

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 440 265**

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: **21.07.2008 E 08775243 (2)**

97 Fecha y número de publicación de la concesión europea: **18.09.2013 EP 2173884**

54 Título: **Plantas que tienen un aumento de características relacionadas con el rendimiento y un método para elaboración de las mismas**

30 Prioridad:

20.07.2007 EP 07112908

20.07.2007 EP 07112902

20.07.2007 EP 07112903

27.07.2007 EP 07113319

05.09.2007 US 970065 P

06.11.2007 US 985688 P

12.11.2007 US 987252 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.01.2014

73 Titular/es:

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

72 Inventor/es:

**HATZFELD, YVES y
FRANKARD, VALERIE**

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 440 265 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 un aumento de características relacionadas con el rendimiento y un método para elaboración de las mismas.

5 La presente invención se refiere en general al campo de la biología molecular y se relaciona con un método para incrementar diferentes características de la planta relacionadas con el rendimiento, mediante el incremento de la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido que incrementa el rendimiento, por lo cual el polipéptido que incrementa el rendimiento es uno 19/20 localizado en el núcleo con el motivo AT-hook (AHL 19/20). La presente invención también se relaciona con plantas que tienen una mayor expresión de una secuencia de ácido nucleico que codifica dicho polipéptido que incrementa el rendimiento, en donde dichas plantas tienen más características relacionadas con el rendimiento con respecto a las plantas que sirven de control. La invención también provee construcciones útiles en los métodos de la invención.

15 La población mundial siempre en crecimiento y la oferta cada vez más escasa de tierras cultivables disponibles para la agricultura impulsan la investigación para lograr una mayor eficiencia de la agricultura. Los medios convencionales para la mejora hortícola y de los cultivos utilizan técnicas selectivas de fitomejoramiento para identificar las plantas que tienen rasgos deseables. Sin embargo, dichas técnicas selectivas de fitomejoramiento tiene varias desventajas, a saber, que estas técnicas son típicamente de trabajo intensivo y traen como resultado plantas que contienen a menudo componentes genéticos heterogéneos que no siempre resultan en la transmisión de la característica deseable a partir de las plantas progenitoras. Los avances en biología molecular le han permitido a la humanidad modificar el germoplasma de animales y de plantas. La modificación por ingeniería genética de las plantas supone el aislamiento y manipulación de material genético (típicamente en la forma de ADN o ARN) y la posterior introducción de ese material genético en una planta. Dicha tecnología tiene la capacidad de proporcionar cultivos o plantas que tienen diferentes características económicas, agronómicas u hortícolas mejoradas.

25 Una característica de particular interés económico es un mayor rendimiento. El rendimiento normalmente se define como la producción medible de valor económico de un cultivo. Esto puede ser definido en términos de cantidad y/o de calidad. El rendimiento depende directamente de varios factores, por ejemplo, de la cantidad y tamaño de los órganos, de la arquitectura de la planta (por ejemplo, la cantidad de ramas), de la producción de semillas, de la senectud de la hoja y más. El desarrollo de la raíz, la captación de nutrientes, la tolerancia al estrés y el vigor inicial también pueden ser factores importantes para determinar el rendimiento. Por lo tanto, la optimización de uno o más de los factores mencionados anteriormente puede contribuir al incremento del rendimiento de los cultivos.

30 El rendimiento de semillas es una característica particularmente importante, ya que las semillas de muchas plantas son importantes para la nutrición humana y animal. Cultivos tales como maíz, arroz, trigo, canola y soja dan cuenta de más de la mitad de la ingesta total de calorías por parte de la humanidad, ya sea a través del consumo directo de las semillas en si mismas o a través del consumo de productos alimenticios producidos con semillas procesadas. Son además una fuente de azúcares, aceites y muchas clases de metabolitos utilizados en procesos industriales. Las semillas contienen un embrión (la fuente de nuevos brotes y raíces) y un endospermo (la fuente de nutrientes para el desarrollo del embrión durante la germinación y durante el desarrollo temprano de las plántulas). El desarrollo de una semilla involucra muchos genes, y requiere de la transferencia de metabolitos a partir de las raíces, hojas y tallos dentro de la semilla en desarrollo. El endospermo, en particular, asimila los precursores metabólicos de carbohidratos, aceites y proteínas y los sintetiza en macromoléculas de almacenamiento para llenar el grano.

45 El índice de cosecha, la relación del rendimiento de semillas con respecto al peso seco por encima del suelo, es relativamente estable bajo muchas condiciones ambientales y de este modo se puede obtener una muy buena correlación entre el tamaño de la planta y el rendimiento de granos (por ejemplo, Rebetzke et al., 2002, Crop Science 42: 739). Estos procesos están intrínsecamente enlazados porque la mayor parte de la biomasa del grano depende de la productividad fotosintética presente o almacenada por parte de las hojas y el tallo de la planta (Gardener et al., 1985, Physiology of Crop Plants. Iowa State University Press, páginas 68 - 73). Por lo tanto, la selección del tamaño de la planta, incluso en etapas tempranas de desarrollo, ha sido utilizada como un indicador del rendimiento potencial futuro (por ejemplo, Tittonell et al., 2005, Agric Ecosys & Environ 105: 213). Cuando se analiza el impacto de las diferencias genéticas sobre tolerancia al estrés, la capacidad de estandarizar las propiedades del suelo, la temperatura, la disponibilidad de agua y de nutrientes y la intensidad de la luz es una ventaja intrínseca del invernadero o de los ambientes con cámaras de crecimiento de plantas en comparación con el campo. Sin embargo, las limitaciones artificiales en el rendimiento debido a una pobre polinización provocada por la ausencia de viento o de insectos, o insuficiente espacio para la maduración de la raíz o el desarrollo del dosel, pueden restringir el uso de estos ambientes controlados para analizar las diferencias de rendimiento. Por lo tanto, las mediciones del tamaño de la planta en el desarrollo temprano, bajo condiciones estándar en una cámara de crecimiento o invernadero, son prácticas estándar para proporcionar una indicación de las ventajas potenciales de rendimiento genético.

5 Otra característica importante es aquella de la tolerancia mejorada al estrés abiótico. El estrés abiótico es una de las causas principales de pérdida de cultivos en el mundo, reduciendo los rendimientos promedio para la mayoría de las plantas de cultivo en más del 50% (Wang et al. (2003) *Planta* 218: 1 - 14). Los estreses abióticos pueden ser provocados por sequía, salinidad, extremos de temperaturas, toxicidad química, exceso o deficiencia de nutrientes (macroelementos y/o microelementos), radiación y estrés oxidativo. La capacidad de incrementar la tolerancia de una planta al estrés abiótico sería una ventaja económica importante para los granjeros a nivel mundial y permitiría el cultivo de granos en condiciones adversas y en territorios donde el cultivo no puede ser posible de otro modo.

10 Otra característica importante para muchos cultivos es el vigor inicial. La mejora del vigor inicial es un objetivo importante de los programas modernos de fitomejoramiento de arroz tanto en zonas de cultivo templadas como tropicales de arroz. Raíces largas son importantes para el anclaje apropiado del suelo en arroz sembrado en agua. Cuando el arroz se siembra directamente en campos inundados, y donde las plantas deben emerger rápidamente a través del agua, los brotes más largos están asociados con el vigor. Donde se practica el cultivo por medio de sembradoras, mesocotilos y coleoptilos más largos son importantes para una buena aparición de las plántulas. La capacidad para modificar por ingeniería genética el vigor inicial en las plantas sería de gran importancia en la agricultura. Por ejemplo, un pobre vigor inicial ha sido una limitación para la introducción de híbridos de maíz (*Zea mays* L.) con base en el germoplasma del Cinturón de Maíz en el Atlántico Europeo.

20 Otra característica importante adicional es aquella de las características mejoradas relacionadas con el rendimiento de plantas cultivadas bajo condiciones de estrés abiótico. El estrés abiótico es una de las causas principales de pérdida de cultivos en el mundo, reduciendo los rendimientos promedio para la mayoría de las plantas de cultivo en más del 50% (Wang et al. (2003) *Planta* 218: 1 - 14). Los estreses abióticos pueden ser provocados por sequía, salinidad, extremos de temperaturas, toxicidad química, exceso o carencia de nutrientes (macroelementos y/o microelementos), radiación y estrés oxidativo. La capacidad para mejorar las características relacionadas con el rendimiento de plantas cultivadas bajo condiciones de estrés abiótico sería una ventaja económica importante para los granjeros a nivel mundial y permitiría la siembra de cultivos en condiciones adversas y en territorios donde el cultivo no puede ser posible de otro modo.

25 Se puede incrementar por lo tanto el rendimiento de un cultivo por medio de la optimización de uno de los factores anteriormente mencionados.

30 Dependiendo del uso final, se puede favorecer a modificación de ciertas características de rendimiento con respecto a otras. Por ejemplo para aplicaciones tales como la producción de forraje o madera, o recursos de biocombustibles, puede ser deseable un incremento en las partes vegetativas de una planta, y para aplicaciones tales como la producción de harina, almidón, o de aceite, puede ser particularmente deseable un incremento en los parámetros de la semilla. Incluso entre los parámetros de la semilla, algunos pueden verse favorecidos sobre otros, dependiendo de la aplicación. Diferentes mecanismos pueden contribuir para incrementar el rendimiento de semilla, ya sea en la forma de mayor tamaño de la semilla o el incremento en el número de semillas.

35 Un enfoque para incrementar las características relacionadas con el rendimiento (rendimiento de semilla y/o de biomasa) en plantas puede ser a través de la modificación de los mecanismos inherentes al crecimiento de una planta, tal como el ciclo celular o diferentes rutas de señalización involucradas en el desarrollo de la planta o en mecanismos de defensa.

40 Ahora se ha encontrado que se pueden incrementar diferentes características relacionadas con el rendimiento en las plantas con respecto a plantas que sirven de control, sin retraso de la floración, por medio del aumento de la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido 19/20 localizado en el núcleo con el motivo AT-hook (AHL 19/20). El aumento de características relacionadas con el rendimiento de semillas comprende uno o más de: aumento del número de flores por panícula, aumento del rendimiento total de semillas por planta, aumento de la cantidad de semillas llenas, y aumento del índice de cosecha.

45 Antecedentes

50 Las proteínas que se enlazan al ADN son proteínas que contienen cualquiera entre muchos dominios de enlazamiento al ADN y por lo tanto tienen una especificidad o afinidad general por el ADN. Las proteínas que se enlazan al ADN incluyen por ejemplo factores de transcripción que modulan el proceso de transcripción, nucleasas que rompen las moléculas de ADN, e histonas que están involucradas en el empaquetamiento del ADN en el núcleo de la célula.

55 El motivo AT-hook es un motivo corto de proteína que se enlaza al ADN que fue descrito primero en las proteínas cromosómicas diferentes a la histona del grupo de alta movilidad, HMG- I/Y (Reeves y Nissen (1990) *J Biol Chem* 265: 8573 - 8582). Se sabe que AT-hook interactúa con la ranura menor de las secuencias de ácido nucleico ricas en AT (Huth et al. (1997) *Nat Struct Biol* 4: 657 - 665). Los motivos AT-hook han sido identificados en una gran variedad de proteínas que se enlazan al ADN de animales, plantas y microorganismos. A diferencia de varios

motivos de enlazamiento al ADN bien caracterizados, el motivo AT-hook es corto, hasta de 13 residuos de aminoácidos, y tiene una secuencia tripeptídica típica con una glicina- arginina- prolina (Gly- Arg- Pro o GRP) en su centro.

5 En *Arabidopsis thaliana*, aproximadamente 30 polipéptidos, que contienen al menos un motivo AT-hook, contienen además un dominio conservado de una planta y de procariotas (PPC), que se describe como DUF296 (dominio de función desconocida 296) en la base de datos de dominios InterPro del European Bioinformatics Institute (EBI) (Fujimoto et al. (2004) Plant Molec Biol 56: 225 - 239). Se encontró que una de estas proteínas estaba localizada en el nucleoplasma, y por lo tanto llamada proteína 1 localizada en el núcleo con el motivo At-hook (AHL1; Fujimoto et al., ver más arriba). Se designaron en forma similar los polipéptidos parálogos, es decir, AHL, y se numeraron consecutivamente.

10 En la patente de los Estados Unidos No. 7.193.129, y en la solicitud de patente de los Estados Unidos No. 2005/0097638, se transformó un polipéptido AHL de *Arabidopsis thaliana*, AHL19 (de acuerdo con Fujimoto et al., ver más arriba) (identificado como G2153) en *Arabidopsis*, y expresó utilizando el promotor 35S del CaMV. Las plantas transgénicas mostraron características modificadas, tales como mayor resistencia al estrés por salinidad, mayor resistencia al estrés osmótico, mayor resistencia a la sequía, mayor tolerancia a la congelación y mayor respuesta de la planta a los azúcares. En la solicitud de patente de los Estados Unidos No. 2005/0097638, la sobreexpresión (bajo el control de un promotor 35S del CaMV) del polipéptido AHL19, así como de varios polipéptidos AHL parálogos, retrasó significativamente la floración en las plantas transgénicas comparado con las plantas de control, incrementando por lo tanto el rendimiento. La publicación internacional WO2007/028165A2 se refiere al análisis de diferentes polipéptidos por la capacidad para conferir resistencia a una enfermedad o tolerancia al estrés abiótico.

Resumen

25 De acuerdo con una forma de realización, se provee un método para incrementar las características relacionadas con el rendimiento de semillas en plantas con relación a las plantas de control, que comprende incrementar la expresión de una secuencia de ácido nucleico que codifica un polipéptido AHL19/20 en una planta. El incremento de características relacionadas con el rendimiento de semillas, comprende uno o más de: aumento del número de flores por panícula, aumento del rendimiento total de semillas por planta, aumento de la cantidad de semillas llenas, y aumento del índice de cosecha.

Definiciones

30 Polipéptido(s) / Proteína(s)

Los términos "polipéptido" y "proteína" se utilizan indistintamente aquí y se refieren a aminoácidos en una forma polimérica de cualquier longitud, unidos por medio de enlaces peptídicos.

Polinucleótido(s) / Ácido(s) nucleico(s) / Secuencia(s) de ácido nucleico / Secuencia(s) de nucleótidos

35 Los términos "polinucleótido(s)", "secuencia(s) de ácido nucleico", "secuencia(s) de nucleótidos", "ácido(s) nucleico(s)" se utilizan indistintamente aquí y se refieren a nucleótidos, ya sea ribonucleótidos o desoxirribonucleótidos o una combinación de ambos, en una forma polimérica no ramificada de cualquier longitud.

Planta(s) de control

40 La elección de plantas control adecuadas es una parte rutinaria de la programación de un experimento y puede incluir las plantas correspondientes de tipo silvestre o las plantas correspondientes sin el gen de interés. La planta control es típicamente de la misma especie de planta o incluso de la misma variedad que la planta que va a ser evaluada. La planta control puede ser también un nulicigoto de la planta que va a ser evaluada. Una "planta de control" tal como se utiliza aquí se refiere no sólo a plantas enteras, sino también a partes de plantas, incluyendo semillas y partes de la semilla.

Homólogo(s)

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

Una supresión se refiere a la remoción 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 incluir fusiones en el terminal N y/o en el terminal C así como inserciones dentro de la secuencia de uno solo o de múltiples aminoácidos. En general, las inserciones dentro de la secuencia de aminoácidos serán más pequeñas que las fusiones en el terminal N o en el terminal C, del orden de aproximadamente 1 a 10 residuos. Los ejemplos de proteínas o péptidos de fusión del terminal N o C incluyen el dominio de enlazamiento o el dominio de activación de un activador transcripcional tal como el que se utiliza en el sistema de doble híbrido de levadura, las proteínas de recubrimiento de fago, la etiqueta de (histidina)₆, la etiqueta de glutatona S transferasa, la proteína A, la proteína de enlazamiento de maltosa, la dihidrofolato reductasa, el epítipo Tag-100, el epítipo c-myc, el epítipo FLAG®, lacZ, CMP (péptido de enlazamiento de calmodulina), el epítipo HA, el epítipo de proteína C y el epítipo VSV.

Una sustitución se refiere al reemplazo de aminoácidos de la proteína por otros aminoácidos que tienen propiedades similares (tales como hidrofobicidad, hidrofiliidad, antigenicidad, propensión a formar o romper estructuras helicoidales α o estructuras de láminas β similares). Las sustituciones de aminoácidos son típicamente de residuos individuales, pero pueden agruparse dependiendo de las restricciones funcionales colocadas sobre el polipéptido; las inserciones usualmente serán del orden de aproximadamente 1 a 10 residuos de aminoácidos. Las sustituciones de aminoácidos son preferiblemente sustituciones conservadoras de aminoácidos. Las tablas de sustitución conservadora son bien conocidas en la técnica (véase por ejemplo Creighton (1984) Proteins. W. H. Freeman and Company (Eds.) y la Tabla 1 siguiente).

Tabla 1: Ejemplos de sustituciones conservadoras de aminoácidos

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 aminoácidos se pueden realizar fácilmente utilizando técnicas de síntesis de péptidos bien conocidas en la técnica, tales como la síntesis de péptidos en fase sólida y similares, o por medio de manipulación de ADN recombinante. Los métodos para la manipulación de las secuencias de ADN para producir las variantes de sustitución, inserción o supresión de una proteína son bien conocidos en la técnica. Por ejemplo, las técnicas para realizar mutaciones de sustitución en sitios predeterminados en el ADN son bien conocidas por aquellos ordinariamente capacitados en la técnica e incluyen la mutagénesis M13, mutagénesis *in vitro* del Gen T7 (USB, Cleveland, OH), mutagénesis dirigida al sitio QuickChange (Stratagene, San Diego, CA), mutagénesis dirigida al sitio mediada por PCR u otros protocolos de mutagénesis dirigida al sitio.

Derivados

Los “derivados” incluyen péptidos, oligopéptidos, polipéptidos los cuales pueden, en comparación con la secuencia de aminoácidos de la forma de la proteína de origen natural, tal como la proteína de interés, incluir sustituciones de aminoácidos con residuos de aminoácidos que no son de origen natural, o adiciones de residuos de aminoácidos que no son de origen natural. Los “derivados” de una proteína también abarcan péptidos, oligopéptidos, polipéptidos los cuales incluyen residuos de aminoácidos de origen natural alterados (glicosilados, acilados, prenilados, fosforilados, miristoilados, sulfatados, etc.) o alterados de origen no natural comparado con la secuencia de aminoácidos de una forma de origen natural del polipéptido. Un derivado puede incluir también uno o más sustituyentes o adiciones que no son de aminoácidos comparado con la secuencia de aminoácidos de la cual se deriva, por ejemplo una molécula reportera u otro ligando, enlazado en forma covalente o no covalente a la secuencia de aminoácidos, tal como una molécula reportera que se enlaza para facilitar su detección, y residuos de aminoácidos de origen no natural con relación a la secuencia de aminoácidos de una proteína de origen natural.

Además, “derivados” también incluye fusiones de la forma de origen natural de la proteína con péptidos marcadores tales como FLAG, HIS6 o tiorredoxina (para una revisión de los péptidos marcadores, véase Terpe, App. Microbiol. Biotechnol. 60, 523 - 533, 2003).

Ortólogo(s) / Parálogo(s)

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

5 Dominio

El término "dominio" se refiere a un conjunto de aminoácidos conservados en posiciones específicas a lo largo de una alineación de secuencias de proteínas relacionadas por evolución. Mientras que los aminoácidos en otras posiciones pueden variar entre homólogos, los aminoácidos que son altamente conservados 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 homólogos de proteína, se pueden utilizar como identificadores para determinar si cualquier polipéptido en cuestión pertenece a una familia de polipéptidos previamente identificada.

Motivo / Secuencia de consenso / Firma

15 El término "motivo" o "secuencia de consenso" o "firma" se refieren a una región conservada corta en la secuencia de proteínas relacionadas por evolución. Los motivos son frecuentemente partes de dominios altamente conservados, pero pueden incluir también solamente parte del dominio, o estar localizados fuera del dominio conservado (si todos los aminoácidos del motivo caen fuera de un dominio definido).

Hibridación

20 El término "hibridación" como se define aquí es un proceso en donde las secuencias nucleotídicas complementarias sustancialmente homólogas se aparean entre sí. El proceso de hibridación puede ocurrir completamente en solución, es decir ambas moléculas de ácido nucleico complementarias están en solución. El proceso de hibridación puede ocurrir también con una de las moléculas complementarias de ácido nucleico inmovilizadas a una matriz tal como perlas magnéticas, perlas de Sefarosa o cualquier otra resina. El proceso de hibridación puede ocurrir además con una de las moléculas complementarias de ácido nucleico inmovilizadas a un soporte sólido tal como una membrana de nitrocelulosa o de nailon o inmovilizada, por ejemplo, por medio de fotolitografía, por ejemplo a un soporte de vidrio síliceo (éste último conocido como arreglos de secuencias de ácido nucleico o microarreglos o como chips de secuencias de ácido nucleico). Con el propósito de permitir que se produzca la hibridación, las moléculas de ácido nucleico son generalmente desnaturalizadas térmica o químicamente para fundir una hebra bicatenaria en dos hebras individuales y/o para remover horquillas u otras estructuras secundarias de las moléculas de ácido nucleico monocatenario.

35 El término "rigurosidad" se refiere a las condiciones bajo las cuales tiene lugar una hibridación. La rigurosidad de la hibridación está influenciada por condiciones tales como la temperatura, la concentración salina, la fuerza iónica y la composición del amortiguador de hibridación. En general, las condiciones de baja rigurosidad se seleccionan para que sean aproximadamente 30°C menores al punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y pH definidos. Las condiciones de rigurosidad media se dan en los casos donde la temperatura está 20°C por debajo de la T_m , y las condiciones de alta rigurosidad se dan en los casos cuando la temperatura está 10°C por debajo de la T_m . Las condiciones de hibridación de alta rigurosidad se utilizan típicamente para aislar las secuencias de hibridación que tienen una gran similitud de secuencia con la secuencia de ácido nucleico objetivo. Sin embargo, las secuencias de ácido nucleico pueden desviarse en secuencia e incluso codificar un polipéptido sustancialmente idéntico, debido a la degeneración del código genético. Por lo tanto, algunas veces pueden ser necesarias condiciones de hibridación de rigurosidad media, para identificar tales moléculas de secuencia de ácido nucleico.

45 La T_m es la temperatura en condiciones de una fuerza iónica y pH definidos, a la cual 50% de la secuencia objetivo hibrida a una sonda perfectamente emparejada. La T_m depende de las condiciones de la solución y de la composición base y la longitud de la sonda. Por ejemplo, las secuencias más largas hibridan específicamente a temperaturas más elevadas. La velocidad máxima de hibridación se obtiene aproximadamente desde 16°C hasta 32°C por debajo de la T_m . La presencia de cationes monovalentes en la solución de hibridación reduce la repulsión electrostática entre las dos hebras de la secuencia de ácido nucleico promoviendo de este modo la formación del híbrido; este efecto es visible para concentraciones de sodio de hasta 0,4 M (para concentraciones más elevadas, este efecto puede ignorarse). La formamida reduce la temperatura de fusión de los dúplex ADN-ADN y ADN-ARN con 0,6 hasta 0,7°C para cada por ciento de formamida, y la adición de 50% de formamida permite que la hibridación se lleve a cabo entre 30 a 45°C, si bien la velocidad de la hibridación se disminuirá. La falta de emparejamiento de los pares de bases reduce la velocidad de hibridación y la estabilidad térmica de los dúplex. En promedio y para sondas largas, la T_m disminuye aproximadamente 1°C por % de falta de emparejamiento de las bases. La T_m puede calcularse utilizando las siguientes ecuaciones, dependiendo de los tipos de híbridos:

55 1) híbridos de ADN - ADN (Meinkoth y Wahl, Anal. Biochem, 138: 267 - 284, 1984):

ES 2 440 265 T3

$$T_m = 81,5^\circ\text{C} + 16,6 \times \log_{10} [\text{Na}^+]^a + 0,41 \times \% [\text{G/C}^b] - 500 \times [\text{L}^c]^{-1} - 0,61 \times \% \text{ de formamida}$$

2) híbridos de ADN - ARN o ARN - ARN:

$$T_m = 79,8 + 18,5 (\log_{10} [\text{Na}^+]^a) + 0,58 (\% \text{ de G/C}^b) + 11,8 (\% \text{ de G/C}^b)^2 - 820/\text{L}^c$$

3) híbridos de oligo - ADN u oligo - ARN^d:

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

Para 20 - 35 nucleótidos: $T_m = 22 + 1,46 (I_n)$

^a o para otro catión monovalente, pero sólo preciso en el rango de 0,01 - 0,4 M.

^b sólo preciso para % de GC en el rango de 30% a 75%.

^c L = longitud del dúplex en pares de bases.

10 ^d oligo, oligonucleótido; I_n = longitud efectiva del cebador = 2 x (no. de G/C) + (no. de A/T).

15 El enlazamiento no específico se puede controlar utilizando cualquiera entre una cantidad de técnicas conocidas tales como, por ejemplo, bloqueo de la membrana con soluciones que contienen proteína, adiciones de ARN heterólogo, ADN, y SDS al amortiguador de hibridación, y tratamiento con ARNasa. Para las sondas no homólogas, se puede llevar a cabo una serie de hibridaciones variando uno de (i) disminución progresiva de la temperatura de apareamiento (por ejemplo de 68°C hasta 42°C) o (ii) disminución progresiva de la concentración de formamida (por ejemplo desde 50% hasta 0%). La persona ordinariamente capacitada conoce diferentes parámetros que pueden ser alterados durante la hibridación y los cuales o bien mantendrán o cambiarán las condiciones de rigurosidad.

20 Además de las condiciones de hibridación, la especificidad de la hibridación típicamente también depende de la función de los lavados posteriores a la hibridación. Para eliminar el medio resultante de la hibridación no específica, se lavan las muestras con soluciones salinas diluidas. Los factores críticos de tales lavados incluyen la fuerza iónica y la temperatura de la solución final de lavado: cuánto más baja la concentración de sal y más alta la temperatura de lavado, mayor la rigurosidad del lavado. Las condiciones de lavado se llevan a cabo típicamente en o por debajo de la rigurosidad de la hibridación. Una hibridación positiva produce una señal que es al menos dos veces aquella de la señal de fondo. En general, las condiciones rigurosas adecuadas para los ensayos de hibridación de la secuencia de ácido nucleico o los procedimientos de detección por amplificación génica son como se expusieron anteriormente. Las condiciones más o menos rigurosas también se pueden seleccionar. Una persona normalmente capacitada en la técnica conoce diferentes parámetros que pueden ser alterados durante el lavado y que o bien mantendrán o cambiarán las condiciones de rigurosidad.

30 Por ejemplo, las condiciones típicas de hibridación de alta rigurosidad para los híbridos de ADN de longitud superior a 50 nucleótidos abarcan la hibridación a 65°C en 1x SSC o a 42°C en 1x SSC y 50% de formamida, seguido por un lavado a 65°C en 0,3x SSC. Los ejemplos de condiciones de hibridación de rigurosidad media para los híbridos de ADN de longitud superior a 50 nucleótidos abarcan la hibridación a 50°C en 4x SSC o a 40°C en 6x SSC y 50% de formamida, seguido por un lavado a 50°C en 2x SSC. La longitud del híbrido es de la longitud anticipada para la hibridación del ácido nucleico. Cuando se hibridan moléculas de ácido nucleico de secuencia conocida, se puede determinar la longitud del híbrido por medio de la alineación de las secuencias y la identificación de las regiones conservadas descritas aquí. 1x SSC es NaCl 0,15 M y citrato sódico 15 mM; la solución de hibridación y las soluciones de lavado pueden incluir además 5 x del reactivo de Denhardt, 0,5 - 1,0% de SDS, 100 µg/ml de ADN de esperma de salmón fragmentado y desnaturalizado, 0,5% de pirofosfato de sodio.

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

Variante de Empalme

45 El término "variante de empalme" tal como se utiliza aquí abarca las variantes de una secuencia de ácido nucleico en las cuales se han suprimido, reemplazado, desplazado o agregado los intrones y/o exones seleccionados, o en las cuales se han acortado o alargado los intrones. Tales variantes serán aquellas en las cuales se retiene sustancialmente la actividad biológica de la proteína; esto se puede lograr reteniendo selectivamente segmentos funcionales de la proteína. Tales variantes de empalme pueden hallarse en la naturaleza o pueden ser artificiales. Los métodos para predecir y aislar tales variantes de empalme son bien conocidos en la técnica (ver por ejemplo

Foissac y Schiex (2005) BMC Bioinformatics 6: 25).

Variante alélica

5 Los alelos o las variantes alélicas son formas alternativas de un gen dado, localizados en la misma posición cromosómica. Las variantes alélicas abarcan los Polimorfismos de un Solo Nucleótido (SNP por sus siglas en inglés), así como también los Polimorfismos de Inserción/Supresión pequeña (INDEL por sus siglas en inglés). El tamaño de los INDEL es usualmente inferior a 100 pb. Los SNP y los INDEL forman el conjunto más grande de variantes de secuencia en las cadenas polimórficas de origen natural de la mayoría de los organismos.

Transposición de genes / Evolución dirigida

10 La transposición de genes o la evolución dirigida consiste de iteraciones de la transposición del ADN seguida por la detección y/o selección apropiada para generar variantes de secuencias de ácido nucleico o porciones de las mismas que codifican las proteínas que tienen una actividad biológica modificada (Castle et al., (2004) Science 304(5674): 1151 - 4; patentes estadounidenses Nos. 5.811.238 y 6.395.547).

Elemento regulador / Secuencia de control / Promotor

15 Los términos "elemento regulador", "secuencia de control" y "promotor" se utilizan todos en forma intercambiable en la presente descripción y se deben tomar en un contexto amplio para referirse a secuencias reguladoras de ácido nucleico capaces de efectuar la expresión de las secuencias a las cuales están ligadas. El término 'promotor' típicamente se refiere a una secuencia control de la secuencia de ácido nucleico localizada secuencia arriba del inicio de la transcripción de un gen y la cual está involucrada en el reconocimiento y enlazamiento de la ARN polimerasa y otras proteínas, dirigiendo de este modo la transcripción de un de ácido nucleico operativamente enlazado. Abarcados por los términos antes mencionados, se hallan las secuencias reguladoras de la transcripción que se derivan de un gen genómico eucariota clásico (que incluye la caja TATA la cual es necesaria para la iniciación precisa de la transcripción, con o sin una secuencia de la caja CCAAT) y elementos reguladores adicionales (es decir, secuencias activadoras secuencia arriba, reforzadoras y silenciadoras) los cuales alteran la expresión génica en respuesta a estímulos de desarrollo y/o externos, o en una forma específica del tejido. También se incluyen dentro del término una secuencia reguladora de la transcripción de un gen procarionota clásico, en cuyo caso puede incluir una secuencia de caja -35 y/o secuencias reguladoras de la transcripción de caja -10. El término "elemento regulador" también abarca una molécula sintética de fusión o un derivado que confiere, activa o mejora la expresión de una molécula de la secuencia de ácido nucleico en una célula, tejido u órgano.

30 Un "promotor de una planta" incluye elementos reguladores, los cuales median la expresión de un segmento de una secuencia de codificación en células de plantas. El "promotor de la planta" preferiblemente se origina a partir de una célula de la planta, por ejemplo, a partir de la planta la cual es transformada con la secuencia de ácido nucleico que va a ser expresada en el proceso de la invención y descrita aquí. Esto también aplica a otras señales reguladoras de la "planta", tales como terminadores de la "planta". Los promotores secuencia arriba de las secuencias nucleotídicas útiles en los métodos de la presente invención pueden ser modificados mediante una o más sustituciones, inserciones y/o supresiones de nucleótidos sin interferir con la funcionalidad o actividad ya sea de los promotores, el marco de lectura abierto (ORF por sus siglas en inglés) o la región reguladora 3' tal como terminadores u otras regiones reguladoras 3' las cuales se localizan lejos del ORF. Además es posible que la actividad de los promotores se incremente por la modificación de su secuencia, o que sean completamente reemplazados por promotores más activos, incluso promotores de organismos heterólogos. Para la expresión en plantas, la molécula de la secuencia de ácido nucleico debe, como se describió anteriormente, estar operativamente enlazada a un promotor adecuado, o contener un promotor adecuado, el cual expresa al gen en el momento adecuado y con el patrón de expresión espacial requerido.

45 Para la identificación de promotores funcionalmente equivalentes, se pueden analizar la fuerza del promotor y/o el patrón de expresión de un promotor candidato por ejemplo enlazando operativamente el promotor a un gen reportero y ensayando el nivel de expresión y el patrón del gen reportero en diferentes tejidos de la planta. Los genes reporteros adecuados bien conocidos incluyen por ejemplo a la beta-glucuronidasa o a la beta-galactosidasa. La actividad del promotor se ensaya mediante la medición de la actividad enzimática de la beta-glucuronidasa o la beta-galactosidasa. La fuerza del promotor y/o el patrón de expresión pueden ser comparados luego con la de un promotor de referencia (tal como el utilizado en los métodos de la presente invención). En forma alternativa, la fuerza del promotor se puede ensayar cuantificando los niveles de ARNm o comparando los niveles de ARNm de la secuencia de ácido nucleico utilizada en los métodos de la presente invención, con niveles de ARNm de los genes de mantenimiento tales como el ARNr 18S, utilizando métodos conocidos en la técnica, tal como las transferencias tipo Northern con análisis densitométrico de autorradiogramas, PCR cuantitativa en tiempo real o RT-PCR (Heid et al, 1996, Genome Methods 6: 986 - 994). En general, por "promotor débil" se entiende un promotor que dirige la expresión de una secuencia de codificación a un nivel bajo. Por "nivel bajo" se entiende a niveles de aproximadamente 1/10.000 de transcritos hasta aproximadamente 1/100.000 de transcritos, hasta

5 aproximadamente 1/500.0000 de transcriptos por célula. Por el contrario, un "promotor fuerte" dirige la expresión de una secuencia de codificación a nivel alto, o aproximadamente a 1/10 de transcriptos hasta aproximadamente 1/100 de transcriptos hasta aproximadamente 1/1.000 de transcriptos por célula. En general, por "promotor de fuerza media" se entiende un promotor que dirige la expresión de una secuencia de codificación a un nivel que está en todos los casos por debajo de aquel obtenido bajo el control de un promotor 35S del CaMV.

Operativamente enlazado

El término "operativamente enlazado" tal como se utiliza aquí, se refiere a un enlace funcional entre la secuencia promotora y el gen de interés, de modo que la secuencia promotora pueda iniciar la transcripción del gen de interés.

Promotor constitutivo

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

Tabla 2a: Ejemplos de promotores constitutivos de la planta

Fuente de Genes	Referencia
Actina	McElroy et al, Plant Cell, 2: 163 - 171, 1990
HMGB	WO 2004/070039
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
Ciclofilina de arroz	Buchholz et al, Plant Mol Biol. 25(5): 837 - 43, 1994
Histona H3 de maíz	Lepetit et al, Mol. Gen. Genet. 231: 276 - 285, 1992
Histona H3 de alfalfa	Wu et al. Plant Mol. Biol. 11: 641 - 649, 1988
Actina 2	An et al, Plant J. 10(1); 107 - 121, 1996
Subunidad pequeña de Rubisco	Patente de los Estados Unidos No. 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
V-ATPasa	WO 01/14572
Proteínas de caja G	WO 94/12015

15 Promotor ubicuo

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

Promotor regulado por desarrollo

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

20 Promotor inducible

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 es expuesta a diferentes condiciones de estrés, o un promotor "inducible por patógenos", es decir activado cuando una planta es expuesta a diferentes patógenos.

25

Promotor específico de un órgano / Específico de un tejido

Un promotor específico de un órgano o específico de un tejido es uno que es capaz de iniciar Preferiblemente la transcripción en ciertos órganos o tejidos, tales como las hojas, las raíces, el tejido de la semilla, etc. Por ejemplo, un "promotor específico de la raíz" es un promotor que es transcripcionalmente activo predominantemente en las raíces de la planta, sustancialmente con la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión defectuosa en estas otras partes de la planta. Los promotores capaces de iniciar la transcripción en ciertas células únicamente se denominan aquí como "específicos de la célula".

30

Ejemplos de promotores específicos de la raíz se enumeran en la Tabla 2b a continuación:

Tabla 2b: Ejemplos de promotores específicos de la raíz

Fuente del gen	Referencia
RCc3	Plant Mol Biol. 1995 Jan; 27(2): 237 - 48
PHT1 de Arabidopsis	Kovama et al., 2005; Mudge et al. (2002, Plant J. 31: 341)
Transportador de fosfato de Medicago	Xiao et al., 2006
Pyk10 de Arabidopsis	Nitz et al. (2001) Plant Sci 161(2): 337 - 346
Genes expresable en la raíz	Tingey et al., EMBO J. 6: 1, 1987.
Gen inducible por la auxina del tabaco	Van der Zaal et al., Plant Mol. Biol. 16, 983, 1991.
β -tubulina	Oppenheimer, et al., Gene 63: 87, 1988.
Genes específicos de la raíz del tabaco	Conkling, et al., Plant Physiol. 93: 1203, 1990.
Gen G1-3b de B. napus	United States Patent 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
Lede AHL 19/201 (tomate)	Lauter et al. (1996, PNAS 3:8139)
El LeNRT1-1 (tomate)	Lauter et al. (1996, PNAS 3:8139)
Gen de la patatina clase I (patata)	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, EUA
OsRAB5a (arroz)	Wang et al. 2002, Plant Sci. 163: 273
ALF5 (Arabidopsis)	Diener et al. (2001, Plant Cell 13: 1625)
NRT2;1 Np (N. plumbaginifolia)	Quesada et al. (1997, Plant Mol. Biol. 34:265)
Lectina específica de la raíz de la cebada	Lerner & Raikhel (1989) Plant Phys 91: 124 - 129
Proteína rica en hidroxiprolina específica de la raíz	Keller & Lamb (1989) Genes & Dev 3: 1639 - 1646
CDC27B/hobbit de Arabidopsis	Blilou et al. (2002) Genes & Dev 16: 2566 - 2575

5 Un promotor específico de semilla es transcripcionalmente activo predominantemente en el tejido de la semilla, pero no necesariamente exclusivamente en el tejido de la semilla (en casos de expresión defectuosa). El promotor específico de la semilla puede estar activo durante el desarrollo de la semilla y/o durante la germinación. Ejemplos de promotores específicos de semilla se muestran en la Tabla 2c a continuación. Otros ejemplos de promotores específicos de semilla se presentan en Qing Qu y Takaiwa (Plant Biotechnol. J. 2, 113 - 125, 2004), cuya divulgación se incorpora aquí por referencia como si se la hubiera expuesto en su totalidad.

10 Tabla 2c: Ejemplos de promotores específicos de la semillas

Fuente de Genes	Referencia
Gene específicos de la semilla	Simon et al., Plant Mol. Biol. 5: 191, 1985.
	Scofield et al., J. Biol. Chem. 262: 12202, 1987.
	Baszczynski et al., Plant Mol. Biol. 14: 633, 1990.
Albúmina de la nuez del 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.
Zeina	Matzke et al Plant Mol Biol, 14(3): 323 - 32 1990
NapA	Stalberg et al, Planta 199: 515 - 519, 1996.
Glutenina-1 LMW y HMW 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 , γ -gliadinas de trigo	EMBO J. 3: 1409 - 15, 1984
Promotor ltr1 de cebada	Diaz et al. (1995) Mol Gen Genet 248(5): 592 - 8
Hordeína B1, C, D 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-Carbajosa et al., Plant J. 13: 629 - 640, 1998.
Prolamina NRP33 de arroz	Wu et al, Plant Cell Physiology 39(8) 885 - 889, 1998
α -globulina G1b-1 de arroz	Wu et al, Plant Cell Physiology 39(8) 885 - 889, 1998

(continuación)

Fuente de Genes	Referencia
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-glucosa pirofosforilasa de arroz	Trans Res 6: 157 - 68, 1997
Familia de genes ESR del maíz	Plant J 12: 235 - 46, 1997
α -kafirina del sorgo	DeRose et al., Plant Mol. Biol 32: 1029 - 35, 1996
KNOX	Postma-Haarsma et al, Plant Mol. Biol. 39: 257 - 71, 1999
Oleosina del arroz	Wu et al, J. Biochem. 123: 386, 1998
Oleosina de girasol	Cummins et al., Plant Mol. Biol. 19: 873 - 876, 1992
Proteína ribosomal 40S putativa del arroz, PRO0117	WO 2004/070039
Alanina aminotransferasa del arroz, PRO0136	No publicada
Inhibidor ITR1 de tripsina (cebada), PRO0147	No publicada
WSI18 de arroz, PRO0151	WO 2004/070039
RAB21 de arroz, PRO0175	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 tipo catepsina β	Cejudo et al, Plant Mol Biol 20: 849 - 856, 1992
Cebada Ltp2	Kalla et al., Plant J. 6: 849 - 60, 1994
Chi26	Leah et al., Plant J. 4: 579 - 89, 1994
Maíz B-Perú	Selinger et al., Genetics 149; 1125 - 38, 1998

5 Un “promotor activo en partes aéreas” se refiere a un promotor que es capaz de iniciar preferencialmente la transcripción en partes aéreas de una planta sustancialmente con la exclusión de cualquier otra de las partes de la planta (específicamente las partes por debajo de las partes aéreas), mientras que todavía permite cualquier expresión defectuosa en esas otras partes de la planta. La Tabla 2d a continuación muestra ejemplos de tales promotores, que son transcripcionalmente activos predominantemente en tejido verde.

10 Un promotor específico del tejido verde como se define aquí, es un promotor que es transcripcionalmente activo predominantemente en tejido verde, sustancialmente con la exclusión de cualquier otra de las partes de una planta, mientras que aún permite cualquier expresión defectuosa en estas otras partes de la planta.

Ejemplos de promotores específicos del tejido verde que pueden ser utilizados para llevar a cabo los métodos de la invención son presentados en la Tabla 2d a continuación.

Tabla 2d: Ejemplos de promotores específicos del tejido verde

Gen	Expresión	Referencia
Ortofosfato diquinasa del maíz	Específica de la hoja	Fukavama et al., 2001
Fosfoenolpiruvato carboxilasa del maíz	Específica de la hoja	Kausch et al., 2001
Fosfoenolpiruvato carboxilasa del arroz	Específica de la hoja	Liu et al., 2003
Pequeña subunidad de Rubisco del arroz	Específica de la hoja	Nomura et al., 2000
Beta expansina EXBP9 del arroz	Específica del brote	WO 2004/070039
Pequeña subunidad de Rubisco del guandú	Específica de la hoja	Panguluri et al., 2005
RBCS3A del guisante	Específica de la hoja	

15 Otro ejemplo de un promotor específico del tejido es un promotor específico del meristema, el cual es transcripcionalmente activo predominantemente en el tejido meristemático, sustancialmente con la exclusión de cualquier otra de las partes de una planta, mientras que aún permite cualquier expresión defectuosa en estas otras partes de la planta. Ejemplos de promotores específicos del meristema que pueden ser utilizados para llevar a cabo los métodos de la invención se presentan en la Tabla 2e a continuación.

20

Tabla 2e: Ejemplos de promotores específicos del meristema

Fuente del gen	Patrón de expresión	Referencia
OSH1 del arroz	Meristema apical del brote, desde la etapa globular embrionaria hasta la etapa de plántula	Sato et al. (1996) Proc. Natl. Acad. Sci. USA, 93: 8117 - 8122
Metalotioneína del arroz	Específico del meristema	BAD87835.1
WAK1 & WAK 2	Meristemas apicales de raíz y de brote, y en hojas y sépalos en expansión	Wagner & Kohorn (2001) Plant Cell 13(2): 303 - 318

Tabla 2f: Ejemplos de promotores específicos del endospermo

Fuente del 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
glutenina-1 de alto y de bajo peso molecular del 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 del trigo	Rafalski et al. (1984) EMBO 3: 1409 - 15
Promotor ltr1 de cebada	Diaz et al. (1995) Mol Gen Genet 248(5): 592 - 8
Hordeína de cebada B1, C, D	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 et al. (1999) J Biol Chem 274(14): 9175 - 82

5

(continuación)

Fuente del gen	Referencia
Promotor sintético	Vicente-Carbajosa et al. (1998) Plant J 13: 629 - 640
Prolamina NRP33 de arroz	Wu et al. (1998) Plant Cell Physiol 39(8) 885 - 889
Globulina Glb-1 de arroz	Wu et al. (1998) Plant Cell Physiol 39(8) 885 - 889
Globulina REB/OHP-1 de arroz	Nakase et al. (1997) Plant Molec Biol 33: 513 - 522
ADP-glucosa pirofosforilasa de arroz	Russell et al. (1997) Trans Res 6: 157 - 68
Familia del gen ESR de maíz	Opsahl-Ferstad et al. (1997) Plant J 12: 235 - 46
Kafirina de sorgo	DeRose et al. (1996) Plant Mol Biol 32: 1029 - 35

Tabla 2g: Ejemplos de promotores específicos del embrión

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

10

Tabla 2h: Ejemplos de promotores específicos de la aleurona

Fuente del gen	Referencia
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 del tipo 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
B-Perú de maíz	Selinger et al., Genetics 149; 1125 - 38, 1998

Terminador

15

El término "terminador" abarca una secuencia de control que es una secuencia de ADN al extremo de una unidad de transcripción la cual señala el procesamiento 3' y la poliadenilación de un transcripto primario y la terminación de la

transcripción. El terminador puede derivarse del gen natural, a partir de una variedad de otros genes de planta, o del ADN-T. El terminador que va a ser añadido puede derivarse de, por ejemplo, los genes de la nopalina sintasa o de la octopina sintasa, o en forma alternativa de otro gen de la planta, o con menor preferencia a partir de cualquier otro gen eucariota.

5 Modulación

El término "modulación" significa en relación con expresión o con expresión génica, un proceso en el cual el nivel de expresión es cambiado por dicha expresión génica en comparación con la planta de control, Preferiblemente el nivel de expresión se incrementa. La expresión original, no modulada puede ser de cualquier tipo de expresión de un ARN estructural (ARNr, ARNt) o de ARNm con posterior traducción. El termino "que modula la actividad" significa cualquier cambio de la expresión de las secuencias de ácido nucleico de la invención o las proteínas codificadas, lo cual conduce a un mayor rendimiento y/o a un mayor crecimiento de las plantas.

Expresión

El término "expresión" o "expresión génica" significa la transcripción de un gen específico o genes específicos o constructo genético específico. El término "expresión" o "expresión génica" significa en particular la transcripción de un gen o genes o un constructo genético en un ARN estructural (ARNr, ARNt) o ARNm con o sin traducción posterior de este último en una proteína. El proceso incluye la transcripción de ADN y el procesamiento del producto de ARNm resultante.

Expresión incrementada / sobreexpresión

El término "expresión incrementada" o "sobreexpresión" tal como se utiliza aquí significa cualquier forma de expresión que sea adicional al nivel original de expresión de tipo silvestre.

Los métodos para incrementar la expresión de los genes o de los productos génicos están bien documentados en la técnica e incluyen, por ejemplo, la sobreexpresión conducida por promotores apropiados, el uso de incrementadores de la transcripción o incrementadores de la traducción. Las secuencias de ácido nucleico aisladas que sirven como elementos promotores o incrementadores pueden ser introducidas en una posición apropiada (típicamente secuencia arriba) de una forma no heteróloga de un polinucleótido para favorecer la expresión de una secuencia de ácido nucleico que codifica al polipéptido de interés. Por ejemplo, los promotores endógenos pueden ser alterados *in vivo* por medio de mutación, supresión, y/o sustitución (véase, Kmiec, patente de los Estados Unidos No. 5.565.350; Zarlring et al., WO9322443), o se pueden introducir los promotores aislados en una célula de una planta en la orientación y distancia apropiados de un gen de la presente invención para controlar la expresión del gen.

Si se desea una expresión polipeptídica, en general es deseable incluir una región de poliadenilación en el extremo 3' de un región de codificación del polinucleótido. La región de poliadenilación puede derivarse del gen natural, a partir de una variedad de otros genes de la planta, o a partir del ADN-T. La secuencia del extremo de 3' que va a ser añadida puede derivarse, por ejemplo, de los genes de la nopalina sintasa o de la octopina sintasa, o en forma alternativa a partir de otro gen de la planta, o menos preferiblemente a partir de cualquier otro gen eucariota.

También se puede añadir una secuencia del intrón a la región no traducida (UTR) 5' o en la secuencia codificadora de la secuencia de codificación parcial para incrementar la cantidad del mensaje de maduración que se acumula en el citosol. La inclusión de un intrón que puede empalmarse en la unidad de transcripción tanto en los constructos de expresión de plantas como de animales ha demostrado que incrementa la expresión génica tanto en los niveles de ARNm como de proteína hasta 1000 veces (Buchman y Berg (1988) Mol. Cell Biol. 8: 4395 - 4405; Callis et al. (1987) Genes Dev. 1: 1183 - 1200). Tal incremento del intrón de la expresión génica es típicamente mayor cuando se localiza cerca del extremo 5' de la unidad de transcripción. El uso de los intrones del maíz, intrón Adh1-S 1, 2 y 6 y el intrón de Bronze-1 es conocido en la técnica. Para información general, véase: The Maize Handbook, Capítulo 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

Gen endógeno

La referencia aquí a un gen "endógeno" no sólo se refiere al gen en cuestión tal como se encuentra en una planta en su forma natural (es decir, sin que exista intervención humana alguna), sino que también se refiere al mismo gen (o un ácido nucleico/gen sustancialmente homólogo) en un forma aislada posteriormente (re)introducido en una planta (un transgén). Por ejemplo, una planta transgénica que contiene dicho transgén puede encontrar una reducción sustancial de la expresión del transgén y/o una reducción sustancial de la expresión del gen endógeno.

El gen aislado se puede aislar de un organismo o ser elaborado por el hombre, por ejemplo por medio de síntesis química.

Expresión disminuida

5 La referencia que se hace aquí a una “expresión disminuida” o a una “reducción o eliminación sustancial” de la expresión se pretende que signifique una disminución en la expresión del gen endógeno y/o de los niveles de polipéptido y/o de la actividad del polipéptido con relación a las plantas de control. La reducción o la eliminación sustancial es en orden creciente de preferencia una reducción de al menos 10%, 20%, 30%, 40% o 50%, 60%, 70%, 80%, 85%, 90%, o 95%, 96%, 97%, 98%, 99% o más en comparación con las de las plantas de control.

10 Para la reducción o la eliminación sustancial de la expresión de un gen endógeno en una planta, se requiere de una longitud suficiente de nucleótidos sustancialmente contiguos de una secuencia de ácido nucleico. Con el fin de llevar a cabo el silenciamiento génico, este puede ser tan pequeño como de 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 o menos nucleótidos, alternativamente este puede ser tanto como el gen completo (incluyendo la UTR de 5' y/o 3', ya sea en parte o completa). El alargamiento de nucleótidos sustancialmente contiguos puede derivarse de la secuencia de ácido nucleico que codifica la proteína de interés (gen objetivo), o de cualquier secuencia de ácido nucleico capaz de codificar un ortólogo, un parálogo o un homólogo de la proteína de interés. Preferiblemente, el alargamiento de los nucleótidos sustancialmente contiguos es capaz de formar enlaces de hidrógeno con el gen objetivo (ya sea una cadena sentido o antisentido), más preferiblemente, el alargamiento de los nucleótidos sustancialmente contiguos tiene, en orden creciente de preferencia, una identidad de secuencia del 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 100% con respecto al gen objetivo (ya sea una cadena sentido o antisentido). Una secuencia de ácido nucleico que codifica un polipéptido (funcional) no es un requerimiento para los diferentes métodos discutidos aquí para la reducción o eliminación sustancial de la expresión de un gen endógeno.

20 Esta reducción o eliminación sustancial de la expresión puede lograrse mediante el uso de herramientas y técnicas de rutina. Un método para la reducción o eliminación sustancial de la expresión del gen endógeno es mediante el silenciamiento mediado por ARN utilizando una repetición invertida de una secuencia de ácido nucleico o una parte de la misma (en este caso un alargamiento de nucleótidos sustancialmente contiguos derivados del gen de interés, o de cualquier secuencia de ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), preferiblemente capaz de formar una estructura de horquilla. Otro ejemplo de un método de silenciamiento del ARN involucra la introducción de secuencias de ácido nucleico o de partes de las mismas (en este caso un alargamiento de nucleótidos sustancialmente contiguos que derivan del gen de interés, o de cualquier secuencia de ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), en una orientación sentido en una planta. Otro ejemplo de un método de silenciamiento de ARN involucra el uso de secuencias de ácido nucleico antisentido. El silenciamiento génico también puede lograrse mediante mutagénesis de inserción (por ejemplo, inserción de ADN-T o inserción de un transposón) o mediante estrategias como las descritas, entre otros, por Angeli y Baulcombe ((1999) Plant J 20(3): 357 - 62), (Amplición VIGS WO 98/36083), o Baulcombe (WO 99/15682). Otros métodos, tal 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 cual está involucrado un polipéptido, serán bien conocidos por aquellos normalmente capacitados en el arte. Se pueden utilizar microARN artificiales (miARNa) y/o naturales para desactivar la expresión génica y/o la traducción del ARNm. Los miARN endógenos son los ARN monocatenarios pequeños típicamente de 19 - 24 nucleótidos de largo. Los microARN artificiales (miARNa), que tienen típicamente 21 nucleótidos de longitud, pueden ser modificados por ingeniería genética específicamente para regular en forma negativa la expresión génica de genes individuales o múltiples de interés. Los determinantes de la selección objetivo del ARN micro de una planta son bien conocidos en la técnica. Los parámetros empíricos para el reconocimiento del objetivo han sido definidos y pueden ser utilizados para ayudar en el diseño de los ARNm específicos (Schwab et al., (2005) Dev Cell 8(4): 517 - 27). Las herramientas convenientes para el diseño y generación de los miARNa y sus precursores también se encuentran disponibles al público (Schwab et al., (2006) Plant Cell 18(5): 1121 - 33).

45 En forma más detallada:

50 Esta reducción o eliminación sustancial de la expresión puede ser alcanzada utilizando herramientas y técnicas de rutina. Un método preferido para la reducción o eliminación sustancial de la expresión génica endógena es mediante la introducción y expresión en una planta de una construcción genética en la que el ácido nucleico (en este caso un tramo de nucleótidos sustancialmente contiguo derivado del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo o un homólogo de una cualquiera de las proteínas de interés) se clona como una repetición invertida (en parte o completamente), separada por un espaciador (ADN no codificador).

55 En un método preferido tal, la expresión del gen endógeno se reduce o elimina sustancialmente a través silenciamiento mediado por ARN usando una repetición invertida de un ácido nucleico o una parte del mismo (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), 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 ácido nucleico no codificante de ADN (un espaciador, por ejemplo, un fragmento de la región de unión de la matriz (MAR), un intrón, un polienlazador, etc.) se encuentra entre los dos

- ácidos nucleicos invertidos que forman la repetición invertida. Después de la transcripción de la repetición invertida, se forma un ARN quimérico con una estructura auto-complementaria (parcial o completa). Esta estructura de ARN bicatenaria se denomina como el ARN de horquilla (ARNh). El ARNh es procesado por la planta en los ARNi cortos que se incorporan en un complejo de silenciamiento inducido por ARN (RISC). El RISC escinde aún más los transcritos de ARNm, lo que reduce sustancialmente el número de transcritos de ARNm que se traducen en polipéptidos. Para más detalles generales véase, por ejemplo, Grierson et al. (1998) publicación internacional WO 98/53083; Waterhouse et al. (1999) publicación internacional WO 99/53050).
- El funcionamiento de los métodos de la invención no se basa únicamente en introducir y expresar en una planta una construcción genética en la cual se clona el ácido nucleico como una repetición invertida, sino en uno cualquiera o más de los diferentes métodos de "silenciamiento de genes" bien conocidos que pueden ser utilizados para conseguir los mismos efectos.
- Uno de tales métodos para la reducción de la expresión de genes endógenos es el silenciamiento mediado por ARN de la expresión génica (subregulación). El silenciamiento en este caso se activa en una planta por medio de una secuencia de ARN bicatenario (ARNbc) que es sustancialmente similar a la del gen endógeno objetivo. Este ARNbc es procesado adicionalmente por la planta en aproximadamente 20 hasta aproximadamente 26 nucleótidos llamados ARN cortos de interferencia (ARNi cortos). Los ARNi cortos se incorporan en un complejo de silenciamiento inducido por ARN (RISC) que escinde el transcrito de ARNm del gen endógeno objetivo, lo que reduce sustancialmente el número de transcripciones de ARNm que se traducen en un polipéptido. Preferiblemente, la secuencia de ARN bicatenario corresponde a un gen objetivo.
- Otro ejemplo de un método de silenciamiento de ARN implica la introducción de secuencias de ácido nucleico o partes de las mismas (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) en una orientación sentido en una planta. "Orientación sentido" se refiere a una secuencia de ADN que es homóloga a un transcrito de ARNm del mismo.
- Introducido en una planta sería por lo tanto al menos una copia de la secuencia de ácido nucleico. La secuencia de ácido nucleico 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 varias copias adicionales de una secuencia de ácido nucleico en la planta, ya que existe una correlación positiva entre altos niveles de transcripción y la activación de la co-supresión.
- Otro ejemplo de un método de silenciamiento de ARN implica el uso de secuencias de ácido nucleico antisentido. Una secuencia de ácido nucleico "antisentido" comprende una secuencia de nucleótidos que es complementaria a una secuencia de ácido nucleico "sentido" que codifica una proteína, es decir, complementaria a la cadena codificadora de una molécula de ADNc bicatenario o complementaria a una secuencia de transcripción de ARNm. La secuencia de ácido nucleico antisentido es preferentemente complementaria a la del gen endógeno que va a ser silenciado. La complementariedad se puede localizar en la "región codificadora" y / o en la "región no codificadora" de un gen. El término "región codificadora" se refiere a una región de la secuencia de nucleótidos que comprende codones que son traducidos en residuos de aminoácidos. El término "región no codificadora" se refiere a secuencias 5' y 3' que flanquean la región codificadora que se transcribe pero no se traduce en aminoácidos (también conocidas como regiones no traducidas 5' y 3').
- Las secuencias de ácido nucleico antisentido se pueden diseñar de acuerdo con las reglas de Watson y Crick de apareamiento de bases. La secuencia de ácido nucleico antisentido puede ser complementaria a toda la secuencia de ácido nucleico (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 pueden ser un oligonucleótido que es antisentido a sólo una parte de la secuencia de ácido nucleico (incluyendo las UTR 5' y 3' del ARNm). Por ejemplo, la secuencia antisentido del oligonucleótido puede ser complementaria a la región que rodea el sitio de inicio de la traducción de un transcrito de ARNm que codifica una polipéptido. La longitud de una secuencia de oligonucleótidos antisentido adecuada es conocida en la técnica y puede comenzar desde aproximadamente 50, 45, 40, 35, 30, 25, 20, 15 o 10 nucleótidos de longitud o menos. Una secuencia de ácido nucleico antisentido de acuerdo con la invención puede ser construida usando síntesis química y reacciones de ligación enzimática utilizando métodos conocidos en la arte. Por ejemplo, una secuencia de ácido nucleico antisentido (por ejemplo, una secuencia de oligonucleótido antisentido) puede ser químicamente sintetizada utilizando nucleótidos de origen natural o nucleótidos modificados en diferentes formas diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre las secuencias de ácido nucleico antisentido y sentido, por ejemplo, se pueden utilizar derivados de fosforotioato y nucleótidos sustituidos con acridina. Ejemplos de nucleótidos modificados que pueden utilizarse para generar las secuencias de ácido nucleico antisentido son bien conocidos en la técnica. Modificaciones de nucleótidos conocidos incluyen metilación, ciclización y 'tapas' y sustitución de uno o más nucleótidos de origen natural que ocurren con un análogo tal como inosina. Otras modificaciones de nucleótidos son bien conocidas en la técnica.

La secuencia de ácido nucleico antisentido puede producirse biológicamente usando un vector de expresión en el que una secuencia de ácido nucleico ha sido subclonada en una orientación antisentido (es decir, el ARN transcrito a partir del ácido nucleico insertado será de una orientación antisentido con un ácido nucleico objetivo de interés). Preferiblemente, la producción de secuencias de ácido nucleico antisentido en las plantas se produce por medio de una construcción de ácido nucleico integrada de forma estable que comprende un promotor, un oligonucleótido antisentido enlazado de forma operativa, y un terminador.

Las moléculas de ácido nucleico utilizadas para silenciamiento en los métodos de la invención (ya sea introducidas en una planta o generadas in situ) hibridan con o se enlazan con transcripciones de ARNm y/o ADN genómico que codifica un polipéptido para inhibir de este modo la expresión de la proteína, por ejemplo, mediante la inhibición de la transcripción y/o la traducción. La hibridación puede ser por medio de complementariedad convencional de nucleótidos para formar un dúplex estable, o, por ejemplo, en el caso de una secuencia de ácido nucleico antisentido que se enlaza con dúplex de ADN, a través de interacciones específicas en el surco mayor de la doble hélice. Se pueden introducir secuencias de ácido nucleico antisentido en una planta por medio de transformación o inyección directa en un sitio específico del tejido. Alternativamente, las secuencias de ácido nucleico antisentido se pueden modificar con células objetivo seleccionadas y después administrarse sistémicamente. Por ejemplo, para administración sistémica, se pueden modificar las secuencias de ácido nucleico antisentido de tal manera que se unan específicamente a receptores o antígenos expresados en una superficie celular seleccionada, por ejemplo, por enlazamiento de una secuencia de ácido nucleico antisentido con péptidos o anticuerpos que se enlazan con receptores de superficie celular o antígenos. Las secuencias de ácido nucleico antisentido también se pueden suministrar a las células utilizando los vectores descritos en este documento.

De acuerdo con un aspecto adicional, la secuencia de ácido nucleico antisentido es una secuencia de ácido nucleico α -anomérico. Una secuencia de ácido nucleico α -anomérico forma híbridos bicatenarios específicos con ARN complementario en el que, contrariamente a las unidades usuales, las cadenas corren paralelas entre sí (Gaultier et al. (1987) Nucl Ac Res 15: 6625 - 6641). La secuencia de ácido nucleico antisentido también puede comprender un 2'-o-metilribonucleótido (Inoue et al. Nucl Ac Res 15, (1987) 6131 - 6148) o un análogo quimérico de ARN - ADN (Inoue et al. (1987) FEBS Lett. 215, 327 - 330).

La reducción o eliminación sustancial de la expresión génica endógena también se pueden llevar a cabo usando ribozimas. Las ribozimas son moléculas catalíticas de ARN con actividad de ribonucleasa que son capaces de escindir una secuencia monocatenaria de ácido nucleico, tal como un ARNm, con la que tienen una región complementaria. Por lo tanto, las ribozimas (por ejemplo, ribozimas cabeza de martillo (descritas en Haselhoff y Gerlach (1988) Nature 334, 585 - 591) se puede utilizar para escindir catalíticamente transcritos de ARNm que codifican un polipéptido, reduciendo así sustancialmente el número de transcritos de ARNm que son traducidos en un polipéptido. Se puede diseñar una ribozima que tiene especificidad por una secuencia de ácido nucleico (véase, por ejemplo: Cech et al., patente de los Estados Unidos No. 4.987.071, y Cech et al., patente de los Estados Unidos No. 5.116.742). Alternativamente, se pueden utilizar los transcritos de ARNm correspondientes a una secuencia de ácido nucleico para seleccionar un ARN catalítico que tiene una actividad específica de ribonucleasa de una reserva de moléculas de RNA (Bartel y Szostak (1993) Science 261, 1411 - 1418). El uso de ribozimas para silenciamiento génico en plantas es conocido en la técnica (por ejemplo, Atkins et al. (1994) publicación internacional WO 94/00012; Lenne et al. (1995) publicación internacional WO 95/03404; Lutziger et al. (2000) publicación internacional WO 00/00619; Prinsen et al. (1997) publicación internacional WO 97/13865 y Scott et al. (1997) publicación internacional WO 97/38116).

El silenciamiento génico se puede lograr también por medio de mutagénesis de inserción (por ejemplo, inserción de ADN-T o inserción de transposón) o por medio de estrategias como las descritas, entre otros, por Angell y Baulcombe ((1999) Plant J 20 (3): 357 - 62), (Amplificón VIGS publicación internacional WO 98/36083), o Baulcombe (publicación internacional WO 99/15682).

El silenciamiento génico también puede producirse si hay una mutación en un gen endógeno y / o una mutación en un gen / ácido nucleico aislado introducido posteriormente en una planta. La reducción o eliminación sustancial pueden ser causadas por un polipéptido no funcional. Por ejemplo, un polipéptido puede enlazarse con varias proteínas que interactúan; una o más mutaciones y / o truncamientos puede por lo tanto proporcionar un polipéptido que sea todavía capaz de enlazarse con proteínas que interactúan (tales como proteínas receptoras) pero que no puede exhibir su función normal (tales como un ligando de señalización).

Otro enfoque para el silenciamiento de genes es por direccionamiento de secuencias de ácido nucleico 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 la inhibición de su función

in planta, o la interferencia en la ruta de señalización en la que está involucrado un polipéptido, serán bien conocidos por una persona capacitada. En particular, se puede prever que las moléculas artificiales pueden ser útiles para la inhibición de la función biológica de un polipéptido objetivo, o para interferir con la ruta de señalización en la que está involucrado el polipéptido objetivo.

- 5 Por otra parte, se puede configurar un programa de cribado para identificar en una población de plantas variantes naturales de un gen, variantes que codifican polipéptidos con actividad reducida. Tales variantes naturales también se pueden utilizar, por ejemplo, para llevar a cabo la recombinación homóloga.

10 Se pueden utilizar microARN artificiales y/o naturales (miARN) para desactivar la expresión génica y/o la traducción del ARNm. Los miARN endógenos son pequeños ARN monocatenarios de típicamente 19 a 24 nucleótidos de longitud. Ellos funcionan principalmente para regular la expresión génica y/o la traducción del ARNm. La mayoría de los microARN (los miARN) tienen una complementariedad perfecta o casi perfecta con sus secuencias objetivo. Sin embargo, existen objetivos naturales con hasta cinco faltas de correspondencia. Se procesan a partir de los ARN no codificadores más largos con estructuras plegadas características por medio de ARNasas específicas bicatenarias de la familia Dicer. Tras el procesamiento, se incorporan en el complejo de silenciamiento inducido por ARN (RISC) mediante el enlazamiento con su componente principal, una proteína Argonauta. Los miARN sirven como los componentes de especificidad de RISC, ya que aparean las bases con ácidos nucleicos objetivo, en su mayoría ARNm, en el citoplasma. Eventos reguladores posteriores incluyen escisión y destrucción de ARNm objetivo y/o inhibición translacional. Los efectos de la sobreexpresión de miARN se reflejan a menudo por lo tanto en disminución de los niveles de ARNm de los genes objetivo.

20 Los microARN artificiales (miARNa), que son típicamente de 21 nucleótidos de longitud, pueden ser genéticamente modificados específicamente para regular negativamente la expresión génica de genes individuales o múltiples de interés. Los determinantes de la selección objetivo de microARN de la planta son bien conocidos en la técnica. Los parámetros empíricos para el reconocimiento objetivo han sido definidos y pueden ser utilizados para ayudar en el diseño de los miARNa específicos, (Schwab et al., Dev. Cell 8, 517 - 527, 2005). También se encuentran disponibles al público herramientas convenientes para el diseño y la generación de los miARNa y sus precursores (Schwab et al. Plant Cell 18, 1121 - 1133, 2006).

30 Para un desempeño óptimo, las técnicas de silenciamiento génico utilizadas para reducir la expresión en una planta de un gen endógeno requieren del uso de secuencias de ácido nucleico de plantas monocotiledóneas para la transformación de plantas monocotiledóneas, y de plantas dicotiledóneas para la transformación de plantas dicotiledóneas. Preferiblemente, se introduce una secuencia de ácido nucleico de cualquier especie de planta dada en esa misma especie. Por ejemplo, una secuencia de ácido nucleico de arroz se transforma en una planta de arroz. Sin embargo, no es un requerimiento absoluto que la secuencia de ácido nucleico que va a ser introducida se origine a partir de la misma especie de planta que la planta en la cual será introducida. Es suficiente que exista homología sustancial entre el gen objetivo endógeno y la secuencia de ácido nucleico que va a ser introducida.

35 Se describieron anteriormente ejemplos de diferentes métodos para la reducción o eliminación sustancial de la expresión en una planta de un gen endógeno. Un persona normalmente capacitada en la técnica podría adaptar fácilmente los métodos anteriormente mencionados para el silenciamiento 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, por ejemplo, del uso de un promotor apropiado.

40 Marcador seleccionable (gen) / Gen reportero

45 Los términos "marcador seleccionable", "gen marcador seleccionable" o "gen reportero" incluyen a cualquier gen que confiera un fenotipo a una célula en la cual se exprese para facilitar la identificación y/o la selección de las células que son transfectadas o transformadas con un constructo de una secuencia de ácido nucleico de la invención. Estos genes marcadores permiten la identificación de una transferencia exitosa de las moléculas de la secuencia de ácido nucleico mediante una serie de principios diferentes. Los marcadores adecuados se pueden seleccionar de los marcadores que confieren resistencia antibiótica o herbicida, que introducen una nueva característica metabólica o que permiten una selección visual. Los ejemplos de genes marcadores seleccionables incluyen a los genes que confieren resistencia a los antibióticos (tales como nptII que fosforilan la neomicina y la kanamicina, o hpt, que fosforila higromicina, o genes que confieran resistencia, por ejemplo, a la bleomicina, estreptomycin, tetraciclina, cloramfenicol, ampicilina, gentamicina, genética (G418), espectinomycin o blastidina), a los herbicidas (por ejemplo bar el cual proporciona resistencia a Basta®; aroA o gox que proporcionan resistencia contra el glifosato, o los genes que confieren resistencia, por ejemplo, a la imidazolinona, fosfinotricina o sulfonilurea), o genes que proporcionen una característica metabólica (tal como manA que le permite a las plantas utilizar manosa como única fuente de carbono o xilosa isomerasa para la utilización de xilosa, o marcadores antinutritivos tales como la resistencia a la 2-desoxiglucosa). La expresión de los genes marcadores visuales resulta en la formación de color (por ejemplo β-glucuronidasa, GUS o β-galactosidasa con sus sustratos coloreados, por ejemplo X-Gal), luminiscencia (tal como el sistema de luciferina / luciferasa) o fluorescencia (Proteína Fluorescente Verde, GFP, y

sus derivados). Esta lista representa sólo una pequeña cantidad de posibles marcadores. La persona ordinariamente capacitada en la técnica está familiarizada con tales marcadores. Se prefieren marcadores diferentes, dependiendo del organismo y del método de selección.

5 Se sabe que después de una integración estable o transitoria de secuencias de ácido nucleico en las células de una planta, sólo una minoría de las células incorpora el ADN foráneo 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
10 estos integrantes, se introduce usualmente un gen que codifica un marcador seleccionable (tal como aquellos descritos anteriormente) en las células huésped junto con el gen de interés. Estos marcadores pueden ser utilizados, por ejemplo, en mutantes en los cuales estos genes no son funcionales, por ejemplo, por medio de la supresión
15 mediante métodos convencionales. Además, se pueden introducir las moléculas de la secuencia de ácido nucleico que codifican un marcador seleccionable en una célula huésped en el mismo vector que contiene la secuencia que codifica los polipéptidos de la invención o que se utilizan en los métodos de la invención, o también en un vector separado. Las células que han sido transfectadas en forma estable con la secuencia de ácido nucleico introducida pueden ser identificadas por ejemplo mediante selección (por ejemplo, las células que han integrado el marcador seleccionable sobreviven mientras que las otras células mueren).

Ya que los genes marcadores, particularmente los genes para resistencia a los antibióticos y herbicidas, no se requieren más o no son deseables en la célula huésped transgénica una vez que se han sido introducidos con éxito las secuencias de ácido nucleico, el proceso de acuerdo con la invención para introducir las secuencias de ácido
20 nucleico emplea en forma conveniente técnicas que permiten la remoción o escisión de estos genes marcadores. Un de tales métodos es el que se conoce como cotransformación. El método de cotransformación emplea dos vectores en forma simultánea para la transformación, un vector que porta la secuencia de ácido nucleico de acuerdo con la invención y un segundo vector que porta el gen o los genes marcadores. Una gran proporción de transformantes recibe, o, en el caso de las plantas, incluye (hasta 40% o más de los transformantes), ambos vectores. En el caso de la transformación con *Agrobacterias*, los transformantes usualmente reciben solamente una parte del vector, es decir
25 la secuencia flanqueada por el ADN-T, la cual usualmente representa el casete de expresión. Los genes marcadores pueden ser removidos posteriormente de la planta transformada por medio de cruzamientos. En otro método, los genes marcadores integrados en un transposón se utilizan para la transformación junto con la secuencia de ácido nucleico deseada (conocida como la tecnología *Ac/Ds*). Los transformantes se pueden cruzar con una fuente de transposasa o se transforman los transformantes con un constructo de la secuencia de ácido nucleico que confiere la expresión de una transposasa, en forma transitoria o estable. En algunos casos (aproximadamente el 10%), el transposón salta fuera del genoma de la célula huésped una vez a tenido lugar la transformación en forma exitosa y se pierde. En otro número de casos, el transposón salta a una ubicación diferente. En estos casos el gen marcador debe ser eliminado mediante la realización de cruzamientos. En microbiología, se desarrollaron técnicas que hacen posible, o facilitan, la detección de tales eventos. Un método ventajoso adicional se fundamenta en lo que se conoce como sistemas de recombinación, cuya ventaja es que se puede prescindir de la eliminación por cruzamiento. El sistema más conocido de este tipo es el que se conoce como sistema *Cre/lox*. *Cre1* es una recombinasa que remueve las secuencias localizadas entre la secuencias de *loxP*. Si el gen marcador se integra entre las secuencias de *loxP*, es eliminado una vez que ha tenido lugar con éxito la transformación, por medio de la expresión de la recombinasa. Otros sistemas de recombinación son los sistemas *HIN/HIX*, *FLP/FRT* y *REP/STB* (Tribble et al., *J. Biol. Chem.*, 275, 2000: 22255 - 22267; Velmurugan et al., *J. Cell Biol.*, 149, 2000: 553 - 566). Es posible una
40 integración específica del sitio en el genoma de la planta de las secuencias de ácido nucleico de acuerdo con la invención. Naturalmente, estos métodos se pueden aplicar también a microorganismos tales como levaduras, hongos o bacterias.

Transgénico / Transgén / Recombinante

45 Para los propósitos de la invención, "transgénico", "transgén" o "recombinante" significan con respecto, por ejemplo, a una secuencia de ácido nucleico, un casete de expresión, una construcción génica o un vector que comprende la secuencia de ácido nucleico o un organismo transformado con las secuencias de ácido nucleico, los casetes de expresión o los vectores de acuerdo con la invención, todos aquellos constructos logrados por métodos recombinantes en los cuales o bien

50 (a) las secuencias de ácido nucleico que codifican las proteína útiles en los métodos de la invención, o

(b) la(s) secuencia(s) genética(s) de control que está(n) operativamente enlazada(s) con la secuencia de ácido nucleico de acuerdo con la invención, por ejemplo un promotor, o

(c) a) y b)

55 no están localizadas en su ambiente genético natural o han sido modificadas por métodos recombinantes, siendo posible que la modificación tomara la forma, por ejemplo, de una sustitución, adición, supresión, inversión o inserción de uno o más residuos nucleótidos. Se entiende que el ambiente genético natural significa el locus

genómico o cromosómico natural en la planta original o la presencia en una biblioteca genómica. En el caso de una biblioteca genómica, el ambiente genético natural de la secuencia de ácido nucleico es preferiblemente retenida, al menos en parte. El ambiente flanquea la secuencia de ácido nucleico al menos en un costado y tiene una longitud de secuencia de al menos 50 pb, Preferiblemente de al menos 500 pb, especialmente preferiblemente de al menos 100 pb, lo más preferible al menos 5000 pb. Un casete de expresión de origen natural - por ejemplo la combinación de origen natural del promotor natural de las secuencias de ácido nucleico con la correspondiente secuencia de ácido nucleico que codifica un polipéptido útil en los métodos de la presente invención, como se definió anteriormente - convierte un casete de expresión transgénica cuando este casete de expresión es modificado por métodos sintéticos no naturales ("artificiales") tales como, por ejemplo, tratamiento mutagénico. Los métodos adecuados son descritos, por ejemplo, en la patente de los Estados Unidos No. 5.565.350 o en la publicación internacional WO 00/15815.

Una planta transgénica para los propósitos de la invención se entiende entonces que significa, como se indicó anteriormente, que las secuencias de ácido nucleico utilizadas en el método de la invención no están en su locus natural en el genoma de dicha planta, siendo posible para las secuencias de ácido nucleico que sean expresadas en forma homóloga o heteróloga. Sin embargo, como se mencionó anteriormente, transgénico también significa que, mientras las secuencias de ácidos nucleicos de acuerdo con la invención o utilizadas en el método de la invención, estén en su posición natural en el genoma de una planta, la secuencia ha sido modificada y con respecto a la secuencia natural, y/o que las secuencias reguladoras de las secuencias naturales han sido modificadas. Transgénico Preferiblemente significa que la expresión de las secuencias de ácidos nucleicos de acuerdo con la invención en un locus no natural en el genoma, es decir, que tiene lugar una expresión homóloga o, Preferiblemente heteróloga de las secuencias de ácido nucleico. Las plantas transgénicas preferidas se mencionan en la presente invención.

Transformación

El término "introducción" o "transformación" a los cuales se hace referencia aquí, abarcan la transferencia de un polinucleótido exógeno dentro de una célula huésped, independientemente del método utilizado para la transferencia. El tejido de la planta capaz de propagación clonal posterior, ya sea por medio de organogénesis o de embriogénesis, puede ser transformado con una construcción genética de la presente invención y se puede regenerar una planta entera a partir del mismo. El tejido particular escogido variará dependiendo de los sistemas de propagación clonal disponibles y más adecuados, para la especie particular que va a ser transformada. Los ejemplos de tejidos objetivo incluyen discos de hoja, polen, embriones, cotiledones, hipocotiledones, mega gametofitos, tejido del callo, tejido meristemático existente (por ejemplo, meristema apical, yemas axilares, y meristemas de la raíz), y tejido meristemático inducido (por ejemplo, meristema del cotiledón y meristema del hipocotiledón). El polinucleótido puede ser introducido en forma transitoria o estable en una célula huésped y puede mantenerse no integrado, por ejemplo, como un plásmido. En forma alternativa, puede integrarse al genoma huésped. La célula de la planta transformada resultante puede ser luego utilizada para regenerar una planta transformada en una forma conocida por las personas ordinariamente capacitadas en la técnica.

La transferencia de genes foráneos dentro del genoma de una planta se denomina transformación. La transformación de una especie de planta es hoy en día una técnica bastante rutinaria. En forma conveniente, se puede utilizar cualquiera de los diferentes métodos de transformación para introducir el gen de interés en una célula ancestral adecuada. Los métodos descritos para la transformación y regeneración de plantas a partir de tejidos de la planta o de células de la planta pueden ser utilizados para una transformación transitoria o estable. Los métodos de transformación incluyen el uso de liposomas, electroporación, productos químicos que aumentan la incorporación de ADN libre, inyección del ADN directamente a la planta, bombardeo de partículas, transformación que utiliza virus o polen y microproyección. Los métodos pueden ser seleccionados a partir del método de calcio/poli-etilén glicol para protoplastos (Krens, F. A. et al., (1982) Nature 296, 72 - 74; Negrutiu 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 de la planta (Crossway A et al, (1986) Mol. Gen Genet 202: 179 - 185); bombardeo con partículas recubiertas con ADN o ARN (Klein T. M. et al, (1987) Nature 327: 70) infección con virus (que no se integran) y similares. Las plantas transgénicas, que incluyen plantas de cultivo transgénicas, son producidas Preferiblemente mediante transformación mediada por *Agrobacterium*. Un método de transformación conveniente es la transformación *in planta*. Para este propósito, es posible, por ejemplo, permitir que las agrobacterias actúen en las semillas de la planta o inocular el meristema de la planta con las agrobacterias. Ha probado ser particularmente conveniente de acuerdo con la invención, permitir que una suspensión de agrobacterias transformadas actúen sobre la planta intacta o al menos sobre los primordios de la flor. Posteriormente se deja que se desarrolle la planta hasta obtener las semillas de la planta tratada (Clough y Bent, Plant J. (1998) 16, 735 - 743). Los métodos para transformación del arroz mediados por *Agrobacterium* incluyen métodos bien conocidos para la transformación del arroz, tales como los descritos en cualquiera de las siguientes documentos: solicitud de patente europea EP 1198985 A1, Aldemita y Hodges (Planta 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). En el caso de la transformación del maíz, el método preferido es el descrito ya sea en Ishida et al. (Nat. Biotechnol 14(6): 745 - 50, 1996) o en Frame et al. (Plant Physiol 129(1): 13 - 22, 2002). Dichos métodos son

5 descritos además a modo de ejemplo en B. Jenes et al., Techniques for Gene Transfer, en: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S. D. Kung y R. Wu, Academic Press (1993) 128 - 143 y en Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205 - 225). Las secuencias de ácido nucleico o el constructo que va a ser expresado es clonado preferiblemente en un vector, el cual es adecuado para la transformación de *Agrobacterium tumefaciens*, por ejemplo pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Las agrobacterias transformadas por dicho vector pueden ser luego utilizadas en una forma conocida para la transformación de plantas, tales como las plantas utilizadas como modelo, del tipo *Arabidopsis thaliana* no es considerada dentro del alcance de la presente invención como una planta de cultivo), o plantas de cultivo tales como, por ejemplo, plantas de tabaco, por ejemplo, mediante la inmersión de las hojas magulladas o picadas en una solución de agrobacterias y luego se cultivan en medios adecuados. La transformación de las plantas por medio del *Agrobacterium tumefaciens* es descrita, por ejemplo, en Höfgen y Willmitzer en Nucl. Acid Res. (1988) 16, 9877 o se conoce, entre otros, a partir de F. F. White, Vectors for Gene Transfer in Higher Plants; en Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S. D. Kung y R. Wu, Academic Press, 1993, páginas 15 - 38.

15 Además de la transformación de células somáticas, las cuales tienen que ser luego regeneradas en plantas intactas, también es posible transformar las células de meristemas de la planta 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. De este modo, por ejemplo, se tratan las semillas de *Arabidopsis* con agrobacterias y se obtienen las semillas de las plantas en desarrollo de las cuales una cierta proporción se transforma y de este modo se vuelven transgénicas. [Feldman, K. A. y Marks M. D. (1987). Mol Gen Genet 208: 274 - 289; Feldmann K (1992). En: C Koncz, N-H Chua y J Shell, eds, Methods in Arabidopsis Research. World Scientific, Singapur, páginas 274 - 289]. Los métodos alternativos se basan en la remoción repetida de las inflorescencias y la incubación del sitio de escisión en el centro de la roseta con las agrobacterias transformadas, por lo que pueden obtenerse igualmente semillas transformadas en un momento posterior en el tiempo (Chang (1994). Plant J. 5: 551 - 558; Katavic (1994). Mol Gen Genet, 245: 363 - 370). Sin embargo, un método especialmente efectivo 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). C R Acad Sci Paris Life Sci, 316: 1194 - 1199], mientras que en el caso del método de la "inmersión floral" se incubaba el tejido floral en desarrollo brevemente con una suspensión de agrobacterias tratada con un tensoactivo [Clough, S. J. y Bent A. F. (1998). The Plant J. 16, 735 - 743]. Se recolecta una cierta proporción de semillas transgénicas en ambos casos, y estas semillas pueden distinguirse de las semillas no transgénicas mediante el cultivo bajo las condiciones selectivas anteriormente descritas. Además, la transformación estable de los plástidos es ventajosa porque los plástidos se heredan del progenitor en la mayoría de los cultivos reduciendo o eliminando el riesgo de flujo transgénico a través del polen. La transformación del genoma del cloroplasto se logra generalmente mediante un proceso que ha sido presentado esquemáticamente en Klaus et al., 2004 [Nature Biotechnology 22 (2), 225 - 229]. En resumen, las secuencias que van a ser transformadas se clonan junto con un gen marcador seleccionable entre secuencias de flaqueo homólogas al genoma del cloroplasto. Estas secuencias de flaqueo homólogas dirigen la integración específica del sitio en el plastoma. La transformación plastidial ha sido descrita para muchas especies de plantas diferentes y se presenta una visión de conjunto en Bock (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol. 2001 Sep 21; 312 (3): 425 - 38 o en Maliga, P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotechnol. 21, 20 - 28. Otro avance biotecnológico ha sido informado recientemente en la forma de transformantes del plástido libres de marcadores, que pueden producirse mediante un gen marcador cointegrado en forma transitoria (Klaus et al., 2004, Nature Biotechnology 22(2), 225 - 229).

Marcación de la activación del ADN-T

45 La marcación de la activación del ADN-T (Hayashi et al. Science (1992) 1350 - 1353), involucra la inserción del ADN-T, que contiene usualmente un promotor (puede ser también un mejorador de la traducción o un intrón), en la región genómica del gen de interés o 10 kb secuencia arriba o secuencia abajo de la región de codificación de un gen en una configuración tal que el promotor dirige la expresión del gen objetivo. Típicamente, 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 está típicamente embebido en un ADN-T. Este ADN-T se inserta al azar en el genoma de la planta, por ejemplo, a través de la infección con *Agrobacterium* y conduce a la expresión modificada de los genes cerca del ADN-T insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debido a la expresión modificada de los genes próximos al promotor introducido.

TILLING

55 El término "TILLING" (por sus siglas en inglés) 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 secuencias de ácido nucleico que codifican proteínas con expresión y/o actividad modificada. TILLING también permite la selección de plantas que transportan tales variantes mutantes. Estas variantes mutantes pueden exhibir expresión modificada, ya sea en intensidad o en localización o en sincronización (si las mutaciones afectan al promotor por ejemplo). Estas variantes

mutantes pueden exhibir mayor actividad que la exhibida por el gen en su forma natural. TILLING combina mutagénesis de alta densidad con métodos de selección de alto rendimiento. Las etapas típicamente seguidas en TILLING son: (a) mutagénesis EMS (Redei G. P. y Koncz C (1992) en *Methods in Arabidopsis Research*, Koncz C, Chua NH, Schell J, eds. Singapur, World Scientific Publishing Co, páginas 16 - 82; Feldmann et al., (1994) en Meyerowitz E. M., Somerville C. R., eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, páginas 137 - 172; Lightner J y Caspar T (1998) en J Martínez-Zapater, J Salinas, eds. *Methods on Molecular Biology*, Vol. 82. Humana Press, Totowa, NJ, páginas 91 - 104); (b) preparación y reunión del ADN de individuos; (c) amplificación por PCR de una región de interés; (d) desnaturalización y apareamiento para permitir la formación de heterodúplex; (e) DHPLC, donde la presencia de un heterodúplex en un consolidado se detecta 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 el arte (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 una secuencia de ácido nucleico seleccionada en una posición seleccionada definida. La recombinación homóloga es una tecnología estándar utilizada en forma rutinaria en ciencias biológicas para organismos inferiores tales como levadura o el musgo *Physcomitrella*. Los métodos para llevar a cabo la recombinación homóloga en plantas han sido descritos no sólo para las plantas modelo (Offringa et al. (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; Lida y Terada (2004) *Curr Opin Biotech* 15(2): 132 - 8).

Rendimiento

El término "rendimiento" en general significa una producción medible de valor económico, típicamente relacionado con un cultivo específico, con un área, y con un periodo de tiempo. Las partes de plantas individuales contribuyen en forma directa al rendimiento con base en su cantidad, tamaño y/o peso, o el rendimiento real es el rendimiento por acre para un cultivo y el año, que se determina mediante la división de la producción total (incluye tanto producción cosechada como tasada) por los acres plantados. El término 'rendimiento' de una planta puede referirse a la biomasa vegetal, a órganos reproductivos, y/o a propágulos (tales como semillas) de esa planta.

El término "rendimiento" de una planta puede relacionarse con biomasa vegetativa (biomasa de raíz y/o de brotes), con órganos reproductivos, y/o con propágulos (tales como semilla) de esa planta.

Vigor temprano

El término "vigor temprano" se refiere a un crecimiento bien balanceado, activo y saludable, especialmente durante las etapas tempranas del desarrollo de la planta, y puede resultar de una mayor aptitud de la planta, debido, por ejemplo, a que las plantas se adaptan mejor a su ambiente (es decir, la optimización del uso de los recursos energéticos y la repartición entre brote y raíz). Las plantas que tienen vigor temprano también muestran una mayor supervivencia de las plántulas y un mejor establecimiento del cultivo, lo que a menudo resulta en campos muy uniformes (con el desarrollo del cultivo en forma uniforme, es decir, donde la mayoría de las plantas alcanzan las diferentes etapas del desarrollo sustancialmente al mismo tiempo), y a menudo con un rendimiento mejor y superior. Por lo tanto, el vigor temprano puede ser determinado midiendo diferentes factores, tales como el peso de mil granos, el porcentaje de germinación, el porcentaje de emergencia, el desarrollo de las plántulas, la altura de las plántulas, la longitud de la raíz, la biomasa de la raíz y de los brotes y muchos más.

Incremento / Mejoramiento / Intensificación

Los términos "incremento", "mejoramiento" o "intensificación" son intercambiables y se entienden en el sentido de la aplicación de al menos 5%, 6%, 7%, 8%, 9% o 10%, Preferiblemente de al menos 15% o 20%, más preferiblemente 25%, 30%, 35% o 40% de más rendimiento y/o desarrollo en comparación con las plantas de control según se define aquí.

Rendimiento de semillas

El incremento en el rendimiento de semillas puede manifestarse en sí mismo en una o más de las formas: a) un incremento en la biomasa de la semilla (peso total de la semilla) el cual puede ser con base en una semilla individual y/o por planta y/o por hectárea o acre, b) mayor número de flores por panícula y/o por planta; c) mayor número de semillas (llenas); d) mayor proporción de llenado de las semillas (la cual se expresa como la relación entre la cantidad de semilla llena dividido por la cantidad total de semillas); e) mayor índice de cosecha, el cual se expresa como una relación del rendimiento de las partes cosechables, tal como las semillas, dividido por la biomasa total; y f) mayor número de panículas primarias, (g) mayor peso de mil granos (PMG), el cual se extrapola a partir del recuento

de la cantidad de semillas llenas y su peso total. Un mayor PMG puede ser el resultado de un mayor tamaño de la semilla y/o del peso de la semilla, y también puede ser el resultado de un incremento en el tamaño del embrión y/o del endospermo.

- 5 Un incremento en el rendimiento de semillas puede manifestarse también como un incremento en el tamaño y/o el volumen de la semilla. Además, un incremento en el rendimiento puede manifestarse también como un incremento en el área de la semilla y/o de la longitud de la semilla y/o del ancho de la semilla y/o del perímetro de la semilla. Un mayor rendimiento puede dar como resultado también una arquitectura modificada, o puede ocurrir debido a una arquitectura modificada.

Índice de Verdor

- 10 El "índice de verdor" tal como se utiliza aquí descripción se calcula a partir de las imágenes digitales de las plantas. Para cada pixel perteneciente a la planta objetivo sobre la imagen, se calcula la relación del valor verde versus el valor rojo (en el modelo RGB para codificación del color). El índice de verdor se expresa como el porcentaje de píxeles para los cuales la relación verde con respecto a rojo excede un umbral dado. Bajo condiciones normales de desarrollo, bajo condiciones de desarrollo con estrés por salinidad, y bajo condiciones de desarrollo con una disponibilidad reducida de nutrientes, el índice de verdor de la planta se mide en la última imagen antes del florecimiento. Por el contrario, bajo condiciones de desarrollo con estrés por sequía, el índice de verdor de las plantas se mide en la primera imagen después de la sequía.

Planta

- 20 El término "planta" tal como se utiliza aquí, abarca plantas enteras, ancestros y progenie de las plantas y partes de la planta, que incluyen semillas, brotes, tallos, hojas, raíces (incluyendo tubérculos), flores, y tejidos y órganos, en donde cada uno de los anteriormente mencionados incluye la secuencia del gen/ácido nucleico de interés. El término "planta" abarca además células de la planta, cultivos en suspensión, tejido del callo, embriones, regiones meristemáticas, gametofitos, esporofitos, polen y microesporas, nuevamente en donde cada uno de los mencionados anteriormente incluye la secuencia del gen/ácido nucleico de interés.
- 25 Las plantas que son particularmente útiles en los métodos de la invención incluyen a todas las plantas que pertenezcan a la superfamilia Viridiplantae, en particular plantas monocotiledóneas y dicotiledóneas incluyendo forraje y leguminosas forrajeras, plantas ornamentales, cultivos alimenticios, 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., *Anfocarpus* 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 excelsea*, *Beta vulgaris*, *Brassica* spp. (por ejemplo, *Brassica napus*, y *Brassica rapa* spp. [canola, colza, nabo]), *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* spp., *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*, *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 culinans*, *Linum usitatissimum*, *Litchi chinensis*, *Lotus* spp., *Luffa acutangula*, *Lupinus* spp., *Luzula sylvatica*, *Lycopersicon* spp. (por ejemplo, *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon pyriforme*), *Macrotyloma* spp., *Matus* 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., *Ornithopus* 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*), *Sorgo bicolor*, *Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica*, *Theobroma cacao*, *Trifolium* spp., *Triticosecale rimpai*, *Triticum* spp. (por ejemplo, *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybemum*, *Triticum macha*, *Triticum sativum* o *Triticum vulgare*), *Tropaeolum minus*, *Tropaeolum majus*, *Vaccinium* spp., *Vicia* spp., *Vigna* spp., *Viola odorata*, *Vitis* spp., *Zea mays*, *Zizania piauistris*, *Ziziphus* spp., entre otras.

Descripción detallada de la invención

- 5 Sorprendentemente, se ha encontrado ahora que al incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20, produce plantas que tienen más características relacionadas con el rendimiento de semillas, sin retrasar la floración, con relación a plantas que sirven de control. De acuerdo con una primera forma de realización, la presente invención provee un método para incrementar las características relacionadas con el rendimiento de semillas en plantas con relación a las plantas que sirven como control, que comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20.
- 10 Un método preferido para incrementar la expresión de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 es mediante la introducción y expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20.
- 15 En una forma de realización cualquier referencia en adelante a una "proteína útil en los métodos de la invención" se entiende que significa un polipéptido AHL 19/20 como se define aquí. Cualquier referencia que se haga en adelante a una "secuencia de ácido nucleico útil en los métodos de la invención" se entiende que significa una secuencia de ácido nucleico capaz de codificar dicho polipéptido AHL 19/20. La secuencia de ácido nucleico que va a ser introducida en una planta (y por lo tanto útil para la realización de los métodos de la invención) es cualquier secuencia de ácido nucleico que codifica el tipo de polipéptido, que será descrito ahora, en adelante también denominado "secuencia de ácido nucleico para AHL 19/20" o "gen AHL 19/20".
- 20 Un "polipéptido AHL 19/20" como se define aquí se refiere a cualquier polipéptido que contiene un dominio que tiene al menos 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con un Dominio Conservado (DC) como se representa mediante la SEQ ID NO: 36 (comprendida en la SEQ ID NO: 2).
- 25 Alternativa o adicionalmente, un "polipéptido AHL 19/20" como se define aquí se refiere a cualquier polipéptido que comprende: (i) un motivo que tiene al menos 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con un motivo AT-hook como el representado por medio de la SEQ ID NO: 37; y (ii) un dominio que tiene al menos 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con un dominio conservado de planta y de procariota (PPC) como el representado mediante la SEQ ID NO: 38.
- 30 Alternativa o adicionalmente, un "polipéptido AHL 19/20" como se describe aquí se refiere a cualquier polipéptido que comprende: (i) una señal de localización nuclear; (ii) un motivo de enlazamiento del ADN de AT-hook con una entrada InterPro IPR014476; y (iii) un dominio conservado de planta y de procariota (PPC) con una entrada InterPro IPR005175.
- 35 Alternativa o adicionalmente, un "polipéptido AHL 19/20" como se define aquí se refiere a cualquier secuencia de polipéptidos que cuando se utiliza en la construcción de un árbol filogenético de AHL, tal como aquel descrito en la Figura 1 y en la Figura 2, se agrupa con el grupo de polipéptidos AHL 19/20 que comprende la secuencia de polipéptidos como la reportada por la SEQ ID NO: 2, en vez de con cualquier otro grupo de AHL 19/20.
- 40 Alternativa o adicionalmente, un "polipéptido AHL 19/20" como se define aquí se refiere a cualquier polipéptido que tenga en orden creciente de preferencia al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con el polipéptido AHL 19/20 como el representado por la SEQ ID NO: 2 o con cualquiera de las secuencias de polipéptidos de longitud completa dadas aquí en la Tabla A.
- 45 Un "polipéptido AHL 19/20" como se describe aquí es capaz de complementar una cepa de levadura MLY131 que carece de todos los tres transportadores de amonio de levadura nativa (Hildebrand (2005) J Phycol 41: 105 - 113).
- 50 Los términos "dominio" y "motivo" se definen aquí en la sección de "definiciones". Existen bases de datos especializadas para la identificación de dominios, por ejemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. EUA 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 y Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., páginas 53 - 61, AAAIPress, 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 conjunto de herramientas para el análisis *in silico* de secuencias de proteína está disponible en el servidor de proteómicos 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)). El análisis de la secuencias de polipéptidos de la SEQ ID NO: 2 se presenta más abajo aquí en los Ejemplos 2 y 4. Por ejemplo, un polipéptido AHL 19/20 como se representa mediante la SEQ ID NO: 2 comprende un motivo de enlazamiento del ADN de AT-hook con una entrada InterPro, IPR014476 y un dominio

conservado de planta y de procariotas (PPC), descrito como DUF296 (dominio de función desconocida 296) con una entrada InterPro IPR005175, en la base de datos de dominios InterPro. Los dominios se pueden identificar también utilizando técnicas de rutina, tales como mediante la alineación de secuencias. Uno de tales dominios es el Dominio Conservado (DC) de la SEQ ID NO: 2, como el representado por medio de la SEQ ID NO: 36. El CD comprende un NLS predicho, un motivo de enlazamiento del ADN de AT-hook, y un dominio de PPC, como el representado esquemáticamente en la Figura 3, mostrado en la Figura 4.

Los métodos para el alineamiento de secuencias para comparación son bien conocidos en la técnica; tales métodos incluyen GAP, BESTFIT, BLAST, FASTA y TFASTA. GAP utiliza el algoritmo de Needleman & Wunsch ((1970) J Mol Biol 48: 443 - 453) para hallar el alineamiento global (es decir, expansión de las secuencias completas) de dos secuencias que maximizan la cantidad de coincidencias y minimizan el número de brechas. 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 el análisis con BLAST está disponible al público a través del Centro Nacional para la Información Biotecnológica (NCBI por sus siglas en inglés). Los homólogos se pueden identificar fácilmente utilizando, por ejemplo, el algoritmo de alineación de secuencias múltiples ClustalW (versión 1.83), con los parámetros predeterminados de alineamiento por pares, y un método de asignación de puntajes en porcentaje. Los porcentajes globales de similitud y de identidad también se pueden determinar utilizando uno de los métodos disponibles en el paquete de software MatGAT (Campanella et al., BMC Bioinformatics, 10: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando secuencias de proteínas o de ADN). Se puede realizar una edición manual menor para optimizar el alineamiento entre los motivos conservados, como sería evidente para una persona normalmente capacitada en la técnica. Además, en lugar de utilizar secuencias de longitud completa para la identificación de homólogos, se pueden utilizar también dominios específicos. Los valores de identidad de secuencia se pueden determinar sobre la secuencia completa del ácido nucleico o la secuencia del polipéptido o sobre dominios seleccionados o el(los) motivo(s) conservado(s), utilizando los programas mencionados anteriormente que utilizan los parámetros predeterminados. El Ejemplo 3 describe aquí en la Tabla B el porcentaje de identidad entre el polipéptido AHL 19/20 como el representado mediante la SEQ ID NO: 2 y los polipéptidos AHL 19/20 enumerados en la Tabla A, que está en el intervalo entre 50 y 99% de identidad de secuencia de aminoácidos. En la Tabla B1, se muestra el porcentaje de identidad entre el CD como el representado por la SEQ ID NO: 36 (comprendida en la SEQ ID NO: 2) y el CD de los polipéptidos AHL 19/20 enlistados en la Tabla A del Ejemplo 1, en el intervalo de 70 a 99% de identidad de secuencia de aminoácidos.

La tarea de la predicción de la localización de la proteína subcelular es importante y bien estudiada. Conocer la localización de la proteína ayuda a elucidar su función. Los métodos experimentales para la localización de la proteína varían desde inmunolocalización hasta marcación de proteínas utilizando proteína fluorescente verde (GFP). Tales métodos son precisos si bien son muy laboriosos en comparación con los métodos computacionales. Recientemente se ha realizado un avance importante en la predicción computacional de la localización de la proteína a partir de los datos de la secuencia. Entre los algoritmos bien conocidos por una persona normalmente capacitada en la técnica se encuentran disponibles en las herramientas ExPASy Proteomics de propiedad del Swiss Institute for Bioinformatics, por ejemplo, PSort, TargetP, ChloroP, LocTree, Predotar, LipoP, MITOPROT, PATS, PTS1, SignalP, TMHMM y otras. La identificación de la localización subcelular del polipéptido de la invención se muestra en el Ejemplo 6. Una señal de localización nuclear predicha (NLS) se encuentra en el polipéptido AHL 19/20 de la SEQ ID NO: 2. Una NLS es una o más secuencias cortas de lisinas o argininas positivamente cargadas. En particular, se predice que la SEQ ID NO: 2 de la presente invención se localiza en el compartimiento nuclear de células eucariotas.

Además, los polipéptidos AHL 19/20 útiles en los métodos de la presente invención (al menos en su forma nativa) típicamente, pero no necesariamente, tienen actividad reguladora transcripcional y capacidad para interactuar con otras proteínas. Por lo tanto, los péptidos de AHL 19/20 con actividad reguladora transcripcional reducida, sin actividad reguladora transcripcional, con capacidad de interacción reducida proteína - proteína, o sin capacidad de interacción proteína - proteína, pueden igualmente ser útiles en los métodos de la presente invención. La actividad de enlazamiento del ADN y las interacciones proteína - proteína se pueden determinar fácilmente *in vitro* o *in vivo* usando técnicas bien conocidas en el arte (por ejemplo en Current Protocols in Molecular Biology, Volúmenes 1 y 2, Ausubel et al. (1994), Current Protocols). Para determinar la actividad de enlazamiento del ADN de los polipéptidos AHL 19/20, se encuentran disponibles varios ensayos, tales como ensayos de desplazamiento en gel del enlazamiento del ADN (o ensayos de retardo en gel; Korfhage et al. (1994) Plant C 6: 695 - 708), ensayos de enlazamiento del ADN *in vitro* (Schindler et al. (1993) Plant J 4 (1) : 137 - 150), o activación transcripcional de polipéptidos AHL 19/20 en células de levadura, animal y planta (Halbach et al. (2000) Nucleic Acid Res 28 (18): 3542 - 3550). Las secuencias específicas de enlazamiento del ADN se pueden determinar utilizando la técnica de selección aleatoria de nucleótidos (Viola & González (Mayo 26, 2007) Biochemistry) .

En una forma de realización, la presente invención se ilustra por medio de la transformación de plantas con la secuencia de ácido nucleico representada por la SEQ ID NO: 1, que codifica la secuencia del polipéptido AHL 19/20 de la SEQ ID NO: 2. Sin embargo, la realización de la invención no está restringida a estas secuencias; los métodos de la invención pueden ser llevados a cabo convenientemente utilizando cualquier secuencia de ácido nucleico que

codifica un polipéptido AHL 19/20 como se define en la presente invención.

5 Ejemplos de secuencias de ácido nucleico que codifican polipéptidos AHL 19/20 se exponen en la Tabla A del Ejemplo 1 aquí. Tales secuencias de ácido nucleico son útiles para llevar a cabo los métodos de la invención. Las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1 son ejemplos de secuencias de ortólogos y parálogos del polipéptido AHL 19/20 representadas por la SEQ ID NO: 2, siendo los términos "ortólogos" y "parálogos" como se los define aquí. Otros ortólogos y parálogos pueden ser fácilmente identificados llevando a cabo una búsqueda denominada blast recíproco. Típicamente, esto implica un primer BLAST que involucra el someter una secuencia de consulta a BLAST (por ejemplo utilizando cualquiera de las secuencias enlistadas en la Tabla A del Ejemplo 1) contra cualquier base de datos de secuencias, tal como la base de datos del NCBI disponible al público. BLASTN o TBLASTX (que utilizan valores estándar predeterminados) se utilizan generalmente cuando se parte de una secuencia nucleotídica, y BLASTP o TBLASTN (que utilizan valores estándar predeterminados) cuando se parte de una secuencia de proteína. Los resultados de BLAST opcionalmente se pueden filtrar. Las secuencias de longitud completa ya sea de los resultados filtrados o de los resultados no filtrados, se someten luego nuevamente a BLAST (segundo BLAST) contra las secuencias del organismo a partir del cual se deriva la secuencia de consulta (donde la secuencia de consulta es la SEQ ID NO: 1 o la SEQ ID NO: 2, el segundo BLAST sería por lo tanto contra las secuencias de *Arabidopsis thaliana*). Se comparan luego los resultados del primero y el segundo BLAST. Se identifica un parálogo si una respuesta pertinente de alto rango del primer blast es de la misma especie que aquella de la cual se deriva la secuencia de consulta, un BLAST posterior resulta idealmente luego en la secuencia de consulta como una respuesta pertinente más alta; se identifica un ortólogo si una respuesta pertinente de alto rango en el primer BLAST no proviene de la misma especie que aquella de la cual se deriva la secuencia de consulta, y preferiblemente resulta del BLAST posterior estando la secuencia de consulta entre las respuestas pertinentes más altas.

25 Las respuestas pertinentes de alto rango son aquellas que tienen un valor E bajo. Cuanto más bajo es el valor E, más significativo es el puntaje (o en otras palabras, es más baja la posibilidad de que la respuesta pertinente sea hallada por casualidad). La computación del valor E es bien conocida en la técnica. Además de los valores E, las comparaciones también se les da puntaje por identidad porcentual. La identidad porcentual se refiere a la cantidad de nucleótidos idénticos (o aminoácidos) entre las dos secuencias de ácidos nucleicos (o polipéptidos) comparadas sobre una longitud particular. En el caso de grandes familias, se puede utilizar ClustalW, seguido de un árbol filogenético (*neighbour joining tree*), para ayudar a visualizar el agrupamiento de genes relacionados y para identificar ortólogos y parálogos. Cualquier agrupamiento de la secuencia dentro del grupo que comprende la SEQ ID NO: 2 (polipéptido AHL 19; encerrado en un círculo en las Figuras 1 y 2) se consideraría que cae dentro de la definición anteriormente mencionada de un polipéptido AHL 19/20, y se consideraría adecuado para uso en los métodos de la invención.

35 Las variantes de ácido nucleico también pueden ser útiles para llevar a la práctica los métodos de la invención. Los ejemplos de tales variantes incluyen secuencias de ácidos nucleicos que codifican homólogos y derivados de cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1, siendo los términos "homólogo" y "derivado" como se definieron aquí. También son útiles en los métodos de la invención las secuencias de ácido nucleico que codifican homólogos y derivados de ortólogos o parálogos de cualquiera de las secuencias de polipéptidos que se exponen en la Tabla A del Ejemplo 1. Los homólogos y derivados útiles en los métodos de la presente invención tienen sustancialmente la misma actividad biológica y funcional que la proteína no modificada de la cual se derivan.

45 Variantes de ácido nucleico también pueden ser útiles en la práctica de los métodos de la invención. Los ejemplos de tales variantes incluyen secuencias de ácido nucleico que codifican homólogos y derivados de cualquiera de las secuencias de polipéptidos dadas en la Tabla A del Ejemplo 1, siendo los términos "homólogo" y "derivado" como se define aquí. También útiles en los métodos de la invención son las secuencias de ácido nucleico que codifican homólogos y derivados de ortólogos o parálogos de cualquiera de las secuencias de polipéptidos dadas en la Tabla A del Ejemplo 1. Los homólogos y derivados útiles en los métodos de la presente invención tienen sustancialmente la misma actividad biológica y funcional que la proteína no modificada a partir de la cual se derivan.

50 Otras variantes de ácido nucleico útiles en la práctica de los métodos de la invención incluyen porciones de ácido nucleico que codifican polipéptidos AHL 19/20, secuencias de ácido nucleico que hibridan a secuencias de ácido nucleico que codifican polipéptidos AHL 19/20, variantes de empalme de secuencias de ácido nucleico que codifican polipéptidos AHL 19/20, variantes alélicas de secuencias de ácido nucleico que codifican polipéptidos AHL 19/20 y variantes de secuencias de ácido nucleico que codifican polipéptidos AHL 19/20 obtenidos por medio de transposición de genes. Los términos secuencias de hibridación, variante de empalme, variante alélica y transposición de genes son como se describe aquí.

Las secuencias de ácido nucleico que codifican polipéptidos AHL 19/20 no necesitan ser secuencias de ácido nucleico de longitud completa, ya que el desempeño de los métodos de la invención no confía en el uso de secuencias de ácido nucleico de longitud completa. De acuerdo con la presente invención, se provee un método

para aumentar las características relacionadas con el rendimiento de semillas, en las plantas, que comprende la introducción y la expresión en una planta de una porción de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o una porción de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1.

5 Variantes de la secuencia de ácidos nucleicos que codifican homólogos y derivados de la SEQ ID NO: 46 también pueden ser útiles en la práctica de los métodos de la invención, siendo los términos "homólogo" y "derivado" como se define en este documento. También útiles en los métodos de la invención son secuencias de ácido nucleico que codifican homólogos y derivados de ortólogos o parálogos de la SEQ ID NO: 46. Los homólogos y derivados útiles en los métodos de la presente invención tienen sustancialmente la misma actividad biológica y funcional que la proteína no modificada de la que se derivan.

10 Una porción de una secuencia de ácido nucleico se puede preparar, por ejemplo, haciendo una o más supresiones a la secuencia de ácido nucleico. Las porciones se pueden utilizar en forma aislada o se pueden fusionar a otras secuencias codificadoras (o no codificadoras) con el propósito de producir, por ejemplo, una proteína que combina diferentes actividades. Cuando se fusionan a otras secuencias codificadoras, el polipéptido resultante producido después de la traducción puede ser más grande que el predicho para la porción de proteína.

15 Las porciones útiles en los métodos de la invención, codifican un polipéptido AHL 19/20 como se define aquí, y tienen sustancialmente la misma actividad biológica que las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1. Preferiblemente, la porción es una porción de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o es una porción de una secuencia de ácido nucleico que codifica un ortólogo o un parálogo de cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1. Preferiblemente la porción es, en orden creciente de preferencia al menos de 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 940 nucleótidos consecutivos de longitud, siendo los nucleótidos consecutivos de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o de una secuencia de ácido nucleico que codifica un ortólogo o un parálogo de cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1. Preferiblemente, la porción es una porción de una secuencia nucleica que codifica un polipéptido de una secuencia de polipéptidos que cuando se utiliza en la construcción de un árbol filogenético de AHL, tal como aquella descrita en la Figura 1 o en la Figura 2, agrupaciones con el grupo de polipéptidos AHL 19/20 que comprende la secuencia de polipéptidos representada por la SEQ ID NO: 2 en vez de con cualquier otro grupo de AHL. Más preferiblemente, la porción es una porción de la secuencia de ácido nucleico de la SEQ ID NO: 1.

25 Otra variante de la secuencia de ácido nucleico útil en los métodos de la invención es una secuencia de ácido nucleico capaz de hibridar, en condiciones de rigurosidad reducida, preferiblemente bajo condiciones rigurosas, con una secuencia de ácido nucleico que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20), o con una porción como se define aquí.

30 De acuerdo con la presente invención, en una forma de realización se provee un método para incrementar las características relacionadas con el rendimiento de semillas en plantas, que comprende la introducción y expresión en una planta de una secuencia de ácido nucleico capaz de hibridar con cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o que comprende la introducción y expresión en una planta de una secuencia de ácido nucleico capaz de hibridar con una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1.

35 Las secuencias de hibridación útiles en los métodos de la invención codifican un polipéptido AHL 19/20 como se define aquí, y tienen sustancialmente la misma actividad biológica que las secuencias polipeptídicas presentadas en la Tabla A del Ejemplo 1. Preferiblemente, la secuencia de hibridación es capaz de hibridar con cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o a una porción de cualquiera de estas secuencias, siendo una porción como se definió anteriormente, o donde la secuencia de hibridación es capaz de hibridar con una secuencia de ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1. Preferiblemente, la secuencia de hibridación es capaz de hibridar con una secuencia de ácido nucleico que codifica una secuencia de polipéptidos que cuando se utiliza en la construcción de un árbol filogenético de AHL, tal como el descrito en la Figura 1 o en la Figura 2, se agrupa con el grupo de polipéptidos AHL 19/20 que comprende la secuencia de polipéptidos representada por la SEQ ID NO: 2 en vez de con cualquier otro grupo de AHL. Más preferiblemente, la secuencia de hibridación es capaz de hibridar con una secuencia de ácido nucleico como la representada mediante la SEQ ID NO: 1 o con una porción de la misma.

40 Otra variante de la secuencia de ácido nucleico útil en los métodos de la invención es una variante de empalme que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20) o una variante de empalme como se definió aquí.

5 De acuerdo con la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento de semillas, que comprende introducir y expresar en una planta una variante de empalme de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o una variante de empalme de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1.

10 En una forma de realización, las variantes de empalme son variantes de empalme de una secuencia de ácido nucleico representada por la SEQ ID NO: 1, o una variante de empalme de una secuencia de ácido nucleico que codifica un ortólogo o un parálogo de la SEQ ID NO: 2. Preferiblemente, la variante de empalme es una variante de empalme de una secuencia de ácido nucleico que codifica una secuencia de polipéptidos que cuando se usa en la construcción de un árbol filogenético de AHL, tal como el descrito en la Figura 1 o en la Figura 2, se agrupa con el grupo de polipéptidos AHL 19/20 que comprende la secuencia de polipéptidos representada por la SEQ ID NO: 2 en vez de con cualquier otro grupo de AHL.

15 Otra variante de la secuencia de ácido nucleico útil en la realización de los métodos de la invención es una variante alélica de una secuencia de ácido nucleico que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20) o una variante alélica como se definió aquí.

20 En una forma de realización de acuerdo con la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento de semillas, que comprende introducir y expresar en una planta una variante alélica de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o que comprende introducir y expresar en una planta una variante alélica de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias polipeptídicas presentadas en la Tabla A del Ejemplo 1.

25 Las variantes alélicas útiles en los métodos de la presente invención tienen sustancialmente la misma actividad biológica que el polipéptido AHL 19/20 de la SEQ ID NO: 2 y cualquiera de las secuencias polipeptídicas descritas en la Tabla A del Ejemplo 1. Existen variantes alélicas en la naturaleza, y comprendido dentro de los métodos de la presente invención está 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 una secuencia de ácido nucleico que codifica un ortólogo o un parálogo de la SEQ ID NO: 2. Preferiblemente, la variante alélica es una variante alélica de una secuencia de polipéptidos que cuando se usa en la construcción de un árbol filogenético de AHL, tal como el descrito en la Figura 1 o en la Figura 2, se agrupa con el grupo de polipéptidos AHL 19/20 que comprende la secuencia de polipéptidos representada por la SEQ ID NO: 2 en vez de con cualquier otro grupo de AHL.

30 La transposición de genes o la evolución dirigida también se puede utilizar para generar variantes de secuencias de ácido nucleico que codifican polipéptidos que incrementan el rendimiento seleccionados del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20), siendo el término "transposición de genes" como se definió aquí.

35 En una forma de realización de acuerdo con la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento de semillas, que comprende introducir y expresar en una planta una variante de cualquiera de las secuencias de ácido nucleico presentadas en la Tabla A del Ejemplo 1, o que comprende introducir y expresar en una planta una variante de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de polipéptidos presentadas en la Tabla A del Ejemplo 1, en donde la variante de la secuencia de ácido nucleico se obtiene por medio de transposición de genes.

40 Preferiblemente, la variante de la secuencia de ácido nucleico obtenida por transposición de genes codifica una secuencia de polipéptidos que cuando se usa en la construcción de un árbol filogenético de AHL, tal como el descrito en la Figura 1 o en la Figura 2, se agrupa con el grupo de polipéptidos AHL 19/20 que comprende la secuencia de polipéptidos representada por la SEQ ID NO: 2 en vez de con cualquier otro grupo de AHL.

45 Además, también se pueden obtener variantes de la secuencia de ácido nucleico por mutagénesis dirigida al sitio. Varios métodos se encuentran disponibles para lograr una mutagénesis dirigida al sitio, siendo los más comunes los métodos basados en PCR (Current Protocols in Molecular Biology. Wiley Eds.).

50 Las secuencias de ácido nucleico que codifican polipéptidos AHL 19/20 pueden derivarse de cualquier fuente natural o artificial. La secuencia de ácido nucleico puede ser modificada a partir de su forma nativa en composición y/o ambiente genómico a través de manipulación humana deliberada. Preferiblemente, la secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 es de una planta, preferiblemente además de una planta dicotiledónea, más preferiblemente de la familia Brassicaceae, lo más preferible, la secuencia de ácido nucleico es de *Arabidopsis thaliana*.

El desempeño de los métodos de la invención producen plantas que tienen un incremento de características relacionadas con el rendimiento con relación a las plantas que sirven de control. Los términos "rendimiento" y "rendimiento de semilla" se describen con más detalle en la sección de "definiciones" aquí.

5 Rendimiento de los métodos de la invención produce plantas cultivadas bajo condiciones de estrés abiótico que tienen rasgos mejorados relacionados con el rendimiento en relación con las plantas de control. En particular, el rendimiento de los métodos de la invención produce plantas cultivadas bajo condiciones de estrés abiótico que tienen mayor vigor temprano y mayor rendimiento, especialmente aumento de la biomasa y mayor rendimiento de semilla, en relación con las plantas de control. Los términos "rendimiento" y "rendimiento de semillas" se describen en más detalle en la sección "definiciones" de este documento.

10 La referencia aquí a los rasgos relacionados con rendimiento mejorado se considera que significa un aumento en el vigor temprano y / o en la biomasa (peso) de una o más partes de una planta, que puede incluir partes aéreas (cosechables) y / o partes por debajo del suelo (cosechables). En particular, tales partes cosechables son biomasa y / o semillas, y el rendimiento de los métodos de la invención resulta en plantas que crecen bajo condiciones de estrés abiótico que tienen mayor vigor inicial, biomasa y / o rendimiento de semillas con relación al vigor inicial, biomasa o rendimiento de semillas de las plantas de control cultivadas en condiciones comparables.

15 Rendimiento de los métodos de la invención produce plantas que tienen rasgos mejorados relacionados con el rendimiento. En particular, el rendimiento de los métodos de la invención produce plantas que tienen mayor rendimiento, especialmente mayor rendimiento de semillas con relación a las plantas de control. Los términos "rendimiento" y "rendimiento de semillas" se describen en más detalle en la sección "definiciones" de este documento.

La referencia aquí a rasgos mejorados relacionados con el rendimiento se entiende que significa un incremento en biomasa (peso) de una o más partes de una planta, que puede incluir partes aéreas (cosechables) y / o partes por debajo del suelo (cosechable). En particular, tales partes cosechables son semillas.

25 La referencia aquí a los rasgos mejorados relacionados con el rendimiento se entiende que significa un incremento en biomasa (peso) de una o más partes de una planta, que puede incluir partes aéreas (cosechables) y / o partes por debajo del suelo (cosechable). En particular, dichas partes cosechables son semillas, y el rendimiento de los métodos de la invención resulta en plantas que tienen mayor rendimiento de semillas con relación al rendimiento de semillas de las plantas de control.

30 Tomando el maíz como ejemplo, un incremento de rendimiento se puede manifestar como uno o más de lo siguiente: aumento en el número de plantas establecidas por hectárea o acre, un aumento en el número de mazorcas por planta, un aumento en el número de filas, en el número de granos por fila, peso del grano, peso de mil granos, longitud/diámetro de la mazorca, aumento en la tasa de llenado de la semilla (que es el número de semillas llenas dividido por el número total de semillas y multiplicado por 100), entre otros. Tomando el arroz como ejemplo un aumento en el rendimiento se puede manifestar en sí mismo como un aumento en uno o más de lo siguiente:

35 número de plantas por hectárea o acre, número de panículas por planta, número de espiguillas por panícula, número de flores (floretes) por panícula, que se expresa como una relación del número de semillas llenas sobre el número de panículas primarias), aumento en la tasa de llenado de semilla (que es el número de semillas llenas dividido por el número total de semillas y multiplicado por 100), aumento en el peso de mil granos, entre otros.

40 En una forma de realización, la presente invención proporciona un método para incrementar las características relacionadas con el rendimiento de semillas de las plantas con relación a las plantas que sirven de control, cuyo método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como se definió aquí.

45 Ya que las plantas transgénicas de acuerdo con la presente invención tienen un aumento de características relacionadas con el rendimiento, es probable que estas plantas exhiban una mayor tasa de crecimiento (durante al menos parte de su ciclo de vida), con relación a la tasa de crecimiento de las plantas que sirven como control en una etapa correspondiente en su ciclo de vida.

50 En una forma de realización, la presente invención proporciona un método para mejorar las características relacionadas con el rendimiento de una planta bajo condiciones de crecimiento de estrés abiótico, especialmente el rendimiento de biomasa y / o de semillas de las plantas, en relación con las plantas de control cultivadas en condiciones comparables, cuyo método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido de GRP como se define aquí.

Ya que las plantas transgénicas de acuerdo con la presente invención que crecen bajo condiciones de estrés abiótico tienen un aumento de características relacionadas con el rendimiento, es probable que estas plantas exhiban una mayor tasa de crecimiento (durante al menos parte de su ciclo de vida), con relación a la tasa de

crecimiento de las plantas que sirven como control en una etapa correspondiente en su ciclo de vida, y bajo condiciones de crecimiento comparables.

5 Ya que las plantas transgénicas de acuerdo con la presente invención tienen mayor rendimiento, es probable que estas plantas exhiban una mayor tasa de crecimiento (durante al menos parte de su ciclo de vida), con relación a la tasa de crecimiento de las plantas que sirven como control en una etapa correspondiente en su ciclo de vida.

10 La mayor tasa de crecimiento puede ser específica para una o más partes de una planta (incluyendo las semillas), o puede ser sustancialmente a través de la planta completa. Las plantas con una mayor velocidad de crecimiento pueden tener un ciclo vital más corto. El ciclo de vida de una planta puede entenderse como el tiempo necesario para desarrollarse desde una semilla madura seca hasta la etapa donde la planta ha producido semillas maduras secas, en forma similar al material de partida. Este ciclo de vida puede estar influido por factores tales como el vigor temprano, la velocidad de crecimiento, el índice de verdor, el tiempo de floración y la velocidad de maduración de la semilla. El incremento en la velocidad del 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. La mayor velocidad de crecimiento durante las etapas tempranas en el ciclo de vida de una planta puede reflejar un mayor vigor (temprano). El incremento en la velocidad 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 de otro modo (se puede obtener un efecto similar con un tiempo de floración más temprano; la demora en la floración no es usualmente una característica deseable en los cultivos). Si la velocidad del crecimiento se incrementa lo suficiente, puede permitir el sembrado posterior de semillas de la misma especie de planta (por ejemplo el sembrado y cosecha de plantas de arroz seguido por el sembrado y cosecha de otras plantas de arroz todo dentro de un periodo de crecimiento convencional). De manera similar, si se incrementa suficientemente la velocidad de crecimiento, permitirá el posterior sembrado de diferente especies de planta (por ejemplo sembrado y cosecha de plantas de maíz, seguido por ejemplo, del sembrado y de cosecha opcional de soja, patata o cualquier otra planta adecuada). También se puede cosechar varias veces a partir del mismo rizoma en el caso de algunas plantas de cultivo. La alteración del ciclo de cosecha de una planta puede conducir a un incremento en la producción anual de biomasa por acre (debido a un incremento en el número de veces (digamos en un año) que cualquier planta particular puede ser cultivada y cosechada). Un incremento en la velocidad de crecimiento también puede permitir el cultivo de plantas transgénicas en un área geográfica más amplia que sus contrapartes de tipo silvestre, dado que las limitaciones territoriales para el crecimiento de un cultivo a menudo están determinadas por condiciones ambientales adversas ya sea en el momento de la plantación (a principios de la estación) o en el momento de la cosecha (finales de la estación). Estas condiciones adversas se pueden evitar si se acorta el ciclo de la cosecha. La velocidad de crecimiento puede determinarse por medio de la derivación de diferentes parámetros a partir de las curvas de crecimiento; dichos parámetros pueden ser: T-Mid (el tiempo que les toma a las plantas alcanzar 50% de su tamaño máximo) y T-90 (el tiempo que les toma a las plantas alcanzar 90% de su tamaño máximo), entre otros. La tasa de crecimiento definida aquí no pretende significar retraso en la floración.

40 De acuerdo con una característica preferida de la presente invención, el desempeño de los métodos de la invención produce plantas que tienen una mayor tasa de crecimiento con respecto a las plantas de control. Por lo tanto, de acuerdo con la presente invención, se provee un método para incrementar la tasa de crecimiento de las plantas, en donde dicho método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como se definió aquí.

45 Un incremento de las características relacionadas con el rendimiento de semillas se presenta si la planta está bajo condiciones no estresadas o si la planta es expuesta a varios tipos de estrés en comparación con las plantas control cultivadas bajo condiciones comparables. Una mejora de las características relacionadas con el rendimiento (un incremento en el rendimiento de semillas y/o en la tasa de crecimiento) se presenta si la planta está bajo condiciones no estresadas o si la planta es expuesta a diferentes tipos de estrés comparado con las plantas de control.

50 En una forma de realización particularmente preferida, los métodos de la presente invención se llevan a cabo bajo condiciones no estresadas. Sin embargo, un incremento en el rendimiento y/o en la tasa de crecimiento se presenta si la planta está bajo condiciones no estresadas o si la planta es expuesta a diferentes tipos de estrés comparado con las plantas de control.

55 Las plantas típicamente responden a la exposición al estrés creciendo en forma más lenta. En condiciones de estrés severo, la planta puede incluso detener el crecimiento completamente. Un estrés moderado por otro lado se define aquí como cualquier estrés al cual se expone una planta que no produce como resultado que la planta cese su crecimiento completamente sin la capacidad de reanudarlo. Un estrés moderado en el sentido de la invención conduce a una reducción en el crecimiento de las plantas estresadas de menos del 40%, 35% o 30%, preferiblemente menos del 25%, 20% o 15%, más preferiblemente menos de 14%, 13%, 12%, 11% o 10% o menos en comparación con las plantas control bajo condiciones no estresantes. Debido a los avances en las prácticas agrícolas (irrigación, fertilización, tratamientos con pesticidas) no se encuentran a menudo estreses severos en las

5 plantas de cultivo sembradas. En consecuencia, el crecimiento comprometido inducido por un estrés moderado a menudo es una característica indeseable para la agricultura. Los estreses moderados son los estreses bióticos y/o
 10 abióticos (ambientales) de cada temporada a los cuales está expuesta una planta. Los estreses abióticos pueden deberse a sequía o exceso de agua, estrés anaeróbico, estrés por salinidad, toxicidad química, estrés oxidativo y temperaturas calurosas, frías o congelantes. El estrés abiótico puede ser un estrés osmótico causado por un estrés por agua (particularmente debido a sequía), estrés por salinidad, estrés oxidativo o un estrés iónico. Los estreses bióticos son típicamente aquellos estreses causados por patógenos, tales como bacterias, virus, hongos, nemátodos, e insectos. El término condiciones "no estresantes" como se usa aquí son aquellas condiciones ambientales que permiten el óptimo crecimiento de las plantas. Las personas normalmente capacitadas en el arte son conscientes de las condiciones normales de suelo y las climáticas para una ubicación dada.

15 De acuerdo a lo reportado en Wang et al. (Planta (2003) 218: 1 - 14), un estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan adversamente el crecimiento y la productividad de la planta. Se sabe que la sequía, salinidad, temperaturas extremas y estrés oxidativo están interconectados y pueden inducir daño celular y de crecimiento a través de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755 - 1767) describe un grado particularmente elevado de "interferencia" entre el estrés por sequía y el estrés por alta salinidad. Por ejemplo, la sequía y/o la salinización se manifiestan principalmente como estrés osmótico, produciendo la interrupción de la homeostasis y la distribución de iones en la célula. El estrés oxidativo, que con frecuencia acompaña a temperaturas altas o bajas, al estrés por salinidad o por sequía, puede causar desnaturalización de proteínas funcionales y estructurales. Como consecuencia, estos diversos estreses ambientales a menudo activan rutas de señalización celular y respuestas celulares similares, tales como la producción de proteínas por estrés, favorecimiento de la expresión de antioxidantes, acumulación de solutos compatibles y detención del crecimiento. El término condiciones "no estresantes" como se usa aquí son aquellas condiciones ambientales que permiten el óptimo crecimiento de las plantas. Las personas normalmente capacitadas en el arte son conscientes de las condiciones normales de suelo y las climáticas para una ubicación dada.

25 En una forma de realización, el desempeño de los métodos de la invención produce plantas cultivadas bajo condiciones no estresantes o bajo condiciones de estrés moderado que tienen más características relacionadas con el rendimiento, con respecto a plantas de control cultivadas bajo condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento en plantas cultivadas bajo condiciones no estresantes o bajo condiciones de estrés moderado, en donde el método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20.

35 En una forma de realización, el desempeño de los métodos de la invención produce plantas cultivadas bajo condiciones de estrés moderado que tienen más características relacionadas con el rendimiento, con respecto a plantas de control cultivadas bajo condiciones comparables. Por lo tanto, de acuerdo con una forma de realización de la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento en plantas cultivadas bajo condiciones de estrés moderado, en donde el método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido de GRP.

40 El desempeño de los métodos de acuerdo con la presente invención da como resultado plantas que crecen bajo condiciones de estrés abiótico que tienen más características relacionadas con el rendimiento para controlar a las plantas que crecen bajo condiciones comparables de estrés. De acuerdo a lo reportado en Wang et al. (Planta (2003) 218: 1 - 14), un estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan adversamente el crecimiento y la productividad de la planta. Se sabe que la sequía, salinidad, temperaturas extremas y estrés oxidativo están interconectados y pueden inducir daño celular y de crecimiento a través de mecanismos similares. Rabbani et al. (Plant Physiol (2003) 133: 1755 - 1767) describe un grado particularmente elevado de "interferencia" entre el estrés por sequía y el estrés por alta salinidad. Por ejemplo, la sequía y/o la salinización se manifiestan principalmente como estrés osmótico, produciendo la interrupción de la homeostasis y la distribución de iones en la célula. El estrés oxidativo, que con frecuencia acompaña a temperaturas altas o bajas, al estrés por salinidad o por sequía, puede causar desnaturalización de proteínas funcionales y estructurales. Como consecuencia, estos diversos estreses ambientales a menudo activan rutas de señalización celular y respuestas celulares similares, tales como la producción de proteínas por estrés, favorecimiento de la expresión de antioxidantes, acumulación de solutos compatibles y detención del crecimiento. Ya que diversos tipos de estrés ambiental activan rutas similares, los ejemplos de la presente invención con estrés por sequía no deben ser considerados como una limitación al estrés por sequía, sino más como una selección para indicar la participación de polipéptidos AHL 19/20 como se definió anteriormente, en el incremento de las características relacionadas con el rendimiento con relación a las plantas que sirven como control cultivadas en condiciones de estrés comparables, en estreses abióticos en general.

Ya que diversos estreses ambientales activan rutas similares, los ejemplos de la presente invención con estrés por sequía y estrés por salinidad no deben ser considerados como una limitación al estrés por sequía o al estrés por salinidad, sino más como una selección para indicar la participación de polipéptidos de GRP como se definió

anteriormente, en el mejoramiento de las características relacionadas con el rendimiento con relación a las plantas que sirven como control cultivadas en condiciones de estrés comparables, en estreses abióticos en general.

5 El término "estrés abiótico" como se define aquí se entiende que significa uno o más de: estrés por agua (causado por sequía o exceso de agua), estrés anaeróbico, estrés por salinidad, estrés por temperatura (causada por temperaturas elevadas, frías o congelantes), estrés por toxicidad química y estrés oxidativo. De acuerdo con un aspecto de la invención, el estrés abiótico es un estrés osmótico, seleccionado de estrés por agua, estrés por salinidad, estrés oxidativo y estrés iónico. Preferiblemente, el estrés por agua es estrés por sequía. El término estrés por salinidad no se restringe a la sal común (NaCl), sino que puede ser cualquier estrés causado por una o más de: NaCl, KCl, LiCl, MgCl₂, CaCl₂, entre otras.

10 Preferiblemente, el estrés abiótico es estrés por salinidad. Alternativamente, el estrés abiótico es estrés por salinidad.

15 En una forma de realización, el desempeño de los métodos de la invención produce plantas que tienen más características relacionadas con el rendimiento de semillas, bajo condiciones de estrés abiótico con respecto a las plantas de control cultivadas bajo condiciones de estrés comparables. Por lo tanto, de acuerdo con la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento de semillas, en plantas cultivadas bajo condiciones de estrés abiótico, en donde el método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20. De acuerdo con un aspecto de la invención, el estrés abiótico es un estrés osmótico, seleccionado de uno o más de los siguientes: estrés hídrico, estrés por salinidad, estrés oxidativo y estrés iónico.

20 En una forma de realización el desempeño de los métodos de la invención produce plantas que crecen bajo condiciones de estrés abiótico que tienen características mejoradas relacionadas con el rendimiento en relación con las plantas de control cultivadas en condiciones de estrés comparable. Por lo tanto, de acuerdo con la presente invención, se proporciona un método para mejorar las características relacionados con el rendimiento en plantas cultivadas bajo condiciones de estrés abiótico, cuyo método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido de GRP. De acuerdo con un aspecto de la invención, el estrés abiótico es un estrés osmótico, seleccionado de uno o más de los siguientes: estrés hídrico, estrés por salinidad, estrés oxidativo y estrés iónico. Preferiblemente, el estrés abiótico es estrés por sequía. Alternativa o adicionalmente, el estrés abiótico es estrés por salinidad.

30 Otro ejemplo de estrés ambiental abiótico es la disponibilidad reducida de uno o más nutrientes que deben ser asimilados por las plantas para su crecimiento y desarrollo. Debido a la fuerte influencia de la eficiencia de la utilización de la nutrición sobre el rendimiento de la planta y la calidad del producto, se vierte una enorme cantidad de fertilizante en los campos para optimizar el crecimiento y calidad de las plantas. La productividad de las plantas en general está limitada por tres nutrientes principales, fósforo, potasio y nitrógeno, donde usualmente entre estos tres se encuentra el elemento limitante de la velocidad de crecimiento de las plantas. Por lo tanto el elemento nutricional principal requerido para el crecimiento de las plantas es el nitrógeno (N). Es un constituyente de numerosos compuestos importantes hallados en las células vivientes, incluidos aminoácidos, proteínas (enzimas), ácidos nucleicos y clorofila. 1,5% a 2% del material seco de una planta es nitrógeno y aproximadamente 16% de la proteína total de la planta. Por lo tanto, la disponibilidad de nitrógeno es un factor limitante principal para el crecimiento y producción de las plantas de cultivo (Frink et al. (1999) Proc Natl Acad Sci EUA 96(4): 1175 - 1180), y tiene también un impacto principal en la acumulación de proteínas y composición de los aminoácidos. Por lo tanto, son de gran interés las plantas de cultivo con más características relacionadas con el rendimiento, cuando se cultivan bajo condiciones limitantes de nitrógeno.

45 En una forma de realización, el desempeño de los métodos de la invención produce plantas que crecen bajo condiciones de disponibilidad reducida de nutrientes, particularmente bajo condiciones de disponibilidad reducida de nitrógeno, que tienen más características relacionadas con el rendimiento de semillas con respecto a las plantas de control cultivadas bajo condiciones de estrés comparables. Por lo tanto, de acuerdo con la presente invención, se provee un método para incrementar las características relacionadas con el rendimiento en plantas cultivadas bajo condiciones de disponibilidad reducida de nutrientes, preferiblemente disponibilidad reducida de nitrógeno, en donde el método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20. La disponibilidad reducida de nutrientes puede resultar de una deficiencia o exceso de nutrientes tales como nitrógeno, fosfatos y otros compuestos que contienen fósforo, potasio, calcio, cadmio, magnesio, manganeso, hierro y boro, entre otros. Preferiblemente la disponibilidad reducida de nutrientes es disponibilidad reducida de nitrógeno.

55 En una forma de realización, el desempeño de los métodos de la invención produce plantas que crecen bajo condiciones de disponibilidad reducida de nutrientes, particularmente bajo condiciones de disponibilidad reducida de nitrógeno, que tienen características mejoradas relacionadas con el rendimiento con respecto a las plantas de control cultivadas bajo condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se provee un

5 método para mejorar las características relacionadas con el rendimiento en plantas cultivadas bajo condiciones de disponibilidad reducida de nutrientes, en donde el método comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido de GRP. La disponibilidad reducida de nutrientes puede comprender una disponibilidad reducida de nutrientes tales como nitrógeno, fosfatos y otros compuestos que contienen fósforo, potasio, calcio, cadmio, magnesio, manganeso, hierro y boro, entre otros.

10 La presente invención abarca plantas o partes de las mismas (incluidas semillas) o células de las mismas que pueden ser obtenidas mediante los métodos de acuerdo con la presente invención. Las plantas o partes de las mismas o células de las mismas comprenden un transgén de ácido nucleico que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20) como se definió anteriormente.

15 La invención también provee construcciones y vectores genéticos para facilitar la introducción y/o mayor expresión en plantas de secuencias de ácido nucleico que codifican polipéptidos que incrementan el rendimiento seleccionados del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20). Las construcciones génicas se pueden insertar en vectores, que pueden encontrarse comercialmente disponibles, adecuados para la transformación en plantas y para la expresión del gen de interés en las células transformadas. La invención también provee el uso de una construcción génica como se define aquí en los métodos de la invención.

En forma más específica, la presente invención provee una construcción que comprende:

20 (a) una secuencia de ácido nucleico que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: uno localizado en el núcleo 19/20 con el motivo AT-hook (AHL 19/20) como se definió anteriormente;

(b) una o más secuencias de control capaces de incrementar la expresión de la secuencia del ácido nucleico de (a); y opcionalmente

(c) una secuencia de terminación de la transcripción.

25 En una forma de realización, la secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 es como se definió anteriormente. El término "secuencia control" y "secuencia de terminación" son como se definen aquí.

Preferiblemente, una de las secuencias control de una construcción es un promotor constitutivo aislado de un genoma de una planta. Un ejemplo de un promotor constitutivo es un promotor GOS2, preferiblemente un promotor GOS2 de arroz, más preferiblemente un promotor GOS2 como el representado mediante la SEQ ID NO: 35.

30 Las plantas se transforman con un vector que comprende cualquiera de las secuencias de ácido nucleico descritas anteriormente. La persona calificada en la técnica es muy consciente de los elementos genéticos que deben estar presentes en el vector a fin de transformar, seleccionar y propagar exitosamente las células huéspedes que contienen la secuencia de interés. La secuencia de interés está operativamente enlazada a una o más secuencias de control (al menos a un promotor).

35 Convenientemente, se puede utilizar cualquier tipo de promotor, ya sea natural o sintético, para incrementar la expresión de la secuencia de ácido nucleico. Un promotor constitutivo es particularmente útil en los métodos, preferiblemente un promotor constitutivo aislado de un genoma de una planta. El promotor constitutivo de una planta dirige la expresión de una secuencia codificadora en un nivel que está en todos los casos por debajo de aquel obtenido bajo el control de un promotor viral 35S del CaMV.

40 Otros promotores específicos del órgano, por ejemplo para la expresión preferida en hojas, tallos, tubérculos, meristemas, semillas (embrión y/o endospermo), son útiles para llevar a cabo los métodos de invención. Véase la sección "Definiciones" aquí para las definiciones de los diferentes tipos de promotores.

45 Debe aclararse que la aplicabilidad de la presente invención no está restringida a una secuencia de ácido nucleico que codifica el polipéptido AHL 19/20, como el representado mediante la SEQ ID NO: 1, ni la aplicabilidad de la invención está restringida a la expresión de una secuencia de ácido nucleico que codifica el polipéptido AHL 19/20 cuando está dirigida por un promotor constitutivo.

El promotor constitutivo es preferiblemente un promotor GOS2, preferiblemente un promotor GOS2 de arroz. Preferiblemente además, el promotor GOS2 está representado por una secuencia de ácido nucleico sustancialmente similar a la SEQ ID NO: 47, lo más preferible, el promotor GOS2 es como el representado por la SEQ ID NO: 47. Véase la Tabla 2 en la sección de "Definiciones" aquí para otros ejemplos de promotores constitutivos.

- 5 Véase aquí la sección "Definiciones" para las definiciones de los distintos tipos de promotores. Particularmente útil en los métodos de la invención es un promotor específico de la raíz, en particular un promotor específico de la epidermis de la raíz. El promotor específico de la raíz es preferiblemente un promotor transportador de nitrato, más preferiblemente de arroz (promotor Os NRT1 como lo describe Lin, 2000). El promotor está representado por la SEQ ID NO: 59. Una secuencia de ácido nucleico sustancialmente similar a la SEQ ID NO: 59 también sería útil en los métodos de la invención. Ejemplos de otros promotores específicos de la raíz, que también pueden ser utilizados para llevar a cabo los métodos de la invención se muestran en la Tabla 2b en la sección "Definiciones" más arriba.
- 10 Opcionalmente, se pueden utilizar una o más secuencias terminadoras en la construcción introducida en una planta. Los elementos reguladores adicionales pueden incluir mejoradores de la transcripción así como también de traducción. Las personas normalmente capacitadas en la técnica serán conscientes de las secuencias terminadoras y mejoradores que puedan ser adecuadas para uso en la realización de la invención. También se puede añadir una secuencia de intrón a la región no traducida (UTR) 5' o en la secuencia de codificación para incrementar la cantidad del mensaje maduro que se acumula en el citosol, como se describe en la sección de las definiciones. Otras
- 15 secuencias control (además de las secuencias promotoras, mejoradoras, silenciadoras, del intrón, regiones 3'UTR y/o 5'UTR) pueden ser proteínas y/o elementos estabilizantes del ARN. Tales secuencias serían conocidas o pueden ser fácilmente obtenidas por una persona normalmente capacitada en la técnica.
- 20 Las construcciones genéticas divulgadas en la invención pueden incluir además un origen de secuencia de replicación que se requiere para el mantenimiento y/o replicación en un tipo de célula específico. Un ejemplo es cuando se requiere mantener una construcción genética en una célula bacteriana como un elemento genético episomal (por ejemplo, una molécula de plásmido o de cósmido). Los orígenes de replicación preferidos incluyen, pero no se limitan al f1-ori y colE1.
- 25 Para la detección de la transferencia exitosa de las secuencias de ácido nucleico tal como se utiliza en los métodos de la invención y/o selección de planta transgénicas que comprende estas secuencias de ácido nucleico, es conveniente utilizar genes marcadores (o genes reporteros). 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.
- 30 Los genes marcadores pueden ser removidos o cortados de la célula transgénica una vez no se los necesite más. Las técnicas para remoción del marcador se conocen en la técnica. Se describen técnicas útiles más arriba en la sección de definiciones.
- 35 Se sabe que tras la integración estable o transitoria de secuencias de ácido nucleico en las células de la planta, sólo una minoría de las células incorpora el ADN foráneo 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 para un marcador seleccionable (tal como los descritos anteriormente) es usualmente introducido en las células huésped junto con el gen de interés. Estos marcadores, por ejemplo, pueden utilizarse en mutantes en los cuales estos genes no son funcionales, por ejemplo, por medio de la supresión mediante métodos convencionales. Además, se pueden introducir moléculas de secuencias de ácido nucleico que codifican un
- 40 marcador seleccionable en una célula huésped sobre el mismo vector que comprende la secuencia que codifica los polipéptidos de la invención o que se utiliza en los métodos de la invención, o bien en un vector separado. Las células que han sido transfectadas en forma estable con la secuencia introducida de ácido nucleico se pueden identificar 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 removidos o escindidos de la célula transgénica una vez que no se los necesita. Las técnicas para remoción del gen marcador son conocidas en el arte; las técnicas útiles fueron descritas anteriormente en la sección de definiciones.
- 45 En una forma de realización, la invención también provee un método para la producción de plantas transgénicas que tienen más características relacionadas con el rendimiento de semillas con relación a las plantas que sirven como control, que comprende la introducción y expresión en una planta de cualquier secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como se definió aquí anteriormente.
- 50 Más específicamente, la presente invención provee un método para la producción de plantas transgénicas que tienen más características relacionadas con el rendimiento de semillas con respecto a las plantas que sirven como control, el cual comprende:
- (i) la introducción y expresión en una planta, parte de una planta, o una célula de una planta, de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20, bajo el control de un promotor constitutivo de una planta; y
- (ii) el cultivo de la célula de la planta, parte de la planta o la planta en condiciones que promuevan el crecimiento y desarrollo de la planta.

La secuencia de ácido nucleico de (i) puede ser cualquiera de las secuencias de ácido nucleico capaces de codificar un polipéptido AHL 19/20 como se definió aquí.

5 La secuencia de ácido nucleico se puede introducir directamente en una célula de una planta o dentro de la planta misma (incluyendo la introducción en un tejido, órgano o cualquier otra parte de una planta). De acuerdo con una característica preferida de la presente invención, se introduce preferiblemente la secuencia de ácido nucleico en una planta mediante transformación. El término "transformación" se describe en forma más detallada, en la sección "definiciones" en la presente descripción.

10 Las células de la planta genéticamente modificadas se pueden regenerar mediante todos los métodos con los cuales está familiarizado la persona capacitada en la técnica. Los métodos adecuados pueden encontrarse en las publicaciones antes mencionadas por S. D. Kung y R. Wu, Potrykus o Höfgen y Willmitzer.

15 En general después de la transformación, las células de la planta o los agrupamientos celulares se seleccionan para determinar la presencia de uno o más marcadores que son codificados por genes que pueden expresarse en una planta transferidos junto con el gen de interés, después de lo cual se regenera el material transformado en una planta entera. Para seleccionar las plantas transformadas, el material vegetal obtenido en la transformación es sometido, como regla, a condiciones selectivas de modo que las plantas transformadas puedan distinguirse de las plantas no transformadas. Por ejemplo, las semillas obtenidas en la forma antes descrita se pueden plantar y, después de un período de crecimiento inicial, ser sometidas a una selección adecuada por medio de rociado. Otra posibilidad adicional consiste en el crecimiento de las semillas, si procede después de esterilización, sobre placas de agar utilizando un agente de selección adecuado de modo que únicamente las semillas transformadas pueden convertirse en plantas. En forma alternativa, se seleccionan las plantas transformadas para determinar la presencia de un marcador seleccionable tal como las descritas anteriormente.

25 Después de la transferencia y regeneración del ADN, las plantas putativamente transformadas también se pueden evaluar, por ejemplo utilizando el análisis tipo Southern, para determinar la presencia del gen de interés, el número de copias y/o organización genómica. En forma alternativa o adicional, se pueden monitorear los niveles de expresión del ADN recientemente introducido utilizando análisis tipo Northern y/o tipo Western, siendo ambas técnicas bien conocidas por las personas ordinariamente capacitadas en la técnica.

30 Las plantas transformadas generadas se pueden propagar mediante una variedad de medios, tales como por medio de técnicas de propagación clonal o de fitomejoramiento clásico. Por ejemplo, una primera generación de la planta transformada (o T1) puede ser autofecundada y los transformantes homocigotos de segunda generación (o T2) seleccionados, y las plantas T2 pueden ser luego adicionalmente propagadas a través de las técnicas clásicas de fitomejoramiento. Los organismos transformados generados pueden adoptar una variedad de formas. Por ejemplo, pueden ser quimeras de células transformadas y de células no transformadas; transformantes clonales (por ejemplo, todas las células transformadas para contener el casete de expresión); injertos de tejidos transformados y sin transformar (por ejemplo, en plantas, un rizoma transformado injertado a un vástago sin transformar).

35 La presente invención claramente se extiende a cualquier célula de una planta o planta producida mediante uno cualquiera de los métodos descritos aquí, y a todas las partes de la planta y propágulos de la misma. La presente invención se extiende además para abarcar la progenie de una célula primaria transformada o transfectada, tejido, órgano o planta completa que ha sido producida mediante cualquiera de los métodos antes mencionados, siendo el único requisito que la progenie exhiba la(s) misma(s) característica(s) fenotípica(s) y/o genotípica(s) que las producidas por la progenie en los métodos de acuerdo con la invención.

40 En una forma de realización, la invención también incluye células huésped que contienen una secuencia aislada de ácido nucleico que codifica un polipéptido AHL 19/20 como se definió aquí anteriormente, operativamente enlazada a un promotor constitutivo.

45 Las células huésped preferidas de acuerdo con la invención son las células de la planta. Las plantas huésped para los ácidos nucleicos o el vector utilizado en el método de acuerdo con la invención, el casete de expresión o la construcción o el vector están, en principio, en forma conveniente en todas las plantas, las cuales son capaces de sintetizar los polipéptidos utilizados en el método de la invención.

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

55

- La invención también se extiende a las partes cosechables de una planta que comprende una secuencia aislada de ácido nucleico que codifica un AHL 19/20 (como se definió aquí anteriormente) operativamente enlazado a un promotor constitutivo de una planta, tal como, pero sin limitarse a semillas, hojas, frutos, flores, tallos, rizomas, tubérculos y bulbos. La invención se refiere además a productos derivados, preferiblemente directamente derivados de una parte cosechable de dicha planta, tal como gránulos o polvos secos, aceite, grasa y ácidos grasos, almidón o proteínas.
- Los métodos para incrementar la expresión de las secuencias de ácido nucleico o los genes, o productos génicos, están bien documentados en la técnica y se proveen ejemplos en la sección de definiciones.
- Como se mencionó anteriormente, un método preferido para incrementar la expresión de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 es por medio de la introducción y expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20; sin embargo, los efectos de llevar a cabo el método, es decir, de incrementar las características relacionadas con el rendimiento, también pueden lograrse utilizando otras técnicas bien conocidas, que incluyen pero no se limitan a marcación de la activación del ADN-T, TILLING, recombinación homóloga. Una descripción de estas técnicas es suministrada en la sección de definiciones.
- La presente invención abarca también el uso de secuencias de ácido nucleico que codifican polipéptidos AHL 19/20 como se describe aquí; y el uso de estos polipéptidos AHL 19/20 para incrementar cualquiera de las características relacionadas con el rendimiento de semillas anteriormente mencionadas en plantas, bajo condiciones de crecimiento normales, y bajo condiciones de disponibilidad reducida de nutrientes, preferiblemente bajo condiciones de disponibilidad reducida de nitrógeno.
- Las variantes alélicas de una secuencia de un gen/ácido nucleico que codifican un polipéptido que incrementa el rendimiento pueden encontrar uso también en programas de fitomejoramiento asistidos por marcadores. Tales programas de fitomejoramiento algunas veces requieren la introducción de una variante alélica por medio de tratamiento mutagénico de las plantas, utilizando por ejemplo mutagénesis EMS; alternativamente, el programa puede iniciarse con un grupo de variantes alélicas de las así llamadas de origen "natural" originadas en forma no intencional. La identificación de variantes alélicas tiene lugar entonces, por ejemplo, por medio de PCR. Esto es seguido por una etapa para la selección de variantes alélicas superiores de la secuencia en cuestión y que producen más características relacionadas con el rendimiento. La selección se lleva a cabo típicamente monitoreando el desempeño del crecimiento de las plantas que contienen diferentes variantes alélicas de la secuencia en cuestión. El desempeño del crecimiento puede ser monitoreado en un invernadero o en el campo. Otras etapas opcionales incluyen cruzamiento de plantas en las cuales se identificó la variante alélica superior con otra planta. Esto podría ser utilizado, por ejemplo, para elaborar una combinación de características fenotípicas interesantes.
- Los ácidos nucleicos que codifican un polipéptido que incrementa el rendimiento pueden ser utilizados también como sondas para mapear genética y físicamente los genes de los cuales hacen parte, y como marcadores para características enlazadas a esos genes. Tal información puede ser útil en fitomejoramiento de plantas con el propósito de desarrollar líneas con fenotipos deseados. Tal uso de las secuencias de ácido nucleico que codifican polipéptidos que incrementan el rendimiento requiere únicamente una secuencia de ácido nucleico de al menos 15 nucleótidos de longitud. Los ácidos nucleicos que codifican un polipéptido que incrementa el rendimiento se pueden utilizar como marcadores de polimorfismo con la longitud del fragmento de restricción (RFLP). Las transferencias tipo Southern (Sambrook J, Fritsch EF y Maniatis T (1989) *Molecular Cloning, A Laboratory Manual*) del ADN genómico de la planta digerido con enzimas de restricción pueden ser sondeadas con los ácidos nucleicos del tipo AAT. Los patrones de bandas resultantes pueden ser luego sometidos a análisis genético utilizando programas de computador tales como MapMaker (Lander et al. (1987) *Genomics* 1: 174 - 181) con el propósito de construir un mapa genético. Además, los ácidos nucleicos pueden ser utilizados para examinar las transferencias tipo Southern que contienen los ADN genómicos tratados con endonucleasa de restricción de un conjunto de individuos que representan al progenitor y a la progenie de un cruce genético definido. Se observa la segregación de los polimorfismos de ADN y se utilizan para calcular la posición del ácido nucleico que codifica un polipéptido que incrementa el rendimiento en el mapa genético previamente obtenido utilizando esta población (Botstein et al. (1980) *Am. J. Hum. Genet.* 32: 314 - 331).
- La producción y el uso de las sondas derivadas de genes de la planta para uso en mapeo genético se describe en Bernatzky y Tanksley (1986) *Plant Mol. Biol. Reporter* 4: 37 - 41. Numerosas publicaciones describen el mapeo genético de clones específicos de ADNc utilizando la metodología expuesta anteriormente o variantes de la misma. Por ejemplo, poblaciones de entrecruzamientos F2, poblaciones de retrocruzamiento, poblaciones apareadas al azar, líneas isogénicas cercanas, y otros conjuntos de individuos pueden ser utilizados para el mapeo. Estas metodologías son bien conocidas por aquellos ordinariamente capacitados en la técnica.
- Las sondas de las secuencias de ácido nucleico también se pueden utilizar para mapeo físico (es decir, colocación de secuencias sobre mapas físicos; véase Hoheisel et al. En: *Non-mammalian Genomic Analysis: A Practical Guide*, Academic Press 1996, páginas 319 - 346, y las referencias citadas allí).

- 5 En otra realización, las sondas de secuencias de ácido nucleico pueden ser utilizadas en el mapeo de la hibridación *in situ* de fluorescencia directa (FISH por su siglas en inglés) (Trask (1991) Trends Genet. 7: 149 - 154). Aunque los métodos actuales de mapeo FISH favorecen el uso de clones grandes (varios kb hasta varios cientos de kb, véase Laan et al. (1995) Genome Res. 5: 13 - 20), las mejoras en sensibilidad pueden permitir la realización del mapeo FISH utilizando sondas más cortas.
- 10 Una variedad de métodos basados en la amplificación de las secuencias de ácido nucleico para mapeo genético y físico pueden ser llevados a cabo utilizando las secuencias de ácido nucleico. Los ejemplos incluyen amplificación específica de alelos (Kazazian (1989) J. Lab. Clin. Med 11: 95 - 96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16: 325 - 332), ligación específica de alelos (Landegren et al. (1988) Science 241: 1077 - 1080), reacciones de extensión del nucleótido (Sokolov (1990) Nucleic Acid Sequence Res. 18: 3671), Mapeo Híbrido por Radiación (Walter et al. (1997) Nat. Genet. 7: 22 - 28) y Mapeo Happy (Dear and Cook (1989) Nucleic Acid Sequence Res. 17: 6795 - 6807). Para estos métodos, se utiliza la secuencia de una secuencia de ácido nucleico para diseñar y producir pares de cebadores para uso en la reacción de amplificación o en reacciones de extensión del cebador. El diseño de tales cebadores es bien conocido por aquellos ordinariamente bien capacitados en la técnica. En los métodos que emplean mapeo genético con base en PCR, puede ser necesario identificar las diferencias en la secuencia de ADN entre los progenitores del mapeo entrecruzado en la región correspondiente a la presente secuencia de ácido nucleico. Esto, sin embargo, generalmente, no es necesario para métodos de mapeo.
- 15 Los métodos de acuerdo con la presente invención dan como resultado plantas que tienen más características relacionadas con el rendimiento de semillas, como se describió anteriormente aquí. Estas características pueden ser combinadas también con otras características económicamente ventajosas, tales como las características que incrementan adicionalmente el rendimiento, tolerancia a otros estreses abióticos y bióticos, características que modifican diferentes características arquitectónicas y/o características bioquímicas y/o fisiológicas.
- 20 Los métodos de acuerdo con la presente invención dan como resultado plantas cultivadas bajo condiciones de estrés abiótico que tienen características mejoradas relacionadas con el rendimiento, como se describió aquí anteriormente. Estas características pueden ser combinadas también con otras características económicamente ventajosas, tales como las características que incrementan adicionalmente el rendimiento, tolerancia a otros estreses abióticos y bióticos, características que modifican diferentes características arquitectónicas y/o características bioquímicas y/o fisiológicas.
- 25 En una forma de realización, la invención se relaciona con la materia resumida de la siguiente manera:
- 30 Ítem 1: Un método para aumentar las características relacionadas con el rendimiento de semillas en plantas en relación con las plantas de control, que comprende incrementar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido 19/20 localizado en el núcleo con el motivo AT-hook (AHL19/20), donde el polipéptido AHL19/20 comprende un dominio que tiene al menos 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% o más de identidad de secuencia de aminoácidos de un dominio conservado (CD) como se representa por la SEQ ID NO: 36, y, opcionalmente, seleccionar las plantas que tienen un aumento de características relacionados con el rendimiento de semillas.
- 35 Ítem 2: Método de acuerdo con el ítem 1, en donde dicho polipéptido AHL 19/20 comprende: (i) un motivo que tiene al menos 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con un motivo AT-hook como el representado por la SEQ ID NO: 37, y (ii) un dominio que tiene al menos 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con un dominio conservado de planta y de procarionta (PPC), tal como el representado por la SEQ ID NO: 38 .
- 40 Ítem 3: Método de acuerdo con el ítem 1 o 2 , en el que dicho polipéptido AHL 19/20 comprende: (i) una señal de localización nuclear; (ii) un motivo de enlazamiento de ADN AT-hook con una entrada InterPro IPR014476, y (iii) un dominio conservado de planta y de procarionta (PPC) con una entrada InterPro IPR005175.
- 45 Ítem 4: Método de acuerdo con cualquier ítem anterior, en el que dicho polipéptido AHL 19/20, cuando se utiliza en la construcción de un árbol filogenético AHL, tal como el representado en la Figura 1 o en la Figura 2, se agrupa con el grupo de polipéptidos AHL 19/20 que comprenden la secuencia de polipéptido tal como la representada por la SEQ ID NO: 2, en lugar de con cualquier otro grupo de AHL.
- 50 Ítem 5: Método de acuerdo con cualquier ítem anterior, en el que dicho polipéptido AHL 19/20 tiene en orden creciente de preferencia al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99% o más identidad de secuencia de aminoácidos con el polipéptido AHL 19/20 tal como el representado por la SEQ ID NO: 2 o con cualquiera de las secuencias de polipéptidos presentados en la Tabla A en este documento.
- Ítem 6: Método de acuerdo con cualquier ítem anterior, en el que dicha secuencia de ácido nucleico que codifica un

polipéptido AHL 19/20 está representado por una cualquiera de las secuencias de ácidos nucleicos SEQ ID NOs presentadas en la Tabla A o una parte de las mismas, o una secuencia capaz de hibridarse con una cualquiera de las secuencias de ácidos nucleicos SEQ ID NOs presentadas en la Tabla A.

- 5 Ítem 7: Método de acuerdo con cualquier ítem anterior, en el que dicha secuencia de ácido nucleico codifica un ortólogo o parálogo de cualquiera de las secuencias de polipéptido de las SEQ ID NOs presentadas en la Tabla A.
- Ítem 8: Método de acuerdo con cualquier ítem anterior, en el que dicho aumento de la expresión es llevado a cabo por uno o más cualquiera de: marcado de la activación del ADN-T, TILLING, o recombinación homóloga.
- 10 Ítem 9: Método de acuerdo con cualquier ítem anterior, en el que dicho aumento de la expresión se lleva a cabo mediante la introducción y expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20.
- Ítem 10: Método de acuerdo con cualquier ítem anterior, en el que dicho aumento de características relacionadas con el rendimiento de semillas es uno o más de: (i) aumento del número de flores por panícula; (ii) aumento del peso total de las semillas por planta; (iii) aumento del número de semillas llenas; o (iv) aumento de índice de cosecha.
- 15 Ítem 11: Método de acuerdo con cualquier ítem anterior, en el que dicho aumento de características relacionadas con el rendimiento de semillas se produce en plantas cultivadas bajo condiciones de disponibilidad reducida de nutrientes, preferiblemente bajo condiciones de disponibilidad reducida de nitrógeno, en relación con las plantas de control.
- 20 Ítem 12: Método según el ítem 11, en el que dicho aumento de características relacionadas con el rendimiento de semillas es uno o más de: (i) aumento del rendimiento total de semillas por planta, (ii) aumento del número de semillas llenas, o (iii) aumento de índice de cosecha.
- Ítem 13: Método de acuerdo con cualquier ítem anterior, en el que dicha secuencia de ácido nucleico está operativamente enlazada a un promotor constitutivo, preferiblemente a un promotor constitutivo de la planta, más preferiblemente a un promotor GOS2, lo más preferible a un promotor GOS2 de arroz como el representado por la SEQ ID NO: 35.
- 25 Ítem 14: Método de acuerdo con cualquier ítem anterior, en el que dicha secuencia de ácido nucleico que codifica una polipéptido AHL 19/20 es de origen vegetal, preferiblemente de una planta dicotiledónea, adicionalmente preferiblemente de la familia Brassicaceae, lo más preferible de *Arabidopsis thaliana*.
- 30 Ítem 15: Plantas, sus partes (incluidas las semillas), o células de la planta que se pueden obtener por un método de acuerdo con cualquier ítem anterior, en el que dicha planta, parte o célula de la misma contienen un transgén de ácido nucleico aislado que codifica una polipéptido AHL 19/20 operativamente enlazado a un promotor constitutivo de la planta.
- Ítem 16: construcción que comprende:
- (a) Una secuencia de ácido nucleico que codifica una polipéptido AHL 19/20 como se define en cualquiera de los ítems 1 a 5;
- 35 (b) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (a); y opcionalmente;
- (c) una secuencia de terminación de la transcripción.
- 40 Ítem 17: Construcción de acuerdo con el ítem 16, en el que dicha secuencia de control es un promotor constitutivo de la planta, preferiblemente un promotor GOS2, más preferiblemente un promotor GOS2 como el representado por la SEQ ID NO: 35.
- 45 Ítem 18: El uso de una construcción de acuerdo con los ítems 16 o 17 en un método para la elaboración de plantas que tienen más características relacionadas con el rendimiento de semillas en relación con las plantas de control, en donde dicho aumento de características relacionadas con el rendimiento de semillas son una o más de: (i) aumento del número de flores por panícula; (ii) aumento del peso total de las semillas por planta; (iii) aumento en el número de semillas llenas; o (iv) aumento del índice de cosecha.
- Ítem 19: Planta, parte de la planta o célula de la planta transformada con una construcción de acuerdo con el ítem 16 o 17.

Ítem 20: Método para la producción de plantas transgénicas que tienen mayores características relacionadas con el rendimiento de semillas en relación con las plantas de control, que comprende:

5 (i) introducir y expresar en una planta, parte de la planta, o célula de la planta, una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como se define en uno cualquiera de los ítems 1 a 6, bajo el control del promotor constitutivo de la planta; y

(ii) cultivar la célula de la planta, parte de la planta, o planta en condiciones que promuevan el crecimiento y desarrollo de la planta.

10 Ítem 21: Planta transgénica que tiene un aumento de características relacionadas con el rendimiento de semillas en relación con las plantas de control, como resultado de un aumento de la expresión de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como se define en cualquiera de los ítems 1 a 5, operativamente enlazada a un promotor constitutivo de la planta, o una célula de una planta transgénica o parte de la planta transgénica derivada de dicha planta transgénica.

15 Ítem 22: Planta transgénica de acuerdo con el ítem 15, 19 o 21, en donde dicha planta es una planta de cultivo o una monocotiledónea o un cereal, tal como arroz, maíz, trigo, cebada, mijo, centeno, triticale, sorgo y avena, o una célula de la planta transgénica derivada de dicha planta transgénica.

Ítem 23: Partes cosechables que contienen una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 de una planta de acuerdo con el ítem 22, en donde dichas partes cosechables son preferiblemente semillas.

Ítem 24: Productos derivados de una planta de acuerdo con el ítem 22 y / o de partes cosechables de una planta de acuerdo al ítem 23.

20 Ítem 25: Uso de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como se define en cualquiera de los ítems 1 a 6 en el aumento de las características relacionadas con el rendimiento de semillas, que comprende uno o más de: (i) aumento del número de flores por panícula; (ii) aumento del peso total de las semillas por planta; (iii) aumento en el número de semillas llenas; o (iv) aumento del índice de cosecha.

25 Ítem 26: El uso de acuerdo con el ítem 25, en el que dicho aumento de características relacionadas con el rendimiento de semillas se presenta en condiciones de disponibilidad reducida de nutrientes, preferiblemente bajo condiciones de disponibilidad reducida de nitrógeno.

Descripción de las figuras

La presente invención será descrita a continuación con referencia a las siguientes figuras en las cuales:

30 La Figura 1 representa un árbol filogenético construido después de una alineación de todos los polipéptidos que pertenecen a la familia AHL (descrito en Fujimoto et al., (2004) Plant Molec Biol., 56: 225 - 239) llevado a cabo utilizando el algoritmo Clustal (1.83) de alineación progresiva, utilizando valores predeterminados. El grupo de interés, que comprende los dos parálogos de Arabidopsis AHL19 (SEQ ID NO: 2 o AT3G04570) y AHL20 (SEQ ID NO: 4 o AT4G14465) ha sido encerrado en un círculo .

35 La Figura 2 representa un árbol filogenético construido después de una alineación de todos los polipéptidos que pertenecen a la familia AHL (descrito en Fujimoto et al., (2004) Plant Molec Biol., 56: 225 - 239), y todos los ortólogos y parálogos AHL 19/20 de la Tabla A en el Ejemplo 1 en este documento, llevado a cabo utilizando el algoritmo Clustal (1.83) de alineación progresiva, utilizando valores predeterminados.

40 La Figura 3 representa un boceto de un polipéptido AHL 19/20 tal como el representado por la SEQ ID NO: 2, que comprende la siguientes características: una NLS predicha, un motivo de enlazamiento de ADN AT-hook (del que el núcleo es el tripéptido GRP; comprendido en la entrada InterPro IPR014476 (enlazamiento predicho de ADN AT-hook)), un dominio de PPC (dominio conservado de planta y de procariotas; comprendido en la entrada InterPro IPR005175 (proteína de función desconocida DUF296)), y Dominio Conservado (CD) que comprende tanto un motivo de enlazamiento de ADN AT-hook y un dominio de PPC.

45 La Figura 4 muestra una alineación de múltiples secuencias por CLUSTAL W (1.83) del Dominio Conservado de polipéptidos AHL 19/20 de la Tabla A (como el representado por la SEQ ID NO: 38 para la SEQ ID NO: 2), donde se identifican una serie de características. Desde el extremo del terminal N hasta el extremo del terminal C de los polipéptidos, estos son: (i) una señal de localización nuclear predicha (NLS), (ii) un motivo de enlazamiento de ADN AT-hook, con el núcleo tripéptido GRP, y (iii) un dominio de PPC (DUF 296) .

La Figura 5 muestra el vector binario para aumento de la expresión en *Oryza sativa* de una secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 bajo el control de un promotor GOS2 (pGOS2) de arroz.

La Figura 6 detalla ejemplos de secuencias útiles en la realización de los métodos de acuerdo con la presente invención.

5 Ejemplos

La presente invención será descrita a continuación con referencia a los siguientes y ejemplos, que se citan sólo para efectos ilustrativos. Los siguientes ejemplos no pretenden definir completamente o limitar de algún otro modo el alcance de la invención.

10 Manipulación del ADN: a menos que se establezca otra cosa, las técnicas de ADN recombinante se llevan a cabo de acuerdo con protocolos estándar descritos en (Sambrook (2001) *Molecular Cloning: a laboratory manual*, 3^o 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 estándar y los métodos para el trabajo molecular de la planta se describen en *Plant Molecular Biology Labfax* (1993) de R. D. D. Croy, publicado por BIOS Scientific Publications Ltd (RU) y Blackwell Scientific Publications (RU).

15 **Ejemplo 1:** Identificación de secuencias relacionadas con la secuencia de ácido nucleico utilizada en los métodos de la invención

20 Las secuencias (ADNc de longitud completa, los EST o genómicas) relacionadas con la secuencia de ácido nucleico usadas en los métodos de la presente invención fueron identificadas entre aquellas mantenidas en la base de datos de nucleótidos Entrez en el National Center for Biotechnology Information (NCBI) usando herramientas de búsqueda de secuencias en la base de datos, tales como la 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 secuencias mediante la comparación de secuencias de ácido nucleico o de secuencias de polipéptidos con las bases de datos de secuencias y calculando el significado estadístico de las coincidencias. Por ejemplo, el polipéptido codificado por la secuencia de ácido nucleico de la presente invención se utilizó para el algoritmo TBLASTN, con parámetros predeterminados y el filtro desactivado para ignorar secuencias de baja complejidad. El resultado del análisis se obtuvo a través de la comparación por pares, y se clasificó de acuerdo con el puntaje de la probabilidad (valor E), donde el puntaje refleja la probabilidad de que ocurra una alineación particular al azar (entre más bajo el valor E, más significativa la coincidencia). Además de los valores E, también se dio puntaje a las comparaciones mediante el porcentaje de identidad. El porcentaje de identidad se refiere a la cantidad de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de ácido nucleico (o polipeptídicas) comparadas sobre una longitud particular. En algunos casos, se pueden ajustar los parámetros por defecto para modificar la rigurosidad de la búsqueda. Por ejemplo, se puede incrementar el valor E para mostrar menos coincidencias estrictas. De este modo, se pueden identificar coincidencias cortas casi exactas.

35 La Tabla A proporciona un listado de secuencias de ácidos nucleico relacionadas con la secuencia de ácido nucleico utilizada en los métodos de la presente invención.

Tabla A: Ejemplos de polipéptidos AHL 19/20:

Nombre	Organismo fuente	Ácido nucleico SEQ ID NO:	Polipéptido SEQ ID NO:	Número de acceso a la base de datos	Estatus
Arath_AHL19	<i>Arabidopsis thaliana</i>	1	2	AT3G04570 NP_566232	Longitud completa (FL)
Arath_AHL20	<i>Arabidopsis thaliana</i>	3	4	AT4G14465 NP_567432	FL
Aqufo_AHL19/20	<i>Aquilegia formosa x Aquilegia pubescens</i>	5	6	Cóntigo de DT758489, DT758488.1	FL
Brana_AHL19/20	<i>Brassica napus</i>	7	8	CS226287	FL
Brara_AHL 19/20	<i>Brassica rapa</i>	9	10	AC189468	FL
Glyma_AHL19/20	<i>Glycine max</i>	11	12	CS137412	FL
Goshi_AHL19/20	<i>Gossypium hirsutum</i>	13	14	DW519458	FL
Lacsa_AHL19/20	<i>Lactuca sativa</i>	15	16	DW047323	FL
Lotja_AHL19/20	<i>Lotus japonicus</i>	17	18	AP004971	FL

(continuación)

Nombre	Organismo fuente	Ácido nucleico SEQ ID NO:	Polipéptido SEQ ID NO:	Número de acceso a la base de datos	Estatus
Orysa_AHL19/2 0	<i>Oryza sativa</i>	19	20	AK110263 Os08g056320 0	FL
Orysa_AHL19/2 0 II	<i>Oryza sativa</i>	21	22	CT837915 Os02g082080 0	FL
Poptr_AHL19/20	<i>Populus tremuloides</i>	23	24	scaff XIII.441	FL
Soltu_AHL19/20	<i>Solanum tuberosum</i>	25	26	Cóntigo de CN215397.1 CK276075.1	FL
Thlca_AHL19/2 0	<i>Thlaspi caerulescens</i>	27	28	DQ022564	FL
Vitvi-AHL19/20	<i>Vitis vinifera</i>	29	30	AM463589	FL
Vitvi_AHL19/20 II	<i>Vitis vinifera</i>	31	32	AM429692	FL
Zeama_AHL19/ 20	<i>Zea mays</i>	33	34	AC190270	FL

5 En algunos casos, las secuencias relacionadas han sido ensambladas tentativamente y divulgadas al público por parte de instituciones de investigación, tales como The Institute for Genomic Research (TIGR). La base de datos Eukaryotic Gene Orthologs (EGO) puede ser utilizada para identificar tales secuencias relacionadas, ya sea por medio de una búsqueda de palabras clave o por medio del uso del algoritmo BLAST con la secuencia de ácido nucleico o la secuencias de polipéptidos de interés. En otros casos, se han creado bases de datos especiales de secuencias de ácido nucleico para organismos particulares, tales como por el Joint Genome Institute, por ejemplo para *Ostreococcus tauri*.

10 Ejemplo 2: Alineación de secuencias de polipéptidos de la invención

15 La alineación de secuencias de polipéptidos se lleva a cabo utilizando el programa AlignX del paquete Vector NTI (Invitrogen), que se basa en el algoritmo popular ClustalW 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 predeterminados son para la penalización por abertura de espacio de 10, para penalización por extensión de espacio de 0,1 y la matriz de peso seleccionada es Blosom 62 (si se alinean polipéptidos). Se puede hacer una edición manual menor se puede hacer para optimizar aún más la alineación. Se construye un árbol filogenético de polipéptidos utilizando un algoritmo de agrupamiento unión de vecinos (*neighbour-joining*) conforme a lo dispuesto en el programa AlignX de Vector NTI (Invitrogen) .

20 La alineación de todas las secuencias de polipéptidos AHL de *Arabidopsis thaliana* identificadas en Fujimoto et al. (2004; Tableau más adelante) se realizó usando el algoritmo Clustal (1.83) de alineación progresiva, con valores predeterminados (Thompson et al. (1997) Nucleic Acids Res 25: 4876 - 4882; Chenna et al. (2003). Nucleic Acids Res 31: 3497 - 3500). Se construyó un árbol de unión de vecinos después de eso, y se representa en la Figura 1 de la presente solicitud. El grupo de interés, que comprende los dos parálogos AHL19 de *Arabidopsis* (SEQ ID NO: 2 o AT3G04570), y AHL20 (SEQ ID NO: 4 o AT4G14465) encerrado en un círculo. Cualquier polipéptido que caiga dentro de este grupo de AHL19/20 (después de una nueva etapa de alineación múltiple como se describe anteriormente en esta memoria) se considera que es útil en la realización de los métodos de la invención como se describió en este documento más arriba.

30 Tabla A1: Polipéptidos AHL identificados en *Arabidopsis thaliana*

Número AHL	Número de acceso de Tair	Número de acceso del NCBI
AHL1	At4g12080	NP_192945
AHL2	At4g22770	NP_194008
AHL3	At4g25320	NP_194262
AHL4	At5g51590	NP_199972
AHL5	At1g63470	NP_176536
AHL6	At5g62260	NP_201032
AHL7	At4g00200	NP_191931
AHL8	At5g46640	NP_199476

(continuación)

Número AHL	Número de acceso de Tair	Número de acceso del NCBI
AHL9	At2g45850	NP_182109
AHL10	At2g33620	NP_565769
AHL11	At3g61310	NP_191690
AHL12	At1g63480	NP_176537
AHL13	At4g17950	NP_567546
AHL14	At3g04590	NP_187109
AHL15	At3g55560	NP_191115
AHL16	At2g42940	NP_181822
AHL17	At5g49700	NP_199781
AHL18	At3g60870	NP_191646
AHL19	At3g04570	NP_566232
AHL20	At4g14465	NP_567432
AHL21	At2g35270	NP_181070
AHL22	At2g45430	NP_182067
AHL23	At4g17800	NP_193515
AHL24	At4g22810	NP_194012
AHL25	At4g35390	NP_195265
AHL26	At4g12050	NP_192942
AHL27	At1g20900	NP_173514
AHL28	At1g_14490	NP_172901
AHL29	At1g76500	NP_177776

- 5 Una segunda alineación de secuencias múltiples se realizó incluyendo todos los polipéptidos ortólogos AHL19/20 de la Tabla A y todas las secuencias de AHL de la Tabla A1. Se construyó un árbol de unión de vecinos después de eso, y se representa en la Figura 2 de la presente solicitud. El grupo de interés, que comprende los dos parálogos AHL19 de Arabidopsis (SEQ ID NO: 2 o AT3G04570), y AHL20 (SEQ ID NO: 4 o AT4G14465) ha sido encerrado en un círculo. Cualquier polipéptido que caiga dentro de este grupo de AHL19/20 se considera que es útil en la realización de los métodos de la invención como se describe en este documento.
- 10 El Dominio Conservado (CD) de la SEQ ID NO: 2, tal como la representada por la SEQ ID NO: 36, fue identificado después de la alineación de múltiples secuencia de los polipéptidos AHL19/20 de la Tabla A. En una segunda etapa, se seleccionaron los CD de los polipéptidos AHL19/20 de la Tabla A (fuera de la secuencia del polipéptido de longitud completa) y alinearon, utilizando el algoritmo Clustal (1.83) de alineación progresiva, utilizando valores predeterminados. Se pueden identificar una cantidad de características, y están marcadas en la Figura 4. Desde el extremo del terminal N hasta el extremo del terminal C de los polipéptidos, estos son: (i) una señal de localización nuclear predicha (NLS) , (ii) un motivo de enlazamiento de ADN AT-hook, con el tripéptido nuclear GRP, y (iii) un dominio de PPC (DUF 296) .
- 15

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

- 20 Se determinaron los porcentajes globales de similitud e identidad entre secuencias de polipéptidos de longitud completa útiles para llevar a cabo los métodos de la invención utilizando uno de los métodos disponibles en el arte, el software MatGAT (Matrix Global Alignment Tool) (BMC Bioinformatics. 2003 4: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando secuencias de proteína o de ADN. Campanella J. J., Bitincka L, Smalley J; software administrado por Ledion Bitincka). El software MatGAT genera matrices de similitud/identidad para secuencias de ADN o de proteína sin necesidad de una alineación previa para los datos. El programa realiza una serie de alineaciones de pares utilizando el algoritmo de alineación global de Myers y Miller (con una penalización por apertura de brecha de 12, y una penalización por extensión de brecha de 2), calcula la similitud y la identidad utilizando por ejemplo Blosum 62 (para polipéptidos), y luego coloca los resultados en una matriz de distancia. La similitud de secuencia se muestra en la mitad inferior de la línea divisoria y la identidad de la secuencia se muestra en la mitad superior de la línea divisoria en diagonal.
- 25
- 30

Los parámetros utilizados en la comparación fueron:

Matriz de puntuación: Blosum62

Primera brecha: 12

Extensión de la brecha: 2

También se pueden generar una tabla de MatGAT para alineación local de un dominio específico, o datos sobre en % de identidad / similitud entre dominios específicos.

5 Los resultados del análisis del software se muestran en la Tabla B para la similitud e identidad global sobre la longitud completa de las secuencias de polipéptidos (excluyendo las secuencias parciales de polipéptidos). El porcentaje de identidad se presenta encima de la diagonal y el porcentaje de similitud se presenta debajo de la diagonal.

El porcentaje de identidad entre las secuencias de polipéptidos útiles en la realización de los métodos de la invención puede ser tan bajo como un 52% de identidad de aminoácidos comparado con la SEQ ID NO: 2.

10 Tabla B: Resultados de MatGAT para la similitud e identidad globales sobre la longitud completa de las secuencias de polipéptidos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Aqufo_AHL19_20		63	64	61	61	60	73	66	71	59	61	66	70	63	73	71	58
2. Arath_AHL19	72		56	94	94	56	59	61	59	48	55	61	63	97	61	57	52
3. Arath_AHL20	76	70		54	54	56	63	63	61	54	58	59	61	57	61	63	54
4. Brana_AHL19_20	70	96	67		100	56	59	61	58	49	55	60	61	94	60	56	52
5. Brara_AHL19_20	70	96	67	100		56	59	61	58	49	55	60	61	94	60	56	52
6. Glyma_AHL19_20	75	68	73	68	68		60	64	59	52	58	66	63	58	67	59	55
7. Goshi_AHL19_20	82	69	75	67	67	77		64	70	61	66	64	67	59	71	79	61
8. LacsA_AHL19_20	79	72	77	72	72	81	80		63	56	59	69	73	61	72	64	56
9. Lotja_AHL19_20	82	72	75	71	71	74	80	77		59	61	63	69	58	71	72	62
10. Orysa_AHL19_20	67	58	65	58	58	65	68	69	66		66	52	55	48	56	62	70
11. Orysa_AHL19_20II	71	64	69	63	63	73	76	73	72	70		59	60	56	60	66	67
12. Poptr_AHL19_20	81	71	73	71	71	80	80	81	80	64	74		72	61	75	62	54
13. Soltu_AHL19_20	82	75	76	75	75	79	77	81	80	61	70	87		63	78	66	58
14. Thlca_AHL19_20	73	98	69	95	95	70	68	72	73	58	66	72	75		61	58	52
15. Vitvi_AHL19_20	85	72	74	71	71	81	79	82	82	65	70	85	90	71		69	57
16. Vitvi_AHL19_20_II	80	66	74	66	66	73	84	78	79	70	76	75	76	66	78		63

15 Se puede aumentar considerablemente el porcentaje de identidad si el cálculo de identidad se lleva a cabo entre el Dominio Conservado (CD) de la SEQ ID NO: 2 (como el representado por la SEQ ID NO: 36) y los CD de los polipéptidos útiles en la realización de la invención. Un CD comprende un motivo de enlazamiento de ADN AT-hook (tal como el representado por la SEQ ID NO: 37 para la SEQ ID NO: 2) y un dominio de PPC (como el representado por la SEQ ID NO: 38 para la SEQ ID NO: 2). El porcentaje de identidad sobre los CD entre las secuencias de polipéptidos útiles en la realización de los métodos de la invención oscila entre 75% y 99% de identidad de aminoácidos, como se muestra en la Tabla B1. Esto es significativamente más alto que el porcentaje de identidad de aminoácidos calculado entre las secuencias de polipéptidos AHL19/20 de longitud completa.

20 Tabla B1: Resultados de MatGAT para similitud e identidad globales sobre el dominio de los CD entre las secuencias de polipéptidos

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CD_Aqufo_AHL19_20		81	81	80	80	77	92	83	87	86	87	81	84	81	84	92	86
2. CD Arath_AHL19	91		78	98	98	75	82	77	78	76	80	78	80	99	81	81	77
3. CD_Arath_AHL20	93	88		77	77	73	84	80	79	77	81	76	79	78	79	82	78
4. CD Brana_AHL19_20	91	99	88		100	75	82	77	78	76	79	78	79	98	80	80	78
5. CD_Brara_AHL19_20	91	99	88	100		75	82	77	78	76	79	78	79	98	80	80	78
6. CD_Glyma_AHL19_20	92	88	89	88	88		80	78	80	74	79	81	78	75	82	80	78
7. CD_Goshi_AHL19_20	98	91	93	91	91	93		86	92	87	89	85	88	82	90	94	88
8. CD_LacsA_AHL19_20	96	90	92	90	90	92	96		81	78	83	86	87	77	88	83	78

(continuación)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
9. CD_ Lotja_ AHL19_20	96	90	93	90	90	94	98	95		86	86	81	86	78	86	94	89
10. CD_ Orysa_ AHL19_20	94	87	89	87	87	89	92	92	93		89	78	84	76	81	86	95
11. CD_ Orysa_ AHL19_20/II	94	88	89	88	88	93	95	93	96	95		81	87	80	86	89	90
12. CD_ Poptr_ AHL19_20	96	88	91	88	88	93	97	95	96	92	93		86	78	87	84	78
13. CD_ Soltu_ AHL19_20	94	89	91	89	89	93	96	95	96	90	93	96		80	90	89	83
14. CD_ Thlca_ AHL19_20	91	100	88	99	99	88	91	90	90	87	88	89	89		81	81	77
15. CD_ Vitvi_ AHL19_20	94	90	91	90	90	94	96	96	96	90	93	95	96	90		88	81
16. CD_ Vitvi_ AHL19_20/II	96	90	92	90	90	93	98	94	99	92	95	95	95	90	95		87
17. CD_ Zeama_ AHL19_20	94	88	90	88	88	91	94	93	95	96	96	92	90	88	90	93	

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

10 La base de datos de Recursos Integrados de Familias de Proteínas, Dominios y Sitios (InterPro) es una interfaz integrada para las bases de datos de firma comúnmente utilizadas para búsquedas con base en textos y secuencias. La base de datos InterPro combina estas bases de datos, que utilizan diferentes metodologías y diferentes grados de información biológica sobre proteínas bien caracterizadas para derivar firmas de proteína. Las bases de datos colaboradoras incluyen SWISS-PROT, PROSITE, TrEMBL, PRINTS, Panther, ProDom y Pfam, Smart y TIGRFAMs. InterPro se encuentra alojada en el European Bioinformatics Institute en el Reino Unido.

Los resultados del barrido con InterPro de la secuencia de polipéptidos representadas por la SEQ ID NO: 2 se presentan en la Tabla C.

15 Tabla C: Resultados del barrido con InterPro de la secuencia de polipéptidos como la representada por la SEQ ID NO: 2

Nombre y número de acceso del InterPro	Nombre en la base de datos integrada	Número de acceso en la base de datos integrada	Nombre de acceso en la base de datos integrada	Coordenadas del aminoácido en la SEQ ID NO: 2
IPR005175 Dominio: Proteína de función desconocida DUF296	PFAM	PF03479	DUF 296	107 - 232
InterPro IPR014476 Familia: Motivo de enlazamiento predicho de ADN AT-hook	PIR	PIRSF016021	ESCAROLA	1 - 315

Ejemplo 5: Predicción de la localización subcelular de las secuencias de polipéptidos útiles para llevar a cabo los métodos de la invención

TargetP 1.1 predice la localización subcelular de las proteínas eucariotas. La asignación de ubicación se basa en la

5 presencia predicha de cualquiera de las secuencias previas del terminal N: péptido de tránsito al cloroplasto (cTP), péptido mitocondrial objetivo (mTP) o péptido señal de la ruta secretora (SP). Los puntajes en los que se basa la predicción final no son realmente probabilidades, y no necesariamente se añaden a uno. Sin embargo, la ubicación con el puntaje más alto es la más probable de acuerdo con TargetP, y la relación entre los puntajes (la clase de confiabilidad) puede ser una indicación de qué tan cierta es la predicción. La clase de confiabilidad (RC) varía de 1 a 5, donde 1 indica la predicción más fuerte. TargetP es mantiene en el servidor de la Universidad Técnica de Dinamarca.

Para las secuencias que se predice que contienen una secuencia previa en el terminal N se puede predecir también un sitio de escisión potencial.

10 Se seleccionaron una cantidad de parámetros, tales como el grupo del organismo (vegetal o no vegetal), conjuntos de puntos de corte (ninguno, conjuntos de puntos de corte predefinidos, o conjunto de puntos de corte especificados por el usuario), y el cálculo de la predicción de los sitios de escisión (sí o no).

15 Los resultados del análisis con TargetP 1.1 de la secuencia polipeptídica como la representada por la SEQ ID NO: 2 se presentan en la Tabla D7. El grupo del organismo "planta" ha sido seleccionado, y no se definieron puntos de corte. La localización subcelular predicha de la secuencia polipeptídica como la representada por la SEQ ID NO: 2 no es del cloroplasto, no es mitocondrial y no es la ruta secretora, sino más probablemente del núcleo.

20 Una señal de localización nuclear prevista (NLS) se encuentra (por medio de la alineación de secuencias múltiples, seguida por inspección ocular, por ejemplo) en el polipéptido AHL19/20 de la SEQ ID NO: 2. Una NLS es una o más secuencias cortas lisinas o argininas positivamente cargadas. Se predice que la SEQ ID NO: 2 de la presente invención se localiza en el compartimento nuclear de las células eucariotas.

Tabla D. Análisis con TargetP 1.1 de la secuencia de polipéptidos como la representada por la SEQ ID NO: 2

Longitud (AA)	315
Péptido de tránsito en el cloroplasto	0.100
Péptido de transito mitocondrial	0.278
Péptido señal de la ruta secretora	0.033
Otro objetivo subcelular	0.703
Ubicación predicha	Otra
Clase de confiabilidad	3

Los resultados del análisis con TargetP 1.1 de la secuencia de polipéptido como la representada por la SEQ ID NO: 51 se presentan a continuación.

25 Predicción de TargetP: mitocondrial (0,837, calidad 2)

Se pueden utilizar muchos algoritmos para llevar a cabo los análisis de predicción de la ubicación subcelular, incluyendo:

- ChloroP 1.1 alojado en el servidor de la Universidad Técnica de Dinamarca;
- 30 • Protein Prowler Subcellular Localisation Predictor versión 1.2 alojado en el servidor del Institute de Biociencia Molecular, Universidad de Queensland, Brisbane, Australia;
- PENCE Proteome Analyst 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 una secuencia de ácido nucleico de la invención.

35 A menos que se establezca otra cosa, se llevan a cabo técnicas de ADN recombinante de acuerdo con protocolos estándar que se describen en (Sambrook (2001) Molecular Cloning: a laboratory manual, 3° edición Cold Spring Harbor Laboratory Press, CSH, Nueva York) o en los Volúmenes 1 y 2 de Ausubel et al. (1994), Current Protocols in Molecular Biology, Current Protocols. Los materiales y métodos estándar para trabajo molecular en plantas se describen en Plant Molecular Biology Labfax (1993) por R. D. D. Croy, publicado por BIOS Scientific Publications Ltd (RU) y Blackwell Scientific Publications (RU).

40

Se amplificó el ADNc de *Arabidopsis thaliana* que codifica el polipéptido AHL 19/20 por medio de PCR utilizando como molde un banco de ADNc de *Arabidopsis* sintetizado a partir del ARNm extraído de tejidos mezclados de plantas. Se utilizaron el cebador prm8135 (SEQ ID NO: 41, sentido: 5'-ggggacaagtgtgtacaaaaaagcaggcttaaacaatggcgaatccatggtg-3') y el cebador prm08136 SEQ ID NO: 42; inverso, complementario: 5' - ggggaccactttgtacaagaaagctgggttaaaaaccattttaacgcacg -3'), que incluyen los sitios AttB para recombinación Gateway, para amplificación por PCR. Se llevó a cabo la PCR utilizando Taq ADN polimerasa Hifi en condiciones estándar. Se amplificó y purificó también un fragmento PCR de la longitud esperada (incluyendo sitios attB), empleando métodos estándar. Luego se llevó a cabo la primera etapa del procedimiento Gateway, la reacción BP, durante la cual el fragmento de la PCR recombinado *in vivo* con el plásmido pDONR201 para producir, de acuerdo con la terminología Gateway, un "clon de entrada". Se adquirió el plásmido pDONR201 a través de Invitrogen, como parte de la tecnología Gateway®.

Ejemplo 7: Construcción del vector de expresión utilizando una secuencia de ácido nucleico como la divulgada

El clon de entrada que comprende la SEQ ID NO: 1 se utilizó posteriormente en una reacción LR con un vector de destinación usado para la transformación de *Oryza sativa*. Este vector contenía como elementos funcionales dentro de los bordes del ADN-T: un marcador seleccionable de una planta; un casete de expresión de un marcador cribable; y un casete Gateway destinado a la recombinación *in vivo* de LR con la secuencia de ácido nucleico de interés ya clonada en el clon de entrada. Se localizó un promotor GOS2 de arroz (SEQ ID NO: 35) para expresión constitutiva secuencia arriba de este casete Gateway.

Después de la etapa de recombinación de LR, se transformó el vector de expresión resultante pGOS2::de AHL 19/20 y (Figura 5) en una cepa LBA4044 de *Agrobacterium*, de acuerdo con los métodos bien conocidos en el arte.

Ejemplo 8: Transformación de plantas

Transformación de arroz

Se utilizó *Agrobacterium* que contiene los vectores de expresión en forma independiente para transformar plantas de *Oryza sativa*. Se descascarillaron semillas secas maduras de la variedad de cultivo Japónica de arroz Nipponbare. Se realizó la esterilización incubando durante un minuto en etanol al 70%, seguido por 30 minutos en HgCl₂ al 0,2%, seguido de 6 lavados durante 15 minutos con agua destilada estéril. Las semillas estériles fueron luego germinadas en un medio que contenía 2,4-D (medio de inducción de callos). Después de la incubación en la oscuridad durante cuatro semanas, se cortaron y propagaron los callos embriogénicos derivados de escutelo sobre el mismo medio. Al cabo de dos semanas, los callos se multiplicaron o propagaron mediante un subcultivo en el mismo medio durante otras 2 semanas, se multiplicaron o propagaron los callos por medio de subcultivo sobre el mismo medio durante otras dos semanas. Se subcultivaron piezas de callo sobre medio fresco 3 días antes del cocultivo (para reforzar la actividad de la división celular).

Se utilizó la cepa LBA4404 de *Agrobacterium* que contenía cada vector de expresión individual en forma independiente para el cocultivo. Se inoculó *Agrobacterium* sobre medio AB con los antibióticos apropiados y se cultivó durante 3 días a 28°C. Se recolectaron luego las bacterias y se suspendieron en un medio de cocultivo líquido hasta una densidad (OD₆₀₀) de aproximadamente 1. Se transfirió luego la suspensión a una placa de Petri y se sumergieron los callos en la suspensión durante 15 minutos. Los tejidos de los callos fueron transferidos secos sobre un papel de filtro y transferidos a un medio de cocultivo solidificado e incubados durante 3 días en la oscuridad a 25°C. Los callos cocultivados crecieron en un medio que contenía 2,4-D durante 4 semanas en la oscuridad a 28°C en presencia de un agente de selección. Durante este período, se desarrollaron rápidamente islas de callos resistentes. Luego de transferir este material a un medio de regeneración e incubarlo a la luz, se liberó el potencial embriogénico y se desarrollaron brotes en las próximas cuatro a cinco semanas. Se cortaron los brotes de los callos y se incubaron durante 2 a 3 semanas en un medio que contenía auxina, desde el cual fueron transferidos al suelo. Se cultivaron los brotes encallecidos bajo humedad alta y días cortos en un invernadero.

Se generaron aproximadamente 35 transformantes independientes de arroz T0 para cada construcción. Los transformantes primarios se transfirieron desde una cámara de cultivo de tejido a un invernadero. Luego de un análisis por PCR cuantitativo para verificar la cantidad de copias del inserto de ADN-T, se conservaron únicamente copias simples de plantas transgénicas que exhibieron tolerancia al agente de selección para la cosecha de semillas T1. Luego se cosecharon las semillas tres a cinco meses después de ser trasplantadas. El método produjo transformantes de un solo locus en una proporción superior al 50 % (Aldemita y Hodges1996, Chan et al. 1993, Hiei et al. 1994).

Ejemplo 9: Procedimiento de evaluación fenotípica

10.1 Disposición para la evaluación

- 5 Se generaron aproximadamente 35 nuevos transformantes independientes de arroz T0. Se transfirieron los transformantes primarios desde una cámara de cultivo de tejidos hasta un invernadero para crecimiento y cosecha de semilla T1. Se retuvieron seis eventos, de los cuales la progenie T1 segregó 3:1 por la presencia /ausencia del transgén. Para cada uno de estos eventos se seleccionaron aproximadamente 10 plantas de semillero T1 que contenían el transgén (hetero y homocigotos) y aproximadamente 10 plantas de semillero T1 que carecían del transgén (nulicigotos) y por medio de control visual de expresión del marcador. Las plantas transgénicas y los correspondiente nulicigotos se cultivaron uno al lado del otro en posiciones aleatorias. Las condiciones del invernadero fueron de días cortos (12 horas de luz), 28°C en la luz y 22°C en la oscuridad, y una humedad relativa del 70%.
- 10 Desde la etapa de siembra hasta la etapa de maduración, se pasaron varias veces las plantas a través de una cabina de formación de imágenes digitales. En cada instante de tiempo se tomaron imágenes digitales (2048 x 1536 pixeles, 16 millones de colores) de cada planta desde al menos 6 ángulos diferentes.
- Cribado de la disponibilidad reducida de nutriente (nitrógeno)
- 15 Las plantas de seis eventos (semillas T2) fueron cultivadas en tierra para macetas en condiciones normales excepto por la solución de los nutrientes. Las macetas se regaron desde el trasplante hasta la maduración con una solución nutriente específica que tenía un contenido reducido de nitrógeno (N), por lo general entre 7 a 8 veces menos. El resto del cultivo (maduración de la planta, la cosecha de semillas) fue el mismo que para las plantas no cultivadas bajo estrés abiótico. Los parámetros de crecimiento y el rendimiento se registraron como se detalla para el crecimiento en condiciones normales.
- 20 10.2 Análisis estadístico: prueba F
- 25 Se utilizó un ANOVA de dos factores (análisis de variantes) como un modelo estadístico para la evaluación general de las características fenotípicas de las plantas. Se realizó una prueba F sobre todos los parámetros medidos de todas las plantas de todos los eventos transformados con el gen de la presente invención. Se realizó la prueba F para controlar el efecto del gen sobre todos los eventos de transformación y para verificar el efecto total del gen, también conocido como efecto global del gen. El umbral de significación para un efecto global verdadero del gen se fijó en un nivel de probabilidad del 5% para la prueba F. Un valor significativo de la prueba F indica un efecto del gen, lo que significa que no es sólo la mera presencia o la posición del gen lo que está causando las diferencias en el fenotipo.
- 30 Cuatro eventos T1 fueron evaluados en la generación T2 siguiendo el mismo procedimiento de evaluación que para la generación T1 pero con más individuos por evento.
- 35 Cuando se llevaron a cabo dos experimentos con eventos de solapamiento, se realizó un análisis combinado. Es útil comprobar la consistencia de los efectos sobre los dos experimentos, y si este es el caso, acumular evidencia a partir ambos experimentos con el fin de aumentar la confianza en la conclusión. El método utilizado es un enfoque de modelo mixto que tiene en cuenta la estructura multinivel de los datos (es decir, experimento - evento - segregantes). Los valores de p fueron obtenidos por comparación de la prueba de tasa de probabilidad con las distribuciones de chi cuadrado.
- Desde la etapa de siembra hasta la etapa de maduración, se pasaron varias veces las plantas a través de una cabina de formación de imágenes digitales. En cada instante de tiempo se tomaron imágenes digitales (2048 x 1536 pixeles, 16 millones de colores) de cada planta desde al menos 6 ángulos diferentes.
- 40 10.3 Parámetros medidos
- Medición de parámetros relacionados con la biomasa
- Desde la etapa de siembra hasta la etapa de maduración, se pasaron las plantas varias veces a través de una cabina de formación de imágenes digitales. En cada instante de tiempo se tomaron imágenes digitales (2048 x 1536 pixeles, 16 millones de colores) de cada planta desde al menos 6 ángulos diferentes.
- 45 Se determinó el área de la planta sobre el suelo (o biomasa foliar) contando el número total de píxeles en la imágenes digitales de partes de la planta aéreas discriminadas desde el fondo. Se promedió este valor para las fotos tomada en el mismo instante de tiempo desde los diferentes ángulos y fue convertido a un valor de superficie física expresado en mm cuadrados por 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 aéreas. El área por encima del suelo es el área medida en el instante de tiempo en el que la planta había alcanzado su biomasa foliar máxima. El vigor temprano es el área por encima del suelo de la planta (planta de semillero) tres semanas después de la
- 50

germinación. El aumento de la biomasa de la raíz se expresa como un aumento en la biomasa total de la raíz (medida como biomasa máxima de las raíces observada durante la vida útil de una planta); o como un aumento en el índice raíz / brote (medido como la relación entre los masa de la raíz y la masa de los brotes en el período de crecimiento activo de las raíces y los brotes).

- 5 El vigor inicial se determinó contando el número total de píxeles de partes de la planta sobre el suelo discriminadas del fondo. Este valor fue el promedio de las fotografías tomadas en el mismo instante desde diferentes ángulos y se convirtió en un valor de superficie física expresado en milímetros cuadrados por medio de calibración. Los resultados relativos al vigor inicial descritos a continuación son para plantas tres semanas después de la germinación.

Mediciones de parámetros relacionados con la semilla

- 10 Se cosecharon las panículas primarias maduras, se contaron, se empacaron en bolsas, se etiquetaron con un código de barras y luego se secaron durante tres días en un horno a 37° C. Se trillaron luego las panículas, y se recogieron y contaron todas las semillas. Se separaron las vainas llenas de las vacías utilizando un dispositivo de soplado de aire. Las vainas vacías se descartaron y la fracción restante se contó de nuevo. Las vainas llenas se pesaron en una balanza analítica. El número de semillas llenas se determinó contando el número de vainas llenas que quedó después de la etapa de separación. El peso total de semillas por planta se midió pesando todas las vainas llenas cosechadas de una planta. Se midió la cantidad total de semillas por planta contando el número de vainas cosechadas de una planta. Se extrapoló el Peso de Mil Granos (PMG) a partir del número de semillas llenas contadas y su peso total. El Índice de Cosecha (IC) en la presente invención se define como la relación entre el peso total de semillas por planta y el área por encima del suelo (mm²), multiplicado por un factor de 10⁶. El número total de flores por panícula tal como se define en la presente invención es la relación entre el número total de semillas y el número de panículas primarias maduras. La tasa de llenado de semilla tal como se define en la presente invención es la proporción (expresada como %) del número de semillas llenas sobre el número total de semillas (o floretes).

Ejemplo 10: Resultados de la evaluación fenotípica de las plantas de arroz transgénicas bajo condiciones de crecimiento normales

- 25 Los resultados de la evaluación de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como el representado por la SEQ ID NO: 2, bajo el control del promotor GOS2 para expresión constitutiva, y cultivadas bajo condiciones normales de crecimiento, se presentan a continuación.

- 30 Hubo un incremento significativo en el número de floretes por panícula, en el rendimiento total de semillas por planta, en el número de semillas llenas, y en el índice de cosecha de las plantas transgénicas comparado con los correspondientes nulicigotos (controles), como se muestra en la Tabla E.

Tabla E: Resultados de la evaluación de las plantas de arroz transgénico que expresan la secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como el representado por la SEQ ID NO: 2, bajo el control del promotor GOS2 para expresión constitutiva, bajo condiciones de crecimiento normales.

Característica	Incremento promedio en % en 6 eventos en la generación T1
Número de flores por panícula	14%
Rendimiento total de semillas por planta	17%
Número total de semillas llenas	17%
Índice d cosecha	17%

- 35 **Ejemplo 11:** Resultados de la evaluación fenotípica de las plantas de arroz transgénicas bajo condiciones de crecimiento con disponibilidad reducida de nutrientes (nitrógeno)

- 40 Los resultados de la evaluación de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como el representado por la SEQ ID NO: 2, bajo el control del promotor GOS2 para expresión constitutiva, y cultivadas bajo condiciones de crecimiento con disponibilidad reducida de nutrientes (nitrógeno), se presentan a continuación.

Hubo un incremento significativo en el rendimiento total de semillas por planta, en el número de semillas llenas, y en el índice de cosecha de las plantas transgénicas comparado con los correspondientes nulicigotos (controles), como se muestra en la Tabla F.

Tabla F: Resultados de la evaluación de las plantas de arroz transgénico que expresan la secuencia de ácido nucleico que codifica un polipéptido AHL 19/20 como el representado por la SEQ ID NO: 2, bajo el control del promotor GOS2 para expresión constitutiva, bajo condiciones de crecimiento con disponibilidad reducida de nutrientes (nitrógeno).

5

Característica	Incremento promedio en % en 2 eventos en la generación T1
Vigor temprano	18%
Rendimiento total de semillas por planta	26%
Número total de semillas llenas	27%
Número total de semillas	24%

Ejemplo 12: Ejemplos de transformación de otros cultivos

Transformación de maíz

10 La transformación de maíz (*Zea mays*) se realiza con una modificación del método que describió por Ishida et al. (1996) Nature Biotech 14(6): 745 - 50. La transformación depende del genotipo en el maíz y únicamente genotipos
 15 específicos son responsables por la transformación y regeneración. La línea endogámica A188 (Universidad de Minnesota) o híbridos con A188 como progenitores son buenas fuentes de material donante para transformación, pero también pueden utilizarse otros genotipos exitosamente. Se cosecharon las mazorcas de plantas de maíz
 20 aproximadamente 11 días después de la polinización (DAP) cuando la longitud del embrión inmaduro es de aproximadamente 1 a 1,2 mm. Se cocultivarón los embriones inmaduros con *Agrobacterium tumefaciens* que contenía el vector de expresión, y se recuperaron las plantas transgénicas a través de organogénesis. Se cultivaron los embriones cortados en un medio de inducción de callos, luego en un medio de regeneración de maíz, que contiene el agente de selección (por ejemplo imidazolinona, pero pueden utilizarse diferentes marcadores de selección). Las placas de Petri se incubaron a la luz a 25°C durante 2 - 3 semanas, o hasta que crecieron brotes. Los brotes verdes se transfieren desde cada embrión a un medio enraizador de maíz y se incuban a 25°C durante 2 - 3 semanas, hasta que brotan raíces. Los brotes con raíces se trasplantan a suelo en el invernadero. Se producen semillas T1 a partir de las plantas que exhiben tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

25 Transformación de trigo

La transformación de trigo se realiza con el método descrito por Ishida et al. (1996) Nature Biotech 14(6): 745 - 50. La variedad cultivada Bobwhite (disponible a través de CIMMYT, Méjico) es comúnmente utilizada en la transformación. Los embriones inmaduros se cocultivan con *Agrobacterium tumefaciens* que contienen el vector de expresión, y se recuperan las plantas transgénicas a través de organogénesis. Después de incubación con
 30 *Agrobacterium*, se cultivan los embriones *in vitro* en un medio de inducción de callos, luego en un medio de regeneración que contiene el agente de selección (por ejemplo imidazolinona, pero pueden utilizarse diferentes marcadores de selección). Las placas de Petri se incuban a la luz a 25°C durante 2 - 3 semanas, o hasta que se desarrollaron brotes. Los brotes verdes se transfieren de cada embrión al medio enraizador y se incuban a 25°C durante 2 - 3 semanas, hasta que brotaron raíces. Los brotes con raíces se trasplantan a suelo en el invernadero. Se producen semillas T1 a partir de las plantas que exhiben tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

35

Transformación de soja

La soja se transforma de acuerdo con una modificación del método que se describe en la patente de los Estados Unidos No. 5.164.310 de Texas A&M. Diferentes variedades comerciales de soja son adecuadas para la transformación mediante este método. La variedad cultivada Jack (disponible a través de la fundación de Semillas de Illinois) es comúnmente utilizada para la transformación. Las semillas de soja se esterilizan para el sembrado *in vitro*. Se cortan el hipocotilo, la radícula y un cotiledón de plántulas jóvenes de 7 días. Se cultivan luego el epicotilo y el cotiledón restante hasta desarrollar nodos axiales. Estos nodos axiales se cortan y se incuban con *Agrobacterium tumefaciens* que contenían el vector de expresión. Después del tratamiento de cocultivo, se lavaron los explantes y se los transfirió a un medio de selección. Los brotes regenerados se cortan y se colocan en un medio para alargamiento del brote. Los brotes no mayores a 1 cm se colocan en un medio de enraizamiento hasta desarrollar raíces. Los brotes con raíces se trasplantan a suelo en el invernadero. Se producen semillas T1 a partir de las plantas que exhiben tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

45

Transformación de colza / canola

50 Se usaron pecíolos e hipocotilos de cotiledones de una plántula joven de 5 - 6 días como explantes para el cultivo de tejidos, y se transformaron de acuerdo con Babic et al. (1998, Plant Cell Rep 17: 183 - 188). La variedad cultivada

comercial Westar (Agriculture Canada) es la variedad estándar usada para la transformación, pero también pueden usarse otras variedades. Las semillas de canola se esterilizan en la superficie para la siembra *in vitro*. Se cortan los explantes de pecíolo de cotiledón con el cotiledón unido a partir de las plántulas *in vitro*, y se inoculan con *Agrobacterium* (que contienen el vector de expresión) sumergiendo el extremo cortado del explante del pecíolo en la suspensión bacteriana. Se cultivan luego los explantes durante 2 días sobre medio MSBAP-3 que contiene 3 mg/l de BAP, 3% de sacarosa, 0,7% de Phytagar a 23°C, con 16 h de luz. Después de dos días de cocultivo con *Agrobacterium*, los explantes de pecíolos se transfieren a un medio MSBAP-3 que contiene 3 mg/l de BAP, cefotaxima, carbenicilina o Timentina (300 mg/l) durante 7 días, y luego se cultivaron sobre medio MSBAP-3 con cefotaxima, carbenicilina o Timentina y un agente de selección hasta la regeneración de los brotes. Cuando los brotes tienen 5 - 10 mm de longitud, se cortan y se transfieren al medio para alargamiento de los brotes (MSBAP-0,5, que contiene 0,5 mg/l de BAP). Se transfieren los brotes de aproximadamente 2 cm de longitud al medio de enraizado (MS0) para la inducción de la raíz. Los brotes con raíces se trasplantan a suelo en el invernadero. Se producen semillas T1 a partir de las plantas que exhiben tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

15 Transformación de alfalfa

Se transforma un clon de regeneración de alfalfa (*Medicago sativa*) utilizando el método de (McKersie et al., 1999 Plant Physiol 119: 839 - 847). La regeneración y transformación de alfalfa depende del genotipo y por lo tanto se requiere una planta regeneradora. Ya se han descrito métodos para obtener plantas que se regeneran. Por ejemplo, estas pueden seleccionarse de la variedad cultivada Rangelander (Agriculture Canada) o de cualquier otra variedad comercial de alfalfa como lo describe Brown D. C. W. y A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111 - 112). Alternativamente, se ha seleccionado la variedad RA3 (Universidad de Wisconsin) para ser usada en el cultivo de tejidos (Walker et al., 1978 Am J Bot 65: 654 - 659). Los explantes de pecíolos se cocultivan con un cultivo de C58C1 pMP90 de *Agrobacterium tumefaciens* durante la noche (McKersie et al., 1999 Plant Physiol 119: 839 - 847) o LBA4404 que contiene el vector de expresión. Los explantes se cocultivan durante 3 días en la oscuridad en un medio de inducción SH que contiene 288 mg/L de Pro, 53 mg/L de tioprolina, 4,35 g/L de K₂SO₄, y 100 µm de acetosiringinona. Los explantes se lavan en un medio de Murashige-Skoog de fuerza media (Murashige y Skoog, 1962) y se siembran en placa sobre el mismo medio de inducción SH sin acetosiringinona, pero con un agente de selección adecuado y un antibiótico adecuado para inhibir el crecimiento de *Agrobacterium*. Después de varias semanas, se transfieren los embriones somáticos a un medio de desarrollo BOi2Y que no contiene reguladores de crecimiento ni antibióticos, y con 50 g/L de sacarosa. Luego se germinan los embriones somáticos en un medio de Murashige-Skoog de fuerza media. Las plántulas enraizadas se trasplantaron a macetas y se desarrollaron en un invernadero. Se producen semillas T1 a partir de las plantas que exhiben tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

Transformación de algodón

La transformación del algodón (*Gossypium hirsutum* L.) se lleva a cabo utilizando *Agrobacterium tumefaciens*, sobre explantes de hipocotilos. Las variedades de cultivo comerciales tales como Coker 130 o Coker 312 (SeedCo, Lubbock, TX) son variedades estándares utilizadas para la transformación, pero también se pueden utilizar otras variedades. Las semillas se esterilizan en la superficie y se germinan en la oscuridad. Los explantes de hipocotilos se cortan de las plántulas germinadas hasta longitudes de aproximadamente 1 - 1,5 centímetros. El explante de hipocotilo se sumerge en el inóculo de *Agrobacterium tumefaciens* que contiene el vector de expresión, durante 5 minutos y luego se cocultivan durante 48 horas sobre MS + 1,8 mg/l de KNO₃ + 2% de glucosa a 24° C, en la oscuridad. Los explantes se transfieren al mismo medio que contiene los marcadores seleccionables apropiados bacterianas y vegetales (renovados varias veces), hasta que se observan callos embriogénicos. Los callos se separan y se subcultivan hasta que aparecen embriones somáticos. Las plántulas derivadas de los embriones somáticos se maduran sobre el medio de enraizamiento hasta que se desarrollan las raíces. Los brotes con raíz se trasplantan a suelo en macetas en el invernadero. Se producen semillas T1 a partir de las plantas que exhiben tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

Ejemplo 13: Ejemplos de cribas de estrés abiótico

Criba por sequía

Se cultivan plantas de un número seleccionado de eventos en tierra para macetas bajo condiciones normales hasta que se acercan a la etapa de retiro de las mismas. A continuación, se transfieren a una sección "seca", donde se suprime el riego. Se insertan sondas de humedad en macetas elegidas al azar para controlar el contenido de agua del suelo (CAS). Cuando el CAS va por debajo de ciertos umbrales, se riega nuevamente las plantas en forma automática y continua hasta que se vuelva a alcanzar un nivel normal. Las plantas son entonces transferidas de nuevo a condiciones normales. El resto del cultivo (maduración de la planta, cosecha de semillas) es el mismo que para las plantas no cultivadas bajo condiciones de estrés abiótico. Los parámetros de crecimiento y rendimiento se registran como se detalla para el crecimiento bajo condiciones normales.

Criba por estrés salino

- 5 Las plantas se cultivan en un sustrato elaborado a base de fibras de coco y Argex (relación 3 a 1). Se utiliza una solución nutriente normal durante las dos primeras semanas después del trasplante de las plántulas en el invernadero. Después de las dos primeras semanas, se añade sal 25 mM (NaCl) a la solución nutriente, hasta que se cosechan las plantas. Los parámetros de crecimiento y rendimiento se registran como se detalla para el crecimiento bajo condiciones normales.

Ejemplo 14: Cribas por estrés abiótico

Cribado de eficiencia con uso de nitrógeno

- 10 Se cultivan plantas de arroz de generaciones T1, T2 o posteriores en tierra para macetas bajo condiciones normales excepto para la solución nutriente. Las macetas se riegan desde el trasplante hasta la maduración con una solución nutriente específica que tiene un contenido reducido de nitrógeno (N), usualmente entre 7 a 8 veces menos. El resto del cultivo (maduración de la 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 cultivo bajo condiciones normales.

15 LISTADO DE SECUENCIAS

<110> BASF Plant Science GmbH

<120> Plantas que tienen un aumento de características relacionadas con el rendimiento y un método para elaboración de las mismas

<130> PF59353

- 20 <160> 59

<170> PatentIn versión 3.3

<210> 1

<211> 948

<212> ADN

- 25 <213> Arabidopsis thaliana

<400> 1

```

atggcgaatc catggtggac aggacaagtg aacctatccg gcctcgaaac gacgccgcct      60
ggttcctctc agttaaagaa accagatctc cacatctcca tgaacatggc catggactca      120
ggtcacaata atcatcacca tcaccaagaa gtcgataaca acaacaacga cgacgataga      180
gacaacttga gtggagacga ccacgagcca cgtgaaggag ccgtagaagc cccacgcgcg      240
cgtccacgtg gacgtcctgc tggttccaag aacaaaccaa agccaccgat ctctcgtcact      300
cgcgattctc caaatgctct caagagccat gtcatggaga tcgctagtgg gactgacgtc      360
atcgaaaccc tagctacttt tgctaggcgg cgtcaacgtg gcatctgcat cttgagcgga      420
aatggcacag tggctaacgt caccctccgt caaccctcga ccgctgccgt tggggcggt      480
cctggtggtg cggctgtttt ggctttacaa gggagggttg agattctttc tttaacgggt      540
tctttcttgc caggaccggc tccacctggt tccaccgggt taacgattta cttagccggt      600
ggtcaaggtc aggttggtgg aggaagcgtg gtggggccat tgatggcagc aggtccggtg      660
atgctgatcg ccgccacgtt ctctaacgcg acttacgaga gattgccatt ggaggaggaa      720
gaggcagcag agagagcgcg tggtgaggcg agcggaggag tggttccggg gcagctcgga      780
ggcggagggt gccactaag cagcgggtgct ggtggaggcg acggtaacca aggactccg      840
gtgtataata tgcccgghaa tcttggttct aatggtggca gtggtggagg aggacagatg      900
agcggccaag aagcttatgg ttgggctcaa gctaggtcag gatthtaa      948

```

ES 2 440 265 T3

<210> 2

<211> 315

<212> PRT

<213> Arabidopsis thaliana

5 <400> 2

```

Met Ala Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Gly Leu Glu
1      5      10      15
Thr Thr Pro Pro Gly Ser Ser Gln Leu Lys Lys Pro Asp Leu His Ile
      20      25      30
Ser Met Asn Met Ala Met Asp Ser Gly His Asn Asn His His His
      35      40      45
Gln Glu Val Asp Asn Asn Asn Asn Asp Asp Asp Arg Asp Asn Leu Ser
      50      55      60
Gly Asp Asp His Glu Pro Arg Glu Gly Ala Val Glu Ala Pro Thr Arg
65      70      75      80
Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro Pro
      85      90      95
Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His Val Met
      100      105      110
Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr Phe Ala
      115      120      125
Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly Thr Val
      130      135      140
Ala Asn Val Thr Leu Arg Gln Pro Ser Thr Ala Ala Val Ala Ala Ala
145      150      155      160
Pro Gly Gly Ala Ala Val Leu Ala Leu Gln Gly Arg Phe Glu Ile Leu
      165      170      175
Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly Ser Thr
      180      185      190
Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly
      195      200      205
Ser Val Val Gly Pro Leu Met Ala Ala Gly Pro Val Met Leu Ile Ala
      210      215      220
Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu Glu Glu
225      230      235      240
Glu Ala Ala Glu Arg Gly Gly Gly Gly Gly Ser Gly Gly Val Val Pro
      245      250      255
Gly Gln Leu Gly Gly Gly Gly Ser Pro Leu Ser Ser Gly Ala Gly Gly
      260      265      270
Gly Asp Gly Asn Gln Gly Leu Pro Val Tyr Asn Met Pro Gly Asn Leu
      275      280      285
Val Ser Asn Gly Gly Ser Gly Gly Gly Gln Met Ser Gly Gln Glu
      290      295      300
Ala Tyr Gly Trp Ala Gln Ala Arg Ser Gly Phe
305      310      315

```

10 <210> 3

<211> 846

<212> ADN

<213> Arabidopsis thaliana

ES 2 440 265 T3

<400> 3

atggcaaacc	cttgggtggac	gaaccagagt	ggtttagcgg	gcatggtgga	ccattcggtc	60
tcctcaggcc	atcaccaaaa	ccatcaccac	caaagtcttc	ttaccaaaag	agatcttggg	120
atagccatga	atcagagcca	agacaacgac	caagacgaag	aagatgatcc	tagagaagga	180
gccgttgagg	tgggtcaaccg	tagaccaaga	ggtagaccac	caggatccaa	aaacaaacct	240
aaagctccaa	tctttgtgac	aagagacagc	cccaacgcac	tccgtagcca	tgtcttggag	300
atctccgacg	gcagtgacgt	cgccgacaca	atcgctcact	tctcaagacg	caggcaacgc	360
ggcgtttgcg	ttctcagcgg	gacaggctca	gtcgctaacg	tcaccctccg	ccaagccgcc	420
gcaccaggag	gtgtggtctc	tctccaaggc	aggtttgaaa	tcttatcttt	aaccggtgct	480
ttctccctg	gaccttcccc	accgggtca	accggtttaa	cggtttactt	agccggggtc	540
cagggtcagg	tcgttggagg	tagcgttgta	ggcccactct	tagccatagg	gtcgggtcatg	600
gtgattgctg	ctactttctc	taacgctact	tatgagagat	tgcccatgga	agaagaggaa	660
gacggtggcg	gctcaagaca	gattcacgga	ggcggtgact	caccgcccag	aatcggtagt	720
aacctgcctg	atctatcagg	gatggccggg	ccaggctaca	atatgccgcc	gcatctgatt	780
ccaaatgggg	ctggtcagct	agggcacgaa	ccatatacat	gggtccacgc	aagaccacct	840
tactga						846

<210> 4

<211> 281

5 <212> PRT

<213> Arabidopsis thaliana

<400> 4

ES 2 440 265 T3

atggcaaatc	catggtggac	tgggcaggtg	ggactgcctg	gtggttttaga	aacaggagcg	60
ggttcacctg	cgtttagaaa	acgcgatcga	gatttatcga	tgaatgaaag	tgtaagtgg	120
ggtagaggag	gtgaggatga	cgatgaaaga	gataacgggtg	atgagcctaa	agaagggtcg	180
gtagagatag	gtaaccgccg	tccaaggggc	cgaccacctg	ggtcaaagaa	caagccaaaa	240
ccaccgattt	ttgtgactcg	cgatagccca	aacgcgctta	ggagccatgt	gatggaggtc	300
tcaagtggga	ctgatgtagc	cgaaagtgta	gcccaatttg	ctaggaggcg	acaagaggt	360
gtttgtgtac	ttagtggtag	tggcgtagtg	gccaatgtaa	cattgcgaca	accttcagct	420
ccaagtgcag	ttgtggctct	gcaaggtcga	ttcgaaatat	tgtctctaac	tggttcattc	480
ttgcctgggc	cggcaccccc	aggatcaact	gggctgacgg	tctacttggc	aggcggtcag	540
gggcaagtgg	taggcggtag	ctggttggt	actcttattg	cagctggtcc	agttattgtg	600
attgcagcaa	catttgcaaa	tgcaacatat	gagagactac	caattgagga	ggaggaggat	660
gcaggaagtg	gaggtcaggg	acaactccag	ggcgggtgcag	gaagctcacc	accaccaatt	720
ggaagcagta	ccgggcaaca	gcaaccaggg	atgccagacc	tatcctcttt	gccagtgtat	780
aatatgccac	caaacctact	caaaaatgga	gggcagatga	accagcaaga	agcatatgct	840
tgggctcatg	ctcggccacc	gtattga				867

<210> 6

<211> 288

5 <212> PRT

<213> *Aquilegia formosa* x *Aquilegia pubescens*

<400> 6

ES 2 440 265 T3

Met Ala Asn Pro Trp Trp Thr Gly Gln Val Gly Leu Pro Gly Gly Leu
1 5 10 15
Glu Thr Gly Ala Gly Ser Pro Ala Phe Arg Lys Arg Asp Arg Asp Leu
20 25 30
Ser Met Asn Glu Ser Val Ser Gly Gly Arg Gly Gly Glu Asp Asp Asp
35 40 45
Glu Arg Asp Asn Gly Asp Glu Pro Lys Glu Gly Ala Val Glu Ile Gly
50 55 60
Asn Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro Lys
65 70 75 80
Pro Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Arg Ser His
85 90 95
Val Met Glu Val Ser Ser Gly Thr Asp Val Ala Glu Ser Val Ala Gln
100 105 110
Phe Ala Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Gly Ser Gly
115 120 125
Val Val Ala Asn Val Thr Leu Arg Gln Pro Ser Ala Pro Ser Ala Val
130 135 140
Val Ala Leu Gln Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly Ser Phe
145 150 155 160
Leu Pro Gly Pro Ala Pro Pro Gly Ser Thr Gly Leu Thr Val Tyr Leu
165 170 175
Ala Gly Gly Gln Gly Gln Val Val Gly Gly Ser Val Val Gly Thr Leu
180 185 190
Ile Ala Ala Gly Pro Val Ile Val Ile Ala Ala Thr Phe Ala Asn Ala
195 200 205
Thr Tyr Glu Arg Leu Pro Ile Glu Glu Glu Glu Asp Ala Gly Ser Gly
210 215 220
Gly Gln Gly Gln Leu Gln Gly Gly Ala Gly Ser Ser Pro Pro Pro Ile
225 230 235 240
Gly Ser Ser Thr Gly Gln Gln Gln Pro Gly Met Pro Asp Leu Ser Ser
245 250 255
Leu Pro Val Tyr Asn Met Pro Pro Asn Leu Leu Gln Asn Gly Gly Gln
260 265 270
Met Asn Gln Gln Glu Ala Tyr Ala Trp Ala His Ala Arg Pro Pro Tyr
275 280 285

- 5 <210> 7
- <211> 948
- <212> ADN
- <213> Brassica napus
- <400> 7

ES 2 440 265 T3

atggcgaatc	catggtggac	aggacaagtg	aatctctccg	gcctcgaaac	gacgccgccg	60
agttcctctc	agttaaagac	accagatctc	cacatctcca	tgaatatggc	catggactca	120
ggtcataaca	accaccacca	tcatcaccaa	gaagtcaaca	ccaacaacaa	caacgaagac	180
gatagagaca	acttgagcgg	cgacgaccac	gagccacgtg	aaggagccgt	ggaagctccc	240
acgcgccgac	cacgtggacg	tcctgctggt	tccaagaaca	aaccaaagcc	accaatcttt	300
gtcacgcgtg	actctccaaa	cgctctcaag	agccatgtca	tggagatcgc	tagtgggact	360
gacgtcatcg	aaaccctagc	tactttcgct	aggcggcgcc	aacgtggcat	ctgcatcttg	420
agcggtaacg	gcacggtggc	taacgtcaca	ctcctgcaac	catcagtggc	tcccgttgca	480
gctgcccttg	gtggtgcggc	tgtattggcg	ttacaagggg	ggtttgagat	tctttctcta	540
accggttctt	tcttacctgg	accggctcca	cctggatcca	ctggtttaac	tatttactta	600
gctggtggtc	aaggtcaggt	tgttgaggga	agcgtggtgg	ggccattgat	ggctgctggt	660
ccggtgatgc	taatcgctgc	cacgttttct	aatgcgactt	atgagagatt	acctttggat	720
gaggaagaag	cggctgaaag	aggtggcggg	ggaagcgacg	gaggagtggg	tccagggcag	780
ctcggggggc	taggttcccc	gctgagtagt	ggtggcgggt	gaggccatgg	gaaccaagga	840
cttcccgcgt	ataatatgcc	cggaaatctt	gcttctaata	gcggtggagg	aggacagatg	900
agcggccaag	aagcttacgg	ttgggctcaa	gctaggtcag	gattttaa		948

<210> 8

<211> 315

5 <212> PRT

<213> Brassica napus

<400> 8

ES 2 440 265 T3

Met	Ala	Asn	Pro	Trp	Trp	Thr	Gly	Gln	Val	Asn	Leu	Ser	Gly	Leu	Glu
1				5					10					15	
Thr	Thr	Pro	Pro	Ser	Ser	Ser	Gln	Leu	Lys	Thr	Pro	Asp	Leu	His	Ile
			20					25					30		
Ser	Met	Asn	Met	Ala	Met	Asp	Ser	Gly	His	Asn	Asn	His	His	His	His
		35					40					45			
His	Gln	Glu	Val	Asn	Thr	Asn	Asn	Asn	Asn	Glu	Asp	Asp	Arg	Asp	Asn
	50					55					60				
Leu	Ser	Gly	Asp	Asp	His	Glu	Pro	Arg	Glu	Gly	Ala	Val	Glu	Ala	Pro
65					70					75					80
Thr	Arg	Arg	Pro	Arg	Gly	Arg	Pro	Ala	Gly	Ser	Lys	Asn	Lys	Pro	Lys
				85					90					95	
Pro	Pro	Ile	Phe	Val	Thr	Arg	Asp	Ser	Pro	Asn	Ala	Leu	Lys	Ser	His
			100					105					110		
Val	Met	Glu	Ile	Ala	Ser	Gly	Thr	Asp	Val	Ile	Glu	Thr	Leu	Ala	Thr
		115					120					125			
Phe	Ala	Arg	Arg	Arg	Gln	Arg	Gly	Ile	Cys	Ile	Leu	Ser	Gly	Asn	Gly
	130					135					140				
Thr	Val	Ala	Asn	Val	Thr	Leu	Arg	Gln	Pro	Ser	Val	Ala	Pro	Val	Ala
145					150					155					160
Ala	Ala	Pro	Gly	Gly	Ala	Ala	Val	Leu	Ala	Leu	Gln	Gly	Arg	Phe	Glu
				165					170					175	
Ile	Leu	Ser	Leu	Thr	Gly	Ser	Phe	Leu	Pro	Gly	Pro	Ala	Pro	Pro	Gly
			180					185					190		
Ser	Thr	Gly	Leu	Thr	Ile	Tyr	Leu	Ala	Gly	Gly	Gln	Gly	Gln	Val	Val
		195					200					205			
Gly	Gly	Ser	Val	Val	Gly	Pro	Leu	Met	Ala	Ala	Gly	Pro	Val	Met	Leu
	210					215					220				
Ile	Ala	Ala	Thr	Phe	Ser	Asn	Ala	Thr	Tyr	Glu	Arg	Leu	Pro	Leu	Asp
225					230					235					240
Glu	Glu	Glu	Ala	Ala	Glu	Arg	Gly	Gly	Gly	Gly	Ser	Asp	Gly	Gly	Val
				245					250					255	
Val	Pro	Gly	Gln	Leu	Gly	Gly	Val	Gly	Ser	Pro	Leu	Ser	Ser	Gly	Gly
			260					265					270		
Gly	Gly	Gly	His	Gly	Asn	Gln	Gly	Leu	Pro	Ala	Tyr	Asn	Met	Pro	Gly
		275					280					285			
Asn	Leu	Ala	Ser	Asn	Gly	Gly	Gly	Gly	Gly	Gln	Met	Ser	Gly	Gln	Glu
	290				295						300				
Ala	Tyr	Gly	Trp	Ala	Gln	Ala	Arg	Ser	Gly	Phe					
305					310					315					

<210> 9

<211> 978

5 <212> ADN

<213> Brassica rapa

<400> 9

ES 2 440 265 T3

ataatcagat	acaatctatt	tagggtttta	atggcgaatc	catggtggac	aggacaagtg	60
aatctctccg	gcctcgaaac	gacgccgccg	agttcctctc	agttaaagac	accagatctc	120
cacatctcca	tgaatatggc	catggactca	ggtcataaca	accaccacca	tcatcaccaa	180
gaagtcaaca	ccaacaacaa	caacgaagac	gatagagaca	acttgagcgg	cgacgaccac	240
gagccacgtg	aaggagccgt	ggaagctccc	acgcgccgac	cacgtggacg	tcttctgtgt	300
tccaagaaca	aaccaaagcc	accaatcttt	gtcacgcgtg	actctccaaa	cgctctcaag	360
agccatgtca	tggagatcgc	tagtgggact	gacgtcatcg	aaaccctagc	tactttcgct	420
aggcggcgcc	aacgtggcat	ctgcatcttg	agcggtaacg	gcacgggtggc	taacgtcaca	480
ctccgtcaac	catcagtggc	tcccgttgca	gctgcccctg	gtggtgcggc	tgtattggcg	540
ttacaagggg	ggtttgagat	tctttctcta	accggttctt	tcttacctgg	accggctcca	600
cctggatcca	ctggtttaac	tatttactta	gctgggtggc	aaggtcaggt	tgttgaggga	660
agcgtggtgg	ggccattgat	ggctgctggt	ccggtgatgc	taatcgctgc	cacgttttct	720
aatgcgactt	atgagagatt	acctttggat	gaggaagaag	cggctgaaag	aggtggcggt	780
ggaagcgacg	gaggagtggg	tccagggcag	ctcggggggc	taggttcccc	gctgagtagt	840
ggtggcggtg	gaggccatgg	gaaccaagga	cttcccgcgt	ataatatgcc	cggaaatctt	900
gcttctaata	gcggtggagg	aggacagatg	agcggccaag	aagcttacgg	ttgggctcaa	960
gctaggtcag	gattttaa					978

<210> 10

<211> 315

5 <212> PRT

<213> Brassica rapa

<400> 10

ES 2 440 265 T3

Met Ala Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Gly Leu Glu
1 5 10 15
Thr Thr Pro Pro Ser Ser Ser Gln Leu Lys Thr Pro Asp Leu His Ile
20 25 30
Ser Met Asn Met Ala Met Asp Ser Gly His Asn Asn His His His His
35 40 45
His Gln Glu Val Asn Thr Asn Asn Asn Asn Glu Asp Asp Arg Asp Asn
50 55 60
Leu Ser Gly Asp Asp His Glu Pro Arg Glu Gly Ala Val Glu Ala Pro
65 70 75 80
Thr Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys
85 90 95
Pro Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His
100 105 110
Val Met Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr
115 120 125
Phe Ala Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly
130 135 140
Thr Val Ala Asn Val Thr Leu Arg Gln Pro Ser Val Ala Pro Val Ala
145 150 155 160
Ala Ala Pro Gly Gly Ala Ala Val Leu Ala Leu Gln Gly Arg Phe Glu
165 170 175
Ile Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly
180 185 190
Ser Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val
195 200 205
Gly Gly Ser Val Val Gly Pro Leu Met Ala Ala Gly Pro Val Met Leu
210 215 220
Ile Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Leu Asp
225 230 235 240
Glu Glu Glu Ala Ala Glu Arg Gly Gly Gly Gly Ser Asp Gly Gly Val
245 250 255
Val Pro Gly Gln Leu Gly Gly Val Gly Ser Pro Leu Ser Ser Gly Gly
260 265 270
Gly Gly Gly His Gly Asn Gln Gly Leu Pro Ala Tyr Asn Met Pro Gly
275 280 285
Asn Leu Ala Ser Asn Gly Gly Gly Gly Gly Gln Met Ser Gly Gln Glu
290 295 300
Ala Tyr Gly Trp Ala Gln Ala Arg Ser Gly Phe
305 310 315

<210> 11

<211> 843

5 <212> ADN

<213> Glycine max

<400> 11

ES 2 440 265 T3

```

atggccaacc ggtggtggac cgggtcgggtg ggtctagaga actctggcca ctcgatgaaa      60
aaaccggatc tggggttttc catgaacgag agtacgggtga cggggaacca tataggagaa      120
gaagatgagg acagagaaaa cagcgacgag ccaagagagg gagctattga cgtcgccacc      180
acgcgccgcc ctaggggacg tccaccgggc tccagaaaca agccgaaacc gccgatattc      240
gtcacccgag acagccctaa cgcgctgagg agccacgtca tggagattgc cgtcggagcc      300
gacatcgccg actgctggc gcagttcgtc cggaggcgcc agcgcggggt ttccattctc      360
agcggcagcg ggaccgtcgt caacgtcaat ctccggcaac ccacggcacc cggcgcctc      420
atggcgctcc acggcgcgtt cgacatcctc tccctcaccg gctcctttct cctggggccg      480
tcccctcccg gcgccaccgg gtcacaatc tacctcgccg gaggccaggg gcagatcgtc      540
ggcggcggag tgggtgggccc gctcgtggcg gcgggccccg tattggtaat ggcggtact      600
ttttccaatg ctacgtatga aagattgcct ttagaggatg atgatcagga acaacacggc      660
ggcggaggcg gaggaggttc gccgcaggaa aaaaccgggg gtcccggcga ggcgtcgtcg      720
tcgatttcgg tttataacaa taatgttctt ccgagtttag gtcttccgaa tgggcaacat      780
ctgaaccatg aagcttattc ttctccttgg ggtcattctc ctcatgccag acctccttc      840
taa

```

<210> 12

<211> 280

5 <212> PRT

<213> Glycine max

<400> 12

```

Met Ala Asn Arg Trp Trp Thr Gly Ser Val Gly Leu Glu Asn Ser Gly
 1                    5                    10                    15
His Ser Met Lys Lys Pro Asp Leu Gly Phe Ser Met Asn Glu Ser Thr
    20                    25                    30
Val Thr Gly Asn His Ile Gly Glu Glu Asp Glu Asp Arg Glu Asn Ser
    35                    40                    45
Asp Glu Pro Arg Glu Gly Ala Ile Asp Val Ala Thr Thr Arg Arg Pro
    50                    55                    60
Arg Gly Arg Pro Pro Gly Ser Arg Asn Lys Pro Lys Pro Pro Ile Phe
    65                    70                    75                    80
Val Thr Arg Asp Ser Pro Asn Ala Leu Arg Ser His Val Met Glu Ile
    85                    90                    95
Ala Val Gly Ala Asp Ile Ala Asp Cys Val Ala Gln Phe Ala Arg Arg
    100                    105                    110
Arg Gln Arg Gly Val Ser Ile Leu Ser Gly Ser Gly Thr Val Val Asn
    115                    120                    125
Val Asn Leu Arg Gln Pro Thr Ala Pro Gly Ala Val Met Ala Leu His
    130                    135                    140
Gly Arg Phe Asp Ile Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro
    145                    150                    155                    160
Ser Pro Pro Gly Ala Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln
    165                    170                    175
Gly Gln Ile Val Gly Gly Gly Val Val Gly Pro Leu Val Ala Ala Gly
    180                    185                    190
Pro Val Leu Val Met Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg
    195                    200                    205
Leu Pro Leu Glu Asp Asp Asp Gln Glu Gln His Gly Gly Gly Gly
    210                    215                    220
Gly Gly Ser Pro Gln Glu Lys Thr Gly Gly Pro Gly Glu Ala Ser Ser
    225                    230                    235                    240
Ser Ile Ser Val Tyr Asn Asn Asn Val Pro Pro Ser Leu Gly Leu Pro

```

ES 2 440 265 T3

				245					250					255	
Asn	Gly	Gln	His	Leu	Asn	His	Glu	Ala	Tyr	Ser	Ser	Pro	Trp	Gly	His
			260					265					270		
Ser	Pro	His	Ala	Arg	Pro	Pro	Phe								
		275					280								

<210> 13

<211> 834

5 <212> ADN

<213> Gossypium hirsutum

<400> 13

atggaccg	caggcaattc	accagcttta	aacaaaogtg	accttgaat	ttctatgaac	60
gatgctaaca	aaagtagaag	caacggaaga	ggggatgatg	atgatgaaga	tagagacacc	120
ggcgatgagc	ctaaagaagg	agcggtcgag	gtcggtaacc	gaagaccccg	aggtcgtcca	180
ccgggatcca	aaaacaagcc	taaaccaccc	atTTTTgtga	caagggatag	ccctaacgcg	240
ctccgtagtc	atgttatgga	agtcgcaagt	ggaaccgatg	tagccgagag	tatagcccaa	300
ttcgctcgga	gaagacaacg	tggagtttgt	ttgcttagcg	gcagcggctc	ggtcgccaac	360
gttactctaa	gacaaccggc	agcaccggc	gcggtggttg	cccttcatgg	aaggtttgaa	420
atTTTgtctt	tgaccggggc	TTTTctccc	ggaccggctc	caccgggatc	gacagggctc	480
accgtgtact	tagctggtgg	tcaaggacaa	gTTgTTggag	gaagtgttgt	cggctcactt	540
atagcagcag	ggcctgttat	ggtcattgca	gcaactTTTT	ccaacgcaac	ttatgaaaga	600
ctgccttttag	aagatgaaga	agaagttgta	agcgcgggtc	acggtggacc	gatgcaaggc	660
ggagcaaacg	attcaccgcc	ggaaattggg	agtagcggag	gcggcggttc	acacacaggt	720
ctgcctgatc	catcttact	tccaatatac	aatTTgcctc	ctaatttact	ctcaaatgga	780
gggcaactag	ggcatgaacc	ctatggttgg	acacatggga	gaccacccta	ttaa	834

10

<210> 14

<211> 277

<212> PRT

<213> Gossypium hirsutum

15

<400> 14

ES 2 440 265 T3

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

Val Val Ser Ala Gly His Gly Gly Pro Met Gln Gly Gly Ala Asn Asp
 210 215 220
 Ser Pro Pro Glu Ile Gly Ser Ser Gly Gly Gly Ser His Thr Gly
 225 230 235 240
 Leu Pro Asp Pro Ser Ser Leu Pro Ile Tyr Asn Leu Pro Pro Asn Leu
 245 250 255
 Leu Ser Asn Gly Gly Gln Leu Gly His Glu Pro Tyr Gly Trp Thr His
 260 265 270
 Gly Arg Pro Pro Tyr
 275

<210> 15

<211> 813

5 <212> ADN

<213> Lactuca sativa

<400> 15

ES 2 440 265 T3

atgtctaacc	gatggtggac	cggccaggtc	aacgtggcag	gcgtagaaac	atcatctcag	60
gcgatcaaga	aaccagatct	gggtatctca	atgaatgata	ccaccacagg	aagtgaagaa	120
gatgaaagag	acaacaacag	cgatgatcca	agagaagggtg	caattgaccc	ttctaaccgt	180
aggccacgag	gccgacctcc	gggatccaaa	aacaaaccaa	agccaccgat	tttcgtcacc	240
agagacagcc	ctaacgccct	ccgcagccac	gtcatggagg	tagcgagtgg	tacagatata	300
gcagaaagta	tagctcaatt	cagccgaaaa	cgacaacgcg	gtgtgtgtgt	gatgagtgct	360
agcggcacag	tcatgaatgt	aaccctaaga	caaccttcgg	cacctggctc	agtcattggct	420
ctacaaggcc	ggttcgagat	tttatcccta	accggtgcct	tcttaccggg	tccttctcct	480
cctggatcca	ccgggctcac	tatatattta	gctggtggcc	agggccaggc	tgtgggaggc	540
agcgtggtgg	gatcattggt	ggcatcagga	ccagtgatgg	ttatagcagc	cacgttctcc	600
aacgccacat	atgaaagact	cccggttgag	gaagaggagg	aagcagatac	cgtgacacct	660
gggctagggtg	gtggtggatc	accaccgcaa	ctcggaatgg	gtgatcagaa	tccgatggca	720
gggtataata	tgcagccgaa	tttgatcccg	aatggtgggtg	gacagatgaa	ccatgaagct	780
tttgctttgg	ctcatggccg	gcccacgtac	tag			813

<210> 16

<211> 270

5 <212> PRT

<213> Lactuca sativa

<400> 16

ES 2 440 265 T3

Met	Ser	Asn	Arg	Trp	Trp	Thr	Gly	Gln	Val	Asn	Val	Ala	Gly	Val	Glu
1				5					10					15	
Thr	Ser	Ser	Gln	Ala	Ile	Lys	Lys	Pro	Asp	Leu	Gly	Ile	Ser	Met	Asn
			20					25					30		
Asp	Thr	Thr	Thr	Gly	Ser	Glu	Glu	Asp	Glu	Arg	Asp	Asn	Asn	Ser	Asp
		35					40					45			
Asp	Pro	Arg	Glu	Gly	Ala	Ile	Asp	Pro	Ser	Asn	Arg	Arg	Pro	Arg	Gly
	50					55					60				
Arg	Pro	Pro	Gly	Ser	Lys	Asn	Lys	Pro	Lys	Pro	Pro	Ile	Phe	Val	Thr
65					70						75				80
Arg	Asp	Ser	Pro	Asn	Ala	Leu	Arg	Ser	His	Val	Met	Glu	Val	Ala	Ser
				85					90					95	
Gly	Thr	Asp	Ile	Ala	Glu	Ser	Ile	Ala	Gln	Phe	Ser	Arg	Lys	Arg	Gln
			100					105					110		
Arg	Gly	Val	Cys	Val	Met	Ser	Ala	Ser	Gly	Thr	Val	Met	Asn	Val	Thr
		115					120					125			
Leu	Arg	Gln	Pro	Ser	Ala	Pro	Gly	Ser	Val	Met	Ala	Leu	Gln	Gly	Arg
	130					135					140				
Phe	Glu	Ile	Leu	Ser	Leu	Thr	Gly	Ala	Phe	Leu	Pro	Gly	Pro	Ser	Pro
145					150					155					160
Pro	Gly	Ser	Thr	Gly	Leu	Thr	Ile	Tyr	Leu	Ala	Gly	Gly	Gln	Gly	Gln
				165					170					175	
Val	Val	Gly	Gly	Ser	Val	Val	Gly	Ser	Leu	Val	Ala	Ser	Gly	Pro	Val
			180					185					190		
Met	Val	Ile	Ala	Ala	Thr	Phe	Ser	Asn	Ala	Thr	Tyr	Glu	Arg	Leu	Pro
		195					200					205			
Val	Glu	Glu	Glu	Glu	Glu	Ala	Asp	Thr	Val	Thr	Pro	Gly	Leu	Gly	Gly
	210					215					220				
Gly	Gly	Ser	Pro	Pro	Gln	Leu	Gly	Met	Gly	Asp	Gln	Asn	Pro	Met	Ala
225					230					235					240
Gly	Tyr	Asn	Met	Gln	Pro	Asn	Leu	Ile	Pro	Asn	Gly	Gly	Gly	Gln	Met
				245					250					255	
Asn	His	Glu	Ala	Phe	Ala	Leu	Ala	His	Gly	Arg	Pro	Thr	Tyr		
			260					265					270		

<210> 17

<211> 882

5 <212> ADN

<213> Lotus japonicus

<400> 17

ES 2 440 265 T3

```

atggctaatac cttggtggac aagccagga gggttctctg gggttgacct aggaacccat      60
tcacctggct tgagcaaacg tcacacggac cttgtgatca atgaaaacag cagcgggtgt      120
aatagagatg aagatgaaga tgataacagg gaagatgagc caaaagaagg tgcagttgag      180
gttggaactc ggagaccaag gggaagacca cgggatcca agaacaagcc aagaccacc      240
atctttgtaa caagggacag cccaaacgcc ctgaggagtc atgttatgga ggttgcagga      300
ggagctgatg tcgcagaaag cgtggcccag tttgcgagga ggcgccagcg tggggtttgt      360
gtgatgagcg ggagtggctc tgtggcaaac gttaccctga gacaacctgc ggctccgggt      420
gctgtttag cactccatgg caggtttag atcttatccc taactggggc gttcctacct      480
ggccctgctc ctccaggatc cactggtcta acagtgtatc tttctggagg acagggtcag      540
gtagtgggag ggagtgtggt ggggtctcta gttgcagcag gaccagttat ggtcattgct      600
gcaacttttg ctaatgcaac atatgagagg ttgccacttg atgatgatga tgagggacct      660
agtggggccg ctacggcggc aagcggagga ggaagtggat cgtctcctcc acctggaatt      720
ggaattggca gtggtggggg tcatcaactg caggctggac tggttccaga tccatcatcc      780
atgccgttgt ataatctgcc accaaatctg ttgtccaatg gaggaggagg acaagtgggg      840
catgatgctc ttgcttgggc tcatggaaga acaccttact ga                          882

```

<210> 18

<211> 293

5 <212> PRT

<213> Lotus japonicus

<400> 18

```

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

```

ES 2 440 265 T3

	130					135					140				
Leu	His	Gly	Arg	Phe	Glu	Ile	Leu	Ser	Leu	Thr	Gly	Ala	Phe	Leu	Pro
145					150					155					160
Gly	Pro	Ala	Pro	Pro	Gly	Ser	Thr	Gly	Leu	Thr	Val	Tyr	Leu	Ser	Gly
				165					170					175	
Gly	Gln	Gly	Gln	Val	Val	Gly	Gly	Ser	Val	Val	Gly	Ser	Leu	Val	Ala
			180					185					190		
Ala	Gly	Pro	Val	Met	Val	Ile	Ala	Ala	Thr	Phe	Ala	Asn	Ala	Thr	Tyr
		195					200					205			
Glu	Arg	Leu	Pro	Leu	Asp	Asp	Asp	Asp	Glu	Gly	Pro	Ser	Gly	Ala	Ala
210					215						220				
Thr	Ala	Ala	Ser	Gly	Gly	Gly	Ser	Gly	Ser	Ser	Pro	Pro	Pro	Gly	Ile
225					230						235				240
Gly	Ile	Gly	Ser	Gly	Gly	Gly	His	Gln	Leu	Gln	Ala	Gly	Leu	Val	Pro
				245					250					255	
Asp	Pro	Ser	Ser	Met	Pro	Leu	Tyr	Asn	Leu	Pro	Pro	Asn	Leu	Leu	Ser
			260					265					270		
Asn	Gly	Gly	Gly	Gly	Gln	Val	Gly	His	Asp	Ala	Leu	Ala	Trp	Ala	His
		275					280					285			
Gly	Arg	Thr	Pro	Tyr											
	290														

<210> 19

<211> 708

5 <212> ADN

<213> Oryza sativa

<400> 19

atggcggtcca	aggagccaag	cgggcgaccac	gaccacgaga	tgaacgggac	cagcgccggg	60
ggcggcgagc	ccaaggacgg	cgcggtggtg	accggccgca	accggcgccc	ccgcggacgg	120
ccgccgggct	ccaagaacaa	gcccagccg	cccattctcg	tgacgcggga	cagcccgaac	180
gcgctgcgca	gccacgtcat	ggaggtggcc	ggcggcgccg	atgtcgccga	gtccatcgcg	240
cacttcgcgc	ggcggcgcca	gcgcggcgtc	tgcgtgctca	gcggggccgg	caccgtgacc	300
gacgtggccc	tgcgccagcc	ggccgcgccg	agcgcctggg	tggcgctccg	tgggcggttc	360
gagatcctgt	ccctgacggg	gacgttcctg	ccggggccgg	cgccgcgggg	ctccaccggg	420
ctgaccgtgt	acctcgccgg	cgggcagggg	caggtggtgg	gcggcagcgt	ggtggggacg	480
ctcaccgagg	cggggcgggt	catggtgatc	gcctccacct	tcgccaacgc	cacctacgag	540
aggctgccc	tggatcagga	ggaggaggaa	gcagcggcag	gcggcatgat	ggcgccggcg	600
ccactcatgg	ccggcgccgc	cgatccacta	cttttcggcg	ggggaatgca	cgacgcgggg	660
cttgcctcat	ggcaccatgc	ccgcctccg	ccggcgccgc	cctactag		708

10

<210> 20

<211> 235

<212> PRT

<213> Oryza sativa

15

<400> 20

ES 2 440 265 T3

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

Gly	Thr	Val	Thr	Asp	Val	Ala	Leu	Arg	Gln	Pro	Ala	Ala	Pro	Ser	Ala
			100					105					110		
Val	Val	Ala	Leu	Arg	Gly	Arg	Phe	Glu	Ile	Leu	Ser	Leu	Thr	Gly	Thr
		115					120					125			
Phe	Leu	Pro	Gly	Pro	Ala	Pro	Pro	Gly	Ser	Thr	Gly	Leu	Thr	Val	Tyr
		130				135					140				
Leu	Ala	Gly	Gly	Gln	Gly	Gln	Val	Val	Gly	Gly	Ser	Val	Val	Gly	Thr
145					150					155					160
Leu	Thr	Ala	Ala	Gly	Pro	Val	Met	Val	Ile	Ala	Ser	Thr	Phe	Ala	Asn
				165					170					175	
Ala	Thr	Tyr	Glu	Arg	Leu	Pro	Leu	Asp	Gln	Glu	Glu	Glu	Glu	Ala	Ala
			180					185					190		
Ala	Gly	Gly	Met	Met	Ala	Pro	Pro	Pro	Leu	Met	Ala	Gly	Ala	Ala	Asp
		195				200						205			
Pro	Leu	Leu	Phe	Gly	Gly	Gly	Met	His	Asp	Ala	Gly	Leu	Ala	Ala	Trp
	210					215					220				
His	His	Ala	Arg	Pro	Pro	Pro	Pro	Pro	Pro	Tyr					
225					230					235					

<210> 21

<211> 801

5 <212> ADN

<213> Oryza sativa

<400> 21

atgggcttgc	cggagcagcc	gtccggctcg	tggggcccca	aggcggagct	cccggtggcc	60
aaggagccgg	aggcgagccc	gacggggggc	gcggcggcgg	accacgccga	cgagaacaac	120
gaatccggcg	gcggcgagcc	gcgggagggc	gccgtggtgg	cggcgcccaa	ccggcgcccc	180
cgcgcccgcc	cgccgggctc	caagaacaag	ccgaagccgc	ccatcttcgt	gacgcgcgac	240
agccccaacg	cgctgcgcag	tcacgtcatg	gaggtggccg	gcggcgccga	cgtcgcgcgac	300
gccategcgc	agttctcgcg	cgcgccagcg	cgcggcgtct	gcgtgctcag	cggcgcgggg	360
acggtcgcca	acgtcgcgct	gcgccagccg	tggcgccccg	gcgccgtcgt	cgccctgcac	420
ggccgcttcg	agatcctctc	cctcaccggc	accttcctcc	caggcccggc	gcctccgggt	480
tccacggggc	tcaccgtcta	cctcgccggc	ggccagggcc	aggttgctcg	cggcagcgtc	540
gtggggctgc	tcategcgcg	gggcccggtc	atggtgatcg	cgccacggtt	cgccaacgcc	600
acctacgagc	gcctgccact	ggaggaagaa	gaggagggtc	caggcccggc	catgcccggc	660
ggcgcggagc	ccctcatggc	cggcgccac	ggcctcgccg	acccttcggc	gctgccaatg	720
ttcaacctgc	cgccgagcaa	cgggctcggc	ggcggcgggc	acggcttccc	atgggcggcg	780
caccctgcc	caccgtactg	a				801

ES 2 440 265 T3

<210> 22

<211> 266

<212> PRT

<213> Oryza sativa

5 <400> 22

```

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

          100          105          110
Val Cys Val Leu Ser Gly Ala Gly Thr Val Ala Asn Val Ala Leu Arg
115      120      125
Gln Pro Ser Ala Pro Gly Ala Val Val Ala Leu His Gly Arg Phe Glu
130      135      140
Ile Leu Ser Leu Thr Gly Thr Phe Leu Pro Gly Pro Ala Pro Pro Gly
145      150      155      160
Ser Thr Gly Leu Thr Val Tyr Leu Ala Gly Gly Gln Gly Gln Val Val
165      170      175
Gly Gly Ser Val Val Gly Ser Leu Ile Ala Ala Gly Pro Val Met Val
180      185      190
Ile Ala Ser Thr Phe Ala Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu
195      200      205
Glu Glu Glu Glu Gly Ser Gly Pro Pro Met Pro Gly Gly Ala Glu Pro
210      215      220
Leu Met Ala Gly Gly His Gly Ile Ala Asp Pro Ser Ala Leu Pro Met
225      230      235      240
Phe Asn Leu Pro Pro Ser Asn Gly Leu Gly Gly Gly Gly Asp Gly Phe
245      250      255
Pro Trp Ala Ala His Pro Cys Pro Pro Tyr
260      265

```

<210> 23

10 <211> 855

<212> ADN

<213> Populus tremuloides

<400> 23

ES 2 440 265 T3

```

atggcaaacc ggtggtggac agggcaagtg ggattgccgg ggatggacac atcaaccagt      60
tcacatctc caatgaaaaa gccagatcta ggtatatcca tgtccaacaa caatagagaa      120
gccaccgaga gtggtgctgg caaagaagat gagcaagaag acgaaagaga aaatagcgac      180
gagcctagag aaggcgctat agatatcgcc tctcgccgcc ctagaggccg tccaccaggg      240
tccaagaaca agcctaagcc accaattttc gttactcgag acagccctaa tgcactcaag      300
agtcatgtga tggagatagc tagtggatct gatatagctg aaaatttagc ttgttttgca      360
aggaagagac aaagaggagt ttgtgtgctt agtggaagtg gtatggtaac caatgtaacc      420
ctcaagcaac cttctgcctc aggtgctggt atggctctcc atggtagggt tgagattttg      480
tcactcactg gagcgttctt gcctggacca gcccacctg gagcgacagg actaactata      540
tatttagccg gagggcaagg acaagtggta ggaggcagtg tggtaggata actagttgca      600
tcaggaccgg taatggttat tgctgcaaca ttttcaaatg ctacttatga gagattgcca      660
ctagaagatg aagaggaagg cagtgggtggc gcacaagggc agctcgggtg cggcaacggg      720
agcggtgagg gtaatggtgg gggcatgggg gatccagcaa catcaatgcc agtttatcaa      780
ttgccaaata tggtgacctaa tggacaattg aaccatgaag gatatgggtg ggctcacggc      840
agaccaccct attag

```

<210> 24

5 <211> 284

<212> PRT

<213> Populus tremuloides

<400> 24

```

Met Ala Asn Arg Trp Trp Thr Gly Gln Val Gly Leu Pro Gly Met Asp
1          5          10          15
Thr Ser Thr Ser Ser Ser Ser Pro Met Lys Lys Pro Asp Leu Gly Ile
          20          25          30
Ser Met Ser Asn Asn Asn Arg Glu Ala Thr Glu Ser Gly Ala Gly Lys
          35          40          45
Glu Asp Glu Gln Glu Asp Glu Arg Glu Asn Ser Asp Glu Pro Arg Glu
          50          55          60
Gly Ala Ile Asp Ile Ala Ser Arg Arg Pro Arg Gly Arg Pro Pro Gly

```

10

ES 2 440 265 T3

65					70					75					80
Ser	Lys	Asn	Lys	Pro	Lys	Pro	Pro	Ile	Phe	Val	Thr	Arg	Asp	Ser	Pro
				85					90					95	
Asn	Ala	Leu	Lys	Ser	His	Val	Met	Glu	Ile	Ala	Ser	Gly	Ser	Asp	Ile
			100					105						110	
Ala	Glu	Asn	Leu	Ala	Cys	Phe	Ala	Arg	Lys	Arg	Gln	Arg	Gly	Val	Cys
		115					120					125			
Val	Leu	Ser	Gly	Ser	Gly	Met	Val	Thr	Asn	Val	Thr	Leu	Lys	Gln	Pro
	130					135					140				
Ser	Ala	Ser	Gly	Ala	Val	Met	Ala	Leu	His	Gly	Arg	Phe	Glu	Ile	Leu
145					150					155					160
Ser	Leu	Thr	Gly	Ala	Phe	Leu	Pro	Gly	Pro	Ala	Pro	Pro	Gly	Ala	Thr
				165						170					175
Gly	Leu	Thr	Ile	Tyr	Leu	Ala	Gly	Gly	Gln	Gly	Gln	Val	Val	Gly	Gly
			180					185							190
Ser	Val	Val	Gly	Ser	Leu	Val	Ala	Ser	Gly	Pro	Val	Met	Val	Ile	Ala
		195					200					205			
Ala	Thr	Phe	Ser	Asn	Ala	Thr	Tyr	Glu	Arg	Leu	Pro	Leu	Glu	Asp	Glu
	210					215					220				
Glu	Glu	Gly	Ser	Gly	Gly	Ala	Gln	Gly	Gln	Leu	Gly	Gly	Gly	Asn	Gly
225					230						235				240
Ser	Gly	Glu	Gly	Asn	Gly	Gly	Gly	Met	Gly	Asp	Pro	Ala	Thr	Ser	Met
				245					250						255
Pro	Val	Tyr	Gln	Leu	Pro	Asn	Met	Val	Pro	Asn	Gly	Gln	Leu	Asn	His
			260					265							270
Glu	Gly	Tyr	Gly	Trp	Ala	His	Gly	Arg	Pro	Pro	Tyr				
		275					280								

<210> 25

<211> 885

5 <212> ADN

<213> Solanum tuberosum

<400> 25

atgtcaaacc	catggtggac	aggccaagta	ggtttacaag	gagttgaaac	atcatcatcc	60
gcgggttcgc	cttctctcaa	gaagccagat	ctaggcgtat	caatgaacga	tatagtgggt	120
ggtagtggta	gtcatgatga	agatagggac	catagcgcag	accctaaaga	gggtgcagtc	180
gaagtagcca	ctcgtcgacc	cagaggtcga	ccagctggct	caaagaacaa	acctaaacca	240
ccaatatttg	ttacaagggg	tagccctaac	gcacttagaa	gccacgtaat	ggaagttgct	300
aatggagctg	atgtggcgga	aagtatagct	caatttgcta	ggaaaagaca	aagaggtggt	360
tgtgttttga	gtgctactgg	aactgttact	aatgtaacc	taagacaacc	atctgctcct	420
ggagctgtca	tggcattaca	cggccggttc	gagatcttat	cgttgaccgg	agctttctta	480
cctggaccgg	cccctcctgg	atcaacaggg	ttgactatat	acctagcagg	aggacaagga	540
caagttgtgg	gaggaagtgt	agttaggtct	ttagtggctt	cgggaccagt	tatggtaatt	600
gcatcaactt	tttttaatgc	aacatgatgag	aggctacctt	tggaggagga	ggaagaaggc	660
ggtggaacgg	tggcccaagg	acaacttggg	ggtggtggat	cgccaccggg	aatgggagga	720
agtgggtgtg	gtggtggagg	acaacaacaa	caaggtggtg	gtggtatggg	tgatattcca	780
tcatcaaata	tgccagtata	taatttgcca	ccaaatttgc	taccaaattg	tggacaaatg	840
aaccatgaag	catttggttg	ggcacatgga	cgccctcctt	tttaa		885

10

<210> 26

<211> 294

<212> PRT

<213> Solanum tuberosum

ES 2 440 265 T3

<400> 26

```

Met Ser Asn Pro Trp Trp Thr Gly Gln Val Gly Leu Gln Gly Val Glu
1      5      10
Thr Ser Ser Ser Ala Gly Ser Pro Ser Leu Lys Lys Pro Asp Leu Gly

                20                25                30
Val Ser Met Asn Asp Ile Val Gly Gly Ser Gly Ser His Asp Glu Asp
35
Arg Asp His Ser Asp Asp Pro Lys Glu Gly Ala Val Glu Val Ala Thr
50
Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro
65      70      75      80
Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Arg Ser His Val
85      90      95
Met Glu Val Ala Asn Gly Ala Asp Val Ala Glu Ser Ile Ala Gln Phe
100
Ala Arg Lys Arg Gln Arg Gly Val Cys Val Leu Ser Ala Thr Gly Thr
115
Val Thr Asn Val Thr Leu Arg Gln Pro Ser Ala Pro Gly Ala Val Met
130
Ala Leu His Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly Ala Phe Leu
145      150      155      160
Pro Gly Pro Ala Pro Pro Gly Ser Thr Gly Leu Thr Ile Tyr Leu Ala
165      170      175
Gly Gly Gln Gly Gln Val Val Gly Gly Ser Val Val Gly Ser Leu Val
180      185      190
Ala Ser Gly Pro Val Met Val Ile Ala Ser Thr Phe Phe Asn Ala Thr
195      200      205
Tyr Glu Arg Leu Pro Leu Glu Glu Glu Glu Gly Gly Gly Thr Val
210      215      220
Ala Gln Gly Gln Leu Gly Gly Gly Ser Pro Pro Gly Met Gly Gly
225      230      235      240
Ser Gly Gly Gly Gly Gly Gly Gln Gln Gln Gln Gly Gly Gly Gly Met
245      250      255
Gly Asp Ile Pro Ser Ser Asn Met Pro Val Tyr Asn Leu Pro Pro Asn
260      265      270
Leu Leu Pro Asn Gly Gly Gln Met Asn His Glu Ala Phe Gly Trp Ala
275      280      285
His Gly Arg Pro Pro Phe
290

```

5 <210> 27

<211> 939

<212> ADN

<213> *Thlaspi caerulescens*

<400> 27

ES 2 440 265 T3

atggcgaatc	catggtggac	aggacaagtg	aatctctccg	gccttgaaac	gacgccgcct	60
ggttcctctc	agttaaagaa	atcagatctc	cacatctcca	tgaacatggc	catggactca	120
ggtcataaca	accatcatca	tcaccaagaa	gtcgcacaaca	ataacaacaa	cgatgacgac	180
agagataact	tgagcggcga	tgaacacgag	ccacgtgaag	gagccgtaga	agccccacg	240
cgccgtccac	gtggacgtcc	tgctggttcc	aagaacaaac	caaagccacc	gatctttgtc	300
acgcgcgatt	ctccaaacgc	tctcaagagc	catgtcatgg	agatcgctag	tgggactgac	360
gtcatcgaaa	ccctagctac	tttcgctagg	cggcgccaac	gtggcatctg	catcttgagc	420
ggcaacggca	cggtggtctaa	cgtaactctc	cgccaaccat	catctgccgc	agttgctgcg	480
gctcccgggg	gtgcgggcgt	tttggcttta	caagggaggt	ttgagattct	ctctttaaca	540
ggatcgttct	tgctgggacc	tgctccacct	ggatccaccg	gtttaaccat	ctacttagcc	600
ggtggtcaag	gtcaggtcgt	tggaggaagt	gtggtggggc	cattgatggc	ggctggtccg	660
gttatgttaa	tcgcggccac	gttttctaata	gcgacttacg	agagattgcc	tttgaggag	720
gaagaggcgg	ctgagagagg	cgggtggagga	ggcagcgtcc	caggacaact	cggaggggga	780
ggctcgccgc	tgagtagcgg	tgggtgtgga	ggggatggca	atcaaggact	tccggtgtac	840
aatatgcccg	gaaatcttgt	ttctaattgt	ggcggaggcg	gaggacagat	gagtgggcaa	900
gaagcttatg	gttgggctca	agctaggtca	ggattttaa			939

<210> 28

<211> 312

5 <212> PRT

<213> *Thlaspi caerulescens*

<400> 28

ES 2 440 265 T3

Met Ala Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Gly Leu Glu
 1 5 10 15
 Thr Thr Pro Pro Gly Ser Ser Gln Leu Lys Lys Ser Asp Leu His Ile
 20 25 30
 Ser Met Asn Met Ala Met Asp Ser Gly His Asn Asn His His His His
 35 40 45
 Gln Glu Val Asp Asn Asn Asn Asn Asn Asp Asp Asp Arg Asp Asn Leu
 50 55 60
 Ser Gly Asp Glu His Glu Pro Arg Glu Gly Ala Val Glu Ala Pro Thr
 65 70 75 80
 Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro
 85 90 95
 Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His Val
 100 105 110
 Met Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr Phe
 115 120 125
 Ala Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly Thr
 130 135 140
 Val Ala Asn Val Thr Leu Arg Gln Pro Ser Ser Ala Ala Val Ala Ala
 145 150 155 160
 Ala Pro Gly Gly Ala Ala Val Leu Ala Leu Gln Gly Arg Phe Glu Ile
 165 170 175
 Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly Ser
 180 185 190
 Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly
 195 200 205
 Gly Ser Val Val Gly Pro Leu Met Ala Ala Gly Pro Val Met Leu Ile
 210 215 220
 Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu Glu
 225 230 235 240
 Glu Glu Ala Ala Glu Arg Gly Gly Gly Gly Gly Ser Val Pro Gly Gln
 245 250 255
 Leu Gly Gly Gly Gly Ser Pro Leu Ser Ser Gly Gly Gly Gly Gly Asp
 260 265 270
 Gly Asn Gln Gly Leu Pro Val Tyr Asn Met Pro Gly Asn Leu Val Ser
 275 280 285
 Asn Gly Gly Gly Gly Gly Gly Gln Met Ser Gly Gln Glu Ala Tyr Gly
 290 295 300
 Trp Ala Gln Ala Arg Ser Gly Phe
 305 310

<210> 29

<211> 876

5 <212> ADN

<213> Vitis vinifera

<400> 29

atggcgaacc ggtggtgggc tgggcaggtg ggtctgcaag gtgtagatac ctcatcagct 60
 tcacctgcaa tgaagaaacc agatctggga atatccatga atgaaaatgg aggaagcggg 120
 agcggaggcg gaggagagga agaagaggaa aaagaaaaca gtgatgagcc cagagagggg 180
 gcaattgagg tggctacgcg caggcctagg ggccggccgc ctggctccaa gaacaagcca 240
 aaacctccga tttttgtgac aagggacagc cctaacgctc tgcgcagcca cgttatggag 300
 gtggcaaacg gctccgacat cacagaaagc atagcccaat tcgcgagaag gcggcaacga 360

ES 2 440 265 T3

```

ggcgtctgcg tgctcagcgc aagtgggaca gtcatagaacg taacgcttcg ccagccttct 420
gcccttggtg gtgcagttat ggcacttcat ggccgattcg aaattctttc cttaccgggc 480
gcgttctctac cgggaccagc gccaccaggc tccactggac taaccatata cctagcaggc 540
ggtcaagctc aggtcgtggg tggtagcgtg gtgggttcac tcatagcggc aggtccagtt 600
atggtgattg cagctacctt ttcgaatgca acctacgaga ggctccccct agaagacgaa 660
gaagaggcgg gcagcgcagc acaggagcag ctgcgtggcg gcggaggcgg tgggtgggtca 720
ccgccagga ttggcggcag tggggggcag cagcaggcag ggatggcaga tccttctcc 780
atgccggttt ataatttgcc accaaatttg cttccaaatg gtggacaact gaacatgat 840
gcttatggtt gggcacatgg gcgccagcct tactag 876

```

<210> 30

<211> 291

5 <212> PRT

<213> Vitis vinifera

<400> 30

```

Met Ala Asn Arg Trp Trp Ala Gly Gln Val Gly Leu Gln Gly Val Asp
1          5          10          15
Thr Ser Ser Ala Ser Pro Ala Met Lys Lys Pro Asp Leu Gly Ile Ser
20          25          30
Met Asn Glu Asn Gly Gly Ser Gly Ser Gly Gly Gly Gly Glu Glu Glu
35          40          45
Glu Glu Lys Glu Asn Ser Asp Glu Pro Arg Glu Gly Ala Ile Glu Val
50          55          60
Ala Thr Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro
65          70          75          80
Lys Pro Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Arg Ser
85          90          95
His Val Met Glu Val Ala Asn Gly Ser Asp Ile Thr Glu Ser Ile Ala
100          105          110
Gln Phe Ala Arg Arg Arg Gln Arg Gly Val Cys Val Leu Ser Ala Ser
115          120          125
Gly Thr Val Met Asn Val Thr Leu Arg Gln Pro Ser Ala Pro Gly Gly
130          135          140
Ala Val Met Ala Leu His Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly
145          150          155          160
Ala Phe Leu Pro Gly Pro Ala Pro Pro Gly Ser Thr Gly Leu Thr Ile
165          170          175
Tyr Leu Ala Gly Gly Gln Ala Gln Val Val Gly Gly Ser Val Val Gly
180          185          190
Ser Leu Ile Ala Ala Gly Pro Val Met Val Ile Ala Ala Thr Phe Ser
195          200          205
Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu Asp Glu Glu Glu Ala Gly
210          215          220
Ser Ala Ala Gln Glu Gln Leu Ala Gly Gly Gly Gly Gly Gly Ser
225          230          235          240
Pro Pro Gly Ile Gly Gly Ser Gly Gly Gln Gln Ala Gly Met Ala
245          250          255
Asp Pro Ser Ser Met Pro Val Tyr Asn Leu Pro Pro Asn Leu Leu Pro
260          265          270
Asn Gly Gly Gln Leu Asn His Asp Ala Tyr Gly Trp Ala His Gly Arg
275          280          285
Gln Pro Tyr
290

```

ES 2 440 265 T3

<210> 31

<211> 783

<212> ADN

5 <213> Vitis vinifera

<400> 31

```

atggaccgag cagctgtttc gccgatgcta aataaacgag atcgcgagat atcaatcaac    60
gataaccccc gcacaggaga cgatgaagaa gagaaagaca acgaaggcga gcccacggag    120
ggtgcagtag aagtcggcac tcgtagacca agaggtcgcc cgcctggatc caaaaacaag    180
cccaaaccac ctattttcgt cacgcgcgac agcccgaacg cccttcggag ccacgtgatg    240
gaggtggccg gcggccacga cgttgccgaa agcgtcgccc agttcgcccg taggcgtcaa    300
cgaggggtct gcgtcctcag cggcagcggc tccgtagcca acgtgactct gagacagccc    360
gccgcgcctg gcgcccgtgt ggcactccat ggaagattcg agattctgtc cctaacagga    420
gcattcctcc ccggacctgc ccctcccggc tccactggac tcaccgtgta cctcgccgga    480
ggtcagggcc aggttggtggg aggaagtgtg gttggatcac tggtagcggc aggcccggtg    540
atagtgatag ccgccacttt tgcgaacgca acatacgaaa gactgcctct ggaagaagaa    600
gaagaagggt ggcaggcgcc gccgccgagt ggttcgcccg ctgcaattgg aagcagtggg    660
ggacagcadc actctggcct gccggagctg cccatataca atctgccacc gaacctactc    720
cctaacggcg gcccaattgag tcatgacccc tactcatggg ctcatgctcg gcccccttac    780
tga

```

10 <210> 32

<211> 260

<212> PRT

<213> Vitis vinifera

<400> 32

ES 2 440 265 T3

Met	Asp	Pro	Ala	Ala	Val	Ser	Pro	Met	Leu	Asn	Lys	Arg	Asp	Arg	Glu
1				5					10					15	
Ile	Ser	Ile	Asn	Asp	Asn	Pro	Gly	Thr	Gly	Asp	Asp	Glu	Glu	Glu	Lys
			20					25					30		
Asp	Asn	Glu	Gly	Glu	Pro	Thr	Glu	Gly	Ala	Val	Glu	Val	Gly	Thr	Arg
		35					40					45			
Arg	Pro	Arg	Gly	Arg	Pro	Pro	Gly	Ser	Lys	Asn	Lys	Pro	Lys	Pro	Pro
	50					55					60				
Ile	Phe	Val	Thr	Arg	Asp	Ser	Pro	Asn	Ala	Leu	Arg	Ser	His	Val	Met
65				70						75					80
Glu	Val	Ala	Gly	Gly	His	Asp	Val	Ala	Glu	Ser	Val	Ala	Gln	Phe	Ala
				85					90					95	
Arg	Arg	Arg	Gln	Arg	Gly	Val	Cys	Val	Leu	Ser	Gly	Ser	Gly	Ser	Val
			100				105						110		
Ala	Asn	Val	Thr	Leu	Arg	Gln	Pro	Ala	Ala	Pro	Gly	Ala	Val	Val	Ala
		115					120					125			
Leu	His	Gly	Arg	Phe	Glu	Ile	Leu	Ser	Leu	Thr	Gly	Ala	Phe	Leu	Pro
	130					135					140				
Gly	Pro	Ala	Pro	Pro	Gly	Ser	Thr	Gly	Leu	Thr	Val	Tyr	Leu	Ala	Gly
145					150						155				160
Gly	Gln	Gly	Gln	Val	Val	Gly	Gly	Ser	Val	Val	Gly	Ser	Leu	Val	Ala
				165					170					175	
Ala	Gly	Pro	Val	Ile	Val	Ile	Ala	Ala	Thr	Phe	Ala	Asn	Ala	Thr	Tyr
			180					185					190		
Glu	Arg	Leu	Pro	Leu	Glu	Glu	Glu	Glu	Glu	Gly	Gly	Gln	Ala	Pro	Pro
		195					200					205			
Pro	Ser	Gly	Ser	Pro	Pro	Ala	Ile	Gly	Ser	Ser	Gly	Gly	Gln	His	His
	210					215					220				
Ser	Gly	Leu	Pro	Glu	Leu	Pro	Ile	Tyr	Asn	Leu	Pro	Pro	Asn	Leu	Leu
225					230					235					240
Pro	Asn	Gly	Gly	Gln	Leu	Ser	His	Asp	Pro	Tyr	Ser	Trp	Ala	His	Ala
				245					250					255	
Arg	Pro	Pro	Tyr												
			260												

<210> 33

<211> 810

5 <212> ADN

<213> Zea mays

<400> 33

ES 2 440 265 T3

```
atggcacctt cctccaagga cggcgccacc gccaccgagc agccgacgag cggcgacgac 60
gaccgggaga acggcggcac gggcgagccc aaggaaggcg cggtggtggc gggcaaccgg 120
cgccccgcg ggcggccgcc ggggtccaag aacaagccca agccgcccac cttcgtgacg 180
cgcgacagcc ccaacgcgct gcgcagccac gtgatggagg tggccggcgg cgccgacgtg 240
gccgagtcca tcgcccactt cgcgcgccgc aggcagcgcg gcgtgtgcgt gctcagcggc 300
gcgggcaccg tcgccgacgt ggcgctccgc cagcccgcgg ctccggggcg cgtggtcgcc 360
ctccgcggcc gcttcgagat cctctcgctc accggcacgt tcctgccggg ccccgcgccg 420
ccgggctcca cggggctcac cgtgtacctc gcgggcggcc aggggcaggt cgtcggcggc 480
agcgtcgtcg gcacgctcac cgcggcgggg cccgtcatgg tgatggcgtc cacgttcgcc 540
aacgccacct acgagaggct gccgctggac gacgccgacg aggagcccgc cgggcagcag 600
gcggcgcagc tgcctcccgg accgggcgga gggcagccta tggtaatggg cgggatggcc 660
gaccctcag cggtgccaat gttcggcggc gccggcggtg tgccgccaag cctcatgcca 720
gcaggggccc cagccgcctc ctccggtgcg ggctgcagc tcgggcacga ccgacttgca 780
tgggctcatg cacggccacc gccatactag 810
```

<210> 34

<211> 269

5 <212> PRT

<213> Zea mays

<400> 34

ES 2 440 265 T3

Met	Ala	Pro	Ser	Ser	Lys	Asp	Gly	Ala	Thr	Ala	Thr	Glu	Gln	Pro	Thr
1				5					10					15	
Ser	Gly	Asp	Asp	Asp	Arg	Glu	Asn	Gly	Gly	Thr	Gly	Glu	Pro	Lys	Glu
			20					25					30		
Gly	Ala	Val	Val	Ala	Gly	Asn	Arg	Arg	Pro	Arg	Gly	Arg	Pro	Pro	Gly
		35				40						45			
Ser	Lys	Asn	Lys	Pro	Lys	Pro	Pro	Ile	Phe	Val	Thr	Arg	Asp	Ser	Pro
	50					55					60				
Asn	Ala	Leu	Arg	Ser	His	Val	Met	Glu	Val	Ala	Gly	Gly	Ala	Asp	Val
65					70					75					80
Ala	Glu	Ser	Ile	Ala	His	Phe	Ala	Arg	Arg	Arg	Gln	Arg	Gly	Val	Cys
				85					90					95	
Val	Leu	Ser	Gly	Ala	Gly	Thr	Val	Ala	Asp	Val	Ala	Leu	Arg	Gln	Pro
			100					105					110		
Ala	Ala	Pro	Gly	Ala	Val	Val	Ala	Leu	Arg	Gly	Arg	Phe	Glu	Ile	Leu
		115					120					125			
Ser	Leu	Thr	Gly	Thr	Phe	Leu	Pro	Gly	Pro	Ala	Pro	Pro	Gly	Ser	Thr
	130					135					140				
Gly	Leu	Thr	Val	Tyr	Leu	Ala	Gly	Gly	Gln	Gly	Gln	Val	Val	Gly	Gly
145					150					155					160
Ser	Val	Val	Gly	Thr	Leu	Thr	Ala	Ala	Gly	Pro	Val	Met	Val	Met	Ala
				165					170					175	
Ser	Thr	Phe	Ala	Asn	Ala	Thr	Tyr	Glu	Arg	Leu	Pro	Leu	Asp	Asp	Ala
			180					185					190		
Asp	Glu	Glu	Pro	Ala	Gly	Gln	Gln	Ala	Ala	Gln	Leu	Pro	Pro	Gly	Pro
		195					200					205			
Gly	Gly	Gly	Gln	Pro	Met	Val	Met	Gly	Gly	Met	Ala	Asp	Pro	Ser	Ala
	210					215					220				
Val	Pro	Met	Phe	Gly	Gly	Ala	Gly	Gly	Val	Pro	Pro	Ser	Leu	Met	Pro
225						230				235					240
Ala	Gly	Ala	Ala	Ala	Ala	Ser	Ser	Gly	Ala	Gly	Leu	Gln	Leu	Gly	His
				245					250					255	
Asp	Arg	Leu	Ala	Trp	Ala	His	Ala	Arg	Pro	Pro	Pro	Tyr			

260

265

<210> 35

<211> 2194

5 <212> ADN

<213> Oryza sativa

<400> 35

ES 2 440 265 T3

```

aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct      60
aaatataaaa tgagacctta tataatgtagc gctgataact agaactatgc aagaaaaact     120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt     180
tttcccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc     240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaacctc ttcttgaata     300
aaaaaatcct tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga     360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt     420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat     480
ttagtaatta aagacaattg acttattttt attatttatc ttttttcgat tagatgcaag     540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctcct caatacacgt     600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaaatc     660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaaaaaat     720
aattttacag aatgatgata aaagtatgaa acgaactatt taggtttttc acatacaaaa     780
aaaaaaaagaa ttttgctcgt gcgcgagcgc caatctocca tattgggcac acaggcaaca     840
acagagtggc tgcccacaga acaaccaca aaaaacgatg atctaacgga ggacagcaag     900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa     960
aaccaagcat cctccttctc ccatctataa attcctcccc ccttttcccc tctctatata    1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag    1080
cgaccgcctt ctcgatccat atcttccggt cgagttcttg gtcgatctct tccctcctcc    1140
acctcctcct cacagggtat gtgcctcctt tcggttgttc ttggatttat tgttctaggt    1200
tgtgtagtac gggcgttgat gttaggaaag gggatctgta tctgtgatga ttctgttct    1260
tggatttggg atagaggggt tcttgatggt gcatgttatc ggttcggttt gattagtagt    1320
atggttttca atcgtctgga gagctctatg gaaatgaaat ggtttaggga tcggaatcct    1380
gcgattttgt gagtaccttt tgtttgaggt aaaatcagag caccggtgat tttgcttggg    1440
gtaataaagt acggttgttt ggtcctcgat tctggtagtg atgcttctcg atttgacgaa    1500
gctatccttt gtttattccc tattgaacaa aaataatcca actttgaaga cggccccggt    1560
gatgagattg aatgattgat tcttaagcct gtccaaaatt tcgcagctgg cttgtttaga    1620
tacagtagtc cccatcacga aattcatgga aacagttata atcctcagga acaggggatt    1680
cctgttctt ccgatttgct ttagtcccag aatttttttt cccaaatata ttaaaaagtc    1740
actttctggt tcagttcaat gaattgattg ctacaaataa tgcttttata gcgttatcct    1800
agctgtagtt cagttaatag gtaatacccc tatagtttag tcaggagaag aacttatccg    1860
atctctgata tccattttta attatatgaa atgaactgta gcataagcag tattcatttg    1920
gattattttt tttattagct ctcaccctt cattattctg agctgaaagt ctggcatgaa    1980
ctgtcctcaa ttttgttttc aaattcacat cgattatcta tgcattatcc tcttgtatct    2040
acctgtagaa gtttcttttt ggttattcct tgactgcttg attacagaaa gaaatttatg    2100
aagctgtaat cgggatagtt atactgcttg ttcttatgat tcatttcctt tgtgcagttc    2160
ttggtgtagc ttgccacttt caccagcaaa gttc

```

<210> 36

<211> 173

5 <212> PRT

<213> Secuencia artificial

<220>

<223> dominio conservado comprendido en la SEQ ID NO: 2

<400> 36

10

```

Glu Pro Arg Glu Gly Ala Val Glu Ala Pro Thr Arg Arg Pro Arg Gly
1           5           10           15
Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Pro Pro Ile Phe Val Thr
                20                25                30

```

ES 2 440 265 T3

```

Arg Asp Ser Pro Asn Ala Leu Lys Ser His Val Met Glu Ile Ala Ser
   35                               40                               45
Gly Thr Asp Val Ile Glu Thr Leu Ala Thr Phe Ala Arg Arg Arg Gln
   50                               55                               60
Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly Thr Val Ala Asn Val Thr
65                               70                               75                               80
Leu Arg Gln Pro Ser Thr Ala Ala Val Ala Ala Ala Pro Gly Gly Ala
   85                               90                               95
Ala Val Leu Ala Leu Gln Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly
   100                              105                              110
Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly Ser Thr Gly Leu Thr Ile
   115                              120                              125
Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Ser Val Val Gly
   130                              135                              140
Pro Leu Met Ala Ala Gly Pro Val Met Leu Ile Ala Ala Thr Phe Ser
145                              150                              155                              160
Asn Ala Thr Tyr Glu Arg Leu Pro Leu Glu Glu Glu Glu
   165                              170

```

<210> 37

<211> 14

5 <212> PRT

<213> Secuencia artificial

<220>

<223> AT hook

<220>

10 <221> VARIANTE

<222> (8) .. (8)

<223> /reemplazo = "Ala"

<220>

<221> VARIANTE

15 <222> (11) .. (11)

<223> /reemplazo = "Arg"

<400> 37

```

Arg Arg Pro Arg Gly Arg Pro Pro Gly Ser Lys Asn Lys Pro
1                               5                               10

```

20 <210> 38

<211> 125

<212> PRT

<213> Secuencia artificial

ES 2 440 265 T3

<220>

<223> dominio de PPC (DUF296) comprendido en la SEQ ID NO: 2

<400> 38

```

Leu Lys Ser His Val Met Glu Ile Ala Ser Gly Thr Asp Val Ile Glu
1      5      10      15
Thr Leu Ala Thr Phe Ala Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu
20      25      30
Ser Gly Asn Gly Thr Val Ala Asn Val Thr Leu Arg Gln Pro Ser Thr
35      40      45
Ala Ala Val Ala Ala Ala Pro Gly Gly Ala Ala Val Leu Ala Leu Gln
50      55      60
Gly Arg Phe Glu Ile Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro

65      70      75      80
Ala Pro Pro Gly Ser Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln
85      90      95
Gly Gln Val Val Gly Gly Ser Val Val Gly Pro Leu Met Ala Ala Gly
100     105     110
Pro Val Met Leu Ile Ala Ala Thr Phe Ser Asn Ala Thr
115     120     125

```

5

<210> 39

<211> 52

<212> ADN

<213> Secuencia artificial

10

<220>

<223> cebador: prm8135

<400> 39

ggggacaagt ttgtacaaaa aagcaggctt aaacaatggc gaatccatgg tg 52

<210> 40

15

<211> 50

<212> DNA

<213> Secuencia artificial

<220>

<223> cebador: prm8136

20

<400> 40

ggggaccact ttgtacaaga aagctgggtt aaaaccatt ttaacgcacg 50

<210> 41

<211> 948

ES 2 440 265 T3

<212> ADN

<213> Brassica oleracea

<400> 41

```

atgcgaaatc catggtggac aggacaagtg aatctctcca gtctcgaaac gacgccgccg      60
agttcctctc agttaaagac accagatctc cacatctcca tgaacatggc catggtctca      120
ggtcataaca accacatca tcatcaccaa gaagtcaaca ccaacaaca caacgaagac      180
gatagagaca acttgagcgg cgacgaccgc gagccacgtg aaggagccgt ggaagctccc      240
acgcgccgac cacgtggacg tcctgctggt tccaagaaca aaccaaagcc accaatcttt      300
gtcacgcgtg attctccaaa cgctctcaag agccatgtca tggagatcgc tagtgggact      360
gatgtcatag aaaccctagc tactttcgtc agggcggcgc aacgtggcat ctgcatcttg      420
agcggtaacg gcacggtggc taacgtcaca ctccgtcaac catcagtggc tcccgttgca      480
gctgccctg gtgggtgcggc tgtattggcg ttacaaggga ggtttgagat tctttctcta      540
accggttctt tcttacctgg accggctcca cctggatcca ctggtttaac tatttactta      600
gctggtggtc aaggtcaggt tgttgaggga agcgtgggtg gggcattgat ggctgctggt      660
ccggtgatgc taatcgctgc cacgttttct aatgcgactt atgagagatt acctttggat      720

gaggaagaag cggctgaaag aggtggcggg ggaagcgacg gaggagtggg tccagggcag      780
ctcgggggcg taggttcccc gctgagtagt ggtggcggtg gaggccacgg gaaccaagga      840
cttcccgcac ataatatgcc cggaaacctt gcttctaata gcggtggagg aggacagatg      900
agcagccaag aagcgtacgg ttgggctcaa gctaggtcag gattttaa      948

```

5

<210> 42

<211> 315

<212> PRT

10 <213> Brassica oleracea

<400> 42

ES 2 440 265 T3

Met Arg Asn Pro Trp Trp Thr Gly Gln Val Asn Leu Ser Ser Leu Glu
1 5 10 15
Thr Thr Pro Pro Ser Ser Ser Gln Leu Lys Thr Pro Asp Leu His Ile
20 25 30
Ser Met Asn Met Ala Met Val Ser Gly His Asn Asn His His His His
35 40 45
His Gln Glu Val Asn Thr Asn Asn Asn Asn Glu Asp Asp Arg Asp Asn
50 55 60
Leu Ser Gly Asp Asp Arg Glu Pro Arg Glu Gly Ala Val Glu Ala Pro
65 70 75 80
Thr Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys
85 90 95
Pro Pro Ile Phe Val Thr Arg Asp Ser Pro Asn Ala Leu Lys Ser His
100 105 110
Val Met Glu Ile Ala Ser Gly Thr Asp Val Ile Glu Thr Leu Ala Thr
115 120 125
Phe Ala Arg Arg Arg Gln Arg Gly Ile Cys Ile Leu Ser Gly Asn Gly
130 135 140
Thr Val Ala Asn Val Thr Leu Arg Gln Pro Ser Val Ala Pro Val Ala
145 150 155 160
Ala Ala Pro Gly Gly Ala Ala Val Leu Ala Leu Gln Gly Arg Phe Glu
165 170 175
Ile Leu Ser Leu Thr Gly Ser Phe Leu Pro Gly Pro Ala Pro Pro Gly
180 185 190
Ser Thr Gly Leu Thr Ile Tyr Leu Ala Gly Gly Gln Gly Gln Val Val
195 200 205
Gly Gly Ser Val Val Gly Ala Leu Met Ala Ala Gly Pro Val Met Leu
210 215 220
Ile Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Leu Asp
225 230 235 240
Glu Glu Glu Ala Ala Glu Arg Gly Gly Gly Gly Ser Asp Gly Gly Val
245 250 255
Val Pro Gly Gln Leu Gly Gly Val Gly Ser Pro Leu Ser Ser Gly Gly
260 265 270
Gly Gly Gly His Gly Asn Gln Gly Leu Pro Ala Tyr Asn Met Pro Gly
275 280 285
Asn Leu Ala Ser Asn Gly Gly Gly Gly Gly Gln Met Ser Ser Gln Glu
290 295 300
Ala Tyr Gly Trp Ala Gln Ala Arg Ser Gly Phe
305 310 315

ES 2 440 265 T3

<210> 43

<211> 918

<212> ADN

<213> Medicago truncatula

5 <400> 43

```

atggcgaaca ggtggtggac cggaccggtt ggtctaggag ggatggacaa ctcagtaacc      60
tcctctccac taggaaaacc ggatctgggt ttctccatga atcaaagtgc tgtaacagga      120
gtgaacaaca tgaacaacaa caacaatgaa gaagaagaag atgagaaaga aacagcgcac      180
gaacacaaag gaggtgcaat agaaacaaac acctccacgc gccgcccaag aggccgtcca      240
tcaggttcaa aaaacaaacc aaaaccacca atattcataa caagagatag ccctaacgcg      300
ctacgaagcc atgtcatgga agtagcaaca ggaacagata tatcagatag catcgttcag      360
tttgcaagaa aaagacagag aggtatattgc attctaagcg caagtggaac cgtcgttaac      420
gtttctctcc ggcaacctac aggtcccgga gctgtggtag cgcttccagg gagatttgat      480
atactctctt tgactggttc tgtgcttctt ggaccttcac cgccgggagc tactggtttg      540
actatattatc tttctggagg acaaggacag gtggttggcg gcggagttgt tggteccctt      600
gtggcggcag gaccagttat gttgatggcg gcgacathtt cgaatgctac gtatgagagg      660
ctgccggttg aggatggtga tgatcaagaa gggcatcagg gtggtggtgg tgatgatgag      720
tctccgacgc gtgcagcggg gatgggacag ttagcgattg gatctggttg agaaggttct      780
tcaattccac caggctataa caatgttggt ggtaatttgg gtgtttcaaa tggaggacaa      840
caacaattgt tgaataatca tgaggcttat aataattctc cttgggggtca tgctagtcat      900
ggtagaccac catactaa

```

<210> 4

10 <211> 305

<212> PRT

<213> Medicago truncatula

<400> 44

ES 2 440 265 T3

Met Ala Asn Arg Trp Trp Thr Gly Pro Val Gly Leu Gly Gly Met Asp
1 5 10 15
Asn Ser Val Thr Ser Ser Pro Leu Gly Lys Pro Asp Leu Gly Phe Ser
20 25 30
Met Asn Gln Ser Ala Val Thr Gly Val Asn Asn Met Asn Asn Asn Asn
35 40 45
Asn Glu Glu Glu Glu Asp Glu Lys Glu Asn Ser Asp Glu His Lys Gly
50 55 60
Gly Ala Ile Glu Thr Asn Thr Ser Thr Arg Arg Pro Arg Gly Arg Pro
65 70 75 80
Ser Gly Ser Lys Asn Lys Pro Lys Pro Pro Ile Phe Ile Thr Arg Asp
85 90 95
Ser Pro Asn Ala Leu Arg Ser His Val Met Glu Val Ala Thr Gly Thr
100 105 110
Asp Ile Ser Asp Ser Ile Val Gln Phe Ala Arg Lys Arg Gln Arg Gly
115 120 125
Ile Cys Ile Leu Ser Ala Ser Gly Thr Val Val Asn Val Ser Leu Arg
130 135 140
Gln Pro Thr Gly Pro Gly Ala Val Val Ala Leu Pro Gly Arg Phe Asp
145 150 155 160
Ile Leu Ser Leu Thr Gly Ser Val Leu Pro Gly Pro Ser Pro Pro Gly
165 170 175
Ala Thr Gly Leu Thr Ile Tyr Leu Ser Gly Gly Gln Gly Gln Val Val
180 185 190
Gly Gly Gly Val Val Gly Pro Leu Val Ala Ala Gly Pro Val Met Leu
195 200 205
Met Ala Ala Thr Phe Ser Asn Ala Thr Tyr Glu Arg Leu Pro Val Glu
210 215 220
Asp Gly Asp Asp Gln Glu Gly His Gln Gly Gly Gly Gly Asp Asp Glu
225 230 235 240
Ser Pro Thr Arg Ala Ala Gly Met Gly Gln Leu Ala Ile Gly Ser Val

ES 2 440 265 T3

```

                245                250                255
Gly Glu Gly Ser Ser Ile Pro Pro Gly Tyr Asn Asn Val Gly Gly Asn
                260                265                270
Leu Gly Val Ser Asn Gly Gly Gln Gln Gln Leu Leu Asn Asn His Glu
                275                280                285
Ala Tyr Asn Asn Ser Pro Trp Gly His Ala Ser His Gly Arg Pro Pro
                290                295                300
Tyr
305

```

<210> 45

<211> 632

5 <212> ADN

<213> Arabidopsis thaliana

<400> 45

```

gcagttccct actctcgcgt taacgctagc atggatctcg ggccccaat aatgatttta      60
tttgactga tagtgacctg ttcggtgcaa caaattgatg agcaatgctt ttttataatg     120
ccaactttgt acaaaaaagc aggcttcaca atgtcttgct gtggaggaaa ctgcggatgt     180
ggatctggct gcaagtgcgg caacggttgt ggaggttgca aaatgtaccg tgacttgagg     240
ttctccggcg agacaaccac aactgagact tttgtcttgg gcggtgcacc ggcgatgaag     300
aatcagtagc aggcttcagg ggagagtaac aacgctgaga acgatgcttg caagtgtgga     360
tctgactgca agtgtgatcc ttgcacctgc aagtgaaacc cagctttctt gtacaaagtt     420
ggcattataa gaaagcattg cttatcaatt tgttgcaacg aacaggtcac tatcagtcaa     480
aataaaatca ttatttgcca tccagctgca gctctggccc gtgtctcaaa atctctgatg     540
ttacattgca caagataaaa atatatcatc atgaacaata aaactgtctg cttacataaa     600
cagtaataca aggggtggtt tgagccatat tc                                     632

```

10

<210> 46

<211> 81

<212> PRT

<213> Arabidopsis thaliana

15 <400> 46

```

Met Ser Cys Cys Gly Gly Asn Cys Gly Cys Gly Ser Gly Cys Lys Cys
1          5          10          15
Gly Asn Gly Cys Gly Gly Cys Lys Met Tyr Pro Asp Leu Gly Phe Ser
                20                25                30
Gly Glu Thr Thr Thr Glu Thr Phe Val Leu Gly Val Ala Pro Ala
                35                40                45
Met Lys Asn Gln Tyr Glu Ala Ser Gly Glu Ser Asn Asn Ala Glu Asn
                50                55                60
Asp Ala Cys Lys Cys Gly Ser Asp Cys Lys Cys Asp Pro Cys Thr Cys
65          70          75          80
Lys

```

- <210> 47
- <211> 2194
- <212> ADN
- <213> Oryza sativa

5 <400> 47

aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60

```

aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact 120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtctcgt 180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc 240
tctgtcatga agttaaatta ttcgaggtag ccataattgt catcaaactc ttcttgaata 300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga 360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480
ttagtaatta aagacaattg acttattttt attatttatc ttttttcgat tagatgcaag 540
gtacttacgc acacactttg tgctcatgtg catgtgtgag tgcacctct caatacacgt 600
tcaactagca acacatctct aatatcactc gcctatttaa tacatttagg tagcaatc 660
tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa taaaaaaat 720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa 780
aaaaaaagaa ttttgctcgt gcgagagcgc caatctcca tattgggcac acaggcaaca 840
acagagtggc tgccacaga acaaccaca aaaaacgatg atctaacgga ggacagcaag 900
tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaagaa 960
aaccaagcat cctccttctc ccatctataa attcctcccc ccttttcccc tctctatata 1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
cgaccgcctt ctcgatccat atcttccggt cgagttcttg gtcgatctct tccctcctcc 1140
acctcctcct cacaggggat gtgcctccct tcggttggtc ttggatttat gttctaggt 1200
tgtgtagtac gggcgttgat gttaggaaag gggatctgta tctgtgatga ttctgttct 1260
tggatttggg atagaggggt tcttgatgtt gcatgttatc ggttcgggtt gatttagtagt 1320
atggttttca atcgtctgga gagctctatg gaaatgaaat ggtttaggga tcggaatctt 1380
gcgattttgt gagtacctt tgtttgaggt aaaatcagag caccggtgat tttgcttgg 1440
gtaataaagt acggttggtt ggtcctcgat tctggtagtg atgcttctcg atttgacgaa 1500
gctatccttt gtttattccc tattgaacaa aaataatcca actttgaaga cggcctcgtt 1560
gatgagattg aatgattgat tcttaagcct gtccaaaatt tcgcagctgg ctgttttaga 1620
tacagtagtc cccatcacga aattcatgga aacagttata atcctcagga acaggggatt 1680
cctgttctt cggatttgct ttagtcccag aattttttt cccaaatc ttaaaaagtc 1740
actttctggt tcagttcaat gaattgattg ctacaaataa tgcttttata gcgttatcct 1800
agctgtagtt cagttaatag gtaatcccc tatagtttag tcaggagaag aacttatccg 1860
atctctgac tccattttta attatagaa atgaactgta gcataagcag tattcatttg 1920
gattattttt tttattagct ctcaccctt cattattctg agctgaaagt ctggcatgaa 1980
ctgtcctcaa ttttgttttc aaattccat cgattatcta tgcattatcc tcttgtatct 2040
acctgtagaa gtttcttttt ggttattcct tgactgcttg attacagaaa gaaatttatg 2100
aagctgtaat cgggatagtt atactgctg ttcttatgat tcatttcctt tgtgcagttc 2160
ttggtgtagc ttgccacttt caccagcaaa gttc
    
```

- <210> 48
- <211> 53
- 10 <212> ADN
- <213> Secuencia artificial
- <220>
- <223> cebador: prm03240

ES 2 440 265 T3

<400> 48

ggggacaagt ttgtacaaaa aagcaggctt cacaatgtct tgctgtggag gaa 53

<210> 49

<211> 47

5 <212> ADN

<213> Secuencia artificial

<220>

<223> cebador: prm03241

<400> 49

10 ggggaccact ttgtacaaga aagctgggtt tcactgacg gtgcaag 47

<210> 50

<211> 1566

<212> ADN

<213> Chlamydomonas reinhardtii

15 <400> 50

atg	cggaagg	aagc	gactcg	tctt	gtgtcc	gcct	gctgc	gggc	gggcaa	caat	ggcgtg	60
tct	acgtcgt	gggc	tgttg	tggc	actcgc	ctca	agtcgg	cgat	gcccc	gcct	gatgag	120
aaga	aggacg	aggac	ctgca	tgcc	aaggag	ggca	agggtgc	tgcac	ccctca	cctt	ctgaac	180
gaga	acgtgg	tgaag	actca	gtat	gcogtc	cgtg	gcgagc	tttac	ctgcg	cgct	gagcag	240
ctcc	gcaagg	aggg	caagga	gatc	attttc	acaaa	cgctcg	gaaac	cccgca	cgcg	ctgggt	300
gcc	aagcccc	tgac	cttcac	ccgt	cagggtg	ctag	ccctgt	gcg	ccgcgc	cttc	ctgctg	360
gatc	accccc	aggt	ggagga	catg	ttcccc	gccg	acgcca	tcgc	gcgtgc	caag	aagatc	420
ctag	ctcct	tcaag	ggcgg	tgtg	ggcgcc	tacac	cgact	cgcg	tggcaa	cccg	ctggtg	480
cgcg	aggagg	tggc	ccgctt	catc	gagaag	cgtg	acggcg	ttcc	ctgaa	cccc	gaccac	540
atct	ctctga	cggac	ggcgc	ctcg	gtggcc	gtgc	gcttgt	gcct	gaacgc	catg	atccgc	600
cacg	accgcg	actc	cgctg	ggtg	ccccatc	ccgc	agttacc	cgct	gtacag	cgct	ccatc	660
cgct	gtacg	gcgg	cacgct	ggtg	ggctac	ttct	ggatg	agcg	ccgcgg	ctgg	ggcctg	720
tccg	tggagg	agct	gcagcg	cgcg	ctgcag	gagg	gcgcgc	agg	agggcaa	gctg	gtgcgc	780
ggc	ctggtgt	ttat	caaccc	cggt	aacccc	accg	gccagt	gctt	gagcaa	ggag	aacctg	840
cagg	agctga	tcaag	tttgc	gtac	caggag	aagat	tgtgc	tcat	ggcgg	tgag	gtgtac	900
cagg	agaacg	tgtac	cagga	tgag	cgcccg	ttt	gtgagcg	cca	agaaggt	gatg	tgggag	960
atg	ggcgagc	cctac	cgag	ccac	gtggag	ctgt	gtcct	tcc	acaccgt	gtcc	aagggc	1020
actg	ccggcg	agtg	cgccct	gcgc	ggcggc	tacg	tggaga	tgac	taacat	ccac	ccccggc	1080
gcc	attgagg	aggt	gtgcaa	gtgc	gcctcc	atta	acctgt	cgcc	caacac	catg	gggccag	1140
atcg	cgctgt	ccgt	gctcgt	caac	ccgccc	aagc	ccggcg	atcc	ctctta	cgac	cagtac	1200
acca	aggaga	aggc	ctcgga	gctg	gtgtcg	ctgc	gcgcgc	gcgc	gcacat	ggtg	acggac	1260
ggct	taacg	cgct	ggacgg	cgta	ccctgc	aact	tcaccg	aggg	gcctat	gtac	agcttc	1320
cccc	agatta	agct	gcccggc	caagg	cgctg	gagg	ccggca	aggg	ccggcg	aaag	gcgggc	1380
gacg	tgttct	actg	ccctcaa	actt	ctggag	gcc	accggca	tctc	accgt	gccc	ggcagc	1440
ggct	tcggcc	agg	aggagg	cacct	tccac	ctgc	gcacca	ccat	tctgcc	tcgc	gaggag	1500
gtg	atgacgc	actt	cgctgga	gaag	ttcgac	aagt	ttcaca	agg	acttcat	gaag	cagtat	1560
tcg	taa											1566

<210> 51

ES 2 440 265 T3

<211> 521

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 51

5

```

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

```

ES 2 440 265 T3

Lys Gly Gly Val Gly Ala Tyr Thr Asp Ser Arg Gly Asn Pro Leu Val
 145 150 155 160
 Arg Glu Glu Val Ala Arg Phe Ile Glu Lys Arg Asp Gly Val Pro Ser
 165 170 175
 Asn Pro Asp His Ile Phe Leu Thr Asp Gly Ala Ser Val Ala Val Arg
 180 185 190
 Leu Cys Leu Asn Ala Met Ile Arg His Asp Arg Asp Ser Val Leu Val
 195 200 205
 Pro Ile Pro Gln Tyr Pro Leu Tyr Ser Ala Ser Ile Arg Leu Tyr Gly
 210 215 220
 Gly Thr Leu Val Gly Tyr Phe Leu Asp Glu Arg Arg Gly Trp Gly Leu
 225 230 235 240
 Ser Val Glu Glu Leu Gln Arg Ala Leu Gln Glu Ala Arg Glu Glu Gly
 245 250 255
 Lys Leu Val Arg Gly Leu Val Phe Ile Asn Pro Gly Asn Pro Thr Gly
 260 265 270
 Gln Cys Leu Ser Lys Glu Asn Leu Gln Glu Leu Ile Lys Phe Ala Tyr
 275 280 285
 Gln Glu Lys Ile Val Leu Met Ala Asp Glu Val Tyr Gln Glu Asn Val
 290 295 300
 Tyr Gln Asp Glu Arg Pro Phe Val Ser Ala Lys Lys Val Met Trp Glu
 305 310 315 320
 Met Gly Glu Pro Tyr Arg Ser His Val Glu Leu Leu Ser Phe His Thr
 325 330 335
 Val Ser Lys Gly Thr Ala Gly Glu Cys Gly Leu Arg Gly Gly Tyr Val
 340 345 350
 Glu Met Thr Asn Ile His Pro Gly Ala Ile Glu Glu Val Cys Lys Cys
 355 360 365
 Ala Ser Ile Asn Leu Ser Pro Asn Thr Met Gly Gln Ile Ala Leu Ser
 370 375 380
 Val Leu Val Asn Pro Pro Lys Pro Gly Asp Pro Ser Tyr Asp Gln Tyr
 385 390 395 400
 Thr Lys Glu Lys Ala Ser Glu Leu Val Ser Leu Arg Arg Arg Ala His
 405 410 415
 Met Val Thr Asp Gly Phe Asn Ala Leu Asp Gly Val Thr Cys Asn Phe
 420 425 430
 Thr Glu Gly Ala Met Tyr Ser Phe Pro Gln Ile Lys Leu Pro Ala Lys
 435 440 445
 Ala Leu Glu Ala Ala Lys Ala Ala Gly Lys Ala Gly Asp Val Phe Tyr
 450 455 460
 Cys Leu Lys Leu Leu Glu Ala Thr Gly Ile Ser Thr Val Pro Gly Ser
 465 470 475 480
 Gly Phe Gly Gln Glu Glu Gly Thr Phe His Leu Arg Thr Thr Ile Leu
 485 490 495
 Pro Arg Glu Glu Val Met Thr His Phe Val Glu Lys Phe Asp Lys Phe
 500 505 510
 His Lys Asp Phe Met Lys Gln Tyr Ser
 515 520

<210> 52

<211> 1416

5 <212> ADN

<213> Oryza sativa

<400> 52

ES 2 440 265 T3

cccacgcgtc cgcccacgcg tccgggacac cagaaacata gtacacttga gctcactcca 60
 aactcaaaaca ctcacaccaa tggctctcca agttcaggcc gcactcctgc cctctgctct 120
 ctctgtcccc aagaagggtg acttgagcgc ggtggtgaag gagccggggt tccttagcgt 180
 gagcagaagg ccaagaagcc gtcgctggtg gtgagggcgg tggcgacgcg gcgggcccgt 240

ggcgagcccc ggcgcgggca cgtcgaaggc ggacgggaag aagacgctgc ggcagggggt 300
 ggtggtgatc accggcgcgt cgtcggggct cgggctcgcg gcggcgaagg cgcttggcgg 360
 agacggggaa gtggcacgtg gtgatggcgt tccgcgactt tcctgaaggc ggcgacggcg 420
 gcgaaggcgg cgggatggc ggccgggagc tacaccgtca tgcacctgga cctcgctcc 480
 ctcgacagcg tccgccagtt cgtggacaac ttccggcgt ccggcatgcc gctcgacgcg 540
 ctggtgtgca acgcccaca tctaccggcc gacggcgcgg caaccgacgt tcaacgcgga 600
 cgggtacgag atgagcgtcg ggtgaacca cctggggccac ttctcctcg cccgctcat 660
 gctcgacgac ctcaagaaat ccgactacc cgtcgcggcg ctcatcatcc tcggctccat 720
 caccggcaac accaacacct tcgcccga cgtccctccc aaggccgggc taggcgacct 780
 ccgggggctc gccggcgggc tccgcgggca gaacgggtcg gcgatgatcg acggcgcgga 840
 gagcttcgac ggcccaagg cgtacaagga cagcaagatc tgtaacatgc tgacgatgca 900
 ggagttccac cggagattcc acgaggagac cgggatcacg ttccgctcgc tgtaccggg 960
 gtgcatcgcg acgacgggct tgttcgcga gcacatccc ctgttcggc tgtgttccc 1020
 gccgttccag cgttcgtga cgaaggggtt cgtgtcggag gcggagtccg ggaagcggct 1080
 ggcgaggtg gtggcgacc cgagcctgac caagtccggc gtgtactgga gctggaacaa 1140
 ggactcggcg tcgttcgaga accagctctc gcaggaggcc agcgaccggc agaaggccag 1200
 gaagctctgg gacctcagcg agaagctcgt cggcctcgtc tgagtttatt atttaccat 1260
 tcgtttcaac tgttaatttc ttccgggttt aggggtttc agctttcagt gagagaggcc 1320
 tgtcaagtga tgtacaatta gtaattttt tttaccgcac aaatcatgca ataaaaccac 1380
 aggcttacat tatcgatttg tccacctaaa ttaagt 1416

<210> 53

<211> 52

5 <212> ADN

<213> Secuencia artificial

<220>

<223> cebador: prm08408

<400> 53

10 ggggacaagt ttgtacaaa aagcaggctt aaacaatgcg gaaggaagcg ac 52

<210> 54

<211> 50

<212> ADN

<213> Secuencia artificial

15 <220>

<223> cebador: prm08409

<400> 54

ggggaccact ttgtacaaga aagctgggtc gaattgctaa gctgttacga 50

<210> 55

ES 2 440 265 T3

<211> 1453

<212> ADN

<213> Oryza sativa

<400> 55

5

```

atggctgctc ccagcgtcgc cgtcgacaac ctcaacccca aggttttgaa ttgtgagtat    60
gcagtgcgtg gagagattgt gatccatgct cagcgcctgc agcaacagct acagactcaa    120
ccagggtctc ttccctttga tgagatccta tactgcaaca ttgggaatcc ccagtctctt    180
ggtcagaagc cagttacatt cttcagggag gttattgctc tttgtgatca tccatgcttg    240
ttggaaaagg aggaaaccaa atcattgttc agtgctgatg ccatttctcg agcaacaaca    300
attcttgcct cgattcctgg aagagcaact ggagcataca gccacagcca gggcatcaaa    360
gggctgcgtg atgcaattgc tgcctggaatt gcatcacgtg atggataccc tgcaaatgca    420
gacgacattt tccttactga cggagcaagc cctggagttc acatgatgat gcagttactg    480

```

```

ataaggaacg agaaagatgg cattctctgc ccaattcctc aatatccttt gtactcagcc    540
tccattgctc ttcattggtg agctcttgct ccgtattatc ttaatgaatc aacaggctgg    600
ggtttgagag tctctgacct taagaagcaa ctgcaagatt ctcggttgaa aggcattgat    660
gttagggctt tggtagttaa caatccagga aatccaactg ggcaggttct tgctgaggaa    720
aaccaacggg acatagtgaa gttctgcaa aatgaggac ttgttcttct ggctgatgag    780
gtgtaccaag agaacatcta tgttgacaac aagaaattta actctttcaa gaagatagcg    840
agatccatgg gatacaacga ggatgatctc ccttttagtat catttcaatc tgtttctaag    900
ggatattatg gtgaatgtgg caaaagagga ggctacatgg agattactgg cttcagtgtc    960
ccagttagag agcagatcta caaagtggcg tcagtgaact tatgttccaa tatcactggc   1020
cagatccctg ccagcctcgt catgaatcca ccaaaggctg gagatgcata atatgcttca   1080
tacaaggcag agaaagatgg aatcctccaa tcattagctc gccgtgcaa ggcattggag   1140
aatgctttca acagtcttga ggaattaca tgcaacaaaa ctgaaggagc aatgtacctc   1200
ttccctcagc ttagtctgcc aaaaaggca attgacgtg ctaaagctgc taacaaagca   1260
cctgatgctt tctatgcctc tcgtctcctc gaggcaaccg gaattgttgt tgtccctgga   1320
tctggatttg gccaaagtcc tggcacatgg cacatcagat gcacaatcct gccacaggag   1380
gagaagatcc ccgcgatcat ctcccgttc aaggcattcc atgagggctt catggcagcg   1440
taccgcgact gaa                                     1453

```

<210> 56

<211> 483

10 <212> PRT

<213> Oryza sativa

<400> 56

ES 2 440 265 T3

Met Ala Ala Pro Ser Val Ala Val Asp Asn Leu Asn Pro Lys Val Leu
1 5 10 15
Asn Cys Glu Tyr Ala Val Arg Gly Glu Ile Val Ile His Ala Gln Arg
20 25 30
Leu Gln Gln Gln Leu Gln Thr Gln Pro Gly Ser Leu Pro Phe Asp Glu
35 40 45
Ile Leu Tyr Cys Asn Ile Gly Asn Pro Gln Ser Leu Gly Gln Lys Pro
50 55 60
Val Thr Phe Phe Arg Glu Val Ile Ala Leu Cys Asp His Pro Cys Leu
65 70 75 80
Leu Glu Lys Glu Glu Thr Lys Ser Leu Phe Ser Ala Asp Ala Ile Ser
85 90 95
Arg Ala Thr Thr Ile Leu Ala Ser Ile Pro Gly Arg Ala Thr Gly Ala
100 105 110
Tyr Ser His Ser Gln Gly Ile Lys Gly Leu Arg Asp Ala Ile Ala Ala
115 120 125
Gly Ile Ala Ser Arg Asp Gly Tyr Pro Ala Asn Ala Asp Asp Ile Phe
130 135 140
Leu Thr Asp Gly Ala Ser Pro Gly Val His Met Met Met Gln Leu Leu
145 150 155 160
Ile Arg Asn Glu Lys Asp Gly Ile Leu Cys Pro Ile Pro Gln Tyr Pro
165 170 175
Leu Tyr Ser Ala Ser Ile Ala Leu His Gly Gly Ala Leu Val Pro Tyr
180 185 190
Tyr Leu Asn Glu Ser Thr Gly Trp Gly Leu Glu Ile Ser Asp Leu Lys
195 200 205
Lys Gln Leu Glu Asp Ser Arg Leu Lys Gly Ile Asp Val Arg Ala Leu
210 215 220
Val Val Ile Asn Pro Gly Asn Pro Thr Gly Gln Val Leu Ala Glu Glu
225 230 235 240
Asn Gln Arg Asp Ile Val Lys Phe Cys Lys Asn Glu Gly Leu Val Leu
245 250 255
Leu Ala Asp Glu Val Tyr Gln Glu Asn Ile Tyr Val Asp Asn Lys Lys
260 265 270
Phe Asn Ser Phe Lys Lys Ile Ala Arg Ser Met Gly Tyr Asn Glu Asp

ES 2 440 265 T3

		275					280					285			
Asp	Leu	Pro	Leu	Val	Ser	Phe	Gln	Ser	Val	Ser	Lys	Gly	Tyr	Tyr	Gly
	290					295					300				
Glu	Cys	Gly	Lys	Arg	Gly	Gly	Tyr	Met	Glu	Ile	Thr	Gly	Phe	Ser	Ala
305					310					315					320
Pro	Val	Arg	Glu	Gln	Ile	Tyr	Lys	Val	Ala	Ser	Val	Asn	Leu	Cys	Ser
				325					330					335	
Asn	Ile	Thr	Gly	Gln	Ile	Leu	Ala	Ser	Leu	Val	Met	Asn	Pro	Pro	Lys
			340					345					350		
Ala	Gly	Asp	Ala	Ser	Tyr	Ala	Ser	Tyr	Lys	Ala	Glu	Lys	Asp	Gly	Ile
		355					360					365			
Leu	Gln	Ser	Leu	Ala	Arg	Arg	Ala	Lys	Ala	Leu	Glu	Asn	Ala	Phe	Asn
	370					375					380				
Ser	Leu	Glu	Gly	Ile	Thr	Cys	Asn	Lys	Thr	Glu	Gly	Ala	Met	Tyr	Leu
385					390					395					400
Phe	Pro	Gln	Leu	Ser	Leu	Pro	Gln	Lys	Ala	Ile	Asp	Ala	Ala	Lys	Ala
				405					410					415	
Ala	Asn	Lys	Ala	Pro	Asp	Ala	Phe	Tyr	Ala	Leu	Arg	Leu	Leu	Glu	Ala
			420					425					430		
Thr	Gly	Ile	Val	Val	Val	Pro	Gly	Ser	Gly	Phe	Gly	Gln	Val	Pro	Gly
		435					440					445			
Thr	Trp	His	Ile	Arg	Cys	Thr	Ile	Leu	Pro	Gln	Glu	Glu	Lys	Ile	Pro
	450					455					460				
Ala	Ile	Ile	Ser	Arg	Phe	Lys	Ala	Phe	His	Glu	Gly	Phe	Met	Ala	Ala
465					470					475					480
Tyr	Arg	Asp													

<210> 57

<211> 2559

5 <212> ADN

<213> Oryza sativa

<400> 57

ES 2 440 265 T3

gaaaggggag	agaaagagag	agaagggaga	gagagagaga	gagaaggatg	aggaagaaga	60
agggatgggg	cgctggcgag	ctcctctctg	cgggtgaacg	gccgacaagc	tcctcccccg	120
cgcgtggacg	gccagcgacc	tccttccttg	tgcgttgctg	ccgcccggcc	gcgctctagt	180
gattgaaggt	gagaggagag	gaaaagatga	gagagagggg	agaggggtga	gaatgatagc	240
tggggccata	tgtcgggtgg	tcccactatt	tttttttggt	aatgacatgt	tggctctaca	300
aatTTTTgtt	tttactctaa	tgccacctaa	gcgacacgtc	gacgacacgt	ggaacgaaga	360
ccgggtcaac	accgccacgt	aggtgccacg	tcagccaaaa	ccaattccaa	aaccacctag	420
gatatagttt	gcaccggttt	tgtagtttag	aagagtcgat	atatccgggt	ttgtggttgg	480
aggatcatgaa	tcgtactctg	gccatagttg	agggagttaa	agtatatattt	ttccaaggaa	540
aaaatgaatc	gagtgtgtca	aactgaactg	aagacttaaa	aaggttgaat	ggcagtttga	600
ctgctagtcg	attaatcaga	tttaacttta	caatactact	tatttttttc	cctctcgagg	660
aatgctctagc	agtatatttg	cttgacagct	caaaaatata	aaggatttgc	agtaccatcc	720
aaatttagga	acaacataca	tggaaaagac	aaatcgctcg	gcgcatgagg	cgcttacgtg	780
caggaaaaat	aaaaggaaac	tgaagctgga	aaaaagagag	acattataat	ttgccgttgc	840
tcattttcta	tttttagtgag	agttacatgc	gggtgcagtg	gtgctgtgga	gttgtgactc	900
tccacttccg	tgtaatcggg	aaaagaagta	aaaaagaaaa	gaaaagggga	gtcggagaga	960
gcaccggtag	cattattcca	agcaggtgga	cccgcgtgtc	atccccactc	tacaaagcgc	1020
aaaatcatca	agggccttcg	cctcggcgtg	gaggagagtg	aggacggccc	acgcgagca	1080
gcagagagtc	gggaggtggc	tccgcttcca	cagctctact	ccatctctct	cagtgtcggg	1140
ctcgcgggag	tccggccaat	ccagccggtt	catgcttcat	tctctcgggt	cgtgatttct	1200
ccgattttcg	tctccatcta	gtacctgaag	cgaggcaaat	ttaattgccc	ccttttcggg	1260
gcaaactatc	tcgtcagatt	agtcgcatgc	atgttccttc	gttgaatttt	gcaaagttag	1320
ttgtagagag	aagtctcttg	gaggggtggat	gctacggctc	catctctctc	cTTTTcccc	1380
aacaagcgag	ctagcgaagg	ggaaaatggg	gggagcagaa	gaatatccat	gttaggttcg	1440
cgtgcttgcc	tctcggctga	gctctagctg	ttacggcggt	cgtcaggatg	gctaaccggt	1500
ctcgccaatt	agaagatgga	taggtcgtag	cgtagatggt	attacttgat	ggttgatgcg	1560

ctgcccattt	attgttctta	gcaggttctg	tcttctcagt	ccgtgtgagt	gtttcatcat	1620
attggctacc	aagatgatca	ctcttcgttt	atcaagagag	tagggtgaga	tctcaatccg	1680
ttgcaactga	tgagtacttc	cTTTgtctca	gaatgtaagt	atTTTtgagt	tagacacaga	1740
tattaagaaa	gtaggtagag	atgattggag	gagagttgtg	attgatgggg	aagagaaagt	1800
aggtgaaaaa	aaatggttgt	gattggttaa	gaggacagag	taggtgaata	aatagcttca	1860
ttttgagaca	agttactgtg	ctaaaaatag	ctacattttg	agacggagat	agtagtatac	1920
ttcacttact	accgagtacg	gctttagttt	tgctacctcc	gtcctaaaaat	atagcaacct	1980
aggatcggat	gtagcatggt	actactaatc	tagataggca	gcatgtctaa	attcatagta	2040
atatggtgac	tcgtttagta	gaatgttgat	atatttttagg	atggaagaaa	tatataaata	2100
ctgTTTTttt	attcgaagta	gTTggcccat	cattttctgaa	atagatgatt	gatgccatga	2160
cgccgcttgc	tttctagaac	tactagtaat	tttaggtgag	agctagtact	gatgcgtcag	2220
tctaagataa	tggacaaaaa	agggctacag	gctactattg	attatcacat	taaaactctg	2280
tacgacagat	ttttctgatt	aaatgatagc	catatgcccc	acgtgctgct	tgtctaaact	2340
gaaaacctgac	atcactcaca	gtatgccag	ttgTTgggtg	gtctattatt	atttataaat	2400
tataactctg	gcattTTTTt	tattgtaggg	caatatgttt	tccattattt	tccattaaaa	2460
cctctaactc	gcacttccac	tatctgctca	aaatctcagg	ctactttctt	tcctcttctt	2520
caggacatta	acctggttta	cttgtaagaa	agtaaagcc			2559

<210> 58

<211> 49

5 <212> ADN

<213> Secuencia artificial

<220>

<223> cebador: prm001646

<400> 58

ggggacaagt ttgtacaaaa aagcaggctt cacaatggct gctcccagc 49

<210> 59

<211> 46

<212> ADN

5 <213> Secuencia artificial

<220>

<223> cebador: prm001647

<400> 59

ggggaccact ttgtacaaga aagctgggta attcagtcgc ggtacg 46

10

REIVINDICACIONES

1. Un método para incrementar las características relacionadas con el rendimiento de semillas bajo condiciones de cultivo normales, condiciones de cultivo reducidas en nitrógeno y/o condiciones de cultivo con disponibilidad reducida de nutrientes en plantas de arroz con relación a las plantas que sirven como control,
- 5 que comprende incrementar la expresión en una planta de arroz de una secuencia de ácido nucleico que codifica el polipéptido que incrementa el rendimiento 19/20 localizado en el núcleo con el motivo AT-hook (AHL 19/20),
- en donde se logra el incremento de la expresión por medio de la introducción y expresión en una planta de un ácido nucleico que codifica un polipéptido AHL 19/20, y
- 10 en donde dicho polipéptido AHL 19/20 comprende un dominio que tiene al menos 80% o más identidad de secuencia de aminoácidos con un Dominio Conservado (DC) como el representado por la SEQ IS NO: 36,
- y opcionalmente seleccionar plantas de arroz que tienen más características relacionadas con el rendimiento de semillas,
- dicha característica relacionada con el rendimiento de semillas es una o más de: (i) aumento del número de flores por panícula; (ii) aumento del peso total de las semillas por planta; (iii) aumento del número de semillas llenas; o (iv) aumento de índice de cosecha.
- 15
2. Un método de acuerdo con la reivindicación 1, en donde dicho polipéptido que incrementa el rendimiento tiene al menos 50% o más identidad de secuencia de aminoácidos con el polipéptido AHL 19/20 como el representado por la SEQ ID NO: 2 o con cualquiera de las secuencias polipeptídicas suministradas en la Tabla A aquí.
3. Un método de acuerdo con cualquier reivindicación precedente, en donde dicha mayor expresión se logra por una cualquiera o más de: marcación de la activación de ADN-T, TILLING, o recombinación homóloga.
- 20
4. Construcción que comprende:
- (a) Una secuencia de ácido nucleico que codifica una polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: un 19/20 localizado en el núcleo (AHL 19/20) con un motivo AT-hook como se define en cualquiera de las reivindicaciones 1 a 2;
- 25 (b) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (a); en donde dicha secuencia de control es un promotor GOS2 de arroz; y opcionalmente;
- (c) una secuencia de terminación de la transcripción.
5. Una construcción de acuerdo con la reivindicación 4, en donde la secuencia de ácido nucleico está operativamente enlazada a un promotor GOS2 de arroz como el representado por la SEQ ID NO: 35..
- 30
6. El uso de una construcción que comprende:
- (a) Una secuencia de ácido nucleico que codifica una polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: un 19/20 localizado en el núcleo (AHL 19/20) con un motivo AT-hook como se define en cualquiera de las reivindicaciones 1 a 2;
- 35 (b) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (a); y opcionalmente;
- (c) una secuencia de terminación de la transcripción
- en un método para elaborar plantas de arroz que tienen más características relacionadas con el rendimiento de semillas con relación a plantas que sirven de control bajo condiciones de cultivo normales, condiciones de cultivo reducidas en nitrógeno y/o condiciones de cultivo con disponibilidad reducida de nutrientes, cuyas características relacionada con el rendimiento de semillas son una o más de: (i) aumento del número de flores por panícula; (ii) aumento del peso total de las semillas por planta; (iii) aumento del número de semillas llenas; o (iv) aumento de índice de cosecha.
- 40
7. Planta de arroz, parte de la planta o célula de la planta de arroz, por lo cual la planta de arroz, parte de la planta o

célula de la planta de arroz es

(i) transformada con una construcción que comprende:

5 (a) una secuencia de ácido nucleico que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: un 19/20 localizado en el núcleo (AHL 19/20) con un motivo AT-hook como se define en cualquiera de las reivindicaciones 1 a 2 en donde la secuencia de ácido nucleico está

(aa) operativamente enlazada con un promotor constitutivo; o

(bb) operativamente enlazada con un promotor constitutivo de la planta; o

(cc) operativamente enlazada con un promotor GOS2; o

(dd) operativamente enlazada con un promotor GOS2 de arroz como el representado por la SEQ ID NO: 35;

10 (b) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (a); y opcionalmente;

(c) una secuencia de terminación de la transcripción

o

(ii) se obtiene por medio de un método de acuerdo con las reivindicaciones 1 - 2;

15 por lo cual las partes de la planta o las células de la planta comprenden la construcción de acuerdo con el ítem (i).

8. Un método para la producción de plantas de arroz transgénico que tienen más características relacionadas con el rendimiento de semillas con relación a plantas que sirven de control bajo condiciones de cultivo normales, condiciones de cultivo reducidas en nitrógeno y/o condiciones de cultivo con disponibilidad reducida de nutrientes, que comprende:

20 (i) la introducción y expresión en una planta de arroz, parte de una planta de arroz, o una célula de una planta de arroz, de una secuencia de ácido nucleico que codifica un polipéptido para incrementar el rendimiento seleccionado del grupo que consiste de: uno 19/20 localizado en el núcleo con un motivo AT-hook (AHL 19/20) como se define en cualquiera de las reivindicaciones 1 a 2, bajo el control de un promotor constitutivo de una planta; y

25 (ii) el cultivo de la célula de la planta, parte de la planta de arroz, o la planta de arroz bajo condiciones que promuevan el crecimiento y desarrollo de la planta.

9. Partes cosechables de arroz que comprenden una secuencia de ácido nucleico que codifica el polipéptido que incrementa el rendimiento 19/20 localizado en el núcleo con un motivo AT-hook (AHL 19/20) como se define en las reivindicaciones 1 o 2 de una planta de arroz de acuerdo con la reivindicación 7, en donde dichas partes cosechables son semillas y en donde dichas partes cosechables contienen una construcción que comprende:

30 (a) una secuencia de ácido nucleico que codifica un polipéptido que incrementa el rendimiento seleccionado del grupo que consiste de: un 19/20 localizado en el núcleo (AHL 19/20) con un motivo AT-hook como se define en cualquiera de las reivindicaciones 1 a 2 en donde la secuencia de ácido nucleico está

(aa) operativamente enlazada con un promotor constitutivo; o

(bb) operativamente enlazada con un promotor constitutivo de la planta; o

35 (cc) operativamente enlazada con un promotor GOS2; o

(dd) operativamente enlazada con un promotor GOS2 de arroz como el representado por la SEQ ID NO: 35;

(b) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (a); y opcionalmente;

(c) una secuencia de terminación de la transcripción

- 5 10. El uso de una secuencia de ácido nucleico que codifica el polipéptido que incrementa el rendimiento 19/20 localizado en el núcleo con un motivo AT-hook (AHL 19/20) como se define en las reivindicaciones 1 o 2, en incrementar las características relacionadas con el rendimiento de semillas bajo condiciones de cultivo normales, condiciones de cultivo reducidas en nitrógeno y/o condiciones de cultivo con disponibilidad reducida de nutrientes en plantas de arroz, que comprende una o más de: (i) aumento del número de flores por panícula; (ii) aumento del peso total de las semillas por planta; (iii) aumento del número de semillas llenas; o (iv) aumento de índice de cosecha.
11. El método de acuerdo con la reivindicación 2, en donde la secuencia de ácido nucleico está
- (a) operativamente enlazada con un promotor constitutivo; o
- (b) operativamente enlazada con un promotor constitutivo de la planta; o
- 10 (c) operativamente enlazada con un promotor GOS2; o
- (d) operativamente enlazada con un promotor GOS2 de arroz como el representado por la SEQ ID NO: 35.

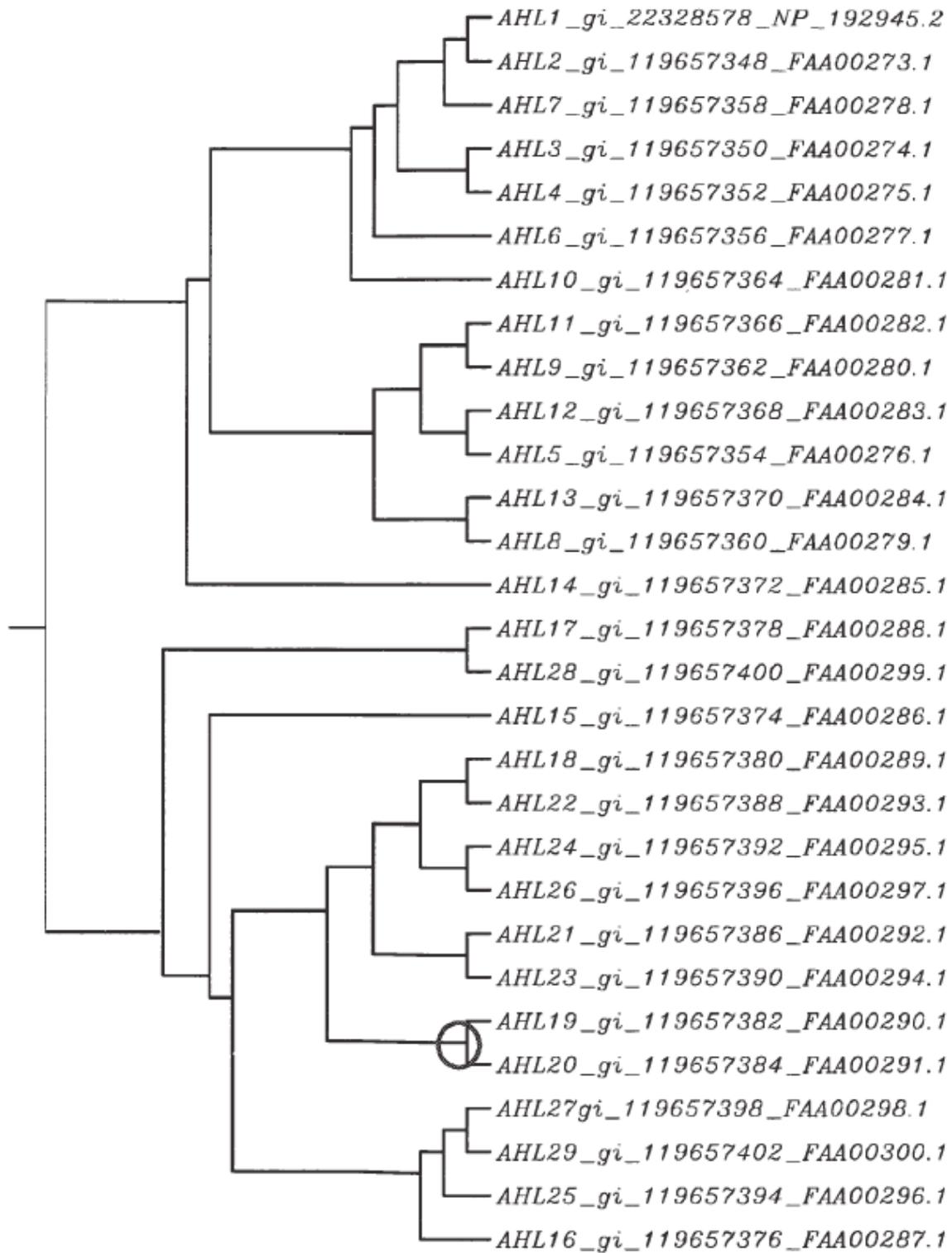


FIGURA 1

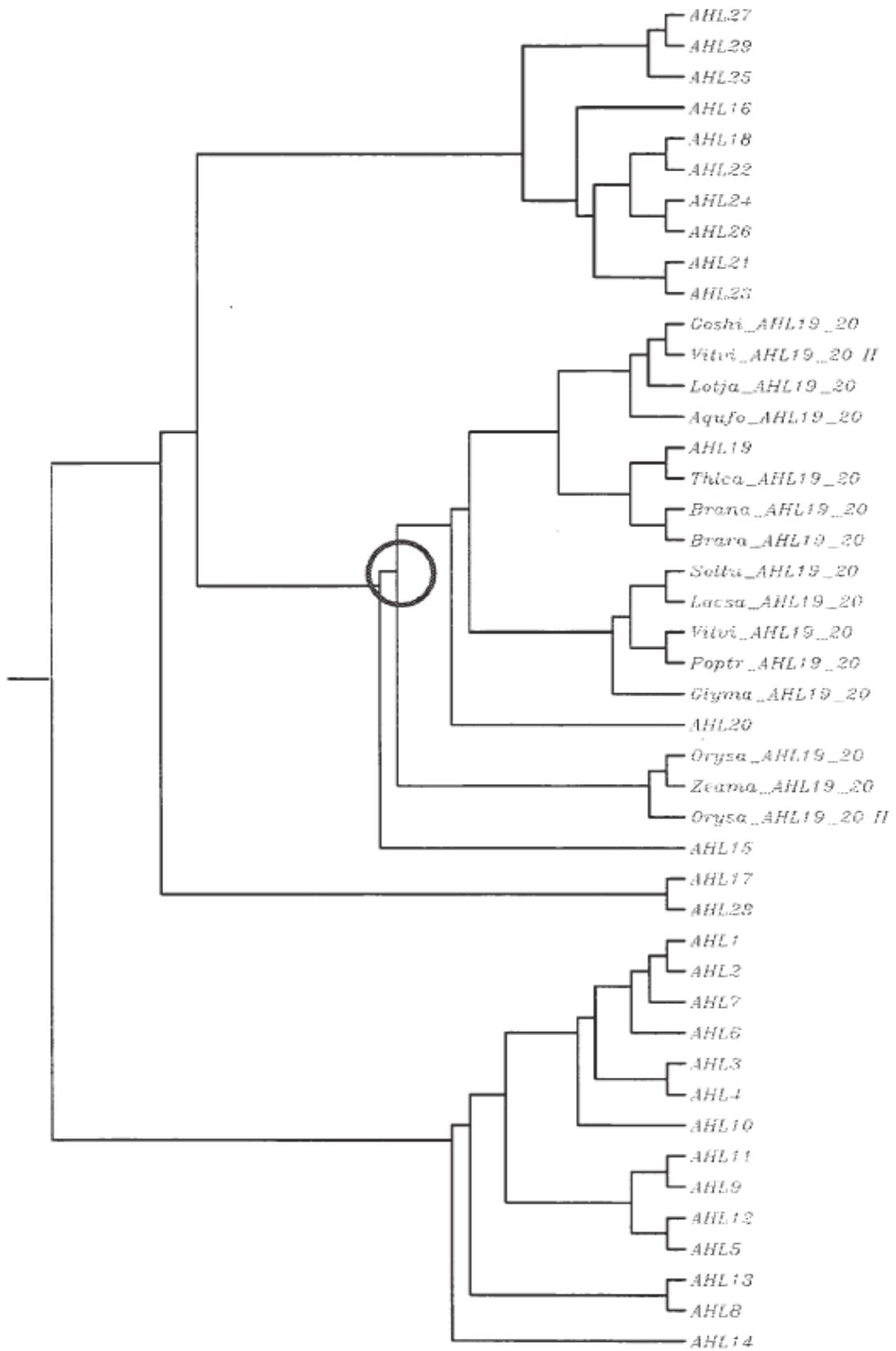


FIGURA 2

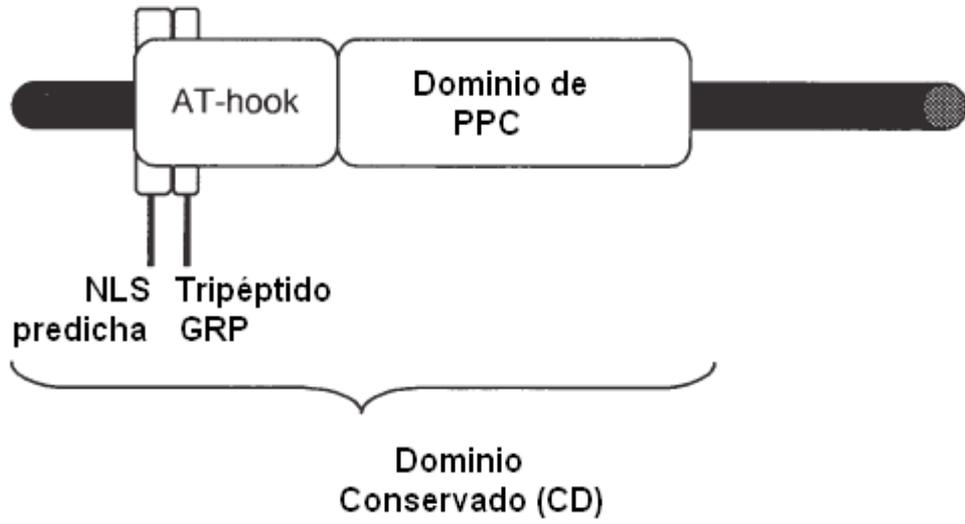


FIGURA 3

Alineación de secuencias múltiples del Dominio Conservado (CD) con CLUSTAL W (1.83) (continuación)

CD_Lotja_AHL19_20	ARRRQGVCMSSGGSVANVTLRQP	-----AAPGAVVALHGRFEILSLTGAFLLPGPA
CD_Vitvi_AHL19_20\II	ARRRQGVCVLSGSGSVANVTLRQP	-----AAPGAVVALHGRFEILSLTGAFLLPGPA
CD_Goshi_AHL19_20	ARRRQGVCLLSGSGSVANVTLRQP	-----AAPGAVVALHGRFEILSLTGAFLLPGPA
CD_Aqufo_AHL19_20	ARRRQGVCLSGGSVANVTLRQP	-----SAPSAVVALQGRFEILSLTGSFLPGPA
CD_Orysa_AHL19_20	ARRRQGVCLSGAGTVDVALRQP	-----APSAVVALRGRFEILSLTGTFFLLPGPA
CD_Zeama_AHL19_20	ARRRQGVCLSGAGTVADVALRQP	-----AAPGAVVALRGRFEILSLTGTFFLLPGPA
CD_Orysa_AHL19_20\II	SRRRQGVCLSGAGTVANVALRQP	-----SAPGAVVALHGRFEILSLTGTFFLLPGPA
CD_Arath_AHL19	ARRRQVICILSGNGTVANVTLRQP	STAAVAAPGGAAVLALQGRFEILSLTGSFLPGPA
CD_Thlca_AHL19_20	ARRRQVICILSGNGTVANVTLRQP	SSAAVAAPGGAAVLALQGRFEILSLTGSFLPGPA
CD_Brana_AHL19_20	ARRRQVICILSGNGTVANVTLRQP	SVAPVAAAPGGAAVLALQGRFEILSLTGSFLPGPA
CD_Brara_AHL19_20	ARRRQVICILSGNGTVANVTLRQP	SVAPVAAAPGGAAVLALQGRFEILSLTGSFLPGPA
CD_Arath_AHL20	SRRRQGVCLSGTGSVANVTLR	-----QAAAPG--GVVSLQGRFEILSLTGAFLLPGPS
CD_Lacsa_AHL19_20	SRKRQGVCMSSASGTMNVTLRQP	-----SAPG-SVMALQGRFEILSLTGAFLLPGPS
CD_Soltu_AHL19_20	ARKRQGVCLSATGTVTNVTLRQP	-----SAPG-AVMALHGRFEILSLTGAFLLPGPA
CD_Vitvi_AHL19_20	ARRRQGVCLVLSASGTMNVTLRQP	-----SAPGGAVMALHGRFEILSLTGAFLLPGPA
CD_Glyma_AHL19_20	ARRRQGVSIILSGSGTVVNVNLRQP	-----TAPGAVMALHGRFDILSLTGSFLPGPS
CD_Poptr_AHL19_20	ARKRQGVCLVLSGSGMVTNVTLKQP	-----SASGAVMALHGRFEILSLTGAFLLPGPA
:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:	*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:

Dominio de PPC (continuación)

FIGURA 4 (continuación)

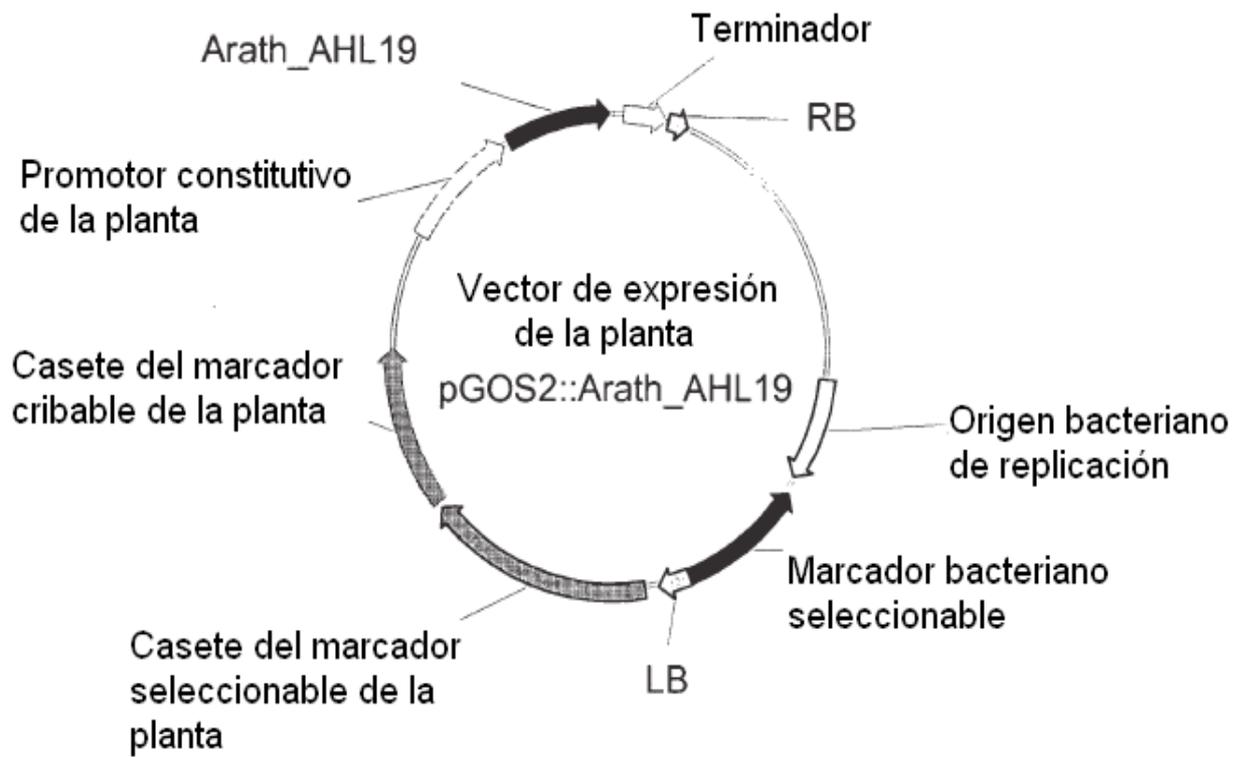


FIGURA 5

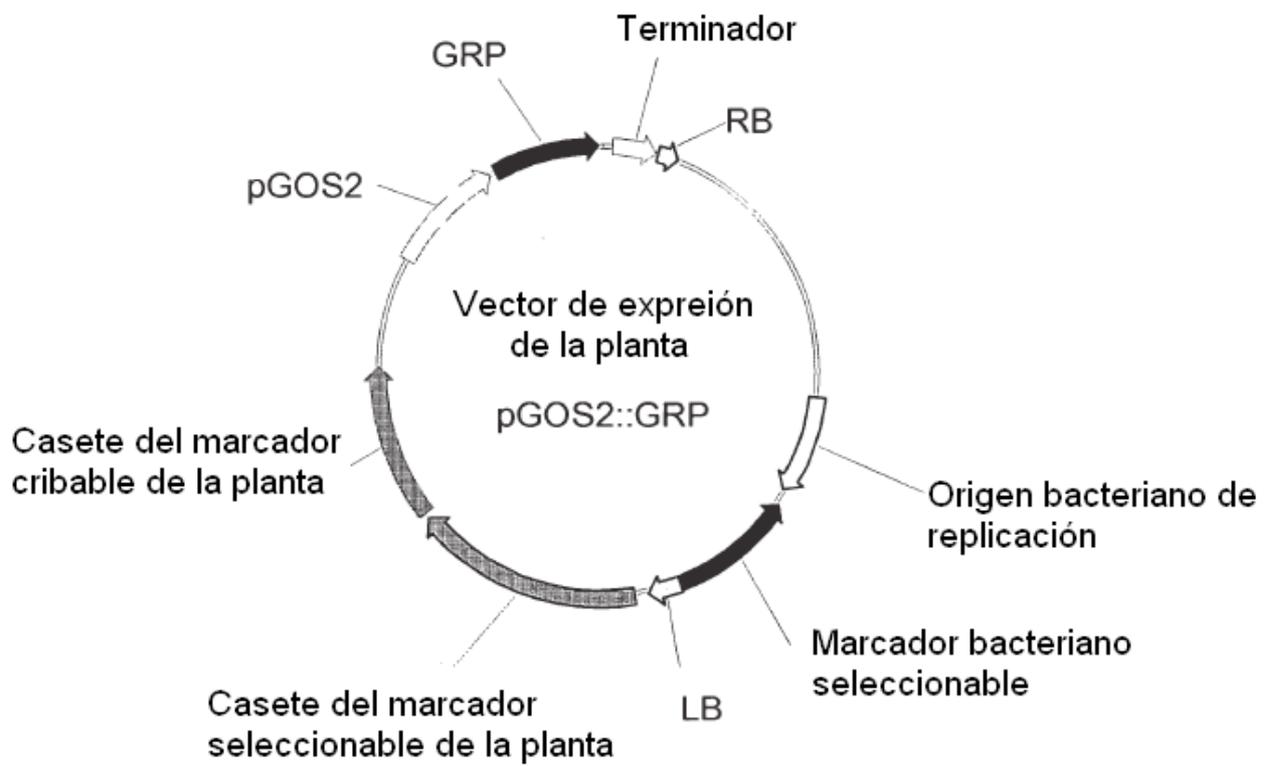


FIGURA 7

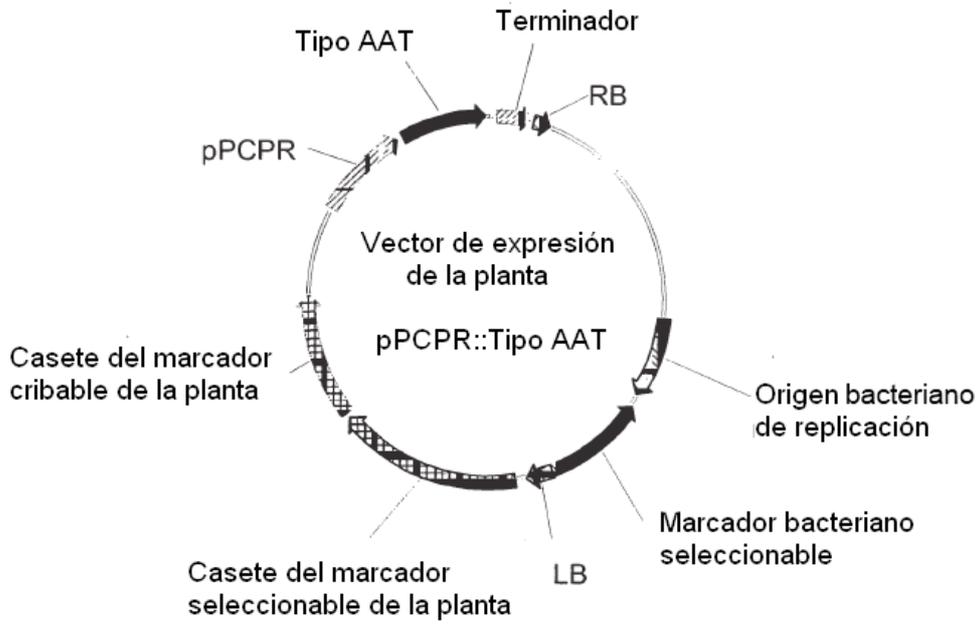


FIGURA 9

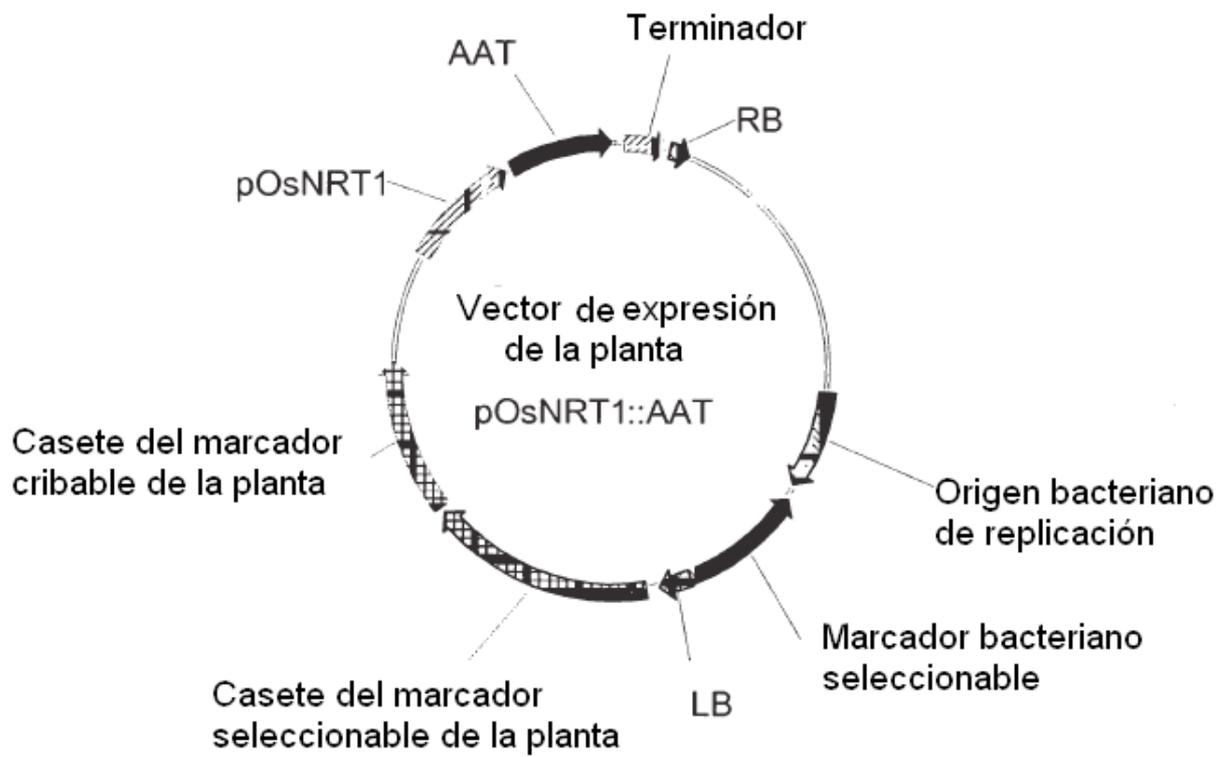


FIGURA 8