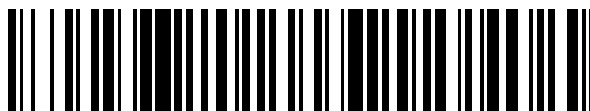


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 531 376**

51 Int. Cl.:

C07K 14/415 (2006.01)

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.04.2008** **E 08749937 (2)**

97 Fecha y número de publicación de la concesión europea: **10.12.2014** **EP 2144925**

54 Título: **Plantas que tienen mejores características de crecimiento en condiciones de disponibilidad reducida de nutrientes y un procedimiento para producirlas**

30 Prioridad:

30.04.2007 US 799083

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
13.03.2015

73 Titular/es:

CROPDESIGN N.V. (100.0%)
Technologiepark 21C
9052 Zwijnaarde, BE

72 Inventor/es:

SANZ MOLINERO, ANA ISABEL

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 531 376 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 mejores características de crecimiento en condiciones de disponibilidad reducida de nutrientes y un procedimiento para producirlas

La población mundial en crecimiento y el suministro menguante de tierra arable disponible para agricultura está estimulando la investigación hacia el incremento de la eficiencia de la agricultura. Los medios convencionales para mejoras de cultivos y horticultura usan técnicas de siembra selectivas para identificar plantas que tengan características deseables. No obstante, dichas técnicas selectivas tienen varios inconvenientes, es decir, estas técnicas normalmente son muy laboriosas y dan lugar a plantas que a menudo contienen componentes genéticos heterogéneos que no siempre pueden hacer que el rasgo deseable se pase de las plantas parentales. Los avances en biología molecular han permitido que la humanidad modifique el plasma germinal de los animales y las plantas. La modificación genética de las plantas abarca el aislamiento y la manipulación de material genético (normalmente en forma de ARN o ADN) y la posterior introducción de dicho material genético en una planta. Dicha tecnología tiene la capacidad de liberar cultivos o plantas que tengan varios rasgos económicos, agronómicos u hortícolas mejorados.

Un rasgo de interés económico concreto es el rendimiento. El rendimiento normalmente se define como el producto mensurable de valor económico de un cultivo. Este se puede definir en términos de cantidad y/o de calidad. El rendimiento depende directamente de diversos factores, por ejemplo, número y tamaño de los órganos, arquitectura de la planta (por ejemplo, del número de ramas), rendimiento de semillas y más factores. El desarrollo de raíces, la absorción de nutrientes y la eficiencia del uso y la tolerancia al estrés, pueden ser factores importantes en la determinación del rendimiento. Por lo tanto, optimizar uno de los factores mencionados anteriormente puede contribuir a incrementar el rendimiento de una cosecha.

El rendimiento de semillas es un rasgo particularmente importante, ya que las semillas de muchas plantas son importantes para la nutrición de seres humanos y animales. Cosechas tales como maíz, arroz, trigo, canola y soja representan más de la mitad del aporte calórico total de los seres humanos, ya sea a través del consumo directo de las propias semillas o a través del consumo de productos alimenticios generados con semillas procesadas. También son una fuente de azúcares, aceites y muchos tipos de metabolitos usados en procedimientos industriales. Las semillas contienen un embrión (la fuente de nuevos vástagos y raíces) y un endospermo (la fuente de nutrientes para el crecimiento del embrión durante la germinación y durante el crecimiento temprano de las plántulas). El desarrollo de una nueva semilla implica muchos genes y requiere la transferencia de metabolitos desde las raíces, las hojas y los tallos hasta la semilla en crecimiento. El endospermo, en particular, asimila los precursores metabólicos de carbohidratos, aceites y proteínas y los sintetiza como macromoléculas de almacenamiento para engordar el grano.

La biomasa de la planta es rendimiento para cultivos de forraje como la alfalfa, ensilaje y heno. En concreto, la biomasa de la raíz es el rendimiento para cultivos como la patata, la mandioca o la remolacha azucarera. Se han usado muchas aproximaciones para los cultivos de granos. Entre los principales se encuentran las estimaciones del tamaño de la planta. El tamaño de la planta se puede medir de muchos modos dependiendo de la especie y su estado de desarrollo, pero incluyen el peso total de la planta, el peso seco de la planta encima del suelo, el peso fresco de la planta encima del suelo, el área de la hoja, el volumen del tallo, la altura de la planta, el diámetro de la roseta, la longitud de la hoja, la longitud de la raíz, la masa de la raíz, el número de brotes y el número de hojas. Muchas especies mantienen una proporción conservada entre el tamaño de diferentes partes de la planta en un estadio del desarrollo dado. Estas relaciones alométricas se usan para extrapolar desde una de estas medidas del tamaño a la otra (p. ej., Tittmonell et al., (2005) *Agric Ecosys & Environ* 105: 213). El tamaño de la planta en una etapa del desarrollo temprana normalmente se correlacionará con el tamaño de la planta más tarde en el desarrollo. Una planta más grande con un área de la hoja más grande normalmente absorberá más luz y dióxido de carbono que una planta más pequeña y, por tanto, probablemente ganará mayor peso durante el mismo periodo (Fasoula y Tollenaar (2005) *Maydica* 50: 39). Esto se añade a la potencial continuación del microambiente o la ventaja genética que la planta tenía que alcanzar el mayor tamaño inicialmente. Existe un fuerte componente genético con respecto al tamaño de la planta y la tasa de crecimiento (por ejemplo, ter Steege et al. 2005 *Plant Physiology* 139: 1078) y por eso para un intervalo de tamaño de planta de genotipos diversos en una condición ambiental se correlaciona posiblemente con el tamaño en otra (Hittalmani et al. (2003) *Theoretical Applied Genetics* 107:679). De este modo se usa un ambiente estándar como aproximación para los diversos y dinámicos ambientes que se las cosechas en el campo se encuentran en diferentes localizaciones y tiempos.

Otro rasgo importante para muchos cultivos es el vigor temprano. La mejora del vigor temprano es un importante objetivo de los modernos programas de cultivo del arroz en variedades cultivadas de arroz tanto templados como tropicales. Las raíces largas son importantes para el anclaje apropiado al suelo en arroz sembrado en agua. Cuando el arroz se siembra directamente en campos anegados y cuando las plantas deben emerger rápidamente a través del agua, los vástagos más largos se asocian al vigor. Cuando se practica la siembra en línea, los mesocótilos y los coleóptilos más largos son importantes para una buena emergencia de las plántulas. La capacidad para manipular genéticamente el vigor temprano en plantas sería de gran importancia en la agricultura. Por ejemplo, un vigor temprano escaso ha sido una limitación para la introducción de híbridos de maíz (*Zea mays* L.) basados en el germoplasma de Corn Belt en la Europa atlántica.

El índice de cosecha, la relación entre el rendimiento de semillas con respecto al peso seco por encima del suelo es relativamente estable en muchas condiciones ambientales y, de este modo, se puede obtener una muy buena correlación entre el tamaño de la planta y el rendimiento de los granos (por ejemplo, Rebetzke et al., (2002) Crop Science 42:739). Estos procesos están intrínsecamente relacionados porque la mayor parte de la biomasa del grano depende de la productividad fotosintética presente o almacenada por las hojas y el tallo de la planta (Gardener et al., (1985) Physiology of Crop Plants. Iowa State University Press, pp68 - 73). Por tanto, la selección del tamaño de la planta, incluso en etapas tempranas de desarrollo, se ha utilizado como un indicador del rendimiento potencial futuro (por ejemplo, Tittone et al., (2005) Agric Ecosys & Environ 105: 213). Cuando se analiza el impacto de las diferencias genéticas sobre la tolerancia al estrés, la capacidad para 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 follaje, 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, en condiciones estándar en una cámara de cultivo o en invernadero, son prácticas convencionales para proporcionar indicaciones de las posibles ventajas de rendimiento genético.

Otro rasgo importante es el de la tolerancia mejorada al estrés abiótico. El estrés abiótico es una causa principal de pérdida de cosechas en todo el mundo, reduciendo los rendimientos medios para la mayoría de las plantas de cultivo principales en más de 50% (Wang et al., (2003) Planta 218: 1 - 14). El estrés abiótico puede producirse por sequía, salinidad, temperaturas extremas, toxicidad química, exceso o falta de nutrientes (macroelementos y/o microelementos), radiación y estrés oxidativo. La capacidad para aumentar la tolerancia de las plantas al estrés abiótico sería de un gran beneficio económico para los agricultores de todo el mundo y permitiría el cultivo de cosechas en condiciones adversas y en territorios en los que el cultivo de las cosechas no es posible de otro modo.

La capacidad para aumentar el rendimiento de las plantas tendría muchas aplicaciones en áreas tales como agricultura, incluyendo la producción de plantas decorativas, arboricultura, horticultura y silvicultura. El incremento del rendimiento también puede ser útil en la producción de algas para usar en biorreactores (para la producción biotecnológica de sustancias tales como productos farmacéuticos, anticuerpos o vacunas, o para la bioconversión de residuos orgánicos) y otras de estas áreas.

Antecedentes

Las proteínas homeodominio de tipo cremallera de leucina (HDZip) constituyen una familia de factores de transcripción caracterizados por la presencia de un dominio de unión al ADN y un motivo de cremallera de leucina (Zip) adyacente. El homeodominio normalmente consiste en 60 residuos de aminoácidos conservados que forman una hélice 1-bucle-hélice 2-giro-hélice 3 que se une al ADN. Este sitio de unión al ADN normalmente es pseudopalindrómico. La cremallera de leucina, adyacente al extremo C terminal del homeodominio, consiste en varias repeticiones en héptada (al menos cuatro) en las que normalmente aparece una leucina (en ocasiones un valina o una isoleucina) cada siete aminoácidos. La cremallera de leucina es importante para la dimerización de las proteínas. Esta dimerización es un requisito previo para la unión a ADN ((1993) EMBO J 12(9): 3507 - 3517), y puede proceder entre dos proteínas HDZip idénticas (homodímero) o entre dos proteínas HDZip diferentes (heterodímero).

Los genes del homeodominio están presentes en todos los eucariotas y constituyen una familia génica de al menos 89 miembros en *Arabidopsis thaliana*. La cremallera de leucina también se encuentra en eucariotas distintos a las plantas. No obstante, la presencia de un homeodominio y una cremallera de leucina es específica de la planta (se encuentra en al menos 47 de las 89 proteínas en *Arabidopsis*) y se ha descubierto en musgo además de en plantas vasculares, (Sakakibara et al. (2001) Mol Biol Evol 18(4): 491 - 502). La cremallera de leucina se localiza en el extremo C terminal del homeodominio, estas dos características se solapan en tres aminoácidos.

Los genes de HDZip de *Arabidopsis* se han clasificado en cuatro clases diferentes, HDZip I a IV, en base a los criterios de similitud de secuencia (Sessa et al. (1994) En Plant Molec Biol, pp412 - 426). Como las proteínas HD-Zip de las tres otras clases, las proteínas HDZip de clase I son bastante divergentes en su estructura amino primaria fuera el homeodominio y la cremallera de leucina. Dentro del homeodominio y de la cremallera de leucina, las proteínas HDZip de clase I se caracterizan adicionalmente por dos características específicas:

1) en el homeodominio, además de los aminoácidos invariables Leu₁₆Trp₄₈Phe₄₉Asn₅₁Arg₅₃, la posición 46 está ocupada por una Ala (A) y la posición 56 por un Try (W) (o en ocasiones por una Phe (F)) (Sessa et al. (1997) J Mol Biol 274(3):303 - 309; véase la Figura 1), al que se hace referencia como homeodominio de clase I, y

2) la cremallera de leucina comprende seis héptadas para el helecho *Ceratopteris richardii* que presenta siete héptadas (dentro de cada héptada, las posiciones se denominan a, b, c, d, e, f and g, estando la leucina conservada en la posición d; Sakakibara et al (2001) Mol Biol Evol 18(4): 491 - 502; véase la Figura 1). HDZip II, III y IV presentan una cremallera de leucina con solo cinco héptadas.

Concerniente a sus propiedades de unión a ADN, las proteínas HDZip de clase I preferentemente se unen a

semisitos de 5 pb que solapan en una posición central CAA(A/T)ATTG (Sessa et al. (1993) EMBO J 12(9): 3507 - 3517).

Se ha demostrado que diferentes proteínas HDZip activan o reprimen la transcripción. En Arabidopsis, se ha demostrado que las HDZip de clase I -5, -6 y -16 actúan como activadores de la transcripción en ensayos de expresión transitoria en hojas de Arabidopsis usando un gen indicador (luciferasa; Henriksson et al. (2005) Plant Phys 139: 509 - 518). Dos proteínas HDZip de clase I de arroz, Oshox4 y Oshox5, actuaron como activadores en ensayos de expresión transitoria en cultivos en suspensión de células de arroz usando otro gen indicador (glucuronidasa; Meijer et al. (2000) Mol Gen Genet 263:12 - 21). Por el contrario, dos proteínas HDZip de clase I de arroz, Oshox1 y Oshox3, actuaron como represores de la transcripción en los mismos experimentos (Meijer et al. (1997) Plant J 11: 263 - 276; Meijer et al. (2000) ant.).

Se ha demostrado que varias proteínas HDZip de clase I están implicadas en la respuesta a la luz y en la respuesta relacionada con el déficit de ácido abscísico (ABA)agua (Hjellström et al., (2003) Plant Cell Environ 26: 1127 - 1136). Las HDZip de clase I -3, -13, -20, y -23 que sobreexpresan Arabidopsis transgénico sugieren que estos genes están implicados en la regulación del desarrollo de cotiledones y de hojas (Aoyama et al. (1995) Plant Cell 7: 1773 - 1785; Hanson (2000) en Comprehensive summaries of Uppsala Dissertations from the Faculty of Science and Technology, Uppsala). Los genes de ATHB3, -13, -20, y -23 son similares y forman una subclase distinta dentro de las HDZip de clase I. Dado que estos genes producen alteraciones similares en la forma de los cotiledones cuando se expresan de forma constitutiva, se denominan genes de HDZip dirigidos a los cotiledones (POC). Hanson concluye que las proteínas de clase I que están estrechamente relacionadas filogenéticamente también están relacionadas funcionalmente, en la mayoría de los casos.

El documento WO2007/051866 es una solicitud internacional que divulga, entre otros, usos del polipéptido HDZip hox5 en plantas. No obstante, no se divulgan resultados para el crecimiento en condiciones de poca disponibilidad de nitrógeno.

El documento US2007/0022495 divulga plantas transgénicas transformadas con factores de transcripción de plantas.

Sorprendentemente, ahora se ha descubierto que la modulación de la expresión en una planta, una secuencia de ácido nucleico que codifica un polipéptido de HDZip hox5 de clase I o un homólogo de la misma da plantas que tienen un mayor rendimiento en condiciones de disponibilidad reducida de nutrientes, respecto a las correspondientes plantas de tipo silvestre.

De acuerdo con una realización de la presente invención se proporciona un procedimiento para aumentar el rendimiento en plantas cultivadas en condiciones de disponibilidad reducida de nutrientes, respecto a las correspondientes plantas silvestres, que comprende modular la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo, en el que dicha secuencia de ácido nucleico que codifica dicho polipéptido hox5 homeodominio con cremallera de leucina (HDZip) u homólogo del mismo se selecciona del grupo que consiste en:

- a) una secuencia de ácido nucleico representada por la SEC ID N° 1;
- b) una secuencia de ácido nucleico que codifica un polipéptido representada por la SEC ID N° 2,
- c) una secuencia de ácido nucleico que es capaz de hibridar en condiciones rigurosas con una secuencia de ácido nucleico representada por la SEC ID N° 1 y
- d) una secuencia de ácido nucleico que codifica un polipéptido que tiene una identidad de secuencia de al menos un 70 % con un polipéptido, representada por la SEC ID N° 2.

La elección de plantas control adecuadas es una parte rutinaria de una situación experimental y puede incluir las correspondientes plantas silvestres o las correspondientes plantas sin el gen de interés. La planta de control es típicamente de la misma especie de planta o incluso de la misma variedad que la planta que se va a evaluar. La planta de control también puede ser un nulicigoto de la planta que se va a evaluar. Una "planta de control" como se usa en el presente documento no solo se refiere a plantas completas, sino también a partes de las plantas, que incluyen semillas y partes de semilla

De forma ventajosa, el funcionamiento de los procedimientos de acuerdo con la presente invención tiene como resultado plantas que tienen un rendimiento mayor cuando se cultivan en condiciones de disponibilidad reducida de nutrientes, respecto a las correspondientes plantas silvestres.

La expresión "rendimiento aumentado" como se define en el presente documento quiere decir un incremento en uno cualquiera o más de los siguientes, cada uno con respecto a las plantas silvestres correspondientes: (i) aumento en biomasa (peso) de una o más partes de una planta, las partes sobre la superficie (cosechables), aumento de la biomasa de la raíz, aumento del volumen de la raíz, aumento del número de raíces, aumento del diámetro de la raíz o aumento de la longitud de la raíz (de raíces gruesas o finas) o aumento de la biomasa de cualquier otra parte

- cosechable; (ii) aumento en el rendimiento total de semilla, que incluye un aumento en la biomasa de las semillas (peso de la semilla) y que puede ser un incremento del peso de la semilla por planta o en base a una semilla individual; (iii) aumento del número de flores (floretes) por panícula, que se expresa como una proporción entre el número de semillas llenas y el número de panículas primarios; (iv) aumento de la tasa de relleno de las semillas; (v) número de semillas (llenas); (vi) aumento del tamaño de las semillas, que también puede influir sobre la composición de las semillas; (vii) aumento del volumen de las semillas, que puede también influir sobre la composición de las semillas (incluyendo contenido y composición total de aceite, proteína y carbohidrato); (viii) aumento del área de las semillas (individual o promedio); (ix) aumento de la longitud de las semillas (individual o promedio); (x) aumento de la anchura de las semillas (individual o promedio); (xi) aumento del perímetro de las semillas (individual o promedio); (xii) aumento del índice de recolección (IR), que se expresa como la proporción entre el rendimiento de partes cosechables, tales como las semillas, y la biomasa total; (xiii) aumento del peso de mil granos (TKW), que se extrapola del número de semillas llenas contadas y su peso total. Un aumento del TKW puede ser el resultado de un aumento del tamaño de las semillas y/o del peso de las semillas. Un aumento del TKW puede ser el resultado de un aumento del tamaño del embrión y/o del tamaño del endospermo.
- 15 Preferentemente, el rendimiento aumentado se selecciona de uno o más de los siguientes: aumento del número total de semillas, aumento del número de semillas llenas, aumento del rendimiento total de las semillas, aumento del número de flores por panícula, aumento de la tasa de llenado de las semillas, aumento del IR, aumento del TKW, aumento de la longitud de la raíz o aumento del diámetro de la raíz, cada uno respecto a las correspondientes plantas silvestres.
- 20 Por tanto, de acuerdo con la presente invención se proporciona un procedimiento para aumentar el rendimiento en plantas cultivadas en condiciones de disponibilidad reducida de nitrógeno respecto a las correspondientes plantas silvestres, en el que el procedimiento comprende modular la expresión en una planta de un ácido nucleico como se define en la reivindicación 1 que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo.
- 25 Tomando el maíz como un ejemplo, un aumento en el rendimiento puede manifestarse como uno o más de los siguientes: aumento en el número de plantas por hectárea o acre, un aumento en el número de espigas por planta, un aumento en el número de filas, número de granos por fila, peso de grano, peso de mil granos, longitud/diámetro de la espiga, entre otros. Tomando el arroz como un ejemplo, un aumento en el rendimiento puede manifestarse como un aumento en uno o más de los siguientes: número de plantas por hectárea o acre, número de panículas por planta, número de espiguillas por panícula, número de flores por panícula, aumento en el índice de llenado de la semilla, aumento en peso de mil granos, entre otros. Un aumento en el rendimiento también puede ser el resultado de una arquitectura modificada o puede producirse como un resultado de arquitectura modificada.
- 30 Dado que las plantas transgénicas preparadas mediante el procedimiento de acuerdo con la presente invención tienen un rendimiento aumentado en condiciones de disponibilidad reducida del nitrógeno, es probable que estas plantas presenten una tasa de crecimiento aumentada (durante al menos parte de su ciclo de vida), con respecto a la tasa de crecimiento de las correspondientes plantas silvestres o plantas de control en una etapa correspondiente en su ciclo de vida. La tasa de crecimiento aumentada puede ser específica de una o más partes de una planta (incluyendo raíces o semillas) o puede ser sustancialmente en toda la planta. Una planta que tiene una tasa de crecimiento aumentada puede incluso exhibir una floración temprana. El aumento en la tasa de crecimiento puede tener lugar en una o más etapas en el ciclo de vida de una planta o durante sustancialmente el ciclo de vida completo de la planta. Una tasa de crecimiento aumentada durante las etapas tempranas en el ciclo de vida de una planta puede reflejar un vigor potenciado. El aumento en la tasa de crecimiento puede alterar el ciclo de la cosecha de una planta, lo que permite que las plantas se siembren más tarde y/o se cosechen más pronto de lo que de otra manera sería posible. Si la tasa de crecimiento se aumenta suficientemente, puede permitir la siembra adicional de semillas de la misma especie de la planta (por ejemplo siembra y cosecha de plantas de arroz, seguido de siembra y cosecha de plantas de arroz adicionales todas dentro de un periodo de crecimiento convencional). De manera similar, si la tasa de crecimiento se aumenta suficientemente, puede permitir la siembra adicional de semillas de diferentes especies de plantas (por ejemplo, la siembra y cosecha de plantas de maíz seguidas de, por ejemplo, la siembra y cosecha opcional de plantas de soja, patata o cualquier otra planta adecuada). Los tiempos de cosecha adicionales del mismo rizoma en el caso de algunas plantas de cultivo también pueden ser posibles. Alterar el ciclo de cosecha de una planta puede conducir a un aumento en la producción de biomasa anual por acre (debido a un aumento en el número de tiempos (digamos en un año) que se puede cultivar y cosechar cualquier planta particular). Un aumento en la tasa de crecimiento también puede permitir el cultivo de plantas transgénicas en un área geográfica más amplia que sus homólogos silvestres, debido a que las limitaciones territoriales para cosechar un cultivo a menudo se determinan por condiciones ambientales adversas en el momento de plantar (estación temprana) o en el momento de la recolección (estación tardía). Dichas condiciones adversas pueden evitarse si se acorta el ciclo de cosecha. La tasa de crecimiento puede determinarse derivando diversos parámetros a partir de las curvas de crecimiento, dichos parámetros pueden ser: T-Medio (el tiempo que tarda una planta en alcanzar el 50 % de su tamaño máximo) y T-90 (el tiempo que tarda una planta en alcanzar el 90 % de su tamaño de máximo), entre otros.
- 60 La realización de los procedimientos de la divulgación proporciona plantas que tienen una tasa de crecimiento aumentada. Por tanto, de acuerdo con la presente divulgación se proporciona un procedimiento para aumentar la tasa de crecimientos en plantas cultivadas en condiciones de disponibilidad reducida de nitrógeno respecto a las

correspondientes plantas silvestres, en el que el procedimiento comprende modular la expresión en una planta de un ácido nucleico como se define en la reivindicación 1 que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo.

Un aumento en el rendimiento y/o la tasa de crecimiento se produce cuando la planta está en condiciones de cero estrés o cuando la planta está expuesta a diversos tipos de estrés en comparación con las plantas silvestres correspondientes en condiciones comparables. Las plantas con condiciones de crecimiento óptimas (cultivadas en condiciones de estrés cero) normalmente producen, en orden creciente de preferencia, al menos un 90 %, 87 %, 85 %, 83 %, 80 %, 77 % o 75 % de la producción promedio de dicha planta en un ambiente dado. La producción promedio de dicha planta se puede calcular en base a la recolección y/o la estación y/o localización. Los expertos en la técnica conocen las producciones de rendimiento promedio de un cultivo. Típicamente las plantas responden a exposición frente al estrés creciendo más lentamente. En condiciones de estrés intenso, la planta puede incluso detener por completo el crecimiento. Por otro lado, estrés leve se define en el presente documento como cualquier estrés al que se expone una planta que no da como resultado el cese total del crecimiento de la planta sin la capacidad de reanudar el crecimiento. El estrés leve conduce a una reducción en el crecimiento de la planta sometida a estrés de menos de 40 %, 35 % o 30 %, preferentemente menos del 25 %, 20 % o 15 %, más preferentemente menos del 14 %, 13 %, 12 %, 11 % o 10% o menos en comparación con la planta de control en condiciones de estrés cero. Debido a los avances en las prácticas agrícolas (riego, fertilización, tratamientos con pesticidas), las formas de estrés intenso no se encuentran frecuentemente en plantas cultivadas. Como una consecuencia, el crecimiento comprometido inducido por estrés leve es a menudo una característica indeseable para la agricultura. El estrés leve (como se usa en el presente documento) es un estrés abiótico diario (ambiental) al cual está expuesta la planta. El estrés abiótico puede deberse a sequía o a exceso de agua, estrés anaeróbico, estrés salino, toxicidad química, estrés oxidativo y temperaturas altas, frías o heladas, exceso o reducida disponibilidad de nutrientes (macroelementos y/o microelementos). El estrés abiótico puede ser un estrés osmótico provocado por estrés al agua (particularmente debido a sequía), estrés salino, estrés oxidativo y estrés iónico. El estrés biótico es típicamente aquel estrés provocado por patógenos, tales como bacterias, virus, hongos, nematodos e insectos. La expresión condiciones de "estrés cero" como se usa en el presente documento son las condiciones ambientales que permiten el crecimiento de las plantas.

La realización de los procedimientos como se describe en el presente documento da plantas cultivadas en condiciones de estrés cero o en condiciones de estrés abiótico leve que tienen un rendimiento aumentado con respecto a las correspondientes plantas silvestres.

La realización de los procedimientos como se describe de acuerdo con los resultados de la memoria descriptiva en plantas que tienen una tolerancia aumentada al estrés abiótico. Como describen Wang et al. (Planta (2003) 218: 1 - 14), el estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan adversamente al crecimiento de las plantas y a su productividad. Se sabe que la sequía, la salinidad, las temperaturas extremas y el estrés oxidativo se interconectan y que pueden inducir daño celular y en el crecimiento a través de mecanismos similares. Por ejemplo, la sequía y/o salinización se manifiestan principalmente como estrés osmótico, que produce la interrupción de la homeostasis y la distribución de iones en la célula. El estrés oxidativo, que frecuentemente acompaña a estrés por temperaturas altas o bajas, por salinidad o por sequía, puede producir desnaturalización de las proteínas funcionales y estructurales. Como una consecuencia, este diverso estrés ambiental a menudo activa rutas de señalización celular y respuestas celulares similares, tales como la producción de proteínas de estrés, regulación por aumento de antioxidantes, acumulación de solutos compatibles y paralización del crecimiento.

Diversos estreses ambientales activan rutas similares para el crecimiento de las plantas en condiciones de estrés por sequía y estrés salino. Estos ejemplos deben verse como una selección para indicar la afectación de los polipéptidos HDZip hox5 de clase I u homólogos de los mismos en el incremento de la tolerancia al estrés abiótico en general. Se ha indicado un grado particularmente alto de "interferencia" entre estrés por sequía y estrés por alta salinidad (Rabbani et al. (2003) Plant Physiol 133: 1755 - 1767). Por tanto, sería evidente que un polipéptido HDZip hox5 de clase I o un homólogo del mismo, junto con su utilidad en el incremento de la tolerancia a la sequía y la tolerancia a la sal en plantas, también será útil en la protección de la planta contra otros diversos estreses abióticos.

Por la expresión "estrés abiótico" como se define en el presente documento se entiende uno cualquiera o más de: estrés hídrico (debido a sequía o exceso de agua), estrés anaeróbico, estrés salino, estrés por temperatura (debido calor, frío o temperaturas de congelación), estrés por toxicidad química y estrés oxidativo. De acuerdo con un aspecto de la especificación, el estrés abiótico es un estrés osmótico, seleccionado de estrés hídrico, estrés salino, estrés oxidativo y estrés iónico. Preferentemente, el estrés hídrico es estrés por sequía. La expresión estrés salino no está restringido al estrés resultante de un exceso de sal común (NaCl), pero puede ser de uno o más de: NaCl, KCl, LiCl, MgCl₂, CaCl₂, entre otros.

Otro ejemplo de estrés ambiental abiótico se reduce 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 en el rendimiento de la planta y la calidad del producto, una enorme cantidad de fertilizante se vierte sobre campos para optimizar el crecimiento de la planta y la calidad. La productividad de las plantas normalmente está limitada por tres nutrientes principales, fósforo, potasio y nitrógeno, que suele ser el elemento

limitante de la velocidad en el crecimiento de plantas de estos tres. Por tanto, el elemento nutricional importante requerido para el crecimiento de la planta es nitrógeno (N). Es un constituyente de numerosos compuestos importantes que se encuentran en las células vivas, incluyendo aminoácidos, proteínas (enzimas), ácidos nucleicos, y clorofila. Del 1,5 % al 2 % de la materia seca de la planta es nitrógeno y aproximadamente el 16 % de la proteína total de la planta. Por lo tanto, la disponibilidad de nitrógeno es un factor fundamental limitante para el crecimiento y producción de plantas de cultivos (Frink et al. (1999) Proc Natl Acad Sci USA 96(4): 1175 - 1180), y tiene así un impacto importante en la acumulación de proteínas y la composición de aminoácidos. Por lo tanto, son de gran interés las plantas de cultivo con un rendimiento aumentado cuando se cultivan en condiciones de limitación de nutrientes, preferentemente condiciones limitantes de nitrógeno.

La realización de los procedimientos de la memoria descriptiva da lugar a plantas que tienen un rendimiento aumentado cuando se cultivan en condiciones de disponibilidad reducida de nutrientes con respecto a las correspondientes plantas silvestres. Por tanto, de acuerdo con la presente invención se proporciona un procedimiento para aumentar el rendimiento en plantas cultivadas en condiciones de disponibilidad reducida de nutrientes, preferentemente disponibilidad reducida de nitrógeno, respecto a las correspondientes plantas silvestres, en el que el procedimiento comprende modular la expresión en una planta de un ácido nucleico que tiene una secuencia como se define en la reivindicación 1 que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo.

El aumento de la tolerancia al estrés abiótico se manifiesta por las plantas con un mayor rendimiento, en relación a las correspondientes plantas silvestres. En particular, dicho rendimiento aumentado puede incluir uno o más de los siguientes: aumento del número total de semillas, aumento del número de semillas llenas, aumento del rendimiento total de las semillas, aumento del número de flores por panícula, aumento de la tasa de llenado de las semillas, aumento del IR, aumento del TKW, aumento de la longitud de la raíz o aumento del diámetro de la raíz, cada uno respecto a las correspondientes plantas silvestres. Preferentemente, el aumento de la tolerancia al estrés abiótico es la tolerancia aumentada a la reducción de la disponibilidad de nutrientes, más preferiblemente la tolerancia aumentada a la reducción de la disponibilidad de nitrógeno.

Ventajosamente, la realización de los procedimientos de la invención da lugar a plantas que tienen un índice de verdor aumentado en condiciones de disponibilidad reducida de nutrientes con respecto a las correspondientes plantas silvestres. El índice de verdor se calcula a partir de las imágenes digitales de las plantas. Para cada píxel que pertenece al objeto de la planta sobre la imagen se calcula la proporción del valor verde frente al valor rojo (en el modelo RGB por el que se codifica el color). El índice de verdor como se define en el presente documento se expresa como el porcentaje de píxeles para el que la proporción de verde con respecto a rojo supera un umbral determinado. Un aumento del índice de verdor puede indicar una senescencia reducida o retardada que a su vez permite la prolongación de la actividad fotosintética de una planta, que a su vez conduce a varios efectos beneficiosos bien conocidos en la técnica.

La realización de los procedimientos de la invención da lugar a plantas que tienen un índice de verdor aumentado en condiciones de disponibilidad reducida de nutrientes con respecto a las correspondientes plantas silvestres. Por tanto, de acuerdo con la presente invención se proporciona un procedimiento para aumentar el índice de verdor en plantas cultivadas en condiciones de disponibilidad reducida de nitrógeno respecto a las correspondientes plantas silvestres, en el que el procedimiento comprende modular la expresión en una planta de un ácido nucleico como se establece en la reivindicación 1 que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo. Preferiblemente, las condiciones de disponibilidad reducida de nutrientes son condiciones de disponibilidad reducida de nitrógeno.

Rabbani et al. (2003, Plant Physiol 133: 1755 - 1767) indican que existen mecanismos moleculares similares de la tolerancia al estrés y respuestas entre dicotiledóneas y monocotiledóneas. Los procedimientos de la invención son, por lo tanto, ventajosamente aplicables a cualquier planta

El término "planta" como se usa en el presente documento incluye plantas completas, ancestros y progenie de las plantas y partes de las plantas, incluyendo semillas, brotes, tallos, hojas, raíces (incluidos tubérculos), flores y tejidos y órganos, donde cada uno de los elementos mencionados anteriormente comprende el gen/secuencia de ácido nucleico de interés. El término "planta" también abarca células vegetales, cultivos en suspensión, tejido de callo, embriones, regiones meristemáticas, gametofitos, esporofitos, polen y microesporas, de nuevo donde cada uno de los elementos mencionados anteriormente comprende el gen/secuencia ácido nucleico de interés. Por lo tanto, el término "planta" tal como se usa en el presente documento abarca una planta, parte de planta (incluidas las semillas), o célula vegetal, en el que cada uno de los mencionados anteriormente comprende una secuencia de ácido nucleico recombinante que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo.

Las plantas que son particularmente útiles en los procedimientos de la invención incluyen todas las plantas que pertenecen a la superfamilia Viridiplantae, en particular plantas monocotiledóneas y dicotiledóneas incluyendo forraje o 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., *Artocarpus* spp., *Asparagus officinalis*, *Avena* spp. (p. ej., *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua*

var. sativa, *Avena hybrida*), *Averrhoa carambola*, *Bambusa* spp., *Benincasa hispida*, *Bertholletia excelsa*, *Beta vulgaris*, *Brassica* spp. (p. ej., *Brassica napus*, *Brassica rapa* ssp. [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*,
 5 *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* (p.ej., *Elaeis guineensis*, *Elaeis oleifera*), *Eleusine coracana*, *Erianthus* spp., *Eriobotrya japonica*, *Eucalyptus* spp., *Eugenia uniflora*, *Fagopyrum* spp., *Fagus* spp., *Festuca arundinacea*, *Ficus carica*, *Fortunella* spp., *Fragaria* spp., *Ginkgo biloba*, *Glycine* spp. (p. ej.,
 10 *Glycine max*, *Soja hispida* or *Soja max*), *Gossypium hirsutum*, *Helianthus* spp. (p. ej., *Helianthus annuus*), *Hemerocallis fulva*, *Hibiscus* spp., *Hordeum* spp. (p. ej., *Hordeum vulgare*), *Ipomoea batatas*, *Juglans* spp., *Lactuca sativa*, *Lathyrus* spp., *Lens culinaris*, *Linum usitatissimum*, *Litchi chinensis*, *Lotus* spp., *Luffa acutangula*, *Lupinus* spp., *Luzula sylvatica*, *Lycopersicon* spp. (p. ej., *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon pyriforme*), *Macrotyloma* spp., *Malus* spp., *Malpighia emarginata*, *Mammea americana*, *Mangifera indica*, *Manihot* spp., *Manilkara zapota*, *Medicago sativa*, *Melilotus* spp., *Mentha* spp., *Miscanthus sinensis*, *Momordica* spp., *Morus nigra*, *Musa* spp., *Nicotiana* spp., *Olea* spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. (p. ej., *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.,
 15 *Solanum* spp. (p. ej., *Solanum tuberosum*, *Solanum integrifolium* o *Solanum lycopersicum*), *Sorghum bicolor*, *Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica*, *Theobroma cacao*, *Trifolium* spp., *Triticosecale rimpai*, *Triticum* spp. (p. ej., *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum* or *Triticum vulgare*), *Tropaeolum minus*, *Tropaeolum majus*, *Vaccinium* spp., *Vicia* spp.,
 20 *Vigna* spp., *Viola odorata*, *Vitis* spp., *Zea mays*, *Zizania palustris*, *Ziziphus* spp., amaranto, alcachofa, espárrago, brécol, coles de Bruselas, repollo, canola, zanahoria, coliflor, apio, berza, lino, col rizada, lentejas, colza, okra, cebolla, patata, arroz, soja, fresa, remolacha azucarera, caña de azúcar, girasol, tomate, calabacín, té y algas, entre otros. De acuerdo con una realización preferida del procedimiento de la invención, la planta es una planta de cultivo
 25 tal como soja, girasol, canola, alfalfa, colza, algodón, tomate, patata o tabaco. Adicionalmente preferentemente, la planta es una planta monocotiledónea, tal como caña de azúcar. Más preferentemente la planta es un cereal, tal como arroz, maíz, trigo, cebada, triticale, mijo, centeno, sorgo o avena.

El término "polipéptido HDZip hox5 de clase I u homólogo del mismo" de acuerdo con la memoria descriptiva se refiere a un polipéptido que comprende de N-terminal a C-terminal: (i) una caja ácida; y (ii) un homeodominio de
 35 clase I; y (iii) una cremallera de leucina con más de 5 héptadas.

Adicionalmente, el polipéptido hox5 HDZip de la clase I o un homólogo del mismo como se describe en el presente documento puede comprender una cualquiera o ambas de las siguientes: (a) una cola DE Trp; y (b) el motivo de aminoácidos RPFF, en el que R es Arg, Pro P y F Phe. El motivo de (b) precede a la caja ácida, al examinar la proteína desde el N-terminal al C-terminal.

40 Un ejemplo de un polipéptido HDZip hox5 de clase como se ha definido anteriormente en el presente documento que comprende desde N-terminal a C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas; y, adicionalmente, que comprende: (a) una cola DE Trp; y (b) el motivo de aminoácidos RPFF, en el que R es Arg, Pro P y F Phe, está representado como en la SEC ID N° 2. Adicionalmente, dichos ejemplos se proporcionan en la Tabla A del ejemplo 1 en el presente documento.

45 Un polipéptido HDZip hox5 de clase I u homólogo del mismo como se describe en el presente documento está codificado por un gen/secuencia de ácido nucleico de HDZip hox5 de clase I. Por lo tanto el término "gen/secuencia de ácido nucleico de HDZip hox5 de clase I" tal como se define en el presente documento es cualquier gen / secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo como se define anteriormente en el presente documento.

50 Los polipéptidos HDZip hox5 de clase I u homólogos de los mismos pueden identificarse fácilmente utilizando técnicas de rutina bien conocidas en la técnica, tales como por la alineación de secuencias. Los procedimientos para el alineamiento de las secuencias para comparación son bien conocidos en la técnica, dichos procedimientos incluyen GAP, BESTFIT, BLAST, FASTA y TFASTA. GAP utiliza el algoritmo de Needleman y Wunsch ((1970) J Mol Biol 48: 443 - 453) para encontrar el alineamiento de las dos secuencias completas que maximizan el número de emparejamientos y minimiza el número de espacios. El algoritmo BLAST (Altschul et al. (1990) J Mol Biol 215: 403 -
 55 10) calcula el porcentaje de identidad de secuencia y desarrolla un análisis estadístico de la similitud entre las dos secuencias. El software para realizar el análisis BLAST está públicamente disponible a través del National Centre for Biotechnology Information. Los homólogos de los polipéptidos HDZip hox5 de clase I que comprenden un homeodominio de clase I y una cremallera de leucina con más de 5 héptadas pueden identificarse fácilmente
 60 utilizando, por ejemplo, el algoritmo de alineamiento de secuencia múltiple ClustalW (versión 1.83) , con los parámetros de alineamiento por pares predeterminados, y un procedimiento de puntuación en porcentaje. Puede realizarse una edición manual menor para optimizar el alineamiento entre los motivos conservados, como sería obvio

para un experto en la técnica (véase el ejemplo 2 y la figura 2 en el presente documento)

Los diferentes dominios estructurales en un polipéptido HDZip hox5 de clase I, tal como el homeodominio y la cremallera de leucina, se pueden identificar usando bases de datos especializadas por ejemplo, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857 - 5864; Letunic et al. (2002) Nucl Acids Res 30, 242 - 244; alojados por EMBL en Heidelberg), 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. (En) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53 - 61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)) o Pfam (Bateman et al., Nucl Acids Res 30(1):276 - 280 (2002)). La predicción de la cremallera de leucina y la identificación de héptadas pueden realizarse usando software especializado, tal como ZZIP, que combina un algoritmo de predicción de bobina en espiral estándar con una búsqueda aproximada para la repetición de la leucina característica (Bornberg-Bauer et al. (1998) Computational Approaches to Identify Leucine Zippers, Nucl Acids Res, 26(11): 2740 - 2746). Los resultados de la identificación de dominios de las secuencias del polipéptido HDZip hox5 de clase I se presentan en el Ejemplo 4 de esta solicitud.

Además, la presencia de una caja ácida puede también identificarse fácilmente. La composición primaria de aminoácidos (en %) para determinar si un dominio de polipéptido es rico en aminoácidos específicos se puede calcular usando programas de software del servidor ExPASy, en particular, la herramienta ProtParam (Gasteiger E et al. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784 - 3788). La composición de la proteína de interés puede compararse después con la composición de aminoácidos media (en %) en el banco de datos de secuencias de proteínas Swiss-Prot. Dentro de este banco de datos, el contenido promedio de Asp (D) y Glu (E) es del 5,3 % y del 6,6 % respectivamente, siendo el promedio combinado de 11,9 %. Como ejemplo, la caja ácida de la SEC ID N° 2 comprende el 9,1 % de D y el 54,5 % de E, siendo el promedio combinado de 63,6 % (véase el Ejemplo 4 en el presente documento). Como se define en el presente documento, una caja rica en ácido tiene un contenido combinado de Asp (D) y Glu (E) contenido (en términos de %) por encima de lo que se encuentran en la composición promedio de aminoácidos (en términos de %) de las proteínas en la base de datos de secuencias proteicas Swiss-Prot. Una caja ácida puede ser parte de un dominio de activación de la transcripción. Los dominios de activación de la transcripción eucariótica se han clasificado en función de su contenido de aminoácidos, y las principales categorías incluyen dominios de activación ácidos ricos en glutamina y ricos en prolina. (2005) Plant J. 43(5): 769 - 88, y referencias citadas en la misma).

Un número seleccionado de polipéptidos entre los polipéptidos HDZip hox5 de clase I u homólogos de los mismos como se describe en el presente documento comprende además el motivo de aminoácidos RPFF, en el que R es Arg, Pro P y F Phe. Este motivo precede a la caja ácida, al examinar la proteína desde el N -terminal al C-terminal (véase la Figura 2). La presencia del RPFF se pueden identificar usando procedimientos para la alineación de secuencias para la comparación como se ha descrito anteriormente en el presente documento. En algunos casos, los parámetros por defecto se pueden ajustar para modificar la rigurosidad de la búsqueda. Por ejemplo utilizando BLAST, el umbral de significación estadística (denominado valor "previsto") para indicar las coincidencias contra secuencias de bases de datos se puede aumentar para mostrar coincidencias menos estrictas. De esta manera, se pueden identificar las coincidencias cortas casi exactas.

Un número seleccionado de polipéptidos contra los polipéptidos HDZip hox5 de clase I u homólogos de los mismos tal como se describe en el presente documento puede comprender además una cola de Trp. Una cola de Trp como se define en el presente documento es los últimos 10 aminoácidos del C-terminal del polipéptido que comprende al menos un residuo de Trp (véase la Figura 2).

Ejemplos de polipéptidos HDZip hox5 de clase I u homólogos de los mismos, como se describe en el presente documento (codificados por la secuencia de polinucleótido con número de acceso entre paréntesis) se dan en la Tabla A de los ejemplos.

Se ha de entender que las secuencias como se describen en el presente documento que entran en la definición de "polipéptido HDZip hox5 de clase I u homólogo del mismo" no se limitan a las secuencias proporcionadas en la Tabla A, pero que cualquier polipéptido que comprende desde N-terminal a C- terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas, puede ser adecuado para su uso en la realización de los procedimientos de la memoria descriptiva.

Los polipéptidos HDZip hox5 de clase I u homólogos de los mismos tienen actividad de unión a ADN, preferiblemente a semisitos de 5 pb que se superponen en una posición central, CAA (A / T) ATTG, como se detecta en los ensayos de un híbrido en levaduras (Meijer et al. (2000) Mol Gen Genet 263:12 - 21). En ensayos transitorios en suspensiones celulares de arroz, el cobombardeo de un polipéptido HDZip hox5 de clase I con el gen indicador GUS tuvo como resultado un aumento del número de puntos teñidos, que también fueron más intensos de color (Meijer et al, ant.). Este ensayo es útil para demostrar la función de activador de la los polipéptidos HDZip hox5 de clase I u homólogos hox5

Ejemplos de secuencias de ácidos nucleicos de HDZip hox5 de clase I incluyen, pero no se limitan a, los enumerados en la Tabla A de los Ejemplos. Los genes / secuencias de ácidos nucleicos de HDZip hox5 de clase I y

variantes de los mismos pueden ser adecuados en la práctica de los procedimientos de la memoria descriptiva. Las variantes de los genes/secuencias de ácido nucleico de HDZip hox5 incluyen porciones de un gen/secuencia de ácido nucleico de HDZip hox5 de clase I capaces de hibridar con un gen/secuencia de ácido nucleico de HDZip hox5 de clase I.

- 5 El término porción, como se define en el presente documento, se refiere a un segmento de ADN que codifica un polipéptido que comprende de N-terminal a C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Una porción puede prepararse, por ejemplo, realizando una o más deleciones en la secuencia de ácido nucleico de HDZip hox5 de clase I. Las porciones pueden usarse en forma aislada o pueden fusionarse con otras secuencias codificantes (o no codificantes) para, por ejemplo, producir una
- 10 proteína que combine varias actividades. Cuando se fusiona con otras secuencias codificantes, el polipéptido resultante producido después de la traducción puede ser más grande que el esperado para la porción HDZip hox5 de clase I. Preferentemente, la porción es una porción de una secuencia de ácido nucleico como se representa en una cualquiera de las secuencias enumeradas en la Tabla A del Ejemplo 1 en el presente documento. Lo más preferentemente, la porción de una secuencia de ácido nucleico como se representa en la SEC ID N° 1.
- 15 Otra variante de un gen/ secuencia de ácido nucleico de HDZip hox5 de clase I es una secuencia de ácido nucleico capaz de hibridar en condiciones de rigurosidad reducida, preferiblemente en condiciones rigurosas, con un gen/ secuencia de ácido nucleico de HDZip hox5 de clase I como se ha definido anteriormente en el presente documento, en el que la hibridación de la secuencia codifica un polipéptido que comprende desde N-terminal a C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas.
- 20 Preferentemente, la secuencia de hibridación es una que puede hibridarse con una secuencia ácido nucleico como se representa mediante una cualquiera de las secuencias de ácido nucleico enumeradas en la Tabla A del ejemplo 1 en el presente documento o con una porción de cualquiera de las secuencias mencionadas anteriormente como se ha definido anteriormente en el presente documento. Más preferentemente, la secuencia de hibridación es una que es capaz de hibridarse con una secuencia de ácido nucleico representada por la SEC ID N° 1.
- 25 El término "hibridación" como se define en el presente documento es un proceso en el que las secuencias de nucleótidos complementarias sustancialmente homólogas hibridan entre sí. El proceso de hibridación puede producirse completamente en la solución, es decir, ambas moléculas de ácidos nucleicos complementarios están en la solución. El proceso de hibridación también puede producirse con una de las moléculas de ácidos nucleicos complementarias inmovilizada en una matriz tal como perlas magnéticas, perlas de Sefarosa o cualquier otra resina.
- 30 El proceso de hibridación puede producirse adicionalmente con una de las moléculas de ácido nucleico complementaria inmovilizada en un soporte sólido, tal como una membrana de nitrocelulosa o de nailon, o inmovilizada mediante, por ejemplo, fotolitografía a, por ejemplo, un soporte vítreo silíceo (el último conocido como matrices o micromatrices de ácido nucleico o como microplacas de ácido nucleico). Con el fin de permitir que se produzca la hibridación, las moléculas de ácido nucleico de manera general se desnaturalizan térmica o
- 35 químicamente para fundir una doble cadena en dos cadenas únicas y/o para retirar horquillas u otras estructuras secundarias de ácidos nucleicos de cadena sencilla. La rigurosidad de la hibridación está influenciada por condiciones tales como temperatura, concentración de sales, fuerza iónica y composición del tampón de hibridación.

"Condiciones de hibridación rigurosas" y "condiciones de lavado de hibridación rigurosas" en el contexto de experimentos de hibridación de ácidos nucleicos como hibridaciones de tipo Southern y Northern son dependientes de la secuencia y son diferentes con diferentes parámetros ambientales. El experto en la materia conoce los diversos parámetros que pueden alterarse durante la hibridación y el lavado y que, o bien mantener o cambiar las condiciones de rigurosidad.

La T_m es la temperatura a la fuerza iónica y el pH definidos, a la cual el 50 % de una secuencia diana hibrida con una sonda perfectamente apareada. La T_m depende de las condiciones de solución y de la composición base y la longitud de la sonda. Por ejemplo, las secuencias más largas hibridan específicamente a temperaturas más altas. El índice máximo de hibridación se obtiene de aproximadamente 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 cadenas de ácido nucleico, de modo que se estimula la formación del híbrido; este efecto es visible para concentraciones de sodio de hasta 0,4 M. La formamida reduce la temperatura de fusión de los dúplex ADN-ADN y ADN-ARN con 0,6 a 0,7 °C para cada porcentaje de formamida y la adición de formamida al 50 % permite que se realice la hibridación a de 30 a 45 °C, aunque disminuirá el índice de hibridación. Los emparejamientos erróneos de los pares de bases reducen el índice de hibridación y la estabilidad térmica de los dúplex. De media y para sondas grandes, la T_m disminuye aproximadamente 1 °C por % de emparejamiento erróneo de 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):

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

2. Híbridos de ADN-ARN o ARN-ARN

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

3. Híbridos de oligo-ADN u oligo ARN^d

Para <20 nucleótidos: $T_m = 2 (\ln)$

Para 20-25 nucleótidos: $T_m = 22 + 1,46 (\ln)$

^a o para otro catión monovalente, pero solo exacto en el intervalo 0,01-0,4 M.

5 ^b solo es exacto para % GC en el intervalo de 30 % a 75 %.

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

^d oligo, oligonucleótido; \ln = longitud eficaz del cebador = $2x$ (nº de G/C) + (nº de A/T).

Nota: para cada 1 % de formamida, la T_m se reduce en aproximadamente 0,6 a 0,7 °C, mientras que la presencia de urea 6 M reduce la T_m en aproximadamente 30 °C.

10 La especificidad de la hibridación es típicamente función de los lavados posthibridación. Para retirar el fondo resultante de la hibridación no específica, las muestras se lavan con soluciones salinas diluidas. Los factores críticos de dichos lavados incluyen la fuerza iónica y la temperatura de la solución de lavado final: cuanto menor es la concentración de sal y mayor es la temperatura de lavado, mayor es la rigurosidad de lavado. Las condiciones de lavado se realizan típicamente a o por debajo de la rigurosidad de hibridación. Generalmente, las condiciones de

15 rigurosidad adecuadas para los ensayos de hibridación de ácido nucleico o los procedimientos de detección de amplificación de genes son como se han establecido anteriormente. También pueden seleccionarse condiciones más o menos rigurosas. Generalmente, se seleccionan condiciones de baja rigurosidad de aproximadamente 50 °C menor que el punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y pH definidos. Las condiciones de rigurosidad media son cuando la temperatura es 20 °C menor que la T_m , y las condiciones de alta

20 rigurosidad son cuando la temperatura es 10 °C menor que la T_m . Por ejemplo, las condiciones rigurosas son las que son por lo menos tan rigurosas como, por ejemplo, las condiciones de A-L; y las condiciones de rigurosidad reducida son al menos tan rigurosas como, por ejemplo, las condiciones de M-R. La unión no específica se puede controlar mediante uno cualquiera de un número de técnicas conocidas tales como, por ejemplo, bloquear la membrana con soluciones que contienen proteínas, adiciones de ARN heterólogo, ADN, y SDS al tampón de hibridación, y tratamiento con RNasa. Ejemplos de condiciones de hibridación y de lavado se enumeran en la Tabla 1 dada a

25 continuación.

Tabla 1: Ejemplos de condiciones de hibridación y de lavado

Condiciones de rigurosidad	Híbrido polinucleotídico [±]	Longitud del híbrido (pb)n ^{o±}	Temperatura de hibridación y tampón [†]	Temperatura de lavado y tampón [†]
A	ADN: ADN	> o igual a 50	65° C 1 × SSC; o 42° C, 1 × SSC y 50 % de formamida	65° C; 0,3 × SSC
B	ADN: ADN	< 50	Tb*; 1 × SSC	Tb*; 1 × SSC
C	ADN: ARN	> o igual a 50	67° C 1 × SSC; o 45° C, 1 × SSC y 50 % de formamida	67° C; 0,3 × SSC
D	ADN: ARN	< 50	TD*; 1 × SSC	TD*; 1 × SSC
E	ARN: ARN	> o igual a 50	70° C 1 × SSC; o 50° C, 1 × SSC y 50 % de formamida	70° C; 0,3 × SSC
F	ARN: ARN	< 50	Tf*; 1 × SSC	Tf*; 1 × SSC
G	ADN: ADN	> o igual a 50	65° C 4 × SSC; o 45° C, 4 × SSC y 50 % de formamida	65° C; 1×SSC
H	ADN: ADN	< 50	Th*; 4 ×SSC	Th*; 4 x SSC
I	ADN: ARN	> o igual a 50	67° C 4 × SSC; o 45° C, 4 × SSC y 50 % de formamida	67° C; 1 × SSC
J	ADN: ARN	< 50	Tj*; 4 × SSC	Tj*; 4 × SSC

Condiciones de rigurosidad	Híbrido polinucleotídico [‡]	Longitud del híbrido (pb)n [‡]	Temperatura de hibridación y tampón [†]	Temperatura de lavado y tampón [†]
K	ARN: ARN	> o igual a 50	70° C 4 × SSC; o 40° C, 6 × SSC y 50 % de formamida	67° C; 1 × SSC
L	ARN: ARN	< 50	Tl*; 2 × SSC	Tl*; 2 × SSC
M	ADN: ADN	> o igual a 50	50° C 4 × SSC; o 40° C, 6 × SSC y 50 % de formamida	50° C; 2 × SSC
N	ADN: ADN	< 50	Tn*; 6 × SSC	Tn*; 6 × SSC
O	ADN: ARN	> o igual a 50	55° C 4 × SSC; o 42° C, 6 × SSC y 50 % de formamida	55° C; 2 × SSC
P	ADN: ARN	<50	Trp*; 6 × SSC	Trp*; 6 × SSC
Q	ARN: ARN	> o igual a 50	60° C 4 × SSC; o 45° C, 6 × SSC y 50 % de formamida	60° C.; 2 × SSC
R	ARN: ARN	< 50	Tr*; 4 × SSC	Tr*; 4 × SSC

[‡] La "longitud del híbrido" es la duración prevista para la secuencia de ácido nucleico de hibridación. Cuando los ácidos nucleicos de la secuencia conocida se hibridan, la longitud del híbrido puede determinarse alineando las secuencias e identificando las regiones conservadas descritas en el presente documento.

[†] SSPE (1× SSPE es NaCl 0,15 M , NaH₂PO₄ 10 mM y EDTA 1,25 mM, pH 7,4) puede sustituirse por SSC (1 × SSC es NaCl 0,15M y citrato sódico 15 mM) en los tampones de hibridación y lavado; los lavados se realizan durante 15 minutos una vez completada la hibridación. Las hibridaciones y lavados pueden incluir adicionalmente reactivo de Denhard 5x, SDS 0,5-1,0 %, 100 µg/ml de ADN de espera de salmón fragmentado, desnaturalizado, pirofosfato sódico al 0, 5 % y formamida al 50%.

* Tb-Tr: La temperatura de hibridación para híbridos que se ha previsto que tengan una longitud inferior a 50 pares de bases debe ser de 5 - 10 °C menor que la temperatura de fusión T_m de los híbridos; la T_m se determina de acuerdo a las ecuaciones mencionadas anteriormente.

[‡] La presente memoria descriptiva también abarca la sustitución de una cualquiera, o más de parejas de híbridos de ADN o ARN con un PNA, o un ácido nucleico modificado.

Con objeto de definir el nivel de rigurosidad, se puede hacer 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).

- 5 La secuencia del ácido nucleico de HDZip hox5 de clase I puede derivar de cualquier fuente natural o artificial. El gen/secuencia de ácido nucleico puede aislarse de una fuente microbiana, como levaduras u hongos, o de una planta, alga o animal (incluido el ser humano). Este secuencia ácido nucleico puede modificarse con respecto a su forma nativa en composición y/o en medio genómico a través de manipulación humana deliberada. La secuencia de ácido nucleico es de origen vegetal, ya sea de la misma especie de planta (por ejemplo, una en la que se va a introducir) o de una especie de planta diferente. Preferentemente, la secuencia de ácido nucleico se puede aislar de una especie monocotiledónea, además preferentemente de la familia Poaceae, más preferentemente del género *Oryza*, lo más preferentemente de *Oryza sativa*. Más preferentemente, la secuencia de ácido nucleico de HDZip hox5 de clase I aislada de *Oryza sativa* está representada por la SEC ID N° 1 y la secuencia polipeptídica de de HDZip hox5 de clase I es como la representada por la SEC ID N° 2.
- 10
- 15 La expresión de una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo puede modularse mediante la introducción de una modificación genética, por uno cualquiera (o más) de los siguientes procedimientos: activación de ADN-T, TILLING, mutagénesis dirigida a sitio, evolución dirigida y recombinación homóloga o introduciendo y expresando en una planta una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo. Después de la introducción de la modificación genética, viene una etapa de selección de la expresión modificada de una secuencia ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo, cuya modificación en la expresión da lugar a plantas con rendimiento aumentado en condiciones de disponibilidad reducida de nutrientes con respecto a las
- 20

correspondientes plantas silvestres.

La activación de ADN-T (Hayashi et al. Science (1992) 1350 - 1353) implica la inserción de T-DNA, que normalmente contiene un promotor (también puede ser un potenciador de la traducción o un intrón), en la región genómica del gen de interés o 10 kb aguas arriba o abajo de la región codificante de un gen en una configuración tal que el promotor dirige la expresión del gen diana. Típicamente, la regulación de la expresión del gen diana mediante su promotor natural se interrumpe y el gen cae bajo el control del promotor recién introducido. El promotor normalmente está embebido en un ADN-T. Este ADN-T se inserta al azar en el genoma de la planta, por ejemplo, a través de infección por infección con *Agrobacterium* y conduce a la expresión modificada de genes cerca al ADN-T insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debido a la sobreexpresión de los genes cercanos al promotor introducido. El promotor a introducir puede ser cualquier promotor capaz de dirigir la expresión de un gen en el organismo deseado, en este caso una planta. Por ejemplo, los promotores constitutivos preferidos de tejido, preferidos de tipo celular e inducibles son todos adecuados para uso en la activación de T-ADN.

También se puede introducir una modificación genética en el locus de un gen de HDZip hox5 de clase I utilizando la técnica de TILLING (Inducción Dirigida de Lesiones Locales en el Genoma). Es una tecnología de mutagénesis que es útil para generar y/o identificar, y en última instancia aislar, variantes mutagenizadas de una secuencia de ácido nucleico HDZip hox5 de clase I capaces de exhibir actividad de HDZip hox5 de clase I. El TILLING también permite la selección de plantas que llevan dichas variantes mutantes. Estas variantes mutantes pueden presentar mayor actividad de HDZip hox5 de clase I que la exhibida por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de detección de alto rendimiento. Las etapas que típicamente se siguen en el TILLING son: (a) mutagénesis EMS (Redei GP y Koncz C (1992) En Methods in Arabidopsis Research, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, pp. 16 - 82; Feldmann et al., (1994) en Meyerowitz EM, Somerville CR, eds, Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137 - 172; Lightner J y Caspar T (1998) en J Martinez-Zapater, J Salinas, eds, Methods on Molecular Biology, Vol. 82. Humana Press, Totowa, NJ, pp 91 - 104); (b) preparación y agrupamiento de ADN en individuos; (c) amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) DHPLC, donde la presencia de un heterodúplex en un grupo se detecta como un pico adicional en el cromatograma; (f) identificación del mutante individual; y (g) secuenciación del producto PCR mutante. En la técnica se conocen bien procedimientos de TILLING (McCallum y col., (2000) Nat Biotechnol 18: 455-457; revisado por Stemple (2004) Nat Rev Genet 5 (2): 145-50) 145 - 50).

La recombinación homóloga permite la introducción en un genoma de una secuencia de ácido nucleico seleccionado en una posición definida seleccionada. La recombinación homóloga es una tecnología convencional utilizada rutinariamente en ciencias biológicas para organismos inferiores tales como levaduras o el musgo *Physcomitrella*. Se han descrito procedimientos para realizar recombinación homóloga en plantas no solo para el modelo de plantas (*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*). La secuencia de ácido nucleico que va a ser la diana (que puede ser una secuencia de ácido nucleico de HDZip hox5 de clase I o variante del mismo como se ha definido en el presente documento anteriormente) no tiene que dirigirse al locus de un gen de HDZip hox5 de clase I, pero puede introducirse en, por ejemplo, regiones de alta expresión. La secuencia de ácido nucleico que va a ser la diana puede ser un alelo mejorado utilizado para reemplazar el gen endógeno o puede introducirse además en el gen endógeno.

Puede usarse mutagénesis dirigida a sitio para generar variantes de las secuencias de ácidos nucleicos de HDZip hox5 de clase I. Se dispone de diversos procedimientos para realizar mutagénesis dirigida a sitio; siendo los más frecuentes los procedimientos basados en PCR (Current Protocols in Molecular Biology. Wiley Eds.).

También puede usarse evolución dirigida para generar variantes de las secuencias de ácidos nucleicos de HDZip hox5 de clase I. Esta consiste en iteraciones de barajado de ADN, seguido de detección y/o selección apropiada para generar variantes de las secuencias de ácido nucleico de HDZip hox5 de clase I o partes de las mismas que codifican polipéptidos de HDZip hox5 de clase I u homólogos o porciones de los mismos que tienen una actividad biológica modificada (Castle y col., (2004) Science 304 (5674): 1151 - 4; Patentes de Estados Unidos 5.811.238 y 6.395.547).

La activación de ADN-T, TILLING, recombinación homóloga, mutagénesis dirigida a sitio y la evolución dirigida son procedimientos para introducir una modificación genética para modular la expresión de una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo. Por tanto, de acuerdo con la presente memoria descriptiva se proporciona un procedimiento para modular la expresión de una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo, que comprende introducir una modificación genética mediante cualquiera de uno o más) de: activación de ADN-T, TILLING, recombinación homóloga, mutagénesis dirigida a sitio y evolución dirigida.

Un procedimiento preferido para la introducción de una modificación genética (que en este caso no tiene por qué estar en el locus de un gen de HDZip hox5 de clase I) es introducir y expresar en una planta de una secuencia de ácido nucleico que codifica para un polipéptido HDZip hox5 de clase I o un homólogo del mismo. Un polipéptido HDZip hox5 de clase I o un homólogo del mismo se define como un polipéptido que comprende de N-terminal a C-

terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. La secuencia de ácido nucleico a introducir en una planta puede ser una secuencia de ácido nucleico de longitud completa o puede ser una porción o una secuencia de hibridación como se ha definido anteriormente en el presente documento.

5 Los "homólogos" de una proteína incluyen péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones, delecciones 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 derivan. Para producir tales homólogos, los aminoácidos de la proteína pueden sustituirse por otros aminoácidos que tienen propiedades similares (tales como hidrofobicidad, hidrofiliidad, antigenicidad, propensión a formar o romper estructuras en α hélice o estructuras en lámina β similares). En la técnica se conocen bien las tablas de sustituciones conservadoras (véase, por ejemplo, Creighton (1984) Proteins. W.H. Freeman y Company y la Tabla 2 más adelante). Los homólogos útiles en los procedimientos de acuerdo con la invención son preferiblemente polipéptidos HDZip hox5 de clase I como se ha definido en el presente documento anteriormente.

15 También están abarcadas por el término "homólogos" dos formas especiales de homología, que incluyen secuencias ortólogas y secuencias parálogas, que abarcan conceptos evolutivos utilizados para describir las relaciones ancestrales de los genes. El término "parólogo" se refiere a duplicaciones de genes dentro del genoma de una especie que conduce a genes parálogos. El término "ortólogo" se refiere a genes homólogos en diferentes organismos debido a la especiación.

20 Los ortólogos en, por ejemplo, especies de plantas monocotiledóneas, pueden encontrarse fácilmente realizando la denominada búsqueda BLAST recíproca. Esto se puede hacer mediante un primer BLAST que involucra la secuencia en cuestión (por ejemplo, SEC ID N° 1 o SEC ID N° 2) contra cualquier base de datos de secuencias, tal como la base de datos NCBI públicamente disponible. Se pueden usar BLASTN o TBLASTX cuando se parte de una secuencia de nucleótidos, o BLASTP o TBLASTN pueden usarse cuando se parte del polipéptido, con valores estándar predeterminados. Los resultados BLAST pueden filtrarse. Las secuencias de longitud completa de los resultados filtrados o los resultados no filtrados vuelven después a buscarse con BLAST (segundo BLAST) contra las secuencias del organismo del cual deriva la secuencia en cuestión. Después se comparan los resultados del primer y segundo BLAST. Cuando los resultados del segundo BLAST dan como coincidencia con la mayor similitud una secuencia de ácido nucleico de HDZip hox5 de clase I o un polipéptido HDZip hox5 de clase I, se ha encontrado un parólogo, si se origina del mismo organismo y para la secuencia usada en el primer BLAST. En caso de que se origina de un organismo distinto del de la secuencia utilizada en el primer BLAST, se ha encontrado un ortólogo. En el caso de familias grandes, se puede utilizar ClustalW, seguido de un árbol de unión del vecino, para ayudar a visualizar el agrupamiento. Preferentemente, dichos polipéptidos de HDZip hox5 de clase I tienen, en orden creciente de preferencia, una identidad o similitud de secuencia de al menos el 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % (identidad funcional) con un polipéptido de HDZip hox5 de clase I sin modificar (preferentemente, SEC ID N° 2; véase el Ejemplo 3 en el presente documento). El porcentaje de identidad entre los homólogos de HDZip hox5 de clase I fuera del homeodominio y la cremallera de leucina es supuestamente bajo (véase el ejemplo 3 en el presente documento). Ejemplos de ortólogos y parálogos de un polipéptido HDZip hox5 de clase I como se representa mediante la SEC ID N° 2 se pueden encontrar en la Tabla del Ejemplo 1 en el presente documento.

40 Un homólogo puede estar en forma de una "variante de sustitución" de una proteína, es decir, en la que al menos un residuo en una secuencia de aminoácidos se ha eliminado y un residuo diferente se ha insertado en su lugar. Las sustituciones de aminoácidos son típicamente de residuos individuales, pero pueden agruparse dependiendo de restricciones funcionales sobre el polipéptido; las inserciones normalmente serán del orden de aproximadamente 1 a 10 residuos de aminoácidos. Preferentemente, las sustituciones de aminoácidos comprenden sustituciones de aminoácidos conservadoras. Las tablas de sustituciones conservadoras están fácilmente disponibles en la técnica. La siguiente tabla muestra ejemplos de sustituciones de aminoácidos conservados.

Tabla 2: Ejemplos de sustituciones de aminoácidos conservados

Residuo	Sustituciones conservadoras	Residuo	Sustituciones conservadoras
Ala	Ser	Leu	Ile; Val
Arg	Lys	Lys	Arg; Gln
Asn	Gln; His	Met	Leu; Ile
Asp	Glu	Phe	Met; Leu; Tyr

Residuo	Sustituciones conservadoras	Residuo	Sustituciones conservadoras
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		

Un homólogo también puede estar en la forma de una "variante de inserción" de una proteína, es decir, en la que uno o más residuos de aminoácidos se introducen en un sitio predeterminado en una proteína. Las inserciones pueden comprender fusiones en N-terminal y / o en C-terminal, así como inserciones dentro de la secuencia de aminoácidos individuales o múltiples. Generalmente, las inserciones dentro de la secuencia de aminoácidos serán más pequeñas que las fusiones en N- o C-terminal, del orden de aproximadamente 1 a 10 residuos. Ejemplos de proteínas o péptidos de fusión en N- o C-terminal incluyen el dominio de activación o dominio de unión de un activador de la transcripción como se utiliza en el sistema en levadura de dos híbridos, las proteínas de recubrimiento de fago, cola de 6 (histidinas), marcador de glutatión S-transferasa, proteína A, proteína de unión a maltosa, dihidrofolato reductasa, epítipo Tag 100, epítipo de c-myc, epítipo FLAG®, lacZ, CMP (péptido de unión a calmodulina), epítipo HA, epítipo de proteína C y epítipo VSV.

Los homólogos en la forma de "variantes de delección" de una proteína se caracterizan por la eliminación de uno o más aminoácidos de una proteína.

Las variantes de aminoácidos de una proteína pueden fabricarse fácilmente utilizando técnicas de síntesis de péptidos bien conocidas en la materia, tales como la síntesis de péptidos en fase sólida y similares, o por manipulaciones de ADN recombinante. En la técnica se conocen bien los procedimientos para la manipulación de secuencias de ADN para producir variantes de sustitución, inserción o delección de una proteína. Por ejemplo, técnicas para realizar mutaciones de sustitución en sitios predeterminados en el ADN son bien conocidas para los expertos en la técnica e incluyen mutagénesis M13, mutagénesis in vitro T7-Gen (USB, Cleveland, OH) , mutagénesis dirigida a sitio QuickChange (Stratagene, San Diego, CA), mutagénesis dirigida a sitio mediada por PCR u otros protocolos de mutagénesis dirigida a sitio.

El polipéptido HDZip hox5 de clase I u homólogo del mismo puede ser un derivado. "Derivados" incluyen péptidos, oligopéptidos, polipéptidos que pueden, en comparación con la secuencia de aminoácidos de la forma natural de la proteína, tal como la proteína de interés, comprender sustituciones de aminoácidos con restos de aminoácido de origen no natural o adiciones de restos de aminoácido de origen no natural. Los "derivados" de una proteína también incluyen péptidos, oligopéptidos, polipéptidos que comprenden restos de aminoácido modificados de origen natural (glucosilados, acilados, prenilados, fosforilados, miristoilados, sulfatados, etc.) o modificados no origen no natural en comparación con la secuencia de aminoácidos de forma no natural del polipéptido. Un derivado también puede comprender uno o más sustituyentes o adiciones no aminoacídicas en comparación con la secuencia de aminoácidos de la que deriva, por ejemplo, una molécula indicadora u otro ligando, unido covalente o no covalentemente a la secuencia de aminoácidos, tal como una molécula indicadora que se une para facilitar su detección, y restos de aminoácido que no son de origen natural con respecto a la secuencia de aminoácidos de una proteína de origen natural.

El polipéptido HDZip hox5 de clase I u homólogo del mismo puede estar codificado por una variante de corte y empalme alternativa de un gen/secuencia de ácido nucleico de HDZip hox5 de clase I. El término "variante de corte y empalme" como se usa en el presente documento incluye variantes de una secuencia de ácido nucleico en la que los intrones y/o exones seleccionados se han escindido, reemplazado, desplazado o añadido, o en la que los intrones se han acortado o alargado. Dichas variantes serán unas en las que la actividad biológica de la proteína se conserva sustancialmente; esto puede conseguirse conservando selectivamente los segmentos funcionales de la proteína. Dichas variantes de corte y empalme pueden encontrarse en la naturaleza o pueden ser fabricadas por el hombre. Los procedimientos para fabricar dichas variantes de corte y empalme son bien conocidos en la técnica. Variantes de corte y empalme preferidas son variantes de corte y empalme de una secuencia de ácido nucleico que codifica un polipéptido que comprende de N-terminal a C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Adicionalmente, un polipéptido hox5 HDZip de la clase I o un homólogo del mismo puede comprender una o ambas de las siguientes: (a) una cola DE Trp; y (b) el motivo de

aminoácidos RPFF, en el que R es Arg, Pro P y F Phe. El motivo de (b) precede a la caja ácida, al examinar la proteína desde el N-terminal al C-terminal. Más preferidas son las variantes de corte y empalme de secuencias de ácidos nucleicos como se enumeran en la tabla A del Ejemplo 1 en el presente documento. La más preferida es una variante de corte y empalme de una secuencia de ácido nucleico como se representa en la SEC ID N° 1.

El homólogo también puede estar codificado por una variante alélica de una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo, preferiblemente una variante alélica de una secuencia de ácido nucleico que codifica un polipéptido que comprende de N-terminal a C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Adicionalmente, un polipéptido hox5 HDZip de la clase I o un homólogo del mismo puede comprender una o ambas de las siguientes: (a) una cola DE Trp; y (b) el motivo de aminoácidos RPFF, en el que R es Arg, Pro P y F Phe. El motivo de (b) precede a la caja ácida, al examinar la proteína desde el N-terminal al C-terminal. Más preferidas son las variantes alélicas de secuencias de ácidos nucleicos como se enumeran en la tabla A del Ejemplo 1 en el presente documento. La más preferida es una variante alélica de una secuencia de ácido nucleico como se representa en la SEC ID N° 1. Existen variantes alélicas en la naturaleza y englobado dentro de los procedimientos de la memoria descriptiva es el uso de estos alelos naturales. Las variantes alélicas incluyen Polimorfismos de nucleótidos sencillos (SNP), así como Polimorfismos de Inserción/Delección pequeños (INDEL). El tamaño de INDEL es normalmente menor de 100 pb. Los SNP y los INDEL forman el grupo de variantes de secuencia más grande en las cepas polimórficas de origen natural de la mayoría de los organismos.

De acuerdo con un aspecto preferido de la presente invención, se prevé la expresión modulada de la secuencia de ácido nucleico HDZip hox5 de clase I. Los procedimientos para la modulación de la expresión de genes o productos génicos están bien documentados en la técnica e incluyen, por ejemplo, la sobreexpresión dirigida por promotores apropiados, el uso de potenciadores de la transcripción o potenciadores de traducción. Las secuencias de ácidos nucleicos aislados que sirven como elementos promotores o potenciadores pueden introducirse en una posición adecuada (típicamente cadena arriba) de una forma no heteróloga de un polinucleótido para regular por aumento la expresión de una secuencia de ácido nucleico de HDZip hox5 de clase I. Por ejemplo, los promotores endógenos pueden alterarse in vivo por mutación, delección y/o sustitución (véase, Kmiec, patente de EE.UU. 5.565.350; Zarlign y col., PCT/US93/03868), o pueden introducirse promotores aislados en una célula vegetal en la orientación y distancia apropiada de un gen como se divulga en el presente documento de tal manera que se controle la expresión del gen. Los procedimientos para la reducción de la expresión de genes o productos génicos están bien documentados en la técnica.

Si se desea la expresión del polipéptido, generalmente es deseable incluir una región de poliadenilación en el extremo 3' de una región codificante de polinucleótido. La región de poliadenilación puede derivar del gen natural, de una diversidad de otros genes de planta, o de ADN-T. La secuencia en el extremo 3' a añadir puede derivar de, por ejemplo, los genes nopalina sintasa u octopina sintasa, o, como alternativa, de otro gen vegetal, o, menos preferentemente de cualquier otro gen eucariota

También puede añadirse una secuencia intrónica a la región no traducida en 5' o la secuencia codificante de la secuencia de codificación parcial para aumentar la cantidad del mensaje maduro que se acumula en el citosol. Se ha observado que la inclusión de un intrón de corte y empalme en la unidad de transcripción en construcciones de expresión tanto de plantas como de animales aumenta la expresión génica tanto a nivel de proteína como de ARNm hasta 1.000 veces (Buchman y Berg (1988) Mol. Cell Biol. 8: 4395 - 4405; Callis et al. (1987) Genes Dev 1:1183 - 1200). Dicha potenciación intrónica de la expresión génica es típicamente mayor cuando se pone cerca del extremo 5' de la unidad de transcripción. En la técnica se conoce el uso de los intrones de maíz, el intrón Adh1-S 1, 2 y 6, el intrón Bronze-1. Véase The Maize Handbook, capítulo 116, Freeling y Walbot, Eds., Springer, N.Y. (1994).

La divulgación también proporciona construcciones genéticas y vectores para facilitar la introducción y / o expresión de las secuencias de nucleótidos útiles en los procedimientos como se describe en el presente documento.

Se divulga una construcción génica que comprende:

- (i) una secuencia de ácido nucleico HDZip hox5 de clase I como se ha definido anteriormente en el presente documento;
- (ii) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (i); y, opcionalmente
- (iii) Una secuencia de terminación de transcripción.

Las construcciones útiles en los procedimientos de acuerdo con la presente invención se pueden construir utilizando tecnología de ADN recombinante bien conocida por las personas expertas en la técnica. Las construcciones génicas se pueden insertar en vectores, que pueden estar comercialmente disponibles, adecuados para transformar en plantas y adecuados para la expresión del gen de interés en las células transformadas.

Las plantas se transforman con un vector que comprende la secuencia de interés (es decir, una secuencia de ácido nucleico que codifica un polipéptido de HDZip hox5 de clase I u homólogo del mismo como se ha definido

anteriormente). La secuencia de interés está unida operablemente a una o más secuencias de control (al menos a un promotor). Las expresiones “elemento regulador”, “secuencia de control” y “promotor” se usan de forma intercambiable en el presente documento y se han tomado en un contexto amplio para hacer referencia a secuencias de ácido nucleico capaces de efectuar la expresión de las secuencias a las que están unidas. Incluidas en los términos mencionados anteriormente se encuentran las secuencias reguladoras de la transcripción derivadas de un gen genómico eucariota clásico (que incluye la caja TATA que es necesaria para el inicio adecuado de la transcripción, con o sin una secuencia de caja CCAAT) y elementos reguladores adicionales (es decir, secuencias activadoras en dirección 5', potenciadores y silenciadores) que alteran la expresión génica en respuesta a estímulos del desarrollo y/o externos, o en una forma específica de tejido. También se incluye dentro del término una secuencia reguladora de la transcripción de un gen procariota clásico, en cuyo caso puede incluir una secuencia de caja -35 y/o secuencias reguladoras transcripcionales de caja -10. La expresión “elemento regulador” también incluye una molécula de fusión sintética o un derivado que confiere, activa o potencia la expresión de una molécula de secuencia ácido nucleico en una célula, tejido u órgano. La expresión “unido operablemente” como se usa en el presente documento se refiere a un enlace funcional entre la secuencia promotora y el gen de interés, de tal manera que la secuencia promotora es capaz de iniciar la transcripción del gen de interés.

Ventajosamente, se puede usar cualquier tipo de promotor para dirigir la expresión de la secuencia de ácido nucleico como se divulga en el presente documento. El promotor puede ser un promotor inducible, es decir, que ha inducido o incrementado el inicio de la transcripción en respuesta a un estímulo del desarrollo, químico, ambiental o físico. Un ejemplo de un promotor inducible que es un promotor inducible por estrés, es decir un promotor activado cuando una planta se expone a diversas condiciones de estrés. Adicional o alternativamente, el promotor puede ser un promotor preferido del tejido, es decir uno que es capaz de iniciar la transcripción preferentemente en ciertos tejidos, tales como las hojas, raíces, tejido de las semillas, etc. Los promotores capaces de iniciar la transcripción en ciertos tejidos sólo se denominan en el presente documento como “específicos de tejido”.

En una realización, una secuencia de ácido nucleico HDZip hox5 de clase I está unida operablemente a un promotor constitutivo. Un promotor constitutivo es transcripcionalmente activo durante la mayor parte, pero no necesariamente todas las fases de su crecimiento y desarrollo y se expresa sustancialmente de forma ubicua. El promotor constitutivo es preferiblemente un promotor GOS2, más preferentemente el promotor constitutivo es un promotor GOS2 de arroz, más preferentemente el promotor constitutivo está representado por una secuencia de ácido nucleico sustancialmente similar a la SEC ID N° 33 o la SEC ID N° 52, lo más preferentemente, el promotor constitutivo es como se representa mediante la SEC ID N° 33 o la SEC ID N° 52. Ejemplos de otros promotores constitutivos que también pueden utilizarse para realizar los procedimientos de la invención se muestran en la Tabla 3 a continuación.

Tabla 3: Ejemplos de promotores constitutivos

Fuente del gen	Patrón de expresión	Referencia
Actina	Constitutivo	McElroy et al., Plant Cell, 2: 163 - 171, 1990
CAMV 35S	Constitutivo	Odell et al., Nature, 313: 810 - 812, 1985
CaMV 19S	Constitutivo	Nilsson et al., Physiol. Plant. 100:456 - 462, 1997
GOS2	Constitutivo	de Pater et al., Plant J Nov;2(6):837 - 44, 1992
Ubiquitina	Constitutivo	Christensen et al., Plant Mol. Biol. 18: 675 - 689, 1992
Ciclofilina del arroz	Constitutivo	Buchholz et al., Plant Mol Biol. 25(5): 837 - 43, 1994
Histona H3 del maíz	Constitutivo	Lepetit et al., Mol. Gen. Genet. 231:276 - 285, 1992
Actina 2	Constitutivo	An et al., Plant J. 10(1); 107 - 121, 1996
HMGB	Constitutivo	Documento WO 2004/070039

Opcionalmente, también se pueden utilizar una o más secuencias terminadoras en la construcción introducida en una planta. El término “terminador” incluye una secuencia de control que es una secuencia de ADN en el extremo de una unidad de transcripción que señala el procesamiento y la poliadenilación en 3' de un transcrito primario y la

terminación de la transcripción. Elementos reguladores adicionales pueden incluir potenciadores tanto de la transcripción como de la traducción. Los expertos en la técnica conocerán las secuencias terminadoras de terminación y potenciadoras que pueden ser adecuados para su uso en la realización de la invención. Tales secuencias serían conocidas o pueden obtenerlas fácilmente un experto en la técnica.

- 5 Las construcciones genéticas de la divulgación puede incluir además una secuencia del origen de secuencia de replicación que se requiere para el mantenimiento y / o replicación en un tipo celular específico. Un ejemplo es cuando se requiere mantener una construcción genética para en una célula bacteriana como un elemento genético episomal (por ejemplo, una molécula de plásmido o de cósmido). Los orígenes preferidos de replicación incluyen, entre otros, el f1-ori y colE1.
- 10 La construcción genética puede comprender opcionalmente un gen marcador seleccionable. Tal como se utiliza en el presente documento, la expresión "gen marcador seleccionable" incluye cualquier gen que confiere un fenotipo sobre una célula en la que se expresa para facilitar la identificación y / o selección de células que son transfectadas o transformadas con una construcción de la secuencia de ácido nucleico. Los marcadores adecuados se pueden seleccionar de marcadores que confieren resistencia a antibióticos o a herbicidas, que introducen un nuevo rasgo metabólico o que permiten selección visual. Ejemplos de genes marcadores seleccionables incluyen genes que confieren resistencia a antibióticos (tales como nptII que fosforila la neomicina y la kanamicina, o hpt, que fosforila la higromicina), a herbicidas (por ejemplo bar que proporciona resistencia a Basta; aroA o gox que proporcionan resistencia contra el glifosato, o los genes que confieren resistencia a, por ejemplo, imidazolinona, fosfotricina o sulfonilurea) o genes que proporcionan un rasgo metabólico (tal como manA que permite a las plantas utilizar manosa como única fuente de carbono). La expresión de genes marcadores visuales tiene como resultado la formación de color (por ejemplo β -glucuronidasa, GUS), luminiscencia (tal como luciferasa) o fluorescencia (Proteína Fluorescente Verde, GFP, y derivados de los mismos).
- 15 20

También se divulga una construcción génica que comprende:

- 25 (i) una secuencia de ácido nucleico HDZip hox5 de clase I como se ha definido anteriormente en el presente documento;
- (ii) un promotor constitutivo capaz de dirigir la expresión de la secuencia de ácido nucleico de (i); y, opcionalmente
- (iii) Una secuencia de terminación de transcripción.

- 30 El promotor constitutivo es preferiblemente un promotor GOS2, más preferentemente el promotor constitutivo es el promotor GOS2 de arroz, más preferentemente el promotor constitutivo está representado por una secuencia de ácido nucleico sustancialmente similar a la SEC ID N° 33 o la SEC ID N° 52, lo más preferentemente, el promotor constitutivo es como se representa mediante la SEC ID N° 33 o la SEC ID N° 52.

- 35 La invención también proporciona un procedimiento para la producción de plantas transgénicas que tienen rendimiento aumentado en condiciones de disponibilidad reducida de nutrientes, respecto a las correspondientes plantas silvestres, que comprende la introducción y expresión en una planta de una secuencia de ácido nucleico como se define en la reivindicación 1, en el que dicha secuencia de ácido nucleico codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo.

- 40 Más específicamente, la presente invención proporciona un procedimiento para la producción de plantas transgénicas que tienen mayor rendimiento reducido en condiciones de disponibilidad reducida de nutrientes, en relación a las correspondientes plantas silvestres, en el que el procedimiento comprende:

- (i) introducir y expresar en una planta, parte de la planta o célula de planta una secuencia de ácido nucleico HDZip hox5 de clase I, como se define en la reivindicación 1; y
- (ii) cultivar la célula vegetal en condiciones que estimulan el crecimiento y desarrollo de la planta.

- 45 La secuencia de ácido nucleico puede introducirse directamente en una célula de planta o en la propia planta (incluyendo la introducción en un tejido, órgano o en cualquier otra parte de una planta). De acuerdo con una característica preferida, la secuencia de ácido nucleico se introduce preferentemente en una planta por transformación.

- 50 El término "transformación" como se refiere en el presente documento abarca la transferencia de un polinucleótido exógeno en una célula huésped, independientemente del procedimiento utilizado para la transferencia. El tejido vegetal capaz de la propagación clonal posterior, ya sea por organogénesis o embriogénesis, puede transformarse con una construcción genética y una planta entera regenerada a partir del mismo. El tejido particular escogido variará dependiendo de los sistemas de propagación clonal disponibles para, y más adecuados para, las especies particulares que se están transformado. Ejemplos de tejidos objetivo incluyen discos de hojas, polen, embriones, cotiledones, hipocótilos, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, yemas axilares, y meristemos radiculares), y tejido meristemático inducido (por ejemplo, meristemo de
- 55

cotiledón y meristemo del hipocotiledón). polinucleótico puede introducirse transitoria o establemente en una célula huésped y puede mantenerse no integrado, por ejemplo, como un plásmido. Como alternativa, puede integrarse en el genoma huésped. La célula vegetal transformada resultante puede después utilizarse para regenerar una planta transformada de una manera conocida por los expertos en la técnica.

- 5 La transformación de especies de plantas es actualmente una técnica bastante habitual. Ventajosamente, puede usarse cualquiera de los diversos procedimientos de transformación para introducir el gen de interés en una célula antecesora adecuada. Los procedimientos de transformación incluyen el uso de liposomas, electroporación, productos químicos que aumentan la captación de ADN libre, inyección del ADN directamente en la planta, pistola de bombardeo de partículas, transformación utilizando virus o polen y microproyección. Pueden seleccionarse procedimientos del procedimiento de calcio/poli(etilenglicol) para protoplastos (Krens, F.A. et al. (1982) *Nature* 296, 72 - 74; Negrotiu 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 el material de planta (Crossway A et al. (1986) *Mol. Gen Genet* 202: 179 - 185); bombardeo de partículas recubiertas con ADN o ARN (Klein TM et al. (1987) *Nature* 327: 70) infección con virus (no integrantes) y similares. Las plantas transgénicas de arroz que expresan una secuencia de ácido nucleico/gen de HDZip hox5 de clase I se producen, preferentemente, mediante transformación mediada por *Agrobacterium* usando cualquiera de los procedimientos conocidos para la transformación del arroz, tal como se describe en cualquiera de los siguientes: la 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 de maíz, el procedimiento preferido es como se describe en Ishida et al. (*Nat. Biotechnol* 14(6): 745 - 50, 1996) o Frame et al. (*Plant Physiol* 129(1): 13 - 22, 2002).

Generalmente después de la transformación, se seleccionan células vegetales o agrupaciones de células para determinar la presencia de uno o más marcadores que están codificados por genes expresables en plantas cotransferidos con el gen de interés, tras lo cual el material transformado se regenera en una planta entera.

- 25 Después de la transferencia de ADN y la regeneración, las plantas transformadas putativamente se pueden evaluar, por ejemplo utilizando análisis de tipo Southern, para detectar la presencia del gen de interés, el número de copias y / o la organización genómica. Alternativamente o adicionalmente, los niveles de expresión del ADN recién introducido pueden controlarse utilizando análisis de tipo Northern y/o Western, o PCR cuantitativa, siendo todas las técnicas bien conocidas por los expertos habituales en la técnica.

- 30 Las plantas transformadas generadas se pueden propagar mediante diversos medios, tales como por propagación clonal o técnicas de reproducción clásicas. Por ejemplo, una primera generación (o T1) de planta transformada puede autofecundarse para dar transformantes de la segunda generación homocigota (o T2) y las plantas T2 se propagan adicionalmente a través de técnicas clásicas de reproducción.

- 35 Los organismos transformados generados pueden adoptar diversas formas. Por ejemplo, pueden ser quimeras de células transformadas y de células no transformadas; transformantes clonales (por ejemplo, todas las células transformadas que contengan el casete de expresión); injertos de tejidos transformados y no transformados (por ejemplo, en plantas, una planta madre transformada injertada a un vástago no transformado).

- 40 La divulgación se extiende a cualquier célula vegetal o planta producida mediante cualquiera de los procedimientos descritos en el presente documento, y a todas las partes de plantas y propágulos de las mismas. La divulgación se extiende además se para incluir la progenie de una célula, tejido u órgano primario transformado o transfectado o la planta completa que se ha producido mediante cualquiera de los procedimientos anteriormente mencionados, siendo el único requisito que la progenie presente las mismas características genotípicas y/o fenotípicas que las producidas por el precursor en los procedimientos de acuerdo con la memoria descriptiva.

Se divulgan células huésped que contienen una secuencia de ácido nucleico HDZip hox5 de clase I aislado. Las células huésped preferidas son células vegetales.

- 45 También se divulgan partes cosechables de una planta, tales como, pero sin limitación, semillas, hojas, frutos, flores, tallos, raíces, rizomas, tubérculos y bulbos. También se divulgan productos derivados de una parte cosechable de dicha planta, tales como gránulos secos o polvos, aceite, grasa y ácidos grasos, almidón o proteínas.

- 50 La presente invención también abarca el uso de ácidos nucleicos HDZip hox5 de clase I y el uso de polipéptidos de HDZip hox5 de clase I como se define en la reivindicación 1 en un rendimiento mejorado en plantas cultivadas en condiciones de disponibilidad reducida de nutrientes con respecto a plantas silvestres. Preferentemente, el aumento del rendimiento es uno o más de: aumento del rendimiento total de semillas por planta, aumento en el número de semillas llenas, aumento de la tasa de llenado de semillas, mayor número de flores por panícula, o mayor índice de cosecha.

- 55 Las secuencias de ácidos nucleicos de HDZip hox5 de clase I o variantes de las mismas, o polipéptidos de HDZip hox5 de clase I u homólogos de los mismos, como se divulga en el presente documento pueden encontrar uso en programas de cultivo en los que se identifica un marcador de ADN que puede estar genéticamente vinculado a un gen de HDZip hox5 de clase I variante del mismo. Las secuencias de ácidos nucleicos / genes de HDZip hox5 de clase I o variantes de las mismas, o polipéptidos I de HDZip hox5 de clase I u homólogos de los mismos se puede

utilizar para definir un marcador molecular. Este marcador de ADN o de proteína puede después usarse en programas de cultivo para seleccionar plantas con mayor rendimiento cuando se cultivan en condiciones de disponibilidad reducida de nutrientes, como se ha definido anteriormente en el presente documento en los procedimientos de la invención. El gen de HDZip hox5 de clase I puede ser, por ejemplo, una secuencia de ácido nucleico como se enumera en la Tabla A del ejemplo 1 en el presente documento.

Las variantes alélicas de una secuencia de ácido nucleico/gen de HDZip hox5 de clase I como se divulga en el presente documento también pueden encontrar uso en programas de reproducción asistida mediante marcador. Dichos programas de reproducción a veces requieren la introducción de variación alélica por tratamiento mutagénico de las plantas, usando, por ejemplo, mutagénesis con EMS; como alternativa, el programa puede comenzar con una colección de variantes alélicas de origen denominado "natural" producida involuntariamente. A continuación, la identificación de variantes alélicas se realiza, por ejemplo, por PCR. A esto le sigue una etapa de selección de variantes alélicas superiores de la secuencia en cuestión y que dará un rendimiento aumentado en condiciones de disponibilidad reducida de nutrientes, con respecto a las plantas silvestres correspondientes. La selección se realiza típicamente supervisando el rendimiento del crecimiento de las plantas que contienen diferentes variantes alélicas de la secuencia en cuestión, por ejemplo, diferentes variantes alélicas de uno cualquiera de los ácidos nucleicos indicados en la Tabla A del Ejemplo 1 en el presente documento. El rendimiento del crecimiento puede supervisarse en un invernadero o en el campo. Etapas opcionales adicionales incluyen cruzamiento de plantas, en el que se identifica la variante alélica superior, con otra planta. Esto podría usarse, por ejemplo, para realizar una combinación de características fenotípicas interesantes.

Una secuencia de ácido nucleico HDZip hox5 de clase I o una variante de la misma también puede usarse como una sonda para mapear, genética y físicamente, genes que forman parte de, y como marcadores para rasgos ligados a estos genes. Dicha información puede ser útil en reproducción de plantas para desarrollar líneas con fenotipos deseados. Dicho uso de secuencias de ácido nucleico HDZip hox5 de clase I o variantes de las mismas requiere solo una secuencia de ácido nucleico de al menos 15 nucleótidos de longitud. Las de secuencias de ácido nucleico HDZip hox5 de clase I o variantes de las mismas pueden usarse como marcadores de polimorfismos de longitud de fragmentos de restricción (RFLP). Las transferencias Southern (Sambrook J, Fritsch EF y Maniatis T (1989) Molecular Cloning, A Laboratory Manual) del ADN genómico de planta digerido con enzimas de restricción se pueden sondear con las secuencias de ácido nucleico de HDZip hox5 de clase I o variantes de las mismas. Los patrones de banda resultantes pueden después someterse a análisis genético usando programas informáticos tales como MapMaker (Lander et al. (1987) Genomics 1: 174 - 181) para construir un mapa genético. Además, las secuencias de ácidos nucleicos pueden usarse para sondear las transferencias de Southern que contienen ADN genómicos tratados con endonucleasas de restricción de un conjunto de individuos que representan a los padres y a la progenie de un cruce genético definido. La segregación de los polimorfismos de ADN se observa y se usa para calcular la posición de la secuencia de ácido nucleico de HDZip hox5 de clase I en el mapa genético previamente obtenido usando esta población (Botstein et al. (1980) Am. J. Hum. Genet. 32:314 - 331).

La producción y el uso de sondas derivadas de genes de plantas para su 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 de ADNc específicos utilizando la metodología indicada anteriormente o variaciones de la misma. Por ejemplo, para el mapeo pueden utilizarse poblaciones de intercrucamiento F2, poblaciones de retrocruzamiento, poblaciones apareadas al azar, líneas isogénicas cercanas y otros grupos de individuos. Dichas metodologías son bien conocidas por los expertos en la técnica.

Las sondas de secuencias de ácidos nucleicos también pueden usarse para mapeo físico (es decir, colocación de secuencias en mapas físicos; véase Hoheisel et al. En: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, pp. 319 - 346, y referencias citadas en la misma).

Las sondas de secuencias de ácidos nucleicos pueden usarse en la hibridación in situ con fluorescencia (FISH) (Trask (1991) Trends Genet. 7:149 - 154). Aunque los procedimientos actuales de mapeo con FISH favorecen el uso de clones grandes (de varios kb a varios cientos de kb; véase Laan et al. (1995) Genome Res. 5:13 - 20), mejoras en sensibilidad pueden permitir la realización de mapeo FISH usando sondas más cortas.

Usando las secuencias de ácidos nucleicos pueden realizarse diversos procedimientos basados en amplificación de ácidos nucleicos para el mapeo genético y físico. Ejemplos incluyen amplificación específica de alelo (Kazazian (1989) J. Lab. Clin. Med 11:95 - 96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield et al. (1993) Genomics 16:325 - 332), ligamiento específico de alelo (Landegren et al. (1988) Science 241:1077 - 1080), reacciones de extensión de nucleótidos (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 y Cook (1989) Nucleic acid sequence Res. 17:6795 - 6807). Para estos procedimientos, se utiliza la secuencia de una secuencia de ácido nucleico para diseñar y producir pares de cebadores para su uso en la reacción de amplificación o en reacciones de extensión con cebador. El diseño de dichos cebadores es bien conocido por los expertos en la técnica. En los procedimientos que emplean mapeo genético basado en PCR puede ser necesario identificar diferencias de secuencias de ADN entre los padres del cruce del mapeo en la región correspondiente a la secuencia de ácido nucleico instantánea. Sin embargo, generalmente esto no es necesario para procedimientos de mapeo.

Los procedimientos de acuerdo con la presente invención dan como resultado plantas que tienen rendimiento aumentado en condiciones de disponibilidad reducida de nutrientes, como se ha descrito anteriormente en el presente documento. Este rendimiento aumentado también puede combinarse con otros rasgos económicamente ventajosos, tales como rasgos que potencian adicionalmente el rendimiento, tolerancia a estrés abiótico y biótico, rasgos que modifican diversas características arquitectónicas y/o características bioquímicas y/o fisiológicas.

Descripción de las figuras

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

La **Figura 1** muestra un alineamiento múltiple de homeodominios de HDZip de clase I de diferentes fuentes vegetales, mediante el programa de alineamiento múltiple VNTI AlignX, basado en un algoritmo ClustalW modificado (InforMax, Bethesda, MD), con parámetros predeterminados para la penalización por apertura de espacio de 10 y una extensión de hueco de 0,05. Los aminoácidos invariables del homeodominio L₁₆, W₄₈, F₄₉, N₅₁ y R₅₃ están encuadradas verticalmente. Los aminoácidos HDZip de clase I preferidos A₄₆ and W₅₆ también están encuadrados verticalmente. Las tres hélices de ADN necesarias para la unión a ADN están marcados con recuadros negros por encima de la alineación. Las seis héptadas están separadas por una línea vertical. Las siete posiciones dentro de cada héptada se denominan a, b, c, d, e, f y g. La Leu ocupa la posición d dentro de cada héptada, y están encuadradas verticalmente.

La **Figura 2** muestra un alineamiento múltiple de varios polipéptidos de HDZip hox5 de clase I, mediante el programa de alineamiento múltiple VNTI AlignX, basado en un algoritmo ClustalW modificado (InforMax, Bethesda, MD), con parámetros predeterminados para la penalización por apertura de espacio de 10 y una extensión de hueco de 0,05). Los tres dominios principales caracterizados, de N-terminal a C-terminal, están encuadrados fuertemente y se identificaron como la caja ácida, el homeodominio de clase I y la cremallera leucina-seis héptadas. Además, la cola Trp y el motivo de aminoácidos RPFF estén ligeramente encuadrados.

La **Figura 3** muestra un vector binario para la expresión en *Oryza sativa* de un HDZip hox5 de clase de I de *Oryza sativa* bajo el control de un promotor GOS2.

La **Figura 4** detalla ejemplos de las secuencias de la cremallera de leucina de homeodominio de clase I (HDZip) útiles en la realización de los procedimientos de acuerdo con la presente invención. Varias secuencias resultan de conjuntos de EST públicos (véase la Tabla A del Ejemplo 1 en el presente documento), con la secuenciación de menor calidad. Como consecuencia, pueden esperarse unas pocas sustituciones de ácido nucleico. Los codones de iniciación (ATG) y de terminación delimitan las secuencias de ácidos nucleicos que codifican los polipéptidos HDZip hox5 de clase I de longitud completa.

Ejemplos

La presente invención se describirá a continuación con referencia a los siguientes ejemplos, que son únicamente ilustrativos y no están destinados a definir completamente o limitar de otro modo el alcance de la invención.

A menos que se indique lo contrario, las técnicas de ADN recombinante se realizaron de acuerdo con protocolos estándar descritos en (Sambrook (2001) Molecular Cloning: a laboratory manual, 3ª 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. Materiales y procedimientos estándar para el trabajo molecular en plantas se describen en Plant Molecular Biology Labfase (1993) de R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (Reino Unido) y Blackwell Scientific Publications (Reino Unido).

Ejemplo 1: Identificación de secuencias relacionadas con la SEC ID N° 1 y la SEC ID N° 2

Las secuencias (ADNc de longitud completa, EST o genómico) se identificaron entre las conservadas en la base de datos de nucleótidos Entrez en el National Center for Biotechnology Information (NCBI) utilizando las herramientas de búsqueda de secuencia de bases de datos, tal como la Herramienta de Alineamiento Local Básica (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403 - 410; y Altschul et al. (1997) Nucleic Acids Res. 25:3389 - 3402). El programa se utiliza para encontrar regiones de similitud local entre las secuencias comparando las secuencias de ácidos nucleicos o polipéptidos con las bases de datos de secuencia y calculando el significado estadístico de los emparejamientos. Por ejemplo, el polipéptido codificado por el ácido nucleico de la SEC ID N° 1 se utilizó para el algoritmo TBLASTN, con configuraciones predeterminadas y el filtro para ignorar las secuencias de baja complejidad desechadas. El resultado de los análisis se observó por comparación apareadas y se clasificó de acuerdo con la puntuación de probabilidad (valor E), en el que la puntuación refleja la probabilidad de que se produzca un alineamiento particular por casualidad (cuanto menor sea el valor E, más significativa es la coincidencia). Además de valores E, también se puntuaron las comparaciones por porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de ácidos nucleicos (o polipéptido) comparadas sobre una longitud particular. En algunos casos, los parámetros por defecto se pueden ajustar para modificar la rigurosidad de la búsqueda. Por ejemplo, el valor E puede aumentarse para mostrar menos emparejamientos exigentes. De esta manera, se pueden identificar las coincidencias cortas casi exactas.

La Tabla A dada a continuación proporciona una lista de secuencias de ácidos nucleicos relacionadas con la secuencia de ácidos nucleicos de SEC ID N° 1.

Tabla A: Ejemplos de secuencias relacionadas con la secuencia de ácido nucleico de SEC ID N° 1

Nombre	Número de acceso de nucleótidos en NCBI	SEC ID N° del nucleótido	SEC ID N° del polipéptido traducido	Fuente
Orysa_hox5	XM_482406	1	2	<i>Oryza sativa</i>
Orysa_hox16	XM_467603	3	4	<i>Oryza sativa</i>
Zeama_hox5*	CO458693 DV024016	5	6	<i>Zea mays</i>
Zeama_hox16	AY105265	7	8	<i>Zea mays</i>
Sacof_hox5*	CA088615 CA115362 CA142506	9	10	<i>Saccharum officinarum</i>
Sorbi_hox5*	BE363386 CD432381	11	12	<i>Sorghum bicolor</i>
Triae_hox16*	DR735359 DR741379 CD916488	13	14	<i>Triticum aestivum</i>
Arath_ATHB1	X58821	15	16	<i>Arabidopsis thaliana</i>
Dauca_CHB3**	D26575	17	18	<i>Daucus carota</i>
Glyma_HD157**	AF184278	19	20	<i>Glycine max</i>
Crapl_CPHB-5	AF443621	21	22	<i>Craterostigma plantagineum</i>
Goshi_hox5*	DT465649 CD486134	23	24	<i>Gossypium hirsutum</i>
Lycles_hox5	BT014213.1	25	26	<i>Lycopersicon esculentum</i>
Lycles_VaHOX1	X94947	27	28	<i>Lycopersicon esculentum</i>
Medsa_hox16*	CB892061 CA858059	29	30	<i>Medicago sativa</i>
Aqufo_hox5	DT758247	31	32	<i>Aquilegia formosa</i> x <i>Aquilegia pubescens</i>
Poptr_hox16_1	scaff_XV.439	40	41	<i>Populus tremuloides</i>

Nombre	Número de acceso de nucleótidos en NCBI	SEC ID N° del nucleótido	SEC ID N° del polipéptido traducido	Fuente
Poptr_hox16_2	scaff_XII.649	42	43	<i>Populus tremuloides</i>
Poptr_hox16_3	lcljscaff_VIII.1839	44	45	<i>Populus tremuloides</i>
Medtr_hox16_1	CR954197.2	46	47	<i>Medicago truncatula</i>
Phavu_hox16	AF402605	48	49	<i>Phaseolus vulgaris</i>
Lotco_hox16	AP006364	50	51	<i>Lotus corniculatus</i>
<p>*Contig recopilados de varias consultas EST (las principales de muestran); siendo calidad de la secuenciación de EST normalmente más baja, se puede esperar un par de sustituciones de ácidos nucleicos.</p> <p>**Las secuencias de <i>Daucus carota</i> y <i>Glycine max</i> se han corregido en comparación con su número de acceso.</p>				

Ejemplo 2: Alineación de las secuencias de polipéptidos HDZip hox5 de clase I

Se usó AlignX del Vector NTI (Invitrogen) basado en el algoritmo Clustal popular de alineación progresiva (Thompson et al. (1997) Nucleic Acids Res 25:4876 - 4882; Chenna et al. (2003). Nucleic Acids Res 31:3497 – 3500). Se puede construir un árbol filogenético usando un algoritmo de agrupamiento de unión al vecino. Los valores por defecto son para la penalización por apertura de hueco de 10, la penalización por extensión de hueco de 0,1 y la matriz de peso seleccionado es Blosum 62 (si polipéptidos están alineados).

El resultado de la alineación de secuencias múltiples se muestra en la Figura 2. Los tres dominios principales caracterizados, de N-terminal a C-terminal, están encuadrados fuertemente y se identificaron como la caja ácida, el homeodominio de clase I y la cremallera leucina-seis héptadas. El "dominio conservado" comprende estos tres dominios. Además, la cola Trp y el motivo de aminoácidos RPFF estén ligeramente encuadrados.

Ejemplo 3: Cálculo del porcentaje de identidad global entre la secuencias de los polipéptidos HDZip hox5 de clase I

Los porcentajes globales de similitud e identidad entre las secuencias de polipéptidos HDZip hox5 de clase I se determinaron utilizando el software Matrix Global Alignment Tool (MatGAT) (BMC Bioinformatics. 2003 4:29. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. Campanella JJ, Bitincka L, Smalley J; software de Ledion Bitincka). El software MatGAT genera matrices de similitud / identidad de secuencias de ADN o de proteína sin necesidad de pre-alineación de los datos. El programa realiza una serie de alineaciones apareadas utilizando el algoritmo de alineación global de Myers y Miller (con una penalización por apertura de hueco de 12, y una penalización de extensión de hueco de 2), calcula la similitud y la identidad utilizando, por ejemplo Blosum 62 (para los polipéptidos), y luego coloca los resultados en una matriz de distancia. La similitud de secuencia se muestra en la mitad inferior de la línea divisoria y la identidad de secuencia se muestra en la mitad superior de la línea divisoria diagonal.

Los parámetros utilizados en la comparación fueron:

Matriz de puntuación: Blosum62

Primer hueco: 12

Extensión de hueco: 2

Los resultados del análisis de software se muestran en la Tabla B 1 para la similitud global y de identidad sobre la longitud completa de las secuencias polipeptídicas (excluyendo las secuencias de polipéptidos parciales). El porcentaje de identidad se proporciona encima de la diagonal y el porcentaje de similitud se proporciona por debajo de la diagonal.

El porcentaje de identidad entre las secuencias de polipéptido mostradas puede ser tan bajo como del 29 % de identidad de aminoácidos en comparación con la SEC ID N° 2.

Tabla B1: Resultados de MatGAT para similitud e identidad globales en toda la longitud de las secuencias de polipéptidos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Aqufo_Hox5		56	54	37	34	64	36	64	35	34	36	47	37	36	38	38	36	42	41	46	49	42
2. Arath_ATHB1	73		52	34	34	59	36	57	33	34	36	44	34	35	40	35	35	39	39	41	43	39
3. Crapl_CPHB-5	69	66		33	35	56	37	59	33	33	33	45	39	34	37	36	36	41	39	41	44	41
4. Dauca_CHB3	52	52	48		44	39	53	35	46	49	46	30	33	47	58	56	43	32	33	31	33	33
5. Glyma_HD157	50	47	48	58		33	44	32	43	43	72	33	31	84	48	48	47	32	31	31	31	32
6. Goshi_Hox5	79	74	71	53	49		38	64	36	36	37	46	38	35	39	36	35	40	39	46	49	40
7. Lotco_Hox16	51	53	51	66	62	53		35	45	66	50	29	31	49	62	59	49	30	31	30	32	31
8. Lyces_Hox5	75	70	72	51	45	75	50		34	34	36	46	38	34	37	36	33	41	41	45	47	41
9. Lyces_VaHOX1	49	48	47	63	58	48	62	47		45	44	31	32	47	53	49	44	33	33	32	33	33
10. Medtr_Hox16	48	48	50	65	64	49	78	48	63		46	30	30	45	59	55	42	31	30	31	31	30
11. Medtr_Hox16_1	52	49	50	61	81	49	67	49	61	64		33	28	77	51	48	50	32	31	29	29	32
12. Orysa_Hox16	62	59	58	50	50	60	47	58	45	50	51		49	34	32	31	30	46	45	73	76	45
13. Orysa_Hox5	53	47	52	48	48	52	48	50	44	45	46	59		32	32	32	30	66	66	49	50	65
14. Phavu_HOX16	51	51	48	64	89	49	65	47	63	65	88	49	48		56	55	51	34	32	31	32	33
15. Poptr_HOX16_1	54	54	52	71	66	52	75	50	66	73	69	48	49	71		92	48	35	35	32	34	34
16. Poptr_HOX16_2	51	49	51	70	66	50	73	49	65	70	66	47	46	71	96		47	34	33	32	32	34
17. Poptr_HOX16_3	52	51	47	59	59	52	63	45	59	59	62	44	44	65	63	63		34	33	30	31	33
18. Sacof_Hox5	62	58	57	47	44	60	48	57	44	45	46	56	69	46	48	47	47		95	46	46	94
19. Sorbi_Hox5	62	57	55	46	45	58	51	58	45	44	46	56	69	47	50	47	46	97		43	46	94
20. Triae_Hox16	62	54	56	48	48	59	47	58	47	49	47	82	61	48	52	51	46	56	55		72	45
21. Zeama_Hox16	63	58	59	51	49	62	51	60	49	50	47	81	62	51	49	48	46	56	57	81		45
22. Zeama_Hox5	62	58	56	46	44	59	49	57	45	45	45	55	68	46	50	46	48	96	96	57	56	

- 5 El "dominio conservado" de las secuencias de polipéptidos HDZip hox5 de clase I comprende de N-terminal a C-terminal, una caja ácida, un homeodominio de clase I y la cremallera de leucina-seis héptadas (véase la Figura 2), como se ha definido anteriormente en el presente documento. Cuando el análisis del porcentaje de identidad se realiza en los dominios conservados en lugar de en las secuencias de polipéptido de longitud completa, se observa un aumento en el porcentaje de identidad, como se muestra en la Tabla B 2. Los valores más bajos están ahora por encima del 50 % de identidad de los aminoácidos en comparación con la SEC ID N° 2.

Tabla B2: Resultados de MatGAT para similitud e identidad globales en el “Dominio conservado” de las secuencias de polipéptidos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Aqufo_hox5_CD		81	74	66	82	62	82	62	61	75	66	61	67	63	60	68	67	73	76	67
2. Arath_ATHB1_CD	93		74	59	82	61	85	61	61	71	66	63	65	62	59	67	66	71	71	67
3. Crapl_CPHB3_5_CD	85	85		61	78	62	77	57	60	69	63	62	62	58	65	65	64	66	68	65
4. Dauca_CHB3_CD	81	79	75		64	70	62	66	69	57	57	64	80	75	66	58	59	57	57	58
5. Goshi_hox5_CD	94	95	89	81		66	83	63	63	74	66	63	68	63	64	67	66	73	75	67
6. Lotco_hox16_CD	80	78	74	81	80		62	67	85	54	52	62	77	73	64	53	52	54	54	53
7. Lyces_hox5_CD	91	92	88	79	92	77		57	61	75	68	63	66	62	60	71	69	75	76	70
8. Lyces_VaHox1_CD	77	75	71	79	76	84	74		71	56	57	62	73	68	60	58	58	57	57	58
9. Medtr_hox16_CD	77	75	74	81	77	93	75	84		59	57	61	75	71	62	57	57	58	59	57
10. Orysa_hox16_CD	93	92	85	77	92	79	92	74	75		84	58	60	58	58	82	82	94	96	82
11. Orysa_hox5_CD	90	87	83	78	88	77	88	71	74	91		59	58	57	57	92	93	81	84	94
12. Phavu_hox16_CD	79	79	75	80	79	81	76	75	80	76	75		71	71	65	60	60	57	58	60
13. Poptr_hox16_1_CD	81	80	75	88	82	90	79	84	87	79	80	86		93	65	59	60	58	61	59
14. Poptr_hox16_2_CD	79	77	74	85	79	87	76	82	84	75	75	85	98		62	57	57	56	58	57
15. Poptr_hox16_3_CD	74	75	73	80	75	76	71	76	77	73	71	78	77	77		57	57	56	58	58
16. Sacof_hox5_CD	89	86	80	77	87	77	86	71	73	88	96	74	79	75	72		98	79	82	98
17. Sorbi_hox5_CD	89	86	81	77	87	77	87	71	73	89	97	74	79	76	72	99		78	82	98
18. Triae_hox16_CD	93	92	85	79	92	80	92	74	75	98	91	77	80	77	73	88	89		95	79
19. Zeama_hox16_CD	93	92	86	78	92	81	92	75	77	98	92	79	80	76	74	90	91	98		82
20. Zeama_hox5_CD	89	86	81	77	87	77	87	71	73	89	97	74	79	75	72	99	100	89	91	

Ejemplo 4: Identificación de los dominios comprendidos en las secuencias de polipéptidos de HDZip hox5 de clase I

- 5 La base de datos del Recurso integrado de las Familias de proteínas, Dominios y Sitios (InterPro) es una interfaz integrada para las bases de datos de firmas de uso común para búsquedas de texto y de secuencias. La base de datos de InterPro combina estas bases de datos, que utilizan diferentes metodologías y diferentes grados de información biológica sobre las proteínas bien caracterizadas para obtener firmas de proteínas. Bases de datos colaboradoras incluyen SWISS-PROT, PROSITE, TrEMBL, IMPRESIONES, ProDom y Pfam, inteligente y TIGRFAM. Interpro está alojada en el Instituto Europeo de Bioinformática en el Reino Unido.
- 10 Los resultados de la exploración en InterPro de la secuencia polipeptídica como la representada por la SEC ID N° 2 se presentan en la Tabla C.

Tabla C: Resultados de la exploración en InterPro de la secuencia polipeptídica como la representada por la SEC ID N° 2

InterPro	IPR000047	Motivo hélice-giro-hélice, represor de tipo lambda	
	PRINTS	PR00031	HTHREPRESSR
InterPro	IPR001356	Homeocaja	
	PRODOM	PD000010	Homeocaja
	PRINTS	PR00024	HOMEOCAJA
	PFAM	PF00046	Homeocaja
	SMART	SM00389	HOX
	PROFILE	PS00027	HOMEOCAJA_1
	PROFILE	PS50071	HOMEOCAJA_2
InterPro	IPR003106	Cremallera de leucina, asociada a homeocaja	
	PFAM	PF02183	HALZ
InterPro	IPR009057	De tipo homeodominio	
	SUPERFAMILIA	SSF46689	De tipo homeodominio
InterPro	IPR012287	Relacionado con el homeodominio	
	GENE3D	G3DSA:1.10.10.60	Relacionado con el homeodominio

- 15 La composición primaria de aminoácidos (en %) para determinar si un dominio de polipéptido es rico en aminoácidos específicos (por ejemplo en una caja ácida) se puede calcular usando programas de software del servidor ExPASy, en particular, la herramienta ProtParam (Gasteiger E et al. (2003) ExPASy: the proteomics server for in-depth protein

knowledge and analysis. Nucleic Acids Res 31:3784 - 3788). La composición de la secuencia polipeptídica de interés puede compararse después con la composición de aminoácidos media (en %) en el banco de datos de secuencias de proteínas Swiss-Prot.

En la siguiente tabla (Tabla D), se comparan el % Asp (D), el % de Glu (E) y su contenido combinado en la caja ácida de la SEC ID N° 2 con el promedio de la base de datos de secuencias de proteínas Swiss-Prot.

Tabla D:

	% Asp (D)	% Glu (E)	% Asp (D) + % Glu (E)
Promedio de la base de datos de secuencias de proteínas Swiss-Prot	5,3 %	6,6 %	11,9 %
Caja ácida de la SEC ID N° 2	9,1 %	54,5 %	63,6 %

Una caja ácida puede ser parte de un dominio de activación de la transcripción. Los dominios de activación de la transcripción eucariótica se han clasificado en función de su contenido de aminoácidos, y las principales categorías incluyen dominios de activación ácidos ricos en glutamina y ricos en prolina. (2005) Plant J. 43(5): 769 - 88, y referencias citadas en la misma).

El Consorcio de Ontología Génica (GO) es una colaboración internacional entre científicos de diferentes bases de datos biológicos, con una oficina editorial con sede en el Instituto Europeo de Bioinformática. El objetivo de GO es proporcionar vocabularios controlados para la descripción de la función molecular, proceso biológico y los componentes celulares de productos génicos. Al realizar una exploración InterPro como se ha descrito anteriormente, también se realizan búsquedas en la base de datos GO. Las secuencias de polipéptidos HDZip hox5 de clase I tienen como función molecular de factor transcripción y actividad de unión a ADN específica de secuencia, y se localizan en el núcleo de la célula vegetal (véase la tabla a continuación (Tabla E)).

Tabla E:

	Entrada en ontología génica
<i>Homeodominio</i>	Función Molecular: actividad de factor de transcripción (GO: 0003700) Componente celular: núcleo (GO:0005634) Función Molecular: Unión a ADN específica de secuencia (GO:0043565)
<i>Cremallera de leucina, asociada a homeocaja</i>	Función Molecular: Unión a ADN (GO:0003677) Componente celular: núcleo (GO:0005634)

Ejemplo 5: Predicción de la topología de las secuencias de polipéptidos HDZip hox5 de clase I

La predicción de la cremallera de leucina y la identificación de héptadas se llevó a cabo usando software especializado, tal como 2ZIP, que combina un algoritmo de predicción de bobina en espiral estándar con una búsqueda aproximada para la repetición de la leucina característica (Bornberg-Bauer et al. (1998) Nucleic Acids Res 26(11): 2740 - 2746; alojados en Max Planck Institut, Golm en Alemania). Usando este software se puede identificar Una cremallera de leucina potencial, una repetición de leucinas o una bobina en espiral.

Las secuencias de polipéptidos HDZip hox5 de clase I comprenden una predicción de cremallera de leucina, con al menos 5, preferiblemente 6 héptadas. Cuando el polipéptido de la SEC ID N° 2 se somete a este algoritmo, una potencial cremallera de leucina está entre las posiciones 143 y 178, como se muestra en el resultado a continuación (los números reflejan la posición del amino ácido, la región C superenrollada, y la L leucina dentro de la héptada):

[illegible]

Ejemplo 6: Ensayo para las secuencias de polipéptidos HDZip hox5 de clase I

Los polipéptidos HDZip *hox5* de clase I u homólogos de los mismos tienen actividad de unión a ADN, preferiblemente a semisitos de 5 pb que se superponen en una posición central, CAA (A / T) ATTG, como se detecta en los ensayos de un híbrido en levaduras (Meijer et al. (2000) Mol Gen Genet 263:12 - 21). En ensayos transitorios en suspensiones celulares de arroz, el cobombardeo de un polipéptido HDZip *hox5* de clase I con el gen indicador GUS tuvo como resultado un aumento del número de puntos teñidos, que también fueron más intensos de color (Meijer et al, ant.). Este ensayo es útil para demostrar la función de activador de la los polipéptidos HDZip *hox5* de clase I u homólogos *hox5*

Ejemplo 7: Clonación de la secuencia de ácido nucleico de HDZip hox5 de clase I de *Oryza sativa*

La secuencia de ácidos nucleicos HDZip *hox5* de clase I de *Oryza sativa* se amplificó por PCR usando como molde una biblioteca de ADNc de *Oryza sativa* (Invitrogen, Paisley, RU). Después de transcripción inversa de ARN extraído de las plántulas, los ADNc se clonaron en pCMV Sport 6.0. El tamaño medio del inserto del banco fue de 1,6 kb, y el número original de clones fue del orden de $11,67 \times 10^7$ ufc. El título original se determinó que era de $3,34 \times 10^6$ ufc/ml, después de una primera amplificación de 6×10^{10} ufc/ml. Después de la extracción del plásmido, se utilizaron 200 ng de molde en una mezcla PCR de 50 μ l. Se utilizaron los cebadores prm06000 (SEC ID N° 34; sentido, codón de iniciación en **negrita**, sitio AttB1 en cursiva: 5' – GGGGACAAAGTTTGTACAAAAAGCAGGCTTAAACAATGGATCCCGGCCG-3') y prm06001 (SEC ID N° 35; inverso, complementario, sitio AttB2 en cursiva: 5'-GGGGACCACTTTGTACAAG AAAGCTGGGTGATCAGCTCCAGAACCAGG-3'), que incluyen los sitios AttB para la recombinación Gateway, para la amplificación por PCR. La PCR se realizó usando la ADN polimerasa de Taq Hifi en condiciones convencionales. Se amplificó un fragmento de PCR de 1116 pb (incluyendo los sitios attB, desde el codón de iniciación al codón de terminación 1050 pb) y se purificó también utilizando procedimientos convencionales. Después, se realizó la primera etapa del procedimiento Gateway, la reacción BP, durante la cual el fragmento de PCR se recombina in vivo con el plásmido pDONR201 para producir, de acuerdo con la terminología Gateway, un “clon de entrada”. El plásmido pDONR201 se adquirió en Invitrogen, como parte de la

tecnología Gateway®.

Ejemplo 8: Construcción de vectores

El clon de entrada que comprende las secuencias de ácido nucleico se usó posteriormente en una reacción LR con un vector “destinatario” utilizado para transformación de *Oryza sativa*. Este vector contenía como elementos funcionales dentro de los límites de ADN-T: un marcador de selección de plantas; un casete de expresión marcador detectable; y un casete Gateway diseñado para recombinación LR in vivo con la secuencia de interés ya clonada en el clon de entrada. Un promotor GOS2 de arroz (SEC ID N° 33 o SEC ID N° 52) para la expresión específica se localizó cadena arriba de este casete Gateway.

Después de la etapa de recombinación LR, el vector de expresión resultante (Figura 3) se transformó en la cepa de *Agrobacterium* LBA4044 de acuerdo con procedimientos bien conocidos en la técnica.

Ejemplo 9: Transformación de plantas

Transformación de arroz

El *Agrobacterium* que contenía el vector de expresión se usó para transformar plantas de *Oryza sativa*. Semillas secas maduras de la variedad de cultivo japonesa de arroz Nipponbare se descascarillaron. La esterilización se realizó incubando durante un minuto en etanol al 70 %, seguido de 30 minutos en HgCl₂ al 0,2 %, seguido de un lavado de 6 veces durante 15 minutos con agua destilada estéril. Las semillas estériles germinaron después en un medio que contenía 2, 4-D (medio de inducción de callo). Después de la incubación en la oscuridad durante cuatro semanas, los callos embriogénicos derivados de escutelo se cortaron y se propagaron en el mismo medio. Después de dos semanas, los callos se multiplicaron o se propagaron mediante subcultivo en el mismo medio durante otras 2 semanas. Piezas de callos embriogénico se subcultivaron en medio reciente 3 días antes de cocultivo (para reforzar la actividad de la división celular).

La cepa LBA4404 de *Agrobacterium* que contenía el vector de expresión se utilizó para el co-cultivo. Se inoculó *Agrobacterium* en medio AB con los antibióticos apropiados y se cultivó durante 3 días a 28 °C. Después las bacterias se recogieron y se suspendieron en medio de cocultivo líquido a una densidad (DO600) de aproximadamente 1. Después, la suspensión se transfirió a una placa de Petri y los callos se sumergieron en la suspensión durante 15 minutos. A continuación, los tejidos de callo se secaron por transferencia a un papel de filtro y se transfirieron a medio de cocultivo solidificado y se incubaron durante 3 días en la oscuridad a 25 °C. Los callos cocultivados crecieron sobre 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 periodo, rápidamente se desarrollaron islas de callos resistentes a crecimiento. Después de la transferencia de este material a un medio de regeneración e incubación a la luz, el potencial embriogénico se liberó y se desarrollaron brotes las siguientes cuatro a cinco semanas. Los brotes se extirparon de los callos y se incubaron durante 2 a 3 semanas en un medio que contenía auxina desde el cual se transfirieron al suelo. Los brotes endurecidos se cultivaron a alta humedad y días cortos en un invernadero.

Se generaron aproximadamente 35 transformantes de arroz T0 independientes para una construcción. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero. Después de un análisis PCR cuantitativo para verificar el número de copias del inserto de ADN-T, solo una única copia de plantas transgénicas que presentaban tolerancia al agente de selección se mantuvieron para cosecha de la semilla T1. Después las semillas se cosecharon de tres a cinco meses después del trasplante. El procedimiento produjo transformantes de sitio único a una tasa de alrededor de 50 % (Aldemita and Hodges1996, Chan et al. 1993, Hiei et al. 1994).

Ejemplo 10: Procedimiento de evaluación fenotípica

10.1 1 Configuración de evaluación

Se generaron aproximadamente 35 transformantes independientes de arroz T0. Los transformantes primarios se transfirieron de una cámara de cultivo de tejido a un invernadero para el cultivo y la cosecha de las semillas T1. Siete eventos, de los cuales la progenie T1 segrega 3: 1 para la presencia / ausencia del transgén, se retuvieron. Para cada uno de estos eventos, aproximadamente 10 plántulas T1 que contienen el transgén (hetero y homocigotos) y aproximadamente 10 plántulas T1 que carecen del transgén (nulicigotos) fueron seleccionados mediante el control de la expresión del marcador visual. Las plantas transgénicas y los nulicigotos correspondientes se hicieron crecer lado a lado en posiciones aleatorias. Condiciones de invernadero fueron días cortos (12 horas de luz), 28 °C en la luz y 22 °C en la oscuridad, y una humedad relativa de 70 %.

Todos los eventos T1 se evaluaron después en la generación T2 siguiendo el mismo procedimiento de evaluación como para la generación T1. Desde la etapa de siembra hasta la etapa de madurez se pasaron las plantas varias veces a través de un gabinete de imágenes digitales. En cada punto de tiempo, se tomaron imágenes digitales (2048x1536 píxeles, 16 millones de colores) de cada planta desde por lo menos 6 ángulos diferentes.

Selección de disponibilidad reducida de nutrientes (nitrógeno)

Las plantas de seis eventos (semillas T2) se cultivaron en tierra para macetas en condiciones normales a excepción de la solución nutritiva. Las macetas se regaron desde el trasplante a la maduración con una solución nutriente específico que contiene nitrógeno reducido N (N) contenido, por lo general entre 7 a 8 veces menos. El resto del cultivo (maduración de la planta, cosecha de semilla) fue el mismo que para las plantas no cultivadas bajo estrés abiótico. Después se midieron los parámetros relacionados con las semillas.

10.2 Análisis estadístico: Prueba F

Un ANOVA de dos factores (análisis de variantes) se utilizó como un modelo estadístico para la evaluación global de las características fenotípicas de la planta. Se llevó a cabo una prueba F en todos los parámetros medidos de todas las plantas de todos los eventos transformados con el gen de la presente divulgación. La prueba F se llevó a cabo para comprobar si el efecto del gen sobre todos los eventos de transformación y para verificar un efecto general del gen, también conocido como un efecto de gen global. El umbral de significación para un verdadero efecto génica global se fijó en un nivel de probabilidad del 5 % para la prueba F. Un valor significativo de la prueba F apunta a un efecto génico, lo que significa que no es sólo la mera presencia o la posición del gen que causa las diferencias en el fenotipo.

10.3 Parámetros medidos

10.3.1 Medición de parámetros relacionados con la biomasa

Desde la etapa de siembra hasta la etapa de madurez se pasaron las plantas varias veces a través de un gabinete de imágenes digitales. En cada punto de tiempo, se tomaron imágenes digitales (2048x1536 píxeles, 16 millones de colores) de cada planta desde por lo menos 6 ángulos diferentes.

El área por encima del suelo de la planta (o biomasa foliar) se determinó contando el número total de píxeles en las imágenes digitales de partes de la planta sobre el suelo discriminadas desde el fondo. Este valor se promedió para las fotografías tomadas en el mismo punto de tiempo desde los diferentes ángulos y se convierte en un valor superficie física expresada en mm cuadrados por calibración. Los experimentos muestran que el área de la planta por encima del suelo medida de esta manera se correlaciona con la biomasa de las partes de la planta por encima del suelo. El área por encima del suelo es el punto de tiempo en el que la planta había alcanzado su máxima biomasa foliar. El vigor temprano es la planta (plántulas) área sobre el suelo tres semanas después de la germinación.

Un parámetro adicional se calcula a partir de las imágenes digitales de las plantas: el índice de verdor. Para cada píxel que pertenece al objeto de la planta sobre la imagen se calcula la proporción del valor verde frente al valor rojo (en el modelo RGB por el que se codifica el color). El índice de verdor se expresa como el porcentaje de píxeles para el que la proporción de verde con respecto a rojo supera un umbral determinado. En condiciones de crecimiento con disponibilidad reducida de nutrientes se midió el índice de verdor de las plantas en la última imagen antes de la floración.

Para medir los parámetros relacionados con la raíz-, las plantas se cultivaron en macetas especialmente diseñados con fondos transparentes para permitir la visualización de las raíces. Una cámara digital grabó imágenes a través de la parte inferior de la olla durante el crecimiento de la planta. Las características de la raíz tales como el área total proyectada (que puede correlacionarse con el volumen total de la raíz), el diámetro medio y la longitud de las raíces por encima de cierto umbral de espesor (longitud de raíces gruesas, o longitud de la raíz de espesor) se dedujeron de la imagen utilizando de software apropiado. El aumento de la biomasa de la raíz se expresa como un incremento en la biomasa total de las raíces (medida como máximo biomasa de raíces observadas durante la vida útil de una planta); o como un aumento en el índice de la raíz / tallo (medido como la relación entre la masa de raíces y disparar masa en el período de crecimiento activo de la raíz y brote).

10.3.2 Mediciones de parámetros relacionados con las semillas

Las panículas primarias maduras se cosecharon, contaron, embolsaron, marcaron con código de barras y después se secaron durante tres días en un horno a 37 °C. Las panículas se trillaron y todas las semillas se recogieron y se contaron. Las cáscaras llenas se separan de las vacías utilizando un dispositivo de soplado de aire. Las cáscaras vacías se desecharon 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 quedaron después de la etapa de separación. El rendimiento total de semillas se midió pesando todas las vainas llenas cosechadas de una planta. Número total de semillas por planta se midió contando el número de cáscaras cosechadas de una planta. El peso de mil granos (TKW) se extrapola a partir del número de semillas llenas contadas y su peso total. El índice de recolección (IR) es la relación entre el rendimiento de semilla total y el área por encima del suelo (mm^2), multiplicado por un factor de 106. El número total de flores por panícula es la relación entre el número total de semillas y el número de maduro panículas primarias. La tasa de llenado de semilla es la proporción (expresado en %) del número de semillas llenas sobre el número total de semillas (o floretes).

Ejemplo 11: Resultados de las plantas de arroz transgénicas que expresan la secuencia de ácido nucleico HDZip hox5 de clase I cultivadas bajo condiciones de disponibilidad reducida de nutrientes

Los resultados de la evaluación de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico útil en la realización de los procedimientos de la invención en condiciones de disponibilidad reducida de nutrientes en condiciones de estrés se presentan en la Tabla F. Se muestra la diferencia porcentual entre los transgénicos y los nulicigotos correspondientes.

- 5 **Tabla F:** Resultados de la evaluación de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico útil en la realización de los procedimientos de la invención, cultivadas en condiciones de disponibilidad reducida de nutrientes.

Rasgo	% Diferencia en T1
Rendimiento total de semillas por planta	20
Número de semillas llenas	19
Tasa de llenado de semillas	4
Número de flores por panícula	9
Índice de recolección	12
Índice de verdor antes de la floración	11

10 **Ejemplo 12: Transformación de maíz, trigo, soja, canola, alfalfa, algodón, con secuencias útiles en los procedimientos de la invención**

Transformación de maíz

La transformación de maíz (*Zea mays*) se realizó con una modificación del procedimiento descrito por Ishida et al. (1996) Nature Biotech 14(6): 745 - 50. La transformación es dependiente del genotipo en maíz y solo son genotipos específicos susceptibles a transformación y regeneración. Como progenitores, la línea endogámica A188 (Universidad de Minnesota) o híbridos con A188, son buenas fuentes de material donante para la transformación, pero también pueden usarse otros genotipos satisfactoriamente. Se cosecharon espigas de plantas de maíz aproximadamente 11 días después de la polinización (DDP) cuando la longitud del embrión inmaduro era de aproximadamente 1 a 1, 2 mm. Los embriones inmaduros se co-cultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través organogénesis. Los embriones cortados se cultivaron en medio de inducción de callo y después en medio de regeneración de maíz que contenía el agente de selección (por ejemplo imidazolinona, pero pueden usarse diversos marcadores de selección). Las placas de Petri se incubaron a la luz a 25 °C durante 2-3 semanas, o hasta que se desarrollasen los brotes. Los brotes verdes se transfirieron desde cada embrión al medio de enraizamiento de maíz y se incubaron a 25 °C durante 2-3 semanas, hasta el desarrollo de las raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de trigo

La transformación de trigo se realizó con el procedimiento descrito por Ishida et al. (1996) Nature Biotech 14(6): 745 - 50. La variedad de cultivo Bobwhite (disponible de CIMMYT, México) se utiliza normalmente para transformación. Embriones inmaduros se co-cultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través de organogénesis. Después de la incubación con *Agrobacterium*, los embriones cortados se cultivan in vitro en medio de inducción de callo y después en medio de regeneración que contenía el agente de selección (por ejemplo imidazolinona, pero pueden usarse diversos marcadores de selección). Las placas de Petri se incubaron a la luz a 25 °C durante 2-3 semanas, o hasta que se desarrollasen los brotes. Los brotes verdes se transfieren desde cada embrión al medio de enraizamiento y se incuban a 25 °C durante 2-3 semanas, hasta el desarrollo de las raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de soja

40 La transformación de soja se realiza de acuerdo con una modificación del procedimiento descrito en el Texas A&M patente de Estados Unidos 5.164.310. Diversas variedades comerciales de soja son susceptibles a transformación mediante este procedimiento. La variedad de cultivo Jack (disponible de la fundación Illinois Seed) se utiliza normalmente para transformación. Las semillas de soja se esterilizan para siembra in vitro. El hipocótilo, la radícula y

un cotiledón se extirpan de plántulas jóvenes de siete días de vida. El epicótilo y el cotiledón restante se cultivan adicionalmente para desarrollar nodos axilares. Estos nodos axilares se extirpan y se incuban con *Agrobacterium tumefaciens* que contenía el vector de expresión. Después de tratamiento de co-cultivo, los explantes se lavan y se transfieren al medio de selección. Los brotes regenerados se extirpan y se colocan en un medio de elongación de brote. Los brotes no mayores de 1 cm se colocaron se colocan en medio de enraizamiento hasta que se desarrollaron raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación colza/canola

Peciolos cotiledonarios e hipocótilos de plántulas jóvenes de 5-6 días de vida se utilizan como explantes para el cultivo tisular y se transforman de acuerdo con Babic et al. (1998, Plant Cell Rep 17: 183 - 188). La variedad de cultivo comercial Westar (Agriculture Canada) es la variedad convencional usada para la transformación, pero también pueden usarse otras variedades. Las semillas de canola se esterilizan en la superficie para la siembra in vitro. Explantes de peciolos cotiledonarios con el cotiledón unido se extirpan de plántulas in vitro y se inoculan con *Agrobacterium* (que contenía el vector de expresión) sumergiendo el extremo cortado del explante del peciolo en la suspensión bacteriana. Después los explantes se cultivan durante 2 días en medio MSBAP-3 que contenía BAP 3 mg/l, sacarosa al 3 %, Fitagar al 0, 7 % a 23 °C, 16 h de luz. Después de dos días de co-cultivo con *Agrobacterium*, los explantes de peciolo se transfieren a medio MSBAP-3 que contenía BAP 3 mg/l, cefotaxima, carbenicilina o timentina (300 mg/l) durante 7 días y después se cultivan en medio MSBAP-3 con cefotaxima, carbenicilina o timentina y agente de selección hasta la regeneración del brote. Cuando los brotes tenían una longitud de 5-10 mm, se cortaron y se transfirieron al medio de elongación de brote (MSBAP-0, 5 que contenía BAP 0, 5 mg/l). Los brotes de aproximadamente 2 cm de longitud se transfieren al medio de enraizamiento (MS0) para la inducción de raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de alfalfa

Un clon de regeneración de alfalfa (*Medicago sativa*) se transforma utilizando el procedimiento de (McKersie y col., 1999 Plant Physiol 119:) 839 - 847). La regeneración y transformación de la alfalfa es dependiente del genotipo y por lo tanto se requiere una planta regenerada. Se han descrito procedimientos para obtener plantas regeneradas. Por ejemplo, éstas pueden seleccionarse de variedades de cultivo Rangelander (Agriculture Canada) o de cualquier otra variedad comercial de alfalfa como describen Brown DCW y A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111 - 112). Como alternativa, la variedad RA3 (Universidad de Wisconsin) se ha seleccionado para su uso en cultivo tisular (Walker et al., 1978 Am J Bot 65:654 - 659). Explantes de peciolo se co-cultivan con un cultivo durante una noche de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie y col., 1999 839-847) o LBA4404 que contiene el vector de expresión. Los explantes se co-cultivan durante 3 días en la oscuridad en un medio de inducción SH que contenía Pro 288 mg/l, tioprolina 53 mg/l, K2SO4 4, 35 g/l y 100 µm de acetosiringinona. Los explantes se lavan en medio Murashige-Skoog de fuerza media (Murashige y Skoog, 1962) y se siembran en placas en el mismo medio de inducción SH con acetosiringinona pero con un agente de selección adecuado y un antibiótico adecuado para inhibir el crecimiento de *Agrobacterium*. Después de varias semanas, los embriones somáticos se transfirieron al medio de desarrollo BOi2Y que no contenía reguladores de crecimiento, sin antibióticos, y sacarosa 50 g/l. Posteriormente, los embriones somáticos germinaron en medio Murashige-Skoog de fuerza media. Las plántulas enraizadas se trasplantaron en macetas y se cultivaron en un invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de algodón

La transformación de algodón (*Gossypium hirsutum* L.) se realiza utilizando *Agrobacterium tumefaciens*, en explantes de hipocótilos. Las variedades de cultivo comerciales tales como Coker 130 o Coker 312 (Seedco, Lubbock, TX) son variedades estándar utilizadas para la transformación, pero también se pueden utilizar otras variedades. Se esteriliza la superficie de las semillas y germinan en la oscuridad. Explantes de hipocótilo se cortan de las plántulas germinadas a longitudes de aproximadamente 1 - 1,5 centímetro. El explante de hipocótilo se sumerge en el inóculo de *Agrobacterium tumefaciens* que contiene el vector de expresión, durante 5 minutos y luego se cocultiva durante aproximadamente 48 horas en MS +1,8 mg / l KNO3 + 2% de glucosa a 24 ° C, en la oscuridad. Los explantes se transfieren al mismo medio que contiene marcadores de selección de bacterias y vegetales apropiadas (renovado varias veces), hasta que se vean callos embriogénicos. Los callos se separan y se subcultivaron hasta que aparecen los embriones somáticos. Las plántulas derivadas de los embriones somáticos se maduran en un medio de enraizamiento hasta que se desarrollan las raíces. Los brotes enraizados se trasplantaron al suelo para macetas en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

LISTADO DE SECUENCIAS

<110> CropDesign N.V.

5 <120> Plantas que tienen características mejoradas de crecimiento en condiciones de disponibilidad reducida de nutrientes y un procedimiento para fabricar las mismas

<130> PF60795

<150> US 11/799.083

10 <151> 2007 - 04 - 30

<160> 52

<170> PatentIn versión 3.3

15

<210> 1

<211> 1050

<212> ADN

<213> Oryza sativa

20

<400> 1

atggateccg	gccgcgtcgt	gttcgaactcc	ggcgtggcgc	ggcggggcgtg	ccccggcggc	60
gogcagatgc	ttctcttcgg	cggcggcggc	agcgccaaca	gcggcggtt	cttcgaggc	120
gtgcgcgcgc	cggtgctggg	gatggatgaa	tcgcggtcgt	cgctcgtcgc	ggcgggggcg	180
ggggcgaaac	ggcgttctt	caagacgcac	gaggagctcc	tggaggagga	gtactacgac	240
gagcagcgcc	cggagaagaa	gcggcggtg	acggcggaac	aggtgcagat	gtcggagcgg	300
agcttcgagg	aggagaacaa	gctggagccg	gagcggaaga	cggagctcgc	ccgccgcctc	360
ggcatgcgcc	cccggcaggt	cgccgtctgg	ttccagaacc	gccgcgcccg	ctggaagacc	420
aagcagctcg	agcagcactt	cgaccgcctc	aaggccgcct	acgacgcctt	cgccgcgcac	480
caccatgcgc	tcctctccga	caacgaccgc	ctccgcgcgc	aggtaatctc	attaaccgag	540
aagctgcaag	acaaggagac	gtcgcgcgtg	tcggcgacca	tcaccaccgc	ggcgaggag	600
gtcgaccagc	cggacgaaca	cacggaggcc	gcgtcaacca	ccggttctgc	caccgtcgac	660
ggcgcatctg	cggcgccacc	gcccggccac	cagcagccgc	cgcataaaga	tgatcttgtg	720
agcagcgccg	gcaccaacga	cgacggcgat	ggcggcgcgc	ccgtggtggt	cttcgacgtc	780
accgagggcg	ccaacgaccg	cctcagctgc	gagtcggcgt	acttcgcgga	cgccgcggag	840
gcgtaccagc	gcgactgcgc	cgggcactac	gccctctcgt	cggaggagga	ggaecgcggc	900
gcggtcagcg	acgagggcctg	cagcttcgac	ctccccgacg	ccgcgcgcgc	cgccgcgcgc	960
atgttcgcgc	ccgcgcgagt	tgtgcaccac	gacgcgcgcg	acgacgagga	ggcgcgagctc	1020
ggcagctcga	ccgcctgggt	ctggagctga				1050

25 <210> 2

<211> 349

<212> PRT

<213> Oryza sativa

ES 2 531 376 T3

<400> 2

```

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

Met Leu Glu Arg Ser Phe Glu Glu Glu Asn Lys Leu Glu Pro Glu Arg
100      105      110
Lys Thr Glu Leu Ala Arg Arg Leu Gly Met Ala Pro Arg Gln Val Ala
115      120      125
Val Trp Phe Gln Asn Arg Arg Ala Arg Trp Lys Thr Lys Gln Leu Glu
130      135      140
His Asp Phe Asp Arg Leu Lys Ala Ala Tyr Asp Ala Leu Ala Ala Asp
145      150      155      160
His His Ala Leu Leu Ser Asp Asn Asp Arg Leu Arg Ala Gln Val Ile
165      170      175
Ser Leu Thr Glu Lys Leu Gln Asp Lys Glu Thr Ser Pro Ser Ser Ala
180      185      190
Thr Ile Thr Thr Ala Ala Gln Glu Val Asp Gln Pro Asp Glu His Thr
195      200      205
Glu Ala Ala Ser Thr Thr Gly Phe Ala Thr Val Asp Gly Ala Leu Ala
210      215      220
Ala Pro Pro Pro Gly His Gln Gln Pro Pro His Lys Asp Asp Leu Val
225      230      235      240
Ser Ser Gly Gly Thr Asn Asp Asp Gly Asp Gly Gly Ala Ala Val Val
245      250      255
Val Phe Asp Val Thr Glu Gly Ala Asn Asp Arg Leu Ser Cys Glu Ser
260      265      270
Ala Tyr Phe Ala Asp Ala Ala Glu Ala Tyr Glu Arg Asp Cys Ala Gly
275      280      285
His Tyr Ala Leu Ser Ser Glu Glu Glu Asp Gly Gly Ala Val Ser Asp
290      295      300
Glu Gly Cys Ser Phe Asp Leu Pro Asp Ala Ala Ala Ala Ala Ala
305      310      315      320
Met Phe Gly Ala Ala Gly Val Val His His Asp Ala Ala Asp Asp Glu
325      330      335
Glu Ala Gln Leu Gly Ser Trp Thr Ala Trp Phe Trp Ser
340      345

```

5

<210> 3

<211> 1197

<212> ADN

<213> Oryza sativa

10

<400> 3

ES 2 531 376 T3

```

atggagtcg gcgggtcat cttcagcag gcgggtccg gcgcgggca gatgctcttc 60
ttggactcg gcgctggcg cggcggcgtc ggcgcgggg ccatgttcca tcgagggggc 120
agaccgctg tcggcatgga ggaaggagg cgcggcgca agcggccctt cttcaccacc 180
cccgaccagc tcttcgaaga ggaglaclac gacgagcagc lcccgagaa gaagcggcgc 240
ctcaccgczg agcagglgca tctgctggag aggaagcttc aggaggagaa caagctggag 300
ccggagcgga agacggagct ggcgcggaag ctagggtctg acccgcgga ggtcgccgtg 360
tggttccaga accgcgcgc gcgctggaag accaagcagc tcgagcgga cttcgaccgc 420
ctcaaggcgt cgttcgacgc cctccgcgc gaccacgag cctcctcca ggacaaccac 480
cgctccact ctcaggtcat gtcgttgacc gagaagctgc aagagaagga gacgacgacc 540
gagggcagcg ccggcgcggc cgttgacgtc ccgggcttgc ctgcggcggc cgacgtgaag 600
gtcgccgtcc cggacgcga ggaaccggcg ctggaggagg cggcgcggc gtcgaggag 660
cagcaggagc agcaggtgaa ggccgaggac aggtgagca cggcgagcg cgggagcgcg 720
gtggtggaca cggacgcga actggtggtc ggtgcggc ccgaagcgc tcgcccgcgt 780
ggacagcagc gtggagtgt acttcccggg cggcgacgag taccacgact gcgtgatggg 840
ccccatggc cagcccgcg ggggcatcca gtcgaggag gacgacgcg ccggcagcga 900
cgagggtgc agctactac ccgacgacgc cggcgctctc ttcgcccacc acggccacca 960
ccaccaccac caacacgcg acgacgacga ggaggacggc cagcagatca gctgctggtg 1020
gatgtgcaac tagatttct cgcgcgcgc gtcgtcgtgc attcaattct cgtgttaaaa 1080
aaatcgttct ctttttcatt tttcgccttc ttgtctgta atgttgagtt tcgacgggt 1140
atgagaagga aggaggtgta tgcattgtga tggatatgta gggtaacaca tcggtga 1197

```

<210> 4

<211> 343

5 <212> PRT

<213> Oryza sativa

<400> 4

ES 2 531 376 T3

Met Glu Ser Gly Arg Leu Ile Phe Ser Thr Ala Gly Ser Gly Ala Gly
1 5 10 15
Gln Met Leu Phe Leu Asp Cys Gly Ala Gly Gly Val Gly Gly
20 25 30
Gly Ala Met Phe His Arg Gly Ala Arg Pro Val Leu Gly Met Glu Glu
35 40 45
Gly Gly Arg Gly Val Lys Arg Pro Phe Phe Thr Thr Pro Asp Glu Leu
50 55 60
Leu Glu Glu Glu Tyr Tyr Asp Glu Gln Leu Pro Glu Lys Lys Arg Arg
65 70 75 80
Leu Thr Pro Glu Gln Val His Leu Leu Glu Arg Ser Phe Glu Glu Glu
85 90 95
Asn Lys Leu Glu Pro Glu Arg Lys Thr Glu Leu Ala Arg Lys Leu Gly
100 105 110
Leu Gln Pro Arg Gln Val Ala Val Trp Phe Gln Asn Arg Arg Ala Arg
115 120 125
Trp Lys Thr Lys Gln Leu Glu Arg Asp Phe Asp Arg Leu Lys Ala Ser
130 135 140
Phe Asp Ala Leu Arg Ala Asp His Asp Ala Leu Gln Asp Asn His
145 150 155 160
Arg Leu His Ser Gln Val Met Ser Leu Thr Glu Lys Leu Gln Glu Lys
165 170 175
Glu Thr Thr Thr Glu Gly Ser Ala Gly Ala Ala Val Asp Val Pro Gly
180 185 190
Leu Pro Ala Ala Ala Asp Val Lys Val Ala Val Pro Asp Ala Glu Glu
195 200 205
Pro Ala Leu Glu Glu Ala Ala Ala Phe Glu Glu Gln Gln Glu Gln
210 215 220
Gln Val Lys Ala Glu Asp Arg Leu Ser Thr Gly Ser Gly Gly Ser Ala
225 230 235 240
Val Val Asp Thr Asp Ala Gln Leu Val Val Gly Cys Gly Arg Gln His
245 250 255
Leu Ala Ala Val Asp Ser Ser Val Glu Ser Tyr Phe Pro Gly Gly Asp
260 265 270
Glu Tyr His Asp Cys Val Met Gly Pro Met Asp His Ala Ala Gly Gly
275 280 285
Ile Gln Ser Glu Glu Asp Asp Gly Ala Gly Ser Asp Glu Gly Cys Ser
290 295 300
Tyr Tyr Ala Asp Asp Ala Gly Val Leu Phe Ala Asp His Gly His His
305 310 315 320
His His His Gln His Ala Asp Asp Asp Glu Glu Asp Gly Gln Gln Ile
325 330 335
Ser Cys Trp Trp Met Trp Asn
340

<210> 5

<211> 819

5 <212> ADN

<213> Zea mays

<400> 5

10 atggatccga ggcgggtcag tttegactct ggccggcgcc gccggggcgg ccggcgccag 60
atgctgctct tcggcgccgg aggcagcgcc aacagcaacg gcttcttcgg aggtgttccg 120
atggcgctcc tggcatgga ccacgcgacg ccggtgggca agcgccctt cttcacgaca 180
cacgaggagc tctagagga ggagtactac gacgagcagg ccgggagaa gaagcgccga 240
ctgacggcgg agcaggtgca gctgctggag ccgagcttcg aagaagagaa caagctggag 300
ccggagcgca agaccgagct ggcctgcgcg ctggggatgg cgcgccgcca ggtagctgtt 360
tggttccaga accgcccgcg gcgctggaag accaagcaac tcgagacga ctatgaccgc 420
ctcaagcctg cttacgacgc actcgcgcgc gaccaccagg gcctcctggc cgacaacgat 480
aacctccggg cacaggtgat ctcctgacg gagaagctgc aaggcaagga gacatccccg 540
tcggcaacca ctgctgccc agaggtcgac cagccagacg aacacacgc tgtgtcaggc 600
acggaagaac tgcggcgca gcagctcaag gacaacctcc acagcagcg cgactgcact 660
ggccatgcca cctctcttc ggaagaagac gacggtggcg tggtcagtga ccaggcctgc 720
agcttcgctc tccgggatgc catgttcgct gccgggttca cccaccatgg ccggcaggag 780
gtgcagctgg ccaactggac atccatgttc tggaactga 819

ES 2 531 376 T3

<210> 6

<211> 272

<212> PRT

5 <213> Zea mays

<400> 6

```

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

```

10

<210> 7

<211> 993

<212> ADN

<213> Zea mays

15

<400> 7

ES 2 531 376 T3

```

atggagtctg gacggctcat cttcaacggc cggggtcttg ggcgggggca gatgctcttc      60
ctcgactcgc ggcgaggcgg cgggtccggc ggcggttgtt tccatcgagg cgggagaccg      120
atgcttgccc ttgaagaagg ggcgggggta aaacggccct tcttcacctc gcccgacgag      180
ctcctcgagg aagagtacta cgacgagcag ctgccggaga agaagcgcgc cctcacccca      240
gagcagctgc ttctgctgga gaggagcttc gaggaggaga acaagctgga gccggagcgc      300
aagacgggagc tggcgcgcaa gctgggcctg cagcctcgcc aggtggccgt ctggttccag      360
aaccgcgcgc ccgggtgga gaccaagcag ctcgagcgcg acttcgacgc cctcaaggcc      420
tccttcgacg ctctccgagc ggaccacgac gccctcctcc aggacaacaa ccgcctccgc      480
tcacagcttg tctcgttgac cgagaagctg caagagaagg aggatgcgac ggaggcgggc      540
gccaccgctg acaccgcgcg gccggcggtg gacgtcgagg ctccctgggc cgacgacgtc      600
gaggagccag cagagcctgc ggcgacgttc gagggtgctg aggaggtgaa gtcggaggac      660
aggctgagca ccggcagcgg cgggagcgcg gtggtggacg cggacgcgct gctgtacggc      720
aggttcgcgc cggcagttga tagcagcgtg gagtcgtact tccccggcgg cgaggaccac      780
taccacgact gcgggacgat gggccccgtg aatcatggcg ccggaggagg catccagtcg      840
gacgacgacg gcgcggcgag cgacgagggg tgcagctact acgcgcgacg agccgcgcgc      900
gccgcgcgcg cgttcttcgc cggacacgac acccaccacc acggggacga ggacgaggac      960
gccggccaga tcagctggcg gatgtggaac tag                                     993

```

<210> 8

<211> 330

5 <212> PRT

<213> Zea mays

<400> 8

```

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

```

10

ES 2 531 376 T3

260										265				270			
Gly	Ala	Gly	Gly	Gly	Ile	Gln	Ser	Asp	Asp	Asp	Asp	Gly	Ala	Gly	Ser	Asp	
275										280				285			
Glu	Gly	Cys	Ser	Tyr	Tyr	Ala	Asp	Glu	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	
290										295				300			
Phe	Phe	Ala	Gly	His	Ala	Thr	His	His	His	Ala	Asp	Glu	Asp	Glu	Asp	Asp	
305										310				315			
Ala	Gly	Gln	Ile	Ser	Trp	Trp	Met	Trp	Asn								
325										330							

<210> 9

<211> 825

5 <212> ADN

<213> Saccharum officinarum

<400> 9

atggtaccca	gctgggtcag	tttcaactcc	ggcgggcgcg	ggcggggcgg	cggcggcagc	60
cagatgtctc	tcttcggcgg	cggaggcagc	gccaacagca	acggtctctt	ccgagtggtt	120
ccgatgcggg	tctctggcat	ggacgcagcg	atcgctctgg	gcaagcggcc	ctctcttacc	180
acacacacgg	agctctctga	ggagaggtac	cgcgcgcagc	ggcgcgccga	gaagaagcgc	240
cgtctgaacg	cggagcaggt	gcagctgctg	gagcgcagct	tcgagggaag	gaacaagctg	300
gagcccgacg	gcgaagacga	cctgctctgc	cgctctcgga	tggccgcccg	ccagctgggc	360
gtctggttcc	agaaacgcgc	cgctgcctgg	aagaccgaac	agctcgagac	cgaactgac	420
cacctcaagg	ctgcctacga	cgcgctcggc	ggcagccacc	aggcctgcat	ggccgacaac	480
gatagcctcc	gggcacaggt	gtctctccct	acagagaagc	tcgaaggcaa	ggagacatcc	540
ccgtgcggca	ccaactgcgc	ccaagaggtc	caccagccag	acgaacacac	cgcgcgctca	600
ggcactgaga	aactgctlgc	gcagcagctc	aagcagcacc	tcacacagcag	cggcgcactgc	660
actggccctc	gtgccctcgc	ctcagagcaa	gaagatgtgt	ctgtgtgtcag	cgacgagggc	720
agcttctcgc	tccgggatac	catgtttgtc	gcgggggtca	cccaccatgg	cgccgcagcc	780
gaggagcgac	agctggccaa	ctggacatcc	tgtttctgga	actga		825

$\langle 210 \rangle$ 10

<211> 274

<212> PRT

15 <213> Saccharum officinarum

<400> 10

ES 2 531 376 T3

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

<210> 11

5 <211> 828

<212> ADN

<213> Sorghum bicolor

<400> 11

10

atgggaccca ggcgggtcag ttctgactcc ggccggcgcc ggccggggcg cgccggcgcc 60
ggcgccgaga tgctgctctt cggcgccgga ggcagcgcca acagcaacgg cttcttccga 120
ggtgttccga tggcggtcct gggcatggac gacgcgacgc gcgtgggcaa gcggcctttc 180
ttcaccacgc acgaggagct cctggaggag gactactacg acgagcaggc gcccgagaag 240
aagcgccctc tgacggcgga gcaggtgcag ctgctggagc gtagcttcga ggaagagaac 300
aagctggagc cggagcgcaa gaccgagctg gctcgccgcc tcgggatggc gcctcgccag 360
gtggccgtct ggttccagaa ccgcccgcgc cgtgggaaga ctaagcagct cgagaccgac 420
tatgaccgcc tcaaggctgc ctacgacgag ctgcgccgac accaccaggc cctcctggcc 480
gacaacgata gctccgggc acaggtgac tccctaacgg ataagctgca acgcaaggag 540
acatccccgt cggcgaccac tgcctcccaa gaggtcgacc agccagacga acacacgct 600
gcgtcaggca ctgagaaact gctggtgcag cagctcaagg acgacctcca cagcagcgcc 660
gacttcaact gccatggcgc cctctcttca gaggaagagg atggtggtgt ggtcagcgac 720
gagggctgca gctttgatct ccgggatgac atgttcgctg ccggggtcac ccaccatggc 780
gocgagcagg cgcagctggc caactggaca tctgtgttct ggaactga 828

<210> 12

<211> 275

15 <212> PRT

ES 2 531 376 T3

<213> Sorghum bicolor

<400> 12

5

```

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

```

<210> 13

<211> 1023

10

<212> ADN

<213> Triticum aestivum

<400> 13

ES 2 531 376 T3

```

atggagcccg gccggctcat cttcaacacg tcgggctccg gcaacggaca gatgctcttc 60
atggactcgg gcggggggcg catcgccggc gcggccggca tgttccatcg aggggtgaga 120
ccggtcctcg gcggcatgga agaagggcgc ggcgtagaag gcccttctt cacctcgccg 180
gatgacatgc tcgaggagga gtactacgac gagcagctcc cggagaagaa gcggcgccctc 240
accccgagc aggtccacct gctggagagg agcttcgagg aggagaacaa gctggagccg 300
gagaggaaga cggagctggc ccgcaagctc gggtcgagc cagccaggt gcccgctctg 360
ttccagaacc gccggcgccg gtggaagaca aagacgtcg agcgcgactt cgaaccgctc 420
aagcgctct tcgacgacct ccggggcgac cagcagcccc tcttccagga caaccaccgg 480
ctccggtcac aggtggtaac gttgacggag aagatgcaag ataaggaggc gccggaaggc 540
agcttcggcg cagccgcca cgcctcgag ccggagcagg ccggcgcgga gccgaaggct 600
tcttgcccg acccgagga gcaggccgg gcaggagg cgctcgaggt ggtgcagcag 660
cagctgcacg tgaaggacga ggagaggctg agcccgggga gcggcgggag ccgggtgctg 720
gacgcagagg acgcgctgct cgggagcgga tgcggcctcg ccggcggtgt ggacagcagc 780
gtggactcgt actgcttccc gggggggccc gcggcgacg agtaccacga g-gcgtggtg 840
ggccccgtgg cggggcgcat ccagtcggag gaggacgacg gcgcgggcag cgaacgaggc 900
tgcagctact accccgacga cggcgccgct ttcttcgccc ccgcgcaag qcacggccac 960
catcgcaagg acgacgacga tcagcaggac gacggccaga tcagctact gatgtggaac 1020
tag 1023

```

<210> 14

<211> 340

5 <212> PRT

<213> Triticum aestivum

<400> 14

```

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

```

10

ES 2 531 376 T3

Glu Glu Glu Tyr Tyr Asp Glu Gln Leu Pro Glu Lys Lys Arg Arg Leu
65 70 75 80
Thr Pro Glu Gln Val His Leu Leu Glu Arg Ser Phe Glu Glu Glu Asn
85 90 95
Lys Leu Glu Pro Glu Arg Lys Thr Glu Leu Ala Arg Lys Leu Gly Leu
100 105 110
Gln Pro Arg Gln Val Ala Val Trp Phe Gln Asn Arg Arg Ala Arg Trp
115 120 125
Lys Thr Lys Thr Leu Glu Arg Asp Phe Asp Arg Leu Lys Ala Ser Phe
130 135 140
Asp Ala Leu Arg Ala Asp His Asp Ala Leu Leu Gln Asp Asn His Arg
145 150 155 160
Leu Arg Ser Gln Val Val Thr Leu Thr Glu Lys Met Gln Asp Lys Glu
165 170 175
Ala Pro Glu Gly Ser Phe Gly Ala Ala Ala Asp Ala Ser Glu Pro Glu
180 185 190
Gln Ala Ala Ala Glu Ala Lys Ala Ser Leu Ala Asp Ala Glu Glu Gln
195 200 205
Ala Ala Ala Ala Glu Ala Phe Gln Val Val Gln Gln Gln Leu His Val
210 215 220
Lys Asp Glu Glu Arg Leu Ser Pro Gly Ser Gly Gly Ser Ala Val Leu
225 230 235 240
Asp Ala Arg Asp Ala Leu Leu Gly Ser Gly Cys Gly Leu Ala Gly Val
245 250 255
Val Asp Ser Ser Val Asp Ser Tyr Cys Phe Pro Gly Gly Ala Gly Gly
260 265 270
Asp Glu Tyr His Glu Cys Val Val Gly Pro Val Ala Gly Gly Ile Gln
275 280 285
Ser Glu Glu Asp Asp Gly Ala Gly Ser Asp Glu Gly Cys Ser Tyr Tyr
290 295 300
Pro Asp Asp Ala Ala Val Phe Phe Ala Ala Ala Gln Gly His Gly His
305 310 315 320
His Arg Thr Asp Asp Asp Asp Gln Gln Asp Asp Gly Gln Ile Ser Tyr
325 330 335
Trp Met Trp Asn
340

<210> 15

<211> 819

5 <212> ADN

<213> Arabidopsis thaliana

<400> 15

atggaatcca attcgttttt cttcgatcca tctgcttcac acggcaacag catgtttctt 60
cttgggaatc tcaatccgt cgtccaagga ggaggagcaa gatcgatgat gaacatggag 120
gaaacttoga agcgaaggcc cttctttagc tcccttgagg atctctacga cgatgacttt 180
tacgacgacc agttgcctga aaagaagcgt cgcctcacta ccgaacaagt gcatctgctg 240
gagaaaaagt tccgagacaga gaacaagcta gagcctgaac gcaagactca gcttgccaag 300
aagcttgctc tacagccaag gcaagtggct gtctggtttc agaatgcccg agctcgttgg 360
aaaacaaaac agcttgagag agactacgat cttctcaagt ccacttacga ccaacttctt 420
tctaactacg actccatcgt catggacaac gataagctca gatecgaggt taattccctg 480
accgaaaagc ttcaggggcaa acaagagaca gctaataaac cacttggtca agtgcccgaa 540
ccaaaccaac ttgatccggt ttacattaat gcggcagcaa tcaaaaccga ggaccggtta 600
agttcaggga gcgttgggag cgcggtacta gacgacgacg cacttcaact actagacagc 660
tgtgactctt acttcccaag catcgtaccc atccaagaca acagcaacgc cagtgatcat 720
gacaatgacc ggagctgttt cgcgcagctc ttgtgtccca ccacttcacc gtgcgacgat 780
catcacgctg aatcattggc tttctgggga tggccttag 819

10

<210> 16

<211> 272

<212> PRT

<213> Arabidopsis thaliana

<400> 16

```

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

```

5

<210> 17

<211> 969

<212> ADN

10 <213> Daucus carota

<400> 17

```

atggcggttc ggagggtgtt ctatggggag ggagccaata cgacgtcggc tagcctgttg      60
tttcatagtc aaagaccaga gcccttcttt cttctctgcac cttctccttc tctaattggg      120
tcaaaatcca tggttagcct tcaagatgct aagcgaaaaa atccctacga tgggttcttt      180
atgcggtcat atgatgaaga agaaattggg gatgaagaat atgatgaata ctttcagcag      240
cctgagaaga agaggaggtt caaggtctgat caaatccagt ttcttgagaa aagttttgag      300
actgataaca agcttgagcc tgaagaaaaa gtccagcttg caaaagaact cggcttgacag      360
ccaagacagg ttgcgatatg gtttcagaac cgtcgagcac ggtggaagac caaaacacta      420
gaaaaagatt atgatgtatt gcaaaatagc tacaacagcc tcaaggctga ctatgacaat      480
ctacttgccg agaaagaaaa acttaaagcc gaggttctcg acctgacaga caagctactt      540
ctcaaaagag ataaggggag caagacagta gtttttgata agcaaaaggt gctctgacga      600
ttccaacaag aacgtgttag taatgacata tctgtgggtg aagtactcag taactcagtt      660
atggactgca agcaagaaga tcataactct gtgaaaagtg atgcagttga tcttgacagt      720
ccacactaca gtgatgaagt ctactccagt tttatggagc cagtggatcg ctttatgtt      780

```

ES 2 531 376 T3

```

tttgaacctg ctcagtcgga tatatctcaa gatgaagaag atgacatggg gaacaactta      840
ttttcccat catatcatgt tttctcaaag actgaagacg gtagttactc cgaccagcct      900
tcgaactctt cgtactttgg cttccagttt gaagatcata cgtttggctt ttggggtaact      960
gaattataa                                     969

```

<210> 18

<211> 322

5 <212> PRT

<213> *Daucus carota*

<400> 18

```

Met Ala Gly Arg Arg Val Phe Tyr Gly Glu Gly Ala Asn Thr Thr Ser
1          5          10          15
Ala Ser Leu Leu Phe His Ser Gln Arg Pro Glu Pro Phe Phe Leu Ser
20          25          30
Ala Pro Ser Pro Ser Leu Ile Gly Ser Lys Ser Met Val Ser Phe Gln
35          40          45
Asp Ala Lys Arg Lys Asn Pro Tyr Asp Gly Phe Phe Met Arg Ser Tyr
50          55          60
Asp Glu Glu Glu Ile Gly Asp Glu Glu Tyr Asp Glu Tyr Phe Gln Gln
65          70          75          80
Pro Glu Lys Lys Arg Arg Leu Lys Ala Asp Gln Ile Gln Phe Leu Glu
85          90          95
Lys Ser Phe Glu Thr Asp Asn Lys Leu Glu Pro Glu Arg Lys Val Gln
100          105          110
Leu Ala Lys Glu Leu Gly Leu Gln Pro Arg Gln Val Ala Ile Trp Phe
115          120          125
Gln Asn Arg Arg Ala Arg Trp Lys Thr Lys Thr Leu Glu Lys Asp Tyr
130          135          140
Asp Val Leu Gln Asn Ser Tyr Asn Ser Leu Lys Ala Asp Tyr Asp Asn
145          150          155          160
Leu Leu Ala Glu Lys Glu Lys Leu Lys Ala Glu Val Leu Asp Leu Thr
165          170          175
Asp Lys Leu Leu Leu Lys Glu Asp Lys Gly Ser Lys Thr Val Val Phe
180          185          190
Asp Lys Gln Lys Val Ser Ala Ala Phe Gln Gln Glu Arg Val Ser Asn
195          200          205
Asp Ile Ser Val Gly Glu Val Leu Ser Asn Ser Val Met Asp Cys Lys
210          215          220
Gln Glu Asp His Asn Ser Val Lys Ser Asp Ala Val Asp Ser Asp Ser
225          230          235          240
Pro His Tyr Ser Asp Glu Val Tyr Ser Ser Phe Met Glu Pro Val Asp
245          250          255
Arg Ser Tyr Val Phe Glu Pro Ala Gln Ser Asp Ile Ser Gln Asp Glu
260          265          270
Glu Asp Asp Met Gly Asn Asn Leu Phe Leu Pro Ser Tyr His Val Phe
275          280          285
Ser Lys Thr Glu Asp Gly Ser Tyr Ser Asp Gln Pro Ser Asn Ser Ser
290          295          300
Tyr Phe Gly Phe Pro Val Glu Asp His Thr Phe Gly Phe Trp Gly Thr
305          310          315          320
Glu Leu

```

10

<210> 19

<211> 1038

<212> ADN

15 <213> *Glycine max*

ES 2 531 376 T3

<400> 19

```

atggcgagt gcaagcttta tgcgggttca aacatgtcac ttctctccca aaacgaaagg      60
ctcccttgct cctctgaagt ccttgagtct ctttgggctc agacctctaa ccttgccttc      120
ttccaagggt caaaaccggt ggttgatttt gagaatgtaa gtgggagcag gatgacggat      180
aggcctttct ttcaagcgtt ggagaaggaa gagaactgtg atgaggatta cgaggggtgt      240
ttccaccaac cggggaagaa aaggaggctc acaagcgaac aagttcagtt ccttgaaagg      300
aactttgagg tagagaacaa gcttgaacct gaaaggaaag tccaacttgc aaaagagctt      360
ggcttgacgc caaggcaagt tgctatatgg ttccaaaacc gaagggaag gttcaagacc      420
aagcagctag aaaaagacta tggcgtgttg aaagctagtt atgacagact caaaagtgc      480
tatgaaagtc ttgttcaaga gaatgacaag ttaaaagcag aggtgaattc tctggagagc      540
aaattgattc ttagagataa agagaaggag gagaattcgg atgacaagtc atctcctgat      600
gatgctgtca attcttcttc accccacaac aacaaggagc ctatggattt attaatatt      660
tcaaaaaatg caacaacaac aacaacatct gaaaatggga ccaaagtgtt gtcaccactc      720
ccactcccta ttatggtaac atgctgcaag caagaagatg ccaactcagc caaaagtgat      780
gtccttgatt cggatagccc acattgcact tcattcgttg agcctgtctg ttcctctcat      840
gcctttgaac cagaagacca ctcaagaagc ttctccaag atgaaggaga taaccttagt      900
gaaaaccttt tgatgacctt ccttctctct tgttgcttac ctaagggttg agaacactgc      960
tatgacggcc ctctgaaaa ctcttgtaat ttgggcttcc aggttgagga tcaaaccttc     1020
tgtttctggc cctattga

```

5 <210> 20

<211> 345

<212> PRT

<213> Glycine max

10 <400> 20

```

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

```

ES 2 531 376 T3

			260					265					270			
Val	Glu	Pro	Ala	Asp	Ser	Ser	His	Ala	Phe	Glu	Pro	Glu	Asp	His	Ser	
		275					280					285				
Glu	Asp	Phe	Ser	Gln	Asp	Glu	Glu	Asp	Asn	Leu	Ser	Glu	Asn	Leu	Leu	
	290					295					300					
Met	Thr	Phe	Pro	Ser	Ser	Cys	Cys	Leu	Pro	Lys	Val	Gln	Glu	His	Cys	
305				310						315					320	
Tyr	Asp	Gly	Pro	Pro	Glu	Asn	Ser	Cys	Asn	Phe	Gly	Phe	Gln	Val	Glu	
			325						330					335		
Asp	Gln	Thr	Phe	Cys	Phe	Trp	Pro	Tyr								
			340					345								

<210> 21

<211> 858

5 <212> ADN

<213> Craterostigma plantagineum

<400> 21

atgaactctg	ctcggaattt	cttcgaccca	tcttcccacg	gcaacatgct	gcagttttct	60
gggaaacgcc	gcggcgattc	atccgttttc	cgaggaaaca	gctcgtcgtc	ggtgctgaac	120
atggagagca	gctcgtttaa	acgacagatt	ttcagcggcg	gcggcggcga	tgaattctac	180
gacgagcaat	actacgacga	cgagttgttg	ctcgagaaag	agcgcgcgac	caaccgcgag	240
caggttcact	tgcttgagaa	gagcttcgag	gctgagaaca	agcttgagcc	tgagcgaaag	300
gctgagctgc	cgaagaagct	cggattgcag	cgaggccaag	tgcgctattg	gtcccaaac	360
cgccgacac	ggtgaaagac	taagcagtta	cgagggact	acgacaagct	taagtcttcc	420
tatgattctc	ttctctcaac	ctacgactct	attcgccagg	aaacgacaa	gctcaaaagc	480
gagctccttt	cttcgaccca	gaatttgcaa	cccaagagcg	acgacgaccc	acggcgcgaa	540
alagglogaa	acttcagttc	atcgtcgctg	ctcgtcgagc	cgccctgagcc	ggccgtgctt	600
aagctgcagg	tgaaggttgc	ggaccgcttg	agcagcgggg	ggaacggcag	cgcgataatg	660
gacggcgacg	gacctcagca	gctcctcgac	gacagcggcg	actcgtactt	cgagaacgac	720
gagggaatcg	actgcgcctc	cgcaagtttg	gctgctgcga	aggaggacga	cggcagcgat	780
gaggggccgt	gttaacttcac	cgaggetctc	goggcggagg	aggaggaggc	gcgcgtttgct	840
tggtgtattt	ggtcttaa					858

<210> 22

<211> 285

<212> PRT

15 <213> Craterostigma plantagineum

<400> 22

ES 2 531 376 T3

```

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

```

<210> 23

5 <211> 813

<212> ADN

<213> Gossypium hirsutum

<220>

10 <221> misc_feature

<222> (808)..(808)

<223> n e s a, c, g, o t

<400> 23

15

ES 2 531 376 T3

```

atggagtctg gccgtctctt tttcaatccc tccactaccc accgcaacat gttgcttctc   60
gggaaacactg aacccatctt tcgaggggca agaacaatgg ttagcatgga ggaaaaacca  120
aagaagcgac tgttcttcag ctgcgcggag gatttgtagc acgaagagta ctaccgagag  180
cagttgcccg agaaaaagcg tcgccttacg tcggagcagg tctatctgct agagaagagc  240
tttgagcgag agaaacaagct ggagcgggag aggaagagcc agttggccaa gaagtttagga  300
ctgcaaccaa ggcaggtggc ggtatgggtc cagaaccgcc ctgcaaggct gaagacaaag  360
cagcttgaaa gggactatga cctcctcaaa tcttcctttg attcccttca gtccaattat  420
gacctattc tcaaagaaaa tgagaagctc aaatctgagg tagcttcctt gactgaaaaa  480
ctacaagcca aagatgtggc aacagaagca atagcaggcg aaaaggatga agggtttagca  540
gctgagatgg cctccgcctt ccaattcagt atgaaggtag aggaaccgtc tagtagcggc  600
agtgtcggaa gcgcgggtgt ggatgaggat gcccacagc tggtaggacag cggcaattcc  660
tactttccaa gcgatgaata ctccagagc attggccctt tcgatggggt ccagtcggaa  720
gatgaggatg gcagtgataa ttgcgggagt tactttctcg atgtgttcgc aaccacagag  780
cagggagcat taggattgtg ggcttgntc taa                               813

```

<210> 24

<211> 270

5 <212> PRT

<213> *Gossypium hirsutum*

<220>

<221> misc_feature

10 <222> (270)..(270)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 24

```

Met Glu Ser Gly Arg Leu Phe Phe Asn Pro Ser Thr Thr His Arg Asn
1           5           10           15
Met Leu Leu Leu Gly Asn Thr Glu Pro Ile Phe Arg Gly Ala Arg Thr
                20                25                30

```

15

ES 2 531 376 T3

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

<210> 25

<211> 858

5 <212> ADN

<213> Lycopersicon esculentum

<220>

<221> misc_feature

10 <222> (294) .. (294)

<223> n es a, c, g, or t

<400> 25

atgggacatc ggcatalatt ttccgacccg tctctgtgtc acggcaacat gctgttccct 60
 gggagccgag atcctgttct ccgaggacca agatcgacqa tgatgaagat ggaggactcc 120
 togaagaggc gacccttctt tagctcgccg gaggatctat atgacgagga atactacgac 180
 gagcagtcac cggagaagaa ggcgcgtctc actcctgagc aggtgcactt gttggagaag 240
 agctttgaga cagaaaacaa gctggagccc gagcgcaaaa ccagctggc ctanaagctg 300
 gggctgcagc ccagacaggt gctgttatgg ttccaaaacc gccgtgcccg gaggagacc 360
 aagcagctcg agagggatta tgatcagctc aaatcctctt atgactccct tctctctgat 420
 tttgactccg ttccgaaaaga taacgataag ctcaaactcg aggttgtttc attgatggaa 480
 aagttacagg ggaaagtggg tggaggagca gggggaaatg aaaaatctga catcttggag 540
 gtggatgcta tgacgatcct tcaagtgaag gtgaaggctg gggaccggtt gaggagtggc 600
 agtggctgga gcgcgggtgg agatgagcat agttcacagc tgggtggacag tggggactca 660
 tattttcaca ctgatcatga ggagtatcca gggcctggag gatgcaatgt tctccacccc 720
 atggatggtt tacaatcgga ggaagatgat ggtagtgatg atcatggcag ttgccatggc 780
 tacttctcta acgtcttctt ggcagaagag cagcaccatg aacaaggaga agagcctatt 840
 ggatggttct ggtcttaa 858

15

<210> 26

<211> 285

<212> PRT

<213> Lycopersicon esculentum

5

<220>

<221> misc_feature

<222> (98)..(98)

<223> Xaa puede ser cualquier aminoácido de origen natural

10

<400> 26

```

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

```

15 <210> 27

<211> 972

<212> ADN

<213> Lycopersicon esculentum

20 <400> 27

ES 2 531 376 T3

atggctccag	ggattctcta	tgettggttct	tctaatttcg	atggcgtttt	tactcaaaaa	60
cagagagacg	tgttttcttc	atctactgca	ccgaaagggc	atcttggttc	cctttttgcc	120
cctgcctctt	cttcttctaa	tttcttgga	tccagttcta	tggtaggttt	tcgcggtgtt	180
aatggaggga	agagatcatt	ctttgattcg	ttcgatcagg	atgacaatga	agctgatgaa	240
ttgggggaat	atcttcatca	agcggagaag	aagaggcgac	ttactgacaa	ccaagttcag	300
tttcttgaga	agagtttttg	ggaagagaac	aaacttgaac	cagaaagaaa	agttcagctt	360
gctaagaac	ttggtctgca	gcctcgccaa	attgcaattt	ggtttcagaa	tcgtcgtgcg	420
cgatggaaga	ctaagcagct	cgagaaagat	tatgatgaat	tgaggaatag	atacgatact	480
ctgaaatcaa	attacaataa	tcttctcaag	gaaaaagaag	atcttcgaac	tgaagttttc	540
cgtctcaccg	gtaagctgtt	tatcaaagag	aaaggaaatg	ggcaattgga	tttgcgcgat	600
gaacacaaac	actccaatgc	attggcaaaa	gaaaccgtgg	ttgatccaat	gtccaatgta	660
ccagctctgg	ttgttaagca	ccagcaggaa	gatttaagct	ctgctaagag	tcgatgttttc	720
gactcagaaa	gcccacgtta	caccagtaga	atgcattcct	cagtcgtaga	tcaggatgat	780
tctgctcgcg	catttgaaac	tgatcagtcg	gattcatctc	aggatgatga	tgaaaacttc	840
agcaagaata	tgctttctac	tgccaacctc	cttggcaaag	acgcggatga	tgattatccc	900
gcgacatcat	caaatttgag	ttactttgga	tttccagttg	aagaccaagg	ttttggtttc	960
tggaattatt	aa					972

5 <210> 28

<211> 323

<212> PRT

<213> Lycopersicon esculentum

10 <400> 28

ES 2 531 376 T3

Met Ala Pro Gly Ile Leu Tyr Gly Gly Ser Ser Asn Phe Asp Gly Val
1 5 10 15
Phe Thr Gln Lys Gln Arg Asp Val Phe Ser Ser Ser Thr Ala Pro Lys
20 25 30
Gly His Leu Gly Ser Leu Phe Ala Pro Ala Ser Ser Ser Ser Asn Phe
35 40 45
Leu Gly Ser Ser Ser Met Val Ser Phe Arg Gly Val Asn Gly Gly Lys
50 55 60
Arg Ser Phe Phe Asp Ser Phe Asp Gln Asp Asp Asn Glu Ala Asp Glu
65 70 75 80
Leu Gly Glu Tyr Leu His Gln Ala Glu Lys Lys Arg Arg Leu Thr Asp
85 90 95
Asn Gln Val Gln Phe Leu Glu Lys Ser Phe Gly Glu Glu Asn Lys Leu
100 105 110
Glu Pro Glu Arg Lys Val Gln Leu Ala Lys Glu Leu Gly Leu Gln Pro
115 120 125
Arg Gln Ile Ala Ile Trp Phe Gln Asn Arg Arg Ala Arg Trp Lys Thr
130 135 140
Lys Gln Leu Glu Lys Asp Tyr Asp Glu Leu Arg Asn Arg Tyr Asp Thr
145 150 155 160
Leu Lys Ser Asn Tyr Asn Asn Leu Leu Lys Glu Lys Glu Asp Leu Arg
165 170 175
Thr Glu Val Phe Arg Leu Thr Gly Lys Leu Phe Ile Lys Glu Lys Gly
180 185 190
Asn Gly Gln Leu Asp Leu Arg Asp Glu His Lys His Ser Asn Ala Leu
195 200 205
Ala Lys Glu Thr Val Val Asp Pro Met Ser Asn Val Pro Ala Leu Val
210 215 220
Val Lys His Gln Gln Glu Asp Leu Ser Ser Ala Lys Ser Asp Val Phe
225 230 235 240
Asp Ser Glu Ser Pro Arg Tyr Thr Ser Arg Met His Ser Ser Val Val
245 250 255
Asp Gln Asp Asp Ser Ala Arg Ala Phe Glu Thr Asp Gln Ser Asp Ser
260 265 270
Ser Gln Asp Asp Asp Glu Asn Phe Ser Lys Asn Met Leu Ser Thr Ala
275 280 285
Asn Leu Leu Gly Lys Asp Ala Asp Asp Asp Tyr Pro Ala Thr Ser Ser
290 295 300
Asn Leu Ser Tyr Phe Gly Phe Pro Val Glu Asp Gln Gly Phe Gly Phe
305 310 315 320
Trp Thr Tyr

<210> 29

<211> 1014

5 <212> ADN

<213> *Medicago sativa*

<400> 29

atggcgggtg ggagagtgtt ttcaaatggt cctgcaaata tttcaaatat aaatatgaat 60
atattgtctt agaatcaaca acaaaactct cgtggaaact ctcttcaaca acctcttgat 120
tctcttttcc ttctctctcc tgcctcttct tttgggttcaa gatctatggt gaggttttgaa 180
gatgttcaag gaaggaaaag gcgcaacagg tctttctttg gaggatttga ccttgacgaa 240
aacggacagc atgagatgga tgaqlactll catcaalccq agaaqaaaac qctctctcca 300
gtggatcaag ttcatgttct tgagaaaagc tttagaggagg acaacaaact tgaaccagag 360
aggaaaacca agctagctaa agaccttggt ttgcagccac ggcaagttgc tatttggttt 420
caaaaccgtc gtgcaagggt gaagactaaa cagcttgaga aggattatga tctcttaat 480
gatggttatg agtctcttaa gacagagtat gacaaccttc tcaaagagaa agatagggtta 540
caatctcagg tggcaagcct aactgaaaag gtacttgaaa gagagaaaaca agagggaaaa 600
ttcaaacaaag gtgaaagtga aacaaaggaa ttcttgaggg aaccaacaat taataagcct 660
ttggttgatt cagtttctga ggggtgaagg tccaaatgtt caattggtga ggcttotaat 720
aataataata ataataacaa acttgaagat attagtctag caaggagtga catattggat 780
tgtgaagtc cagctacac tgatggagtg ttagagacat gtgattcttc ctatgtatct 840
gaacctgaat atcaatcgga cctatcaca gatgaagaag atcacaattt attgcctcct 900
tacctcttta caaaactga agatgtgaat tactccgacc cgccacataa tcaacaagct 960
tatggatttc aagaggaaga tcatcatcaa gctctttggc cttggtctta cttag 1014

<210> 30

<211> 337

<212> PRT

5 <213> *Medicago sativa*

<400> 30

```

Met Ala Gly Gly Arg Val Phe Ser Asn Gly Pro Ala Asn Ile Ser Asn
1      5      10      15
Ile Asn Met Asn Ile Leu Leu Gln Asn Gln Gln Thr Pro Arg Gly
20      25      30
Asn Ser Ser Gln Gln Pro Leu Asp Ser Leu Phe Leu Ser Ser Ala
35      40      45
Ser Phe Phe Gly Ser Arg Ser Met Val Ser Phe Glu Asp Val Gln Gly
50      55      60
Arg Lys Arg Arg Asn Arg Ser Phe Phe Gly Gly Phe Asp Leu Asp Glu
65      70      75      80
Asn Gly Glu Asp Glu Met Asp Glu Tyr Phe His Gln Ser Glu Lys Lys
85      90      95
Arg Arg Leu Ser Val Asp Gln Val Gln Phe Leu Glu Lys Ser Phe Glu
100     105     110
Glu Asp Asn Lys Leu Glu Pro Glu Arg Lys Thr Lys Leu Ala Lys Asp
115     120     125
Leu Gly Leu Gln Pro Arg Gln Val Ala Ile Trp Phe Gln Asn Arg Arg
130     135     140
Ala Arg Trp Lys Thr Lys Gln Leu Glu Lys Asp Tyr Asp Ser Leu Asn
145     150     155     160
Asp Gly Tyr Glu Ser Leu Lys Thr Glu Tyr Asp Asn Leu Leu Lys Glu
165     170     175
Lys Asp Arg Leu Gln Ser Glu Val Ala Ser Leu Thr Glu Lys Val Leu
180     185     190
Glu Arg Glu Lys Gln Glu Gly Lys Phe Lys Gln Gly Glu Ser Glu Thr
195     200     205
Lys Glu Phe Leu Lys Glu Pro Thr Ile Asn Lys Pro Leu Val Asp Ser
210     215     220
Val Ser Glu Gly Glu Gly Ser Lys Leu Ser Ile Val Glu Ala Ser Asn
225     230     235     240
Asn Asn Asn Asn Asn Asn Lys Leu Glu Asp Ile Ser Ser Ala Arg Ser
245     250     255
Asp Ile Leu Asp Cys Glu Ser Pro Arg Tyr Thr Asp Gly Val Leu Glu
260     265     270
Thr Cys Asp Ser Ser Tyr Val Phe Glu Pro Glu Tyr Gln Ser Asp Leu
275     280     285
Ser Gln Asp Glu Glu Asp His Asn Leu Leu Pro Pro Tyr Ile Phe Thr
290     295     300
Lys Leu Glu Asp Val Asn Tyr Ser Asp Pro Pro His Asn Ser Thr Ser
305     310     315     320
Tyr Gly Phe Gln Glu Glu Asp His His Gln Ala Leu Trp Pro Trp Ser
325     330     335
Tyr

```

10

<210> 31

<211> 843

<212> ADN

15 <213> *Aquilegia formosa* x *Aquilegia pubescens*

<400> 31

ES 2 531 376 T3

```

atggattcaa caacaagccg tcttttcttt gatgggttct gccatgggaa catggtgctt    60
ttagggaagt gagatcccgt tcttcgaggt tcaagatcat tcattaatat ggaagattct    120
ttgaaaagac gtccctttta tagttcaaca gatgaactaa ttgaagagga gttttatgat    180
gaacagctac ctgaaaagaa acgtcgtctt acttctgagc aggttcatct atgggagaag    240
agcttttgaga cagagaacaa gctggaacca gatcgtaaga cccagcttgc taagaagctt    300
gggttgcaac cgagacaagt tgcagtttgg ttccagaata gacgagctcg ttggaagact    360
aagcaactag agagagatta tgatcttctt aaagcttctt atgattccct tcgttctgat    420
tacgatgaca ttgttaaaga gaatgagaag ctcaaactctg aggtgggttc ctaactggg    480
aagttgcagg tcaaggaggg agctgggatg gatttaaact agatatctga cccaccactc    540
tccactgaag aaaatgttga tgaactacg atgcaattta atgttaaggt agaggatcgc    600
ttgagctctg gcagtggggt aagtgtgtgt gttgatgagg aatgtcgaca gcttgttgac    660
agtgttgatt cctatttccc tggcgatgac tatggtcaat gcataggccc agtagatgga    720
gtccagtcag aagaagatga cattagtgcg gacagccgga gctatttctc agatgtcttt    780
ccagctgcac cagagcagaa ccaccaggag agtgagacat tgggttggtg ggactgggct    840
taa

```

<210> 32

5 <211> 280

<212> PRT

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

<400> 32

10

```

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

```

<210> 33

ES 2 531 376 T3

<211> 2193

<212> ADN

<213> Oryza sativa

5 <400> 33

```

aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaaact 120
catccaccta ctttagtggc aatcgggcta aataaaaaag agtcgtaca ctagtttcgt 180
tttctcttagt aatttaaggg gaaaatgaaa tcattattgc ttagaatata cgtttcacatc 240
totgtcatga agttaaatga ttogaggtag ccataattgt catcaaaactc ttcttgaata 300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga 360
atgaagatat tctgaacgta ttggcaaaaga tttaaacata taattatata attttatagt 420
ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480
ttagtaatta aagacaattg acctattttt attattttatc ttttttcgat tagatgcaag 540
gtactttacgc acacacttgg tgctcatgtg catgtgtgag tgcacctcct caatacacgt 600
tcaactagca acacatctct aatalcactc gctatlttaa tacatttagg tagcaatabc 660
tgaattcaag cactccacca tcaccagacc acctttaata atatctaaaa taaaaaaaat 720
aattttacag aatagcatga aaagtatgaa acgaactatt taggtttttc acatacaaaa 780
aaaaaaagaa ttttgctcgt gcgcgagcgc caatctccca tattggggac acaggcaaca 840
acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900
tccgcaacaa ctttttaaca gcaggctttg cggccaggag agaggaggag agggcaagaa 960
aaccaagcat cctctctctc ccctctataa attctctccc ctttttccc ttctctatata 1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
cgaccgcctt ctctgatcca tatcttccgg tcgagttctt ggtcgatctc ttccctctc 1140
cacctcctcc tcacagggta tgtgccttcc ggttggtctt ggattttatt ttctaggttg 1200
tgtagtacgg gcgttgatgt taggaaaggg gatctgtatc tgtgatgatt cctgttcttg 1260
gattttgggat agaggggttc ttgatgttgc atgttatcgg ttccggtttga ttagttagat 1320
ggttttcaat cgtctggaga gctctatgga aatgaaatgg tttagggtac ggaatcttgc 1380
gattttgtga gtacctttg tttgaggtaa aatcagagca ccggtgattt tgccttggtg 1440
aataaaaagta cggttgttgg gtoctcgatt ctggtagtga tgcctctcga ttgacgaag 1500
ctatcctttg tttattccct attgaacaaa aataatccaa ctttgaagac ggtcccgttg 1560
atgagattga atgattgatt cttaagcctg tccaaaattt cgcagctggc ttggttagat 1620
acagtgtcc ccatacagaa attcatgaa acagttataa tcctcaggaa caggggattc 1680
cctgttcttc cgatttgctt tagtcccaga atttttttcc ccaaatatct taaaaagtca 1740

ctttctcgtt cagttcaatg aattgattgc tacaaataat gcttttatag cgtttatccta 1800
gctgtagttc agttaatagg taataccctt atagtttagt caggagaaga acttatccga 1860
tttctgatct ccatttttaa ttatatgaaa tgaactgtag cataagcagt attcatttgg 1920
attatttttt ttattagctc tcaccttcc attattctga gctgaaagtc tggcatgaac 1980
tgtctcfaat tttgttttca aattcacatc gatttatctat gcattatcct ctgttatcta 2040
cctgtagaag tttcttttgg gttattcctt gactgottga ttacagaaag aaatttatga 2100
agctgtaatc gggatagcca tactgottgt tottatgatt catttctctt gtcgagttct 2160
tgggttagct tgccacttcc accagcaaaag ttc

```

10 <210> 34

<211> 49

<212> ADN

<213> Artificial sequence

15 <220>

<223> cebador: prm6000

<400> 34

20

ggggacaagt ttgtacaaaa aagcaggctt aaacaatgga tccggccg

49

ES 2 531 376 T3

<210> 35

<211> 48

<212> ADN

5 <213> Secuencia artificial

<220>

<223> cebador: prm6001

10 <400> 35

ggggaccact ttgtacaaga aagctgggtg atcagctcca gaaccagg 48

<210> 36

15 <211> 834

<212> ADN

<213> Oryza sativa

<400> 36

20

atgaagcgcac	ccggcggtgc	cgccggcggc	ggaggcagcc	catcgctcgt	cacgatggct	60
aattctagt	atgatggata	tggaggggtt	gggatggagg	cggaggggga	cgtggaggag	120
gagatgatgg	cgtgcggcgg	cgccggggag	aaqaagcggc	ggctgagcgt	ggagcagggt	180
cgcgcgctgg	agcggagcct	cgaggtggag	aacaagcttg	agcctgagcg	gaaggcgcg	240
ctggcgcccg	acctcgccct	gcagccggcg	caggtcgccg	lctgggtcca	gaaccgccc	300
gcgcggtgga	agaccaagca	gctcgagcgc	gactacgcgc	cgctccggca	ttcctacgac	360
tcctcgccgc	tcgatcacga	cgcgctccgc	cgcgacaagg	acgccctcct	cgcgcgagac	420
aaggagctga	aggcgaagct	cggggacgag	gaggcgccgg	cgagcttcac	gtcgggtgaag	480
gaggagcccg	cgccctccga	cgggccaccg	gcggcgggat	ttgggtcgtc	cgacagcgac	540
tcaagcccg	tgcgtgaacga	cgtggacgag	gcggcgcccg	cgcccgccgc	gacggacgag	600
ctggctcccg	aggcgtgcac	gtttctcggt	gcgcgcggcg	cgccggcgcc	ggcgcgggcg	660
gcagcgcccg	cgccgagcca	cgaggaggtg	ttcttccacg	gcaatttccg	caagggtggag	720
gaggaccaga	cggggttcct	cgacgacgac	gagccgtgcg	gcgggttcct	cgcgcgagat	780
cagcccccgc	cgtctcgcgc	gtggtggggc	gaaccacagg	agcactggaa	ctga	834

<210> 37

<211> 277

25 <212> PRT

<213> Oryza sativa

<400> 37

30

Met Lys Arg Pro Gly Gly Ala Gly Gly Gly Gly Gly Ser Pro Ser Leu

ES 2 531 376 T3

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

<210> 38

<211> 750

5 <212> ADN

<213> Oryza sativa

<400> 38

atggatcggg aggaaggacaq cgaatggatg atgatggacg ttggaggqaa gggcgqgaag	60
ggcgcgcccg ggcggcgccg ggcggaacag aagaagcggg tcagcgagga gcagatcaag	120
tcgtcgaggt ccatgttcgc gacgcagacc aagctggagc cgaggcagaa gctgcagctc	180
gccaggagac tcggcctgca gctcgccag gtcgccatct ggttccagaa caagcgcgcg	240
cggtggaagt ccaagcagct cgagcgcgag tactccgccc tcgcgcagca ctacgacgcc	300
ctcctctgca gctacgagtc cctcaagaag gagaagctcg cctcatcaa gcagctggag	360
aagctggcgg agatgctgca ggagccacgg gggaagtacg gcgataatgc cggggacgac	420
gcgcggtcgg ggcggctcgc cggcatgaag aaggaggagt tcgtcgcgcg gggcgcgcc	480
gccacgctct actcgtcggc cgagggtggc gggaacgca cgacgacgac ggccaagtgt	540
atgccccact tcggcagcga cgacgtcgac gcggggctct tcctccggcc gtcgtcgag	600
catcatccgc cgcgcgcgca cgcgggtgcc ggcttcacgt cctccgagcc ggccgcccac	660
caccagtcct tcaacttcca ctcgagctgg ccgtcgccca cggagcagac ctgcagcagc	720
acgccatggt gggaattcga gagcgagtga	750

10

<210> 39

<211> 249

<212> PRT

15 <213> Oryza sativa

<400> 39

```

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

```

5

<210> 40

<211> 945

<212> ADN

<213> Populus tremuloides

10

<400> 40

```

atggcgcgctg glaccgggag ttctaattcc aatllgctctg llllgcttca aaqccaaaaga      60
ggccctttgtg ctgcttcaca acctcttgaa tcttttttcc tttctggctc ttctcctttct      120
tttcttggtt caagatccat gatgagtttt gaagatgttc atcaagcaaa cggatcaacc      180
aggccttttt tccgctcgtt tgatcacgaa gacaatggag acgatgatct ggatgaatat      240
tttcatcaac ctgaaaagaa gaggagactt actgttgatc aagttcagtt tcttgaaaag      300
agttttgagc ttgagaacaa gcttgaacct gaaaggaaaa tccagcttgc aaaggatctt      360
ggccttcagc cgcgtcaggt tgctatatgg ttccaaaacc gccgagcaag atggaagact      420
aaacagctgg aaaaggatta tgacgttttg caatctagct acaatagcct taaggctgac      480
tatgacaacc tctcaagga gaaggagaaa ctaaaagctg aggttaatct tctcacgcac      540
aagttgctcc tcaaagagaa agagaaggga atctcagaat tgtctgataa agatgcatta      600
tcgcaagagc cacctaaaag gctatatagc gattcagctt ccgaggggtg agtgctgaaa      660
atctcaacag tggcctgtaa gcaggaagat atcagctcag ccaaaagcga catatttgat      720
tcagacagcc cacattacgc tgatgggggtg cattcctcac tcttagaggc aggagattct      780
tcatatgttt tcgaaccoga tcaatcagat ttgtcacaag atgaagaaga taacttttagc      840
aagagcttat tgccctcata cgtctttccg aagcttgaag atgacgatta cctcgaccgc      900
cctgcaagtt ttgaagatca tgccttttgg tcttggtcat actaa      945

```

15 <210> 41

ES 2 531 376 T3

<211> 314

<212> PRT

<213> Populus tremuloides

5 <400> 41

```

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

```

<210> 42

10 <211> 954

<212> ADN

<213> Populus tremuloides

<400> 42

15

```

atggcggcctt gtgggtggcgg tggtgggtggt totaatccca atttgtctgt ttagttcaa      60
agccaaagag gcccttgatgc tgctttctcaa cctcttgaag cttttttcct tctggctct      120
tctccttctt ttcttggttc aagatccatg atgagttttg cagatgttca ccaagcaaatt      180
ggatcaacta gaccgtttct ccgccatat gatcacgaag acaacggcga cgatgatttg      240
gatgaatatt ttcatcaacc tgaaaagaag aggagactta ctgttgatca agttcagttt      300
cttgaaagaa gttttgaggt tgagaacaag cttgaacccg aaaggaaaat ccagctggcg      360

```

ES 2 531 376 T3

```

aaggatcttg gcttgcagcc tcggcaggtt gccatatggt ttcaaaaccg ccgggcaaga 420
tggaagacga aacagcttga aaaagattat gaggttctgc aatctagcta caatggcctt 480
aaggctgact acgacaacct cttcaaggag aaggagaaac taaaagctga ggtaaacttt 540
ctcaccaacg agttgctcct taaagagaaa gaaaaaggaa gctcagaatt gctcgataaa 600
gatgcattat ctcaagagcc acccaaaaag gcaatagccg attcagcttc agagggtgaa 660
gtgtcgaaaa ctccaacogt ggctgcccag caggaagata ttagctcagc caaaagtgat 720
atgtttgatt cagacagccc acattttgcg gatggggtag attcctcact cttagaggca 780
ggtgattcct cacatgtcct cgagcccgac caatcggatt tatcacaaga tgaagaagat 840
aacttgagca agagtctttt gcctccgtac gtctttccaa agcttgaaga tgggtgattac 900
tctgaccgcc cagcaagttt tgaagatcat gccttttggg cctggtcata ctaa 954

```

<210> 43

<211> 317

5 <212> PRT

<213> Populus tremuloides

<400> 43

```

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

```

10

<210> 44

<211> 909

<212> ADN

<213> Populus tremuloides

<400> 44

```

atggcgggtg ataaagactg tggcagttct aaaatgacca tttttcttcg aaacggcagg      60
ctccctcctt gtgaatctct ctgtattctc acctctttta gcactcttca tgggtgcaaaa    120
totatggtta attttaggaa tgatggagga gacactgtag acatgtcttt tttccaacca    180
catgtcaaaag aagaaaagtag cgaatgaggat tatgatgcgc accttaagcc atctgaaaag    240
aaaaggcggc ttacagctgc tcaagtcag tttcttgaga agagctttga ggcgagaat     300
aagcttgaac cagagaggaa gatgcagctt gctaaagaac tcggcttgca gcctcgccag    360
glttgaatal ggtttcaaaa ccglagagct cggttcaaga acaagcagct ggaacgggac    420
lacgaactccl tgagaalccag ctlllgacaaa ctcaaggctg attatgacaa actccctctc    480
gagaagcaga atttgaaaaa cgaagcttctt tcaactgaaag aaaaatttct tagcagagag    540
gaaagtatgg aaagttcaga accatttgat gtcattccatt caccggatgc agaacttgag    600
cctattctctg atacagtgtc tgaaaatggt tcggccattg tgccaatggt gacacccaaa    660
caagaagaaa gttcagctaa aaatgatggt ttcaactcag acagcccacg ttcatttttg    720
gagcccccgtg attggtatcg tgttttcgag tcagaccaac cagatttttc ccaagttgaa    780
gaagataatc tcaccaggag cttttctacc cctccgtact ttccaaaact ctaccgagag    840
ccactgcaa gttcacgtaa ttttgaattc tcagcggaag atcagccctt ttggtcctgg    900
atttactga

```

5

<210> 45

<211> 302

<212> PRT

10 <213> Populus tremuloides

<400> 45

```

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

```

ES 2 531 376 T3

245 250 255
 Ser Gln Val Glu Glu Asp Asn Leu Thr Arg Ser Phe Leu Pro Pro Pro
 260 265 270
 Tyr Phe Pro Lys Leu Tyr Arg Glu Pro Pro Ala Ser Ser Arg Asn Phe
 275 280 285
 Glu Phe Ser Ala Glu Asp Gln Pro Phe Trp Ser Trp Ile Tyr
 290 295 300

<210> 46

<211> 975

5 <212> ADN

<213> *Medicago truncatula*

<400> 46

atggcagggtg	gcaagctctt	tggtggttct	aatatgtcac	ttttgcttca	aaatgaaaga	60
ctcccttgta	cttctgaagt	ccttgaatct	ctttgggttc	acacccctgc	ttcttttcaa	120
ggttcaaat	cagtggttaa	ttttgagaat	ggtaggtgga	gcaacagagt	ggtaacagat	180
agacccttct	ttcaacaact	tgagaaagaa	gagaattgtg	gtgatgaaga	ttatgaagca	240
tgtaccatc	aacaaggaaa	gaaaaggagg	ctttcaagtg	aacaagttca	atttcttgaa	300
aagagttttg	aggtagaaaa	caagcttgaa	cctgatagga	aagttcaact	tgcaaaagag	360
cttggtttgc	aaccaagaca	agttgtata	tggtttcaaa	acagaagggc	aaggttcaaa	420
actaaacagc	ttgaaaaaga	ttatggcaca	ttgaaagcta	gctttgatag	ttctcaaagat	480
gattatgata	atcttcttca	agagaatgac	aagttaaaag	aagaggtgaa	ttctctcaag	540
aacaaattga	tcccaagaga	taaagaaaaa	gtgaattcag	aagacaaatc	atcaccagaa	600
gcaatcaatt	cacctcataa	caacatagat	ccaatggata	taatttcaat	tacaaattca	660
gaaaatgggt	ccaaaatgct	actcccta	atgggtactaa	aatgtaagca	agaagatgcc	720
aattcagcta	aaagtgatgt	gcttgattct	gatagcccac	attgcaatga	tgqgaacaat	780
ctttcttcll	lcalagagcc	lacaqallca	qallllclac	aagalgaaga	qqataalgal	840
aacttgagtc	ataatctctt	gactcttctt	tgcttaccaa	aagttgaaga	tglttgctat	900
gatgaccac	atgaaaattc	ttgtaatttt	gggttcctctg	ttgaagatca	aaccttttgt	960
ttctggcctt	attga					975

10

<210> 47

<211> 324

<212> PRT

15 <213> *Medicago truncatula*

<400> 47

ES 2 531 376 T3

```

Met Ala Gly Gly Lys Leu Phe Gly Gly Ser Asn Met Ser Leu Leu Leu
1      5      10      15
Gln Asn Glu Arg Leu Pro Cys Thr Ser Glu Val Leu Gln Ser Leu Trp
20      25      30
Val His Thr Pro Ala Ser Phe Gln Gly Ser Asn Ser Val Val Asn Phe
35      40      45
Glu Asn Gly Gly Gly Ser Asn Arg Val Val Thr Asp Arg Pro Phe Phe
50      55      60
Gln Gln Leu Glu Lys Glu Glu Asn Cys Gly Asp Glu Asp Tyr Glu Ala
65      70      75      80
Cys Tyr His Gln Gln Gly Lys Lys Arg Arg Leu Ser Ser Gln Gln Val
85      90      95
Gln Phe Leu Glu Lys Ser Phe Glu Val Glu Asn Lys Leu Glu Pro Asp
100     105     110
Arg Lys Val Gln Leu Ala Lys Glu Leu Gly Leu Gln Pro Arg Gln Val
115     120     125
Ala Ile Trp Phe Gln Asn Arg Arg Ala Arg Phe Lys Thr Lys Gln Leu
130     135     140
Glu Lys Asp Tyr Gly Thr Leu Lys Ala Ser Phe Asp Ser Leu Lys Asp
145     150     155     160
Asp Tyr Asp Asn Leu Leu Gln Glu Asn Asp Lys Leu Lys Glu Glu Val
165     170     175
Asn Ser Leu Lys Asn Lys Leu Ile Pro Arg Asp Lys Glu Lys Val Asn
180     185     190
Ser Glu Asp Lys Ser Ser Pro Glu Ala Ile Asn Ser Pro His Asn Asn
195     200     205
Ile Asp Pro Met Asp Ile Ile Ser Ile Thr Asn Ser Glu Asn Gly Ser
210     215     220
Lys Met Ser Leu Pro Asn Met Val Leu Lys Cys Lys Gln Glu Asp Ala
225     230     235     240
Asn Ser Ala Lys Ser Asp Val Leu Asp Ser Asp Ser Pro His Cys Asn
245     250     255
Asp Gly Asn Asn Leu Ser Ser Phe Ile Glu Pro Thr Asp Ser Asp Phe
260     265     270
Ser Gln Asp Glu Glu Asp Asn Asp Asn Leu Ser His Asn Leu Leu Thr
275     280     285
Leu Pro Cys Leu Pro Lys Val Glu Asp Val Cys Tyr Asp Asp Pro His
290     295     300
Glu Asn Ser Cys Asn Phe Gly Phe Pro Val Glu Asp Gln Thr Phe Cys
305     310     315     320
Phe Trp Pro Tyr

```

<210> 48

5 <211> 984

<212> ADN

<213> Phaseolus vulgaris

<400> 48

10

ES 2 531 376 T3

```

atggcgggtg gcaagcttca tcttggttca aacatgtcac ttctcttcca aaacgacag 60
ctcccttgct cctctgaagt ccttgagtct ctttgggctc acacctctaa cgtgcttcc 120
ttccaagctt caaatctat ggttgatttt gagaatgtta gtgggggcag ggtgacggat 180
aggccctttt ttcaagcgtt ggagaaggaa gataactgtg atgatgatta tgagggttgc 240
ttccatcaac ogggttaagaa aaggaggctc acaagcgaac aagttcagtt ccttgaaagg 300
aacctttgag tcgagaacaa gcttgaacct gaaagggaag tccaacttgc aaaggagctt 360
ggcttgacgc caaggcaagt ggttatatgg ttccaaaacc gaagggcag gttcaagacc 420
aagcagctag aaaaagatta tggcacattg aaagctagct atgacagact caaagggtgac 480
tatgaaagtc ttcttcaaga gaatgacaag ttaaaagcag aggtgaatto tctggagagc 540
aaattgattc ttagagataa agagaaggag aattcggacg acaagtcato tcttgatgct 600
gtcaattcac ccacaaaaga gcttatggat ttaatttcaa attcaacato tgaaaatggg 660
accaaagtgt cactccctat tatggtaaca tgcaagcaag aagatgcaa ttcagccaaa 720
agtgatgtgc ttgattcgga cagccacat tgcactgatg ggaaccatcc ctttcatc 780
gtggagcctg ctgattccct ccatgctttt gaaccagacc actccgactt cttccaagat 840
gaagaggata atctttagta aagccttttg accctccctt gcttaccaaa ggttgaagaa 900
gcttgctatg atgacctcc tgaaaacct tgtaattttg gcttccatgt cgaggatcaa 960
acctctctgt tctggcccta ttga 984

```

<210> 49

<211> 327

5 <212> PRT

<213> Phaseolus vulgaris

<400> 49

```

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

```

10

ES 2 531 376 T3

50	55	60
Gln Ala Leu Glu Lys	Glu Asp Asn Cys Asp	Asp Asp Tyr Glu Gly Cys
65	70	75
Phe His Gln Pro Gly	Lys Lys Arg Arg Leu	Thr Ser Glu Gln Val Gln
85	90	95
Phe Leu Glu Arg Asn	Phe Glu Val Glu Asn	Lys Leu Glu Pro Glu Arg
100	105	110
Lys Val Gln Leu Ala	Lys Glu Leu Gly Leu	Gln Pro Arg Gln Val Ala
115	120	125
Ile Trp Phe Gln Asn	Arg Arg Ala Arg Phe	Lys Thr Lys Gln Leu Glu
130	135	140
Lys Asp Tyr Gly Thr	Leu Lys Ala Ser Tyr	Asp Arg Leu Lys Gly Asp
145	150	155
Tyr Glu Ser Leu Leu	Gln Glu Asn Asp Lys	Leu Lys Ala Glu Val Asn
165	170	175
Ser Leu Glu Ser Lys	Leu Ile Leu Arg Asp	Lys Glu Lys Glu Asn Ser
180	185	190
Asp Asp Lys Ser Ser	Pro Asp Ala Val Asn	Ser Pro His Lys Glu Pro
195	200	205
Met Asp Leu Ile Ser	Asn Ser Thr Ser Glu	Asn Gly Thr Lys Val Ser
210	215	220
Leu Pro Ile Met Val	Thr Cys Lys Gln Glu	Asp Ala Asn Ser Ala Lys
225	230	235
Ser Asp Val Leu Asp	Ser Asp Ser Pro His	Cys Thr Asp Gly Asn His
245	250	255
Pro Ser Ser Phe Val	Glu Pro Ala Asp Ser	Ser His Ala Phe Glu Pro
260	265	270
Asp His Ser Asp Phe	Ser Gln Asp Glu Glu	Asp Asn Leu Ser Glu Ser
275	280	285
Leu Leu Thr Leu Pro	Cys Leu Pro Lys Val	Glu Glu Ala Cys Tyr Asp
290	295	300
Asp Pro Pro Glu Asn	Pro Cys Asn Phe Gly	Phe His Val Glu Asp Gln
305	310	315
Thr Phe Cys Phe Trp	Pro Tyr	
325		

<210> 50

<211> 957

5 <212> ADN

<213> Lotus corniculatus

<400> 50

alggcggcag	ggaggggtctt	tggcggcggt	tctgtctctc	ctgcaaatgt	ctccgatacc	60
agtctttttg	ttcagaatca	acctcctgat	tctttctctc	tcctctctac	ctctgcttct	120
ttttctcggt	caagatccat	ggtgagcttc	gcagataata	aattagggca	aacgcggctg	180
ttttctcccg	cgtttgacct	cgatgagaac	ggcgatgagg	tcattggacg	gtactttcac	240
caatcggaga	agaagcgccg	tctctctgtt	gaccaagtgc	agtttctgga	gaagagcttc	300
gagggtggata	acaagctcga	acctgacagg	aaaaccaaga	ttgccaagga	ccttggtttg	360
cagccacgcc	aagtgcgaat	ctggttccag	aacgcgcgtg	cacggtggaa	gacgaaacag	420
cttgagaagg	attatgattc	tctgcatagt	agctttgaga	gtctcaaato	caactatgat	480
aatctttctc	aggagaaaaga	catgttaaaa	gctgaggttg	caagtctcac	tgagaagggtg	540
cttgcaagag	agaatttgaa	acaagttgaa	agtgaacaaa	agggatttgt	tgaaccaccc	600
caaaggcctt	tacttgattc	agtttcagag	ggtgaagaat	ctaaagtctc	tggtggggct	660
tgtaaacatg	aggatatcag	ttcagccagg	agtgagagtt	tggattctga	tggcccacgt	720
tacagggatg	gatatggagt	taactcagca	gtgctagaga	catgtgatto	ttcttatgtg	780
ggtgaacctg	atcaatcgga	tatgtcacag	gatgaggaag	acaacctgac	caagaccttg	840
ttgcctccat	acatgttttc	caaacttgga	gatatggatt	actccgaccc	gcctgaaagt	900
tcattgtaatt	tggattttcc	ggaggaagat	catgcccttt	ggtcatggtc	ttactga	957

<210> 51

<211> 318

<212> PRT

<213> Lotus corniculatus

<400> 51

5

```

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

```

<210> 52

<211> 2194

10 <212> ADN

<213> Oryza sativa

<400> 52

```

aatccgaaaa gtttctgcac cgttttcacc cccaaactaa caatataggg aacgtgtgct      60
aatataaaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaaact      120
catccacctt ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtttcgt      180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cgttcacatc      240
tctgtcatga agttaatta ttcgaggtag ccataattgt catcaaaactc ttcttgaata      300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga      360

```

15

ES 2 531 376 T3

atgaagatat	tctgaacgta	ttggcaaaga	tttaaacata	taattatata	attttatagt	420
ttgtgcattc	gtcatatcgc	acatcattaa	ggacatgtct	tactccatcc	caatttttat	480
ttagtaatta	aagacaattg	acttattttt	attattttatc	ttttttcgat	tagatgcaag	540
gtacttaacg	acacaacttg	tgtcatgtg	catgtgtgag	tgcacctcct	caatacacgt	600
tcaactagca	acacatctct	aatatcactc	gcctatttaa	tacatttagg	tagcaatatc	660
tgaattcaag	cactccacca	tcaccagacc	acttttaata	atatctaaaa	tacaaaaaat	720
aattttacag	aatagcatga	aaagtatgaa	acgaactatt	taggtttttc	acatacaaaa	780
aaaaaaagaa	ttttgtctgt	gcgcgagcgc	caatctccca	tattgggcac	acaggcaaca	840
acagagtggc	tgcacacaga	acaacccaca	aaaaacgatg	atctaacgga	ggacagcaag	900
tcgcgaacaa	ccttttaaca	gcaggctttg	cgccacggag	agaggaggag	aggcaagaa	960
aaccaagcat	cctcctctcc	ccatctataa	attcctcccc	ccttttcccc	ctctctatata	1020
ggagggcatcc	aagccaaqaa	gagggagagc	accaaggaca	cgcgactagc	agaagccgag	1080
cqaccgcctt	ctcquatccat	atcttccggt	cgaqttcttg	gtcgatctct	ccccctctcc	1140
acctcctcct	cacagggtat	gtgcctcctt	tcggttggtc	ttggatttat	tgttctaggt	1200
tgtgtagtac	gggcgttgat	gttaggaaag	gggatctgta	tctgtgatga	ttcctgttct	1260
tggatttggg	atagagggtt	tcttgatgtt	gcattgttate	ggttcgggtt	gatttagtagt	1320
atggttttca	atcgtctgga	gagctctatg	gaaatgaaat	ggttttaggga	tcggaatctt	1380
gcgattttgt	gagtaacctt	tgtttgaggt	aaaatcacag	caccgggtgat	tttgccttgg	1440
gtaataaagt	acggttgttt	ggctcctgat	tctggtagtg	atgcttctcg	atttgacgaa	1500
gcataccttt	gtttattccc	tattgaacaa	aaataatcca	actttgaaga	cggctccggt	1560
gatgagattg	aatgattgat	tottaagcct	gtccaaaatt	tcgcagctgg	cctgttttaga	1620
tacagttagc	ccatcacga	aattcatgga	aacagttata	atcctcagga	acaggggatt	1680
cctgtttctt	cagatttgct	ttagtcccag	aatttttttt	cccaaatatc	ttaaaaagtc	1740
actttctcgt	tcagttcaat	gaattgattg	ctacaaataa	tgccttttata	gcgttatcct	1800
agctgtagtt	cagttaatag	gtaatacccc	tatagtttag	tcaggagaag	aaacttatccg	1860
atttctgata	tccattttta	attatatgaa	atgaactgta	gcataagcag	tattcatttg	1920
gattattttt	tttattagct	ctcacccttt	cattattctg	agctgaaagt	ctggcatgaa	1980
ctgtcctcaa	llllqlllcc	aaallccact	cqallalala	lqcallalcc	ctllqlalcl	2040
acctgtagaa	gtttctctct	ggttattcct	tgactgcttg	attacagaaa	gaaatttatg	2100
aagctgtaat	cgggatagtt	atactgcttg	ttcttatgat	tcatttccct	tgtgcagttc	2160
ttggtgtagc	ttgccacttt	caccagcaaa	gttc			2194

REIVINDICACIONES

1. Procedimiento para incrementar el rendimiento en plantas cultivadas en condiciones de disponibilidad reducida de nutrientes, en relación con las correspondientes plantas silvestres, que comprende modular la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido hox5 de cremallera de leucina de homeodominio (HDZip) de clase I u homólogo del mismo, y opcionalmente seleccionar las plantas que tienen mayor rendimiento, en el que dicha secuencia de ácido nucleico que codifica dicho polipéptido hox5 de cremallera de leucina de homeodominio (HDZip) de clase I u homólogo del mismo se selecciona del grupo que consiste en:
 - a) una secuencia de ácido nucleico representada por la SEC ID N° 1;
 - b) una secuencia de ácido nucleico que codifica un polipéptido representada por la SEC ID N° 2,
 - c) una secuencia de ácido nucleico que es capaz de hibridar en condiciones rigurosas con una secuencia de ácido nucleico representada por la SEC ID N° 1 y
 - d) una secuencia de ácido nucleico que codifica un polipéptido que tiene una identidad de secuencia de al menos un 70 % con un polipéptido, representada por la SEC ID N° 2.
2. El procedimiento de la reivindicación 1, en el que dicho polipéptido HDZip hox5 de clase I o un homólogo del mismo, comprende de N-terminal a C-terminal: (i) un homeodominio de clase I; y (ii) una cremallera de leucina con más de 5 héptadas.
3. El procedimiento de la reivindicación 2, en el que dicho polipéptido HDZip hox5 de clase I o un homólogo del mismo, comprende además uno o ambos de los siguientes: (i) al menos un residuo de Trp en los últimos 10 aminoácidos del C-terminal de dicho polipéptido; y (ii) un motivo de aminoácidos RPFF, en el que R es Arg, Pro P y F Phe, y dentro de este motivo, permitiendo que uno o más cambio(s) conservador (s) en cualquier posición, y / o uno o dos cambio(S) no conservador (s) en cualquier posición.
4. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que dicha secuencia de ácido nucleico HDZip hox5 de clase I es de origen vegetal, preferiblemente de una planta monocotiledónea, más preferiblemente de la familia Poaceae, más preferiblemente del género Oryza, más preferiblemente de Oryza sativa .
5. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que dicha secuencia de ácido nucleico codifica un ortólogo o parálogo del polipéptido HDZip hox5 de clase I de la SEC ID N° 2.
6. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que dicha modulación de la expresión es una expresión incrementada en una planta de una secuencia de ácido nucleico que codifica el polipéptido hox5 de cremallera de leucina de homeodominio (HDZip) de clase I o un homólogo del mismo.
7. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que dicha expresión es modulada mediante la introducción de una modificación genética, preferiblemente introduciendo y expresando en una planta una secuencia de ácido nucleico que codifica el polipéptido HDZip hox5 de clase I o un homólogo del mismo.
8. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en el que dicha secuencia de ácido nucleico HDZip hox5 de clase I está unido operativamente a un promotor constitutivo.
9. Procedimiento de acuerdo con la reivindicación 8, en el que dicho promotor es un promotor GOS2, más preferentemente el promotor constitutivo es un promotor GOS2 de arroz, más preferentemente el promotor constitutivo está representado por una secuencia de ácido nucleico sustancialmente similar a la SEC ID N° 33 o SEC ID N° 52, lo más preferentemente, el promotor constitutivo es como está representado mediante la SEC ID N° 33 o SEC ID N° 52.
10. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 9, en el que dicho rendimiento aumentado es uno o más de: aumento del rendimiento total de semillas por planta, aumento en el número de semillas llenas, aumento de la tasa de llenado de semillas, mayor número de flores por panícula, o mayor índice de cosecha.
11. Procedimiento de acuerdo con una de las reivindicaciones 1 a 10, en el que dicha disponibilidad reducida de nutrientes es disponibilidad reducida de nitrógeno.
12. Procedimiento para aumentar el índice de verdor en las plantas cultivadas en condiciones de disponibilidad reducida de nutrientes, en relación a las correspondientes plantas silvestres, en el que el procedimiento comprende introducir y expresar en una planta una secuencia de ácido nucleico que codifica un polipéptido HDZip hox5 de clase I o un homólogo del mismo como se define en la reivindicación 1.
13. El uso de una construcción que comprende:
 - (i) una secuencia de ácido nucleico de HDZip hox5 de clase I que codifica un polipéptido hox5 de cremallera de leucina de homeodominio (HDZip) de clase I como se define en la reivindicación 1;
 - (ii) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (i); y,

opcionalmente

(iii) Una secuencia de terminación de transcripción.

en un procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 12.

5 14. Procedimiento para la producción de una planta transgénica que tiene un mayor rendimiento en condiciones de disponibilidad reducida de nutrientes reducida, en relación a las plantas silvestres correspondientes, en el que el procedimiento comprende:

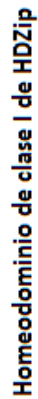
(i) introducir y expresar en una planta, parte de planta o célula de planta una secuencia de ácido nucleico HDZip hox5 de clase I que codifica un polipéptido hox5 de cremallera de leucina de homeodominio (HDZip) de clase I o un homólogo del mismo como se define en la reivindicación 1,

10 (ii) cultivar la célula vegetal en condiciones que estimulan el crecimiento y desarrollo de la planta.

15. Procedimiento de acuerdo con la reivindicación 14, en el que dicho rendimiento aumentado es uno o más de: aumento del rendimiento total de semillas por planta, aumento en el número de semillas llenas, aumento de la tasa de llenado de semillas, mayor número de flores por panícula, o mayor índice de cosecha.

15 16. Uso de una secuencia de ácido nucleico/gen de HDZip hox5 de clase I que codifica un polipéptido hox5 de cremallera de leucina de homeodominio (HDZip) de clase I u homólogo del mismo como se define en la reivindicación 1 en el aumento del rendimiento en las plantas cultivadas en condiciones de disponibilidad reducida de nutrientes, en relación a las correspondientes plantas silvestres.

20 17. Uso de acuerdo con la reivindicación 16, en el que dicho rendimiento aumentado es uno o más de: aumento del rendimiento total de semillas por planta, aumento en el número de semillas llenas, aumento de la tasa de llenado de semillas, mayor número de flores por panícula, o mayor índice de cosecha.

**FIGURA 1**

Cremallera de leucina de HDZip de clase I			
Zeama_hox5	(SEQ ID NO 06)	TKQLETDYDRLKAAADALAAADHQGLLADNDNLRAQVISLTK	
Glyma_HD157	(SEQ ID NO 20)	TKQLEKDYGVLLKASYDRLKSDYESLVQENDKLAENSLKESK	
Lycyes_VaHOX1	(SEQ ID NO 28)	TKQLEKDYDELNRNRYDTILKSNYNLLKEKEDLRTVFRLTGK	
Dauca_CHB3	(SEQ ID NO 18)	TKTLEKDYDVLLQNSYNLSKADYDNLAEKEKLKAEVLDLTDK	
Medtr_hox16	(SEQ ID NO 30)	TKQLEKDYDSLNDGYESLKTEDYDNLKEKDRLQSEVASLTK	
Triae_hox16	(SEQ ID NO 14)	TKTLEKDYDRLKASFDALRADHDALLQDNHRLRSQVTLTK	
Goshi_hox5	(SEQ ID NO 24)	TKQLEKDYDILLKSSFDLSQSNYDTILKENEKLKSEVASLTK	
Sorbi_hox5	(SEQ ID NO 12)	TKQLETDYDRLKAAADALAAADHQGLLADNDNLRAQVISLTDK	
Sacof_hox5	(SEQ ID NO 10)	TKQLETDYDHLKAAADALAAADHQGLLADNDNLRAQVSLTK	
Agufo_hox5	(SEQ ID NO 32)	TKQLEKDYDILLKASYDSLRSYDDIVKENEKLKSEVSLTGK	
Orysa_hox5	(SEQ ID NO 02)	TKQLEHDFDRLKAAADALAAADHALLSDNDRLRAQVISLTK	
Zeama_hox16	(SEQ ID NO 08)	TKQLEKDYDRLKASFDALRADHDALLQDNHRLRSQVSLTK	
Orysa_hox16	(SEQ ID NO 04)	TKQLEKDYDRLKASFDALRADHDALLQDNHRLHSHQVMSLTK*	
Lycyes_hox5	(SEQ ID NO 26)	TKQLEKDYDQLKSSYDSLSDFSVRKDNKDKLSEVSLMEK	
Arath_ATHB1	(SEQ ID NO 16)	TKQLEKDYDILLKSTYDQLLSNYDSIVMDNDKLSEVSLTK	
Crapl_CPHB-5	(SEQ ID NO 22)	TKQLEKDYDKLKSSYDSLSTYDSIRQENDKLKAEVSLNEK	
Orysa_hox6	(SEQ ID NO 39)	SKQLEKDYDALRDDYDALLCSESLKKEKLALIKQLEKLAEM	
Orysa_hox4	(SEQ ID NO 37)	TKQLEKDYAALRHSYDSLRLDHDALLRRDKDALLAEIKELKAK	
Consenso		TKQLEKDYD LKASYDALRADYDALL DNDKLRAEVVSLTK	
		abcdefgabcdeffgabcdeffgabcdeffgabcdeffg	

heptad1heptad2heptad3heptad4heptad5heptad6

FIGURA 1 (continuación)

	1	50
Zeama_hox5	(1) ---MDPSAVSFDSGGARRGGG-----AQMLLFGGGGGSAN	
Aqufo_hox5	(1) ---MDSTTSRLFFDGSCHGN-----MLLLGSGDPV	
Arath_ATHB1	(1) ---MESNSFFFDPSASHGNS-----MFFLGNLNPV	
Orysa_hox5	(1) ---MDPGRVVFDSGVARRACPGG-----AQMLLFGGGGGSAN	
Crapl_CPHB-5	(1) ---MNSARIFFDPSHGNMLQ-----FLGNAGGDSSV	
Dauca_CHB3	(1) ---MAGR RVFYG---EGANTTSAS-----LLFHSQRPEPFLLSAPSPS	
Glyma_HD157	(1) ---MASGKLYAGSNMSLLLQNER-----LPCSSEVLES LWQTSNPAS	
Goshi_Hox5	(1) ---MESGRLLFFNPSTTHRN-----MLLLGNTPEPI	
Lyces_hox5	(1) ---MGSGHIFFDPSCHGN-----MLFLGSGDPV	
Lyces_VaHOX1	(1) ---MAPGILYGGSSNFDGVFTQKQRDVFSSTAPKHLGSLFAPASSSSN	
Medtr_HOX16_1	(1) ---MAGGKLFGGSNMSLLLQNER-----LPCTSEVLES LWVHT--PAS	
Orysa_hox16	(1) ---MESGRLLIFSTAGS---GAGQ-----MLFLDCGAGGGGVG	
Sacof_hox5	(1) ---MDPSAVSFNSGGARRGGGG-----TQMLLFGGGGGSAN	
Sorbi_hox5	(1) ---MDPSAVSFDSGGARRGGGG-----GAQMLLFGGGGGSAN	
Triae_hox16	(1) ---MEPGRLLIFNTSGS---GNGQ-----MLFMDCGAGGIA-G	
Zeama_hox16	(1) ---MESGRLLIFNAPGS---GAGQ-----MLFLDCGAGGG--P	
Poptr_HOX16_1	(1) ---MAGGTGGSNNSLVLLQSQRG-----PCAASQPLESFFLSGS-SPS	
Poptr_HOX16_2	(1) MAACGGGGGGGNPNLSVLVQSQRG-----PCAASQPLEAFFLSGS-SPS	
Poptr_HOX16_3	(1) ---MAGDKDCGSSKMTIFLRNGR-----LPPCESLCILTS-FST	
Phavu_hox16	(1) ---MAGGKLFHPSNMSLLLQNR-----LPCSSEVLES LWHTSNAAS	
Lotco_HOX16	(1) ---MAGGRVFSGGSAAPANVSDTS-----LLLQNPDDSLFLSTSAS	
Medtr_HOX unknown	(1) MHEMAFFQANFMLQTPHHDDHHQ-----PSSLNSILPQD	
Piclg_hox unknown	(1) ---MACDRSALYTSSVIMNTEDN-----SSAHAIAMIASSCTPAT	
Orysa_hox4	(1) -----MKRPGGAGG	
Medtr_HOX unknown2	(1) -----MKRLNNTSDS	
Orysa_hox6	(1) -----	
Consenso	(1) M G F	G

	51	100
Zeama_hox5	(32) SNGFFRGVPMVAVLGMDDATRVG-----KRPFFTTH--ELLEEEYYDE	
Aqufo_hox5	(28) LRGSR----SFINMEDSL-----KRPFFYSST--DELIEEEFYDE	
Arath_ATHB1	(28) VQGGGAR---SMMNMEETS-----KRPFFSS--EDLYDDDFYDD	
Orysa_hox5	(34) SGGFFRGVPAVAVLGMDSESRSSSAAGAGARPFFFTTH--ELLEEEYYDE	
Crapl_CPHB-5	(30) FRGTRSS---SVLNMEESS-----LKRQIFSGGGDEFYDEEYYDE	
Dauca_CHB3	(38) LIGS-----KSMVSFQDAKRKNP-----YDGFMRSYDEEEIGDEEYDE	
Glyma_HD157	(41) FQGS-----KPVVDFENVSGSR-----MTLRPFQADEKEENC-DEEYEG	
Goshi_Hox5	(27) FRGAR-----TMVSMENP-----KRLFFSS--EDLYDEEYYDE	
Lyces_hox5	(27) FRGPRS---TMMKMEDSS-----KRPFFSS--EDLYDEEYYDE	
Lyces_VaHOX1	(48) FLGS-----SSMVSFRGVNGG-----KRSFFDSH--QDDNEADELGE	
Medtr_HOX16_1	(39) FQGS-----NSVNFENGSGSNRV---VTRPFQOIEKEENCDEDEYEA	
Orysa_hox16	(32) GGAMFHRGARPVLGMEEG-----RGVRRPFFTTHD--ELLEEEYYDE	
Sacof_hox5	(33) SNGFFRGVPMVAVLGMDDATRVG-----KRPFFTTH--ELLEEEYYDE	
Sorbi_hox5	(35) SNGFFRGVPMVAVLGMDDATRVG-----KRPFFTTH--ELLEEEYYDE	
Triae_hox16	(31) AAGMFHRGVPRVPLGMEEG-----RGVRRPFFTSHD--DMLEEEYYDE	
Zeama_hox16	(30) GGGLFHRGGRPMLGLEEG-----RGVRRPFFTSHD--ELLEEEYYDE	
Poptr_HOX16_1	(41) FLGS-----RSMMSFEDVHQANG-----STRPFFRSHDHEDNG-DDDLDE	
Poptr_HOX16_2	(44) FLGS-----RSMMSFADVHQANG-----STRPFFRSHDHEDNG-DDDLDE	
Poptr_HOX16_3	(36) LHGA-----KSMVNFRNDGGDT-----VMSFFQPHVKEESS-DEYDA	
Phavu_hox16	(41) FQGS-----KSMVDFENVSGGR---VTRPFQADEKEDNC-DDDYEG	
Lotco_HOX16	(41) FLGS-----RSMVSFADNKLQ-----TRSFSAHDLDENG-DEVMDE	
Medtr_HOX unknown	(36) YHGG-----PSFLGKRCMS-----SSGIELGEEANIPEEDLSDD	
Piclg_hox unknown	(40) FQGTR-----SISVFETGNERKRP---AGNSYSALELSDDIGDEGSD	
Orysa_hox4	(10) GGG-----PSLVTMANS-----DDGYGGVGMEAGDVEEEMMA	
Medtr_HOX unknown2	(11) FSTP-----LITISPSTEEHSPRNKHVYGMFQSMMLDGFEEEGCVVE	
Orysa_hox6	(1) -----MDGEEDS-----BWMMDVG--GKGG--KGGG	
Consenso	(51) G S E	RPF E EE YDE

dominio conservado

FIGURA 2

homeodominio de clase I			101	150
Zeama_hox5	(73)	CAP--EKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAV		
Aqufo_hox5	(62)	QLP--EKKRRLTSEQVHLLLEKSFETENKLEPDRKTQLAKKLGQPRQVAV		
Arath_ATHB1	(64)	QLP--EKKRRLTSEQVHLLLEKSFETENKLEPERKTQLAKKLGQPRQVAV		
Orysa_hox5	(82)	CAP--EKKRRLTAEQVQMLERSFEEENKLEPERKTELARRLGMAPRQVAV		
Crapl_CPHB-5	(68)	QLLP-EKKRRLTAEQVHLLLEKSFEEENKLEPERKAEAKKLGQPRQVAV		
Dauca_CHB3	(77)	YFQPEKKRRLKADQIQFLEKSFETDNKLEPERKVQLAKELGLQPRQVAV		
Glyma_HD157	(80)	CFHQPGKKRRLTSEQVQFLERNFEVENKLEPERKVQLAKELGLQPRQVAV		
Goshi_Hox5	(61)	QLP--EKKRRLTSEQVYLLLEKSFEEENKLEPERKSQLAKKLGQPRQVAV		
Lyces_hox5	(62)	CSP--EKKRRLTPEQVHLLLEKSFETENKLEPERKTQLAXKLGQPRQVAV		
Lyces_VaHOX1	(84)	YLHQAEEKRRLTDNQVFLEKSFEEENKLEPERKVQLAKELGLQPRQVAV		
Medtr_HOX16_1	(81)	CYHQGGKKRRLSSEQVQFLEKSFEEENKLEPDRKVQLAKELGLQPRQVAV		
Orysa_hox16	(73)	QLP--EKKRRLTPEQVHLLERSFEEENKLEPERKTELARKLGLQPRQVAV		
Sacof_hox5	(74)	CAP--EKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAV		
Sorbi_hox5	(76)	CAP--EKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAV		
Triae_hox16	(72)	QLP--EKKRRLTPEQVHLLERSFEEENKLEPERKTELARKLGLQPRQVAV		
Zeama_hox16	(70)	QLP--EKKRRLTPEQVLLERSFEEENKLEPERKTELARKLGLQPRQVAV		
Poptr_HOX16_1	(80)	YFHQPEKKRRLTVDQVQFLEKSFEEENKLEPERKIQLAKDLGLQPRQVAV		
Poptr_HOX16_2	(83)	YFHQPEKKRRLTVDQVQFLEKSFEEENKLEPERKIQLAKDLGLQPRQVAV		
Poptr_HOX16_3	(74)	HLKPSEKKRRLTAAQVQFLEKSFEEENKLEPERKMQLAKELGLQPRQVAV		
Phavu_hox16	(80)	CFHQPGKKRRLTSEQVQFLERNFEVENKLEPERKVQLAKELGLQPRQVAV		
Lotco_HOX16	(78)	YFHQSEKKRRLSVDQVQFLEKSFEEENKLEPERKTKIAKDLGLQPRQVAV		
Medtr_HOX unknown	(71)	CSQAGEKKRRLNMEQVKTLEKSFELGNKLEPERKMQLARALNGLQPRQVAV		
Piclg_hox unknown	(81)	CIHLGEKKRRLTLEQVRALEKNFEMANKLEPERKMQLAKALGLQPRQVAV		
Orysa_hox4	(45)	CGGGGEKKRRLSVEQVRLERSFEEENKLEPERKARLARDLGLQPRQVAV		
Medtr_HOX unknown2	(54)	TGHHSEKKRRLRVDQVKALEKNFEVENKLEPERKEKLAIELGLQPRQVAV		
Orysa_hox6	(24)	GGGAADRKKRFSEEQIKSLESMFATQTKLEPRQKLQALREGLQPRQVAV		
Consenso	(101)	EKKRRLT EQVQ LEKSFE ENKLEPERK QLAK LGLQPRQVAV		
dominio conservado in				
homeodominio de clase I			cremallera de leucina	
			151	200
Zeama_hox5	(121)	WFQNRRLRWKTKQLETDYDRLKAAYDALAADHQGLLADNDNLRAQVLSLT		
Aqufo_hox5	(110)	WFQNRRLRWKTKQLERDYDLLKASYDSLRSYDDIVKENKLEKSEVVSLSLT		
Arath_ATHB1	(112)	WFQNRRLRWKTKQLERDYDLLKSTYDQLLSNYDSIVMDNDKLRSEVTSLSLT		
Orysa_hox5	(130)	WFQNRRLRWKTKQLEHDFDRLKAAYDALAADHALLSDNDRLRAQVLSLT		
Crapl_CPHB-5	(117)	WFQNRRLRWKTKQLERDYDKLKSSYDSLSTYDSIRQENDKLKAEVLSLT		
Dauca_CHB3	(127)	WFQNRRLRWKTKQLEKDYDVLQNSYNSLKADYDNLAEKELKAEVLDLSLT		
Glyma_HD157	(130)	WFQNRRLRWKTKQLEKDYGVLKASYDRLKSDYESLVQENDKLKAEVNSLE		
Goshi_Hox5	(109)	WFQNRRLRWKTKQLERDYDLLKSSFDLSQSNYDTILKENKLEKSEVASLSLT		
Lyces_hox5	(110)	WFQNRRLRWKTKQLERDYDQLKSSYDSLSDFDSDVRKDNKLEKSEVVSLSLT		
Lyces_VaHOX1	(134)	WFQNRRLRWKTKQLEKDYDELNRNYDTLKSNNYNLLKEKEDLRTVEFRLT		
Medtr_HOX16_1	(131)	WFQNRRLRWKTKQLEKDYDVLKASFDLSKDDYDNLQENDKLKAEVNSLSLT		
Orysa_hox16	(121)	WFQNRRLRWKTKQLERDFDRLKASFDALRADHDALLQDNHRLHSQVMSLT		
Sacof_hox5	(122)	WFQNRRLRWKTKQLETDYDHLKAAYDALAADHQGLLADNDLSRAQVVSLSLT		
Sorbi_hox5	(124)	WFQNRRLRWKTKQLETDYDRLKAAYDALAADHQGLLADNDLSRAQVLSLT		
Triae_hox16	(120)	WFQNRRLRWKTKQLERDFDRLKASFDALRADHDALLQDNHRLHSQVVTLSLT		
Zeama_hox16	(118)	WFQNRRLRWKTKQLERDFDRLKASFDALRADHDALLQDNHRLHSQVVSLSLT		
Poptr_HOX16_1	(130)	WFQNRRLRWKTKQLEKDYDVLQSSYNSLKADYDNLLEKEKELKAEVNSLSLT		
Poptr_HOX16_2	(133)	WFQNRRLRWKTKQLEKDYEVLLQSSYNSLKADYDNLLEKEKELKAEVNSLSLT		
Poptr_HOX16_3	(124)	WFQNRRLRWKTKQLEKDYDVLQSSYNSLKADYDNLLEKEKELKAEVNSLSLT		
Phavu_hox16	(130)	WFQNRRLRWKTKQLEKDYDVLQSSYNSLKADYDNLLEKEKELKAEVNSLSLT		
Lotco_HOX16	(128)	WFQNRRLRWKTKQLEKDYDVLQSSYNSLKADYDNLLEKEKELKAEVNSLSLT		
Medtr_HOX unknown	(121)	WFQNRRLRWKTKQLEKDYDVLQSSYNSLKADYDNLLEKEKELKAEVNSLSLT		
Piclg_hox unknown	(131)	WFQNRRLRWKTKQLEKDFNVLLKQDYDALQDYDNLMEENNNLQAMIERMS		
Orysa_hox4	(95)	WFQNRRLRWKTKQLERDYAALRHSYDSLRLDHDALRRDKDALLAEIKELK		
Medtr_HOX unknown2	(104)	WFQNRRLRWKTKQLERDYGVLLKANYDALKLFDAIAQDNKAFHKEIKELK		
Orysa_hox6	(74)	WFQNRRLRWKTKQLEREYSALRDDYDALLCSESLKKEKLALIKOLEKLA		
Consenso	(151)	WFQNRRLRWKTKQLE DYD LK SYD L DYD LL EN L AEV SLT		
dominio conservado				

FIGURA 2 (continuación)

		201	250
Zeama_hox5	(171)	EKLQKETSPTS-----ATTAAQEVDPDEHTAVSGTEELL	
Aqufo_hox5	(160)	GKLQVKEGAGM-----ELNQISDPPLS--TEENVDTVTT	
Arath_ATHB1	(162)	EKLQKQETAN-----EPPGQVPEP-----NQLDPVYI	
Orysa_hox5	(180)	EKLQDKETSPTS-----SATITTAQEVDPDEHTAASTTGFA	
Crapl_CPHB-5	(167)	EKLQPKDDDDP-----SAEIGRNLSSSSPVDAAEPPC	
Dauca_CHB3	(177)	DKLLLKEDKGS-----KTVVFDKQKVSAAQQERVSNDIS	
Glyma_HD157	(180)	SKLILRDKEKEE-----NSDDKSSPDADVNSSSPHNNKEPMDLLI	
Goshi_Hox5	(159)	EKLQAKDVATE-----AIAGEKDEG-----LAAEMASA	
Lyces_hox5	(160)	EKLQKVVVGA-----GGNEKSD-----ILEVDAMTI	
Lyces_VaHOX1	(184)	GKLFKEKNG-----QLDLRDEHKHSNALAKETVVDPM	
Medtr_HOX16_1	(181)	NKLI PRDKEKV-----NSEDKSS-PEAINSP--HNNIDPMDIIS	
Orysa_hox16	(171)	EKLQEKETTTEGSAGAAVDVPGL-PAAADVAVPDAEPALEAAAAAFE	
Sacof_hox5	(172)	EKLQKETSPTS-----ATTAAQEVDPDEHTAASGTEKLL	
Sorbi_hox5	(174)	DKLQKETSPTS-----ATTAAQEVDPDEHTAASGTEKLL	
Triae_hox16	(170)	EKMQDKEAPEGSFGAAADASEPE-QAAAEAKASLADAEQAAAAEAEFVV	
Zeama_hox16	(168)	EKLQEKEDATEGGATADTAAP----AVDVEASLADDVEEPAEPAATFEV	
Poptr_HOX16_1	(180)	DKLLKEKEKEG-----ISELSDKDALSQEPKRAIADSAS	
Poptr_HOX16_2	(183)	NELLKEKEKEG-----SSELSDKDALSQEPKRAIADSAS	
Poptr_HOX16_3	(174)	EKLLSREESME-----SSEPFDVHS-PDAELEPIPTVS	
Phavu_hox16	(180)	SKLILRDKEKE-----NSDDKSS-PDAVNSP----HKEPMDLIS	
Lotco_HOX16	(178)	EKVLARENKQ-----VESETKGLVEPPQRPPLDVS	
Medtr_HOX unknown	(171)	NREPTESINLN-----KETEGSSNRSENSEIKLDM	
Piclg_hox unknown	(181)	SKSQSCNDQKFOAN-----SSKLQKDDQDLQLLMSATKVDCA	
Orysa_hox4	(145)	AKLGDEEAAAS-----FTSVKEEPAASDGPPAAGFGSSDS	
Medtr_HOX unknown2	(154)	SKLGEEEKSTIN-----VLVKEELTMLESCDEDKHNPSS	
Orysa_hox6	(124)	EMLQEPGRKYG-----DNAGDDARSGGVAGMKKEEFVG	
Consenso	(201)	EKLQ KE	
		dominio	
		conservado	
		251	300
Zeama_hox5	(206)	AQQ-----LKDNLHSSG-----	
Aqufo_hox5	(191)	MQFN-----VKVEDRLSSGSGVSAVVDEEC	
Arath_ATHB1	(190)	NAAA-----IKTEDRLSSGSGVSAVLDLDDA	
Orysa_hox5	(218)	TVDGALA-----APPPGHQPPHKDDLVSSTGNDGDDGAA	
Crapl_CPHB-5	(200)	LKLT-----VKVEDRLSTGNSGSAVMDGDG	
Dauca_CHB3	(212)	VGE-----VLSNSVMDCKQEDHNSVK--SDAVDS	
Glyma_HD157	(220)	ISKNATTTTTSENGTKVLSPLPLPIMVTCCKQEDANSK--SDVLDSDS	
Goshi_Hox5	(187)	LQFS-----MKVEDRLSSGSGVSAVVEDA	
Lyces_hox5	(187)	LQVK-----VKAGDRLSSGSGSAVVDEHS	
Lyces_VaHOX1	(219)	NVP-----ALVVKHQQEDLSSAK--SDVFDSES	
Medtr_HOX16_1	(217)	ITN-----SENGSKMS--LPNMVLKCKQEDANSK--SDVLDSDS	
Orysa_hox16	(220)	EQQE-----QQVKAEDRLSTGSGGSAVVDTDA	
Sacof_hox5	(207)	AQQ-----LKDDLHSSG-----	
Sorbi_hox5	(209)	VQQ-----LKDDLHSSG-----	
Triae_hox16	(219)	QQQ-----LHVKDEERLSPGSGGSAVLDARD	
Zeama_hox16	(213)	LQ-----EVKSEDRLSTGSGGSAVVADADA	
Poptr_HOX16_1	(215)	EGE-----VSKISTVACKQEDISSAK--SDIFDSDS	
Poptr_HOX16_2	(218)	EGE-----VSKTSTVACQEDISSAK--SDMFDSDS	
Poptr_HOX16_3	(208)	EN-----VSAIVPMVTPKQEESSAK--NDVFNDS	
Phavu_hox16	(214)	N-----STSENGTKVS--LPIMVT-CKQEDANSK--SDVLDSDS	
Lotco_HOX16	(210)	EGEE-----SKVSVGACKHEDISSAR--SESLDSDS	
Medtr_HOX unknown	(204)	RTP-----ASDSPLSTHQHTTSRTFFPPS	
Piclg_hox unknown	(219)	DKEN-----NNEGSSIGSEGSSVLDMDS	
Orysa_hox4	(180)	DSS-----AVLNDVDAAGAAPAATDALAPEA	
Medtr_HOX unknown2	(189)	TSNP-----SSESKDHLDDYDCIINNNDVGIG	
Orysa_hox6	(157)	AGG-----AATLYSSAEGGGTTTTAKL	
Consenso	(251)	K ED S G S V D D	

FIGURA 2 (continuación)

		301		350
Zeama_hox5	(218)	-----DCTGHGTLSE		
Aqufo_hox5	(216)	RQ-----LVDSVDSYFPGD----	DYQCIG-----	PVDGVQS
Arath_ATHB1	(215)	PQ-----LLDSCDSYFP-----		SIVPI
Orysa_hox5	(255)	VVVF-----DVTEGANDRLSCESAYFADAAEAYERDCAGHYALS--		SE
Crapl_CPHB-5	(225)	PQQL-----LDDSGDSYFEND----	EEYDCAA-----	ASLAA
Dauca_CHB3	(241)	PHYSD---EVYSSFMPEVDRSYVFEPQSS-----		DISQD
Glyma_HD157	(267)	PHCTS-----FVEPADSSHAPEPEDHSE-----		DFSQD
Goshi_Hox5	(212)	PQ-----LVDSGNSYFSPD----	EYSRGIG-----	PFDGVQS
Lyces_hox5	(212)	SQ-----LVDSGDSYFHTD----	HEEYPGPGGCNVPPMDGLQS	
Lyces_VaHOX1	(245)	PRYTSR---MHSSVVDQDDSSARAFETDQS-----		DSSQD
Medtr_HOX16_1	(253)	PHCNDG--NNLSSFIEPTDSDFS-----		QD
Orysa_hox16	(247)	QLVVGCGRQHAAVDSSVESYFPGG----	DEYHDC-VMGPMDDHAAGGIQS	
Sacof_hox5	(219)	-----DCTGHGALS--		SE
Sorbi_hox5	(221)	-----DFTGHGALS--		SE
Triae_hox16	(245)	ALLG-SGCGLAGVVDSSVDSYCFPGGAGGDEYHEC--	VVGPVAGG-IQS	
Zeama_hox16	(237)	LLYG---RFAAAVDSSVESYFPGGE---	DHYHDCGTMGPNHAGGGGIQ	
Poptr_HOX16_1	(244)	PHYAD---GVHSSLLEAGDSSYVFEPDQS-----		DLSQD
Poptr_HOX16_2	(247)	PHFAD---GVHSSLLEAGDSSHVFEPDQS-----		DLSQD
Poptr_HOX16_3	(236)	PRSEL-----EPRDCYRVFESDQP-----		DFSQV
Phavu_hox16	(249)	PHCTDG--NHPSFVEPADSSHAPEPDHS-----		DFSQD
Lotco_HOX16	(239)	PRYRDG-YGVNSAVLETCDSSYVVEPDQS-----		DMSQD
Medtr_HOX unknown	(228)	ARPS-----SGIAQLFQTSSRPE-----		IQC
Piclg_hox unknown	(243)	PGTIDS-----QQNIDSIGFSN-----		VKA
Orysa_hox4	(206)	CTFLGAPPAAGAGAGAAAAASHEEVFFHG-----		NFLKV
Medtr_HOX unknown2	(215)	ETSSLFPVDLKGSSSDSSAISSSGVLQSQ-----		QHLLL
Orysa_hox6	(181)	MPHFG-----SDDVDAGLFLRPS-----		SQHHP
Consenso	(301)	P	DS	Q
		351		400
Zeama_hox5	(229)	EDDGGVVSDEGCS-----	FALPDAMFAAGFTHHG-----	AEEVQ
Aqufo_hox5	(244)	EEDDISDDS---R-----	SYFSDVFPAPEQNHQ-----	ES-E
Arath_ATHB1	(232)	QDNSNASDHDNDR-----	SCFADVFPVPTTSPSHD-----	HHGE
Orysa_hox5	(296)	EEDGGAVSDEGCS-----	FDLPDAAAAAAMFGAAGVVHDDAADDEEAQ	
Crapl_CPHB-5	(254)	KEDDGSDEGG-----	CYFTEALAAEEEE-----	A
Dauca_CHB3	(272)	EEDDMGNLFLPS---YHVFSKTEDGSYSQPSNSSYF-----		GFPVE
Glyma_HD157	(295)	EEDNLSENLLMTFSSCCLPKVEEHCHYDGPENSCNF-----		GFQVE
Goshi_Hox5	(240)	EEDDGSNDNG-----	SYFSDVFATTEQG-----	
Lyces_hox5	(247)	EEDDGSDDHGSCH-----	GYFSNVFVAEEQHHEQ-----	GE-E
Lyces_VaHOX1	(276)	DDENFSKNMLSTAN---LLGKDADDDYPATSSNLSYFG-----		FPVE
Medtr_HOX16_1	(276)	EEDNDNLSHNLLTLP--CLPKVEDVCYDDPHENSCNF-----		GFPVE
Orysa_hox16	(292)	EEDDGAGSDEGCS-----	YYADDAGVLFADHGHHHHHQHADDDEEDGQQ	
Sacof_hox5	(230)	EEDGGVVSDEG-S-----	FDLPDAMFAAGVTHHGA-----	DAEEAQ
Sorbi_hox5	(232)	EEDGGVVSDEGCS-----	FDLPDAMFAAGVTHHGA-----	AEEAQ
Triae_hox16	(290)	EEDDGAGSDEGCS-----	YYPDDAAVFFAAAQGHGHHRTDDDDQDDGQ	
Zeama_hox16	(280)	SDDDGAGSDEGCS-----	YYADEAAAAAAFFAGHATHHHHADEDEDAGQ	
Poptr_HOX16_1	(275)	EEDNFSKSLLP---YVFPKLEDDDYSDPPAS-----		FE
Poptr_HOX16_2	(278)	EEDNLSKSLLP---YVFPKLEDDDYSDPPAS-----		FE
Poptr_HOX16_3	(260)	EEDNLTRSFLPP-----	PYFPKLYREPPASSRNF-----	EFSAE
Phavu_hox16	(281)	EEDNLSESLTLP-----	CLPKVEEACYDDPPENPCNF-----	GFHVE
Lotco_HOX16	(272)	EEDNLTKTLPPY---MFSKLGMDYSDPPESSCNF-----		GFPEE
Medtr_HOX unknown	(249)	QKIDQMVKEES-----	LSNMFCGMDDQAG-----	
Piclg_hox unknown	(263)	RDLRLCENFRP-----	KVEENVSQADEPCNYLFYN-----	NLETGP
Orysa_hox4	(240)	EEDETGFLLDDDEP-----	CGGFFADDQPPP-----	LSSWW
Medtr_HOX unknown2	(251)	SPSSSMNCFQYQKSYHVKMEEHNFSLADEACN-----		FFSDE
Orysa_hox6	(204)	PPPHAGAGFTSSE-----	PAADHQSFNFHSSWPSS-----	TEQTCS
Consenso	(351)	EED	D	E

FIGURA 2 (continuación)

			Cola de Trp	
			401	414
Zeama_hox5	(263)		LANWTSMEFWN----	
Aqufo_hox5	(273)		TLGWWDWA-----	
Arath_ATHB1	(265)		SLAFWGWP-----	
Orysa_hox5	(340)		IGSWTAWFWS----	
Crapl_CPHB-5	(278)		FFAWCIWS-----	
Dauca_CHB3	(312)		DHTFGFWGTEL---	
Glyma_HD157	(337)		DQTFCFWPY-----	
Goshi_Hox5	(263)		ALGLWAWX-----	
Lyces_hox5	(279)		HIGWFWS-----	
Lyces_VaHOX1	(315)		DQGFGFWTY-----	
Medtr_HOX16_1	(316)		DQTFCFWPY-----	
Orysa_hox16	(336)		ISCWWMWN-----	
Sacof_hox5	(265)		LANWTSWFWN----	
Sorbi_hox5	(266)		LANWTSWFWN----	
Triae_hox16	(334)		IS-YWMWN-----	
Zeama_hox16	(324)		IS-WWMWN-----	
Poptr_HOX16_1	(306)		DHAFWSWSY-----	
Poptr_HOX16_2	(309)		DHAFWCWSY-----	
Poptr_HOX16_3	(294)		DQPFWSWIY-----	
Phavu_hox16	(319)		DQTFCFWPY-----	
Lotco_HOX16	(310)		DHALWSWSY-----	
Medtr_HOX desconocido	(273)		---FWPWLEQQHFN	
Piclg_hox desconocido	(299)		ILWDYNWSSGL---	
Orysa_hox4	(270)		AEPTEHWN-----	
Medtr_HOX desconocido 2	(289)		CAPTLOWYCPDQWS	
Orysa_hox6	(240)		STPWWEFESE----	
Consenso	(401)		W	

FIGURA 2 (continuación)

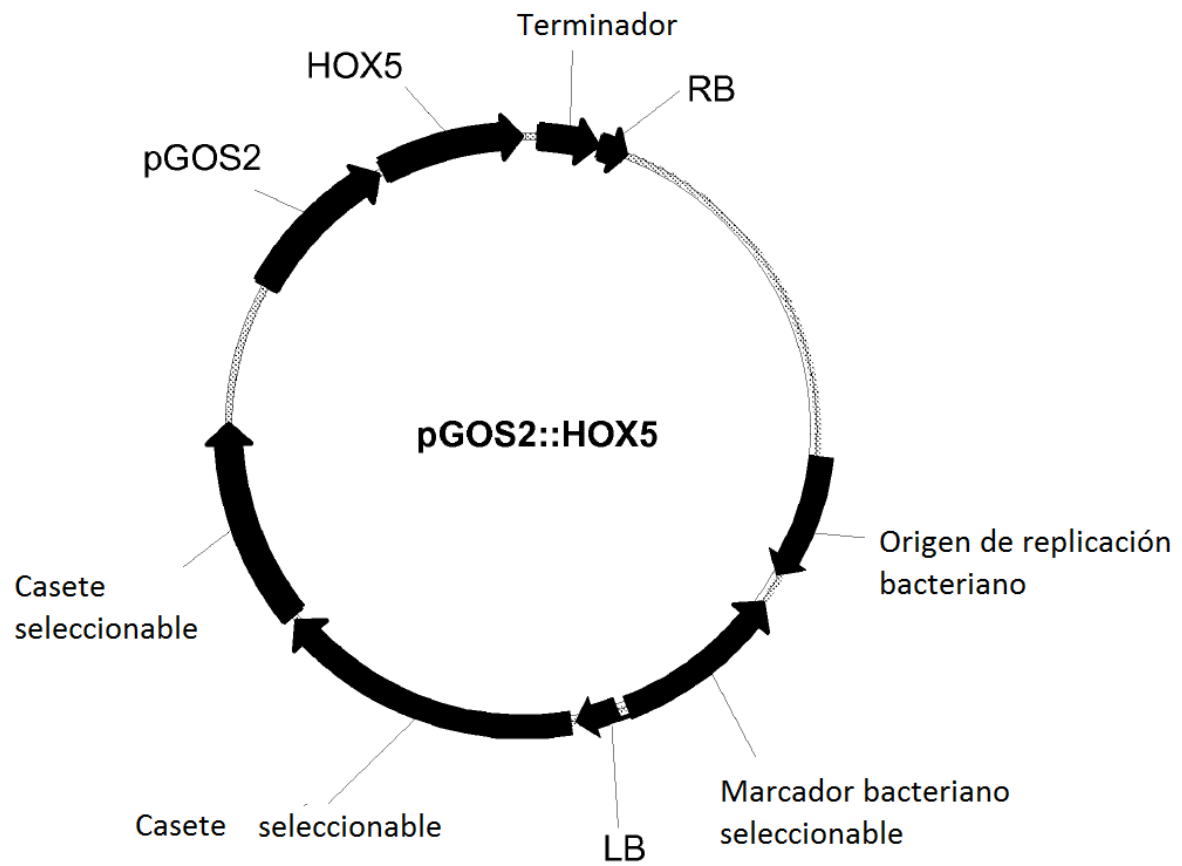


FIGURA 3

SEC ID Nº 1 secuencia de ADNc de *Oryza sativa* *Oryza hox5* XM 482406.1

ATGGATCCCGGCCGCGTCGTGTTCTGACTCCGGCGTGCGCGGGCGGGCGTGCCCCGGCGGGCGCGCAG
ATGCTTCTCTTCGGCGGGCGGGCGGCAGCGCCAACAGCGGGCGGCTTCTTCCGAGGCGTGCCGGCGGGCG
GTGCTGGGGATGGATGAATCGCGGTCTGTCGTCTGCGCGGGCGGGGGCGGGGGCGAAGCGGCCGTTT
TTCACGACGCACGAGGAGCTCCTGGAGGAGGAGTACTACGACGAGCAGGCGCCGGAGAAGAAGCGG
CGGCTGACGGCGGAGCAGGTGCAGATGCTGGAGCGGAGCTTCGAGGAGGAGAACAAGCTGGAGCCG
GAGCGGAAGACGGAGCTCGCCCCGCCCTCGGCATGGCCCCCGGCAGGTCGCCGTCTGGTTCCAG
AACCGCCGCGCCCGCTGGAAGACCAAGCAGCTCGAGCACGACTTCGACCGCCTCAAGGCCGCCTAC
GACGCCCTCGCCCGCCGACCACCATGCCCTCCTCTCCGACAACGACCGCCTCCGCGCGCAGGTAATC
TCATTAACCGAGAAGCTGCAAGACAAGGAGACGTCGCCGTCGTCGCGCACCATCACCACCGCGGCG
CAGGAGGTGACACGCGGACGAACACACGAGGCCGCGTCAACCACCGGCTTCGCCACCGTCGAC
GGCGCATTTGGCGGCGCCACCGCCCGGCCACCAGCAGCCCGCATAAAGATGATCTTGTGAGCAGC
GGCGGCACCAACGACGACGGCGATGGCGGCGCGGCCGTTGGTGGTCTTCGACGTCACCGAGGGCGCC
AACGACCGCCTCAGCTGCGAGTCGGCGTACTTCGCCGACGCCGCGGAGGCGTACGAGCGCGACTGC
GCCGGGCACTACGCCCTCTCGTCGGAGGAGGAGGACGGCGGCGCGGTACGCGACGAGGGCTGCAGC
TTCGACCTCCCGACGCCGCCGCCGCCGCCGATGTTTCGGCGCGCGGAGTTGTGCACCAC
GACGCCGCGGACGACGAGGAGGCGCAGCTCGGCAGCTGGACCGCCTGGTTCTGGAGCT**TGA**

SEC ID Nº 2 secuencia de aminoácidos traducidos *Oryza sativa* *Oryza_hox5* sequence

MDPGRVVFDSGVARRACPGBAQMLLFGGGGSANSGGFFRGVPAAVLGMDESRRSSSSAAGAGAKRPF
FTTHEELLEEEYYDEQAPEKKRRLTAEQVQMLERSFEEENKLEPERKTELARRLGMAPRQVAVWFQ
NRRARWKTQLEHDFDRLKAAYDALAADHALLSDNDRLRAQVISLTEKLQDKETSPSSATITTA
QEVDQPDHTEAASTTGATVDGALAAPPPGHQPPHKDDLVSSTGNDGDDGAHVVFVDVTEGA
NDRLSCESAYFADAAEAYERDCAGHYALSSEEDDGAHSVDEGCSFDLPDAAAAAAMFGAAGVVHH
DAADDEEAQLGSWTAWFWS

SEC ID Nº 3 secuencia de ADNc de *Oryza sativa* *Oryza_hox16* XM 467603.1

ATGGAGTCCGGCCGGCTCATCTTCAGCACGGCGGGCTCCGGCGCCGGGCAGATGCTCTTCTTGGAC
TGCGGCGCTGGCGGCGGGCGTCTGGCGGCGGGGCCATGTTCCATCGAGGGGCGAGACCGGTGCTC
GGCATGGAGGAAGGAGGGCGCGGCGTCAAGCGGCCCTTCTTACCACCCCCGACGAGCTCCTCGAA
GAGGAGTACTACGACGAGCAGCTCCCGGAGAAGAAGCGGCGCCTCACGCCGAGCAGGTGCATCTG
CTGGAGAGGAGCTTCGAGGAGGAGAACAAGCTGGAGCCGGAGCGGAAGACGGAGCTGGCGCGGAAG
CTAGGGCTGCAGCCGCGCAGGTGCGCGTGTGGTTCCAGAACC GCCGCGCGCTGGAAGACCAAG
CAGCTCGAGCGCGACTTCGACCGCCTCAAGGCGTCGTTTCGACGCCCTCCGCGCCGACCACGACGCC
CTCCTCCAGGACAACCACCGCCTCCACTCTCAGGTCATGTCGTTGACCGAGAAGCTGCAAGAGAAG
GAGACGACGACCGAGGGCAGCGCCGGCGGCGGCGTTGACGTCCCGGGCTTGCTGCGCGGCCGAC
GTGAAGGTCGCCGTCCCGGACGCCGAGGAACCGGCGCTGGAGGAGGCGGCGGCGGCTTCGAGGAG
CAGCAGGAGCAGCAGGTGAAGGCCGAGGACAGGCTGAGCACGGGACGCGCGGAGCGCGGTGGTG
GACACGGACGCGCAACTGGTGGTCGGGTGCGGCCGGCAAGCATCTCGCCGCCGTGGACAGCAGCGT
GGAGTCGTACTTCCCGGGCGGCGACGAGTACCACGACTGCGTGATGGGCCCCATGGACCACGCCGC
GGGGGGCATCCAGTCGGAGGAGGACGACGGCGCCGGCAGCGACGAGGGCTGCAGTACTACGCCGA
CGACGCCGGCGTCTTCTCGCCGACCACGGCCACCACCACCACCAACACGCGGACGACGACGA
GGAGGACGGCCAGCAGATCAGCTGCTGGTGGATGTGGAAGTAGATTTCTCGCGCGCGCGCTGCTC
GTGCATTCAATTCTCGTGTTAAAAAATCGTTCTCTTTTTCATTTTTCGCTTCTTTGTCTGTAAT
GTTGAGTTTCGATCGGCTATGAGAAGGAAGGAGGTGTATGCATGTGCATGGTATGGTAGGGTAACA
CATCGG**TGA**

FIGURA 4 (continuación)

SEC ID Nº 4 secuencia de aminoácidos traducidos Oryza sativa Orysa_hox16

MESGRILFSTAGSGAGQMLFLDCGAGGGGVGGGAMFHRGARPVLGMEEGGRGVKRPFFFTTPDELLE
EEYYDEQLPEKKRRLTPEQVHLLERSFEEENKLEPERKTELARKLGLQPRQVAVWFQNRARWTK
QLERDFDRLKASFDALRADHDALLQDNHRLHSQVMSLTEKLQEKETTTEGSAGAAVDVPGLPAAAD
VKVAVPDAAEPALAEAAAAFEEQQEQVKAEDRLSTGSGGSAVVDTAQLVVGCGRQHLLAAVDSSV
ESYFPGGDEYHDCVMGPMMDHAAGGIQSEEDDAGSDEGCSYYADDAGVLFADHGHHHHHQHADDDE
EDGQQISCWMMWN

**SEC ID Nº 5 Contig. de la secuencia de ADNc de Zea mays Zeama_hox5
de esencialmente CO458693 & DV024016**

ATGGATCCGAGCGCGGTCACTTTTCGACTCTGGCGGCGCGCGCGGGGCGGCGGCGCGCAGATGCTG
CTCTTCGGCGGCGGAGGCAGCGCCAACAGCAACGGCTTCTTCCGAGGTGTTCCGATGGCGGTCTCTG
GGCATGGACGACGCGACGCGCGTGGGCAAGCGGCCCTTCTTCACGACACACGAGGAGCTCCTAGAG
GAGGAGTACTACGACGAGCAGGCGCCGAGAGAAGAAGCGCCGACTGACGGCGGAGCAGGTGCAGCTG
CTGGAGCGGAGCTTCGAAGAAGAGAACAAGCTGGAGCCGGAGCGCAAGACCGAGCTGGCTCGCCGC
CTGGGGATGGCGCCCCGCCAGGTAGCTGTTTGGTTCCAGAACCGCCGCGCGCTGGAAGACCAAG
CAACTCGAGACCGACTATGACCGCCTCAAGGCTGCTTACGACGCACTCGCCGCCGACCACCAGGGC
CTCCTGGCCGACAACGATAACCTCCGGGCACAGGTGATCTCCCTGACGGAGAAGCTGCAAGGCAAG
GAGACATCCCCGTCGGCAACCACTGCTGCCCAAGAGGTGACACGACGACGAACACACCGCTGTG
TCAGGCACGGAAGAACTGCTGGCGCAGCAGCTCAAGGACAACCTCCACAGCAGCGCGACTGCACT
GGCCATGGCACCCCTCTCTTCGGAAGAAGACGAGGTGGCGTGGTCACTGACGAGGGCTGCAGCTTC
GCTCTCCCGGATGCCATGTTTCGCTGCCGGGTTACCCACCATGGCGCCGAGGAGGTGCAGCTGGCC
AACTGGACATCCATGTTCTGGAAC**TGA**

SEC ID Nº 6 secuencia de aminoácidos traducidos Zea mays Zeama_hox5

MDPSAVSFDSGGARRGGGAQMLLFGGGGSANSNGFFRGPVMAVLGMDDATRVGKRPFFTTHEELLE
EEYYDEQAPEKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAVWFQNRARWTK
QLETQDRLKAAYDALADHQGLLADNDNLRAQVLSLTEKLQKETSPTSATTAQEVDPDEHTAV
SGTEELLAQQLKDNLHSSGDCGTGHGTLSSSEEDDGGVVSDEGCSFALPDAMFAAGFTTHGAEEVQLA
NWTSMFWN

SEC ID Nº 7 secuencia de ADNc de Zea mays Zeama_hox16 AY105265

ATGGAGTCTGGACGGCTCATCTTCAACGCGCGGGGCTCTGGCGCCGGGAGATGCTCTTCTCGAC
TGCGGCGCAGGCGGCGGTCCCGGCGGCGGCTTGTTCATCGAGGCGGGAGACCGATGCTTGGCCTT
GAAGAAGGGCGCGCGTAAAACGGCCCTTCTTCACTCGCCCGACGAGCTCCTCGAGGAAGAGTAC
TACGACGAGCAGCTGCCGGAGAAGAAGCGCCGCTCACCCAGAGCAGGTGCTTCTGCTGGAGAGG
AGCTTCGAGGAGGAGAACAAGCTGGAGCCGGAGCGCAAGACGGAGCTGGCGCGCAAGCTGGGCCTG
CAGCCTCGCCAGGTGGCCGTCTGGTTCCAGAACCAGCGCGCCGGTGGGAAGACCAAGCAGCTCGAG
CGCGACTTCGACCGCCTCAAGGCCTCCTTCGACGCTCTCCGAGCGGACCACGACGCCCTCCTCCAG
GACAACAACCGCCTCCGCTCACAGGTTGTGTCTGTTGACCGAGAAGCTGCAAGAGAAGGAGGATGCG
ACGGAGGGCGGCGCCACCGCTGACACCGCGCGCGGCGGTGGACGTCGAGGCTTCCCTGGCCGAC
GACGTCGAGGAGCCAGCAGAGCCTGCGGCGACGTTTCGAGGTGCTGCAGGAGGTGAAGTCCGAGGAC
AGGCTGAGCACCGGCAGCGGCGGAGCGCGGTGGTGGACGCGGACGCGCTGCTGTACGGCAGGTTT
GCCGCGGCAGTTGATAGCAGCGTGGAGTCGTACTTCCCCGGCGGCGAGGACCACTACCACGACTGC
GGGACGATGGGCCCCGTGAATCATGGCGCCGAGGAGGCATCCAGTCGGACGACGACGGCGCCGGC
AGCGACGAGGGGTGCAGCTACTACGCCGACGAAGCCGCCGCCGCCGCCGCGTTCTTCGCCGGA
CACGCCACCCACCACCGCGGACGAGGACGAGGACGCCGCCGAGATCAGCTGGTGGATGTGGAAC
TAG

FIGURA 4 (continuación)

SEC ID Nº 8 secuencia de aminoácidos traducidos Zea mays Zeama_hox16

MESGRILFNAPGSGAGQMLFLDCGAGGGPGGGLFHRGGRPMLGLEEGRGVVRPFFTSPELLEEEY
YDEQLPEKKRRLTPEQVLLERSFEEENKLEPERKTELARKLGLQPRQVAVWFQNRRLRWKTKQLE
RDFDRLKASFDALRADHDALLQDNNRLRSQVVSLEKLEKEDATEGGATADTAAPAVDVEASLAD
DVEEPAEPAATFEVLQEVKSEDLSTGSGGSAVVDADALLYGRFAAAVDSSVESYFPGGEDHYHDC
GTMGPVNHGAGGGIQSDDDGAGSDEGCSYADEAAAAAAFFAGHATHHHADEDEDAGQISWWMWN

SEC ID Nº 9 Contig. de la secuencia de ADNc de Saccharum officinarum Sacof_hox5 de esencialmente CA088615, CA115362 & CA142506

ATGGATCCGAGCGCGGTACGTTTCAACTCCGGCGGCGCGCGCGGGGCGGCGGCGGCACGCAGATG
CTGCTCTTCGGCGGCGGAGGCAGCGCCAACAGCAACGGCTTCTTCCGAGGTGTTCCGATGGCGGTC
CTGGGCATGGACGACGCGACGCGGTGGGCAAGCGGCCCTTCTTACCACACACGAGGAGCTCCTG
GAGGAGGAGTACTACGACGAGCAGGCGCCGAGAAGAAGCGCCGTCTGACGGCGGAGCAGGTGCAG
CTGCTGGAGCGGAGCTTCGAGGAAGAGAACAAGCTGGAGCCCGAGCGCAAGACCGAGCTGGCTCGC
CGCCTCGGGATGGCGCCCCCGCCAGGTGGCCGTCTGGTTCCAGAACCGCCGCGCGCTGGAAGACC
AAGCAGCTCGAGACCGACTATGACCACCTCAAGGCTGCCTACGACGCGCTCGCCGCCGACCACCAG
GGCCTCCTGGCCGACAACGATAGCCTCCGGGCACAGGTGGTCTCCCTAACAGAGAAGCTGCAAGGC
AAGGAGACATCCCCGTGGCCACCACTGCTGCCAAGAGGTGACCGAGCCAGACGAACACACCGCG
GCGTCAGGCACTGAGAACTGCTGGCGCAGCAGCTCAAGGACGACCTCCACAGCAGCGGCGACTGC
ACTGGCCATGGTGCCCTCTCCTCAGAGGAAGAAGATGGTGGTGTGGTCAGTGACGAGGGCAGCTTT
GATCTCCCGATGCCATGTTTGCTGCCGGGTACCCACCATGGCGCCGACGCCGAGGAGGCACAG
CTGGCCAACCTGGACATCCTGGTTCTGGAAC**TGA**

SEC ID Nº 10 secuencia de aminoácidos traducidos Saccharum officinarum Sacof_hox5

MDPSAVSFNSGGARRGGGTQMLLFGGGGSANSNGFFRGVPMVLMDDATRVGKRPFFTTHEELL
EEYYDEQAPEKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAVWFQNRRLRWKT
KQLETDYDHLKAAYDALADHQGLLADNDSLRAQVVSLEKLEKEDATEGGATADTAAPAVDVEASLAD
ASGTEKLLAQQLKDDLHSSGDCGTHGALSSEEDGGVVSDEGSFDLPDAMFAAGVTHHGADAEAAQ
LANWTSWFWN

SEC ID Nº 11 Sorghum bicolor Sorbi_hox5 BE363386, CD432381

ATGGATCCGAGCGCGGTACGTTTCAACTCCGGCGGCGCGCGGGGCGGCGGCGGCGGCGGCGCGC
CAGATGCTGCTCTTCGGCGGCGGAGGCAGCGCCAACAGCAACGGCTTCTTCCGAGGTGTTCCGATG
GCGGTCTTGGGCATGGACGACGCGACGCGGTGGGCAAGCGGCCCTTCTTACCACGCACGAGGAG
CTCCTGGAGGAGGAGTACTACGACGAGCAGGCGCCCGAGAAGAAGCGCCGTCTGACGGCGGAGCAG
GTGCAGCTGCTGGAGCGGAGCTTCGAGGAAGAGAACAAGCTGGAGCCGAGCGCAAGACCGAGCTG
GCTCGCCGCTCGGGATGGCGCCTCGCCAGGTGGCCGTCTGGTTCCAGAACCGCCGCGCGCTGG
AAGACTAAGCAGCTCGAGACCGACTATGACCGCCTCAAGGCTGCCTACGACGCGCTCGCCGCCGAC
CACCAGGGCCTCCTGGCCGACAACGATAGCCTCCGGGCACAGGTGATCTCCCTAACGGATAAGCTG
CAACGCAAGGAGACATCCCCGTGGCGACCACTGCTGCCAAGAGGTGACCGAGCCAGACGAACAC
ACCGCTGCGTCAGGCATGAGAACTGCTGGTGACGAGCTCAAGGACGACCTCCACAGCAGCGGC
GACTTCACTGGCCATGGTGCCCTCTCTTCAGAGGAAGAGGATGGTGGTGTGGTCAGCGACGAGGGC
TGCAGCTTTGATCTCCCGATGCCATGTTGCTGCCGGGTACCCACCATGGCGCCGAGGAGGCG
CAGCTGGCCAACCTGGACATCCTGGTTCTGGAAC**TGA**

FIGURA 4 (continuación)

SEC ID Nº 12 Secuencia de aminoácidos traducidos Sorghum bicolor Sorbi_hox5

MDPSAVSFDSGGARRGGGGGAQMLLFGGGGSANSNGFFRGVPMVAVLGMDATRVGKRPFFTTHEE
LLEEEYYDEQAPEKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAVWFQNRARW
KTKQLETDYDRLKAAYDALAADHQGLLADNDSLRAQVISLTDKLQRKETSPSATTAQEVDPDEH
TAASGTEKLLVQQLKDDLHSSGDFGTGHGALSSEEDDGGVVSDEGCSFDLPDAMFAAGVTHHGAEAA
QLANWTSWFVN

**SEC ID Nº 13 Triticum aestivum Triae_hox16 DR735359,DR741379,
CD916488**

ATGGAGCCCCGGCGCGCTCATCTTCAACACGTCGGGCTCCGGCAACGGACAGATGCTCTTCATGGAC
TGCGGCGCGGGCGGCATCGCCGGCGCGGCCGCGCATGTTCCATCGAGGGGTGAGACCGGTCTCTCGGC
GGCATGGAAGAAGGGCGCGCGTGAAGCGGCCCTTCTTCACCTCGCCGGATGACATGCTCGAGGAG
GAGTACTACGACGAGCAGCTCCCGGAGAAGAAGCGGCGCCTCACCCCGGAGCAGGTCCACCTGCTG
GAGAGGAGCTTCGAGGAGGAGAACAAGCTGGAGCCGGAGAGGAAGACGGAGCTGGCCCCGAAGCTC
GGGCTGCAGCCACGCCAGGTGGCCGTCTGGTTCCAGAACC GCCCGCGCCCGGTGGAAGACAAAGACG
CTGGAGCGCAGCTTCGACCGCCTCAAGGCGTCTTCGACGCCCTCCGGGCCGACCACGACGCCCTC
CTCCAGGACAACCACCGGTCCGGTCACAGGTGGTAACGTTGACCGAGAAGATGCAAGATAAGGAG
GCGCCGAAGGCAGCTTCGGTGCAGCCGCCGACGCCCTCGGAGCCGGAGCAGGCGGCGCGGAGGCG
AAGGCTTCCTTGGCCGACGCCGAGGAGCAGGCCGCGGCAGCGAGGCGTTCGAGGTGGTGCAGCAG
CAGCTGCACGTGAAGGACGAGGAGAGGCTGAGCCCGGGGAGCGGCGGGAGCGCGGTGCTGGACGCG
AGGGACGCGCTGCTCGGGAGCGGATGCGGCCTCGCCGGCGTGGTGGACAGCAGCGTGGACTCGTAC
TGCTTCCCGGGGGCGCCGGCGGCGACGAGTACCACGAGTGCCTGGTGGGCCCCGTGGCGGGCGGC
ATCCAGTCGGAGGAGGACGACGGCGCGGCGACGAGGGCTGCAGCTACTACCCCGACGACGCC
GCCGTCTTCTTCGCCGCCGCGCAAGGGCACGCCACCATCGCACGGACGACGACGATCAGCAGGAC
GACGCCCAGATCAGCTACTGGATGTGGAAC**TAG**

SEC ID Nº 14 Secuencia de aminoácidos traducidos Triticum aestivum Triae_hox16

MEPGRLIFNTSGSNGQMLFMDCGAGGIAGAAGMFHRGVPRVPLGGMEEGRGVGRPFFTSPPDMLLE
EYYDEQLPEKKRRLTPEQVHLLERSFEEENKLEPERKTELARKLGLQPRQVAVWFQNRARWKT
LERDFDRLKASFDALRADHDALLQDNHRLRSQVVTLETKMQDKEAPEGSFGAAADASEPEQAAAEA
KASLADAEQAAAAEAFEVQQLHVKDEERLSPGSGGSAVLDAARDALLGSGCLAGVVDSSVDSY
CFPGGAGGDEYHECVVGPVAGGIQSEEDDGAGSDEGCSYYPDDAAVFFAAAQGHGHHRTDDDDQD
DGQISYWMWN

SEC ID Nº 15 secuencia de ADNc de Arabidopsis thaliana Arath_ATH1

ATGGAATCCAATTCGTTTTTCTTCGATCCATCTGCTTCACACGGCAACAGCATGTTCTTCTTGGG
AATCTCAATCCCGTCGTCCAAGGAGGAGGAGCAAGATCGATGATGAACATGGAGGAACTTCGAAG
CGAAGGCCCTTCTTTAGCTCCCTGAGGATCTCTACGACGATGACTTTTACGACGACCAGTTGCCT
GAAAAGAAGCGTCGCCTCACTACCGAACAGTGCATCTGCTGGAGAAAAGCTTCGAGACAGAGAAC
AAGCTAGAGCCTGAACGCAAGACTCAGCTTGCCAAGAAGCTTGGTCTACAGCCAAGGCAAGTGGCT
GTCTGGTTTCAGAATCGCCGAGCTCGTTGGAAAACAAAACAGCTTGAGAGAGACTACGATCTTCTC
AAGTCCACTTACGACCAACTTCTTTCTAACTACGACTCCATCGTCATGGACAACGATAAGCTCAGA
TCCGAGGTTACTTCCCTGACCGAAAAGCTTCAGGGCAAACAAGAGACAGCTAATGAACCACTGGT
CAAGTGCCCGAACCAAACTTGATCCGGTTTACATTAATGCGGCAGCAATCAAAACCGAGGAC
CGGTTAAGTTCAGGGAGCGTTGGGAGCGCGGTACTAGACGACGACGCACCTCACTACTAGACAGC

FIGURA 4 (continuación)

TGTGACTCTTACTTCCCAAGCATCGTACCCATCCAAGACAACAGCAACGCCAGTGATCATGACAAT
GACCGGAGCTGTTTCGCCGACGTCTTTGTGCCACCACTTCACCGTCGCACGATCATCACGGTGAA
TCATTGGCTTTCTGGGGATGGCCTTA

SEC ID Nº 16 Secuencia de aminoácidos traducidos Arabidopsis thaliana Arath_ATHB1

MESNSFFFDPSASHGNSMFFLGNLNPVVQGGGARSMMMEETSKRRPFFSSPEDLYDDDFYDDQLP
EKKRRLTTEQVHLLLEKSFETENKLEPERKTQLAKKLGLQPRQVAVWFQNRRLRWKTKQLERDYDLL
KSTYDQLLSNYDSIVMDNDKLRSEVTSLETKLQKGQETANEPPGQVPEPNQLDPVYINAAAIKTED
RLSSGSVGSVAVLDDDDAPQLLDSCDSYFPSIVPIQDNSNASDHDNDRSCFADVFPVPTTSPSHDHHGE
SLAFWGW

SEC ID Nº 17 secuencia de ADNc de Daucus carota Dauca CHB3 D26575

ATGGCGGGTTCGGAGGGTGTCTATGGGGAGGGAGCCAATACGACGTCGGCTAGCCTGTTGTTTCAT
AGTCAAAGACCTGAGCCTTTCTTTCTTTCTGCACCTTCTCCTTCTCTAATTGGTTCAAAATCCATG
GTTAGCTTTCAAGATGCTAAGCGAAAAAATCCCTACGATGGGTCTTTATGCGGTCATATGATGAA
GAAGAAATTGGGGATGAAGAATATGATGAATACTTTCAGCAGCCTGAGAAGAAGAGGAGGCTCAAG
GCTGATCAAATCCAGTTTCTTGAGAAAAGTTTGAGACTGATAACAAGCTTGAGCCTGAAAGAAAA
GTTTCAGCTTGCAAAAGAACTCGGCTTGCAGCCAAGACAGGTTGCGATATGGTTTCAGAACCGTCGA
GCACGGTGGAAGACCAAAACACTAGAAAAAGATTATGATGTATTGCAAAATAGCTACAACAGCCTC
AAGGCTGACTATGACAATCTACTTGCCGAGAAAGAAAACTTAAAGCCGAGGTCTCGACCTGACA
GACAAGCTACTTCTCAAAGAAGATAAGGGGAGCAAGACAGTAGTTTTTGATAAGCAAAAGGTGTCT
GCAGCATTTCCAACAAGAACGTGTTAGTAATGACATATCTGTGGGTGAAGTACTCAGTAACCTCAGTT
ATGGACTGCAAGCAAGAAGATCATAACTCTGTGAAAAGTGATGCAGTTGATTCTGACAGTCCACAC
TACAGTGATGAAGTCTACTCCAGTTTTATGGAGCCAGTGGATCGCTCTTATGTTTTGAACCTGCT
CAGTCGGATATATCTCAAGATGAAGAAGATGACATGGGGAACAACCTATTCTCCCATCATATCAT
GTTTTCTCAAAGACTGAAGACGGTAGTTACTCCGACCAGCCTTCGAACCTCTCGTACTTTGGCTTC
CCAGTTGAAGATCATACGTTTGGCTTTTGGGGTACTGAATTATA

SEC ID Nº 18 Secuencia de aminoácidos traducidos Daucus carota Dauca_CHB3

MAGRRVFYEGEANTTSASLLFHSQRPEPFFLSAPSPSLIGSKSMVSFQDAKRKNPYDGGFFMRSYDE
EEIGDEEYDEYFQQPEKKRRLKADQIQFLEKSFETDNKLEPERKVQLAKELGLQPRQV
AIWFQNRRLRWKTKLEKDYDVLQNSYNSLKADYDNLAEKEKLKAEVLDLTKLLLKEDKG
SKTVVFDKQKVSAAFQQQERVSNDISVGEVLSNSVMDCKQEDHNSVKSDAVIDSDSPHYSDEVYSSFM
EPVDRSYVFEPAQSDISQDEEDDMGNLFLPSYHVFSKTEDGSYSDQPSNSSYFGFPVEDHTFGFW
GTEL

SEC ID Nº 19 secuencia de ADNc de Glycine max Glyma_HD157 AF184278

ATGGCGAGTGGCAAGCTTTATGCGGGTCAACATGTCACCTTCTCCTCCAAAACGAAAGGCTCCCT
TGCTCCTCTGAAGTCCTTGAGTCTCTTTGGGCTCAGACCTCTAACCCTGCTTCCTTCCAAGGTTCA
AAACCCGTGGTTGATTTTGAGAATGTAAGTGGGAGCAGGATGACGGATAGGCCTTTCTTTCAAGCG
TTGGAGAAGGAAGAGAACTGTGATGAGGATTACGAGGGGTGTTTCCACCAACCGGGGAAGAAAAGG
AGGCTCACAAGCGAACAAGTTTCAGTTCCTTGAAAGGAACCTTTGAGGTAGAGAACAAGCTTGAACCC
GAAAGGAAAGTCCAACCTTGCAAAAGAGCTTGCTTGCAGCCAAGGCAAGTTGCTATATGGTTCCAA
AACCAGAGGGCAAGGTTCAAGACCAAGCAGCTAGAAAAAGACTATGGCGTGTGAAAGCTAGTTAT
GACAGACTCAAAAGTGACTATGAAAGTCTTGTTCAGAGAATGACAAGTTAAAGCAGAGGTGAAT

FIGURA 4 (continuación)

TCTCTGGAGAGCAAATTGATTCTTAGAGATAAAGAGAAGGAGGAGAATTCGGATGACAAGTCATCT
CCTGATGATGCTGTCAATTCTTCTTACCCCCACAACAACAAGGAGCCTATGGATTTATTAATTATT
TCAAAAAATGCAACAACAACAACAACATCTGAAAATGGGACCAAAGTGTTGTCACCACTCCCACTC
CCTATTATGGTAACATGCTGCAAGCAAGAAGATGCCAACTCAGCCAAAAGTGATGTCCTTGATTGCG
GATAGCCCACATTGCACTTCATTTCGTGGAGCCTGCTGATTCTCTCATGCCTTTGAACCAGAAGAC
CACTCAGAAGACTTCTCCCAAGATGAAGAGGATAACCTTAGTGAAAACCTTTTGATGACCTTCCCT
TCTTCTTGTGCTTACCTAAGGTTGAAGAAGACTGCTATGACGGCCCTCCTGAAAACCTCTTGTAAT
TTTGGCTTCCAGGTTGAGGATCAAACCTTCTGTTTCTGGCCCTAT**TGA**

SEC ID N° 20 Secuencia de aminoácidos traducidos de Glycine maiz Glyma_HD157

MASGKLYAGSNMSLLLQNERLPCSEVLESLSWAQTSNPASFQGSKPVVDFFENVSGSRMTDRPFFQA
LEKEENCDEDEYEGCFHQPGKKRRLTSEQVQFLERNFEVENKLEPERKVQLAKELGLQPRQVAIWFO
NRRARFKTKQLEKDYGLKASYDRLKSDYESLVQENDKLKAEVNSLESKLILRDKEKEENSDDKSS
PDDAVNSSSPHNKEPMDLLIIISKNATTTTTSENGTKVLSPLPLPIMVTCCKQEDANSKSDVLDSD
SPHCTSFVEPADSSHAPEPEDHSEDFSQDEEDNLSENLLMTFPSSCCLPKVEEHCYDGPPENSCN
FGFQVEDQTFCFWPY

SEC	ID	N°	21	Craterostigma	plantagineum	CPHB-5
AF443621						

ATGAACTCTGCTCGGATTTTCTTCGACCCATCTTCCCACGGCAACATGCTGCAGTTTCTTGGAAC
GCCGGCGGCGATTTCATCCGTTTTCCGAGGAACAAGATCGTCGTCGGTGCTGAACATGGAGGAGAGC
TCGTTAAACGACAGATTTTCAGCGGCGGCGGCGGCGGATGAATTCTACGACGAGGAATACTACGAC
GAGCAGTTGTTGCCTGAGAAGAAGCGCCGACTCACCGCCGAGCAGGTTCACTTGCTTGAGAAGAGC
TTCGAGGCTGAGAACAAGCTTGAGCCTGAGCGAAAGGCTGAGCTGGCGAAGAAGCTCGGATTGCAG
CCGAGGCAAGTCGCCATTTGGTTCCAAAACCGCCGAGCACGGTGGAAGACTAAGCAGTTAGAGAGG
GACTACGACAAGCTTAAGTCTTCTATGATTCTTCTCTCAACCTACGACTCTATTCGCCAGGAA
AACGACAAGCTCAAAGCCGAGCTCCTTTCCCTGAACGAGAAATTGCAACCCAAAGACGACGACGAC
CCATCGGCGCAAATAGGTGCAAAATCTCAGTTCATCGTCGCCGCTGTGACGCGGCTGAGCCGCCG
TGCCTGAAGCTGACGGTGAAGGTGGAGGACCGCCTGAGCACGGGGAGCAACGGCAGCGCAGTAATG
GACGGCGACGGACCTCAGCAGCTCCTCGACGACAGCGGCGACTCGTACTTCGAGAACGACGAGGAA
TACGACTGCGCCGCCGCAAGTTTGGCTGCTGCGAAGGAGGACGACGGCAGCGATGAGGGCGGGTGT
TACTTCACCGAGGCTCTCGCGGCGGAGGAGGAGGAGGCGCCGTTTGCTTGGTGTATTTGGTCT**TAA**

SEC ID N° 22 Secuencia de aminoácidos traducidos de Craterostigma plantagineum CPHB-5

MNSARIFFDPSSHGNMLQFLGNAGGDSSVFRGTRSSSVLNMEESSLKRQIFSGGGGDEFYDEEYYD
EQLLPEKKRRLTAEQVHLLKSF EAENKLEPERKAELAKKLGLQPRQVAIWFO NRRARWKTKQLER
DYDKLKSSYDSLSTYDSIRQENDKLKAELLSLNEKLQPKDDDDPSAEIGRNLSSSSPPVDAAEPP
CLKLTVKVEDRLSTGSNGSAVMDGDGPQQLLDDSGDSYFENDEEYDCAAASLAAAKEDDGSDEGGC
YFTEALAAEEEEAPFAWCIS

SEC ID N° 23 Secuencia de ADNc de Gossypium hirsutum Goshi_hox5 DR465649, CD486134

ATGGAGTCTGGCCGTCTTTTTTTTCAATCCCTCCACTACCCACCGCAACATGTTGCTTCTCGGGAAC
ACTGAACCCATCTTTTCGAGGGGCAAGAACAATGGTTAGCATGGAGGAAAACCCAAAGAAGCGACTG
TTCTTCAGCTCGCCGAGGATTTGTACGACGAAGAGTACTACGACGAGCAGTTGCCCGAGAAAAAG
CGTCGCCCTTACGTCGGAGCAGGTGTATCTGCTAGAGAAGAGCTTTGAGGCAGAGAACAAGCTGGAG

FIGURA 4 (continuación)

CCGGAGAGGAAGAGCCAGTTGGCCAAGAAGTTAGGACTGCAACCAAGGCAGGTGGCGGTATGGTTC
CAGAACCGCCGTGCAAGGTGGAAGACAAAGCAGCTTGAAAGGGACTATGACCTCCTCAAATCTTCC
TTTGATTCCCTTCAGTCCAATTATGACACTATTCTCAAAGAAAATGAGAAGCTCAAATCTGAGGTA
GCTTCCTTGACTGAAAACTACAAGCCAAAGATGTGGCAACAGAAGCAATAGCAGGTGAAAAGGAT
GAAGGGTTAGCAGCTGAGATGGCCTCCGCCCTCCAATTCAGTATGAAGGTGGAGGACCGTCTTAGT
AGCGGCAGTGTCGGAAGCGCGGTGGTGGATGAGGATGCCCCACAGCTGGTGGACAGCGGCAATTCC
TACTTTCCAAGCGATGAATACTCCAGAGGCATTGGCCCTTTTCGATGGGGTTTCAGTCGGAAGATGAG
GATGGCAGTGATAATTGCGGGAGTTACTTCTCCGATGTGTTTCGCAACCACAGAGCAGGGAGCATT
GGATTGTGGGCCTGGNTCTAA

SEC ID N° 24 Secuencia de aminoácidos traducidos de *Gossypium hirsutum* Goshi_hox5

MESGRLFFNPSTTHRNMLLLGNTEPIFRGARTMVSMEENPKKRLFFSSPEDLYDEEYYDEQLPEKK
RRLTSEQVYLLEKSFEAKENKLEPERKSQALAKKLGLQPRQVAVWFQNRARRWKTQLERDYDLLKSS
FDSLQSNYDITLKENELKSEVASLTKLQAKDVATEAIAAGEKDEGLAAEMASALQFSMKVEDRLS
SGSVGSAVVDEDAPQLVDSGNSYFPSDEYSRGIGPFDGVQSEDEDGSDNCGSYFSDVFATTEQAL
GLWAWX

SEC ID N° 25 Secuencia de ADNc de *Lycopersicon esculentum* Lyces_hox5 BT014213.1

ATGGGATCTGGGCATATATTTTTTCGACCCGTCGTCGTGTACGGCAACATGCTGTTCTTGGGAGC
GGAGATCCTGTTTTCCGAGGACCAAGATCGACGATGATGAAGATGGAGGACTCCTCGAAGAGGCGA
CCCTTCTTTAGCTCGCCGGAGGATCTATATGACGAGGAATACTACGACGAGCAGTCACCGGAGAAG
AAGCGCCGTCTCACTCCTGAGCAGGTGCACTTGTTGGAGAAGAGCTTTGAGACAGAAAACAAGCTG
GAGCCCGAGCGCAAAACCCAGCTGGCCTANAAGCTGGGGCTGCAGCCAGACAGGTGGCTGTATGG
TTCCAAAACCGCCGTGCCCGGTGGAAGACCAAGCAGCTCGAGAGGGATTATGATCAGCTCAAATCC
TCTTATGACTCCCTTCTCTCTGATTTTGAATCCGTTTCGCAAAGATAACGATAAGCTCAAATCTGAG
GTTGTTTCATTGATGGAAGTTACAGGGGAAAGTGGTTGGAGGAGCAGGGGGAAATGAAAAATCT
GACATCTTGGAGGTGGATGCTATGACGATCCTTCAAGTGAAGGTGAAGGTGGGGACCGGTTGAGC
AGTGGCAGTGGTGGGAGCGCGGTGGTAGATGAGCATAGTTTACAGCTGGTGGACAGTGGGGACTCA
TATTTTCACACTGATCATGAGGAGTATCCAGGGCCTGGAGGATGCAATGTTCTCCACCCATGGAT
GGTTTACAATCGGAGGAAGATGATGGTAGTGATGATCATGGCAGTTGCCATGGCTACTTCTCTAAC
GTCTTTGTGGCAGAAGAGCAGCACCATGAACAAGGAGAAGAGCCTATTGGATGGTCTGTTCT**TAA**

SEC ID N° 26 Secuencia de aminoácidos traducidos de *Lycopersicon esculentum* Lyces_hox5

MMSGHIFFDPSCHGNMLFLGSGDPVFRGPRSTMMKMEDSSKRRPFFSSPEDLYDEEYYDEQSPEK
KRRLTPEQVHLLLEKSFEAKENKLEPERKTQLAXKLGLQPRQVAVWFQNRARRWKTQLERDYDQLKS
SYDSLSDFDVSRKDNKDKSEVSLMEKLQGVVGGAGGNEKSDILEVDAMTILQVKVKAGDRLS
SGSGGSAVVDEHSSQLVDSGDSYFHTDHEEYPGPGGCNVPPMDGLQSEDDGSDDHGSGCHGYFSN
VFVAEEQHHEQGEEPIGWFS

SEC ID N° 27 Secuencia de ADNc de *Lycopersicon esculentum* VaHOX1 X94947

ATGGCTCCAGGGATTCTCTATGGTGGTTCTTCTAATTTTCGATGGCGTTTTTACTCAAAAACAGAGA
GACGTGTTTTCTTCATCTACTGCACCGAAAGGGCATCTTGGTTCCCTTTTTGCCCCTGCCTCTTCT
TCTTCTAATTTCTTGGGATCCAGTTCTATGGTGAATTTTCGCGGTGTTAATGGAGGGAAGAGATCA
TTCTTTGATTCTGTCGATCAGGATGACAATGAAGCTGATGAATTGGGGGAATATCTTCATCAAGCG
GAGAAGAAGAGGCGACTTACTGACAACCAAGTTCAGTTTCTTGAGAAGAGTTTTGGGGAAGAGAAC

FIGURA 4 (continuación)

AAACTTGAACCAGAAAGAAAAGTTTCAGCTTGCTAAAGAACTTGGTCTGCAGCCTCGCCAAATTGCA
 ATTTGGTTTTCAGAATCGTCGTGCGCATGGAAGACTAAGCAGCTCGAGAAAGATTATGATGAATTG
 AGGAATAGATACGATACTCTGAAATCAAATTACAATAATCTTCTCAAGGAAAAAGAAGATCTTCGA
 ACTGAAGTTTTCCGTCTCACCAGTAAGCTGTTTATCAAAGAGAAAGGAAATGGGCAATTGGATTTG
 CGCGATGAACACAAACACTCCAATGCATTGGCAAAAGAAACCGTGGTTGATCCAATGTCCAATGTA
 CCAGCTCTGGTTGTTAAGCACCAGCAGGAAGATTTAAGCTCTGCTAAGAGTGATGTTTTCGACTCA
 GAAAGCCACGTTACACCAGTAGAATGCATTCTCAGTCGTAGATCAGGATGATTCTGCTCGCGCA
 TTTGAAACTGATCAGTCGGATTCATCTCAGGATGATGATGAAAACCTTCAGCAAGAATATGCTTTCT
 ACTGCCAACCTACTTGGCAAAGACGCGGATGATGATTATCCC GCGACATCATCAAATTTGAGTTAC
 TTTGGATTTCCAGTTGAAGACCAAGGTTTTGGTTTCTGGACTTAT**TAA**

SEC ID N° 28 Secuencia de aminoácidos traducidos de *Lycopersicum esculentum* VaHOX1 X94947

MAPGILYGGSSNFDGVFTQKQRDVFSSTAPKGLGSLFAPASSSSNFLGSSSMVSFRGVNNGKRS
 FFDSFDQDDNEADELGEYLHQAEKKRRLTDNQVQFLEKSFGEENKLEPERKVQLAKELGLQPRQIA
 IWFQNRARRWKTQLEKDYDELRNRYDTLKSNNLLKEKEDLRTEVFRLTGKLFIKEKNGQLDL
 RDEHKHSNALAKETVVDPMSPALVVKHQEDLSSAKSDVFDSESPRYTSRMHSSVVDQDDSARA
 FETDQSDSSQDDDENFSKNMLSTANLLGKDADDDYPATSSNLSYFGFPVEDQGFQGFWTY

SEC ID N° 29 Secuencia de ADNc de *Medicago sativa* Medsa_hox16 CB892031, CA858059

ATGGCGGGTGGGAGAGTTTTTCAAATGGTCCTGCAAATATTTCAAATATAAATATGAATATTTTG
 CTTCAGAATCAACAACAACTCCTCGTGGAACCTCTTCTCAACAACCTCTTGATTCTCTTTTCTT
 TCTTCTTCTGCTTCTTTCTTTGGTTCAAGATCTATGGTGAGTTTTGAAGATGTTCAAGGAAGGAAA
 AGGCGCAACAGGTCTTTCTTTGGAGGATTTGATCTTGACGAAAACGGAGAGGATGAGATGGATGAG
 TACTTTTCATCAATCCGAGAAGAAACGGCGTCTCTCAGTGGATCAAGTTCAGTTTCTTGAGAAAAGC
 TTTGAGGAGGACAACAACTTGAACCAGAGAGGAAAACCAAGCTAGCTAAAGACCTTGGTTTGCAG
 CCACGGCAAGTTGCTATTTGGTTTCAAACCGTCTGCAAGGTGGAAGACTAAACAGCTTGAGAAG
 GATTATGATTCTCTTAATGATGGTTATGAGTCTCTTAAGACAGAGTATGACAACCTTCTCAAAGAG
 AAAGATAGGTTACAATCTGAGGTGGCAAGCCTAACTGAAAAGGTACTTGAAAAGAGAGAAACAAGAG
 GAAAATTCAAACAAGGTGAAAGTGAAACAAAGGAATTCTTGAAGGAACCAACAATTAATAAGCCT
 TTGGTTGATTGAGTTTCTGAGGGTGAAGGATCCAAATTGTCAATTGTTGAGGCTTCTAATAATAAT
 AATAATAATAACAACTTGAAGATATTAGTTCAGCAAGGAGTGACATATTGGATTGTGAAAGTCCA
 CGCTACACTGATGGAGTGTTAGAGACATGTGATTCTTCCTATGTATTTGAACCTGAATATCAATCG
 GACCTATCACAAGATGAAGAAGATCACAATTTATTGCCTCCTTACATCTTTACAAAACCTGAAGAT
 GTGAATTACTCCGACCCGCCACATAATTCAAACAAGTTATGGATTTCAGAGGAAGATCATCATCAA
 GCTCTTTGGCCTTGGTCTTAT**TAG**

SEC ID N° 30 Secuencia de aminoácidos traducidos de *Medicago sativa* Medsa_hox16

MAGGRVFSNGPANISNINMNILLQNQQQTPRGNSSQQPLDSLFLSSSASFSGSRSMVSFEDVQGRK
 RRNRSFFGGFDLDENGEDEMDEYFHQSEKKRRLSVDQVQFLEKSFEEFNKLEPERKTKLAKDLGLQ
 PRQVAIWQNRARRWKTQLEKDYDSLNDGYESLKTEYDNNLLKEKDRLQSEVASLTEKVLEREKQE
 GKFKQGESETKEFLKEPTINKPLVDSVSEGEKSLIVEASNNNNNNKLEDISSARSDILDCESP
 RYTDGVLETCDSSVFEPEYQSDLSQDEEDHNNLLPPYIFTKLEDVNYSDPPHNSTSYGFQEEDHHQ
 ALWPWSY

FIGURA 4 (continuación)

SEC ID N° 31 Secuencia de ADNc de *Aquilegia formosa* x *Aquilegia pubescens* Aqufo_hox5 DT58247

ATGGATTCAACAACAAGCCGCTTTTCTTTGATGGTTCCCTGCCATGGGAACATGTTGCTTTTAGGG
AGTGGAGATCCCGTTCTTCGAGGTTCAAGATCATTCAATTAATATGGAAGATTCTTTGAAAAGACGT
CCTTTTATAGTTCAACAGATGAACTAATTGAAGAGGAGTTTATGATGAACAGCTACCTGAAAAG
AAACGTCGTCTTACTTCTGAGCAGGTTTCTATTTGGAGAAGAGCTTTGAGACAGAGAACAAGCTG
GAACCAGATCGTAAGACCCAGCTTGCTAAGAAGCTTGGGTGCAACCGAGACAAGTTGCAGTTTGG
TTTCAGAATAGACGAGCTCGTTGGAAGACTAAGCAACTAGAGAGAGATTATGATCTTCTTAAAGCT
TCTTATGATTCCCTTCGTTCTGATTACGATGACATTGTTAAAGAGAATGAGAAGCTCAAATCTGAG
GTGGTTTCTTAACTGGGAAGTTGCAGGTCAAGGAGGGAGCTGGGATGGAGTTAAATCAGATATCT
GACCCACCACTCTCCACTGAAGAAAATGTTGATGTAACACGATGCAATTTAATGTTAAGGTTGAG
GATCGCTTGAGCTCTGGCAGTGGGGTAAGTGCTGTGGTTGATGAGGAATGTCGACAGCTTGTTGAC
AGTGTTGATTCTTATTTCCCTGGCGATGACTATGGTCAATGCATAGGCCAGTAGATGGAGTCCAG
TCAGAAGAAGATGACATTAGTGACGACAGCCGGAGCTATTTCTCAGATGTCTTCCAGCTGCACCA
GAGCAGAACCACCAGGAGAGTGAGACATTGGGTGTTGGTGGGACTGGGCT**TAA**

SEC ID N° 32 Secuencia de aminoácidos traducidos de *Aquilegia formosa* x *Aquilegia pubescens* Aqufo_hox5

MDSTTSRLFFDGSGHGNMLLLGSGDPVLRGSRFINMEDSLKRRPFYSSTDELIEEEFYDEQLPEK
KRRLTSEQVHLLKSFETENKLEPDRKTQLAKKLGLQPRQVAVWFQNRARRWKTKQLERDYLKKA
SYDSLRSYDDIVKENKLEKSEVVSITGKLQVKEGAGMELNQISDPPLSTEENVDTTMMQFNVKVE
DRLSSGSGVSAVVDEECRQLVDSVDSYFPGDDYGQCIGPVDGVQSEEDDISDDSRSYFSDVFPAP
EQNHQESETLGWWDWA

SEC ID N° 33 Promotor GOS2 de *Oryza sativa*

AATCCGAAAAGTTTCTGCACCGTTTTTACCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATACTAGAACTATGCAAGAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACCTCTTCTTGAATAAAAAAATCTTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTTTAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCCTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAGAAATT
TTGCTCGTGCGCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCACA
GAACAACCCACAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGGCAAGAAAACCAAGCATCCTCCTCCTCCCATCTATAA
ATTCTCCTCCCCCTTTTCCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTTCGATCCATATCTTCCGGTCGAGTTCTTGG
TCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCCTTCGGTTGTTCTTGGATTTAT
TGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCCTG
TTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTTCGGTTTGATTAGTAGTAT
GGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGTACGGAATCTTGCGATTTT
GTGAGTACCTTTTGTGTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAATAAAGTACG
GTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC

FIGURA 4 (continuación)

CTATTGAACAAAAATAATCCAACCTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTTA
AGCCTGTCCAAAATTTTCGCAGCTGGCTTGTTTAGATACAGTAGTCCCATCACGAAATTCATGGAA
ACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTGCTTTAGTCCCAGAATTTTT
TTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATGC
TTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTGTAGTCAGGAGAA
GAACTTATCCGATTTCTGATCTCCATTTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATTC
ATTTGGATTATTTTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAAC
TGTCTCAATTTTGTGTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGTA
GAAGTTTCTTTTTGGTTATTCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCGG
GATAGTTATACTGCTTGTTCTTATGATTCATTTCTTTGTGCAGTTCTTGGTGTAGCTTGCCACTT
TCACCAGCAAAGTTC

SEC ID N° 34 Cebador prm6000

GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACAATGGATCCCGGCCG

SEC ID N° 35 Cebador prm6001

GGGGACCACCTTTGTACAAGAAAGCTGGGTGATCAGCTCCAGAACCAGG

SEC ID N° 36 Secuencia de ADNc de *Oryza sativa* Orysa_hox4 AD145728

ATGAAAGCGACCCGGCGGTGCCGGCGGGCGGAGGCAGCCCATCGCTCGTCACGATGGCTAATTCT
AGTGATGATGGATATGGAGGGGTGGGATGGAGGCGGAGGGGGACGTGGAGGAGGAGATGATGGCG
TGCGGCGGGCGGGGAGAAGAAGCGGCGGCTGAGCGTGAGCAGGTTTCGCGCGCTGGAGCGGAGC
TTCGAGGTGGAGAACAAGCTTGAGCCTGAGCGGAAGGCGCGGCTGGCGCGCGACCTCGGCCTGCAG
CCGCGCCAGGTGCGCGTCTGGTTCCAGAACCAGCGCGCGCGGTGGAAGACCAAGCAGCTCGAGCGC
GACTACGCCGCGCTCCGCCATTCTACGACTCCCTGCGCCTCGATCACGACGCGCTCCGCCGCGAC
AAGGACGCCCTCCTCGCCGAGATCAAGGAGCTGAAGGCGAAGCTCGGGGACGAGGAGGCGGCGGCG
AGCTTCACGTCGGTGAAGGAGGAGCCGGCGGCCTCCGACGGGCCACCGGCGGCGGGATTTGGGTCTG
TCCGACAGCGACTCAAGCGCGGTGCTGAACGACGTGGACGCGGCCGCGCGCCGCGCGGCGGCGGCG
GACGCGCTGGCTCCGGAGGCGTGCACGTTTCTCGGTGCGCCGCCGCGCGGGCGCGGGCGCGGGC
GCAGCGGCGGCGGCGAGCCACGAGGAGGTGTTCTTCCACGGCAATTTCTCAAGGTGGAGGAGGAC
GAGACGGGGTTCTCGACGACGACGAGCCGTGCGGCGGGTTCTTCGCCGACGATCAGCCCCCGCGC
CTCTCGTCTGTGGTGGGCCGAACCCACGGAGCACTGGAAC**TGA**

SEC ID N° 37 Secuencia de aminoácidos traducidos de *Oryza sativa*

MKRPGGAGGGGSGPSLVTMANSSDDGYGGVGM EAEGDVEEEMMACGGGGGEKKRRLSVEQVRALERS
FEVENKLEPERKARLARDLGLQPRQVAVWFQNRRLRWTKQLERDYAALRHSYDSLRLDHDALRRD
KDALLAEIKELKAKLGDEEAAASFTSVKEEPAASDGPPAAGFGSSSDSDSSAVLNDVDAAGAAPAAT
DALAPEACTFLGAPPAAGAGAGAAAAASHEEVFFHGNFLKVEEDETGFLLDDDEPCGGFFADDQPPP
LSSWWAEPTEHWN

SEC ID N° 38 Secuencia de ADNc de *Oryza sativa* Orysa_hox6 AK103160

ATGATGGGGAGGAGGACAGCGAGTGGATGATGATGGACGTTGGAGGGAAGGGCGGGAAGGGCGGC
GGCGGCGGCGGCGGCGGCGGACAGGAAGAAGCGGTTTCAGCGAGGAGCAGATCAAGTCGCTGGAGTCC
ATGTTTCGCGACGCAGACCAAGCTGGAGCCGAGGCAGAAGCTGCAGCTCGCCAGGGAGCTCGGCCTG
CAGCCTCGCCAGGTGCGCATCTGGTTCCAGAACAAGCGCGCGCGGTGGAAGTCCAAGCAGCTCGAG
CGCGAGTACTCCGCCCTCCGCGACGACTACGACGCCCTCCTCTGCAGCTACGAGTCCCTCAAGAAG
GAGAAGCTCGCCCTCATCAAGCAGCTGGAGAAGCTGGCGGAGATGCTGCAGGAGCCACGGGGGAAG

FIGURA 4 (continuación)

TACGGCGATAATGCCGGGGACGACGCGCGGTTCGGGCGGCGTCCCGGCATGAAGAAGGAGGAGTTCTC
GTCGGCGCGGGCGGCGCCGCCACGCTCTACTCGTCGGCCGAGGGTGGCGGGACGACGACGACGACG
ACGGCCAAGTTGATGCCCCACTTCGGCAGCGACGACGTCGACGCGGGGCTCTTCCTCCGGCCGTCG
TCGCAGCATCATCCGCCGCCGCCGACGCCGGTGCCGGCTTCACGTCCTCCGAGCCGGCCGCCGAC
CACCAGTCCTTCAACTTCCACTCGAGCTGGCCGTTCGTCCACGGAGCAGACCTGCAGCAGCACGCCA
TGGTGGGAATTCGAGAGCGAGTGA

SEC ID N° 39 Secuencia de aminoácidos traducidos de *Oryza sativa* Orysa_hox6

MDGEEDSEWMMMDVGGKGGKGGGGGAADRKKRFSEEQIKSLESMFATQTKLEPRQKLQLARELGL
QPRQVAIWQFNKRARWKSQLEREYSALRDDYDALLCSESLKKEKLALIKQLEKLAEMLQEPGRK
YGDNAGDDARSGGVAGMKKEEFVGAGGAATLYSSAEGGGTTTTTTAKLMPHFGSDDVDAGLFLRPS
SQHHPPPPHAGAGFTSSEPAADHQSFNFHSSWPSSTEQTCSSTPWWFEFESE

SEC ID N° 40 Secuencia de ADNc de *Populus tremuloides* Poptr_Hox16_1

ATGGCGGGTGGTACCGGTGGTTCTAATTCCAATTTGTCTGTTTTGCTTCAAAGCCAAAGAGGCCCT
TGTGCTGCTTACAACCTCTTGAATCTTTTTCTTTCTGGCTCTTCTCCTTCTTTTCTTGGTTCA
AGATCCATGATGAGTTTTGAAGATGTTTCATCAAGCAAACGGATCAACCAGGCCTTTTTTCCGCTCG
TTTGATCACGAAGACAATGGAGACGATGATCTGGATGAATATTTTCATCAACCTGAAAAGAAGAGG
AGACTTACTGTTGATCAAGTTCAGTTTCTTGAAAAGAGTTTTGAGCTTGAGAACAAGCTTGAACCT
GAAAGGAAAATCCAGCTTGCAAAGGATCTTGGCCTTCAGCCGCGTCAGGTTGCTATATGGTTTCAA
AACC GCCGAGCAAGATGGAAGACTAAACAGCTGGAAAAGGATTATGACGTTTTGCAATCTAGCTAC
AATAGCCTTAAGGCTGACTATGACAACCTCCTCAAGGAGAAGGAGAACTAAAAGCTGAGGTTAAT
CTTCTCACCGACAAGTTGCTCCTCAAAGAGAAAAGAGAAGGGAATCTCAGAATTGTCTGATAAAGAT
GCATTATCGCAAGAGCCACCTAAAAGGGCTATAGCTGATTACGCTTCCGAGGGTGAAGTGTGCGAAA
ATCTCAACAGTGGCCTGTAAGCAGGAAGATATCAGCTCAGCCAAAAGCGACATATTTGATTACAGAC
AGCCACATTACGCTGATGGGGTGCATTCCTCACTCTTAGAGGCAGGAGATTCTTCATATGTTTTTC
GAACCCGATCAATCAGATTTGTGACAAGATGAAGAAGATAACTTTAGCAAGAGCTTATTGCCTCCA
TACGTCTTTCCGAAGCTTGAAGATGACGATTACTCTGACCCGCCTGCAAGTTTTGAAGATCATGCC
TTTTGGTCTGGTCATACTAA

SEC ID N° 41 Secuencia de aminoácidos traducidos de *Populus tremuloides* Poptr_Hox16_1

MAGGTGGSNSNSLVLLQSQRGPCAASQPLESFFLSGSSPSFLGSRSMMSFEDVHQANGSTRPFFRS
FDHEDNGDDDLDEYFHQPEKKRRLTVQVQFLEKSFLENKLEPERKIQLAKDLGLQPRQVAIWQ
NRRARWTKQLEKDYDVLQSSYNLSLKADYDNLKEKEKLKAEVNLLTDKLLLKEKEKGISELSDKD
ALSQEPPKRAIADSASEGEVSKIISTVACKQEDISSAKSDIFDSDSPHYADGVHSSLLEAGDSSYVF
EPDQSDLSQDEEDNFSKSLPPYVFPKLEDDDYSDPPASFEDHAFWSWSY

SEC ID N° 42 Secuencia de ADNc de *Populus tremuloides* Poptr_Hox16_2

ATGGCGGCTTGTGGTGGTGGTGGTGGTGGTCTAATCCCAATTTGTCTGTTTTAGTTCAAAGCCAA
AGAGGCCCTTGTGCTGCTTCTCAACCTCTTGAAGCTTTTTCTTTCTGGCTCTTCTCCTTCTTTT
CTTGGTTCAAGATCCATGATGAGTTTTGCAGATGTTACCAAGCAAATGGATCAACTAGACCGTTT
TTCCGCCCATATGATCACGAAGACAACGGCGACGATGATTTGGATGAATATTTTCATCAACCTGAA
AAGAAGAGGAGACTTACTGTTGATCAAGTTCAGTTTCTTGAAAGAAGTTTTGAGGTTGAGAACAAG
CTTGAACCCGAAAGGAAAATCCAGCTGGCGAAGGATCTTGGCTTGCAGCCTCGGCAGGTTGCCATA
TGGTTTCAAACCGCCGGGCAAGATGGAAGACGAAACAGCTTGAAAAAGATTATGAGGTTCTGCAA

FIGURA 4 (continuación)

TCTAGCTACAATGGCCTTAAGGCTGACTACGACAACCTCTTCAAGGAGAAGGAGAACTAAAAGCT
GAGGTTAATCTTCTCACCAACGAGTTGCTCCTTAAAGAGAAAGAGAAAGGAAGCTCAGAATTGTCT
GATAAAGATGCATTATCTCAAGAGCCACCCAAAAAGGCAATAGCCGATTCAGCTTCAGAGGGTGAA
GTGTGCGAAAACCTCAACCGTGGCCTGCCAGCAGGAAGATATTAGCTCAGCCAAAAGTGATATGTTT
GATTCAGACAGCCACATTTTGC GGATGGGGTACATTCCTCACTCTTAGAGGCAGGTGATTCTTCA
CATGTCTTCGAGCCCGACCAATCGGATTTATCACAAGATGAAGAAGATAACTTGAGCAAGAGTCTT
TTGCCTCCGTACGTCTTTCCAAAGCTTGAAGATGGTGATTACTCTGACCCGCCAGCAAGTTTTGAA
GATCATGCCTTTTGGTGCTGGTCATACTAA

SEC ID N° 43 Secuencia de aminoácidos traducidos de *Populus tremuloides* Poptr_Hox16_2

MAACGGGGGGSNPNLSVLVQSQRGPCAASQPLEAFFLSGSSPSFLGSRSMMSFADVHQANGSTRPF
FRPYDHEDNGDDDLDEYFHQPEKKRRLTVQVQFLERSFEVENKLEPERKIQIAKDLGLQPRQVAI
WFQNRARRWKTQLEKDYEVLQSSYNGLKADYDNLFKEKEKLAEVNLLTNEELLKEKEKGSSELS
DKDALSQEPKKAIAADSASEGEVSKTSTVACQQEDISSAKSDMFDSDSPHFADGVHSSLLEAGDSS
HVFEPDQSDLSQDEEDNLSKSLPPYVFPKLEDGDYSDPPASFEDHAFWCWSY

SEC ID N° 44 Secuencia de ADNc de *Populus tremuloides* Poptr_Hox16_3

ATGGCGGGTGATAAAGACTGTGGCAGTTCTAAAATGACCATTTTTCTTCGAAACGGCAGGCTCCCT
CCTTGTGAATCTCTCTGTATTCTCACCTCTTTTAGCACTCTTCATGGTGCAAAATCTATGGTTAAT
TTTAGGAATGATGGAGGAGACACTGTAGACATGTCTTTTTTCCAACCACATGTCAAAGAAGAAAGT
AGCGATGAGGATTATGATGCGCACCTTAAGCCATCTGAAAAGAAAAGGCGGCTTACAGCTGCTCAA
GTCCAGTTTCTTGAGAAGAGCTTTGAGGCGGAGAATAAGCTTGAACCAGAGAGGAAGATGCAGCTT
GCTAAAGAACTCGGCTTGCAGCCTCGCCAGGTTGCAATATGGTTTCAAACCGTAGAGCTCGGTTC
AAGAACAAGCAGCTGGAAAGGACTACGACTCCTTGAGAATCAGCTTTGACAACTCAAGGCTGAT
TATGACAACTCCTCCTCGAGAAGCAGAATTTGAAAAACGAGCTTCTTTCAGTGAAGAAAAATTG
CTTAGCAGAGAGGAAAGTATGGAAAGTTTCAAGACCATTTGATGTCATCCATTCACCGGATGCAGAA
CTTGAGCCTATTCTTGATACAGTGTCTGAAAATGTTTCCGCCATTGTGCCAATGGTGACACCCAAA
CAAGAAGAAAGTTTCAAGCTAAAAATGATGTTTTCACTCAGACAGCCACGTTTCAATTTTTGGAGCCC
CGTGATTGTTATCGTGTTTTTCGAGTCAGACCAACCAGATTTTTTCCCAAGTTGAAGAAGATAATCTC
ACCAGGAGCTTTCTACCCCTCCGTACTTTCCAAACTCTACCGAGAGCCACCTGCAAGTTTACGT
AATTTTGAATTCTCAGCGGAAGATCAGCCCTTTTGGTCCTGGATTTACTGA

SEC ID N° 45 Secuencia de aminoácidos traducidos de *Populus tremuloides* Poptr_Hox16_3

MAGDKDCGSSKMTIFLRNGRLPPCESLCILTSFSTLHGAKSMVNFNRNDGGDTVDMSFFQPHVKEES
SDEDYDAHLKPSEKKRRLTAAQVQFLEKSFEAKNKLEPERKMQIAKELGLQPRQVAIWFQNRARRF
KNKQLERDYDSLRI SFDKLADYDKLLLEKQNLKNELL SLKEKLLSREESMESSEPFQDVIHSPDAE
LEPIPDTVSENVSAIVPMVTPKQEESSAKNDVFNSDSPRSFLEPRDCYRVFESDQPDFSQVEEDNL
TRSFLLPPYFPKLYREPPASSRNFESAEQPFWSWIY

SEC ID N° 46 Secuencia de ADNc de *Medicago truncatula* Medr_HOX16_1

ATGGCAGGTGGCAAGCTTTTTGGTGTTCTAATATGTCACTTTTGCTTCAAAATGAAAGACTCCCT
TGTAATCTTGAAGTCTTGAATCTCTTTGGGTTTACACCCCTGCTTCTTTTCAAGGTTCAAATTCA
GTGGTTAATTTTGAAGAATGGTGGTGGTAGCAACAGAGTGGTAACAGATAGACCCTTCTTTCAACAA
CTTGAGAAAGAAGAGAATTGTGGTGATGAAGATTATGAAGCATGCTACCATCAACAAGGAAAGAAA
AGGAGGCTTTCAAGTGAACAAGTTCAATTTCTTGAAAAGAGTTTTGAGGTAGAAAACAAGCTTGAA
CCTGATAGGAAAGTTCAACTTGCAAAAGAGCTTGGTTTGCAA

FIGURA 4 (continuación)

CCAAGACAAGTTGCTATATGGTTTCAAAACAGAAGGGCAAGGTTCAAACTAAACAGCTTGAAAA
GATTATGGCACATTGAAAGCTAGCTTTGATAGTCTCAAAGATGATTATGATAATCTTCTTCAAGAG
AATGACAAGTTAAAAGAAGAGGTGAATTCTCTCAAGAACAAATTGATCCCAAGAGATAAAGAAAA
GTGAATTCAGAAGACAAATCATCACCAGAAGCAATCAATTCACCTCATAACAACATAGATCCAATG
GATATAATTTCAATTACAAATTCAGAAAATGGGTCCAAAATGTCACCTCCCTAATATGGTACTAAAA
TGTAAGCAAGAAGATGCCAATTCAGCTAAAAGTGATGTGCTTGATTCTGATAGCCACATTGCAAT
GATGGGAACAATCTTTCTTCTTTCATAGAGCCTACAGATTTCAGATTCTCACAAGATGAAGAGGAT
AATGATAACTTGAGTCATAATCTTTTGACTCTTCCTTGCTTACCAAAGTTGAAGATGTTTGCTAT
GATGACCCACATGAAAATTCTTGTAATTTTGGGTTCCCTGTTGAAGATCAAACCTTTTGTCTGG
CCTTATTGA

SEC ID N° 47 Secuencia de aminoácidos traducidos de *Medicago truncatula* Medr_HOX16_1

MAGGKLFGGSNMSLLLQNERLPCTSEVLESLWVHTPASFQGSNSVNFENGSGSNRVVTDPRPFFQQ
LEKEENCDEDEYACVHQQGGKRRRLSSEQVQFLEKSFEEVENKLEPDRKVQLAKELGLQPRQVAIWF
QNRARFRTKQLEKDYGTLLKASFDLKDNDLLQENDKLKEEVNSLKNKLI PRDKEKVNSEDKSS
PEAINSPHNNIDPMDIISITNSENGSKMSLPNMVLKCKQEDANSKSDVLDSDSPHCNDGNNLSSF
IEPTDSDFSQDEEDNDNLSHNLLTLPCLPKVEDVCYDDPHENSCNFGFPVEDQTFCFWPY

SEC ID N° 48 Secuencia de ADNc de *Phaseolus vulgaris* Phavu_HOX16

ATGGCGGGTGGCAAGCTTCATCCTGGTTCAAACATGTCACCTTCTCCTCCAAAACGACAGGCTCCCT
TGCTCCTCTGAAGTCCTTGAGTCTCTTTGGGCTCACACCTCTAACGCTGCTTCCTTCCAAGGTTCA
AAATCTATGGTTGATTTTGAGAATGTTAGTGGGGGCAGGGTGACGGATAGGCCCTTTTTTCAAGCG
TTGGAGAAGGAAGATAACTGTGATGATGATTATGAGGGTTGCTTCCATCAACCGGGTAAGAAAAGG
AGGCTCACAAGCGAACAAGTTCAGTTCCTTGAAAGGAACCTTTGAGGTGCGAGAACAAGCTTGAACCT
GAAAGGAAGGTCCAACCTTGCAAAGGAGCTTGGCTTGCAAGCAAGGCAAGTGGCTATATGGTTCCAA
AACCGAAGGGCAAGGTTCAAGACCAAGCAGCTAGAAAAAGATTATGGCACATTGAAAGCTAGCTAT
GACAGACTCAAAGGTGACTATGAAAGTCTTCTTCAAGAGAATGACAAGTTAAAAGCAGAGGTGAAT
TCTCTGGAGAGCAAATTGATTCTTAGAGATAAAGAGAAGGAGAATTCGGACGACAAGTCATCTCCT
GATGCTGTCAATTCACCCACAAAGAGCCTATGGATTTAATTTCAAATTCACATCTGAAAATGGG
ACCAAAGTGTCACTCCCTATTATGGTAACATGCAAGCAAGAAGATGCCAATTCAGCCAAAAGTGAT
GTGCTTGATTCGGACAGCCACATTGCACTGATGGGAACCATCCCTCTTCATTCTGAGGCTGCT
GATTCCTCCCATGCTTTTGAACCAGACCACTCCGACTTCTCCCAAGATGAAGAGGATAATCTTAGT
GAAAGCCTTTTGACCCTCCCTTGCTTACCAAAGGTTGAAGAAGCCTGCTATGATGACCCTCCTGAA
AACCTTGTAATTTTGGCTTCCATGTGCGAGGATCAAACCTTCTGTTTCTGGCCCTATTGA

SEC ID N° 49 Secuencia de aminoácidos traducidos de *Phaseolus vulgaris* Phavu_HOX16

MAGGKLHPGSNMSLLLQNDRLPCSSEVLESLWAHTSNAASFQGSKSMVDFENVSGGRVTDPRPFFQA
LEKEDNCDDDYEGCFHQPGKKRRLTSEQVQFLERNFEVENKLEPERKVQLAKELGLQPRQVAIWFQ
NRRARFRTKQLEKDYGTLLKASYDRLKGDYESLLQENDKLKAEVNSLESKLILRDKEKENSDDKSSP
DAVNSPHKEPMDLISNSTSENGTKVSLPIMVTCKQEDANSKSDVLDSDSPHCTDGNHPSSFVEPA
DSSHAPEPDHSDFSQDEEDNLSESLTLPCLPKVEEACYDDPPENPCNFGFHVEDQTFCFWPY

SEC ID N° 50 Secuencia de ADNc de *Lotus corniculatus* Lotco_Hox16

ATGGCGGGAGGGAGGGTCTTTAGCGGCGGTCTGCTGCTCCTGCAAATGTTTCCGATACCAGTCTT
TTGCTTCAGAATCAACCTCCTGATTCTTCTCTCTTCTCTACCTCTGCTTCTTTTCTCGGTTCA
AGATCCATGGTGAGCTTCGCAGATAATAAATTAGGGCAAACGCGGTCGTTCTTCTCCGCGTTTGAC

FIGURA 4 (continuación)

CTCGATGAGAACGGCGATGAGGTCATGGACGAGTACTTTCACCAATCGGAGAAGAAGCGCCGTCTC
TCTGTTGACCAAGTTTCAGTTTCTGGAGAAGAGCTTCGAGGTGGATAACAAGCTCGAACCTGACAGG
AAAACCAAGATTGCCAAGGACCTTGGTTTGCAGCCACGCCAAGTCGCAATCTGGTTCCAGAACCGC
CGTGCACGGTGGAAAGACGAAACAGCTTGAGAAGGATTATGATTCTCTGCATAGTAGCTTTGAGAGT
CTCAAATCCAACCTATGATAATCTTCTCAAGGAGAAAGACATGTTAAAAGCTGAGGTGGCAAGTCTC
ACTGAGAAGGTGCTTGCAAGAGAGAATTTGAAACAAGTTGAAAGTGAAACAAAGGGATTGGTTGAA
CCACCCCAAAGGCCTTTACTTGATTTCAGTTTCAGAGGGTGAAGAATCTAAAGTCTCTGTTGGGGCT
TGTAACCATGAGGATATCAGTTTCAGCCAGGAGTGAGAGTTTGGATTCTGATAGCCACGTTACAGG
GATGGATATGGAGTTAACTCAGCAGTGCTAGAGACATGTGATTCTTCTTATGTGGTTGAACCTGAT
CAATCGGATATGTCACAGGATGAGGAAGACAACCTGACCAAGACCCTGTTGCCTCCATACATGTTT
TCCAAACTTGAGATATGGATTACTCCGACCCGCCTGAAAGTTCATGTAATTTCCGGATTTCGGAG
GAAGATCATGCCCTTTGGTCATGGTCTTACTGA

SEC ID N° 51 Secuencia de aminoácidos traducidos de *Lotus corniculatus* Lotco_Hox16

MAGGRVFSGGSAAPANVSDTSLLLQNQPPDSSLFLSTSASFLGSRSMVSFADNKLQTRSFFSAFD
LDENGDEVMDEYFHQSEKKRRLSVDQVQFLEKSFEVDNKLPEPRKTKIAKDLGLQPRQVAIWQNR
RARWKTQLEKDYDSLHSSFESLKSNDNLLKEKDMLKAEVASLTEKVLARENKQVESETKGLVE
PPQRPLLDVSVEGEESKVSVGACKHEDISSARSESLDSDSPRYRDGYGVNSAVLETCDSSYVVEPD
QSDMSQDEEDNLTKTLLPPYMFSLGDMDYSDPPESSCNFGFPEEDHALWSWSY

SEC ID N° 52 Promotor GOS2 de *Oryza sativa*

AATCCGAAAAGTTTCTGCACCGTTTTACCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAACTCATCCACCTACT
TTAGTGGAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTTCGTTTTCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTTACATCTCTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACTCTTCTTGAATAAAAAAATCTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTTAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTTCGTCATATCGCACATCATTAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCCTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTACATACAAAAAAGAAATT
TTGCTCGTGCGGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCCACA
GAACAACCCACAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGGCAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT
CGATCTCTTCCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCGATT
TTGTGAGTACCTTTTGTGTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAAATAAAGTAC
GGTTGTTTGGTCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACCTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTTCGAGCTGGCTTGTTTAGATACAGTAGTCCCATCACGAAATTCATGGA

FIGURA 4 (continuación)

AACAGTTATAATCCTCAGGAACAGGGGATTCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTGTAGTCAGGAGA
AGAACTTATCCGATTTCTGATCTCCATTTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCTCAATTTTGTTCCTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTTCTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTTCTTATGATTCATTTCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

FIGURA 4 (continuación)