptido comprende por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 6.

Otro aspecto de la invención se refiere a una construcción de gen recombinante que comprende un polinucleótido seleccionado del grupo que consiste de:

- 15 a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma; y

b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma, en el que el polinucleótido se puede expresar en una célula anfitriona para producir un homólogo de polipéptido homebox4 relacionado con WUSCHEL en las plantas de *C. olitorius* y *C. capsularis*.

- 20 En ciertas realizaciones, dicha construcción comprende adicionalmente una región promotora unida de forma operable a

a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma; o

- 25 b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma, en la que dicho promotor mejora la transcripción o expresión de la molécula de ácido nucleico.

Otro aspecto de la invención se refiere a un transformante que comprende una construcción de gen recombinante capaz de expresar un polinucleótido seleccionado del grupo que consiste de:

- 30 a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma; y

- 35 b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma, en el que dicho transformante produce un homólogo de polipéptido homebox4 relacionado con WUSCHEL.

Otro aspecto de la invención se refiere a un método para producir una planta o planta transgénica que tiene rendimiento de fibra mejorado o potenciado en relación con las plantas de control, que comprenden:

- 40 a) introducir en una célula vegetal una construcción de gen recombinante que comprende un polinucleótido seleccionado del grupo que consiste de: i) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma; y ii) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma,

b) cultivar la célula vegetal bajo condiciones para promover el crecimiento y desarrollo de la planta; y

- 45 c) expresar un polipéptido seleccionado del grupo que consiste de: i) un polipéptido que comprende una secuencia de aminoácidos establecida en la SEQ ID: NO: 3 o SEQ ID NO: 6, o un fragmento biológicamente activo de la

misma; y ii) un polipéptido que tiene por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 3 o SEQ ID NO: 6.

5 Las secuencias proporcionadas por la presente invención también se pueden utilizar como materiales preparatorios para la modificación racional o el diseño de nuevas enzimas con características que permiten a las enzimas funcionar mejor en procesos exigentes.

Aunque esta invención se ha descrito en su forma preferida con un grado de particularidad, se entiende que la presente divulgación de la forma preferida se ha realizado solo a modo de ejemplo y que se puede recurrir a numerosos cambios en los detalles de construcción y la combinación y disposiciones de partes sin apartarse del alcance de la invención y las reivindicaciones.

10 Ejemplos

Los siguientes ejemplos están destinados a ilustrar adicionalmente la invención, sin ningún intento de que la invención se limite a las realizaciones específicas descritas en la misma.

Ejemplo 1. Diseño y síntesis de cebadores

15 Los cebadores utilizados en el estudio se diseñaron a partir del transcriptoma curado manualmente y los “modelos de genes” predichos a partir de las secuencias genómicas de *C. olitorius* y *C. capsularis*, al elegir las secuencias manualmente con ORF completos o utilizando bases de datos similares. los genes se han aislado con éxito de otras plantas. El análisis bioinformático comparativo de las secuencias de nucleótidos obtenidas a partir del transcriptoma se llevó a cabo utilizando NCBI BLAST, BLASTP, RPS-BLAST, BLASTX y PSI-BLAST para identificar homólogos de los genes relacionados y para la identificación adecuada del gen. Las alineaciones de secuencia de nucleótidos se realizaron a través de clustalW versión 1.82 cuando se encontraron múltiples secuencias del “conjunto de genes”. La alineación fue editada. Los cebadores específicos de genes (tanto directos como inversos) se seleccionaron manualmente o mediante la herramienta Primer 3 plus y los cebadores se sintetizaron de forma personalizada.

25 Se sintetizaron todos los oligonucleótidos utilizados en este estudio y purificaron por HPLC por el proveedor y adquiridos de Integrated DNA Technologies (IDT). Se preparó una solución madre de aproximadamente 100 pmoles en ddH₂O de autoclave y se almacenó a aproximadamente -20 °C, en alícuotas para uso.

Secuencias de oligonucleótidos utilizadas como cebadores para PCR

Nombre del cebador	SEQ ID NO	Secuencia de oligonucleótidos	Amplificado a partir de ADNc	producto
COL F	1	CCATGGGAAACATGAAGGTGC	682	
COL R	1	TGAAACGTCCATCATCTGCCT		
CCA F	4	CCATGGGAAACATGAAGGTGC	675	
CCA R	4	TTCATGATCTGCCTTCCGGG		

30 Ejemplo 2 Amplificación, clonación y secuenciación de homeobox4 relacionada con WUSCHEL de *C. olitorius* y *C. capsularis*

35 El ARN total se aisló a partir de plántulas de tres días de edad cultivadas en medio MS como se describió previamente por Chotnczynski P and Saechi N, método de etapa única de aislamiento de ARN mediante extracción con ácido guanidinio tiocianato-fenol-cloroformo, (Anal Biochem 1987, 162: 156-159). La calidad o la integridad del ARN se verificó mediante electroforesis en gel de agarosa y se cuantificó utilizando Thermo Scientific Nano Drop 2000 según procedimientos estándar. La primera cadena de ADNc se sintetizó utilizando la transcriptasa inversa Superscript III (Invitrogen) siguiendo las instrucciones del fabricante. El gen se amplificó a partir del ADNc mediante PCR utilizando los cebadores específicos del gen. La reacción de PCR (50 µL) contenía 1 µL, de ADNc, 20 pmoles de cada cebador, 5 µL de tampón de PCR 10X, 5 µL de mezcla de dNTP 2.5 mM y 1.0 unidad de ADN polimerasa PfuTaq. La PCR se llevó a cabo en Thermal Cycler (Applied Biosystems) utilizando las siguientes condiciones: 40 desnaturalización inicial durante aproximadamente 5 minutos a aproximadamente 95°C, seguido de 35 ciclos de

desnaturalización a aproximadamente 95°C durante aproximadamente 30 segundos, hibridación a aproximadamente 59-61°C. C durante aproximadamente 30 segundos y la extensión a aproximadamente 72°C durante aproximadamente 1 minuto, con una extensión final a aproximadamente 72°C durante aproximadamente 7 minutos. El producto de la PCR se analizó en gel de agarosa al 1% utilizando tampón IX TAE y el amplicón se eluyó del gel utilizando el kit de extracción de gel QIAGEN siguiendo las instrucciones del fabricante. El producto de PCR purificado se ligó en el kit de clonación pCR®8/GW/TOPO® TA (Invitrogen) y se transformó en células competentes de *E. coli* (Invitrogen). Los plásmidos se aislaron de colonias putativas utilizando el kit QIAprep Spin Miniprep (QIAGEN) siguiendo las instrucciones del fabricante. La presencia del inserto se verificó utilizando los cebadores específicos del gen y los plásmidos positivos se sometieron a secuenciación.

10 Ejemplo 3 Análisis de la secuencia

La secuencia de nucleótidos y la secuencia de aminoácidos se analizaron mediante los programas BLASTN y BLASTP, respectivamente. Las secuencias informadas de otras plantas se alinearon con ClustalW. El análisis filogenético se llevó a cabo utilizando Neighbor Joining (NJ).

Ejemplo 4 Construcción de ruta de biosíntesis de fibra

15 La reconstrucción de la ruta metabólica automática que muestra el papel de homeobox4 relacionada con WUSCHEL (WOX4) en la biosíntesis de la fibra se construyó al identificar ortólogos de la proteína de *C. olitorus* y *C. capsularis* en comparación con el genoma de *Arbidopsis*. Las reacciones enzimáticas catalizadas por homeobox4 relacionada con WUSCHEL (WOX4) codificadas en el genoma de *C. olitorus* y *C. capsularis* se construyeron utilizando reacciones enzimáticas disponibles en la base de datos Resnet-Plant 3.0 para Pathway Studio, así como de bases de datos de ruta metabólica.

20 Aunque se han descrito diversas realizaciones de la presente invención e ilustrado en este documento, los expertos en la técnica fácilmente imaginarán una variedad de otros medios y/o estructuras para realizar las funciones y/o obtener los resultados y/o una o más de las ventajas descritas en este documento. Por lo tanto, se debe entender que las realizaciones anteriores se presentan solo a modo de ejemplo y que la invención se define por el alcance de las reivindicaciones adjuntas.

25 Aunque se han descrito diversas realizaciones de la presente invención e ilustrado en este documento, los expertos en la técnica fácilmente imaginarán una variedad de otros medios y/o estructuras para realizar las funciones y/o obtener los resultados y/o una o más de las ventajas descritas en este documento. Aquellos expertos en la técnica reconocerán, o serán capaces de determinar utilizando no más que la experimentación de rutina, muchos equivalentes a las realizaciones específicas de la invención descritas en el presente documento. Por lo tanto, se debe entender que las realizaciones anteriores se presentan solo a modo de ejemplo y que, dentro del alcance de las reivindicaciones adjuntas y sus equivalentes; la invención se puede poner en práctica de otra manera que la específicamente descrita y reivindicada.

30

ES 2 672 349 T3

SEQ ID NO : 1
LONGITUD : 1250 bp
TIPO : ADN
NOMBRE/CLAVE CARACTERÍSTICA : Intrón, exón que excluye 5' UTR de 150 bp y 3' UTR de 150 bp
ORGANISMO : Jute, *C. olitorius*

```
ACCTCATCATTGGTCTTGGTTCTTCAAAGCACCACTCATATCCCACCTTCTTTTTCTTTCTTTT
TTTTAAAATTTTATTTTTTGGGTTAAATTTATTTTTGTTATTATATAATCAGAGCATTCTCCCAT
CATTCACTCCTTCACCATGGGAAACATGAAGGTGCATCAGTTGGCACGTGGCTTATGGGAGCATGAA
CCCTCCCTCTCCCTTGGTTGCAAGCGCTTACGCCCTCTTGCTCCCAAGCTCCACCCTTCTCTTCCC
CTGATCATACTTCCGTCTCCTCTTTGACCTTAAGACCTTCATTTCGTCCCGAAAGTGGCCCCGAAA
ACTTTGCCCTTCTGACGACAAGCGAGATTCTCATTCTCCCAGGTACTTACAAATATCCATGAAATG
AAACCCTTTTATTTTTCCAATGATGTATATAATTAGTGGGGAAGGATAAGATTTTGAACCTTACGATG
AAATAACAGTGTGGAAAGTTGATAATCAAAGGTTAAGAATGTGAAGCTGAAACTTTTTCTTTTT
TTTGGTTATGTAGGTGGAAACGCACCCAGGGGAACCGGTGGAATCCGACGCAAGAGCAGATAGGG
ATATTGGAGATGCTGTATAGAGGTGGGATGCGAACTCCAAATGCACAGCAAATAGAACAGATCACTG
CACAGTTAGGCAAGTACGGGAAGATCGAAGGCAAAAACGTTTTCTATTGGTTCCAAAACCAAAAGC
ACGCGAAAGGCAAAGCAGAAGCGTAACAGTCTTGGTCTTAGCCATTCTCCAGAACTCTGCTCCC
ATTACCACCATAACTTTGGACTCTAGGGTAAGTTCAAACCAACAAAACCTTTCTTTGTATATATAT
AACGGTTAGTTTTTAGTTTTACTTCTTATAAACAGCAATTAACATTAATGTTTTTGTATATATA
TAGGGGAAGTAATGGAGAGAGAGGAGGATAGTCCATATAAGAGAAAGTGTAGGAGCTGGTCATTTG
AGTACTTAGAAGAAGAAAGCAGATCATCATCGTTCGAGTCAAGAGGAGGAAAACAGAACTCTGGAGCT
TTTCCATTGCACCCGGAAGGCAGATGATGGACGTTTCAACTTTGAAAAACAAGGAAAAAGGGAAGC
TTAACCCAAAACAAAAGACTGCTACAAAACCCAAAACCTCTGTTCCCATTTATGAAATGATAAACA
TATGCTTTGATGATCCATGATGATGATGATGATAATGAAGCTGA
```

SEQ ID NO : 2
LONGITUD : 669 bp
TIPO : ADN
ORGANISMO : Jute, *C. olitorius*
NOMBRE/CLAVE CARACTERÍSTICA : CDS

ES 2 672 349 T3

UBICACIÓN : (1).....(669)

ATGGGAAACATGAAGGTGCATCAGTTGGCACGTGGCTTATGGGAGCATGAACC
CTCCCTCTCCCTTGGTTGCAAGCGCTTACGCCCTCTTGCTCCCAAGCTCCACCCT
TCCTCTTCCCCTGATCATACTTCCGTCTCCTCTTTCGACCTTAAGACCTTCATTCG
TCCCGAAAGTGGCCCCGAAAACCTTTGCCCTTCTGACGACAAGCGAGATTCTCA
TTCTCCCCAGGTGGAAACGCACCCAGGGGGAACGCGGTGGAATCCGACGCAAG
AGCAGATAGGGATATTGGAGATGCTGTATAGAGGTGGGATGCGAACTCCAAAT
GCACAGCAAATAGAACAGATCACTGCACAGTTAGGCAAGTACGGGAAGATCGA
AGGCAAAAACGTTTTCTATTGGTTCCAAAACCACAAAGCACGCGAAAGGCAAA
AGCAGAAGCGTAACAGTCTTGGTCTTAGCCATTCTCCAGAACTCTGCTCCCA
TTACCACCATAACTTTGGACTCTAGGGGGGAAGTAATGGAGAGAGAGGAGGAT
AGTCCATATAAGAGAAAGTGTAGGAGCTGGTCATTTGAGTACTTAGAAGAAGA
AAGCAGATCATCATCGTCGAGTCAAGAGGAGGAAAACAGAACTCTGGAGCTTT
TCCCATTGCACCCGGAAGGCAGATGA

SEQ ID NO : 3
LONGITUD : 222
TIPO : PROTEÍNA
ORGANISMO : Jute, C. olitorius

MGNMKVHQLARGLWEHEPSLSLGCKRLRPLAPKLHPSSSPDHTSVSSFDLKT FIRPESGPRKLCPSD
DKRDSHSPQVETHPGGTRWNPTQEIQIGILEMLYRGGMRTPNAQQIEQIT AQLGKYGKIEGKNV FYWF
QNHKARERQKQRNSLGLSHSPRNSAPI TTTITLDSRGEVMEREEDSPYKRKCRSWSFEYLEEESRSS
SSSQEEENRTLELFLHPEGR*

SEQ ID NO : 4
LONGITUD : 1237 bp
TIPO : ADN
NOMBRE/CLAVE : Intrón, exón que incluye 5' UTR de 150 bp y 3' UTR de 150 bp
CARACTERÍSTICA

ES 2 672 349 T3

ORGANISMO : Jute, *C. olitorius*

TCACAAGTCAACCTCACCTCATCATTTGGTCTTGGTCTTCAAAGCACACCTCATATCCCCTTCC
TTTTTCAATTTTAAATTTTTTTTTGGGGTTAAATTTATTTGGTTATATAATCAGAGCATTCTCCCCAT
CATTCACTCCTTCACCATGGGAAACATGAAGGTGCATCAGTTGGCACGTGGCTTATGGGAGCATGAA
CCCTCCCTCTCCCTTGGTTGCAAGCGCTTACGCCCTCTTGCTCCCAAGCTCCACCCTTCCCTTTCCC
CTGATCATACTTCCGTCTCCTCTTTTCGACCTTAAGACCTTCATTTCGTCGCCGAAAGTGGCCCCGGAA
ACTTTGCCCTTCTGACGACAAGCGAGATTCTCATTTCTGCCAGGTACTTAAAAATTAATATCCATGA
AATAATTTGTGGGGAAGGATAAGTTTTGAACTTAATGCATAATAACAGTGTGGAACTTAATAGGTT
AAGAATATTTGAAGAACTTCTATATATATGAAGCTGAAACTTTTTTTGTGGTGTGTAGGTGGAAACG
CACCCAGGGGAAACGCGGTGGAATCCGACGCAAGAGCAGATAGGGATACTGGAGATGCTGTATAGAG
GTGGGATGCGAACTCCAAATGCACAGCAAATAGAACAGATCACTGCACAGCTAGGCAAGTACGGCAA
GATCGAAGGCCAAAAACGTTTTCTATTGGTCCAAAACCAAAAGCACGCGAAAGGCCAAAAGCAGAAG
CGTAACAGTCTTGGTCTTAGCCATTCTCCAGAACTCAGCTCCATTACCACTATAACTTTGGACA
CTAGGGTAAGTTCAAACCAACAAAACCTTTCTGTGTATATATATATAACGGTTAGTTTTTAGTT
TTTACTTCTTATAAACAGAGAAAATTAACAATGTTGTGTTTTATATATATATAGGGGGAAGTAATGG
AAAGAGAGGAGGATAGTCCATATAAGAGAAAAGTGTAGGAGCTGGTCTTTTGAGTACTTAGAAGAAGA
AAGCAGATCATCATCGTCGAGTCAAGAGGAGGAAAAACAGAACTCTGGAGCTTTTCCCATGACCCCG
GAAGGCAGATCATGAAGGGGGTTTTCAACTTTCAACTTTCAACTTTCAACTTTCAAAATGAAGGGAAA
AGGGAAGCTTAACCCAAAACAAAAGACTGCTACAAAACCCAAAACCTCTGTTCCCATTTATGAAAT
GATAAACTTATGCTTTGATGATCGATCCATG

SEQ ID NO : 5
LONGITUD : 672
TIPO : ADN
ORGANISMO : Jute, *C. olitorius*
NOMBRE/CLAVE :CDS
CARACTERÍSTICA
UBICACIÓN : (1)... (672)

ATGGGAAACATGAAGGTGCATCAGTTGGCACGTGGCTTATGGGAGCATGAACCCCTCCCTCTCCCTTG
GTTGCAAGCGCTTACGCCCTCTTGCTCCCAAGCTCCACCCTTCCCTCTTCCCCTGATCATACTTCCGT
CTCCTCTTTTCGACCTTAAGACCTTCATTTCGTCGCCGAAAGTGGCCCCGGAAACTTTGCCCTTCTGAC

ES 2 672 349 T3

GACAAGCGAGATTCTCATTCTCGCCAGGTGGAAACGCACCCAGGGGGAACGCGGTGGAATCCGACGC
AAGAGCAGATAGGGATACTGGAGATGCTGTATAGAGGTGGGATGCGAACTCCAAATGCACAGCAAAT
AGAACAGATCACTGCACAGCTAGGCAAGTACGGCAAGATCGAAGGCAAAAACGTTTTCTATTGGTTC
CAAAACCACAAAGCACGCGAAAAGGCAAAAGCAGAAGCGTAACAGTCTTGGTCTTAGCCATTCTCCCA
GAAACTCAGCTCCCATTACCACTATAACTTTGGACACTAGGGGGGAAGTAATGGAAAAGAGAGGAGGA
TAGTCCATATAAGAGAAAAGTGTAGGAGCTGGTCTTTTGAGTACTTAGAAGAAGAAAAGCAGATCATCA
TCGTGAGTCAAGAGGAGGAAAACAGAACTCTGGAGCTTTTCCCATTGCACCCGGAAGGCAGATCAT
GA

SEQ ID NO : 6
LONGITUD : 223
TIPO : PRT
ORGANISMO : Jute, *C. olitorius*

MGNMKVHQLARGLWEHEPSLSLGCCKRLRPLAPKLHPSSSPDHTSVSSFDLKTFFIRPESGPRKLCPSD
DKRDSHSRQVETHPGGTRWNPTQEIQIGILEMLYRGGMRTPNAQQIEQITAQLGKYGKIEGKNVIFYWF
QNHKARERQKQKRNSLGLSHSPRNSAPI TTTITLDTRGEVMEREEDSPYKRKCRSWSFEYLEEESRSS
SSSQEEENRTLELFPLHPEGRS*

REIVINDICACIONES

1. Un polinucleótido aislado que codifica un polipéptido homebox4 relacionado con WUSCHEL derivado de la planta *C. olitorius* o *C. capsularis*, que comprende una molécula de ácido nucleico seleccionada del grupo que consiste de:
- 5 a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos establecida en la SEQ ID NO 2 o SEQ ID NO: 5, o un complemento de la misma; y
- b) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO: 2 o SEQ ID NO: 5, o un complemento de la misma.
- 10 2. El polinucleótido aislado de codificación de la reivindicación 1, en el que la planta de *C. olitorius* es la variedad O-4.
3. El polinucleótido aislado de la reivindicación 1, en el que la planta de *C. capsularis* es la variedad CVL-1.
4. Un polipéptido homebox4 relacionado con WUSCHEL aislado que comprende una secuencia de aminoácidos seleccionada del grupo que consiste en:
- 15 (a) una secuencia de aminoácidos establecida en la SEQ ID: NO: 3 o SEQ ID NO. 6, o fragmento biológicamente activo de la misma;
- (b) una secuencia de aminoácidos que comprende una secuencia de aminoácidos que tiene por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 3 o 6; dicho polipéptido comprende una función seleccionada del grupo que consiste en la catalización de la iniciación, formación, mejora, y variación, para de esta manera modificar la composición de fibra de floema en la planta de *C. olitorius* o *C. capsularis*.
- 20 5. El polipéptido aislado de la reivindicación 4, en el que la planta de *C. olitorius* es la variedad O-4.
6. El polipéptido aislado de la reivindicación 4, en el que la planta de *C. capsularis* es la variedad CVL-1.
7. Una construcción de gen recombinante que comprende el polinucleótido aislado de una cualquiera de las reivindicaciones 1 a 3, en la que el polinucleótido se puede expresar en una célula anfitriona para producir un homólogo de polipéptido homebox4 relacionado con WUSCHEL de las plantas de *C. olitorius* y *C. capsularis*.
- 25 8. La construcción de gen recombinante de la reivindicación 7, que comprende adicionalmente una región promotora unida de forma operable a una molécula de ácido nucleico establecida en a) o b), en la que dicho promotor mejora la transcripción o expresión de la molécula de ácido nucleico.
9. Un transformante que comprende una construcción de gen recombinante de la reivindicación 7 capaz de expresar un polinucleótido, en el que dicho transformante produce un homólogo de polipéptido homebox4 relacionado con WUSCHEL.
- 30 10. Un método para producir una planta o planta transgénica que tiene rendimiento de fibra mejorado o potenciado relacionado para controlar plantas, que comprende:
- a) introducir en una célula vegetal una construcción de gen recombinante que comprende un polinucleótido de la reivindicación 7; y
- 35 b) cultivar la célula vegetal bajo condiciones para promover el crecimiento y desarrollo de la planta; y
- c) expresar un polipéptido seleccionado del grupo que consiste de:
- i) un polipéptido que comprende una secuencia de aminoácidos establecida en la SEQ ID: NO: 3 o SEQ ID NO: 6, o un fragmento biológicamente activo de la misma; y
- 40 ii) un polipéptido que tiene por lo menos 95% de identidad de secuencia a la secuencia de aminoácidos establecida en la SEQ ID NO: 3 o SEQ ID NO: 6.
11. El polinucleótido aislado de la reivindicación 1, en el que dicho polinucleótido tiene:

por lo menos 98% de identidad de secuencia a la secuencia de nucleótidos establecida en la SEQ ID NO 2 o SEQ ID NO: 5.

12. El polipéptido aislado de la reivindicación 4, en el que dicho polipéptido tiene:

5 por lo menos 98% de identidad de secuencia a dicha secuencia de aminoácidos establecida en la SEQ ID: NO: 3 o SEQ ID NO: 6.

13. Una semilla de un transformante de la reivindicación 9, en la que dicha semilla comprende el polinucleótido aislado de una cualquiera de las reivindicaciones 1-3 y 11.

10 14. Un método para inducir, iniciar, mejorar, o potenciar el crecimiento de la planta, altura de la planta, fibra y rendimiento de semilla en una planta de yute, que comprende incorporar en una planta de yute la construcción de gen recombinante de la reivindicación 8.

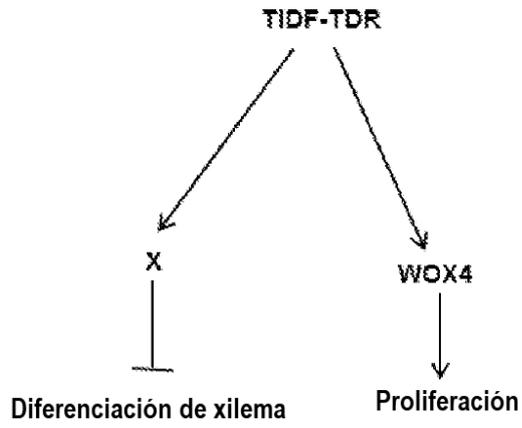


Figura 1

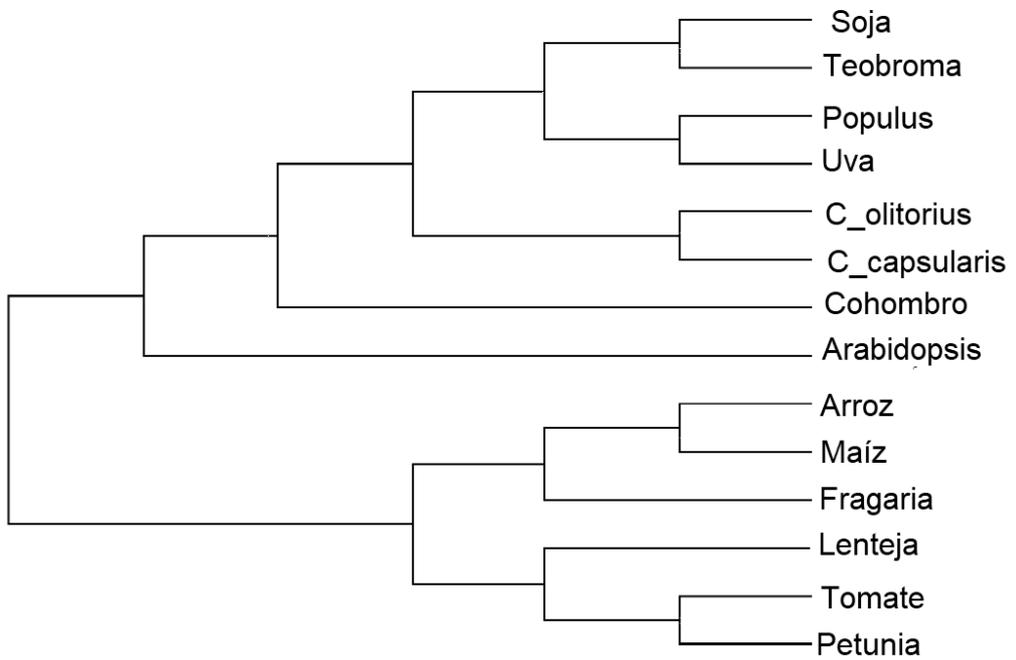


Figura 2