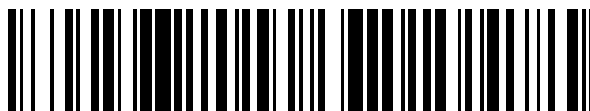


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 555 659**

51 Int. Cl.:

**C07K 14/415** (2006.01)

**C12N 9/12** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.11.2006** **E 06819308 (5)**

97 Fecha y número de publicación de la concesión europea: **30.09.2015** **EP 1948807**

54 Título: **Plantas que tienen características de crecimiento mejoradas y un procedimiento de producción de las mismas**

30 Prioridad:

**07.11.2005 EP 05110413**

**07.11.2005 EP 05110429**

**14.11.2005 US 736194 P**

**17.11.2005 EP 05110900**

**23.11.2005 US 739686 P**

**24.11.2005 EP 05111260**

**05.12.2005 US 742287 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**07.01.2016**

73 Titular/es:

**CROPDESIGN N.V. (100.0%)**

**Technologiepark 21C**

**9052 Zwijnaarde, DE**

72 Inventor/es:

**SANZ MOLINERO, ANA ISABEL;**

**REUZEAU, CHRISTOPHE;**

**HATZFELD, YVES;**

**BOUDOLF, VERONIQUE;**

**DE VEYLDER, LIEVEN;**

**INZE, DIRK y**

**MIRONOV, VLADIMIR**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 555 659 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 características de crecimiento mejoradas y un procedimiento de producción de las mismas

El aumento constante de la población mundial y la cantidad menguante de tierra arable para la agricultura potencia la investigación hacia el aumento de la eficacia de la agricultura. Los medios convencionales para las mejoras de cultivos y hortícolas utilizan técnicas de cultivo selectivo para identificar plantas que tengan características deseables. Sin embargo, dichas técnicas de cultivo selectivo tienen varios inconvenientes, a saber, estas técnicas son típicamente muy laboriosas y dan como resultado plantas que a menudo contienen componentes genéticos heterogéneos que no siempre dan como resultado que el rasgo deseable se transmita desde las plantas progenitoras. Los avances en biología molecular han permitido a la humanidad modificar el germoplasma de animales y plantas. La modificación por ingeniería genética de plantas implica el aislamiento y la manipulación del material genético (típicamente en forma de ADN o ARN) y la posterior introducción de ese material genético en una planta. Dicha tecnología tiene la capacidad de proporcionar cultivos o plantas que tengan diversos rasgos económicos, agronómicos y hortícolas mejorados.

Un rasgo de particular interés económico es el rendimiento. El rendimiento se define normalmente como el producto medible de valor económico de un cultivo. Este puede definirse en términos de cantidad y/o de calidad. El rendimiento depende directamente de varios factores, por ejemplo, el número y tamaño de los órganos, la arquitectura de la planta (por ejemplo, el número de ramas), la producción de semillas, la senescencia de la hoja y más. El desarrollo de la raíz, la captación de nutrientes y la tolerancia al estrés y el vigor temprano también pueden ser factores importantes para determinar el rendimiento. La optimización de uno de los factores anteriormente mencionados puede, por lo tanto, contribuir a aumentar el rendimiento del cultivo.

El rendimiento de la semilla es un rasgo particularmente importante ya que las semillas de muchas plantas son importantes para la nutrición humana y animal. Los cultivos tales como maíz, arroz, trigo, canola y soja suponen más de la mitad de la ingesta calórica humana, ya sea mediante consumo directo de las semillas en sí o mediante el consumo de productos cárnicos criados 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 brotes 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 semilla implica muchos genes, y requiere de la transferencia de metabolitos desde las raíces, hojas y tallos a la semilla en crecimiento. El endospermo, en particular, asimila los precursores metabólicos de los hidratos de carbono, aceites y proteínas y los sintetiza en macromoléculas de almacenamiento para rellenar el grano.

Otro rasgo de particular interés económico es el de tolerancia al estrés abiótico mejorada. El estrés abiótico es una causa principal de pérdida de cultivos en todo el mundo, reduciendo los rendimientos medios de la mayoría de plantas de cultivo en más de un 50 % (Wang y col., Planta (2003) 218: 1-14). El estrés abiótico puede estar causado por sequía, salinidad, extremos de temperatura, toxicidad química y estrés oxidativo. La capacidad para mejorar la tolerancia de las plantas al estrés abiótico podría ser una ventaja económica para los granjeros en todo el mundo y podría permitir cultivar cultivos durante condiciones adversas y en territorios donde el cultivo de cultivos no sería posible de otro modo.

La capacidad para aumentar el rendimiento de las plantas tendría muchas aplicaciones en áreas tales como la agricultura, incluyendo la producción de plantas ornamentales, arboricultura, horticultura y silvicultura. El aumento del rendimiento también podría ser útil en la producción de algas para su uso en biorreactores (para la producción biotecnológica de sustancias tales como agentes farmacéuticos, anticuerpos o vacunas, y para la bioconversión de residuos orgánicos) y otras áreas similares.

### **Antecedentes**

#### **Proteínas de homeodominio de cremallera de leucina (HDZip)**

Las proteínas de homeodominio de cremallera de leucina (HDZip por sus siglas en inglés, homeodomain leucine zipper) constituyen una familia de factores de transcripción caracterizada por la presencia de un dominio de unión a ADN (HD) y un motivo de cremallera de leucina (Zip) adyacente. El homeodominio consiste normalmente en 60 restos de aminoácidos conservados que forman un hélice1-bucle-hélice2-giro-hélice3 que se une a ADN. Este sitio de unión a ADN es normalmente pseudopalindrómico. La cremallera de leucina, adyacente al extremo C-terminal del homeodominio, consiste en varias repeticiones de héptada (al menos cuatro) en las que normalmente aparece una leucina (en ocasiones una valina o una isoleucina) cada siete aminoácidos. La cremallera de leucina es importante para la dimerización de proteínas. Esta dimerización es un requisito previo para la unión a ADN (Sessa y col., (1993) EMBO J 12(9): 3507-3517), y puede aparecer 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 de al menos 89 miembros en *Arabidopsis thaliana*. La cremallera de leucina también se encuentra por sí sola en eucariotas distintos de plantas. Sin embargo, la presencia tanto de un homeodominio como de una cremallera de leucina, es específica de plantas (encontrada en al menos 47 de las 89 proteínas en *Arabidopsis*), y se ha encontrado en musgos además de en plantas vasculares (Sakakibara y col., (2001) Mol Biol Evol 18(4): 491-502). Por lo tanto, la cremallera de leucina se

encuentra en el extremo C-terminal del homeodominio, solapándose estas dos características por tres aminoácidos.

Los genes de HDZip de *Arabidopsis* se han clasificado en cuatro clases diferentes, HDZip I a IV, basándose en criterios de similitud de secuencia (Sessa y col., (1994) en Plant Molec Biol, págs. 412-426). Al igual que las proteínas HDZip de las otras tres clases, las proteínas HDZip de clase I son bastante divergentes en cuanto a su estructura de amino primaria fuera del homeodominio y de la cremallera de leucina. Dentro tanto del homeodominio como 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 invariantes Leu<sub>16</sub>Trp<sub>48</sub>Phe<sub>49</sub>Asn<sub>51</sub>Arg<sub>53</sub>, la posición 46 está ocupada por una Ala (A) y la posición 56 por una Try (W) (u ocasionalmente por una Phe (F)) (Sessa y col., (1997) J Mol Biol 274(3):303-309; véase la figura 1), citado como un homeodominio de clase I, y

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

En referencia a sus propiedades de unión a ADN, las proteínas HDZip de clase I se unen preferentemente a semisitos de 5 pb que se solapan en una posición central, CAA(A/T)ATTG (Sessa y col., (1993) EMBO J 12(9): 3507-3517).

Se ha demostrado que diferentes proteínas HDZip activan o bien reprimen la transcripción. En *Arabidopsis*, se demostró que ATHB1, -5, -6 y -16 de HDZip de clase I actúan como activadores transcripcionales en ensayos de expresión transitoria en hojas de *Arabidopsis* usando un gen indicador (luciferasa; Henriksson y col., (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 y col., (2000) Mol Gen Genet 263:12-21). Por el contrario, dos proteínas HDZip de clase II de arroz, Oshox1 y Oshox3, actuaron como represores transcripcionales en los mismos experimentos (Meijer y col., (1997) Plant J 11: 263-276; Meijer y col., (2000) anteriormente citado).

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 y col., (2003) Plant Cell Environ 26: 1127-1136). Las *Arabidopsis* transgénicas que sobreexpresan ATHB1, -3, -13, -20 y -23 de HDZip de clase I sugieren que estos genes están implicados en la regulación del desarrollo del cotiledón y de la hoja (Aoyama y col., (1995) Plant Cell 7: 1773-1785; Hanson (2000) en Comprehensive summaries of Uppsala Dissertations from the Faculty of Science and Technology, Uppsala). Los genes ATHB3, -13, -20 y -23 son similares y forman una subclase distinta dentro de HDZip de clase I. Ya que estos genes provocan alteraciones similares en la forma del cotiledón cuando se expresan de manera constitutiva, se citan como los genes de HDZip de cotiledón puntiagudo (POC). Hanson llega a la conclusión de que las proteínas HDZip de clase I que están estrechamente relacionadas filogenéticamente también están relacionadas funcionalmente, en la mayoría de los casos.

El documento WO 2004/099365 desvela una secuencia de HDZip de girasol denominada Hahb-4, construcciones, vectores y plantas transgénicas que comprenden esta secuencia o variantes de la misma.

El documento WO 03/007699 desvela factores de transcripción potencialmente relacionados con el estrés en el arroz que se identificaron basándose en su identidad de secuencia con las secuencias de HDZip de *Arabidopsis*.

Deng Xin y col., (Plant Mol Biol vol. 49, n.º 6, agosto de 2002, páginas 601-610) re refiere al aislamiento de genes de HDZip de la planta de la resurrección, *Craterostigma*.

El n.º de referencia de la base de datos Uniprot Q6ZA74 desvela la secuencia de la proteína de la caja homeostática-cremallera de leucina HOX5 del arroz.

Se ha descubierto que la modulación de la expresión en una planta de un ácido nucleico que codifica un polipéptido de Hox5 de HDZip de clase I que tiene una secuencia tal como se muestra en SEC ID N.º: 2 proporciona plantas que tienen características de crecimiento mejoradas en relación a las plantas de tipo silvestre correspondientes u otras plantas de control.

La presente invención se refiere, por lo tanto, a un procedimiento para aumentar el rendimiento de plantas en plantas que crecen en condiciones sin estrés en relación a las plantas de tipo silvestre correspondientes, que comprende introducir y sobreexpresar en una planta un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de SEC ID N.º: 2, y seleccionar plantas que tienen rendimiento aumentado, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, número total de semillas aumentado, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, aumento del índice de cosecha (IC), aumento del peso de mil granos (PMG), aumento de la biomasa de la raíz, cada uno en relación a las plantas de tipo silvestre correspondientes.

La presente invención se refiere además a un procedimiento para aumentar el rendimiento de plantas cultivadas en condiciones de estrés abiótico, comprendiendo dicho procedimiento introducir y sobreexpresar en una planta un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de SEC ID N°: 2, y seleccionar plantas que tienen rendimiento aumentado, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, número total de semillas aumentado, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, aumento del índice de cosecha (IC), aumento del peso de mil granos (PMG), aumento de la biomasa de la raíz, cada uno en relación a las plantas de tipo silvestre correspondientes.

Además, la presente invención se refiere al uso de una construcción que comprende

- (i) un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de clase I de SEC ID N°: 2;
- (ii) un promotor de GOS2 capaz de dirigir la expresión de la secuencia de ácido nucleico de (i), y
- (iii) una secuencia de terminación de la transcripción,

en el aumento del rendimiento de las plantas cultivadas en condiciones sin estrés en relación a las plantas de tipo silvestre correspondientes, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado.

La presente invención también abarca el uso de un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de SEC ID N°: 2, o el uso del polipéptido hox5 de HDZip de clase I de SEC ID N°: 2, en el aumento del rendimiento de las plantas cultivadas en condiciones sin estrés en relación a las plantas de tipo silvestre correspondientes, en el que dicho ácido nucleico de hox5 de HDZip de clase I se sobreexpresa en una planta, y en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado, cada uno en relación a las plantas de tipo silvestre correspondientes.

La presente invención también prevé el uso de un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de SEC ID N°: 2, o el uso del polipéptido hox5 de HDZip de clase I de SEC ID N°: 2, en el aumento del rendimiento de plantas en plantas cultivadas en condiciones de estrés abiótico, en el que dicho ácido nucleico de hox5 de HDZip de clase I se sobreexpresa en una planta, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, aumento del índice de cosecha (IC), aumento del peso de mil granos (PMG), aumento de la biomasa de la raíz, cada uno en relación a las plantas de tipo silvestre correspondientes.

Por lo tanto, la presente invención proporciona un procedimiento para mejorar características de crecimiento de plantas que comprende modular la expresión en una planta de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo.

La selección de las plantas de control adecuadas es una parte rutinaria de una configuración experimental y puede incluir plantas de tipo silvestre correspondientes o 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", tal como se usa en el presente documento no solo se refiere a plantas completas, sino también a partes de planta, incluyendo semillas y partes de semilla.

Ventajosamente, la ejecución de los procedimientos de acuerdo con la presente invención da como resultado plantas que tienen características de crecimiento mejoradas, en particular uno o más de rendimiento aumentado, crecimiento mejorado, biomasa mejorada, arquitectura mejorada, división celular mejorada y tolerancia mejorada al estrés abiótico en relación a las plantas de tipo silvestre correspondiente o a otras plantas de control.

La expresión "rendimiento aumentado" se entiende como un aumento en uno o más de los siguientes, cada uno en relación a las plantas de tipo silvestre correspondientes u otras plantas de control: (i) biomasa aumentada (peso) de una o más partes de una planta, particularmente de partes por encima del suelo (cosechables), o 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 biomasa aumentada de cualquier otra parte cosechable; (ii) rendimiento total de semillas aumentado, lo que incluye un aumento en la biomasa de la semilla (peso de la semilla) y que puede ser un aumento en el peso de la semilla por planta basándose en semillas individuales y/o por hectárea o acre; (iii) número aumentado de flores (floretes) por panícula, que se expresa como una proporción del número de semillas cargadas dividido entre el número de panículas primarias; (iv) aumento de la tasa de carga de semillas (expresado en términos de porcentaje como la proporción del número de semillas cargadas entre el número de floretes); (v) número aumentado de semillas (cargadas); (vi) tamaño aumentado de la semilla, que también puede influenciar a la composición de las semillas; (vii) volumen aumentado de la semilla, que también puede influenciar a la

composición de las semillas (incluyendo el contenido total y composición de aceites, proteínas e hidratos de carbono); (viii) área de la semilla aumentada (individual o media); (ix) longitud de la semilla aumentada (individual o media); (x) anchura de la semilla aumentada (individual o media); (xi) perímetro de la semilla aumentado (individual o medio); (xii) aumento del índice de cosecha (IC), que se expresa como una proporción del rendimiento de partes cosechables, tales como semillas, entre la biomasa total; y (xiii) aumento del peso de mil granos (PMG), que se extrapola a partir del número de semillas cargadas contadas y de su peso total. Un PMG aumentado puede ser el resultado de un tamaño de semilla y/o de un peso de semilla aumentados. Un PMG aumentado puede ser el resultado de un aumento en el tamaño del embrión y/o del tamaño del endospermo.

Tomando al maíz como ejemplo, un aumento del rendimiento puede manifestarse como uno o más de los siguientes: aumento en el número de plantas establecidas 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 del grano, peso de mil granos, longitud/diámetro de la espiga, aumento en la tasa de carga de semillas (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), entre otros. Tomando al arroz como ejemplo, un aumento en el rendimiento puede manifestarse en sí mismo 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 (floretes) por panícula (que se expresa como una proporción del número de semillas llenas dividido entre el número de panículas primarias), aumento en la tasa de carga de semillas (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), aumento en el peso de mil granos, entre otros.

Ya que las plantas transgénicas de acuerdo con la presente invención tienen rendimiento aumentado, es probable que estas plantas muestren una tasa de crecimiento aumentada (durante al menos parte de su ciclo vital), en relación a la velocidad de crecimiento de las plantas de control en un estado correspondiente de su ciclo vital. La velocidad de crecimiento aumentada puede ser específica para una o más partes de una planta (incluyendo las semillas), o puede ser sustancialmente a lo largo de la planta completa. Las plantas que tienen una velocidad de crecimiento aumentada pueden tener un ciclo vital más corto. El ciclo vital de una planta puede entenderse como el tiempo necesario para crecer desde una semilla madura seca hasta el estado en el que la planta ha producido semillas maduras secas, similares al material de partida. Este ciclo vital puede estar influenciado por factores tales como el vigor temprano, la velocidad de crecimiento, en índice de verdor, el tiempo de floración y la velocidad de maduración de la semilla. El aumento en la velocidad de crecimiento puede tener lugar en uno o más estados en el ciclo vital de una planta o durante sustancialmente todo el ciclo vital de la planta. La velocidad de crecimiento aumentada durante las etapas tempranas del ciclo vital de una planta puede reflejar un vigor potenciado. El aumento en la velocidad de crecimiento puede alterar el ciclo de cosecha de una planta permitiendo que se siembren las plantas más tarde y/o se cosechen más pronto de lo que sería posible de otro modo (puede obtenerse un efecto similar con el tiempo de floración temprana). Si la velocidad de crecimiento aumenta lo suficiente, también puede permitir la siembra adicional de semillas de la misma especie de planta (por ejemplo, siembra y cosecha de plantas de arroz seguida de siembra y cosecha de plantas de arroz adicional todas dentro de un periodo de crecimiento convencional). De igual forma, si la velocidad de crecimiento aumenta lo suficiente, puede permitir la siembra adicional de semillas de diferentes especies de plantas (por ejemplo, la siembra y cosecha de plantas de arroz seguida de, por ejemplo, la siembra y opcional cosecha de soja, patata o cualquier otra planta adecuada). También puede ser posible cosechar un número de veces adicionales el mismo rizoma en el caso de algunas plantas de cultivo. Alterar el ciclo de cosecha de una planta puede dar lugar a un aumento en la producción de biomasa anual por metro cuadrado (debido a un aumento en el número de veces (por ejemplo, un año) que se puede cultivar y cosechar cualquier planta particular). Un aumento en la velocidad de crecimiento también puede permitir el cultivo de plantas transgénicas en un área geográfica mayor que la de sus homólogos de tipo silvestre, debido a que las limitaciones territoriales para cosechar un cultivo se determinan a menudo por condiciones ambientales adversas bien en el momento de plantar (estación temprana) o bien en el momento de cosechar (estación tardía). Dichas condiciones adversas pueden evitarse si se acorta el ciclo de la cosecha. La velocidad de crecimiento puede determinarse derivando diversos parámetros de curvas de crecimiento, pudiendo ser dichos parámetros: T-Medio (el tiempo que tardan las plantas en alcanzar el 50 % de su tamaño máximo) y T-90 (el tiempo que tardan las plantas en alcanzar el 90 % de su tamaño máximo), entre otros.

El rendimiento de los procedimientos de la divulgación proporciona plantas que tienen una velocidad de crecimiento aumentada. Por lo tanto, de acuerdo con la presente divulgación, se proporciona un procedimiento para aumentar la velocidad de crecimiento de plantas, comprendiendo dicho procedimiento modular la expresión en una planta de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o de un homólogo del mismo.

Un aumento en el rendimiento y/o velocidad de crecimiento se produce cuando la planta está en condiciones sin estrés o cuando la planta está expuesta a diversos tipos de estrés en comparación con las plantas de control. Típicamente, las plantas responden a la exposición frente al estrés creciendo más lentamente. En condiciones de estrés intenso, la planta puede incluso detener por completo su crecimiento. Por otro lado, el 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 sin la capacidad de reanudar el crecimiento. El estrés leve, en el sentido de la invención, da lugar a una reducción en el crecimiento de las plantas estresadas de menos del 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 sin estrés. Debido a los avances en las prácticas agrícolas (riego, fertilización, tratamiento con pesticidas) las formas de estrés intenso no se encuentran frecuentemente en las plantas cultivadas. Como consecuencia, el crecimiento comprometido inducido por estrés leve es a menudo una característica no deseable para

la agricultura. Las formas de estrés leve son formas de estrés biótico y/o abiótico (ambiental) comunes a las que está expuesta la planta. Las formas de estrés abiótico pueden deberse a sequía o a exceso de agua, a estrés anaeróbico, a estrés salino, toxicidad química, estrés oxidativo y temperaturas cálidas, frías o de congelación. El estrés abiótico puede ser un estrés osmótico causado por un estrés hídrico (particularmente debido a sequía), a estrés salino, estrés oxidativo o un estrés iónico. Las formas de estrés biótico son típicamente aquellas causadas por patógenos, tales como bacterias, virus, hongos e insectos.

La ejecución de los procedimientos de acuerdo con la presente invención da como resultado plantas que tienen tolerancia aumentada al estrés abiótico. Tal como se describe en Wang y col., (Planta (2003) 218: 1-14), el estrés abiótico da lugar a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan adversamente al crecimiento y productividad de la planta. Se sabe que la sequía, las temperaturas extremas y el estrés oxidativo están interconectados y pueden inducir daño celular y al crecimiento mediante mecanismos similares. Por ejemplo, la sequía y/o la salinización se manifiestan principalmente como estrés osmótico, dando como resultado la interrupción de la homeostasis y de la distribución de iones en la célula. El estrés oxidativo, que frecuentemente viene acompañado de estrés por altas o bajas temperaturas, salinidad o sequía puede causar la desnaturalización de proteínas funcionales y estructurales. Como consecuencia, estas diversas formas de estrés ambiental activan a menudo rutas de señalización celular y respuestas celulares similares, tales como la producción de proteínas de estrés, regulación positiva de antioxidantes, acumulación de solutos compatibles y detención del crecimiento.

Ya que diversas formas de estrés ambientales activan rutas similares, la ejemplificación de la presente invención con estrés por sequía (en la medida en que la invención se refiere al uso de polipéptidos de hox4 de HDZip de clase I tal como se definen en la reivindicación 1 y de sus ácidos nucleicos codificantes) no debe entenderse como una limitación al estrés por sequía, sino más como una exploración para indicar la implicación de polipéptidos de hox5 de HDZip de clase I o de homólogos de los mismos en las formas de estrés abiótico en general. Además, los procedimientos de la presente invención pueden llevarse a cabo en condiciones sin estrés o en condiciones de sequía leve para proporcionar plantas que tengan características de crecimiento mejoradas (en particular rendimiento aumentado) en relación a las plantas de tipo silvestre correspondientes u otras plantas de control. La expresión condiciones "sin estrés", tal como se usa en el presente documento, se refiere preferentemente a aquellas condiciones ambientales que no sobrepasan de manera significativa las condiciones climáticas diarias y otras condiciones abióticas que puedan encontrarse las plantas, más preferentemente aquellas condiciones que permitan el crecimiento óptimo de las plantas. Los expertos en la materia son conscientes de las condiciones normales del suelo y de las condiciones climáticas para una localización dada.

Se ha comunicado un grado particularmente alto de "intercomunicación" entre el estrés por sequía y el estrés por alta salinidad (Rabbani y col., (2003) Plant Physiol 133: 1755-1767). Por lo tanto, será evidente que un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo podría, junto con su utilidad para conferir tolerancia a la sequía en plantas, ser de utilidad para proteger a la planta contra otras diversas formas de estrés abiótico. De igual forma, será evidente que una cinasa similar a *shaggy* de grupo I (tal como se define en el presente documento) podría, junto con su utilidad para conferir tolerancia a la sal en plantas, ser de utilidad para proteger a la planta contra otras diversas formas de estrés abiótico. Además, Rabbani y col., (2003, Plant Physiol 133: 1755-1767) comunican que existen mecanismos moleculares similares de tolerancia y respuestas al estrés entre dicotiledóneas y monocotiledóneas. Por lo tanto, los procedimientos de la invención pueden aplicarse de manera ventajosa a cualquier planta.

La expresión "estrés abiótico", tal como se define en el presente documento, se entiende como uno cualquiera o más de: estrés hídrico (debido a sequía o exceso de agua), a estrés anaeróbico, a estrés salino, estrés por temperatura (debido a temperaturas cálidas, frías o de congelación), estrés por toxicidad química y estrés oxidativo. De acuerdo con un aspecto de la invención, el estrés abiótico es un estrés osmótico, seleccionado entre estrés hídrico, estrés salino, estrés oxidativo y estrés iónico. Preferentemente, el estrés hídrico es estrés por sequía. La expresión estrés hídrico no está restringida a la sal común (NaCl), sino que puede ser una cualquiera o más de: NaCl, KCl, LiCl, MgCl<sub>2</sub>, CaCl<sub>2</sub>, entre otras.

La tolerancia aumentada al estrés abiótico se manifiesta por un rendimiento aumentado de la planta en condiciones de estrés abiótico. Particularmente en la medida en que la invención se refiere al uso de polipéptidos de hox5 de HDZip de clase I y de sus ácidos nucleicos codificantes, dicho rendimiento aumentado puede incluir uno o más de los siguientes: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, número aumentado de flores por panícula, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado, diámetro de las raíces aumentado, cada uno en relación a las plantas de tipo silvestre correspondientes.

La ejecución de los procedimientos de la invención proporciona plantas que tienen tolerancia aumentada al estrés abiótico. La ejecución de los procedimientos de la invención proporciona plantas cultivadas en condiciones sin estrés o en condiciones de sequía leve con características de crecimiento mejoradas (particularmente rendimiento aumentado) en relación a las plantas de tipo silvestre correspondientes u otras plantas cultivadas en condiciones comparables.

De acuerdo con la presente divulgación, se proporciona un procedimiento para aumentar la tolerancia al estrés abiótico en plantas, comprendiendo dicho procedimiento modular la expresión en una planta de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o de un homólogo del mismo. De acuerdo con un aspecto de la divulgación, el estrés abiótico es un estrés osmótico, seleccionado entre uno o más de los siguientes: estrés hídrico, a

estrés salino, estrés oxidativo y estrés iónico. Preferentemente, el estrés hídrico es estrés por sequía.

Particularmente en la medida en que la divulgación se refiere al uso de polipéptidos de hox5 de HDZip de clase I y de sus ácidos nucleicos codificantes, la ejecución de los procedimientos de la divulgación proporciona un índice de verdor aumentado en relación a plantas de tipo silvestre correspondientes. El índice de verdor, tal como se define en el presente documento, es la proporción (expresada como %) de píxeles amarillos en imágenes de plantas registradas por una cámara digital. Un índice de verdor aumentado puede indicar una senescencia reducida o retrasada que a su vez permite prolongar la actividad fotosintética de una planta, lo que a su vez da lugar a diversos efectos beneficiosos bien conocidos en la técnica.

Por lo tanto, la divulgación proporciona un procedimiento para aumentar el índice de verdor en plantas, comprendiendo dicho procedimiento modular la expresión en una planta de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo. Preferentemente, el índice de verdor aumenta en condiciones de estrés abiótico, más preferentemente en condiciones de estrés hídrico, aún más preferentemente en condiciones de estrés por sequía.

Preferentemente, en los casos donde el procedimiento de la invención comprenda modular la expresión en una planta de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo, el rendimiento aumentado incluye uno o más de los siguientes: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, número aumentado de flores por panícula, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado, longitud de las raíces aumentada, cada uno en relación a las plantas de tipo silvestre correspondientes u otras plantas de control.

De acuerdo con una característica preferida de la presente divulgación, se proporciona un procedimiento para aumentar el rendimiento de una planta en relación a plantas de tipo silvestre correspondientes u otras plantas de control, comprendiendo dicho procedimiento modular la expresión en una planta de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo.

Por lo tanto, los procedimientos de la invención pueden aplicarse de manera ventajosa a cualquier planta.

El término "planta", tal como se usa en el presente documento, abarca plantas completas, ancestros y descendencia de plantas y partes de plantas, incluyendo semillas, brotes, tallos, hojas, raíces (incluyendo tubérculos), flores, y tejidos y órganos, en la que cada uno de los anteriormente mencionados comprende el gen/ácido nucleico de interés. El término "planta" también abarca células de planta, cultivos en suspensión, tejido calloso, embriones, regiones meristemáticas, gametofitos, esporofitos, polen y microesporas, de nuevo en la que cada uno de los anteriormente mencionados comprende el gen/ácido nucleico de interés.

Las plantas que son particularmente útiles en los procedimientos de la invención incluyen todas las plantas que pertenecen a la superfamilia *Viridiplantae*, en particular a plantas monocotiledóneas y dicotiledóneas, incluyendo forraje y leguminosas de forraje, plantas ornamentales, cultivos alimenticios, árboles o arbustos seleccionados entre la lista que comprende *Acacia* spp., *Acer* spp., *Actinidia* spp., *Aesculus* spp., *Agathis australis*, *Albizia amara*, *Alsophila tricolor*, *Andropogon* spp., *Arachis* spp., *Areca catechu*, *Astelia fragrans*, *Astragalus cicer*, *Baikiaea plurijuga*, *Betula* spp., *Brassica* spp., *Bruguiera gymnorrhiza*, *Burkea africana*, *Butea frondosa*, *Cadaba farinosa*, *Calliandra* spp., *Camellia sinensis*, *Canna indica*, *Capsicum* spp., *Cassia* spp., *Centroema pubescens*, *Chaenomeles* spp., *Cinnamomum cassia*, *Coffea arabica*, *Colophospermum mopane*, *Coronilla varia*, *Cotoneaster serotina*, *Crataegus* spp., *Cucumis* spp., *Cupressus* spp., *Cyathia dealbata*, *Cydonia oblonga*, *Cryptomeria japonica*, *Cymbopogon* spp., *Cynthea dealbata*, *Cydonia oblonga*, *Dalbergia monetaria*, *Davallia divaricata*, *Desmodium* spp., *Dicksonia squarrosa*, *Diheteropogon amplexans*, *Dioclea* spp., *Dolichos* spp., *Dorycnium rectum*, *Echinochloa pyramidalis*, *Ehretia* spp., *Eleusine coracana*, *Fragaria* spp., *Erythrina* spp., *Eucalyptus* spp., *Euclea schimperi*, *Eulalia villosa*, *Fagopyrum* spp., *Feijoa sellowiana*, *Fragaria* spp., *Flemingia* spp., *Freyinetia banksii*, *Geranium thunbergii*, *Ginkgo biloba*, *Glycine javanica*, *Gliricidia* spp., *Gossypium hirsutum*, *Grevillea* spp., *Guibourtia coleosperma*, *Hedysarum* spp., *Hemarthria altissima*, *Heteropogon contortus*, *Hordeum vulgare*, *Hyparrhenia rufa*, *Hypericum erectum*, *Hyperthelia dissoluta*, *Indigo incarnata*, *Iris* spp., *Leptarrhena pyrolifolia*, *Lespedeza* spp., *Lettuca* spp., *Leucaena leucocephala*, *Loudetia simplex*, *Lotonus bainesii*, *Lotus* spp., *Macrotyloma axillare*, *Malus* spp., *Manihot esculenta*, *Medicago sativa*, *Metasequoia glyptostroboides*, *Musa sapientum*, *Nicotianum* spp., *Onobrychis* spp., *Ornithopus* spp., *Oryza* spp., *Peltophorum africanum*, *Pennisetum* spp., *Persea gratissima*, *Petunia* spp., *Phaseolus* spp., *Phoenix canariensis*, *Phormium cookianum*, *Photinia* spp., *Picea glauca*, *Pinus* spp., *Pisum sativum*, *Podocarpus totara*, *Pogonarthria fleckii*, *Pogonarthria squarrosa*, *Populus* spp., *Prosopis cineraria*, *Pseudotsuga menziesii*, *Pterolobium stellatum*, *Pyrus communis*, *Quercus* spp., *Rhaphiolepis umbellata*, *Rhopalostylis sapida*, *Rhus natalensis*, *Ribes grossularia*, *Ribes* spp., *Robinia pseudoacacia*, *Rosa* spp., *Rubus* spp., *Salix* spp., *Schyzachyrium sanguineum*, *Sciadopitys verticillata*, *Sequoia sempervirens*, *Sequoiadendron giganteum*, *Sorghum bicolor*, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus alopecuroides*, *Stylosanthes humilis*, *Tadehagi* spp., *Taxodium distichum*, *Themeda triandra*, *Trifolium* spp., *Triticum* spp., *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vitis vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*, *Zea mays*, amaranto, alcachofa, espárrago, brécol, coles de Bruselas, repollo, canola, zanahoria, coliflor, apio, berza, lino, col rizada, lenteja, colza, okra, cebolla, patata, arroz, soja, fresa, remolacha azucarera, caña de azúcar, girasol, tomate, calabaza, té y algas, entre otros.

De acuerdo con una realización preferida de la presente invención, la planta es una planta de cultivo, tal como soja, girasol, canola, alfalfa, colza, algodón, tomate, patata o tabaco. De manera preferentemente adicional, la planta es una planta monocotiledónea, tal como caña de azúcar. Más preferentemente la planta es un cereal, tal como arroz, maíz, trigo, cebada, mijo, centeno, sorgo o avena.

**5 Polipéptidos de hox5 de HDZip de clase I y homólogos de los mismos, y ácidos nucleicos/genes de hox5 de HDZip de clase I útiles en los procedimientos de la invención**

La expresión "polipéptido de hox5 de HDZip de clase I u homólogo del mismo", tal como se define en el presente documento, se refiere a 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.

10 Además, el polipéptido de hox5 de HDZip de clase I o un homólogo del mismo puede comprender uno cualquiera o ambos de los siguientes: (a) una cola de Trp; y (b) el motivo de aminoácidos RPFF, donde R es Arg, P es Pro y F es Phe. Dentro de este motivo, se permiten uno o más cambios conservativos en cualquier posición, y uno o dos cambios no conservativos en cualquier posición. El motivo de (b) precede a la caja ácida, cuando se examina la proteína desde N-terminal hasta C-terminal.

15 Un ejemplo de un polipéptido de hox5 de HDZip de clase I, tal como se define anteriormente, 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; y que comprende además: (a) una cola de Trp; y (b) el motivo de aminoácidos RPFF, donde R es Arg, P es Pro y F es Phe, se representa como SEC ID N°: 2. Además, dichos ejemplos se proporcionan en la tabla A del ejemplo 1 en el presente documento.

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

25 Los polipéptidos de hox5 de HDZip de clase I u homólogos de los mismos pueden identificarse fácilmente usando técnicas rutinarias bien conocidas en la materia, tal como mediante alineamiento de secuencias. Los procedimientos para el alineamiento de secuencias para comparación se conocen bien en la técnica, incluyendo dichos métodos GAP, BESTFIT, BLAST, FASTA y TFASTA. GAP usa el algoritmo de Needleman y Wunsch ((1970) J Mol Biol 48: 443-453) para hallar el alineamiento de dos secuencias completas que maximiza el número de coincidencias y minimiza el número de huecos. El algoritmo BLAST (Altschul y col., (1990) J Mol Biol 215: 403-10) calcula el porcentaje de identidad de secuencia y efectúa un análisis estadístico de la similitud entre las dos secuencias. El programa informático para efectuar un análisis BLAST está disponible públicamente a través del National Centre for Biotechnology Information. Los homólogos de hox5 de HDZip de clase I que comprenden un homeodominio de clase I y una cremallera de leucina con más de 5 héptadas pueden identificarse fácilmente usando, por ejemplo, el algoritmo de alineamiento de múltiples secuencias ClustalW (versión 1.83) disponible en  
35 <http://clustalw.genome.jp/sit-bin/nph-ClustalW>, con los parámetros de alineamiento por pares por defecto, y un procedimiento de puntuación en porcentaje. Puede efectuarse una edición manual menor para optimizar el alineamiento entre motivos conservados, tal como será evidente para un experto en la materia.

Los diversos dominios estructurales en una proteína hox5 de HDZip de clase I, tales como el homeodominio y la cremallera de leucina, pueden identificarse usando bases de datos especializadas, por ejemplo, SMART (Schultz y col., (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic y col. (2002) Nucleic Acids Res 30, 242-244; <http://smart.embl-heidelberg.de/>), InterPro (Mulder y col., (2003) Nucl. Acids. Res. 31, 315-318; <http://www.ebi.ac.uk/interpro/>), Prosite (Bucher y Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB- 94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., págs 53-61, AAAIPress, Menlo Park; Hulo y col., Nucl. Acids. Res. 32:D134-D137, (2004), <http://www.expasy.org/prosite/>) o Pfam (Bateman y col., Nucleic Acids Research 30(1):276-280 (2002), <http://www.sanger.ac.uk/Software/Pfam/>). La predicción de la cremallera de leucina y la identificación de la héptada puede efectuarse usando programas informáticos especializados, tales como 2ZIP, que combina un algoritmo convencional de predicción hélices superenrolladas con una búsqueda aproximada para la repetición de leucina  
50 característica (Bomberg-Bauter y col., (1998) Computational Approaches to Identify Leucine Zippers, Nucleic Acids Res., 26(11): 2740-2746, <http://2zip.molgen.mpg.de>).

Además, también puede identificarse fácilmente la presencia de una caja ácida. Puede calcularse la composición primaria de aminoácidos (en %) para determinar si un dominio de polipéptido es rico en aminoácidos específicos usando programas informáticos del servidor ExPASy, en particular la herramienta ProtParam (Gasteiger E y col., (2003) ExPASy: the proteomics server for in- depth protein knowledge and analysis. Nucleic Acids Res 31:3784-3788). Entonces puede compararse la composición de la proteína de interés con la composición media de aminoácidos (en %) en el banco de datos de secuencias de proteínas Swiss-Prot. En este banco de datos, el contenido medio de Asp (D) y de Glu (E) son del 5,3 % y del 6,6 %, respectivamente, siendo la media combinada del 11,9 %. Como ejemplo, la caja ácida de SEC ID N°: 2 contiene un 9,1 % de D y un 54,5 % de E, siendo la media combinada del 63,6 %. Tal como

se define en el presente documento, una caja rica en ácido tiene un contenido combinado de Asp (D) y de Glu (E) (en términos de %) por encima del encontrado en la composición de aminoácidos media (en términos de %) de las proteínas en la base de datos de secuencias de proteínas Swiss-Prot. Una caja ácida puede ser parte de un dominio de activación de la transcripción. Los dominios de activación de la transcripción de eucariotas se han clasificado de acuerdo con su contenido de aminoácidos, y las categorías principales incluyen dominios de activación ácidos, ricos en glutamina y ricos en prolina (Rutherford y col., (2005) Plant J. 43(5):769-88, y las referencias citadas en el mismo).

Un número seleccionado de proteínas entre los polipéptidos de hox5 de HDZip de clase I u homólogos de los mismos comprenden además el motivo de aminoácidos RPFF, donde R es Arg, P es Pro y F es Phe. Dentro de este motivo, se permiten uno o más cambios conservativos en cualquier posición, y uno o dos cambios no conservativos en cualquier posición. Este motivo precede a la caja ácida, cuando se examina la proteína desde N-terminal hasta C-terminal (véase la figura 2). La presencia del RPFF puede identificarse usando procedimientos para el alineamiento de secuencias para comparación tal como se ha descrito anteriormente en el presente documento. En algunos casos, pueden ajustarse los parámetros por defecto para modificar la rigurosidad de la búsqueda. Por ejemplo, usando BLAST, puede aumentarse el umbral de significación estadística (denominado valor de "expectativa") para comunicar coincidencias contra secuencias de la base de datos para mostrar coincidencias menos rigurosas. De este modo, pueden identificarse coincidencias cortas prácticamente exactas.

Un número seleccionado de proteínas entre los polipéptidos de hox5 de HDZip de clase I u homólogos de los mismos pueden comprender además una cola de Trp. Una cola de Trp, tal como se define en el presente documento es los últimos 10 aminoácidos del extremo C-terminal de la proteína que comprende al menos un resto de Trp (véase la figura 2).

Los ejemplos de polipéptidos de hox5 de HDZip de clase I u homólogos de los mismos (codificados por el número de referencia de secuencia entre paréntesis) se proporcionan en la tabla A. Las SEC ID N°: 3 a 32, 34 a 51 se muestran con fines ilustrativos.

Debe entenderse que las secuencias que se encuentran dentro de la definición de "polipéptido de hox5 de HDZip de clase I u homólogo del mismo" no deben limitarse a las secuencias proporcionadas en la tabla A, sino que cualquier polipéptido que comprende desde N-terminal hasta 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 ejecución de los procedimientos de la divulgación.

Los polipéptidos de hox5 de HDZip de clase I u homólogos de los mismos tienen actividad de unión a ADN, preferentemente a semisitios de 5 pb que se solapan en una posición central, CAA(A/T)ATTG, tal como se detecta en ensayos de un híbrido de levadura (Meijer y col., (2000) Mol Gen Genet 263:12-21). En ensayos transitorios en suspensiones de células de arroz, el bombardeo conjunto de un polipéptido de hox5 de HDZip de clase I con el gen indicador GUS dio como resultado un aumento en el número de puntos teñidos, que también tenían un color más intenso (Meijer y col., anteriormente citado). Este ensayo es útil para demostrar la función activadora de los polipéptidos u homólogos de hox5 de HDZip de clase I.

Las secuencias útiles en los procedimientos de la divulgación no están limitadas a los ácidos nucleicos y polipéptidos de hox5 de homeodominio de cremallera de leucina (HDZip) de clase I anteriormente mencionados. Los procedimientos de acuerdo con la presente divulgación también pueden llevarse a cabo usando variantes de ácidos nucleicos que codifican polipéptidos de hox5 de homeodominio de cremallera de leucina (HDZip) de clase I u homólogos de los mismos.

Los ejemplos de dichas variantes incluyen porciones, secuencias hibridantes, variantes alélicas, variantes de corte y empalme y variantes obtenidas mediante reordenamiento génico.

Puede prepararse una porción, por ejemplo, produciendo una o más eliminaciones en un ácido nucleico. Las porciones pueden usarse en forma aislada o pueden fusionarse a otras secuencias codificantes (o no codificantes) para, por ejemplo, producir una proteína que combine varias actividades. Cuando se fusionan a otras secuencias codificantes, el polipéptido resultante producido tras la traducción puede ser mayor que aquel predicho para la porción.

En los casos donde la secuencia útil en los procedimientos de la divulgación sea un hox5 de HDZip de clase I, la porción codifica un polipéptido que comprende desde N-terminal hasta C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Preferentemente, la porción es una porción de uno de los ácidos nucleicos proporcionados en la tabla A. Más preferentemente, la porción es una porción de un ácido nucleico tal como se representa por la SEC ID N°: 1.

Por lo tanto, la divulgación proporciona un procedimiento para mejorar características de crecimiento de plantas que comprende modular la expresión en una planta o una parte de un ácido nucleico que codifica un polipéptido de hox5 de homeodominio de cremallera de leucina (HDZip) de clase I o un homólogo del mismo.

Otra variante de ácido nucleico es un ácido nucleico capaz de hibridar en condiciones de rigurosidad reducida, preferentemente en condiciones rigurosas, con un ácido nucleico que codifica un ácido nucleico/gen de hox5 de HDZip de clase I u homólogo del mismo.

El término "hibridación", tal como se define en el presente documento, es un procedimiento en el que las secuencias nucleotídicas complementarias sustancialmente homólogas hibridan entre sí. El procedimiento de hibridación puede suceder completamente en solución, es decir, ambos ácidos nucleicos complementarios se encuentran en solución. El procedimiento de hibridación también puede suceder con uno de los ácidos nucleicos complementarios inmovilizados a una matriz, tal como perlas magnéticas, perlas de sefarsa o cualquier otra resina. El procedimiento de hibridación puede suceder además con uno de los ácidos nucleicos complementarios inmovilizados a un soporte sólido, tal como una membrana de nitrocelulosa o de nylon o inmovilizados mediante por ejemplo fotolitografía a, por ejemplo, un soporte de vidrio silíceo (este último conocido como matrices o micromatrices de ácidos nucleicos o como placas de ácidos nucleicos). Para permitir que tenga lugar la hibridación, las moléculas de ácido nucleico se desnaturalizan generalmente de manera térmica o química para separar una doble hebra en dos hebras individuales y/o para eliminar horquillas u otras estructuras secundarias de los ácidos nucleicos monocatenarios. La rigurosidad de la hibridación está influenciada por condiciones tales como la temperatura, la concentración de sal, la fuerza iónica y la composición del tampón de hibridación.

Las "condiciones de hibridación rigurosas" y las "condiciones de lavado de hibridación rigurosas" en el contexto de los experimentos de hibridación de ácidos nucleicos, tales como hibridaciones de Southern y de Northern son dependientes de la secuencia y son diferentes con parámetros ambientales diferentes. El experto en la materia será consciente de diversos parámetros que pueden alterarse durante la hibridación y el lavado y que bien mantendrán o cambiarán las condiciones de rigurosidad.

La  $T_f$  es la temperatura a una fuerza iónica y pH definidos, a la cual el 50 % de la secuencia diana hibrida con una sonda perfectamente emparejada. La  $T_f$  depende de las condiciones de solución y de la composición base y longitud de la sonda. Por ejemplo, las secuencias más largas hibridan específicamente a temperaturas mayores. La velocidad máxima de hibridación se obtiene desde aproximadamente 16 °C hasta 32 °C por debajo de la  $T_f$ . La presencia de cationes monovalentes en la solución de hibridación reduce la repulsión electrostática entre las dos hebras de ácido nucleico, promoviendo de este modo la formación de híbridos; este efecto es visible para concentraciones de sodio de hasta 0,4 M. La formamida reduce la temperatura de fusión de los dúplex de ADN-ADN y ADN-ARN en 0,6 a 0,7 °C para cada porcentaje de formamida, y la adición de formamida al 50 % permite que se efectúe la hibridación a 30 a 45 °C, aunque disminuirá la velocidad de hibridación. Los emparejamientos erróneos de pares de bases reducen la velocidad de hibridación y la estabilidad térmica de los dúplex. De media y para todas las sondas, la  $T_f$  disminuye aproximadamente 1 °C por % de emparejamiento erróneo de bases. La  $T_f$  puede calcularse usando las siguientes ecuaciones, dependiendo de los tipos de híbridos:

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

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

2. Híbridos de ADN-ARN o de ARN-ARN:

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

3. Híbridos de oligo-ADN u oligo-ARN<sup>d</sup>:

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

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

<sup>a</sup> o para otro catión monovalente, pero solo preciso en el intervalo de 0,01-0,4 M.

<sup>b</sup> solo preciso para %GC en el intervalo del 30 % al 75 %.

<sup>c</sup> L = Longitud del dúplex en pares de bases.

<sup>d</sup> Oligo, oligonucleótido;  $I_n$ , longitud eficaz del cebador =  $2 \times (\text{n.º de G/C}) + (\text{n.º de A/T})$ .

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

La especificidad de la hibridación es típicamente la función de los lavados post-hibridación. Para eliminar el fondo resultante de la unión 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 más baja sea la concentración de sal y mayor sea la temperatura de lavado, mayor será la rigurosidad del lavado. Las condiciones de lavado se llevan a cabo típicamente a, o por debajo de, la rigurosidad de la hibridación. En general, las condiciones rigurosas adecuadas para ensayos de hibridación de ácido nucleico o procedimientos de detección de la amplificación génica son tal como se han expuesto anteriormente. También pueden seleccionarse condiciones de mayor o menor rigurosidad. En general, las condiciones de rigurosidad baja se seleccionan para que sean aproximadamente 50 °C menores que el punto de fusión térmico ( $T_f$ ) para la secuencia específica a una fuerza iónica y pH definidos. Las condiciones de rigurosidad media se producen cuando la temperatura es 20 °C menor que la  $T_f$ , y las condiciones de rigurosidad alta se producen cuando la temperatura es 10 °C menor que la  $T_f$ . Por ejemplo, son condiciones rigurosas aquellas que son al menos tan rigurosas como, por ejemplo, las condiciones A-L; y las condiciones de rigurosidad reducida son al menos tan rigurosas como, por ejemplo, las condiciones M-R. La unión no específica puede controlarse usando una cualquiera o una serie de técnicas conocidas tales como, por ejemplo, bloqueo de la membrana con soluciones que contienen proteína, adiciones de ARN, ADN heterólogo y SDS al tampón de

hibridación, y tratamiento con RNasa. Los ejemplos de condiciones de lavado e hibridación se enumeran en la tabla 2 a continuación.

**Tabla 2:** Ejemplos de condiciones de hibridación y lavado

| Condición de rigurosidad | Híbrido de polinucleótido <sup>‡</sup> | Longitud de híbrido (pb) <sup>‡</sup> | Temperatura y tampón de hibridación <sup>†</sup>  | Temperatura y tampón de lavado <sup>†</sup> |
|--------------------------|----------------------------------------|---------------------------------------|---------------------------------------------------|---------------------------------------------|
| A                        | ADN:ADN                                | > o igual a 50                        | 65 °C SSC 1X; o 42 °C, SSC 1X y formamida al 50 % | 65 °C; SSC 0,3X                             |
| B                        | ADN:ADN                                | <50                                   | Tb*; SSC 1X                                       | Tb*; SSC 1X                                 |
| C                        | ADN:ARN                                | > o igual a 50                        | 67 °C 1XSSC; o 45 °C, SSC 1X y formamida al 50 %  | 67 °C; 0,3XSSC                              |
| D                        | ADN:ARN                                | <50                                   | Td*; SSC 1X                                       | Td*; SSC 1X                                 |
| E                        | ARN:ARN                                | > o igual a 50                        | 70 °C 1XSSC; o 50 °C, SSC 1X y formamida al 50 %  | 70 °C; 0,3XSSC                              |
| F                        | ARN:ARN                                | <50                                   | Tf*; SSC 1X                                       | Tf*; SSC 1X                                 |
| G                        | ADN:ADN                                | > o igual a 50                        | 65 °C 4XSSC; o 45 °C, SSC 4X y formamida al 50 %  | 65 °C; SSC 1X                               |
| H                        | ADN:ADN                                | <50                                   | Th*; SSC 4X                                       | Th*; SSC 4X                                 |
| I                        | ADN:ARN                                | > o igual a 50                        | 67 °C 4XSSC; o 45 °C, SSC 4X y formamida al 50 %  | 67 °C; SSC 1X                               |
| J                        | ADN:ARN                                | <50                                   | Tj*; SSC 4X                                       | Tj*; SSC 4X                                 |
| K                        | ARN:ARN                                | > o igual a 50                        | 70 °C 4XSSC; o 40 °C, SSC 6X y formamida al 50 %  | 67 °C; SSC 1X                               |
| L                        | ARN:ARN                                | <50                                   | T1, SSC 2X                                        | T1, SSC 2X                                  |
| M                        | ADN:ADN                                | > o igual a 50                        | 50 °C 4XSSC; o 40 °C, SSC 6X y formamida al 50 %  | 50 °C; SSC 2X                               |
| N                        | ADN:ADN                                | <50                                   | Tn*; SSC 6X                                       | Tn*; SSC 6X                                 |
| O                        | ADN:ARN                                | > o igual a 50                        | 55 °C 4XSSC; o 42 °C, SSC 6X y formamida al 50 %  | 55 °C; SSC 2X                               |
| P                        | ADN:ARN                                | <50                                   | Tp*; SSC 6X                                       | Tp*; SSC 6X                                 |
| Q                        | ARN:ARN                                | > o igual a 50                        | 60 °C 4XSSC; o 45 °C, SSC 6X y formamida al 50 %  | 60 °C.; SSC 2X                              |
| R                        | ARN:ARN                                | <50                                   | Tr*; SSC 4X                                       | Tr*; SSC 4X                                 |

<sup>‡</sup> La "longitud del híbrido" es la longitud anticipada para el ácido nucleico hibridante. Cuando se hibridan ácidos nucleicos de secuencia conocida, la longitud del híbrido puede determinarse alineando las secuencias e identificando las regiones conservadas descritas en el presente documento.

<sup>†</sup> SSPE (SSPE 1X es NaCl 0,15 M, NaH<sub>2</sub>PO<sub>4</sub> 10 mM, y EDTA 1,25 mM, pH 7,4) puede sustituirse con SSC (SSC 1x es NaCl 0,15 M y citrato de sodio 15 mM) en los tampones de hibridación y lavado; los lavados se llevan a cabo 15 minutos después de completar la hibridación. Las hibridaciones y lavados pueden incluir adicionalmente reactivo de Denhardt 5X, SDS al 0,5 - 1,0 %, 100 µg/ml de ADN de esperma de salmón fragmentado desnaturalizado, pirofosfato de sodio al 0,5 %, y formamida hasta el 50 %.

\* Tb-Tr: La temperatura de hibridación para los híbridos que se anticipa que tienen menos de 50 pares de bases de longitud debe ser 5-10 °C menos que la temperatura de fusión T<sub>f</sub> de los híbridos; la T<sub>f</sub> se determina de acuerdo con las ecuaciones anteriormente mencionadas. <sup>‡</sup> La presente invención abarca la sustitución de uno cualquiera o más de los compañeros híbridos de ADN o ARN con bien un APN o un ácido nucleico modificado.

Con el fin de definir el nivel de rigurosidad, puede hacerse referencia a Sambrook y col., (2001) Molecular Cloning: a laboratory manual, 3ª edición, Cold Spring Harbor Laboratory Press, CSH, Nueva York o a Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989).

- 5 En los casos donde la secuencia útil en los procedimientos de la divulgación sea un *hox5* de HDZip de clase I, la secuencia hibridante codifica un polipéptido que comprende desde N-terminal hasta C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Preferentemente, la secuencia hibridante es capaz de hibridar en condiciones de rigurosidad reducida, preferentemente en condiciones rigurosas, con uno de los ácidos nucleicos proporcionados en la tabla A o una parte de los mismos tal como se define en el presente documento. Más preferentemente, la porción es una porción de un ácido nucleico representado por SEC ID N°: 1.
- 10 Por lo tanto, la divulgación proporciona un procedimiento para mejorar las características de crecimiento de plantas que comprende modular la expresión en una planta de un ácido nucleico capaz de hibridar en condiciones de rigurosidad reducida, preferentemente en condiciones rigurosas, a un ácido nucleico que codifica un polipéptido de *hox5* de cremallera de leucina de homeodominio (HDZip) de clase I o un homólogo del mismo; o que comprende modular la expresión en una planta de un ácido nucleico capaz de hibridar a rigurosidad reducida.
- 15 Los ácidos nucleicos o variantes de los mismos pueden proceder de cualquier fuente natural o artificial. El ácido nucleico/gen o variante del mismo puede aislarse a partir de una fuente microbiana, tal como una levadura u hongo, o de una fuente de planta, de alga o de animal (incluyendo seres humanos). Este ácido nucleico puede modificarse a partir de su forma nativa en cuanto a su composición y/o ambiente genómico mediante manipulación humana deliberada. El ácido nucleico es preferentemente de origen vegetal, ya sea de la misma especie de planta (por ejemplo, de aquella en la que se va a introducir) o sea de una especie de planta diferente. El ácido nucleico puede aislarse a partir de una especie de monocotiledónea, preferentemente de la familia *Poaceae*, más preferentemente del género *Oryza*, lo más preferentemente de *Oryza sativa*. El ácido nucleico puede aislarse a partir de una especie de dicotiledónea, preferentemente de la familia *Brassicaceae*, más preferentemente de *Arabidopsis thaliana*.
- 20 La expresión de un ácido nucleico puede modularse introduciendo una modificación genética (preferentemente en el locus del gen en cuestión). Tal como se define en el presente documento, el locus de un gen se entiende como una región genómica, que incluye el gen de interés y 10 kb cadena arriba o cadena abajo de la región codificante.
- 25 Puede introducirse la modificación genética, por ejemplo, mediante uno cualquiera (o más) de los siguientes procedimientos: activación de ADN-T, TILLING, mutagénesis de sitio dirigido, evolución dirigida y recombinación de homólogos o introduciendo y expresando en una planta un ácido nucleico que codifica un polipéptido de *hox5* de HDZip de clase I o un homólogo del mismo. Después de la introducción de la modificación genética, sigue una etapa de selección para la expresión modulada del ácido nucleico, dando dicha modulación de la expresión plantas que tienen características de crecimiento mejoradas, en particular, rendimiento aumentado.
- 30 El marcado de activación de ADN-T (Hayashi y col., Science (1992) 1350-1353) implica la inserción de ADN-T, 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 cadena arriba o cadena abajo de la región codificante de un gen en una configuración tal que el promotor dirige la expresión del gen diana. Normalmente, la regulación de la expresión del gen diana por su promotor natural se altera y el gen queda bajo el control del nuevo promotor introducido. El promotor se incluye típicamente en un ADN-T. Este ADN-T se inserta al azar en el genoma de la planta, por ejemplo, mediante infección con *Agrobacterium* y da lugar a la sobreexpresión de genes próximos al ADN-T insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debido a la sobreexpresión de genes próximos al promotor introducido. El promotor que va a introducirse 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 y preferidos de tipo de célula e inducibles son todos adecuados para su uso en la activación de ADN-T.
- 35 El marcado de activación de ADN-T (Hayashi y col., Science (1992) 1350-1353) implica la inserción de ADN-T, 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 cadena arriba o cadena abajo de la región codificante de un gen en una configuración tal que el promotor dirige la expresión del gen diana. Normalmente, la regulación de la expresión del gen diana por su promotor natural se altera y el gen queda bajo el control del nuevo promotor introducido. El promotor se incluye típicamente en un ADN-T. Este ADN-T se inserta al azar en el genoma de la planta, por ejemplo, mediante infección con *Agrobacterium* y da lugar a la sobreexpresión de genes próximos al ADN-T insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debido a la sobreexpresión de genes próximos al promotor introducido. El promotor que va a introducirse 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 y preferidos de tipo de célula e inducibles son todos adecuados para su uso en la activación de ADN-T.
- 40 También puede introducirse una modificación genética en el locus del gen en cuestión usando la técnica del TILLING (*Targeted Induced Local Lesions in Genomes*, por sus siglas en inglés, Inducción Dirigida de Lesiones Locales en el Genoma). Esta es una tecnología de mutagénesis útil para generar y/o identificar y aislar variantes mutadas. El TILLING también permite la selección de plantas que portan dichas variantes mutantes. Estas variantes mutantes pueden incluso mostrar mayor actividad que la mostrada por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de exploración de alto rendimiento. Las etapas seguidas típicamente 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. Singapur, World Scientific Publishing Co, págs. 16-82; Feldmann y col., (1994) en Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, págs. 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, págs. 91-104); (b) preparación de ADN y agrupamiento de individuos; (c) Amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) DHPLC, donde se detecta la presencia de un heterodúplex en un grupo como un pico adicional en el cromatograma; (f) identificación del individuo mutante; y (g) secuenciación del producto de PCR mutante. Los procedimientos para el TILLING se conocen bien en la técnica (McCallum y col., (2000) *Nat Biotechnol* 18: 455-457; revisado por Stemple (2004) *Nat Rev Genet* 5(2): 145-50).
- 45 También puede introducirse una modificación genética en el locus del gen en cuestión usando la técnica del TILLING (*Targeted Induced Local Lesions in Genomes*, por sus siglas en inglés, Inducción Dirigida de Lesiones Locales en el Genoma). Esta es una tecnología de mutagénesis útil para generar y/o identificar y aislar variantes mutadas. El TILLING también permite la selección de plantas que portan dichas variantes mutantes. Estas variantes mutantes pueden incluso mostrar mayor actividad que la mostrada por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de exploración de alto rendimiento. Las etapas seguidas típicamente 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. Singapur, World Scientific Publishing Co, págs. 16-82; Feldmann y col., (1994) en Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, págs. 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, págs. 91-104); (b) preparación de ADN y agrupamiento de individuos; (c) Amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) DHPLC, donde se detecta la presencia de un heterodúplex en un grupo como un pico adicional en el cromatograma; (f) identificación del individuo mutante; y (g) secuenciación del producto de PCR mutante. Los procedimientos para el TILLING se conocen bien en la técnica (McCallum y col., (2000) *Nat Biotechnol* 18: 455-457; revisado por Stemple (2004) *Nat Rev Genet* 5(2): 145-50).
- 50 También puede introducirse una modificación genética en el locus del gen en cuestión usando la técnica del TILLING (*Targeted Induced Local Lesions in Genomes*, por sus siglas en inglés, Inducción Dirigida de Lesiones Locales en el Genoma). Esta es una tecnología de mutagénesis útil para generar y/o identificar y aislar variantes mutadas. El TILLING también permite la selección de plantas que portan dichas variantes mutantes. Estas variantes mutantes pueden incluso mostrar mayor actividad que la mostrada por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de exploración de alto rendimiento. Las etapas seguidas típicamente 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. Singapur, World Scientific Publishing Co, págs. 16-82; Feldmann y col., (1994) en Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, págs. 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, págs. 91-104); (b) preparación de ADN y agrupamiento de individuos; (c) Amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) DHPLC, donde se detecta la presencia de un heterodúplex en un grupo como un pico adicional en el cromatograma; (f) identificación del individuo mutante; y (g) secuenciación del producto de PCR mutante. Los procedimientos para el TILLING se conocen bien en la técnica (McCallum y col., (2000) *Nat Biotechnol* 18: 455-457; revisado por Stemple (2004) *Nat Rev Genet* 5(2): 145-50).
- 55 También puede introducirse una modificación genética en el locus del gen en cuestión usando la técnica del TILLING (*Targeted Induced Local Lesions in Genomes*, por sus siglas en inglés, Inducción Dirigida de Lesiones Locales en el Genoma). Esta es una tecnología de mutagénesis útil para generar y/o identificar y aislar variantes mutadas. El TILLING también permite la selección de plantas que portan dichas variantes mutantes. Estas variantes mutantes pueden incluso mostrar mayor actividad que la mostrada por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de exploración de alto rendimiento. Las etapas seguidas típicamente 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. Singapur, World Scientific Publishing Co, págs. 16-82; Feldmann y col., (1994) en Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, págs. 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, págs. 91-104); (b) preparación de ADN y agrupamiento de individuos; (c) Amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) DHPLC, donde se detecta la presencia de un heterodúplex en un grupo como un pico adicional en el cromatograma; (f) identificación del individuo mutante; y (g) secuenciación del producto de PCR mutante. Los procedimientos para el TILLING se conocen bien en la técnica (McCallum y col., (2000) *Nat Biotechnol* 18: 455-457; revisado por Stemple (2004) *Nat Rev Genet* 5(2): 145-50).

Puede usarse la mutagénesis de sitio dirigido para generar ácidos nucleicos variantes. Hay disponibles varios procedimientos para lograr la mutagénesis de sitio dirigido, siendo los más comunes los procedimientos basados en la PCR (Current Protocols in Molecular Biology, Wiley Eds <http://www.4ulr.com/products/currentprotocols/index.html>).

- 5 También puede usarse evolución dirigida o reordenamiento génico para generar variantes de ácido nucleico. Esto consiste en iteraciones de reordenamiento de ADN seguidas de la exploración y/o selección adecuada para generar variantes que tienen una actividad biológica modificada (Castle y col., (2004) Science 304(5674): 1151-4; Patentes de Estados Unidos N° 5.811.238 y 6.395.547).

La activación de ADN-T, el TILLING, la mutagénesis de sitio dirigido y la evolución dirigida son ejemplos de tecnologías que permiten la generación de nuevos alelos y variantes.

- 10 La recombinación de homólogos permite la introducción en un genoma de un ácido nucleico seleccionado en una posición seleccionada definida. La recombinación de homólogos es una tecnología convencional usada de manera rutinaria en las ciencias biológicas para organismos inferiores, tales como levaduras o el musgo *Physcomitrella*. Los procedimientos para llevar a cabo la recombinación de homólogos en plantas se han descrito no solo para plantas modelo (Offringa y col., (1990) EMBO J 9(10): 3077-84) sino también para plantas de cultivo, por ejemplo, arroz (Terada y col., (2002) Nat Biotech 20(10): 1030-4; Iida y Terada (2004) Curr Opin Biotech 15(2): 132-8). El ácido nucleico que se va a dirigir no necesita dirigirse al locus del gen en cuestión, sino que puede introducirse en, por ejemplo, regiones de alta expresión. El ácido nucleico que se va a dirigir puede ser un alelo mejorado usando para reemplazar al gen endógeno o puede introducirse además del gen endógeno.

- 20 Un procedimiento preferido para introducir una modificación genética es introducir y expresar en una planta un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo. El ácido nucleico que se va a introducir en una planta puede ser un ácido nucleico de longitud completa o puede ser una porción o una secuencia hibridante, tal como se ha definido anteriormente en el presente documento.

- 25 Los "homólogos" de una proteína abarcan péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones, eliminaciones y/o inserciones de aminoácidos en relación a la proteína en cuestión sin modificar y que tienen actividad biológica y funcional similar a la de la proteína no modificada de la que proceden. Para producir dichos homólogos, pueden reemplazarse los aminoácidos de la proteína por otros aminoácidos que tengan propiedades similares (tales como una similar hidrofobicidad, hidrofiliadad, antigenicidad, propensión a formar o romper estructuras  $\alpha$ -helicoidales o estructuras  $\beta$ -laminares). Las tablas de sustituciones conservativas se conocen bien en la técnica (véase, por ejemplo, Creighton (1984) Proteins. W.H. Freeman and Company y la tabla 4 a continuación). Los
- 30 homólogos útiles en los procedimientos de acuerdo con la divulgación son preferentemente polipéptidos que tienen en orden creciente de preferencia al menos un 30 %, 40 % 50 % 60 % 70 % 80 % 85 % 90 % 95 % 96 % 97 % 98 % o 99 % de identidad o similitud de secuencia (identidad funcional) con las secuencias mencionadas en el presente documento como útiles en los procedimientos de la divulgación, por ejemplo, las secuencias proporcionadas en la tabla A en el presente documento.

- 35 También se encuentran abarcadas por el término "homólogos" dos formas especiales de homología que incluyen secuencias ortólogas y secuencias parálogas, que abarcan conceptos evolutivos usados para describir relaciones ancestrales de genes. El término "parálogo" se refiere a duplicaciones génicas en el genoma de una especie que dan lugar a genes parálogos. El término "ortólogo" se refiere a genes homólogos en diferentes organismos debido a la especiación.

- 40 Los ortólogos en, por ejemplo, especies de plantas monocotiledóneas pueden encontrarse fácilmente efectuando una búsqueda blast denominada recíproca. Tomando el ejemplo de hallar un ortólogo o parálogo de un ácido nucleico de hox5 de HDZip de clase I o de un polipéptido de hox5 de HDZip de clase I, esto puede efectuarse mediante un primer BLAST que implica efectuar una búsqueda BLAST de 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 pública del NCBI que puede
- 45 encontrarse en <http://www.ncbi.nlm.nih.gov>. Pueden usarse BLASTN o TBLASTX cuando se parte de una secuencia nucleotídica, o BLASTP o TBLASTN cuando se parte de la proteína, con valores por defecto convencionales. Los resultados del BLAST pueden filtrarse. Las secuencias de longitud completa de los resultados filtrados o de los resultados no filtrados vuelven a usarse para un BLAST (segundo BLAST) contra las secuencias del organismo del que procede la secuencia en cuestión. Los resultados del primer y del segundo BLAST se comparan entonces.
- 50 Cuando los resultados del segundo BLAST proporcionan como coincidencias con la máxima similitud un ácido nucleico de hox5 de HDZip de clase I o un polipéptido de hox5 de HDZip de clase I, se ha hallado un parálogo, si tiene su origen en el mismo organismo que para la secuencia usada en el primer BLAST. En el caso de que tenga su origen en un organismo distinto al de la secuencia usada en el primer BLAST, se ha hallado un ortólogo. En el caso de grandes familias, puede usarse ClustalW, seguido de un árbol de unión del vecino más próximo, para ayudar a
- 55 visualizar el agrupamiento.

Un homólogo puede estar en forma de una "variante sustitucional" de una proteína, es decir, donde al menos un resto en una secuencia de aminoácidos se ha eliminado y se ha insertado un resto diferente en su lugar. Las sustituciones de aminoácidos son típicamente de restos individuales, pero pueden agruparse dependiendo de las restricciones funcionales dispuestas sobre el polipéptido; las inserciones estarán en el orden de aproximadamente 1 a 10 restos de

aminoácidos. Preferentemente, las sustituciones de aminoácidos comprenden sustituciones conservativas de aminoácidos. Las tablas de sustituciones conservativas están fácilmente disponibles en la técnica. La tabla a continuación proporciona ejemplos de sustituciones de aminoácidos conservadas.

**Tabla 4:** Ejemplos de sustituciones de aminoácidos conservados

| Resto | Sustituciones conservativas | Resto | Sustituciones conservativas |
|-------|-----------------------------|-------|-----------------------------|
| Ala   | Ser                         | Leu   | Ile; Val                    |
| Arg   | Lys                         | Lys   | Arg; Gln                    |
| Asn   | Gln; His                    | Met   | Leu; Ile                    |
| Asp   | Glu                         | Phe   | Met; Leu; Tyr               |
| Gln   | Asn                         | Ser   | Thr; Gly                    |
| Cys   | Ser                         | Thr   | Ser; Val                    |
| Glu   | Asp                         | Trp   | Tyr                         |
| Gly   | Pro                         | Tyr   | Trp; Phe                    |
| His   | Asn; Gln                    | Val   | Ile; Leu                    |
| Ile   | Leu, Val                    |       |                             |

5

Un homólogo también puede estar en forma de una "variante insercional" de una proteína, es decir, donde se introducen uno o más restos de aminoácidos en un sitio predeterminado en una proteína. Las inserciones pueden comprender fusiones en N-terminal y/o C-terminal así como inserciones intrasecuencia de aminoácidos individuales o múltiples. En general, las inserciones dentro de la secuencia de aminoácidos serán menores que las fusiones en N- o C-terminal, del orden de aproximadamente 1 a 10 restos. Los ejemplos de proteínas de fusión en N- o C-terminal de péptidos incluyen el dominio de unión o el dominio de activación de un activación transcripcional tal como se usa en el sistema de dos híbridos de levadura, proteínas de envuelta de fago, marcador 6-(histidina), marcador de glutatión S-transferasa, proteína A, proteína de unión a maltosa, dihidrofolato reductasa, epítipo Tag 100, epítipo c-myc, epítipo FLAG®, lacZ, CMP (péptido de unión a calmodulina), epítipo HA, epítipo de proteína C y epítipo de VSV.

10

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

Las variantes de aminoácidos de una proteína pueden producirse fácilmente usando técnicas de síntesis peptídica bien conocidas en la técnica, tales como la síntesis peptídica en fase sólida y similares, o mediante manipulaciones de ADN recombinante. Los procedimientos para la manipulación de secuencias de ADN para producir variantes de sustitución, inserción o eliminación de una proteína se conocen bien en la técnica. Por ejemplo, las técnicas para producir mutaciones por sustitución en sitios predeterminados en el ADN se conocen bien por los expertos en la materia e incluyen mutagénesis M13, mutagénesis *in vitro* T7-Gen (USB, Cleveland, OH), mutagénesis de sitio dirigido QuickChange (Stratagene, San Diego, CA), mutagénesis de sitio dirigido mediada por la PCR u otros protocolos de mutagénesis de sitio dirigido.

20

25 El polipéptido de hox5 de HDZip de clase I o el homólogo del mismo pueden ser un derivado. Los "derivados" incluyen péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que pueden comprender sustituciones, eliminaciones o adicionales de restos de aminoácidos de origen natural y no natural en comparación con la secuencia de aminoácidos de una forma de origen natural de la proteína. Los derivados pueden comprender restos de aminoácidos de origen natural alterados, glucosilados, acilados, prenilados o de origen no natural en comparación con la secuencia de aminoácidos de una forma de origen natural del polipéptido. Un derivado también puede comprender uno o más sustituyentes no aminoacídicos en comparación con la secuencia de aminoácidos de la que procede, por ejemplo, una molécula indicadora u otro ligando, unido de manera covalente o no covalente a la secuencia de aminoácidos, tal como una molécula indicadora que está unida para facilitar su detección, y restos de aminoácido de origen no natural en relación a la secuencia aminoacídica de la proteína de origen natural.

30

35 El polipéptido de hox5 de HDZip de clase I o el homólogo del mismo pueden estar codificados por una variante de corte y empalme alternativo.

En los casos donde la secuencia útil en los procedimientos de la divulgación sea un ácido nucleico que codifica a un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo, la variante de corte y empalme codifica un polipéptido que comprende desde N-terminal hasta C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Además, el polipéptido de hox5 de HDZip de clase I o un homólogo del mismo puede comprender cualquiera o ambos de los siguientes: (a) una cola de Trp; y (b) el motivo de

40

aminoácidos RPFF, donde R es Arg, P es Pro y F es Phe. Dentro de este motivo, se permiten uno o más cambios conservativos en cualquier posición, y uno o dos cambios no conservativos en cualquier posición. El motivo de (b) precede a la caja ácida, cuando se examina la proteína desde N-terminal hasta C-terminal. Además se prefieren variantes de corte y empalme de secuencias de ácido nucleico proporcionadas en la tabla A. Lo más preferido es una variante de corte y empalme de una secuencia de ácido nucleico como la representada por la SEC ID N°: 1.

Por lo tanto, la presente divulgación proporciona un procedimiento para mejorar las características de crecimiento de plantas que comprende modular la expresión en una planta de una variante de corte y empalme alternativo de un ácido nucleico que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) o un homólogo del mismo.

La expresión "variante de corte y empalme alternativo", tal como se usa en el presente documento, abarca variantes de una secuencia de ácido nucleico en la que se han cortado, reemplazado o añadido intrones y/o exones seleccionados, o en la que se han acortado o alargado intrones. Dichas variantes serán aquellas en las que se mantenga la actividad biológica de la proteína, lo que puede lograrse manteniendo de manera selectiva segmentos funcionales de la proteína. Dichas variantes de corte y empalme pueden encontrarse en la naturaleza o pueden producirse por el hombre. Los procedimientos para producir dichas variantes de corte y empalme se conocen bien en la técnica.

El homólogo también puede estar codificado por una variante alélica de un ácido nucleico que codifica un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo.

En los casos donde la secuencia útil en los procedimientos de la divulgación sea un ácido nucleico que codifica a un polipéptido de hox5 de HDZip de clase I o un homólogo del mismo, la variante alélica codifica un polipéptido que comprende desde el N-terminal hasta el C-terminal: (i) una caja ácida; y (ii) un homeodominio de clase I; y (iii) una cremallera de leucina con más de 5 héptadas. Además, el polipéptido de hox5 de HDZip de clase I, o un homólogo del mismo, puede comprender cualquiera o ambos de los siguientes: (a) una cola de Trp; y (b) el motivo de aminoácidos RPFF, donde R es Arg, P es Pro y F es Phe. Dentro de este motivo, se permiten uno o más cambios conservativos en cualquier posición, y uno o dos cambios no conservativos en cualquier posición. El motivo de (b) precede a la caja ácida, cuando se examina la proteína desde el N-terminal hasta el C-terminal. Además se prefieren variantes alélicas de secuencias de ácido nucleico proporcionadas en la tabla A. Lo más preferido es una variante alélica de una secuencia de ácido nucleico como la representada por la SEC ID N°: 1.

Por lo tanto, la presente divulgación proporciona un procedimiento para mejorar características de crecimiento de plantas que comprende modular la expresión en una planta o una variante alélica de un ácido nucleico que codifica un polipéptido de hox5 de homeodominio de cremallera de leucina (HDZip) de clase I o un homólogo del mismo.

Las variantes alélicas existen en la naturaleza, y el uso de estos alelos naturales está abarcado en los procedimientos de la presente divulgación. Las variantes alélicas abarcan polimorfismos de nucleótidos individuales (SNP), así como polimorfismos de inserción/eliminación pequeña (INDEL). El tamaño de los INDEL es normalmente menor de 100 pb. Los SNP y los INDEL forman el mayor conjunto de variantes de secuencias en cepas polimórficas de origen natural de la mayoría de organismos.

De acuerdo con la presente divulgación, se contempla la expresión modulada del ácido nucleico o variante del mismo. En la técnica se documentan bien procedimientos para aumentar la expresión de genes o productos génicos e incluyen, por ejemplo, la sobreexpresión conducida por promotores apropiados, el uso de potenciadores de la transcripción o potenciadores de la traducción. Los ácidos nucleicos aislados que sirven como promotores o elementos potenciadores pueden introducirse en una posición apropiada (típicamente cadena arriba) de una forma no heteróloga de un polinucleótido para regular positivamente la expresión del ácido nucleico o variante del mismo. Por ejemplo, pueden alterarse promotores in vivo por mutación, delección, y/o sustitución (véase, Kmiec, la Patente de Estados Unidos No. 5,565,350; Zarling y col., PCT/US93/03868), o pueden introducirse promotores aislados en una célula de planta en la orientación y distancia apropiada de un gen de la presente divulgación para controlar la expresión del gen.

En la técnica se documentan bien procedimientos para reducir la expresión de genes o productos génicos.

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

También puede añadirse una secuencia intrónica a la región no traducida (UTR, *untranslated region*) 5' o la secuencia codificante de la secuencia codificante 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 1000 veces (Buchman y Berg (1988) Mol. Cell Biol. 8: 4395-4405; Callis y col., (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. Para una información general véase: The Maize Handbook, Capítulo 116, Freeling y

Walbot, Eds., Springer, N. Y. (1994).

La divulgación también proporciona construcciones génicas y vectores para facilitar la introducción y/o expresión de las secuencias de nucleótidos útiles en los procedimientos de acuerdo con la divulgación.

Por lo tanto, se desvela una construcción génica que comprende:

- 5 (i) un ácido nucleico de hox5 de HDZip de clase I o una variante del mismo;
- (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 la transcripción.

10 Usando tecnología de ADN recombinante, bien conocida por los expertos en la técnica, pueden construirse construcciones útiles en los procedimientos de acuerdo con la presente invención. Las construcciones génicas pueden insertarse en vectores, que se encuentran disponibles en el comercio, adecuados para la transformación en plantas y adecuados para la expresión del gen de interés en las células transformadas. Por lo tanto, la invención proporciona el uso de una construcción génica, como se define anteriormente en el presente documento, en los procedimientos de la invención.

15 Las plantas se transforman con un vector que comprende la secuencia de interés. La secuencia de interés está unida operativamente a una o más secuencias de control (al menos a un promotor). Las expresiones "elemento regulador", "secuencia de control" y "promotor" se usan todas indistintamente en el presente documento y se toman en un contexto amplio para referirse a secuencias de ácido nucleico reguladoras capaces de efectuar la expresión de las secuencias a las que se unen. En los términos anteriormente mencionados se incluyen las secuencias reguladoras transcripcionales procedentes de un gen genómico eucariota clásico (incluyendo la caja TATA que es necesaria para el inicio de la transcripción adecuado, con o sin secuencia de caja CCAAT) y elementos reguladores adicionales (es decir, secuencias activadoras cadena arriba, potenciadores y silenciadores) que alteran la expresión génica en respuesta al desarrollo y/o estímulo externo, o de una manera específica de tejido. También se incluye en el término una secuencia reguladora transcripcional 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 ácido nucleico en una célula, tejido u órgano. La expresión "unido operativamente" 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 puede iniciar la transcripción del gen de interés.

30 Ventajosamente, para dirigir la expresión de la secuencia del ácido nucleico puede usarse cualquier tipo de promotor. El promotor puede ser un promotor inducible, es decir, que tiene un inicio de la transcripción inducido o aumentado en respuesta a un estímulo evolutivo, químico, ambiental o físico. Siendo un ejemplo de un promotor inducible un promotor inducible por estrés, es decir, un promotor activado cuando una planta se expone a diversas condiciones de estrés. Además, o como alternativa, el promotor, que puede ser un promotor preferido de tejido, es uno que puede iniciar preferentemente la transcripción en determinados tejidos, tales como hojas, raíces, tejido de semillas, etc. En el presente documento, los promotores capaces de iniciar la transcripción solo en determinados tejidos reciben el nombre de "específicos de tejido".

40 Preferentemente, el ácido nucleico de hox5 de HDZip de clase I o variante del mismo, está unido operativamente a un promotor constitutivo. Un promotor constitutivo es transcripcionalmente activo durante la mayor parte de, pero no necesariamente en todas, las fases del crecimiento y desarrollo y se expresa sustancialmente de manera ubicua. El promotor constitutivo es preferentemente un promotor de GOS2, más preferentemente el promotor constitutivo es un promotor de GOS2 de arroz, adicionalmente de manera preferida el promotor constitutivo está representado por una secuencia de ácido nucleico sustancialmente similar a la SEC ID N°: 33, más preferentemente el promotor constitutivo es como se representa en la SEC ID N°: 33. Debe aclararse que la aplicabilidad de la presente invención no está limitada al ácido nucleico de hox5 de HDZip de clase I o a una variante del mismo. En la Tabla 5, más adelante, se muestran ejemplos de otros promotores constitutivos que también pueden usarse para realizar los procedimientos de la invención.

**Tabla 5:** Ejemplos de promotores constitutivos:

| Fuente del Gen | Referencia                                                     |
|----------------|----------------------------------------------------------------|
| Actina         | McElroy y col, Plant Cell, 2: 163-171, 1990                    |
| CAMV 35S       | Odell y col., Nature, 313: 810-812, 1985                       |
| CaMV 19S       | Nilsson y col., Physiol. Plant. 100:456-462, 1997              |
| GOS2           | de Pater y col, Plant J Nov;2(6):837-44, 1992, WO 2004/065596, |
| Ubiquitina     | Christensen y col, Plant Mol. Biol. 18: 675-689, 1992          |

(continuación)

| Fuente del Gen            | Referencia                                             |
|---------------------------|--------------------------------------------------------|
| Ciclofilina de arroz      | Buchholz y col., Plant Mol Biol. 25(5): 837-43, 1994   |
| Histona H3 de maíz        | Lepetit y col., Mol. Gen. Genet. 231:276-285, 1992     |
| Histona H3 de alfalfa     | Wu y col. Plant Mol. Biol. 11:641-649, 1988            |
| Actina 2                  | An y col. Plant J. 10:9411; 107-121, 1996              |
| 34S FMV                   | Sanger y col., Plant. Mol. Biol., 14, 1990: 433-443    |
| Subunidad pequeña Rubisco | US 4.962.028;                                          |
| OCS                       | Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553      |
| SAD1                      | Jain y col., Crop Science, 39 (6), 1999: 1696          |
| SAD2                      | Jain y col., Crop Science, 39 (6), 1999: 1696          |
| Nos                       | Shaw y col. (1984) Nucleic Acids Res. 12(20):7831-7846 |
| V-ATPasa                  | WO 01/14572,                                           |
| Súper promotor            | WO 95/14098,                                           |
| Proteínas de la caja G    | WO 94/12015                                            |

Un promotor específico de semilla es uno que es transcripcionalmente activo predominantemente en tejidos de semilla. Un promotor específico de semilla también puede tener alguna actividad residual en otras partes. Preferentemente, el promotor específico de semilla es un promotor específico de endospermo y/o de aleurona, lo que significa que el promotor es transcripcionalmente activo predominantemente en tejido de endospermo y/o en las capas de aleurona de una semilla, aunque puede haber alguna actividad residual o expresión débil en alguna otra parte.

5

**Tabla 6:** Ejemplos de promotores específicos de semilla

| Fuente del Gen                            | Patrón de expresión | Referencia                                                                                                                                             |
|-------------------------------------------|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| genes específicos de semilla              | semilla             | Simon, y col., Plant Mol. Biol. 5: 191, 1985; Scofield, y col., J. Biol. Chem. 262: 12202, 1987.; Baszczynski, y col., Plant Mol. Biol. 14: 633, 1990. |
| albúmina de Nuez del Brasil               | semilla             | Pearson, y col., Plant Mol. Biol. 18: 235-245, 1992.                                                                                                   |
| Legumina                                  | semilla             | Ellis, y col., Plant Mol. Biol. 10: 203-214, 1988.                                                                                                     |
| Glutelina (arroz)                         | semilla             | Takaiwa, y col., Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa, y col., FEBS Lett. 221: 43-47, 1987.                                                      |
| Zeína                                     | semilla             | Matzke y col Plant Mol Biol, 14(3):323-32 (1990)                                                                                                       |
| napA                                      | semilla             | Stalberg, y col, Planta 199: 515-519, 1996.                                                                                                            |
| Gluteína-1 de BPM y APM de trigo          | endospermo          | Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989                                                                                                      |
| SPA de trigo                              | semilla             | Albani y col, Plant Cell, 9: 171-184, 1997                                                                                                             |
| $\alpha$ , $\beta$ , y gliadinas de trigo | endospermo          | EMBO 3:1409-15, 1984                                                                                                                                   |
| Promotor ltr1 de cebada                   | endospermo          |                                                                                                                                                        |
| Hordeína B1, C, D de cebada               | endospermo          | Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996                                                                |
| DOF de cebada                             | endospermo          | Mena y col, The Plant Journal, 116(1): 53-62, 1998                                                                                                     |

(continuación)

| Fuente del Gen                           | Patrón de expresión              | Referencia                                                  |
|------------------------------------------|----------------------------------|-------------------------------------------------------------|
| promotor sintético                       | endospermo                       | Vicente-Carbajosa y col., Plant J. 13: 629-640, 1998.       |
| Prolamina NRP33 de arroz                 | endospermo                       | Wu y col., Plant Cell Physiology 39(8) 885-889, 1998        |
| □ globulina Glb-1 de arroz               | endospermo                       | Wu y col., Plant Cell Physiology                            |
|                                          |                                  | 39(8) 885-889, 1998                                         |
| OSH1 de arroz                            | embrión                          | Sato y col, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996 |
| D globulina REB/OHP-1 de arroz           | endospermo                       | Nakase y col. Plant Mol. Biol. 33: 513-522, 1997            |
| ADP-glucosa PP de arroz                  | endospermo                       | Trans Res 6:157-68, 1997                                    |
| familia del gen ESR de maíz              | endospermo                       | Plant J 12:235-46, 1997                                     |
| □-kafirina de sorgo                      | endospermo                       | PMB 32:1029-35,1996                                         |
| KNOX                                     | embrión                          | Postma-Haarsma y col, Plant Mol. Biol. 39:257-71, 1999      |
| Oleosina de arroz                        | embrión y aleurona               | Wu y col, J. Biochem, 123:386, 1998                         |
| Oleosina de girasol                      | semilla (embrión y semilla seca) | Cummins, y col., Plant Mol. Biol. 19: 873-876, 1992         |
| Supuesta proteína ribosomal 40S de arroz | débil en endospermo              |                                                             |
| alfa-globulina de arroz                  | fuerte en endospermo             |                                                             |
| alanina aminotransferasa de arroz        | débil en endospermo              |                                                             |
| inhibidor de tripsina ITR1 (cebada)      | débil en endospermo              |                                                             |
| WSI18 de arroz                           | embrión + estrés                 |                                                             |
| RAB21 de arroz                           | embrión + estrés                 |                                                             |
| Oleosina de arroz de 18kd                | aleurona + embrión               |                                                             |

Para la identificación de promotores funcionalmente equivalentes, la fuerza del promotor y/o el patrón de expresión de un promotor candidato pueden analizarse, por ejemplo, uniendo operativamente el promotor a un gen indicador y evaluando el nivel de expresión y el patrón del gen indicador en diversos tejidos de la planta. Los genes indicadores bien conocidos adecuados incluyen, por ejemplo, beta-glucuronidasa o beta-galactosidasa. La actividad del promotor se evalúa midiendo la actividad enzimática de la beta-glucuronidasa o beta-galactosidasa. La fuerza del promotor y/o el patrón de expresión pueden después compararse con la de un promotor de referencia (tal como el que se usa en los procedimientos de la presente invención). Como alternativa, la fuerza del promotor puede evaluarse cuantificando niveles de ARNm o comparando niveles de ARNm del ácido nucleico usado en los procedimientos de la presente invención, con niveles de ARNm de genes constitutivos tal como ARNr 18S, usando procedimientos bien conocidos

en la técnica, tales como transferencia de Northern con análisis densitométrico de autorradiogramas, PCR cuantitativa en tiempo real o RT-PCR (Heid y col., 1996 *Genome Methods* 6: 986-994). Generalmente por "promotor débil" se entiende un promotor que dirige la expresión de una secuencia codificante a un nivel bajo. Por "nivel bajo" se entiende niveles de aproximadamente 1/10.000 transcritos a aproximadamente 1/100.000 transcritos, a aproximadamente 1/5000000 transcritos por célula. Por el contrario, un "promotor fuerte" dirige la expresión de una secuencia codificante a alto nivel, o a aproximadamente 1/10 transcritos a aproximadamente 1/100 transcritos a aproximadamente 1/1000 transcritos por célula.

Opcionalmente, en la construcción introducida en una planta, también pueden usarse unas o más secuencias terminadoras. El término "terminador" incluye una secuencia de control que es una secuencia de ADN en el extremo de una unidad transcripcional que señala el procesamiento y la poliadenilación en dirección 3' de un transcrito primario y la terminación de la transcripción. Otros elementos reguladores pueden incluir potenciadores transcripcionales así como traduccionales. Los expertos en la técnica conocerán secuencias terminadoras y potenciadoras que pueden ser adecuadas para su uso a la hora de realizar la presente invención. Un experto en la materia conocerá o sabrá cómo obtener dichas secuencias.

Las construcciones genéticas de la divulgación pueden incluir adicionalmente un origen de secuencia de replicación que es necesario para la conservación y/o replicación en un tipo de célula específico. Un ejemplo es cuando en una célula bacteriana se requiere conservar una construcción genética como un elemento genético episomal (por ejemplo, una molécula de plásmido o de cósmido). Como orígenes de replicación preferidos se incluyen, pero sin limitación, f1-ori y colE1.

La construcción genética puede comprender opcionalmente un gen marcador de selección. Tal como se usa en el presente documento, la expresión "gen marcador de selección" incluye cualquier gen que confiere un fenotipo en una célula en la que se exprese para facilitar la identificación y/o selección de células que se transfectan o transforman con una construcción de ácido nucleico de la divulgación.

Para la detección de la transferencia satisfactoria de las secuencias de ácidos nucleicos como se usa en los procedimientos de la invención y/o selección de plantas transgénicas que comprenden estos ácidos nucleicos, es ventajoso usar genes marcadores (o genes indicadores). Por lo tanto, la construcción genética puede comprender opcionalmente un gen marcador de selección. Tal como se usa en el presente documento, las expresiones "marcador de selección", "gen marcador de selección" o "gen indicador" incluyen cualquier gen que confiera un fenotipo en una célula en la que se exprese para facilitar la identificación y/o selección de células que se transfectan o transforman con una construcción de ácido nucleico de la divulgación. Estos genes marcadores permiten la identificación de una transferencia satisfactoria de las moléculas de ácido nucleico mediante una serie de principios diferentes. Los marcadores adecuados pueden seleccionarse de marcadores que confieren resistencia a antibióticos o a herbicidas, que introducen un nuevo rasgo metabólico o que permiten selección visual. Los ejemplos de genes marcadores de selección incluyen genes que confieren resistencia a antibióticos (tales como nptII que fosforila neomicina y kanamicina, o hpt, que fosforila higromicina, o genes que confieren resistencia a, por ejemplo, bleomicina, estreptomycin, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomycin o blasticidina), a herbicidas (por ejemplo bar que proporciona resistencia a Basta®; aroA o gox que proporciona resistencia contra glifosato, o los genes que confieren resistencia a, por ejemplo, imidazolinona, fosfinotricina o sulfonilurea), o genes que proporcionan un rasgo metabólico (tal como manA que permite a las plantas usar manosa como única fuente de carbono o xilosa isomerasa para la utilización de xilosa, o marcadores antinutritivos tales como la resistencia a 2-desoxiglucosa). La expresión de genes marcadores visuales da como resultado la formación de color (por ejemplo  $\beta$ -glucuronidasa, GUS o  $\beta$ -galactosidasa con sus sustratos de color, por ejemplo X Gal), luminiscencia (tal como el sistema luciferina/luciferasa) o fluorescencia (Proteína Fluorescente Verde, GFP y derivados de los mismos). Esta lista solo representa una pequeña cantidad de marcadores posibles. El experto está familiarizado con dichos marcadores. Se prefieren diferentes marcadores dependiendo del organismo y del procedimiento de selección.

Se sabe que, después de la integración estable o transitoria de los ácidos nucleicos en células de plantas, solo una minoría de las células capta el ADN exógeno y, si se desea, lo integra en su genoma, dependiendo del vector de expresión usado y de la técnica de transfección usada. Para identificar y seleccionar estos integrantes, normalmente se introduce un gen que codifica un marcador de selección (tal como los descritos anteriormente) en las células huésped junto con el gen de interés. Estos marcadores pueden usarse, por ejemplo, en mutantes en los que estos genes no son funcionales mediante, por ejemplo, delección por procedimientos convencionales. Además, las moléculas de ácido nucleico que codifican un marcador de selección pueden introducirse en una célula huésped en el mismo vector que comprende la secuencia que codifica los polipéptidos de la invención o usarse en los procedimientos de la invención, o incluso en un vector distinto. Las células que se han transfectado de manera estable con el ácido nucleico introducido pueden identificarse, por ejemplo, células que tienen integrado el marcador de selección sobreviven mientras que las otras mueren).

Dado que los genes marcadores, particularmente genes para resistencia a antibióticos y herbicidas, ya no se requieren o no se desean en la célula huésped transgénica una vez que se han introducido satisfactoriamente los ácidos nucleicos, el proceso de acuerdo con la invención para introducir los ácidos nucleicos emplea ventajosamente técnicas que permiten la retirada o escisión de estos genes marcadores. Un procedimiento de este tipo es lo que se conoce como co-transformación. El procedimiento de co-transformación emplea dos vectores simultáneamente para la

transformación, un vector que lleva el ácido nucleico de acuerdo con la invención y un segundo que lleva el gen (o genes) marcador. Una proporción grande de transformantes recibe o, en el caso de plantas, comprende (hasta el 40 % o más de los transformantes), ambos vectores. En el caso de transformación con *Agrobacteria*, los transformantes normalmente solo reciben una parte del vector, es decir, la secuencia flanqueada por el ADN-T, que normalmente representa el casete de expresión. Posteriormente los genes marcadores pueden retirarse de la planta transformada realizando cruzamientos. En otro procedimiento, se usan genes marcadores integrados en un transposón para la transformación junto con el ácido nucleico deseado (conocido como tecnología Ac/Ds). Los transformantes pueden cruzarse con una fuente transposasa o los transformantes se transforman con una construcción de ácido nucleico que confiere expresión de una transposasa, de manera transitoria o estable. En algunos casos (aprox. 10 %), el transposón salta del genoma de la célula huésped una vez que se ha producido la transformación satisfactoriamente y se pierde. En una serie de casos adicionales, el transposón salta a una localización diferente. En estos casos, el gen marcador debe eliminarse realizando cruzamientos. En microbiología, se desarrollan técnicas que hacen posible, o facilitan, la detección de dichos acontecimientos. Un procedimiento ventajoso adicional se basa en lo que se conoce como sistemas de recombinación; cuya ventaja es que la eliminación puede dispensarse por cruzamiento. El sistema mejor conocido de este tipo es lo que se conoce como sistema Cre/lox. Cre1 es una recombinasa que retira las secuencias localizadas entre las secuencias loxP. Si el gen marcador se integra entre las secuencias loxP, éste se elimina una vez se haya producido la transformación satisfactoriamente, por expresión de la recombinasa. Otros sistemas de recombinación son el sistema HIN/HIX, FLP/FRT y REP/STB (Tribble y col., J. Biol. Chem., 275, 2000: 22255-22267; Velmurugan y col., J. Cell Biol., 149, 2000: 553-566). De acuerdo con la invención, es posible una integración específica de sitio en el genoma de la planta de la secuencia de ácidos nucleicos. Naturalmente, estos procedimientos también pueden aplicarse a microorganismos tales como levaduras, hongos o bacterias.

La divulgación se refiere a una construcción génica que comprende:

- (i) un ácido nucleico de hox5 de HDZip de clase I o una variante del mismo;
- (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 la transcripción.

El promotor constitutivo es preferentemente un promotor de GOS2, más preferentemente el promotor constitutivo es el promotor de GOS2 de arroz, adicionalmente de manera preferida el promotor constitutivo está representado por una secuencia de ácido nucleico sustancialmente similar a la SEC ID N°: 33, más preferentemente el promotor constitutivo es como se representa en la SEC ID N°: 33. Adicionalmente la invención proporciona el uso de una construcción como se define en la reivindicación 7 en el aumento del rendimiento de una planta en plantas que crecen en condiciones sin estrés en comparación con las plantas de tipo silvestre correspondientes, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado.

La presente divulgación también incluye plantas obtenibles por los procedimientos de acuerdo con la presente invención. Por lo tanto, la presente divulgación proporciona plantas, partes de plantas (incluyendo células de plantas) obtenibles por los procedimientos de acuerdo con la presente divulgación, cuyas plantas o partes (incluyendo células) comprenden un transgén de ácido nucleico de hox5 de HDZip de clase I.

La invención también proporciona un procedimiento para la producción de plantas transgénicas que tienen características de crecimiento mejoradas, en particular, rendimiento aumentado, cada uno en relación a las plantas de tipo silvestre correspondientes u otras plantas de control, que comprende la introducción y expresión en una planta de cualquier ácido nucleico descrito en el presente documento que sea útil en los procedimientos de la divulgación.

Para los fines de la invención, "transgénico", "transgén" o "recombinante" significa con respecto a, por ejemplo, una secuencia de ácido nucleico, un casete de expresión, una construcción génica o un vector que comprende la secuencia de ácido nucleico o un organismo transformado con las secuencias de ácido nucleico, casetes o vectores de expresión, todas estas construcciones realizadas por procedimientos recombinantes en los que

- a) las secuencias de ácido nucleico que codifican las proteínas útiles en los procedimientos de la invención, o
- b) la secuencia (o secuencias) de control genético que se une operativamente con la secuencia de ácido nucleico de acuerdo con la divulgación, por ejemplo un promotor, o
- c) a) y b) no se localizan en su ambiente genético natural o se han modificado por procedimientos recombinantes, siendo posible que se produzca la modificación de, por ejemplo, una sustitución, adición, delección, inversión o inserción de uno o más restos de nucleótido. Se entiende que, ambiente genético natural, significa el locus genómico o cromosómico natural en la planta original o la presencia en una genoteca. En el caso de una genoteca, el ambiente genético natural de la secuencia de ácidos nucleicos se conserva preferentemente, al menos en parte. El ambiente flanquea la secuencia de ácidos nucleicos al menos en un lado y tiene una longitud de secuencia de al menos 50 pb, preferentemente de al menos 500 pb, en especial, preferentemente de al menos 1000 pb, más preferentemente de al menos 5000 pb. Un casete de expresión de origen natural - por ejemplo la combinación de origen natural del promotor natural de las secuencias de ácidos nucleicos con la secuencia de ácidos nucleicos correspondiente que codifica un polipéptido útil en los procedimientos de la presente invención, como se ha definido anteriormente - llega a ser un casete de expresión transgénica cuando esta casete de expresión se modifica mediante procedimientos sintéticos no naturales ("artificiales") tal como, por ejemplo, tratamiento

mutagénico. Se describen procedimientos adecuados, por ejemplo, en los documentos US 5.565.350 o WO 00/15815.

Más específicamente, la presente divulgación proporciona un procedimiento para la producción de plantas transgénicas que tienen características de crecimiento mejoradas (particularmente rendimiento aumentado) cuyo procedimiento comprende:

- (i) introducir y expresar en una planta, en una parte de la planta o en una célula de planta, un ácido nucleico de hox5 de HDZip de clase I o una variante del mismo; y
- (ii) cultivar la célula de planta en condiciones que promuevan el crecimiento y desarrollo de la planta.

El 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 de la presente divulgación, el ácido nucleico se introduce preferentemente en una planta por transformación.

El término "transformación" como se denomina en el presente documento incluye la transferencia de un polinucleótido exógeno en una célula hospedadora, independientemente del procedimiento usado para la transferencia. El tejido de planta capaz de propagación clonal posterior, mediante organogénesis o embriogénesis, puede transformarse con una construcción genética y a partir de la misma regenerarse una planta completa. El tejido particular seleccionado variará dependiendo de los sistemas de propagación clonales disponibles para, y mejor adaptados a, la especie particular que se va a transformar. Las dianas tisulares ejemplares incluyen discos foliares, polen, embriones, cotiledones, hipocotiledones, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, brotes axilares, y meristemas radiculares), e inducen tejido meristemático (por ejemplo, meristemo de cotiledón y meristemo de hipocotiledón). El polinucleótido puede introducirse de manera estable o transitoria en una célula huésped y puede mantenerse no integrado, por ejemplo, como un plásmido. Como alternativa, éste puede integrarse en el genoma huésped. La célula de la planta transformada resultante puede después usarse para regenerar una planta transformada de una manera conocida por los expertos en la técnica.

La transformación de especies plantas es actualmente una técnica muy habitual. Ventajosamente, cualquiera de los diversos procedimientos de transformación puede usarse para introducir el gen de interés en una célula antecesora adecuada. Los métodos de transformación también 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 usando virus o polen y microproyección. Pueden seleccionarse procedimientos entre el procedimiento de calcio/polietylenglicol para protoplastos (Krens, F.A. y col. (1982) *Nature* 296, 72-74; Negrutiu I y col. (1987) *Plant Mol Biol* 8: 363-373); electroporación de protoplastos (Shillito R. D. y col. (1985) *Bio/Technol* 3, 1099-1102); microinyección en el material de planta (Crossway A y col. (1986) *Mol. Gen Genet* 202: 179-185); bombardeo de partículas revestido con ADN o ARN (Klein TM y col., (1987) *Nature* 327: 70) infección con virus (no integrantes) y similares. Las plantas transgénicas de arroz, se producen preferentemente mediante transformación mediada por *Agrobacterium* usando cualquiera de los procedimientos bien conocidos para la transformación del arroz, tales como los descritos en cualquiera de las siguientes referencias: solicitud de patente europea publicada EP 1198985 A1, Aldemita y Hodges (*Planta* 199: 612-617, 1996); Chan y col. (*Plant Mol Biol* 22 (3): 491-506, 1993), Hiei y col. (*Plant J* 6 (2): 271-282, 1994. En el caso de la transformación del maíz, el procedimiento preferido es como se describe en Ishida y col. (*Nat. Biotechnol* 14(6): 745-50, 1996) o en Frame y col. (*Plant Physiol* 129(1): 13-22, 2002).

Generalmente, después de la transformación, las células o grupos de células de planta se seleccionan para detectar la presencia de uno o más marcadores que están codificados por genes que pueden expresarse en plantas co-transferidos con el gen de interés, después de lo cual el material transformado se regenera en una planta completa.

Después de la transferencia y regeneración de ADN, las plantas supuestamente transformadas pueden evaluarse, usando, por ejemplo, análisis de Southern, para determinar la presencia del gen de interés, número de copias y/o organización genómica. Como alternativa o adicionalmente, los niveles de expresión del ADN recién introducido pueden supervisarse usando análisis de Northern y/o Western, o PCR cuantitativa; siendo todas estas técnicas bien conocidas por los expertos habituales en la materia.

Las plantas transformadas generadas pueden propagarse mediante diversos medios, tales como mediante técnicas clásicas de propagación clonal o de reproducción. Por ejemplo, una primera generación (o T1) de plantas transformadas puede reproducirse asexualmente para dar transformantes de segunda generación (o T2) homocigotos y después las plantas T2 pueden propagarse adicionalmente a través de técnicas de reproducción clásicas.

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).

La presente divulgación se amplía claramente a cualquier célula de planta o planta producida mediante cualquiera de los procedimientos descritos en el presente documento, y a todas las partes de la planta y propágulos de la misma. La presente invención se amplía adicionalmente 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 la misma característica, o características, genotípicas y/o fenotípicas que las producidas por el precursor en los procedimientos de acuerdo con la divulgación.

5 La divulgación también incluye células huésped que comprenden un ácido nucleico de hox5 de HDZip de clase I aislado o una variante del mismo.

La divulgación también se amplía a partes cosechables de una planta, tales como, pero sin limitación, semillas, hojas, frutos, flores, tallos, rizomas, tubérculos y bulbos. Adicionalmente la divulgación se refiere a productos derivados de una parte cosechable de dicha planta, tales como gránulos deshidratados o polvos, aceite, grasa y ácidos grasos, almidón o proteínas.

10 La presente invención también incluye el uso de ácidos nucleicos de hox5 de HDZip de clase I y variantes de los mismos y el uso de polipéptidos de hox5 de HDZip de clase I como se expone en las reivindicaciones 8 y 9. Dichos usos se refieren a mejorar cualquiera de las características del crecimiento de las plantas como se expone en las reivindicaciones 8 y 9.

15 Los ácidos nucleicos de hox5 de HDZip de clase I o variantes de los mismos, o los polipéptidos de hox5 de HDZip de clase I u homólogos de los mismos pueden encontrar uso en programas de reproducción en los que se identifica un marcador de ADN, que puede estar unido genéticamente a uno de los ácidos nucleicos o variantes mencionados anteriormente. Los ácidos nucleicos o variantes, o polipéptidos u homólogos de los mismos pueden usarse para definir un marcador molecular. Este marcador de ADN o de proteína puede después usarse en programas de reproducción para seleccionar plantas que tengan rasgos potenciados relacionados con el rendimiento (tales como rendimiento aumentado). Los ácidos nucleicos o variantes son como se definen anteriormente.

20 Las variantes alélicas de un ácido nucleico/gen que codifica a hox5 de HDZip de clase I pueden encontrar uso en programas de reproducción asistida mediante marcador. Dichos programas de reproducción a veces requieren la introducción de variación alélicas 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. Después, se realiza la identificación de variantes alélicas, por ejemplo, por PCR. A esto le sigue una etapa de selección de variantes alélicas superiores de la secuencia en cuestión y que dará características de crecimiento mejoradas, tal como rendimiento aumentado. 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 cualquiera de los ácidos nucleicos /genes descritos en el presente documento que sean útiles en los procedimientos de la invención. El rendimiento de crecimiento puede supervisarse en un invernadero o en el campo. Etapas opcionales adicionales incluyen cruzamiento de plantas en el que se identificó la variante alélica superior con otra planta. Esto podría usarse, por ejemplo, para realizar una combinación de características fenotípicas interesantes.

35 Los ácidos nucleicos y variantes mencionados anteriormente también pueden usarse como sondas 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 los ácidos nucleicos o variantes requiere solo una secuencia de ácido nucleico de al menos 15 nucleótidos de longitud. Los ácidos nucleicos o variantes pueden usarse como marcadores de polimorfismos de longitud de fragmentos de restricción (PLFR). Con los ácidos nucleicos o variantes pueden explorarse transferencias de Southern (Sambrook J, Fritsch EF y Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de ADN genómico de planta digerido por enzimas de restricción. Los patrones de banda resultantes pueden después someterse a análisis genético usando programas informáticos tales como MapMaker (Lander y col. (1987) Genomics 1: 174-181) para construir un mapa genético. Además, los ácidos nucleicos pueden usarse para explorar 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 polimorfismos de ADN se observa y se usa para calcular la posición del ácido nucleico que codifica el polipéptido JMJC en el mapa genético previamente obtenido usando esta población (Botstein y col. (1980) Am. J. Hum. Genet. 32:314-331).

45 La producción y 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.

55 Las sondas de ácido nucleico también pueden usarse para mapeo físico (es decir, colocación de secuencias en mapas físicos; véase Hoheisel y col. In: Non-mammalian Genomic Analysis: A Practical Guide, Academic press 1996, páginas 319-346, y referencias citadas en su interior).

En otra realización, pueden usarse las sondas de ácidos nucleicos en mapeo directo de hibridación in situ con fluorescencia (HISF, Fluorescent In Situ Hybridization) (Trask (1991) Trends Genet. 7: 149-154). Aunque los

procedimientos actuales del mapeo HSF favorecen el uso de clones grandes (de varios kb a varios cientos de kb; véase Laan y col. (1995) *Genome Res.* 5: 13-20), mejoras en sensibilidad pueden permitir la realización de mapeo HSF usando sondas más cortas.

Usando los ácidos nucleicos pueden realizarse diversos procedimientos basados en amplificación de ácidos nucleicos para el mapeo genético y físico. Como ejemplos se incluyen la amplificación específica de alelo (Kazazian (1989) *J. Lab. Clin. Med.* 11: 95-96), el polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield y col. (1993) *Genomics* 16: 325-332), el ligamiento específico de alelo (Landegren y col. (1988) *Science* 241: 1077-1080), las reacciones de extensión de nucleótidos (Sokolov (1990) *Nucleic Acid Res.* 18: 3671), el Mapeo Híbrido por Radiación (Walter y col. (1997) *Nat. Genet.* 7: 22-28) y el Mapeo Happy (por las siglas en inglés mapping based on the analysis of approximately HAPloid DNA samples using the PolYmerase chain reaction) (Dear y Cook (1989) *Nucleic Acid Res.* 17: 6795-6807). Para estos procedimientos, se utiliza la secuencia de un ácido nucleico para diseñar y producir pares de cebadores para su uso en la reacción de amplificación o en reacciones de extensión con cebador. El diseño de dichos cebadores es muy 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 características de crecimiento mejoradas, como se expone en las reivindicaciones. Estas características de crecimiento mejoradas también pueden combinarse con otros rasgos ventajosos desde el punto de vista económico, tales como rasgos que potencian adicionalmente el rendimiento, la tolerancia a diversos estreses abióticos y bióticos, 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 Fig.1 muestra un alineamiento múltiple de homeodominios HDZip de clase I de diferentes fuentes de plantas, usando un programa de alineamiento múltiple VNTI AlignX, basado en un algoritmo ClustalW modificado (InforMax, Bethesda, MD, <http://www.infor-maxinc.com>), con parámetros determinados de penalización por apertura de hueco de 10 y una extensión de hueco de 0,05. Los aminoácidos invariantes L16, W48, F49, N51 y R53, del homeodominio, se indican en un cuadro vertical. Los aminoácidos preferidos A46 y W56 de HDZip de Clase I también se indican en un cuadro vertical. Las tres hélices necesarias para la unión al ADN se marcan como cajas negras encima del alineamiento. Las seis héptadas se separan con 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 se indica en un cuadro vertical.

La Fig. 2 muestra un alineamiento múltiple de polipéptidos de hox5 de HDZip de clase I de diversas plantas, usando un programa de alineamiento múltiple VNTI AlignX, basado en un algoritmo ClustalW modificado (InforMax, Bethesda, MD, <http://www.informaxinc.com>), con parámetros determinados de penalización por apertura de hueco de 10 y una extensión de hueco de 0,05). Los tres dominios principales caracterizados, desde el extremo N al extremo C, están fuertemente empaquetados y se identifican como la caja ácida, el homeodominio de clase I y la cremallera de leucina de seis héptadas. Además, la cola de Trp y el motivo de aminoácidos RPFF están ligeramente empaquetados.

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

La figura 4 detalla ejemplos de secuencias de hox5 de cremallera de leucina de homeodominio de clase I (HDZip). Se obtienen como resultado varias secuencias a partir de ensamblajes EST públicos (véase la tabla A), con una menor calidad de secuenciación. Como consecuencia, pueden esperarse algunas sustituciones de ácidos nucleicos. Los codones de inicio (ATG) y de parada delimitan las secuencias de ácido nucleico.

### **Ejemplos**

A menos que se afirme lo contrario, las técnicas de ADN recombinante se llevaron a cabo de acuerdo con protocolos convencionales 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 y col. (1994), *Current Protocols in Molecular Biology*, Current Protocols. Los materiales y procedimientos convencionales para el trabajo molecular con plantas se describen en *Plan Molecular Biology Labfase* (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (R.U.) y Blackwell Scientific Publications (R.U.).

### ***EJEMPLOS - Polipéptidos de hox5 de HDZip de clase I y secuencias codificantes***

#### ***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 de entre aquellas mantenidas en la base de datos Entrez Nucleotides en el National Center for Biotechnology Information (NCBI) usando herramientas de búsqueda de secuencias en bases de datos, tales como la Basic Local Alignment Tool (BLAST) (Altschul y col., (1990) J. Mol. Biol. 215:403-410; y Altschul y col. (1997) Nucleic Acids Res. 25:3389-3402). El programa se usó para encontrar regiones de similitud local entre secuencias comparando secuencias de ácido nucleico o polipeptídicas con bases de datos de secuencias y calculando la significación estadística de las coincidencias. Por ejemplo, el polipéptido codificado por el ácido nucleico de la SEC ID N°: 1 se usó para el algoritmo TBLASTN, con configuraciones por defecto y el filtro para ignorar el conjunto de secuencias de baja complejidad. La salida del análisis se visualizó mediante comparación por pares, y se clasificó de acuerdo con la puntuación de probabilidad (valor de E), donde la puntuación refleja la probabilidad de que suceda un alineamiento particular por casualidad (cuanto más bajo sea el valor de E, más significativa será la coincidencia). Además de los valores de E, las comparaciones también se puntuaron por el porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos idénticos (o de aminoácidos) entre las dos secuencias de ácido nucleico (o polipeptídicas) comparadas sobre una longitud concreta. En algunos casos, pueden ajustarse los parámetros por defecto para modificar la rigurosidad de la búsqueda. Por ejemplo, el valor de E puede aumentarse para mostrar coincidencias menos rigurosas. De este modo, pueden identificarse coincidencias cortas prácticamente exactas.

A continuación, en la tabla A se proporciona una lista de secuencias de ácidos nucleicos relacionadas con la secuencia de ácidos nucleicos de la SEC ID N°: 1.

Tabla A: Ejemplos de secuencias relacionadas con la secuencia de ácidos nucleicos de SEC ID N°: 1

| Nombre                                                                                                                                                                                                                                                                                                                                        | Número de referencia del NCBI del nucleótido | 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<br>DV024016                         | 5                        | 6                                   | <i>Zea mays</i>                                       |
| Zeama_hox16                                                                                                                                                                                                                                                                                                                                   | AY105265                                     | 7                        | 8                                   | <i>Zea mays</i>                                       |
| Sacof_hox5*                                                                                                                                                                                                                                                                                                                                   | CA088615<br>CA115362<br>CA142506             | 9                        | 10                                  | <i>Saccharum officinarum</i>                          |
| Sorbi_hox5*                                                                                                                                                                                                                                                                                                                                   | BE363386<br>CD432381                         | 11                       | 12                                  | <i>Sorghum bicolor</i>                                |
| Triae_hox16*                                                                                                                                                                                                                                                                                                                                  | DR735359<br>DR741379<br>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<br>CD486134                         | 23                       | 24                                  | <i>Gossypium hirsutum</i>                             |
| Lyces_hox5                                                                                                                                                                                                                                                                                                                                    | BT014213.1                                   | 25                       | 26                                  | <i>Lycopersicon esculentum</i>                        |
| Lyces_VaHOX1                                                                                                                                                                                                                                                                                                                                  | X94947                                       | 27                       | 28                                  | <i>Lycopersicon esculentum</i>                        |
| Medsa_hox16*                                                                                                                                                                                                                                                                                                                                  | CB892061<br>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>                            |
| Poptr_hox16_2                                                                                                                                                                                                                                                                                                                                 | scaff_XII.649                                | 42                       | 43                                  | <i>Populus tremuloides</i>                            |
| Poptr_hox16_3                                                                                                                                                                                                                                                                                                                                 | lcl scaff_VI II.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>                             |
| *Código compilado a partir de varias referencias de EST (se muestran las principales); Siendo la calidad de secuenciación EST normalmente menor, pueden esperarse algunas sustituciones de ácidos nucleicos.<br>**Las secuencias de <i>Daucus carota</i> y de <i>Glycine max</i> se han corregido en comparación con su número de referencia. |                                              |                          |                                     |                                                       |

**Ejemplo 2: Alineamiento de secuencias polipeptídicas de hox5 de HDZip de clase I**

Se usó AlignX a partir del vector NTI (Invitrogen) basado en el popular algoritmo Clustal de alineamiento progresivo (Thompson y col. (1997) Nucleic Acids Res 25:4876-4882; Chenna y col. (2003). Nucleic Acids Res 31:3497-3500). Puede construirse un árbol filogenético usando un algoritmo de agrupamiento del vecino más próximo. Los valores por defecto son para la penalización por hueco abierto de 10, para la penalización por extensión de hueco de 0,1 y la matriz de ponderación seleccionada es Blosum 62 (en caso de alinear polipéptidos).

El resultado de los múltiples alineamientos de secuencia se muestra en la figura 2. Los tres dominios principales caracterizados, desde el extremo N al extremo C, están fuertemente empaquetados y se identifican como la caja ácida, el homeodominio de clase I y la cremallera de leucina de seis héptadas. El "dominio conservado" comprende estos tres dominios. Además, la cola de Trp y el motivo de aminoácidos RPFF están ligeramente empaquetados.

**Ejemplo 3: Cálculo del porcentaje de identidad global entre secuencias polipeptídicas de hox5 de HDZip de clase I**

Los porcentajes de similitud e identidad globales entre las secuencias polipeptídicas de hox5 de HDZip de clase I se determinaron usando el programa informático 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; programa alojado por Ledion Bitincka). El programa MatGAT genera matrices de similitud/identidad para secuencias de ADN o proteína sin la necesidad del alineamiento previo de los datos. El programa efectúa una serie de alineamientos por parejas usando el algoritmo de alineamiento global de Myers y Miller (con una penalización por hueco abierto de 12, y una penalización por extensión de hueco de 2), calcula la similitud y la identidad usando, por ejemplo Blosum 62 (para polipéptidos), y después sitúa 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 usados en la comparación fueron:

**Matriz de puntuación: Blosum62**

Primer hueco: 12

Hueco extendido: 2

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

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

Tabla B 1: Resultados de MatGAT para la similitud e identidad global sobre la longitud completa de las secuencias polipeptídicas.

|                   | 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 | 61 |    | 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 |    |

El "dominio conservado" de las secuencias polipeptídicas de hox5 de HDZip de clase I comprende desde el extremo N al extremo C, una caja ácida, un homeodominio de clase II y la cremallera de leucina de seis héptadas (véase la figura 2), tal como se ha definido anteriormente en el presente documento. Cuando se efectúa el análisis del porcentaje de identidad en los dominios conservados en lugar de en las secuencias polipeptídicas de longitud completa, se observa un aumento en el porcentaje de identidad, tal como se muestra en la tabla B 2. Los valores inferiores son ahora por encima del 50 % de identidad de aminoácidos en comparación con la SEC ID N°: 2.

**Tabla B 2: Resultados de MatGAT para la similitud e identidad global sobre el "dominio conservado" de las secuencias polipeptídicas.**

|                      | 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_CPHB-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 dominios comprendidos en las secuencias polipeptídicas de hox5 de HDZip de clase I**

La base de datos Integrated Resource of Protein Families, Domains and Sites (InterPro) es una interfaz integrada para las bases de firma usadas de manera común para búsquedas basadas en texto y secuencias. La base de datos combina estas bases de datos, que usan metodologías diferentes y diferentes grados de información biológica acerca de proteínas bien caracterizadas para derivar firmas de proteínas. Las bases de datos colaboradoras incluyen SWISS-PROT, PROSITE, TrEMBL, PRINTS, ProDom y Pfam, Smart y TIGRFAMs. InterPro está alojada en el European Bioinformatics Institute en el Reino Unido.

Los resultados del barrido InterPro de la secuencia polipeptídica representada por la SEC ID N°: 2 se presentan en la tabla C.

**Tabla C: Resultados del barrido InterPro de la secuencia polipeptídica representada por la SEC ID N°: 2**

|                 |                  |                                                           |                   |
|-----------------|------------------|-----------------------------------------------------------|-------------------|
| <b>InterPro</b> | <b>IPR000047</b> | <b>Motivo hélice-giro-hélice, represor de tipo lambda</b> |                   |
|                 | PRINTS           | PR00031                                                   | HTHREPRESSR       |
| <b>InterPro</b> | <b>IPR001356</b> | <b>Caja homeostática</b>                                  |                   |
|                 | PRODOM           | PD000010                                                  | Caja homeostática |
|                 | PRINTS           | PR00024                                                   | HOMEBOX           |
|                 | PFAM             | PF00046                                                   | Caja homeostática |
|                 | SMART            | SM00389                                                   | HOX               |
|                 | PROFILE          | PS00027                                                   | HOMEBOX_1         |
|                 | PROFILE          | PS50071                                                   | HOMEBOX_2         |

(continuación)

|                 |                  |                                                            |                        |
|-----------------|------------------|------------------------------------------------------------|------------------------|
| <b>InterPro</b> | <b>IPR003106</b> | <b>Cremallera de leucina, asociada a caja homeostática</b> |                        |
|                 | PFAM             | PF02183                                                    | HALZ                   |
| <b>InterPro</b> | <b>IPR009057</b> | <b>Similar a homeodominio</b>                              |                        |
|                 | SUPERFAMILY      | SSF46689                                                   | Similar a homeodominio |
| <b>InterPro</b> | <b>IPR012287</b> | <b>Relacionado con homeodominio</b>                        |                        |
|                 | GENE3D           | G3DSA:1.10.10.60                                           | Rel. con homeodominio  |

Puede calcularse la composición primaria de aminoácidos (en %) para determinar si un dominio de polipéptido es rico en aminoácidos específicos (por ejemplo, en una caja ácida) usando programas informáticos del servidor ExPASy, en particular la herramienta ProtParam (Gasteiger E y col., (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784-3788). Entonces puede compararse la composición de la secuencia polipeptídica de interés con la composición media de aminoácidos (en %) en el banco de datos de secuencias de proteínas Swiss-Prot.

En la tabla a continuación (tabla D), se comparan el % de Asp (D), % de Glu (E) y su contenido combinado en la caja ácida de SEC ID N°: 2 con la media en el banco de datos de secuencias de proteínas Swiss-Prot.

**Tabla D**

|                                                                  | % de Asp (D) | % de Glu (E) | % de Asp (D) + % de Glu (E) |
|------------------------------------------------------------------|--------------|--------------|-----------------------------|
| Media en el banco de datos de secuencias de proteínas Swiss-Prot | 5,3 %        | 6,6 %        | 11,9 %                      |
| Caja ácida de 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 de eucariotas se han clasificado de acuerdo con su contenido de aminoácidos, y las categorías principales incluyen dominios de activación ácidos, ricos en glutamina y ricos en prolina (Rutherford y col., (2005) Plant J. 43(5):769-88, y las referencias citadas en el mismo),

El consorcio Genome Ontology (GO) es una colaboración internacional entre científicos en varias bases de datos biológicas, con una oficina editorial radicada en el European Bioinformatics Institute. El objetivo de GO es proporcionar vocabularios controlados par la descripción de funciones moleculares, procesos biológicos y componentes celulares de productos génicos. Cuando se efectúa un barrido de InterPro tal como se describe anteriormente, también se efectúa una búsqueda en la base de datos de GO. Las secuencias polipeptídicas de hox5 de HDZip de clase I tienen como función molecular la de factor de transcripción y actividad de unión a ADN específica de secuencia, y localizada en el núcleo de células vegetales (véase la tabla a continuación (tabla E)).

**Tabla E**

|                                                     | Entrada de Gene Ontology                                                                                                                                                               |
|-----------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Homeodominio                                        | Función molecular: actividad de factor de transcripción (GO:0003700)<br>Componente celular: núcleo (GO:0005634)<br>Función molecular: unión a ADN específica de secuencia (GO:0043565) |
| Cremallera de leucina, asociada a caja homeostática | Función molecular: Unión a ADN (GO:0003677) Componente celular: núcleo (GO:0005634)                                                                                                    |

### **Ejemplo 5: Predicción de la topología de secuencias polipeptídicas de hox5 de HDZip de clase I**

La predicción de la cremallera de leucina y la identificación de la héptada se llevó a cabo usando programas informáticos especializados, tales como 2ZIP, que combina un algoritmo convencional de predicción hélices

superenrolladas con una búsqueda aproximada para la repetición de leucina característica (Bornberg-Bauter y col., (1998) *Nucleic Acids Res* 26(11): 2740-2746; alojado en el Instituto Max Planck, Golm en Alemania). Una potencial cremallera de leucina, una repetición de leucinas o una hélice superenrollada pueden identificarse usando este programa informático.

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

[illegible]

**Ejemplo 6: Ensayo para secuencias polipeptídicas de hox5 de HDZip de clase I**

Los polipéptidos de *hox5* de HDZip de clase I u homólogos de los mismos tienen actividad de unión a ADN, preferentemente a semisitos de 5 pb que se solapan en una posición central, CAA(A/T)ATTG, tal como se detecta en ensayos de un híbrido de levadura (Meijer y col., (2000) Mol Gen Genet 263:12-21). En ensayos transitorios en suspensiones de células de arroz, se ha comunicado que el bombardeo conjunto de un polipéptido de *hox5* de HDZip de clase I con el gen indicador GUS dio como resultado un aumento en el número de puntos teñidos, que también tenían un color más intenso (Meijer y col., anteriormente citado). Este ensayo es útil para demostrar la función activadora de los polipéptidos u homólogos de *hox5* de HDZip de clase I.

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

Se amplificó la secuencia de ácidos nucleicos de *hox5* de HDZip de clase I de *Oryza sativa* mediante PCR usando como molde una biblioteca de ADNc de plántula de *Oryza sativa* (Invitrogen, Paisley, R.U.). Después de la retrotranscripción del ARN extraído de las plántulas, se clonaron los ADNc en pCMV Sport 6.0. En tamaño promedio del inserto del banco fue de 1,6 kb el número original de clones era del orden de  $1,67 \times 10^7$  ufc. Se determinó que el título original era de  $3,34 \times 10^6$  ufc/ml después de la amplificación de  $6 \times 10^{10}$  ufc/ml. Después de la extracción del plásmido, se usaron 200 ng de molde en una mezcla de PCR de 50  $\mu$ l. Los cebadores prm06000 (SEC ID N°: 34; sentido, codón de inicio en negra, sitio de AttB1 en cursiva: 5'-GGGGACAAGTTTGTACAAAAAGCAGGCTTAAACAATGGATCCCG- GCCG 3') y prm06001 (SEC ID N°: 35; inverso, complementario, sitio de AttB2 en cursiva: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTGATCAGCTCCAGAACCAGG 3'), que incluye los sitios de Attb para la recombinación Gateway, se usaron para la amplificación por PCR. La PCR se llevó a cabo usando ADN polimerasa Hifi Taq en condiciones convencionales. Un fragmento de PCR de 1116 pb (que incluyen sitios de attB; a partir de 1050 pb) se amplificó y purificó usando procedimientos convencionales. La primera etapa del procedimiento Gateway, la reacción BP, se llevó a cabo en ese momento, durante el cual el fragmento de PCR se recombina *in vivo* con el plásmido para producir, de acuerdo con la terminología de Gateway, un "clon de entrada". El plásmido pDONR201 se compró a través de Invitrogen, como parte de la tecnología Gateway®.

**Ejemplo 8: Construcción del vector**

Posteriormente se usó el clon de entrada que comprendía las secuencias de ácido nucleico en una reacción LR con un vector de "destino" usado para la transformación de *Oryza sativa*. Este vector contenía como elementos funcionales entre los límites del ADN-T: un marcador de selección en plantas; un casete de expresión marcador explorable; y un casete Gateway previsto para la recombinación de LR *in vivo* con la secuencia de interés ya clonada en el clon de entrada. Se colocó un promotor de GOS2 de ratón (SEC ID N°: 33) para la expresión constitutiva cadena arriba de este casete de Gateway.

Después de la etapa de recombinación de LR, el vector de expresión resultante (figura 3) se transformó en la cepa LBA4044 de *Agrobacterium* 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*. Se descascarillaron las semillas secas maduras de arroz *Japonica* de la variedad Nipponbare. Se llevó a cabo la esterilización incubando durante un minuto en etanol al 70 %, seguido de 30 minutos de HgCl<sub>2</sub> al 0,2 %, seguido de lavados 6 veces 15 minutos con agua destilada estéril. Las semillas estériles germinaron después en un medio que contenía 2,4-D (medio de inducción de callos). Después de incubar 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. Los trozos de callo embriogénico se subcultivaron en medio reciente 3 días antes del cultivo conjunto (para potenciar la actividad de la división celular).

Se usó la cepa LBA4404 de *Agrobacterium* que contenía el vector de expresión para el cultivo conjunto. Se inoculó *Agrobacterium* en medio AB con los antibióticos adecuados y se cultivaron durante 3 días a 28 °C. Después, las bacterias se recogieron y se suspendieron en medio de cultivo conjunto líquido a una densidad (DO600) de aproximadamente 1. Después se transfirió la suspensión a una placa de Petri y se los callos se sumergieron en la suspensión durante 15 minutos. Los tejidos de callos se transfirieron secos sobre un papel de filtro y se transfirieron a un medio sólido de cultivo conjunto y se incubaron durante 3 días en la oscuridad a 25 °C. Los callos cultivados conjuntamente crecieron en 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, se desarrollaron islas de callos resistentes de crecimiento rápido. Después de transferir este material al medio de regeneración y de incubación con luz, se liberó el potencial embriogénico y se desarrollaron brotes en las siguientes cuatro a cinco semanas. Se cortaron los brotes de los callos y se incubaron durante 2 a 3 semanas en un medio que contenía auxina desde el cual se transfirieron al suelo. Los brotes endurecidos se cultivaron con alta humedad y días cortos en un invernadero.

Se generaron aproximadamente 35 transformantes de arroz de T0 para una construcción. Los transformantes primarios se transfirieron de 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 las plantas transgénicas de una sola copia que mostraban tolerancia al agente de selección, se mantuvieron para la cosecha de las semillas de T1. Después, las semillas se recogieron al cabo de tres a cinco meses de su trasplante. El procedimiento produjo transformantes de un solo locus a una tasa de más del 50 % (Aldemita y Hodges, 1996, Chan y col. 1993, Hiei y col. 1994).

**Ejemplo 10: Procedimiento de evaluación fenotípica****10.1 Configuración de la evaluación**

Se generaron aproximadamente 35 transformantes de arroz de T0. Los transformantes primarios se transfirieron de una cámara de cultivo tisular a un invernadero para el crecimiento y la cosecha de la simiente de T1. Se retuvieron siete eventos, de los cuales la progenie de T1 segregó 3:1 respecto de la presencia/ausencia del transgén. Para cada uno de estos eventos, se seleccionaron aproximadamente 10 plántulas de T1 que contenían el transgén (hetero y homocigotos) y aproximadamente 10 plántulas de T1 que carecían del transgén (nulicigotos) controlando la expresión de marcadores visuales. Las plantas transgénicas y los correspondientes nulicigotos se cultivaron unos al lado de otros en posiciones aleatorias. Las condiciones del invernadero fueron de días cortos (12 horas de luz), 28 °C en la luz y 22 °C en la oscuridad, y con una humedad relativa del 70 %.

Todos los eventos de T1 se evaluaron adicionalmente en la generación de T2 siguiendo el mismo procedimiento de evaluación que para la generación de T1. Desde el estado de la siembra hasta el estado de madurez las plantas se pasaron varias veces a través de una cabina de obtención de imágenes digitales. En cada instante, se tomaron imágenes digitales (2048 x 1536 píxeles, 16 millones de colores) de cada planta desde al menos 6 ángulos distintos.

*Exploración de estrés salino*

Se cultivaron plantas de 4 eventos (semillas de T2) en un sustrato compuesto de fibras de coco y argex (proporción de 3 a 1). Se usó una solución de nutrientes normales durante las dos primeras semanas después de trasplantar las

plántulas al invernadero. Después de las dos primeras semanas, se añadió sal (NaCl) 25 mM a la solución de nutrientes, hasta que se cosecharon las plantas.

#### *Exploración de sequía*

- 5 Se cultivaron plantas de cinco eventos (semillas de T2) en suelo de siembra en condiciones normales hasta que alcanzaron el estado de emergencia. Después se transfirieron a una sección "seca" donde se contenía la irrigación. Se insertaron sondas de humedad en macetas al azar para controlar el contenido de agua del suelo (CAS). Cuando el CAS se encontró por debajo de determinados umbrales, las plantas volvieron a regarse automáticamente de manera continua hasta que se alcanzaba de nuevo un nivel normal. Las plantas volvieron a transferirse a condiciones normales. El resto del cultivo (maduración de plantas, cosecha de las semillas) fue igual que para plantas no cultivadas en condiciones de estrés abiótico. Se efectuó una ronda de confirmación que consistía en repetir la exploración con semillas de T2 no cosechadas de plantas de la primera exploración de sequía, sino de plantas cultivadas en condiciones normales.

#### *Exploración de disponibilidad reducida de nutrientes (nitrógeno)*

- 15 Las plantas de arroz se cultivan en suelo de siembra en condiciones normales salvo para la solución de nutrientes. Las macetas se riegan desde el trasplante hasta la maduración con una solución específica que tiene un contenido de nitrógeno (N) reducido, normalmente de 7 a 8 veces menos. El resto del cultivo (maduración de plantas, cosecha de las semillas) fue igual que para plantas no cultivadas con estrés abiótico. Entonces se miden los parámetros relacionados con la semilla.

#### 10.2 Análisis estadístico: Prueba F

- 20 Se usó un ANOVA (análisis de variantes) de dos factores como modelo estadístico para la evaluación general de características fenotípicas de las plantas. 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 invención. La prueba F se llevó a cabo para comprobar un efecto del gen sobre todos los eventos de transformación y para verificar el efecto general del gen, también conocido como efecto génico global. El umbral de significación para un efecto génico global auténtico se situó en un nivel de probabilidad del 5 % para la prueba F. Un valor de la prueba F significativo indica un efecto génico, lo que significa que no es solo la mera presencia o posición del gen la que está provocando las diferencias en el fenotipo.

#### 10.3 Parámetros medidos

##### 10.3.1 Medición de parámetros relacionados con la biomasa

- 30 Desde el estado de la siembra hasta el estado de madurez las plantas se pasaron varias veces a través de una cabina de obtención de imágenes digitales. En cada instante, se tomaron imágenes digitales (2048 x 1536 píxeles, 16 millones de colores) de cada planta desde al menos 6 ángulos distintos.

- 35 El área superficial de la planta (o biomasa foliar) se determinó contando el número total de píxeles en las imágenes digitales a partir de partes de planta superficiales discriminadas del fondo. Este valor se promedió para las fotografías obtenidas en el mismo instante desde los diferentes ángulos y se convirtió en un valor de superficie física expresado en mm cuadrados mediante calibración. Los experimentos muestran que el área superficial de la planta medida de este modo se correlaciona con la biomasa de las partes de planta superficiales. El área superficial es en el instante en el que la planta ha alcanzado su biomasa foliar máxima. El vigor temprano es el área superficial de la planta (plántula) tres semanas después de la germinación.

- 40 Se usó un parámetro adicional para medir la tolerancia al estrés abiótico cuando fue necesario: el índice de verdor tras sequía, que da una indicación de la senescencia de la planta. Este es la proporción (expresada como %) de píxeles amarillos en la primera imagen después de la sequía.

- 45 Para medir los parámetros relacionados con la raíz, las plantas se cultivaron en macetas especialmente diseñadas con fondos transparentes para permitir la visualización de las raíces. Una cámara digital registró imágenes desde el fondo de la maceta durante el crecimiento de la planta. Las características radiculares, tales como el área total proyectada (que puede correlacionarse con el volumen total de raíces), el diámetro medio y la longitud de las raíces por encima de un determinado nivel de espesor (longitud de las raíces gruesas), se dedujeron a partir de la fotografía usando programas informáticos adecuados. El aumento en la biomasa radicular se expresa como un aumento en la biomasa radicular total (medida como la biomasa máxima de las raíces observada durante el periodo de vida de la planta); o como un aumento en el índice raíz/brote (medido como la proporción entre la masa de la raíz y la masa de brotes en el periodo de crecimiento activo de las raíces y brotes).

##### 10.3.2 Medición de parámetros relacionados con las semillas

Se cosecharon las panículas primarias maduras, se contaron, se embolsaron, se marcaron con un código de barras y se secaron durante tres días en un horno a 37 °C. Después se trillaron las panículas y se recogieron y contaron las semillas. Las cáscaras rellenas se separaron de las vacías usando un dispositivo soplador. Las cáscaras vacías se

desecharon y se volvió a contar la fracción restante. Las cáscaras rellenas se pesaron en una balanza analítica. Se determinó el número de semillas cargadas contando el número de cáscaras rellenas que permanecía después de la etapa de separación. El rendimiento total de semilla se midió pesando todas las cáscaras rellenas recogidas de una planta. El número total de semillas por planta se midió contando el número de cáscaras recogidas de una planta. El peso de mil granos (PMG) se extrapola a partir del número de semillas cargadas contadas y de su peso total. El índice de cosecha (IC) en la presente invención se define como la relación entre el rendimiento de semillas total y el área superficial ( $\text{mm}^2$ ), multiplicado por un factor 106. El número total de flores por panícula, tal como se define en la presente invención, es la proporción entre el número total de semillas y el número de panículas primarias maduras. La tasa de carga de semillas tal como se define en la presente invención es la proporción (expresada en %) del número de semillas cargadas frente al número total de semillas (o floretes).

**Ejemplo 11: Resultados de plantas transgénicas que expresan un ácido nucleico de hox5 de HDZip de clase I en condiciones de crecimiento normales**

Los resultados de la evaluación de las plantas de arroz transgénicas que expresan el ácido nucleico de hox5 de HDZip de clase I en condiciones de crecimiento normales se presentan en la tabla F. Se muestra el porcentaje de diferencia entre las plantas transgénicas y los nulicigotos, con un valor de P de la prueba F por debajo de 0,05.

**Tabla F:** Resultados de la evaluación de plantas transgénicas que expresan un ácido nucleico de hox5 de HDZip de clase I en condiciones de crecimiento normales.

| Rasgo                                  | % de diferencia en T1 | % de diferencia en T2 |
|----------------------------------------|-----------------------|-----------------------|
| Diámetro medio de las raíces           | 2                     | 1                     |
| Longitud de raíces gruesas             | 9                     | 17                    |
| Rendimiento total de semillas          | 17                    | 19                    |
| Número de semillas cargadas            | 15                    | 17                    |
| Tasa de carga                          | 7                     | 9                     |
| Número total de semillas               | 6                     | 9                     |
| Número de flores por panícula          | 8                     | 9                     |
| TKW                                    | 2                     | 2                     |
| Índice de cosecha                      | 16                    | 16                    |
| Índice de verdor antes de la floración | 2                     | 4                     |

**Ejemplo 12: Resultados de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico útil para llevar a cabo los procedimientos de la invención, en condiciones de crecimiento con estrés salino**

Los resultados de la evaluación de las plantas de arroz transgénicas que expresan el ácido nucleico que codifica una secuencia polipeptídica de hox5 de HDZip de clase I en condiciones de crecimiento de estrés salino se presentan en la tabla G. Se muestra el porcentaje de diferencia entre las plantas transgénicas y los nulicigotos, con un valor de P de la prueba F por debajo de 0,05.

**Tabla G:** Resultados de la evaluación de plantas transgénicas que expresan el ácido nucleico de hox5 de HDZip de clase I en condiciones de crecimiento de estrés salino.

| Rasgo                                  | % de diferencia en T1 |
|----------------------------------------|-----------------------|
| Rendimiento total de semillas          | 41                    |
| Número de semillas cargadas            | 40                    |
| Tasa de carga                          | 30                    |
| TKW                                    | 1                     |
| Índice de cosecha                      | 36                    |
| Índice de verdor antes de la floración | 7                     |

**Ejemplo 13: Resultados de plantas transgénicas que expresan el ácido nucleico de *hox5* de HDZip de clase I en condiciones de crecimiento de estrés por sequía**

Los resultados de la evaluación de las plantas de arroz transgénicas que expresan la secuencia de ácido nucleico útil para llevar a cabo los métodos de la invención en condiciones de crecimiento con estrés por sequía se presentan en la tabla H. Se muestra el porcentaje de diferencia entre las plantas transgénicas y los nulicigotos, con un valor de P de la prueba F por debajo de 0,05.

**Tabla H:** Resultados de la evaluación de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico útil para llevar a cabo los procedimientos de la invención, en condiciones de crecimiento con estrés por sequía.

| Rasgo                         | % de diferencia en T1 | % de diferencia en T2 |
|-------------------------------|-----------------------|-----------------------|
| Diámetro medio de las raíces  | 2                     | 1                     |
| Longitud de raíces gruesas    | 9                     | 14                    |
| Rendimiento total de semillas | 18                    | 31                    |
| Número de semillas cargadas   | 21                    | 33                    |
| Tasa de carga                 | 15                    | 33                    |
| Índice de cosecha             | 18                    | 34                    |
| Índice de verdor tras sequía  | 2                     | 3                     |

**Transformación de maíz, trigo, soja, canola, alfalfa con secuencias útiles en los procedimientos desvelados****Transformación de maíz**

La transformación del maíz (*Zea mays*) se efectúa con una modificación del procedimiento descrito por Ishida y col., (1996) Nature Biotech 14(6): 745-50. La transformación depende del genotipo en el maíz y solo genotipos específicos son susceptibles de transformación y regeneración. La línea endogámica A188 (University of Minnesota) o los híbridos con A188 como progenitor son buenas fuentes de material donante para transformación, pero también pueden usarse otros genotipos de manera satisfactoria. Las mazorcas se recogen de la planta del maíz aproximadamente a los 11 días después de la polinización (DDP) cuando la longitud del embrión inmaduro es de aproximadamente 1 a 1,2 mm. Los embriones inmaduros se cultivan conjuntamente con *Agrobacterium tumefaciens* que contiene al vector de expresión, y las plantas transgénicas se recuperan mediante organogénesis. Los embriones extraídos se cultivan en medio de inducción de callos, después en medio de regeneración de maíz, que contiene el agente de selección (por ejemplo, imidazolinona, aunque pueden usarse varios marcadores de selección). Las placas de Petri se incuban con luz a 25 °C durante 2-3 semanas, o hasta que se desarrollan brotes. Los brotes verdes se transfieren desde cada embrión a medio de enraizamiento de maíz y se incubaron a 25 °C durante 2-3 semanas, hasta que se desarrollaron las raíces. Los brotes enraizados se trasplantan a suelo en el invernadero. Las semillas de T1 se producen a partir de plantas que muestran tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

**Transformación de trigo**

La transformación del trigo se efectúa con el procedimiento descrito por Ishida y col., (1996) Nature Biotech 14(6): 745-50. La planta de cultivo Bobwhite (disponible a través de CIMMYT, México) se usa de manera común en transformación. Los embriones inmaduros se cultivan conjuntamente con *Agrobacterium tumefaciens* que contiene el vector de expresión, y las plantas transgénicas se recuperan mediante organogénesis. Después de la incubación con *Agrobacterium*, los embriones se cultivan *in vitro* en medio de inducción de callos, después en medio de regeneración, que contiene el agente de selección (por ejemplo, imidazolinona, aunque pueden usarse varios marcadores de selección). Las placas de Petri se incuban con luz a 25 °C durante 2-3 semanas, o hasta que se desarrollan brotes. Los brotes verdes se transfieren desde cada embrión a medio de enraizamiento y se incubaron a 25 °C durante 2-3 semanas, hasta que se desarrollaron las raíces. Los brotes enraizados se trasplantan a suelo en el invernadero. Las semillas de T1 se producen a partir de plantas que muestran tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

**Transformación de soja**

La soja se transforma de acuerdo con una modificación del procedimiento descrito en la Patente US 5.164.310 de Texas A&M. Diversas variedades comerciales de soja son susceptibles de transformación mediante este procedimiento. La variedad de cultivo Jack (disponible a través de la Illinois Seed Foundation) se usa comúnmente para la transformación. Las semillas de soja se esterilizan para la siembra *in vitro*. El hipocotilo, la radícula y un cotiledón se extraen de las plántulas jóvenes de siete días de edad. El epicotilo y el cotiledón restante se cultivan adicionalmente para que desarrollen nódulos axilares. Estos nódulos axilares se extraen y se incuban con *Agrobacterium tumefaciens* que contiene al vector de expresión. Después del tratamiento del cocultivo, los explantes

se lavan y transfieren a medio de selección. Los brotes regenerados se extraen y se colocan en un medio de elongación de brotes. Los brotes no más largos de 1 cm se colocan en medio de enraizamiento hasta que se desarrollan las raíces. Los brotes enraizados se trasplantan a suelo en el invernadero. Las semillas de T1 se producen a partir de plantas que muestran tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

#### **Transformación de colza/canola**

Se usan los peciolo cotiledonarios y los hipocotilos de plántulas jóvenes de 5-6 días como explantes para el cultivo tisular y se transforman de acuerdo con Babic y col., (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 superficialmente para su siembra *in vitro*. Los explantes de peciolo del cotiledón con el cotiledón unido se extraen de las plántulas *in vitro*, y se inoculan con *Agrobacterium* (que contiene el vector de expresión) sumergiendo el extremo cortado del explante de peciolo en la suspensión bacteriana. Los explantes se cultivan entonces durante 2 días en medio MSBAP-3 que contiene 3 mg/l de BAP, sacarosa al 3 %, Phytagar al 0,7 % a 23 °C, 16 h de luz. Después de dos días de cultivo conjunto con *Agrobacterium*, los explantes de peciolo se transfieren a medio MSBAP-3 que contiene 3 mg/l de BAP, 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 de los brotes. Cuando los brotes tienen 5-10 mm de longitud, se cortan y transfieren a medio de elongación de brotes (MSBAP-0.5, que contiene 0,5 mg/l de BAP). Los brotes de aproximadamente 2 cm de longitud se transfieren al medio de enraizamiento (MS0) para la inducción de raíces. Los brotes enraizados se trasplantan a suelo en el invernadero. Las semillas de T1 se producen a partir de plantas que muestran tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

#### **Transformación de alfalfa**

Se transforma un clon de regeneración de alfalfa (*Medicago sativa*) usando el procedimiento de (McKersie y col, 1999 Plant Physiol 119: 839-847). La regeneración y transformación de la alfalfa depende del genotipo y por lo tanto se requiere una planta regeneradora. Se han descrito procedimientos para obtener plantas regeneradoras. Por ejemplo, estas pueden seleccionarse a partir de la variedad de cultivo Rangelander (Agriculture Canada) o cualquier otra variedad de alfalfa comercial, tal como se describe por Brown DCW y A Atanassov (1985, Plant Cell Tissue Organ Culture 4: 111-112). Como alternativa, se ha seleccionado la variedad RA3 (University of Wisconsin) para su uso en cultivo tisular (Walker y col., 1978 Am J Bot 65:654-659). Los explantes de peciolo se cultivan conjuntamente con un cultivo durante toda una noche de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie y col., 1999 Plant Physiol 119: 839-847) o LBA4404 que contienen al vector de expresión. Los explantes se cultivan conjuntamente durante 3 d en la oscuridad en medio de inducción SH que contiene 288 mg/l de Pro, 53 mg/l de tioprolina, 4,35 g/l de K<sub>2</sub>SO<sub>4</sub>, y 100 mm acetosiringinona. Los explantes se lavan en medio de Murashige-Skoog de fuerza media (Murashige y Skoog, 1962) y se plantan en el mismo medio de inducción SH sin acetosiringinona pero con un agente de selección adecuado y un antibiótico adecuado para inhibir el crecimiento de *Agrobacterium*. Después de varias semanas, los embriones somáticos se transfieren a medio de desarrollo BOi2Y que no contiene reguladores del crecimiento, ni antibióticos, y sacarosa a 50 g/l. Los embriones somáticos se germinan posteriormente en medio de Murashige-Skoog de fuerza media. Las plántulas enraizadas se trasplantaron a macetas y se cultivaron en un invernadero. Las semillas de T1 se producen a partir de plantas que muestran tolerancia al agente de selección y que contienen una sola copia del inserto de ADN-T.

#### **Listado de secuencias**

<110> CropDesign N.V.

<120> Plantas que tienen características de crecimiento mejoradas y un procedimiento para producir las mismas

<130> PF57967

<150> EP 05110413.1  
<151> 07-11-2005

<150> EP 05110429.7  
<151> 07-11-2005

<150> US 60/736.194  
<151> 14-11-2005

<150> EP 05110900.7  
<151> 17-11-2005

<150> US 60/739.686  
<151> 23-11-2005

<150> EP 05111260.5  
<151> 24-11-2005

<150> US 60/742.287  
<151> 05-12-2005

5 <160> 51

<170> PatentIn version 3.3

<210> 1  
<211> 1050  
<212> ADN

10 <213> Oryza sativa

<400> 1

```

atggatcccg gccgcgtcgt gttcgactcc ggcgtggcgc ggcgggcgtg ccccggcggc      60
gcgcagatgc ttctcttcgg cggcggcggc agcgccaaca gcggcggtt cttccgaggc      120
gtgccggcgg cgggtgctggg gatggatgaa tcgcggtcgt cgtcgtcggc ggcgggggcg      180
ggggcggaagc ggccgttctt cacgacgcac gaggagctcc tggaggagga gtactacgac      240
gagcaggcgc cggagaagaa gcggcggtcg acggcggagc aggtgcagat gctggagcgg      300
agcttcgagg aggagaacaa gctggagccg gagcggaaga cggagctcgc ccgccgcctc      360
ggcatggccc cccggcaggt cgccgtctgg ttccagaacc gccgcgcccg ctggaagacc      420
aagcagctcg agcacgactt cgaccgcctc aaggccgcct acgacgcctt cgccgccgac      480
caccatgccc tcctctccga caacgaccgc ctccgcgcgc aggtaatctc attaaccgag      540
aagctgcaag acaaggagac gtcgccgtcg tcggcgacca tcaccaccgc ggcgcaggag      600
gtcgaccagc cggacgaaca cacggaggcc gcgtcaacca ccggcttcgc caccgtcgac      660
ggcgcatctg cggcgccacc gcccggccac cagcagccgc cgcataaaga tgatcttggt      720
agcagcggcg gcaccaacga cgacggcgat ggcggcgcgg ccgtggtggt cttcgacgtc      780
accgagggcg ccaacgaccg cctcagctgc gagtcggcgt acttcgccga cgccgcggag      840
gcgtacgagc gcgactgcgc cgggcactac gccctctcgt cggaggagga ggacggcggc      900
gcggtcagcg acgagggctg cagcttcgac ctccccgacg ccgccgccgc cgccgccgcc      960
atgttcggcg ccgccggagt tgtgcaccac gacgccgcgg acgacgagga ggcgagctc     1020
ggcagctgga ccgcctggtt ctggagctga                                1050

```

<210> 2  
<211> 349  
<212> PRT  
<213> Oryza sativa

15

<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 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 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 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

<210> 3  
 <211> 1197  
 <212> ADN  
 <213> Oryza sativa

5

<400> 3

atggagtcgc gccggctcat cttcagcacg gcgggctccg gcgccgggca gatgctcttc 60  
 ttggactgcg gcgctggcgg cgggcggtc ggccggcggg ccattgtcca tcgaggggcg 120  
 agaccggtgc tcggcatgga ggaaggagg cgccgggtca agccggcctt cttcaccacc 180  
 cccgacgagc tcctcgaaga ggagtactac gacgagcagc tcccggagaa gaagcggcgc 240  
 ctcacgccgg agcaggtgca tctgctggag aggagcttcg aggaggagaa caagctggag 300  
 ccggagcggg agacggagct ggcgcggaag ctagggtgct agccgcggca ggtcgcctg 360  
 tggttccaga accgccgcgc gcgctggaag accaagcagc tcgagcgcga cttcagaccg 420  
 ctcaaggcgt cgttcgacgc cctccgcgcc gaccacgacg cctcctcca ggacaaccac 480  
 cgctccact ctcaggtcat gtcgttgacc gagaagctgc aagagaagga gacgacgacc 540  
 gagggcagcg ccggcgcggc cgttgacgtc ccgggcttgc ctgcccgggc cgacgtgaag 600  
 gtcgccgtcc cggacgccga ggaaccggcg ctggaggagg cgccggcggc gttcaggagg 660  
 cagcaggagc agcaggtgaa ggccgaggac aggtctagca cgggcagcgg cgggagcgcg 720  
 gtggtggaca cggacgcgca actggtggtc ggggtcggcc ggcaagcatc tcgccgccgt 780

ggacagcagc gtggagtcgt acttcccggg cggcgacgag taccacgact gcgtgatggg 840  
 ccccatggac cacgccgcgg ggggcatcca gtcggaggag gacgacggcg ccggcagcga 900  
 cgagggctgc agctactacg ccgacgacgc cggcgctctc ttcgccgacc acggccacca 960  
 ccaccaccac caacacgcgg acgacgacga ggaggacggc cagcagatca gctgctggtg 1020  
 gatgtggaac tagattttct gcgcgcgcgc gtcgtcgtgc attcaattct cgtgttaaaa 1080  
 aaatcgttct ctttttcatt tttccgcttc tttgtctgta atgttgagtt tcgatcggct 1140  
 atgagaagga aggaggtgta tgcattgtgca tggatatgga gggtaacaca tcggtga 1197

<210> 4  
 <211> 343  
 <212> PRT  
 <213> Oryza sativa  
 <400> 4

5

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 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 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

# ES 2 555 659 T3

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 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  
 <212> ADN  
 <213> Zea mays

5

<400> 5

atggatccga gcgcggtcag tttcgactct ggcggcgcg cgcggggcgg cggcgcgag 60  
 atgctgctct tcggcggcgg aggcagcgcc aacagcaacg gcttcttccg aggtgttccg 120  
 atggcgggtcc tgggcatgga cgacgcgacg cgcgtgggca agcggccctt cttcacgaca 180  
 cacgaggagc tcctagagga ggagtactac gacgagcagg cgccggagaa gaagcgccga 240  
 ctgacggcgg agcaggtgca gctgctggag cggagcttcg aagaagagaa caagctggag 300  
 ccggagcgca agaccgagct ggctcgccgc ctggggatgg cgccccgcca ggtagctgtt 360  
 tggttccaga accgccgcgc gcgctggaag accaagcaac tcgagaccga ctatgaccgc 420  
 ctcaaggctg cttacgacgc actcgccgcc gaccaccagg gcctcctggc cgacaacgat 480  
 aacctccggg cacaggtgat ctccctgacg gagaagctgc aaggcaagga gacatccccg 540  
 tcggcaacca ctgctgcca agaggtcgac cagccagacg aacacaccgc tgtgtcaggc 600

acggaagaac tgctggcgca gcagctcaag gacaacctcc acagcagcgg cgactgcact 660  
 ggccatggca cctctcttc ggaagaagac gacggtggcg tggtcagtga cgagggctgc 720  
 agcttcgctc tcccggatgc catgttcgct gccgggttca cccaccatgg cgccgaggag 780  
 gtgcagctgg ccaactggac atccatgttc tggaactga 819

<210> 6  
 <211> 272  
 <212> PRT  
 <213> Zea mays

5

<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

# ES 2 555 659 T3

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

<210> 7  
<211> 993  
<212> ADN  
<213> Zea mays

5

<400> 7

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| atggagtctg gacggctcat cttcaacgcg ccgggctctg gcgccgggca gatgctcttc   | 60  |
| ctcgactgcg gcgcaggcgg cgggtcccggc ggcggttctgt tccatcgagg cgggagaccg | 120 |
| atgcttggcc ttgaagaagg gcgcggcgta aaacggccct tcttcacctc gcccgcgag    | 180 |
| ctcctcgagg aagagtacta cgacgagcag ctgccggaga agaagcgccg cctcacccca   | 240 |
| gagcagggtgc ttctgctgga gaggagcttc gaggaggaga acaagctgga gccggagcgc  | 300 |
| aagacggagc tggcgcgcaa gctgggcctg cagcctcgcc aggtggccgt ctggttcag    | 360 |
| aaccgcccgcg cccggtggaa gaccaagcag ctcgagcgcg acttcgaccg cctcaaggcc  | 420 |
| tccttcgacg ctctccgagc ggaccacgac gccctcctcc aggacaacaa ccgcctccgc   | 480 |
| tcacaggttg tgtcgttgac cgagaagctg caagagaagg aggatgcgac ggagggcggc   | 540 |
| gccaccgctg acaccgccgc gccggcggtg gacgtcgagg cttccctggc cgacgacgtc   | 600 |
| gaggagccag cagagcctgc ggcgacgttc gaggtgctgc aggaggtgaa gtccgaggac   | 660 |
| aggctgagca ccggcagcgg cgggagcgcg gtggtggacg cggacgcgct gctgtacggc   | 720 |
| aggttcgccg cggcagttga tagcagcgtg gagtcgtact tccccggcgg cgaggaccac   | 780 |
| taccacgact gcgggacgat gggccccgtg aatcatggcg ccggaggagg catccagtcg   | 840 |
| gacgacgacg gcgccggcag cgacgagggg tgcagctact acgccgacga agccgccgcc   | 900 |
| gccgccgccg cgttcttcgc cggacacgcc acccaccacc acgcggacga ggacgaggac   | 960 |
| gccggccaga tcagctggtg gatgtggaac tag                                | 993 |

<210> 8  
<211> 330  
<212> PRT  
<213> Zea mays

10

<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  
 260 265 270  
 Gly Ala Gly Gly Gly Ile Gln Ser 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  
290 295 300

Phe Phe Ala Gly His Ala Thr His His His Ala Asp Glu Asp Glu Asp  
305 310 315 320

Ala Gly Gln Ile Ser Trp Trp Met Trp Asn  
325 330

<210> 9  
<211> 825  
<212> ADN  
<213> Saccharum officinarum

5

<400> 9

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atggatccga gcgcggtcag tttcaactcc ggcgggcgcgc ggcggggcgg cggcggcacg | 60  |
| cagatgctgc tcttcggcgg cggaggcagc gccaacagca acggcttctt ccgaggtggt  | 120 |
| ccgatggcgg tcctgggcat ggacgacgcg acgcgcgtgg gcaagcggcc cttcttcacc  | 180 |
| acacacgagg agctcctgga ggaggagtac tacgacgagc aggcgcccga gaagaagcgc  | 240 |
| cgtctgacgg cggagcaggt gcagctgctg gagcggagct tcgaggaaga gaacaagctg  | 300 |
| gagcccgagc gcaagaccga gctggctcgc cgcctcggga tggcgccccg ccaggtggcc  | 360 |
| gtctgggttcc agaaccgccg cgcgcgctgg aagaccaagc agctcgagac cgactatgac | 420 |
| cacctcaagg ctgcctacga cgcgctcgc gccgaccacc agggcctcct ggccgacaac   | 480 |
| gatagcctcc gggcacagggt ggtctcccta acagagaagc tgcaaggcaa ggagacatcc | 540 |
| ccgtcggcca ccaactgctgc ccaagaggtc gaccagccag acgaacacac cgcggcgtca | 600 |
| ggcactgaga aactgctggc gcagcagctc aaggacgacc tccacagcag cggcgactgc  | 660 |
| actggccatg gtgccctctc ctcagaggaa gaagatggtg gtgtggtcag tgacgagggc  | 720 |
| agctttgatc tcccggatgc catgtttgct gccggggtca cccaccatgg cgccgacgcc  | 780 |
| gaggaggcac agctggccaa ctggacatcc tggttctgga actga                  | 825 |

<210> 10  
<211> 274  
<212> PRT  
<213> Saccharum officinarum

10

<400> 10

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 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 Glu 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

# ES 2 555 659 T3

<210> 11  
 <211> 828  
 <212> ADN  
 <213> Sorghum bicolor

5 <400> 11

```

atggatccga gcgcggtcag tttcgactcc ggcggcgcg cgcggggcg cggcggcggc 60
ggcgcgcaga tgctgctctt cggcggcgga ggcagcgcca acagcaacgg cttcttccga 120

ggtgttccga tggcggtcct gggcatggac gacgcgacgc gcgtgggcaa gcggcctttc 180
ttcaccacgc acgaggagct cctggaggag gagtactacg acgagcaggc gcccgagaag 240
aagcgccgctc tgacggcgga gcaggtgcag ctgctggagc ggagcttcga ggaagagaac 300
aagctggagc cggagcgcaa gaccgagctg gctcgccgcc tcgggatggc gcctcgccag 360
gtggccgtct ggttccagaa ccgccgcgcg cgctggaaga ctaagcagct cgagaccgac 420
tatgaccgcc tcaaggctgc ctacgacgcg ctcgccgccg accaccaggg cctcctggcc 480
gacaacgata gcctccgggc acaggtgatc tccctaacgg ataagctgca acgcaaggag 540
acatccccgt cggcgaccac tgctgccccaa gaggtcgacc agccagacga acacaccgct 600
gcgtcaggca ctgagaaact gctggtgcag cagctcaagg acgacctcca cagcagcggc 660
gacttcactg gccatggtgc cctctcttca gaggaagagg atggtggtgt ggtcagcgac 720
gagggctgca gctttgatct cccggatgcc atgttcgctg ccggggtcac ccaccatggc 780
gccgaggagg cgcagctggc caactggaca tcctggttct ggaactga 828
  
```

<210> 12  
 <211> 275  
 <212> PRT  
 <213> Sorghum bicolor

10 <400> 12

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 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

# ES 2 555 659 T3

<210> 13  
 <211> 1023  
 <212> ADN  
 <213> Triticum aestivum

5 <400> 13

```

atggagcccg gccggctcat cttcaacacg tcgggctccg gcaacggaca gatgctcttc      60
atggactgcg gcgcgggchg catcgccggc gcggccggca tgttccatcg aggggtgaga      120
ccggtcctcg gcggcatgga agaagggcgc ggcgtgaagc ggcccttctt cacctcgccg      180
gatgacatgc tcgaggagga gtactacgac gaggcagctcc cggagaagaa gcggcgccctc      240
accccgagac aggtccacct gctggagagg agcttcgagg aggagaacaa gctggagccg      300
gagaggaaga cggagctggc ccgcaagctc gggctgcagc cagccaggtt ggccgtcttg      360
ttccagaacc gccgcgcccc gtggaagaca aagacgctgg agcgcgactt cgaccgcctc      420
aaggcgctct tcgacgccct ccggggccgac cagcagcccc tcctccagga caaccaccgg      480
ctccggtcac aggtggtaac gttgaccgag aagatgcaag ataaggaggc gccggaaggc      540
agcttcggtg cagccgcca cgcctcggag ccggagcagg cggcggcgga ggcgaaggct      600
tccttgcccg acgcccagga gcaggccgcg gcagcggagg cgttcgaggt ggtgcagcag      660
cagctgcacg tgaaggacga ggagaggctg agcccgggga gcggcgggag cgcggtgctg      720
gacgcgaggg acgcgctgct cgggagcggg tgcggcctcg ccggcgtggt ggacagcagc      780
gtggactcgt actgcttccc ggggggcgcc ggcggcgacg agtaccacga gtgcgtggtg      840
ggccccgtgg cgggcggcat ccagtcggag gaggacgacg gcgcgggcag cgacgagggc      900
tgcagctact accccgacga cgcgcgctc ttcttcgccg ccgcgcaagg gcacggccac      960
catcgcacgg acgacgacga tcagcaggac gacggccaga tcagctactg gatgtggaac     1020
tag                                                                    1023
  
```

<210> 14  
 <211> 340  
 <212> PRT  
 <213> Triticum aestivum

10 <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

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 Glu 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  
 <212> ADN  
 <213> *Arabidopsis thaliana*  
 <400> 15

5

atggaatcca attcgttttt cttcgatcca tctgcttcac acggcaacag catgttcttc 60  
 cttggaatc tcaatcccgt cgtccaagga ggaggagcaa gatcgatgat gaacatggag 120  
 gaaacttcga agcgaaggcc cttcttttagc tccctgagg atctctacga cgatgacttt 180  
 tacgacgacc agttgcctga aaagaagcgt cgcctcacta ccgaacaagt gcatctgctg 240  
 gagaaaagct tcgagacaga gaacaagcta gagcctgaac gcaagactca gcttgccaag 300  
 aagcttggtc tacagccaag gcaagtggct gtctggtttc agaatcgccg agctcgttgg 360  
 aaaacaaaac agcttgagag agactacgat cttctcaagt ccacttacga ccaacttctt 420  
 tctaactacg actccatcgt catggacaac gataagctca gatccgaggt tacttcctg 480  
 accgaaaagc ttcagggcaa acaagagaca gctaataaac cacctgggtca agtgcccgaa 540  
 ccaaaccaac ttgatccggt ttacattaat gcggcagcaa tcaaaaccga ggaccggtta 600  
 agttcagggg gcgttgggag cgcggtacta gacgacgacg cacctcaact actagacagc 660  
 tgtgactctt acttcccaag catcgtaacc atccaagaca acagcaacgc cagtgatcat 720  
 gacaatgacc ggagctgttt cgccgacgtc tttgtgcca ccacttcacc gtcgcacgat 780  
 catcacggtg aatcattggc tttctgggga tggccttag 819

<210> 16  
 <211> 272  
 <212> PRT  
 <213> *Arabidopsis thaliana*

5 <400> 16

```

Met Glu Ser Asn Ser Phe Phe Phe Asp Pro Ser Ala Ser His Gly Asn
 1          5          10
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
Glu Lys Ser Phe Glu Thr Glu Asn Lys Leu Glu Pro Glu 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

<210> 17  
<211> 969  
<212> ADN  
<213> Daucus carota

5

<400> 17

|            |             |            |            |            |             |     |
|------------|-------------|------------|------------|------------|-------------|-----|
| atggcggttc | ggaggggtgtt | ctatggggag | ggagccaata | cgacgtcggc | tagcctgttg  | 60  |
| tttcatagtc | aaagacctga  | gcctttcttt | ctttctgcac | cttctccttc | tctaattggt  | 120 |
| tcaaaatcca | tggttagctt  | tcaagatgct | aagcgaaaaa | atccctacga | tgggttcttt  | 180 |
| atgcggtcat | atgatgaaga  | agaaattggg | gatgaagaat | atgatgaata | ctttcagcag  | 240 |
| cctgagaaga | agaggaggct  | caaggctgat | caaatccagt | ttcttgagaa | aagttttgag  | 300 |
| actgataaca | agcttgagcc  | tgaaagaaaa | gttcagcttg | caaaagaact | cggcttgag   | 360 |
| ccaagacagg | ttgcgatatg  | gtttcagaac | cgctcagcac | ggtggaagac | caaaacacta  | 420 |
| gaaaaagatt | atgatgtatt  | gcaaaatagc | tacaacagcc | tcaaggctga | ctatgacaat  | 480 |
| ctacttgccg | agaaagaaaa  | acttaaagcc | gaggttctcg | acctgacaga | caagctactt  | 540 |
| ctcaaagaag | ataaggggag  | caagacagta | gtttttgata | agcaaaaggt | gtctgcagca  | 600 |
| ttccaacaag | aacgtgttag  | taatgacata | tctgtgggtg | aagtactcag | taactcagtt  | 660 |
| atggactgca | agcaagaaga  | tcataactct | gtgaaaagtg | atgcagttga | ttctgacagt  | 720 |
| ccacactaca | gtgatgaagt  | ctactccagt | tttatggagc | cagtggatcg | ctcttatggt  | 780 |
| tttgaacctg | ctcagtcgga  | tatatctcaa | gatgaagaag | atgacatggg | gaacaactta  | 840 |
| tttctcccat | catatcatgt  | tttctcaaag | actgaagacg | gtagttactc | cgaccagcct  | 900 |
| tcgaactctt | cgtactttgg  | cttcccagtt | gaagatcata | cgtttggctt | ttgggggtact | 960 |
| gaattataa  |             |            |            |            |             | 969 |

<210> 18  
<211> 322  
<212> PRT  
<213> Daucus carota

10

<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

<210> 19  
<211> 1038  
<212> ADN  
<213> Glycine max  
<400> 19

5

atggcgagtg gcaagcttta tgcgggttca aacatgtcac ttctcctcca aaacgaaagg 60  
ctcccttgct cctctgaagt ccttgagtct ctttgggctc agacctctaa ccctgcttcc 120  
ttccaaggtt caaaaccctg ggttgatttt gagaatgtaa gtgggagcag gatgacggat 180  
aggcctttct ttcaagcgtt ggagaaggaa gagaactgtg atgaggatta cgaggggtgt 240  
ttccaccaac cggggaagaa aaggaggctc acaagcgaac aagttcagtt ccttgaaagg 300  
aactttgagg tagagaacaa gcttgaaccc 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 attaattatt 660  
tcaaaaaatg caacaacaac aacaacatct gaaaatggga ccaaagtgtt gtcaccactc 720  
ccactcccta ttatggtaac atgctgcaag caagaagatg ccaactcagc caaagtgat 780  
gtccttgatt cggatagccc acattgcact tcattcgtgg agcctgctga ttcctctcat 840  
gcctttgaac cagaagacca ctcaagagac ttctcccaag atgaagagga taaccttagt 900  
gaaaaccttt tgatgacctt cccttcttct tgttgcttac ctaagggtga agaactgc 960  
tatgacggcc ctctgaaaa ctcttgtaat tttggcttcc aggttgagga tcaaacttc 1020  
tgtttctggc cctattga 1038

<210> 20  
<211> 345  
<212> PRT  
<213> Glycine max  
<400> 20

10

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  
 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 Glu 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  
<212> ADN  
<213> Craterostigma plantagineum

<400> 21

atgaactctg ctcggatttt cttecgaccca tcttcccacg gcaacatgct gcagtttctt 60  
gggaacgccg gcggcgattc atccgttttc cgaggaacaa gatcgtcgtc ggtgctgaac 120  
atggaggaga gctcgttaaa acgacagatt ttcagcggcg gcggcggcga tgaattctac 180  
gacgaggaat actacgacga gcagttgttg cctgagaaga agcggccgact caccgccgag 240  
caggttcact tgcttgagaa gagcttcgag gctgagaaca agcttgagcc tgagcgaaag 300  
gctgagctgg cgaagaagct cggattgcag ccgaggcaag tcgccatttg gttccaaaac 360  
cgccgagcac ggtggaagac taagcagtta gagagggact acgacaagct taagtcttcc 420  
tatgattctc ttctctcaac ctacgactct attcgccagg aaaacgacaa gctcaaagcc 480  
gagctccttt ccctgaacga gaaattgcaa cccaaagacg acgacgaccc atcggccgaa 540  
ataggtcgaa atctcagttc atcgtcgccg cctgtcgacg cggctgagcc gccgtgcctg 600  
aagctgacgg tgaaggtgga ggaccgcctg agcacgggga gcaacggcag cgcagtaatg 660  
gacggcgacg gacctcagca gctcctcgac gacagcggcg actcgtactt cgagaacgac 720  
gaggaatacg actgcgccgc cgcaagtttg gctgctgcga aggaggacga cggcagcgat 780  
gagggcgggg gttacttcac cgaggctctc gcggcgaggagg aggaggaggc gccgtttgct 840  
tggtgtattt ggtcttaa 858

<210> 22  
<211> 285  
<212> PRT  
<213> Craterostigma plantagineum

<400> 22

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 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

<211> 813

<212> ADN

<213> Gossypium hirsutum

<220>

<221> Variación

<222> (808)..(808)

<223> /reemplazar="c" /reemplazar="g" /reemplazar="t"

<400> 23

atggagtctg gccgtctttt tttcaatccc tccactaccc accgcaacat gttgcttctc 60  
 gggaaactg aacccatctt tcgaggggca agaacaatgg ttagcatgga ggaaaaccca 120  
 aagaagcgac tgttcttcag ctccgaggag gatttgtacg acgaagagta ctacgacgag 180  
 cagttgcccg agaaaaagcg tcgccttacg tcggagcagg tgtatctgct agagaagagc 240  
 tttgaggcag agaacaagct ggagccggag aggaagagcc agttggccaa gaagttagga 300  
 ctgcaaccaa ggcagggtggc ggtatggttc cagaaccgcc gtgcaagggtg gaagacaaag 360  
 cagcttgaaa gggactatga cctcctcaaa tcttcctttg attcccttca gtccaattat 420  
 gacactattc tcaaagaaaa tgagaagctc aaatctgagg tagcttcctt gactgaaaaa 480  
 ctacaagcca aagatgtggc aacagaagca atagcagggtg aaaaggatga agggtaggca 540  
 gctgagatgg cctccgccct ccaattcagt atgaagggtg aggaccgtct tagtagcggc 600  
 agtgtcggaa gcgcggtggt ggatgaggat gccccacagc tgggtggacag cggcaattcc 660  
 tactttccaa gcgatgaata ctccagaggc attggccctt tcgatggggt tcagtcggaa 720  
 gatgaggatg gcagtgataa ttgcgggagt tacttctccg atgtgttcgc aaccacagag 780  
 cagggagcat taggattgtg ggcttgatc taa 813

<210> 24

<211> 270

<212> PRT

5 <213> Gossypium hirsutum

<220>

<221> INCIERTO

<222> (270)..(270)

<223> xaa puede ser cualquier aminoácido de origen natural

10 <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

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  
<212> ADN  
<213> Lycopersicon esculentum

<220>  
<221> Variación  
<222> (294)..(294)  
<223> /reemplazar="c" /reemplazar="g" /reemplazar="t"

<400> 25

atgggatctg ggcataatatt ttctgacccg tcgtcgtgtc acggcaacat gctgttcctt 60  
gggagcggag atcctgtttt ccgaggacca agatcgacga tgatgaagat ggaggactcc 120  
tcgaagaggc gacccttctt tagctcgccg gaggatctat atgacgagga atactacgac 180  
gagcagtcac cggagaagaa gcgccgtctc actcctgagc aggtgcactt gttggagaag 240  
agctttgaga cagaaaacaa gctggagccc gagcgcaaaa cccagctggc ctaaaagctg 300  
gggctgcagc ccagacaggt ggctgtatgg ttccaaaacc gccgtgcccc gtggaagacc 360

aagcagctcg agagggatta tgatcagctc aaatcctctt atgactccct tctctctgat 420  
 ttgactccg ttcgcaaaga taacgataag ctcaaactctg aggttggttc attgatggaa 480  
 aagttacagg ggaaagtggg tggaggagca gggggaaatg aaaaatctga catcttgagg 540  
 gtggatgcta tgacgatcct tcaagtgaag gtgaaggctg gggaccggtt gagcagtggc 600  
 agtgggtggga gcgcggtggg agatgagcat agttcacagc tgggtggacag tggggactca 660  
 tattttcaca ctgatcatga ggagtatcca gggcctggag gatgcaatgt tcctccaccc 720  
 atggatgggt tacaatcgga ggaagatgat ggtagtgatg atcatggcag ttgccatggc 780  
 tacttctcta acgtctttgt ggcagaagag cagcaccatg aacaaggaga agagcctatt 840  
 ggatgggttct ggtcttaa 858

<210> 26

<211> 285

<212> PRT

<213> Lycopersicon esculentum

<220>

<221> INCIERTO

<222> (98)..(98)

<223> Xaa puede ser cualquier aminoácido de origen natural

<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 Gln 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

<210> 27  
 <211> 972  
 <212> ADN  
 <213> Lycopersicon esculentum

5

<400> 27

atggctccag ggattctcta tgggtggttct tctaatttcg atggcgtttt tactcaaaaa 60  
 cagagagacg tgttttcttc atctactgca ccgaaagggc atcttggttc cttttttgcc 120  
 cctgcctctt cttcttctaa tttcttggga tccagttcta tggtagagttt tcgcggtggt 180  
 aatggagggga agagatcatt ctttgattcg ttcgatcagg atgacaatga agctgatgaa 240  
 ttgggggaat atcttcatca agcggagaag aagaggcgac ttactgacaa ccaagttcag 300  
 tttcttgaga agagttttgg ggaagagaac aaacttgaac cagaaagaaa agttcagctt 360  
 gctaaagaac ttggtctgca gcctcgccaa attgcaattt ggtttcagaa tcgtcgtgcg 420  
 cgatggaaga ctaagcagct cgagaaagat tatgatgaat tgaggaatag atacgatact 480  
 ctgaaatcaa attacaataa tcttctcaag gaaaaagaag atcttcgaac tgaagttttc 540  
 cgtctcaccg gtaagctggt tatcaaagag aaaggaaaatg ggcaattgga tttgcgcgat 600  
 gaacacaaac actccaatgc attggcaaaa gaaaccgtgg ttgatccaat gtccaatgta 660  
 ccagctctgg ttgttaagca ccagcaggaa gatttaagct ctgctaagag tgatgttttc 720  
 gactcagaaa gccacgta caccagtaga atgcattcct cagtcgtaga tcaggatgat 780

tctgctcgcg catttgaaac tgatcagtcg gattcatctc aggatgatga tgaaaacttc 840  
agcaagaata tgctttctac tgccaaccta cttggcaaag acgcggatga tgattatccc 900  
gcgacatcat caaatttgag ttactttgga tttccagttg aagaccaagg ttttggtttc 960  
tggacttatt aa 972

<210> 28

<211> 323

<212> PRT

5 <213> Lycopersicon esculentum

<400> 28

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  
 <212> ADN  
 <213> Medicago sativa  
 <400> 29

atggcggggtg ggagagtttt ttcaaattggt cctgcaaata tttcaaatat aaatatgaat 60  
 attttgcttc agaatcaaca acaaactcct cgtggaaact cttctcaaca acctcttgat 120  
 tctcttttcc tttcttcttc tgcttctttc tttggttcaa gatctatggt gagttttgaa 180  
 gatgttcaag gaaggaaaag gcgcaacagg tctttctttg gaggatttga tcttgacgaa 240  
 aacggagagg atgagatgga tgagtacttt catcaatccg agaagaaacg gcgtctctca 300  
 gtggatcaag ttcagtttct tgagaaaagc tttgaggagg acaacaaact tgaaccagag 360  
 aggaaaacca agctagctaa agaccttggg ttgcagccac ggcaagttgc tatttggttt 420  
 caaaaccgtc gtgcaagggt gaagactaaa cagcttgaga aggattatga ttctcttaat 480  
 gatggttatg agtctcttaa gacagagtat gacaaccttc tcaaagagaa agataggtta 540  
 caatctgagg tggcaagcct aactgaaaag gtacttgaaa gagagaaaca agagggaaaa 600  
 ttcaaacaag gtgaaagtga aacaaaggaa ttcttgaagg aaccaacaat taataagcct 660  
 ttggttgatt cagtttctga ggggtgaagga tccaaattgt caattggtga ggcttctaatt 720  
 aataataata ataataacaa acttgaagat attagttcag caaggagtga catattggat 780  
 tgtgaaagtc cacgctacac tgatggagtg ttagagacat gtgattcttc ctatgtattt 840  
 gaacctgaat atcaatcgga cctatcacaa gatgaagaag atcacaattt attgcctcct 900

tacatcttta caaaacttga agatgtgaat taçtccgacc cgc̄cacataa ttcaacaagt 960  
tatggatttc aagaggaaga tcatcatcaa gctctttggc cttgggtctta ttag 1014

<210> 30  
<211> 337  
<212> PRT  
<213> Medicago sativa  
<400> 30

5

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 Gln Thr Pro Arg Gly  
20 25 30  
Asn Ser Ser Gln Gln Pro Leu Asp Ser Leu Phe Leu Ser 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

<210> 31

<211> 843

<212> ADN

<213> Aquilegia formosa x Aquilegia pubescens

<400> 31

atggattcaa caacaagccg tcttttcttt gatggttcct gccatgggaa catgttgctt 60

ttagggagtg gagatcccgt tcttcgaggt tcaagatcat tcattaatat ggaagattct 120

ttgaaaagac gtccttttta tagttcaaca gatgaactaa ttgaagagga gttttatgat 180

gaacagctac ctgaaaagaa acgtcgtctt acttctgagc aggttcatct attggagaag 240

agctttgaga cagagaacaa gctggaacca gatcgtaaga cccagcttgc taagaagctt 300

gggttgcaac cgagacaagt tgcagtttg tttcagaata gacgagctcg ttggaagact 360

aagcaactag agagagatta tgatcttctt aaagcttctt atgattccct tcgttctgat 420

tacgatgaca ttgttaaaga gaatgagaag ctcaaactctg aggtgggttc cttaactggg 480

aagttgcagg tcaaggaggg agctgggatg gagttaaatc agatatctga cccaccactc 540

tccactgaag aaaatgttga tgtaactacg atgcaattta atgttaaggt tgaggatcgc 600

ttgagctctg gcagtggggg aagtgcgtg gttgatgagg aatgtcgaca gcttgttgac 660

agtgttgatt cctatttccc tggcgatgac tatggtcaat gcataggccc agtagatgga 720

gtccagtcag aagaagatga cattagtac gacagccgga gctatttctc agatgtcttt 780

ccagctgcac cagagcagaa ccaccaggag agtgagacat tgggttggtg ggactgggct 840

taa 843

<210> 32  
 <211> 280  
 <212> PRT  
 <213> Aquilegia formosa x Aquilegia pubescens

5 <400> 32

```

Met Asp Ser Thr Thr Ser Arg Leu Phe Phe Asp Gly Ser Cys His Gly
1      5      10
Asn Met Leu Leu Leu Gly Ser Gly Asp Pro Val Leu Arg Gly Ser Arg
20     25
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 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
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 Ser
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  
<211> 2193  
<212> ADN  
<213> Oryza sativa

5

<400> 33

```

aatccgaaaa gtttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct      60
aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact      120
catccacctta ctttagtggt aatcgggcta aataaaaaag agtcgctaca ctagtttcgt      180
tttccttagt aattaagtgg gaaaatgaaa tcattattgc ttagaatata cggtcacatc      240
tctgtcatga agttaaatga ttcgaggtag ccataattgt catcaaactc ttcttgaata      300
aaaaaatctt tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga      360
atgaagatat tctgaacgta ttggcaaaga tttaaacata taattatata attttatagt      420
ttgtgcattc gtcataatgc acatcattaa ggacatgtct tactccatcc caatttttat      480
ttagtaatta aagacaattg acttattttt attatttatt ttttttcgat tagatgcaag      540
gtacttiacgc acacactttg tgctcatgtg 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 ttttgctcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca      840
acagagtggc tgccacaga acaaccaca aaaaacgatg atctaacgga ggacagcaag      900
tcgcgaacaa ctttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa      960
aaccaagcat cctcctctc ccatctataa attcctcccc ctttttcccc tctctatata     1020
ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag     1080
cgaccgcctt cttcgatcca tatcttcgg tcgagttctt ggtcgatctc ttcctcctc     1140
cacctcctcc tcacagggtg tgtgcccttc ggttgttctt ggatttattg ttctagggtg     1200
tgtagtacgg gcgttgatgt taggaaagg gatctgtatc tgtgatgatt cctgttcttg     1260
gatttgggat agaggggttc ttgatgttgc atgttatcgg ttcggtttga ttagtagtat     1320
ggttttcaat cgtctggaga gctctatgga aatgaaatgg tttagggtac ggaatcttgc     1380
gattttgtga gtacctttt tttgaggtaa aatcagagca ccggtgattt tgcttggtgt     1440
aataaaaagta cggttgtttg gtcctcgatt ctggtagtga tgcttctcga tttgacgaag     1500

```

ctatcctttg tttattccct attgaacaaa aaataatccaa ctttgaagac ggtcccgttg 1560  
atgagattga atgattgatt ctttaagcctg tccaaaattt cgcagctggc ttgttttagat 1620  
acagtagtcc ccatcacgaa attcatggaa acagttataa tcctcaggaa caggggattc 1680  
cctgttcttc cgatttgctt tagtcccaga attttttttc ccaaataatct taaaaagtca 1740  
ctttctgggtt cagttcaatg aattgattgc tacaaataat gcttttatag cgttatccta 1800  
gctgtagttc agttaatagg taatacccct atagtttagt caggagaaga acttatccga 1860  
tttctgatct ccatttttaa ttatatgaaa tgaactgtag cataagcagt attcatttgg 1920  
attatttttt ttattagctc tcaccccttc attattctga gctgaaagtc tggcatgaac 1980  
tgtcctcaat tttgttttca aattcacatc gattatctat gcattatcct cttgtatcta 2040  
cctgtagaag tttctttttg gttattcctt gactgcttga ttacagaaaag aaatttatga 2100  
agctgtaatc gggatagtta tactgcttgt tcttatgatt catttccttt gtgcagttct 2160  
tggtgtagct tgccactttc accagcaaag ttc 2193

<210> 34  
<211> 49  
<212> ADN  
<213> Secuencia artificial

5

<220>  
<223> cebador: prm6000

<400> 34  
ggggacaagt ttgtacaaaa aagcaggctt aaacaatgga tcccgccg 49

10

<210> 35  
<211> 48  
<212> ADN  
<213> Secuencia artificial

<220>  
<223> cebador: prm6001

15

<400> 35  
ggggaccact ttgtacaaga aagctgggtg atcagctcca gaaccagg 48

<210> 36  
<211> 834  
<212> ADN  
<213> Oryza sativa

20

<400> 36

atgaagcgac ccggcgggtgc cggcggcggc ggaggcagcc catcgctcgt cacgatggct 60  
aattctagt atgatggata tggaggggtt gggatggagg cggaggggga cgtggaggag 120  
gagatgatgg cgtgcggcgg cggcggggag aagaagcggc ggctgagcgt ggagcaggtt 180  
cgcgcgctgg agcggagctt cgaggtggag aacaagcttg agcctgagcg gaaggcgcgg 240  
ctggcgcgcg acctcggcct gcagccgcgc caggctcgccg tctggttcca gaaccgccgc 300  
gcgcgggtgga agaccaagca gctcgagcgc gactacgccg cgctccgcca ttctacgac 360  
tccctgcgcc tcgatcacga cgcgctccgc cgcgacaagg acgccctcct cgccgagatc 420

aaggagctga aggcgaagct cggggacgag gaggcggcgg cgagcttcac gtcggtgaag 480  
gaggagccgg cggcctccga cgggccaccg gcggcgggat ttgggtcgtc cgacagcgac 540  
tcaagcgcgg tgctgaacga cgtggacgcg gccggcgccg cggccgcggc gacggacgcg 600  
ctggctccgg aggcgtgcac gtttctcggg gcgccgcccg ccgcggggcg gggcgcgggc 660  
gcagcggcgg cggcgagcca cgaggaggtg ttcttcacg gcaatttcct caaggtggag 720  
gaggacgaga cggggttcct cgacgacgac gagccgtgcg gcgggttctt cgccgacgat 780  
cagccccgc cgctctcgtc gtggtgggcc gaaccacgg agcactggaa ctga 834

<210> 37  
<211> 277  
<212> PRT  
<213> Oryza sativa  
<400> 37

5

Met Lys Arg Pro Gly Gly Ala Gly Gly Gly Gly Gly Ser Pro Ser Leu  
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 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 80  
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  
145 150 155 160  
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

# ES 2 555 659 T3

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  
225 230 235 240

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  
<212> ADN  
<213> Oryza sativa  
<400> 38

atggatgggg aggaggacag cgagtggatg atgatggacg ttggagggaa gggcggggaag 60  
ggcggcgggc gcggcgggcg ggcggacagg aagaagcggg tcagcgagga gcagatcaag 120  
tcgctggagt ccatgttcgc gacgcagacc aagctggagc cgaggcagaa gctgcagctc 180  
gccagggagc tcggcctgca gcctcgccag gtcgccatct gggtccagaa caagcgcgcg 240  
cgggtggaagt ccaagcagct cgagcgcgag tactccgccc tccgcgacga ctacgacgcc 300  
ctcctctgca gctacgagtc cctcaagaag gagaagctcg ccctcatcaa gcagctggag 360  
aagctggcgg agatgctgca ggagccacgg gggaagtacg gcgataatgc cggggacgac 420  
gcgcgggtcgg gcggcgtcgc cggcataaag aaggaggagt tcgtcggcgc gggcgggcgcc 480  
gccacgctct actcgtcggc cgagggtggc gggacgacga cgacgacgac ggccaagttg 540  
atgccccact tcggcagcga cgacgtcgac gcggggctct tcctccggcc gtcgtcgcag 600  
catcatccgc cgccgcccga cgccggtgcc ggcttcacgt cctccgagcc ggccgcccga 660  
caccagtcct tcaacttcca ctcgagctgg ccgtcgtcca cggagcagac ctgcagcagc 720  
acgccatggg gggaattcga gagcgagtga 750

<210> 39  
<211> 249  
<212> PRT  
<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 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 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

<210> 40  
 <211> 945  
 <212> ADN  
 <213> *Populus tremuloides*  
 <400> 40

atggcgggtg gtaccggtgg ttctaattcc aatttgctctg ttttgcttca aagccaaaga 60  
ggcccttggtg ctgcttcaca acctcttgaa tcttttttcc tttctggctc ttctccttct 120  
tttcttggtt caagatccat gatgagtttt gaagatgttc atcaagcaaa cggatcaacc 180  
aggccttttt tccgctcggt tgatcacgaa gacaatggag acgatgatct ggatgaatat 240  
tttcatcaac ctgaaaagaa gaggagactt actgttgatc aagttcagtt tcttgaaaag 300  
agttttgagc ttgagaacaa gcttgaacct gaaaggaaaa tccagcttgc aaaggatctt 360  
ggccttcagc cgcgtcaggt tgctatatgg tttcaaaacc gccgagcaag atggaagact 420  
aaacagctgg aaaaggatta tgacgttttg caatctagct acaatagcct taaggctgac 480  
tatgacaacc tcctcaagga gaaggagaaa ctaaaagctg aggttaatct tctcaccgac 540  
aagttgctcc tcaaagagaa agagaaggga atctcagaat tgtctgataa agatgcatta 600  
tcgcaagagc cacctaaaag ggctatagct gattcagctt ccgaggggtga agtgtcgaaa 660  
atctcaacag tggcctgtaa gcaggaagat atcagctcag ccaaagcgga catatttgat 720  
tcagacagcc cacattacgc tgatgggggtg cattcctcac tcttagaggc aggagattct 780  
tcatatgttt tcgaaccgga tcaatcagat ttgtcacaag atgaagaaga taactttagc 840  
aagagcttat tgctccata cgtctttccg aagcttgaag atgacgatta ctctgacccg 900  
cctgcaagtt ttgaagatca tgccttttgg tcctgggtcat actaa 945

<210> 41

<211> 314

<212> PRT

<213> *Populus tremuloides*

<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 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

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 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  
<211> 954  
<212> ADN  
<213> *Populus tremuloides*  
<400> 42

atggcggcctt gtggtggtgg tgggtggtggt tctaataccca atttgctctgt tttagttcaa 60  
agccaaagag gcccttgatgc tgctttctcaa cctcttgaag cttttttcct ttctggctct 120  
tctccttctt ttcttggttc aagatccatg atgagttttg cagatgttca ccaagcaaatt 180  
ggatcaacta gaccgttttt ccgcccatat gatcacgaag acaacggcga cgatgatattg 240  
gatgaatatt ttcatacaacc tgaaaagaag aggagactta ctgttgatca agttcagttt 300

cttgaaagaa gttttgaggt tgagaacaag cttgaacccg aaaggaaaat ccagctggcg 360  
aaggatcttg gcttgagcc tcggcaggtt gccatatggt ttcaaaaccg ccgggcaaga 420  
tggaagacga aacagcttga aaaagattat gaggttctgc aatctagcta caatggcctt 480  
aaggctgact acgacaacct cttcaaggag aaggagaaac taaaagctga ggttaatctt 540  
ctcaccaacg agttgctcct taaagagaaa gagaaaggaa gctcagaatt gtctgataaa 600  
gatgcattat ctcaagagcc acccaaaaag gcaatagccg attcagcttc agagggtgaa 660  
gtgtcgaaaa cttcaaccgt ggctgccag caggaagata ttagctcagc caaaagtgat 720  
atgtttgatt cagacagccc acattttgcg gatgggggtac attcctcact cttagaggca 780  
gggtgattctt cacatgtctt cgagcccgac caatcggatt tatcacaaga tgaagaagat 840  
aacttgagca agagtctttt gcctccgtac gtctttccaa agcttgaaga tgggtgattac 900  
tctgacccgc cagcaagttt tgaagatcat gccttttggt gctgggcata ctaa 954

<210> 43

<211> 317

<212> PRT

<213> *Populus tremuloides*

<400> 43

Met Ala Ala Cys Gly Gly Gly Gly Gly Gly Ser Asn Pro Asn Leu Ser  
1 5 10 15

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 165     | Asn Leu Phe Lys | Glu 170     | Lys Glu Lys Leu | Lys 175 | Ala                 |
| Glu Val Asn     | Leu 180     | Leu Thr Asn Glu | Leu 185     | Leu Lys Glu     | Lys 190 | Glu Lys             |
| Gly Ser         | Ser 195     | Glu Leu Ser Asp | Lys 200     | Asp Ala Leu Ser | Gln 205 | Glu Pro Pro         |
| Lys Lys         | Ala 210     | Ile Ala Asp     | Ser 215     | Ala Ser Glu Gly | Glu 220 | Val Ser Lys Thr     |
| Ser Thr Val     | Ala 225     | Cys Gln Gln Glu | Asp 230     | Ile Ser         | Ser 235 | Ala Lys Ser Asp     |
| Met Phe Asp     | Ser 245     | Asp Ser Pro His | Phe 250     | Ala Asp Gly     | Val 255 | His Ser Ser         |
| Leu Leu Glu     | Ala 260     | Gly Asp Ser Ser | His 265     | Val Phe Glu     | Pro 270 | Asp Gln Ser         |
| Asp Leu         | Ser 275     | Gln Asp Glu Glu | Asp 280     | Asn Leu Ser     | Lys 285 | Ser Leu Leu Pro     |
| Pro Tyr         | Val 290     | Phe Pro Lys     | Leu 295     | Glu Asp Gly     | Asp 300 | Tyr Ser Asp Pro Pro |
| Ala 305         | Ser Phe Glu | Asp 310         | His Ala Phe | Trp Cys         | Trp 315 | Ser Tyr             |

<210> 44  
 <211> 909  
 <212> ADN  
 <213> *Populus tremuloides*

5

<400> 44

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atggcgggtg ataaagactg tggcagttct aaaatgacca tttttcttcg aaacggcagg  | 60  |
| ctccctcctt gtgaatctct ctgtattctc acctctttta gcactcttca tgggtgcaaaa | 120 |
| tctatggtta attttaggaa tgatggagga gacactgtag acatgtcttt tttccaacca  | 180 |
| catgtcaaag aagaaagtag cgatgaggat tatgatgcgc accttaagcc atctgaaaag  | 240 |
| aaaaggcggc ttacagctgc tcaagtccag tttcttgaga agagctttga ggcggagaat  | 300 |
| aagcttgaac cagagaggaa gatgcagctt gctaaagaac tcggcttgca gcctcgccag  | 360 |
| gttgcaatat ggtttcaaaa ccgtagagct cggttcaaga acaagcagct ggaaagggac  | 420 |
| tacgactcct tgagaatcag ctttgacaaa ctcaaggctg attatgacaa actcctcctc  | 480 |
| gagaagcaga atttgaaaaa cgagcttctt tctactgaaag aaaaattgct tagcagagag | 540 |
| gaaagtatgg aaagttcaga accatttgat gtcattccatt caccggatgc agaacttgag | 600 |

cctattcctg atacagtgtc tgaaaatggt tccgccattg tgccaatggt gacacccaaa 660  
 caagaagaaa gttcagctaa aaatgatggt ttcaactcag acagcccacg ttcatttttg 720  
 gagccccgtg attgttatcg tgttttcgag tcagaccaac cagatttttc ccaagttgaa 780  
 gaagataatc tcaccaggag ctttctaccc cctccgtact ttccaaaact ctaccgagag 840  
 ccacctgcaa gttcacgtaa ttttgaattc tcagcggaag atcagccctt ttggtcctgg 900  
 atttactga 909

<210> 45

<211> 302

<212> PRT

<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  
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  
<212> ADN  
<213> Medicago truncatula

5

<400> 46

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atggcaggtg gcaagctttt tgggtggttct aatatgtcac ttttgcttca aaatgaaaga | 60  |
| ctcccttgta cttctgaagt ccttgaatct ctttgggttc acaccctgc ttcttttcaa   | 120 |
| ggttcaaatt cagtggttaa ttttgagaat ggtgggtggt gcaacagagt ggtaacagat  | 180 |
| agacccttct ttcaacaact tgagaaagaa gagaattgtg gtgatgaaga ttatgaagca  | 240 |
| tgctaccatc aacaaggaaa gaaaaggagg ctttcaagt g aacaagttca atttcttgaa | 300 |
| aagagttttg aggtagaaaa caagcttgaa cctgatagga aagttcaact tgcaaaagag  | 360 |
| cttgggttgc aaccaagaca agttgctata tggtttcaaa acagaagggc aaggttcaaa  | 420 |
| actaaacagc ttgaaaaaga ttatggcaca ttgaaagcta gctttgatag tctcaaagat  | 480 |
| gattatgata atcttcttca agagaatgac aagttaaaag aagaggtgaa ttctctcaag  | 540 |
| aacaaattga tccaagaga taaagaaaaa gtgaattcag aagacaaatc atcaccagaa   | 600 |
| gcaatcaatt cacctcataa caacatagat ccaatggata taatttcaat taciaaattca | 660 |
| gaaaatgggt ccaaaatgtc actcccta atgggtactaa aatgtaagca agaagatgcc   | 720 |
| aattcagcta aaagtgatgt gcttgattct gatagccac attgcaatga tgggaacaat   | 780 |
| ctttcttctt tcatagagcc tacagattca gatttctcac aagatgaaga ggataatgat  | 840 |
| aacttgagtc ataatctttt gactcttcct tgcttaccaa aagttgaaga tgtttgctat  | 900 |
| gatgaccac atgaaaattc ttgtaatttt gggttccctg ttgaagatca aaccttttgt   | 960 |
| ttctggcctt attga                                                   | 975 |

<210> 47

<211> 324  
 <212> PRT  
 <213> *Medicago truncatula*  
 <400> 47

```

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 Glu 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 Glu 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  
<211> 984  
<212> ADN  
<213> Phaseolus vulgaris

5

<400> 48

atggcggggtg gcaagcttca tcctggttca aacatgtcac ttctcctcca aaacgacagg 60  
ctcccttgct cctctgaagt ccttgagtct ctttgggctc acacctctaa cgctgcttcc 120  
ttccaagggtt caaaatctat ggttgatttt gagaatgtta gtgggggcag ggtgacggat 180  
aggccctttt ttcaagcgtt ggagaaggaa gataactgtg atgatgatta tgagggttgc 240  
ttccatcaac cgggtaagaa aaggaggctc acaagcgaac aagttcagtt ccttgaaagg 300  
aactttgagg tcgagaacaa gcttgaacct gaaaggaagg tccaacttgc aaaggagctt 360  
ggcttgcagc caaggcaagt ggctatatgg ttccaaaacc gaagggaag gttcaagacc 420  
aagcagctag aaaaagatta tggcacattg aaagctagct atgacagact caaagggtgac 480  
tatgaaagtc ttcttcaaga gaatgacaag taaaagcag aggtgaattc tctggagagc 540  
aaattgattc ttagagataa agagaaggag aattcggacg acaagtcac tcctgatgct 600  
gtcaattcac cccacaaaga gcctatggat ttaatttcaa attcaacatc tgaaaatggg 660  
accaaagtgt cactccctat tatggtaaca tgcaagcaag aagatgccaa ttcagccaaa 720  
agtgatgtgc ttgattcgga cagcccacat tgcactgatg ggaaccatcc ctcttcattc 780  
gtggagcctg ctgattcctc ccatgctttt gaaccagacc actccgactt ctccaagat 840  
gaagaggata atcttagtga aagccttttg accctccctt gcttaccaaa ggttgaagaa 900  
gcctgctatg atgaccctcc tgaaaaccct tgtaattttg gcttccatgt cgaggatcaa 960  
accttctgtt tctggcccta ttga 984

<210> 49  
<211> 327  
<212> PRT  
<213> Phaseolus vulgaris

10

&lt;400&gt; 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  
 50 55 60  
 Gln Ala Leu Glu Lys Glu Asp Asn Cys Asp 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 Thr Leu Lys Ala Ser Tyr Asp Arg Leu Lys Gly Asp  
 145 150 155 160  
 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  
 225 230 235 240  
 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  
305 310 315 320

Thr Phe Cys Phe Trp Pro Tyr  
325

<210> 50  
<211> 957  
<212> ADN  
<213> Lotus corniculatus

5

<400> 50

atggcgggag ggaggggtctt tagcggcggt tctgctgctc ctgcaaagtgt ttccgatacc 60  
agtctttttgc ttcagaatca acctcctgat tcttctctct tcctctctac ctctgcttct 120  
tttctcgggt caagatccat ggtgagcttc gcagataata aattagggca aacgcggctg 180  
ttcttctccg cgtttgacct cgatgagaac ggcgatgagg tcatggacga gtactttcac 240  
caatcggaga agaagcggcg tctctctgtt gaccaagttc agtttctgga gaagagcttc 300  
gaggtggata acaagctcga acctgacagg aaaaccaaga ttgccaagga ccttggtttg 360  
cagccacgcc aagtcgcaat ctggttccag aaccgccgtg cacggtggaa gacgaaacag 420  
cttgagaagg attatgattc tctgcatagt agctttgaga gtctcaaate caactatgat 480  
aatctttctca aggagaaaga catgttaaaa gctgaggtgg caagtctcac tgagaagggtg 540  
cttgcaagag agaatttgaa acaagttgaa agtgaaacaa agggatttgg tgaaccaccc 600  
caaaggcctt tacttgattc agtttcagag ggtgaagaat ctaaagttct tgttggggct 660  
tgtaaaccatg aggatatcag ttcagccagg agtgagagtt tggattctga tagcccacgt 720  
tacagggatg gatattggagt taactcagca gtgctagaga catgtgattc ttcttatgtg 780  
gttgaacctg atcaatcgga tatgtcacag gatgaggaag acaacctgac caagaccctg 840  
ttgcctccat acatgttttc caaacttggga gatatggatt actccgaccc gcctgaaagt 900  
tcatgtaatt tcggatttcc ggaggaagat catgcccttt ggtcatggtc ttactga 957

<210> 51  
<211> 318  
<212> PRT  
<213> Lotus corniculatus

10

<400> 51

ES 2 555 659 T3

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

## REIVINDICACIONES

1. Procedimiento para aumentar el rendimiento de las plantas, en plantas cultivadas en condiciones sin estrés, en relación a las plantas de tipo silvestre correspondientes, que comprende introducir y sobreexpresar en una planta, un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de SEC ID N°: 2, y seleccionar plantas que tienen rendimiento aumentado, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, número total de semillas aumentado, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, aumento del índice de cosecha (IC), aumento del peso de mil granos (PMG), aumento de la biomasa de la raíz, cada uno en relación a las plantas de tipo silvestre correspondientes.
2. Procedimiento para aumentar el rendimiento de las plantas, en plantas cultivadas en condiciones de estrés abiótico, comprendiendo dicho procedimiento introducir y sobreexpresar en una planta un ácido nucleico de hox5 de HDZip de clase I que codifica un polipéptido de hox5 de cremallera de leucina de homeodominio (HDZip) de SEC ID N°: 2, y seleccionar plantas que tienen rendimiento aumentado, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, número total de semillas aumentado, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, aumento del índice de cosecha (IC), aumento del peso de mil granos (PMG), aumento de la biomasa de la raíz, cada uno en relación a las plantas de tipo silvestre correspondientes.
3. Procedimiento de acuerdo con la reivindicación 1 o 2, en el que dicho ácido nucleico de hox5 de HDZip de clase I que se introduce y se expresa en una planta es SEC ID NO: 1.
4. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que dicho ácido nucleico de hox5 de HDZip de clase I está unido operativamente a un promotor constitutivo.
5. Método de acuerdo con la reivindicación 4, en el que dicho promotor constitutivo es un promotor de GOS2.
6. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 2 a 5, en el que dicho estrés abiótico es un estrés osmótico, seleccionado entre uno o más de los siguientes: estrés hídrico, estrés salino, estrés oxidativo y estrés iónico, en el que, preferentemente, dicho estrés hídrico es estrés por sequía.
7. Uso de una construcción que comprende
  - (i) un ácido nucleico de hox5 de HDZip de clase I como se define en una cualquiera de las reivindicaciones 1 a 3;
  - (ii) un promotor de GOS2 capaz de dirigir la expresión de la secuencia de ácido nucleico de (i), y
  - (iii) una secuencia de terminación de la transcripción,
 en el aumento del rendimiento de las plantas cultivadas en condiciones sin estrés en relación a las plantas de tipo silvestre correspondientes, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado.
8. Uso de un ácido nucleico de hox5 de HDZip de clase I como se define en una cualquiera de las reivindicaciones 1 a 3, o uso de un polipéptido de hox5 de HDZip de clase I como se define en una cualquiera de las reivindicaciones 1 a 3, en el aumento del rendimiento de las plantas cultivadas en condiciones sin estrés en relación a las plantas de tipo silvestre correspondientes, en el que dicho ácido nucleico de hox5 de HDZip de clase I se sobreexpresa en una planta, y en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, IC aumentado, PMG aumentado, cada uno en relación a las plantas de tipo silvestre correspondientes.
9. Uso de un ácido nucleico de hox5 de HDZip de clase I como se define en una cualquiera de las reivindicaciones 1 a 3, o uso de un polipéptido de hox5 de HDZip de clase I como se define en una cualquiera de las reivindicaciones 1 a 3, en el aumento del rendimiento de plantas en plantas cultivadas en condiciones de estrés abiótico, en el que dicho ácido nucleico de hox5 de HDZip de clase I se sobreexpresa en una planta, en el que dicho rendimiento aumentado se selecciona entre uno o más de: número aumentado de semillas cargadas, rendimiento total de semillas aumentado, aumento de la tasa de carga de semillas, aumento del índice de cosecha (IC), aumento del peso de mil granos (PMG), aumento de la biomasa de la raíz, cada uno en relación a las plantas de tipo silvestre correspondientes.

Homeodominio de HDZip de Clase I

|                                            |   | Hélice 1         | Hélice 2            | Hélice 3                                                        |                 |
|--------------------------------------------|---|------------------|---------------------|-----------------------------------------------------------------|-----------------|
|                                            | 1 |                  |                     |                                                                 | 60              |
| Zeama_hox5 (SEC ID N° 06)                  |   | APEKRRRLTAEQVQL  | ILERSFEEENKLEPERKTE | LARRLGMAPRQVAVWF                                                | QNRARFWKTKQ     |
| Glyma_HD157 (SEC ID N° 20)                 |   | QPGKRRRLTSEQVQH  | ILERNFEVENKLEPERKVQ | LAKELGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Lyces_VaHOX1 (SEC ID N° 28)                |   | QAEKRRRLTDNQVQH  | LEKSPGEEENKLEPERKVQ | LAKELGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Dauca_CHB3 (SEC ID N° 18)                  |   | QPEKRRRLKADQIQH  | LEKSPFETDNKLEPERKVQ | LAKELGLQPRQVALWF                                                | QNRARFWKTKT     |
| Medtr_hox16 (SEC ID N° 30)                 |   | QSEKRRRLSVDQVQH  | LEKSPFEEENKLEPERKTK | LAKDLGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Triae_hox16 (SEC ID N° 14)                 |   | LPEKRRRLTPEQVHIL | ILERSFEEENKLEPERKTE | LARKLGLQPRQVALWF                                                | QNRARFWKTKT     |
| Goshi_hox5 (SEC ID N° 24)                  |   | LPEKRRRLTSEQVYIL | LEKSPFEEENKLEPERKSQ | LAKLGLQPRQVALWF                                                 | QNRARFWKTKQ     |
| Sorbi_hox5 (SEC ID N° 12)                  |   | APEKRRRLTAEQVQL  | ILERSFEEENKLEPERKTE | LARRLGMAPRQVAVWF                                                | QNRARFWKTKQ     |
| Sacof_hox5 (SEC ID N° 10)                  |   | APEKRRRLTAEQVQL  | ILERSFEEENKLEPERKTE | LARRLGMAPRQVAVWF                                                | QNRARFWKTKQ     |
| Aqufo_hox5 (SEC ID N° 32)                  |   | LPEKRRRLTSEQVHIL | LEKSPFETENKLEPDRKTQ | LAKLGLQPRQVALWF                                                 | QNRARFWKTKQ     |
| Orysa_hox5 (SEC ID N° 02)                  |   | APEKRRRLTAEQVQM  | ILERSFEEENKLEPERKTE | LARRLGMAPRQVAVWF                                                | QNRARFWKTKQ     |
| Zeama_hox16 (SEC ID N° 08)                 |   | LPEKRRRLTPEQVLI  | ILERSFEEENKLEPERKTE | LARKLGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Orysa_hox16 (SEC ID N° 04)                 |   | LPEKRRRLTPEQVHIL | ILERSFEEENKLEPERKTE | LARKLGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Lyces_hox5 (SEC ID N° 26)                  |   | SPEKRRRLTPEQVHIL | LEKSPFETENKLEPERKTQ | LAXKLGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Arath_ATHB1 (SEC ID N° 16)                 |   | LPEKRRRLTTEQVHIL | LEKSPFETENKLEPERKTQ | LAKLGLQPRQVALWF                                                 | QNRARFWKTKQ     |
| Crapl_CPHB-5 (SEC ID N° 22)                |   | LPEKRRRLTAEQVHIL | LEKSPFEEENKLEPERKAE | LAKLGLQPRQVALWF                                                 | QNRARFWKTKQ     |
| Orysa_hox6 (SEC ID N° 39)                  |   | AADRKKRFSEEQIKS  | LESMFATQTKLEPROKLQ  | LARELGLQPRQVALWF                                                | QNRARFWKSKQ     |
| Orysa_hox4 (SEC ID N° 37)                  |   | GGEKRRRLSVEQVRL  | ILERSFEEENKLEPERKAR | LARDLGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Consenso                                   |   | PEKRRRLTAEQV     | ILERSFEEENKLEPERKT  | LARKLGLQPRQVALWF                                                | QNRARFWKTKQ     |
| Aminoácidos invariantes del homeodominio   |   | L <sub>16</sub>  |                     |                                                                 |                 |
| Aminoácidos preferidos de HDZip de Clase I |   |                  | A <sub>46</sub>     |                                                                 | W <sub>56</sub> |
|                                            |   |                  |                     | W <sub>48</sub> F <sub>49</sub> N <sub>51</sub> R <sub>53</sub> |                 |

FIGURA 1

Cremallera de leucina de HDZip de Clase I

|              |                |                                             |      |
|--------------|----------------|---------------------------------------------|------|
| Zeama_hox5   | (SEC ID N° 06) | TKQLETDYDRLKAAVDALAADHQGLADNDNLRAQVISL      | TEK  |
| Glyma_HD157  | (SEC ID N° 20) | TKQLEKDYGVWKASVDRLKSDYESLVQENDKLAKEVNSLES   | ESK  |
| Lyces_VaHOX1 | (SEC ID N° 28) | TKQLEKDYDELRNRDYLKSNYNLLKEKEILLRTEVFRLLTGK  |      |
| Dauca_CHB3   | (SEC ID N° 18) | TKTLEKDYDVLLQNSYNLKAADYNLLAEKEKLAKEVLDELTDK |      |
| Medtr_hox16  | (SEC ID N° 30) | TKQLEKDYDYLNDGYESLKTEYDNLKEKDRLQSEVASL      | TEK  |
| Triae_hox16  | (SEC ID N° 14) | TKTLEKDYDRLKASFDALRADHDALLQDNHRLRSQVVL      | TEK  |
| Goshi_hox5   | (SEC ID N° 24) | TKQLEKDYDILKSSFDLSQSNYDTLLKENEKLKSEVASL     | TEK  |
| Sorbi_hox5   | (SEC ID N° 12) | TKQLETDYDRLKAAVDALAADHQGLADNDSLRAQVISL      | TDK  |
| Sacof_hox5   | (SEC ID N° 10) | TKQLETDYDHLKAAVDALAADHQGLADNDSLRAQVVS       | TEK  |
| Aqufo_hox5   | (SEC ID N° 32) | TKQLEKDYDILKASVDLSRSDYDILLKENEKLKSEVVS      | LTGK |
| Orysa_hox5   | (SEC ID N° 02) | TKQLEHDFDRLKAAVDALAADHALLSDNDRLRAQVISL      | TEK  |
| Zeama_hox16  | (SEC ID N° 08) | TKQLEKDYDRLKASFDALRADHDALLQDNHRLRSQVVS      | TEK  |
| Orysa_hox16  | (SEC ID N° 04) | TKQLEKDYDRLKASFDALRADHDALLQDNHRLHSQVMS      | TEK* |
| Lyces_hox5   | (SEC ID N° 26) | TKQLEKDYDQLKSSYDLSLSDFDVSRKNDKLSKVVS        | MEK  |
| Arath_ATHB1  | (SEC ID N° 16) | TKQLEKDYDILKSTYDQLLSNYDSIVMDNDKLRSEVTS      | TEK  |
| Crapl_CPHB-5 | (SEC ID N° 22) | TKQLEKDYDKLKSSYDLSLSTYDSIRQENDKLAELSL       | NEK  |
| Orysa_hox6   | (SEC ID N° 39) | SKQLEKREYSALRDDYDALLCSEYSLKKEKLALIKQLEK     | LAEM |
| Orysa_hox4   | (SEC ID N° 37) | TKQLEKDYAALRHSYDLSRLDHDALRRDKDALAEIKEL      | KAK  |
| Consenso     |                | TKQLEKDYD LKASYDALRADYDALL DNDKLRAEVVS      | TEK  |
|              |                | abcdefgabcdefgabcdefgabcdefgabcdefgabcdefg  |      |

heptad1heptad2heptad3heptad4heptad5heptad6

FIGURA 1 (continuación)

|                        |     |                                                 |     |                       |
|------------------------|-----|-------------------------------------------------|-----|-----------------------|
|                        |     | 1                                               |     | 50                    |
| Zeama_hox5             | (1) | ---MDPSAVS---DSGGARRGGG-----                    | --- | AQMLLFGCGGSAN         |
| Aqufo_hox5             | (1) | ---MDSTTSRLFDGSGCHGN-----                       | --- | MLLLCGSDPV            |
| Arath_ATHB1            | (1) | ---VESNSFFDPSASHGNS-----                        | --- | ---MFFLENLNPV         |
| Orysa_hox5             | (1) | ---MDPFRVVDSGVARRACPGG-----                     | --- | AQMLLFGCGGSAN         |
| Crapl_CPHB-5           | (1) | ---MNSARIFDPSHGNMLQ-----                        | --- | ---FLGNAEGDSSV        |
| Dauca_CHB3             | (1) | ---MAGRRVFG---EGANTTSAS-----                    | --- | LLFHSQRPEPFLLSAPSPS   |
| Glyma_HD157            | (1) | ---VASEKLYAGSNMSLLQNER-----                     | --- | LPCSSEVLESLSWAQTSNPAS |
| Goshi_Hox5             | (1) | ---VESRLFPNPSTTHRN-----                         | --- | ---MLLLENTEPI         |
| Lyces_hox5             | (1) | ---MGSCHIFDPSCHGN-----                          | --- | ---MLFLCGSDPV         |
| Lyces_VaHOX1           | (1) | ---MAGPILYGGSSNFDGVFTQKQDVFSSTAPKGLGSLFAPASSSSN | --- |                       |
| Medtr_HOX16_1          | (1) | ---MAGEKLFSGSNMSLLQNER-----                     | --- | LPCTSEVLESLSWVHT--PAS |
| Orysa_hox16            | (1) | ---VESRLISTAGS---GAGQ-----                      | --- | MLFLDCGAGGGGVG        |
| Sacof_hox5             | (1) | ---MDPSAVS---NSGGARRGGG-----                    | --- | ---TQMLLFGCGGSAN      |
| Sorbi_hox5             | (1) | ---MDPSAVS---DSGGARRGGGG-----                   | --- | GAQMLLFGCGGSAN        |
| Triae_hox16            | (1) | ---MEPERLIINTSGS---GNGQ-----                    | --- | MLFMDCGAGGIA-G        |
| Zeama_hox16            | (1) | ---VESRLINAPGS---GAGQ-----                      | --- | MLFLDCGAGG--P         |
| Poptr_HOX16_1          | (1) | ---MAGTGGSNNSLSVLLQSQRG-----                    | --- | PCAASQPLESFFLSGS-SPS  |
| Poptr_HOX16_2          | (1) | MAACGGGGGGSNPNLSVLVQSQRG-----                   | --- | PCAASQPLEAFFLSGS-SPS  |
| Poptr_HOX16_3          | (1) | ---MAGDKDCGSSKMTIFLRNGR-----                    | --- | LPCESLCILTS-FST       |
| Phavu_hox16            | (1) | ---MAGEKLHPGSNMSLLQNR-----                      | --- | LPCSSEVLESLSWAHTSNAAS |
| Lotco_HOX16            | (1) | ---MAGERVFGGSAAPANVSDTS-----                    | --- | LLLQNPDPSSSLFSTAS     |
| Medtr_HOX desconocido  | (1) | MHEVAPFOANMLQTPHHDDHHQ-----                     | --- | PSSLNSILPQD           |
| Piclg_hox desconocido  | (1) | ---MACDRSALYTSSVIMNTEDN-----                    | --- | SSAHAIAAMIASSCTPPAT   |
| Orysa_hox4             | (1) | -----                                           | --- | ---MKRPGGAGG          |
| Medtr_HOX desconocido2 | (1) | -----                                           | --- | ---MKRLNNTSDS         |
| Orysa_hox6             | (1) | -----                                           | --- | -----                 |
| Consenso               | (1) | M G F                                           |     | G                     |

|                        |      |                                   |     |                         |
|------------------------|------|-----------------------------------|-----|-------------------------|
|                        |      | 51                                |     | 100                     |
| Zeama_hox5             | (32) | SN---FFRGVPMVAVLGMDDATRVG-----    | --- | KRPFITTHE---ELLEEEYDE   |
| Aqufo_hox5             | (28) | LR---SR---SFINMDSL-----           | --- | KRPFYSST---DLIEEFYDE    |
| Arath_ATHB1            | (28) | VQGGAR---SMNMETS-----             | --- | KRPFSSS---EDLYDDFYDD    |
| Orysa_hox5             | (34) | SG---FFRGVPAVAVLGMDESRSSSSAAGAGAR | --- | KRPFITTHE---ELLEEEYDE   |
| Crapl_CPHB-5           | (30) | FR---TRSS---SVLNMESS-----         | --- | LRQIFSGGGDFYDEEYDE      |
| Dauca_CHB3             | (38) | LI---S---KSMVSFQDAKRKNP-----      | --- | YDCFFMRSYDEBEIGDEYDE    |
| Glyma_HD157            | (41) | FO---S---KPVVDFENVSGSR-----       | --- | MTTRFFFALEKEBNC-DEDYEG  |
| Goshi_Hox5             | (27) | FR---AR---TMVSMENP-----           | --- | KRFLFFSSS---EDLYDEEYDE  |
| Lyces_hox5             | (27) | FR---PRS---TMMKMDSS-----          | --- | KRPFSSS---EDLYDEEYDE    |
| Lyces_VaHOX1           | (48) | FL---S---SMVSFRGVNGG-----         | --- | KRSFFDSED-QDDNEADELGE   |
| Medtr_HOX16_1          | (39) | FO---S---NVVNFENGGGNSRV-----      | --- | VTRRFFQKEKEBNCDEDEYEA   |
| Orysa_hox16            | (32) | GGAMFHRGAREVVLGMDEGG-----         | --- | RGVTRFFITTD---ELLEEEYDE |
| Sacof_hox5             | (33) | SN---FFRGVPMVAVLGMDDATRVG-----    | --- | KRPFITTHE---ELLEEEYDE   |
| Sorbi_hox5             | (35) | SN---FFRGVPMVAVLGMDDATRVG-----    | --- | KRPFITTHE---ELLEEEYDE   |
| Triae_hox16            | (31) | AA---MFHRGVRPVLGGMEEG-----        | --- | RGVTRFFITSD---DMLEEEYDE |
| Zeama_hox16            | (30) | GG---LFHRRGGRPMLGLEG-----         | --- | RGVTRFFITSD---ELLEEEYDE |
| Poptr_HOX16_1          | (41) | FL---S---RSMMSFQDVHQANG-----      | --- | STRFFRSEDHEDNG-DDDLDE   |
| Poptr_HOX16_2          | (44) | FL---S---RSMMSFADVHQANG-----      | --- | STRFFRSPDHEDNG-DDDLDE   |
| Poptr_HOX16_3          | (36) | LH---CA---KSMVNFRRNDGGDT-----     | --- | VTRMSFFQPHVKESS-DEDYDA  |
| Phavu_hox16            | (41) | FO---S---KSMVDFENVSGGR-----       | --- | VTRRFFQALEKEDNC-DEDYEG  |
| Lotco_HOX16            | (41) | FL---S---RSMVSFADNKLQ-----        | --- | TRRFFSADDLDENG-DEVMDE   |
| Medtr_HOX desconocido  | (36) | YH---G---PSFLGKRCMS-----          | --- | FSSGIELGEEANIPEDLSDD    |
| Piclg_hox desconocido  | (40) | FO---CTR---SISVFETGNERKRP-----    | --- | ACNSYSALELSDDIGEDGSD    |
| Orysa_hox4             | (10) | GG---S---PSLVTMANSS-----          | --- | IDGYGGVGMEAEGDVBEEMMA   |
| Medtr_HOX desconocido2 | (11) | FSTP-----LITISPSTEEHSPRNKHVY      | --- | GMESQSMMLDGFEBEGCVEE    |
| Orysa_hox6             | (1)  | -----MDGDEDS-----                 | --- | HWMMMDVG--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--EKKRRLTAEOVQLERSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Aqufo_hox5             | (62)  | QLP--EKKRRLTSEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Arath_ATHB1            | (64)  | QLP--EKKRRLTPEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Orysa_hox5             | (82)  | CAP--EKKRRLTAEOVQLERSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Crapl_CPHB-5           | (68)  | QLP--EKKRRLTAEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Dauca_CHB3             | (77)  | YFQPEKKRRLKADQIQFLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Glyma_HD157            | (80)  | CFHQPEKKRRLTSEOVQLERNFE    | ENKLEPERKTELARRLGMAPROVAV |
| Goshi_Hox5             | (61)  | QLP--EKKRRLTSEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Lyces_hox5             | (62)  | CSP--EKKRRLTPEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Lyces_VaHOX1           | (84)  | YLHQAEKKRRLTDNOVQFLEKSFE   | ENKLEPERKTELARRLGMAPROVAV |
| Medtr_HOX16_1          | (81)  | QYHQCEKKRRLTSEOVQLERSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Orysa_hox16            | (73)  | QLP--EKKRRLTPEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Sacof_hox5             | (74)  | CAP--EKKRRLTAEOVQLERSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Sorbi_hox5             | (76)  | CAP--EKKRRLTAEOVQLERSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Triae_hox16            | (72)  | QLP--EKKRRLTPEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Zeama_hox16            | (70)  | QLP--EKKRRLTPEOVHLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Poptr_HOX16_1          | (80)  | YFHQPEKKRRLTVDVOVLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Poptr_HOX16_2          | (83)  | YFHQPEKKRRLTVDVOVLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Poptr_HOX16_3          | (74)  | HLKPSKKRRLTAAQVQFLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Phavu_hox16            | (80)  | CFHQPEKKRRLTSEOVQLERNFE    | ENKLEPERKTELARRLGMAPROVAV |
| Lotco_HOX16            | (78)  | YFHQPEKKRRLTVDVOVLEKSFE    | ENKLEPERKTELARRLGMAPROVAV |
| Medtr_HOX desconocido  | (71)  | CSQACEKKRRLNMEQVKTLEKSFE   | ENKLEPERKTELARRLGMAPROVAV |
| Piclg_hox desconocido  | (81)  | CIHLGEKKRRLTLEOVRALEKNFE   | ENKLEPERKTELARRLGMAPROVAV |
| Orysa_hox4             | (45)  | CGGGCEKKRRLTSEOVRALEKSFE   | ENKLEPERKTELARRLGMAPROVAV |
| Medtr_HOX desconocido2 | (54)  | TGHHSKKRRLTVDVOVLEKSFE     | ENKLEPERKTELARRLGMAPROVAV |
| Orysa_hox6             | (24)  | CGGAADKKRRLTSEOVRALEKSFE   | ENKLEPERKTELARRLGMAPROVAV |
| Consenso               | (101) | EKKRRLT EOVQ LEKSFE        | ENKLEPERK OLAK LGLOPROVAV |
|                        |       | dominio conservado         |                           |
|                        |       | Homeodominio de cremallera |                           |
|                        |       | de leucina de Clase I      |                           |
|                        |       | 151                        | 200                       |
| Zeama_hox5             | (121) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Aqufo_hox5             | (110) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Arath_ATHB1            | (112) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Orysa_hox5             | (130) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Crapl_CPHB-5           | (117) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Dauca_CHB3             | (127) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Glyma_HD157            | (130) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Goshi_Hox5             | (109) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Lyces_hox5             | (110) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Lyces_VaHOX1           | (134) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Medtr_HOX16_1          | (131) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Orysa_hox16            | (121) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Sacof_hox5             | (122) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Sorbi_hox5             | (124) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Triae_hox16            | (120) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Zeama_hox16            | (118) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Poptr_HOX16_1          | (130) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Poptr_HOX16_2          | (133) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Poptr_HOX16_3          | (124) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Phavu_hox16            | (130) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Lotco_HOX16            | (128) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Medtr_HOX desconocido  | (121) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Piclg_hox desconocido  | (131) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Orysa_hox4             | (95)  | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Medtr_HOX desconocido2 | (104) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Orysa_hox6             | (74)  | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
| Consenso               | (151) | NEONRRARNTKQLEKQYDYLKASFE  | DAIRADHDAITQDNHRSQVMSLT   |
|                        |       | dominio conservado         |                           |

FIGURA 2 (continuación)

|                        |       | 201                                              | 250                                    |
|------------------------|-------|--------------------------------------------------|----------------------------------------|
| Zeama_hox5             | (171) | ERLQKETSPTS-----                                 | ATTAAQEVDPDEHTAVSGTEELL                |
| Aqufo_hox5             | (160) | CKLQVKEGAGM-----                                 | ELNQISDPPLS--TEENVDTVIT                |
| Arath_ATHB1            | (162) | EKLGKQETAN-----                                  | EPPGQVPEP----NQLDPVYI                  |
| Orysa_hox5             | (180) | EKLGKETSPTS-----                                 | SATITTAQEVDPDEHTAASTTGFA               |
| Crapl_CPHB-5           | (167) | EKLGKEDDDDP-----                                 | SAEIGRNLSSSSPFVDAEPPC                  |
| Dauca_CHB3             | (177) | DKLLREDKGS-----                                  | KTVVPDKQKVSAAFQQERVNDIS                |
| Glyma_HD157            | (180) | SKLLRDKEKEE-----                                 | NSDDKSSPDDAVNSSSPHNNKEPMDLLI           |
| Goshi_hox5             | (159) | EKLGKQDVATE-----                                 | AIAGEKDEG----LAEEMASA                  |
| Lyces_hox5             | (160) | EKLGKQVVGGA-----                                 | GGNEKSD-----ILEVDAMTI                  |
| Lyces_VaHOX1           | (184) | CKLFIKEKNG-----                                  | QLDLRDEHKHSNALAKETVVDPM                |
| Medtr_HOX16_1          | (181) | NKLLIPRDKEKV-----                                | NSEDKSS-PEAINSP--HNNIDPMDIIS           |
| Orysa_hox16            | (171) | EKLGKETTTEGSAGAAVDVPG-PAADVKVAVPDAAEPALEAAAE     |                                        |
| Sacof_hox5             | (172) | EKLGKETSPTS-----                                 | ATTAAQEVDPDEHTAASGTEKLL                |
| Sorbi_hox5             | (174) | DKLQKETSPTS-----                                 | ATTAAQEVDPDEHTAASGTEKLL                |
| Triae_hox16            | (170) | EKLGKQEAPEGSFGAAADASEPE-QAAAEAKASLADAEQAAAEAFEVV |                                        |
| Zeama_hox16            | (168) | EKLGKEDATEGGATADTAAP----                         | AVDVEASLADDVEEPAEPAATFEV               |
| Poptr_HOX16_1          | (180) | DKLLKEKEKG-----                                  | ISELSDKDALSQEPFKRAIADSAS               |
| Poptr_HOX16_2          | (183) | NEILLKEKEKG-----                                 | SSELGDKDALSQEPFKRAIADSAS               |
| Poptr_HOX16_3          | (174) | EKLLSRDESME-----                                 | SSEPFVDVHS-PDALEPIPDTVS                |
| Phavu_hox16            | (180) | SKLLRDKEKE-----                                  | NSDDKSS-PDAVNSP--HKEPMDIIS             |
| Lotco_HOX16            | (178) | EKVLARENKQ-----                                  | VESETKGLVEPPQRPPLDVS                   |
| Medtr_HOX desconocido  | (171) | NREPTESINLN-----                                 | KETEGSSSNRSENSEIKLDM                   |
| Piclg_hox desconocido  | (181) | SKSSCNDQKFOAN-----                               | SSKLQKDDQLQLLMSATKVDCA                 |
| Orysa_hox4             | (145) | AKIGDESAAS-----                                  | FTSVKEEPAASDGPFAAGFGSSDS               |
| Medtr_HOX unknown2     | (154) | SKLGEEKSTIN-----                                 | VLVKEELTMLESCEDEKHNPSSE                |
| Orysa_hox6             | (124) | EMLGEPGRKYG-----                                 | DNAGDDARSGGVAGMKKEEFVG                 |
| Consenso               | (201) | EKLQ KE                                          |                                        |
|                        |       | dominio                                          |                                        |
|                        |       | conservado                                       |                                        |
|                        |       | 251                                              | 300                                    |
| Zeama_hox5             | (206) | AQQ-----                                         | LKDNLHSSG-----                         |
| Aqufo_hox5             | (191) | MQFN-----                                        | VKVEDRLSSGSGVSAVVDDEEC                 |
| Arath_ATHB1            | (190) | NAAA-----                                        | IKTEDRLSSGSGVSAVLDDEA                  |
| Orysa_hox5             | (218) | TVDGALA-----                                     | APPPGHQPPHKDELVSSSGTNDGDDGGA           |
| Crapl_CPHB-5           | (200) | LKLT-----                                        | VKVEDRLSTGSGSAVMDGCG                   |
| Dauca_CHB3             | (212) | VGE-----                                         | VLSNSVMDCKOEDHNSVK--SDAVDSSES          |
| Glyma_HD157            | (220) | ISKNATTTTTSENGTKVLSPLPLPIMVTCCOEDANSK--SOVLDSSES |                                        |
| Goshi_hox5             | (187) | LQFS-----                                        | MKVEDRLSSGSGVSAVVDDEA                  |
| Lyces_hox5             | (187) | LQVK-----                                        | VKAGORLSSGSGSAVVDDEHS                  |
| Lyces_VaHOX1           | (219) | NVP-----                                         | ALVVKHQOEDLSSAK--SOVFDSES              |
| Medtr_HOX16_1          | (217) | ITN-----                                         | SENGSKMS--LPNMVLCKOEDANSK--SOVLDSSES   |
| Orysa_hox16            | (220) | EQQE-----                                        | QOVKAEDRLSTGSGSAVVDDEA                 |
| Sacof_hox5             | (207) | AQQ-----                                         | LKDDLHSSG-----                         |
| Sorbi_hox5             | (209) | VQQ-----                                         | LKDDLHSSG-----                         |
| Triae_hox16            | (219) | QQQ-----                                         | LHVKDEERLSPGSGSAVLDLARD                |
| Zeama_hox16            | (213) | LQ-----                                          | EVKSEERLSTGSGSAVVDDEA                  |
| Poptr_HOX16_1          | (215) | EGE-----                                         | VSKISTVACQOEDISSAK--SOIFDSSES          |
| Poptr_HOX16_2          | (218) | EGE-----                                         | VSKISTVACQOEDISSAK--SOMFDSSES          |
| Poptr_HOX16_3          | (208) | EN-----                                          | VSAIVPMVTPKQEESSAK--NDVFNSES           |
| Phavu_hox16            | (214) | N-----                                           | STSENGTKVS--LPIMVTCKOEDANSK--SOVLDSSES |
| Lotco_HOX16            | (210) | EGEE-----                                        | SKVSVGACQOEDISSAK--SESLSSES            |
| Medtr_HOX desconocido  | (204) | RTP-----                                         | ASDSPLSTHQHTTSTRTFFPPS                 |
| Piclg_hox desconocido  | (219) | DKEN-----                                        | NNEGPPSICSEGSULDMS                     |
| Orysa_hox4             | (180) | DSS-----                                         | AVLNDVDAAGAAPAATDALAPEA                |
| Medtr_HOX desconocido2 | (189) | TSNP-----                                        | SSESKDHLDYDCIINNNDVIG                  |
| Orysa_hox6             | (157) | AGG-----                                         | AATLYSSAEGGTTTTTAKL                    |
| Consenso               | (251) |                                                  | K E D S G S V D D                      |

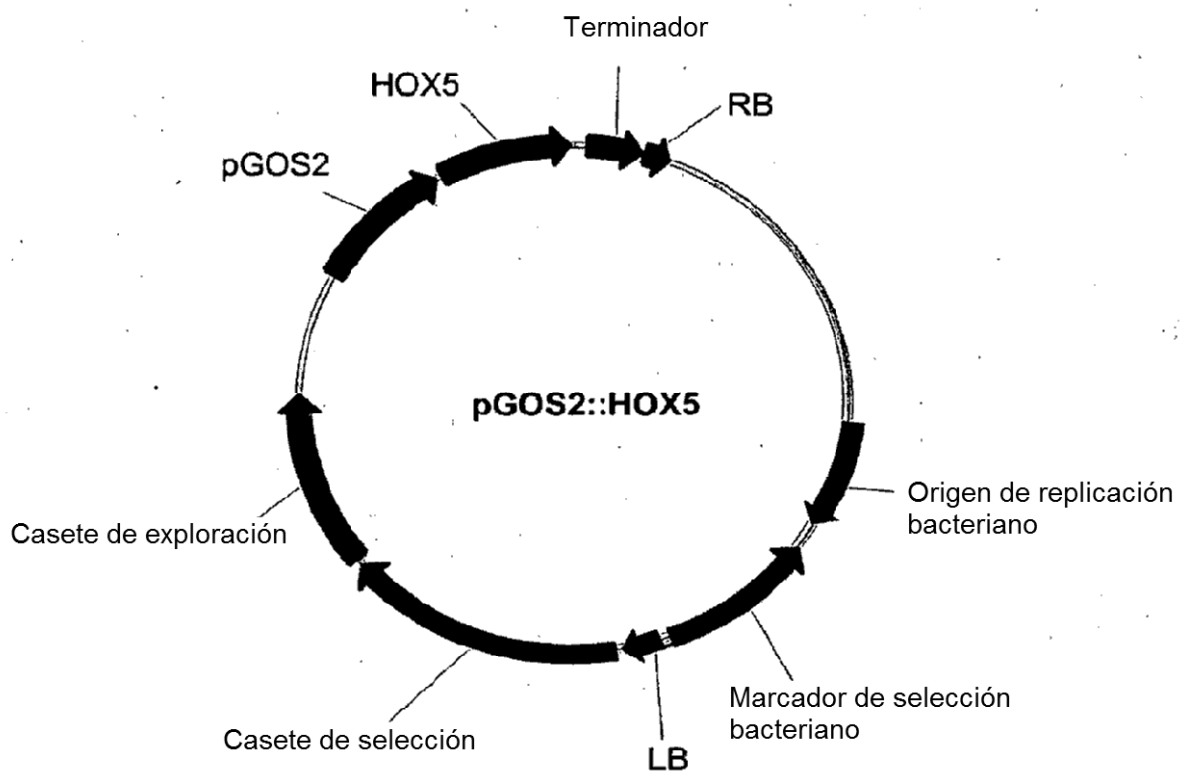
FIGURA 2 (continuación)

|                        |       |                                                      |  |     |
|------------------------|-------|------------------------------------------------------|--|-----|
|                        |       | 301                                                  |  | 350 |
| Zeama_hox5             | (218) | -----DCTGHGTLSS--SE                                  |  |     |
| Aqufo_hox5             | (216) | RQ-----LVDSVDSYFPGD---DYQCIG--PVDGVQS                |  |     |
| Arath_ATHB1            | (215) | EQ-----LLDSCDSYFP-----SIVPI                          |  |     |
| Orysa_hox5             | (255) | VVVF-----DVTEGANRLSCESAYFADAAEAYERDCAGHYALS--SE      |  |     |
| Crapl_CPHB-5           | (225) | EQQL-----LDDSGDSYFEND---EEYDCAA--ASLAAA              |  |     |
| Dauca_CHB3             | (241) | DHYS-----EVYSSFMEFVDRSYVFEPAS-----DISD               |  |     |
| Glyma_HD157            | (267) | DHCTS-----FVEPADSSHAPEPEDHSE-----DFSOD               |  |     |
| Goshi_Hox5             | (212) | EQ-----LVDSGNSYFPSD---EYSRGIG-----PFDGVQS            |  |     |
| Lyces_hox5             | (212) | SQ-----LVDSGDSYFHTD---HEEYPGPGGCNVPPMDGLCS           |  |     |
| Lyces_VaHox1           | (245) | DRYTSR---MHSSVVDQDDSAFAFETDQS-----DSSOD              |  |     |
| Medtr_HOX16_1          | (253) | DHCNDG--NNLSSFIEPTDSDPS-----D                        |  |     |
| Orysa_hox16            | (247) | QLVVGCGRQHLLAAVDSSVESYFPGG---DEYHDC--VMGPMMDHAAGGICS |  |     |
| Sacof_hox5             | (219) | -----DCTGHGALS--SE                                   |  |     |
| Sorbi_hox5             | (221) | -----DFTGHGALS--SE                                   |  |     |
| Triae_hox16            | (245) | ALLG--SGCGLAGVVDSSVDSYCFPGGAGGDEYHEC---VVGFPVAGG-IQS |  |     |
| Zeama_hox16            | (237) | LLYG---RFAAAVDSSVESYFPGGE---DHYHDCGTMGPMVNHGAGGGIQ   |  |     |
| Poptr_HOX16_1          | (244) | DHYAD---GVHSSLLEAGDSYVFEPDQS-----DLSOD               |  |     |
| Poptr_HOX16_2          | (247) | DHFAD---GVHSSLLEAGDSHVFEPDQS-----DLSOD               |  |     |
| Poptr_HOX16_3          | (236) | DRSFL-----EPRDCYRVFESDQP-----DFSOD                   |  |     |
| Phavu_hox16            | (249) | DHCTDG--NHPSSFVEPADSSHAPEPDHS-----DFSOD              |  |     |
| Lotco_HOX16            | (239) | DRYRDG--YGVNSAVLETCDSYVVEPDQS-----DMSOD              |  |     |
| Medtr_HOX desconocido  | (228) | ARPS-----SGIAQLFQTSRPE-----IDC                       |  |     |
| Piclg_hox desconocido  | (243) | DGTIDS-----QQNIDSIGFSN-----VKA                       |  |     |
| Orysa_hox4             | (206) | CTFLGAPPAAGAGAGAAAAAHEEVFFHG-----NFLKV               |  |     |
| Medtr_HOX desconocido2 | (215) | ETSSLFPVDLKGSSSDSDSAISSGVLSQ-----QHLLL               |  |     |
| Orysa_hox6             | (181) | MPHFG-----SDDVDAGLFLRPS-----SQHHP                    |  |     |
| Consenso               | (301) | P DS Q                                               |  |     |
|                        |       | 351                                                  |  | 400 |
| Zeama_hox5             | (229) | EDGGVVSDEGCS-----FALPDAMFAAGFTTHG-----AEEVQ          |  |     |
| Aqufo_hox5             | (244) | EDDIISDDS---R-----SYFSDVFPAPEQNHQ-----ES-E           |  |     |
| Arath_ATHB1            | (232) | QDNNSASDHNDNR-----SCFADVFVPTTSPSHD-----HHG-E         |  |     |
| Orysa_hox5             | (296) | EDGGAVSDEGCS-----FDLPAAAAAAMFGAAGVVHDAADDEEAQ        |  |     |
| Crapl_CPHB-5           | (254) | KEDGGSDEGG-----CYFTEALAAEEEE-----A                   |  |     |
| Dauca_CHB3             | (272) | EDDMGNLFLPS---YHVFSKTEGYSYSDQPSNSSF-----GFPVE        |  |     |
| Glyma_HD157            | (295) | EDNLSENLLMTFPSSCCLPKVEEHICYDGPENSCNF-----GFQVE       |  |     |
| Goshi_Hox5             | (240) | EDDGSDNCG-----SYFSDVFATTEQG-----                     |  |     |
| Lyces_hox5             | (247) | EDDGSDDHGSC---GYFSNVFVAEEQHHEQ-----GE-E              |  |     |
| Lyces_VaHox1           | (276) | DDENFSKNMLSTAN---LLGKDAEDDYPATSSNLSYFG-----FPVE      |  |     |
| Medtr_HOX16_1          | (276) | EDNDNLSHNLLTLP---CLPKVEDVCYDDPHENSCNF-----GFPVE      |  |     |
| Orysa_hox16            | (292) | EDDGAGSDEGCS-----YYADAGVLFADHGHHHHHQHADDDEEDGQQ      |  |     |
| Sacof_hox5             | (230) | EDGGVVSDEG-S-----FDLPAMFAAGVTHGA-----DAEEAQ          |  |     |
| Sorbi_hox5             | (232) | EDGGVVSDEGCS-----FDLPAMFAAGVTHG-----AEEAQ            |  |     |
| Triae_hox16            | (290) | EDDGAGSDEGCS-----YYPDAAEVFAAAQGHGHHRTDDDDQDDGQ       |  |     |
| Zeama_hox16            | (280) | SDDGAGSDEGCS-----YYADEAAAAAAFFAGHATHHHHADEDEDAGQ     |  |     |
| Poptr_HOX16_1          | (275) | EDNFSKSLPP---YVFPKLEDYSDPPAS-----FE                  |  |     |
| Poptr_HOX16_2          | (278) | EDNLSKSLPP---YVFPKLEDYSDPPAS-----FE                  |  |     |
| Poptr_HOX16_3          | (260) | EDNLTRSLPP-----PYFPKLYREPPASSRNF-----EFSAE           |  |     |
| Phavu_hox16            | (281) | EDNLSESLTLP---CLPKVEEACYDDPPENPCNF-----GFHVE         |  |     |
| Lotco_HOX16            | (272) | EDNLTKTLLPPY---MFSKLGMDYSDPPESSCNF-----GFPEE         |  |     |
| Medtr_HOX desconocido  | (249) | QKIDQMVKEES-----LSNMFCGMDQAG-----                    |  |     |
| Piclg_hox desconocido  | (263) | RDRLLECNRFP-----KVEENVSADEPCNYLFYN-----NLETGP        |  |     |
| Orysa_hox4             | (240) | EDDETGLDDDEP-----CGGFFADDQPPP-----LSSWW              |  |     |
| Medtr_HOX desconocido2 | (251) | SPSSSMNCFQYQKSYHVKMEHNFLSADEACN-----FFSDE            |  |     |
| Orysa_hox6             | (204) | PPPHAGAGFTSSE-----PAADHQSNFHSWPSS-----TEQTCS         |  |     |
| Consenso               | (351) | EED D E                                              |  |     |

FIGURA 2 (continuación)

| Cola de Trp                  |                 |     |
|------------------------------|-----------------|-----|
|                              | 401             | 414 |
| Zeama_hox5 (263)             | IANWTSMFVN----  |     |
| Aqufo_hox5 (273)             | TLGWWDNA-----   |     |
| Arath_ATHB1 (265)            | SLAFWGP-----    |     |
| Orysa_hox5 (340)             | IGSWTAMFWS----  |     |
| Crapl_CPHB-5 (278)           | HFAWCIMS-----   |     |
| Dauca_CHB3 (312)             | DHTFGFVGTTEL--- |     |
| Glyma_HD157 (337)            | IQTFCFMPY-----  |     |
| Goshi_Hox5 (263)             | ALGLWAX-----    |     |
| Lycès_hox5 (279)             | HIGWFWS-----    |     |
| Lyces_VaHOX1 (315)           | DQGFVFTY-----   |     |
| Medtr_HOX16_1 (316)          | IQTFCFMPY-----  |     |
| Orysa_hox16 (336)            | ISCWWMNN-----   |     |
| Sacof_hox5 (265)             | IANWTSMFVN----  |     |
| Sorbi_hox5 (266)             | IANWTSMFVN----  |     |
| Triae_hox16 (334)            | IS-YWMNN-----   |     |
| Zeama_hox16 (324)            | IS-WWMNN-----   |     |
| Poptr_HOX16_1 (306)          | DHAFWWSY-----   |     |
| Poptr_HOX16_2 (309)          | DHAFWWSY-----   |     |
| Poptr_HOX16_3 (294)          | DQPFWSYIY-----  |     |
| Phavu_hox16 (319)            | IQTFCFMPY-----  |     |
| Lotco_HOX16 (310)            | DHALWWSY-----   |     |
| Medtr_HOX desconocido (273)  | ---FWPQLEQQHFN  |     |
| Piclg_hox desconocido (299)  | LLWDYNTSSGL---  |     |
| Orysa_hox4 (270)             | AEPTEHNN-----   |     |
| Medtr_HOX desconocido2 (289) | CAPTLOMYCPDQWS  |     |
| Orysa_hox6 (240)             | STPWWEFESE----  |     |
| Consenso (401)               | W               |     |

FIGURA 2 (continuación)



**FIGURA 3**

**SEC ID N°: 1 Secuencia de ADNc XM\_482406.1 de Orysa hox5 de *Oryza sativa***

ATGGATCCCGGCCGCGTCTGTTCGACTCCGGCGTGGCGCGGCGGGCGTGCCCCGGCGCGCG  
GCAGATGCTTCTCTTCGGCGCGCGCGGCAGCGCCAACAGCGCGGCTTCTTCCGAGGCGTG  
CGGCGCGCGTGTGTTGGGATGGATGAATCGCGGTCGTCTGTCTCGGCGCGGCGGGGGCGGGGGCG  
AAGCGGCCGTTCTTACGACGACGAGGAGCTCCTGGAGGAGGAGTACTACGACGAGCAGGC  
GCCGGAGAAGAAGCGGCGGCTGACGGCGGAGCAGGTGCAGATGCTGGAGCGGAGCTTCGAGG  
AGGAGAACAAGCTGGAGCCCGAGCGGAAGACGGAGCTCGCCCGCCGCTCGGCATGGCCCCC  
CGGCAGGTCGCCGTCTGGTTCCAGAACC CGCGCGCCCGCTGGAAGACCAAGCAGCTCGAGCA  
CGACTTCGACCGCCTCAAGGCCGCTACGACGCCCTCGCCGCGGACCAACCATGCCCTCCTCT  
CCGACAACGACCGCCTCCGCGCGCAGGTAATCTCATTAAACCGAGAAGCTGCAAGACAAGGAG  
ACGTCGCCGTCGTCTGGCGACCATCACCACCGCGCGCAGGAGGTCGACCAGCCGGACGAACA  
CACGGAGGCCGCGTCAACCAACCGGCTTCGCCACCGTCGACGGCGCATTTGGCGGCGCCACCGC  
CCGGCCACCAGCAGCCGCGCCGATAAAGATGATCTTGTGAGCAGCGGCGGCACCAACGACGAC  
GGCGATGGCGGCGCGGCCGTGGTGGTCTTCGACGTCACCGAGGGCGCCAACGACCGCCTCAG  
CTGCGAGTCGGCGTACTTTCGCCGACGCGCGGAGGCGTACGAGCGCGACTGCGCCGGGCACT  
ACGCCCTCTCGTCGGAGGAGGAGGACGGCGCGCGGTGAGCGACGAGGGCTGCAGCTTCGAC  
CTCCCCGACGCGCGCGCGCGCGCGCCGCGCCATGTTCCGGCGCGCGCGGAGTTGTGCACCA  
CGCGCGGACGACGAGGAGGCGCAGCTCGGCAGCTGGACCGCCTGGTTCTGGAGCTGA

**SEC ID N°: 2 Secuencia de aminoácidos traducida de Orysa\_hox5 de *Oryza sativa***

MDPGRVVFDSGVARRACPGGAQMLLFGGGGSANSGGFFRGVPAAVLGMDSESRSSSSAAGAGA  
KRPFFTTHEELLEEEYYDEQAPEKKRRLTAEQVQLMERSFEEENKLEPERKTELARRLGMAP  
RQVAVWFQNRARRWKTQLEHDFDRLKAAYDALAADHHALLSDNDRRLRAQVLSLTEKLQDKE  
TSPSSATITTAQEVDPDEHTEAASTTGATVDGALAAPP PGHQPPHKDDLVS SGGTND  
GDGGA AVVFDVTEGANDRLSCESAYFADAAEAYERDCAGHYALSSEEDGGAVSDEGCSFD  
LPDA AAAAAAMFGAAGVVHDAADDEEAQLGSWTAWFWS

**SEC ID N°: 3 Secuencia de ADNc XM\_467603.1 de Orysa\_hox16 de *Oryza sativa***

ATGGAGTCCGGCCGGCTCATCTTCAGCACGGCGGGCTCCGGCGCCGGGCAGATGCTCTTCTT  
GGA CTGCGGCGCTGGCGGCGGCGGCGTCCGGCGCGGGCCATGTTCCATCGAGGGGCGAGAC  
CGGTGCTCGGCATGGAGGAAGGAGGGCGCGGCGTCAAGCGGCCCTTCTTCACCACCCCGAC  
GAGCTCCTCGAAGAGGAGTACTACGACGAGCAGCTCCCGGAGAAGAAGCGGCGCCTCACGCC  
GGAGCAGGTGCATCTGCTGGAGAGGAGCTTCGAGGAGGAGAACAAGCTGGAGCCGGAGCGGA  
AGACGAGCTGGCGCGGAAGCTAGGGCTGCAGCCGCGGCAGGTGCGCGTGTGGTTCCAGAAC  
CGCCGCGCGCTGGAAGACCAAGCAGCTCGAGCGCGACTTCGACCGCCTCAAGGCGTCGTT  
CGACGCCCTCCGCGCGGACCACGACGCCCTCCTCCAGGACAACCACCGCCTCCACTCTCAGG  
TCATGTCGTTGACCGAGAAGCTGCAAGAGAAGGAGACGACGACCGAGGGCAGCGCCGCGCG  
GCCGTTGACGTCCCGGGCTTGCTTGCCTGCGGCGGCGGACGTGAAGGTGCGCGTCCCGGACGCCGA  
GGAACCGGCGCTGGAGGAGGCGGCGGCGGCGTTCGAGGAGCAGCAGGAGCAGCAGGTGAAGG  
CCGAGGACAGGCTGAGCACGGGCGAGCGGCGGGAGCGCGGTGGTGGACACGGACGCGCAACTG  
GTGGTCGGGTGCGGCCGGAAGCATCTCGCCGCGGTGGACAGCAGCGTGGAGTCGTACTTCC  
CGGGCGGCGACGAGTACCACGACTGCGTGATGGGCCCCATGGACCACGCCGCGGGGGGCATC  
CAGTCGGAGGAGGACGACGGCGCGGCGAGCGAGGGCTGCAGCTACTACGCCGACGACGC  
CGGCGTCCTCTTCGCCGACCACGGCCACCACCACCACCAACACGCGGACGACGAGG  
AGGACGGCCAGCAGATCAGCTGCTGGTGGATGTGGAAGTATAGATTTCTCGCGCGCGCGCTCG  
TCGTGCATTCAATTCTCGTGTAAAAAATCGTTCTCTTTTCATTTTCCGCTTCTTTGTC  
TGTAATGTTGAGTTTCGATCGGCTATGAGAAGGAAGGAGGTGTATGCATGTGCATGGTATGG  
TAGGGTAACACATCGGTGA

FIGURA 4 (continuación)

**SEC ID N°: 4 Secuencia de aminoácidos traducida de *Orysa\_hox16* de *Oryza sativa***

MESGRLLIFSTAGSGAGQMLFLDCGAGGGGVGGGAMFHRGARPVLGMEEGGRGVKRPFFFTTPD  
ELLEEEYYDEQLPEKKRRLTPEQVHLLERSFEEENKLEPERKTELARKLGLQPROVAVWFQ  
RRARWKTQLERDFDRLKASFDALRADHDALLQDNHRLHSQVMSLTEKLQEKETTTEGSAGA  
AVDVPGLPAAADV KVAVPDAEPALEEAFAEEQQEQVKAEDRLSTGSGGS AVVD TDAQ  
LVGCGRQH LA AVDSVESYFPGGDEYHDCVMGPM DHAAGGIQSEEDDGAGSDEGCSY YADDA  
GVL FADHGH HHHHQHADDDEEDGQ QISCWWMWN

**SEC ID N°: 5 Contigo de secuencia de ADNc desde esencialmente C0458693 y DV024016 de *Zeama\_hox5* de *Zea mays***

ATGGATCCGAGCGCGGTTCAGTTTCGACTCTGGCGGCGCGCGCGGGCGCGCGCGCGCAGAT  
GCTGCTCTTCGGCGGCGGAGGACGCGCAACAGCAACGGCTTCTTCGAGGTGTTCCGATGG  
CGGTCTTCGGGATGGAGCGACGCGCGTGGGCAAGCGGCCCTTCTTCACGACACACGAG  
GAGCTCCTAGAGGAGGAGTACTACGACGAGCAGGCGCCGAGAAGAAGCGCCGACTGACGGC  
GGAGCAGGTGCAGCTGCTGGAGCGGAGCTTCGAAGAAGAGAACAAGCTGGAGCCGGAGCGCA  
AGACCGAGCTGGCTCGCCGCTGGGGATGGCGCCCCGCCAGGTAGCTGTTTGGTTCCAGAAC  
CGCCGCGCGCTGGAAGACCAAGCAACTCGAGACCGACTATGACCGCCTCAAGGCTGCTTA  
CGACGCACTCGCGCCGACCACAGGGCCTCTGGCCGACAACGATAACCTCCGGGCACAGG  
TGATCTCCCTGACGAGAAGCTGCAAGGCAAGGAGACATCCCCGTGCGCAACCACTGCTGCC  
CAAGAGGTGACACAGCCAGACGAACACACCGCTGTGTGTCAGGCACGGAAGAAGTGTGGCGCA  
GCAGCTCAAGGACAACCTCCACAGCAGCGGCGACTGCACTGGCCATGGCACCTCTCTTCGG  
AAGAAGACGACGGTGGCGTGGTCAGTGACGAGGGCTGCAGCTTCGCTCTCCCGGATGCCATG  
TTCGCTGCCGGGTTACCCACCATGGCGCCGAGGAGGTGCAGCTGGCCAACCTGGACATCCAT  
GTTCTGGA ACTGA

**SEC ID N°: 6 Secuencia de aminoácidos traducida de *Zeama\_hox5* de *Zea mays***

MDPSAVSFDSGGARRGGGAQMLLFGGGGSANSNGFFRGVFM AVLGMDDATRVGKRPF FTTHE  
ELLEEEYYDEQAPEKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPROVAVWFQ  
RRARWKTQLETDYDRLKAA YDALADHQLLADNDNLRAQV I SLTEKLQ GKETS P SATTA  
QEVDQPD EHTAVSGTEELLAQQLKDN LHSSGDCTGHG TLSSEEDDG VVSDEGCSFALPDAM  
FAAGFTHHGAEVQLANW TSMFWN

**SEC ID N°: 7 Secuencia de ADNc AY105265 de *Zeama\_hox16* de *Zea mays***

ATGGAGTCTGGACGGCTCATCTTCAACGCGCCGGGCTCTGGCGCCGGGCAGATGCTCTTCCT  
CGACTGCGGCGCAGGCGGCGGTCCCGGCGGCGGCTTGTTCATCGAGGCGGGAGACCGATGC  
TTGGCCTTGAAGAAGGGCGCGGCGTAAACGGCCCTTCTTCACCTCGCCCGACGAGCTCCTC  
GAGGAAGAGTACTACGACGAGCAGCTGCCGGAGAAGAAGCGCCGCTCACCCAGAGCAGGT  
GCTTCTGCTGGAGAGGAGCTTCGAGGAGGAGAACAAGCTGGAGCCGGAGCGCAAGACGGAGC  
TGGCGCGCAAGCTGGGCCTGCAGCCTCGCCAGGTGGCCGTCTGGTTCCAGAACCGCCGCGCC  
CGGTGGAAGACCAAGCAGCTCGAGCGGACTTCGACCGCCTCAAGGCCTCCTTCGACGCTCT  
CCGAGCGGACACGACGCCCTCCTCCAGGACAACAACCGCCTCCGCTCACAGGTTGTGTGCT  
TGACCGAGAAGCTGCAAGAGAAGGAGGATGCGACGAGGCGGCGCCACCGCTGACACCGCC  
GCGCCGGCGGTTGGACGTCGAGGCTTCCCTGGCCGACGACGTCGAGGAGCCAGCAGAGCCTGC  
GGCGACGTTTCGAGGTGCTGCAGGAGGTGAAGTCCGAGGACAGGCTGAGCACCGGCAGCGGCG  
GGAGCGCGGTGGTGGACGCGGACGCGCTGCTGTACGGCAGGTTCGCCGCGGCAGTTGATAGC  
AGCGTGGAGTCGTACTTCCCCGGCGGCGAGGACCACTACCACGACTGCGGGACGATGGGCC

FIGURA 4 (continuación)

CGTGAATCATGGCGCCGGAGGAGGCATCCAGTCGGACGACGACGGCGCCGGCAGCGACGAGG  
GGTGCAGCTACTACGCCGACGAAGCCGCCGCCGCCGCCGCGTTCTTCGCCGGACACGCC  
ACCCACCACCACGCGGACGAGGACGAGGACGCCGCCAGATCAGCTGGTGGATGTGGAAC  
TAA

**SEC ID N°: 8** Secuencia de aminoácidos traducida de *Zeamax\_hox16* de *Zea mays*  
MESGRLLFNAPGSGAGQMLFLDCGAGGGPGGGLFHRGGRPMLGLEEGRGVKRPFFTSPELL  
EEYYDEQLPEKKRRLTPEQVLLERSFEEENKLEPERKTELARKLGLQPRQVAVWFQNRRA  
RWKTKQLERDFDLKASFDALRADHDALLQDNNRLRSQVVSLEKLQEKEDATEGGATADTA  
APAVDVEASLADDVEEPAEPAATFEVLQEVKSEDRLSTGSGGSAVVDADALLYGRFAAAVDS  
SVESYFPGGEDHYHDCGTMGPNVNHAGGGIQSDDDGAGSDEGCSYYADEAAAAAAFFAGHA  
THHHADEDEDAGQISWMMWN

**SEC ID N°: 9** Contigo de secuencia de ADNc desde esencialmente CA088615, CA115362 y CA142506 de *Sacof\_hox5* de *Saccharum officinarum*

ATGGATCCGAGCGCGGTCACTTTCAACTCCGGCGGCGCGCGCGGGCGGCGGCGGCACGCA  
GATGCTGCTCTTCGGCGGCGGAGGACGCGCAACAGCAACGGCTTCTTCGAGGTGTTCGGA  
TGGCGGTCTTGGGCATGGACGACGCGACGCGCGTGGGCAAGCGGCCCTTCTTCACCACACAC  
GAGGAGCTCCTGGAGGAGGAGTACTACGACGAGCAGGCGCCCGAGAAGAAGCGCCGTCTGAC  
GGCGGAGCAGGTGCAGCTGCTGGAGCGGAGCTTCGAGGAAGAGAACAAGCTGGAGCCCGAGC  
GCAAGACCGAGCTGGCTCGCCGCCTCGGGATGGCGCCCCGCCAGGTGGCCGTCTGGTTCCAG  
AACCGCCGCGCGCTGGAAGACCAAGCAGCTCGAGACCGACTATGACCACCTCAAGGCTGC  
CTACGACGCGCTCGCCGCCGACCACCAGGGCCTCCTGGCCGACAACGATAGCCTCCGGGCAC  
AGGTGGTCTCCCTAACAGAGAAGCTGCAAGGCAAGGAGACATCCCCGTCGCGCCACCACTGCT  
GCCCAAGAGGTGACCCAGCCAGACGAACACACCGCGGCGTCAGGCACTGAGAACTGCTGGC  
GCAGCAGCTCAAGGACGACCTCCACAGCAGCGGCGACTGCACTGGCCATGGTGGCCCTCTCCT  
CAGAGGAAGAAGATGGTGGTGTGGTCACTGACGAGGGCAGCTTTGATCTCCCGGATGCCATG  
TTTGCTGCCGGGTACCCACCATGGCGCCGACGCGGAGGAGGCACAGCTGGCCAACTGGAC  
ATCCTGGTCTGGAACCTGA

**SEC ID N°: 10** Secuencia de aminoácidos traducida de *Sacof\_hox5* de *Saccharum officinarum*  
MDPSAVSFNSGGARRGGGTQMLLFGGGGSANSNGFFRGVPMVLMDDATRVGKRPFFTTH  
EELLEEEYYDEQAPEKKRRLTAEQVQLLERSFEEENKLEPERKTELARRLGMAPRQVAVWFQ  
NRRARWTKQLETDYDHLKAAYDALADHQLLADNDSLRAQVVSLEKLQKETSPPSATT  
AQEVDQPDEHTAASGTEKLLAQQLKDDLHSSGDCTGHGALSSEEDGGVVSDEGSFDLPDAM  
FAAGVTHHGADAEAAQLANWTSWFWN

**SEC ID N°: 11** BE3 63386, CD432381 de *Sorbi\_hox5* de *Sorghum bicolor*  
ATGGATCCGAGCGCGGTCACTTTGACTCCGGCGGCGCGCGCGGGCGGCGGCGGCGGCGG  
CGCGCAGATGCTGCTCTTCGGCGGCGGAGGACGCGCAACAGCAACGGCTTCTTCGAGGTG  
TTCCGATGGCGGTCTTGGGCATGGACGACGCGACGCGCGTGGGCAAGCGGCCCTTCTTCACC  
ACGCACGAGGAGCTCCTGGAGGAGGAGTACTACGACGAGCAGGCGCCCGAGAAGAAGCGCCG  
TCTGACGGCGGAGCAGGTGCAGCTGCTGGAGCGGAGCTTCGAGGAAGAGAACAAGCTGGAGC  
CGGAGCGCAAGACCGAGCTGGCTCGCCGCCTCGGGATGGCGCCTCGCCAGGTGGCCGTCTGG  
TTCCAGAACCAGCGCGCGCTGGAAGACTAAGCAGCTCGAGACCGACTATGACCGCCTCAA  
GGCTGCCTACGACGCGCTCGCCGCCGACCACCAGGGCCTCCTGGCCGACAACGATAGCCTCC

FIGURA 4 (continuación)

GGGCACAGGTGATCTCCCTAACGGATAAGCTGCAACGCAAGGAGACATCCCCGTCGGCGACC  
 ACTGCTGCCCCAAGAGGTGACCCAGCCAGACGAACACACCGCTGCGTCAGGCACTGAGAACT  
 GCTGGTGACGACGCTCAAGGACGACCTCCACAGCAGCGGCGACTTCACTGGCCATGGTGCCC  
 TCTCTTCAGAGGAAGAGGATGGTGGTGTGGTCAGCGACGAGGGCTGCAGCTTTGATCTCCCG  
 GATGCCATGTTGCTGCCGGGTACCCACCATGGCGCCGAGGAGGCGCAGCTGGCCAACTG  
 GACATCCTGGTTCTGGAAGTGA

**SEC ID N°: 12** Secuencia de aminoácidos traducida de Sorbi\_hox5 de *Sorghum bicolor*  
 MDPSAVSFDSGGARRGGGGGAQMLLFGGGGSANSNGFFRGVPMVLMDDATRVGKRPFFT  
 THEELLEEEYYDEQAPEKKRRLTAEQVOLLERSFEEENKLEPERKTELARRLGMAPRQVAVW  
 FQNRRLRWKTKQLETDYDRLKAAYDALADHQGLLADNDSLRAQVLSLTDKLRKETSPTSAT  
 TAAQEVDQPDHTAASGTEKLLVQQLKDDLHSSGDFTHGHGALSSEEDDGGVVSDEGCSFDLP  
 DAMFAAGVTHHGAEAAQLANWTSWFWN

**SEC ID N°: 13** DR735359, DR741379, CD916488 de Triae\_hox16 de *Triticum aestivum*  
 ATGGAGCCCCGGCCGGCTCATCTTCAACACGTCGGGCTCCGGCAACGGACAGATGCTCTTCAT  
 GGAAGTGGCGCGCGGGCGGCATCGCCGGCGCGGCGGCATGTTCCATCGAGGGGTGAGACCGG  
 TCCTCGGCGGCATGGAAGAAGGGCGCGGCGTGAAGCGGCCCTTCTTCACCTCGCCGGATGAC  
 ATGCTCGAGGAGGAGTACTACGACGAGCAGCTCCCGGAGAAGAAGCGGCGCCTCACCCCGGA  
 GCAGGTCCACCTGCTGGAGAGGAGCTTCGAGGAGGAGAAACAAGCTGGAGCCGGAGAGGAAGA  
 CGGAGCTGGCCCGCAAGCTCGGGCTGCAGCCACGCCAGGTGGCCGTCTGGTTCCAGAACCGC  
 CGCGCCCGGTGGAAGACAAAGACGCTGGAGCGCGACTTCGACCGCCTCAAGGCGTCTCTCGA  
 CGCCCTCCGGGCGGACACGACGCGCTCTCCAGGACAACCACCGGCTCCGGTACAGGTGG  
 TAACGTTGACCGAGAAGATGCAAGATAAGGAGGCGCCGGAAGGCAGCTTCGGTGCAGCCGCC  
 GACGCTTCGGAGCCGAGCAGGCGGCGGCGGAGGCGAAGGCTTCTTGGCCGACGCCGAGGA  
 GCAGGCCGCGGCAGCGGAGGCGTTCGAGGTGGTGCAGCAGCAGCTGCACGTGAAGGACGAGG  
 AGAGGCTGAGCCCGGGGAGCGGCGGGAGCGCGGTGCTGGACGCGAGGGACGCGCTGCTCGGG  
 AGCGGATGCGGCCTCGCCGGCGTGGTGGACAGCAGCGTGGACTCGTACTGCTTCCCGGGGGG  
 CGCCGGCGGCGACGAGTACCACGAGTGCCTGGTGGGCCCCGTGGCGGGCGGCATCCAGTCGG  
 AGGAGGACGACGGCGCGGCGAGCGACGAGGGCTGCAGCTACTACCCGACGACGCCGCCGTC  
 TTCTTCGCCGCCGCGCAAGGGCACGGCCACCATCGCACGGACGACGATCAGCAGGACGA  
 CGGCCAGATCAGCTACTGGATGTGGAAGTGA

**SEC ID N°: 14** Secuencia de aminoácidos traducida de Triae\_hox16 de *Triticum aestivum*  
 MEPGRLLIFNTSGSNGQMLFMDCGAGGIAGAAGMFHRGVRPVLGGMEEGRGVGRPFPTSPDD  
 MLEEEYYDEQLPEKKRRLTPEQVHLLERSFEEENKLEPERKTELARKLGLQPRQVAVWFQNR  
 RARWKTKTLERDFDRLKASFDALRADHDALLQDNHRLRSQVVTLEKMQDKEAPEGSFGAAA  
 DASEPEQAAAEAKASLADAEQAAAAEAFEVVQQLHVKDEERLSPGSGGSAVLDARDALLG  
 SGCGLAGVVDSSVDSYCFPGGAGGDEYHECVVGPVAGGIQSEEDDGAGSDEGCSYYPDDAAV  
 FFAAAQGHGHRTHDDDDQDDGQISYWMWN

FIGURA 4 (continuación)

**SEC ID N°: 15 Secuencia de ADNc X58821 de Arath\_ATHB1 de *Arabidopsis thaliana***

**ATGGAATCCAATTCGTTTTTCTTCGATCCATCTGCTTCACACGGCAACAGCATGTTCTTCCT**  
**TGGGAATCTCAATCCCGTCGTCCAAGGAGGAGGAGCAAGATCGATGATGAACATGGAGGAAA**  
**CTTCGAAGCGAAGGCCCTTCTTTAGCTCCCCTGAGGATCTCTACGACGATGACTTTTACGAC**  
**GACCAGTTGCCTGAAAAGAAGCGTCGCCTCACTACCGAACAAGTGCATCTGCTGGAGAAAAG**  
**CTTCGAGACAGAGAACAAGCTAGAGCCTGAACGCAAGACTCAGCTTGCCAAGAAGCTTGCTC**  
**TACAGCCAAGGCAAGTGGCTGTCTGGTTTCAGAATCGCCGAGCTCGTTGGAAAACAAAACAG**  
**CTTGAGAGAGACTACGATCTTCTCAAGTCCACTTACGACCAACTTCTTTCTAACTACGACTC**  
**CATCGTCATGGACAACGATAAGCTCAGATCCGAGGTTACTTCCCTGACCGAAAAGCTTCAGG**  
**GCAAACAAGAGACAGCTAATGAACCACCTGGTCAAGTGCCCCGAACCAACCAACTTGATCCG**  
**GTTTACATTAATGCGGCAGCAATCAAAACCGAGGACCGGTTAAGTTCAGGGAGCGTTGGGAG**  
**CGCGGTACTAGACGACGACGCACCTCAACTACTAGACAGCTGTGACTCTTACTTCCCAAGCA**  
**TCGTACCCATCCAAGACAACAGCAACGCCAGTGATCATGACAATGACCGGAGCTGTTTCGCC**  
**GACGTCTTTGTGCCCACTTCACCGTCGCACGATCATCACGGTGAATCATTGGCTTTCTG**  
**GGGATGGCCTTAG**

**SEC ID N°: 16 Secuencia de aminoácidos traducida de Arath\_ATHB1 de *Arabidopsis thaliana***

**MESNSFFFDPSASHGNSMFFLGNLNPVVQGGGARSMMNMEETSKRRPFFSPEDLYDDDFYD**  
**DQLPEKKRRLTTEQVHLLLEKSFETENKLEPERKTQLAKKLGLQPRQVAVWFQNRARRWKTKQ**  
**LERDYDLLKSTYDQLLSNYDSIVMDNDKLRSEVTSLTEKLQKQETANEPGQVPEPNQLDP**  
**VYINAAAIKTEDRLSSGSGVGSVAVLDDAPQLLDSCDSYFPSIVPIQDNSNASDHDNDRSCFA**  
**DVFVPTTSPSHDHHGESLAFWGW**

**SEC ID N°: 17 Secuencia de ADNc D26575 de Dauca\_CHB3 de *Daucus carota***

**ATGGCGGGTCGGAGGGTGTTCATGGGGAGGGAGCCAATACGACGTCGGCTAGCCTGTTGTT**  
**TCATAGTCAAAGACCTGAGCCTTCTTTCTTTCTGACCTTCTCCTTCTCTAATTGGTTCAA**  
**AATCCATGGTTAGCTTTCAAGATGCTAAGCGAAAAAATCCCTACGATGGGTCTTTATGCGG**  
**TCATATGATGAAGAAGAAATTGGGGATGAAGAATATGATGAATACTTTCAGCAGCCTGAGAA**  
**GAAGAGGAGGCTCAAGGCTGATCAAATCCAGTTTCTTGAGAAAAGTTTTGAGACTGATAACA**  
**AGCTTGAGCCTGAAAGAAAAAGTTTCAGCTTGCAAAAGAACTCGGCTTGACGCCAAGACAGGTT**  
**GCGATATGGTTTTCAGAACCGTCGAGCACGGTGGAAGACCAAAACACTAGAAAAAGATTATGA**  
**TGTATTGCAAAATAGCTACAACAGCCTCAAGGCTGACTATGACAATCTACTTGCCGAGAAAAG**  
**AAAAACTTAAAGCCGAGGTTCTCGACCTGACAGACAAGCTACTTCTCAAAGAAGATAAGGGG**  
**AGCAAGACAGTAGTTTTTGTATAAGCAAAGGTGTCTGCAGCATTTCAAACAAGAACGTGTAG**  
**TAATGACATATCTGTGGGTGAAGTACTCAGTAAGTCAAGTATGGACTGCAAGCAAGAAGATC**  
**ATAACTCTGTGAAAAGTGATGCAGTTGATTCTGACAGTCCACACTACAGTGATGAAGTCTAC**  
**TCCAGTTTTATGGAGCCAGTGGATCGCTCTTATGTTTTTGAACCTGCTCAGTCGGATATATC**  
**TCAAGATGAAGAAGATGACATGGGGAACAACCTTATTTCTCCCATCATATCATGTTTTCTCAA**  
**AGACTGAAGACGGTAGTTACTCCGACCAGCCTTCGAACTCTTCGTACTTTGGCTTCCAGTT**  
**GAAGATCATACGTTTGGCTTTTGGGGTACTGAATTATAA**

**FIGURA 4 (continuación)**

**SEC ID Nº: 18 Secuencia de aminoácidos traducida de Dauca\_CHB3 de *Daucus carota***

MAGRRVYFYGEGANTTSASLLFHSQRPEPFFLSAPSPSLIGSKSMVSFQDAKRKNPYDGGFFMR  
SYDEEEIGDEEYDEYFQQPEKKRRLKADQIQFLEKSFETDNKLEPERKVQLAKELGLQPRQV  
AIWFQNRARRWKTLEKDYDVLQNSYNSLKADYDNLAEKEKLKAEVLDLTDKLLKEDKG  
SKTVVFDKQKVSAAFQQRVSNDISVGEVLSNSVMDCKQEDHNSVKSDAVDSDSPHYSDEVY  
SSFMEPVDRSYVFEPAQSDISQDEEDDMGNNLFLPSYHVFSKTEDGSYSDQPSNSSYFGFPV  
EDHTFGFWGTEL

**SEC ID Nº: 19 Secuencia de ADNc AF184278 de Glyma\_HDI57 de *Glycine max***

ATGGCGAGTGGCAAGCTTTATGCGGGTTCAAACATGTCACTTCTCCTCCAAAACGAAAGGCT  
CCCTTGCTCCTCTGAAGTCCTTGAGTCTCTTTGGGCTCAGACCTCTAACCTTGCTTCTTCC  
AAGGTTCAAACCCGTGGTTGATTTTGAGAATGTAAGTGGGAGCAGGATGACGGATAGGCCT  
TTCTTTCAAGCGTTGGAGAAGGAAGAGAACTGTGATGAGGATTACGAGGGGTGTTTCCACCA  
ACCGGGGAAGAAAAGGAGGCTCACAAGCGAACAAGTTCAGTTCCTTGAAAGGAACTTTGAGG  
TAGAGAACAAGCTTGAACCCGAAAGGAAAGTCCAACCTTGCAAAAGAGCTTGGCTTGCAGCCA  
AGGCAAGTTGCTATATGGTTCCAAAACCGAAGGGCAAGGTTCAAGACCAAGCAGCTAGAAAA  
AGACTATGGCGTGTGAAAGCTAGTTATGACAGACTCAAAAGTGACTATGAAAGTCTTGTTT  
AAGAGAATGACAAGTTAAAAGCAGAGGTGAATTCTCTGGAGAGCAAATTGATTCTTAGAGAT  
AAAGAGAAGGAGGAGAATTCGGATGACAAGTCATCTCCTGATGATGCTGTCAATTCTTCTTC  
ACCCCAACAACAAGGAGCCTATGGATTTATTAATTATTTCAAAAAATGCAACAACAACA  
CAACATCTGAAAATGGGACCAAAGTGTGTCAACCACTCCCACTCCCTATTATGGTAACATGC  
TGCAAGCAAGAAGATGCCAACTCAGCCAAAAGTGATGTCTTGATTTCGGATAGCCACATTG  
CACTTCATTTCGTGGAGCCTGCTGATTCTCTCATGCCTTTGAACCAGAAGACCACTCAGAAG  
ACTTCTCCCAAGATGAAGAGGATAACCTTAGTGAAACCTTTTGATGACCTTCCCTTCTTCT  
TGTTGCTTACCTAAGGTTGAAGAACAAGTCTATGACGGCCCTCCTGAAAACCTTTGTAATTT  
TGGCTTCCAGGTTGAGGATCAAACCTTCTGTTTCTGGCCCTATTGA

**SEC ID Nº: 20 Secuencia de aminoácidos traducida de Glyma\_HD157 de *Glycine max***

MASGKLYAGSNMSLLQNERLPCSSEVLESLSWAQTSNPASFQGSKPVVDFENVSGSRMTDRP  
FFQALEKEENCDEDEYEGCFHQPGKKRRLTSEQVQFLERNFEVENKLEPERKVQLAKELGLQ  
RQVAIWQNRARRFKTKQLEKDYGVLLKASYDRLKSDYESLVQENDKLKAEVNSLESKLI LRD  
KEKEENSDDKSSPDDAVNSSSPHNNKEPMDLLIISKNAATTTTSENGTKVLSPLPLPIMVTC  
CKQEDANSKSDVLDSDSPHCTSFVEPADSSSHAFEPEDHSEDFSQDEEDNLSENLLMTFPSS  
CCLPKVEEHCYDGPPENSCNFGFQVEDQTFCFWPY

**SEC ID Nº: 21 AF443621 de CPHB-5 de *Craterostigma plantagineum***

ATGAACTCTGCTCGGATTTTCTTCGACCCATCTTCCCACGGCAACATGCTGCAGTTTCTTGG  
GAACGCCGGCGGCGGATTATCCGTTTTCCGAGGAACAAGATCGTCGTCGGTGCTGAACATGG  
AGGAGAGCTCGTTAAAACGACAGATTTTCAGCGGCGGCGGCGGCGATGAATTCTACGACGAG  
GAATACTACGACGAGCAGTTGTTGCCTGAGAAGAAGCGCCGACTCACCGCCGAGCAGGTTCA  
CTTGCTTGAGAAGAGCTTCGAGGCTGAGAACAAGCTTGAGCCTGAGCGAAAGGCTGAGCTGG  
CGAAGAAGCTCGGATTGCAGCCGAGGCAAGTCGCCATTTGGTTCCAAAACCGCCGAGCACGG  
TGGAAGACTAAGCAGTTAGAGAGGGACTACGACAAGCTTAAGTCTTCTATGATTCTCTTCT  
CTCAACCTACGACTCTATTTCGCCAGGAAAACGACAAGCTCAAAGCCGAGCTCCTTTCCCTGA  
ACGAGAAATTGCAACCCAAAGACGACGACGACCCATCGGCCGAAATAGGTCGAAATCTCAGT

FIGURA 4 (continuación)

TCATCGTCGCGCCCTGTGACGCGGCTGAGCCGCGTGCCTGAAGCTGACGGTGAAGGTGGA  
GGACCGCCTGAGCACGGGGAGCAACGGCAGCGCAGTAATGGACGGCGACGGACCTCAGCAGC  
TCCTCGACGACAGCGGCGACTCGTACTTCGAGAACGACGAGGAATACGACTGCGCCGCGCA  
AGTTTGGCTGCTGCGAAGGAGGACGACGGCAGCGATGAGGGCGGGTGTACTTACCGAGGC  
TCTCGCGGCGAGGAGGAGGAGCGCCGTTTGTGTTGGTGTATTTGGTCTTAA

SEC ID N°: 22 Secuencia de aminoácidos traducida de CPHB-5 de *Craterostigma plantagineum*  
MNSARIFFDPSSHGNMLQFLGNAGGDSVFRGTRSSSVLNMESSLKRQIFSGGGGDEFYDE  
EYYDEQLLPEKKRRLTAEQVHLLKSF EAENKLEPERKAELAKKLGLQPRQVAIWQNRARR  
WKTKQLERDYDKLKSSYDSLSTYDSIRQENDKLKAEILLSLNEKLQPKDDDDPSAEIGRNLS  
SSSPVDAAEPPCLKLTVKVEDRLSTGSNGSAVMDGDPQQLLDDSGDSYFENDEEYDCAAA  
SLAAAKEDDGSDEGGCYFTEALAAEEEEAPFAWCIS

SEC ID N°: 23 Secuencia de ADNc DT465649, CD486134 de Goshi\_hox5 de *Gossypium hirsutum*  
ATGGAGTCTGGCCGTCTTTTTTCAATCCCTCCACTACCCACCGCAACATGTTGCTTCTCGG  
GAACACTGAACCCATCTTTCGAGGGGCAAGAACAATGGTTAGCATGGAGGAAAACCCAAAGA  
AGCGACTGTTCTTCAGCTCGCCGGAGGATTTGTACGACGAAGAGTACTACGACGAGCAGTTG  
CCCGAGAAAAAGCGTCGCCCTTACGTGCGGAGCAGGTGTATCTGCTAGAGAAGAGCTTTGAGGC  
AGAGAACAAGCTGGAGCCGGAGAGGAAGAGCCAGTTGGCCAAGAAGTTAGGACTGCAACCAA  
GGCAGGTGGCGGTATGGTTCCAGAACCGCCGTGCAAGGTGGAAGACAAAGCAGCTTGAAAGG  
GACTATGACCTCCTCAAATCTTCTTTGATTCCCTTCAGTCCAATTATGACACTATTCTCAA  
AGAAAATGAGAAGCTCAAATCTGAGGTAGCTTCTTCTGACTGAAAACTACAAGCCAAAGATG  
TGGCAACAGAAGCAATAGCAGGTGAAAAGGATGAAGGGTTAGCAGCTGAGATGGCCTCCGCC  
CTCCAATTTCAGTATGAAGGTGGAGGACCGTCTTAGTAGCGGCAGTGTTCGGAAGCGCGGTGGT  
GGATGAGGATGCCCCACAGCTGGTGGACAGCGCAATTCCTACTTTCCAAGCGATGAATACT  
CCAGAGGCATTGGCCCTTTTCGATGGGGTTTCAGTCGGAAGATGAGGATGGCAGTGATAATTGC  
GGGAGTTACTTCTCCGATGTGTTTCGCAACCACAGAGCAGGGAGCATTAGGATTGTGGGCCTG  
GNTCTAA

SEC ID N°: 24 Secuencia de aminoácidos traducida de Goshi\_hox5 de *Gossypium hirsutum*  
MESGRLFFNPSTTHRNMLLLGNTEPIFRGARTMVSMEENPKKRLFFSSPEDLYDEEYDEQL  
PEKKRRLTSEQVYLLEKSF EAENKLEPERKSQ LAKKLGLQPRQVAVWFQNRARRWKTKQLER  
DYDLLKSSFDSLQSNYDTILKENEKLKSEVASLTEKLQAKDVATEA IAGEKDEGLAAEMASA  
LQFSMKVEDRLSSGSVGSADVDEAPQLVDSGNSYFSPSEYSRGIGPFDGVQSEDEDGSDNC  
GSYFSDVFATTEQGALGLWAWX

SEC ID N°: 25 Secuencia de ADNc BT014213.1 de Lyces\_hox5 de *Lycopersicon esculentum*  
ATGGGATCTGGGCATATATTTTTTCGACCCGTCGTGTCACGGCAACATGCTGTTCTTGG  
GAGCGGAGATCCTGTTTTCCGAGGACCAAGATCGACGATGATGAAGATGGAGGACTCCTCGA  
AGAGGCGACCCTTCTTTAGCTCGCCGGAGGATCTATATGACGAGGAATACTACGACGAGCAG  
TCACCGGAGAAGAAGCGCCGTCTCACTCCTGAGCAGGTGCACTTGTTGGAGAAGAGCTTTGA  
GACAGAAAACAAGCTGGAGCCCGAGCGCAAAACCCAGCTGGCCTANAAGCTGGGGCTGCAGC

FIGURA 4 (continuación)

CCAGACAGGTGGCTGTATGGTTCCAAAACCGCCGTGCCCGGTGGAAGACCAAGCAGCTCGAG  
AGGGATTATGATCAGCTCAAATCCTCTTATGACTCCCTTCTCTCTGATTTTGACTCCGTTTCG  
CAAAGATAACGATAAGCTCAAATCTGAGGTTGTTTCATTGATGGAAAAGTTACAGGGGAAAG  
TGGTTGGAGGAGCAGGGGAAATGAAAAATCTGACATCTTGGAGGTGGATGCTATGACGATC  
CTTCAAGTGAAGGTGAAGGCTGGGGACCGGTTGAGCAGTGGCAGTGGTGGGAGCGCGGTGGT  
AGATGAGCATAGTTTACAGCTGGTGGACAGTGGGGACTCATATTTTACACTGATCATGAGG  
AGTATCCAGGGCCTGGAGGATGCAATGTTTCTCCACCCATGGATGGTTTACAATCGGAGGAA  
GATGATGGTAGTGATGATCATGGCAGTTGCCATGGCTACTTCTCTAACGTCTTTGTGGCAGA  
AGAGCAGCACCATGAACAAGGAGAAGAGCCTATTGGATGGTTCTGGTCTTAA

**SEC ID N°: 26** Secuencia de aminoácidos traducida de *Lyces\_hox5* de *Lycopersicon esculentum*  
MGSGHIFDPSSCHGNMFLGSGDPVFRGPRSTMMKMEDSSKRRPFFSSPEDLYDEEYYDEQ  
SPEKKRRLTPEQVHLLKESFETENKLEPERKTQLAXKLGLQPRQVAVWFQNRARRWTKQLE  
RDYDQLKSSYDLSLSDFDVSRKDNKDKSEVVSLEKLGKVVGGAGGNEKSDILEVDAMTI  
LQVKVKAGDRLSSGSGGSAVVDEHSSQLVDSGDSYFHTDHEEYPGPGGCNVPPMPMDGLQSEE  
DDGSDDHGSCGHYFSNVFVAEEQHHEQGEPIGWFS

**SEC ID N°: 27** Secuencia de ADNc X94947 de *VaHOX1* de *Lycopersicon esculentum*  
ATGGCTCCAGGGATTCTCTATGGTGGTTCTTCTAATTTTCGATGGCGTTTTTACTCAAAAACA  
GAGAGACGTGTTTTCTTCATCTACTGCACCGAAAGGGCATCTTGGTTCCCTTTTTGCCCCTG  
CCTCTTCTTCTTAATTTCTTGGGATCCAGTTCTATGGTGAGTTTTCGCGGTGTTAATGGA  
GGGAAGAGATCATTCTTTGATTTCGTTTCGATCAGGATGACAATGAAGCTGATGAATTGGGGGA  
ATATCTTCATCAAGCGGAGAAGAAGAGGCGACTTACTGACAACCAAGTTCAGTTTCTTGAGA  
AGAGTTTTGGGGAAGAGAACAACTTGAACCAGAAAGAAAAGTTCAGCTTGCTAAAGAACTT  
GGTCTGCAGCCTCGCCAAATTGCAATTTGGTTTCAGAATCGTCGTGCGCGATGGAAGACTAA  
GCAGCTCGAGAAAGATTATGATGAATTGAGGAATAGATACGATACTCTGAAATCAAATTACA  
ATAATCTTCTCAAGGAAAAAGAAGATCTTTCGAACCTGAAGTTTTCCGTCTCACCGGTAAGCTG  
TTTATCAAAGAGAAAGGAAATGGGCAATTGGATTGCGCGATGAACACAAACACTCCAATGC  
ATTGGCAAAGAAACCGTGGTTGATCCAATGTCCAATGTACCAGCTCTGGTTGTTAAGCACC  
AGCAGGAAGATTTAAGCTCTGCTAAGAGTGATGTTTTCGACTCAGAAAGCCACGTTACACC  
AGTAGAATGCATTCTCAGTCGTAGATCAGGATGATTCTGCTCGCGCATTGAAACTGATCA  
GTCGGATTCTCAGGATGATGATGAAAACCTTCAGCAAGAATATGCTTTCTACTGCCAACC  
TACTTGGCAAAGACGCGGATGATGATTATCCGCGACATCATCAAATTTGAGTTACTTTGGA  
TTTCCAGTTGAAGACCAAGGTTTTGGTTTCTGGACTTATTAA

**SEC ID N°: 28** Secuencia de aminoácidos traducida de *VaHOX1* de *Lycopersicon esculentum*  
MAPGILYGGSSNFDGVFTQKQRDVFSSTAPKHLGSLFAPASSSSNFLGSSSMVSFRGVNG  
GKRSFFDSFDQDDNEADELGEYLHQAEEKRRLTDNQVFLEKSFGEENKLEPERKVQLAKEL  
GLQPRQIAIWFQNRARRWTKQLEKDYDELNRNYDTLKSNNYNNLLKEKEDLRTEVFRLTGKL  
FIKEKNGQLDLRDEHKHSNALAKETVVDPMSPNPALVVKHQEDLSSAKSDVFDSESPRYT  
SRMHSSVVDQDD SARAFETDQSDSSQDDDENFSKNMLSTANLLGKDADDDYPATSSNLSYFG  
FPVEDQGFGFTY

FIGURA 4 (continuación)

**SEC ID Nº: 29** Secuencia de ADNc CB892061, CA858059 de Medsa\_hox16 de *Medicago sativa*  
**ATGGCGGGTGGGAGAGTTTTTTCAAATGGTCCTGCAAATATTTCAAATATAAATATGAATAT**  
**TTTGCTTCAGAATCAACAACAACTCCTCGTGGAACCTCTTCTCAACAACCTCTTGATTCTC**  
**TTTTCTTTCTTCTTCTGCTTCTTTCTTTGGTTCAAGATCTATGGTGAGTTTTGAAGATGTT**  
**CAAGGAAGGAAAAGGCGCAACAGGTCTTTCTTTGGAGGATTTGATCTTGACGAAAACGGAGA**  
**GGATGAGATGGATGAGTACTTTTCATCAATCCGAGAAGAAACGGCGTCTCTCAGTGGAATCAAG**  
**TTTCACTTTCTTGAGAAAAGCTTTGAGGAGGACAACAACTTGAACCAGAGAGGAAAACCAAG**  
**CTAGCTAAAGACCTTGGTTTGCAGCCACGGCAAGTTGCTATTTGGTTTCAAACCCGTCGTGC**  
**AAGGTGGAAGACTAAACAGCTTGAGAAGGATTATGATTCTCTTAATGATGGTTATGAGTCTC**  
**TTAAGACAGAGTATGACAACCTTCTCAAAGAGAAAGATAGGTTACAATCTGAGGTGGCAAGC**  
**CTAACTGAAAAGGTACTTGAAAGAGAGAAAACAGAGGGAAAATTCAAACAAGGTGAAAGTGA**  
**AACAAAGGAATTCTTGAAAGGAACCAACAATTAATAAGCCTTTGGTTGATTGAGTTTCTGAGG**  
**GTGAAGGATCCAAATTGTCAATTGTTGAGGCTTCTAATAATAATAATAATAACAACTT**  
**GAAGATATTAGTTTCAAGGAGTGACATATTGGATTGTGAAAGTCCACGCTACACTGATGG**  
**AGTGTAGAGACATGTGATTCTTCCTATGTATTTGAACCTGAATATCAATCGGACCTATCAC**  
**AAGATGAAGAAGATCACAAATTTATTGCCTCCTTACATCTTTACAAAACCTGAAGATGTGAAT**  
**TACTCCGACCCGCCACATAATTCAACAAGTTATGGATTTCAGAGGAAGATCATCATCAAGC**  
**TCTTTGGCCTTGGTCTTATTAG**

**SEC ID Nº: 30** Secuencia de aminoácidos traducida de Medsa\_hox16 de *Medicago sativa*  
**MAGGRVFSNGPANISINIMNILLQNNQQTTPRGNSSQQPLDSLFLSSSASFSGSRSMVSFEDV**  
**QGRKRNRNSFFGGFDLDENGEDEMDEYFHQSEKKRRLSVDQVQFLEKSFEEDNKLEPERKTK**  
**LAKDLGLQPRQVAIWQNRRRARWKTQLEKDYDSLNDGYESLKTEYDNLLKEKDRLQSEVAS**  
**LTEKVLEREKQEGKFKQGESETKEFLKEPTINKPLVDSVSEGEKSKLSIVEASNNNNNNNNKL**  
**EDISSARSDILDCESPRYTDGVLETCDSSYVFEPEYQSDLSQDEEDHNLPPYIFTKLEDVN**  
**YSDPPHNSTSYGFQEEEDHHQALWPWSY**

**SEC ID Nº: 31** Secuencia de ADNc DT758247 de Aqufo\_hox5 de *Aquilegia formosa x Aquilegia pubescens*  
**ATGGATTCAACAACAAGCCGCTTTTTCTTTTGATGGTTCTTGCCATGGGAACATGTTGCTTTT**  
**AGGGAGTGGAGATCCCGTTCTTCGAGGTTCAAGATCATTCATTAATATGGAAGATTCTTTGA**  
**AAAGACGTCCTTTTTATAGTTCAACAGATGAACTAATTGAAGAGGAGTTTATGATGAACAG**  
**CTACCTGAAAAGAAACGTCGTCTTACTTCTGAGCAGGTTCTCTATTGGAGAAGAGCTTTGA**  
**GACAGAGAACAAGCTGGAACCAGATCGTAAGACCCAGCTTGCTAAGAAGCTTGGGTTGCAAC**  
**CGAGACAAGTTGCAGTTTGGTTTCAGAATAGACGAGCTCGTTGGAAGACTAAGCAACTAGAG**  
**AGAGATTATGATCTTCTTAAAGCTTCTTATGATTCCCTTCGTTCTGATTACGATGACATTGT**  
**TAAAGAGAATGAGAAGCTCAAATCTGAGGTGGTTTCCTTAACTGGGAAGTTGCAGGTCAAGG**  
**AGGGAGCTGGGATGGAGTTAAATCAGATATCTGACCCACCACTCTCCACTGAAGAAAATGTT**  
**GATGTAACCTACGATGCAATTTAATGTAAAGGTTGAGGATCGCTTGAGCTCTGGCAGTGGGGT**  
**AAGTGCTGTGGTTGATGAGGAATGTCGACAGCTTGTTGACAGTGTTGATTCCTATTTCCCTG**  
**GCGATGACTATGGTCAATGCATAGGCCAGTAGATGGAGTCCAGTCAGAAGAAGATGACATT**  
**AGTGACGACAGCCGAGCTATTTCTCAGATGTCTTTCCAGCTGCACCAGAGCAGAACCACCA**  
**GGAGAGTGAGACATTGGGTTGGTGGGACTGGGCTTAA**

FIGURA 4 (continuación)

**SEC ID N°: 32 Secuencia de aminoácidos traducida de Aqufo\_hox5 de *Aquilegia formosa* x *Aquilegia pubescens***

MDSTTSRLFFDGSCHGNMLLLGSGDPVLRGSRSFNIMEDSLKRRPFYSSTDELIBEEFYDEQ  
 LPEKKRRLTSEQVHLLKESFETENKLEPDRKTQLAKKLGLQPRQVAVWFQNRARRWKTKQLE  
 RDYDLLKASYDSLRSYDDIVKENEKLKSEVVSILTGLQVKEGAGMELNQISDPPLSTEENV  
 DVTTMQFNVKVEDRLSSGSGVSAVVDEECRLVDSVDSYFPGDDYGQCIGPVDGVQSEDDI  
 SDDRSYFSDVFPAAPEQNHQESETLGWWDWA

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

AATCCGAAAAGTTTCTGCACCGTTTTCCACCCCTAACTAACAATATAGGGAACGTGTGCTAA  
 ATATAAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAACTCATC  
 CACCTACTTTTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTTCGTTTTCT  
 TAGTAATTAAGTGGGAAAATGAAATCATTATTGCTTAGAATATACGTTACATCTCTGTCTAT  
 GAAGTTAAATTATTTCGAGGTAGCCATAATTGTCTCAAACTCTTCTTGAATAAAAAAATCTT  
 TCTAGCTGAACCTCAATGGGTAAAGAGAGAGATTTTTTTTAAAAAATAGAATGAAGATATTC  
 TGAACGTATTGGCAAAGATTTAAACATATAATTATATAATTTTATAGTTTGTGCATTCGTCA  
 TATCGCACATCATTAAGGACATGTCTTACTCCATCCCAATTTTTATTTAGTAATTAAAGACA  
 ATTGACTTATTTTTATTATTTATCTTTTTTCGATTAGATGCAAGGTACTTACGCACACACTT  
 TGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAATACACGTTCAACTAGCAACACATCTCT  
 AATATCACTCGCCTATTTAATACATTTAGGTAGCAATATCTGAATTCAGCACTCCACCATC  
 ACCAGACCCTTTTTAATAATATCTAAAAATACAAAAATAATTTTACAGAATAGCATGAAAAG  
 TATGAAACGAACTATTTAGGTTTTTACATACAAAAAAGAAATTTTGCTCGTGCGCGA  
 GCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCACAGAACAAACCA  
 CAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAGGCTTTG  
 CGGCCAGGAGAGAGGAGGAGGCAAGAAAACCAAGCATCCTCCTCCTCCCATCTATAAAT  
 TCCTCCCCCTTTTCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCA  
 AGGACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTTCGATCCATATCTTCCGGTCGAGT  
 TCTTGGTCGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCCTTCGGTTGTTC  
 TTGGATTTATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATC  
 TGTGATGATTCTGTTCTTGGATTGAGGAGGTTCTTGATGTTGCATGTTATCGGTT  
 CGGTTTGATTAGTAGTATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTA  
 GGGTACGGAATCTTGCGATTTTGTGAGTACCTTTTGTGTTGAGGTAAATCAGAGCACCGGTG  
 ATTTTGCTTGGTGTAATAAAAGTACGGTTGTTTGGTCCTCGATTCTGGTAGTGATGCTTCTC  
 GATTTGACGAAGCTATCCTTTGTTTATTCCTATTGAACAAAAATAATCCAACCTTGAAGAC  
 GGTCCCGTTGATGAGATTGAATGATTGATTCTTAAGCCTGTCCAAAATTCGCAGCTGGCTT  
 GTTTAGATACAGTAGTCCCCATCACGAAATTCATGGAAACAGTTATAATCCTCAGGAACAGG  
 GGATTCCTCGTTCTTCCGATTTGCTTTAGTCCCAGAAATTTTTTTTCCCAAATATCTTAAAA  
 GTCACCTTCTGGTTTCAGTTCAATGAATTGATTGCTACAAATAATGCTTTTATAGCGTTATCC  
 TAGCTGTAGTTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGAAGAACTTATCCGA  
 TTTCTGATCTCCATTTTAAATTATATGAAATGAACTGTAGCATAAGCAGTATTCATTGGAT  
 TATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAACTGTC  
 CTCAATTTTGTGTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGATCTACCTGTA  
 GAAGTTTCTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAA  
 TCGGGATAGTTATACTGCTTGTCTTATGATTCAATTCCTTTGTGTCAGTTCTTGGTGTAGCT  
 TGCCACTTTCACCAGCAAAGTTC

**FIGURA 4 (continuación)**



**SEC ID N°: 39 Secuencia de aminoácidos traducida de *Orysa\_hox6* de *Oryza sativa***

MDGEEDSEWMMMDVGGKGGKGGGGGAADRKKRFSEEQIKSLESMFATQTKLEPRQKLQLAR  
ELGLQPRQVAIWQNKRRARWKSQLEREYSALRDDYDALLCSYESLKKEKLALIKQLEKLAE  
MLQEPRGKYGDNAGDDARSGGVAGMKKEEFVGAGGAATLYSSAEGGGTTTTTTAKLMPHFGS  
DDVDAGLFLRPSSQHPPPPHAGAGFTSSEPAADHQSFNPHSSWPSSTEQTCSSTPWWEFES  
E

**SEC ID N°: 40 Secuencia de ADNc de *Poptr\_HOX16\_1* de *Populus tremuloides***

ATGGCGGGTGGTACCGGTGGTTCTAATTCCAATTTGTCTGTTTTGCTTCAAAGCCAAAGAGG  
CCCTTGCTGCTTCAACCTCTTGAATCTTTTTCTTTCTGGCTCTTCTCCTTCTTTTC  
TTGGTTCAAGATCCATGATGAGTTTTGAAGATGTTTCATCAAGCAAACGGATCAACCAGGCCT  
TTTTTCCGCTCGTTTGATCACGAAGACAATGGAGACGATGATCTGGATGAATATTTTCATCA  
ACCTGAAAAGAAGAGAGACTTACTGTTGATCAAGTTTCTTGAAGAAGAGTTTTGAGC  
TTGAGAACAAGCTTGAACCTGAAAGGAAAATCCAGCTTGCAAAGGATCTTGGCCTTCAGCCG  
CGTCAGGTTGCTATATGGTTTCAAAACCGCCGAGCAAGATGGAAGACTAAACAGCTGGAAAA  
GGATTATGACGTTTTGCAATCTAGCTACAATAGCCTTAAGGCTGACTATGACAACCTCCTCA  
AGGAGAAGGAGAACTAAAAGCTGAGGTTAATCTTCTCACCGACAAGTTGCTCCTCAAAGAG  
AAAGAGAAGGGAATCTCAGAATTGTCTGATAAAGATGCATTATCGCAAGAGCCACCTAAAAG  
GGCTATAGCTGATTAGCTTCCGAGGGTGAAGTGTGCGAAAATCTCAACAGTGGCCTGTAAGC  
AGGAAGATATCAGCTCAGCCAAAAGCGACATATTTGATTGACAGCCACATTACGCTGAT  
GGGGTGCATTCTCTACTCTTAGAGGCAGGAGATTCTTCATATGTTTTCGAACCCGATCAATC  
AGATTTGTCACAAGATGAAGAAGATAACTTTAGCAAGAGCTTATTGCCTCCATACGTCTTTC  
CGAAGCTTGAAGATGACGATTACTCTGACCCGCCTGCAAGTTTTGAAGATCATGCCTTTTGG  
TCCTGGTCATACTAA

**SEC ID N°: 41 Secuencia de aminoácidos traducida de *Poptr\_HOX16\_1* de *Populus tremuloides***

MAGGTGGSNSNLVLLQSQRGPCAASQPLESFFLSGSSPSFLGSRSMMSFEDVHQANGSTRP  
FFRSFDHEDNGDDDLDEYFHQPEKKRRLTVQVQFLEKSFLENKLEPERKIQLAIDLGLQP  
RQVAIWQNRARRWKTQLEKDYDVLQSSYNLSLKADYDNLKEKEKLKAEVNLLTDKLLLKE  
KEKGI SELSDKDALSQEPKRAIADSASEGEVSKI STVACKQEDISSAKSDIFDSDSPHYAD  
GVHSSLLEAGDSSYVFEPDQSDLSQDEEDNFSKSLLPYVFPKLEDDDDYSDPPASFEDHAFW  
WSY

**SEC ID N°: 42 Secuencia de ADNc de *Poptr\_HOX16\_2* de *Populus tremuloides***

ATGGCGGCTTGTGGTGGTGGTGGTGGTCTAATCCCAATTTGTCTGTTTTAGTTCAAAG  
CCAAAGAGGCCCTTGTGCTGCTTCTCAACCTCTTGAAGCTTTTTCTTTCTGGCTCTTCTC  
CTTCTTTTCTTGGTTCAAGATCCATGATGAGTTTTGCAGATGTTACCAAGCAAATGGATCA  
ACTAGACCGTTTTTCCGCCCATATGATCACGAAGACAACGGCGACGATGATTTGGATGAATA  
TTTTTCATCAACCTGAAAAGAAGAGGAGACTTACTGTTGATCAAGTTTCTTGAAGAAGAA  
GTTTTGAGGTTGAGAACAGCTTGAACCCGAAAGGAAAATCCAGCTGGCGAAGGATCTTGGC  
TTGCAGCCTCGGCAGGTTGCCATATGGTTTCAAAACCGCCGGGCAAGATGGAAGACGAAACA  
GCTTGAAAAAGATTATGAGGTTCTGCAATCTAGCTACAATGGCCTTAAGGCTGACTACGACA  
ACCTCTTCAAGGAGAAGGAGAACTAAAAGCTGAGGTTAATCTTCTACCAACGAGTTGCTC  
CTTAAAGAGAAAAGAGAAAGGAAGCTCAGAATTGTCTGATAAAGATGCATTATCTCAAGAGCC  
ACCAAAAAGGCAATAGCCGATTAGCTTCAAGAGGTTGAAGTGTGCGAAACTTCAACCGTGG

FIGURA 4 (continuación)

CCTGCCAGCAGGAAGATATTAGCTCAGCCAAAAGTGATATGTTTGATTTCAGACAGCCCACAT  
TTTGCGGATGGGGTACATTCTCTACTCTTAGAGGCAGGTGATTCTTCACATGTCTTCGAGCC  
CGACCAATCGGATTTATCACAAGATGAAGAAGATAACTTGAGCAAGAGTCTTTTGCCCTCCGT  
ACGTCTTTCCAAAGCTTGAAGATGGTGATTACTCTGACCCGCCAGCAAGTTTTGAAGATCAT  
GCCTTTTGGTGCTGGTCATACTAA

**SEC ID N°: 43** Secuencia de aminoácidos traducida de Poptr\_HOX16\_2 de *Populus tremuloides*  
MAACGGGGGGSNPNSVLVQSQRGPAAASQPLEAFFLSGSSPSFLGSRSMMSFADVHQANGS  
TRPFFRPYDHEDNGDDDLDEYFHQPEKKRRLTVQVQFLERSFEVENKLEPERKIQLAQDLG  
LQPRQVAIWQNRARRWKTQLEKDYEVLLQSSYNGLKADYDNLFKEKEKLKAEVNLLTNELL  
LKEKEKGSSELSKDALSQEPKKAIADSASEGEVSKTSTVACQQEDISSAKSDMFDSDSPH  
FADGVHSSILLEAGDSSHVFEPDQSDLSQDEEDNLSKSLPPYVFPKLEDGDYSDPPASFEDH  
AFWCWSY

**SEC ID N°: 44** Secuencia de ADNc Poptr\_HOX16\_3 de *Populus tremuloides*  
ATGGCGGGTGATAAAGACTGTGGCAGTTCTAAATGACCATTTTTCTTCGAAACGGCAGGCT  
CCCTCCTTGTGAATCTCTGTATTCTCACCTCTTTTAGCACTCTTCATGGTGCAAAATCTA  
TGGTTAATTTTAGGAATGATGGAGGAGACACTGTAGACATGTCTTTTTTCCAACCACATGTC  
AAAGAAGAAAGTAGCGATGAGGATTATGATGCGCACCTTAAGCCATCTGAAAAGAAAAGGCG  
GCTTACAGCTGCTCAAGTCCAGTTTCTTGAGAAGAGCTTTGAGGCGGAGAATAAGCTTGAAC  
CAGAGAGGAAGATGCAGCTTGCTAAAGAACTCGGCTTGACAGCTCGCCAGGTTGCAATATGG  
TTTCAAAACCGTAGAGCTCGGTTCAAGAACAAGCAGCTGGAAAGGGACTACGACTCCTTGAG  
AATCAGCTTTGACAACTCAAGGCTGATTATGACAACTCCTCCTCGAGAAGCAGAATTTGA  
AAAACGAGCTTCTTTCACTGAAAGAAAATGCTTAGCAGAGAGGAAAGTATGGAAAGTTCA  
GAACCATTGATGTCATCCATTACCGGATGCAGAACTTGAGCCTATTCTTGATACAGTGTC  
TGAAAATGTTTCCGCCATTGTGCCAATGGTGACACCCAAACAAGAAGAAAGTTTCAGCTAAAA  
ATGATGTTTTCAACTCAGACAGCCACGTTCATTTTTGGAGCCCCGTGATTGTTATCGTGTT  
TTCGAGTCAGACCAACCAGATTTTTTCCCAAGTTGAAGAAGATAATCTCACCAGGAGCTTTCT  
ACCCCTCCGTACTTTCCAAAACCTTACCGAGAGCCACCTGCAAGTTCACGTAATTTTGAAT  
TCTCAGCGGAAGATCAGCCCTTTTGGTCTGGATTACTGA

**SEC ID N°: 45** Secuencia de aminoácidos traducida de Poptr\_HOX16\_3 de *Populus tremuloides*  
MAGDKDCGSSKMTIFLRNGRLPPCESLCILTSFSTLHGAKSMVNFRNDGGDTVDMSFFQPHV  
KEESSDEDYDAHLKPSEKKRRLTAAQVQFLEKSFEAKNKLEPERKMQLAQELGLQPRQVAIW  
FQNRARRFKNQLERDYDSLRLISFDKLLKADYDKLLLEKQNLKNELLSLKEKLLSREESMESS  
EPFDVIHSPDAELEPIPDTVSENVSAIVPMVTPKQEESSAKNDVFNSDSPRSFLEPRDCYRV  
FESDQPDFSQVEEDNLTRSFLLPPYFPKLYREPPASSRNFEFSAEDQPFWSWIY

**SEC ID N° 46** Secuencia de ADNc de Medr\_HOX16\_1 de *Medicago truncatula*  
ATGGCAGGTGGCAAGCTTTTTGGTGGTTCTAATATGTCACTTTTGCTTCAAATGAAAGACT  
CCCTTGTAATCTCTGAAGTCCCTGAATCTCTTTGGGTTACACCCCTGCTTCTTTTCAAGGTT  
CAAAATTCAGTGGTTAATTTTGAGAATGGTGGTGGTAGCAACAGAGTGGTAACAGATAGACCC  
TTCTTTCAACAACTTGAGAAAGAAGAGAATTGTGGTGATGAAGATTATGAAGCATGCTACCA  
TCAACAAGGAAAGAAAAGGAGGCTTTCAAGTGAACAAGTTCAATTTCTTGAAAAGAGTTTTG  
AGGTAGAAAACAAGCTTGAACCTGATAGGAAAGTTCAACTTGCAAAAGAGCTTGGTTTGCAA

FIGURA 4 (continuación)

CCAAGACAAGTTGCTATATGGTTTCAAAACAGAAGGGCAAGGTTCAAAACTAAACAGCTTGA  
 AAAAGATTATGGCACATTGAAAGCTAGCTTTGATAGTCTCAAAGATGATTATGATAATCTTC  
 TTCAAGAGAATGACAAGTTAAAAGAAGAGGTGAATTCTCTCAAGAACAAATTGATCCCAAGA  
 GATAAAGAAAAAGTGAATTCAGAAGACAAATCATCACCAGAAGCAATCAATTACCTCATAA  
 CAACATAGATCCAATGGATATAATTTCAATTACAAATTCAGAAAATGGGTCCAAAATGTCAC  
 TCCCTAATATGGTACTAAAATGTAAGCAAGAAGATGCCAATTCAGCTAAAAGTGATGTGCTT  
 GATTCTGATAGCCCATTTGCAATGATGGGAACAATCTTTCTTCTTTCATAGAGCCTACAGA  
 TTCAGATTTCTCACAAGATGAAGAGGATAATGATAACTTGAGTCATAATCTTTTGACTCTTC  
 CTTGCTTACCAAAAGTTGAAGATGTTTGCTATGATGACCCACATGAAAATCTTTGTAATTTT  
 GGGTCCCTGTTGAAGATCAAACCTTTTGTCTTCTGGCCTTATTGA

**SEC ID N° 47** Secuencia de aminoácidos traducida de Medtr\_HOX16\_1 de *Medicago truncatula*  
 MAGGKLFGGSNMSLLLQNERLPCTSEVLES LWHTPASFQGSNSVNFENGSGSNRVVTDRP  
 FFQOLEKEENCDEDEYAC YHQGKKRRLSSEQVQFLEKSFEVENKLEPDRKVQLAKELGLQ  
 PRQVAIWQNRRRARFKTKQLEKDYGTLKASFDLSKDDYDNLQENDKLKEEVNSLKNKLI PR  
 DKEKVNSEDKSSPEAINS PHNNIDPMDI I S ITNSENGSKMSLPNMVLKCKQEDANS AKSDVL  
 DSDSPHCNDGNLSSFI EPTDSDFSQDEEDNDNL SHNLLTLPCLPKVEDVCYDDPHENSCNF  
 GFPVEDQTFCFWPY

**SEC ID N° 48** Secuencia de ADNc de Phavu\_HOX16 de *Phaseolus vulgaris*  
 ATGGCGGGTGGCAAGCTTCATCCTGGTTCAAACATGTCACTTCTCCTCCAAAACGACAGGCT  
 CCCTTGCTCCTCTGAAGTCCTTGAGTCTCTTTGGGCTCACACCTCTAACGCTGCTTCCTTCC  
 AAGGTTCAAATCTATGGTTGATTTTGAGAATGTTAGTGGGGGCAGGGTGACGGATAGGCCC  
 TTTTTTCAAGCGTTGGAGAAGGAAGATAACTGTGATGATGATTATGAGGGTTGCTTCCATCA  
 ACCGGGTAAGAAAAGGAGGCTCACAAGCGAACAAGTTCAGTTCCTTGAAAGGAACTTTGAGG  
 TCGAGAACAAGCTTGAACCTGAAAGGAAGGTCCAACCTTGCAAAGGAGCTTGGCTTGCAGCCA  
 AGGCAAGTGGCTATATGGTTCCAAAACCGAAGGGCAAGGTTCAAGACCAAGCAGCTAGAAAA  
 AGATTATGGCACATTGAAAGCTAGCTATGACAGACTCAAAGGTGACTATGAAAGTCTTCTTC  
 AAGAGAATGACAAGTTAAAAGCAGAGGTGAATTCTCTGGAGAGCAAATTGATTCTTAGAGAT  
 AAAGAGAAGGAGAATTCGGACGACAAGTCATCTCCTGATGCTGTCAATTCACCCACAAAGA  
 GCCTATGGATTTAATTTCAAATTCACATCTGAAAATGGGACCAAAGTGTCACTCCCTATTA  
 TGGTAACATGCAAGCAAGAAGATGCCAATTCAGCCAAAAGTGATGTGCTTGATTTCGGACAGC  
 CCACATTGCACTGATGGGAACCATCCCTCTTCATTCTGCGTGGAGCCTGCTGATTCTCCTCCCATGC  
 TTTTGAACCAGACCACTCCGACTTCTCCCAAGATGAAGAGGATAATCTTAGTGAAAGCCTTT  
 TGACCCTCCCTTGCTTACCAAAGGTTGAAGAAGCCTGCTATGATGACCCTCCTGAAAACCCT  
 TGTAATTTTGGCTTCCATGTGCGAGGATCAAACCTTCTGTTTCTGGCCCTATTGA

**SEC ID N°: 49** Secuencia de aminoácidos traducida de Phavu\_HOX16 de *Phaseolus vulgaris*  
 MAGGKLHPGSNMSLLLQNDRLPCSSEVLES LWAHTSNAASFQGSKSMVDFENVSGGRVTDRP  
 FFQALEKEDNCDDEYEGCFHQPGKKRRLTSEQVQFLERNFEVENKLEPERKVQLAKELGLQP  
 RQVAIWQNRRRARFKTKQLEKDYGTLKASYDRLKGDYESLLQENDKLKAEVNSLESKLI LRD  
 KEKENSDDKSSPDVNSPHKEPMDLISNSTSENGTKVSLPIMVTCKQEDANS AKSDVLDSDS  
 PHCTDGNHPSSFVEPADSSHAFEPDHSDFSQDEEDNLSESLLTLPCLPKVEEACYDDPPENP  
 CNFGFHVEDQTFCFWPY

FIGURA 4 (continuación)

**SEC ID Nº: 50 Secuencia de ADNc de Lotco\_Hox16 de *Lotus corniculatus***

ATGGCGGGAGGGAGGGTCTTTAGCGGCGGTTCTGCTGCTCCTGCAAATGTTTCCGATACCAG  
TCTTTTGCTTCAGAATCAACCTCCTGATTCTTCTCTCTTCTCTCTACCTCTGCTTCTTTTC  
TCGGTTCAAGATCCATGGTGAGCTTCGCAGATAATAAATTAGGGCAAACGCGGTTCGTCTTC  
TCCGCGTTTGACCTCGATGAGAACGGCGATGAGGTCATGGACGAGTACTTTCACCAATCGGA  
GAAGAAGCGCCGTCTCTCTGTTGACCAAGTTCAGTTTCTGGAGAAGAGCTTCGAGGTGGATA  
ACAAGCTCGAACCTGACAGGAAAACCAAGATTGCCAAGGACCTTGGTTTGCAGCCACGCCAA  
GTCGCAATCTGGTTCCAGAACCGCCGTGCACGGTGGAAGACGAAACAGCTTGAGAAGGATTA  
TGATTCTCTGCATAGTAGCTTTGAGAGTCTCAAATCCAACCTATGATAATCTTCTCAAGGAGA  
AAGACATGTTAAAAGCTGAGGTGGCAAGTCTCACTGAGAAGGTGCTTGCAAGAGAGAATTTG  
AAACAAGTTGAAAGTGAAACAAAGGGATTGGTTGAACCAACCCCAAAGGCCTTTACTTTGATTC  
AGTTTCAGAGGGTGAAGAATCTAAAGTCTCTGTTGGGGCTTGTAACATGAGGATATCAGTT  
CAGCCAGGAGTGAGAGTTTGGATTCTGATAGCCACGTTACAGGGATGGATATGGAGTTAAC  
TCAGCAGTGCTAGAGACATGTGATTCTTCTTATGTGGTTGAACCTGATCAATCGGATATGTC  
ACAGGATGAGGAAGACAACCTGACCAAGACCTGTTGCCTCCATACATGTTTTCCAACTTG  
GAGATATGGATTACTCCGACCCGCCTGAAAGTTCATGTAATTTCCGATTTCGGGAGGAAGAT  
CATGCCCTTTGGTCATGGTCTTACTGA

**SEC ID Nº 51 Secuencia de aminoácidos traducida de Lotco\_Hox16 de *Lotus corniculatus***

MAGGRVFSGGSAA PANVSDTSLLLQNQPPDSSLFLSTSASFLGSRSMVSFADNKLQTRSF  
SAFDLDENGDEVMDEYFHQSEKKRRLSVDQVQFLEKSFEVDNKLEPDRKTKIAKDLGLQPRQ  
VAIWQNRARRWTKQLEKDYSLSHSSFESLSNYDNLLKEKMLKAEVASLTEKVLARENL  
KQVESETKGLVEPPQRPPLDSVSEGEESKVS VGACKHEDISSARSES LDDSPRYRDGYGVN  
SAVLETCDSSYVVEPDQSDMSQDEEDNLTKTLLPPYMF SKLGMDYSDPPESSCNFGFPEED  
HALWSWSY

**FIGURA 4 (continuación)**