

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 437 621**

(51) Int. Cl.:

**C12N 15/82** (2006.01)

**C07K 14/415** (2006.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **02.05.2008 E 08767542 (7)**

(97) Fecha y número de publicación de la concesión europea: **18.09.2013 EP 2069509**

(54) Título: **Plantas que tienen rasgos mejorados relacionados con el rendimiento y un método para hacer los mismos**

(30) Prioridad:

**03.05.2007 EP 07107448 08.05.2007 US 916575 P**  
**29.05.2007 EP 07109052 29.05.2007 EP 07109068**  
**06.06.2007 US 942214 P 11.06.2007 EP 07109961**  
**19.06.2007 EP 07110548 19.06.2007 EP 07110557**  
**29.06.2007 US 937989 P 05.07.2007 US 948036 P**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:  
**13.01.2014**

(73) Titular/es:

**BASF PLANT SCIENCE GMBH (100.0%)**  
**67056 Ludwigshafen, DE**

(72) Inventor/es:

**SANZ MOLINERO, ANA ISABEL;**  
**HATZFELD, YVES;**  
**VANDENABEELE, STEVEN;**  
**SHIRLEY, AMBER;**  
**DARNIELLE, LALITREE;**  
**MCKERSIE, BRYAN y**  
**FRANKARD, VALERIE**

(74) Agente/Representante:

**CARVAJAL Y URQUIJO, Isabel**

**ES 2 437 621 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

Plantas que tienen rasgos mejorados relacionados con el rendimiento y un método para hacer los mismos

La presente invención se relaciona de manera general con el campo de la biología molecular y se refiere a un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en plantas al aumentar la expresión en

5 una planta de un ácido nucleico que codifica un dominio LOB que comprende proteína (LOB: Límites de Órgano Lateral por sus siglas en inglés), aquí abreviado como polipéptido LBD. La presente invención también se refiere a plantas que tienen aumento de expresión de un ácido nucleico que codifica un polipéptido LBD como se define adelante, cuyas plantas han aumentado el rendimiento de semillas y/o aumentado la biomasa en relación con plantas de tipo silvestre correspondientes u otras plantas de control. La invención también proporciona ácidos nucleicos LBD y polipéptidos LBD así como también construcciones útiles en los métodos de la invención.

10 El aumento de la población mundial y el suministro reducido de la tierra cultivable disponible para agricultura estimulan la investigación para aumentar la eficiencia de la agricultura. Los medios convencionales para mejoras en cultivos y hortícolas utilizan técnicas de reproducción selectiva para identificar plantas con características deseables.

15 Sin embargo, dichas técnicas de reproducción selectivas tienen diversos inconvenientes, a saber, que estas técnicas tienen normalmente mano de obra intensa y resultan en plantas que a menudo contienen componentes genéticos heterogéneos que no siempre pueden resultar en el rasgo deseable que se pasa desde la planta progenitora. Los avances en biología molecular han permitido a la humanidad modificar el germoplasma de animales y plantas. La ingeniería genética de las plantas implica el aislamiento y manipulación de material genético (normalmente en la forma de ADN o ARN) y la posterior introducción de ese material genético en una planta. Dicha tecnología tiene la capacidad de suministrar cultivos o plantas que tienen diversos rasgos económicos, agronómicos u hortícolas mejorados.

20 Un rasgo de particular interés económico es el aumento de rendimiento. El rendimiento se define normalmente como el producto medible de valor económico de un cultivo. Esto se puede definir en términos de cantidad y/o calidad. El rendimiento depende directamente de diversos factores, por ejemplo, el número y tamaño de los órganos, 25 arquitectura de la planta (por ejemplo, el número de ramas), producción de semillas, senescencia de las hojas y más. El desarrollo de las raíces, absorción de nutrientes, tolerancia al estrés y vigor temprano también pueden ser factores importantes en determinar el rendimiento. La optimización de los factores mencionados anteriormente puede por lo tanto contribuir a aumentar el rendimiento de los cultivos.

30 El rendimiento de semillas es un rasgo particularmente importante, debido a que las semillas de muchas plantas son importantes para la nutrición humana y animal. Los cultivos tales como maíz, arroz, trigo, canola y soja representan la mitad de la absorción calórica humana total, ya sea a través del consumo directo de las semillas en sí mismas o a través del consumo de productos de harina que surgen de semillas procesadas. También son una fuente de azúcares, aceites y muchas clases de metabolitos utilizados en los procesos industriales. Las semillas contienen un 35 embrión (la fuente de nuevos brotes y raíces) y un endospermo (la fuente de nutrientes para el crecimiento del embrión durante germinación y durante crecimiento temprano de plántulas). El desarrollo de una semilla involucra muchos genes, y requiere la transferencia de metabolitos desde las raíces, hojas y tallos en el crecimiento de semilla. El endospermo, en particular, asimila los precursores metabólicos de los carbohidratos, aceites y proteínas y los sintetiza en macromoléculas de almacenamiento para llenar el grano.

40 Otro rasgo importante para muchos cultivos es el vigor temprano. El vigor temprano mejorado es un objetivo importante de los programas modernos de reproducción de arroz en cultivos de arroz de zona tropical y templada. Las raíces largas son importantes para el anclaje al suelo adecuado en arroz sembrado en agua. Cuando el arroz se 45 siembra directamente en campos inundados, y cuando las plantas deben emergir rápidamente en agua, los brotes más largos se asocian con el vigor. Cuando se practica plántula por perforación, los mesocotilos y coleoptilos más largos son importantes para la buena emergencia de plántula. La capacidad de modificar por ingeniería el vigor temprano en plantas será de gran importancia en la agricultura. Por ejemplo, el pobre vigor temprano ha sido una limitación para la introducción de híbridos de maíz (*Zea mays* L.) con base en el germoplasma de Maíz en el Atlántico Europeo.

50 Un rasgo importante adicional es el de tolerancia al estrés abiótico mejorado. El estrés abiótico es una causa principal de la pérdida de cultivos en el mundo, reduciendo los cultivos promedio para la mayoría de las plantas de cultivo en más del 50% (Wang et al., Plant (2003) 218: 1-14). El estrés abiótico puede ser provocado por sequía, salinidad, extremos de temperatura, toxicidad química y estrés oxidativo. La capacidad de mejorar la tolerancia de la planta al estrés abiótico sería de gran ventaja económica para los granjeros en el todo el mundo y permitiría la siembra de cultivos durante condiciones adversas y en territorios donde no puede ser posible de otra forma la siembra de cultivos.

55 Por lo tanto, se puede aumentar el rendimiento de las cosecha al optimizar uno de los factores mencionados anteriormente.

Dependiendo del uso final, la modificación de ciertos rasgos de rendimiento puede ser favorable sobre otros. Por ejemplo para aplicaciones tales como producción de forraje o madera, o recursos de biocombustible, puede ser deseable un aumento en las partes vegetativas de una planta, y para aplicaciones tales como producción de harinas, almidón o aceite, puede ser particularmente deseable un aumento en los parámetros de semilla. Incluso entre los

5 parámetros de semilla, algunos pueden ser favorables sobre otros, dependiendo de la aplicación. Diversos mecanismos pueden contribuir a aumentar el rendimiento de semillas, ya sea que esté en la forma de aumento de tamaño de semilla o aumento de número de semillas.

10 Un método para aumentar el rendimiento (rendimiento de semillas y/o biomasa) en plantas puede ser a través de la modificación de los mecanismos de crecimiento inherentes de una planta, tal como el ciclo celular o diversas rutas de señalización involucradas en el crecimiento de planta o en los mecanismos de defensa.

De forma sorprendente, ahora se ha encontrado que modular la expresión de un ácido nucleico que codifica un dominio LOB que comprende proteína (LOB: Límites de órgano Lateral), aquí abreviado como polipéptido LBD, proporciona plantas que tienen rendimiento de semillas y/o aumento de biomasa en relación con plantas de control.

15 De acuerdo con una realización, se proporciona un método para aumentar el rendimiento de semillas y/o biomasa de plantas en relación con plantas de control, que comprende aumentar la expresión en una planta de un ácido nucleico que codifica un dominio LOB que comprende proteína (LOB: Límites de Órgano Lateral), aquí abreviado como polipéptido LBD.

#### Antecedentes

##### Dominio LOB que comprende proteína (LOB: Límites de Órgano Lateral)

20 Las proteínas LBD (o proteínas de dominio LOB) todas comparten un dominio conservado en la región de terminal N, conocida como el dominio LOB. Las proteínas de dominio LOB (Shuai et al., Plant Physiol. 129, 747-761, 2002) se encuentran en diversas especies de plantas y constituyen una familia de genes grande: la *Arabidopsis* se reporta por poseer más de 40 genes que codifican proteínas de dominio LOB, por lo menos 35 genes se encuentran en el arroz y por lo menos 15 genes en el maíz. Los polipéptidos LBD pueden ser reguladores de los factores de transcripción (entre los que se encuentran los factores de transcripción KNOX) y se postulan para cumplir una función en panícula y espiga que se ramifica en el maíz (Bortiri et al., Plant Cell 18, 574-587, 2006), la formación de raíces adventicias (Liu et al., Plant J. 43, 47-56, 2005; Inukai et al., Plant Cell 17, 1387-1396, 2005), proliferación del gametofito femenino (Evans et al., Plant Cell 19, 46-62, 2007), patrón próximo distal en pétalos (Chalfun-Junior et al., Plant Mol. Biol. 57, 559-575, 2005); morfología y venación de hojas (Iwakawa et al., Plant Cell Physiol. 43, 467-478, 2002). Yang et al. (Molecular Phylogenetics and Evolution 39, 248-262, 2006) discriminada en tres clases de proteínas LBD en arroz, con base en la clasificación de Iwakawa et al. (2005) y Shuai et al. (2002). La modulación de la expresión de gen LOB clase I (regulación por aumento o disminución de expresión) a menudo resulta en efectos pleiotrópicos, que llevan a forma de planta anormal e infertilidad. Por ejemplo, el documento US20060218674 describe un método para aumentar el tamaño del endospermo en una semilla de planta al expresar un polipéptido LOB clase I, sin embargo se reduce proporcionalmente el tamaño del embrión.

35 De forma sorprendente, ahora se ha encontrado que aumentar la expresión de un ácido nucleico que codifica un polipéptido LBD proporciona plantas que tienen aumento de biomasa y/o aumento de rendimiento de semillas, en relación con plantas de control.

40 De acuerdo con una realización, se proporciona un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en una planta en relación con plantas de control, que comprenden aumentar la expresión de un ácido nucleico que codifica un polipéptido LBD en una planta.

#### Definiciones

##### Polipéptido(s)/Proteína(s)

45 Los términos “polipéptido” y “proteína” se utilizan de forma intercambiable aquí y se refieren a aminoácidos en una forma polimérica de cualquier longitud, ligados juntos mediante enlaces de péptido.

##### Polinucleótido(s)/Ácido nucleico(s)/Secuencia de ácidos nucleicos/secuencia de nucleótidos

50 Los términos “polinucleótidos”, “secuencia de ácidos nucleicos”, “secuencia de nucleótidos”, “ácido nucleico”, “molécula de ácido nucleico” se utilizan de forma intercambiable aquí y se refiere a nucleótidos, ya sea ribonucleótidos o desoxirribonucleótidos o una combinación de ambos, en una forma no ramificada polimérica de cualquier longitud.

## Plantas de Control

La selección de plantas de control adecuadas es una parte rutinaria de una configuración experimental y puede incluir plantas de tipo silvestre correspondientes o plantas correspondientes sin los genes de interés. La planta de control es normalmente de la misma especie de planta o incluso de la misma variedad que las plantas que se van a evaluar. La planta de control también puede ser un nulicigoto de la planta que se va a evaluar. Los nulicigotos son individuos que pierden el transgén mediante segregación. Una "planta de control" como se utiliza aquí se refiere no solo a plantas completas, sino también a partes de planta, que incluyen semillas y partes de semilla.

## Homólogo(s)

Los "homólogos" de una proteína abarcan péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones, supresiones y/o inserciones de aminoácido en relación con la proteína no modificada en cuestión y que tienen actividad biológica y funcional similar que la proteína no modificada de la cual se derivan.

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

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 también inserciones intra-secuencia de aminoácidos únicos y múltiples. De manera general, 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. Ejemplos de proteínas o péptidos de fusión de terminal N o C incluyen el dominio de unión o dominio de activación de un activador transcripcional como se utiliza en el sistema de dos híbridos de levadura, proteínas de recubrimiento de fago, (histidina)- 6- tag, glutationa S- transferasa- tag, proteína A, maltosa- proteína de unión, reductasa dihidrofolato, epítopo Tag-100, epítopo cmyc, epítopo FLAG®, lacZ, CMP (calmodulina- péptido de unión), epítopo HA, epítopo de proteína C y epítopo VSV.

Una sustitución se refiere al reemplazo de aminoácidos de la proteína con otros aminoácidos que tienen propiedades similares (tales como hidrofobicidad, hidrofilicidad, antigenicidad similar, propensión para formar o descomponer estructuras  $\alpha$ -helicoidales o estructuras de  $\beta$ -lamina). Las sustituciones de aminoácido son normalmente de residuos únicos, pero se pueden agrupar dependiendo de restricciones funcionales colocadas sobre el polipéptido; las inserciones usualmente serán del orden de aproximadamente 1 a 10 residuos de aminoácidos. Las sustituciones de aminoácido son preferiblemente sustituciones conservadoras de aminoácido. Las tablas de sustitución conservadora son bien conocidas en la técnica (véase por ejemplo Creighton (1984) Proteins. W.H. Freeman and Company (Eds) y Tabla 1 adelante).

30 Tabla 1: Ejemplos de sustituciones de aminoácidos conservados

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		

- Las sustituciones, supresiones y/o inserciones de aminoácido se pueden hacer fácilmente utilizando técnicas sintéticas de péptido bien conocidas en la técnica, tal como síntesis de péptido de fase sólida y similares, 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 5 técnicas para realizar mutaciones de sustitución en sitios predeterminados en ADN son bien conocidas por aquellos expertos en la técnica e incluyen mutagénesis M13, mutagénesis T7-Gen in vitro (USB; Cleveland, OH), mutagénesis Dirigida a Sitio QuickChange (Stratagene, San Diego, CA), mutagénesis dirigida a sitio mediada por PCR u otros protocolos de mutagénesis dirigida a sitio.

#### Derivados

- 10 Los "derivados" incluyen péptidos, oligopéptidos, polipéptidos que se pueden comparar con la secuencia de aminoácidos de la forma de origen natural de la proteína, tal como la proteína de interés, comprenden sustituciones de aminoácidos con residuos de aminoácidos de origen no natural, o adiciones de residuos de aminoácidos de origen no natural. Los "derivados" de una proteína también abarcan péptidos, oligopéptidos, polipéptidos que 15 comprenden residuos de aminoácidos de origen natural alterados (glicosilados, acilados, prenilados, fosforilados, miristoilados, sulfatados etc.) o no alterados de forma natural comparados con la secuencia de aminoácidos de una forma de origen natural del polipéptido. Un derivado también puede comprender uno o más sustituyentes de no aminoácido o adiciones en comparación con la secuencia de aminoácidos de las cuales se deriva, por ejemplo una molécula indicadora u otro ligando, se une de forma covalente o no covalente a la secuencia de aminoácidos, tal como una molécula indicadora que se une para facilitar su detección, y residuos de aminoácidos de origen no 20 natural en relación con la secuencia de aminoácidos de una proteína de origen natural. Adicionalmente, "derivados" también incluyen fusiones de la forma de origen natural de la proteína con péptidos etiquetados tales como FLAG, HIS6 o tioredoxina (para una revisión de péptidos etiquetados, véase Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

#### Ortólogo(s)/Parálogo(s)

- 25 Los ortólogos y parálogos abarcan conceptos evolucionados utilizados para describir las relaciones ancestrales de genes. Los parálogos son genes dentro de las mismas especies que se han originado a través de la duplicación de un gen ancestral; los ortólogos son genes de organismos que se han originado a través de especiación, y también se derivan de un gen ancestral común.

#### Dominio

- 30 El término "dominio" se refiere a un grupo de aminoácidos conservados en posiciones específicas a lo largo de una alineación de secuencias de proteínas relacionadas con la evolución. Mientras que los aminoácidos en otras posiciones pueden variar entre homólogos, los aminoácidos que se conservan altamente en posiciones específicas indican aminoácidos que son probablemente esenciales en la estructura, estabilidad o función de una proteína. Identificados por su alto grado de conservación en secuencias alineadas de una familia de proteínas homólogas, se 35 pueden utilizar como identificadores para determinar si cualquier polipéptido en cuestión pertenece a la familia de polipéptido identificada previamente.

#### Motivo/Secuencia consenso/Distintivo

- 40 El término "motivo" o "secuencia consenso" o "distintivo" se refieren a una región conservada corta en la secuencia de proteínas relacionadas con la evolución. Los motivos son partes de dominios frecuentemente altamente conservadas, pero también pueden incluir solo parte del dominio, o se pueden ubicar fuera del dominio conservado (si todos los aminoácidos del motivo caen fuera de un dominio definido).

#### Hibridación

- 45 El término "hibridación" como se define aquí es un proceso en donde las secuencias de nucleótidos sustancialmente complementaria homólogas se desnaturalizan entre sí. El proceso de hibridación puede ocurrir completamente en solución, es decir ambos ácidos nucleicos complementarios están en solución. El proceso de hibridación también puede ocurrir con uno de los ácidos nucleicos complementarios inmovilizados a una matriz tal como perlas magnéticas, perlas de sefarosa o cualquier otra resina. El proceso de hibridación adicionalmente puede ocurrir con uno de los ácidos nucleicos complementarios inmovilizados a un soporte sólido tal como una membrana de 50 nitrocelulosa o nylon o inmovilizado mediante por ejemplo fotolitografía para, por ejemplo, un soporte de vidrio silíceo (la última conocida como disposiciones o microdisposiciones de ácido nucleico o como chips de ácido nucleico). Con el fin de permitir que ocurra hibridación, las moléculas de ácido nucleico generalmente se desnaturalizan térmica o químicamente para fundir una cadena doble en dos cadenas sencillas y/o para eliminar horquillas u otras estructuras secundarias desde los ácidos nucleicos de cadena única.

El término "rigurosidad" se refiere a las condiciones bajo las cuales tiene lugar una hibridación. La rigurosidad de la hibridación es influenciada por las condiciones tales como temperatura, concentración de sal, resistencia iónica y composición reguladora de hibridación. De manera general, se seleccionan condiciones de rigurosidad baja para ser aproximadamente 30° C más baja que el punto de fusión térmico ( $T_m$ ) para la secuencia específica en una resistencia iónica definida y pH. Las condiciones de rigurosidad media son cuando la temperatura está 20° C por debajo del  $T_m$ , y las condiciones de rigurosidad altas son cuando la temperatura está 10° C por debajo del  $T_m$ . Las condiciones de hibridación de alta rigurosidad normalmente se utilizan para aislar las secuencias de hibridación que tienen alta similitud de secuencia con la secuencia objetivo de ácidos nucleicos. Sin embargo, los ácidos nucleicos se puede desviar en secuencia y aún codificar un polipéptido sustancialmente idéntico, debido a la degeneración del código genético. Por lo tanto algunas veces las condiciones de hibridación de rigurosidad media se pueden necesitar para identificar dichas moléculas de ácido nucleico.

El  $T_m$  es la temperatura bajo pH y resistencia iónica definida, en el que 50% de la secuencia objetivo hibrida a una sonda emparejada perfectamente. El  $T_m$  es dependiente de las condiciones de solución y la composición base y longitud de la sonda. Por ejemplo, las secuencias más largas hibridan específicamente en temperaturas mayores. El índice máximo de hibridación se obtiene de aproximadamente 16° C hasta 32° C por debajo de  $T_m$ . La presencia de cationes monovalentes en la solución de hibridación reduce la repulsión electrostática entre las dos cadenas de ácido nucleico promoviendo de esta manera la formación híbrida; este efecto es visible para concentraciones de sodio de hasta 0.4 M (para concentraciones mayores, se puede ignorar este efecto). La formamida reduce la temperatura de fusión de dúplex ADN-ADN y ADN-ARN con 0.6 a 0.7° C para cada porcentaje de formamida, y la adición de 50% de formamida permite que se realice la hibridación a 30 a 45° C, aunque se reducirá el índice de hibridación. Los emparejamientos incorrectos de par base reducen el índice de hibridación y la estabilidad térmica de los dúplex. En promedio y para sondas grandes, el  $T_m$  reduce aproximadamente 1 ° C por % de emparejamiento base incorrecto. Se puede calcular el  $T_m$  utilizando las siguientes ecuaciones, dependiendo de los tipos de híbridos:

1) híbridos de ADN- ADN (Meinkoth and Wahl, Anal. Biochem., 138: 267- 284, 1984).  $T_m = 81.5^{\circ}\text{C} + 16.6x\log_{10}[\text{Na}^+]^a + 0.41x\%[\text{G/C}^b] - 500x[\text{L}^{\circ}]^c - 0.61x\% \text{ de formamida}$

2) híbridos de ADN- ARN o ARN- ARN:  $T_m = 79.8 + 18.5 (\log_{10}[\text{Na}^+]^a) + 0.58 (\% \text{G/C}^b) + 11.8 (\% \text{G/C}^b)^2 - 820/\text{L}^c$

3) híbridos de oligo-ADN o oligo-ARNd:

Para <20 nucleótidos:  $T_m = 2 (I_n)$

Para 20-35 nucleótidos:  $T_m = 22 + 1.46 (I_n)$

<sup>a</sup> o para otro catión monovalente, pero solo exacto en el rango 0.01-0.4 M.

<sup>b</sup> solo exacto para %GC en el rango 30% a 75%.

<sup>c</sup> L = longitud de dúplex en pares base.

<sup>d</sup> oligo, oligonucleótido;  $I_n$  = longitud efectiva del cebador = 23(no. of G/C)+(no. of A/T).

Se puede controlar la unión no específica utilizando una cualquiera de una serie de técnicas conocidas tales como, por ejemplo, bloquear la membrana con soluciones que contienen proteína, adiciones de ARN, ADN, y SDS heterólogo al regulador de hibridación, y tratamiento con Rnasa. Se pueden desarrollar las sondas no homólogas, una serie de hibridaciones al variar uno de (i) reducir progresivamente la temperatura de desnaturización (por ejemplo desde 68° C hasta 42° C) o (ii) reducir progresivamente la concentración de formamida (por ejemplo desde 50% hasta 0%). El experto es consciente de diversos parámetros que se pueden alterar durante hibridación y que ya sea mantendrán o cambiarán las condiciones de rigurosidad.

Aunque las condiciones de hibridación, la especificidad de hibridación normalmente también depende de la función de lavados de post-hibridación. Para eliminar el fondo resultante de la hibridación no específica, las muestras se lavan con soluciones de sal diluida. Los factores críticos de dichos lavados incluyen la resistencia iónica y temperatura de la solución de lavado final: a mayor concentración de sal y mayor temperatura de lavado, mayor rigurosidad del lavado. Las condiciones de lavado se realizan normalmente en o por debajo de rigurosidad de hibridación. Una hibridación positiva proporciona una señal que es por lo menos dos veces de aquella del fondo. De manera general, las condiciones rigurosas adecuadas para ensayos de hibridación de ácido nucleico o procedimientos de detección de amplificación de gen son como se estableció anteriormente. También se pueden seleccionar condiciones más o menos rigurosas. El experto es consciente que se pueden alterar diversos parámetros durante el lavado y que se mantendrán o cambiarán las condiciones de rigurosidad.

Por ejemplo, las condiciones de hibridación de alta rigurosidad típicas para los híbridos de ADN mayores de 50 nucleótidos abarcan hibridación a 65° C en 1x SSC o a 42° C en 1x SSC y 50% de formamida, seguido por lavado a 65° C en 0.3x SSC. Ejemplos de condiciones de hibridación de media rigurosidad para híbridos de ADN más largos de 50 nucleótidos abarcan hibridación a 50° C en 4x SSC o a 40° C en 6x SSC y 50% de formamida, seguido por

- 5 lavado a 50° C en 2x SSC. La longitud del híbrido es la longitud anticipada para la hibridación de ácido nucleico. Cuando se hibridan los ácidos nucleicos de la secuencia conocida, se puede determinar la longitud de híbrido al 10 alinear las secuencias e identificar las regiones conservadas descritas aquí. 1x SSC es NaCl 0.15 M y citrato de sodio 15 mM; la solución de hibridación y las soluciones de lavado pueden incluir adicionalmente 5x de reactivo de Denhardt, 0.5-1.0% de SDS, 100 mg/ml de ADN de esperma de salmón fragmentado, desnaturizado, 0.5% de pirofosfato de sodio.

Para los propósitos de definir el nivel de rigurosidad, se puede hacer referencia a Sambrook et al. (2001) Molecular Cloning: a laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, CSH, New York o a Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989 y actualizaciones anuales).

#### Variante de empalme

- 15 El término “variante de empalme” como se utiliza aquí abarca variantes de una secuencia de ácidos nucleicos en la que se han escindido, reemplazado, desplazado o agregado intrones y/o exones seleccionados, o en los cuales los intrones se han cortado o alargado. Dichas variantes son unas en las que se retiene sustancialmente la actividad biológica de la proteína; esto se puede alcanzar al retener selectivamente los segmentos funcionales de la proteína. 20 Dichas variantes de empalme se pueden encontrar en la naturaleza o pueden ser hechas por el hombre. Los métodos para predecir y aislar dichas variantes de empalme son bien conocidos en la técnica (véase por ejemplo Foissac and Schiex (2005) BMC Bioinformatics 6: 25).

#### Variante alélica

- 25 Los alelos o variantes alélicas son formas alternativas de un gen dado, ubicadas en la misma posición cromosómica. Las variantes alélicas abarcan Polimorfismos de Nucleótido Único (SNPs), así como también Polimorfismos de Inserción/Supresión Pequeña (INDELs). El tamaño de los INDEL es usualmente menor de 100 bp. Los SNP y los INDEL forman los grupos más grandes de secuencias variantes en cepas polimórficas de origen natural de la mayoría de los organismos.

#### Transposiciones genéticas/Evolución dirigida

- 30 Las transposiciones genéticas o evolución dirigida consiste de iteraciones de transposición de ADN seguida por detección apropiada y/o selección para generar variantes de ácidos nucleicos o porciones de los mismos que codifican proteínas que tienen una actividad biológica modificada (Castle et al., (2004) Science 304 (5674): 1151- 4; patentes Estadounidenses 5, 811, 238 y 6, 395, 547).

#### Elemento regulador/Secuencia de control/Promotor

- 35 Los términos “elemento regulador”, “secuencia de control” y “promotor” todos se utilizan de forma intercambiable aquí y se deben tomar en un contexto amplio para referirse a secuencias de ácidos nucleicos reguladoras capaces de efectuar expresión de secuencias a las cuales se ligan. El término “promotor” normalmente se refiere a una secuencia de ácidos nucleicos de control ubicada en dirección 5' desde el inicio transcripcional de un gen que se involucra en el reconocimiento y unión de las polimerasa de ARN y otras proteínas, dirigiendo de esta manera la transcripción de un ácido nucleico ligado operablemente. Se abarcan por los términos mencionados anteriormente 40 secuencias reguladoras transcripcionales derivadas de un gen genómico eucariótico clásico (que incluye el recuadro TATA que se requiere para iniciación de transcripción exacta, con o sin una secuencia de recuadro CCAAT) y elementos reguladores adicionales (es decir secuencias de activación en dirección 5', mejoradores e inactivadores) que alteran la expresión génica en respuesta a estímulos de desarrollo y/o externos, o en una forma específica a tejido. También se incluye dentro del término una secuencia reguladora transcripcional de un gen procariótico 45 clásico, en cuyo caso puede incluir una secuencia reguladora transcripcional del recuadro 35 y/o secuencia reguladora transcripcional del recuadro 10. El término “elemento regulador” también abarca una molécula de fusión sintética o derivado que confiere, activa o mejora la expresión de una molécula de ácido nucleico en una célula, tejido u órgano.

- 50 Un “promotor de planta” comprende elementos reguladores, que median la expresión de un segmento de secuencia codificante en células de planta. De acuerdo con lo anterior, un promotor de planta no necesita ser de origen vegetal, sino que se puede originar de virus o microorganismos, por ejemplo de virus que atacan las células de planta. El “promotor de planta” también se puede originar de una célula de planta, por ejemplo de la planta que se transforma con la secuencia de ácidos nucleicos que se va a expresar en el proceso de la invención y se describe aquí. Esto también aplica a otras señales reguladoras de “planta”, tales como terminadores de “planta”. Los promotores en

dirección 5' de las secuencias de nucleótidos útiles en los métodos de la presente invención se pueden modificar por una o más sustitución (es), inserción (es) y/o supresión (es) de nucleótido sin interferir con la funcionalidad o actividad de ya sea los promotores, el marco de lectura abierto (ORF) o la región reguladora 3' tal como terminadores u otras regiones reguladoras 3' que se ubican lejos del ORF. Es posible adicionalmente que se

5 aumente la actividad de los promotores mediante modificación de su secuencia, o que se puedan reemplazar completamente por más promotores activos, incluso promotores de organismos heterólogos. Para la expresión en plantas, la molécula de ácido nucleico, como se describió anteriormente, se debe ligar operablemente a o comprender un promotor adecuado que exprese el gen en el punto exacto en el tiempo y con el patrón de expresión espacial requerido.

10 Para la identificación de los promotores funcionalmente equivalentes, se puede analizar la resistencia del promotor y/o patrón de expresión de un promotor candidato, por ejemplo, al ligar operablemente el promotor a un gen indicador y ensayar el nivel y el patrón de expresión del gen indicador en diversos tejidos de la planta. Los genes indicadores bien conocidos adecuados incluyen, por ejemplo, beta-glucuronidasa o beta-galactosidasa. La actividad del promotor se ensaya al medir la actividad enzimática de la beta-glucuronidasa o beta-galactosidasa. La resistencia del promotor y/o patrón de expresión entonces se puede comparar con aquella de un promotor de referencia (tal como la utilizada en los métodos de la presente invención). Alternativamente, la resistencia del promotor se puede ensayar al cuantificar los niveles de mARN o al comparar los niveles de mARN del ácido nucleico utilizado en los métodos de la presente invención, con los niveles de mARN de los genes constitutivos tales como 18S rARN, utilizando los métodos conocidos en la técnica, tales como transferencia Northern con análisis densitométrico de autorradiogramas, PCR cuantitativo en tiempo real o RT-PCR (Heid et al., 1996 Genome Methods 6: 986-994). De manera general, por "promotor débil" se entiende un promotor que conduce la expresión de una secuencia codificante a un nivel bajo. Por "nivel" se entiende niveles de aproximadamente 1/10,000 transcripciones a aproximadamente 1/100,000 transcripciones, a aproximadamente 1/500,000 transcripciones por célula. Por el contrario, un "promotor fuerte" conduce la expresión de una secuencia codificante a un alto nivel, o a aproximadamente 1/10 transcripciones a aproximadamente 1/100 transcripciones a aproximadamente 1/1000 transcripciones por célula. De manera general, por "promotor de resistencia media" se entiende un promotor que conduce la expresión de una secuencia codificante a un nivel más bajo que un promotor fuerte, en particular, a un nivel que está en todos los casos por debajo de aquel obtenido cuando está bajo el control de un promotor CaMV 35S.

15

20

25

30 Ligado operablemente

El término "ligado operablemente" tal como se utiliza aquí, se refiere a un enlace funcional entre la secuencia promotora y el gen de interés, de tal manera que la secuencia promotora es capaz de iniciar la transcripción del gen de interés.

Promotor constitutivo

35 Un "promotor constitutivo" se refiere a un promotor que es transcripcionalmente activo durante la mayoría, pero no necesariamente todas, las fases de crecimiento y desarrollo, y bajo la mayoría de condiciones ambientales, en por lo menos una célula, tejido u órgano. La Tabla 2a adelante proporciona ejemplos de promotores constitutivos.

Tabla 2a: Ejemplos de promotores constitutivos

Fuente de gen	Referencia
Actina	McElroy et al, Plant Cell, 2: 163-171, 1990
HMGP	WO 2004/070039
CAMV 35S	Odell et al, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson et al., Physiol. Plant. 100:456-462, 1997
GOS2	de Pater et al, Plant J Nov;2(6):837-44, 1992, WO 2004/065596
Ubiquitina	Christensen et al, Plant Mol. Biol. 18: 675-689, 1992
Arroz ciclofilina	Buchholz et al, Plant Mol Biol. 25(5): 837-43, 1994
Maíz H3 histona	Lepetit et al, Mol. Gen. Genet. 231:276-285, 1992

Fuente de gen	Referencia
Alfalfa H3 histona	Wu et al. Plant Mol. Biol. 11:641-649, 1988
Actina 2	An et al, Plant J. 10(1); 107-121, 1996
34S FMV	Sanger et al., Plant. Mol. Biol., 14, 1990: 433-443
Subunidad pequeña de rubisco	US 4,962,028
OCS	Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553
SAD1	Jain et al., Crop Science, 39 (6), 1999: 1696
SAD2	Jain et al., Crop Science, 39 (6), 1999: 1696
nos	Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846
V-ATPasa	WO 01/14572
Súper-promotor	WO 95/14098
Proteínas de recuadro G	WO 94/12015

#### Promotor ubicuo

Un promotor ubicuo está activo en sustancialmente todos los tejidos o células de un organismo.

#### Promotor regulado por el desarrollo

- 5 Un promotor regulado por el desarrollo está activo durante ciertas etapas de desarrollo o en partes de la planta que experimentan cambios de desarrollo.

#### Promotor inducible

- 10 Un promotor inducible ha inducido o incrementado la iniciación de la transcripción en respuesta a un estímulo químico (para una revisión véase Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48: 89- 108), ambiental o físico, o puede ser “inducible por estrés”, es decir, activado cuando una planta se expone a diversas condiciones de estrés, o un “inducible por patógeno” es decir activado cuando se expone una planta a exposición de diversos agentes patógenos.

#### Promotor específico de órgano/específico a tejido

- 15 Un promotor específico de órgano/específico a tejido es aquel que es capaz de preferiblemente iniciar la transcripción en ciertos órganos o tejidos, tales como hojas, raíces, tejido de la semilla, etc. Por ejemplo, un “promotor específico de raíz” es un promotor que es transcripcionalmente activo predominantemente en las raíces de plantas, sustancialmente para la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión permeable en estas otras partes de la planta. Los promotores capaces de iniciar la transcripción en determinadas células sólo se denominan aquí como “específico a célula”.

- 20 Los ejemplos de promotores específicos de raíz se enumeran en la Tabla 2B adelante:

Tabla 2b: Ejemplos de promotores específicos de raíz

Fuente de gen	Referencia
RCc3	Plant Mol Biol. 1995 Jan;27(2):237-48
PHT1 de Arabidopsis	Kovama et al., 2005; Mudge et al. (2002, planta J. 31:341)
Transportador de fosfato Medicago	Xiao et al., 2006
Pyk10 de Arabidopsis	Nitz et al. (2001) Plant Sci 161(2): 337-346
genes que se pueden expresar en raíz	Tingey et al., EMBO J. 6: 1, 1987.
Gen inducible por auxina de tabaco	Van der Zaal et al., Plant Mol. Biol. 16, 983, 1991.
β-tubulina	Oppenheimer, et al., Gene 63: 87, 1988.
Genes específico a raíz de tabaco	Conkling, et al., Plant physiol. 93: 1203, 1990.
Gen B. napus G1-3b	Patente de los Estados Unidos No. 5, 401, 836
SbPRP1	Suzuki et al., Plant Mol. Biol. 21: 109-119, 1993.
LRX1	Baumberger et al. 2001, Genes & Dev. 15:1128
BTG-26 Brassica napus	US 20050044585
LeAMT1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
El LeNRT1-1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
gen patatin clase I (papa)	Liu et al., Plant Mol. Biol. 153:386-395, 1991.
KDC1 (Daucus carota)	Downey et al. (2000, J. Biol. Chem. 275:39420)
Gen TobRB7	W Song (1997) PhD Thesis, North Carolina State University, Raleigh, NC USA
OsRAB5a (arroz)	Wang et al. 2002, Plant Sci. 163:273
ALF5 (Arabidopsis)	Diener et al. (2001, Plant Cell 13:1625)
NRT2;1Np (N. plumbaginifolia)	Quesada et al. (1997, Plant Mol. Biol. 34:265)

5 Un promotor específico de semilla es transcripcionalmente activo predominantemente en tejido de semilla, pero no necesaria ni exclusivamente en tejido de semilla (en casos de expresión permeable). El promotor específico de semilla puede estar activo durante el desarrollo de la semilla y/o durante germinación. El promotor específico de semilla puede ser específico a endospermo/aleurona/embrión. Ejemplos de promotores específicos de semillas (específicos a endospermo/aleurona/embrión) se muestran en la Tabla 2d, 2e, 2f. Ejemplos adicionales de promotores específicos de semilla se dan en Qing Qu and Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuya descripción se incorpora mediante referencia aquí como si se estableciera completamente.

Tabla 2c: Ejemplos de promotores específicos de semilla

Fuente de gen	Referencia
genes específicos a semilla	Simon et al., Plant Mol. Biol. 5: 191, 1985;
	Scofield et al., J. Biol. Chem. 262: 12202, 1987.;
	Baszcynski et al., Plant Mol. Biol. 14: 633, 1990.
Albúmina de nuez de Brasil	Pearson et al., Plant Mol. Biol. 18: 235-245, 1992.
legumina	Ellis et al., Plant Mol. Biol. 10: 203-214, 1988.
glutelina (arroz)	Takaiwa et al., Mol. Gen. Genet. 208: 15-22, 1986;
	Takaiwa et al., FEBS Letts. 221: 43-47, 1987.
Zeína	Matzke et al Plant Mol Biol, 14(3):323-32 1990
napA	Stalberg et al, Plant 199: 515-519, 1996.
LMW y HMW glutenina-1 de trigo	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
SPA de trigo	Albani et al, Plant Cell, 9: 171-184, 1997
α, β, λ-gliadinas de trigo	EMBO J. 3:1409-15, 1984
Promotor ltr1 de cebada	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
B1, C, D, hordeina de cebada	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996
DOF de cebada	Mena et al, The Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
promotor sintético	Vicente-Carabajosa et al., Plant J. 13: 629-640, 1998.
prolамина NRP33 de arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
α-globulina Glb-1 de arroz	Wu et al, Plant Cell Physiology 39(8) 885-889, 1998
OSH1 de arroz	Sato et al, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
α-globulina REB/OHP-1 de arroz	Nakase et al. Plant Mol. Biol. 33: 513-522, 1997
ADP-pirofosforilasa glucosa de arroz	Trans Res 6:157-68, 1997
familia del gen ESR de maíz	Plant J 12:235-46, 1997
α-kafirina de sorgo	DeRose et al., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39:257-71, 1999
oleosina de arroz	Wu et al, J. Biochem. 123:386, 1998
oleosina de girasol	Cummins et al., Plant Mol. Biol. 19: 873-876, 1992

Fuente de gen	Referencia
proteína ribosómica 40S PRO0117, putativa de arroz	WO 2004/070039
PRO0136, alanina aminotransferasa de arroz	no publicada
Inhibidor de tripsina PRO0147, ITR1 (cebada)	no publicada
PRO0151, WSI18 de arroz	WO 2004/070039
PRO0175, RAB21 de arroz	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039
α-amilasa (Amy32b)	Lanahan et al, Plant Cell 4:203-211, 1992; Skriver et al, Proc Natl Acad Sci USA 88:7266-7270, 1991
Gen similar a catepsina β	Cejudo et al, Plant Mol Biol 20:849-856, 1992
Ltp2 de Cebada	Kalla et al., Plant J. 6:849-60, 1994
Chi26	Leah et al., Plant J. 4:579-89, 1994
Maíz B- de Perú	Selinger et al., Genetics 149:1125-38, 1998

Tabla 2d: Ejemplos de promotores específicos de endospermo

Fuente de gen	Referencia
glutelina (arroz)	Takaiwa et al. (1986) Mol Gen Genet 208:15-22; Takaiwa et al. (1987) FEBS Letts. 221:43-47
Zeína	Matzke et al., (1990) Plant Mol Biol 14(3): 323-32
LMW y HMW glutenina-1 de trigo	Colot et al. (1989) Mol Gen Genet 216:81-90, Anderson et al. (1989) NAR 17:461-2
SPA de trigo	Albani et al. (1997) Plant Cell 9:171-184
gliadinas de trigo	Rafalski et al. (1984) EMBO 3:1409-15
promotor ltr1 de cebada	Diaz et al. (1995) Mol Gen Genet 248(5):592-8
B1, C, D, hordeina de cebada	Cho et al. (1999) Theor Appl Genet 98:1253-62; Muller et al. (1993) Plant J 4:343-55; Sorenson et al. (1996) Mol Gen Genet 250:750-60
DOF de cebada	Mena et al, (1998) Plant J 116(1): 53-62
blz2	Onate e al. (1999) J Biol Chem 274(14):9175-82
promotor sintético	Vicente-Carabajosa et al. (1998) Plant J 13:629-640

Fuente de gen	Referencia
prolамина NRP33 de arroz	Wu et al, (1998) <i>Plant Cell Physiol</i> 39(8) 885-889
globulina Glb-1 de arroz	Wu et al. (1998) <i>Plant Cell Physiol</i> 39(8) 885-889
globulina REB/OHP-1 de arroz	Nakase et al. (1997) <i>Plant Molec Biol</i> 33: 513-522
ADP-pirofosforilasa glucosa de arroz	Russell et al. (1997) <i>Trans Res</i> 6:157-68
familia del gen ESR de maíz	Opsahl-Ferstad et al. (1997) <i>Plant J</i> 12:235-46
kafirina de sorgo	DeRose et al. (1996) <i>Plant Mol Biol</i> 32:1029-35

Tabla 2e: Ejemplos de promotores específicos de embrión:

Fuente de gen	Referencia
arroz OSH1	Sato et al, <i>Proc. Natl. Acad. Sci. USA</i> , 93: 8117-8122, 1996
KNOX	Postma-Haarsma et al, <i>Plant Mol. Biol.</i> 39:257-71, 1999
PRO0151	WO 2004/070039
PRO0175	WO 2004/070039
PRO005	WO 2004/070039
PRO0095	WO 2004/070039

Tabla 2f: Ejemplos de promotores específicos de aleurona:

Fuente de gen	Referencia
α-amilasa (Amy32b)	Lanahan et al, <i>Plant Cell</i> 4:203-211, 1992; Skriver et al, <i>Proc Natl Acad Sci USA</i> 88: 7266-7270, 1991
Gen similar a catepsina β	Cejudo et al, <i>Plant Mol Biol</i> 20:849-856, 1992
Ltp2 de Cebada	Kalla et al., <i>Plant J.</i> 6:849-60, 1994
Chi26	Leah et al., <i>Plant J.</i> 4:579-89, 1994
Maíz B de Perú	Selinger et al., <i>Genetics</i> 149:1125-38, 1998

5

Un promotor específico de tejido verde tal como se define en aquí es un promotor que es transcripcionalmente activo predominantemente en el tejido verde, sustancialmente para la exclusión de cualquier otra parte de una planta, mientras que todavía permite cualquier expresión permeable en estas otras partes de la planta. I

10 Los ejemplos de promotores específicos de tejido verde que se pueden utilizar para realizar los métodos de la invención se muestran en la Tabla 2g.

Tabla 2g: Ejemplos de promotores específicos de tejido verde

Gen	Expresión	Referencia
Ortofosfato diquinsa de Maíz	Específico a hoja	Fukavama et al., 2001
Fosfoenolpiruvato carboxilasa de Maíz	Específico a hoja	Kausch et al., 2001
Fosfoenolpiruvato carboxilasa de Arroz	Específico a hoja	Liu et al., 2003
subunidad de Rubisco pequeña de Arroz	Específico a hoja	Nomura et al., 2000
beta expansina EXBP9 de arroz	Específico a raíz	WO 2004/070039
subunidad de Rubisco pequeña de guandú	Específico a hoja	Panguluri et al., 2005
RBCS3A de guisante	Específico a hoja	

- 5 Otro ejemplo de un promotor específico de tejido es un promotor específico de meristemo, que es transcripcionalmente activo predominantemente en el tejido meristemático, sustancialmente para la exclusión de cualquier otra parte de una planta, mientras que aún permite para cualquier expresión permeable en estas otras partes. Los ejemplos de promotores específicos de meristemo verdes que se pueden utilizar para llevar a cabo los métodos de la invención se muestran en la Tabla 2 h adelante.

Tabla 2h: Ejemplos de promotores específicos de meristemo

Fuente de gen	Patrón de Expresión	Referencia
OSH1 de arroz	Meristemo apical de brote, desde etapa embrioglobal hasta etapa de plántula	Sato et al. (1996) Proc. Natl. Acad. Sci. USA, 93: 8117-8122
Metalotioneina de arroz	Específico a Meristemo	BAD87835.1
WAK1 & WAK 2	Meristemos apicales de brote y raíz, y en hojas de expansión y sépalos	Wagner & Kohorn (2001) Plant Cell 13(2): 303-318

## 10 Terminador

- El término “terminador” abarca una secuencia de control que es una secuencia de ADN en el extremo de una unidad transcripcional que señala el procesamiento 3' y la poliadenilación de una transcripción primaria y la terminación de la transcripción. El terminador se puede derivar a partir del gen natural, a partir de una variedad de otros genes de plantas, o a partir de T-ADN. El terminador que se agrega se puede derivar de, por ejemplo, genes de nopalina sintasa o sintasa octopina, o alternativamente de otro gen de planta, o menos preferiblemente de cualquier otro gen eucariótico.

## 15 Modulación

- El término “modulación” significa en relación con la expresión o la expresión génica, un proceso en el que el nivel de expresión se cambia por dicha expresión génica en comparación con la planta de control, el nivel de expresión se puede aumentar o disminuir. La expresión original, no modulada puede ser de cualquier clase de expresión de un ARN estructural (rARN, tARN) o mARN con la traducción posterior. El término “modular la actividad” significará cualquier cambio de la expresión de las secuencias de ácidos nucleicos de la invención o proteínas codificadas, lo que conduce a un aumento del rendimiento y/o aumento del crecimiento de las plantas.

## 20 Expresión

- El término “expresión” o “expresión génica” se refiere a la transcripción de un gen específico o genes específicos o construcciones genéticas específicas. El término “expresión” o “expresión génica”, en particular, significa la

transcripción de un gen o genes o construcción genética en el ARN estructural (rARN, tARN) o mARN con o sin la posterior traducción de este último en una proteína. El proceso incluye la transcripción de ADN y el procesamiento del producto de mARN resultante.

#### Aumento de expresión/sobreexpresión

5 El término “aumento de expresión” o “sobreexpresión” como se usa aquí, significa cualquier forma de expresión que sea adicional al nivel de expresión de tipo silvestre original.

10 Los métodos para aumentar la expresión de genes o productos génicos están bien documentados en la técnica e incluyen, por ejemplo, sobreexpresión conducida por promotores apropiados, el uso de potenciadores de transcripción o potenciadores de traducción. Los ácidos nucleicos aislados que sirven como promotores o 15 potenciadores de los elementos se pueden introducir en una posición apropiada (normalmente en dirección 5') de una forma no heteróloga de un polinucleótido con el fin de regular de forma ascendente la expresión de un ácido nucleico que codifica el polipéptido de interés. Por ejemplo, los promotores endógenos se pueden alterar *in vivo* por mutación, supresión, y/o sustitución (véase, Kmiec, US 5,565,350; Zarling et al., WO9322443), o se pueden introducir promotores aislados en una célula de planta en la orientación y distancia apropiada desde un gen de la 20 presente invención con el fin de controlar la expresión del gen.

Si se desea la expresión del polipéptido, generalmente es deseable incluir una región de poliadenilación en el extremo 3' de una región codificante de polinucleótido. La región de poliadenilación se puede derivar del gen natural, de una variedad de otros genes de plantas, o de T-ADN. Se puede derivar la secuencia de extremo 3' que se va a 25 agregar de, por ejemplo, genes de nopalina sintasa o sintasa octopina, o alternativamente de otro gen de planta, o menos preferiblemente de cualquier otro gen eucariótico.

También se puede agregar una secuencia de intrón a la región 5' no traducida (UTR) o la secuencia de codificación de la secuencia de codificación parcial para aumentar la cantidad del mensaje de madurez que se acumula en el citosol. Se ha demostrado que la inclusión de un intrón empalmable en la unidad de transcripción en las 30 construcciones de expresión de plantas y animales aumenta la expresión génica tanto los niveles de mARN y proteína hasta 1000 veces (Buchman and Berg (1988) Mol. Cell Biol. 8: 4395 -4,405; Callis et al. (1987) Genes Dev. 1:1183-1200). Dicha mejora de intrón de la 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 35 son conocidos en la técnica. Para información general, véase: The Maize Handbook, Capítulo 116, Freeling and Walbot, Eds, Springer, N.Y. (1994).

#### 30 Gen endógeno

La referencia aquí a un gen “endógeno” no sólo se refiere al gen en cuestión como se encuentra en una planta en su forma natural (es decir, sin que exista ninguna intervención humana), sino también se refiere a ese mismo gen (o un ácido nucleico/gen sustancialmente homólogo) en una forma aislada posteriormente (re) introducida en una planta (un transgén). Por ejemplo, una planta transgénica que contiene dicho transgén se puede encontrar con una 35 reducción sustancial de la expresión del transgen y/o reducción sustancial de la expresión del gen endógeno. Se puede aislar el gen aislado de un organismo o puede ser hecho por el hombre, por ejemplo mediante síntesis química.

#### Disminución de la expresión

40 La referencia aquí a “disminución de la expresión” o “eliminación o reducción sustancial” de expresión se considera que significa una disminución en la expresión génica endógena y/o niveles de polipéptido y/o actividad del polipéptido en relación con las plantas de control. La reducción o eliminación sustancial tiene un orden creciente de preferencia de por lo menos 10%, 20%, 30%, 40% o 50%, 60%, 70%, 80%, 85%, 90%, o 95%, 96%, 97%, 98%, 99% o más reducido en comparación con aquel de las plantas de control.

45 Para la reducción o eliminación sustancial de la expresión se requiere un gen endógeno en una planta, una longitud suficiente de nucleótidos sustancialmente contiguos de una secuencia de ácido nucleico. Con el fin de realizar la inactivación de los genes, esto puede ser tan poco como 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 o menos nucleótidos, alternativamente, esto puede ser tanto como el gen completo (que incluye el UTR 5' y/o 3', ya sea en parte o en todo). El tramo de nucleótidos sustancialmente contiguos se puede derivar a partir del ácido nucleico que codifica la proteína de interés (gen objetivo), o de cualquier ácido nucleico capaz de codificar un ortólogo, parólogo u 50 homólogo de la proteína de interés. Preferiblemente, el tramo de nucleótidos sustancialmente contiguos es capaz de formar enlaces de hidrógeno con el gen objetivo (ya sea en cadena sentido o antisentido), más preferiblemente, el tramo de nucleótidos sustancialmente contiguos tiene, en orden creciente de preferencia, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 100% de identidad de secuencia para el gen objetivo (ya sea cadena sentido o antisentido). Una secuencia de ácido nucleico que codifica un polipéptido (funcional) no es un requisito

para los diversos métodos descritos aquí para la reducción o eliminación sustancial de la expresión de un gen endógeno.

Se puede lograr esta reducción o eliminación sustancial de la expresión utilizando herramientas y técnicas de rutina. Un método preferido para la reducción o eliminación sustancial de la expresión génica endógena es al introducir y expresar en una planta una construcción genética en la que el ácido nucleico (en este caso un tramo de nucleótidos sustancialmente contiguos derivados a partir del gen de interés, o de cualquier nucleico ácido capaz de codificar un ortólogo, parálogo u homólogo de una cualquiera de la proteína de interés) se clona como una repetición invertida (en parte o completamente), separados por un separador (ADN no codificante).

En dicho un método preferido, la expresión del gen endógeno se reduce o se elimina sustancialmente a través de inactivación mediada por ARN utilizando una repetición invertida de un ácido nucleico o una parte del mismo (en este caso un tramo de nucleótidos sustancialmente contiguos derivado del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), preferiblemente capaz de formar una estructura de horquilla. La repetición invertida se clona en un vector de expresión que comprende secuencias de control. Una secuencia de ácidos nucleicos de ADN no codificante (un separador, por ejemplo, un fragmento de región de adhesión de matriz (MAR), un intrón, un polilíquido, etc.) se ubica entre los dos ácidos nucleicos invertidos que forman la repetición invertida. Después de la transcripción de la repetición invertida, se forma un ARN químico con una estructura autocomplementaria (parcial o total). Esta estructura de ARN de doble cadena se refiere como la horquilla de ARN (hpARN). El hpARN se procesa por la planta en siARNs que se incorporan en un complejo de inactivación inducido por ARN (RISC). El RISC adicionalmente divide las transcripciones de mARN, reduciendo sustancialmente de esta manera el número de transcripciones de mARN que se van a traducir en polipéptidos. Para detalles generales adicionales véase, por ejemplo, Grierson et al. (1998) documento WO 98/53083; Waterhouse et al. (1999) documento WO 99/53050).

El desempeño de los métodos de la invención no se basa en introducir y expresar en una planta una construcción genética en la que el ácido nucleico se clona como una repetición invertida, sino que se puede utilizar uno o más de los diversos métodos de "inactivación génica" bien conocidos para lograr los mismos efectos.

Uno de dichos métodos para la reducción de expresión génica endógena es inactivación mediada por ARN de la expresión génica (regulación por disminución). La inactivación en este caso se activa en una planta por una secuencia de ARN de doble cadena (dsARN) que es sustancialmente similar al gen endógeno objetivo. Este dsARN se procesa adicionalmente por la planta en aproximadamente 20 a aproximadamente 26 nucleótidos llamados ARN de interferencia corta (siARNs). Los siARNs se incorporan en un complejo de inactivación inducido por ARN (RISC) que divide la transcripción de mARN del gen objetivo endógeno, reduciendo sustancialmente de esta manera el número de transcripciones de mARN que se van a traducir en un polipéptido. Preferiblemente, la secuencia de ARN de doble cadena corresponde a un gen objetivo.

Otro ejemplo de un método de inactivación de ARN implica la introducción de secuencias de ácidos nucleicos o partes de las mismas (en este caso un tramo de nucleótidos sustancialmente contiguos derivados del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés) en una orientación del sentido en una planta. "Orientación del sentido" se refiere a una secuencia de ADN que es homóloga a una transcripción de mARN de la misma. Por lo tanto se podría introducir en una planta por lo menos una copia de la secuencia de ácidos nucleicos. La secuencia de ácidos nucleicos adicional reducirá la expresión del gen endógeno, dando lugar a un fenómeno conocido como co-supresión. La reducción de la expresión génica será más pronunciada si se introducen diversas copias adicionales de una secuencia de ácidos nucleicos en la planta, ya que existe una correlación positiva entre los altos niveles de transcripción y la activación de la co-supresión.

Otro ejemplo de un método de inactivación del ARN implica el uso de secuencias de ácidos nucleicos antisentido. Una secuencia de ácidos nucleicos "antisentido" comprende una secuencia de nucleótidos que es complementaria a una secuencia de ácidos nucleicos "sentido" que codifica una proteína, es decir, complementaria a la cadena codificante de una molécula de cADN de doble cadena o complementaria a una secuencia de transcripción de mARN. La secuencia de ácidos nucleicos antisentido es preferiblemente complementaria al gen endógeno que se va a inactivar. La complementariedad se puede ubicar en la "región codificante" y/o en la "región no codificante" de un gen. El término "región codificante" se refiere a una región de la secuencia de nucleótidos que comprende codones que se traducen en residuos de aminoácidos. El término "región no codificante" se refiere a secuencias 5' y 3' que flanquean la región codificante que se transcriben pero no se traducen en aminoácidos (también, denominados regiones no traducidas 5' y 3').

Se pueden diseñar las secuencias de ácidos nucleicos antisentido de acuerdo con las reglas de Watson y Crick. La secuencia de ácidos nucleicos antisentido puede ser complementaria a la secuencia de ácidos nucleicos completa (en este caso un tramo de nucleótidos sustancialmente contiguos derivados del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), pero también puede ser un oligonucleótido que es antisentido solo a una parte de la secuencia de ácidos nucleicos (que incluye el mARN 5' y 3' UTR). Por ejemplo, la secuencia de oligonucleótido antisentido puede ser complementaria a la región que rodea el

sitio de inicio de traducción de una transcripción de mARN que codifica un polipéptido. La longitud de una secuencia de oligonucleótidos antisentido adecuados se conoce en la técnica y puede iniciar desde aproximadamente 50, 45, 40, 35, 30, 25, 20, 15 o 10 nucleótidos de longitud o menos. Una secuencia de ácidos nucleicos antisentido de acuerdo con la invención se puede construir utilizando síntesis química y reacciones de ligado enzimático utilizando métodos conocidos en la técnica. Por ejemplo, se puede sintetizar una secuencia de ácidos nucleicos antisentido (por ejemplo, una secuencia de oligonucleótidos antisentido) químicamente utilizando nucleótidos de origen natural o nucleótidos diversamente modificados diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre las secuencias de ácidos nucleicos antisentido y sentido, por ejemplo, se pueden utilizar derivados de fosforotioato y nucleótidos sustituidos con acridina. Ejemplos de nucleótidos modificados que se pueden utilizar para generar las secuencias de ácidos nucleicos antisentido son bien conocidos en la técnica. Las modificaciones de nucleótidos conocidas incluyen metilación, ciclización y 'tapas' y sustitución de uno o más de los nucleótidos de origen natural con un análogo tal como inosina. Otras modificaciones de nucleótidos son bien conocidas en la técnica.

Se puede producir la secuencia de ácidos nucleicos antisentido biológicamente utilizando un vector de expresión en el que se ha subclonado una secuencia de ácidos nucleicos en una orientación antisentido (es decir, el ARN se transcribe del ácido nucleico insertado será de una orientación antisentido a un ácido nucleico objetivo de interés). Preferiblemente, la producción de secuencias de ácidos nucleicos antisentido en plantas ocurre por medio de una construcción de ácido nucleico integrado de forma estable que comprende un promotor, un oligonucleótido antisentido ligado operablemente, y un terminador.

Las moléculas de ácido nucleico utilizadas para inactivar en los métodos de la invención (ya sea introducidas en una planta o generadas *in situ*) hibridan con o se unen a transcripciones de mARN y/o ADN genómico que codifica un polipéptido para inhibir de esta manera la expresión de la proteína, por ejemplo, al inhibir la transcripción y/o traducción. La hibridación puede ser mediante complementariedad de nucleótido convencional para formar un dúplex estable, o, por ejemplo, en el caso de una secuencia de ácidos nucleicos antisentido que se une a dúplex de ADN, a través de interacciones específicas en el surco mayor de doble hélice. Se pueden introducir secuencias de ácidos nucleicos antisentido en una planta mediante transformación o inyección directa en un sitio de tejido específico. Alternativamente, se pueden modificar las secuencias de ácidos nucleicos antisentido para dirigirse a células seleccionadas y luego se administran sistémicamente. Por ejemplo, para administración sistémica, las secuencias de ácidos nucleicos antisentido se pueden modificar de tal manera que se unen específicamente a receptores o antígenos expresados en una superficie celular seleccionada, por ejemplo, al ligar la secuencia de ácidos nucleicos antisentido a péptidos o anticuerpos que se unen a receptores de superficie celular o antígenos. Las secuencias de ácidos nucleicos antisentido también se pueden suministrar a las células utilizando los vectores descritos aquí.

De acuerdo con un aspecto adicional, la secuencia de ácidos nucleicos antisentido es una secuencia de ácidos nucleicos a-anomérica. Una secuencia de ácidos nucleicos a-anomérica forma híbridos de doble cadena específicos con ARN complementario en los cuales, contrariamente a las unidades b usuales, las cadenas corren paralelas entre sí (Gaultier et al. (1987) Nucl Ac Res 15: 6625 a 6641). La secuencia de ácidos nucleicos antisentido también puede comprender un 2'-o-metilribonucleótido (Inoue et al. (1987) Nucl Ac Res 15, 6131-6.148) o un análogo de ARN - ADN químico (Inoue et al. (1987) FEBS Lett. 215, 327-330).

La reducción o eliminación sustancial de la expresión génica endógena también se pueden realizar utilizando ribozimas. Las ribozimas son moléculas de ARN catalíticas con actividad de ribonucleasa que son capaces de dividir una secuencia de ácidos nucleicos de cadena única, tal como un mARN, para el cual tienen una región complementaria. Por lo tanto, las ribozimas (por ejemplo, ribozimas cabeza de martillo (descritas en Haselhoff and Gerlach (1988) Nature 334, 585-591) se puede utilizar para dividir catalíticamente las transcripciones de mARN que codifican un polipéptido, reduciendo sustancialmente de esta manera el número de transcripciones de mARN que se van a traducir en un polipéptido. Se puede diseñar una ribozima que tiene especificidad para una secuencia de ácidos nucleicos puede ser diseñado (véase, por ejemplo: Cech et al. Patente Estadounidense 4,987,071, y Cech et al. patente Estadounidense No. 5,116,742). Alternativamente, se pueden utilizar las transcripciones de mARN que corresponden a una secuencia de ácidos nucleicos para seleccionar un ARN catalítico que tiene una actividad ribonucleasa específica de un grupo de moléculas de ARN (Bartel and Szostak (1993) Science 261, 1411-1418). El uso de ribozimas para la Atkins et al. (1994) WO 94/00012; Lenne et al. (1995) WO 95/03404; Lutziger et al. (2000) WO 00/00619; Prinsen et al. (1997) WO 97/13865 y Scott et al. (1997) WO 97/38116).

También se puede conseguir la inactivación génica mediante mutagénesis de inserción (por ejemplo, inserción de T-ADN o inserción de transposón) o mediante estrategias como se describe por, entre otros, Angell and Baulcombe ((1999) Plant J 20(3): 357 - 62), (Amplicón VIGS WO 98/36083), o Baulcombe (WO 99/15682).

También puede ocurrir inactivación génica si se presenta una mutación en un gen endógeno y/o una mutación en un gen aislado/ácido nucleico posteriormente introducido en una planta. La reducción o eliminación sustancial pueden ser provocadas por un polipéptido no funcional. Por ejemplo, el polipéptido se puede unir a diversas proteínas que interactúan; una o más mutaciones y/o truncamiento por lo tanto pueden proporcionar un polipéptido que es todavía

capaz de unirse a proteínas que interactúan (tales como proteínas receptoras), pero que no pueden exhibir su función normal (por ejemplo, ligando de señalización).

Un método adicional para inactivar genes es al objetivar secuencias de ácidos nucleicos complementarias a la región reguladora del gen (por ejemplo, el promotor y/o potenciadores) para formar estructuras de triple hélice que evitan la transcripción del gen en las células objetivo. Véase Helene, C., Anticancer Drug Res. 6, 569-84, 1991; Helene et al., Ann. N.Y. Acad. Sci. 660, 27-36 1992; y Maher, L.J., Bioassays 14, 807-15, 1992.

Otros métodos, tales como el uso de anticuerpos dirigidos a un polipéptido endógeno para inhibir su función en la planta, o la interferencia en la ruta de señalización en la que se implica un polipéptido, serán bien conocidos para el experto. En particular, se puede prever que las moléculas hechas por el hombre pueden ser útiles para inhibir la función biológica de un polipéptido objetivo, o para interferir con la ruta de señalización en el que está involucrado el polipéptido objetivo.

Alternativamente, se puede establecer un programa de detección para identificar en una población de plantas variantes naturales de un gen, cuyas variantes codifican polipéptidos con actividad reducida. También se pueden utilizar dichas variantes naturales, por ejemplo, para realizar la recombinación homóloga.

15 Se pueden utilizar microARN artificiales y/o naturales (miARNs) para inactivar la expresión génica y/o traducción del mARN. Los miARNs endógenos son ARN pequeños de cadena simple de normalmente 19 a 24 nucleótidos de longitud. Funcionan principalmente para regular la expresión génica y/o traducción del mARN. La mayoría de los microARNs de plantas (miARNs) tienen complementariedad perfecta o casi perfecta, con sus secuencias objetivo. Sin embargo, existen objetivos naturales con hasta cinco emparejamientos incorrectos. Se procesan de ARNs no 20 codificantes más largos con estructuras de retro-plegado características mediante RNAsas específicas de doble cadena de la familia Dicer. Luego del procesamiento, se incorporan en el complejo de inactivación inducido por ARN (RISC) al unirse a su componente principal, una proteína Argonaute. Los miARN sirven como los componentes de especificidad de RISC, ya que los pares de bases objetivas ácidos nucleicos, en su mayoría, mARNs en el citoplasma. Los eventos reguladores posteriores incluyen división de mARN objetivo y la destrucción y/o inhibición 25 de la traducción. Los efectos de la sobreexpresión de los genes miARN por lo tanto se reflejan a menudo en niveles de mARN reducidos de los genes objetivo.

Los microARN artificiales (amiARNs), que tienen normalmente 21 nucleótidos de longitud, se pueden modificar genéticamente específicamente para regular negativamente la expresión génica de genes de interés múltiples o únicos. Los determinantes de la selección objetivo de microARN de planta son bien conocidos en la técnica. Se han 30 definido los parámetros empíricos para el reconocimiento objetivo y se pueden utilizar para ayudar en el diseño de amiARNs específicos, (Schwab et al., Dev. Cell 8, 517-527, 2005). Las herramientas convenientes para el diseño y generación de amiARNs y sus precursores también están disponibles para el público (Schwab et al., Plant Cell 18, 1121-1133, 2006 al).

35 Para un desempeño óptimo, las técnicas de inactivación de genes utilizados para reducir la expresión en una planta de un gen endógeno requiere el uso de secuencias de ácidos nucleicos a partir de plantas monocotiledóneas para la transformación de plantas monocotiledóneas, y de plantas dicotiledóneas para la transformación de las mismas. Preferiblemente, se introduce una secuencia de ácidos nucleicos de cualquier especie de planta dada en esa misma especie. Por ejemplo, una secuencia de ácidos nucleicos de arroz se transforma en una planta de arroz. Sin embargo, no es un requisito absoluto que la secuencia de ácidos nucleicos que se va a introducir se origine de las 40 mismas especies de plantas como la planta en la que se introducirá. Es suficiente que se presente una homología sustancial entre el gen objetivo endógeno y el ácido nucleico que se va a introducir.

45 Se describieron anteriormente ejemplos de diversos métodos para la reducción o eliminación sustancial de la expresión en una planta de un gen endógeno. Una persona experta en la técnica fácilmente sería capaz de adaptar los métodos mencionados anteriormente para inactivar a fin de lograr la reducción de la expresión de un gen endógeno en una planta completa o en partes de la misma a través de por ejemplo el uso de un promotor adecuado.

#### Marcador seleccionable (gen)/gen indicador

50 "Marcador seleccionable", "gen marcador seleccionable" o "gen indicador" incluye cualquier gen que confiere un fenotipo sobre una célula en la que se expresa para facilitar la identificación y/o selección de células que se transfecan o transforman con una construcción de ácido nucleico de la invención. Estos genes marcadores permiten la identificación de una transferencia exitosa de las moléculas de ácido nucleico a través de una serie de diferentes principios. Se pueden seleccionar marcadores adecuados a partir de marcadores que confieren resistencia antibiótica o herbicida, que introducen un nuevo rasgo metabólico o que permiten selección visual. Ejemplos de genes marcadores seleccionables incluyen genes que confieren resistencia a los antibióticos (tales como nptII que fosforilan neomicina y kanamicina, o hpt, fosforilar higromicina, o genes que confieren resistencia a, por ejemplo, bleomicina, estreptomicina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espechinomicina o

blasticidina), a los herbicidas (por ejemplo barra que proporciona resistencia a Basta®; aroA o gox que proporcionan resistencia contra glifosato, o los genes que confieren resistencia a, por ejemplo, imidazolinona, fosfinotricina o sulfonilurea), o genes que proporcionan un rasgo metabólico (tal como manA que permite a las plantas utilizar manosa como única fuente de carbono o isomerasa de xilosa para la utilización de xilosa, o marcadores antinutritivos tales como la resistencia a la 2-desoxiglucosa). La expresión de genes marcadores visuales resulta en la formación de color (por ejemplo  $\beta$ -glucuronidasa, GUS o  $\beta$ -galactosidasa con sus sustratos de color, por ejemplo X-Gal), luminiscencia (tal como el sistema luciferina/luceferasa) o fluorescencia (Proteína Verde Fluorescente, GFP, y derivados de la misma). Esta lista representa sólo un pequeño número de posibles marcadores. El experto está familiarizado con dichos marcadores. Se prefieren diferentes marcadores, dependiendo del organismo y el método de selección.

Se sabe que luego de la integración estable o transitoria de ácidos nucleicos en células de planta, sólo una minoría de las células se lleva al ADN externo y, si se desea, se integra en su genoma, dependiendo del vector de expresión utilizado y de la técnica de transfección utilizada. Para identificar y seleccionar estos integrantes, un gen que codifica un marcador seleccionable (tal como aquellos descritos anteriormente) se introduce usualmente en las células anfitrionas junto con el gen de interés. Estos marcadores se pueden utilizar, por ejemplo en mutantes en los que estos genes no son funcionales, por ejemplo, mediante supresión mediante métodos convencionales. Adicionalmente, las moléculas de ácido nucleico que codifican un marcador seleccionable se pueden introducir en una célula anfitriona en el mismo vector que comprende la secuencia que codifica los polipéptidos de la invención o que se utilizan en los métodos de la invención, o bien se pueden identificar en un vector separado. Las células que se han transfecido de manera estable con el ácido nucleico introducido, por ejemplo, mediante selección (por ejemplo, las células que han integrado el marcador seleccionable sobreviven mientras que las otras células mueren).

En razón a que los genes marcadores, particularmente los genes para resistencia a antibióticos y herbicidas, ya no se requieren o no se desean en la célula anfitriona transgénica que una vez se han introducido los ácidos nucleicos con éxito, el proceso de acuerdo con la invención para introducir los ácidos nucleicos ventajosamente emplea técnicas que permiten la eliminación o división de estos genes marcadores. Un dicho método es lo que se conoce como co-transformación. El método de co-transformación emplea simultáneamente dos vectores para la transformación, un vector que lleva el ácido nucleico de acuerdo con la invención y un segundo que lleva del gen marcador. Una gran proporción de los transformantes recibe o, en el caso de las plantas, comprende (hasta 40% o más de los transformantes), ambos vectores. En caso de transformación con Agrobacteria, los transformantes usualmente reciben sólo una parte del vector, es decir, la secuencia flanqueada por el ADN-T, que usualmente representa el casete de expresión. Los genes marcadores se pueden eliminar posteriormente de la planta transformada al realizar cruces. En otro método, los genes marcadores integrados en un transposón se utilizan para la transformación junto con el ácido nucleico deseado (conocido como la tecnología de Ac/Ds). Los transformantes se pueden cruzar con una fuente de transposasa o los transformantes se transforman con una construcción de ácido nucleico que confiere la expresión de una transposasa, de forma transitoria o estable. En algunos casos (aprox. 10%), el transposón salta del genoma de la célula anfitriona una vez que ha tenido lugar la transformación con éxito y se pierde. En una serie de casos adicional, el transposón salta a una ubicación diferente. En estos casos, el gen marcador debe ser eliminado al realizar cruces. En microbiología, se desarrollan técnicas que hacen posible o facilitan la detección de este tipo de eventos. Un método ventajoso adicional se basa en lo que se conoce como sistemas de recombinación; cuya ventaja es que se puede prescindir de la eliminación al cruzarla. El sistema más conocido de este tipo es aquel que se conoce como el sistema Cre/lox. CRE1 es una recombinasa que elimina las secuencias situadas entre las secuencias loxP. Si el gen marcador se integra entre las secuencias loxP, se elimina una vez que ha tenido lugar la transformación con éxito, mediante la expresión de la recombinasa. Sistemas de recombinación adicionales son los sistemas HIN/HIX, FLP/FRT y REP/STB (Tribble et al., J. Biol. Chem., 275, 2000: 22255 a 22267; Velmurugan et al., J. Cell Biol., 149, 2000: 553-566). Es posible una integración específica a sitio en el genoma de la planta de las secuencias de ácidos nucleicos de acuerdo con la invención. Naturalmente, estos métodos también se pueden aplicar a los microorganismos tales como levaduras, hongos o bacterias.

#### Transgénicos/transgén/Recombinante

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 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 aquellas construcciones resultan de métodos recombinantes en cualquiera de

- (a) las secuencias de ácidos nucleicos que codifican proteínas útiles en los métodos de la invención, o
- (b) secuencia genética de control que se liga operativamente con la secuencia de ácidos nucleicos de acuerdo con la invención, por ejemplo un promotor, o
- (c) a) y b)

no se ubican en su ambiente genético natural o se han modificado por métodos recombinantes, es posible que la modificación tome la forma de, por ejemplo, una sustitución, adición, supresión, inversión o inserción de uno o más residuos de nucleótidos. Se entiende que el ambiente genético natural significa el sitio genómica natural o cromosómico en la planta original o la presencia en una colección genómica. En el caso de una colección genómica, el ambiente genético natural de la secuencia de ácidos nucleicos se retiene preferiblemente, por lo menos en parte. El ambiente flanquea la secuencia de ácidos nucleicos por lo menos en un lado y tiene una longitud de secuencia de por lo menos 50 pb, preferiblemente por lo menos 500 pb, especialmente preferiblemente por lo menos 1000 pb, más preferiblemente por lo menos 5000 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 codifican un polipéptido útil en los métodos de la presente invención, tal como se definió anteriormente - se convierte en un casete de expresión transgénico cuando este casete de expresión se modifica mediante métodos no naturales, sintéticos ("artificiales") tales como, por ejemplo, el tratamiento mutagénico. Los métodos adecuados se describen, por ejemplo, en los documentos US 5.565.350 o WO 00/15815.

Por lo tanto una planta transgénica para los propósitos de la invención se entiende que significa, como anteriormente, que los ácidos nucleicos utilizados en el método de la invención no están en su lugar natural en el genoma de dicha planta, siendo posible para los ácidos nucleicos ser expresados de forma 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, se ha modificado la secuencia con respecto a la secuencia natural, y/o que se han modificado las secuencias reguladoras de las secuencias naturales. Se entiende preferiblemente que transgénico significa la expresión de los ácidos nucleicos de acuerdo con la invención en un sitio no natural en el genoma, es decir, tiene lugar la expresión homóloga o, preferiblemente, heteróloga de los ácidos nucleicos. Se mencionan aquí las plantas transgénicas preferidas.

#### Transformación

El término "introducción" o "transformación" como se denomina aquí abarca la transferencia de un polinucleótido exógeno en una célula anfitriona, independientemente del método utilizado para la transferencia. El tejido de planta 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 allí. El tejido particular escogido variará dependiendo de los sistemas de propagación clonal disponibles para, y se adapta mejor a, las especies particulares que se están transformando. Ejemplos de tejidos objetivo incluyen discos de hojas, polen, embriones, cotiledones, hipocotilos, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, yemas axilares, y meristemos radiculares) y tejido meristemático inducido (por ejemplo, meristemo de cotiledón y meristemo de hipocotilo). El polinucleótido se puede introducir transitoriamente o establemente en una célula anfitriona y se puede mantener no integrado, por ejemplo, como un plásmido. Alternativamente, se puede integrar en el genoma anfitrión. La célula de planta transformada resultante luego se puede utilizar para regenerar una planta transformada de una manera conocida por los expertos en la técnica.

La transferencia de genes extraños en el genoma de una planta se denomina transformación. La transformación de especies de plantas es ahora una técnica bastante rutinaria. Ventajosamente, se pueden utilizar cualquiera de los 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 de plantas o células de plantas se pueden utilizar para transformación transitoria o estable. Los métodos de transformación incluyen el uso de liposomas, electroporación, productos químicos que aumentan la captación de ADN libre, inyección del ADN directamente en la planta, bombardeo de partículas, transformación utilizando virus o polen y microproyección. Los métodos se pueden seleccionar del método de calcio/polietilenglicol para protoplastos ((Krens, F.A. et al., (1982) Nature 296, 72- 74; Negruțiu I et al. (1987) Plant Mol Biol 8: 363- 373); electroporación de protoplastos (Shillito R.D. et al. (1985) Bio/ Technol 3, 1099- 1102); microinyección en material vegetal ((Crossway A et al., (1986) Mol. Gen Genet 202: 179- 185); bombardeo de partículas recubierto de ADN o ARN (Klein TM et al., (1987) Nature 327: 70) infección con virus (no integrador) y similares. Las plantas transgénicas, que incluyen plantas de cultivo transgénicas, se producen preferiblemente a través de transformación mediada por Agrobacterium. Un método de transformación ventajoso es la transformación en planta. Para este fin, es posible, por ejemplo, para permitir que las agrobacterias actúen sobre las semillas de plantas o que inoculen el meristemo de planta con agrobacterias. Ha demostrado ser especialmente conveniente de acuerdo con la invención permitir una suspensión de agrobacterias transformadas para actuar en la planta intacta o por lo menos en los primordios de flores. La planta se cultiva posteriormente hasta que se obtienen las semillas de la planta tratada (Clough and Bent, Plant J. (1998) 16, 735- 743). Los métodos para la transformación mediada por Agrobacterium de arroz incluyen métodos bien conocidos para la transformación de arroz, tales como aquellos descritos en cualquiera de los siguientes: solicitud de patente europea EP 1198985 A1, Aldemita and Hodges (Plant 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271- 282, 1994), cuyas descripciones se incorporan aquí mediante referencia como si se establecen completamente. En el caso de la transformación del maíz, el método preferido es como se describe ya sea en Ishida et al. (Nat. Biotechnol 14 (6): 745-50, 1996) o Frame et al. (Plant Physiol 129 (1): 13-22, 2002), cuyas descripciones se incorporan aquí mediante referencia como si se establecen completamente. Dichos

métodos se describen adicionalmente por vía de ejemplo en B. Jenes et al., *Techniques for Gene Transfer*, en: *Transgenic Plants*, Vol. 1, *Engineering and Utilization*, eds. S.D. Kung and R. Wu, Academic Press (1993) 128- 143 y en Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205- 225). Los ácidos nucleicos o construcciones que se van a expresar se clonian preferiblemente en un vector, que es adecuado para transformar *Agrobacterium tumefaciens*, por ejemplo pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Las agrobacterias transformadas por dicho vector luego se puede utilizar de una forma conocida para la transformación de plantas, tales como plantas utilizadas como un modelo, similar a *Arabidopsis thaliana* que está dentro del alcance de la presente invención no se considera como una planta de cultivo), o plantas de cultivo tales como, a modo de ejemplo, plantas de tabaco, por ejemplo al sumergir las hojas picadas o las hojas cortadas en una solución de agrobacterias y luego cultivarlas en medios adecuados. La transformación de plantas por medio de *Agrobacterium tumefaciens* se describe, por ejemplo, por Höfgen and Willmitzer en Nucl. Acid Res. (1988) 16, 9877 o se conoce, entre otras cosas de F. F. White, *Vectors for Gene Transfer in Higher Plants*; in *Transgenic Plants*, Vol. 1, *Engineering and Utilization*, eds. S.D. Kung and R. Wu, Academic Press, 1993, pp. 15- 38.

Adicionalmente a la transformación de las células somáticas, que luego se han regenerado en plantas intactas, también es posible transformar las células de los meristemos de plantas y, en particular, aquellas células que se desarrollan en gametos. En este caso, los gametos transformados siguen el desarrollo natural de la planta, dando lugar a plantas transgénicas. Así, por ejemplo, las semillas de *Arabidopsis* se tratan con agrobacterias y las semillas se obtienen de las plantas en desarrollo de las que una cierta proporción se transforma y por lo tanto es transgénica [Feldman, KA and Marks MD (1987). Mol Gen Genet 208: 274- 289; Feldmann K (1992). In: C Koncz, N- H Chua y J Shell, eds, *Methods in Arabidopsis Research*. Word Scientific, Singapore, pp. 274- 289]. Los métodos alternativos se basan en la eliminación repetida de las inflorescencias y la incubación del sitio de división en el centro de la roseta con agrobacterias transformadas, mediante el cual se pueden obtener las semillas transformadas del mismo modo en un punto posterior en el tiempo (Chang (1994). Plant J. 5: 551-558; Katavic (1994) Mol Gen Genet, 245: 363-370). Sin embargo, un método especialmente eficaz es el método de infiltración al vacío con sus modificaciones, tales como el método de "inmersión floral". En el caso de infiltración al vacío de *Arabidopsis*, se tratan plantas intactas bajo presión reducida, con una suspensión de agrobacterias [Bechthold, N (1993). CR Acad Sci Paris Life Sci., 316: 1194-1199], mientras que en el caso del método de "inmersión floral" el tejido floral en desarrollo se incuba brevemente con una suspensión de agrobacterias tratada con tensoactivo [Clough, SJ and Bent AF (1998) The Plant J. 16, 735-743]. Una cierta proporción de las semillas transgénicas se cosechan en ambos casos, y estas semillas se pueden distinguir de semillas no transgénicas al cultivar en las condiciones selectivas mencionadas anteriormente. Adicionalmente la transformación estable de plástidos es de ventaja porque los plástidos que se heredan por vía materna es la mayoría de los cultivos que reducen o eliminan el riesgo de flujo de transgenes a través del polen. La transformación del genoma del cloroplasto se logra generalmente por un proceso que se ha mostrado esquemáticamente en Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Brevemente las secuencias que se van a transformar se clonian junto con un gen marcador seleccionable entre las secuencias de flanqueo homólogas al genoma del cloroplasto. Estas secuencias de flanqueo homólogas dirigen la integración específica de sitio en el plastoma. La transformación de plástidos se ha descrito para muchas especies diferentes de plantas y se proporciona una visión general en [Feldman, KA and Marks MD (1987). Mol Gen Genet 208: 274- 289; Feldmann K (1992). In: C Koncz, N- H Chua y J Shell, eds, *Methods in Arabidopsis Research*. Word Scientific, Singapore, pp. 274- 289]. Progresos hacia la comercialización de la tecnología de transformación de plástidos. Trends Biotechnol. 21, 20 - 28. Recientemente se ha informado progreso biotecnológico adicional en forma de marcadores libres de transformantes de plástidos, que se pueden producir por un gen marcador co-integrado transitorio (Klaus et al., 2004, Nature Biotechnology 22 (2), 225-229).

#### Marcado de activación de T-ADN

El marcado de activación de T-ADN (Hayashi et al., Science (1992) 1350 - 1353), implica la inserción de T - ADN, que usualmente contiene un promotor (puede ser también un potenciador de la traducción o un intrón), en la región genómica del gen de interés o 10 kb en dirección 5' o 3' de la región codificante de un gen en una configuración tal que el promotor dirige la expresión del gen objetivo. Normalmente, la regulación de la expresión del gen objetivo por su promotor natural se interrumpe y el gen cae bajo el control del promotor recientemente introducido. El promotor normalmente se incorpora en un T-ADN. Este T - ADN se inserta aleatoriamente en el genoma de la planta, por ejemplo, a través de infección de *Agrobacterium* y conduce a la expresión modificada de genes cerca al T-ADN insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debido a la expresión modificada de genes cerca al promotor introducido.

#### TILLING

El término "TILLING" es una abreviatura de "Detección de Lesiones Locales Inducidas en Genomas" y se refiere a una tecnología de mutagénesis útil para generar y/o identificar ácidos nucleicos que codifican proteínas con expresión y/o actividad modificada. El TILLING también permite la selección de las plantas portadoras de dichas variantes mutantes. Estas variantes mutantes pueden exhibir expresión modificada, ya sea en la resistencia o en la ubicación o en el tiempo (si por ejemplo las mutaciones afectan el promotor). Estas variantes mutantes pueden exhibir actividad más alta que aquella exhibida por el gen en su forma natural. El TILLING combina mutagénesis de

alta densidad con métodos de selección de alto rendimiento. Las etapas que normalmente siguen al TILLING son: (a) mutagénesis EMS Redei GP and Koncz C (1992) In *Methods in Arabidopsis Research*, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, pp. 16-82; Feldmann et al., (1994) In Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; 5 Lightner J and Caspar T (1998) In J Martinez- Zapater, J Salinas, eds, *Methods on Molecular Biology*, Vol. 82. Humana Press, Totowa, NJ, pp 91-104); (b) Preparación de ADN y agrupamiento de individuos; (c) amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) 10 DHPLC, donde se detecta la presencia de un heterodúplex en un grupo como un pico extra en el cromatograma; (f) identificación del individuo mutante; y (g) secuenciación del producto de PCR mutante. Los métodos para TILLING son bien conocidos en la técnica (McCallum et al., (2000) *Nat Biotechnol* 18: 455-457; revisado por Stemple (2004) *Nat Rev Genet* 5 (2): 145-50).

#### Recombinación homóloga

La recombinación homóloga permite la introducción en un genoma de un ácido nucleico seleccionado en una 15 posición seleccionada definida. La recombinación homóloga es una tecnología estándar utilizada rutinariamente en ciencias biológicas para organismos inferiores tales como levadura o el musgo *Physcomitrella*. Los métodos para realizar la recombinación homóloga en plantas se han descrito no solamente para plantas modelo (Offringa et al. 20 (1990) *EMBO J* 9 (10): 3077-84) sino también para plantas de cultivo, por ejemplo arroz (Terada et al. (2002.) *Nat Biotech* 20 (10): 1030-4; Iida and Terada (2004) *Curr Opin Biotech* 15 (2): 132-8), y existen métodos que son generalmente aplicables con independencia del organismo objetivo (Miller et al, *Nature Biotechnol*. 25, 778-785, 2007).

#### Rendimiento

El término "rendimiento" en general significa un producto medible de valor económico, por lo general relacionado con un cultivo especificado, para un área, y para un período de tiempo. Las partes individuales de plantas contribuyen directamente con el rendimiento sobre la base de su número, tamaño y/o peso, o el rendimiento real es el 25 rendimiento por metro cuadrado para un cultivo y año, que se determina al dividir la producción total (incluye la producción cosechada y valorada) por metros cuadrados plantados. El término "rendimiento" de una planta se puede relacionar con la biomasa de planta (biomasa de raíz y/o brote), para los órganos reproductivos, y/o para los propágulos (tal como semillas) de esa planta.

#### Vigor temprano

30 El "vigor temprano" se refiere al crecimiento bien equilibrado activo y saludable especialmente durante las primeras etapas de crecimiento de las plantas, y puede resultar del aumento de la aptabilidad de la planta debido a, por ejemplo, las plantas que se adaptan mejor a su ambiente (es decir, optimizar el uso de los recursos energéticos y dividir en partes entre brote y raíz). Las plantas que tienen vigor temprano también muestran aumento en la supervivencia de plántulas y un mejor establecimiento del cultivo, que a menudo resulta en campos altamente 35 uniformes (con el crecimiento de cultivo de manera uniforme, es decir, con la mayoría de las plantas que alcanzan diversas etapas de desarrollo en sustancialmente el mismo tiempo), y a menudo mejor y más alto rendimiento. Por lo tanto, el vigor temprano se puede determinar al medir diversos factores, tales como el peso de mil granos, porcentaje de germinación, porcentaje de emergencia, crecimiento de plántulas, altura de plántulas, longitud de raíz, biomasa de raíz y brotes y muchos más.

#### 40 Aumento/Mejora/Potencia

Los términos "aumento", "mejora" o "potencia" son intercambiables y se entenderán en el sentido de la aplicación de por lo menos un 3%, 4%, 5%, 6%, 7%, 8%, 9% o 10%, preferiblemente por lo menos 15% o 20%, más preferiblemente 25%, 30%, 35% o 40% más de rendimiento y/o crecimiento en comparación con las plantas de control como se define aquí.

#### 45 Rendimiento de semilla

El aumento de rendimiento de semilla se puede manifestar como uno o más de los siguientes: a) un aumento en la biomasa de semilla (peso de semilla total), que puede ser sobre una base de semilla individual y/o por planta y/o por metro cuadrado; b) aumento del número de flores por planta; c) aumento del número de semillas (cargadas); d) índice de aumento de carga de semilla (que se expresa como la relación entre el número de semillas cargadas, dividido por el número total de semillas); e) aumento de índice de cosecha, que se expresa como una relación del rendimiento de las partes cosechables, tales como semillas, dividido por el total de biomasa, y f) aumento de peso de mil granos (PMG), que se extrae a partir del número de semillas cargadas contadas y su peso total. Un aumento de TKW puede resultar de un aumento de tamaño de semilla y/o peso de las semillas, y también puede resultar de un aumento de tamaño del embrión y/o endospermo.

Un aumento en el rendimiento de semilla también se puede manifestar como un aumento de tamaño de la semilla y/o volumen de semillas. Adicionalmente, un aumento en el rendimiento de semilla también se puede manifestar como un incremento en el área de semilla y/o longitud de semilla y/o ancho de la semilla y/o perímetro de semilla. El aumento de rendimiento también puede resultar en arquitectura modificada, o puede ocurrir debido a la arquitectura modificada.

5

#### Índice de verdor

El "índice de verdor" tal como se utiliza aquí se calcula a partir de imágenes digitales de plantas. Para cada píxel que pertenece a la planta objeto en la imagen, se calcula la relación entre el valor de verde versus el valor de rojo (en el modelo RGB para el color de codificación). El índice de verdor se expresa como el porcentaje de píxeles para los que la relación de verde a rojo excede un umbral dado. Bajo condiciones de crecimiento normales, bajo condiciones de crecimiento de estrés salino, y bajo condiciones de crecimiento de disponibilidad de nutrientes reducidos, el índice de verdor de las plantas se mide en la última formación de imagen antes de la floración. En contraste, bajo condiciones de crecimiento de estrés por sequía, el índice de verdor de las plantas se mide en la primera formación de imagen después de la sequía.

10

#### Planta

El término "planta" como se utiliza aquí abarca las plantas enteras, ancestros y progénesis de las plantas y partes de las mismas, que incluyen semillas, brotes, tallos, hojas, raíces (que incluyen tubérculos), flores y tejidos y órganos, en donde cada uno de los anteriormente mencionados comprenden el gen/ácido nucleico de interés. El término "planta" también abarca células de planta, cultivos en suspensión, tejido de callo, embriones, regiones meristemáticas, gametofitos, esporofitos, polen y microesporas, de nuevo en donde cada uno de los anteriormente mencionados comprenden el gen/ácido nucleico de interés.

Las plantas que son particularmente útiles en los métodos de la invención incluyen todas las plantas que pertenecen a la superfamilia de Viridiplantae, en particular plantas monocotiledóneas y dicotiledóneas, que incluyen forraje o leguminosas forrajeras, plantas ornamentales, cultivos de alimentos, árboles o arbustos seleccionados de la lista que comprende Acer spp., Actinidia spp., Abelmoschus spp., Agave sisalana, Agropyron spp., Agrostis stolonifera, Allium spp., Amaranthus spp., Ammophila arenaria, Ananas comosus, Annona spp., Apium graveolens, Arachis spp., Artocarpus spp., Asparagus officinalis, Avena spp. (por ejemplo Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida), Averrhoa carambola, Bambusa sp., Benincasa hispida, Bertholletia excelsa, Beta vulgaris, Brassica spp. (por ejemplo Brassica napus, Brassica rapa spp. (por ejemplo, Brassica napus, Brassica rapa spp. [canola, colza, nabina]), Cadaba farinosa, Camellia sinensis, Canna indica, Cannabis sativa, Capsicum spp., Carex elata, Carica papaya, Carissa macrocarpa, Carya spp., Carthamus tinctorius, Castanea spp., Ceiba pentandra, Cichorium endivia, Cinnamomum spp., Citrullus lanatus, Citrus spp., Cocos spp., Coffea spp., Colocasia esculenta, Cola spp., Corchorus sp., Coriandrum sativum, Corylus spp., Crataegus spp., Crocus sativus, Cucurbita spp., Cucumis spp., Cynara spp., Daucus carota, Desmodium spp., Dimocarpus longan, Dioscorea spp., Diospyros spp., Echinochloa spp., Elaeis (por ejemplo Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Eragrostis tef, Erianthus sp., Eriobotrya japonica, Eucalyptus sp., Eugenia uniflora, Fagopyrum spp., Fagus spp., Festuca arundinacea, Ficus carica, Fortunella spp., Fragaria spp., Ginkgo biloba, Glycine spp. (por ejemplo Glycine max, Soja hispida o Soja max), Gossypium hirsutum, Helianthus spp. (por ejemplo Helianthus annuus), Hemerocallis fulva, Hibiscus spp., Hordeum spp. (por ejemplo Hordeum vulgare), Ipomoea batatas, Juglans spp., Lactuca sativa, Lathyrus spp., Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus spp., Luffa acutangula, Lupinus spp., Luzula sylvatica, Lycopersicon spp. (por ejemplo Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma spp., Malus spp., Malpighia emarginata, Mammea americana, Mangifera indica, Manihot spp., Manilkara zapota, Medicago sativa, Melilotus spp., Mentha spp., Miscanthus sinensis, Momordica spp., Morus nigra, Musa spp., Nicotiana spp., Olea spp., Opuntia spp., Omphithopus spp., Oryza spp. (por ejemplo Oryza sativa, Oryza latifolia), Panicum miliaceum, Panicum virgatum, Passiflora edulis, Pastinaca sativa, Pennisetum sp., Persea spp., Petroselinum crispum, Phalaris arundinacea, Phaseolus spp., Phleum pratense, Phoenix spp., Phragmites australis, Physalis spp., Pinus spp., Pistacia vera, Pisum spp., Poa spp., Populus spp., Prosopis spp., Prunus spp., Psidium spp., Punica granatum, Pyrus communis, Quercus spp., Raphanus sativus, Rheum rhabarbarum, Ribes spp., Ricinus communis, Rubus spp., Saccharum spp., Salix sp., Sambucus spp., Secale cereale, Sesamum spp., Sinapis sp., Solanum spp. (por ejemplo Solanum tuberosum, Solanum integrifolium o Solanum lycopersicum), Sorghum bicolor, Spinacia spp., Syzygium spp., Tagetes spp., Tamarindus indica, Theobroma cacao, Trifolium spp., Tripsacum dactyloides, Triticosecale rimpau, Triticum cum spp. (por ejemplo Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum, Triticum monococcum o Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium spp., Vicia spp., Vigna spp., Viola odorata, Vitis spp., Zea mays, Zizania palustris, Ziziphus spp., entre otros.

#### Descripción detallada de la invención

Dominio LOB que comprende proteína (LOB: Límites de Órgano Lateral)

5 Sorprendentemente, se ha encontrado ahora que el aumento de expresión en una planta de un ácido nucleico que codifica un polipéptido LBD da a las plantas que tienen mayor rendimiento de semilla y/o el aumento de la biomasa con relación a las plantas de control. De acuerdo con una primera forma de realización, la presente invención proporciona un método para aumentar el rendimiento de semillas y/o la biomasa en las plantas en relación con las plantas de control, que comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD.

El aumento de la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD es al introducir y expresar en una planta de un ácido nucleico que codifica un polipéptido LBD.

10 Cualquier referencia en lo sucesivo a una "proteína útil en los métodos de la invención" se entiende que significa un polipéptido LBD como se define aquí. Cualquier referencia en lo sucesivo a un "ácido nucleico útil en los métodos de la invención" se entiende un ácido nucleico capaz de codificar dicho polipéptido LBD. El ácido nucleico que se va a introducir en una planta (y por lo tanto útil en la realización de los métodos de la invención) es cualquier ácido nucleico que codifica el tipo de proteína que se describirá ahora, en lo sucesivo también denominado "ácido nucleico LBD" o "gen LBD".

15 20 Un "polipéptido LBD" tal como se define aquí se refiere a cualquier polipéptido que comprende un dominio DUF260 (Pfam número de acceso PF03195, número de acceso Interpro IPR004883, véase Figura 1). Preferiblemente, la secuencia del polipéptido LBD, cuando se utiliza en la construcción de un árbol filogenético, tal como la representada en la Figura 3 (o tal como se define por Yang et al., 2006, utilizando el método de HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), se agrupan con el grupo de polipéptidos de dominio LOB clase II que comprenden la secuencia de aminoácidos representada por la SEQ ID NO: 2 más que con cualquier otro grupo.

La proteína LBD también comprende por lo menos uno de los siguientes motivos conservados:

Motivo 1: MSCNGCRXLRKGX (SEQ ID NO: 5); en la que X en la posición 8 puede ser cualquier aminoácido, pero preferiblemente V o I y en donde X en la posición 14 puede ser cualquier aminoácido, pero preferiblemente es uno de S, G o N.

25 Motivo 2: QXXATXFXAKFXGR (SEQ ID NO: 6), en donde X en la posición 2 puede ser cualquier aminoácido, pero preferiblemente uno de A, S, o G; en donde X en la posición 3 puede ser cualquier aminoácido, pero preferiblemente uno de N, Q, o H; donde X en la posición 6 puede ser cualquier aminoácido, pero preferiblemente uno de V, L, o I; en donde X en la posición 8 puede ser cualquier aminoácido, pero preferiblemente uno de L, V, A, o I, y en donde X en la posición 12 puede ser cualquier aminoácido, pero preferiblemente uno de Y o F.

30 Motivo 3: FXSLLXEAXG (SEQ ID NO: 7); en donde X en la posición 2 puede ser cualquier aminoácido, pero preferiblemente uno de R, S, K o Q; en donde X en la posición 6 puede ser cualquier aminoácido, pero preferiblemente uno de Y, H, o F, y en donde X en la posición 9 puede ser cualquier aminoácido, pero preferiblemente C o A.

35 Preferiblemente, la proteína LBD comprende más de siete, más preferiblemente por lo menos nueve residuos de Cys, y no comprende el distintivo DP (V/I) YG (SEQ ID NO: 8).

El dominio DUF260 se caracteriza por la presencia de un motivo C-x(2)-C-x(6)-C-x(3)-C (SEQ ID NO: 9), en donde X puede ser cualquier aminoácido.

40 45 El término "dominio" y "motivo" se define en la sección "definiciones" aquí. 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) Nucleic Acids Res 30, 242- 244, InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher and Bairoch (1994), Una sintaxis perfil generalizado de motivos de secuencias biomoleculares y su función en la interpretación de secuencia automática (In) 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), o Pfam (Bateman et al., Nucleic Acids Research 30 (1): 276- 280 (2002). Un grupo de herramientas en el análisis in silico de las secuencias de proteínas está disponible en el servidor proteómico ExPASy (Instituto suizo de Bioinformática (Gasteiger et al., ExPASy: the proteomics server for in- depth protein knowledge and analysis, Nucleic Acids Res. 31: 3784- 3788 (2003)). También se pueden identificar dominios utilizando técnicas de rutina, tales como por la alineación de secuencias.

50 Los métodos de alineación de secuencias para comparación son bien conocidos en la técnica, dichos métodos incluyen GAP, BESTFIT, BLAST, FASTA y TFASTA. El GAP utiliza el algoritmo de Needleman and Wunsch ((1970) J. Mol Biol. 48: 443-453) para encontrar la alineación global (es decir, abarca las secuencias completas) de dos secuencias que maximizan el número de emparejamientos y minimizan el número de espacios. El algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403-10) calcula el porcentaje de identidad de secuencia y realiza un análisis

estadístico de la similitud entre las dos secuencias. El software para realizar análisis BLAST está disponible públicamente a través del Centro Nacional de Información sobre Biotecnología (NCBI). Los homólogos se pueden identificar fácilmente utilizando, por ejemplo, el algoritmo de alineación de secuencias múltiples ClustalW (versión 1.83), con los parámetros de alineación por parejas por defecto, y un método de puntuación en porcentaje. Los 5 porcentajes de similitud e identidad globales también se pueden determinar utilizando uno de los métodos disponibles en el paquete de software MatGAT (Campanella et al., BMC Bioinformatics, 2003 Jul 10, 04:29. MatGAT: una aplicación que genera matrices de similitud/identidad que utilizan proteínas o secuencias de ADN). Se puede 10 realizar edición manual menor para optimizar la alineación entre motivos conservados, como sería evidente para un experto en la técnica. Adicionalmente, en lugar de utilizar secuencias de longitud completa para la identificación de homólogos, también se pueden utilizar dominios específicos. Los valores de identidad de secuencia se pueden 15 determinar sobre el ácido nucleico completo o secuencia de aminoácidos o sobre los dominios seleccionados o motivos conservados, utilizando los programas mencionados anteriormente que utilizan parámetros por defecto.

Adicionalmente, los polipéptidos LBD (por lo menos en su forma nativa), como factores de transcripción, 20 normalmente tienen actividad de unión a ADN. Las herramientas y técnicas para medir la actividad de unión de ADN 15 son bien conocidas en la técnica e incluyen, por ejemplo, ensayos de retardo en gel. Los métodos experimentales para la actividad de caracterización de factores de transcripción se pueden encontrar, por ejemplo, en Sambrook et al. (2001) Molecular Cloning: A Laboratory Manual, tercera Edición Cold Spring Harbor Laboratory Press, CSH, New York o en los Volúmenes 1 y 2 de Ausubel et al. (actualizados anualmente), Current Protocols in Molecular Biology, Current Protocols.

20 La presente invención se ilustra mediante la transformación de plantas con la secuencia de ácidos nucleicos representada por la SEQ ID NO: 1, que codifica la secuencia de polipéptido de SEQ ID NO: 2. Sin embargo, el desempeño de la invención no se limita a estas secuencias; los métodos de la invención ventajosamente se pueden 25 realizar utilizando cualquier ácido nucleico que codifica el LBD o polipéptido LBD como se define aquí.

Los ejemplos de ácidos nucleicos que codifican polipéptidos LBD se dan en la Tabla A1 del Ejemplo 1 aquí. Dichos 30 ácidos nucleicos son útiles en la realización de los métodos de la invención. Las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1 son secuencias de ejemplos de ortólogos y parálogos del polipéptido LBD representado por la SEQ ID NO: 2, los términos "ortólogos" y "parálogos" son como se define aquí. Se pueden 35 identificar fácilmente ortólogos y parálogos al realizar una denominada búsqueda de blast recíproca. Normalmente, esto implica un primer BLAST que implica BLASTing de una secuencia de consulta (por ejemplo, utilizando cualquiera de las secuencias enumeradas en la Tabla A1 del Ejemplo 1) contra cualquier base de datos de 40 secuencia, tal como la base de datos NCBI disponible públicamente. El BLASTN o TBLASTX (que utilizan valores estándar por defecto) se utilizan generalmente cuando se parte de una secuencia de nucleótidos, y BLASTP o TBLASTN (que utilizan valores estándar por defecto) cuando se parte de una secuencia de proteína. Se pueden filtrar optionalmente los resultados de BLAST. Las secuencias de longitud completa de cualquiera de los resultados 45 filtrados o los resultados no filtrados luego se someten nuevamente a análisis BLAST de nuevo (segundo BLAST) contra secuencias del organismo del que se deriva la secuencia de consulta (en donde la secuencia de consulta es la SEQ ID NO: 1 o SEQ ID NO: 2, el segundo BLAST por lo tanto, estaría en contra de las secuencias de Arabidopsis). Luego se comparan los resultados del primero y segundo BLAST. Se identifica un parálogo si un 50 acierto de alta clasificación del primer blast es de la misma especie que aquel de la que se deriva la secuencia de consulta, un BLAST entonces idealmente resulta en la secuencia de consulta entre los más altos aciertos; un ortólogo se identifica si un acierto de alta clasificación en el primer BLAST no es de la misma especie del que se deriva la secuencia de consulta, y preferiblemente sobre los resultados BLAST de nuevo en la secuencia de consulta que está entre los mayores aciertos.

Los aciertos de alta clasificación son aquellos que tienen valor E bajo. Cuanto menor sea el valor de E, más 55 significativa es la puntuación (o en otras palabras, menor será la posibilidad de que se encuentre acierto por posibilidad). Se conoce bien el cálculo del valor E en la técnica. Además de los valores E, las comparaciones también se califican por el porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos idénticos (o aminoácidos) entre las dos secuencias de ácidos nucleicos comparadas (o polipéptido) sobre una longitud particular. En el caso de familias grandes, se puede utilizar ClustalW, seguido por un árbol de unión vecino, para ayudar a visualizar la agrupación de los genes relacionados e identificar ortólogos y parálogos.

Las variantes de ácidos nucleicos también pueden ser útiles en la práctica de los métodos de la invención. Ejemplos 60 de dichas variantes incluyen ácidos nucleicos que codifican homólogos y derivados de cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1, los términos "homólogo" y "derivado" son como se define aquí. También útiles en los métodos de la invención son los ácidos nucleicos que codifican homólogos y derivados de 65 ortólogos o parálogos de una cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1. Los homólogos y derivados útiles en los métodos de la presente invención tienen sustancialmente la misma actividad biológica y funcional como la proteína no modificada de la que se derivan.

Otras variantes de ácido nucleico útiles en la práctica de los métodos de la invención incluyen porciones de ácidos nucleicos que codifican polipéptidos LBD, ácidos nucleicos que hibridan a ácidos nucleicos que codifican

polipéptidos LBD, variantes de empalme de ácidos nucleicos que codifican polipéptidos LBD, variantes alélicas de ácidos nucleicos que codifican polipéptidos LBD y variantes de ácidos nucleicos que codifican polipéptidos LBD obtenidos por transposición genética. Los términos secuencia de hibridación, variante de empalme, variante alélica y transposiciones genéticas son como se describen aquí.

5 Ventajosamente, la presente invención proporciona ácido nucleico LBD y secuencias de polipéptidos desconocidos hasta ahora.

De acuerdo con una realización adicional de la presente invención, se proporciona una molécula de ácido nucleico aislada que comprende:

(i) un ácido nucleico representado por la SEQ ID NO: 69;

10 (ii) un ácido nucleico que es complementario a cualquiera de las SEQ ID NOs dadas en (i);

(iii) un ácido nucleico que codifica un polipéptido LBD que tiene en orden creciente de preferencia, por lo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia con la SEQ ID NO: 70;

(iv) un ácido nucleico capaz de hibridar en condiciones rigurosas con uno cualquiera de los ácidos nucleicos dados en (i), (ii) o (iii) anteriormente.

15 De acuerdo con una realización adicional de la presente invención, se proporciona por lo tanto un polipéptido aislado que comprende:

(i) una secuencia de aminoácidos que tiene, en orden creciente de preferencia, por lo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos dada en SEQ ID NO: 70.

20 (ii) derivados de cualquiera de las secuencias de aminoácidos dadas en (i).

Los ácidos nucleicos que codifican polipéptidos LBD no necesitan ser ácidos nucleicos de longitud completa, ya que el desempeño de los métodos de la invención no se basa en el uso de secuencias de ácidos nucleicos de longitud completa. De acuerdo con la presente invención, se proporciona un método para aumentar el rendimiento de semillas y/o biomasa en plantas, que comprende introducir y expresar en una planta una porción de una cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A1 del Ejemplo 1, o una porción de un ácido nucleico que codifica un ortólogo, parólogo u homólogo de cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1.

25 Se puede preparar una porción de un ácido nucleico, por ejemplo, al hacer una o más supresiones en el ácido nucleico. Las porciones se pueden utilizar en forma aislada o pueden estar fusionadas a otras secuencias codificantes (o no codificantes) con el fin de, por ejemplo, producir una proteína que combina diversas actividades. Cuando se fusiona a otras secuencias codificantes, el polipéptido resultante producido luego de traducción puede ser más grande que el previsto para la porción de proteína.

30 Las porciones útiles en los métodos de la invención, codifican un polipéptido LBD como se define aquí, y tienen sustancialmente la misma actividad biológica que las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1. Preferiblemente, la porción es una porción de uno cualquiera de los ácidos nucleicos dados en la Tabla A1 del Ejemplo 1, o es una porción de un ácido nucleico que codifica un ortólogo o parólogo de una cualquiera de las secuencias de aminoácidos dada en la Tabla A1 del Ejemplo 1. Preferiblemente, la porción es de por lo menos 500,

35 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 nucleótidos consecutivos de longitud, los nucleótidos consecutivos son de una cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A1 del Ejemplo 1, o de

40 un ácido nucleico que codifica un ortólogo o parólogo de una cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1. Más preferiblemente, la porción es una porción del ácido nucleico de la SEQ ID NO: 1. Preferiblemente, la porción codifica una secuencia de aminoácidos que cuando se utiliza en la construcción de un

45 árbol filogenético, tal como el representada en la Figura 3 (o como se define por Yang et al., 2006, utilizando el método de HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), se agrupa con el grupo de polipéptidos de dominio LOB clase II que comprenden la secuencia de aminoácidos representada por la SEQ ID NO: 2 más que con cualquier otro grupo.

Otra variante de ácido nucleico útil en los métodos de la invención es un ácido nucleico capaz de hibridar, bajo condiciones de rigurosidad reducidas, preferiblemente bajo condiciones rigurosas, con un ácido nucleico que codifica un polipéptido LBD como se define aquí, o con una porción como se define aquí.

De acuerdo con la presente invención, se proporciona un método para aumentar el rendimiento de semillas y/o biomasa en plantas, que comprende introducir y expresar en una planta un ácido nucleico capaz de hibridar con uno cualquiera de los ácidos nucleicos dados en la Tabla A1 del Ejemplo 1, o que comprende introducir y expresar en una planta de un ácido nucleico capaz de hibridarse a un ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A1 del Ejemplo 1.

5 Las secuencias de hibridación útiles en los métodos de la invención codifican un polipéptido LBD como se define aquí, y tienen sustancialmente la misma actividad biológica que las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1. Preferiblemente, la secuencia de hibridación es capaz de hibridar a uno cualquiera de los ácidos nucleicos dados en la Tabla A1 del Ejemplo 1, o a una porción de cualquiera de estas secuencias, una porción como 10 se define anteriormente, o en la que la secuencia de hibridación es capaz de hibridar a un ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1. Más preferiblemente, la secuencia de hibridación es capaz de hibridar a un ácido nucleico como el representado por la SEQ ID NO: 1 o a una porción del mismo.

15 La secuencia de hibridación puede codificar una secuencia de aminoácidos que cuando se utiliza en la construcción de un árbol filogénico, tal como el representada en la Figura 3 (o tal como se define por Yang et al., 2006, utilizando el método de HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), se agrupa con el grupo de polipéptidos de dominio LOB clase II que comprenden la secuencia de aminoácidos representada por la SEQ ID NO: 2 más que con cualquier otro grupo.

20 Otra variante de ácido nucleico útil en los métodos de la invención es una variante de empalme que codifica un polipéptido LBD como se definió anteriormente, una variante de empalme es como se define aquí.

25 De acuerdo con la presente invención, se proporciona un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en las plantas, que comprende introducir y expresar en una planta de una variante de empalme de una cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A1 del Ejemplo 1, o una variante de empalme de un ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1.

30 Las variantes de empalme preferidas son variantes de empalme de un ácido nucleico representado por la SEQ ID NO: 1, o una variante de empalme de un ácido nucleico que codifica un ortólogo o parálogo de la SEQ ID NO: 2. La secuencia de aminoácidos codificada por la variante de empalme, cuando se utiliza en la construcción de un árbol filogénico, tal como el representado en la Figura 3 (o tal como se define por Yang et al., 2006, utilizando el método de HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), se pueden agrupar con el grupo de polipéptidos de dominio LOB clase II que comprenden la secuencia de aminoácidos representada por la SEQ ID NO: 2 más que con cualquier otro grupo.

35 Otra variante de ácido nucleico útil en la realización de los métodos de la invención es una variante alélica de un ácido nucleico que codifica un polipéptido LBD como se definió anteriormente, una variante alélica es como se define aquí.

40 De acuerdo con la presente invención, se proporciona un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en las plantas, que comprende introducir y expresar en una planta una variante alélica de cualquiera de los ácidos nucleicos dada en la Tabla A1 del Ejemplo 1, o que comprende introducir y expresar en una planta una variante alélica de un ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1.

45 Las variantes alélicas útiles en los métodos de la presente invención tienen sustancialmente la misma actividad biológica que el polipéptido LBD de la SEQ ID NO: 2 y cualquiera de los aminoácidos descritos en la Tabla A1 del Ejemplo 1. Las variantes alélicas existen en la naturaleza, y abarcadas dentro de los métodos de la presente invención es el uso de estos alelos naturales. Preferiblemente, la variante alélica es una variante alélica de la SEQ ID NO: 1 o una variante alélica de un ácido nucleico que codifica un ortólogo o parálogo de la SEQ ID NO: 2. La secuencia de aminoácidos codificada por la variante alélica, cuando se utiliza en la construcción de un árbol filogénico, tal como se representa en la Figura 3 (o tal como se define por Yang et al., 2006, utilizando el método de HKY (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), se pueden agrupar con el grupo de polipéptidos de dominio LOB clase II que comprenden la secuencia de aminoácidos representada por la SEQ ID NO: 2 más que con cualquier otro grupo.

50 También se pueden utilizar transposiciones genéticas o evolución dirigida para generar variantes de ácidos nucleicos que codifican polipéptidos LBD como se definió anteriormente; el término "transposiciones genéticas", tal como se define aquí.

De acuerdo con la presente invención, se proporciona un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en las plantas, que comprende introducir y expresar en una planta de una variante de una cualquiera de las secuencias de ácidos nucleicos dada en la Tabla A1 del Ejemplo 1, o que comprende introducir y expresar en una planta de una variante de un ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de aminoácidos dadas en la Tabla A1 del Ejemplo 1, la variante de ácido nucleico se obtiene por transposición genética.

5 La secuencia de aminoácidos codificada por el ácido nucleico variante obtenido por transposición genética, cuando se utiliza en la construcción de un árbol filogenético, tal como se representa en la Figura 3 (o tal como se define por Yang et al., 2006, utilizando la HKY método (Hasegawa et al., J. Mol. Evol. 22, 160-174, 1985)), se puede agrupar con el grupo de polipéptidos de dominio LOB clase II que comprenden la secuencia de aminoácidos representada por la SEQ ID NO: 2 en lugar de con cualquier otro grupo.

10 Adicionalmente, las variantes de ácidos nucleicos también se pueden obtener a través de mutagénesis dirigida al sitio. Diversos métodos están disponibles para alcanzar mutagénesis dirigida al sitio, los métodos con base en PCR son más comunes (Current Protocols in Molecular Biology. Wiley Eds.).

15 15 Los ácidos nucleicos que codifican polipéptidos LBD se pueden derivar de cualquier fuente natural o artificial. El ácido nucleico puede ser modificado a partir de su forma nativa en composición y/o ambiente genómico a través de manipulación humana deliberada. Preferiblemente, el ácido nucleico que codifica el polipéptido LBD es de una planta, más preferiblemente de una planta dicotiledónea, más preferiblemente de la familia Brassicaceae, más preferiblemente el ácido nucleico es de *Arabidopsis thaliana*.

20 20 El desempeño de los métodos de la invención proporciona plantas que tienen rasgos relacionados con el rendimiento mejorados. En particular, el desempeño de los métodos de la invención da plantas que tienen mayor rendimiento, especialmente aumento de rendimiento de semilla en relación con plantas de control. Los términos "rendimiento" y "rendimiento de semilla" se describen en más detalle en la sección "definiciones" aquí.

25 25 La referencia aquí a un aumento en la biomasa (peso) de una o más partes de una planta, puede incluir partes por encima de la tierra (cosechables) y/o partes bajo tierra(cosechable). En particular, dichas partes cosechables sobre el suelo son la biomasa en brotes y semillas, y el desempeño de los métodos de la invención resulta en plantas que tienen aumento de biomasa de brotes y rendimiento de semilla con relación a la biomasa de brotes y el rendimiento de semillas de plantas de control.

30 30 Tomando el maíz como un ejemplo, un aumento de rendimiento de semilla o aumento de biomasa se puede manifestar como uno o más de los siguientes: aumento en el número de plantas establecidas por hectárea o por acre, un aumento en el número de mazorcas por planta, un aumento en el número de surcos, número de granos por surco, peso del grano, peso de mil granos, longitud /diámetro de la mazorca, aumento del índice de llenado de semilla (que es el número de semillas cargadas, dividido por el número total de semillas y multiplicado por 100), entre otros. Tomando el arroz como un ejemplo, un rendimiento de semilla o aumento de la biomasa se puede manifestar como un aumento en uno o más de los siguientes: número de plantas por hectárea o por acre, número de panículas por planta, número de espiguillas por panícula, número de flores (floretes) por panícula (que se expresa como una relación entre el número de semillas cargadas más el número de panículas primarias), aumento en el índice de llenado de semilla (que es el número de semillas cargadas, dividido por el número total de semillas y multiplicado por 100), el aumento de peso de mil granos, entre otros.

35 40 La presente invención proporciona un método para aumentar el rendimiento de semilla y de biomasa, especialmente biomasa de brote con relación a las plantas de control, cuyo método comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD como se define aquí.

45 45 Ya que las plantas transgénicas de acuerdo con la presente invención tienen aumento de rendimiento de semillas o aumento de biomasa, es probable que estas plantas exhiban un mayor índice de crecimiento (por lo menos durante parte de su ciclo de vida), con relación al índice de crecimiento de las plantas de control en una etapa correspondiente en su ciclo de vida.

50 55 El aumento de índice de crecimiento puede ser específico para una o más partes de una planta (que incluyen semillas), o puede ser a través de sustancialmente toda la planta. Las plantas que tienen un aumento de índice de crecimiento pueden tener un ciclo de vida más corto. Se puede tomar que el ciclo de vida de una planta significa el tiempo necesario para crecer desde una semilla madura seca hasta la etapa donde la planta ha producido semillas maduras secas, similar al material de partida. Este ciclo de vida puede estar influenciado por factores tales como vigor temprano, índice de crecimiento, índice de verdor, tiempo de floración y velocidad de maduración de la semilla. El aumento del índice de crecimiento puede tener lugar en una o más etapas en el ciclo de vida de una planta o sustancialmente durante todo el ciclo de vida de la planta. El aumento de índice de crecimiento durante las primeras etapas en el ciclo de vida de una planta puede reflejar vigor mejorado. El aumento en el índice de crecimiento puede

- alterar el ciclo de cosecha de una planta permitiendo que las plantas sean sembradas más tarde y/o cosechadas más pronto de lo que sería posible (un efecto similar se puede obtener con tiempo de floración temprano). Si se incrementa lo suficiente el índice de crecimiento el índice de crecimiento, puede permitir la siembra de más semillas de la misma especie de planta (por ejemplo la siembra y cosecha de plantas de arroz seguido por siembra y cosecha de plantas de arroz adicionales todo dentro de un período de crecimiento convencional). Del mismo modo, si se incrementa lo suficiente el índice de crecimiento, puede permitir la siembra de más semillas de diferentes especies de plantas (por ejemplo la siembra y cosecha de plantas de maíz seguido, por ejemplo, por la siembra y cosecha opcional de soja, papas o cualquier otra planta adecuada). También puede ser posible temporadas adicionales de cosecha del mismo rizoma en el caso de algunas plantas de cultivo. La alteración del ciclo de cosecha de una planta
- 5 puede llevar a un aumento de la producción anual de biomasa por hectárea (debido a un aumento en el número de veces (digamos en un año) que cualquier planta en particular se puede sembrar y cosechar). Un aumento en el índice de crecimiento puede permitir también el cultivo de plantas transgénicas en un área geográfica más amplia que sus contrapartes de tipo silvestre, ya que las limitaciones territoriales para el crecimiento de un cultivo son a menudo determinadas por condiciones ambientales adversas ya sea en el momento de la siembra (principios de 10 temporada) o en el momento de la cosecha (finales de temporada). Dichas condiciones adversas se pueden evitar si se acorta el ciclo de cosecha. El índice de crecimiento se puede determinar al derivar diferentes parámetros a partir de curvas de crecimiento, tales parámetros pueden ser: T - Mid (el tiempo tomado para que las plantas alcancen 15 50% de su tamaño máximo) y T - 90 (tiempo tomado para que plantas alcancen el 90% de su tamaño máximo), entre otros.
- 15 20 El desempeño de los métodos de la invención puede dar a las plantas que tienen un mayor índice de crecimiento relativo a las plantas de control. Se describe aquí un método para aumentar el índice de crecimiento de las plantas, cuyo método comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD como se define aquí. En particular, se observa un aumento en el índice de crecimiento durante las primeras etapas de crecimiento de la planta (vigor temprano).
- 25 30 35 40 45 50 55 60 Un aumento en el rendimiento de semilla y/o biomasa y/o índice de crecimiento se produce si la planta está bajo condiciones de no estrés o si la planta está expuesta a diversos tipos de estrés en comparación con las plantas de control. Las plantas responden normalmente a la exposición al estrés al crecer más lentamente. En condiciones de estrés severas, la planta puede incluso dejar de crecer por completo. El estrés leve por otro lado se define aquí como cualquier estrés al que está expuesta una planta que no resulta en que la planta deje de crecer completamente sin la capacidad para reanudar el crecimiento. El estrés leve en el sentido de la invención conduce a una reducción en el crecimiento de las plantas estresadas de menos de 40%, 35% o 30%, preferiblemente menos de 25%, 20% o 15%, más preferiblemente menos de 14%, 13%, 12%, 11% o 10% o menos en comparación con la planta de control en condiciones de no estrés. Debido a los avances en las prácticas agrícolas (riego, fertilización, tratamientos fitosanitarios), no se encuentran a menudo estrés severo en las plantas sembradas. Como consecuencia, el crecimiento comprometido inducido por estrés leve es a menudo una característica indeseable para la agricultura. El estrés leve es el estrés cotidiano biótico y/o abiótico (ambientales) al que está expuesta una planta. Los estreses abióticos pueden ser debido a la sequía o exceso de agua, el estrés anaeróbico, estrés salino, toxicidad química, el estrés oxidativo y las temperaturas calientes, frías o de congelación. El estrés abiótico puede ser un estrés osmótico causado por un estrés hídrico (particularmente debido a la sequía), estrés salino, estrés oxidativo o un estrés iónico. Los estreses bióticos son normalmente aquellos estreses provocados por patógenos, tales como bacterias, virus, hongos e insectos.
- En particular, se pueden realizar los métodos de la presente invención en condiciones de no estrés o bajo condiciones de sequía moderadas para obtener plantas que tienen mayor rendimiento con relación a las plantas de control. Como se reportó en Wang et al. (Plant (2003) 218: 1-14), el estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan adversamente el crecimiento de plantas y la productividad. Se sabe que la sequía, salinidad, temperaturas extremas y el estrés oxidativo están interconectados y pueden inducir el crecimiento y el daño celular a través de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755-1767) describe un grado particularmente elevado de "comunicación entre cascadas" entre estrés de sequía y estrés de alta salinidad. Por ejemplo, la sequía y/o salinización se manifiestan principalmente como estrés osmótico, dando como resultado la perturbación de la homeostasis y la distribución de iones en la célula. El estrés oxidativo, que acompaña con frecuencia alta o baja temperatura, salinidad o estrés por sequía, puede provocar la desnaturalización de las proteínas estructurales y funcionales. Como consecuencia, estos diversos estreses ambientales a menudo activan rutas de señalización celular similares y respuestas celulares, tales como la producción de proteínas de estrés, sobre regulación de antioxidantes, acumulación de solutos compatibles y detención del crecimiento. El término "condiciones sin estrés" como se utiliza aquí son las condiciones ambientales que permitan un crecimiento óptimo de las plantas. Las personas expertas en la técnica son conscientes de las condiciones normales del suelo y las condiciones climáticas para una ubicación dada. Las plantas con las condiciones óptimas de crecimiento, (cultivadas en condiciones de no estrés) generalmente producen en orden creciente de preferencia por lo menos 90%, 87%, 85%, 83%, 80%, 77% o 75% de la producción promedio de dicha planta en un ambiente dado. Se puede calcular la producción promedio sobre la cosecha y/o las bases de temporada. Los expertos en la técnica son conscientes de las producciones de rendimiento promedio de un cultivo.

El desempeño de los métodos de la invención da plantas cultivadas en condiciones de no estrés o bajo condiciones de sequía leves aumenta el rendimiento en relación con las plantas de control cultivadas en condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se proporciona un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en cultivos de plantas en condiciones no de estrés o bajo condiciones de sequía leves, cuyo método comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD.

5 El desempeño de los métodos de la invención da plantas que crecen bajo condiciones de deficiencia de nutrientes, particularmente bajo condiciones de deficiencia de nitrógeno invención, el aumento de vigor temprano y el aumento de rendimiento en relación con las plantas de control cultivadas en condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se proporciona un método para aumentar el vigor inicial y el rendimiento en las plantas cultivadas en condiciones de deficiencia de nutrientes, cuyo método comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD. La deficiencia de nutrientes puede resultar de una falta o exceso de nutrientes tales como nitrógeno, fosfatos y otros compuestos que contienen fósforo, potasio, calcio, cadmio, magnesio, manganeso, hierro y boro, entre otros.

10 15 La presente invención abarca plantas o partes de las mismas (que incluyen semillas) que se pueden obtener por los métodos de acuerdo con la presente invención. Las plantas o partes de las mismas comprenden un transgén de ácido nucleico que codifica un polipéptido LBD como se definió anteriormente.

20 25 La invención también proporciona construcciones genéticas y vectores para facilitar la introducción y/o la expresión en plantas de ácidos nucleicos que codifican los polipéptidos LBD. Las construcciones génicas se pueden insertar en vectores, que pueden estar disponibles comercialmente, adecuados para transformarse en plantas y adecuados para la expresión del gen de interés en las células transformadas. La invención también proporciona el uso de una construcción génica como se define aquí en los métodos de la invención.

Más específicamente, la presente invención proporciona una construcción que comprende:

- (a) un ácido nucleico que codifica un polipéptido LBD como se definió anteriormente;
- 25 (b) una o más secuencias de control capaces de conducir la expresión de la secuencia de ácidos nucleicos de (a); y opcionalmente
- (c) una secuencia de terminación de transcripción.

30 Preferiblemente, el ácido nucleico que codifica un polipéptido LBD es como se definió anteriormente. El término "secuencia de control" y "secuencia de terminación" son como se definen aquí.

35 35 Las plantas se transforman con un vector que comprende cualquiera de los ácidos nucleicos descritos anteriormente. El experto en la técnica es muy consciente de los elementos genéticos que deben estar presentes en el vector para transformar, seleccionar y propagar con éxito células anfitrionas que contienen la secuencia de interés. La secuencia de interés se liga operativamente a una o más secuencias de control (por lo menos a un promotor).

40 45 Ventajosamente, cualquier tipo de promotor, ya sea natural o sintético, se puede utilizar para conducir la expresión de la secuencia de ácidos nucleicos. Un promotor constitutivo es particularmente útil en los métodos. Véase la sección "Definiciones" aquí para las definiciones de los distintos tipos de promotores.

40 Debe ser claro que la aplicabilidad de la presente invención no está restringida al ácido nucleico que codifica el polipéptido LBD representado por la SEQ ID NO: 1, ni la aplicabilidad de la invención se restringe a la expresión de un ácido nucleico que codifica el polipéptido LBD cuando es accionado por un promotor constitutivo.

45 50 El promotor constitutivo es preferiblemente un promotor GOS2, preferiblemente un promotor GOS2 de arroz. Adicionalmente, preferiblemente el promotor constitutivo se representa por una secuencia de ácidos nucleicos sustancialmente similar a la SEQ ID NO: 10, más preferiblemente el promotor constitutivo es como el representado por la SEQ ID NO: 10. Véase sección "Definiciones" aquí para otros ejemplos de promotores constitutivos.

50 Opcionalmente, se pueden utilizar una o más secuencias de terminación en la construcción introducida en una planta. Los elementos reguladores adicionales pueden incluir potenciadores de transcripción y traducción. Aquellos expertos en la técnica serán conscientes de secuencias terminadoras y potenciadores que pueden ser adecuados para su uso en la realización de la invención. Una secuencia de intrón también se puede agregar a la región 5' no traducida (UTR) o en la secuencia de codificación para aumentar la cantidad del mensaje de madurez que se acumula en el citosol, como se describe en la sección de definiciones. Otras secuencias de control (además de secuencias de promotor, potenciador, silenciador, intrones, regiones 3'UTR y/o 5'UTR) pueden ser elementos de

estabilización de proteína y/o ARN. Dichas secuencias serían conocidas o se pueden obtener fácilmente por un experto en la técnica.

Las construcciones genéticas de la invención pueden incluir además un origen de secuencia de replicación que se requiere para el mantenimiento y/o replicación en un tipo celular específico. Un ejemplo es cuando se requiere una construcción genética que se va a mantener en una célula bacteriana como un elemento genético episomal (por ejemplo molécula de plásmido o cósmido). Los orígenes preferidos de replicación incluyen, pero no se limitan a, el f1-ori y colE1.

Para la detección de la transferencia exitosa de las secuencias de ácidos nucleicos tal como se utiliza en los métodos de la invención y/o selección de las plantas transgénicas que comprenden estos ácidos nucleicos, es ventajoso el uso de genes marcadores (o genes indicadores). Por lo tanto, la construcción genética puede comprender opcionalmente un gen marcador seleccionable. Los marcadores seleccionables se describen en más detalle en la sección "definiciones" aquí.

Se sabe que luego de la integración estable o transitoria de ácidos nucleicos en células vegetales, sólo una minoría de las células lleva el ADN extraño y, si se desea, lo integra en su genoma, dependiendo del vector de expresión utilizado y de la técnica de transfección utilizada. Para identificar y seleccionar estos integrantes, un gen que codifica un marcador seleccionable (tal como los descritos anteriormente) se introduce generalmente en las células anfitrionas junto con el gen de interés. Estos marcadores se pueden utilizar, por ejemplo en mutantes en los que estos genes no son funcionales, por ejemplo, supresión por métodos convencionales. Adicionalmente, las moléculas de ácido nucleico que codifican un marcador seleccionable se pueden introducir en una célula anfitriona en el mismo vector que comprende la secuencia que codifica los polipéptidos de la invención o se utilizan en los métodos de la invención, o también en un vector separado. Se pueden identificar las células que han sido transfectadas de manera estable con el ácido nucleico introducido, por ejemplo, mediante selección (por ejemplo, las células que han integrado el marcador seleccionable sobreviven mientras que las otras células mueren). Los genes marcadores pueden ser retirados o extirpados de la célula transgénica, una vez que ya no son necesarios. Las técnicas para eliminación del gen marcador se conocen en el arte, las técnicas útiles se describen anteriormente la sección de definiciones.

La invención también proporciona un método para la producción de plantas transgénicas que tienen aumento de rendimiento de semilla y/o aumento de biomasa con relación a las plantas de control, que comprende la introducción y expresión en una planta de cualquier ácido nucleico que codifica un polipéptido LBD como se definió anteriormente.

Más específicamente, la presente invención proporciona un método para la producción de plantas transgénicas que tienen aumento de rendimiento de semilla y/o aumento de biomasa particularmente aumento de biomasa de brotes, cuyo método comprende:

- (i) introducir y expresar en una célula de panta o planta un ácido nucleico que codifica el polipéptido LBD, y
- 35 (ii) cultivar la célula vegetal en condiciones que promueven el crecimiento y desarrollo de plantas.

El ácido nucleico de (i) puede ser cualquiera de los ácidos nucleicos capaces de codificar un polipéptido LBD como se define aquí.

Se puede introducir el ácido nucleico directamente en una célula de planta o en la planta misma (que incluye la introducción en un tejido, órgano o cualquier otra parte de una planta). De acuerdo con una característica preferida de la presente invención, el ácido nucleico se introduce preferiblemente en una planta mediante transformación. El término "transformación" se describe con más detalle en la sección "definiciones" aquí.

Las células de las plantas modificadas genéticamente se pueden regenerar a través de todos los métodos con los que el experto está familiarizado. Los métodos adecuados se pueden encontrar en las publicaciones anteriormente mencionadas por SD Kung and R. Wu, Potrykus o Höfgen and Willmitzer.

45 De manera general después de la transformación, las células de planta o agrupaciones de células se seleccionan para la presencia de uno o más marcadores que son 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 entera. Para seleccionar las plantas transformadas, el material de planta obtenido en la transformación, como una regla, se somete 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 inicial de crecimiento, se someten a una selección adecuada por rociado. Una posibilidad adicional consiste en hacer crecer semillas, si es apropiado, después de esterilización, en placas de agar utilizando un agente de selección adecuado, de modo que sólo las semillas transformadas pueden crecer en las plantas.

Alternativamente, las plantas transformadas se examinan para detectar la presencia de un marcador seleccionable, tal como los descritos anteriormente.

Después de la transferencia de ADN y regeneración, las plantas putativamente transformadas también se pueden evaluar, por ejemplo utilizando análisis Southern, para la presencia del gen de interés, número de copia y/o organización genómica. Alternativamente o adicionalmente, los niveles de expresión del ADN recientemente introducido se pueden monitorear utilizando análisis Northern y/o Western, siendo ambas técnicas bien conocidas por las personas medianamente versadas en la técnica.

Las plantas transformadas generadas se pueden propagar mediante una variedad de medios, tales como mediante propagación clonal o técnicas clásicas de reproducción. Por ejemplo, una primera generación (o T1) de planta transformada se puede autopolinizar y los transformantes homocigotos de segunda generación (o T2) seleccionados, y las plantas T2 entonces se pueden propagar aún más a través de técnicas clásicas de reproducción. Los organismos transformados generados pueden tomar una variedad de formas. Por ejemplo, pueden ser quimeras de células transformadas y células no transformadas; los 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 a un descendiente no transformado).

La presente invención se extiende claramente a cualquier célula de planta o planta producida por cualquiera de los métodos descritos aquí, y para todas las partes de la planta y propágulos de las mismas. La presente invención se extiende adicionalmente para abarcar la progénesis de una célula, tejido, órgano o planta entera transformada o transfectada primaria que ha sido producida por cualquiera de los métodos anteriormente mencionados, siendo el único requisito que la progénesis exhiba la misma característica genotípica y/o fenotípica como aquellas producidas por los progenitores en los métodos de acuerdo con la invención.

La invención también incluye células anfitrionas que contienen un ácido nucleico aislado que codifica un polipéptido LBD como se definió anteriormente. Las células anfitrionas preferidas de acuerdo con la invención son células de planta. Las plantas anfitrionas para los ácidos nucleicos o los vectores utilizados en el método de acuerdo con la invención, el casete de expresión o construcción o vector son, en principio, de forma ventajosa todas las plantas, que son capaces de sintetizar los polipéptidos utilizados en el método de la invención.

Los métodos de la invención son ventajosamente aplicables a cualquier planta. Las plantas que son particularmente útiles en los métodos de la invención incluyen todas las plantas que pertenecen a la superfamilia de Viridiplantae, en particular plantas monocotiledóneas y dicotiledóneas, que incluyen forraje o leguminosas forrajeras, plantas ornamentales, cultivos de alimentos, árboles o arbustos. De acuerdo con una realización preferida de la presente invención, la planta es una planta de cultivo. Ejemplos de plantas de cultivo incluyen soja, girasol, canola, alfalfa, colza, algodón, tomate, papa y tabaco. Más preferiblemente, la planta es una planta monocotiledónea. Ejemplos de plantas monocotiledóneas son la caña de azúcar. Más preferiblemente, la planta es un cereal. Ejemplos de cereales incluyen arroz, maíz, trigo, cebada, mijo, centeno, triticale, sorgo y avena.

La invención también se extiende a partes cosechables de una planta tales como, pero sin limitarse a semillas, hojas, frutos, flores, tallos, rizomas, tubérculos y bulbos. La invención se relaciona adicionalmente con productos derivados, derivados preferiblemente de forma directa, a partir de una parte cosechable de dicha planta, tales como gránulos o polvos secos, aceite, grasa y ácidos grasos, almidón o proteínas.

Los métodos para aumentar la expresión de ácidos nucleicos o genes o productos genéticos, están bien documentados en la técnica y se proporcionan ejemplos en la sección de definiciones.

De acuerdo con los métodos de la invención, aumentar la expresión de un ácido nucleico que codifica un polipéptido LBD es al introducir y expresar en una planta de un ácido nucleico que codifica un polipéptido LBD. Se describen aquí otras técnicas bien conocidas, que incluyen pero no se limitan a T-ADN de marcado de activación, TILLING, recombinación homóloga. Una descripción de estas técnicas se proporciona en la sección definiciones.

La presente invención también abarca el uso de ácidos nucleicos que codifican polipéptidos LBD como se describe aquí y el uso de estos polipéptidos LBD en la mejora de cualquiera de los rasgos antes mencionados relacionados con el rendimiento en plantas.

Los ácidos nucleicos que codifican el polipéptido LBD descrito aquí, o los polipéptidos LBD en sí mismos, pueden encontrar uso en los programas de reproducción en los que se identifica un marcador de ADN que puede estar genéticamente ligado a un gen que codifica un polipéptido LBD. Los ácidos nucleicos/genes, o los polipéptidos LBD en sí mismos se pueden utilizar para definir un marcador molecular. Este ADN o marcador de proteína se puede utilizar luego en programas de reproducción para seleccionar plantas que tienen rasgos relacionados con el rendimiento mejorados como se definió anteriormente en los métodos de la invención.

Las variantes alélicas de un gen/ácido nucleico que codifica el polipéptido LBD también pueden encontrar uso en los programas de selección asistida por marcadores. Dichos programas de reproducción algunas veces requieren la introducción de una variación alélica mediante tratamiento mutagénico de las plantas, utilizando por ejemplo mutagénesis EMS; alternativamente, el programa puede iniciar con una colección de variantes alélicas denominadas de origen "natural" provocadas involuntariamente. Tiene lugar entonces la identificación de variantes alélicas, por ejemplo, mediante PCR. Esto se sigue por una etapa para la selección de variantes alélicas superiores de la secuencia en cuestión y que dan un mayor rendimiento. La selección se lleva a cabo normalmente al monitorear el desempeño del crecimiento de las plantas que contienen diferentes variantes alélicas de la secuencia en cuestión. Se puede monitorear el crecimiento en un invernadero o en el campo. Etapas opcionales adicionales incluyen plantas de cruce en las que la variante alélica superior, se identifica con otra planta. Esto podría ser utilizado, por ejemplo, para hacer una combinación de características fenotípicas interesantes.

Los ácidos nucleicos que codifican polipéptidos LBD también se pueden utilizar como sondas para mapear genéticamente y físicamente los genes que son una parte de, y como marcadores para los rasgos ligados a esos genes. Esta información puede ser útil en la reproducción de la planta para desarrollar estirpes con fenotipos deseados. Dicho uso de ácidos nucleicos que codifican el polipéptido LBD requiere sólo una secuencia de ácidos nucleicos de por lo menos 15 nucleótidos de longitud. Los ácidos nucleicos que codifican el polipéptido LBD se pueden utilizar como marcadores de polimorfismo de longitud de fragmentos de restricción (RFLP). Se pueden sondear la transferencia Southern (Sambrook J, Fritsch EF y Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de ADN genómico de planta digerido con restricción con los ácidos nucleicos que codifican LBD. Los 15 patrones de bandas resultantes entonces se pueden someter a análisis genéticos utilizando programas de ordenador tales como MapMaker (Lander et al (1987) Genomics 1: 174-181) con el fin de construir un mapa genético. Adicionalmente, los ácidos nucleicos se pueden utilizar para sondear transferencias Southern que 20 contienen ADN genómicos tratados con endonucleasa de restricción de un grupo de individuos que representan a los progenitores y la progénesis de un cruce genético definido. La segregación de los polimorfismos de ADN se 25 observa y se utiliza para calcular la posición del ácido nucleico que codifica el polipéptido LBD en el mapa genético previamente obtenido utilizando esta población Botstein et al. (1980) Am. J. Hum. Genet. 32:314-331).

La producción y uso de sondas derivadas de gen de planta para uso en mapeo genético se describe en Bematzky y Tanksley (1986) Plant Mol. Biol. Reporter 4: 37-41. Numerosas publicaciones describen el mapeo genético de clones de cADN específicos utilizando la metodología representada anteriormente o variaciones de la misma. Por ejemplo, 30 poblaciones de entrecruzamiento F2, poblaciones de retrocruzamiento, poblaciones apareadas aleatoriamente, estirpes isogénicas cercanas, y otros grupos de individuos se pueden utilizar para la proyección. Dichas metodologías son bien conocidas por los expertos en la técnica.

También se pueden utilizar sondas de ácidos nucleicos para el mapeo físico (es decir, colocación de secuencias sobre mapas físicos; véase Hoheisel et al en: Analysis: A Practical Guide, Academic press 1996, pp. 319-346, y 35 referencias citadas en la misma).

En otra realización, las sondas de ácido nucleico se pueden utilizar en la fluorescencia directa de hibridación in situ (FISH) de mapeo (Trask (1991) Trends Genet. 7:149-154). Aunque los métodos actuales de mapeo FISH favorecen el uso de grandes clones (varios kb a varios cientos kb; véase Laan et al. (1995) Genome Res. 5:13-20), las mejoras en la sensibilidad pueden permitir la realización de mapeo FISH utilizando sondas más cortas.

40 Una variedad de métodos basados en la amplificación de ácido nucleico para mapeo genético y físico se puede llevar a cabo utilizando los ácidos nucleicos. Los ejemplos incluyen amplificación específica a alelo (Kazazian (1989) J. Lab. Clin. Med 11: 95- 96), Polymorphism of PCR- amplified fragments (CAPS; Sheffield et al. (1993) Genomics 16: 325- 332), ligado específico a alelo (Landegren et al. (1988) Science 241: 1077- 1080), reacciones de extensión de nucleótidos (Sokolov (1990) Nucleic Acid Res. 18: 3671.) mapeo de híbridos de radiación (Walter et al. (1997) 45 Nat. Genet. 7: 22-28) y mapeo Happy (Dear and Cook (1989) Nucleic Acid Res. 17: 6.795 a 6807). Para estos métodos, se utiliza la secuencia de un ácido nucleico para diseñar y producir pares de cebadores para su uso en la reacción de amplificación o en reacciones de extensión de cebador. El diseño de dichos cebadores es bien conocido por aquellos expertos en la técnica. En los métodos que emplean el mapeo genético basado en PCR, puede ser necesario identificar las diferencias de secuencia de ADN entre los progenitores del mapeo cruzado en la región 50 correspondiente a la secuencia de ácidos nucleicos instantánea. Esto, sin embargo, generalmente no es necesario para métodos de mapeo.

55 Los métodos de acuerdo con la presente invención resultan en plantas que tienen aumento de rendimiento de semilla y/o aumento de biomasa. Estos rasgos se pueden combinar también con otros rasgos económicamente ventajosos, tales como rasgos adicionales que mejoran el rendimiento, la tolerancia a otros tipos de estrés abiótico y biótico, rasgos de modificación de diversas características arquitectónicas y/o características bioquímicas y/o fisiológicas.

Descripción de las figuras

La presente invención se describirá ahora con referencia a las siguientes figuras en las que:

5 Dominio LOB que comprende proteína (LOB: Límites de Órgano Lateral)

La Figura 1 representa la secuencia de la SEQ ID NO: 2, con el dominio DUF260 mostrado en negrita, los motivos conservados 1, 2 y 3 están subrayados y los residuos de Cys conservados (motivo de la SEQ ID NO: 9) se muestran en cursiva.

10 La Figura 2 representa una alineación múltiple de secuencias de diversas proteínas LBD útiles en los métodos de la presente invención.

La Figura 3 muestra un árbol filogénico de proteínas LBD clase II y clase I. La secuencia de SEQ ID NO: 2 se representa por AtLBD37.

15 La figura 4 representa el vector binario para aumento de expresión en *Oryza sativa* de un ácido nucleico que codifica el LBD bajo el control de un promotor GOS2 de arroz (pGOS2)

20 La figura 5 detalla ejemplos de secuencias de LBD útiles en la realización de los métodos de acuerdo con la presente invención.

### Ejemplos

25 La presente invención se describirá ahora con referencia a los siguientes ejemplos, que son solamente por vía de ilustración. Los siguientes ejemplos no están destinados a definir completamente o limitar de otro modo el alcance de la invención.

30 Manipulación del ADN: a menos que se indique lo contrario, se realizan técnicas de ADN recombinante de acuerdo con protocolos estándar descritos en Sambrook (2001) Molecular Cloning: A Laboratory Manual, tercera Edición Cold Spring Harbor Laboratory Press, CSH, New York o en los Volúmenes 1 y 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Los materiales y métodos para trabajo molecular de plantas estándar se describen en Plant Molecular Biology Labfax (1993) por el R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (Reino Unido) y Blackwell Scientific Publications (Reino Unido).

35 Dominio LOB que comprende proteína (LOB: Límites de Órgano Lateral)

40 **Ejemplo 1:** Identificación de secuencias relacionadas con la secuencia de ácidos nucleicos utilizada en los métodos de la invención

45 Las secuencias (cADN de longitud completa, EST o genómicas) relacionadas con la secuencia de ácidos nucleicos utilizada en los métodos de la presente invención se identificaron entre aquellas mantenidas en la base de datos de nucleótidos Entrez en el Centro Nacional de Información sobre Biotecnología (NCBI) utilizando herramientas de búsqueda de base de datos de secuencia, tales como el Basic Local Alignment Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; y Altschul et al. (1997) Nucleic Acids Res. 25:3389-3402). El programa se utiliza para encontrar regiones de similitud local entre las secuencias al comparar el ácido nucleico o secuencias de polipéptidos con bases de datos de secuencias y al calcular el significado estadístico de emparejamientos. Por ejemplo, el polipéptido codificado por el ácido nucleico utilizado en la presente invención se utiliza para el algoritmo TBLASTN, 50 con la configuración predeterminada y el filtro para ignorar secuencias de baja complejidad desactivadas. El resultado del análisis se observa mediante comparación por pares, y se clasifica de acuerdo con la probabilidad (valor E), donde la puntuación refleja la probabilidad de que una se produzca una alineación particular aleatoria (cuanto menor sea el valor E, es más significativo el cierto). Además de los valores E, las comparaciones también se anotaron por el porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos idénticos (o aminoácidos) entre las dos secuencias de ácidos nucleicos comparadas (o polipéptido) de más de una longitud particular.

55 La Tabla A1 proporciona una lista de secuencias de ácidos nucleicos relacionadas con la secuencia de ácidos nucleicos LBD utilizado en los métodos de la presente invención.

Tabla A1: Ejemplos de polipéptidos LBD: agregar secuencias de ADN de arroz, maíz, trigo, canola, patata, soja, *Arabidopsis*

Fuente de Planta*	Ácido nucleico SEQ ID NO:	Proteína SEQ ID NO:
Arabidopsis thaliana LBD proteína	1	2
Os01g03890, <i>Oryza sativa</i>	59	11
Os01g32770, <i>Oryza sativa</i>	60	12
Os03g33090, <i>Oryza sativa</i>	61	13
Os03g41330, <i>Oryza sativa</i>	62	14
Os07g40000, <i>Oryza sativa</i>	63	15
TC9404, <i>Nicotiana benthamiana</i>		16
TC227562, <i>Glycine max</i>		17
TC216138, <i>Glycine max</i>		18
TC147776, <i>Hordeum vulgare</i>		19
TC104758, <i>Sorgo bicolor</i>		20
TC18561, <i>Aquilegia sp.</i>		21
TC60668, <i>Vitis vinifera</i>		22
TC15459, <i>Lotus japonicus</i>		23
TC30552, <i>Gossypium hirsutum</i>		24
TC235711, <i>Triticum aestivum</i>	71	25
TC133081, <i>Solanum tuberosum</i>		26
TC107091, <i>Medicago truncatula</i>		27
TC147808, <i>Hordeum vulgare</i>		28
TC55931, <i>Vitis vinifera</i>		29
TC162239, <i>Solanum tuberosum</i>		30
TC69225, <i>Pinus taeda</i>		31
TC67269, <i>Pinus taeda</i>		32
TC220806, <i>Glycine max</i>		33
TC270332, <i>Triticum aestivum</i>		34
TC18329, <i>Aquilegia sp.</i>		35
TC137193, <i>Solanum tuberosum</i>		36

Fuente de Planta*	Ácido nucleico SEQ ID NO:	Proteína SEQ ID NO:
TC133385, Solanum tuberosum		37
TC140088, Solanum tuberosum		38
TC14656, Picea alba		39
TC59178, Pinus taeda		40
TC67974, Pinus taeda		41
TC178827, Lycopersicon esculentum		42
Pt-III.589, Populus tremuloides		43
Pt-V.543, Populus tremuloides		44
Pt-XIV.94, Populus tremuloides		45
Pt-II105, Populus tremuloides		46
Pt-123.86, Populus tremuloides		47
Pt-X180, Populus tremuloides		48
Pt-XII.481, Populus tremuloides		49
DQ787782, Caragana korshinskii		50
AAP37970, Brassica napus		51
ABE82505, Medicago truncatula	72	52
ABE78739, Medicago truncatula		53
Q9SN23, Arabidopsis thaliana	64	54
Q9SZE8, Arabidopsis thaliana	65	55
Q9ZW96, Arabidopsis thaliana	66	56
Q9M886, Arabidopsis thaliana	67	57
Q9CA30, Arabidopsis thaliana	68	58
Ls_LBD, Linum usitatissimum	69	70

\*: números de acceso de base de datos GenBank o SwissProt se proporcionan cuando estén disponibles, los códigos de TC son de TIGR.

#### Ejemplo 2: Alineación de secuencias de polipéptidos LBD

La alineación de secuencias de polipéptidos se realiza utilizando el programa AlignX del Vector NTI (Invitrogen) que se basa en el popular algoritmo Clustal W de alineación progresiva (Thompson et al (1997) Nucleic Acids Res. 25:4876-4882; Chenna. et al. (2003). Nucleic Acids Res 31:3497-3500). Los valores por defecto son para la penalización de espacio abierto de 10, por la penalización de extensión de espacio de 0,1 y la matriz de peso seleccionada es Blosum 62. Se realiza edición manual menor para optimizar la alineación. La conservación de la

secuencia entre el polipéptidos LBD es esencialmente en el dominio de terminal N DUF260 de los polipéptidos, la región de terminal C por lo general es más variable en longitud de secuencia y composición. Los polipéptidos LBD se alinean en la Figura 2.

5 Un árbol filogénico de polipéptidos LBD (Figura 3) se construye utilizando un algoritmo de unión de vecino como se proporciona en el programa AlignX del Vector NTI (Invitrogen). Las secuencias de proteínas LBD clase I utilizadas en la construcción del árbol están disponibles públicamente y se indican con sus números de acceso de GenBank o SwissProt.

**Ejemplo 3:** Cálculo del porcentaje de identidad global entre secuencias de polipéptidos útiles en la realización de los métodos de la invención

10 Se determinan los porcentajes globales de similitud e identidad entre secuencias de polipéptidos de longitud completa útiles en la realización de los métodos de la invención utilizando uno de los métodos disponibles en la técnica, el software MatGAT (Matrix Global Alignment Tool) (BMC - Bioinformáticos. 2003 4: 29 MatGAT: una aplicación que genera matrices de similitud/identidad con proteínas o secuencias de ADN Campanella JJ, Bitincka L, J Smalley; el software ofrecido por Ledion Bitincka). El Software de MatGAT genera matrices de similitud/identidad de secuencias de ADN o proteína sin necesidad de pre-alineación de los datos. El programa realiza una serie de alineaciones por pares utilizando el algoritmo de alineación global Myers y Miller (con una penalización por apertura de espacio de 12, y una penalización de extensión de espacio de 2), calcula la similitud y la identidad utilizando, por ejemplo, Blosum 62 (para los polipéptidos), y luego coloca los resultados en una matriz de distancia. La similitud de secuencia se muestra en la mitad inferior de la línea de división y de identidad de secuencia se muestra en la parte superior de la línea de división diagonal.

Los parámetros utilizados en la comparación son:

Matriz de clasificación: Blosum62

Primer espacio: 12

Espacio de extensión: 2

25 Los resultados del análisis de software se muestran en la Tabla B para similitud e identidad global sobre la longitud completa de las secuencias de polipéptidos. El porcentaje de identidad se da por encima de la diagonal en negrita y el porcentaje de similitud se da por debajo de la diagonal (cara normal).

30 El porcentaje de identidad entre las secuencias de polipéptidos LBD útiles en la realización de los métodos de la invención puede ser tan bajo como 30.2% de identidad de aminoácidos en comparación con la SEQ ID NO: 2 (representado por A - LBD37, fila 44). El% de identidad probablemente será mayor sólo cuando se comparan las secuencias del dominio DUF206. Para identificar el dominio DUF260 (tal como se expone en la Figura 1) en otras proteínas LBD, se puede utilizar la alineación múltiple de la Figura 2.

Tabla A2: resultados MatGAT de similitud e identidad global en toda la longitud de las secuencias de polipéptidos LBD.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Os- Os01g03890		43.2	35.8	34.5	34.4	32.9	33.0	35.5	35.2	34.5	31.5	32.2
2. Os- Os01g32770	55.5		31.0	31.1	31.5	30.3	32.4	30.7	30.6	30.9	29.8	31.2
3. Os- Os03g33090	46.9	40.3		48.3	65.0	56.7	53.9	52.1	66.7	71.3	51.5	55.0
4. Os- Os03g41330	46.3	43.0	60.3		48.3	43.0	43.2	41.2	46.4	48.6	40.2	45.1
5. Os- Os07g40000	45.9	40.9	73.0	59.9		53.0	50.6	49.6	68.0	63.1	47.6	52.3

# ES 2 437 621 T3

	1	2	3	4	5	6	7	8	9	10	11	12
6. Nb- TC9404	45.9	42.4	66.5	60.3	67.7		60.7	60.3	51.7	51.5	56.7	66.5
7. Gm- TC227562	45.0	44.5	65.2	59.1	65.2	75.2		73.7	50.6	49.4	55.2	64.8
8. Gm- TC216138	47.2	41.5	63.3	58.8	62.1	70.4	79.6		50.0	48.8	53.3	65.3
9. Hv- TC147776	45.9	41.5	76.1	58.6	79.6	69.2	67.4	63.8		58.7	47.3	52.2
10. Sb- TC104758	45.3	40.6	77.0	59.6	71.1	63.0	62.6	58.8	68.5		49.0	51.7
11. Aq- TC18561	47.2	43.3	61.5	56.9	62.3	67.5	69.7	63.8	67.1	61.7		58.1
12. Vv- TC60668	45.0	43.3	65.4	59.1	66.7	79.4	77.4	77.1	66.2	64.7	72.7	
13. Lj- TC15459	45.6	40.9	64.1	59.1	64.1	72.2	81.4	88.8	67.5	61.6	70.5	79.3
14. Gh- TC30552	46.3	43.0	66.8	63.4	64.7	75.9	78.4	75.4	66.8	64.3	72.8	82.8
15. Ta- TC235711	56.7	61.5	42.9	43.2	46.3	46.7	46.3	47.4	47.0	44.6	48.4	45.6
16. St- TC133081	59.0	55.8	47.4	48.5	47.1	50.7	49.6	47.4	49.3	48.9	50.4	50.0
17. Mt- TC107091	58.6	60.0	45.5	46.2	43.5	46.5	46.2	47.2	42.5	44.9	49.2	49.2
18. Hv- TC147808	59.0	64.8	42.1	42.1	44.7	44.4	46.7	46.0	45.0	44.7	44.4	45.0
19. Vv- TC55931	57.0	55.2	49.4	47.8	46.2	53.0	51.4	49.8	48.2	46.2	53.8	51.8
20. St- TC162239	49.5	48.2	50.4	52.6	51.7	54.7	54.7	55.0	46.6	51.1	57.7	56.8
21. Pta- TC69225	53.0	50.0	38.9	41.6	40.4	44.9	44.9	43.1	43.1	41.9	44.9	44.6
22. Pta- TC67269	51.8	52.7	39.6	41.4	39.3	42.6	42.9	41.4	43.2	40.8	44.0	42.6
23. Gm- TC220806	45.9	46.1	56.4	49.1	51.3	54.5	52.6	51.7	52.3	48.5	52.4	53.5
24. Ta- TC270332	41.4	40.6	56.5	84.5	59.1	61.3	59.1	57.1	60.0	57.9	59.7	58.7
25. Aq-	62.5	61.5	46.6	47.9	44.1	50.0	50.7	47.6	44.8	45.2	51.4	49.7

# ES 2 437 621 T3

	1	2	3	4	5	6	7	8	9	10	11	12
TC18329												
26. St- TC137193	45.6	42.1	68.7	57.3	65.9	89.3	72.6	70.4	68.9	61.3	74.0	77.6
27. St- TC133385	45.6	41.8	70.8	56.9	66.4	87.9	71.3	68.3	66.2	62.1	70.6	76.3
28. St- TC140088	44.3	42.1	64.5	51.3	61.5	71.0	68.3	65.8	65.8	60.9	68.4	70.2
29. Pa- TC14656	48.5	45.2	47.7	46.9	48.4	51.6	52.7	52.7	47.3	48.0	53.4	53.8
30. Pta- TC59178	47.7	45.2	40.5	40.2	43.5	44.1	43.3	44.6	44.1	40.5	46.8	46.6
31. Pta- TC67974	48.2	45.2	50.0	48.2	46.4	52.5	52.2	54.3	48.6	49.3	54.0	54.3
32. Le- TC178827	45.9	41.8	68.3	57.8	66.8	89.7	73.9	71.7	67.6	62.1	73.6	78.1
33. Pt-III.589	45.0	43.9	63.0	60.1	62.6	73.5	76.5	79.6	68.5	64.3	69.7	79.0
34. Pt-V.543	48.2	43.9	64.5	60.3	60.7	74.0	76.4	77.7	66.5	63.6	68.6	77.3
35. Pt.XIV.94	41.7	39.4	64.4	56.0	58.0	63.4	61.3	58.8	61.3	55.7	64.9	64.5
36. Pt-II105	43.0	39.1	63.9	54.7	58.0	60.7	63.0	58.8	65.3	56.6	63.6	65.8
37. Pt-123.86	57.7	57.9	47.6	46.3	46.3	46.6	48.6	47.6	43.6	45.9	47.3	45.9
38. Pt-X180	52.1	49.4	50.2	51.0	49.8	54.8	55.6	55.2	48.5	50.2	56.0	57.7
39. Pt-XII.481	51.5	47.3	43.6	45.8	44.7	48.4	46.5	49.5	46.5	46.9	46.9	50.9
40. Ck-LOB- DQ787782	48.5	43.3	66.7	59.5	60.6	73.2	82.3	82.1	65.8	61.7	71.4	77.5
41. Bn- AAP37970	56.0	56.1	48.9	49.6	46.2	49.6	50.4	51.9	50.0	47.0	50.0	51.5
42. Mt- ABE82505	47.2	44.2	65.2	62.2	65.2	72.5	79.8	80.8	64.8	64.3	71.7	76.8
43. Mt- ABE78739	54.7	56.7	45.7	48.9	44.6	50.7	50.4	49.6	47.1	47.8	51.1	50.4
44. At- LBD37- Q9FN11	45.0	46.1	61.6	56.4	64.0	68.8	68.8	69.6	64.0	63.6	64.0	68.8
45. At- LBD38- Q9SN23	45.9	45.2	61.9	55.5	63.2	61.9	70.4	70.4	62.8	63.6	67.6	70.9
46. At- LBD39-	45.0	41.8	61.3	56.7	63.3	65.8	70.8	67.9	60.0	62.5	70.4	71.7

# ES 2 437 621 T3

	1	2	3	4	5	6	7	8	9	10	11	12
Q9SZE8												
47. At-LBD40-Q9ZW96	53.1	49.1	54.5	54.1	55.8	55.8	54.9	52.5	53.2	56.2	60.9	56.7
48. At-LBD41-Q9M886	52.8	56.7	44.9	49.0	47.5	49.8	50.6	47.9	49.8	47.1	51.0	50.2
49. At-LBD42-Q9CA30	52.4	49.4	54.1	50.2	51.9	58.4	57.5	52.9	54.1	52.8	55.4	55.4

	13	14	15	16	17	18	19	20	21	22	23	24
1. Os-Os01g03890	33.7	33.0	42.5	46.6	42.9	43.5	42.8	37.7	36.0	36.8	38.1	32.2
2. Os-Os01g32770	30.1	32.4	52.2	46.5	47.1	54.4	41.7	37.0	37.1	38.2	39.7	30.9
3. Os-Os03g33090	54.0	54.5	32.4	37.0	34.4	31.3	37.2	37.6	30.2	28.3	41.3	47.4
4. Os-Os03g41330	41.2	44.1	32.1	33.1	30.2	31.4	36.4	37.0	29.9	28.4	35.5	78.7
5. Os-Os07g40000	50.0	51.7	34.7	35.8	33.9	33.2	35.2	36.5	28.7	28.3	38.7	46.3
6. Nb-TC9404	60.8	63.7	33.7	35.8	33.2	32.1	38.7	38.1	32.3	29.8	39.6	44.0
7. Gm-TC227562	72.4	67.6	33.1	34.7	33.2	33.1	36.6	36.2	32.3	30.7	39.3	42.8
8. Gm-TC216138	81.5	66.5	34.4	34.2	33.2	33.2	37.0	35.7	31.1	30.7	37.6	42.3
9. Hv-TC147776	50.8	51.9	34.9	35.4	31.6	32.6	36.8	34.6	29.3	28.6	37.9	45.4
10. Sb-TC104758	49.4	51.3	33.8	34.5	32.5	33.1	35.3	38.0	29.6	30.1	37.4	46.9
11. Aq-TC18561	54.8	58.5	33.4	35.0	34.2	32.1	37.5	37.6	32.3	30.8	37.3	41.1
12. Vv-TC60668	65.6	68.3	33.1	35.6	35.9	32.1	36.2	38.5	33.3	32.0	39.6	43.5
13. Lj.TC15459		64.9	34.5	33.2	33.8	33.4	36.0	36.5	34.1	29.4	36.8	39.6
14. Gh-TC30552	76.8		36.2	35.4	33.9	34.1	37.5	36.9	32.0	30.7	40.1	44.2

# ES 2 437 621 T3

	13	14	15	16	17	18	19	20	21	22	23	24
15. Ta- TC235711	47.4	49.5		45.5	49.5	89.9	49.1	41.3	37.8	36.4	46.0	31.8
16. St- TC133081	50.0	50.4	57.5		56.6	47.4	52.5	45.4	37.4	40.1	50.0	32.5
17. Mt- TC107091	48.2	47.5	61.5	66.4		50.5	48.8	41.6	36.4	39.5	51.7	31.2
18. Hv- TC147808	45.4	47.7	92.4	59.6	63.2		48.3	39.0	37.8	39.1	43.7	31.8
19. Vv- TC55931	51.8	51.4	60.3	63.1	61.1	57.9		47.0	35.8	38.1	51.2	36.4
20. St- TC162239	54.9	55.1	55.7	58.8	52.8	52.3	63.2		32.9	36.0	48.7	37.7
21. Pta- TC69225	46.4	44.0	50.9	51.5	53.0	51.2	47.9	43.1		44.5	32.9	27.8
22. Pta- TC67269	42.3	44.3	50.9	54.8	52.4	52.4	50.0	46.4	63.1		35.3	28.6
23. Gm- TC220806	51.1	54.7	51.6	55.8	55.5	49.0	57.3	60.7	40.1	42.0		35.8
24. Ta- TC270332	54.9	61.2	44.3	45.3	44.2	42.7	47.8	50.9	41.0	40.2	51.7	
25. Aq- TC18329	49.7	51.0	63.8	70.0	70.8	65.2	65.9	55.5	52.1	53.6	54.1	45.9
26. St- TC137193	71.3	74.6	46.0	48.9	48.5	45.4	52.6	52.1	43.1	41.7	57.1	58.7
27. St- TC133385	70.0	73.3	43.9	48.9	46.5	43.0	52.6	53.0	41.9	41.4	57.1	57.0
28. St- TC140088	66.7	70.3	45.3	47.4	46.8	43.0	53.0	52.6	42.5	41.1	55.3	52.2
29. Pa- TC14656	51.6	53.4	50.2	49.8	48.2	50.0	46.6	52.3	47.3	46.1	44.8	44.8
30. Pta- TC59178	44.4	45.7	41.6	44.9	46.6	44.9	44.4	42.7	47.4	49.6	39.1	40.2
31. Pta- TC67974	50.7	54.3	49.1	50.0	48.2	47.0	47.8	51.8	46.4	48.2	46.7	47.8
32. Le- TC178827	72.6	75.4	46.7	48.9	48.2	43.7	53.0	52.1	42.5	43.2	56.9	58.3
33. Pt-III.589	78.6	81.9	50.5	51.5	48.5	47.0	55.3	57.1	44.3	44.6	53.4	58.0
34. Pt-V.543	76.9	79.8	49.1	52.2	49.2	48.0	53.0	56.2	45.2	45.5	53.3	57.0

# ES 2 437 621 T3

	13	14	15	16	17	18	19	20	21	22	23	24
35. Pt-XIV.94	59.9	62.5	44.9	45.6	42.5	44.0	48.2	51.7	41.6	41.4	56.3	57.8
36. Pt-II105	61.6	62.5	43.9	47.4	43.5	42.1	47.4	55.1	38.9	40.8	56.7	55.2
37. Pt-123.86	47.6	48.3	56.1	67.2	75.4	55.3	61.8	54.7	53.3	51.5	55.1	43.9
38. Pt-X180	56.0	56.8	53.0	56.6	55.1	51.0	66.4	59.8	48.8	47.6	60.2	49.8
39. Pt-XII.481	48.4	50.2	52.3	58.0	53.8	49.7	58.2	54.6	45.5	49.1	51.3	44.0
40. Ck-LOB-DQ787782	80.2	79.3	49.1	51.8	46.5	47.0	53.0	54.3	43.7	44.9	53.2	58.0
41. Bn-AAP37970	47.0	51.1	58.9	65.7	67.4	56.0	60.2	57.6	47.6	48.5	56.8	47.3
42. Mt-ABE82505	78.9	77.7	50.2	50.7	47.8	48.7	54.2	54.3	44.9	45.8	52.8	61.8
43. Mt-ABE78739	53.6	50.4	59.2	66.9	71.8	60.3	64.4	54.0	52.4	49.7	56.5	46.4
44. At-LBD37-Q9FN11	67.6	75.2	48.8	54.0	49.5	47.0	54.9	55.6	45.2	44.6	52.0	56.4
45. At-LBD38-Q9SN23	63.6	73.7	48.4	51.1	48.8	47.0	51.8	55.1	46.4	46.1	52.2	53.0
46. At-LBD39-Q9SZE8	73.8	70.0	48.8	50.4	47.8	47.4	52.6	53.8	45.2	43.8	51.7	61.7
47. At-LBD40-Q9ZW96	53.2	56.2	54.4	61.7	62.1	51.3	60.9	61.5	48.5	48.2	62.2	52.8
48. At-LBD41-Q9M886	49.8	52.1	56.8	67.5	67.4	57.0	61.2	58.9	49.1	49.1	57.4	46.0
49. At-LBD42-Q9CA30	56.1	54.5	54.0	54.7	54.5	51.3	60.5	64.5	47.0	45.8	62.7	51.1
	25	26	27	28	29	30	31	32	33	34	35	36
1. Os-Os01g03890	44.0	32.9	33.6	32.2	32.4	32.5	32.1	33.5	33.7	33.7	33.2	33.2
2. Os-Os01g32770	46.3	30.0	29.4	29.7	28.1	29.4	30.4	30.9	30.6	29.1	30.0	31.5
3. Os-Os03g33090	33.3	61.0	60.2	50.0	35.7	33.9	38.7	60.0	52.9	52.9	50.2	51.7
4. Os-Os03g41330	31.7	44.0	44.0	39.0	36.7	30.2	36.1	44.7	41.0	41.0	41.0	39.7
5. Os-Os07g40000	34.1	52.6	52.5	46.8	35.3	33.3	37.5	51.9	49.6	48.6	46.5	46.0
6. Nb-TC9404	36.6	83.2	82.3	57.5	38.9	36.9	41.9	83.2	64.2	64.9	50.2	49.1

# ES 2 437 621 T3

	13	14	15	16	17	18	19	20	21	22	23	24
7. Gm- TC227562	35.2	61.1	60.3	51.0	39.8	36.6	41.1	62.1	65.3	64.5	48.5	48.5
8. Gm- TC216138	34.1	61.4	60.2	51.7	38.2	34.9	40.2	61.7	67.2	66.8	47.5	46.9
9. Hv- TC147776	32.8	54.2	54.6	50.2	33.8	32.7	35.7	53.6	50.0	49.8	46.2	47.3
10. Sb- TC104758	31.4	51.7	51.5	46.7	35.6	33.1	37.5	49.0	52.5	50.2	43.8	45.1
11. Aq: TC18561	35.5	59.3	58.8	52.8	38.3	36.4	40.5	58.0	56.8	59.4	52.6	48.5
12. Vv- TC60668	35.1	67.1	67.2	54.5	39.6	36.6	42.3	66.5	69.4	66.7	50.9	50.4
13. Lj- TC15459	33.4	59.7	59.2	51.0	37.4	37.6	39.2	60.7	65.6	64.9	47.9	47.9
14. Gh- TC30552	35.2	63.9	63.4	57.3	39.4	37.2	41.6	64.2	71.4	70.9	50.0	50.0
15. Ta- TC235711	50.3	33.8	32.8	32.4	29.5	31.0	30.8	34.5	36.2	34.8	33.4	34.5
16. St- TC133081	57.4	33.2	34.7	32.5	30.9	32.1	32.6	35.4	36.4	36.0	35.0	34.2
17. Mt- TC107091	56.8	34.2	33.6	33.1	31.5	32.0	31.0	33.9	33.8	35.2	29.6	30.8
18. Hv- TC147808	50.6	34.4	31.5	29.1	31.4	31.6	30.3	33.1	34.8	34.8	32.1	32.5
19. Vv- TC55931	56.3	37.5	39.2	36.6	33.8	31.2	34.9	38.4	37.3	35.5	36.8	35.9
20. St- TC162239	41.8	36.3	37.2	35.0	35.6	32.0	35.7	37.6	39.0	37.3	37.6	37.9
21. Pta- TC69225	36.9	30.5	31.6	29.6	30.5	30.7	31.2	29.9	31.7	31.8	31.1	30.2
22. Pta- TC67269	43.1	30.1	29.2	29.2	28.5	30.0	32.2	31.5	29.5	31.0	29.8	29.8
23. Gm- TC220806	50.0	43.4	42.1	38.4	34.7	28.1	35.5	42.8	39.6	38.5	40.5	41.9
24. Ta- TC270332	32.1	45.5	43.5	39.2	35.7	30.3	36.8	43.3	41.3	40.9	40.9	39.7
25. Aq- TC18329		35.5	35.2	31.4	30.2	33.1	31.5	34.8	33.4	35.5	33.4	34.1
26. St-	48.6		96.8	58.4	38.8	36.1	41.9	95.0	64.3	62.5	50.2	50.2

# ES 2 437 621 T3

	13	14	15	16	17	18	19	20	21	22	23	24
TC137193												
27. St- TC133385	48.6	96.8		59.7	38.1	35.8	40.1	92.8	63.9	60.9	50.7	50.9
28. St- TC140088	47.9	72.4	71.9		38.8	33.2	39.8	58.7	54.1	53.8	46.3	45.6
29. Pa- TC14656	49.7	51.6	51.6	54.2		38.6	88.1	38.7	42.1	40.3	33.5	33.5
30. Pta- TC59178	46.0	44.9	44.6	43.3	54.3		39.4	36.9	36.4	38.4	32.5	31.4
31. Pta- TC67974	49.3	52.2	51.4	52.9	92.8	53.2		41.7	42.5	42.8	35.0	36.1
32. Le- TC178827	48.3	97.2	95.0	71.6	50.9	45.2	51.4		64.6	62.2	49.5	49.5
33. Pt-III.589	49.7	73.5	72.3	65.5	54.2	44.9	56.5	75.2		83.3	50.4	48.7
34. Pt-V.543	50.3	73.6	71.5	64.0	54.2	46.8	56.2	73.1	88.8		50.8	48.3
35. Pt.XIV.94	47.9	63.1	63.2	62.7	47.7	39.7	51.1	61.0	63.4	62.0		86.1
36. Pt-II105	47.2	63.1	64.6	60.4	48.4	39.4	51.8	61.5	62.2	59.9	91.8	
37. Pt-123.86	70.9	48.0	46.6	48.6	49.7	47.1	50.7	46.6	49.0	50.0	43.9	42.9
38. Pt-X180	61.0	51.9	51.9	53.5	46.2	41.6	46.7	53.5	58.5	58.3	50.2	48.5
39. Pt-XII.481	60.0	50.5	49.5	46.9	49.5	41.6	50.0	50.9	49.5	49.5	45.1	42.1
40. Ck.LOB- DQ787782	47.2	72.3	71.4	68.8	53.4	44.1	54.7	71.9	79.4	78.9	60.6	62.3
41. Bn- AAP37970	65.2	52.3	50.8	47.0	47.7	44.1	50.0	50.8	49.6	51.5	50.4	49.2
42. Mt- ABE82505	50.3	70.4	69.1	67.0	57.4	46.8	55.1	71.2	76.9	78.9	62.7	62.2
43. Mt- ABE78739	67.2	50.0	50.0	50.7	54.7	43.3	53.2	48.2	49.6	48.6	50.0	48.9
44. At-LBD37- Q9FN11	51.4	66.8	65.6	64.4	54.2	47.9	52.9	67.2	76.0	74.0	60.8	60.4
45. At-LBD38- Q9SN23	49.7	65.6	64.4	60.3	52.3	44.9	53.6	66.0	75.3	72.1	58.7	61.5
46. At-LBD39- Q9SZE8	51.0	66.3	64.6	62.1	54.5	46.8	53.3	67.9	71.7	71.1	59.6	60.8
47. At-LBD40- Q9ZW96	61.7	56.7	55.4	54.9	49.1	41.0	49.6	55.8	55.9	55.4	50.6	51.5

# ES 2 437 621 T3

	13	14	15	16	17	18	19	20	21	22	23	24
48. At-LBD41-Q9M886	66.9	50.6	49.8	48.7	50.5	45.5	49.3	52.1	48.7	51.3	49.0	47.1
49. At-LBD42-Q9CA30	58.3	53.6	54.1	55.4	47.7	41.3	48.2	54.5	54.6	56.2	54.9	54.5

	37	38	39	40	41	42	43	44	45	46
1. Os-Os01g03890	44.0	39.2	34.8	35.5	44.4	34.9	40.5	31.9	33.3	32.4
2. Os-Os01g32770	45.3	38.5	32.3	32.1	42.7	33.9	44.8	30.2	32.1	29.7
3. Os-Os03g33090	34.7	36.9	32.7	53.0	36.4	54.9	35.1	50.0	51.2	49.6
4. Os-Os03g41330	33.7	35.2	31.2	43.6	36.2	44.6	32.6	43.9	42.4	40.7
5. Os-Os07g40000	32.8	35.7	31.9	49.8	32.6	51.3	33.1	49.2	47.4	48.1
6. Nb-TC9404	33.8	37.3	32.6	58.6	32.5	61.1	35.1	56.9	54.0	56.3
7. Gm-TC227562	32.8	36.4	32.4	76.6	34.7	74.9	37.1	56.2	57.8	57.3
8. Gm-TC216138	31.6	36.7	31.6	74.3	36.0	74.0	34.9	53.9	55.0	54.1
9. Hv-TC147776	31.1	38.2	33.5	50.6	34.3	50.6	34.5	49.2	48.6	45.8
10. Sb-TC104758	33.1	35.2	31.3	48.7	33.7	50.2	33.1	50.6	48.2	46.7
11. Aq-TC18561	33.4	36.2	32.5	56.7	36.0	57.8	36.0	51.6	55.6	57.1
12. Vv-TC60668	33.4	39.7	34.3	65.3	33.7	66.8	36.0	54.2	57.4	58.8
13. Lj-TC15459	32.1	35.9	31.0	70.4	32.8	70.7	35.3	52.8	53.0	54.8
14. Gh-TC30552	33.7	38.6	33.7	66.2	35.2	69.0	32.9	62.4	63.5	58.7
15. Ta-TC235711	45.3	42.7	36.2	35.4	46.0	38.3	46.6	34.1	35.9	33.7
16. St-TC133081	55.7	45.6	37.3	36.5	55.8	37.2	54.9	37.1	35.8	36.5
17. Mt-TC107091	63.7	43.0	37.3	34.2	54.3	33.9	60.6	33.4	37.2	34.2
18. Hv-TC147808	45.0	41.9	34.3	34.3	45.2	37.4	46.8	32.4	35.4	32.8

# ES 2 437 621 T3

	37	38	39	40	41	42	43	44	45	46
19. Vv-TC55931	52.0	51.4	40.9	36.8	49.8	38.7	48.9	37.4	34.9	36.7
20. St-TC162239	41.6	45.1	39.7	36.3	41.8	36.0	42.9	35.4	35.7	35.0
21. Pta-TC69225	37.6	33.4	29.4	33.4	34.7	32.3	38.0	31.7	31.1	32.5
22. Pta-TC67269	38.8	36.8	31.2	30.4	35.7	31.2	38.1	31.3	32.3	31.3
23. Gm-TC220806	48.7	48.8	40.1	39.7	48.1	37.9	51.3	36.5	38.0	38.4
24. Ta-TC270332	31.8	33.3	29.3	42.5	32.5	44.2	32.4	41.5	40.4	43.7
25. Aq-TC18329	61.1	48.3	40.6	33.4	51.2	36.6	55.9	35.1	32.8	34.4
26. St-TC137193	34.1	37.3	35.7	60.9	33.0	60.0	34.4	57.3	55.7	54.1
27. St-TC133385	34.0	37.3	36.4	60.1	34.5	59.6	33.8	56.5	54.9	53.1
28. St-TC140088	33.3	36.9	33.5	53.0	31.4	51.9	31.9	48.8	47.0	49.6
29. Pa-TC14656	32.1	33.3	33.2	41.1	29.3	41.7	32.8	39.0	40.0	39.9
30. Pta-TC59178	34.0	30.3	28.3	37.4	32.8	38.4	30.2	36.5	35.7	37.6
31. Pta-TC67974	33.2	34.2	34.1	42.2	31.9	41.9	34.2	40.9	42.6	40.4
32. Le-TC178827	32.8	38.9	36.5	60.8	34.1	60.7	33.1	56.5	55.5	55.3
33. Pt-III.589	34.7	37.9	33.5	67.1	35.5	65.7	32.7	64.1	63.9	57.5
34. Pt-V.543	34.3	36.6	33.3	67.4	34.5	66.7	32.0	62.6	62.9	59.3
35. Pt-XIV.94	31.8	35.8	30.4	49.4	36.0	51.7	37.1	45.0	45.3	46.5
36. Pt-II105	32.4	35.0	30.4	48.5	36.7	51.1	35.6	43.6	45.7	48.1
37. Pt-123.86		45.8	39.3	33.1	56.2	34.3	55.9	34.8	34.1	33.0
38. Pt-X180	56.4		43.1	34.6	42.3	34.9	44.1	37.2	37.3	39.9
39. Pt-XII.481	55.4	56.8		32.5	37.5	33.6	37.4	31.9	30.5	29.9
40. Ck-LOB-DQ787782	49.0	53.9	47.3		35.5	80.9	33.8	54.0	55.8	57.8

# ES 2 437 621 T3

	37	38	39	40	41	42	43	44	45	46
41. Bn-AAP37970	70.3	56.8	55.7	51.5		37.7	51.6	33.2	34.1	34.3
42. Mt-ABE82505	47.0	55.2	47.3	87.1	51.9		36.7	55.6	57.3	56.7
43. Mt-ABE78739	67.6	57.9	55.4	50.7	64.7	51.4		31.9	32.7	34.3
44. At-LBD37-Q9FN11	53.0	53.2	48.0	69.2	50.4	66.8	50.0		76.9	55.3
45. At-LBD38-Q9SN23	47.0	55.9	49.5	70.0	51.5	69.2	48.6	86.4		56.8
46. At-LBD39-Q9SZE8	49.0	57.3	49.1	72.5	53.8	71.3	51.1	61.2	68.8	
47. At-LBD40-Q9ZW96	65.5	62.7	52.0	55.8	68.9	56.2	66.2	55.2	55.5	56.3
48. At-LBD41-Q9M886	70.9	58.6	57.1	47.9	90.5	50.2	65.5	53.2	51.7	52.1
49. At-LBD42-Q9CA30	55.4	62.7	52.0	58.4	56.8	55.8	57.2	53.6	52.6	52.9

	47	48	49							
1. Os-Os01g03890	41.5	41.3	38.2							
2. Os-Os01g32770	41.5	44.8	39.0							
3. Os-Os03g33090	40.4	35.0	40.8							
4. Os-Os03g41330	38.2	33.7	36.0							
5. Os-Os07g40000	40.2	33.8	36.3							
6. Nb-TC9404	39.4	34.6	40.7							
7. Gm-TC227562	40.8	35.7	39.8							
8 Gm-TC216138	36.5	34.0	35.5							
9. Hv-TC147776	38.2	34.6	37.7							
10. Sb-TC104758	36.2	34.5	38.6							
11. Aq-TC18561	42.1	37.2	36.6							

# ES 2 437 621 T3

	47	48	49
12. Vv-TC60668	40.2	36.1	38.2
13. Lj-TC15459	38.1	35.1	36.4
14. Gh-TC30552	39.8	34.6	36.6
15. Ta- TC235711	44.3	44.3	41.0
16. St-TC133081	50.7	55.9	45.7
17. Mt- TC107091	52.6	56.1	40.2
18. Hv- TC147808	42.1	46.5	39.7
19. Vv-TC55931	49.2	49.3	47.3
20. St-TC162239	46.5	44.5	47.5
21. Pta-TC69225	36.4	36.2	34.1
22. Pta-TC67269	36.0	36.3	33.3
23. Gm- TC220806	52.8	50.8	49.8
24. Ta- TC270332	39.2	30.4	36.3
25. Aq-TC18329	49.8	53.2	43.7
26. St-TC137193	39.2	34.2	38.6
27. St-TC133385	40.1	34.1	38.4
28. St-TC140088	37.8	35.7	36.6
29. Pa-TC14656	33.8	30.4	31.3
30. Pta-TC59178	30.2	32.5	27.2
31. Pta-TC67974	34.8	31.8	33.0
32. Le- TC178827	39.0	34.6	38.5
33. Pt-III.589	36.6	35.0	35.3
34. Pt-V.543	38.2	35.4	39.2
35. Pf-XIV.94	37.8	33.8	38.3
36. Pt-II105	40.2	35.7	37.4
37. Pt-123.86	53.8	59.1	43.6

	47	48	49
38. Pt-X180	46.8	41.5	49.6
39. Pt-XII.481	38.5	37.6	36.4
40. Ck-LOB-DQ787782	40.2	31.6	38.2
41. Bn-AAP37970	55.5	83.5	42.1
42. Mt-ABE82505	40.3	33.8	36.8
43. Mt-ABE78739	53.3	51.6	43.8
44. At-LBD37-Q9FN11	37.5	34.5	34.9
45. At-LBD38-Q9SN23	40.5	33.5	34.8
46. At-LBD39-Q9SZE8	38.3	35.0	37.1
47. At-LBD40-Q9ZW96		56.7	46.5
48. At-LBD41-Q9M886	68.4		42.6
49. At-LBD42-Q9CA30	61.8	56.3	

**Ejemplo 4:** Identificación de dominios comprendidos en secuencias de polipéptidos útiles en la realización de los métodos de la invención

- 5 Los Recursos Integrados de base de datos de Familias de Proteínas, Dominios y Sitios (InterPro) es una interfaz integrada para las bases de datos distintivas de uso común para las búsquedas con base en texto y en secuencia. La base de datos de InterPro combina estas bases de datos, que utilizan diferentes metodologías y distintos grados de información biológica acerca de proteínas bien caracterizadas para obtener proteínas distintivas. Las bases de datos de colaboradores incluyen SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom y Pfam, Smart y TIGREAMS. Pfam es una gran colección de alineaciones de secuencias múltiples y modelos Markov ocultos que cubren muchos dominios y familias de proteínas comunes. Pfam está alojado en el servidor del Instituto Sanger en el Reino Unido. Interpro está alojado en el Instituto Europeo de Bioinformática en el Reino Unido.
- 10 Los resultados de la exploración InterPro de la secuencia de polipéptidos como se representa por la SEQ ID NO: 2 se presentan en la Tabla A3.

Tabla A3: Resultados de escaneo de InterPro (principales números de acceso) de la secuencia de polipéptido como se representa por la SEQ ID NO: 2.

base de datos	Número de acceso	Nombre de adhesión	coordenadas de aminoácidos en la SEQ ID NO 2
PFAM	PF03195	DUF260	2-107
PROFILE	PS50891	LOB	1-107

5 **Ejemplo 5:** Topología de predicción de las secuencias de polipéptidos útiles en la realización de los métodos de la invención

El TargetP 1.1 predice la localización subcelular de las proteínas eucarióticas. La asignación de ubicación se basa en la presencia predicha de cualquiera de las pre-secuencias terminales N: péptido de tránsito al cloroplasto (CTP), péptido de direccionamiento mitocondrial (MTP) o péptido señal de ruta secretora (SP). Las clasificaciones en la que se basa la predicción final en realidad no son probabilidades, y no necesariamente se agregan a una. Sin embargo, la ubicación con la puntuación más alta es la más probable de acuerdo con TargetP, y la relación entre las clasificaciones (la clase de fiabilidad) puede ser una indicación de cómo la predicción es cierta. La clase de fiabilidad (RC) varía de 1 a 5, donde 1 indica la predicción más fuerte. TargetP se mantiene en el servidor de la Universidad Técnica de Dinamarca.

10 15 También se pueden predecir las secuencias predichas que contienen una pre-secuencia de terminal N de un sitio de división potencial.

Se selecciona un número de parámetros, tales como el grupo organismo (no de planta o planta), conjuntos de corte (ninguno, conjunto predefinido de corte, o conjuntos de corte específicos a usuario), y el cálculo de la predicción de sitios de división (sí o no).

20 Los resultados de análisis TargetP 1.1 de la secuencia de polipéptidos como se representa por la SEQ ID NO: 2 se presentan la Tabla A4. El grupo organismo “planta” se ha seleccionado, no hay corte definidos, y se requiere la longitud predicha del péptido de tránsito.

Tabla A4: Análisis TargetP 1.1 de la secuencia de polipéptido como la representada por la SEQ ID NO: 2

Longitud (AA)	250
Cloroplásticos péptido de tránsito	0.026
Péptido de tránsito mitocondrial	0.401
Péptido señal de ruta secretora	0.019
Otra obetivación subcelular	0.573
Ubicación predich	/
Clase de Fiabilidad	5
Longitud de péptido de tránsito predicha	/

25 La localización subcelular de la secuencia de polipéptido como se representa por la SEQ ID NO: 2 puede ser el citoplasma o el núcleo, no se prevé ningún péptido de tránsito. SubLoc (Hua & Sun, Bioinformatics, 17, 721-728, 2001) predice una localización nuclear (índice de fiabilidad: 2, precisión: 74%); esta predicción está de acuerdo con los datos de Liu et al (2005).

Muchos otros algoritmos se pueden utilizar para llevar a cabo este tipo de análisis, que incluyen:

- ChloroP 1.1 alojado en el servidor de la Universidad Técnica de Dinamarca;
- Predictor de Localización Subcelular de Proteínas Prowler versión 1.2 alojado en el servidor del Instituto de Biociencia Molecular de la Universidad de Queensland, Brisbane, Australia;
- 5      • analista de proteoma PENCE PA- GOSUB 2.5 alojado en el servidor de la Universidad de Alberta, Edmonton, Alberta, Canadá;
- TMHMM, alojado en el servidor de la Universidad técnica de Dinamarca

**Ejemplo 6:** Clonación de la secuencia de ácidos nucleicos LBD utilizado en los métodos de la invención

10 La secuencia de ácidos nucleicos utilizada en los métodos de la invención se amplifica por PCR utilizando como plantilla una colección de cADN de plántula de *Arabidopsis thaliana* personalizada (en pCMV Sport 6.0; Invitrogen, Paisley, Reino Unido). Se realiza PCR utilizando polimerasa ADN Hifi Taq en condiciones estándar, utilizando 200 ng de plantilla en una mezcla de 50 ml de PCR. Los cebadores utilizados fueron prm009067 (SEQ ID NO: 3; sentido, codón de inicio en negrita):

15 5' - ggggacaagttgtacaaaaaaggcaggcttaacaatgagctgcaatgggtgc - 3' y prm009068 (SEQ ID NO: 4; inversa, complementaria):

20 5' - ggggaccactttgtacaagaaaagctgggtactaactctgagaaaaaccgcc - 3',

que incluye los sitios AttB para recombinación Gateway. El fragmento de PCR amplificado se purifica utilizando también métodos estándar. Luego, se realiza la primera etapa del procedimiento Gateway, la reacción BP, durante la cual los fragmentos de PCR se recombinan *in vivo* con el plásmido pDONR201 para producir, de acuerdo con la terminología Gateway, un “clon de entrada”, pLBD. El plásmido pDONR201 se compra de Invitrogen, como parte de la tecnología Gateway®.

25 El clon de entrada que comprende la SEQ ID NO: 1 se utiliza luego en una reacción LR con un vector de destino utilizado para la transformación de *Oryza sativa*. Este vector contiene como elementos funcionales dentro de los límites de T - ADN: un marcador seleccionable de planta; un casete de expresión de marcador detectable, y un casete Gateway destinado a recombinación *in vivo* LR con la secuencia de ácidos nucleicos de interés ya clonada en el clon de entrada. Un promotor GOS2 de arroz (SEQ ID NO: 10) para la expresión constitutiva se ubica en dirección 5' de este casete Gateway.

Después de la etapa de recombinación LR, el vector de expresión resultante pGOS2::LBD (Figura 4) se transforma en la cepa de *Agrobacterium* LBA4044 de acuerdo con métodos bien conocidos en la técnica.

**Ejemplo 7:** Transformación de Planta

30 Transformación de arroz

35 El *Agrobacterium* que contiene el vector de expresión se utiliza para transformar plantas de *Oryza sativa*. Semillas secas maduras del arroz japonica cultivar Nipponbare se descascarillan. La esterilización se lleva a cabo al incubar durante un minuto en etanol al 70%, seguido de 30 minutos en 0.2% de HgCl<sub>2</sub>, seguido de unas 6 veces de lavado de 15 minutos con agua destilada estéril. Las semillas estériles luego se germinan en un medio que contiene 2,4 - D (medio de inducción de callo). Después de la incubación en la oscuridad durante cuatro semanas, se dividen los callos embriogénicos, derivado del escutelo y se propagan en el mismo medio. Después de dos semanas, los callos se multiplican o propagan mediante subcultivo en el mismo medio durante otras 2 semanas. Las piezas de callos embriogénicos se subcultivan en medio fresco 3 días antes de co-cultivo (para reforzar la actividad de división celular).

40 La cepa de *Agrobacterium* LBA4404 que contiene el vector de expresión se utiliza para co-cultivo. El *Agrobacterium* se inocula en un medio AB con los antibióticos apropiados y se cultiva durante 3 días a 28° C. Luego se recogen las bacterias y se suspenden en medio de co -cultivo líquido a una densidad (DO<sub>600</sub>) de aproximadamente 1. La suspensión se transfiere después a una placa Petri y los callos sumergido en la suspensión durante 15 minutos. Los tejidos de callo se transfieren luego en seco en un papel de filtro y se transfieren a solidificado, medio de co-cultivo y se incuban durante 3 días en la oscuridad a 25° C. Co-cultivada callos se cultivan en 2, 4-D que contiene medio durante 4 semanas en la oscuridad a 28° C en presencia de un agente de selección. Durante este período, islas de callos resistentes desarrolladas crecen rápidamente. Después de la transferencia de este material a un medio de regeneración y la incubación a la luz, el potencial embriogénico se libera y se desarrollan brotes en las próximas cuatro a cinco semanas. Los brotes se extirpan de los callos y se incuban durante 2 a 3 semanas en un medio que

contiene auxina desde el cual se transfieren al suelo. Los brotes endurecidos se hacen crecer bajo alta humedad y días cortos en un invernadero.

5 Aproximadamente 35 transformantes independientes de arroz T0 se generan por una construcción. Los transformantes primarios se transfieren de una cámara de cultivo de tejidos a un invernadero. Después se transfieren los transformantes primarios desde una cámara de cultivo de tejidos a un invernadero. Después de un análisis cuantitativo de PCR para verificar el número de copias de la inserción de T-ADN, sólo copias de plantas transgénicas individuales que exhiben tolerancia al agente de selección se mantienen para la cosecha de las semillas T1. Las semillas luego se recogen a continuación tres a cinco meses después del trasplante. El método produce transformantes de sitios únicos en un índice de más del 50% (Aldemita and Hodges, 1996, Chan et al. 10 1993, Hiei et al. 1994).

#### Transformación del maíz

15 La transformación de maíz (*Zea mays*) se realiza con una modificación del método descrito por Ishida et al. (1996) *Nature Biotech* 14 (6): 745-50. La transformación es dependiente del genotipo en el maíz y sólo los genotipos específicos son susceptibles de transformación y regeneración. La estirpe A188 consanguínea (Universidad de Minnesota) o híbridos con A188 como un progenitor son una buena fuente de material donante para transformación, pero también se pueden utilizar otros genotipos con éxito. Las mazorcas se cosechan de la planta de maíz aproximadamente 11 días después de la polinización (DAP) cuando la longitud del embrión inmaduro es de aproximadamente 1 a 1.2 mm. Los embriones inmaduros se cocultivan con *Agrobacterium tumefaciens* que contiene el vector de expresión, y las plantas transgénicas se recuperan por organogénesis. Los embriones extirpados se 20 cultivan en medio de inducción de callo, luego medio de regeneración, el maíz, que contiene el agente de selección (por ejemplo imidazolinona pero se pueden utilizar diferentes marcadores de selección). Las placas de Petri se incuban a la luz a 25° C durante 2-3 semanas, o hasta que se desarrollan los brotes. Los brotes verdes se transfieren de cada embrión a medio de enraizamiento de maíz y se incuban a 25° C durante 2-3 semanas, hasta 25 que se desarrollan las raíces. Los brotes enraizados se trasplantan al suelo en el invernadero. Las semillas T1 se producen a partir de plantas que exhiben tolerancia al agente de selección y que contienen una única copia de inserción de T-ADN.

#### Transformación de trigo

30 La transformación de trigo se realiza con el método descrito por Ishida et al. (1996) *Nature Biotech* 14 (6): 745-50. Se utiliza habitualmente cultivar Bobwhite (disponible de CIMMYT, México) en la transformación. Los embriones inmaduros se co-cultivan con *Agrobacterium tumefaciens* que contiene el vector de expresión, y las plantas transgénicas se recuperan por organogénesis. Después de incubación con *Agrobacterium*, los embriones se cultivan in vitro en medio de inducción de callo, a continuación, luego medio de regeneración, que contiene el agente de selección (por ejemplo imidazolinona pero se pueden utilizar diferentes marcadores de selección). Las placas de Petri se incuban a la luz a 25° C durante 2-3 semanas, o hasta que se desarrollan brotes. Los brotes verdes se 35 transfieren de cada embrión a un medio de enraizamiento y se incuban a 25° C durante 2-3 semanas, hasta que se desarrollan raíces. Los brotes enraizados se transplantan al suelo en el invernadero. Las semillas T1 se producen a partir de plantas que exhiben tolerancia al agente de selección y que contienen una única copia de inserción de T-ADN.

#### Transformación de soja

40 La soja se transforma de acuerdo con una modificación del método descrito en la patente Estadounidense de Texas A & M 5,164,310. Diversas variedades de soja comerciales son susceptibles a transformación por este método. Se utiliza comúnmente cultivar Jack (disponible de la base de semillas de Illinois) para transformación. Las semillas de soja se sintetizan en la siembra in vitro. El hipocotilo, la radícula y un cotiledón se dividen de plántulas jóvenes de siete días. El epicotilo y el cotiledón restante se cultivan adicionalmente para desarrollar ganglios axilares. Estos 45 ganglios axilares se extirpan y se incuban con *Agrobacterium tumefaciens* que contiene el vector de expresión. Después del tratamiento de co-cultivo, los explantes se lavan y se transfieren a medio de selección. Los brotes regenerados se extirpan y se colocan en un medio de alargamiento de brotes. Los brotes no mayores de 1 cm se colocan en un medio de enraizamiento hasta que se desarrollan las raíces. Los brotes enraizados se trasplantan a suelo en el invernadero. Las semillas T1 se producen a partir de plantas que exhiben tolerancia al agente de selección y que contienen una única copia de inserción de T-ADN.

#### Transformación de colza/canola

50 Los pecíolos cotiledonares e hipocotilos de plántulas jóvenes de 5-6 días se utilizan como explantes para el cultivo de tejidos y se transforma de acuerdo a Babic et al. (1998, *Plant Cell Rep.* 17: 183-188). El cultivar Westar comercial (Agriculture Canada) es la variedad estándar que se utiliza para la transformación, pero también se pueden utilizar otras variedades. Las semillas de canola se esterilizan en la superficie para siembra in vitro. Los explantes de

pecíolo de cotiledones con el cotiledón adjunto se dividen de plántulas in vitro, y se inoculan con Agrobacterium (que contiene el vector de expresión) al sumergir el extremo cortado del explante de pecíolo en la suspensión bacteriana. Los explantes luego se cultivan durante 2 días en medio MSBAP - 3 que contiene 3 mg/l de BAP, 3% de sacarosa, 0.7% de Phytagar a 23° C, 16 h de luz. Después de dos días de co-cultivo con Agrobacterium, los explantes de pecíolo se transfieren a medio MSBAP - 3 que contiene 3 mg/l de BAP, cefotaxima, carbenicilina o timentina (300 mg/l) durante 7 días, y luego se cultivan en medio MSBAP - 3 con cefotaxima, carbenicilina o timentina y agente de selección hasta la regeneración de brotes. Cuando los brotes son de 5 - 10 mm de longitud, se cortan y se transfieren a medio de alargamiento de brote (MSBAP - 0.5, que contiene 0.5 mg/l de BAP). Los brotes de alrededor de 2 cm de longitud se transfieren al medio de enraizamiento (MS0) para la inducción de la raíz. Los brotes enraizados se transplantan al suelo en el invernadero. Las semillas T1 se producen a partir de plantas que exhiben tolerancia al agente de selección y que contienen una única copia de la inserción de T-ADN.

#### Transformación de Alfalfa

Un clon de regeneración de alfalfa (*Medicago sativa*) se transforma utilizando el método de (McKersie et al., 1999 Plant Physiol 119: 839-847). La regeneración y transformación de la alfalfa es dependiente del genotipo y por lo tanto se requiere una planta de regeneración. Se han descrito métodos para obtener la regeneración de plantas. Por ejemplo, éstos se pueden seleccionar de la variedad cultivar Rangelander (Agriculture Canada) o cualquier otra variedad de alfalfa comercial como se describe por Brown DCW and A Atanassov (1985 Plant Cell Tissue Organ Culture 4: 111-112). Alternativamente, se ha seleccionado la variedad RA3 (Universidad de Wisconsin) para su uso en cultivo de tejidos (Walker et al., 1978 Am J Bot 65:654-659). Los explantes de pecíolo se cocultan con un cultivo de una noche de Agrobacterium tumefaciens C58C1 pMP90 (McKersie et al, 1999 Plant Physiol 119: 839 a 847) o LBA4404 que contiene el vector de expresión. Los explantes se cocultan durante 3 días en oscuridad en medio de inducción SH que contiene 288 mg/L de Pro, 53 mg/L de tioprolina, 4,35 g/L de K<sub>2</sub>SO<sub>4</sub>, y 100 µm de acetosiringinona. Los explantes se lavan en medio Murashige Skoog de resistencia media (Murashige and Skoog, 1962) y se siembran en el mismo medio de inducción SH sin acetosiringinona pero con un agente de selección adecuado y antibiótico adecuado para inhibir el crecimiento de Agrobacterium. Después de varias semanas, los embriones somáticos se transfieren a un medio de desarrollo BO2Y que no contiene reguladores de crecimiento, ni antibióticos, y 50 g/l de sacarosa. Los embriones somáticos posteriormente se germinan en medio de Murashige – Skoog de resistencia media. Se trasplantan las plántulas enraizadas a macetas y se hacen crecer en un invernadero. Las semillas T1 se producen a partir de plantas que exhiben tolerancia al agente de selección y que contienen una única copia de inserción de T- ADN.

#### Transformación de algodón

La transformación de algodón (*Gossypium hirsutum* L.) se realiza utilizando Agrobacterium tumefaciens, en explantes de hipocótilos. Los cultivares comerciales tales como Coker 130 o Coker 312 (SeedCo, Lubbock, TX) son variedades estándar utilizadas para la transformación, pero también se pueden utilizar otras variedades. Las semillas se esterilizan en superficie y se germinan en la oscuridad. Los explantes de hipocótilo se cortan de las plántulas germinadas a longitudes de alrededor de 1-1.5 centímetro. El explante de hipocótilo se sumerge en el inóculo de Agrobacterium tumefaciens que contiene el vector de expresión, durante 5 minutos y después se cocultan durante aproximadamente 48 horas en MS + 1.8 mg/l de KNO<sub>3</sub> + 2% de glucosa a 24° C, en la oscuridad. Los explantes se transfieren al mismo medio que contiene marcadores seleccionables bacterianos y de planta apropiados (renovados varias veces), hasta que se ven los callos embriogénicos. Los callos se separan y se subcultivan hasta que aparecen plántulas de embriones somáticas derivadas de embriones somáticos que se maduran en un medio de enraizamiento hasta que se desarrollan las raíces. Los brotes enraizados se trasplantan a tierra para macetas en el invernadero. Las semillas T1 se producen a partir de plantas que exhiben tolerancia al agente de selección y que contienen una única copia de inserción de T-ADN.

#### 45 Ejemplo 8: Procedimiento de Evaluación fenotípica

##### 8.1 Configuración de evaluación

Aproximadamente se generan 35 transformantes independientes T0 de arroz. Los transformantes primarios se transfieren desde una cámara de cultivo de tejidos hasta un invernadero para el cultivo y se cosechan de semillas T1. Seis eventos, de los cuales se retiene la progénesis T1 segregada 3:1 para la presencia/ausencia del transgén. 50 Para cada uno de estos eventos, aproximadamente 10 plántulas T1 que contienen el transgén (hetero y homocigotos) y aproximadamente 10 plántulas T1 que carecen del transgén (nulicigotos) se seleccionan al monitorear la expresión del marcador visual. Las plantas transgénicas y los correspondientes nulicigotos crecen lado a lado en posiciones aleatorias. Las condiciones de invernadero de día corto (12 horas de luz), 28° C en la luz y 22° C en la oscuridad, y una humedad relativa de 70%. A las plantas que crecen en condiciones de no estrés se les proporciona agua a intervalos regulares para asegurar que el agua y los nutrientes no son limitantes para satisfacer las necesidades de la planta para completar el crecimiento y desarrollo.

#### Detección de eficiencia del uso de nitrógeno

Las plantas de arroz de semillas T2 se cultivan en tierra para macetas bajo condiciones normales excepto para la solución de nutrientes. Las macetas se riegan desde el trasplante hasta la maduración con una solución de nutriente específico que contiene contenido de nitrógeno N reducido (N), usualmente entre 7 a 8 veces menos. El resto del cultivo (maduración de planta, cosecha de semillas) es el mismo que para las plantas no cultivadas bajo estrés abiótico. Los parámetros de crecimiento y rendimiento se registran como se detalla para el crecimiento bajo condiciones normales.

### 8.2 Análisis estadístico: prueba F

Se utiliza un ANOVA de dos factores (análisis de variantes) como un modelo estadístico para la evaluación total de las características fenotípicas de la planta. Una prueba F se lleva a cabo en todos los parámetros medidos de todas las plantas de todos los eventos transformados con el gen de la presente invención. La prueba F se lleva a cabo para verificar un efecto del gen sobre todos los eventos de transformación y para verificar un efecto general del gen, también conocido como un efecto de gen global. El umbral de significación para un verdadero efecto génico global se fija en un nivel de probabilidad del 5% para la prueba F. Un valor de prueba F significativo apunta a un efecto del gen, lo que significa que no es únicamente la sola presencia o la posición del gen lo que provoca las diferencias en el fenotipo.

### 8.3 Parámetros medidos

#### Medición de parámetros relacionados con biomasa

Desde la etapa de siembra hasta la etapa de madurez se pasan las plantas varias veces a través de una cabina de imágenes digitales. En cada punto de tiempo las imágenes digitales (2048x1536 píxeles, 16 millones de colores) se toman de cada planta desde por lo menos 6 ángulos diferentes. El área de la planta sobre el suelo (o biomasa foliar) se determina contando el número total de píxeles en las imágenes digitales de partes de la planta sobre el suelo discriminadas desde el fondo. Este valor se promedia de las fotografías tomadas en el mismo punto de tiempo desde los diferentes ángulos y se convierte a un valor superficie física expresada en mm cuadrados mediante calibración. Los experimentos muestran que el área de la planta sobre el suelo medida de esta manera se correlaciona con la biomasa de las partes de la planta por encima del suelo. El área por encima del suelo es el área medida en el punto de tiempo en el que la planta ha alcanzado su máxima biomasa foliar. El vigor temprano se determina al contar el número total de píxeles de partes de la planta sobre el suelo discriminadas desde el fondo. Este valor se promedia de las fotografías tomadas en el mismo punto de tiempo desde diferentes ángulos y se convierte en un valor de superficie física expresada en mm cuadrados mediante calibración. Los resultados descritos adelante son de plantas de tres semanas después de la germinación.

#### Mediciones de parámetros relacionados con semilla

Las panículas primarias maduras se cosechan, cuentan, se colocan en bolsas, que llevan la etiqueta de código de barras y luego se secan durante tres días en un horno a 37° C. Las panículas luego se trillan, y se recogen y se cuentan todas las semillas. Las cáscaras cargadas se separan de las vacías utilizando un dispositivo de soplado de aire. Las cáscaras vacías se desechan y la fracción restante se cuenta de nuevo. Las cáscaras cargadas se pesan en una balanza analítica. El número de semillas cargadas se determina al contar el número de cáscaras cargadas que quedan después de la etapa de separación. El rendimiento total de la semilla se mide al pesar todas las cáscaras cargadas cosechadas de una planta. Se mide el número total de semillas por planta al contar el número de hojas cosechadas de una planta. El peso de Mil Granos (PMG) se extrae del número de semillas cargadas contadas y su peso total. El Índice de Cosecha (HI) en la presente invención se define como la relación entre el rendimiento total de semillas y el área por encima del suelo ( $mm^2$ ), multiplicado por un factor de  $10^6$ . La semilla del índice de relleno tal como se define en la presente invención es la proporción (expresado en %) del número de semillas cargadas sobre el número total de semillas (o floretes).

#### Ejemplo 9: Resultados de la evaluación fenotípica de las plantas transgénicas que comprenden la SEQ ID NO: 1

La evaluación de las plantas de arroz transgénicas que expresan un ácido nucleico LBD en condiciones de no estrés muestra que hubo un aumento de más del 5% de la biomasa por encima de la tierra (AreaMax), rendimiento total de la semilla, número de semillas cargadas, el índice de relleno, la cosecha índice, y más de 3% para el peso de mil granos.

La evaluación de las plantas de arroz transgénicas que expresan un ácido nucleico LBD en la detección de eficiencia del uso de nitrógeno revela un aumento de más de 5% para el vigor de la emergencia (vigor temprano).

#### Ejemplo 10: Resultados de la evaluación fenotípica de las plantas transgénicas que comprenden la SEQ ID NO: 71

La región de codificación comprendida en la SEQ ID NO: 71 se clona bajo el control del promotor GOS2 de arroz en un vector de transformación de arroz como se describe en el Ejemplo 6. Las plantas de arroz transgénicas que comprenden la región codificante de la SEQ ID NO: 71 se generan siguiendo los procedimientos del Ejemplo 7. Las plantas se evalúan de acuerdo con el procedimiento descrito en el Ejemplo 8. La SEQ ID NO: 71 codifica la proteína LBD representada por la SEQ ID NO: 25

5 La evaluación de plantas de arroz transgénicas que expresan la SEQ ID NO: 71 en condiciones de no estrés muestran que hubo un aumento de más del 5% de la biomasa por encima de la tierra (AreaMax), el número de flores por panícula, el número total de semillas por planta, y más del 3% para el peso de mil granos.

**Ejemplo 11:** Resultados de la evaluación fenotípica de las plantas transgénicas que comprenden la SEQ ID NO: 72

10 La región de codificación comprendida en la SEQ ID NO: 72 se clona bajo el control del promotor GOS2 de arroz en un vector de transformación de arroz como se describe en el Ejemplo 6. Las plantas de arroz transgénicas que comprenden la región codificante de la SEQ ID NO: 72 se generan siguiendo los procedimientos del Ejemplo 7. Las plantas se evalúan de acuerdo con el procedimiento descrito en el Ejemplo 8. La SEQ ID NO: 72 codifica la proteína LBD representada por la SEQ ID NO: 52.

15 La evaluación de las plantas de arroz transgénicas que expresan la SEQ ID NO: 72 en condiciones de no estrés muestra que hubo un aumento de más del 5% de la biomasa por encima del suelo (AreaMax), el peso de semilla por planta, número de semillas cargadas, el número de flores por panícula, el número total de semillas por planta, índice de cosecha y más del 3% para el peso de mil granos.

20 La evaluación de las plantas de arroz transgénicas que expresan la SEQ ID NO: 72 en la detección de eficiencia del uso de nitrógeno revela un aumento de más del 5% de la biomasa por encima del suelo (AreaMax).

## REIVINDICACIONES

1. Un método para aumentar el rendimiento de semillas y/o aumentar la biomasa en plantas en relación con plantas de control, que comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido LBD al introducir y expresar en una planta un ácido nucleico que codifica un polipéptido LBD que comprende un dominio DUF260 y uno o más de los siguientes motivos:
- 5 (i) Motivo 1: MSCNGCRXLRKG CX (SEQ ID NO: 5),  
 (ii) Motivo 2: QXXATXF FXAKFXGR (SEQ ID NO: 6),  
 (iii) Motivo 3: FXSLLXEAXG (SEQ ID NO: 7).
2. El método de acuerdo con la reivindicación 1, en donde dicho ácido nucleico que codifica un polipéptido LBD codifica una cualquiera de las proteínas enumeradas en la Tabla A1 o es una porción de dicho un ácido nucleico, o un ácido nucleico capaz de hibridar bajo condiciones rigurosas con dicho un ácido nucleico, en donde dicha porción es por lo menos 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 de nucleótidos consecutivos en longitud.
- 10 3. El método de acuerdo con cualquier reivindicación precedente, en donde dicha secuencia de ácidos nucleicos codifica un ortólogo o parálogo de cualquiera de las proteínas dadas en la Tabla A1.
4. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en donde dicho aumento de rendimiento de semillas y/o biomasa se obtienen bajo condiciones de no estrés.
- 15 5. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en donde dicho aumento de rendimiento de semillas y/o biomasa se obtienen bajo condiciones de deficiencia de nitrógeno.
6. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en donde dicho ácido nucleico se liga operablemente a un promotor constitutivo.
- 20 7. El método de acuerdo con la reivindicación 6, en donde dicho promotor constitutivo es un promotor GOS2.
8. El método de acuerdo con cualquier reivindicación precedente, en donde dicho ácido nucleico que codifica un polipéptido LBD es de origen de planta, preferiblemente de una planta dicotiledónea, adicional y preferiblemente de la familia Brassicaceae, más preferiblemente del gen *Arabidopsis*, más preferiblemente de *Arabidopsis thaliana*.
- 25 9. Una molécula aislada de ácido nucleico que comprende una cualquiera de las siguientes características:
- i) un ácido nucleico representado por la SEQ ID NO: 69;  
 (ii) un ácido nucleico que es complementario a la SEQ ID NO dado en (i);  
 (iii) un ácido nucleico que codifica un polipéptido LBD que tiene por lo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia con la SEQ ID NO: 70;  
 (iv) un ácido nucleico capaz de hibridar bajo condiciones rigurosas a uno cualquiera de los ácidos nucleicos dados en (i), (ii) o (iii) anteriormente.
- 30 10. Un polipéptido LBD aislado que comprende:
- 35 (i) una secuencia de aminoácidos que tiene por lo menos 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos dada en la SEQ ID NO: 70.
11. Planta o parte de la misma, que incluye semillas, que se pueden obtener mediante un método como se define en cualquiera de las reivindicaciones 1 a 8, en donde dicha planta o parte de la misma comprende un ácido nucleico recombinante que codifica un polipéptido LBD introducido en dicha planta y en donde dicha planta se selecciona de transformantes homocigotos y en donde dicho polipéptido LBD comprende un dominio DUF260 y uno o más de los siguientes motivos:
- 40 (i) Motivo 1: MSCNGCRXLRKG CX (SEQ ID NO: 5),

(ii) Motivo 2: QXXATXFXAKFXGR (SEQ ID NO: 6),

(iii) Motivo 3: FXSLLXEAXG (SEQ ID NO: 7).

12. Construcción que comprende:

(i) ácido nucleico que codifica un polipéptido LBD como se define en la reivindicación 1 o 10;

5 (ii) una o más secuencias de control capaces de conducir la expresión de la secuencia de ácidos nucleicos de (a); y opcionalmente

(iii) una secuencia de terminación de transcripción.

13. Construcción de acuerdo con la reivindicación 12, en donde una de dichas secuencias de control es un promotor GOS2.

10 14. Uso de una construcción de acuerdo con la reivindicación 12 o 13 en un método para hacer plantas que tienen aumento de biomasa y/o aumento de rendimiento de semillas en relación con plantas de control.

15. 15. Planta, parte de planta o célula de planta transformada con una construcción de acuerdo con la reivindicación 12 o 13.

16. 15. Método para la producción de una planta transgénica que tiene aumento de biomasa y/o aumento de rendimiento de semillas en relación con plantas de control, que comprende:

(i) introducir y expresar en una planta un ácido nucleico que codifica un polipéptido LBD como se define en la reivindicación 1 o 10; y

(ii) cultivar la célula de planta bajo condiciones que promueven el crecimiento y desarrollo de la planta.

20 17. Planta transgénica que tiene aumento de biomasa y/o aumento de rendimiento de semillas, en relación con plantas de control, que resultan de la aumento de expresión de un ácido nucleico que codifica un polipéptido LBD como se define en la reivindicación 1 o 10 introducido en dicha planta, o una célula de planta transgénica derivada de dicha planta transgénica, y en donde dicha planta se selecciona de transformantes homocigotos.

25 18. Planta transgénica de acuerdo con la reivindicación 11, 15 o 17, o una célula de planta transgénica derivada de la misma, en donde dicha planta es un cultivo o una monocotiledónea o un cereal, tal como arroz, maíz, trigo, cebada, maíz, centeno, triticale, sorgo y avena.

19. 19. Partes cosechadas de una planta de acuerdo con la reivindicación 17, en donde dichas partes cosechadas son preferiblemente biomasa de brote y/o semillas y en donde dichas partes cosechadas comprenden una construcción de acuerdo con la reivindicación 12 o 13.

ES 2 437 621 T3

CLUSTAL W (1.83) alineación de secuencia múltiple

MS CNG CRV LRG CS EN CLR P C IQ WI ET DA Q GH AT VFA K FF G R A GL M SF IS A V P DS  
QRP AL F Q S L L Y EA C G RT V NP VNG A G M L W TG N W N C Q A A V E T V L RG G S L R P I E L T H  
1 2

Os-0501903890	St-TC-62339
72s-TC-235711	
8W-TC-143806	
Os-0501932270	
At-05042-QIC430	
Pt-121-85	
At-128040-Q28996	
Bn-NA3391970	
At-128041-QHHA86	
He-TC3-070921	
St-TC3-070921	
Gn-TC1208086	
He-AB57878139	
Vv-TC55321	
Aq-TC18229	
Pt-01010	
Pt-01111-481	
Pt-TC67269	
Pta-TC67269	
Os-0501932090	
Sb-TC101758	
Qa-0501960000	
Hv-TC147716	
Os-0501941330	
St-TC130322	
St-TC137193	
St-TC133385	
La-TC176827	
Nb-TC2406	
Gn-TC216138	
Lj-TC215459	
Gn-TC217562	
Ck-100-00777782	
Wt-NA8822505	
Pt-111-569	
Pt-9 543	
Wt-PEC00568	
Qh-TC100552	
At-00007-097911	
At-12038-Q5823	
At-12039-Q5823	
Aq-TC18561	
Fr-TC559179	
St-TC140088	
Fr-TC14656	
Fr-TC67974	
Pt-KIV-94	
Pt-11105	

FIGURA 1

FIGURA 2

**FIGURA 2** (continuación)

FIGURA 2 (continuación)

ES 2 437 621 T3

FIGURA 2 (continuación)

FIGURA 2 (continuación)

**FIGURA 2** (continuación)

FIGURA 2 (continuación)

ES 2 437 621 T3

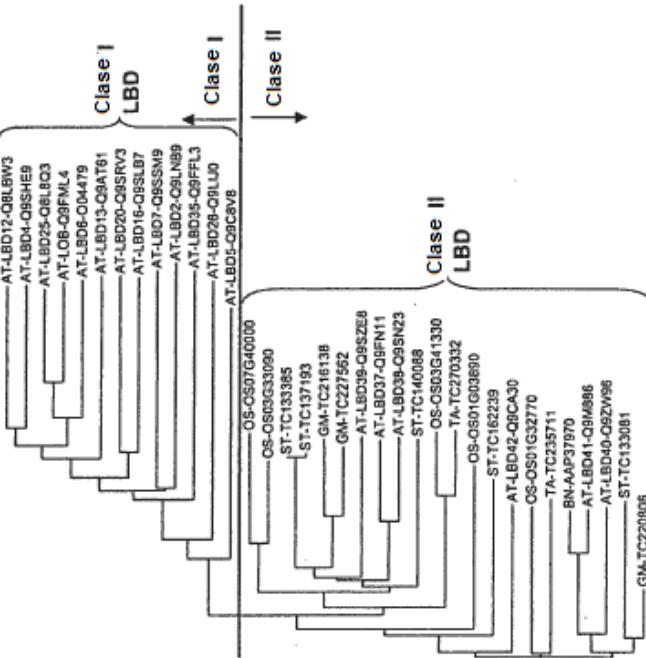


FIGURA 3

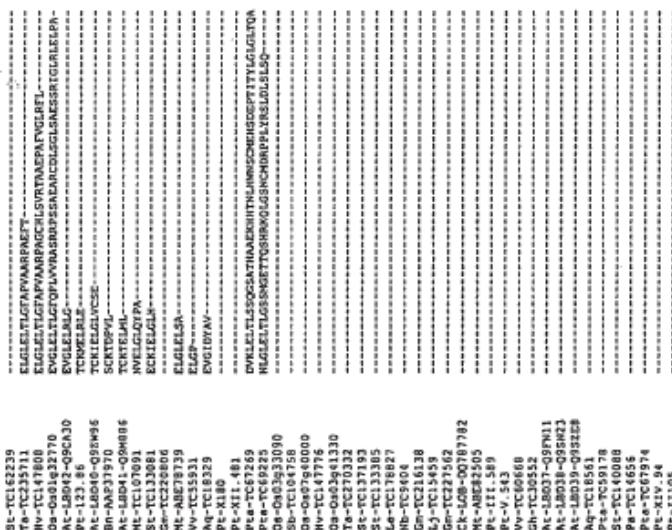


FIGURA 2 (continuación)

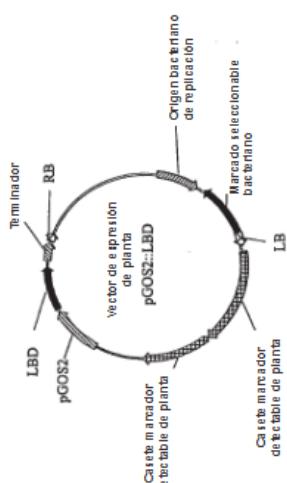


FIGURA 4

FIGURA 5

CATATATTATATATATTTATAGTTTGCAATTGGCATATCGCACATCATTAAGGMCATGGCTTA  
CTCACCCATTTCATTATTTATTAAGGCAATGATTTTATTTATTCAT  
ATTAGTCAGAGTCAGTCATTCAGCTGCTGACTGCACTGCATCTCCAT  
ACAGTCACACTAGCAACACATCTCATATATCAGTCATAGTCAGATAC  
TGAATTCAGACACTCAGCATCACCAAGACACTTTCATATATCATAAA  
ACAGATAGTCAGTAAAGATAGTCAGAACAGACACTTTCATACACACAA  
ATGTCAGTCAGAACAGACACTTTCATACACAGACACTTTCATACACAA  
GCAACACCAACAAACAGATAGTCAGAACAGACACTTTCATACACAA  
GCTTTCGGCCAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG  
ATTCTCCCCCCTTCCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCC  
ATTCTCCCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCC  
GACACCGGACTAGCAGAACGACCCATCTCCATCCATCCATCCATCC  
CGATCTCTCCCTCTCCACCTCCACCTCCACCTCCACCTCCACCTCC  
ATGTCCTAGGTTGGTAGTCAGGGGGCTCTGGTTGGGGGGGGGGGGGG  
ATGGTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  
TTGGTAGTACCTTGTGGTAGTAGGGTAAATATCAGGGCAGGGGGGGGG  
GG  
CCTATGGACACAAATAACCAACTGAGGAGGAGGAGGAGGAGGAGGAG  
ANGCTCTCCAAATTTCAGGTTGGTAGTCAGGGGGCTCTGGTTGGGGGG  
AACAGTAAATTCCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
TTTCCCAAAATTCCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG  
CTTTATAGGGTAACTCTAGGGTAACTCTAGGGTAACTCTAGGGTAACT  
AGACCTTATCGGATTCGATCTGATCTGATCTGATCTGATCTGATCTG  
CATTATGGACACAAATAACCAACTGAGGAGGAGGAGGAGGAGGAGGAG  
CTGGCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT  
AGAGGTTCTTGTGTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  
GGATAGTTATCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  
TTCACCCAGAACATC

SEQ ID NO: 11, Os-0803941330, *Oryza sativa*

MRMSNCGRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
RSLLYEAERGMMLNPYGSVGLLWGNHQLQCAQVSEAAAQTPPLRPAFV  
RAGDVGRRAGPVALRHLDSIRAKERKRPGGAAMAHRSOFLVSEKAA  
LNWAPGALSHESASDAAAPPESEGHSHDADTYDGSYHOSOPEPRA  
PPPPIPQKTEPADSSGGSQQHHRKRPVPLGLAISTKAAQ

SEQ ID NO: 12, Os-0801932710, *Oryza sativa*

MRMSNCGRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
RSLLYEAERGMMLNPYGSVGLLWGNHQLQCAQVSEAAAQTPPLRPAFV  
RHSKIDAEADAAANLVRGGTRFKAASSNSNSKIGLTLAAGAARAA  
VYVGDODHHHPALSHEVHEESAGSIDDDDDYDQNDNDIAIDVTP  
SQAEQSPVPPVPEHEEGEEEVGLLTGQPLVYRASSRPPSAEAR  
DLSLQGSAESSRIGRLPA

SEQ ID NO: 13, Os-0803933080, *Oryza sativa*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLATGNNPQAAVETVRLGGAGP  
GWSITFTAKVRKAEVPEAPSCDGLCILSPGSPPAVGERKPA  
LNLFLV

SEQ ID NO: 14, Os-0803941330, *Oryza sativa*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLATGNNPQAAVETVRLGGAGP  
AGTSSSPONDINSASAAPRERPRUNGGAGGGVQQLPHAVAV  
GGGGASDEPSTTCEASDGGAGATPLMLFS

SEQ ID NO: 15, Os-0807940000, *Oryza sativa*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLGGSHLICQAVADVLRLGGG  
MNGCSTFSTAKRSTTTTNNGTPTDAAAAPQPEPSCDGL  
SHNSGGSVTIGVUVDGEREPELNLFLV

SEQ ID NO: 16, Ag-TC9404, *Nicotiana benthamiana*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLWGNHVCQAVETVRLGGAGL  
QHPSQNNIQTMQKRRREPEASVNLQQLADLSLIPFQS  
TTTCPSGTTGHNQNEPKLILFLV

SEQ ID NO: 17, Gm-TC271562, *Glycine max*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLWGNHVCQAVETVRLGGAGL  
EVGCTDTRTRTPNTPNCTFTSRSSGGGKBRSEELAKLQE  
ESPWTSESEGTTEENRGDQNSHWRVNLFLV

SEQ ID NO: 18, Gm-TC271618, *Glycine max*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLWGNHVCQAVETVRLGGAGL  
CTDKURVRDNPNEFRPQSRDQSGGGKBRKRFELAKLQ  
TRHPSPSNNESEGGTTACLESGIDYHADGDKRVLFLV

SEQ ID NO: 19, Hv-TC147776, *Hordeum vulgare*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLWGNHVCQAVETVRLGGAGL  
RRAVGCSTYSHAKRVTFRKTAPEASHHQEPSCDGL  
GSVTTAGAGDAGKUEPELNLFLV

SEQ ID NO: 20, Sb-TC104758, *Sorghum bicolor*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLWGNHVCQAVETVRLGGAGL  
PTGTSITSDDSVTTTGTGGDREPELNLFLV

SEQ ID NO: 21, Ag-TC10561, *Aquilegia sp.*  
MSCNGCRVLRKGCGSEACTLRCPLQWVTPAQRPAFV  
LLEEAERGRTINPVGAVGLLWGNHVCQAVETVRLGGAGL  
TTOQPTYRPKMHHKEKEIHENYNNKTHDOLSLTS555P  
SGAGYWRDHHYLRSGEDNNNINGDQKLNLFLV

FIGURA 5 (continuación)

FIGURA 5 (continuación)

SEQ ID NO: 22, <i>Vv-Tc60668</i> , <i>Vitis vinifera</i> MSNCNGCRVLRKGSESCILRPECLQWIESAQSGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT SHNSEBSEVTTCPETPKPSGERKLNLFLF	SEQ ID NO: 29, <i>Vv-Tc55931</i> , <i>Vitis vinifera</i> MSNCNGCRVLRKGSESCILRPECLQWIESAQSGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIDTAAEQGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIDTAAEQGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT
SEQ ID NO: 23, <i>Lj-Tc15459</i> , <i>Lotus japonicus</i> MSNCNGCRVLRKGSESCILRPECLQWIDTAAEQGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT PGPSMNSSESVTTACILGSGIDHCSRDEDRKVILKLF	SEQ ID NO: 30, <i>St-Tc162239</i> , <i>Solanum tuberosum</i> MSNCNGCRVLRKGSESCILRPECLQWIDTAAEQGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIDTAAEQGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT
SEQ ID NO: 24, <i>Gri-Tc30552</i> , <i>Gossypium hirsutum</i> MSNCNGCRVLRKGSESCILRPECLQWIESPEAQSGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIESPEAQSGHATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT	SEQ ID NO: 31, <i>Pta-Tc69225</i> , <i>Pinus taeda</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT
SEQ ID NO: 25, <i>Ta-Tc23571</i> , <i>Triticum aestivum</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT	SEQ ID NO: 32, <i>Pta-Tc67269</i> , <i>Pinus taeda</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT
SEQ ID NO: 26, <i>St-Tc133081</i> , <i>Solanum tuberosum</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT	SEQ ID NO: 33, <i>Gm-Tc220806</i> , <i>Glycine max</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT
SEQ ID NO: 27, <i>St-Tc107091</i> , <i>Medicago truncatula</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT	SEQ ID NO: 34, <i>Ta-Tc270332</i> , <i>Triticum aestivum</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT
SEQ ID NO: 28, <i>Hy-Tc147808</i> , <i>Hordeum vulgare</i> MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT	SEQ ID NO: 35, <i>Ag-Tc18329</i> , <i>Aquilegia</i> sp. MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT MSNCNGCRVLRKGSESCILRPECLQWIKTSSEQAQATVYAKFYGRAGLMLNINAGPEHLRPAIF LILYACGTVNPNGAVLGLTGNWICQAAETVAKFYGRAGLMLNINAGPEHLRPAIF TNEFKLROPNSSSRSKPKRKRLEAKVQCTDILRFLPLGKISRGIGMSYLPETRRPGT

FIGURA 5 (continuación)

SEQ ID NO: 36, St-TC137193, *Solanum tuberosum*  
 MSCNGCRVLRKGSDNCILRPLQHNTETPENQHATVAKFFRAGLMSFISAVPENORPALSQF  
 LLYEAQGRTVNPNGAVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 QDPSLNHBPKMKQKRREPEETSMLAIDLSDLTGPNQKVNYSPLAENHRRGTPSPMSNSEESEGTIT  
 CFESTAVIGEEPKLSSLFN

SEQ ID NO: 37, St-TC13386, *Solanum tuberosum*  
 MSCNGCRVLRKGSDNCILRPLQHNTETPENQHATVAKFFRAGLMSFISAVPENORPALSQF  
 LLYEAQGRTVNPNGAVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 QDPSLNHBPKMKQKRREPEETSMLAIDLSDLTGPNQKVNYSPLAENHRRGTPSPMSNSEESEGTIT  
 CFESTAVIGEEPKLSSLFN

SEQ ID NO: 44, Pt-v-543, *Populus tremuloides*  
 MSCNGCRVLRKGCSENCILRPLQHNTPEAOGHATVAKFFRAGLMSFISAVPENORPALSQF  
 LLYEAQGRTVNPNGAVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 EVACIDIKLQDNPBNP1HHSRFSNSRSRSVSKRKGTEEPVAVNVRHDDDLIUTPSTSQKGEAEI  
 CRPGTPSNHNSAESVTNTACFDSAGLGEDEGTKLNLFL

SEQ ID NO: 45, Pt-XIV.94, *Populus tremuloides*  
 MSCNGCRVLRKGCSSETCTV1RSLCHHHTPEAOGNATLFLAKFFGRAGLMSFISAVPENORPALSQF  
 LLYEAQGRTVNPNGAVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 AVASDMSMTHNQSKFDKNNQVSDHVNCLARSGRDKRGRDAYSFTYGESETTSESSGA  
 DRKLNLFLV

SEQ ID NO: 46, Pt-XI105, *Populus tremuloides*  
 MSCNGCRVLRKGCSSETCTV1RSLCHHHTPEAOGNATLFLAKFFGRAGLMSFISAVPENORPALSQF  
 LLYEAQGRTVNPNGAVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 AVAVHNMARNQSRFAKENNKEYVSDRINCLARSGRDKRGRDAYSFTYGESETTSESSGA  
 DRKLNLFLV

SEQ ID NO: 47, Pt-123.86, *Populus tremuloides*  
 MMSNGCRVLRKGCSENCILRPLQHNTPEAOGHATVAKFFRAGLMSFISAVPENORPALSQF  
 RSLLYEAQGRIVNPYGVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 VSRLAVIDGSKEDESAVPSTAEALPFPNPNESKNICKAHVASHGTSIGELGLDTLQLEPV  
 SRAHIVPVKVRKRVAYGSQDVTCKHRLV

SEQ ID NO: 48, Pt-X180, *Populus tremuloides*  
 MSCNGCRVLRKGCSENCILRPLQHNTPEAOGHATVAKFFRAGLMSFISAVPENORPALSQF  
 RSLLYEAQGRIVNPYGVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 SRDQNSPELNUKVXERTRKSTGTRPSSLAETYGRNGNOEGELFEPRESMWLGNGDSRIMEDS  
 IIAAENVEDLSNSRVEPDVKFENGQDDSDLSLTLAIVE

SEQ ID NO: 49, Pt-XI.481, *Populus tremuloides*  
 MPPSALSFSLPHHITLILIGTNSSPSLRHFSTISFRTEYTMMSNGCR1LRKGODGNSIKPC  
 LOWIEPDSDSANATLFLAKFFRAGLMSFISAVPENORPALSQF  
 LLYEAQGRIVNPVIGSVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 GTRPSSLAETYGRNGNOEGELFEPRESMWLGNGDSRIMEDS  
 DLDLTGHY

SEQ ID NO: 50, CK-10B-DQ787782, *Carex sibirica*  
 MSCNGCRVLRKGCSENCILRPLQHNTPEAOGHATVAKFFRAGLMSFISAVPENORPALSQF  
 LLYEAQGRTVNPNGAVGLLWTGNNHVCQAVENTVLRGALRITSEFAGASWEIDEVSCTDVFKL  
 QDPSLNHBPKMKQKRREPEETSMLDLSLTGPNQKVNYSPLAENHRRGTPSPMSNSEESEGTIT  
 SSAGVGDQGKREPKLSSLFN

FIGURA 5 (continuación)

FIGURA 5 (continuación)

SEQ ID NO: 51, *Brassica napus*  
 MRSNCNGCRVLRKGCSNCSIRPLQMLTNSQVGLWNSHINWOLCOSAENVNNGEPTTENMDATTSQGPPLHMDRIF  
 RSLLHEACGRIVNPYIGCAGCAGGTTGCGACGCGTGGCGAGGTTGCTGGCAAGGGGTGACGGGACAAACNGCGCCATCGGC  
 IADENS PARAAGAGS TDRKGTRRKA RAVAAVAKPASGGGEAASHISLUSHQSEVVAFHEGESK  
 EESNSNSEWHTSPPAVQGSEIEKLDTLUGLEPVSRADHVV PVVKRMGVFSTHQEESSEKTD PVL

SEQ ID NO: 52, *Ms-ABE02505, Medicago truncatula*  
 M\$CNGCRVLRKGCSSECTIRPLQMLTPEA VAKFGRAGLMSFISI NVPPEPQRPALEQS  
 LIEEACGRIVNPYNGAVGLMTGNNHVCORAETVLRGSLTRELPELTLI DAPLTADESEAEDG  
 GADMNMTREPNFLASSRSNSNKVCURSRSRSLPQSSKRRVPAIAINLRLTIPFQOKAVERRHGS P  
 SMTSESVTTTACLETGIGSRMWSHGDRKVANLET

SEQ ID NO: 53, *Ms-ABE78739, Medicago truncatula*  
 M\$CNGCRVLRKGCSIRPLQMLTPEA VAKFGRAGLMSFISI NVPPEPQRPALEQS  
 RSLLHEACGRIVNPYIGCAGGTTGCGAGCAGGTTGCGAGCAGGTTGCGAGCAGGTTGCGAGCAGGTTGCGAGCAGGTTGCGAGC  
 VSADENLLETKVORTSRSRFRAMMKPKSEVSGSISGSSSAKEVDRSISQESLISQPKDSEGVSV  
 ETILIFDSESEPEPEQAKYVGRAGRAEGLSDGVLERLIGEPIVSRETFMVPIKRRKTVLDC  
 SASSKTVLGLLESA

SEQ ID NO: 54, *At-LBD38-09SN23, Arabidopsis thaliana*  
 M\$CNGCRVLRKGCSENCSIRPLQMLTPEA VAKFGRAGLMSFISI NVPPEPQRPALEQS  
 LIEEACGRIVNPYNGAVGLMTGNNHVCORAETVLRGSLTRELPELTLI DAPLTADESEAEDG  
 ICTEMLNURKAODSGDRNITHCREFSSRSRSLPQSSKRRVPAIAINLRLTIPFQOKAVERRHGS P  
 LPPKEDTPSMYSESESTTVSFSNNNAGDRYVRCGGGGGATKMLNLFNA

SEQ ID NO: 55, *At-LBD39-09S28, Arabidopsis thaliana*  
 M\$CNGCRVLRKGCSSECTIRPLQMLTPEA VAKFGRAGLMSFISI NVPPEPQRPALEQS  
 LIEEACGRIVNPYNGAVGLMTGNNHVCORAETVLRGSLTRELPELTLI DAPLTADESEAEDG  
 ICPNQTHCREFSSRSRSLPQSSKRRVPAIAINLRLTIPFQOKAVERRHGS P  
 KGRGPGEPESESVTTWCEWNGRGNKQERKKEKKMLNLFNA

SEQ ID NO: 56, *At-LBD40-092W96, Arabidopsis thaliana*  
 M\$CNGCRVLRKGCSNCSIRPLQMLTNSQVGLWNSHINWOLCOSAENVNNGEPTTENMDATTSQGPPLHMDRIF  
 RSLLHEACGRIVNPYIGCAGGTTGCGACGCGTGGCGAGGTTGCTGGCAAGGGGTGACGGGACAAACNGCGCCATCGGC  
 RYVSSRDENYBRSRSEACKERNUVSLISHESSLSHESPVSSSETTIEBKTWIGELTLEPLAR  
 GHNVWVPPMKRNLERCTSEDEDTCTIEGLVCE

SEQ ID NO: 57, *At-LBD41-09MBB6, Arabidopsis thaliana*  
 M\$CNGCRVLRKGCSEDCSIRPLQMLTPEA VAKFGRAGLMSFISI NVPPEPQRPALEQS  
 RSLLHEACGRIVNPYIGCAGGTTGCGACGCGTGGCGAGGTTGCTGGCAAGGGGTGACGGGACAAACNGCGCCATCGGC  
 ISDINPRAAATGSTDLKAKTRRKA RVSQDVTAOESEIGSKSDEA SHESSTSHOSITVAHEGESEK  
 SESNSNVEVLA FSPPAVNGSGEIEKLDTLUGLEPVSRADHVV PVVKRMGVFSTHQEESSEKTD PVL

SEQ ID NO: 58, *At-LBD42-09CA30, Arabidopsis thaliana*  
 M\$CNGCRVLRKGCSNCSIRPLQMLTNSQVGLWNSHINWOLCOSAENVNNGEPTTENMDATTSQGPPLHMDRIF  
 RSLLHEACGRIVNPYIGCAGGTTGCGACGCGTGGCGAGGTTGCTGGCAAGGGGTGACGGGACAAACNGCGCCATCGGC  
 RYVAKDPPGCGSSELA TRVNANNSKTOGTRKRESYNOLGE SHMMWLPSSSATHGTYHFTLE  
 NVESEREAPFNOSSPNLGFDDQVDNEVGELELRG

FIGURA 5 (continuación)

FIGURA 5 (continuación)

FIGURA 5 (continuación)

SEQ ID NO: 70, *Linum usitatissimum*  
MIRELRAVIPMDNSISSTROPLFLAKFGRSISDLSSTISSVDPORPALFOSLIFCAGRTVNP  
VNGAVGULLMSRHHNCQAVETVLAGSISLHTPGAAVGAAPNFDODSDDCSAAGFANOLRPAIIL  
ATGCGNRLNLSSSS1SEYRRHMA1VGGHERRNNRDA5FFAESETTTESTGDRKLIN  
LTV

FIGURA 5 (continuación)

FIGURA 5 (continuación)