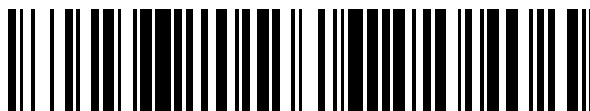


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 558 745**

51 Int. Cl.:

C12N 15/82 (2006.01)

C07K 14/415 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.05.2008 E 11192693 (7)**

97 Fecha y número de publicación de la concesión europea: **14.10.2015 EP 2505652**

54 Título: **Plantas que tienen rasgos potenciados relacionados con el rendimiento y un procedimiento de producción de las mismas**

30 Prioridad:

08.05.2007 US 91657507 P
06.06.2007 US 94221407 P
29.06.2007 US 93798907 P
05.07.2007 US 94803607 P 03.05.2007 EP
07107448 29.05.2007 EP 07109052 29.05.2007
EP 07109068 11.06.2007 EP 07109961
19.06.2007 EP 07110548 19.06.2007 EP 07110557

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.02.2016

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 558 745 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Plantas que tienen rasgos potenciados relacionados con el rendimiento y un procedimiento de producción de las mismas

5 La presente invención se refiere, en líneas generales, al campo de la biología molecular y atañe a un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla mediante el aumento de la expresión en una planta de un ácido nucleico que codifica un polipéptido dedo de homeodominio de planta-homeodominio (PHDf-HD, por sus siglas en inglés *plant homeodomain finger-homeodomain*). La presente invención también atañe a plantas que tienen expresión aumentada de un ácido nucleico que codifica un polipéptido PHDf-HD, cuyas plantas tienen rasgos relacionados con el rendimiento de semilla potenciados, con respecto a las correspondientes plantas de tipo silvestre u otras plantas de control. La invención también proporciona construcciones útiles para los procedimientos de la invención.

15 El continuo aumento de la población mundial y la disminución de suministro de tierra cultivable disponible para la agricultura incentivan la investigación hacia el aumento de la eficiencia de la agricultura. Los medios convencionales para mejoras de cultivo y hortícolas utilizan técnicas de siembra selectiva para identificar plantas que tienen características deseables. Sin embargo, dichas técnicas de siembra selectiva tienen diversos inconvenientes, en concreto, que estas técnicas son típicamente laboriosas y producen plantas que a menudo contienen componentes genéticos heterogéneos que no siempre producen el rasgo deseable que se transmite desde las plantas progenitoras. Los avances en biología molecular han permitido a la humanidad modificar el plasma germinal de animales y plantas. La ingeniería genética de las plantas conlleva el aislamiento y la manipulación de material genético (típicamente en forma de ADN o ARN) y la introducción posterior de ese material genético en una planta. Dicha tecnología tiene la capacidad de producir cultivos o plantas que tienen diversos rasgos económicos, agronómicos u hortícolas mejorados.

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

30 El rendimiento de semilla es un rasgo particularmente importante, ya que las semillas de muchas plantas son importantes para la nutrición humana y animal. Cultivos tales como maíz, arroz, trigo, canola y soja representan más de la mitad de la ingesta calórica humana total, bien a través del consumo directo de las propias semillas o a través del consumo de productos cárnicos generados sobre semillas procesadas. También son una fuente de azúcares, aceites y muchos tipos de metabolitos utilizados en los procesos 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 la transferencia de metabolitos desde las raíces, hojas y tallos al interior de la semilla en crecimiento. El endospermo, en particular, asimila los precursores metabólicos de los carbohidratos, aceites y proteínas y los sintetiza en macromoléculas de almacenamiento para llenar el grano.

40 Otro rasgo importante para muchos cultivos es el vigor temprano. Mejorar el vigor temprano es un objetivo importante de los programas de siembra de arroz modernos en variedades de cultivo de arroz tanto de clima templado como tropical. Las raíces largas son importantes para el anclaje adecuado al suelo en arroz sembrado en agua. Cuando el arroz se siembra directamente en campos inundados, y cuando las plantas deben emerger rápidamente a través del agua, los brotes más alargados se asocian con el vigor. Cuando se practica siembra con perforación, para la buena emergencia de las plántulas son importantes mesocótilos y coleótilos más alargados. La capacidad de modificar genéticamente el vigor temprano en plantas sería de gran importancia en la agricultura. Por ejemplo, un vigor temprano pobre ha sido una limitación para la introducción de híbridos de maíz (*Zea mays* L.) a base de plasma germinal del Cinturón de Maíz en el Atlántico Europeo.

50 Un rasgo adicional de importancia es el de tolerancia mejorada al estrés abiótico. El estrés abiótico es la causa principal de pérdidas de cultivo en todo el mundo, reduciendo en más de un 50 % los rendimientos promedio para la mayoría de las plantas de cultivo principales (Wang y col. *Planta* (2003) 218: 1-14). El estrés abiótico puede producirse por sequía, salinidad, temperaturas extremas, toxicidad química, y estrés oxidativo. La capacidad de mejorar la tolerancia de la planta al estrés abiótico sería de gran ventaja económica para los granjeros de todo el mundo y permitiría la cosecha de cultivos durante condiciones adversas y en territorios cuando la cosecha de cultivos no pueda ser posible de otra forma.

Por lo tanto el rendimiento del cultivo puede aumentarse optimizando uno de los factores mencionados anteriormente.

Dependiendo del uso final, la modificación de determinados rasgos de rendimiento puede estar favorecida sobre otros. Por ejemplo, para aplicaciones tales como forraje o producción de madera, o fuente de biocombustible, puede ser deseable un aumento en las partes vegetativas de una planta, y para aplicaciones tales como producción de harina, almidón o aceite, el aumento en los parámetros de semilla puede ser particularmente deseable. Incluso entre los parámetros de semilla, algunos pueden estar favorecidos sobre otros, dependiendo de la aplicación. Diversos mecanismos pueden contribuir a aumentar el rendimiento de las semillas, ya sea que esté en la forma de tamaño de semilla aumentado o número de semilla aumentado.

Una estrategia para aumentar el rendimiento (rendimiento de semillas y/o biomasa) en las plantas puede ser a través de la modificación de los mecanismos de crecimiento intrínsecos de una planta, tal como el ciclo celular o diversas rutas de señalización implicados en el crecimiento de la planta o en los mecanismos de defensa.

De forma sorprendente, ahora se ha encontrado que el aumento de la expresión de un ácido nucleico que codifica un polipéptido dedo de homeodominio de planta-homeodominio (PHDf-HD), proporciona plantas que tienen rasgos potenciados relacionados con el rendimiento de semilla con respecto a las plantas de control.

De acuerdo con una realización, se proporciona un procedimiento para mejorar o potenciar rasgos relacionados con el rendimiento de semilla de una planta con respecto a las plantas de control, que comprende aumentar en una planta la expresión de un ácido nucleico que codifica un polipéptido dedo de homeodominio de planta-homeodominio (PHDf-HD).

Antecedentes

Los factores de transcripción normalmente se definen como proteínas que muestran afinidad de unión a ADN específica de secuencia y que son capaces de activar y/o reprimir la transcripción. El genoma de *Arabidopsis thaliana* codifica al menos 1533 reguladores transcripcionales, lo que explica el ~ 5,9 % de su número total estimado de genes (Riechmann y col. (2000) Science 290: 2105-2109). La base de datos *The Database of Rice Transcription Factors* (DRTF), es una colección de factores de transcripción conocidos y predichos de *Oryza sativa L. ssp. indica* y *Oryza sativa L. ssp. japonica* y, actualmente, contiene 2.025 supuestos modelos génicos de factores de transcripción (FT) en *indica* y 2.384 in *japonica*, distribuidos en 63 familias (Gao y col. (2006) Bioinformatics 2006, 22(10):1286-7).

Una de estas familias es la superfamilia de los factores de transcripción homeodominio (HD), implicados en muchos aspectos de los procesos del desarrollo. Los factores de transcripción HD se caracterizan por la presencia de un homeodominio (HD), que es un dominio de unión (DU) a ADN de 60 aminoácidos. *Arabidopsis thaliana* y el arroz contienen aproximadamente 100 factores de transcripción HD, los que adicionalmente pueden clasificarse en subfamilias a base de la identidad de secuencia de aminoácidos (Richardt y col. (2007) Plant Phys 143(4): 1452-1466). Algunas de estas subfamilias se caracterizan por la presencia de dominios conservados adicionales que facilitan la unión a ADN y/o las interacciones proteína-proteína.

Uno de estos dominios es el dedo PHD, llamado dedo de homeodominio de planta (PHDf, por sus siglas en inglés *plant homeodomain finger*) debido a su asociación en un mismo polipéptido con un HD de unión a ADN, que se identificó originalmente mediante identidad de secuencia de aminoácidos entre un factor de transcripción homeosecuencia de maíz, ZmHOX1a, (Bellman y Werr (1992) EMBO J 11:3367-3374) y su pariente de *Arabidopsis* ATHAT3.1 (Schindler y col. (1993) Plant J 4: 137-150). EL PHDf es un motivo similar a dedo de zinc Cys₄-His-Cys₃, capaz de quelar dos iones zinc. Los PHDfs se encuentran en proteínas nucleares y se cree que están implicados en la regulación transcripcional mediada por cromatina (Halbach y col. (2000) Nucleic Acid Res 28(18): 3542-3550).

Por lo tanto, los factores transcripcionales que combinan un PHDf y un HD son llamados PHDf-HD (Halbach y col. (2000), citado anteriormente). En plantas tales PHDf-HD se caracterizan, adicionalmente, por la presencia de una cremallera de leucina (ZIP, abreviatura del inglés *zipper*: cremallera) cadena arriba del PHDf. Ambos dominios (el ZIP y el PHDf) juntos forman una región de 180 aminoácidos altamente conservada llamada el dominio ZIP/PHDf (Halbach y col. (2000), citado anteriormente).

Plantas de tabaco transgénicas sobreexpresando fuertemente cualquiera de los dos polipéptidos PHDf-HD (ZmHOX1a o ZmHOX1b) de maíz, utilizando un promotor 35S del virus del mosaico de la coliflor, combinado con un potenciador omega, mostraron cambios morfológicos idénticos: reducción del tamaño, formación de raíces adventicias, y transformaciones florales homeóticas (Uberlacker y col. (1996) Plant Cell 8: 349-362). Plantas de arroz y de tabaco transgénicas sobreexpresando fuertemente el polipéptido PHDf-HD HAZ1 de *Oryza sativa*, utilizando el promotor 35S del virus del mosaico de la coliflor, no mostraron crecimiento anormal o cambio fenotípico en comparación con los tipos silvestres (Ito y col. (2004) Gene 331: 9-15).

En la patente de Estados Unidos 7.196.245, un polipéptido PHDf-HD de *Arabidopsis thaliana* (identificado como G416) se transformó en *Arabidopsis*, y mostró que promovía la floración temprana en las plantas transgénicas, en comparación con las plantas de control, sin impacto en el rendimiento de semilla.

Sorprendentemente, ahora se ha encontrado que la expresión aumentada de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD proporciona plantas que tienen rasgos relacionados con el rendimiento de semilla potenciados, con respecto a las plantas de control.

De acuerdo con una realización, se proporciona un procedimiento para potenciar en plantas rasgos relacionados con el rendimiento de semilla, con respecto a las plantas de control, que comprende aumentar la expresión de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD en una planta. Los rasgos relacionados con el rendimiento de semilla potenciados comprenden uno o más de: número de panículas primarias aumentado, rendimiento total de semilla por planta aumentado, número de semillas (llenas) aumentado, peso de mil granos (PMG) aumentado, índice de cosecha aumentado.

Definiciones

Polipéptido(s)/proteína(s)

Los términos “polipéptido” y “proteína” se usan indistintamente en el presente documento y se refieren a aminoácidos en una forma polimérica de cualquier longitud, unidos entre sí por enlaces peptídicos.

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

En el presente documento, los términos “polinucleótido(s)”, “secuencia(s) de ácidos nucleicos”, “secuencia(s) de nucleótidos”, “ácido(s) nucleico(s)” y “molécula de ácido nucleico” se usan indistintamente y se refieren a nucleótidos, ya sean ribonucleótidos o desoxirribonucleótidos o una combinación de ambos, en una forma polimérica no ramificada de cualquier longitud.

Planta(s) de control

La elección de plantas de control adecuadas es una parte rutinaria de un montaje experimental y puede incluir plantas de tipo silvestre correspondientes o plantas correspondientes sin el gen de interés. La planta de control es típicamente de la misma especie de planta o incluso de la misma variedad de planta 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. Los nulicigotos son individuos que perdieron el transgén por segregación. Una “planta de control” 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.

Homólogo(s)

Los “homólogos” de una proteína incluyen péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones, deleciones y/o inserciones de aminoácidos con respecto a la proteína no modificada en cuestión y que tienen actividad funcional y biológica similar como la proteína no modificada de la cual proceden.

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

Una inserción se refiere a uno o más restos de aminoácido que se introducen en un sitio predeterminado en una proteína. Las inserciones pueden comprender fusiones de N y/o C terminal, así como inserciones intra-secuencia de un solo aminoácido o múltiples aminoácidos. Generalmente, las inserciones en la secuencia de aminoácidos serán más pequeñas que las fusiones de N o C terminal, del orden de aproximadamente 1 a 10 restos. Los ejemplos de proteínas de fusión de N o C terminal o péptidos incluyen el dominio de unión o dominio de activación de un activador transcripcional como se usa en el sistema de doble híbrido de levadura, proteínas de cubierta de fago, etiqueta de (histidina)-6, etiqueta 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, epítipo VSV.

Una sustitución se refiere a un reemplazo de aminoácidos de la proteína con otros aminoácidos que tienen propiedades semejantes (tales como hidrofobicidad, hidrofiliidad, antigenicidad, propensidad para formar o romper estructuras α -helicoidales o β laminares semejantes). Las sustituciones de aminoácidos son típicamente de restos únicos, pero pueden agruparse dependiendo de las limitaciones funcionales puestas al polipéptido y variar desde 1 a 10 aminoácidos; las inserciones normalmente serán del orden de aproximadamente 1 a 10 restos de aminoácido. Las sustituciones de aminoácidos son preferentemente sustituciones de aminoácidos conservativas. En la técnica se conocen bien las tablas de sustitución conservativa (véase, por ejemplo, Creighton (1984) Proteins. W. H. Freeman and Company (Eds) y la Tabla 1 a continuación).

Tabla 1: Ejemplos de sustituciones de aminoácidos conservativas

Resto	Sustituciones conservativas	Resto	Sustituciones conservativas
Ala	Ser	Leu	Ile; Val
Arg	Lys	Lys	Arg; Gln
Asn	Gln; His	Met	Leu; Ile

(continuación)

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

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

10 Derivados

Los “derivados” incluyen péptidos, oligopéptidos, polipéptidos que pueden, en comparación con la secuencia de aminoácidos de forma natural de la proteína, tal como la proteína de interés, comprender sustituciones de aminoácidos con restos de aminoácido de origen no natural, o adiciones de restos de aminoácido de origen no natural. Los “derivados” de una proteína también incluyen péptidos, oligopéptidos, polipéptidos, que comprenden restos de aminoácido modificados de origen natural (glucosilado, acilado, prenilado, fosforilado, miristoilado, sulfatado, etc.) o modificados de origen no natural en comparación con la secuencia de aminoácidos de forma no natural del polipéptido. Un derivado también puede comprender uno o más sustituyentes o adiciones que no son de aminoácido en comparación con la secuencia de aminoácidos de la que procede, por ejemplo una molécula indicadora u otro ligando, unido covalente o no covalentemente a la secuencia de aminoácidos, tal como una molécula indicadora que se une para facilitar su detección, y restos de aminoácido que no son de origen natural con respecto a la secuencia de aminoácidos de una proteína de origen natural. Además, los “derivados” también incluyen fusiones de la forma de origen natural de la proteína con péptidos etiquetadores, tales como FLAG, HIS6 o tiorredoxina (para una revisión de péptidos etiquetadores, véase Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

25 Ortólogo(s)/Parólogo(s)

Los ortólogos y parálogos incluyen conceptos evolutivos usados para describir las relaciones antecesoras de los genes. Los parálogos son genes dentro de la misma especie que se han originado a través de la duplicación de un gen antecesor; los ortólogos son genes de diferentes organismos que se han originado a través de la especiación, y también proceden de un gen antecesor común.

30 Dominio

El término “dominio” se refiere a un conjunto de aminoácidos conservados en posiciones específicas a lo largo de un alineamiento de secuencias de las proteínas evolutivamente relacionadas. Aunque los aminoácidos en otras posiciones pueden variar entre los homólogos, los aminoácidos que están muy conservados en posiciones específicas indican aminoácidos que son probablemente esenciales en la estructura, estabilidad o función de una proteína. Identificado por su alto grado de conservación en las secuencias alineadas de una familia de homólogos de proteína, éstas pueden usarse como identificadores para determinar si cualquier polipéptido en cuestión pertenece a una familia polipeptídica previamente identificada.

Motivo/Secuencia consenso/Firma

Las expresiones “motivo” o “secuencia consenso” o “firma” se refieren a una región conservada corta en la secuencia de las proteínas evolutivamente relacionadas. Los motivos son partes de dominios frecuentemente muy conservadas, pero también pueden incluir solo parte del dominio, o localizarse fuera del dominio conservado (si todos los aminoácidos del motivo se encuentran fuera de un dominio definido).

Hibridación

El término "hibridación", como se define en el presente documento, es un proceso en el que secuencias de nucleótidos complementarias, sustancialmente homólogas, se hibridan entre sí. El proceso de hibridación puede producirse completamente en solución, es decir, ambos ácidos nucleicos complementarios están en solución. El proceso de hibridación también puede producirse con uno de los ácidos nucleicos complementarios inmovilizado en una matriz, tal como perlas magnéticas, perlas de Sefarosa o cualquier otra resina. El proceso de hibridación puede producirse adicionalmente con uno de los ácidos nucleicos complementario inmovilizado en un soporte sólido, tal como una membrana de nitrocelulosa o de nailon, o inmovilizado, por ejemplo, mediante fotolitografía a, por ejemplo, un soporte vítreo silíceo (el último conocido como matrices o micromatrices de ácido nucleico o como microplacas de ácido nucleico). Para permitir que se produzca la hibridación, las moléculas de ácido nucleico están, generalmente, térmica o químicamente desnaturalizadas para fundir una doble cadena en dos cadenas únicas y/o retirar horquillas u otras estructuras secundarias de ácidos nucleicos monocatenarios.

El término "rigurosidad" se refiere a las condiciones en las que se realiza la hibridación. La rigurosidad de la hibridación está influenciada por condiciones tales como temperatura, concentración salina, fuerza iónica y composición del tampón de hibridación. Generalmente, se seleccionan condiciones de baja rigurosidad por ser aproximadamente 30 °C menores que el punto de fusión térmico (T_f) para la secuencia específica a una fuerza iónica definida y pH. Las condiciones de rigurosidad media son cuando la temperatura es 20 °C por debajo de la T_f , y las condiciones de rigurosidad alta son cuando la temperatura es 10 °C por debajo de la T_f . Las condiciones de hibridación de rigurosidad alta se usan típicamente para aislar secuencias de hibridación que tiene alta similitud de secuencia con la secuencia de ácidos nucleicos diana. Sin embargo, los ácidos nucleicos pueden desviarse en cuanto a la secuencia y aún codificar un polipéptido sustancialmente idéntico, debido a la generación del código genético. Por lo tanto, en ocasiones, pueden necesitarse condiciones de hibridación de rigurosidad media para identificar dichas moléculas de ácidos nucleicos.

La T_f es la temperatura bajo fuerza iónica definida y pH, en la que el 50 % de la secuencia diana se 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 más altas. La tasa máxima de hibridación se obtiene de 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 cadenas de ácido nucleico por lo cual se promueve la formación del híbrido; este efecto es visible para concentraciones de sodio de hasta 0,4 M (para mayores concentraciones, este efecto puede ignorarse). La formamida reduce la temperatura de fusión de dúplex de ADN-ADN y ADN-ARN con 0,6 a 0,7 °C para cada porcentaje de formamida, y la adición de formamida al 50 % permite que se realice la hibridación de 30 a 40 °C, aunque la tasa de hibridación disminuirá. Los emparejamientos erróneos de los pares de bases reducen la tasa de hibridación y la estabilidad térmica de los dúplex. Por término medio y para sondas grandes, 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):

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

2) híbridos de ADN-ARN o ARN-ARN:

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

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

$$\text{Para } <20 \text{ nucleótidos: } T_f = 2 (I_n)$$

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

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

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

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

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

La unión no específica puede controlarse usando una cualquiera de las numerosas técnicas conocidas, tales como, por ejemplo, bloqueo de la membrana con soluciones que contienen la proteína, adiciones del ARN heterólogo, ADN y SDS al tampón de hibridación y tratamiento con ARNasa. Para sondas no homólogas, puede realizarse una serie de hibridaciones al variar uno de (i) disminuir progresivamente la temperatura de hibridación (por ejemplo de 68 °C a 42 °C) o (ii) disminuir progresivamente la concentración de formamida (por ejemplo del 50 % al 0 %). El experto en la técnica es consciente de que durante la hibridación pueden alterarse diversos parámetros y que se mantendrán o cambiarán las condiciones de rigurosidad.

Además de las condiciones de hibridación, la especificidad de hibridación también depende típicamente de la función de los lavados post-hibridación. Para retirar el fondo resultante de la hibridación no específica, las muestras se lavan con soluciones salinas diluidas. Los factores críticos de dichos lavados incluyen la fuerza iónica y la

temperatura de la solución de lavado final: a menor concentración salina y mayor temperatura de lavado, mayor rigurosidad de lavado. Las condiciones de lavado se realizan típicamente a, o por debajo de, la rigurosidad de hibridación. Una hibridación positiva da lugar a una señal que es al menos dos veces la del fondo. Generalmente, las condiciones de rigurosidad adecuadas para los ensayos de hibridación de ácidos nucleicos o los procedimientos de detección de amplificación de genes son como se expusieron anteriormente. También pueden seleccionarse condiciones más o menos rigurosas. El experto en la técnica es consciente de que, durante el lavado, pueden alterarse diversos parámetros y que mantendrán o cambiarán las condiciones de rigurosidad.

Por ejemplo, las condiciones de hibridación típicas de alta rigurosidad para híbridos de ADN mayores de 50 nucleótidos incluyen una hibridación a 65 °C en 1x SSC o a 42 °C en 1x SSC y formamida al 50 %, seguido de lavado a 65 °C en 0,3x SSC. Los ejemplos de condiciones de hibridación de rigurosidad media para híbridos de ADN más largos de 50 nucleótidos incluyen una hibridación a 50 °C en 4x SSC o a 40 °C en 6x SSC y formamida al 50 %, seguido de lavado a 50 °C en 2x SSC. La longitud del híbrido es la longitud prevista para la hibridación del ácido nucleico. Cuando los ácidos nucleicos de secuencia conocida se hibridan, la longitud del híbrido puede determinarse alineando las secuencias e identificando las regiones conservadas descritas en el presente documento. 1x SSC es NaCl 0,15 M y citrato sódico 15 mM, la solución de hibridación y las soluciones de lavado pueden incluir adicionalmente reactivo de Denhardt 5x, SDS al 0,5-1,0 %, 100 µg/ml de ADN de esperma de salmón fragmentado, desnaturalizado 100 µg/ml, pirofosfato sódico al 0,5 %.

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 y actualizaciones anuales).

Variante de corte y empalme

La expresión “variante de corte y empalme” como se usa en el presente documento incluye variantes de una secuencia de ácidos nucleicos en la que los intrones y/o exones seleccionados se han escindido, reemplazado, desplazado o añadido, o en la que los intrones se han acortado o alargado. Dichas variantes serán aquellas en las que la actividad biológica de la proteína esté sustancialmente conservada; esto puede conseguirse conservando selectivamente los segmentos funcionales de la proteína. Dichas variantes de corte y empalme pueden encontrarse en la naturaleza o pueden ser sintetizadas por el hombre. En la técnica se conocen bien procedimientos para predecir y aislar dichas variantes de corte y empalme (véase, por ejemplo, Foissac y Schiex (2005) *BMC Bioinformatics* 6: 25).

Variante alélica

Los alelos o variantes alélicas son formas alternativas de un gen determinado, localizado en la misma posición cromosómica. Las variantes alélicas incluyen Polimorfismos Mononucleotídicos (SNP, por las siglas *Single Nucleotide Polymorphisms*), así como Polimorfismos Pequeños de Inserción/Delección (INDELs, por las siglas *Small Insertion/Deletion Polymorphisms*). El tamaño de los INDELs es normalmente menor de 100 pb. Los SNP y los INDELs forman la serie de variantes de secuencia más grande en las cepas polimórficas de origen natural de la mayoría de los organismos.

Combinación de genes/Evolución dirigida

La combinación de genes o evolución dirigida consiste en repeticiones de combinación de ADN seguido de exploración y/o selección apropiada para generar variantes de ácidos nucleicos o partes de las mismas que codifican proteínas que tienen una actividad biológica modificada (Castle y col., (2004) *Science* 304(5674): 1151-4; Patentes de Estados Unidos 5.811.238 y 6.395.547).

Elemento regulador/Secuencia de control/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 una secuencia de ácidos nucleicos reguladora capaz de efectuar la expresión de las secuencias a las que se unen. El término “promotor” se refiere típicamente a una secuencia de control de ácido nucleico localizada cadena arriba del inicio transcripcional de un gen y que está implicado en el reconocimiento y unión de la ARN polimerasa y otras proteínas, dirigiendo así la transcripción de un ácido nucleico unido operativamente. 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 reguladores 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.

Un “promotor de planta” comprende elementos reguladores, que median la expresión de un segmento de secuencia codificante en células de plantas. Por consiguiente, un promotor de planta no tiene que ser de origen vegetal, sino que puede originarse de virus o microorganismos, por ejemplo de virus que atacan a las células de planta. El “promotor de planta” también puede originarse de una célula de planta, por ejemplo, de la planta que se transforma con la secuencia de ácidos nucleicos a expresar en el procedimiento de la invención y descrito en el presente documento. Esto también se aplica a las otras señales reguladoras de “planta”, tales como terminadores de “planta”. Los promotores cadena arriba de la secuencias de nucleótidos útiles en los procedimientos de la presente invención pueden modificarse mediante una o más sustituciones, inserciones y/o deleciones de nucleótido sin interferir con la funcionalidad o actividad de cualquiera de los promotores, la fase de lectura abierta (ORF, *Open Reading Frame*) o la región reguladora 3’ tal como terminadores u otras regiones reguladoras 3’ que se localizan fuera de la ORF. Además es posible que la actividad de los promotores aumente por la modificación de su secuencia o que se reemplacen completamente mediante más promotores activos, incluso promotores de organismos heterólogos. Para la expresión en plantas, la molécula de ácido nucleico debe, como se ha descrito anteriormente, estar unida operativamente a, o comprender, un promotor adecuado que exprese el gen de manera adecuada, en tiempo y con el patrón de expresión espacial requerido.

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 la 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 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. Generalmente, por “promotor de fuerza media” se entiende un promotor que dirige la expresión de una secuencia codificante a un nivel más bajo que el de un promotor fuerte, en particular a un nivel que es, en todos los casos, inferior al obtenido cuando está bajo el control de un promotor CaMV 35S.

Unido operativamente

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.

Promotor constitutivo

Un “promotor constitutivo” se refiere a un promotor que es transcripcionalmente activo durante la mayor parte, aunque no necesariamente en todas, las fases de crecimiento y desarrollo y en la mayor parte de las condiciones ambientales, en al menos una célula, tejido u órgano. La siguiente Tabla 2a proporciona ejemplos de promotores constitutivos.

Tabla 2a: Ejemplos de promotores constitutivos

Fuente del gen	Referencia
Actina	McElroy y col, <i>Plant Cell</i> , 2: 163-171, 1990
HMGP	WO 2004/070039
CAMV 35S	Odell y col, <i>Nature</i> , 313: 810-812, 1985
CaMV 19S	Nilsson y col., <i>Physiol. Plant.</i> 100:456-462, 1997
GOS2	de Pater y col, <i>Plant J Nov</i> ; 2(6):837-44, 1992, WO 2004/065596
Ubiquitina	Christensen y col, <i>Plant Mol. Biol.</i> 18: 675-689, 1992
Ciclofilina de arroz	Buchholz y col, <i>Plant Mol Biol.</i> 25(5): 837-43, 1994
Histona H3 de maíz	Lepetit y col, <i>Mol. Gen. Genet.</i> 231:276-285, 1992
Histona H3 de alfalfa	Wu y col. <i>Plant Mol. Biol.</i> 11:641-649, 1988
Actina 2	An y col, <i>Plant J.</i> 10(1); 107-121, 1996
34S FMV	Sanger y col., <i>Plant. Mol. Biol.</i> , 14, 1990: 433-443

(continuación)

Fuente del gen	Referencia
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

Promotor ubicuo

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

Promotor regulado evolutivamente

- 5 Un promotor regulado evolutivamente es activo durante determinadas etapas del desarrollo o en partes de la planta que experimentan cambios evolutivos.

Promotor inducible

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

Promotor específico de órgano/específico de tejido

- 15 Un promotor específico de órgano o específico de tejido es uno que es capaz de iniciar preferentemente la transcripción en determinados órganos o tejidos, tales como hojas, raíces tejido de semillas, etc. Por ejemplo, un "promotor específico de raíz" es un promotor que es transcripcionalmente activo predominantemente en las raíces las de plantas, sustancialmente con la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de la planta. En el presente documento a los promotores capaces de iniciar la transcripción solo en determinadas células se les denomina promotores "específicos de célula".

A continuación en la Tabla 2b se indican ejemplos de promotores específicos de raíz:

- 20 **Tabla 2b:** Ejemplos de promotores específicos de raíz

Fuente del gen	Referencia
RCc3	Plant Mol Biol. 1995 Jan;27(2):237-48
<i>Arabidopsis</i> PHT1	Kovama y col., 2005; Mudge y col. (2002, Plant J. 31: 341)
Transportador de fosfato en <i>Medicago</i>	Xiao y col., 2006
<i>Arabidopsis</i> Pyk10	Nitz y col. (2001) Plant Sci 161(2): 337-346
genes expresables en raíz	Tingey y col., EMBO J. 6: 1, 1987.
gen inducible por auxina en tabaco	Van der Zaal y col., Plant Mol. Biol. 16, 983, 1991.
β -tubulina	Oppenheimer, y col., Gene 63: 87, 1988.
genes específicos de la raíz del tabaco	Conkling, y col. Plant Physiol. 93: 1203, 1990.
gen G1-3b de <i>B. napus</i>	Patente de Estados Unidos No. 5.401.836
SbPRP1	Suzuki y col., Plant Mol. Biol. 21: 109-119, 1993.
LRX1	Baumberger y col. 2001, Genes & Dev. 15:1128
BTG-26 de <i>Brassica napus</i>	US 20050044585

(continuación)

Fuente del gen	Referencia
LeAMT1 (tomate)	Lauter y col. (1996, PNAS 3:8139)
The LeNRT1-1 (tomate)	Lauter y col. (1996, PNAS 3:8139)
gen de la patatina de clase I (patata)	Liu y col., Plant Mol. Biol. 153:386-395, 1991.
KDC1 (<i>Daucus carota</i>)	Downey y col. (2000, J. Biol. Chem. 275:39420)
gen de TobRB7	W Song (1997) PhD Thesis, North Carolina State University, Raleigh, NC USA
OsRAB5a (raíz)	Wang y col. 2002, Plant Sci. 163:273
ALF5 (<i>Arabidopsis</i>)	Diener y col. (2001, Plant Cell 13:1625)
NRT2;1Np (<i>N. plumbaginifolia</i>)	Quesada y col. (1997, Plant Mol. Biol. 34:265)

5 Un promotor específico de semilla es transcripcionalmente activo de forma predominante en tejidos de semilla, pero no necesariamente exclusivamente en tejidos de semilla (en casos de expresión parcial). El promotor específico de semilla puede estar activo durante el desarrollo y/o durante la germinación de la semilla. El promotor específico de semilla puede ser específico de endospermo/aleurona/embrión. En las Tablas 2d, 2e, 2f, se muestran ejemplos de promotores específicos de semilla (específicos de endospermo/aleurona/embrión). Ejemplos adicionales de promotores específicos de semilla se proporcionan en Qing Qu y Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuya divulgación se incorpora como referencia como si se indicara completamente.

Tabla 2c: Ejemplos de promotores específicos de semilla

Fuente del gen	Referencia
genes específicos de 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	Pearson y col., Plant Mol. Biol. 18: 235-245, 1992.
legumina	Ellis y col., Plant Mol. Biol. 10: 203-214, 1988.
glutelina (arroz)	Takaiwa y col., Mol. Gen. Genet. 208: 15-22, 1986;
	Takaiwa y col., FEBS Letts. 221: 43-47, 1987.
zeína	Matzke y col Plant Mol Biol, 14(3):323-32 1990
napA	Stalberg y col, Planta 199: 515-519, 1996.
gluteína-1 de BPM y APM de trigo	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989
SPA de trigo	Albani y col, Plant Cell, 9: 171-184, 1997
α , β y γ gliadinas de trigo	EMBO J. 3:1409-15, 1984
promotor ltr1 de cebada	Diaz y col. (1995) Mol Gen Genet 248(5):592-8
hordeína B1, C, D de cebada	Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996
DOF de cebada	Mena y col, The Plant Journal, 116(1): 53-62, 1998
blz2	EP99106056.7
promotor sintético	Vicente-Carbajosa y col., Plant J. 13: 629-640, 1998.
prolamina NRP33 de arroz	Wu y col, Plant Cell Physiology 39(8) 885-889, 1998
α -globulina Glb-1 de arroz	Wu y col, Plant Cell Physiology 39(8) 885-889, 1998

ES 2 558 745 T3

(continuación)

Fuente del gen	Referencia
OSH1 de arroz	Sato y col, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996
α globulina REB/OHP-1 de arroz	Nakase y col. Plant Mol. Biol. 33: 513-522, 1997
ADP-glucosa pirofosforilasa de arroz	Trans Res 6:157-68, 1997
familia del gen ESR de maíz	Plant J 12:235-46, 1997
α -kafirina de sorgo	DeRose y col., Plant Mol. Biol 32:1029-35, 1996
KNOX	Postma-Haarsma y col, Plant Mol. Biol. 39:257-71, 1999
oleosina de arroz	Wu y col, J. Biochem. 123:386, 1998
oleosina de girasol	Cummins y col., Plant Mol. Biol. 19: 873-876, 1992
PRO0117, supuesta proteína ribosomal 40S de arroz	WO 2004/070039
PRO0136, alanina aminotransferasa de arroz	No publicado
PRO0147, inhibidor de tripsina ITR1 (cebada)	No publicado
PRO0151, WSI18 de arroz	WO 2004/070039
PRO0175, RAB21 de arroz	WO 2004/070039
PRO005	WO 2004/070039
PRO009	WO 2004/070039
α -amilasa (Amy32b)	Lanahan y col, Plant Cell 4: 203-211, 1992; Skriver y col, Proc Natl Acad Sci USA 88:7266-7270, 1991
gen similar a la β catepsina	Cejudo y col, Plant Mol Biol 20: 849-856, 1992
Ltp2 de cebada	Kalla y col., Plant J. 6: 849-60, 1994
Chi26	Leah y col., Plant J. 4: 579-89, 1994
B-Perú de maíz	Selinger y col., Genetics 149; 1125-38, 1998

Tabla 2d: ejemplos de promotores específicos de endospermo

Fuente del gen	Referencia
glutelina (arroz)	Takaiwa y col. (1986) Mol Gen Genet 208:15-22; Takaiwa y col. (1987) FEBS Letts. 221:43-47
zeína	Matzke y col., (1990) Plant Mol Biol 14(3): 323-32
gluteína-1 de BPM y APM de trigo	Colot y col. (1989) Mol Gen Genet 216:81-90, Anderson y col. (1989) NAR 17:461-2
SPA de trigo	Albani y col. (1997) Plant Cell 9:171-184
gliadinas de trigo	Rafalski y col. (1984) EMBO 3:1409-15
promotor Itr1 de cebada	Diaz y col. (1995) Mol Gen Genet 248(5):592-8
hordeína B1, C, D de cebada	Cho y col. (1999) Theor Appl Genet 98:1253-62; Muller y col. (1993) Plant J 4:343-55; Sorenson y col. (1996) Mol Gen Genet 250:750-60
DOF de cebada	Mena y col. (1998) Plant J 116(1): 53-62
blz2	Onate y col. (1999) J Biol Chem 274(14):9175-82
promotor sintético	Vicente-Carbajosa y col. (1998) Plant J 13:629-640

(continuación)

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

Tabla 2e: Ejemplos de promotores específicos de embrión:

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

Tabla 2f: Ejemplos de promotores específicos de aleurona:

Fuente del gen	Referencia
α -amilasa (Amy32b)	Lanahan y col, Plant Cell 4:203-211, 1992
	Skriver y col, Proc Natl Acad Sci USA 88:7266-7270, 1991
gen similar a β catepsina	Cejudo y col, Plant Mol Biol 20:849-856, 1992
Ltp2 de cebada	Kalla y col., Plant J. 6:849-60, 1994
Chi26	Leah y col., Plant J. 4:579-89, 1994
B-Perú de maíz	Selinger y col., Genetics 149:1125-38,1998

5 Un promotor específico de tejido verde como se define en el presente documento es un promotor que es transcripcionalmente activo predominantemente en tejido verde, sustancialmente con la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de la planta.

Los ejemplos de promotores específicos de tejido verde que pueden usarse para realizar los procedimientos de la invención se muestran a continuación en la Tabla 2g.

Tabla 2g: Ejemplos de promotores específicos de tejido verde

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

Otro ejemplo de un promotor específico de tejido es un promotor específico de meristemo, que es transcripcionalmente activo predominantemente en tejido meristemático, sustancialmente con la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de la planta. Los ejemplos de promotores específicos de meristemo verde que pueden usarse para realizar los procedimientos de la invención se muestran en la siguiente Tabla 2h.

Tabla 2h: Ejemplos de promotores específicos de meristemo

Fuente del gen	Patrón de expresión	Referencia
OSH1 de arroz	Meristemo apical de brote, desde la fase globular embrionaria a la fase de plántula	Sato y col. (1996) Proc. Natl. Acad. Sci. USA, 93: 8117-8122
Metalotioneína de arroz	Específico de meristemo	BAD87835.1
WAK1 y WAK 2	Meristemos apicales de brote y raíz, y en hojas y sépalos en expansión	Wagner y Kohorn (2001) Plant Cell 13(2): 303-318

Terminador

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 transcripto primario y la terminación de la transcripción. El terminador puede proceder del gen natural, de una variedad de otros genes de planta, o de ADN T. El terminador a añadir puede proceder, por ejemplo, de genes de nopalina sintasa u octopina sintasa, o como alternativa, de otro gen de planta, o menos preferentemente, de cualquier otro gen eucariota.

Modulación

El término "modulación" significa en relación a la expresión o expresión de un gen, un proceso en el que en el nivel de expresión se cambia mediante dicho gen de expresión en comparación con la planta de control, el nivel de expresión puede aumentarse o disminuirse. La expresión original, no modulada, puede ser de cualquier tipo de expresión de un ARN estructural (ARNr, ARNt) o ARNm con traducción posterior. La expresión "modulación de la actividad" significará cualquier cambio de la expresión de las secuencias de ácido nucleico de la invención o proteínas codificadas, que conduce a un rendimiento aumentado y/o a un crecimiento aumentado de las plantas.

Expresión

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

Expresión aumentada/sobreexpresión

La frase "expresión aumentada" o "sobreexpresión" como se usa en el presente documento, significa cualquier forma de expresión que es adicional al nivel de expresión original de tipo silvestre.

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 de un ácido nucleico que codifica el polipéptido de interés. Por ejemplo, pueden alterarse promotores endógenos *in vivo* por mutación, delección y/o sustitución (véase, Kmiec, US 5.565.350; Zarlign y col., WO9322443) o pueden introducirse promotores aislados en una célula de planta en la orientación y distancia apropiada de un gen de la presente invención para controlar la expresión del gen.

Si se desea la expresión del polipéptido, generalmente es deseable incluir una región de poliadenilación en el extremo 3' de una región codificante polinucleotídica. La región de poliadenilación puede proceder del gen natural, de una diversidad de otros genes de planta, 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 menos preferentemente 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).

5 Gen endógeno

En el presente documento la referencia a un gen "endógeno", no solo se refiere al gen en cuestión como se encuentra en una planta en su forma natural (es decir, sin existir ninguna intervención humana), sino también se refiere a aquel mismo gen (o un ácido nucleico/gen sustancialmente homólogo) en una forma aislada posteriormente (re)introducida en una planta (un transgén). Por ejemplo, una planta transgénica que contiene dicho transgén puede encontrar una reducción sustancial de la expresión transgénica y/o reducción sustancial de la expresión del gen endógeno. El gen aislado puede aislarse de un organismo o puede fabricarlo el hombre, por ejemplo, mediante síntesis química.

Expresión disminuida

En el presente documento la referencia a "expresión disminuida" o "reducción o eliminación sustancial" de expresión significa una disminución en la expresión del gen endógeno y/o los niveles de polipéptido y/o actividad de polipéptido con respecto a las plantas de control. La reducción o eliminación sustancial está en orden creciente de preferencia al menos 10 %, 20 %, 30 %, 40 % o 50 %, 60 %, 70 %, 80 %, 85 %, 90 % o 95 %, 96 %, 97 %, 98 %, 99 % o más reducido en comparación con el de las plantas de control.

Para la reducción o eliminación sustancial de la expresión de un gen endógeno en una planta, se requiere una longitud suficiente de nucleótidos sustancialmente contiguos de una secuencia de ácidos nucleicos. Para realizar el silenciamiento génico, éste puede ser tan pequeño como 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 o menos nucleótidos, como alternativa éste puede ser tanto como el gen completo (incluyendo la UTR 5' y/o 3', ya sea en parte o en su totalidad). El tramo de los nucleótidos sustancialmente contiguos puede proceder del ácido nucleico que codifica la proteína de interés (gene diana), o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo y homólogo de la proteína de interés. Preferentemente, el tramo de los nucleótidos sustancialmente contiguos es capaz de formar enlaces de hidrógeno con el gen diana (cadena en sentido o antisentido), más preferentemente, el tramo de nucleótidos sustancialmente contiguos tiene, en orden creciente de preferencia, una identidad de secuencia de 50 %, 60 %, 70 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % 100 % con el gen diana (cadena en sentido o antisentido). Una secuencia de ácidos nucleicos que codifica un polipéptido (funcional) no es un requisito para los diversos procedimientos analizados en el presente documento para la reducción o eliminación sustancial de la expresión de gen endógeno.

Esta reducción o eliminación sustancial de la expresión puede realizarse usando herramientas y técnicas rutinarias. Un procedimiento preferente para la reducción o eliminación sustancial de la expresión de un gen endógeno es introducir y expresar en una planta una construcción genética en la que el ácido nucleico (en este caso un tramo de nucleótidos sustancialmente contiguos procedentes del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo y homólogo de cualquiera de las proteínas de interés) se clona como una repetición invertida (en parte o completamente), separado por un espaciador (ADN no codificante).

En dicho procedimiento preferente, la expresión del gen endógeno se reduce o elimina sustancialmente a través de silenciamiento mediado por ARN usando una repetición invertida de un ácido nucleico o una parte del mismo (en este caso un tramo de nucleótidos sustancialmente contiguos procedentes del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), preferentemente capaz de formar una estructura en horquilla. La repetición invertida se clona en un vector de expresión que comprende secuencias de control. Una secuencia de ácidos nucleicos de ADN no codificante (un espaciador, por ejemplo un fragmento de la región de unión de matriz (MAR, *Matrix Attachment Region Fragment*), un intrón, un poliengarce, etc.) se localiza entre los dos ácidos nucleicos invertidos que forman la repetición invertida. Después de la transcripción de la repetición invertida, se forma un ARN quimérico con una estructura auto-complementaria (parcial o completa). Esta estructura de ARN bicatenario se denomina ARN en horquilla (ARNh) El ARNh se procesa en la planta en los ARNip que se incorporan en un complejo de silenciamiento inducido por ARN (RISC). El RISC escinde adicionalmente los transcritos de ARNm, por lo cual se reduce sustancialmente el número de transcritos de ARNm que van a traducirse en los polipéptidos. Para detalles adicionales generales véase, por ejemplo, Grierson y col. (1998) WO 98/53083; Waterhouse y col. (1999) WO 99/53050).

La realización de los procedimientos de la invención no se basa en introducir y expresar en una planta una construcción genética en la que se clona el ácido nucleico como una repetición invertida, sino que pueden utilizarse cualquiera de uno o más de los diversos procedimientos de "silenciamiento génico" bien conocidos para conseguir los mismos efectos.

Un procedimiento de este tipo para la reducción de la expresión del gen endógeno es el silenciamiento mediado por ARN de la expresión génica (regulación negativa). En este caso, el silenciamiento se activa en una planta mediante una secuencia de ARN bicatenario (ARNbc) que es sustancialmente similar al gen endógeno diana. Este ARNbc se

- 5 procesa adicionalmente por la planta de aproximadamente 20 a aproximadamente 26 nucleótidos denominado ARN de interferencia pequeño (ARNip). Los ARNip se incorporan en un complejo de silenciamiento inducido por ARN (RISC) que escinde el transcrito de ARNm del gen endógeno diana, reduciendo por tanto sustancialmente el número de transcritos de ARNm que se traducen en un polipéptido. Preferentemente, la secuencia de ARN bicatenario corresponde a un gen diana.
- Otro ejemplo de un procedimiento de silenciamiento de ARN implica la introducción de secuencias de ácidos nucleicos o partes de las mismas (en este caso un tramo de nucleótidos sustancialmente contiguos procedentes del gen de interés o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés) en una orientación en sentido en una planta. "Orientación en sentido" se refiere a una secuencia de ADN que es homóloga a un transcrito de ARNm de la misma. Por lo tanto, en una planta se introduciría al menos una copia de la secuencia de ácidos nucleicos. La secuencia de ácidos nucleicos adicional reducirá la expresión del gen endógeno, dando lugar a un fenómeno conocido como cosupresión. La reducción de la expresión del gen sería más pronunciada si en una planta se introdujesen varias copias adicionales de una secuencia de ácidos nucleicos, cuando hay una correlación positiva entre altos niveles de transcrito y la activación de cosupresión.
- 15 Otro ejemplo de un procedimiento de silenciamiento de ARN implica el uso de secuencias de ácidos nucleicos antisentido. Una secuencia de "ácidos nucleicos antisentido" comprende una secuencia de nucleótidos que es complementaria a una secuencia de ácidos nucleicos "en sentido" que codifica una proteína, es decir complementaria a la cadena que codifica una molécula de ADNc bicatenario o complementario a una secuencia de transcrito de ARNm. La secuencia de ácidos nucleicos antisentido es preferentemente complementaria a un gen endógeno a silenciar. La complementariedad puede localizarse en la "región codificante" y/o en la "región no codificante" de un gen. La expresión "región codificante" se refiere a una región de la secuencia de nucleótidos que comprende codones que se traducen en restos de aminoácidos. La expresión "región no codificante" se refiere a las secuencias 5' y 3' que flanquean la región codificante que se transcriben pero no se traducen en aminoácidos (también denominadas regiones no traducidas 5' y 3').
- 20
- 25 Las secuencias de ácidos nucleicos antisentido pueden diseñarse de acuerdo con las normas de emparejamiento de bases de Watson y Crick. La secuencia de ácidos nucleicos antisentido puede ser complementaria a la secuencia de ácidos nucleicos completa (en este caso un tramo de nucleótidos sustancialmente contiguos procedentes del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), pero también puede ser un oligonucleótido que sea antisentido para solo una parte de la secuencia de ácidos nucleicos (incluyendo la UTR 5' y 3' del ARNm). Por ejemplo, la secuencia de oligonucleótidos antisentido puede ser complementaria a la región que rodea el sitio de inicio de la traducción de un transcrito de ARNm que codifica un polipéptido. La longitud de una secuencia de oligonucleótidos antisentido adecuada se conoce en la técnica y puede comenzar desde aproximadamente 50, 45, 40, 35, 30, 25, 20, 15 o 10 nucleótidos de longitud o menos. Una secuencia de ácidos nucleicos antisentido, de acuerdo con la invención, puede construirse usando síntesis química y reacciones de ligamiento enzimático usando procedimientos conocidos en la técnica. Por ejemplo, una secuencia de ácidos nucleicos antisentido (por ejemplo, una secuencia de oligonucleótidos antisentido) puede sintetizarse químicamente usando nucleótidos de origen natural o varios nucleótidos modificados diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre las secuencias de ácidos nucleicos antisentido y sentido, por ejemplo, pueden usarse derivados de fosforotioato y nucleótidos sustituidos por acridina. En la técnica se conocen bien ejemplos de nucleótidos modificados que pueden usarse para generar las secuencias de ácidos nucleicos antisentido. Las modificaciones de nucleótido conocidas incluyen metilación, ciclación y 'protecciones' y sustitución de uno o más de los nucleótidos de origen natural con un análogo tal como inosina. En la técnica se conocen bien otras modificaciones de nucleótidos.
- 30
- 35
- 40
- 45 La secuencia de ácidos nucleicos antisentido puede producirse biológicamente usando un vector de expresión en el que se ha subclonado una secuencia de ácidos nucleicos en una orientación antisentido (es decir, el ARN transcrito del ácido nucleico insertado tendrá una orientación antisentido con respecto a un ácido nucleico diana de interés). Preferentemente, la producción de las secuencias de ácidos nucleicos antisentido en plantas se produce mediante una construcción de ácido nucleico establemente integrada que comprende un promotor, un oligonucleótido antisentido unido operativamente y un terminador.
- 50
- 55 Las moléculas de ácido nucleico usadas para el silenciamiento en los procedimientos de la invención (tanto si se introducen en una planta como si se generan *in situ*) se hibridan con, o se unen a, los transcritos de ARNm y/o ADN genómico que codifica un polipéptido para inhibir así la expresión de la proteína, por ejemplo, al inhibir la transcripción y/o traducción. La hibridación puede ser por complementariedad de nucleótidos convencional para formar un dúplex estable o, por ejemplo, en el caso de una secuencia de ácidos nucleicos antisentido que se une a los dúplex de ADN, a través de interacciones específicas en el surco principal de la hélice doble. Las secuencias de ácidos nucleicos antisentido pueden introducirse en una planta por transformación o inyección directa en un sitio de tejido específico. Como alternativa, las secuencias de ácidos nucleicos antisentido pueden modificarse para dirigir las células seleccionadas y después administrarse por vía sistémica. Por ejemplo, para la administración sistémica, las secuencias de ácidos nucleicos antisentido pueden modificarse de tal manera que se unan específicamente a receptores o a antígenos expresados en una superficie de la célula seleccionada, por ejemplo, ligando la secuencia de ácidos nucleicos antisentido a los péptidos o anticuerpos que se unen a los receptores o antígenos de la superficie celular. Las secuencias de ácidos nucleicos antisentido también pueden administrarse a las células
- 60

usando los vectores descritos en el presente documento.

De acuerdo con un aspecto adicional, la secuencia de ácidos nucleicos antisentido es una secuencia de ácidos nucleicos α -anomérica. Una secuencia de ácidos nucleicos α -anomérica forma híbridos bicatenarios específicos con ARN de complementariedad en el que, contrario a las unidades β normales, las cadenas corren paralelas entre sí (Gaultier y col. (1987) Nucl Ac Res 15: 6625-6641). La secuencia de ácidos nucleicos antisentido también pueden comprender un 2'-*o*-metilribonucleótido (Inoue y col. (1987) Nucl Ac Res 15, 6131-6148) o un análogo de ARN-ADN químico (Inoue y col. (1987) FEBS Lett. 215, 327-330).

La reducción o eliminación sustancial de la expresión del gen endógeno también puede realizarse usando ribozimas. Las ribozimas son moléculas de ARN catalíticas con actividad de ribonucleasa que pueden escindir una secuencia de ácidos nucleicos monocatenaria, tal como ARNm, con el que tienen una región complementaria. Por tanto, las ribozimas (por ejemplo, ribozimas de cabeza de martillo (descritas en Haselhoff y Gerlach (1988) Nature 334, 585-591) pueden usarse para escindir catalíticamente los transcritos de ARNm que codifican un polipéptido, reduciendo por tanto sustancialmente el número de transcritos de ARNm que van a traducirse en un polipéptido. Puede diseñarse una ribozima que tenga especificidad por una secuencia de ácidos nucleicos (véase, por ejemplo: Cech y col. Patentes de Estados Unidos N° 4.987.071; y Cech y col. Patente de Estados Unidos N° 5.116.742). Como alternativa, pueden usarse transcritos de ARNm correspondientes a una secuencia de ácidos nucleicos para seleccionar un ARN catalítico que tenga una actividad de ribonucleasa específica de un conjunto de moléculas de ARN (Bartel y Szostak (1993) Science 261, 1411-1418). El uso de ribozimas para el silenciamiento génico en plantas es conocido en la técnica (por ejemplo Atkins y col. (1994) documento WO 94/00012; Lenne y col. (1995) documento WO 95/03404; Lutziger y col. (2000) documento WO 00/00619; Prinsen y col. (1997) documento WO 97/13865 y Scott y col. (1997) documento WO 97/38116).

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

El silenciamiento génico también puede producirse si hay una mutación en un gen endógeno y/o una mutación en un gen aislado/ácido nucleico posteriormente introducido en una planta. La reducción de eliminación sustancial puede estar provocada por un polipéptido no funcional. Por ejemplo, el polipéptido puede unirse a diversas proteínas que interactúan; por lo tanto pueden proporcionarse una o más mutaciones y/o truncamientos para un polipéptido que sea aún capaz de unir las proteínas que interactúan (tal como proteínas receptoras) pero que no pueden mostrar su función normal (tal como ligando de señalización).

Una estrategia adicional para silenciamiento génico es dirigir la secuencia de ácidos nucleicos complementaria a la región reguladora del gen (por ejemplo, el promotor y/o los potenciadores) para formar estructuras helicoidales triples que impiden la transcripción del gen en células diana. Véase Helene, C., Anticancer Drug Res. 6, 569-84, 1991; Helene y col., Ann. N. Y. Acad. Sci. 660, 27-36 1992; y Maher, L. J. Bioassays 14, 807-15, 1992.

Otros procedimientos, tales como el uso de anticuerpos dirigidos contra un polipéptido endógeno para inhibir su función en la planta, o la interferencia en la ruta de señalización en la que está implicado un polipéptido, serán muy conocidos por el experto. En particular, puede esperarse que las moléculas fabricadas por el hombre puedan ser útiles para inhibir la función biológica de un polipéptido diana o para interferir con la ruta de señalización en la que está implicado el polipéptido diana.

Como alternativa, puede establecerse un programa de exploración para identificar en una población de plantas las variantes naturales de un gen, cuyas variantes codifican los polipéptidos con actividad reducida. Dichas variantes naturales también pueden usarse, por ejemplo, para realizar recombinación homóloga.

Pueden usarse microARN (miARN) artificiales y/o naturales para desactivar la expresión génica y/o traducción del ARNm. Los miARN endógenos son ARN pequeños monocatenarios típicamente con una longitud de 19-24 nucleótidos. Actúan principalmente regulando la expresión génica y/o la traducción de ARNm. La mayoría de los microARN (miARN) de planta tienen complementariedad perfecta o casi perfecta con sus secuencias diana. Sin embargo, hay dianas naturales con hasta cinco emparejamientos erróneos. Se procesan a partir de ARN no codificantes más largos con estructuras de plegamiento características mediante RNasas específicas bicatenarias de la familia Dicer. Después del procesamiento, se incorporan en el complejo de silenciamiento inducido por ARN (RISC, *RNA-induced silencing complex*) uniéndose a su componente principal, una proteína Argonauta. Los miARN sirven como componentes de especificidad del RISC, debido al emparejamiento con bases de ácidos nucleicos diana, principalmente ARNm, en el citoplasma. Los acontecimientos reguladores posteriores incluyen la escisión de ARNm diana y la destrucción y/o inhibición traduccional. Por tanto, los efectos de la sobreexpresión de miARN a menudo se reflejan en niveles de ARNm disminuidos de los genes diana.

Los microARN artificiales (amiARN), que tienen típicamente 21 nucleótidos de longitud, pueden modificarse por ingeniería genética específicamente para regular negativamente la expresión génica de un solo gen o de múltiples genes de interés. En la técnica se conocen bien los determinantes de la selección diana de microARN de planta. Se han definido los parámetros empíricos para el reconocimiento diana y pueden usarse para ayudar a diseñar los

amiARN específicos, (Schwab y col., Dev. Cell 8, 517-527, 2005). Herramientas convenientes para el diseño y generación de los amiARN y sus precursores también están disponibles para el público (Schwab y col., Plant Cell 18, 1121-1133, 2006).

5 Para un rendimiento óptimo, las técnicas de silenciamiento génico usadas para reducir la expresión en una planta de un gen endógeno requieren el uso de secuencias de ácidos nucleicos de plantas monocotiledóneas para la transformación de plantas monocotiledóneas, y de plantas dicotiledóneas para la transformación de plantas dicotiledóneas. Preferentemente, una secuencia de ácidos nucleicos de cualquier especie de planta determinada se introduce dentro de esta misma especie. Por ejemplo, una secuencia de ácidos nucleicos de arroz se transforma en una planta de arroz. Sin embargo, no es un requisito absoluto que la secuencia de ácidos nucleicos que se introduce se origine de la misma especie de planta como la planta en la que se introducirá. Basta con que haya una homología sustancial entre el gen diana endógeno y el ácido nucleico a introducir.

10 Anteriormente se han descrito ejemplos de diversos procedimientos para la reducción o eliminación sustancial de la expresión en una planta de un gen endógeno. Un experto en la técnica podrá adaptar fácilmente los procedimientos anteriormente mencionados para el silenciamiento para conseguir la reducción de la expresión de un gen endógeno en una planta completa o en partes de la misma usando un promotor apropiado, por ejemplo.

Marcador de selección (gen)/Gen indicador

Un “marcador de selección”, “gen marcador de selección” o “gen indicador” incluye cualquier gen que confiere un fenotipo a una célula en la que se expresa para facilitar la identificación y/o selección de células que se transfectan o transforman con una construcción de ácido nucleico de la invención. 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), espectinomicina 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 β -glucuronidasa, GUS o β -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, por selección (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 procedimiento 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 cotransformación. El procedimiento de cotransformación 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.

Transgénico/Transgén/Recombinante

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

- (a) las secuencias de ácidos nucleicos 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 ácidos nucleicos de acuerdo con la invención, 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 al menos 500 pb, en especial preferentemente al menos 1000 pb, más preferentemente 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.

Por lo tanto, una planta transgénica, para los fines de la invención, se entiende que significa, como se ha indicado anteriormente, que los ácidos nucleicos usados en el procedimiento de la invención no están en su locus natural en el genoma de dicha planta, siendo posible que los ácidos nucleicos se expresen de manera homóloga o heteróloga. Sin embargo, como se ha mencionado, transgénico también significa que, aunque los ácidos nucleicos de acuerdo con la invención o usados en el procedimiento de la invención están en su posición natural en el genoma de una planta, la secuencia se ha modificado con respecto a la secuencia natural y/o que las secuencias reguladoras de las secuencias naturales se han modificado. Por transgénico se entiende preferentemente que significa la expresión de los ácidos nucleicos de acuerdo con la invención en un locus no natural en el genoma, es decir, se produce la expresión homóloga o, preferentemente, heteróloga de los ácidos nucleicos. En el presente documento se mencionan las plantas transgénicas preferidas.

Transformación

El término “introducción” o “transformación” como se denomina en el presente documento, incluye la transferencia de un polinucleótido exógeno en una célula huésped, 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 de la presente invención y a partir del mismo 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 transferencia de genes externos dentro del genoma de una planta se denomina transformación. 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 procedimientos descritos para la transformación y regeneración de plantas de tejidos de plantas o células de plantas pueden utilizarse para la transformación transitoria o estable. Los procedimientos de transformación incluyen el uso de liposomas, electroporación, productos químicos que aumentan la captación de ADN libre, inyección del ADN directamente en la planta, pistola de bombardeo de partículas, transformación usando virus o polen y microproyección. Pueden seleccionarse procedimientos entre, el procedimiento de calcio/poli(etilenglicol) para protoplastos (Krens, F. A. y col., (1982) *Nature* 296, 72-74; Negruțiu 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, incluyendo plantas de cultivo transgénicas, se producen preferentemente mediante transformación mediada por *Agrobacterium*. Un procedimiento de transformación ventajoso es la transformación en la planta. Para esta finalidad, es posible, por ejemplo, permitir que la agrobacteria actúe en las semillas de la planta o inocular las agrobacterias en el meristemo de la planta. Se ha demostrado de un modo particularmente conveniente, de acuerdo con la invención, permitir que una suspensión de agrobacterias transformadas actúe en la planta intacta o al menos en los primordios florales. Posteriormente la planta crece hasta que se obtienen las semillas de la planta tratada (Clough y Bent, *Plant J.* (1998) 16, 735-743). Los procedimientos para la transformación de arroz mediada por *Agrobacterium* incluyen procedimientos bien conocidos para la transformación de arroz, tales como los descritos en cualquiera de las siguientes referencias bibliográficas: solicitud de patente Europea 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 de 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). Dichos procedimientos se describen adicionalmente, a modo de ejemplo, en B. Jenes y col., *Techniques for Gene Transfer*, en: *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds. S. D. Kung y R. Wu, Academic Press (1993) 128-143 y en Potrykus *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991) 205-225). Los ácidos nucleicos o la construcción a expresar se clona preferentemente en un vector, que es adecuado para transformar *Agrobacterium tumefaciens*, por ejemplo, pBin19 (Bevan y col., *Nucl. Acids Res.* 12 (1984) 8711). Las agrobacterias transformadas por dicho vector pueden después usarse de una manera conocida para la transformación de plantas, tales como plantas usadas como un modelo, como *Arabidopsis* (*Arabidopsis thaliana* que se encuentra dentro del ámbito de la presente invención no se considera como una planta de cultivo) o plantas de cultivo tales como, a modo de ejemplo, plantas de tabaco, por ejemplo, sumergiendo hojas laceradas o picadas en una solución agrobacteriana y después cultivarlas en un medio adecuado. La transformación de plantas mediante *Agrobacterium tumefaciens* la describen, por ejemplo, Höfgen y Willmitzer en *Nucl. Acid Res.* (1988) 16, 9877 o se conoce, entre otros, de F. F. White, *Vectors for Gene Transfer in Higher Plants*; in *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds. S. D. Kung y R. Wu, Academic Press, 1993, págs. 15-38.

Además de la transformación de células somáticas, que después deben regenerarse en plantas intactas, también es posible transformar las células de meristemos de planta y en particular aquellas células que se desarrollan en gametos. En este caso, los gametos transformados siguen el desarrollo natural de la planta, dando lugar a plantas transgénicas. Por tanto, por ejemplo, semillas de *Arabidopsis* se tratan con agrobacterias y se obtienen semillas del desarrollo de plantas de las que se transforma una determinada proporción y por tanto transgénicas [Feldman, KA y Marks MD (1987). *Mol Gen Genet* 208: 274-289; Feldmann K (1992). En: C Koncz, N-H Chua y J Shell, eds, *Methods in Arabidopsis Research*. World Scientific, Singapore, pág. 274-289]. Los procedimientos alternativos son a base de la retirada repetida de las inflorescencias e incubación del sitio de escisión en el centro de la roseta con agrobacterias transformadas, por lo que las semillas transformadas pueden obtenerse de modo similar en un punto de tiempo final (Chang (1994). *Plant J.* 5: 551-558; Katavic (1994). *Mol Gen Genet*, 245: 363-370). Sin embargo, un procedimiento especialmente eficaz es el procedimiento de infiltración por vacío con sus modificaciones tal como el procedimiento de "inmersión floral". En el caso de la infiltración por vacío de *Arabidopsis*, las plantas intactas bajo presión reducida se tratan con una suspensión agrobacteriana [Bechthold, N (1993). *C R Acad Sci Paris Life Sci*, 316: 1194-1199], mientras que en el caso del procedimiento de "inmersión floral" el tejido floral en desarrollo se incubaba brevemente con una suspensión agrobacteriana tratada con tensioactivo [Clough, SJ y Bent AF (1998) *The Plant J.* 16, 735-743]. En ambos casos se recoge una determinada proporción de semillas transgénicas y estas semillas pueden diferenciarse de semillas no transgénicas al cultivar bajo las condiciones selectivas descritas anteriormente. Además, la transformación estable de los plástidos es ventajosa ya que estos se heredan por vía materna en la mayor parte de los cultivos reduciendo o eliminando el riesgo de flujo transgénico a través del polen. La transformación del genoma del cloroplasto generalmente se realiza mediante un procedimiento que han presentado esquemáticamente Klaus y col., 2004 [*Nature Biotechnology* 22 (2), 225-229]. En resumen, las secuencias a transformar se clonan junto con un gen marcador de selección entre las secuencias flanqueantes homólogas al genoma del cloroplasto. Estas secuencias flanqueantes homólogas dirigen la integración específica de sitio en el plastoma. La transformación plastidial se ha descrito para muchas especies de plantas diferentes y se ofrece una revisión en Bock (2001) *Transgenic plastids in basic research and plant biotechnology*. *J Mol Biol.* 2001 Sep 21; 312 (3):425-38 o Maliga, P (2003) *Progress towards commercialization of plastid transformation technology*. *Trends Biotechnol.* 21, 20-28. Recientemente se han descrito progresos biotecnológicos adicionales, en forma de transformantes plastidiales sin marcador, que pueden producirse mediante un gen marcador cointegrado transitorio

(Klaus y col., 2004, Nature Biotechnology 22(2), 225-229).

El etiquetado en activación de ADN T

El etiquetado en 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 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 de tal manera que el promotor dirige la expresión del gen diana. Típicamente, la regulación de la expresión del gen diana mediante su promotor natural se interrumpe y el gen cae bajo el control del promotor nuevamente introducido. El promotor se incorpora típicamente en un ADN T. Este ADN T se inserta al azar el genoma de la planta, por ejemplo, a través de infección por *Agrobacterium* y conduce a expresión modificada de genes cerca al ADN T insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debidos a la expresión modificada de los genes cerca al promotor introducido.

TILLING

El término "TILLING" es la abreviatura de "Targeted Induced Local Lesions In Genomes" (Inducción Dirigida de Lesiones Locales en el Genoma) y se refiere a una tecnología de mutagénesis útil para generar y/o identificar ácidos nucleicos que codifican proteínas con expresión y/o actividad modificada. El TILLING también permite la selección de plantas que llevan dichas variantes mutantes. Estas variantes mutantes pueden presentar expresión modificada, en fuerza o en localización o en tiempo (si las mutaciones afectan, por ejemplo, al promotor). Estas variantes mutantes pueden presentar mayor actividad que la mostrada por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de detección de alto rendimiento. Las etapas seguidas típicamente en el TILLING son: (a) mutagénesis EMS (Redei GP y Koncz C (1992) In Methods in *Arabidopsis* Research, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, páginas 16-82; Feldmann y col., (1994) In Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, páginas 137-172; Lightner J y Caspar T (1998) In J Martinez-Zapater, J Salinas, eds, Methods on Molecular Biology, Vol. 82. Humana Press, Totowa, NJ, páginas 91-104); (b) preparación y agrupamiento de ADN en individuos; (c) amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heteroduplex; (e) DHPLC, en la que la presencia de un heteroduplex en un grupo se detecta como un pico extra en el cromatograma; (f) identificación del mutante individual; y (g) secuenciación del producto PCR mutante. Los procedimientos de TILLING son bien conocidos en la técnica (McCallum y col., (2000) Nat Biotechnol 18: 455-457; revisado por Stemple (2004) Nat Rev Genet 5(2): 145-50).

Recombinación homóloga

La recombinación homóloga permite la introducción en un genoma de un ácido nucleico seleccionado en una posición definida seleccionada. La recombinación homóloga es una tecnología convencional utilizada rutinariamente en ciencias biológicas para organismos inferiores tales como levaduras o el musgo *Physcomitrella*. Se han descrito procedimientos para realizar recombinación homóloga en plantas no solo para el modelo de plantas (Offringa 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; lida y Terada (2004) Curr Opin Biotech 15(2): 132-8), y existen estrategias que son generalmente aplicables independientemente del organismo diana (Miller y col, Nature Biotechnol. 25, 778-785, 2007).

Rendimiento

El término "rendimiento" en general significa un producto medible de valor económico, típicamente relacionado con un cultivo, un área y un periodo de tiempo específicos. Las partes de planta individuales contribuyen directamente al rendimiento a base de su número, tamaño y/o peso, o el rendimiento real es el rendimiento por metro cuadrado de un cultivo y año, que se determina dividiendo la producción total (incluyendo la producción tanto cosechada como valorada) por metro cuadrado plantado. El término "rendimiento" de una planta puede estar relacionado con biomasa vegetativa (biomasa de raíz y/o brote), con órganos reproductores y/o con propágulos (tal como semillas) de esta planta.

Vigor temprano

"Vigor temprano" se refiere a un crecimiento bien equilibrado saludable activo especialmente durante las etapas tempranas del crecimiento de la planta, y puede ser el resultado del aumento de la eficacia biológica de la planta debido a, por ejemplo, plantas que están mejor adaptadas a su ambiente (es decir, optimizando el uso de fuentes de energía y repartiendo entre brotes y raíces). Las plantas que tienen vigor temprano también muestran supervivencia aumentada de plántula y un mejor establecimiento del cultivo, que a menudo produce campos muy uniformes (creciendo el cultivo de una manera uniforme, es decir, alcanzando la mayoría de las plantas las diversas etapas de desarrollo a sustancialmente el mismo tiempo) y a menudo mejor y mayor rendimiento. Por lo tanto, el vigor temprano puede determinarse midiendo diversos factores, tales como, peso de mil granos, porcentaje de germinación, porcentaje de emergencia, crecimiento de plántula, altura de plántula, longitud de raíz, biomasa de brote y raíz y muchos más.

Aumento/Mejora/Potenciación

Los términos “aumento”, “mejora” o “potenciación” son intercambiables y significan en el sentido de la solicitud al menos 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 % o 10 %, preferentemente al menos 15 % o 20 %, más preferentemente 25 %, 30 %, 35 % o 40 % más rendimiento y/o crecimiento en comparación con plantas de control como se define en el presente documento.

Rendimiento de semilla

El rendimiento de semilla aumentado puede manifestarse por sí mismo como uno o más de los siguientes: a) un aumento en biomasa de semilla (peso total de semilla) que puede establecerse sobre una semilla individual y/o por planta y/o por metro cuadrado; b) número de flores por planta aumentado; c) número de semillas (llenas) aumentado; d) tasa de llenado de semilla aumentado (que se expresa como la proporción entre el número de semillas llenas dividido entre el número total de semillas); e) índice de cosecha aumentado, que se expresa como una proporción del rendimiento de partes cosechables, tales como semillas, dividido entre la biomasa total, y f) peso de mil granos (PMG) aumentado, que se extrapola a partir del número de semillas llenas contadas y su peso total. Un PMG aumentado puede ser el resultado de un tamaño y/o peso de semilla aumentado y también puede ser el resultado de un aumento en el tamaño del embrión y/o endospermo.

También puede manifestarse un aumento en el rendimiento de la semilla como un aumento del tamaño de la semilla y/o volumen de la semilla. Además, también puede manifestarse un aumento en el rendimiento de la semilla en sí mismo como un aumento en el área de semilla y/o longitud de semilla y/o anchura de semilla y/o un perímetro de semilla. El rendimiento aumentado puede ser también el resultado de arquitectura modificada, o puede ocurrir debido a la arquitectura modificada.

Índice de verdor

El “índice de verdor” como se usa en el presente documento se calcula a partir de imágenes digitales de plantas. Para cada píxel que pertenece al objeto de la planta sobre la imagen, se calcula la proporción del valor verde frente al valor rojo (en el modelo RGB por el que se codifica el color). El índice de verdor se expresa como el porcentaje de píxeles por el que la proporción con verde con respecto a rojo supera un umbral determinado. En condiciones de crecimiento normales, en condiciones de crecimiento de estrés salino y en condiciones de crecimiento de disponibilidad de nutrientes reducido, el índice de verdor de las plantas se mide en la última imagen antes de la floración. Por otro lado, en condiciones de crecimiento de estrés por sequía, el índice de verdor de las plantas se mide en la primera imagen después de la sequía.

Planta

El término “planta” como se usa en el presente documento incluye plantas completas, antecesoras y progenie de las plantas y partes de las plantas, incluyendo semillas, brotes, tallos, hojas, raíces (incluyendo tubérculos), flores y tejidos y órganos, en el que cada uno de los elementos anteriormente mencionados comprende el gen/ácido nucleico de interés. El término “planta” también incluye células de planta, cultivos en suspensión, tejido de callo, embriones, regiones meristemáticas, gametofitos, esporofitos, polen y microesporas, de nuevo en el que cada uno de los elementos mencionados anteriormente 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 plantas monocotiledóneas y dicotiledóneas incluyendo forraje o leguminosas forrajeras, plantas ornamentales, cultivos alimenticios, árboles o arbustos seleccionados de la lista que comprende *Acer spp.*, *Actinidia spp.*, *Abelmoschus spp.*, *Agave sisalana*, *Agropyron spp.*, *Agrostis stolonifera*, *Allium spp.*, *Amaranthus spp.*, *Ammophila arenaria*, *Ananas comosus*, *Annona spp.*, *Apium graveolens*, *Arachis spp.*, *Artocarpus spp.*, *Asparagus officinalis*, *Avena spp.* (por ejemplo, *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua var. sativa*, *Avena hybrida*), *Averrhoa carambola*, *Bambusa sp.*, *Benincasa hispida*, *Bertholletia excelsa*, *Beta vulgaris*, *Brassica spp.* (por ejemplo, *Brassica napus*, *Brassica rapa ssp.* [canola, colza, nabo]), *Cadaba farinosa*, *Camellia sinensis*, *Canna indica*, *Cannabis sativa*, *Capsicum spp.*, *Carex elata*, *Carica papaya*, *Carissa macrocarpa*, *Carya spp.*, *Carthamus tinctorius*, *Castanea spp.*, *Ceiba pentandra*, *Cichorium endivia*, *Cinnamomum spp.*, *Citrullus lanatus*, *Citrus spp.*, *Cocos spp.*, *Coffea spp.*, *Colocasia esculenta*, *Cola spp.*, *Corchorus sp.*, *Coriandrum sativum*, *Corylus spp.*, *Crataegus spp.*, *Crocus sativus*, *Cucurbita spp.*, *Cucumis spp.*, *Cynara spp.*, *Daucus carota*, *Desmodium spp.*, *Dimocarpus longan*, *Dioscorea spp.*, *Diospyros spp.*, *Echinochloa spp.*, *Elaeis* (por ejemplo, *Elaeis guineensis*, *Elaeis oleifera*), *Eleusine coracana*, *Eragrostis tef*, *Erianthus sp.*, *Eriobotrya japonica*, *Eucalyptus sp.*, *Eugenia uniflora*, *Fagopyrum spp.*, *Fagus spp.*, *Festuca arundinacea*, *Ficus carica*, *Fortunella spp.*, *Fragaria spp.*, *Ginkgo biloba*, *Glycine spp.* (por ejemplo, *Glycine max*, *Soja hispida* o *Soja max*), *Gossypium hirsutum*, *Helianthus spp.* (por ejemplo, *Helianthus annuus*), *Hemerocallis fulva*, *Hibiscus spp.*, *Hordeum spp.* (por ejemplo, *Hordeum vulgare*), *Ipomoea batatas*, *Juglans spp.*, *Lactuca sativa*, *Lathyrus spp.*, *Lens culinaris*, *Linum usitatissimum*, *Litchi chinensis*, *Lotus spp.*, *Luffa acutangula*, *Lupinus spp.*, *Luzula sylvatica*, *Lycopersicon spp.* (por ejemplo, *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon pyriforme*), *Macrotyloma spp.*, *Malus spp.*, *Malpighia emarginata*, *Mammea americana*, *Mangifera indica*, *Manihot spp.*, *Manilkara zapota*, *Medicago sativa*, *Melilotus spp.*, *Mentha spp.*, *Miscanthus sinensis*, *Momordica spp.*, *Morus nigra*, *Musa spp.*, *Nicotiana spp.*, *Olea*

spp., *Opuntia* spp., *Ornithopus* spp., *Oryza* spp. (por ejemplo *Oryza sativa*, *Oryza latifolia*), *Panicum miliaceum*, *Panicum virgatum*, *Passiflora edulis*, *Pastinaca sativa*, *Pennisetum* sp., *Persea* spp., *Petroselinum crispum*, *Phalaris arundinacea*, *Phaseolus* spp., *Phleum pratense*, *Phoenix* spp., *Phragmites australis*, *Physalis* spp., *Pinus* spp., *Pistacia vera*, *Pisum* spp., *Poa* spp., *Populus* spp., *Prosopis* spp., *Prunus* spp., *Psidium* spp., *Punica granatum*, *Pyrus communis*, *Quercus* spp., *Raphanus sativus*, *Rheum rhabarbarum*, *Ribes* spp., *Ricinus communis*, *Rubus* spp., *Saccharum* spp., *Salix* sp., *Sambucus* spp., *Secale cereale*, *Sesamum* spp., *Sinapis* sp., *Solanum* spp. (por ejemplo, *Solanum tuberosum*, *Solanum integrifolium* o *Solanum lycopersicum*), *Sorghum bicolor*, *Spinacia* spp., *Syzygium* spp., *Tagetes* spp., *Tamarindus indica*, *Theobroma cacao*, *Trifolium* spp., *Tripsacum dactyloides*, *Triticosecale rimpai*, *Triticum* spp. (por ejemplo, *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum*, *Triticum monococcum* o *Triticum vulgare*), *Tropaeolum minus*, *Tropaeolum majus*, *Vaccinium* spp., *Vicia* spp., *Vigna* spp., *Viola odorata*, *Vitis* spp., *Zea mays*, *Zizania palustris*, *Ziziphus* spp., entre otras.

Descripción detallada de la invención

Sorprendentemente, ahora se ha descubierto que la expresión aumentada en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD proporciona plantas que tienen rasgos relacionados con rendimiento de semilla potenciados, con respecto a las plantas de control.

De acuerdo con una primera realización, la presente invención proporciona un procedimiento para potenciar en plantas rasgos relacionados con el rendimiento de semilla, con respecto a las plantas de control, que comprende aumentar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD, en el que el polipéptido PHDf-HD comprende:

(i) un dominio que tiene al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más, de identidad de secuencia de aminoácidos con un dominio cremallera de leucina/dedo de homeodominio de planta (ZIP/PHDf), como se representa en SEC ID N°: 233; y

(ii) un dominio que tiene al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más, de identidad de secuencia de aminoácidos con un homeodominio (HD) como se representa en SEC ID N°: 234; y

(iii) en el que dicho polipéptido PHDf-HD, cuando se utiliza en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 13, se agrupa con el grupo de polipéptidos PHDf-HD que comprenden la secuencia de polipéptido como se representa en SEC ID N°: 180, en lugar de con cualquier otro grupo de HD; y/o en el que dicho polipéptido PHDf-HD tiene en orden creciente de preferencia al menos el 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más, de identidad de secuencia con el polipéptido PHDf-HD como se representa en SEC ID N°: 180, o con cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 en el presente documento y, de forma opcional, seleccionar plantas que tengan potenciados rasgos relacionados con el rendimiento.

Un procedimiento preferente para aumentar la expresión de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD es mediante la introducción y expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD.

En lo sucesivo, cualquier referencia en el presente documento a "proteína útil en los procedimientos de la invención" se considera que significa un polipéptido PHDf-HD, como se define en el presente documento. En lo sucesivo, cualquier referencia en el presente documento a una "secuencia de ácido nucleico útil en los procedimientos de la invención" se considera que significa una secuencia de ácido nucleico capaz de codificar tal polipéptido PHDf-HD. La secuencia de ácido nucleico que se introducirá en una planta (y por lo tanto útil en la realización de los procedimientos de la invención) es cualquier secuencia de ácido nucleico que codifique el tipo de polipéptido, que será descrita ahora, en lo sucesivo también llamada "secuencia de ácido nucleico de PHDf-HD" o "gen de PHDf-HD".

Un "polipéptido PHDf-HD" como se define en el presente documento se refiere a cualquier polipéptido que comprende: (i) un dominio que tiene al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más, de identidad de secuencia de aminoácidos con un dominio cremallera de leucina/dedo de homeodominio de planta (ZIP/PHDf) como se representa en SEC ID N°: 233; y (ii) un dominio que tiene al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más, de identidad de secuencia de aminoácidos con un homeodominio (HD) como se representa en SEC ID N°: 234.

De forma alternativa o adicional, un "polipéptido PHDf-HD" como se define en el presente documento, se refiere a cualquier polipéptido que comprende: (i) un dominio PHD como se representa en PFAM00628; y (ii) un HD como se representa en PFAM00046.

De forma adicional, el "polipéptido PHDf-HD" como se define en el presente documento, se refiere a cualquier secuencia de polipéptido, la que cuando se utiliza en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 1, se agrupa con el grupo de polipéptidos PHDf-HD que comprenden la secuencia de

polipéptidos como se representa en SEC ID N°: 180, o con cualquiera de las secuencias de polipéptidos proporcionadas en la Tabla D1, en lugar de con cualquier otro grupo de HD.

De forma alternativa o adicional, un "polipéptido PHDf-HD" como se define en el presente documento, se refiere a cualquier polipéptido que tiene en orden creciente de preferencia al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más, de identidad de secuencia de aminoácidos con el polipéptido PHDf-HD, como se representa en SEC ID N°: 180, o con cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 en el presente documento.

Los términos "dominio" y "motivo" se definen en la sección de "Definiciones" del presente documento. Existen bases de datos especializadas para la identificación de dominios, 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, InterPro (Mulder y col., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher y Bairoch (1994), A generalized profile syntax for biomolecular SECuencias motifs and its function in automatic SECuence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., páginas 53-61, AAAI Press, Menlo Park; Hulo y col., Nucl. Acids. Res. 32:D134-D137, (2004), o Pfam (Bateman y col., Nucleic Acids Research 30(1): 276-280 (2002). En el servidor de proteomics ExpPASy (Swiss Institute of Bioinformatics (Gasteiger y col., ExpPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31: 3784-3788(2003)) se encuentra disponible un conjunto de herramientas para realizar análisis informatizados de secuencias de proteínas. Los dominios también pueden identificarse usando técnicas rutinarias, tales como mediante alineamiento de secuencias. El análisis de la secuencia de polipéptido de SEC ID N°: 180 se presenta a continuación en los Ejemplos 2 y 4 en el presente documento. Por ejemplo, un polipéptido PHDf-HD como se representa en SEC ID N°: 180, comprende un dedo de HDP con entrada Pfam PFA00628, y un HD con entrada Pfam PF00046.

Los procedimientos para el alineamiento de las secuencias para comparación son bien conocidos en la técnica, dichos procedimientos incluyen GAP, BESTFIT, BLAST, FASTA y TFASTA. GAP utiliza el algoritmo de Needleman y Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar el alineamiento global (es decir, extendiéndose por las secuencias completas) de dos secuencias que maximizan el número de emparejamientos y minimiza el número de espacios. El algoritmo BLAST (Altschul y col. (1990) J Mol Biol 215: 403-10) calcula el porcentaje de identidad de secuencia y realiza un análisis estadístico de la similitud entre las dos secuencias. El programa informático para realizar el análisis BLAST está públicamente disponible a través del National Centre for Biotechnology Information (NCBI). Los homólogos pueden identificarse fácilmente usando, por ejemplo, el algoritmo de alineamiento de secuencia múltiple ClustalW (versión 1,83), con los parámetros de alineamiento por pares predeterminados y un procedimiento de clasificación en porcentaje. Los porcentajes globales de similitud e identidad también pueden determinarse utilizando uno de los procedimientos disponibles en el paquete informático MatGAT (Campanella y col., BMC Bioinformatics. Jul 2003 10; 4: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando secuencias de ADN o proteína). Como sería evidente para un experto en la materia, puede realizarse una edición manual menor para optimizar el alineamiento entre los motivos conservados. Adicionalmente, también pueden usarse dominios específicos en lugar de usar secuencias de longitud completa para la identificación de homólogos. Los valores de identidad de secuencia pueden determinarse sobre toda la secuencia de ácido nucleico o secuencia de polipéptido o sobre los dominios seleccionados o uno o más motivos conservados, usando los programas mencionados anteriormente usando los parámetros predeterminados. El Ejemplo 3 en el presente documento describe en la Tabla 2 la identidad porcentual entre ZIP/PHDf como se representa en SEC ID N°: 233 y ZIP/PHDf de los polipéptidos PHDf-HD enumerados en la Tabla D1, y en la Tabla D3 la identidad porcentual entre el HD como se representa en SEC ID N°: 234 y el HD de los polipéptidos enumerados en la Tabla D1 del Ejemplo 1.

Además, las secuencias de aminoácidos enriquecidas en aminoácidos básicos (Lys y Arg) o ácidos (Glu y Asp), llamados respectivamente tramos básicos y ácidos, también pueden identificarse fácilmente simplemente por inspección ocular (Figura 5). Alternativamente, se puede calcular la composición principal de aminoácidos (en %) para determinar si un dominio de polipéptido es rico en aminoácidos específicos, utilizando programas informáticos del servidor ExpPASy, en particular la herramienta ProtParam (Gasteiger E y col. (2003) ExpPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784-3788). Después se puede comparar la composición de la proteína de interés con la composición de aminoácidos promedio (en %) en el banco de datos Swiss-Prot Protein Sequence.

Las súper hélices son importantes para identificar las interacciones proteína-proteína tales como oligomerización, ya sea de proteínas idénticas, de proteínas de la misma familia, o de proteínas no relacionadas. Un polipéptido PHDf-HD presenta al menos una región predicha de súper hélice. Recientemente, se han hecho grandes avances en la predicción computacional de súper hélices a partir de los datos de secuencia. Entre los algoritmos bien conocidos para un experto en la materia están disponibles en las herramientas ExpPASy Proteomics COILS, PAIRCOIL, PAIRCOIL2, MULTICOIL, o MARCOIL, alojado en el Swiss Institute for Bioinformatics. En el Ejemplo 5 y la Figura 4, se muestran respectivamente los resultados numéricos y gráficos de SEC ID N°: 180 producidos mediante el análisis del algoritmo COILS. En una secuencia de polipéptido PHDf-HD como la representada en SEC ID N°: 180 se identifican, con una fuerte probabilidad, dos dominios predichos de súper hélices N-terminales. También se predijo con una menor probabilidad, una súper hélice C-terminal.

La tarea de predicción de localización subcelular de proteínas es importante y está bien estudiada. Conocer la localización de una proteína ayuda a elucidar su función. Los procedimientos experimentales para localización de proteínas varían de inmunolocalización a etiquetado de proteínas utilizando la proteína verde fluorescente (GFP). Dichos procedimientos son precisos aunque laboriosos en comparación con los procedimientos computacionales.

Recientemente se han hecho grandes avances en la predicción computacional de la localización de proteínas a partir de los datos de secuencia. Entre los algoritmos bien conocidos para un experto en la materia, están disponibles en las herramientas ExPASy Proteomics alojadas en el Swiss Institute for Bioinformatics, por ejemplo, PSort, TargetP, ChloroP, LocTree, Predotar, LipoP, MITOPROT, PATS, PTS1, SignalP y otros. La identificación de la localización subcelular del polipéptido de la invención se muestra en el Ejemplo 6. En particular, SEC ID N°: 180 de la presente invención se asigna al compartimento nuclear de células eucariotas.

Además, los polipéptidos PHDf-HD útiles en los procedimientos de la presente invención (al menos en su forma natural) típicamente, pero no necesariamente, tienen actividad regulatoria transcripcional y capacidad para interactuar con otras proteínas. Por lo tanto, los polipéptidos PHDf-HD con actividad regulatoria transcripcional reducida, sin actividad regulatoria transcripcional, con capacidad de interacción proteína-proteína reducida, o sin capacidad de interacción proteína-proteína, pueden ser igualmente útiles en los procedimientos de la presente invención. La actividad de unión a ADN y las interacciones proteína-proteína se pueden determinar fácilmente *in vitro* o *in vivo* utilizando técnicas bien conocidas en la materia (por ejemplo en Current Protocols in Molecular Biology, Volúmenes 1 y 2, Ausubel y col. (1994), Current Protocols). Están disponibles varios ensayos para determinar la actividad de unión a ADN de los polipéptidos PHDf-HD, tales como ensayos de unión a ADN de cambio en la corrida electroforética en gel (o ensayos de retardo en gel; Korfhage y col. (1994) Plant C 6: 695-708), ensayos de unión a ADN *in vitro* (Schindler y col. (1993) Plant J 4(1): 137-150), o activación transcripcional de polipéptidos PHDf-HD en células de levadura, animales y de plantas (Halbach y col. (2000) Nucleic Acid Res 28(18): 3542-3550). Se pueden determinar las secuencias de unión a ADN específicas utilizando la técnica de selección de oligonucleótidos aleatoria (Viola y González (26 mayo 2007) Biochemistry).

La presente invención se ilustra mediante la transformación de plantas con la secuencia de ácidos nucleicos representada por la SEC ID N°: 179, que codifica la secuencia de polipéptidos de la SEC ID N°: 180. Sin embargo, la realización de la invención no está limitada a estas secuencias; pudiendo realizarse los procedimientos de la invención de forma ventajosa usando cualquier secuencia de ácido nucleico que codifique PHDf-HD o el polipéptido PHDf-HD, como se define en el presente documento.

En la Tabla D1 del Ejemplo 1 del presente documento, se proporcionan ejemplos de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD. Dichas secuencias de ácido nucleico son útiles para realizar los procedimientos de la invención. Las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1 son secuencias ejemplo de ortólogos y parálogos del polipéptido PHDf-HD representado en la SEC ID N°: 180, siendo los términos "ortólogos" y "parálogos" como se define en el presente documento. Adicionalmente, los ortólogos y parálogos pueden identificarse fácilmente realizando una búsqueda BLAST denominada recíproca. Típicamente, esto implica un primer BLAST que implica una secuencia de consulta BLASTing (por ejemplo usando cualquiera de las secuencias indicadas en la Tabla D1 del Ejemplo 1) contra cualquier base de datos de secuencia, tal como la base de datos NCBI públicamente disponible. BLASTN o TBLASTX (utilizando valores predeterminados convencionales) pueden usarse cuando se parte de una secuencia de nucleótidos y BLASTP o TBLASTN (usando valores predeterminados convencionales) pueden usarse cuando se parte de una secuencia de proteína. Los resultados BLAST pueden filtrarse opcionalmente. Las secuencias de longitud completa de los resultados filtrados o no filtrados vuelven después a buscarse con BLAST (segundo BLAST) contra las secuencias del organismo del cual procede la secuencia de consulta (en el que la secuencia de consulta es la SEC ID N°: 179 o SEC ID N°: 180, por lo que el segundo BLAST se realizaría contra las secuencias de arroz). Después se comparan los resultados del primer y segundo BLAST. Se identifican un parálogo si un acierto de alta clasificación del primer BLAST es de la misma especie de la que procede la secuencia de consulta, después, un BLAST posterior da como resultado idealmente la secuencia de consulta entre los mayores aciertos; un ortólogo se identifica si un acierto de alta clasificación en el primer BLAST no es de la misma especie de la que procede la secuencia de consulta, y preferentemente los resultados basados en un BLAST posterior en la secuencia de consulta se encuentran entre los mayores aciertos.

Los aciertos de alta clasificación son aquellos que tienen un valor E bajo. Cuanto menor es el valor E, más significativa la clasificación (o en otras palabras menor es la probabilidad de que se encuentre el acierto casualmente). El cálculo del valor E es bien conocido en la técnica. Además de los valores E, las comparaciones también se clasifican mediante el porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de ácido nucleico (o polipéptido) sobre una longitud particular. En el caso de grandes familias, puede usarse ClustalW, seguido de la generación de un árbol de unión del vecino más próximo para ayudar a visualizar el agrupamiento de los genes relacionados y para identificar los ortólogos y parálogos. Cualquier secuencia que se agrupe dentro del grupo que comprende SEC ID N°: 180 (rodeado por un círculo en la Figura 1) debería considerarse que cae dentro de la definición de un polipéptido PHDf-HD mencionada anteriormente, y debería considerarse adecuada para su uso en los procedimientos de la invención.

Las variantes de ácidos nucleicos también pueden ser útiles en la realización práctica de los procedimientos de la invención. Los ejemplos de dichas variantes incluyen secuencias de ácido nucleico que codifican homólogos y derivados de una cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1, siendo

los términos “homólogo” y “derivado” como se definen en el presente documento. También son útiles en los procedimientos de la invención las secuencias de ácido nucleico que codifican homólogos y derivados de ortólogos o parálogos de una cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Los homólogos y derivados útiles en los procedimientos de la presente invención tienen sustancialmente la misma actividad biológica y funcional que la proteína no modificada de la cual proceden.

Las variantes de ácidos nucleicos adicionales útiles en la realización práctica de los procedimientos de la invención incluyen porciones de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD, secuencias de ácido nucleico que hibridan con secuencias de ácido nucleico que codifican polipéptidos PHDf-HD, variantes de corte y empalme de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD, variantes alélicas de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD y variantes de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD obtenidos por combinación de genes. Las expresiones hibridación de secuencias, variante de corte y empalme, variante alélica y combinación de genes son como se describen en el presente documento.

Las secuencias de ácido nucleico que codifican polipéptidos PHDf-HD no tienen por qué ser secuencias de ácido nucleico de longitud completa, dado que la realización de los procedimientos de la invención no se basa en el uso de secuencias de ácido nucleico completas. De acuerdo con la presente invención se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla en plantas, que comprende introducir y expresar en una planta una porción de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla A del Ejemplo 1, o una porción de una secuencia de ácido nucleico que codifica un ortólogo, parálogo y homólogo de cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1.

Una porción de una secuencia de ácido nucleico puede prepararse, por ejemplo, realizando una o más deleciones en la secuencia de ácido nucleico. Las porciones pueden usarse en forma aislada o pueden fusionarse a otras secuencias codificantes (o no codificantes) para producir por ejemplo, una proteína que combine diversas actividades. Cuando se fusiona con otras secuencias codificantes, el polipéptido resultante producido después de la traducción puede ser más grande que el esperado para la porción proteica.

Las porciones útiles en los procedimientos de la invención, codifican un polipéptido PHDf-HD como se define en el presente documento, y tienen sustancialmente la misma actividad biológica que las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Preferentemente, la porción es una porción de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o es una porción de una secuencia de ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Preferentemente la porción es, en orden creciente de preferencia al menos de 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800 nucleótidos consecutivos de longitud, siendo los nucleótidos consecutivos de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o de una secuencia de ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Muy preferentemente, la porción es una porción de la secuencia de ácido nucleico de SEC ID N°: 179. Preferentemente, la porción codifica una secuencia de polipéptido que, cuando se utiliza en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 1, se agrupa con el grupo de polipéptidos PHDf-HD que comprende la secuencia de polipéptido representada en SEC ID N°: 180, en lugar de con cualquier otro grupo HD.

Otra variante de secuencia de ácido nucleico útil en los procedimientos de la invención es una secuencia de ácido nucleico capaz de hibridar, en condiciones de rigurosidad reducida, preferentemente en condiciones rigurosas, con una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define en el presente documento, o con una porción como se define en el presente documento.

De acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla, que comprende introducir y expresar en una planta una secuencia de ácido nucleico capaz de hibridar con una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o que comprende introducir y expresar en una planta una secuencia de ácido nucleico capaz de hibridar con una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de ácidos nucleicos proporcionadas en la Tabla D1 del Ejemplo 1.

Las secuencias de hibridación útiles en los procedimientos de la invención codifican un polipéptido PHDf-HD como se define en el presente documento, y tienen sustancialmente la misma actividad biológica que las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Preferentemente, la secuencia de hibridación es capaz de hibridar con una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o con una porción de cualquiera de estas secuencias, siendo una porción como se define anteriormente, o en la que la secuencia de hibridación es capaz de hibridar con una secuencia de ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Muy preferentemente, la secuencia de hibridación es capaz de hibridar con una secuencia de ácido nucleico como se representa en SEC ID N°: 179 o con una porción del mismo.

Preferentemente, la secuencia de hibridación codifica una secuencia de polipéptido que, cuando se usa en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 4, se agrupa el grupo de polipéptidos

PHDf-HD que comprende la secuencia de polipéptido representada en SEC ID N°: 180, en lugar de con cualquier otro grupo.

5 Otra variante de secuencia de ácido nucleico útil en los procedimientos de la invención es una variante de corte y empalme que codifica un polipéptido PHDf-HD como se define anteriormente en el presente documento, siendo una variante de corte y empalme como se define en el presente documento.

10 De acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla, que comprende introducir y expresar en una planta una variante de corte y empalme de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o una variante de corte y empalme de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1.

15 Las variantes de corte y empalme preferentes son variantes de corte y empalme de una secuencia de ácido nucleico representada en SEC ID N°: 179, o una variante de corte y empalme de una secuencia de ácido nucleico que codifica un ortólogo o parálogo de SEC ID N°: 180. Preferentemente, la secuencia de polipéptido que codifica la variante de corte y empalme, cuando se utiliza en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 1, se agrupa con el grupo de polipéptidos PHDf-HD que comprende la secuencia de polipéptido representada en SEC ID N°: 180 en lugar de con cualquier otro grupo HD.

Otra variante de secuencia de ácido nucleico útil en realizar los procedimientos de la invención es una variante alélica de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define anteriormente en el presente documento, siendo una variante alélica como se define en el presente documento.

20 De acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla, que comprende introducir y expresar en una planta una variante alélica de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o que comprende introducir y expresar en una planta una variante alélica de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1. Las
25 variantes alélicas útiles en los procedimientos de la presente invención tienen sustancialmente la misma actividad biológica que los polipéptidos PHDf-HD de SEC ID N°: 180 y cualquiera de las secuencias de polipéptido ilustradas en la Tabla D1 del Ejemplo 1. Las variantes alélicas existen en la naturaleza, y los procedimientos de la presente invención abarcan el uso de estos alelos naturales. Preferentemente, la variante alélica es una variante alélica de SEC ID N°: 179 o una variante alélica de una secuencia de ácido nucleico que codifica un ortólogo o parálogo de
30 SEC ID N°: 180. Preferentemente, la secuencia de polipéptido que codifica la variante alélica, cuando se usa en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 1, se agrupa con los polipéptidos PHDf-HD que comprenden la secuencia de polipéptido representada en SEC ID N°: 180, antes de con cualquier otro grupo HD.

35 La combinación de genes o evolución dirigida también puede usarse para generar variantes de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD como se define anteriormente; siendo la expresión "combinación de genes" como se define en el presente documento.

40 De acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento en plantas, preferentemente potenciar rasgos relacionados con el rendimiento de semilla, que comprende introducir y expresar en una planta una variante de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla D1 del Ejemplo 1, o que comprende introducir y expresar en una planta una variante de una secuencia de ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 del Ejemplo 1, cuya secuencia de ácido nucleico variante se obtiene mediante combinación de genes.

45 De forma ventajosa, hasta ahora la presente invención proporciona secuencias de ácido nucleico de, y de polipéptido PHDf-HD, desconocidas.

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

50 (i) un ácido nucleico representado en SEC ID N°: 242;
(ii) un ácido nucleico o fragmento del mismo que es complementario a SEC ID N°: 242;
(iii) un ácido nucleico que codifica un polipéptido PHDf-HD que tiene, en orden creciente de preferencia, al menos el 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o más, de identidad de secuencia con la secuencia de aminoácidos proporcionada en SEC ID N°: 243;
(iv) un ácido nucleico capaz de hibridar en condiciones rigurosas con uno cualquiera de los ácidos nucleicos proporcionados anteriormente en (i), (ii) o (iii).

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

- (i) un ácido nucleico representado en SEC ID N°: 242;
- (ii) un ácido nucleico o fragmento del mismo que es complementario a SEC ID N°: 242;
- (iii) un ácido nucleico que codifica un polipéptido PHDf-HD que tiene, en orden creciente de preferencia, al menos el 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o más, de identidad de secuencia con la secuencia de aminoácidos proporcionada en SEC ID N°: 243;
- (iv) un ácido nucleico capaz de hibridar en condiciones rigurosas con uno cualquiera de los ácidos nucleicos proporcionados anteriormente en (i), (ii) o (iii).

Preferentemente, la secuencia de polipéptido que codifica la secuencia de ácido nucleico variante, obtenida mediante combinación de genes, cuando se utiliza en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 1, se agrupa con el grupo de polipéptidos PHDf-HD que comprenden la secuencia de polipéptidos representada en SEC ID N°: 180, en lugar de con cualquier otro grupo HD.

Además, las variantes de secuencias de ácido nucleico se pueden también obtener mediante mutagénesis dirigida. Hay disponibles varios procedimientos para conseguir la mutagénesis dirigida, siendo los más comunes los procedimientos basados en PCR (Current Protocols in Molecular Biology. Wiley Eds.).

Las secuencias de ácido nucleico que codifican los polipéptidos PHDf-HD se pueden obtener a partir de cualquier fuente natural o artificial. La secuencia de ácido nucleico se puede modificar a partir de su forma natural en composición y/o entorno genómico a través de manipulación deliberada por el ser humano. Preferentemente, la secuencia de ácido nucleico que codifica el polipéptido PHDf-HD es de una planta, además, preferentemente de una planta monocotiledónea, más preferentemente de la familia *Poaceae*, muy preferentemente la secuencia de ácido nucleico es de *Oryza sativa*.

La realización de los procedimientos de la invención proporciona plantas que tienen el rendimiento de semilla aumentado con respecto a las plantas de control. Las expresiones “rendimiento” y “rendimiento de semilla” se describen con más detalle en la sección “definiciones” en el presente documento.

En el presente documento la referencia a rasgos potenciados relacionados con el rendimiento se considera que significa un aumento en la biomasa (peso) de una o más partes de una planta, las que pueden incluir partes aéreas (cosechables) y/o partes por debajo del suelo (cosechables). En particular, tales partes cosechables son semillas, y la realización de los procedimientos de la invención da como resultado plantas que tienen rendimiento de semilla aumentado con respecto al rendimiento de semilla de las plantas de control.

Tomando al maíz como ejemplo, un aumento en el rendimiento puede manifestarse como uno o más de los siguientes: el aumento del número de plantas establecidas por hectárea o acre, un aumento en el número de mazorcas por planta, un aumento en el número de hileras, el número de granos por hilera, el peso del grano, el peso de mil granos, el longitud/diámetro de la mazorca, el aumento de la tasa de llenado de semilla (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), entre otros. Tomando el arroz como ejemplo, un incremento en el rendimiento puede manifestarse en sí mismo como un aumento en uno o más de los siguientes: el número de plantas por hectárea o acre, el número de panículas por planta, el número de espigas por panícula, el número de flores (floretes) por panícula (que se expresa como una proporción del número de semillas llenas sobre el número de panículas primarias), el aumento en la tasa de llenado de semilla (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), el aumento en el peso de mil granos, entre otros.

La presente invención proporciona un procedimiento para aumentar el rendimiento de semilla en comparación con las plantas de control, cuyo procedimiento comprende aumentar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define en el presente documento.

Dado que las plantas transgénicas de acuerdo con la presente invención tienen rasgos relacionados con el rendimiento de semilla potenciados, es probable que estas plantas presenten una tasa de crecimiento aumentada (durante al menos parte de su ciclo de vida) con respecto a la tasa de crecimiento de plantas de control en una etapa correspondiente en su ciclo de vida.

La tasa de crecimiento aumentada puede ser específica a una o más partes de una planta (incluyendo semillas), o puede ser sustancialmente en toda la planta. Las plantas que tienen una tasa de crecimiento aumentada pueden tener un ciclo de vida más corto. El ciclo de vida de una planta puede significar el tiempo necesario para crecer a partir de una semilla madura seca hasta la etapa en la que la planta ha producido semillas maduras secas, similar al material de partida. Este ciclo de vida puede estar influenciado por factores tales como el vigor temprano, tasa de crecimiento, índice de verdor, tiempo de floración y velocidad de la maduración de la semilla. El aumento en la tasa de crecimiento puede tener lugar en una o más etapas en el ciclo de vida de una planta o, sustancialmente, durante el ciclo de vida completo de una planta. Una tasa de crecimiento aumentada durante las etapas tempranas en el ciclo de vida de una planta puede reflejar vigor (temprano) potenciado. El aumento en la tasa de crecimiento puede alterar el ciclo de la cosecha de una planta lo que permite que las plantas se siembren más tarde y/o se cosechen más pronto de lo que de otra manera sería posible (puede obtenerse un efecto similar con un tiempo de floración más temprano). Si la tasa de crecimiento se aumenta suficientemente, puede permitir la siembra adicional de

semillas de la misma especie de la planta (por ejemplo siembra y cosecha de plantas de arroz seguido de siembra y cosecha de plantas de arroz adicionales todas dentro de un periodo de crecimiento convencional). De manera similar, si la tasa de crecimiento se aumenta suficientemente, puede permitir la siembra adicional de semillas de diferentes especies de plantas (por ejemplo, la siembra y cosecha de plantas de maíz seguido de, por ejemplo, la siembra y cosecha opcional de plantas de soja, patata o cualquier otra planta adecuada). Los tiempos de cosecha adicionales del mismo rizoma en el caso de algunas plantas de cultivo también pueden ser posibles. Alterar el ciclo de cosecha de una planta puede conducir a un aumento en producción de biomasa anual por acre (debido a un aumento en el número de tiempos (dicho en un año) que se puede cultivar y cosechar cualquier planta particular). Un aumento en la tasa de crecimiento también puede permitir el cultivo de plantas transgénicas en un área geográfica más amplia que sus homólogos de tipo silvestre, debido a que las limitaciones territoriales para cosechar un cultivo a menudo se determinan 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 tasa de crecimiento puede determinarse derivando diversos parámetros de curvas de crecimiento, dichos parámetros pueden ser: 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. La tasa de crecimiento como se define en el presente documento no se entiende que se refiere a floración más temprana.

De acuerdo con una característica preferida de la presente invención, la realización de los procedimientos de la invención proporciona plantas que tienen una tasa de crecimiento aumentada en comparación con plantas de control. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para aumentar la tasa de crecimiento en plantas, cuyo procedimiento comprende aumentar la expresión en una planta una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define en el presente documento.

Los rasgos potenciados relacionados con el rendimiento de semilla se producen ya sea si la planta está en condiciones sin estrés, ya sea si la planta está expuesta a diversos estreses en comparación con las plantas de control crecidas en condiciones comparables. Típicamente, las plantas responden a la exposición al estrés creciendo más lentamente. En condiciones de estrés intenso, la planta puede incluso detener por completo el crecimiento. Por otro lado, estrés moderado se define en el presente documento como cualquier estrés al que se expone la planta, que no da como resultado el cese total del crecimiento de la planta sin la capacidad de reanudar el crecimiento. El estrés leve, en el sentido de la invención, conduce a una reducción en el crecimiento de la planta estresada de menos de 40 %, 35 % o 30 %, preferentemente menos del 25 %, 20 % o 15 %, más preferentemente menos de 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, tratamientos 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 indeseable para la agricultura. El estrés leve es un estrés biótico y/o abiótico (ambiental) común al cual está expuesta la planta. El estrés abiótico puede deberse a sequía o a exceso de agua, a estrés anaeróbico, estrés salino, toxicidad química, estrés oxidativo y a temperaturas altas, frías o heladas. El estrés abiótico puede ser un estrés osmótico causado por estrés hídrico (particularmente debido a sequía), estrés salino, estrés oxidativo y estrés iónico. Típicamente, los estreses bióticos son aquellos estreses provocados por patógenos, tales como bacterias, virus, hongos, nematodos, e insectos. La expresión condiciones "sin estrés", como se usa en el presente documento, son aquellas condiciones ambientales que permiten el crecimiento óptimo de las plantas. Los expertos en la materia conocen las condiciones de suelo y climáticas normales para un emplazamiento dado.

La realización de los procedimientos de la invención proporciona plantas crecidas en condiciones sin estrés o en condiciones de sequía moderada, que tienen rasgos relacionados con el rendimiento de semilla potenciados, con respecto a las plantas de control crecidas en condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla en plantas crecidas en condiciones sin estrés o en condiciones de sequía moderada, cuyo procedimiento comprende modular, preferentemente aumentar, la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD. Las plantas con condiciones de crecimiento óptimas, cultivadas en condiciones sin estrés, producen típicamente, en orden creciente de preferencia, al menos el 90 %, 87 %, 85 %, 83 %, 80 %, 77 % o 75 % de la producción promedio de tal planta en un ambiente dado. La producción promedio se puede calcular basándose en la cosecha y/o temporada en un emplazamiento dado. Los expertos en la materia conocen las producciones de rendimiento promedio de un cultivo.

La realización de los procedimientos de acuerdo con la presente invención da como resultado plantas crecidas en condiciones de estrés abiótico que tienen rasgos relacionados con el rendimiento de semilla potenciados, con respecto a las plantas de control crecidas en condiciones de estrés comparables. Como se informa en Wang y col. (Planta (2003) 218: 1-14), el estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan de forma adversa el crecimiento y la productividad de la planta. Se sabe que la sequía, la salinidad, las temperaturas extremas y el estrés oxidativo están interconectados y pueden inducir daño en el crecimiento y daño celular a través de mecanismos similares. Rabhani y col. (Plant Physiol (2003) 133: 1755-1767) describen un particularmente alto grado de "interferencia" entre estrés por sequía y estrés por alta salinidad. Por ejemplo, la sequía y/o salinización se manifiestan principalmente como estrés osmótico, dando como resultado la alteración de la homeostasis y de la distribución iónica en la célula. El estrés oxidativo, que de forma frecuente

acompaña a alta o baja temperatura, salinidad o estrés por sequía, puede provocar la desnaturalización de proteínas funcionales y estructurales. En consecuencia, estos diversos estreses ambientales a menudo activan rutas de señalización celulares y respuestas celulares semejantes, tales como la producción de proteínas de estrés, la regulación positiva de antioxidantes, la acumulación de solutos compatibles y la detención del crecimiento. Dado que estreses ambientales diversos activan rutas similares, la ejemplificación de la presente invención con estrés por sequía, no debería verse limitada al estrés por sequía, sino más como una detección para indicar la implicación de los polipéptidos PHDf-HD como se define anteriormente, en la potenciación de rasgos relacionados con el rendimiento de semilla, con respecto a las plantas de control crecidas en condiciones de estrés comparables, en estreses abióticos en general.

La expresión “estrés abiótico”, como se define en el presente documento, se entiende que se refiere a uno cualquiera o más de: estrés hídrico (debido a sequía o a exceso de agua), estrés anaeróbico, estrés salino, estrés térmico (debido a temperaturas altas, 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 salino no se restringe a sal común (NaCl), sino que puede ser cualquier estrés provocado por uno o más de: NaCl, KCl, LiCl, MgCl₂, CaCl₂, entre otros.

La realización de los procedimientos de la invención da lugar a plantas que tienen rasgos relacionados con el rendimiento de semilla potenciados, en condiciones de estrés abiótico con respecto a plantas de control que crecen en condiciones de estrés comparables. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semillas, en plantas que crecen en condiciones de estrés abiótico, cuyo procedimiento comprende aumentar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD. De acuerdo con un aspecto de la invención, el estrés abiótico es un estrés osmótico, seleccionado entre uno a más de los siguientes: estrés hídrico, estrés salino, estrés oxidativo y estrés iónico.

Otro ejemplo de estrés ambiental abiótico, es la disponibilidad reducida de uno o más de los nutrientes que las plantas necesitan asimilar para el crecimiento y desarrollo. Se vierte una enorme cantidad de fertilizantes en los campos para optimizar el crecimiento de la planta y la calidad, debido a la fuerte influencia de la eficacia en la utilización de la nutrición en el rendimiento de la planta y la calidad del producto. La productividad de las plantas normalmente se limita por tres nutrientes principales, fósforo, potasio y nitrógeno, el que, entre estos tres, usualmente es el elemento limitante de la tasa en el crecimiento de la planta. Por lo tanto, el principal elemento nutricional necesario para el crecimiento de la planta es el nitrógeno (N). Es el constituyente de numerosos compuestos importantes encontrados en las células vivas, incluyendo aminoácidos, proteínas, (enzimas), ácidos nucleicos, y clorofila. Del 1,5 % al 2 % de la materia seca vegetal es nitrógeno, y es aproximadamente el 16 % de la proteína total de la planta. Por lo tanto, la disponibilidad de nitrógeno es un factor limitante principal del crecimiento y la producción de plantas de cultivo (Frink y col. (1999) Proc Natl Acad Sci USA 96(4): 1175-1180), y tiene también un impacto principal en la acumulación de proteína y en la composición de aminoácidos. Por lo tanto, son de gran interés las plantas de cultivo con rasgos relacionados con el rendimiento potenciados, preferentemente rasgos relacionados con el rendimiento de semilla potenciados, cuando crecen en condiciones limitantes de nitrógeno.

La realización de los procedimientos de la invención da lugar a plantas que crecen en condiciones de déficit de nutrientes, particularmente en condiciones de déficit de nitrógeno, con rasgos relacionados con el rendimiento de semilla potenciados, con respecto a plantas de control que se cultivan en condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para potenciar rasgos relacionados con el rendimiento de semilla en plantas que se cultivan en condiciones de déficit de nutrientes, cuyo procedimiento comprende aumentar la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD. El déficit de nutrientes puede producirse por una carencia o exceso de nutrientes tales como nitrógeno, fosfatos y otros compuestos que contienen fósforo, potasio, calcio, cadmio, magnesio, manganeso, hierro y boro, entre otros.

La presente invención engloba a plantas o partes de las mismas (incluyendo semillas), o células de las mismas, obtenibles por los procedimientos de acuerdo con la presente invención. Las plantas o partes de las mismas o células de las mismas, comprenden un transgén de ácido nucleico que codifica un polipéptido PHDf-HD, como se define anteriormente.

La invención también proporciona construcciones génicas y vectores para facilitar la introducción y/o expresión aumentada en plantas de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD. Las construcciones génicas pueden insertarse en vectores, que pueden estar disponibles en el comercio, adecuados para la transformación en plantas y para la expresión del gen de interés en las células transformadas. La invención también proporciona el uso de una construcción génica como se define en el presente documento en los procedimientos de la invención.

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

- (a) un secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se ha definido anteriormente;

- (b) una o más secuencias de control capaces de aumentar, la expresión de la secuencia de ácidos nucleicos de
- (a) que comprende al menos un promotor GOS2; y opcionalmente
- (c) una secuencia de terminación de la transcripción.

5 Preferentemente, la secuencia de ácido nucleico que codifica un polipéptido PHDf-HD es como se define anteriormente. La expresión "secuencia de control" y "secuencia de terminación" son como se define en el presente documento.

Una de las secuencias de control de la construcción, de acuerdo con la invención, es un promotor constitutivo GOS2 aislados de un genoma de planta. Preferentemente, los promotores GOS2 de arroz, más preferentemente un promotor GOS2 como se representa en SEC ID N°: 235.

10 Las plantas se transforman con un vector que comprende cualquiera de las secuencias de ácido nucleico descritas anteriormente. El experto en la técnica conocerá los elementos genéticos que deben estar presentes en un vector para transformar, seleccionar y propagar satisfactoriamente células hospedadoras que contengan la secuencia de interés. La secuencia de interés está unida operativamente a una o más secuencias de control (al menos a un promotor).

15 De forma ventajosa, puede usarse cualquier tipo de promotor, ya sea natural o sintético, para modular, preferentemente aumentar, la expresión de la secuencia del ácido nucleico. En los procedimientos es particularmente útil un promotor constitutivo, preferentemente un promotor constitutivo aislado de un genoma de planta. El promotor constitutivo de planta dirige la expresión de una secuencia codificante a un nivel que, en todos los casos, está por debajo del obtenido bajo el control de un promotor 35S del CaMV.

20 Son útiles en la realización de los procedimientos de la invención, otros promotores específicos de órgano, por ejemplo para la expresión preferente en hojas, tallos, tubérculos, meristemas, semillas (embrión y/o endospermo). Para definiciones de los diversos tipos de promotores véase la sección "Definiciones" en el presente documento.

25 Debe aclararse que la aplicabilidad de la presente invención no está limitada a la secuencia de ácido nucleico que codifica el polipéptido PHDf-HD representado por la SEC ID N°: 179, ni la aplicabilidad de la invención está limitada a la expresión de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD cuando se dirige por un promotor constitutivo.

30 De forma opcional, en la construcción introducida en una planta se pueden usar una o más secuencias terminadoras. Los elementos reguladores adicionales pueden incluir potenciadores transcripcionales así como traduccionales. Los expertos en la técnica serán conscientes de que pueden ser adecuadas sustancias terminadoras y potenciadoras para su uso en la realización de la invención. También puede añadirse una secuencia intrónica en la región no traducida (UTR) 5' o en la secuencia codificante para aumentar la cantidad del mensaje maduro que se acumula en el citosol, como se describe en la sección de definiciones. Otras secuencias de control (además de las secuencias promotoras, potenciadoras, silenciadoras, intrónicas, regiones UTR3' y/o UTR5') pueden ser elementos estabilizadores de proteína y/o ARN. Un experto en la materia conocería tales secuencias o podría obtenerlas fácilmente.

35 Las construcciones genéticas de la invención 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.

40 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 estas secuencias de ácido nucleico, es ventajoso usar genes marcadores (o genes indicadores). Por lo tanto, la construcción genética puede comprender opcionalmente un gen marcador de selección. Los marcadores de selección se describen con más detalle en la sección "definiciones" del presente documento.

45 Se sabe que después de la integración estable o transitoria de secuencias de ácido nucleico en células de planta, solamente una minoría de las células recoge el ADN extraño y, si se desea, lo integran en sus genomas, dependiendo del vector de expresión usado y de la técnica de transfección usada. Para identificar y seleccionar estos "integrantes", junto con el gen de interés normalmente se introduce en las células hospedadoras, un gen que codifica un marcador de selección (tal como los descritos anteriormente). Estos marcadores pueden usarse, por ejemplo, en mutantes en los que estos genes no son funcionales, mediante, por ejemplo, delección por procedimientos convencionales. Además, pueden introducirse en la célula hospedadora, moléculas de secuencia de ácido nucleico que codifican un marcador de selección, en el mismo vector que comprende la secuencia que codifica los polipéptidos de la invención, o usados en los procedimientos de la invención, o bien en un vector separado. Se pueden identificar las células transfectadas de forma estable con la secuencia de ácido nucleico introducida mediante, por ejemplo, selección (por ejemplo, las células que han integrado el marcador de selección sobreviven, mientras que las otras células mueren). Una vez que ya no son necesarios, los genes marcadores se pueden eliminar o remover de la célula transgénica. Las técnicas para la eliminación de genes marcadores se conocen en la

técnica; anteriormente, en la sección de definiciones, se describen técnicas útiles.

5 La invención también proporciona un procedimiento para la producción de plantas transgénicas que tienen rasgos relacionados con el rendimiento de semillas potenciados, con respecto a las plantas de control, que comprende la introducción y expresión en una planta de cualquier secuencia de ácido nucleico que codifica un polipéptido PHDf-HD, como se define anteriormente en el presente documento.

Más específicamente, la presente invención proporciona un procedimiento para la producción de plantas transgénicas que tienen rasgos relacionados con el rendimiento de semilla potenciados, con respecto a las plantas de control, cuyo procedimiento comprende:

- 10 (i) introducir y expresar en una planta, parte de planta, o célula de planta, una secuencia de ácido nucleico que codifica un polipéptido de PHDf-HD, bajo el control de un promotor constitutivo de planta; y
 (ii) cultivar la célula de planta en condiciones que promueven el crecimiento y el desarrollo de la planta.

La secuencia de ácido nucleico de (i) puede ser cualquiera de las secuencias de ácido nucleico capaces de codificar un polipéptido PHDf-HD, como se define en el presente documento.

15 La secuencia de ácido nucleico puede introducirse directamente en una célula de planta o en la propia planta (incluyendo la introducción en un tejido, órgano o en cualquier otra parte de una planta). De acuerdo con una característica preferida de la presente invención, la secuencia de ácido nucleico se introduce preferentemente en una planta por transformación. El término "transformación" se describe con más detalle en la sección "definiciones" del presente documento.

20 Las células de planta modificadas genéticamente pueden regenerarse mediante todos los procedimientos con los que está familiarizado el experto en la técnica. Pueden encontrarse procedimientos adecuados en las publicaciones anteriormente mencionadas de S. D. Kung y R. Wu, Potrykus o Höfgen y Willmitzer.

25 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 cotransferidos con el gen de interés, después de lo cual el material transformado se regenera en una planta completa. Para seleccionar plantas transformadas, el material de planta obtenido en la transformación, como norma, se somete a condiciones selectivas de tal manera que las plantas transformadas pueden diferenciarse de las plantas no transformadas. Por ejemplo, las señales obtenidas de la manera descrita anteriormente pueden plantarse y, después de un periodo de crecimiento inicial, someterse a una selección adecuada por pulverización. Una posibilidad adicional consiste en cultivar las semillas, si fuera apropiado, después de esterilización, en placas de agar usando un agente de selección adecuado de tal manera que solamente las semillas transformadas puedan desarrollarse en plantas. Como alternativa, las plantas transformadas se exploran para determinar la presencia de un marcador de selección tal como el descrito anteriormente.

30

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

40 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 planta transformada puede reproducirse asexualmente y puede seleccionarse una segunda generación homocigota (o T2) 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).

45

La presente invención se amplía claramente a cualquier célula de planta o planta producida mediante cualquiera de los procedimientos de la divulgación 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 las mismas características, o características, genotípicas y/o fenotípicas que las producidas por el precursor en los procedimientos de acuerdo con la invención.

50

La invención también incluye células huésped que contienen una secuencia de ácido nucleico aislada que codifica un polipéptido PHDf-HD como se ha definido anteriormente en el presente documento, unido operativamente a un promotor constitutivo de planta. De acuerdo con la invención, las células huésped preferentes son células de plantas. Las plantas huésped para las secuencias de ácido nucleico o el vector usados en el procedimiento de acuerdo con la invención, el casete de expresión o construcción o vector son, en principio, ventajosamente todas las

55

plantas, que pueden sintetizar los polipéptidos usados en el procedimiento inventivo.

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

- 5
- 10 La invención también se amplía a partes cosechables de una planta, que comprenden una secuencia de ácido nucleico aislada que codifica un PHDf-HD (como se define anteriormente en el presente documento) ligada operativamente a un promotor constitutivo de planta, tal como, pero sin limitación, de semillas, hojas, frutos, flores, tallos, rizomas, tubérculos y bulbos. También se desvelan productos derivados, preferentemente derivados directamente, de una parte cosechable de tal planta, tal como gránulos o polvos desecados, aceite, grasa o ácidos grasos, almidón o proteínas.
- 15

De acuerdo con la invención, la expresión es aumentada. Los procedimientos para aumentar la expresión de secuencias de ácidos nucleicos o genes, o productos génicos, están bien documentados en la técnica y en la sección definiciones se proporcionan ejemplos.

- 20 Como se menciona anteriormente, un procedimiento preferente para aumentar la expresión de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD es introduciendo y expresando en una planta una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD; sin embargo, también pueden obtenerse los efectos de la realización del procedimiento, es decir potenciar rasgos relacionados con el rendimiento de semilla, usando otras técnicas bien conocidas, incluyendo, pero sin limitación, etiquetado en activación de ADN T, TILLING, recombinación homóloga. Una descripción de estas técnicas se proporciona en la sección definiciones.

- 25 La presente invención también engloba el uso de secuencias de ácido nucleico que codifican polipéptidos PHDf-HD como se describe en el presente documento y el uso de estos polipéptidos PHDf-HD en la potenciación de cualquiera de los anteriormente mencionados rasgos relacionados con rendimiento de semilla en plantas.

- 30 Las secuencias de ácidos nucleicos que codifican el polipéptido PHDf-HD descrito en el presente documento, o los propios polipéptidos PHDf-HD, pueden encontrar uso en programas de reproducción en los que se identifica un marcador de ADN, que puede estar unido genéticamente a un gen que codifica un polipéptido PHDf-HD. Las secuencias de genes/ ácido nucleico, o los propios polipéptidos PHDf-HD 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 relacionados con el rendimiento potenciados, preferentemente rasgos relacionados con el rendimiento de semilla potenciados, como se ha definido en el presente documento anteriormente en los procedimientos de la invención.
- 35

- 40 Las variantes alélicas de una secuencia de gen/ ácido nucleico que codifica un polipéptido PHDf-HD también 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 proporcionan rasgos relacionados con el rendimiento potenciados. La selección se realiza típicamente supervisando el rendimiento del crecimiento de las plantas que contienen diferentes variantes alélicas de la secuencia en cuestión. El rendimiento de crecimiento puede supervisarse en un invernadero o en el campo.
- 45 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.

- 50 Las secuencias de ácido nucleico que codifican los polipéptidos PHDf-HD 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 secuencias de ácido nucleico que codifican un polipéptido PHDf-HD, requiere solo una secuencia de ácidos nucleicos de al menos 15 nucleótidos de longitud. Las secuencias de ácido nucleico que codifican un polipéptido PHDf-HD pueden usarse como marcadores de polimorfismos de longitud de fragmentos de restricción (PLFR). Con las secuencias de ácidos nucleicos que codifican un polipéptido PHDf-HD 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, las secuencias de ácido nucleico pueden usarse para explorar transferencias de Southern que contienen ADN genómicos tratados con endonucleasas de restricción de un conjunto
- 55

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 de la secuencia de ácido nucleico que codifica un polipéptido PHDf-HD en el mapa genético previamente obtenido usando esta población (Botstein y col. (1980) Am. J. Hum. Genet. 32:314-331).

5 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.

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

15 En otra realización, las sondas de secuencias de ácido nucleico pueden también usarse 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 HISF 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 HISF usando sondas más cortas.

20 Usando las secuencias de ácido nucleico, pueden realizarse diversos procedimientos basados en amplificación de secuencias 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 sequence 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 sequence Res. 17: 6795-6807). Para estos procedimientos, se utiliza la secuencia de una secuencia de ácido nucleico para diseñar y producir pares de cebadores para su uso en la reacción de amplificación o en reacciones de extensión con cebador. El diseño de dichos cebadores es 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 ácidos nucleicos instantánea. Sin embargo, generalmente esto no es necesario para procedimientos de mapeo.

35 Los procedimientos de acuerdo con la presente invención dan como resultado plantas que tienen rasgos potenciados relacionados con el rendimiento de semilla, como se ha descrito anteriormente en el presente documento. Estos rasgos también pueden combinarse con otros rasgos económicamente ventajosos, 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:

40 I. Polipéptido dedo de homeodominio de planta-homeodominio (PHDf-HD)

45 **La Figura 1** representa un árbol *neighbour-joining* construido después de un alineamiento de todos los factores de transcripción que pertenecen a la familia HD (descargados de la base de datos TFDB de arroz alojada en el servidor de la University of Potsdam) y de todas las secuencias de polipéptido de la Tabla D1 (cuando son de longitud completa), realizado utilizando el algoritmo Clustal (1.83) de alineamiento progresivo, utilizando valores predeterminados. El grupo de interés, que comprende a los dos parálogos de arroz (SEC ID N°: 180 o Os02g05450.1, y SEC ID N°: 202 o Os06g12400.1), se ha señalado con un círculo. Cualquier polipéptido que cae dentro de este grupo HD (después de una nueva etapa de alineamiento múltiple, como se describe anteriormente en el presente documento) se considera útil en la realización de los procedimientos de la invención, como se describe en el presente documento.

50 **La Figura 2** representa un dibujo de un polipéptido PHDf-HD, como se representa en SEC ID N°: 180, que comprende una o más de las siguientes características: una señal predicha de localización nuclear (SLN), una cremallera de leucina (ZIP), un dedo PHD (PHDf, Pfam00628), un tramo ácido, dos tramos básicos, un homeodominio (HD, Pfam00046).

55 **La Figura 3** muestra el logo de la secuencia del homeodominio (HD) de los polipéptidos PHDf-HD de la Tabla D1, en el que la altura global de la pila indica la conservación de secuencia en esa posición, mientras que la altura de los símbolos dentro de la pila indica la frecuencia relativa de cada aminoácido o ácido nucleico en esa posición. El HD como se representa en SEC ID N°: 234, y comprendido en SEC ID N°: 180, está en conformidad

con el logo de secuencia como se representa en esta figura.

La Figura 4 muestra el resultado gráfico del algoritmo COILS que predice dos dominios súper hélice en la mitad N-terminal del polipéptido como se representa en SEC ID N°: 180. El eje X representa las coordenadas de los restos de aminoácido, el eje Y la probabilidad (que varía de 0 a 1) de que un dominio súper hélice esté presente, y las tres líneas, las tres ventanas (14, 21,28) examinadas.

La Figura 5 muestra un alineamiento de secuencia múltiple CLUSTAL W (1;83) de los polipéptidos PHDf-HD de la Tabla D1 (cuando son de longitud completa), en el que se identifican varias características. Del extremo N al extremo C de los polipéptidos son: (i) una señal predicha de localización nuclear (SLN); (ii) una cremallera de leucina (ZIP), con cuatro héptadas (recuadradas, en las que normalmente una leucina (ocasionalmente una isoleucina, una valina, o una metionina) aparece cada siete aminoácidos); (iii) un dedo PHD (PHDf), con la típica C4HC3 (cuatro cisteínas, una histidina, tres cisteínas) con una cisteína característica espaciando; (iv) un tramo ácido (rico en aminoácidos ácidos D y E); (v) tramos básicos (ricos en los aminoácidos básicos K y R); (vi) un homeodominio (HD).

La Figura 6 muestra el vector binario para la expresión modulada, preferentemente aumentada, en *Oryza sativa*, de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD bajo en control de un promotor GOS2 de arroz (pGOS2).

La Figura 7 detalla ejemplos de secuencias de las PHDf-HD útiles en la realización de los procedimientos de acuerdo con la presente invención.

Ejemplos

La presente invención se describirá ahora con referencia a los siguientes ejemplos, que son solo a modo ilustrativo. Los siguientes ejemplos no pretenden definir completamente o de otra manera limitar el alcance de la invención.

Manipulación de ADN: a menos que se indique otra cosa, se realizarán técnicas de ADN recombinante de acuerdo con los protocolos convencionales descritos en (Sambrook (2001) *Molecular Cloning: a laboratory manual*, 3ª Edición Cold Spring Harbor Laboratory Press, CSH, Nueva York) o en los Volúmenes 1 y 2 de Ausubel y col. (1994), *Current Protocols in Molecular Biology*, Current Protocols. Materiales y procedimientos convencionales de trabajos moleculares en plantas se describen en *Plant Molecular Biology Labfax* (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (RU) y Blackwell Scientific Publications (RU).

Ejemplo 1: Identificación de secuencias relacionadas con la secuencia de ácido nucleico usada en los procedimientos de la invención

Las secuencias (ADNc de longitud completa, EST o genómicas) relacionadas con la secuencia de ácidos nucleicos usadas en los procedimientos de la presente invención se identificaron entre las conservadas en la base de datos de Nucleótidos Entrez en el National Center for Biotechnology Information (NCBI) usando herramientas de búsqueda de secuencias de bases de datos, tales como la Herramienta Básica de Alineamiento Local (BLAST, *Basic Local Alignment Tool*) (Altschul y col. (1990) *J. Mol. Biol.* 215:403-410; and Altschul y col. (1997) *Nucleic Acids Res.* 25:3389-3402). La búsqueda de secuencias también se realizó en otras bases de datos públicas y de patente. Se usaron herramientas de búsqueda de secuencia tales como la Herramienta de Alineamiento Local Básica (BLAST, *Basic Local Alignment Tool*) (Altschul y col. (1990) *J. Mol. Biol.* 215: 403-410; y Altschul y col. (1997) *Nucleic Acids Res.* 25: 3389-3402). El programa se usa para encontrar regiones de similitud local entre las secuencias comparando la secuencia de ácido nucleico o secuencias de polipéptido con las bases de datos de secuencia y calculando el significado estadístico de los emparejamientos. Por ejemplo, el polipéptido que codifica la secuencia de ácido nucleico de la presente invención se usó para el algoritmo TBLASTN, con configuraciones predeterminadas y el filtro para ignorar la activación de secuencias de baja complejidad. El resultado de los análisis se observó por comparación por pares, y se clasificó de acuerdo con la puntuación de probabilidad (valor E), en el que la puntuación refleja la probabilidad de que se produzca un alineamiento particular por casualidad (cuanto menor sea el valor E más significativo es el acierto). Además de los valores E, las comparaciones también se puntuaron por porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de secuencia de ácido nucleico (o de polipéptido) comparadas sobre una longitud particular. En algunos casos, los parámetros predeterminados se pueden ajustar para modificar la rigurosidad de la búsqueda. Por ejemplo, el valor E se puede incrementar para mostrar emparejamientos menos rigurosos. De esta forma, se pueden identificar emparejamientos cortos casi exactos.

La Tabla D1 proporciona una enumeración de secuencias de ácido nucleico relacionadas con la secuencia de ácido nucleico de PHDf-HD utilizada en los procedimientos de la presente invención.

Tabla D1: Ejemplos de polipéptidos PHDf-HD:

ES 2 558 745 T3

Nombre	Organismo fuente	SEC ID N°: del ácido nucleico	SEC ID N°: del polipéptido	N.º de acceso de la base de datos	Estado
Orysa_PHDf_HD	<i>Oryza sativa</i>	179	180	Os02g05450.1 NM_001052422	LC
Arath_PHDf_HD_PRHA	<i>Arabidopsis thaliana</i>	181	182	At4g029940 NM_119140	LC
Eucgr_PHDf_HD	<i>Eucalyptus grandis</i>	183	184	ADW17964	LC
Medtr_PHDf_HD	<i>Medicago Truncatula</i>	185	186	AC 123547	FL
Pinra_PHDf_HD	<i>Pinus radiata</i>	187	188	ADW18458	FL
Poptr_PHDf_HD	<i>Populus tremuloides</i>	189	190	scaff_VI.625	FL
Sacof_PHDf_HD	<i>Saccharum officinarum</i>	191	192	CA157855.1, CA261734.1, CA253314.1, CA220753.1, CA201958.1	FL
Vitvi_PHDf_HD	<i>Vitis vinifera</i>	193	194	AM477372.2, AM488059.1	FL
Zeama_PHDf_HD	<i>Zea mays</i>	195	196	EE162310, DN204182, CF057937	FL
Arath_PHDf_HD_HAT3.1	<i>Arabidopsis thaliana</i>	197	198	AT3G19510 NM_112838	LC
Lotja_PHDf_HD	<i>Lotusjaponicus</i>	199	200	AP006117.1	LC
Orysa_PHDf_HD_HAZ1	<i>Oryza sativa</i>	201	202	AB081340 Os06g12400.1	LC
Petcr_PHDf_HD_PRHP	<i>Petroselinum crispum</i>	203	204	L21975	LC
Zeama_PHDf_HD_HOX1a	<i>Zea mays</i>	205	206	X67561	LC
Zeama_PHDf_HD_HOX1b	<i>Zea mays</i>	207	208	X92428	LC
Zeama_PHDf_HD_HOX2a	<i>Zea mays</i>	209	210	X89760.1	LC
Zeama_PHDf_HD_HOX2b	<i>Zea mays</i>	211	212	X89761	LC
Vitvi_PHDf_HD_II	<i>Vitis vinifera</i>	213	214	AM464161.2 AM478203.2	LC
Poptr_PHDf_HD_II	<i>Populus tremuloides</i>	215	216	scaffJX.730	LC
Poptr_PHDf_HD_III	<i>Populus tremuloides</i>	217	218	LG_I002624	LC
Poptr_PHDf_HD_IV	<i>Populus tremuloides</i>	219	220	LG_XVNI1192	LC

(continuación)

Nombre	Organismo fuente	SEC ID N°: del ácido nucleico	SEC ID N°: del polipéptido	N.º de acceso de la base de datos	Estado
Ostta_PHDf_HD	<i>Ostreococcus tauri</i>	221	222	CR954214.4	LC
Aqufor_PHDf_HD partial	<i>Aquilegia formosa</i> x <i>Aquilegia pubescens</i>	223	224	DR914726.1, DR941696.1, DR943570.1	Parcial
Glyma_PHDf_HD	<i>Glycine max</i>	225	226	Cóntigo GM06LC25006	Parcial
Lotco_PHDf_HD	<i>Lotus corniculatus</i>	227	228	AP004517	Parcial
Sorpr_PHDf_HD 3'	<i>Sorghum propinquum</i>	229	230	BF656332	Parcial
Sorpr_PHDf_HD 5'	<i>Sorghum propinquum</i>	231	232	BF704605	Parcial
Phypa_PHDf_HD	<i>Physcomitrella patens</i>	238	239	XM_001762483	LC
Phypa_PHDf_HD	<i>Physcomitrella patens</i>	240	241	XM_001779822	LC
Zeama_PHDf_HD	<i>Zea mays</i>	242	243		LC

Ejemplo 2: Alineamiento de las secuencias del polipéptido PHDf-HD

La University of Potsdam, Alemania, ha creado bases de datos de factores de transcripción de planta, incluso para *Oryza sativa*, llamada riceTFDB. Se han descargado todas las secuencias de polipéptido correspondientes a los factores de transcripción que pertenecen a la familia HD (120 modelos de genes (91 locus) identificados hasta ahora), incluyendo los dos polipéptidos PHDf-HD (Os02g05450.1 y Os06g12400.1) identificados hasta la fecha. Las secuencias de polipéptido de la Tabla D1 de la presente solicitud se añadieron (cuando son de longitud completa, es decir 21 secuencias de polipéptido) al conjunto de la familia HD.

El alineamiento de todas las secuencias de polipéptido se realizó con el algoritmo Clustal (1.83) de alineamiento progresivo, utilizando valores predeterminados (Thompson y col. (1997) Nucleic Acids Res 25:4876-4882; Chenna y col. (2003). Nucleic Acids Res 31:3497-3500). Posteriormente se construyó un árbol *neighbour-joining*, y se representa en la Figura 1 de la presente solicitud. El grupo de interés, que comprende los dos parálogos de arroz (Os02g05450.1 y Os06g12400.1) se han señalado con un círculo. Cualquier polipéptido que cae dentro de este grupo HD (después de una nueva etapa de alineamiento múltiple, como se describe anteriormente en el presente documento) se considera útil en la realización de los procedimientos de la invención, como se describe en el presente documento.

En un alineamiento de secuencia múltiple de los polipéptidos PHDf-HD de longitud completa de la Tabla D1, se pueden identificar varias características, y están marcadas en la Figura 5. Del extremo N al extremo C de los polipéptidos son: (i) una señal predicha de localización nuclear (SLN); (ii) una cremallera de leucina (ZIP), con cuatro héptadas (recuadradas, en las que normalmente aparece una leucina (ocasionalmente una isoleucina, una valina, o una metionina)) cada siete aminoácidos; (iii) un dedo HD (PHDf), con la típica C4HC3 (cuatro cisteínas, una histidina, tres cisteínas) con una de cisteína característica espaciando; (iv) un tramo ácido (rico en los aminoácidos ácidos D y E); (v) tramos básicos (ricos en los aminoácidos básicos K y R); (vi) un homeodominio (HD).

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

Los porcentajes de similitud e identidad globales entre las secuencias de polipéptidos de longitud completa útiles en la realización de los procedimientos de la invención se determinaron usando uno de los procedimientos disponibles en la técnica, el programa informático MatGAT (Matrix Global Alignment Tool) (BMC Bioinformatics. 2003 4: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando secuencias de proteína o de ADN. Campanella JJ, Bitincka L, Smalley J; programa informático presentado por Ledion Bitincka). El programa informático MatGAT genera matrices de similitud/identidad para secuencias de ADN o proteína sin la necesidad de alinear previamente los datos. El programa realiza una serie de alineamientos por pares usando el algoritmo de

alineamiento global de Myers y Miller (con una penalización por apertura de hueco de 12, y una penalización por extensión de hueco de 2), calcula la similitud e identidad usando, por ejemplo, Blosum 62 (para polipéptidos), y después se pusieron 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.

5 Los parámetros utilizados en la comparación fueron:

Matriz de puntuación: Blosum62

Hueco inicial: 12

Extensión del hueco: 2

10 Los resultados del análisis del programa informático para la similitud e identidad totales a lo largo de la longitud completa de las secuencias de polipéptido (excluyendo las secuencias de polipéptido parciales) se muestran en la Tabla D2. La identidad porcentual se proporciona sobre la diagonal y la similitud porcentual se proporciona debajo de la diagonal.

15 El porcentaje de identidad entre las secuencias de polipéptido PHDf-HD útiles en la realización de los procedimientos de la invención puede ser tan bajo como el 15 % de identidad de aminoácidos en comparación con SEC ID N°: 180.

Tabla D2: Resultados de MatGAT para similitud e identidad totales a lo largo de la longitud completa de las secuencias de polipéptido.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	25	31	40	19	38	26	17	32	40	20	19	40	30	40	23	31	25	25	18	19
1. Araith_PHDf_HD_PRHA	44	23	27	33	26	34	28	27	26	34	35	27	23	28	34	24	36	35	18	20
2. Araith_PHDf_HD_HAT3	46	39	35	18	34	20	18	31	34	17	18	33	56	36	21	57	21	23	15	15
3. Orysa_PHDf_HD	57	45	51	21	47	27	20	36	50	22	21	51	32	52	25	32	26	27	17	17
4. Eucgr_PHDf_HD	36	47	31	35	22	29	32	23	20	39	40	20	18	21	34	18	31	32	20	22
5. Lotja_PHDf_HD	55	44	51	64	36	26	19	35	50	20	20	49	30	50	25	32	27	26	17	17
6. Medtr_PHDf_HD	42	50	35	47	45	43	26	25	25	29	29	26	21	26	32	21	50	46	29	30
7. Orysa_PHDf_HD_HAZ1	31	39	28	32	48	31	38	22	20	32	33	20	18	20	30	18	26	25	19	21
8. Petctr_PHDf_HD_PRHP	53	44	47	51	42	50	44	39	36	23	23	37	28	39	26	30	25	27	19	19
9. Pintra_PHDf_HD	58	45	52	67	36	66	45	33	52	20	21	84	30	55	25	31	25	25	17	17
10. Poptr_PHDf_HD_I	36	47	30	34	56	35	44	46	41	36	80	21	19	22	37	18	31	31	20	20
11. Poptr_PH_Df_HD_III	36	49	31	34	57	34	46	47	42	35	86	21	18	22	38	18	30	30	20	21
12. Poptr_PHDf_HD_II	59	46	51	68	37	67	44	33	54	91	35	35	30	57	25	32	25	25	17	17
13. Poptr_PHDf_HD_IV	48	43	69	51	31	50	38	29	45	51	34	33	51	34	22	83	21	22	15	16
14. Sacof_PHDf_HD	58	47	52	69	36	67	43	32	52	74	36	35	75	52	23	34	27	25	16	17
15. Vitvi_PHDf_HD_I	39	47	36	39	56	39	45	47	43	40	56	57	40	36	41	22	32	31	23	24
16. Vitvi_PHDf_HD_II	48	41	70	51	32	52	38	30	45	54	33	32	53	87	53	34	22	22	15	15
17. Zeama_PHDf_HD	43	54	39	44	43	44	63	38	43	44	44	43	45	42	43	46	42	83	24	24
18. Zeama_hox1a	42	53	40	44	44	42	60	36	43	44	44	43	45	41	44	45	41	88	23	23
19. Zeama_hox1b	27	27	24	25	32	25	36	31	31	26	30	31	27	25	26	34	24	31	30	81
20. Zeama_hox2a	29	28	23	27	33	27	37	34	31	27	32	33	26	25	27	35	25	32	30	86
21. Zeama_Hox2b																				

5 La identidad porcentual puede aumentarse de forma sustancial si el cálculo de identidad se realiza entre el dominio conservado ZIP/PHDf (dominio conservado cremallera de leucina/dedo de homeodominio de planta, que comprende los dominios Zip y PHDf) de SEC ID N°: 180 (como se representa en SEC ID N°: 233), y el dominio conservado ZIP/PHDf de los polipéptidos útiles en la realización de la invención. El ZIP/PHDf conservado de SEC ID N°: 233 es de un total de 180 aminoácidos contiguos de largo. La identidad porcentual a lo largo del dominio conservado ZIP/PHDf entre las secuencias de polipéptido útiles en la realización de los procedimientos de la invención, varía entre el 30 % y el 75 % de identidad de aminoácidos, como se muestra en la Tabla D3.

Tabla D3: Resultados de MatGAT para similitud e identidad totales a lo largo del dominio conservado ZIP/PHDf entre las secuencias de polipeptido.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. ZIP/PHDf_Arath_PHDf_HD_PRHA		52	50	77	48	74	53	48	64	73	48	46	72	51	76	50	52	50	50	53	53
2. ZIP/PHDf_Arath_PHDf-HD_HAT3	70		44	52	73	54	70	69	56	52	75	75	51	43	53	78	46	65	67	64	65
3. ZIP/PHDf_Orysa_PHDf_HD	69	60		56	39	54	37	39	56	53	39	38	53	70	59	41	71	37	38	34	33
4. ZIP/PHDf_Eucgr_PHDf_HD	90	70	73		52	79	48	50	71	82	51	49	81	55	84	52	56	48	48	49	48
5. ZIP/Lotja_PHDf-HD	67	89	59	69		52	64	69	52	47	76	76	47	37	50	78	38	64	66	61	61
6. ZIP/Medtr_PHDf_HD	86	70	72	87	68		52	49	66	74	49	48	75	50	79	51	53	51	51	52	52
7. ZIP/PHDf_Orysa_PHDf-HD_HAZ1	69	83	54	67	83	68		64	51	47	68	67	47	35	47	68	37	79	80	82	84
8. ZIP/PHDf_Petr_PHDf_HD_PRHP	67	80	54	66	83	65	78		53	47	71	73	45	39	48	74	42	62	63	62	62
9. ZIP/PHDf_Pinra_PHDf_HD	82	74	72	83	73	82	70	69		66	53	52	67	52	73	55	54	53	52	53	52
10. ZIP/PHDf_Poptr_PHDf_HD_I	87	70	72	92	69	88	67	66	83		48	47	93	53	82	50	55	49	48	49	49
11. ZIP/PHDf_Poptr_PHDf_HD_III	66	87	57	67	90	65	82	82	70	67		93	48	36	51	82	39	67	70	65	64
12. ZIP/PHDf_Poptr_PHDf_HD_II	65	88	56	66	90	65	81	84	69	66	96		46	35	49	83	38	67	69	64	64
13. ZIP/PHDf_Poptr_PHDf_HD_IV	86	70	72	91	69	88	67	66	82	97	67	66		52	82	48	54	47	47	48	47
14. ZIP/PHDf_Sacof_PHDf_HD	71	60	82	73	57	69	55	57	70	72	56	55	72		57	39	92	35	35	33	33
15. ZIP/PHDf_Vitvi_PHDf_HD_I	89	71	73	90	70	88	67	64	84	91	67	66	89	73		51	58	48	48	48	48
16. ZIP/PHDf_Vitvi_PHDf_HD_II	67	88	58	69	91	68	84	84	73	69	90	93	67	57	69		42	65	67	65	64
17. ZIP/PHDf_Zeama_PHDf_HD	74	61	83	73	58	70	56	58	69	73	57	55	72	96	73	57		38	37	35	34
18. ZIP/PHDf_Zeama_hox1 a	71	84	59	71	85	70	88	79	74	72	83	82	71	61	70	85	62		94	74	76
19. ZIP/PHDf_Zeama_hox1 b	70	84	58	69	85	68	89	79	74	70	84	83	70	60	70	85	60	98		75	78
20. ZIP/PHDf_Zeama_hox2a	67	80	54	67	83	67	89	76	70	67	80	79	66	54	66	82	56	87	87		92
21. ZIP/PHDf_Zeama_Hox2b	67	82	53	66	83	67	90	77	71	66	80	79	67	53	67	82	55	88	88	88	97

La identidad porcentual también puede calcularse entre el HD conservado de SEC ID N°: 180 (como se representa en SEC ID N°: 234), y el HD conservado de los polipéptidos útiles en la realización de la invención. La identidad porcentual a lo largo del HD conservado entre las secuencias de polipéptido útiles en la realización de los procedimientos de la invención, varía entre el 25 % y el 70 % de identidad de aminoácidos, como se muestra en la Tabla D4.

5

Tabla D4: Resultados de MatGAT para similitud e identidad totales a lo largo de HD conservados de las secuencias de polipeptido.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. HD_Arath_PHDCHD_PRHA	34	54	72	36	72	38	36	58	68	38	36	70	54	70	34	54	36	33	29	27
2. HD_Arath_PHDf-HD_HAT3	56		28	32	54	34	40	58	30	32	54	58	32	32	58	28	42	39	40	39
3. HD_Orysa_PHDf_HD	76	60		68	42	58	42	36	46	38	40	66	94	62	42	92	38	35	35	33
4. HD_Eucgr_PHDf_HD	86	58	78		42	72	44	42	60	82	38	84	66	76	42	66	44	41	35	33
5. HD_Loija_PHDf-HD	60	72	66	64		38	52	58	36	42	64	68	42	40	68	40	56	55	50	46
6. HD_Medtr_PHDf_HD	78	60	70	82	66		36	36	62	68	38	36	70	66	36	58	40	37	27	27
7. HD_Orysa_PHDf-HD_HAZ1	56	54	64	62	66	58		44	26	44	50	48	44	40	46	42	70	65	73	71
8. HD_Petr_PHDf_HD_PRHP	58	74	58	56	66	58	58		36	42	54	58	42	38	62	36	50	45	40	39
9. HD_Pinra_PHDf_HD	74	54	64	72	58	72	50	52		58	30	32	60	56	36	44	28	28	21	19
10. HD_Poptr_PHDf_HD_I	86	56	78	92	64	84	64	56	74		36	36	98	72	42	64	42	39	35	33
11. HD_Poptr_PHDf_HD_III	54	72	62	60	82	58	66	66	54	58		88	36	38	72	38	52	49	46	44
12. HD_Poptr_PHDf_HD_II	56	72	64	62	82	60	68	66	56	60	100		36	40	38	78	40	50	47	46
13. HD_Poptr_PHDf_HD_IV	88	56	80	94	64	86	64	56	76	98	60		66	72	42	66	42	39	35	33
14. HD_Sacof_PHDf_HD	74	60	94	76	64	72	62	56	62	76	60	62	78	64	42	98	36	33	35	33
15. HD_Vitvi_PHDf_HD_I	82	60	80	86	70	80	60	58	74	88	62	64	88		38	64	40	37	29	27
16. HD_Vitvi_PHDf_HD_II	64	76	66	66	82	68	64	74	56	66	84	84	66	68		40	48	43	44	44
17. HD_Zeama_PHDf_HD	74	60	94	76	64	72	62	56	62	76	60	62	78	100	64		36	33	35	33
18. HD_Zeama_hox1a	56	60	62	60	66	60	78	60	52	62	62	64	62	60	68	56		86	60	56
19. HD_Zeama_hox1b	57	63	61	61	69	61	80	63	51	63	67	69	63	57	61	71	55	92	59	53
20. HD_Zeama_hox2a	58	60	60	60	65	64	89	60	52	62	64	64	62	60	62	64	58	77	81	90
21. HD_Zeama_Hox2b	56	56	56	56	62	58	85	54	48	58	60	60	58	58	62	54	71	77	92	

Ejemplo 4: Identificación de dominios comprendidos en las secuencias de polipéptido útiles en la realización de los procedimientos de la invención

La base de datos *The Integrated Resource of Protein Families, Domains and Sites* (InterPro) es una interfaz integrada por las bases de datos de firmas usadas comunmente para búsquedas basadas en texto y secuencia. La base de datos InterPro combina estas bases de datos, que usan diferentes metodologías y grados variables de información biológica sobre proteínas bien caracterizadas para obtener firmas proteicas. 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 de la exploración InterPro de la secuencia de polipéptido como se representa en SEC ID N°: 180, se presentan en la Tabla D5.

Tabla D5: Resultados de la exploración InterPro de la secuencia de polipéptido como se representa en SEC ID N°: 180

Número de acceso InterPro	Nombre de la base de datos integrada	Número de acceso a la base de datos integrada	Nombre de acceso a la base de datos integrada	Coordenadas de aminoácidos en SEC ID N°: 180
IPR0013556 Homeosecuencia	ProDom	PD000010	PRH_ARATH_P48785	439-497
IPR0013556 Homeosecuencia	PFAM	PF00046	Homeosecuencia	439-495
IPR0013556 Homeosecuencia	SMART	SM00389	HOX	438-500
IPR0013556 Homeosecuencia	Profile	PS50071	Homeobox_2	436-496
IPR00195 dedo de Zinc, tipo PHD	PFAM	PF00628	PHD	197-251
IPR00195 dedo de Zinc, tipo PHD	SMART	SM00249	PHD	197-249
IPR00195 dedo de Zinc, tipo PHD	Profile	PS50016	ZF_PHD_2	195-251
IPR012287 Relacionado con homeodominio	Gene3D	G3DSA:1.10.10.60	Sin descripción	436-501
IPR013256 Cromatina SPT2	SMART	SM00784	Sin descripción	78-154
Untegrated IPR	PANTHER	PTHR19418	Proteína de homeosecuencia	397-414 431-590

Los resultados de la exploración InterPro identifican claramente las características esenciales de un polipéptido PHDf-HD, es decir, un dedo de zinc de PHDf y una homeosecuencia, como se representa respectivamente, por ejemplo, por las entradas de Pfam PF00628 y PF00046.

Se sabe que los HD tienen restos canónicos. El HD de los polipéptidos PHDf-HD es de secuencia altamente divergente, incluso en posiciones que son casi invariables entre homeodominios. En la Figura 3 se muestra el logo de secuencia del HD de los polipéptidos PHDf-HD de la Tabla D1. Los logos de secuencia son una representación gráfica de un alineamiento de secuencias múltiple de aminoácidos o ácido nucleico. Cada logo consiste de pilas de símbolos, una pila para cada posición en la secuencia. La altura total de la pila indica la conservación de secuencia en esa posición, mientras que la altura de los símbolos dentro de la pila indica la frecuencia relativa de cada aminoácido o ácido nucleico en esa posición. En general, un logo de secuencia proporciona una descripción más rica y más precisa que una secuencia consenso de, por ejemplo, un sitio de unión. El algoritmo (WebLogo) para producir tales logos está disponible en el servidor de la University of California, Berkeley. El HD como se representa en SEC ID N°: 234, y comprendido en SEC ID N°: 180, está en conformidad con el logo de secuencia como se representa en la Figura 3. Los polipéptido útiles en la realización de los procedimientos, de acuerdo con la invención, comprenden un HD comprendido en el logo de secuencia como se muestra en la Figura 3.

Ejemplo 5: Predicción de las características de la estructura secundaria de las secuencias de polipéptido útiles en la realización de los procedimientos de la invención

Las súper hélices normalmente contienen un patrón repetido de siete restos de aminoácidos denominado repeticiones héptadas. Las súper hélices son importantes para identificar interacciones proteína-proteína, tales como oligomerización, de ya sea, proteínas idénticas, proteínas de la misma familia, o proteínas no relacionadas. Recientemente se ha hecho mucho avance en la predicción computacional de súper hélices a partir de datos de secuencia. En las herramientas ExPASy Proteomics están disponibles muchos algoritmos bien conocidos para un experto en la materia. Uno de ellos, COILS, es un programa que compara una secuencia con una base de datos de súper hélices de dos cadenas paralelas conocidas y obtiene una puntuación de similitud. Después el programa calcula la probabilidad de que la secuencia adopte una conformación de súper hélice, mediante la comparación de esta puntuación con la distribución de puntuaciones en proteínas globulares y de súper hélice.

El polipéptido PHDf-HD como se representa en SEC ID N°: 180, tiene dos dominios súper hélice predichos N-terminales, con una alta probabilidad, en las tres ventanas (14, 21 y 28) examinadas. En la Tabla D6, se muestran las coordenadas de restos, restos, las tres ventanas y los valores correspondientes de probabilidad. En la Figura 4, está el resultado gráfico del algoritmo COILS en el polipéptido como se representa en SEC ID N°: 180, en la que las dos súper hélices predichas son claramente visibles en las tres ventanas (representado mediante las tres líneas), en la mitad N-terminal del polipéptido.

Tabla D6: Resultado numérico del algoritmo COILS en el polipéptido como se representa en SEC ID N°: 180. Se muestran las coordenadas de restos (N.º), restos, las tres ventanas y los valores correspondientes de probabilidad. Las probabilidades por encima de 0,09 se muestran el gris.

N.º	Resto	Ventana=14	Prob	Ventana=21	Prob	Ventana=28	Prob
87	P	d	0,001	e	0,001	E	0,001
88	T	e	0,063	f	0,274	F	0,099
89	R	f	0,063	g	0,567	G	0,099
90	R	g	0,063	a	0,567	A	0,099
91	K	b	0,066	b	0,642	B	0,188
92	H	c	0,066	c	0,642	C	0,188
93	K	d	0,069	d	0,642	D	0,591
94	Q	e	0,178	e	0,642	E	0,672
95	K	f	0,335	f	0,642	F	0,672
96	R	g	0,335	g	0,642	G	0,772
97	K	a	0,335	a	0,642	A	0,772
98	N	b	0,347	b	0,642	B	0,772
99	D	c	0,347	c	0,642	C	0,772
100	E	d	0,347	d	0,642	D	0,772
101	S	e	0,347	e	0,642	E	0,808
102	D	f	0,347	f	0,642	F	0,808
103	E	g	0,347	g	0,642	g	0,808
104	V	a	0,347	a	0,642	a	0,808
105	S	b	0,347	b	0,642	b	0,808
106	R	c	0,347	c	0,642	C	0,808
107	M	d	0,347	d	0,642	d	0,808
108	E	e	0,347	e	0,642	e	0,808
109	K	f	0,347	f	0,642	F	0,808
110	R	g	0,347	g	0,642	g	0,808

ES 2 558 745 T3

(continuación)

N.º	Resto	Ventana=14	Prob	Ventana=21	Prob	Ventana=28	Prob
111	A	a	0,347	a	0,642	a	0,808
112	R	b	0,307	b	0,514	b	0,808
113	Y	c	0,066	c	0,503	c	0,808
114	L	d	0,262	d	0,503	d	0,808
115	L	e	0,262	e	0,503	e	0,808
116	I	f	0,262	f	0,503	F	0,808
117	K	g	0,262	g	0,503	g	0,808
118	I	a	0,262	a	0,503	a	0,808
119	K	b	0,262	b	0,503	b	0,808
120	Q	c	0,262	c	0,503	c	0,808
121	E	d	0,262	d	0,503	d	0,808
122	Q	e	0,262	e	0,503	e	0,808
123	N	f	0,262	f	0,503	F	0,808
124	L	g	0,262	g	0,434	g	0,808
125	L	a	0,262	a	0,434	a	0,808
126	D	b	0,262	b	0,434	b	0,808
127	A	c	0,262	c	0,304	c	0,808
128	Y	d	0,145	d	0,304	d	0,808
129	S	e	0,145	e	0,044	e	0,808
130	G	f	0,119	f	0,009	F	0,367
131	D	g	0,078	g	0,007	g	0,136
132	G	a	0,002	a	0,001	a	0,005
133	W	b	0	a	0,001	a	0,001
134	N	b	0,001	b	0,045	b	0,057
135	G	c	0,001	c	0,045	c	0,057
136	H	d	0,003	d	0,07	d	0,219
137	S	e	0,008	e	0,132	e	0,231
138	R	f	0,025	f	0,357	F	0,231
139	E	g	0,025	g	0,357	g	0,231
140	K	a	0,025	a	0,357	a	0,231
141	I	b	0,028	b	0,357	b	0,231
142	K	c	0,07	c	0,357	c	0,462
143	P	d	0,07	d	0,357	d	0,462
144	E	e	0,998	e	0,997	e	0,974
145	K	f	0,998	f	0,997	f	0,974
146	E	g	0,998	g	0,997	g	0,974

(continuación)

N.º	Resto	Ventana=14	Prob	Ventana=21	Prob	Ventana=28	Prob
147	L	a	0,998	a	0,997	a	0,974
148	Q	b	0,998	b	0,997	b	0,974
149	R	c	0,998	c	0,997	c	0,974
150	A	d	0,998	d	0,997	d	0,974
151	K	e	0,998	e	0,997	e	0,974
152	K	f	0,998	f	0,997	f	0,974
153	Q	g	0,998	g	0,997	g	0,974
154	I	a	0,998	a	0,997	a	0,974
155	M	b	0,998	b	0,997	b	0,974
156	K	c	0,998	c	0,997	c	0,974
157	Y	d	0,998	d	0,997	d	0,974
158	K	e	0,989	e	0,997	e	0,974
159	I	f	0,896	f	0,997	f	0,974
160	A	g	0,486	g	0,997	g	0,974
161	I	a	0,409	a	0,997	a	0,974
162	R	b	0,274	b	0,997	b	0,974
163	D	c	0,255	c	0,997	c	0,974
164	V	d	0,095	d	0,997	d	0,974
165	I	e	0,025	e	0,871	e	0,974
166	H	f	0,025	f	0,622	f	0,974
167	Q	g	0,025	g	0,496	g	0,974
168	L	a	0,025	a	0,496	a	0,974
169	D	b	0,025	b	0,468	b	0,974
170	L	c	0,005	c	0,111	c	0,974
171	C	d	0,002	d	0,02	d	0,974
172	S	e	0,001	e	0,007	e	0,719
173	S	f	0,001	f	0,003	f	0,482
174	S	g	0,001	g	0,001	g	0,066
175	G	a	0	a	0	a	0,001

Está predicha otra súper hélice en la mitad C-terminal del polipéptido PHDf-HD como se representa en SEC ID N.º: 180, con una probabilidad menor que los dos dominios hélice comprendidos en la mitad N-terminal del polipéptido.

5 Ejemplo 6: Predicción de la localización subcelular de las secuencias de polipéptido útiles en la realización de los procedimientos de la invención

LOCtree es un algoritmo que puede predecir la localización subcelular y la propensión a la unión de ADN de proteínas que no son de membrana en eucariotas vegetales y que no son vegetales así como en procariotas. LOCtree clasifica proteínas de animal eucariotas en una de cinco clases subcelulares, mientras que las proteínas de planta se clasifican en una de seis clases y las proteínas de procariota se clasifican en una de tres clases.

5 Siempre que esté disponible, LOctree también notifica predicciones a base de lo siguiente: 1) Señales de localización nuclear encontradas por el algoritmo PredictNLS, 2) Localización deducida utilizando motivos Prosite y dominios Pfam encontrados en la proteína, y 3) palabras clave de SWISS-PROT asociadas con una proteína. En los dos últimos casos la localización se deduce utilizando el algoritmo LOCKey basado en entropía. El programa informático está alojado en la University of Columbia, USA.

La predicción de la localización subcelular basada en el motivo y palabra clave, de un polipéptido PHDf-HD como se representa en SEC ID N°: 180, utilizando LOCKey:

Localización Predicha	Confianza	Predicción alternativa	Palabras clave SWISS-PROT utilizadas para asignar localización
Nuclear	100	-	Homeosecuencia, unión a ADN, Regulación de la transcripción, Proteína nuclear, Transcripción, Dedo de zinc

10 La predicción de una señal de localización nuclear (SLN) se hace utilizando, por ejemplo, el algoritmo PredictNLS. El algoritmo también está alojado en la University of Columbia, USA. La predicción de la SLN en el polipéptido de SEC ID N°: 180, utilizando el algoritmo PredictNLS, se muestra en la Tabla a continuación.

Motivo generalizado ([KR]{3,5} entre 3 y 5 veces K o R)	Aciertos en polipéptido de SEC ID N°: 180	Coordenadas en el polipéptido de SEC ID N°: 180
K[RK]{3,5}x{11,18}[RK]Kx{2,3}K	KRRRGSDAATG KSATG PTRRKHKQK	71-95
[KR]{4}x{20,24}K{1,4}xK	KRRRGSDAATG KSATG PTRRKHKQKRK	71-97

La posición de la SLN predicha se muestra en negrita y con doble subrayado en la secuencia de polipéptido a continuación (SEC ID N°: 180):

MNTPEKKPLCYTSRRALQQRTESSSELI SVSKRATRQNTPRKPDSPPKRTRSSANLAKC
 IENKHHSSPL **KRRRGSDAATGKSATGPTRRKHKQKRK** NDESDEVSRMEKRARYLLIKIKQ
 EQNLLDAYSGDGNHGSREKIKPEKELQRAKKQIMKYKIAIRDVIHQDLDCSSSGSKDDS
 VIPPDGCHESVNPEHTICSRCKSHESFPDNNIIFCEGGCKLACHQKCLEPPFDKILPTR
 HGRLCKKCHSSKMKILDAINAHLGTSFTVKCPSSDIFKEAAEHFNSDDGLGQDWLSEYSGD
 EDYDPEENEASSSGEENKSADSNCSGSPLYSPNDDIPDFISADFND AEGFCRESSNLGID
 FGEDGLAEILTHQRPRRDVDYTLQNEQMFGEPIGNDEQSEDEDWGLNKRKRRRTGSTGVG
 TNSVEGRSDVKSNNKAQPRRKLFRIPPAAVEVLRKAFAENELPARSVKENLSTELGISFE
 KIDKWFKNTRCAALDRKGESRYSGPSKRSRTSIEKAETSAKVDQMDNSCFPLPSEIINV
 PTRLQKGLDKKPKSINSPRPQDNETCLSPTDKTKEGTPPTIKPSITDSSQLMNNDIGTE
 ETAVSWVDTWASDALHFLDVSDDHFFDVIEKVCGLNRLQRLKENMLSSSSSTDNNVAA
 ESGLQNEVVLVPAEELKDKAS

15 Pueden utilizarse muchos otros algoritmos para realizar tales análisis, incluyendo:

- CloroP 1.1 alojado en el servidor de la Technical University of Denmark;
- Protein Prowler Subcellular Localisation Predictor versión 1.2 alojado en el servidor del Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia;
- PENCE Proteome Analyst PA-GOSUB 2.5 alojado en el servidor de la University of Alberta, Edmonton, Alberta, Canada;
- TMHMM, alojado en el servidor de la Technical University of Denmark

Ejemplo 7: Clonado de la secuencia de ácido nucleico como se representa en SEC ID N°: 179

25 A menos que se establezca lo contrario, las técnicas de ADN recombinante se realizaron de acuerdo con protocolos convencionales descritos en (Sambrook (2001) Molecular Cloning: a laboratory manual, 3ª Edición Cold Spring Harbor Laboratory Press, CSH, Nueva 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 en

plantas, se describen en Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (RU) y Blackwell Scientific Publications (RU).

5 El gen PHDf-HD de *Oryza sativa* se amplificó por PCR utilizando como molde un banco de ADNc de arroz sintetizado a partir de ARNm extraído de tejidos de planta mezclados. Para la amplificación por PCR se usaron el primer prm09687 (SEC ID N°: 236; en sentido: 5'-ggggacaagttgtacaaaaagcaggcttaacaatgaataccagaaaagaaa-3') y el primer prm09688 SEC ID N°: 237; inverso, complementario: 5'-ggggaccacttgtacaagaagctgggtgatgcaaggtaagatgctt-3'), que incluyen los sitios AttB para la recombinación Gateway. La PCR se realizó utilizando la polimerasa de ADN Taq Hifi en condiciones convencionales. Se amplificó un fragmento de PCR de la longitud esperada (incluyendo sitios AttB) y se purificó utilizando también procedimientos convencionales. Después se realizó la primera etapa del procedimiento Gateway, la reacción BP, durante la cual el fragmento de PCR recombinó *in vivo* con el plásmido pDONR201 para producir, de acuerdo con la terminología Gateway, un "clon de entrada". El plásmido pDONR201 se adquirió en Invitrogen, como parte de la tecnología Gateway®.

15 **Ejemplo 8: Construcción del vector de expresión utilizando la secuencia de ácido nucleico como se representa en SEC ID N°: 179**

El clon de entrada que comprendía la SEC ID N°: 179 se usó posteriormente en una reacción LR con un vector destinatario utilizado para transformación de *Oryza sativa*. Este vector contiene, como elementos funcionales dentro de los límites del ADN T: un marcador de selección de plantas; un casete de expresión marcador detectable y un casete Gateway diseñado para recombinación LR *in vivo* con la secuencia de ácidos nucleicos de interés ya clonada en el clon de entrada. Un promotor GOS2 de arroz (SEC ID N°: 235) para expresión constitutiva se localizaba cadena arriba de este casete Gateway.

Después de la etapa de recombinación LR, el vector de expresión resultante pGOS2::PHDf-HD y (Figura 6) se transformó en la cepa LBA4044 de *Agrobacterium* de acuerdo con procedimientos bien conocidos en la técnica.

Ejemplo 9: Transformación de plantas

25 **Transformación de arroz**

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

35 Para el cocultivo se usó, de forma independiente, la cepa LBA4404 de *Agrobacterium* que contenía cada vector de expresión individual. Se inoculó *Agrobacterium* en medio AB con los anticuerpos apropiados y se cultivó durante 3 días a 28 °C. Después, las bacterias se recogieron y se suspendieron en medio de cocultivo líquido a una densidad (DO₆₀₀) de aproximadamente 1. Después, la suspensión se transfirió a una placa de Petri y los callos se sumergieron en la suspensión durante 15 minutos. Después, los tejidos de callo se secaron por transferencia en un papel de filtro y se transfirieron a medio de cocultivo solidificado y se incubaron durante 3 días en la oscuridad a 25 °C. Los callos cocultivados crecieron sobre un medio que contenía 2,4-D durante 4 semanas en la oscuridad a 28 °C en presencia de un agente de selección. Durante este periodo, rápidamente se desarrollaron islas de callos resistentes a crecimiento. Después de la transferencia de este material a un medio de regeneración e incubación a la luz, el potencial embriogénico se liberó y se desarrollaron brotes las siguientes cuatro a cinco semanas. Los brotes se extirparon de los callos y se incubaron durante 2 a 3 semanas en un medio que contenía auxina desde el cual se transfirieron al suelo. Los brotes endurecidos se cultivaron a alta humedad y en días cortos en un invernadero.

50 Se generaron aproximadamente 35 transformantes de arroz T0 independientes para cada construcción. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero. Después de un análisis de PCR cuantitativa para verificar el número de copias del inserto de ADN-T, solo una única copia de plantas transgénicas que presentaban tolerancia al agente de selección, se guardaron para la cosecha de la semilla T1. Después, las semillas se cosecharon de tres a cinco meses después del trasplante. El procedimiento produjo transformantes de locus único a una tasa de aproximadamente el 50 % (Aldemita y Hodges 1996, Chan y col. 1993, Hiei y col. 1994).

Ejemplo 10: Procedimiento de evaluación fenotípica10.1 Configuración de la evaluación

Se generaron aproximadamente 35 transformantes de arroz T0 independientes. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero para siembra y cosecha de la semilla T1. Se conservaron seis acontecimientos, de los cuales la progenie T1 segregó 3:1 para la presencia/ausencia del transgén. Para cada uno de estos acontecimientos, se seleccionan aproximadamente 10 plántulas T1 que contenían el transgén (hetero y homocigotos) y aproximadamente 10 plántulas T1 que carecían del transgén (nulicigotos) supervisando la expresión con marcador visual. Las plantas transgénicas y los correspondientes nulicigotos se cultivaron a cada lado en posiciones aleatorias. Las condiciones en el invernadero eran días cortos (12 horas de luz), 28 °C con luz y 22 °C en oscuridad, y una humedad relativa del 70 %. Se suministró agua a intervalos regulares a las plantas cultivadas en condiciones sin estrés, para asegurar que el agua y los nutrientes no eran limitantes para satisfacer las necesidades de la planta de agua para completar el crecimiento y desarrollo.

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

10.2 Análisis estadístico: ensayo F

Se utilizó un ANOVA (análisis de varianza) de dos factores como un modelo estadístico para la evaluación global de las características fenotípicas de la planta. Se realizó un ensayo F en todos los parámetros medidos de todas las plantas de todos los acontecimientos transformados con el gen de la presente invención. El ensayo F se realizó para verificar un efecto del gen sobre todos los acontecimientos de transformación y para verificar un efecto global del gen, también conocido como efecto génico global. El umbral de significación de un efecto génico global verdadero se estableció a un nivel de probabilidad del 5 % para el ensayo F. Un valor significativo de los puntos del ensayo F para un efecto génico, significa que no solo es la mera presencia o la posición del gen lo que ocasiona las diferencias en el fenotipo.

10.3 Parámetros medidos

Medición de parámetros relacionados con biomasa

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

El área aérea de la planta (o biomasa foliar) se determinó contando el número total de píxeles en las imágenes digitales de las partes aéreas de la planta, discriminadas del fondo. Este valor se promedió para las imágenes tomadas en el mismo momento desde los diferentes ángulos, y se convirtió a un valor de superficie física expresado en mm cuadrados por calibración. Los experimentos mostraron que el área aérea de la planta, medida de este modo, se correlacionaba con la biomasa de las partes aéreas de la planta. El área aérea es el área medida en el momento en el que la planta ha alcanzado su máxima biomasa foliar. El vigor temprano es el área aérea de la planta (plántula) tres semanas después de la germinación. El aumento de la biomasa radicular se expresa como un aumento en la biomasa radicular total (medido como biomasa radicular máxima observada durante la vida útil de una planta); o como un aumento en el índice de raíz/brote (medido como la relación entre la masa radicular y la masa de brote en el período de crecimiento activo de raíces y brotes).

Mediciones de parámetros relacionados con la semilla

Las panículas primarias maduras se cosecharon, se contaron, se metieron en bolsas, se etiquetaron con código de barras y después se secaron durante tres días en un horno a 37 °C. Después, las panículas se trillaron y todas las semillas se recogieron y se contaron. Las vainas llenas se separaron de las vacías utilizando un dispositivo de soplado de aire. Las vainas vacías se desecharon y la fracción restante volvió a contarse. Las vainas llenas se pesaron en una balanza analítica. El número de semillas llenas se determinó contando el número de vainas llenas que quedaron después de la etapa de separación. El rendimiento total de la semilla se midió pesando todas las vainas llenas cosechadas de una planta. El número total de semillas por planta se midió contando el número de vainas cosechadas de una planta. El peso de mil granos (PMG) se extrapó del número de semillas llenas contadas y de su peso total. El índice de cosecha (IC) en la presente invención se define como la relación entre el rendimiento de semilla total y el área aérea (mm²), multiplicado por un factor de 10⁶. El número total de flores por panícula como se define en la presente invención es la relación entre el número total de semillas y el número de panículas primarias maduras. La tasa de llenado de semilla, como se define en la presente invención, es la proporción (expresada en %) del número de semillas llenas sobre el número total de semillas (o floretes).

Detección de la disponibilidad reducida de nutrientes (nitrógeno)

Se crecieron las plantas de seis acontecimientos (semillas T2) en suelo de recipientes en condiciones normales excepto para la solución de nutrientes. Los recipientes se regaron desde el trasplante hasta la maduración con una solución de nutrientes específica conteniendo contenido reducido de nitrógeno (N), habitualmente entre 7 y 8 veces menos. El resto del cultivo (maduración de la planta, cosecha de semilla) es la misma que la de las plantas no crecidas en estrés abiótico. Los parámetros de crecimiento y rendimiento se registraron como se detalla para el crecimiento en condiciones normales.

Ejemplo 11: Resultados de la evaluación fenotípica de las plantas de arroz transgénicas

Se presentan a continuación los resultados de la evaluación de las plantas de arroz transgénicas que expresan la secuencia de ácido nucleico que codifica un polipéptido PHDf- HD como se representa en SEC ID N°: 180, bajo el control del promotor GOS2 para expresión constitutiva.

En comparación con los nulicigotos correspondientes (controles), en las plantas transgénicas hubo un aumento significativo del número de panículas, del rendimiento total de semilla por planta, del número total de semillas llenas, del número total de semillas, del Peso de Mil Granos (PMG), y del índice de cosecha, como se muestra en la Tabla D7.

Tabla D7: Resultados de la evaluación de plantas de arroz transgénicas que expresan la secuencia de ácido nucleico que codifica un polipéptido PHDf-HD, como se representa en SEC ID N°: 180, bajo el control del promotor GOS2 para expresión constitutiva

	Aumento promedio en % en 3 acontecimientos en la generación T1
Número de panículas iniciales	22 %
Rendimiento total de semilla por planta	30 %
Número total de semillas llenas	24 %
Número de total de semillas	24 %
PMG	4 %
Índice de cosecha	17 %

Ejemplo 12: Resultados de la evaluación fenotípica de las plantas de arroz transgénicas crecidas en disponibilidad reducida de nutrientes

Se presentan a continuación los resultados de la evaluación de las plantas de arroz transgénicas que expresan la secuencia de ácido nucleico que codifica un polipéptido PHDf-HD, como se representa en SEC ID N°: 180, bajo el control del promotor GOS2 para expresión constitutiva, y crecidas en disponibilidad reducida de nutrientes.

En comparación con los nulicigotos correspondientes (controles), en las plantas transgénicas hubo un aumento significativo de la biomasa, del vigor de emergencia, del rendimiento total de semilla por planta, del número total de semillas llenas, de la tasa de llenado de semilla, del número total de semillas, del Peso de Mil Granos (PMG) y del índice de cosecha, como se muestra en la Tabla D8.

Tabla D8: Resultados de la evaluación de las plantas de arroz transgénicas que expresan la secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se representa en SEC ID N°: 180, bajo el control del promotor GOS2 para expresión constitutiva, y crecidas en disponibilidad reducida de nutrientes.

	Aumento promedio total en % (6 acontecimientos en la generación T2)
Biomasa	8%
Vigor de emergencia	20 %
Rendimiento total de semilla por planta	17 %
Número total de semillas llenas	12 %
Tasa de llenado de semilla	3 %
Número total de semillas	9 %
PMG	5 %
Índice de cosecha	9 %

Ejemplo 13: Ejemplos de transformación de otros cultivos

La transformación de maíz, trigo, soja, colza/canola, alfalfa y algodón se llevó a cabo como se describe a continuación.

Transformación de arroz

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

15 Para el cocultivo se usó la cepa LBA4404 de *Agrobacterium*, que contenía el vector de expresión. Se inoculó *Agrobacterium* en medio AB con los antibióticos apropiados y se cultivó durante 3 días a 28 °C. Después, las bacterias se recogieron y se suspendieron en medio de cocultivo líquido a una densidad (DO₆₀₀) de aproximadamente 1. Después, la suspensión se transfirió a una placa de Petri y los callos se sumergieron en la suspensión durante 15 minutos. Después, los tejidos de callo se secaron por transferencia en un papel de filtro y se transfirieron a medio de cocultivo solidificado y se incubaron durante 3 días en la oscuridad a 25 °C. Los callos co-
20 cultivos crecieron sobre un medio que contenía 2,4-D durante 4 semanas en la oscuridad a 28 °C, en presencia de un agente de selección. Durante este periodo, rápidamente se desarrollaron islas de callos resistentes a crecimiento. Después de la transferencia de este material a un medio de regeneración e incubación a la luz, el potencial embriogénico se liberó y se desarrollaron brotes las siguientes cuatro a cinco semanas. Los brotes se cortaron de los callos y se incubaron durante 2 a 3 semanas en un medio que contenía auxina desde el cual se transfirieron al
25 suelo. Los brotes endurecidos se crecieron con alta humedad y en días cortos en un invernadero.

Se generaron aproximadamente 35 transformantes de arroz T0 independientes para una construcción. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero. Después de un análisis de PCR cuantitativa para verificar el número de copias del inserto de ADN-T, solo una única copia de plantas transgénicas que presentaban tolerancia al agente de selección, se guardaron para la cosecha de la semilla
30 T1. Después, las semillas se cosecharon de tres a cinco meses después del trasplante. El procedimiento produjo transformantes de locus único a una tasa de aproximadamente el 50 % (Aldemita y Hodges 1996, Chan y col. 1993, Hiei y col. 1994).

Transformación de maíz

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

Transformación de trigo

50 La transformación de trigo se realizó con el procedimiento descrito por Ishida y col. (1996) Nature Biotech 14(6): 745-50. Normalmente se utiliza la variedad de cultivo Bobwhite (disponible en CIMMYT, México) para la transformación. Embriones inmaduros se cocultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través de organogénesis. Después de la incubación con
55 *Agrobacterium*, los embriones se cultivaron *in vitro* en medio de inducción de callo, después en medio de regeneración que contiene el agente de selección (por ejemplo imidazolinona, aunque pueden utilizarse diversos marcadores de selección). Las placas de Petri se incubaron a la luz a 25 °C durante 2-3 semanas o hasta que se desarrollasen los brotes. Los brotes verdes se transfirieron desde cada embrión al medio de enraizamiento y se incubaron a 25 °C durante 2-3 semanas, hasta el desarrollo de las raíces. Los brotes enraizados se trasplantaron al

suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de soja

5 La soja se transformó de acuerdo con una modificación del procedimiento descrito en el Texas A y M patente de Estados Unidos 5.164.310. Diversas variedades comerciales de soja son susceptibles a transformación mediante este procedimiento. La variedad de cultivo Jack (disponible en la fundación Illinois Seed) se utiliza normalmente para la transformación. Se esterilizaron semillas de soja para la siembra *in vitro*. El hipocotiledóneo, el radículo y un cotiledón se extirparon de plántulas jóvenes de siete días de vida. El epicótilo y el cotiledón restantes se cultivaron adicionalmente para desarrollar nodos axilares. Estos nodos axilares se extirparon y se incubaron con 10 *Agrobacterium tumefaciens* que contenía el vector de expresión. Después del tratamiento de cocultivo, los explantes se lavaron y se transfirieron a medios de selección. Los brotes regenerados se extirparon y se colocaron en un medio de elongación de brote. Los brotes no mayores de 1 cm se colocaron en medio de enraizamiento hasta que se desarrollaron las raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron 15 semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de colza/canola

Peciolos cotiledonarios e hipocótilos de plántulas jóvenes de 5-6 días de vida se utilizaron como explantes para el cultivo tisular y se transformaron de acuerdo con Babic y col. (1998, Plant Cell Rep 17: 183-188). La variedad de cultivo comercial Westar (Agriculture Canadá) es la variedad convencional usada para transformación, pero también 20 pueden usarse otras variedades. Se esterilizaron semillas de canola en la superficie para la siembra *in vitro*. Explantes de peciolos cotiledonarios con el cotiledón unido se extirparon de las plántulas *in vitro* y se inocularon con *Agrobacterium* (que contenía el vector de expresión) sumergiendo el extremo cortado del explante del peciolo en la suspensión bacteriana. Después, los explantes se cultivaron durante 2 días en medio MSBAP-3 que contenía BAP 3 mg/l, sacarosa al 3 %, Fitagar al 0,7 % a 23 °C, 16 h de luz. Después de dos días de cocultivo con *Agrobacterium*, 25 los explantes de peciolos se transfirieron a medio MSBAP-3 que contenía BAP 3 mg/l, cefotaxima, carbenicilina o timentina (300 mg/l) durante 7 días y después se cultivaron en medio MSBAP-3 con cefotaxima, carbenicilina o timentina y agente de selección hasta la regeneración del brote. Cuando los brotes tenían una longitud de 5 - 10 mm, se cortaron y se transfirieron al medio de elongación de brote (MSBAP-0,5 que contenía BAP 0,5 mg/l). Los brotes de aproximadamente 2 cm de longitud se transfirieron al medio de enraizamiento (MS0) para la inducción de raíces. 30 Los brotes enraizados trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de alfalfa

Utilizando el procedimiento de (McKersie y col., 1999 Plant Physiol 119: 839-847) se transformó un clon de regeneración de alfalfa (*Medicago sativa*). La regeneración y transformación de la alfalfa es dependiente del 35 genotipo y por lo tanto se requiere una planta regenerada. Se han descrito procedimientos para obtener plantas regeneradas. Por ejemplo, estas pueden seleccionarse de variedades de cultivo Ranglander (Agriculture Canadá) o de cualquier otra variedad comercial de alfalfa como describen Brown DCW y A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Como alternativa, 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). Explantes de peciolos se cocultivaron durante una noche con un cultivo de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie y col., 1999 Plant Physiol 119: 839- 40 847) o LBA4404 que contenía el vector de expresión. Los explantes se cocultivaron durante 3 días en la oscuridad en un medio de inducción SH que contenía Pro 288 mg/l, tioprolina 53 mg/l, K₂SO₄ 4,35 g/l y acetosiringinona 100 µm. Los explantes se lavaron en medio Murashige-Skoog de fuerza media (Murashige y Skoog, 1962) y se sembraron en placas en el mismo medio de inducción SH con acetosiringinona pero con un agente de selección 45 adecuado y un antibiótico adecuado para inhibir el crecimiento de *Agrobacterium*. Después de varias semanas, los embriones somáticos se transfirieron al medio de desarrollo BOi2Y que no contenía reguladores de crecimiento, sin antibióticos, y sacarosa 50 g/l. Posteriormente, los embriones somáticos germinaron en medio Murashige-Skoog de fuerza media. Las plántulas enraizadas se trasplantaron en macetas y se cultivaron en un invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia 50 del inserto de ADN-T.

Transformación de algodón

La transformación de algodón (*Gossypium hirsutum* L.) se realizó utilizando *Agrobacterium tumefaciens*, sobre explantes de hipocótilos. Las variedades de cultivo comerciales, tales como Coker 130 o Coker 312 (SeedCo, Lubbock, TX), son las variedades convencionales utilizadas para transformación, pero pueden utilizarse otras 55 variedades. Se esteriliza la superficie de las semillas, y se germinan en oscuridad. Se cortan los explantes de hipocótilos de las plántulas germinadas, hasta un largo de aproximadamente 1-1,5 centímetros. El explante de hipocótilo se sumerge en el inóculo de *Agrobacterium tumefaciens* que contiene el vector de expresión, durante 5 minutos, después se cocultiva durante aproximadamente 48 horas en MS + KNO₃ 1,8 mg/l + glucosa al 2 % a 24 °C, en oscuridad. Se transfieren los explantes al mismo medio, que contiene marcadores de selección bacterianos y de

5 planta apropiados (renovados varias veces), hasta que se ven las callos embriogénicos. Los callos se separan y se subcultivan hasta que aparecen los embriones somáticos. Los plantones obtenidos de los embriones somáticos se maduran en medio de enraizamiento hasta que se desarrollan las raíces. Los brotes enraizados se trasplantan a suelo de maceta en el invernadero. Se producen semillas T1 de plantas que presentan tolerancia al agente de selección y que contienen una única copia del inserto de ADN-T.

Ejemplo 14: Ejemplos de detecciones de estrés abiótico

Detección de sequía

10 Se crecen plantas de un número elegido de acontecimientos en suelo de maceta en condiciones normales, hasta que alcanzan la etapa de encabezamiento. Después se transfieren a una sección “seca” en la que se mantiene el riego. Se insertan sondas de humedad en macetas seleccionadas al azar para supervisar el contenido hídrico del suelo (CHS). Cuando el CHS está por debajo de determinados umbrales, las plantas vuelven a regarse automáticamente de manera continua hasta que de nuevo alcanzan un nivel normal. Después las plantas se vuelven a transferir a condiciones normales. El resto del cultivo (maduración de planta, cosecha de semilla) es igual que para las plantas que no se crecen en condiciones de estrés abiótico. Los parámetros de cultivo y rendimiento se registran como se detalla para el cultivo en condiciones normales.

Detección de estrés salino

20 Las plantas se crecieron en un sustrato hecho de fibras de coco y argex (proporción 3 a 1). Se usó una solución de nutrientes normal durante las dos primeras semanas después del trasplante de los plantones en el invernadero. Después de las dos primeras semanas, se añadió a la solución de nutrientes sal 25 mM (NaCl), hasta que las plantas se cosecharon. Para el crecimiento en condiciones normales, los parámetros de crecimiento y rendimiento se registraron como se detalla.

Listados de secuencias, SEC ID N° 179-243

- 25 <210> 179
- <211> 2046
- <212> ADN
- <213> *Oryza sativa*
- <400> 179

30	atgaataccc	cagaaaagaa	accactgtgc	tatacttcta	gaagagcact	gcaacagagg	60
	acagaatcct	cttcggaatt	gatctctgta	tcaaaaagag	caacccgcca	gaatactcca	120
	cgcaagccgg	attctcctcc	caaacgaaca	acaagaagta	gcgctaacct	ggcgaaatgt	180

ES 2 558 745 T3

atagagaaca aacaccatag cagtcctctg aagcggcggc ggggttcaga tgcggctacc 240
 ggaaagtctg caacagggcc aacaaggagg aagcataaac agaaaaggaa aatgatgag 300
 agtgatgagg tcagtcgcat ggaaaaaaga gcaagatatt tacttatcaa aataaacag 360
 gagcagaatt tgctagatgc atactctgga gatggttgga atgggcacag tcgagagaaa 420
 ataaagccag agaaggagct acagcgtgct aagaaacaga tcatgaaata caaatcgct 480
 atccgtgacg tcattcacca acttgatttg tgtagttcca gtgggagcaa ggatgactct 540
 gtgattccac cagatgggtg tcacgagtct gtcaatcctg aacatacaat atgttcaagg 600
 tgcaaatccc acgagtcatt tcctgacaat aacattatat tctgtgaggg gggctgcaaa 660
 ctggcgtgtc atcagaaatg tttggagcct ctttttgata aaattcttcc aactaccgcg 720
 catgggcggc tttgcaaaaca ctgttcttcc aagatgaaaa ttttagacgc cattaatgcc 780
 catctgggga caagttttac agtgaagtgt cctccagtg atattttcaa ggaagcagct 840
 gaacatttca actctgatga tggactgggc caagattggc tatctgagta ttcaggtgat 900
 gaggactatg accctgaaga aatgaggcc agcagcagtg gtgaagagaa taaatctgcc 960
 gattccaatt gttcagggag tccactttat tctccaaatg atgatattcc agacttcata 1020
 tcagcagatt tcaatgatgc tgaaggattc tgcctgtaa gttcaaatct aggaatcgat 1080
 tttggtgaag atggtttggc tgagattctt acccaccaaa ggccaagaag agatgttgac 1140
 tatacacaac ttaatgagca aatgtttgga gagcctattg gcaatgacga acagagtga 1200
 gatgaagatt ggggtcttaa caagagaaag aaaagaagaa ctggttcaac tgggtgttggg 1260
 actaattccg tagagggtcg atcagatgtc aaatccaata agaaagccca acctcggaga 1320
 aaacttttca ggattcctcc tgcggcagtt gaggtgctcc gtaaagcttt tgctgaaaat 1380
 gagcttctg cccggagtgt caaagaaaat ctctcaacag agctgggtat ttcttttgaa 1440
 aagattgata agtggttcaa aaatacacgg tgtgctgctc tcagggatag gaagggtgaa 1500
 agtcgttatt ctggtcccag caaaaggta agaacaagca tagaaaaggc tgaaacttca 1560
 gctaaagtgg atcaaatgga caactcatgc tttcttcctt tatctgaaat aatcaactg 1620
 cccacacgtc tgcagaaagg tcttgataag aagccaaat caattaacag ccctccaaga 1680
 cctcaggaca atgaaacttg ctgtcgcca actgataaaa ccaaggaggg tacaccgcct 1740
 acgatcaagc cttccatcac agattctagt caactgatga ataacgacat cggcactgag 1800
 gagactgctg tttcttgggt ggacacttgg gcctctgacg ctctgcattt cttggacgtg 1860
 agcgatgatg aacatttctt cgatgtgatc gagaagggtg gcggtctcga gaatcggctg 1920
 cagcggttga aggagaacat gctatcatcg tcatcgtcaa ctgacaacaa tgtggctgct 1980
 gagagtggct tgcaaacga agttgtgctt gtaccagctg ctgagctcaa ggataaagca 2040
 tcttaa 2046

5

<210> 180
 <211> 681
 <212> PRT
 <213> *Oryza sativa*
 <400> 180

ES 2 558 745 T3

Met Asn Thr Pro Glu Lys Lys Pro Leu Cys Tyr Thr Ser Arg Arg Ala
1 5 10 15

Leu Gln Gln Arg Thr Glu Ser Ser Ser Glu Leu Ile Ser Val Ser Lys
20 25 30

Arg Ala Thr Arg Gln Asn Thr Pro Arg Lys Pro Asp Ser Pro Pro Lys
35 40 45

Arg Thr Thr Arg Ser Ser Ala Asn Leu Ala Lys Cys Ile Glu Asn Lys
50 55 60

His His Ser Ser Pro Leu Lys Arg Arg Arg Gly Ser Asp Ala Ala Thr
65 70 75 80

Gly Lys Ser Ala Thr Gly Pro Thr Arg Arg Lys His Lys Gln Lys Arg
85 90 95

Lys Asn Asp Glu Ser Asp Glu Val Ser Arg Met Glu Lys Arg Ala Arg
100 105 110

Tyr Leu Leu Ile Lys Ile Lys Gln Glu Gln Asn Leu Leu Asp Ala Tyr
115 120 125

Ser Gly Asp Gly Trp Asn Gly His Ser Arg Glu Lys Ile Lys Pro Glu
130 135 140

Lys Glu Leu Gln Arg Ala Lys Lys Gln Ile Met Lys Tyr Lys Ile Ala
145 150 155 160

Ile Arg Asp Val Ile His Gln Leu Asp Leu Cys Ser Ser Ser Gly Ser
165 170 175

Lys Asp Asp Ser Val Ile Pro Pro Asp Gly Cys His Glu Ser Val Asn
180 185 190

Pro Glu His Thr Ile Cys Ser Arg Cys Lys Ser His Glu Ser Phe Pro
195 200 205

Asp Asn Asn Ile Ile Phe Cys Glu Gly Gly Cys Lys Leu Ala Cys His
210 215 220

Gln Lys Cys Leu Glu Pro Pro Phe Asp Lys Ile Leu Pro Thr Thr Arg
225 230 235 240

His Gly Arg Leu Cys Lys His Cys Ser Ser Lys Met Lys Ile Leu Asp
245 250 255

Ala Ile Asn Ala His Leu Gly Thr Ser Phe Thr Val Lys Cys Pro Ser
260 265 270

ES 2 558 745 T3

Ser Asp **Ile** **Phe** Lys Glu Ala **Ala** Glu His **Phe** Asn Ser Asp Asp Gly
 275 280 285

Leu Gly Gln Asp Trp Leu Ser **Glu** Tyr Ser Gly Asp Glu Asp Tyr Asp
 290 295 300

Pro Glu Glu Asn Glu **Ala** Ser Ser Ser Gly Glu Glu Asn Lys Ser Ala
 305 310 320

Asp Ser Asn Cys Ser **Gly** Ser Pro Leu Tyr Ser Pro Asn Asp Asp **Ile**
 325 330 335

Pro Asp **Phe** **Ile** Ser Ala Asp **Phe** Asn Asp Ala Glu Gly **Phe** Cys Arg
 340 345 350

Glu Ser Ser Asn Leu Gly **Ile** Asp **Phe** Gly Glu Asp Gly Leu Ala Glu
 355 360 365

Ile Leu Thr His Gln Arg Pro Arg Arg Asp Val Asp Tyr Thr Gln Leu
 370 375 380

Asn Glu Gln Met **Phe** Gly Glu Pro **Ile** Gly Asn Asp Glu Gln Ser Glu
 385 390 395 400

Asp Glu Asp Trp **Gly** Leu Asn Lys Arg **Lys** Lys Arg Arg Thr **Gly** Ser
 405 410 415

Thr Gly Val **Gly** Thr Asn Ser Val Glu Gly Arg Ser Asp Val Lys Ser
 420 425 430

Asn Lys **Lys** Ala Gln Pro Arg Arg Lys Leu **Phe** Arg **Ile** Pro Pro Ala
 435 440 445

Ala Val Glu Val Leu Arg **Lys** Ala **Phe** Ala Glu Asn Glu Leu Pro Ala
 450 455 460

Arg Ser Val Lys Glu Asn Leu Ser Thr Glu Leu Gly **Ile** Ser **Phe** Glu
 465 470 475 480

Lys **Ile** Asp Lys Trp **Phe** Lys Asn Thr Arg Cys Ala Ala Leu Arg Asp
 485 490 495

Arg Lys Gly Glu Ser Arg Tyr Ser **Gly** Pro Ser Lys Arg Ser Arg Thr
 500 505 510

Ser **Ile** Glu Lys Ala Glu Thr Ser Ala Lys Val Asp Gln Met Asp Asn
 515 520 525

Ser Cys **Phe** Leu Pro Leu Ser **Glu** **Ile** **Ile** Asn Val Pro Thr Arg Leu
 530 535 540

ES 2 558 745 T3

Gln Lys Gly Leu Asp Lys Lys Pro Lys Ser Ile Asn Ser Pro Pro Arg
 545 550 555 560

Pro Gln Asp Asn Glu Thr Cys Leu Ser Pro Thr Asp Lys Thr Lys Glu
 565 570 575

Gly Thr Pro Pro Thr Ile Lys Pro Ser Ile Thr Asp Ser Ser Gln Leu
 580 585 590

Met Asn Asn Asp Ile Gly Thr Glu Glu Thr Ala Val Ser Trp Val Asp
 595 600 605

Thr Trp Ala Ser Asp Ala Leu His Phe Leu Asp Val Ser Asp Asp Glu
 610 615 620

His Phe Phe Asp Val Ile Glu Lys Val Cys Gly Leu Glu Asn Arg Leu
 625 630 635 640

Gln Arg Leu Lys Glu Asn Met Leu Ser Ser Ser Ser Thr Asp Asn
 645 650 655

Asn Val Ala Ala Glu Ser Gly Leu Gln Asn Glu Val Val Leu Val Pro
 660 670

Ala Ala Glu Leu Lys Asp Lys Ala Ser
 675 680

<210> 181
 <211> 2391
 <212> ADN
 <213> *Arabidopsis thaliana*

<400> 181

atggaggaga gtgaaacaaa ggggagaatc agtcaagaaa ctgataaggc ttgtgtttca 60
 gttgagagaa ttgggtctac tttactctca tcattttgtga agaaaggtaa agaagtatca 120
 aataagagaa attctaagca gaataagcga aaagcggaag aagaactctg ttcgaaatca 180
 agaaccaaga agtactctcg gggatgggtc cgttgtgagg aaatggagga ggagaaggtt 240
 aagaaaacaa gaaagaggaa aagtaagaga cagcaaaagg ataataaggt tgaggttgat 300
 gattctttaa ggttgcaaag gaggacaagg tatttctca ttaagatgaa gatgcagcag 360
 aatcttatcg atgcatacgc aactgaaggt tggaaaggtc agagccggga aaagataaga 420
 ccagacaagg agcttgagag agcccggaaa gagatcttaa actgcaagct tggactacgc 480
 gacgccattc gccaaactgga tctgcttagt tctgtgggaa gcatggaaga gaaagtgata 540
 gcttcagatg gatctattca tcatgatcat atattttgtg cagaatgcaa ttctcgagaa 600
 gcttttccag acaatgatat aatactttgt gatggaacgt gtaatcgagc atttcaccag 660
 aaatgccttg accctccttt ggaaacagaa agcatacctc ccggagatca gggctggttt 720
 tgcaaat ttt gtgattgcaa aattgaaatt atcgacacaa tgaatgcaca aattggaact 780

5

10

ES 2 558 745 T3

catttccctg tggacagcaa ctggcaggat atcttcaatg aagaagctag tcttcctatt 840
 ggatctgaag ctacagttaa caatgaagca gattggcctt cggatgactc caaggatgat 900
 gattatgacc ctgaaatgag ggaaaacggt ggaggaaaca gcagcaacgt tagtggtgat 960
 ggtggtggtg ataacgatga agaaagcatt tcaactagct tgagcttatc atctgatggt 1020
 gttgcccttt caacgggatc atgggaaggc catagactca gcaatatggt ggagcagtgc 1080
 gaaacaagta atgaagaaac tgtatgtggg ccaagacagc gaaggactgt tgactataca 1140
 cagttgtact atgaaatggt tggaaaggat gcagttttgc aagagcaagg tagtgaagat 1200
 gaagactggg gtccgaatga cagaaggaag agaaaaaggg aatctgatgc aggaagcaca 1260
 cttgtaacta tgtgtgaaag cagtaagaaa gatcaagatg ttgtagaaac gctagaacag 1320
 agtgagagag attctgtctc tgtggaaaat aaaggaggtc gaaggcggat gttcagactc 1380
 ccaagaaatg cagttgagaa gctacgtcaa gtattttgcag agactgagct tccctcaaaa 1440
 gctgtgaggg atcgtcttgc aaaagagctg agtcttgatc cagagaaggc caataaatgg 1500
 tttaaaaata ctcggtatat ggcactgcgg aataggaaga ctgagagcgt gaaacaacct 1560
 ggggattcta agacagtctc tggaggagat tctggaccag aagcagtcac ggagaacaac 1620
 actgaaacaa atgaagtcca agatactttg gatgacactg ttccaccggg atttgacgct 1680
 acaaatcaaa acatactctc cccatgtaat aacaaccagg aggaatttca acaggaaaat 1740
 gtctcttttc cttcaccac agatgaaagt caacagtatc tggaacaaaa cgattcttct 1800
 ttcgttctag tgccacatga aaagcaaagc agcgaaatta gcttaaagac ggctgtagag 1860
 gagaatgaga cagaaagcaa gatgatgaaa gagccacatg aagagctaag cagtgaaatg 1920
 agcctaaaga cggctgctga ggagaaggag acagagagca agatgataga agagccacat 1980
 gaagagctaa gcagagaaat gagcttaaag acggctgtag aggagaagga gacagagagc 2040
 aagatgatgg aagagccaca tgatgagcta aacagcgaaa tgagcttaag tacagcggta 2100
 gaggagaagg agacagggag caagatgacg gaagagtcac atgaagagct aagcaacgaa 2160
 atgagcttag aagagaagga gacggggaga aagatgacgg aagaggaaga attggaagca 2220
 gtaatggaga tgctttgtag aacagaaaac aagttactgg acgtgacgca gaggcttgat 2280
 agatttaaaa caccaaaagg ccgaaaaaag ttaggcaatt cttcttctcc tttacttgaa 2340
 gaagactcag tagtgtatgt tccaatagca gagatcaagg agaaaaggta a 2391

<210> 182
 <211> 796
 <212> PRT
 <213> *Arabidopsis thaliana*

5

<400> 182

Met Glu Glu Ser Glu Thr Lys Gly Arg Ile Ser Gln Glu Thr Asp Lys
 1 5 10 15

Ala Cys Val Ser Val Glu Arg Ile Gly Ser Thr Leu Leu Ser Ser Phe
 20 25 30

10

ES 2 558 745 T3

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

ES 2 558 745 T3

Glu Met Arg Glu Asn Gly Gly Gly Asn Ser Ser Asn Val Ser Gly Asp
 305 310 315 320
 Gly Gly Gly Asp Asn Asp Glu Glu Ser Ile Ser Thr Ser Leu Ser Leu
 325 330 335
 Ser Ser Asp Gly Val Ala Leu Ser Thr Gly Ser Trp Glu Gly His Arg
 340 345 350
 Leu Ser Asn Met Val Glu Gln Cys Glu Thr Ser Asn Glu Glu Thr Val
 355 360 365
 Cys Gly Pro Arg Gln Arg Arg Thr Val Asp Tyr Thr Gln Leu Tyr Tyr
 370 375 380
 Glu Met Phe Gly Lys Asp Ala Val Leu Gln Glu Gln Gly Ser Glu Asp
 385 390 400
 Glu Asp Trp Gly Pro Asn Asp Arg Arg Lys Arg Lys Arg Glu Ser Asp
 405 410 415
 Ala Gly Ser Thr Leu Val Thr Met Cys Glu Ser Ser Lys Lys Asp Gln
 420 425 430
 Asp Val Val Glu Thr Leu Glu Gln Ser Glu Arg Asp Ser Val Ser Val
 435 440 445
 Glu Asn Lys Gly Gly Arg Arg Arg Met Phe Arg Leu Pro Arg Asn Ala
 450 455 460
 Val Glu Lys Leu Arg Gln Val Phe Ala Glu Thr Glu Leu Pro Ser Lys
 465 470 475 480
 Ala Val Arg Asp Arg Leu Ala Lys Glu Leu Ser Leu Asp Pro Glu Lys
 485 490 495
 Val Asn Lys Trp Phe Lys Asn Thr Arg Tyr Met Ala Leu Arg Asn Arg
 500 505 510
 Lys Thr Glu Ser Val Lys Gln Pro Gly Asp Ser Lys Thr Val Ser Gly
 515 520 525
 Gly Asp Ser Gly Pro Glu Ala Val Met Glu Asn Asn Thr Glu Thr Asn
 530 535 540
 Glu Val Gln Asp Thr Leu Asp Asp Thr Val Pro Pro Gly Phe Asp Ala
 545 550 555 560
 Thr Asn Gln Asn Ile Leu Ser Pro Cys Asn Asn Asn Gln Glu Glu Phe
 565 570 575
 Gln Gln Glu Asn Val Ser Phe Pro Ser Pro Thr Asp Glu Ser Gln Gln

ES 2 558 745 T3

580 585 590
 Tyr Leu Glu Gln Asn Asp Ser Ser Phe Val Leu Val Pro His Glu Lys
 595 600 605
 Gln Ser Ser Glu Ile Ser Leu Lys Thr Ala Val Glu Glu Asn Glu Thr
 610 615 620
 Glu Ser Lys Met Met Lys Glu Pro His Glu Glu Leu Ser Ser Glu Met
 625 630 635 640
 Ser Leu Lys Thr Ala Ala Glu Glu Lys Glu Thr Glu Ser Lys Met Ile
 645 650 655
 Glu Glu Pro His Glu Glu Leu Ser Arg Glu Met Ser Leu Lys Thr Ala
 660 665 670
 Val Glu Glu Lys Glu Thr Glu Ser Lys Met Met Glu Glu Pro His Asp
 675 680 685
 Glu Leu Asn Ser Glu Met Ser Leu Ser Thr Ala Val Glu Glu Lys Glu
 690 695 700
 Thr Gly Ser Lys Met Thr Glu Glu Ser His Glu Glu Leu Ser Asn Glu
 705 710 715 720
 Met Ser Leu Glu Glu Lys Glu Thr Gly Arg Lys Met Thr Glu Glu Glu
 725 730 735
 Glu Leu Glu Ala Val Met Glu Met Leu Cys Arg Thr Glu Asn Lys Leu
 740 745 750
 Leu Asp Val Thr Gln Arg Leu Asp Arg Phe Lys Thr Pro Lys Gly Arg
 755 760 765
 Lys Lys Leu Gly Asn Ser Ser Ser Pro Leu Leu Glu Glu Asp Ser Val
 770 775 780
 Val Tyr Val Pro Ile Ala Glu Ile Lys Glu Lys Arg
 785 790 795

<210> 183
 <211> 2151
 <212> ADN
 <213> *Eucalyptus grandis*
 <400> 183

5

ES 2 558 745 T3

atgcattcca gaaagaatct tgcccaaaaa gaagactttt tgaaggcaga atctggatcg 60
gaggttcttg cctcgtgat attgagaaat aggagaagga aaaagcacia acagaaacta 120
caatcttctg taaaagcaag aggctcaacc cttgccaaga agagggccaa taactccatc 180
agaaagagac tgggtgcgtag gaacagctct gataaagaca aactcataaa tttaaagct 240
tttcataata agaaaattga aatagttccc gagaagcttt catcccctcc atcttcagaa 300
agagagctgc ctattgcttc tgacgaaaat gtgcaaaaca ccgatgagga tgtcaaagtt 360
aacaagctga gaaggaagag gaagaagaaa atgagaaagc atttttcgaa gttagatgag 420
gcttcacgat tgcagaggag aacaagatac cttcttatta aatgaagct ggaacagaac 480
cttattgatg cttactctgg tgaaggctgg aaaggccaga gtcgggaaaa gatcaagcca 540
acaaaggaat tacaaagggc caggaagcag atactgaatt gcaaactcgg aatacgtgat 600
gctattcgac agctagactc cattagttca gtgggcagca ttgaagactc tgtgattgct 660
cctgatggtt ctgttcacca tgaacatata ttctgtgcaa aatgcaagac aatgaaact 720
tcctcagata atgatatcat actctgcgat gggacatgca attgtgcttt tcatcaaaaa 780
tgtcttgatc ctccactgga aactgaaagc attccccag gagatcaggg ctggttttgc 840
agactctgcg agtgaagat ggaaattata gagtcaacaa atgcacatct ggggaccac 900
ttttctcttg acagcaactg gcaggatata ttcgaggaag aagctgcatt gcctgatggg 960
ggaattcagg tgcttgacct agacgaaaag tggccttcag atgactctga agatgatgac 1020
tatgaccag aaaagatgac gaacaccggc atcaactgtg ccggcaatga tgattctcaa 1080
tcatacgaca caagtagctc taccagcctg agttggctat tagatgggta aatttgtca 1140
ggaccaacaa tatcaggcaa ggaaggggtg aactatcctg agagcattgc aggtctgat 1200
gtattaatgg acacggatac tgatactgta tgtggcccac gacagcgaag tgctgtgat 1260
tataagcagc ttataatga aatgttcggt aaggatactc ctgcctgtga acaagccagt 1320
gaagatgaag actggggtcc taataaaaga agaagaagag aaagggaaac agatgcagcc 1380
agcacctca tgactctctg tgaaagtgag aaaaacaatc cattcattga ccaagcctca 1440
gaaggtgata agaaactctc tgcagaagca agaatgagaa ggcggatfff cagaattcct 1500
ccaactgcag ttgagaagct tgcgcaagta ttgcagaga atgaacttcc ctcaagatct 1560
gtcaaagaaa atcttgctaa agagtttggg cttgaacctg agaaggtag caaatggttc 1620
aagaatgcac gatacatggc actcaaaaca agaaaggctg ggggggctga tcaacctcca 1680
aaatccccgc aaaagcataa agaacttgaa ctagcaaaaa tgatgaaagg ggctgcgct 1740
attttatcta aggatgcctc aacagacatt gaattccatt cctctctaaa tactgagaaa 1800
gtcatccagc aaaagagtac gaagccaata cacgggaagg gttcacctat ggtacctgaa 1860
catagctata tggagaatff tgagttcgat gacaatatgt cttgaaaca aataaagcta 1920
ttgaaaggaa aaccaaaagga caggaagagg gttgacttcg cagccagtga gggattacga 1980
gaaagagaag ctcagttgag ggagctatgt aaaattagaa gtaggctaga tagtctgagg 2040
aagagaacct ttacatttca gtcgagttaa tccaaggctt cagacgaatt caacacttgt 2100
agatcacgtg tagtatatgt tccagtagca gaaataaagg agagagcttg a 2151

<210> 184
<211> 716
<212> PRT
<213> *Eucalyptus grandis*

5

<400> 184

ES 2 558 745 T3

Met His Ser Arg Lys Asn Leu Ala Gln Lys Glu Asp Phe Leu Lys Ala
1 5 10 15

Glu Ser Gly Ser Glu Val Leu Ala Ser Leu Ile Leu Arg Asn Arg Arg
20 25 30

Arg Lys Lys His Lys Gln Lys Leu Gln Ser Ser Val Lys Ala Arg Gly
35 40 45

Ser Thr Leu Ala Lys Lys Arg Ala Asn Asn Ser Ile Arg Lys Arg Leu
50 55 60

Val Arg Arg Asn Ser Ser Asp Lys Asp Lys Leu Ile Asn Leu Lys Ala
65 70 75 80

Phe His Asn Lys Lys Ile Glu Ile Val Pro Glu Lys Leu Ser Ser Pro
85 90 95

Pro Ser Ser Glu Arg Glu Leu Pro Ile Ala Ser Asp Glu Asn Val Gln
100 105 110

Asn Thr Asp Glu Asp Val Lys Val Asn Lys Leu Arg Arg Lys Arg Lys
115 120 125

Lys Lys Met Arg Lys His Phe Ser Lys Leu Asp Glu Ala Ser Arg Leu
130 135 140

Gln Arg Arg Thr Arg Tyr Leu Leu Ile Lys Met Lys Leu Glu Gln Asn
145 150 155 160

Leu Ile Asp Ala Tyr Ser Gly Glu Gly Trp Lys Gly Gln Ser Arg Glu
165 170 175

Lys Ile Lys Pro Thr Lys Glu Leu Gln Arg Ala Arg Lys Gln Ile Leu
180 185 190

Asn Cys Lys Leu Gly Ile Arg Asp Ala Ile Arg Gln Leu Asp Ser Ile
195 200 205

Ser Ser Val Gly Ser Ile Glu Asp Ser Val Ile Ala Pro Asp Gly Ser
210 215 220

Val His His Glu His Ile Phe Cys Ala Lys Cys Lys Thr Asn Glu Thr
225 230 235 240

Ser Ser Asp Asn Asp Ile Ile Leu Cys Asp Gly Thr Cys Asn Cys Ala
245 250 255

Phe His Gln Lys Cys Leu Asp Pro Pro Leu Glu Thr Glu Ser Ile Pro
260 265 270

ES 2 558 745 T3

Pro Gly Asp Gln Gly Trp Phe Cys Arg Leu Cys Glu Cys Lys Met Glu
 275 280 285

Ile Ile Glu Ser Thr Asn Ala His Leu Gly Thr His Phe Ser Leu Asp
 290 295 300

Ser Asn Trp Gln Asp Ile Phe Glu Glu Glu Ala Ala Leu Pro Asp Gly
 305 310 315 320

Gly Ile Gln Val Leu Asp Leu Asp Glu Lys Trp Pro Ser Asp Asp Ser
 325 330 335

Glu Asp Asp Asp Tyr Asp Pro Glu Lys Met Thr Asn Thr Gly Ile Asn
 340 345 350

Cys Ala Gly Asn Asp Asp Ser Gln Ser Tyr Asp Thr Ser Ser Ser Thr
 355 360 365

Ser Leu Ser Trp Ser Leu Asp Gly Glu Ile Leu Ser Gly Pro Thr Ile
 370 375 380

Ser Gly Lys Glu Gly Leu Asn Tyr Pro Glu Ser Ile Ala Gly Ser Asp
 385 390 395 400

Val Leu Met Asp Thr Asp Thr Asp Thr Val Cys Gly Pro Arg Gln Arg
 405 410 415

Asn Ala Val Asp Tyr Lys Gln Leu Tyr Asn Glu Met Phe Gly Lys Asp
 420 425 430

Thr Pro Ala Cys Glu Gln Ala Ser Glu Asp Glu Asp Trp Gly Pro Asn
 435 440 445

Lys Arg Arg Arg Arg Glu Arg Glu Thr Asp Ala Ala Ser Thr Leu Met
 450 455 460

Thr Leu Cys Glu Ser Glu Lys Asn Asn Pro Phe Ile Asp Gln Ala Ser
 465 470 475 480

Glu Gly Asp Lys Lys Leu Ser Ala Glu Ala Arg Met Arg Arg Arg Ile
 485 490 495

Phe Arg Ile Pro Pro Thr Ala Val Glu Lys Leu Arg Gln Val Phe Ala
 500 505

Glu Asn Glu Leu Pro Ser Arg Ser Val Lys Glu Asn Leu Ala Lys Glu
 515 520 525

Phe Gly Leu Glu Pro Glu Lys Val Ser Lys Trp Phe Lys Asn Ala Arg
 530 535 540

ES 2 558 745 T3

Tyr Met Ala Leu Lys Thr Arg Lys Ala Gly Gly Ala Asp Gln Pro Pro
 545 550 555 560
 Lys Ser Pro Gln Lys His Lys Glu Ser Glu Leu Ala Lys Met Met Lys
 565 570 575
 Gly Ala Ala Pro Ile Leu Ser Lys Asp Ala Ser Thr Asp Ile Glu Phe
 580 585 590
 His Ser Ser Leu Asn Thr Glu Lys Val Ile Gln Gln Lys Ser Thr Lys
 595 600 605
 Pro Ile His Gly Lys Gly Ser Pro Met Val Pro Glu His Ser Tyr Met
 610 615 620
 Glu Asn Phe Glu Phe Asp Asp Asn Met Ser Leu Lys Gln Ile Lys Leu
 625 630 635 640
 Leu Lys Gly Lys Pro Lys Asp Arg Lys Arg Val Asp Phe Ala Ala Ser
 645 650 655
 Glu Gly Leu Arg Glu Arg Glu Ala Gln Leu Arg Glu Leu Cys Lys Ile
 660 665 670
 Arg Ser Arg Leu Asp Ser Leu Arg Lys Arg Thr Leu Thr Phe Gln Ser
 675 680 685
 Ser Glu Ser Lys Ala Ser Asp Glu Phe Asn Thr Cys Arg Ser Arg Val
 690 695 700
 Val Tyr Val Pro Val Ala Glu Ile Lys Glu Arg Ala
 705 710 715

<210> 185
 <211> 2124
 <212> ADN
 <213> *Medicago truncatula*
 <400> 185

5

ES 2 558 745 T3

atgcgggata	ctgagagggt	aaataaccaa	ggatctacaa	aaccgagtaa	cacaaaggaa	60
catatcaagt	tgaaggtaga	ttccccacag	ttcaaaacaa	gtagcacaaa	acaccaaagg	120
aagaaacaca	gactgaaatc	aaaatctcat	aaactcggag	gttgcacaaa	tgcgtcagtg	180
aggacagcta	ctgattcttc	caacaaggcg	tcaaccaagg	attcttcaaa	taaaacggat	240
aggaattcaa	cacaggttca	gccttcaaaa	aagatacaag	gaggaaaaac	ttctcttaat	300
aatgatagga	aaggtgagaa	agatgttggt	gatcaagagg	gaaatattca	aaaacgtaag	360
agaaggagaa	agaagaaaag	acaagacat	aatgtagatc	ttgatgatac	tgtacgcctc	420
cagagaagaa	caaggaatat	cttaatcaga	atgaagcaag	aacagaatct	tatcgatgct	480
tacgctggag	aaggttgtaa	aggtcagagt	cgagagaaga	ttaggccaga	aatggagcta	540

ES 2 558 745 T3

caaagagcaa aaaagcagat tttaaaatgc aagcttagta tccgtgatgc cattcaccaa 600
 ctggattctc ttagttcagt tggtagtatt gaaggttctg tcattgctac agatggatct 660
 gtttctcatg aacatatatt ctgtgcaaat tgcaagataa acgaagtttc tccagataat 720
 gatattatac tctgtgatgg cacatgcaat cgtgctgttc accagagatg tcttgatcct 780
 cctttggaga ctgaagacat tctcctgaa gaccaaggct ggttttgcaa gtattgtgac 840
 tgtaagatag aatatattgga ggcaacaaac gcgcatcttg ggactcgttt ccccttagac 900
 agtacttggc aggatgtatt caaggaagaa gctgctattc ctgacggcga tgctgcatta 960
 ctaaatacaag aagaagaatg gccgtcagat gatcctgaag acgatgacta taatccagag 1020
 aggaaagaag aaagccacgg tggcttcaac acagaaggaa atgataaaaa cgcattctgat 1080
 gattcaagca gttcttccag catgtggtct ctaaattggag aatgctctct gttagatgaa 1140
 ggcatcaacc ttgaatatta ttctaacgac cacatagatt ctgatgaatc cggggaaata 1200
 gcatgcggtc gtaggcagcg tagagctggt gactacaaga aactttatga tgaaatgttt 1260
 ggaaaggatg ctccaccttg tgaacaagtg agtgaagatg aagattgggg tctcga meta 1320
 agacggcga gagaaaagga gtctgaggct gttaatactc ttatgactct acatgagagt 1380
 gaaaataaat atccgaataa taagaacaat gatagaataa gaggaattc ttcaggcata 1440
 aaaagacctt gttttagggt ttcacatgag gcagttgaga agcttcgcca agtttttgca 1500
 gagaacgagc ttcctccaaa atccgtgaag gatgcccttt caaaagaatt gggacttgat 1560
 gctgcaaagg ttaacaaatg gttcaaaaat gcacgttaca tggcacttaa aataagaaag 1620
 cttcaagaag gaggacaaca gcttcagagt atcacttcta agacctcaaa ggattctaca 1680
 tcacaacatg tgcaggaaga tgaggttttg aaccggaagt ccgcaaaaat tactgtgatt 1740
 agttccctga agaagtgtga aaatgttact ggcaaagaaa aaacaaaagc atctagcaat 1800
 cctttgaaga aaaagcgacc agaaatacct ccatgcttgg gagaaaatgg caacaaggat 1860
 tccatggagg tcagtgatga tgtttagctta atgaaactat tacaagacag gaaaaagaag 1920
 gtaagttttg catttgaagg agactctgag gcagcagagt tggaaatttga aagactcagc 1980
 aaagtgaaga cgaaaaataga tagattgaag caagattaa ctggagttca aaattgcaga 2040
 tcaaaggggt cagatgaagt tcatttgaat gagccatcta ttatgtatgt tcctatagct 2100
 gagttaaggg aaaaggtgaa atag 2124

<210> 186
 <211> 707
 <212> PRT
 <213> *Medicago truncatula*

<400> 186

Met Arg Asp Thr Glu Arg Leu Asn Asn Gln Gly Ser Thr Lys Pro Ser
 1 5 10 15

Asn Thr Lys Glu His Ile Lys Leu Lys Val Asp Ser Pro Gln Phe Lys
 20 25 30

5

10

ES 2 558 745 T3

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

ES 2 558 745 T3

Asp Val Phe Lys Glu Glu Ala Ala Ile Pro Asp Gly Asp Ala Ala Leu
 305 310 315 320
 Leu Asn Gln Glu Glu Glu Trp Pro Ser Asp Asp Pro Glu Asp Asp Asp
 325 330 335
 Tyr Asn Pro Glu Arg Lys Glu Glu Ser His Gly Gly Phe Asn Thr Glu
 340 345 350
 Gly Asn Asp Lys Asn Ala Ser Asp Asp Ser Ser Ser Ser Ser Ser Met
 355 360 365
 Trp Ser Leu Asn Gly Glu Cys Ser Leu Leu Asp Glu Gly Ile Asn Leu
 370 375 380
 Glu Tyr Tyr Ser Asn Asp His Ile Asp Ser Asp Glu Ser Gly Glu Ile
 385 390 395 400
 Ala Cys Gly Arg Arg Gln Arg Arg Ala Val Asp Tyr Lys Lys Leu Tyr
 405 410 415
 Asp Glu Met Phe Gly Lys Asp Ala Pro Pro Cys Glu Gln Val Ser Glu
 420 425 430
 Asp Glu Asp Trp Gly Pro Arg Lys Arg Arg Arg Arg Glu Lys Glu Ser
 435 440 445
 Glu Ala Val Asn Thr Leu Met Thr Leu His Glu Ser Glu Asn Lys Tyr
 450 455 460
 Pro Asn Asn Lys Asn Asn Asp Arg Ile Arg Gly Asn Ser Ser Gly Ile
 465 470 475 480
 Lys Arg Pro Cys Phe Arg Phe Ser His Glu Ala Val Glu Lys Leu Arg
 485 490 495
 Gln Val Phe Ala Glu Asn Glu Leu Pro Pro Lys Ser Val Lys Asp Ala
 500 505 510
 Leu Ser Lys Glu Leu Gly Leu Asp Ala Ala Lys Val Asn Lys Trp Phe
 515 520 525
 Lys Asn Ala Arg Tyr Met Ala Leu Lys Ile Arg Lys Leu Gln Glu Gly
 530 535 540
 Gly Gln Gln Leu Gln Ser Ile Thr Ser Lys Thr Ser Lys Asp Ser Thr
 545 550 555 560
 Ser Gln His Val Gln Glu Asp Glu Val Leu Asn Pro Lys Ser Ala Lys
 565 570 575
 Ile Thr Val Ile Ser Ser Leu Lys Lys Cys Glu Asn Val Thr Gly Lys

ES 2 558 745 T3

			580					585						590		
Glu	Lys	Thr	Lys	Ala	Ser	Ser	Asn	Pro	Leu	Lys	Lys	Lys	Arg	Pro	Glu	
		595					600					605				
Ile	Pro	Pro	Cys	Leu	Gly	Glu	Asn	Gly	Asn	Lys	Asp	Ser	Met	Glu	Val	
	610					615					620					
Ser	Asp	Asp	Val	Ser	Leu	Met	Lys	Leu	Leu	Gln	Asp	Arg	Lys	Lys	Lys	
625					630					635					640	
Val	Ser	Phe	Ala	Phe	Glu	Gly	Asp	Ser	Glu	Ala	Ala	Glu	Leu	Glu	Phe	
				645					650					655		
Glu	Arg	Leu	Ser	Lys	Val	Lys	Thr	Lys	Ile	Asp	Arg	Leu	Lys	Gln	Arg	
			660					665					670			
Leu	Thr	Gly	Val	Gln	Asn	Cys	Arg	Ser	Lys	Gly	Ser	Asp	Glu	Val	His	
		675					680					685				
Leu	Asn	Glu	Pro	Ser	Ile	Met	Tyr	Val	Pro	Ile	Ala	Glu	Leu	Arg	Glu	
	690					695					700					
Lys	Val	Lys														
705																

- 5
- <210> 187
 - <211> 2556
 - <212> ADN
 - <213> *Pinus radiata*
 - <400> 187

ES 2 558 745 T3

atgccgaaag caggaaaagg ctcttccggt tctctccgaa agaagcaagc gttgtcaaat 60
 acagcatcag agctaaggcc taaagctgca attcagaaga ctaacaatgc cgggagtatt 120
 tctggcagga tgagggaaag gaaaccacgg ttacatgtac agatgattgc ctcaattctc 180
 acgaaaaaaaa ggacaggtga agcatctgaa tcgagaaatg attctgagct agaagacaaa 240
 caatatctga gcaatagttc aaaaacaaat gccaatTTTg tagcggggaa gttagttaaa 300
 aagactttga aaaaagcccc aagattaaac ttgattaaga atgacagtca gaaaaaacc 360
 agtaagagtt ctgcttcgag gtcttatgct gcacagaaac ctgaacggcc acaacagaga 420
 aaaattagga ggacaaagtg gaagggtgat tctaaaaata taacaaaggT gaaaactgtg 480
 agaactagta atatagatgg cttgaagagg aagcagcaaa ggagaaagaa gaaaaagaag 540
 gcgcagaatg ttccacagga tgaagtatca cgtgcaaaga gaaggatcaa atatcttttg 600
 ctcaaaatga aacttgagca gaatcttatt gatgcatatt caggggaggg ctggaaaggc 660
 cagagtcggg agaaaataaa gccagagaag gagttacaga gagctgaaaa gcagattctg 720
 caatgcaaac ttggaattcg tgaagcagtg catcagcttg atcttctaag ttcggaagga 780
 tctatagcag atgccgttat tgatccagaa gggcgtgtat ttcatgagca tatatTTTgt 840

ES 2 558 745 T3

gcaaaatgca aaatgacaga tgctcttccg gataatgaca taatattgtg tgatggagct 900
tgcaatcgtg gcttccatca gaaatgtctg gatccacctt tggctactga aaacattcct 960
cctggagatc agggttggct ctgtaaagt tgcgaatgta agctagagag tttggaagca 1020
ataaatgctc atttgggcac ccagtttaca gttgataata gctgggaggg aatttttgct 1080
gatgcagcta gaattgcgaa tggtgaaagt acagctactg ggaatgggga agaatggcca 1140
tcagatgatt cagaagatga tgactatgat ccagagaaac aagagattag agatgaaacg 1200
ggctatgaga ataatatgag tgattctgga aagtcaagtg gatggtcaga atcatataat 1260
gaaactgaca atgaatcaga tgcttcctta acagataatc taaatgaata catctctaag 1320
ggaaagaaga ggagcagaga cactgatgaa aaaaacggac aagaagctat aatagattta 1380
gatactgatg agaatgatgg ctacagacatg ccagtttcag ggagaaggca gcgaagagat 1440
gttgattata aaaagttgca tgatgaaatg tttggaaaag atggaccttg tgaagatgaa 1500
gtcagtgaag atgaagactg gggacccaat aggaggcgta gaagagcgaa acagccggat 1560
acaaatacca ctaaggattc ttcaaatgaa gagaaagagc atgttagaaa ggaagcagat 1620
gagaaagtag cagagcattc acaagtagc ataaaggata agaaacatat tgtcagactg 1680
ccagctaatg cagttgagaa gctgcgcaaa gtctttgctg aaaatgagct accgtcaagg 1740
tcttacaagg aaaatctgtc acagcaactg ggccttacat ttcgaaaggt ccatatgtgg 1800
ttcaaaaatg cccgttatat gtctctgaag agccgaaagg aatgcctag aatgagaac 1860
tgcaatgaaa ctaatctcat gggagaaggt gatatggggc caaagaaaaa aataccaaaa 1920
gagaagaaag cagaatcaac atgtcttcta tcttctgtta gagcagaata taaatctaga 1980
aatacaaaaa ggaagcatag gagaaaaaac atcaaagcac ccttatctat ctcttccaga 2040
agtgagagaa aaagaactat gtgcagactg cctgccaatg cagttgagag gtttcgccaa 2100
gtctttgctg agaatgagtt gccttcaatg tctgtaaaaa tggcattatc aaagcagttg 2160
gatctacctt atcggcaggt gcatatgtgg ttcaaaaatg ctcgttatat gtctttgaag 2220
aagaaaaaga tgaaccgttt aaaagaatca gctcccgtta tttcctcaca taccagaaga 2280
cttgaagtag aaagtagcaa acccaatcca caaatcaag agccagatga tgctggcaat 2340
tcatacttaa cagagatgga aaaattgggt aatattgagt tgaaacttga aaatatgaga 2400
aagattttag aaacagtctg tcctaaaaga ggtactgatt tttccaatga taagaatcaa 2460
gaaggaatcg ccaaccttct tagtgaggag cttcttgtgt atgtaccagt ggtggagttg 2520
agggagaaga gcttacaggt tgctcctgtc aggtga 2556

<210> 188
<211> 851
5 <212> PRT
<213> *Pinus radiata*
<400> 188

Met Pro Lys Ala Gly Lys Gly Ser Ser Gly Ser Leu Arg Lys Lys Gln
1 5 10 15

10

ES 2 558 745 T3

Ala Leu Ser Asn Thr Ala Ser Glu Leu Arg Pro Lys Ala Ala Ile Gln
 20 25 30

Lys Thr Asn Asn Ala Gly Ser Ile Ser Gly Arg Met Arg Glu Arg Lys
 35 40 45

Pro Arg Leu His Val Gln Met Ile Ala Ser Ile Leu Thr Lys Lys Arg
 50 55 60

Thr Gly Glu Ala Ser Glu Ser Arg Asn Asp Ser Glu Leu Glu Asp Lys
 65 70 75 80

Gln Tyr Leu Ser Asn Ser Ser Lys Thr Asn Ala Asn Phe Val Ala Gly
 85 90 95

Lys Leu Val Lys Lys Thr Leu Lys Lys Ala Pro Arg Leu Asn Leu Ile
 100 105 110

Lys Asn Asp Ser Gln Lys Lys Pro Ser Lys Ser Ser Ala Ser Arg Ser
 115 120 125

Tyr Ala Ala Gln Lys Pro Glu Arg Pro Gln Gln Arg Lys Ile Arg Arg
 130 135 140

Thr Lys Trp Lys Gly Asp Ser Lys Asn Ile Thr Lys Val Lys Thr Val
 145 150 155 160

Arg Thr Ser Asn Ile Asp Gly Leu Lys Arg Lys Gln Gln Arg Arg Lys
 165 170 175

Lys Lys Lys Lys Ala Gln Asn Val Pro Gln Asp Glu Val Ser Arg Ala
 180 185 190

Lys Arg Arg Ile Lys Tyr Leu Leu Leu Lys Met Lys Leu Glu Gln Asn
 195 200 205

Leu Ile Asp Ala Tyr Ser Gly Glu Gly Trp Lys Gly Gln Ser Arg Glu
 210 215 220

Lys Ile Lys Pro Glu Lys Glu Leu Gln Arg Ala Glu Lys Gln Ile Leu
 225 230 235 240

Gln Cys Lys Leu Gly Ile Arg Glu Ala Val His Gln Leu Asp Leu Leu
 245 250 255

Ser Ser Glu Gly Ser Ile Ala Asp Ala Val Ile Asp Pro Glu Gly Arg
 260 265 270

Val Phe His Glu His Ile Phe Cys Ala Lys Cys Lys Met Thr Asp Ala
 275 280 285

ES 2 558 745 T3

Leu Pro Asp Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asn Arg Gly
 290 295 300
 Phe His Gln Lys Cys Leu Asp Pro Pro Leu Ala Thr Glu Asn Ile Pro
 305 310 315 320
 Pro Gly Asp Gln Gly Trp Leu Cys Lys Val Cys Glu Cys Lys Leu Glu
 325 330 335
 Ser Leu Glu Ala Ile Asn Ala His Leu Gly Thr Gln Phe Thr Val Asp
 340 345 350
 Asn Ser Trp Glu Gly Ile Phe Ala Asp Ala Ala Arg Ile Ala Asn Gly
 355 360 365
 Glu Ser Thr Ala Thr Gly Asn Gly Glu Glu Trp Pro Ser Asp Asp Ser
 370 380
 Glu Asp Asp Asp Tyr Asp Pro Glu Lys Gln Glu Ile Arg Asp Glu Thr
 385 390 400
 Gly Tyr Glu Asn Asn Met Ser Asp Ser Gly Lys Ser Ser Gly Trp Ser
 405 410 415
 Glu Ser Tyr Asn Glu Thr Asp Asn Glu Ser Asp Ala Ser Leu Thr Asp
 420 425 430
 Asn Leu Asn Glu Tyr Ile Ser Lys Gly Lys Lys Arg Ser Arg Asp Thr
 435 440 445
 Asp Glu Lys Asn Gly Gln Glu Ala Ile Ile Asp Leu Asp Thr Asp Glu
 450 455 460
 Asn Asp Gly Ser Asp Met Pro Val Ser Gly Arg Arg Gln Arg Arg Asp
 465 470 475 480
 Val Asp Tyr Lys Lys Leu His Asp Glu Met Phe Gly Lys Asp Gly Pro
 485 490 495
 Cys Glu Asp Glu Val Ser Glu Asp Glu Asp Trp Gly Pro Asn Arg Arg
 500 505 510
 Arg Arg Arg Ala Lys Gln Pro Asp Thr Asn Thr Thr Lys Asp Ser Ser
 515 520 525
 Asn Glu Glu Lys Glu His Val Arg Lys Glu Ala Asp Glu Lys Val Ala
 530 535 540
 Glu His Ser Gln Ser Ser Ile Lys Asp Lys Lys His Ile Val Arg Leu
 545 550 555 560
 Pro Ala Asn Ala Val Glu Lys Leu Arg Lys Val Phe Ala Glu Asn Glu

ES 2 558 745 T3

				565						570					575			
Leu	Pro	Ser	Arg	Ser	Tyr	Lys	Glu	Asn	Leu	Ser	Gln	Gln	Leu	Gly	Leu			
			580					585					590					
Thr	Phe	Arg	Lys	Val	His	Met	Trp	Phe	Lys	Asn	Ala	Arg	Tyr	Met	Ser			
		595					600					605						
Leu	Lys	Ser	Arg	Lys	Glu	Met	Pro	Arg	Asn	Glu	Asn	Cys	Asn	Glu	Thr			
	610					615					620							
Asn	Leu	Met	Gly	Glu	Gly	Asp	Met	Gly	Pro	Lys	Lys	Lys	Ile	Pro	Lys			
625					630					635					640			
Glu	Lys	Lys	Ala	Glu	Ser	Thr	Cys	Leu	Leu	Ser	Ser	Val	Arg	Ala	Glu			
				645					650					655				
Tyr	Lys	Ser	Arg	Asn	Thr	Lys	Arg	Lys	His	Arg	Arg	Lys	Asn	Ile	Lys			
			660					665					670					
Ala	Pro	Leu	Ser	Ile	Ser	Ser	Arg	Ser	Glu	Arg	Lys	Arg	Thr	Met	Cys			
		675					680					685						
Arg	Leu	Pro	Ala	Asn	Ala	Val	Glu	Arg	Phe	Arg	Gln	Val	Phe	Ala	Glu			
	690					695					700							
Asn	Glu	Leu	Pro	Ser	Met	Ser	Cys	Lys	Met	Ala	Leu	Ser	Lys	Gln	Leu			
705					710					715					720			
Asp	Leu	Pro	Tyr	Arg	Gln	Val	His	Met	Trp	Phe	Lys	Asn	Ala	Arg	Tyr			
				725					730					735				
Met	Ser	Leu	Lys	Lys	Lys	Lys	Met	Asn	Arg	Leu	Lys	Glu	Ser	Ala	Pro			
			740					745					750					
Val	Ile	Ser	Ser	His	Thr	Arg	Arg	Leu	Glu	Val	Glu	Ser	Ser	Lys	Pro			
		755					760					765						
Asn	Pro	Gln	Ile	Lys	Glu	Pro	Asp	Asp	Ala	Gly	Asn	Ser	Tyr	Leu	Thr			
	770					775					780							
Glu	Met	Glu	Lys	Leu	Gly	Asn	Ile	Glu	Leu	Lys	Leu	Glu	Asn	Met	Arg			
785					790					795					800			
Lys	Ile	Leu	Glu	Thr	Val	Cys	Pro	Lys	Arg	Gly	Thr	Asp	Phe	Ser	Asn			
				805					810					815				
Asp	Lys	Asn	Gln	Glu	Gly	Ile	Ala	Asn	Leu	Leu	Ser	Glu	Glu	Leu	Leu			
			820					825					830					
Val	Tyr	Val	Pro	Val	Val	Glu	Leu	Arg	Glu	Lys	Ser	Leu	Gln	Val	Ala			
		835					840					845						

ES 2 558 745 T3

Pro Val Arg
850

5
<210> 189
<211> 2178
<212> ADN
<213> *Populus tremuloides*

<400> 189

```

atgggtgatt ctggaaagac gtcaaagctg caagactcgc acaaatgttc accttctgac      60
acagtgactg ggtcgttgct gattaaatca ttgaagataa agaaggacag caaaatatct      120
cctagaaaag gccaaaaaac gaaaacaaaa tcaaaaccaa aaccaatacc tcatctaaaa      180
actatcatca gttcagctgt atcaaaaaga aaagtttctc ctaaggcat  tgggaatggc      240
tctaccagta gaaaactgat tcataggaaa attctgcata aagcacttga taagaaggcc      300
tcaaggaagg gggcttcctc agggctacaa ttgtcaacta ttgattctaa gggaaatgga      360
aaaaatggtg atgaaggtgc aattaaanaaa ctcaagaaga ggaagcctaa gaaaaggcaa      420
agggacaagg taaagctaga tgaaccatca cgattgcaga ggagagcaag gtacctgatg      480
attaaatga agctggacca gaaccttata gatgcttact ctggagaagg ctggaaaagg      540
caaagtcgag agaagatcag gccagaaaag gaactactgc gagcaaggaa gcagattttg      600
aagtgtaaac ttggtttacg tgatataatt cggcaggtgg attctcttag tacagtagga      660
tgtattgaag agacagttat ggctccagat ggatctgttt cacatgaaca tatattctgt      720
gcgaagtgca aattgaatga agtttcccca gataatgata ttgtactctg tgatggtaca      780
tgcaactgtg ctttccacca aaaatgcctt gaacctccct tggatactga aagtatacct      840
ccaggagatc agggatgggt ttgcaagttc tgtgaatgca ggatggatat tatagaagct      900
atgaatgctc atttgggaac ccatttctca gaggacagca gttggcagga tattttcaca      960
gaagaagcag caattcctga tgctggaaat gtgctattaa accccgaaga agaatggcct     1020
tcagatgatt ctgaagatga taattacgat ccagagagga gagacaacat catgagtgaa     1080
gcggttactg atgatgatgc atctgatgac atcagcagtt caaccagctt gggttggtct     1140
tcagatggtg aggttttttt gggatccagg aggtgggaga tgcatggctt ggacttcaga     1200
aacaattcta tttatagcag tttagattct gatgaaacca gtgatgggga aattgtttgt     1260
ggccgtaggc agagaagagc aattgactat aagaagttat atgatgaggt gtttgaaag     1320
gatgctcctg cacacgagca agccagtga gacgaagatt ggggtcctgg caaaagaaag     1380
cgacgggaga aggagtcaaa tgcagccagc acccttatga cactgtgtga aagtaaaaaa     1440
aaaagtaaaa gtgatgagac cattgaaggt atgatgaacc tccccctca aactcgaagg     1500
ccaattttca gactcccccc tgatgcagtt gagaaacttc gccaagtatt tgtagaaaat     1560
gaacttccat ctagaactgt caaggagaat ctgtcaaaag aattggcct  tgaacctggg     1620
aaggttagca aatggttcaa gaattcccgt tatttagctc tgaagtcccg gaaggtagag     1680
aaaggagaac aacttcataa ctctagtccc aaagtctctg ctgaaccac  attaaatgtc     1740

```

10

ES 2 558 745 T3

atggagggca aaactgctga tctttcacag gattcttggg aagaaactga ggtatgtatc 1800
 ccaaagaatt tgaaaaggat ccttcagagg aagaagctga agtcaataag cagaagttta 1860
 aggaaaaatg aacagaaaag aggttcctac gaatcaccta ctaaaagcaa tgagatgaat 1920
 gctgagtgcg gtgatgatgt gagcttgaag aagttgttaa aagcaaaacc aaaggagta 1980
 aagaaaaagg ttaatcctat ttcagtagcg gctgaatatg atatggagag actctgtaga 2040
 gcgaaggcta tattagagaa gatgaagcag aaagtgggta agctacaaaa tggcaaagct 2100
 aggaagtcac ccaaaacccg tccacttgaa gaatttatcg tttacattcc cattgctgag 2160
 ttgaggggaa aaaattga 2178

<210> 190
 <211> 725
 <212> PRT
 <213> *Populus tremuloides*

5

<400> 190

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

10

ES 2 558 745 T3

Gly Trp Lys Gly Gln Ser Arg Glu Lys Ile Arg Pro Glu Lys Glu Leu
 180 185 190
 Leu Arg Ala Arg Lys Gln Ile Leu Lys Cys Lys Leu Gly Leu Arg Asp
 195 200 205
 Ile Ile Arg Gln Val Asp Ser Leu Ser Thr Val Gly Cys Ile Glu Glu
 210 215 220
 Thr Val Met Ala Pro Asp Gly Ser Val Ser His Glu His Ile Phe Cys
 225 230 235 240
 Ala Lys Cys Lys Leu Asn Glu Val Ser Pro Asp Asn Asp Ile Val Leu
 245 250 255
 Cys Asp Gly Thr Cys Asn Cys Ala Phe His Gln Lys Cys Leu Glu Pro
 260 265 270
 Pro Leu Asp Thr Glu Ser Ile Pro Pro Gly Asp Gln Gly Trp Phe Cys
 275 280 285
 Lys Phe Cys Glu Cys Arg Met Asp Ile Ile Glu Ala Met Asn Ala His
 290 295 300
 Leu Gly Thr His Phe Ser Glu Asp Ser Ser Trp Gln Asp Ile Phe Thr
 305 310 315 320
 Glu Glu Ala Ala Ile Pro Asp Ala Gly Asn Val Leu Leu Asn Pro Glu
 325 330 335
 Glu Glu Trp Pro Ser Asp Asp Ser Glu Asp Asp Asn Tyr Asp Pro Glu
 340 345 350
 Arg Arg Asp Asn Ile Met Ser Glu Ala Gly Thr Asp Asp Ala Ser
 355 360 365
 Asp Asp Ile Ser Ser Ser Thr Ser Leu Gly Trp Ser Ser Asp Gly Glu
 370 375 380
 Val Phe Leu Gly Ser Arg Arg Trp Glu Met His Gly Leu Asp Phe Arg
 385 390 395 400
 Asn Asn Ser Ile Tyr Ser Ser Leu Asp Ser Asp Glu Thr Ser Asp Gly
 405 410 415
 Glu Ile Val Cys Gly Arg Arg Gln Arg Arg Ala Ile Asp Tyr Lys Lys
 420 425 430
 Leu Tyr Asp Glu Val Phe Gly Lys Asp Ala Pro Ala His Glu Gln Ala
 435 440 445
 Ser Glu Asp Glu Asp Trp Gly Pro Gly Lys Arg Lys Arg Arg Glu Lys

ES 2 558 745 T3

450						455						460					
Glu 465	Ser	Asn	Ala	Ala	Ser 470	Thr	Leu	Met	Thr	Leu 475	Cys	Glu	Ser	Lys	Lys 480		
Lys	Ser	Lys	Ser	Asp 485	Glu	Thr	Ile	Glu	Gly 490	Met	Met	Asn	Leu	Pro 495	Pro		
Gln	Thr	Arg	Arg 500	Pro	Ile	Phe	Arg	Leu 505	Pro	Pro	Asp	Ala	Val 510	Glu	Lys		
Leu	Arg	Gln 515	Val	Phe	Val	Glu	Asn 520	Glu	Leu	Pro	Ser	Arg 525	Thr	Val	Lys		
Glu	Asn 530	Leu	Ser	Lys	Glu	Leu 535	Gly	Leu	Glu	Pro	Gly 540	Lys	Val	Ser	Lys		
Trp 545	Phe	Lys	Asn	Ser	Arg 550	Tyr	Leu	Ala	Leu	Lys 555	Ser	Arg	Lys	Val	Glu 560		
Lys	Gly	Glu	Gln	Leu 565	His	Asn	Ser	Ser	Ser 570	Lys	Val	Ser	Ala	Glu 575	Pro		
Thr	Leu	Asn	Val 580	Met	Glu	Gly	Lys	Thr 585	Ala	Asp	Leu	Ser	Gln 590	Asp	Ser		
Trp	Glu	Glu 595	Thr	Glu	Val	Cys	Ile 600	Pro	Lys	Asn	Leu	Lys 605	Arg	Ile	Leu		
Gln	Arg 610	Lys	Lys	Leu	Lys	Ser 615	Ile	Ser	Arg	Ser	Leu 620	Arg	Lys	Asn	Glu		
Gln 625	Lys	Arg	Gly	Ser	Tyr 630	Glu	Ser	Pro	Thr	Lys 635	Ser	Asn	Glu	Met	Asn 640		
Ala	Glu	Cys	Ser	Asp 645	Asp	Val	Ser	Leu	Lys 650	Lys	Leu	Leu	Lys	Ala 655	Lys		
Pro	Lys	Gly	Val 660	Lys	Lys	Lys	Val	Asn 665	Pro	Ile	Ser	Val	Ala 670	Ala	Glu		
Tyr	Asp	Met 675	Glu	Arg	Leu	Cys	Arg 680	Ala	Lys	Ala	Ile	Leu 685	Glu	Lys	Met		
Lys	Gln 690	Lys	Val	Val	Lys	Leu 695	Gln	Asn	Gly	Lys	Ala 700	Arg	Lys	Ser	Ser		
Lys 705	Thr	Arg	Pro	Leu	Glu 710	Glu	Phe	Ile	Val	Tyr 715	Ile	Pro	Ile	Ala	Glu 720		
Leu	Arg	Gly	Lys	Asn 725													

ES 2 558 745 T3

<210> 191
 <211> 2256
 <212> ADN
 <213> *Saccharum officinarum*

5

<400> 191

atgcattctt cggaaaataa actggtatgc tccaactcgg gaagatcatc aaaagggaaat 60
 gagacctcta tggagttggt tcctgtacca aaaagaccaa ctagacatga tgcttcccgc 120
 caatgcaaat ctgattctcc tttgaagcga tcaccaagga aagctagaaa tgctactctg 180
 gcaaaaagca taaagaataa atatcattgt agtcccctca aacagcggag ggcttcagat 240
 tctgtttctg gaaaagttgc gacaggacta actgtgagga ggaggaagaa aagaaaaatg 300
 caaaatactg atgaggcaac tcgcttgga cgaagagcga gatattttct aatcaagata 360
 aaattggagc agaatttgc agatgcttac tctggagatg gatggaatgg gcaaagccgg 420
 gaaaaaataa agccagataa ggatctgcgg cgtgccagga aacaaatcat aaaatgcaaa 480
 attgctatac gtgatatcat tcgccaaactt ggtttgtata cttctactgg gagtcttaat 540
 gaccagcaa tgctccaga tcagtctacc aatccggaac ataccatgtg ctcaacatgc 600
 aagtctcatg aatcatttcc cagcaataaa tttatcttct gtgaagggcc ctgcaaaagg 660
 gcatatcatg agaaatgttt ggaacctccc ttaaacaaaa gcgtacttcc aacaagtagc 720
 catggatggc tttgcaaatt ctgtttgtgc aaagtgaaaa ttttagaaac tattaatgca 780
 catctaggaa caagttttac agtgaagtgc cccttcgaag atatattcaa ggaagcaacg 840
 gaacaaatag actccgagga tgcactagat gaagattggc tttctgagta ttcaggtgat 900
 gaggactatg atcctgatga aaatgaggac agcggcaact gtatggacag cggggaggaa 960
 attatgtctg atgattccga tggttcagga agccccctt attctccaaa tgacgatatt 1020
 cctgacttca tatcagcaga cttaaatgac gtggaaggggt tttgtcacac caatttagac 1080
 ttaggcattg atgccgttga agatgatttg gcacagatcc ttacctacca aaggccaaga 1140
 agagatgttg attatagaag gcttaatgag gaaatgtttg gaaaaataat aggaaatgaa 1200
 ggacagagtg aggatgaaga ttggggccat gaaaggagaa agaaaaggac ttgttcaggt 1260
 ggtgctgggg ataattctgt gggtttctca aatgttatat ctgaggaaaa gagccaaaag 1320
 aaggggagaa aacttttcag gattcctcct gcagctgtcg aggtacttcg caaagctttt 1380
 gctgaaaatg agcttcacc ccgggatgtc aaagaaaatc tctcaagaga attggaatt 1440
 tcttttgaaa agattgataa atggttcaaa aatacacggt gtgctgctct cagagatcgc 1500
 aaggctgaag ggaacagtca taatacggct cccagcaaaa gctcaagaaa taaaggaaaa 1560
 gctggaatct caagccaggc tgaaaggaat ggtcatgtta ctggtcctg caataactcg 1620
 agaacaaatg aagaaaaatc gttgacacta gctcctaag cccagggaga gaagaatggg 1680
 attgcagcag attatgtctt acaaaaaatc ggtatatcgg ggaaggttga ttcagtagac 1740
 aactcttgct tggttccctt gtctgaaagc atcaatgtgc atacacgatt gcaacagaat 1800
 cttgagatga ggaagatgga atcaactagc agtcctgtgt ggctgcacaa taaaggaggt 1860

ES 2 558 745 T3

tgcttgtttc caacaggcca agccaaggag agtacattac ctacaggcaa gccttgtttg 1920
 cagtcagaaa taaccattc aacaattaat gaagtgagca ctttagtgca agctacttct 1980
 tggatggatg ctgggtcgtg cgctgaggtg caagaagcca ctcttgggtt ggacatcggg 2040
 gcctcagatt accagccttt cctggacgtg atcgatgaga tgtgctgact tgagtgcagg 2100
 ctgcagagac tgaaggagaa catgctctca tctggcatag acggcaaaac cacaggtgag 2160
 agtgacatgg gaaaccaggc tgtggtgctc gtgccgaccg ccgagctcaa ggaaaaagcg 2220
 ccgcatggca gtttttttgg gcattattgc ccatag 2256

<210> 192
 <211> 751
 <212> PRT
 <213> *Saccharum officinarum*
 <400> 192

5

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

10

ES 2 558 745 T3

Glu His Thr Met Cys Ser Thr Cys Lys Ser His Glu Ser Phe Pro Ser
 195 200 205
 Asn Lys Phe Ile Phe Cys Glu Gly Pro Cys Lys Arg Ala Tyr His Glu
 210 215 220
 Lys Cys Leu Glu Pro Pro Leu Asn Lys Ser Val Leu Pro Thr Ser Ser
 225 230 235 240
 His Gly Trp Leu Cys Lys Phe Cys Leu Cys Lys Val Lys Ile Leu Glu
 245 250 255
 Thr Ile Asn Ala His Leu Gly Thr Ser Phe Thr Val Lys Cys Pro Phe
 260 265 270
 Glu Asp Ile Phe Lys Glu Ala Thr Glu Gln Ile Asp Ser Glu Asp Ala
 275 280 285
 Leu Asp Glu Asp Trp Leu Ser Glu Tyr Ser Gly Asp Glu Asp Tyr Asp
 290 295 300
 Pro Asp Glu Asn Glu Asp Ser Gly Asn Cys Met Asp Ser Gly Glu Glu
 305 310 315 320
 Ile Met Ser Asp Asp Ser Asp Gly Ser Gly Ser Pro Leu Tyr Ser Pro
 325 330 335
 Asn Asp Asp Ile Pro Asp Phe Ile Ser Ala Asp Leu Asn Asp Val Glu
 340 345 350
 Gly Phe Cys His Thr Asn Leu Asp Leu Gly Ile Asp Ala Val Glu Asp
 355 360 365
 Asp Leu Ala Gln Ile Leu Thr Tyr Gln Arg Pro Arg Arg Asp Val Asp
 370 375 380
 Tyr Arg Arg Leu Asn Glu Glu Met Phe Gly Lys Ile Ile Gly Asn Glu
 385 390 395 400
 Gly Gln Ser Glu Asp Glu Asp Trp Gly His Glu Arg Arg Lys Lys Arg
 405 410 415
 Thr Cys Ser Gly Gly Ala Gly Asp Asn Ser Val Gly Phe Ser Asn Val
 420 425 430
 Ile Ser Glu Glu Lys Ser Gln Lys Lys Gly Arg Lys Leu Phe Arg Ile
 435 440 445
 Pro Pro Ala Ala Val Glu Val Leu Arg Lys Ala Phe Ala Glu Asn Glu
 450 455 460

ES 2 558 745 T3

Leu Pro Pro Arg Asp Val Lys Glu Asn Leu Ser Arg Glu Leu Gly Ile
 465 470 475 480
 Ser Phe Glu Lys Ile Asp Lys Trp Phe Lys Asn Thr Arg Cys Ala Ala
 485 490 495
 Leu Arg Asp Arg Lys Ala Glu Gly Asn Ser His Asn Thr Ala Pro Ser
 500 505 510
 Lys Ser Ser Arg Asn Lys Gly Lys Ala Gly Ile Ser Ser Gln Ala Glu
 515 520 525
 Arg Asn Gly His Val Thr Gly Pro Cys Asn Asn Ser Arg Thr Asn Glu
 530 535 540
 Glu Lys Ser Leu Thr Leu Ala Pro Asn Ala Gln Gly Glu Lys Asn Gly
 545 550 555 560
 Ile Ala Ala Asp Tyr Val Leu Gln Lys Ser Gly Ile Ser Gly Lys Val
 565 570 575
 Asp Ser Val Asp Asn Ser Cys Leu Val Pro Leu Ser Glu Ser Ile Asn
 580 585 590
 Val His Thr Arg Leu Gln Gln Asn Leu Glu Met Arg Lys Met Glu Ser
 595 600 605
 Thr Ser Ser Pro Val Trp Leu His Asn Lys Gly Gly Cys Leu Phe Pro
 610 615 620
 Thr Gly Gln Ala Lys Glu Ser Thr Leu Pro Thr Gly Lys Pro Cys Leu
 625 630 635 640
 Gln Ser Glu Ile Thr His Ser Thr Ile Asn Glu Val Ser Thr Leu Val
 645 650 655
 Gln Ala Thr Ser Trp Met Asp Ala Gly Ser Cys Ala Glu Val Gln Glu
 660 665 670
 Ala Thr Pro Trp Val Asp Ile Gly Ala Ser Asp Tyr Gln Pro Phe Leu
 675 680 685
 Asp Val Ile Asp Glu Met Cys Gly Leu Glu Cys Arg Leu Gln Arg Leu
 690 695 700
 Lys Glu Asn Met Leu Ser Ser Gly Ile Asp Gly Lys Thr Thr Gly Glu
 705 710 715 720
 Ser Asp Met Gly Asn Gln Ala Val Val Leu Val Pro Thr Ala Glu Leu
 725 730 735

ES 2 558 745 T3

Lys Glu Lys Ala Pro His Gly Ser Phe Phe Gly His Tyr Cys Pro
 740 745 750

5 <210> 193
 <211> 2181
 <212> ADN
 <213> *Vitis vinifera*

<400> 193

atgCGtggaa ctggaaagaa agcaggacat caggaatctg gaaagtcttg ctttcctaaa 60
 agaaatatcg ggcctaagct gaatgcagca ttgcagatta aaaatggtag taaaatatct 120
 cagaccagga aatgcaagcc aaaatcaaaa tctcatgcaa agacaattgg tgcaatcctg 180
 tctaagagaa caactactga ctctccaagc aagggagca ggagtggatc cacaaccaga 240
 aaactgattc acaagaaaac cctgcataaa gccattgata ctgagtcttc caagaaggag 300
 tcttcatcaa agcttaaagg tgagaagccc ccacaaatta gcacaaaca aaatggggaa 360
 actgtggaca agaatgttaa acctcaaaaa ttaaagaaaa gaggaaagcg gaagcggcga 420
 aaggataatt cagagctaga tgaagcttct cgcttgcaaa gaagaaccag gtacctcttg 480
 atcaaaatga aactggagca gaatcttatt gatgcttact ctggggaagg ttggaaaggt 540
 catagtcgag aaaagatcag gccagagaag gaactacaaa gagctacgaa gcaaatattg 600
 aaatgtaaac ttgggatccg tgatgcaatt cgccagctgg agtcaactgag ctccatagga 660
 tgtattgaag atactgcaat tgcttcagat ggatctgttt atcatgaaca cataatctgt 720
 gcaaagtgca agttgCGgga ggctttccca gataatgata ttatactctg tgatgggacg 780
 tgcaattgtg cttccacca gaaatgtctt gatcctccat tggaaactga aaatattcct 840
 ccgggagacc aaggttgggt ttgcaagttc tgCGagtgta agatggaaat actagaagca 900
 atgaatgcac atcttggaac ccgcttctct gtggacagta ctggcagga tattttcaaa 960
 gaagaagctg ctttgCctga tggTgggagt gcactgcCct atccagagga agattgCct 1020
 tctgatgatt cacaagatca tgattatgat cctgaaagaa atgaaaatag ctgcagcatc 1080
 agcactgctg gcactgaagg aaatgcttca gatgatacaa atagtctttt aagcttgagt 1140
 tggtcgtttg aagacgaaat cttttctgga tccaaaagat cagggattat cagtgcagat 1200
 tctgatgaaa catctgattg tgaaatcata agtggtcgta ggcaacgaag agctgttgat 1260
 tatagaaagt tatatgatga aatgtttgga aaggatgctc atgcaaatga acaagttagt 1320
 gaagatgaag actggggTcc tgctaacaaa aggcggagag aaaaggagtc tgatgcagcc 1380
 agcaccctca taactctata tgaaggtgag aaaaagttac ctaatgtgga aaccatggaa 1440
 gccaaacaaa aaatttcttc agatccacaa actaaaaggc cattttccag aattccactt 1500
 gatgcggttg agaagcttcg ccaagcattt ggggagaatg aacttcctc tagagatgtg 1560
 agggagaatc ttgcaaagca gttgggtctt gattatgaga aggtgaacaa atggttcaaa 1620
 aatgcacgat atatagcgtt aaaaacaaga aaggcagaaa gagcaaaaca acttcagact 1680
 tctcccagaa tctccaagga atccagatca gaaatcgtga aggacaaaac tgttgatcta 1740
 gtggcatcaa gggataactc atcagcatca ctggTccgtg cactaaagaa tttgaaaaag 1800

10

ES 2 558 745 T3

gtgctggagga gaaaaaatcc aaagccaata atgaccagcc ctgtaaagaa aaagcatcat 1860
 agaagggctt tacttgagtc acctacaaat gacaaggtta ctatggagtt tgacgatgat 1920
 gtgagcttga agaaacaact gaagcttttg aaagaaaaaa gtaagaggga caagcaaagg 1980
 gtcgacttca aggagggaac tgggtgtacaa gatgcagaga aggagatgga gagactctgc 2040
 caaattaagg ataagataga gaagttgaag caggtataac tgagacttca atgtgacaaa 2100
 actaatcaat ggcaagacca gtcagtgatc tatgttcttg tggcagagct aaggagagaaa 2160
 ggctgatttt gcatcacatc t 2181

<210> 194
 <211> 727
 <212> PRT
 <213> *Vitis vinifera*

5

<400> 194

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

10

ES 2 558 745 T3

Gln Arg Ala Thr Lys Gln Ile Leu Lys Cys Lys Leu Gly Ile Arg Asp
 195 200 205

Ala Ile Arg Gln Leu Glu Ser Leu Ser Ser Ile Gly Cys Ile Glu Asp
 210 215 220

Thr Ala Ile Ala Ser Asp Gly Ser Val Tyr His Glu His Ile Ile Cys
 225 230 235 240

Ala Lys Cys Lys Leu Arg Glu Ala Phe Pro Asp Asn Asp Ile Ile Leu
 245 250 255

Cys Asp Gly Thr Cys Asn Cys Ala Phe His Gln Lys Cys Leu Asp Pro
 260 265 270

Pro Leu Glu Thr Glu Asn Ile Pro Pro Gly Asp Gln Gly Trp Phe Cys
 275 280 285

Lys Phe Cys Glu Cys Lys Met Glu Ile Leu Glu Ala Met Asn Ala His
 290 295 300

Leu Gly Thr Arg Phe Ser Val Asp Ser Thr Trp Gln Asp Ile Phe Lys
 305 310 315 320

Glu Glu Ala Ala Leu Pro Asp Gly Gly Ser Ala Leu Pro Tyr Pro Glu
 325 330 335

Glu Asp Trp Pro Ser Asp Asp Ser Gln Asp His Asp Tyr Asp Pro Glu
 340 345 350

Arg Asn Glu Asn Ser Cys Ser Ile Ser Thr Ala Gly Thr Glu Gly Asn
 355 360 365

Ala Ser Asp Asp Thr Asn Ser Ser Leu Ser Leu Ser Trp Ser Phe Glu
 370 375 380

Asp Glu Ile Leu Ser Gly Ser Lys Arg Ser Gly Ile Ile Ser Ala Asp
 385 390 395 400

ser Asp Glu Thr Ser Asp Cys Glu Ile Ile Ser Gly Arg Arg Gln Arg
 405 410 415

Arg Ala Val Asp Tyr Arg Lys Leu Tyr Asp Glu Met Phe Gly Lys Asp
 420 425 430

Ala His Ala Asn Glu Gln Val Ser Glu Asp Glu Asp Trp Gly Pro Ala
 435 440 445

Asn Lys Arg Arg Arg Glu Lys Glu Ser Asp Ala Ala Ser Thr Leu Ile
 450 455 460

ES 2 558 745 T3

Thr Leu Tyr Glu Gly Glu Lys Lys Leu Pro Asn Val Glu Thr Met Glu
 465 470 475 480
 Ala Lys Gln Lys Ile Ser Ser Asp Pro Gln Thr Lys Arg Pro Phe Ser
 485 490 495
 Arg Ile Pro Leu Asp Ala Val Glu Lys Leu Arg Gln Ala Phe Gly Glu
 500 505 510
 Asn Glu Leu Pro Ser Arg Asp Val Arg Glu Asn Leu Ala Lys Gln Leu
 515 520 525
 Gly Leu Asp Tyr Glu Lys Val Asn Lys Trp Phe Lys Asn Ala Arg Tyr
 530 535 540
 Ile Ala Leu Lys Thr Arg Lys Ala Glu Arg Ala Lys Gln Leu Gln Thr
 545 550 555 560
 Ser Pro Arg Ile Ser Lys Glu Ser Arg Ser Glu Ile Val Lys Asp Lys
 565 570 575
 Thr Val Asp Leu Val Ala Ser Arg Asp Asn Ser Ser Ala Ser Leu Val
 580 585 590
 Arg Ala Leu Lys Asn Leu Lys Lys Val Arg Arg Arg Lys Asn Pro Lys
 595 600 605
 Pro Ile Met Thr Ser Pro Val Lys Lys Lys His His Arg Arg Ala Leu
 610 615 620
 Leu Glu Ser Pro Thr Asn Asp Lys Val Thr Met Glu Phe Asp Asp Asp
 625 630 635 640
 Val Ser Leu Lys Lys Gln Leu Lys Leu Leu Lys Glu Lys Ser Lys Arg
 645 650 655
 Asp Lys Gln Arg Val Asp Phe Lys Glu Gly Thr Gly Val Gln Asp Ala
 660 665 670
 Glu Lys Glu Met Glu Arg Leu Cys Gln Ile Lys Asp Lys Ile Glu Lys
 675 680 685
 Leu Lys Gln Val Ile Leu Arg Leu Gln Cys Asp Lys Thr Asn Gln Trp
 690 695 700
 Gln Asp Gln Ser Val Ile Tyr Val Pro Val Ala Glu Leu Arg Glu Lys
 705 710 715 720
 Gly Arg Phe Cys Ile Thr Ser
 725

<210> 195
<211> 2166
<212> ADN
<213> *Zea mays*

5

<400> 195

ES 2 558 745 T3

atgcattcctt cggaaaataa actggggttg aatgagacct ctatgggggtt ggttcctgta 60
 ccaaaaagac cagctagacc tgatgcttcc caccaatgca aatctgattc ctttatgaag 120
 cgatcaccaa ggaaagttag aaatgctact ctggcaaaaa gcataaagag caaataccac 180
 tatagtcccc tcaaacagcg gaagggttca gattctgttc ctgggaaaat cgtaacagga 240
 ctaaccgcaa ggaaaaagaa gaaaaggaaa atccaaatta cagacgaggc aactcgtttg 300
 gaacgaagag cgagatattt tctaatacaag ataaaactgg agcagaattt gctagatgct 360
 tactctggag atggatggaa tgggcaaagc cgagagaaaa taaagccaga gaaggaactg 420
 caacgtgcca ggaaacaat cataaaatgc aaaattgcta tacgtgatat catccgcca 480
 cttgatttgt atacttctac tgggagtgtt gatgaccac taatgcctac agatcagtcc 540
 accaatcccc aacataccat gtgctcaaca tgcaagtctc atgaatcatt tcccagcaat 600
 aaaattatct tctgcaaagg gccctgcaaa agggcatgtc atgagaaatg tttggaacct 660
 cccttaaaca aaagcgtact tccaacaagt agccatgggt ggctttgcaa attctgtttg 720
 tgcaaggatga ggattttaga aactattaat gcacatctag ggacgagttt tacagtgaag 780
 tgccactttg aagatatatt caaggaaact actgaaactaa tagactctga ggatgacta 840
 gatgaagatt ggctttctga gtattcaggt gatgaggact atgatcctga tgaaaatgag 900
 gccagtggcg actgtatgga cagcggggag aagattatgt ctgatgattc caatggttca 960
 ggaagcccc tttattctcc aaatgacgat attcctgact tcatatcagc agacttaaat 1020
 gttgtggaag ggttttgtca taccaattta gatttaggca ttgatgccgt tgaagatgat 1080
 tttgcacaga tctcaceta ccaaaggcca agaagagatg ttgattatag aaggctaac 1140
 gaggaaatgt ttgggaaaat aaccgggaat gaagaacaga gcgaggacga agattggggc 1200
 catgaaagga gaaagaaaag gacgcattca ggtgttgctg gggataattc tgttggtttc 1260
 ttgaacgtta tatctgatga aaagagccag aagaagggga gaaaactttt caggattcct 1320
 cctgcagctg ttgaggtact tcgcagagct tttgctgaaa atgagcttcc accccgggat 1380
 gttaaagaaa atctctcaag agaattggga atttcttttg aaaagattga taaatggttc 1440
 aaaaatacac ggtgtgctgc tctcagagat cgggaaggctg aaggaaacag tcataataca 1500
 gctcccagca aaagctcaa aaataaagga aaagctggaa tctcaggcaa gactggaagg 1560
 aatggtcatg ttactggtcc ctgcaataat ttgagaacaa atgaagaaaa aacaggtata 1620
 tcggggaagt ttgattcagg agacaactct tgtttggttc ctttctctga agtcatcaat 1680
 gtgcctacgc gattgccaca caactttgag atgaggaaga tggaatcaac tagaagtcc 1740
 gcgaggctac acaataaagg agggttcctg tctgcaacag tccaaattaa ggagagtacg 1800
 ttgttaccca caggcaagcc ttgttggcag tcagaaatga gccatccaac aactaatgaa 1860
 gtgagcactt cagcgaagc tactacttgg atcgacgcg gggcgtgtgc cgaggagcaa 1920

ES 2 558 745 T3

gaacctactc cttgggtgga catcgggggc tcagactacc agcctttcct ggatgtgatc 1980
gacgacatgt gcggaactga gtggaggctg cagagactga aggagaacat gctctcatct 2040
agcatagacg gccaaaccgc cgccagtgag agtgacaaaa gaaaccagac cgtgggtgcta 2100
gtgccaaccg ccgagctcaa ggaaaagctg ccgcatgacg gtttattcgg gcattattgc 2160
ccatag 2166

<210> 196
<211> 721
<212> PRT
<213> *Zea mays*
<400> 196

5

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

10

ES 2 558 745 T3

Cys Lys Arg Ala Cys His Glu Lys Cys Leu Glu Pro Pro Leu Asn Lys
 210 215 220

Ser Val Leu Pro Thr Ser Ser His Gly Trp Leu Cys Lys Phe Cys Leu
 225 230 235 240

Cys Lys Val Arg Ile Leu Glu Thr Ile Asn Ala His Leu Gly Thr Ser
 245 250 255

Phe Thr Val Lys Cys His Phe Glu Asp Ile Phe Lys Glu Thr Thr Glu
 260 265 270

Leu Ile Asp Ser Glu Asp Ala Leu Asp Glu Asp Trp Leu Ser Glu Tyr
 275 280 285

Ser Gly Asp Glu Asp Tyr Asp Pro Asp Glu Asn Glu Ala Ser Gly Asp
 290 295 300

Cys Met Asp Ser Gly Glu Lys Ile Met Ser Asp Asp Ser Asn Gly Ser
 305 310 315 320

Gly Ser Pro Leu Tyr Ser Pro Asn Asp Asp Ile Pro Asp Phe Ile Ser
 325 330 335

Ala Asp Leu Asn Val Val Glu Gly Phe Cys His Thr Asn Leu Asp Leu
 340 345 350

Gly Ile Asp Ala Val Glu Asp Asp Phe Ala Gln Ile Leu Thr Tyr Gln
 355 360 365

Arg Pro Arg Arg Asp Val Asp Tyr Arg Arg Leu Asn Glu Glu Met Phe
 370 375 380

Gly Lys Ile Thr Gly Asn Glu Glu Gln Ser Glu Asp Glu Asp Trp Gly
 385 390 395 400

His Glu Arg Arg Lys Lys Arg Thr His Ser Gly Val Ala Gly Asp Asn
 405 410 415

Ser Val Gly Phe Leu Asn Val Ile Ser Asp Glu Lys Ser Gln Lys Lys
 420 425 430

Gly Arg Lys Leu Phe Arg Ile Pro Pro Ala Ala Val Glu Val Leu Arg
 435 440 445

Arg Ala Phe Ala Glu Asn Glu Leu Pro Pro Arg Asp Val Lys Glu Asn
 450 455 460

Leu Ser Arg Glu Leu Gly Ile Ser Phe Glu Lys Ile Asp Lys Trp Phe
 465 470 475 480

ES 2 558 745 T3

Lys Asn Thr Arg Cys Ala Ala Leu Arg Asp Arg Lys Ala Glu Gly Asn
 485 490 495

Ser His Asn Thr Ala Pro Ser Lys Ser Ser Lys Asn Lys Gly Lys Ala
 500 505 510

Gly Ile Ser Gly Lys Thr Gly Arg Asn Gly His Val Thr Gly Pro Cys
 515 520 525

Asn Asn Leu Arg Thr Asn Glu Glu Lys Thr Gly Ile Ser Gly Lys Phe
 530 535 540

Asp Ser Gly Asp Asn Ser Cys Leu Val Pro Phe Ser Glu Val Ile Asn
 545 550 555 560

Val Pro Thr Arg Leu Pro His Asn Phe Glu Met Arg Lys Met Glu Ser
 565 570 575

Thr Arg Ser Pro Ala Arg Leu His Asn Lys Gly Gly Phe Leu Ser Ala
 580 585 590

Thr Val Gln Ile Lys Glu Ser Thr Leu Leu Pro Thr Gly Lys Pro Cys
 595 600 605

Trp Gln Ser Glu Met Ser His Pro Thr Thr Asn Glu Val Ser Thr Ser
 610 615 620

Ala Gln Ala Thr Thr Trp Ile Asp Ala Gly Ala Cys Ala Glu Glu Gln
 625 630 635 640

Glu Pro Thr Pro Trp Val Asp Ile Gly Val Ser Asp Tyr Gln Pro Phe
 645 650 655

Leu Asp Val Ile Asp Asp Met Cys Gly Leu Glu Trp Arg Leu Gln Arg
 660 665 670

Leu Lys Glu Asn Met Leu Ser Ser Ser Ile Asp Gly Gln Thr Ala Ala
 675 680 685

Ser Glu Ser Asp Lys Arg Asn Gln Thr Val Val Leu Val Pro Thr Ala
 690 695 700

Glu Leu Lys Glu Lys Leu Pro His Asp Gly Leu Phe Gly His Tyr Cys
 705 710 715 720

Pro

<210> 197
 <211> 2172
 <212> ADN
 <213> *Arabidopsis thaliana*

5

<400> 197

ES 2 558 745 T3

atgtataaag cagtaagtaa gcgggtaact agaagtagtg gatctggttt gaagcagaca 60
aatgtagata atggaggaga gattagtcca actgttgaca gagtatctga acaaggcaaa 120
tcatcagagg cagggagtca catgccaact gatgcaaacg gaaatggaca tttgcatcat 180
gagattatgg atcatggaaa ggggaatgaa gaacagaaac caaccctca aaccgtgaag 240
aaggattcta ataccaatac caaattttca gggctcctc gtgagcttgt catcggtcct 300
ccttgccgtg gacaatttga aatccacaac cggtcaggg catctacgag ttcgaagagg 360
ttaggaggag gtggagaaa gaaatgtgtt tttgcaagtc acaagagagc tcagagatct 420
aaggaagatg ctggaccatc ttcggttggt gctaattcca cacctgtagg tcgaccaaag 480
aaaaagaata agactatgaa caaaggacaa gttagagaag atgatgagta cacaagaatc 540
aaaaagaaac ttcgatactt tctaaaccgg attaactatg agcaaagcct aattgacgcc 600
tattcgttag aaggctggaa aggttcaagc ctggaaaaga taaggccaga gaaagagctt 660
gaacgagcca caaaggaaat tttgcggcgc aagcttaaaa ttagagatct tttcagcac 720
cttgacacac tctgtgctga aggaagcctt ccagagtctt tatttgatac tgatggagag 780
atctctagcg aggatatctt ctgtgcaaaa tgtggttcaa aggacttgtc tgttgataat 840
gatatcatac tatgcatggt gttttgcat cgaggcttc accaatactg tcttgaacca 900
cctctgcgga aagaagatat tctccggat gatgagggtt ggctatgccc cggatgtgat 960
tgcaaagatg acagcttggg cttgcttaat gattctttag ggacaaaatt ttcggtcagc 1020
gacagttggg agaaaatatt tctgaggca gcagcagcgc tggttggtgg aggtcaaaat 1080
cttgactgtg atcttccttc agatgattct gatgatgaag agtatgatcc agattgtttg 1140
aatgataatg aaaatgatga agatggcagt gatgacaatg aagaatctga gaacgaagat 1200
ggtagttctg atgaaactga attcacatct gcatctgacg aaatgattga gtcatttaa 1260
gaagggaaag atataatgaa agatgtaatg gcccttccat ctgatgattc tgaggatgat 1320
gactatgatc ctgatgccc gacttgtgat gatgataaag aaagttcaaa ttctgattgt 1380
acatcagaca ctgaagatct tgaaacttcc ttcaaaggag atgagactaa tcagcaagct 1440
gaagatacac cacttgagga tcttggcaga cagacgtctc agctccaagg tgatgcaatt 1500
ttagagtcag atgttgggct tgatgatggt cctgccgggt tatccagaag gaggaatggt 1560
gaaaggctgg attacaaaaa gttgtatgat gaggaatatg acaatgttcc gacttcgtca 1620
agtgatgatg acgattggga taaaactgca agaatgggaa aagaagattc cgagtcggaa 1680
gatgaagggg aactgtacc tctgaaacag tcttcaaagc cagaagatca tacttctaag 1740
aaacttatac ggaagtcgaa aagggcagat aaaaaggata ccttagaaat gccacaagaa 1800
ggctctggag aaaatggtgg ttctggtgaa attgaaaaaa gcagctcttc agcatgtaa 1860
cagacagatc ccaaaactca gagactgtac atatcttttc aagagaatca atatccagac 1920
aaagccacga aggagagctt agccaaagag ctacaaatga cagttaagca agttaataac 1980
tggtttaagc ataggcgttg gtcgataaac agcaaaccct tagtttcaga ggaaaatgct 2040

ES 2 558 745 T3

gagaagtaa aaacaggcaa ggaaggtgaa tgtgagacat ctggtgctgg aagtagcaaa 2100
caactatgg aaaccgagtc agttgcagag aaaccaacca atactggatc acggaaaaga 2160
agaagaaagt ag 2172

5

<210> 198
<211> 723
<212> PRT
<213> *Arabidopsis thaliana*
<400> 198

ES 2 558 745 T3

Met Tyr Lys Ala Val Ser Lys Arg Val Thr Arg Ser Ser Gly Ser Gly
1 5 10 15

Leu Lys Gln Thr Asn Val Asp Asn Gly Gly Glu Ile Ser Pro Thr Val
20 25 30

Asp Arg Val Ser Glu Gln Gly Lys Ser Ser Glu Ala Gly Ser His Met
35 40 45

Pro Thr Asp Ala Asn Gly Asn Gly His Leu His His Glu Ile Met Asp
50 55 60

His Gly Lys Gly Asn Glu Glu Gln Lys Pro Thr Pro Gln Thr Val Lys
65 70 75 80

Lys Asp Ser Asn Thr Asn Thr Lys Phe Ser Gly Ser His Arg Glu Leu
85 90 95

Val Ile Gly Leu Pro Cys Arg Gly Gln Phe Glu Ile His Asn Arg Ser
100 105 110

Arg Ala Ser Thr Ser Ser Lys Arg Leu Gly Gly Gly Gly Glu Arg Asn
115 120 125

Val Leu Phe Ala Ser His Lys Arg Ala Gln Arg Ser Lys Glu Asp Ala
130 135 140

Gly Pro Ser Ser Val Val Ala Asn Ser Thr Pro Val Gly Arg Pro Lys
145 150 155 160

Lys Lys Asn Lys Thr Met Asn Lys Gly Gln Val Arg Glu Asp Asp Glu
165 170 175

Tyr Thr Arg Ile Lys Lys Lys Leu Arg Tyr Phe Leu Asn Arg Ile Asn
180 185 190

Tyr Glu Gln Ser Leu Ile Asp Ala Tyr Ser Leu Glu Gly Trp Lys Gly
195 200 205

Ser Ser Leu Glu Lys Ile Arg Pro Glu Lys Glu Leu Glu Arg Ala Thr
210 215 220

ES 2 558 745 T3

Lys Glu Ile Leu Arg Arg Lys Leu Lys Ile Arg Asp Leu Phe Gln His
 225 230 235 240
 Leu Asp Thr Leu Cys Ala Glu Gly Ser Leu Pro Glu Ser Leu Phe Asp
 245 250 255
 Thr Asp Gly Glu Ile Ser Ser Glu Asp Ile Phe Cys Ala Lys Cys Gly
 260 265 270
 Ser Lys Asp Leu Ser Val Asp Asn Asp Ile Ile Leu Cys Asp Gly Phe
 275 280 285
 Cys Asp Arg Gly Phe His Gln Tyr Cys Leu Glu Pro Pro Leu Arg Lys
 290 295 300
 Glu Asp Ile Pro Pro Asp Asp Glu Gly Trp Leu Cys Pro Gly Cys Asp
 305 310 315 320
 Cys Lys Asp Asp Ser Leu Asp Leu Leu Asn Asp Ser Leu Gly Thr Lys
 325 330 335
 Phe Ser Val Ser Asp Ser Trp Glu Lys Ile Phe Pro Glu Ala Ala Ala
 340 345 350
 Ala Leu Val Gly Gly Gly Gln Asn Leu Asp Cys Asp Leu Pro Ser Asp
 355 360 365
 Asp Ser Asp Asp Glu Glu Tyr Asp Pro Asp Cys Leu Asn Asp Asn Glu
 370 375 380
 Asn Asp Glu Asp Gly Ser Asp Asp Asn Glu Glu Ser Glu Asn Glu Asp
 385 390 395 400
 Gly Ser Ser Asp Glu Thr Glu Phe Thr Ser Ala Ser Asp Glu Met Ile
 405 410 415
 Glu Ser Phe Lys Glu Gly Lys Asp Ile Met Lys Asp Val Met Ala Leu
 420 425 430
 Pro Ser Asp Asp Ser Glu Asp Asp Asp Tyr Asp Pro Asp Ala Pro Thr
 435 440 445
 Cys Asp Asp Asp Lys Glu Ser Ser Asn Ser Asp Cys Thr Ser Asp Thr
 450 455 460
 Glu Asp Leu Glu Thr Ser Phe Lys Gly Asp Glu Thr Asn Gln Gln Ala
 465 470 475 480
 Glu Asp Thr Pro Leu Glu Asp Pro Gly Arg Gln Thr Ser Gln Leu Gln
 485 490 495

ES 2 558 745 T3

Gly Asp Ala Ile Leu Glu Ser Asp Val Gly Leu Asp Asp Gly Pro Ala
 500 505 510

Gly Val Ser Arg Arg Arg Asn Val Glu Arg Leu Asp Tyr Lys Lys Leu
 515 520 525

Tyr Asp Glu Glu Tyr Asp Asn Val Pro Thr Ser Ser Ser Asp Asp Asp
 530 535 540

Asp Trp Asp Lys Thr Ala Arg Met Gly Lys Glu Asp Ser Glu Ser Glu
 545 550 555 560

Asp Glu Gly Asp Thr Val Pro Leu Lys Gln Ser Ser Asn Ala Glu Asp
 565 570 575

His Thr Ser Lys Lys Leu Ile Arg Lys Ser Lys Arg Ala Asp Lys Lys
 580 585 590

Asp Thr Leu Glu Met Pro Gln Glu Gly Pro Gly Glu Asn Gly Gly Ser
 595 600 605

Gly Glu Ile Glu Lys Ser Ser Ser Ser Ala Cys Lys Gln Thr Asp Pro
 610 615 620

Lys Thr Gln Arg Leu Tyr Ile Ser Phe Gln Glu Asn Gln Tyr Pro Asp
 625 630 635 640

Lys Ala Thr Lys Glu Ser Leu Ala Lys Glu Leu Gln Met Thr Val Lys
 645 650 655

Gln Val Asn Asn Trp Phe Lys His Arg Arg Trp Ser Ile Asn Ser Lys
 660 665 670

Pro Leu Val Ser Glu Glu Asn Val Glu Lys Leu Lys Thr Gly Lys Glu
 675 680 685

Gly Glu Cys Glu Thr Ser Val Ala Gly Ser Ser Lys Gln Thr Met Glu
 690 695 700

Thr Glu Ser Val Ala Glu Lys Pro Thr Asn Thr Gly Ser Arg Lys Arg
 705 710 715 720

Arg Arg Lys

<210> 199
 <211> 2901
 <212> ADN
 <213> *Lotus japonicus*
 <400> 199

5

10

atgattagtt caactggaga attgacaagc cataacaaca gtagcattga gcacatggaa 60

ES 2 558 745 T3

actgagcaac ctgaattgtc tgaaaaaaca cctcagatag gctctgaatg tctggacggt 120
 gaacaaaggg tgcttggtac tgtgtcaacc agttctgtta ataataaaaa ttcaaatcaa 180
 gtttctgcca atgtgaccga gaattctgtt tttcaactac cagcaacacc tcaacattgc 240
 ttagaaaaga gtggtcagac tgtaacaagt ccaccctgg aagaaagcac cttgaaacaa 300
 gttgctacta atgtatctaa tggtaaatca gaaaacaaat gtcagacatc ctctcaagat 360
 gttcaaaacg aacttgtgga agcaagtgat gtaataactg gctcttttgt tgcggaccaa 420
 acacaatcca ttctgctca agtgaatagc agttctgtga atgaactatt ggatattcct 480
 tccggagatg ttgagtaga agttggttct gatagtttag aaaaaagtc taaaggggtgc 540
 ttggtagaca gcgaacaaag ggtacttggg actgtgttaa ccaattctat taatgatgga 600
 aattcaaatc aagtttctgc caatgtgacc gagaattctg ttattcaact gccagcacca 660
 cctcaacatt gctttgaaaa gagtggacag actgtaaaaa gttcatctct ggaagaaagt 720
 accttaaac aagtcctccac taatttatct aacagtaaat cagaaaacat atgtcagaca 780
 tcctctcaag atgttcaaaa cgaacttgtg gaaacaagtg atgcagtaac tggctctgtt 840
 gttgcagaca aaacacaatc cattcctgct caagtgaata caagttctgt gaatgaacta 900
 ttggatattc cttccagaga tgttgcagaa gaagttgggt ctgatagttc agaaaggaag 960
 tctaaaagca caactccttc acaattgaga catataggtg aaagtaactc aaagttatcg 1020
 aaaaagaagt atatcctaag gtcgtagga agcagtgaca gagctttgcg gtccaggaca 1080
 aaagaaaaac ctaaagcgcc tgaaccaagt agtaactcag ttgacgttaa taatgatgga 1140
 gtgaagaagg aaaaaggag gaagaagaaa aaaacaggag gggaggaaaa gaaggatcaa 1200
 tttataaaa tcaaagctca cctaagatat ttattaaaca gagtaagcta tgagcaaac 1260
 ttaattgatg ctactctgg tgacggctgg aaaggataca gtaggacaa gttaaagcct 1320
 gataaggaga ttcaacgggc taaatctgaa attcttcgac gtaaattgaa aattagggat 1380
 ctatttcaaa acttggattc cttgtgtgct gaaggaaagt ttccagaatc tttgtttgat 1440
 tctgagggag agattgacag tgaggatata ttctgttcaa ttgcccagac caaagagttg 1500
 agcactgata atgatataat tctctgtgat ggtgcttgtg accgtggatt tcaccagcac 1560
 tgcttggatc ctccattgtt aactgaagac attccacctg gcgatgaggg ttggttatgc 1620
 cccggatgtg attgcaaaga tgactgcgtt gaccttctta acgactcctt aggaacacgt 1680
 ctctctctta gtgacacctg ggagaagggt tttcctgaag cagcagctgt tactggaat 1740
 aatatggatc agggacttcc ttctgatgat tctgatgatg atgattataa ccctaaggt 1800
 aaagaggatg tggaggtga aggaggtgaa tcaagttctg atgaatctga atatgcttct 1860
 gcttctgaga aattggaaga ttcacacatc gaggatccat acttaggctt tccttctgaa 1920
 gattcagagg acaatgatta tgatcctagt gctccagatc ttgacaataa agtcacagaa 1980
 gaaagttcaa gctctgattt cacatctgac tctgaggatc ttgctgctgc tattgaggat 2040
 aacacgtccc ctggacatgt taaacagact aaaagcagac aaaaaagtaa agttggtaag 2100
 aacctatcaa tggttgatga agtgtcatct ttacgagagc cagaacttga ggaagaggg 2160

ES 2 558 745 T3

```

ttcactccag tttctgggaa aagaaatgtg gaaaggctgg actacaaaaa gctgtatgat 2220
gagacataca aaagtgatac tagtgaagat gaagaatgga ctgccagtgc cactccaagt 2280
agaaaaaaga aactctgtgg caaaatgact ccagtgtcat cagatggaaa ggcctcaaat 2340
aattctagac atactcctga aaggaacacc caacaagata aagttgaaaa tacaataat 2400
tcaccacta aatcacttga aggctgtctg gaatctggct caagggataa aaagcctagg 2460
tcttcaacac gtcaaagact tggagacggt gtagtgcaga gacttcacaa atcttttaa 2520
gacaatcagt atccggatcg agccacaaaa gaaagcttgg cagaagaact ggggctcact 2580
ttttccagg ttgacaaatg gtttggaat gctcgatggg gctttcggcg ttcacgcgt 2640
atgggagcta gtccaggcga atatgcttcg ccacaggcca ctggcagcgg acctgaaaat 2700
actggggaga gggaaacgca attagcttcc caggaagtcg gtgaagaaaa attgaaaacc 2760
cccagtccta gaaaaaggaa gcatttgtct gaaccgcagg catctgaagc accgcagatt 2820
attgttcttg gcttagcagc atctcctggc tcaccacggg ctcatgcagg caataaaaag 2880
aagacagtga aaaggaaatg a 2901

```

<210> 200
 <211> 966
 <212> PRT
 <213> *Lotus japonicus*

 <400> 200

5

ES 2 558 745 T3

Met Ile Ser Ser Thr Gly Glu Leu Thr Ser His Asn Asn Ser Ser Ile
1 5 10 15

Glu His Met Glu Thr Gly Gln Pro Glu Leu Ser Glu Lys Thr Pro Gln
20 25 30

Ile Gly Ser Glu Cys Leu Asp Gly Glu Gln Arg Val Leu Gly Thr Val
35 40 45

Ser Thr Ser Ser Val Asn Asn Glu Asn Ser Asn Gln Val Ser Ala Asn
50 55 60

Val Thr Glu Asn Ser Val Phe Gln Leu Pro Ala Thr Pro Gln His Cys
65 70 75 80

Leu Glu Lys Ser Gly Gln Thr Val Thr Ser Pro Pro Leu Glu Glu Ser
85 90 95

Thr Leu Lys Gln Val Ala Thr Asn Val Ser Asn Gly Lys Ser Glu Asn
100 105 110

Lys Cys Gln Thr Ser Ser Gln Asp Val Gln Asn Glu Leu Val Glu Ala
115 120 125

Ser Asp Val Ile Thr Gly Ser Phe Val Ala Asp Gln Thr Gln Ser Ile
130 135 140

ES 2 558 745 T3

Pro Ala Gln Val Asn Thr Ser Ser Val Asn Glu Leu Leu Asp Ile Pro
 145 150 155 160
 Ser Gly Asp Val Ala Val Glu Val Gly Ser Asp Ser Leu Glu Lys Lys
 165 170 175
 Ser Lys Gly Cys Leu Val Asp Ser Glu Gln Arg Val Leu Gly Thr Val
 180 185 190
 Leu Thr Asn Ser Ile Asn Asp Gly Asn Ser Asn Gln Val Ser Ala Asn
 195 200 205
 Val Thr Glu Asn Ser Val Ile Gln Leu Pro Ala Pro Pro Gln His Cys
 210 215 220
 Phe Glu Lys Ser Gly Gln Thr Val Lys Ser Ser Ser Leu Glu Glu Ser
 225 230 235 240
 Thr Leu Lys Gln Val Ser Thr Asn Leu Ser Asn Ser Lys Ser Glu Asn
 245 250 255
 Ile Cys Gln Thr Ser Ser Gln Asp Val Gln Asn Glu Leu Val Glu Thr
 260 265 270
 Ser Asp Ala Val Thr Gly Ser Val Val Ala Asp Lys Thr Gln Ser Ile
 275 280 285
 Pro Ala Gln Val Asn Thr Ser Ser Val Asn Glu Leu Leu Asp Ile Pro
 290 295 300
 Ser Arg Asp Val Ala Glu Glu Val Gly Ser Asp Ser Ser Glu Arg Lys
 305 310 315 320
 Ser Lys Ser Thr Thr Pro Ser Gln Leu Arg His Ile Gly Lys Ser Asn
 325 330 335
 Ser Lys Leu Ser Lys Lys Lys Tyr Ile Leu Arg Ser Leu Gly Ser Ser
 340 345 350
 Asp Arg Ala Leu Arg Ser Arg Thr Lys Glu Lys Pro Lys Ala Pro Glu
 355 360 365
 Pro Ser Ser Asn Ser Val Asp Val Asn Asn Asp Gly Val Lys Lys Glu
 370 375 380
 Lys Arg Arg Lys Lys Lys Lys Thr Gly Gly Glu Glu Lys Lys Asp Gln
 385 390 395 400
 Phe Tyr Lys Ile Lys Ala His Leu Arg Tyr Leu Leu Asn Arg Val Ser
 405 410 415

ES 2 558 745 T3

Tyr Glu Gln Asn Leu Ile Asp Ala Tyr Ser Gly Asp Gly Trp Lys Gly
 420 425 430

Tyr Ser Met Asp Lys Leu Lys Pro Asp Lys Glu Ile Gln Arg Ala Lys
 435 440 445

Ser Glu Ile Leu Arg Arg Lys Leu Lys Ile Arg Asp Leu Phe Gln Asn
 450 455 460

Leu Asp Ser Leu Cys Ala Glu Gly Lys Phe Pro Glu Ser Leu Phe Asp
 465 470 475 480

Ser Glu Gly Glu Ile Asp Ser Glu Asp Ile Phe Cys Ser Ile Cys Gln
 485 490 495

Thr Lys Glu Leu Ser Thr Asp Asn Asp Ile Ile Leu Cys Asp Gly Ala
 500 505 510

Cys Asp Arg Gly Phe His Gln His Cys Leu Asp Pro Pro Leu Leu Thr
 515 520 525

Glu Asp Ile Pro Pro Gly Asp Glu Gly Trp Leu Cys Pro Gly Cys Asp
 530 535 540

Cys Lys Asp Asp Cys Val Asp Leu Leu Asn Asp Ser Leu Gly Thr Arg
 545 550 555 560

Leu Ser Leu Ser Asp Thr Trp Glu Lys Val Phe Pro Glu Ala Ala Ala
 565 570 575

Val Thr Gly Asn Asn Met Asp Gln Gly Leu Pro Ser Asp Asp Ser Asp
 580 585 590

Asp Asp Asp Tyr Asn Pro Asn Gly Lys Glu Asp Val Glu Val Glu Gly
 595 600 605

Gly Glu Ser Ser Ser Asp Glu Ser Glu Tyr Ala Ser Ala Ser Glu Lys
 610 615 620

Leu Glu Asp Ser His His Glu Asp Pro Tyr Leu Gly Leu Pro Ser Glu
 625 630 635 640

Asp Ser Glu Asp Asn Asp Tyr Asp Pro Ser Ala Pro Asp Leu Asp Asn
 645 650 655

Lys Val Thr Glu Glu Ser Ser Ser Ser Asp Phe Thr Ser Asp Ser Glu
 660 665 670

Asp Leu Ala Ala Ala Ile Glu Asp Asn Thr Ser Pro Gly His Val Lys
 675 680 685

ES 2 558 745 T3

Gln Thr Lys Ser Arg Gln Lys Ser Lys Val Gly Lys Asn Pro Ser Met
 690 695 700

Val Asp Glu Val Ser Ser Leu Arg Glu Pro Glu Leu Glu Glu Glu Gly
 705 710 715 720

Phe Thr Pro Val Ser Gly Lys Arg Asn Val Glu Arg Leu Asp Tyr Lys
 725 730 735

Lys Leu Tyr Asp Glu Thr Tyr Lys Ser Asp Thr Ser Glu Asp Glu Glu
 740 745 750

Trp Thr Ala Ser Ala Thr Pro Ser Arg Lys Lys Lys Leu Cys Gly Lys
 755 760 765

Met Thr Pro Val Ser Ser Asp Gly Lys Ala Ser Asn Asn Ser Arg His
 770 775 780

Thr Pro Glu Arg Asn Thr Gln Gln Asp Lys Val Glu Asn Thr Asn Asn
 785 790 795 800

Ser Pro Thr Lys Ser Leu Glu Gly Cys Leu Glu Ser Gly Ser Arg Asp
 805 810 815

Lys Lys Pro Arg Ser Ser Thr Arg Gln Arg Leu Gly Asp Val Val Val
 820 825 830

Gln Arg Leu His Lys Ser Phe Lys Asp Asn Gln Tyr Pro Asp Arg Ala
 835 840 845

Thr Lys Glu Ser Leu Ala Glu Glu Leu Gly Leu Thr Phe Phe Gln Val
 850 855 860

Asp Lys Trp Phe Gly Asn Ala Arg Trp Gly Phe Arg Arg Ser Ser Arg
 865 870 875 880

Met Gly Ala Ser Pro Gly Glu Tyr Ala Ser Pro Gln Ala Thr Gly Ser
 885 890 895

Gly Pro Glu Asn Thr Gly Glu Arg Glu Arg Glu Leu Ala Ser Gln Glu
 900 905 910

Val Gly Glu Glu Lys Leu Lys Thr Pro Ser Pro Arg Lys Arg Lys His
 915 920 925

Leu Ser Glu Pro Gln Ala Ser Glu Ala Pro Gln Ile Ile Val Leu Gly
 930 935 940

Leu Ala Ala Ser Pro Gly Ser Pro Arg Ala His Ala Gly Asn Lys Lys
 945 950 955 960

Lys Thr Val Lys Arg Lys

ES 2 558 745 T3

965

5 <210> 201
 <211> 2379
 <212> ADN
 <213> *Oryza sativa*
 <400> 201

atggataaga caactacttc tgatcttgtt ttggacaacg acaacattgg gagtaatgct 60
 ggttctgcac aggagcctct tactaccaat ggtaagacta gtggggtcag aatagatac 120
 aaacaaactg ttaaaagggg aaggaaaggc tcccaaatat caccaagtaa aacatatccc 180
 ctgagatctt cacatagtaa tgtcaggggtg cttcgctctg cctcaaaaaa gaagaatgag 240
 acacctattg tgcccacaaa tgataaact gctgttcaac gagttgcgaa gaaaaggaaa 300
 agaagcaaac ctttgagacc tgcacctagc aggggtgcttc gctctacctc ggaaaagaag 360
 aataaggcac ataatgagct tctaaatgat ggtgctgggtg ttcaaccagc cgaaaagaaa 420
 agaaaggtag gcagaccccc aaaaggagga actcccaagg acgattacct catgatccga 480
 aagcgagtta gatatgtttt gaatcgaatg aactatgaac aaagttaaat tcaagcctat 540
 gccagtgaag gttggaaagg tcaaagcttg gaaaaataa gacctgagaa ggagcttgaa 600
 cgagccaagg tggaaattct gcgatgcaag tcaagaatac gggaaagcttt tcggaatttg 660
 gattccctcc tatctgaggg aaagctggat gaatctatgt ttgattctgc tggagaaata 720
 tctagtgaag atattttctg tgctgcatgt ggttcaaagg atgttacatt aaaaaatgac 780
 ataattcttt gtgatggaat ctgcgacaga gggttccacc agtattgttt aaacctcct 840
 ttgcttgctg aagatattcc acagggggat gaaggatggc tttgccctgc atgtgattgc 900
 aaaattgact gcatagatgt attaaatgaa ctccaagggg tcaactctc tatccatgac 960
 tcttgggaga aagtttttcc tgaggcagca tcttttttga atggttctaa gcaaattgat 1020
 gcgtccgatc ttccatcaga tgattcagca gacaatgatt atgaccctac tttggctcaa 1080
 gggcacaagg tggatgaaga aaaatcttct ggagaagatg gtggtgaagg attagattct 1140
 gatgactcat cttctgagga ttctgaatct tctgagaagg aaaagtctaa aacttcacag 1200
 aatggaagga cagttgatga tcttggtttg cttctgagg attcagagga tgggtgacttt 1260
 gatccagcag gtccagattc cgacaaagaa caaatgatg aatcaaactc agatcagtca 1320
 gatgaatcag attttacatc tgactctgat gatttttgtg ctgagattgc taaatcctgt 1380
 ggtcaggatg aaatctcagg cccttcatca tcacaaatca gaacagtcga tcgcaccgat 1440
 ggaagtgggt ttgatgggtga gcctaatgca gaaaattcga atcttgatt tatggagaca 1500
 gagctagagc aagacatggt gctaccaatt tctagtaaac gacaagttga acgtttgat 1560
 tacaaaaaac tgtataatga ggcttatggt aaagcatcat ctgattcaag tgacgatgaa 1620
 gagtggatg gaaatagtac acccgaaaaa ggtaatttag aagacagtga aacagattcg 1680
 cttgctgaat cacctcaggg tggaaaagga ttctctagaa gagcaccagt taggtacat 1740
 aataatgaac atactccaca gaatgtaagg cccggtgggt cagtaagcga ccagcaaat 1800

10

ES 2 558 745 T3

gaagtacttt gctccaatag caatggtagc acagccaaaa acagacattt tggctctgca 1860
attaatcaga aactgaaggc acatttcaag gaagatccat accctagtcg tgcaacaaaa 1920
gaaaatctgg cacaagagtt aggcctcaca tttaatcagg tcactaaatg gttttccagt 1980
actcgtcatt acgcaagagt ggcagccaca aaaaaggaga acaacattga aaatcatagc 2040
gctgagaata ataacaatac caacactgtt gatagcatac aactgagagg atccaatgat 2100
attgttagtg tagatagaaa tgatatggtc tccgaggaaa ggacaggaca aagtaatctc 2160
aatgaaggta ctccattaag atcagatacc agctgtgggc agagtgtggc tgtgactcct 2220
atggtacacc cagaaaacca gggcaatgat tcctcaagta atgttagaac accaaatgct 2280
aagtctgagg agaaattgat ccttgactg gaaaattcag atgaggctag gcgaaaggcg 2340
gtacaacggg aactgagaaa gatgaagact ggcagggtga 2379

<210> 202
<211> 792
<212> PRT
<213> *Oryza sativa*

<400> 202

5

ES 2 558 745 T3

Met Asp Lys Thr Thr Thr Ser Asp Leu Val Leu Asp Asn Asp Asn Ile
1 5 10 15

Gly Ser Asn Ala Gly Ser Ala Gln Glu Pro Leu Thr Thr Asn Gly Lys
20 25 30

Thr Ser Gly Val Arg Asn Arg Tyr Lys Gln Thr Val Lys Arg Gly Arg
35 40 45

Lys Gly Ser Gln Ile Ser Pro Ser Lys Thr Tyr Pro Leu Arg Ser Ser
50 55 60

His Ser Asn Val Arg Val Leu Arg Ser Ala Ser Lys Lys Lys Asn Glu
65 70 75 80

Thr Pro Ile Val Pro Thr Asn Asp Asn Thr Ala Val Gln Arg Val Ala
85 90 95

Lys Lys Arg Lys Arg Ser Lys Pro Leu Arg Pro Ala Pro Ser Arg Val
100 105 110

Leu Arg Ser Thr Ser Glu Lys Lys Asn Lys Ala His Asn Glu Leu Leu
115 120 125

Asn Asp Gly Ala Gly Val Gln Pro Ala Glu Lys Lys Arg Lys Val Gly
130 135 140

Arg Pro Pro Lys Gly Gly Thr Pro Lys Asp Asp Tyr Leu Met Ile Arg
145 150 155 160

Lys Arg Val Arg Tyr Val Leu Asn Arg Met Asn Tyr Glu Gln Ser Leu

ES 2 558 745 T3

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

ES 2 558 745 T3

Ser Asp Asp Phe Cys Ala Glu Ile Ala Lys Ser Cys Gly Gln Asp Glu
450 455 460

Ile Ser Gly Pro Ser Ser Ser Gln Ile Arg Thr Val Asp Arg Thr Asp
465 470 475 480

Gly Ser Gly Phe Asp Gly Glu Pro Asn Ala Glu Asn Ser Asn Leu Ala
485 490 495

Phe Met Glu Thr Glu Leu Glu Gln Asp Met Val Leu Pro Ile Ser Ser
500 505 510

Lys Arg Gln Val Glu Arg Leu Asp Tyr Lys Lys Leu Tyr Asn Glu Ala
515 520 525

Tyr Gly Lys Ala Ser Ser Asp Ser Ser Asp Asp Glu Glu Trp Tyr Gly
530 535 540

Asn Ser Thr Pro Glu Lys Gly Asn Leu Glu Asp Ser Glu Thr Asp Ser
545 550 555 560

Leu Ala Glu Ser Pro Gln Gly Gly Lys Gly Phe Ser Arg Arg Ala Pro
565 570 575

Val Arg Tyr His Asn Asn Glu His Thr Pro Gln Asn Val Arg Pro Gly
580 585 590

Gly Ser Val Ser Asp Gln Gln Thr Glu Val Leu Cys Ser Asn Ser Asn
595 600 605

Gly Ser Thr Ala Lys Asn Arg His Phe Gly Pro Ala Ile Asn Gln Lys
610 615 620

Leu Lys Ala His Phe Lys Glu Asp Pro Tyr Pro Ser Arg Ala Thr Lys
625 630 635 640

Glu Asn Leu Ala Gln Glu Leu Gly Leu Thr Phe Asn Gln Val Thr Lys
645 650 655

Trp Phe Ser Ser Thr Arg His Tyr Ala Arg Val Ala Ala Thr Lys Lys
660 665 670

Glu Asn Asn Ile Glu Asn His Thr Ala Glu Asn Asn Asn Asn Thr Asn
675 680 685

Thr Val Asp Ser Ile Gln Leu Arg Gly Ser Asn Asp Ile Val Ser Val
690 695 700

Asp Arg Asn Asp Met Val Ser Glu Glu Arg Thr Gly Gln Ser Asn Leu
705 710 715 720

ES 2 558 745 T3

Asn Glu Gly Thr Pro Leu Arg Ser Asp Thr Ser Cys Gly Gln Ser Val
 725 730 735

Ala Val Thr Pro Met Val His Pro Glu Asn Gln Gly Asn Asp Ser Ser
 740 745 750

Ser Asn Val Arg Thr Pro Asn Ala Lys Ser Ala Glu Lys Leu Ile Pro
 755 760 765

Gly Leu Glu Asn Ser Asp Glu Ala Arg Arg Lys Ala Val Gln Arg Glu
 770 775 780

Leu Arg Lys Met Lys Thr Gly Arg
 785 790

- <210> 203
- <211> 3267
- <212> ADN
- <213> *Petroselinum crispum*
- <400> 203

5

```

atggaagaga tctcagaccc taaaccaaat gcattagagc aagtacttcc tactgtccca    60
aatggtaa atgtactgcacc ggtgcaa atg gaaagtttgg ctggtgatgt tcaaaaagta    120
tcagggtgaag caaagggtgag gatatgttcc tgttgggtgcg agatagtacg ttctccagaa    180
gatttaacta aacttgtgcc ttgta atgat tttgcagaag atataaaatt gtttgattct    240
gatcccatgc aacaagaagc tgaagctcc attggaatac cattgatccc taaacaagtt    300
acaatgtccc ataaccatga ccatgaatct ggttctgaaa tggtcagtaa cgaagta atg    360
caagaaaacc atgtaattgc cacagaaaat acttaccaga aatctgattt tgaccgtata    420
aatatgggac aaaaagaaac catgccagaa gaagtgatcc ataagctttt cttggaatcc    480
tcaacttcct cgatagatat actatta ataat aatcataaca gttatcaatc aggactgcc a    540
ccagaga atg cagtaacaga ttgcaagcaa gtccaattgg gacaccgaag tgatgatgct    600
ataaaga atct ctggtcttgt ggaattggta ataggtcaga aaaatgttgc caagagtcct    660
agccagttgg tagaaacggg gaaaagaggt cgtggtcggc ccaggaaagt acaaactggc    720
cttgagcaat tggta atagg tcagaaaact gctgccaaga gctctagcca gttaggtgac    780
acagggaaaa gatctcgtgg tcggcccaga aaagtacaaa attctccaac ttctttcttg    840
gagaatataa atatggaaca aaaagaaaca ataccagaac aagtgaccca aaattctatc    900
ttggagtcct tgactatccc gacagacaat cagtcaagaa cttataacag tgatcaatca    960
gaactgccac cagaga atgc agcaaaaaat tgcaatcacg cccaatttgg acatcaaagt    1020
gatgatacta caaagatttc tggattcaag gaattggtaa taggtcagga aacagttgca    1080
aagagtccta gccagttggg agacgcagga aaaagaggtc gtggtcggcc cagaaaagta    1140
caaactggtc ttgagcaatt ggtaccagtt caggaaactg cggccaagag ctctagccag    1200
ttgggagaca cagggaaaag atctcgcggc cggcccagga aagtacaaga ttctccaact    1260
    
```

10

ES 2 558 745 T3

tctttagggg gtaatgttaa agtggtagca gagaaaggaa aagattctca ggaattgtcg 1320
gtaaacagca gcagatcatt gcgctcaagg tcacaagaga aatctatcga acccgatgtc 1380
aataatattg tggcagatga aggtgctgac agagaaaaac caaggaagaa gaggaaaaag 1440
cgaatggagg aaaatagggt tgatgaattc tgtagaataa ggacacacct aaggtagtta 1500
ctacacagaa taaaatacga gaagaatttc cttgatgctt actctgggga gggatggaag 1560
ggacaaagcc tagacaaaat aaaacctgag aaggagctta aacgagcaaa agctgaaatt 1620
tttgacgca aattgaaaat tagagatcta tttcagcgcg tagatttagc acgttccgag 1680
ggaaggcttc cagaaattct atttgattct cgaggagaaa ttgacagcga ggatatattc 1740
tgcgcgaaat gtggatccaa ggatgtgaca ctacagtaatg atatcatcct ttgcatggtt 1800
gcatgtgatc gtggatttca ccaattttgt ttggaccac cattgctgaa agaatatatt 1860
cctcctgacg acgaggggtg gctttgcccg ggatgagcgt gcaaaattga ctgcatcaag 1920
ttacttaatg attcccaaga gacaaatatt ttactcgggtg acagctggga gaaagttttt 1980
gctgaggagg cagcagcagc agcttccgga aaaaatctag atgataattc tggactgcct 2040
tcagatgatt ccgaggatga tgattatgat cctgggggccc ctgatttgga tgagaagggtg 2100
caggagatg attcagatc ggatgaatct gattatcagt cggaatcaga tgatatgcaa 2160
gttatacgtc aaaaaaattc gcgcgggctt cttcagatg attccgaaga tgatgaatac 2220
gatcctagtg gcttagttac tgatcagatg tataaggaca gctcatgttc tgattttaca 2280
tctgattctg aggattttac aggtgtgttt gacgattata aggatactgg taaagctcaa 2340
gggccactgg catcgacccc agatcacgtc aggaacaatg aggaaggggtg tgggcatcct 2400
gagcaagggtg atactgcccc cctgtatccg agaagacaag tagagagttt ggactacaaa 2460
aagctaaatg atattgagtt ctctaaaatg tgtgacatcc ttgatattct ttcaagtcaa 2520
cttgatgtaa ttatttgtac tggtaaccag gaagagatg ggaacacatc ttctgattca 2580
agtgatgaag atttataggt tacttcttca ccagataaaa ataattctga taaagaagct 2640
acagcaatgg agcgtggaag agagagcggg gatttggagc tagaccaaaa agcgagagaa 2700
tccactcata ataggagata tataaagaag tttgctgttg aaggcacaga ttcttttctt 2760
tctagatcgt gtgaagactc tgcagcacct gttgctggta gtaaaagtac ttcaaaaaca 2820
ttgcacgggg aacatgcaac tcagagactc ctccaatcct tcaaggaaaa tcaataacct 2880
caacgagctg tgaaagagag ttggctgca gaattagcac tttcagtcg acaggttagc 2940
aactggttta ataatacagc ttggagcttt cgccactcat cacgtattgg atcgatgta 3000
gctaaatttg attcaaacga tactccccgt cagaaaagca ttgacatgct tggaccagc 3060
ttgaaatcgg tcttgacagc tgctacttac agtgaaattg aaaagaagga acaagatac 3120
gcaagccttg gtttgacaga aggttgtagc agatacatga cgctgaatat ggttgagat 3180
gaaggcaatg tgcacactcc ttgtattgca gaaactaggg aagaaaaaac cgaagtcggt 3240
ataagccac aacagaacct actataa 3267

<210> 204
<211> 1088
<212> PRT
<213> *Petroselinum crispum*

5

<400> 204

ES 2 558 745 T3

Met Glu Glu Ile Ser Asp Pro Lys Pro Asn Ala Leu Glu Gln Val Leu
1 5 10 15

Pro Thr Val Pro Asn Gly Lys Cys Thr Ala Pro Val Gln Met Glu Ser
20 25 30

Leu Ala Val Asp Val Gln Lys Val Ser Gly Glu Ala Lys Val Arg Ile
35 40 45

Cys Ser Cys Trp Cys Glu Ile Val Arg Ser Pro Glu Asp Leu Thr Lys
50 55 60

Leu Val Pro Cys Asn Asp Phe Ala Glu Asp Ile Lys Leu Phe Asp Ser
65 70 75 80

Asp Pro Met Gln Gln Glu Ala Glu Ser Ser Ile Gly Ile Pro Leu Ile
85 90 95

Pro Lys Gln Val Thr Met Ser His Asn His Asp His Glu Ser Gly Ser
100 105 110

Glu Met Val Ser Asn Glu Val Met Gln Glu Asn His Val Ile Ala Thr
115 120 125

Glu Asn Thr Tyr Gln Lys Ser Asp Phe Asp Arg Ile Asn Met Gly Gln
130 135 140

Lys Glu Thr Met Pro Glu Glu Val Ile His Lys Ser Phe Leu Glu Ser
145 150 155 160

Ser Thr Ser Ser Ile Asp Ile Leu Leu Asn Asn His Asn Ser Tyr Gln
165 170 175

Ser Gly Leu Pro Pro Glu Asn Ala Val Thr Asp Cys Lys Gln Val Gln
180 185 190

Leu Gly His Arg Ser Asp Asp Ala Ile Lys Asn Ser Gly Leu Val Glu
195 200 205

Leu Val Ile Gly Gln Lys Asn Val Ala Lys Ser Pro Ser Gln Leu Val
210 215 220

Glu Thr Gly Lys Arg Gly Arg Gly Arg Pro Arg Lys Val Gln Thr Gly
225 230 235 240

Leu Glu Gln Leu Val Ile Gly Gln Lys Thr Ala Ala Lys Ser Ser Ser
245 250 255

ES 2 558 745 T3

Gln Leu Gly Asp Thr Gly Lys Arg Ser Arg Gly Arg Pro Arg Lys Val
 260 265 270

Gln Asn Ser Pro Thr Ser Phe Leu Glu Asn Ile Asn Met Glu Gln Lys
 275 280 285

Glu Thr Ile Pro Glu Gln Val Thr Gln Asn Ser Ile Leu Glu Ser Leu
 290 295 300

Thr Ile Pro Thr Asp Asn Gln Ser Arg Thr Tyr Asn Ser Asp Gln Ser
 305 310 315 320

Glu Leu Pro Pro Glu Asn Ala Ala Lys Asn Cys Asn His Ala Gln Phe
 325 330 335

Gly His Gln Ser Asp Asp Thr Thr Lys Ile Ser Gly Phe Lys Glu Leu
 340 345 350

Val Ile Gly Gln Glu Thr Val Ala Lys Ser Pro Ser Gln Leu Val Asp
 355 360 365

Ala Gly Lys Arg Gly Arg Gly Arg Pro Arg Lys Val Gln Thr Gly Leu
 370 375 380

Glu Gln Leu Val Pro Val Gln Glu Thr Ala Ala Lys Ser Ser Ser Gln
 385 390 395 400

Leu Gly Asp Thr Gly Lys Arg Ser Arg Gly Arg Pro Arg Lys Val Gln
 405 410 415

Asp Ser Pro Thr Ser Leu Gly Gly Asn Val Lys Val Val Pro Glu Lys
 420 425 430 435

Gly Lys Asp Ser Gln Glu Leu Ser Val Asn Ser Ser Arg Ser Leu Arg
 435 440 445

Ser Arg Ser Gln Glu Lys Ser Ile Glu Pro Asp Val Asn Asn Ile Val
 450 455 460

Ala Asp Glu Gly Ala Asp Arg Glu Lys Pro Arg Lys Lys Arg Lys Lys
 465 470 475 480

Arg Met Glu Glu Asn Arg Val Asp Glu Phe Cys Arg Ile Arg Thr His
 485 490 495

Leu Arg Tyr Leu Leu His Arg Ile Lys Tyr Glu Lys Asn Phe Leu Asp
 500 505 510

Ala Tyr Ser Gly Glu Gly Trp Lys Gly Gln Ser Leu Asp Lys Ile Lys
 515 520 525

ES 2 558 745 T3

Pro Glu Lys Glu Leu Lys Arg Ala Lys Ala Glu Ile Phe Gly Arg Lys
 530 535 540
 Leu Lys Ile Arg Asp Leu Phe Gln Arg Leu Asp Leu Ala Arg Ser Glu
 545 550 555 560
 Gly Arg Leu Pro Glu Ile Leu Phe Asp Ser Arg Gly Glu Ile Asp Ser
 565 570 575
 Glu Asp Ile Phe Cys Ala Lys Cys Gly Ser Lys Asp Val Thr Leu Ser
 580 585 590
 Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asp Arg Gly Phe His Gln
 595 600 605
 Phe Cys Leu Asp Pro Pro Leu Leu Lys Glu Tyr Ile Pro Pro Asp Asp
 610 615 620
 Glu Gly Trp Leu Cys Pro Gly Cys Glu Cys Lys Ile Asp Cys Ile Lys
 625 630 635 640
 Leu Leu Asn Asp Ser Gln Glu Thr Asn Ile Leu Leu Gly Asp Ser Trp
 645 650 655
 Glu Lys Val Phe Ala Glu Glu Ala Ala Ala Ala Ala Ser Gly Lys Asn
 660 665 670
 Leu Asp Asp Asn Ser Gly Leu Pro Ser Asp Asp Ser Glu Asp Asp Asp
 675 680 685
 Tyr Asp Pro Gly Gly Pro Asp Leu Asp Glu Lys Val Gln Gly Asp Asp
 690 695 700
 Ser Ser Thr Asp Glu Ser Asp Tyr Gln Ser Glu Ser Asp Asp Met Gln
 705 710 715 720
 Val Ile Arg Gln Lys Asn Ser Arg Gly Leu Pro Ser Asp Asp Ser Glu
 725 730 735
 Asp Asp Glu Tyr Asp Pro Ser Gly Leu Val Thr Asp Gln Met Tyr Lys
 740 745 750
 Asp Ser Ser Cys Ser Asp Phe Thr Ser Asp Ser Glu Asp Phe Thr Gly
 755 760 765
 Val Phe Asp Asp Tyr Lys Asp Thr Gly Lys Ala Gln Gly Pro Leu Ala
 770 775 780
 Ser Thr Pro Asp His Val Arg Asn Asn Glu Glu Gly Cys Gly His Pro
 785 790 795 800

ES 2 558 745 T3

Glu Gln Gly Asp Thr Ala Pro Leu Tyr Pro Arg Arg Gln Val Glu Ser
 805 810 815
 Leu Asp Tyr Lys Lys Leu Asn Asp Ile Glu Phe Ser Lys Met Cys Asp
 820 825 830
 Ile Leu Asp Ile Leu Ser Ser Gln Leu Asp Val Ile Ile Cys Thr Gly
 835 840 845
 Asn Gln Glu Glu Tyr Gly Asn Thr Ser Ser Asp Ser Ser Asp Glu Asp
 850 855 860
 Tyr Met Val Thr Ser Ser Pro Asp Lys Asn Asn Ser Asp Lys Glu Ala
 865 870 875 880
 Thr Ala Met Glu Arg Gly Arg Glu Ser Gly Asp Leu Glu Leu Asp Gln
 885 890 895
 Lys Ala Arg Glu Ser Thr His Asn Arg Arg Tyr Ile Lys Lys Phe Ala
 900 905 910
 Val Glu Gly Thr Asp Ser Phe Leu Ser Arg Ser Cys Glu Asp Ser Ala
 915 920 925
 Ala Pro Val Ala Gly Ser Lys Ser Thr Ser Lys Thr Leu His Gly Glu
 930 935 940
 His Ala Thr Gln Arg Leu Leu Gln Ser Phe Lys Glu Asn Gln Tyr Pro
 945 950 955 960
 Gln Arg Ala Val Lys Glu Ser Leu Ala Ala Glu Leu Ala Leu Ser Val
 965 970 975
 Arg Gln Val Ser Asn Trp Phe Asn Asn Arg Arg Trp Ser Phe Arg His
 980 985 990
 Ser Ser Arg Ile Gly Ser Asp Val Ala Lys Phe Asp Ser Asn Asp Thr
 995 1000 1005
 Pro Arg Gln Lys Ser Ile Asp Met Ser Gly Pro Ser Leu Lys Ser
 1010 1015 1020
 Val Leu Asp Ser Ala Thr Tyr Ser Glu Ile Glu Lys Lys Glu Gln
 1025 1030 1035
 Asp Thr Ala Ser Leu Gly Leu Thr Glu Gly Cys Asp Arg Tyr Met
 1040 1045 1050
 Thr Leu Asn Met Val Ala Asp Glu Gly Asn Val His Thr Pro Cys
 1055 1060 1065
 Ile Ala Glu Thr Arg Glu Glu Lys Thr Glu Val Gly Ile Lys Pro

ES 2 558 745 T3

1070

1075

1080

Gln Gln Asn Pro Leu
1085

5

<210> 205
<211> 2160
<212> ADN
<213> *Zea mays*

<400> 205

```

atggaaaaaa atatagctca ctgtcctggt gagggcaatg gtgagattga aaatggtgcg      60
agctctagtc agaatcctga gtcctctgag cattcagttt tgctgtctac atctcaaacg     120
atgccaaata acttggggat aaggaaaaac tataaaaggg cagcaaacag aggtaaaaaa     180
ggttcccaag gactaacagg ccaagcatat actttgatgt catctaatag tgatgtcagg     240
gtgcttcgct ccacatcaag ttcaaagaca acatctactg agcatgttca ggctccagta     300
caaccggctg ccaagagaag aaaaatgagc agagcctcaa acaaaagttc gactgatgag     360
ttttcacaaa ttcgtaaacg agttagatat attttgaacc gaatgaatta tgagcaaagc     420
ctaattgaag cttatgctag cgaaggctgg aaaaatcaaa gtttggataa gataagacct     480
gagaaggagc ttgaacgagc caaatcagaa atcttacgat gcaaattgag aatacgagaa     540
gtcttccgga atattgattc tcttctctcc aaggggaaaa ttgacgaaac tttatttgat     600
tctgaaggag aaatatcttg cgaagatadc tttgttcca cttgtggttc gaatgacgct     660
acattgggta atgatatcat tctctcgcat ggagcttgtg acagagggtt ccaccagaac     720
tgtttaaatc ctcccctgcg tactgaagat atcccatggg gagatgaagg atggctctgc     780
ccagcatgtg attgcaagat agattgtata gatctaataa atgaactcca tgggagtaac     840
atctccattg aggactcttg ggagaaagtt tttccagatg cagctgctat ggcaaatgac     900
tctaaacaag atgacgcatt tgatcttcca tcagatgact cagacgacaa tgactttgac     960
cccaatatgc ctgaagagca tgtggttggg aaggatgaag aatcgtctga agaggatgag    1020
gatggaggat cagactctga tgactcagac tttttgacat gttctgatga ttcggaacct    1080
ttgatagaca agaaggttga tgacctcaga ttaccttctg aagattcaga ggatgatgac    1140
tatgatccag caggtcctga ttcagataaa gatgtcgaag agaagtcaag ttctgatgaa    1200
tctgacttca catctgactc agatgathtt tgtaaggaga tttcaaaatc tggacatgat    1260
gaagtttcat cacctctatt acccgatgcc aagggtgggtg atatggaaaa gattactgct    1320
caagctaaaa caacaagttc tgctgacgac cctatggaga ctgaaataga ccagggtgtg    1380
gttttaccag attcaaggag acggcaagct gaacggttgg actacaaaaa gttgtatgat    1440
gaggcatatg gtgaagcatc atctgattct agtgatgatg aagaatggtc tgggaagaac    1500
acgccataaa taaaaagcaa tgaagagggg gaagccaatt ctctgcggg aaaaggctcc    1560
agagttgtgc accataatga tgaactgact acacaaagca ctaaaaaaag cttgcattct    1620
atccatgggt cagtagatga gaaacctgga gaccttactt ctaatggcag caacagcaca    1680

```

10

ES 2 558 745 T3

gccaggaaag gacattttgg cccagtaatt aatcagaagc tacatgaaca ttttaagaca 1740
caaccatacc ctagtcgttc tgtaaagaa agtctggcag aagaactagg gttaacattt 1800
cgtcaggtta acaaatggtt tgaaactagg cgtcattccg caagggtagc ttcttccagg 1860
aaaggcatca gtctagataa gcatagcccc caaatacta atagtcaagt gacagctagt 1920
atggaaccta aagaacctga ggggactgtg gtggaagaat ctaatgtatg cctaaatggg 1980
ggactacca tctctaaaga agcagtgtct tcaaagtgtg gatcaagaac ccctggaagt 2040
gatgtgggag ggtcaaaagt tgattctgct gaggatcaga atccggggcc tgatcttgca 2100
gagaaggcaa ggcaaaaggc aatacagcag gagttgagga agaaaaaat gggacgatga 2160

<210> 206
<211> 719
<212> PRT
<213> *Zea mays*

<400> 206

5

ES 2 558 745 T3

Met Glu Lys Asn Ile Ala His Cys Pro Val Glu Gly Asn Gly Glu Ile
 1 5 10 15

Glu Asn Gly Ala Ser Ser Ser Gln Asn Pro Glu Ser Leu Glu His Ser
 20 25 30

Val Leu Leu Ser Thr Ser Gln Thr Met Pro Asn Asn Leu Gly Ile Arg
 35 40 45

Lys Asn Tyr Lys Arg Ala Ala Asn Arg Gly Lys Lys Gly Ser Gln Gly
 50 55 60

Leu Thr Gly Gln Ala Tyr Thr Leu Met Ser Ser Asn Ser Asp Val Arg
 65 70 75 80

Val Leu Arg Ser Thr Ser Ser Ser Lys Thr Thr Ser Thr Glu His Val
 85 90 95

Gln Ala Pro Val Gln Pro Ala Ala Lys Arg Arg Lys Met Ser Arg Ala
 100 105 110

Ser Asn Lys Ser Ser Thr Asp Glu Phe Ser Gln Ile Arg Lys Arg Val
 115 120 125

Arg Tyr Ile Leu Asn Arg Met Asn Tyr Glu Gln Ser Leu Ile Glu Ala
 130 135 140

Tyr Ala Ser Glu Gly Trp Lys Asn Gln Ser Leu Asp Lys Ile Arg Pro
 145 150 155 160

Glu Lys Glu Leu Glu Arg Ala Lys Ser Glu Ile Leu Arg Cys Lys Leu
 165 170 175

Arg Ile Arg Glu Val Phe Arg Asn Ile Asp Ser Leu Leu Ser Lys Gly

ES 2 558 745 T3

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

ES 2 558 745 T3

Ser Arg Arg Arg Gln Ala Glu Arg Leu Asp Tyr Lys Lys Leu Tyr Asp
465 470 475 480

Glu Ala Tyr Gly Glu Ala Ser Ser Asp Ser Ser Asp Asp Glu Glu Trp
485 490 495

Ser Gly Lys Asn Thr Pro Ile Ile Lys Ser Asn Glu Glu Gly Glu Ala
500 505 510

Asn Ser Pro Ala Gly Lys Gly Ser Arg Val Val His His Asn Asp Glu
515 520 525

Leu Thr Thr Gln Ser Thr Lys Lys Ser Leu His Ser Ile His Gly Ser
530 535 540

Val Asp Glu Lys Pro Gly Asp Leu Thr Ser Asn Gly Ser Asn Ser Thr
545 550 555 560

Ala Arg Lys Gly His Phe Gly Pro Val Ile Asn Gln Lys Leu His Glu
565 570 575

His Phe Lys Thr Gln Pro Tyr Pro Ser Arg Ser Val Lys Glu Ser Leu
580 585 590

Ala Glu Glu Leu Gly Leu Thr Phe Arg Gln Val Asn Lys Trp Phe Glu
595 600 605

Thr Arg Arg His Ser Ala Arg Val Ala Ser Ser Arg Lys Gly Ile Ser
610 615 620

Leu Asp Lys His Ser Pro Gln Asn Thr Asn Ser Gln Val Thr Ala Ser
625 630 635 640

Met Glu Pro Lys Glu Pro Glu Gly Thr Val Val Glu Glu Ser Asn Val
645 650 655

Cys Leu Asn Gly Gly Thr Thr Ile Ser Lys Glu Ala Val Ser Ser Lys
660 665 670

Val Gly Ser Arg Thr Pro Gly Ser Asp Val Gly Gly Ser Lys Val Asp
675 680 685

Ser Ala Glu Asp Gln Asn Pro Gly Pro Asp Leu Ala Glu Lys Ala Arg
690 695 700

Gln Lys Ala Ile Gln Gln Glu Leu Arg Lys Lys Lys Met Gly Arg
705 710 715

<210> 207
<211> 2079
<212> ADN
<213> *Zea mays*

ES 2 558 745 T3

<400> 207

atgaaaaaa	agacagttca	cgttctggct	gagggtaatg	gtgagattga	aaatggtagc	60
agctctagtc	agaatcctga	gacccttgag	catccagttt	tgctgtctgc	atctcaaact	120
gtgccaata	acttggggat	aaggaaaaac	tataaaaggg	cagcaaacag	aggtaaaaag	180
ggttccaag	gactaacagg	tcaggcatat	acattgaggt	catctgatat	tgatgtcagg	240
gtgcttcgct	ctacatcagg	ttcaaagaca	acatctactg	agcatgtgca	gcctccagga	300
caaccagctg	ccaagagaag	aaaagggagc	agagcctcaa	acaaaaattc	gacagatgag	360
ttctcaciaa	ttcgtaaacg	ggttagatac	atthtgaacc	gaatgaatta	tgagcaaagc	420
ctaattgaag	cttatgctag	tgaaggctgg	aaaaatcaaa	gtttggataa	gataagacct	480
gagaaggagc	ttgaacgagc	caaatcagaa	atcttacgat	gcaaattgag	aatacgggaa	540
gtcttccaaa	atattgattc	tcttctgtcc	aaagggaaaa	ttgacgaatc	tttattttgat	600
tctgaaggag	aaatatcttg	tgaagatata	ttttgtgcta	cttgtggttc	gaaaaatgct	660
acattgggta	atgatatacat	tctctgtgat	ggagcttggtg	acagaggggt	ccaccagaac	720
tgtctaaatc	ctcctctgcg	tactgaagat	attcccatgg	gagatgaagg	atggctctgc	780
ccagcatgtg	attgcaagtt	agactgtata	gatctaataa	atgaactcca	gggcagtgc	840
atctccattg	aggacccttg	ggagaaagtt	ttccagagg	cagctgctat	gaccaatggc	900
tctaagcaag	atgaacgatt	tgatcttcca	tctgatgatt	cggatgacaa	tgattttgac	960
cccaatatgc	ctgaagagca	tgtggctagt	aaggaaagaag	gatcttctga	agaggaagag	1020
gatgatgacg	gaggatcaga	ctctgacgac	tctgacttct	taacctgttc	tgatgatttg	1080
gaacctttga	tagataaaaa	gaaggttgat	gacctgggat	tatcttctga	agattcagag	1140
gatgatgact	atgatccagc	aggtcctgat	tcggataaag	atgtcgaaaa	gaagtcaaat	1200
tctgatgagt	ctgatttcac	gtctgactcg	gatgattttt	gtaaggagat	taaaaaatct	1260
ggtggacatg	atgaagtttc	atcgctcca	ttaccgatg	tcaagggtgg	tgatatggaa	1320
agaacactg	ctcagtctaa	cacaacaagt	tctgtgacg	accctatgga	gactgaaata	1380
gaccagagtg	tggttttacc	agtatcaagg	agacgtcaag	ctgaacggtt	ggactacaaa	1440
aggctgtatg	atgaggcata	tggtgaagca	tcacttgatt	ctagtgatga	agaagaatgg	1500
tctgggaaga	acacaccaat	aaaagcaat	gaagaggggtg	aagttggttc	tcctgcagga	1560
aaaggctcta	gagttgcgca	ccataatgaa	ctgactacac	aaaacactaa	agaaagcttg	1620
cattctcttc	atggttcagt	agatgagaaa	catggagatc	ttacttctaa	tggcagcaat	1680
atcaaagata	ggaaaggaca	ttttgggtcca	gtgattagtc	agaagctaca	tgaacatttt	1740
aagacacaac	cataccctag	tcgttctctt	aaagaaagct	tggcagaaga	actagggcta	1800
acatttcate	aggtcaacag	atggtttgaa	aatagggctc	atthttgcaag	gtttagcttct	1860
tccaggaaaag	gcatcagtc	agataagcat	agccccaaa	atactaatag	cccagtgaca	1920
cctagtatgc	aacctaaaga	accagagggg	actgtgatgg	aagaatctaa	tgatgcata	1980
aacagggatg	ctaccatctc	taaaaaagcg	gtttcttcaa	aagtcggatc	aagaaagaac	2040
cttagcaaga	gttcccctaa	aagccgggtc	aaaagttga			2079

ES 2 558 745 T3

<210> 208
 <211> 692
 <212> PRT
 <213> *Zea mays*

5

<400> 208

Met Glu Lys Lys Thr Val His Val Leu Ala Glu Gly Asn Gly Glu Ile
 1 5 10 15

Glu Asn Gly Thr Ser Ser Ser Gln Asn Pro Glu Thr Leu Glu His Pro
 20 25 30

Val Leu Leu Ser Ala Ser Gln Thr Val Pro Asn Asn Leu Gly Ile Arg
 35 40 45

Lys Asn Tyr Lys Arg Ala Ala Asn Arg Gly Lys Lys Gly Ser Gln Gly
 50 55 60

Leu Thr Gly Gln Ala Tyr Thr Leu Arg Ser Ser Asp Ile Asp Val Arg
 65 70 75 80

Val Leu Arg Ser Thr Ser Gly Ser Lys Thr Thr Ser Thr Glu His Val
 85 90 95

Gln Pro Pro Gly Gln Pro Ala Ala Lys Arg Arg Lys Arg Ser Arg Ala
 100 105 110

Ser Asn Lys Asn Ser Thr Asp Glu Phe Ser Gln Ile Arg Lys Arg Val
 115 120 125

Arg Tyr Ile Leu Asn Arg Met Asn Tyr Glu Gln Ser Leu Ile Glu Ala
 130 135 140

Tyr Ala Ser Glu Gly Trp Lys Asn Gln Ser Leu Asp Lys Ile Arg Pro
 145 150 155 160

Glu Lys Glu Leu Glu Arg Ala Lys Ser Glu Ile Leu Arg Cys Lys Leu
 165 170 175

Arg Ile Arg Glu Val Phe Gln Asn Ile Asp Ser Leu Leu Ser Lys Gly
 180 185 190

Lys Ile Asp Glu Ser Leu Phe Asp Ser Glu Gly Glu Ile Ser Cys Glu
 195 200 205

Asp Ile Phe Cys Ala Thr Cys Gly Ser Lys Asn Ala Thr Leu Gly Asn
 210 215 220

Asp Ile Ile Leu Cys Asp Gly Ala Cys Asp Arg Gly Phe His Gln Asn

ES 2 558 745 T3

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

ES 2 558 745 T3

Gly Glu Val Gly Ser Pro Ala Gly Lys Gly Ser Arg Val Ala His His
 515 520 525

Asn Glu Leu Thr Thr Gln Asn Thr Lys Glu Ser Leu His Ser Leu His
 530 535 540

Gly Ser Val Asp Glu Lys His Gly Asp Leu Thr Ser Asn Gly Ser Asn
 545 550 555 560

Ile Lys Asp Arg Lys Gly His Phe Gly Pro Val Ile Ser Gln Lys Leu
 565 570 575

His Glu His Phe Lys Thr Gln Pro Tyr Pro Ser Arg Ser Leu Lys Glu
 580 585 590

Ser Leu Ala Glu Glu Leu Gly Leu Thr Phe His Gln Val Asn Arg Trp
 595 600 605

Phe Glu Asn Arg Arg His Phe Ala Arg Leu Ala Ser Ser Arg Lys Gly
 610 615 620

Ile Ser Pro Asp Lys His Ser Pro Gln Asn Thr Asn Ser Pro Val Thr
 625 630 635 640

Pro Ser Met Gln Pro Lys Glu Pro Glu Gly Thr Val Met Glu Glu Ser
 645 650 655

Asn Val Cys Ile Asn Arg Asp Ala Thr Ile Ser Lys Lys Ala Val Ser
 660 665 670

Ser Lys Val Gly Ser Arg Lys Asn Leu Ser Lys Ser Ser Pro Lys Ser
 675 680 685

Arg Val Lys Ser
 690

<210> 209
 <211> 4731
 <212> ADN
 <213> *Zea mays*
 <400> 209

5

ES 2 558 745 T3

atgggaagga	aagggtcaga	agtatcacca	agcaaaaggt	atcccttgag	gtctgcacac	60
agtagtggtg	gagttcttcg	ctctgcctta	gccaataata	ataaggacag	tgagcctctg	120
aatgCGGctg	ctgctaccca	agcagctgtg	aaaaaaagaa	gtggcaacca	gctagatagt	180
cctaatagca	ctgtcaggct	acttcgttct	acgtcaaaaa	ataaggacga	ggcacatagc	240
gggcctctta	atgacaggat	cattgataag	ccagctgcaa	ataaaagaaa	acatgccaat	300
ccctcacagg	ttggaagtcc	cagcagcagt	gtcaggggtg	ttcgtctctg	aataaaaaat	360
aaagatatgg	cgcgcagtga	gcctcaaagt	gacaggaccg	ctggtgaggc	agctacaaat	420

ES 2 558 745 T3

aaaagaaaa acgccaattc ctcaaagggtg gggagtctga gctctggtct cagtgtgctt 480
 cggctctgctt caaaaagctaa ggatgagaca tgtatcaaac ctttaaataga tagtcctgct 540
 gttgaaccag cttcaagtaa aagaaaatgc accagtcctg caaagggtggg aagtcccagc 600
 aacagtgtca ggggtcctcg atctacttca aatatataaga atgagacatg taccgagcct 660
 ttaaatagata gtggtgctgt tgaaccagct gcaaatagaa ggaaagggtgt cagtcgttca 720
 aaggagggaa gtcccaatag cagtgtcagg gtgcttcgct ctgcctcaaa taataagatt 780
 gatgtatcta ttgagccact gaataatagt actactggtc aaccggctgc gaaaagaaga 840
 aaaaatggca gttcctttaa gagcagtgcc aggggtgcttt gctctacctc agaaaggaaa 900
 aatgaggcat ctagttagcc tttaaatgat agtactgctg ctcagccagc tgcgaggaaa 960
 aagaaagggtg gtgcaatgcc aaagactgat aacccaaaga ttggtgtcag ggctcctcgc 1020
 tctgcctcag gaaagaagaa tgagacatgt actgagcctg taaatcggag tacttctcgt 1080
 gaaccaactg tgacaaaaaa aagaaattgc aagccttcaa aggatcgcgg tcccaagaag 1140
 gactacttaa aaatttgca aagaattaaa tatatattga atcgaatgaa gtatgaacaa 1200
 gcctttattc aggcttatgc aagtgaagggt tggaaaggcc aaagtttgga aaaaataaga 1260
 cccgagaagg agcttgaacg agccaaggca gaaatcctc aatgcaagtt aagaatccgg 1320
 gaagcttttc gaaatatgga ctctctttta tcgaagggga agcttgaaga atctttgttt 1380
 gattcagcag gacaaaattc aagcgaagat atatattgtg ccatatgtgg ctcaaaggat 1440
 gttacttcac aaaatgacat cattctttgt gatggagcct gtgacagagg attccaccag 1500
 aattgtttga gccctccttt gctaactgaa gaaattcctc cgggggatga aggatggctt 1560
 tgccctgcat gtgtgtgcaa agcagactat atagatgctg taaatgaact tcaaggaagc 1620
 aaactctcca ttcatgactc atgggtgaaa gttttccctg aggcagcttc aactgccaat 1680
 ggttctaaac aagttgatgc ttctgatctg ctgccagatc atataaagga cagtgtctaac 1740
 cttgcattgg ttggaacaca catggtgaat gaaatcaggt tttctgaaga ggatgatagc 1800
 aaagcggatg acctcggctt gccttcggag gactcagggg atggtgactt cgatccagca 1860
 ggtccagatt ctagtgaaga ccaaatgat ggattgaact cagaagagtc agatttcaca 1920
 tctgactctg attatatttg tgctgagatt gctaaatctt gcggccagga tgaagtctcg 1980
 gcatctccat tatcaaatgt gataaacctg actgatagaa tgaaactcag ggcctacagt 2040
 agacgctcta atgaagaaaa ccacaatcat gcatttatgg acatggagct agagcaagac 2100
 atagtgttac cagttccaag caggcggcag gttgaacggt tggattacaa aaaactatac 2160
 gatgaggcat atgaggaaga atcatctaat tcaagtgatg aagaagagtg gtctggaaaa 2220
 gaactcttag aaggcagtgg aacagattca cttgatgagc cttcgtcctg taaaaggctc 2280
 tctagaagag caccagcagg gcagcagaat aatgaactta caccacagag cggcctcat 2340
 ggttctaaca gtgagcagaa aactgaagtt cttcgtccta atggcagtag tagcacagga 2400
 cggaagtatg atcctgaagt taccagaaa ttgaagggtc attttgagaa agatccatat 2460

ES 2 558 745 T3

cctagccgtg aaacaaaaga aaatctatca gaagaactag gcttgacatt taatcaggtc 2520
 tccaaatggt tctctagtagt tctgactat tcaaggggtg cttctgcaa aaaagaaatg 2580
 catcctgaca gccatacttg tgagaataat gatgagacaa ctgtagacag catgcaagca 2640
 aaacaaccca atgctggggg gatggaaaag ttgacggggg atagaaatgg cattgttctt 2700
 gagaaaccga tgggtgcagga caatcttaac caaggagcga tggaaaagtt aactagggat 2760
 aaaaatgaca ttgttcttga gaaaccgggtg atgcagagca gtcttaacca atgtaataat 2820
 gaagatatac cgcttttctg aacagaaatt gagatggagt cttacgaaca agaatcatct 2880
 gactcaagt atgaagagt gtccatactt agcacaccac gaaaaacaag attacaagac 2940
 catggaaaat catcagttac tgaatcactt cgttctgcaa aacgggtgtt cagaagagca 3000
 ccagctaggg agcagaataa tgaacgtatt cagagtgagc agcttcatgg ttctgcaagt 3060
 gagcagcaaa ctgaatttct tctctccaat gtcagcagtg gcaaggtctc aaaatatcat 3120
 tttggcccta tagttagtca gaagcttaag gaacattttg aaaaagatcc atatccttgc 3180
 cgtgcaacaa aagagagtct ggcacaagag ctaggactga catttaatca gatcagtaag 3240
 tggttctctg gcactcgtca ttattcaaga gatgctggtt ccaagaatca gaaacgcccg 3300
 ggagaaaata ctactgtgaa taataatagc acaacctttg atgacataca agtaatagaa 3360
 cccaatggtg ggttgataga taaaccagct gcagatgtaa atgacatgat ttctgaaaag 3420
 ttgatggtcc aaataaatct taacgaaggc attgaagaag atatcccacc gagccaatat 3480
 gccaggtgtg aagagaaaat aactatgact ccagctgcca tctcaagga agctggtctt 3540
 ccaggatatg gtcttgggtc tggagaaaac tttcttcagg tcagttcaag gaatactagc 3600
 tgtgagcaga gtgtgatact ggctcctacc gctatcgcaa gagaagttgg tccaccagga 3660
 tatgcgactg gagaaaacca gggcaatggg gctccatgga atacaagtta tgagcgaaga 3720
 ctatttatga atcctgtgac cagctcaaga gcagtcagtc ctccaggata tgggcctgaa 3780
 gaaagccgag gcagtggtcatt ttcctggaat acaagcagtg acaaaatatt gtttatgagt 3840
 cctacaacca tctcaagaga acttgggtcca ccaggatatg ggcctagaga agaccagtg 3900
 agtggcactt catggaatgc aagctgagcag ttgggaggtat ttatgagttc tgtgaccacc 3960
 tcaagagaag acagtcggcc aggatatggg cctcgagaaa accaggagaa tgacaatatg 4020
 agatgagagt tgagaatggt tacaagtcct acaaccatct ccagagaagt tggcccacca 4080
 ggatatgagc ttactggaaa aaacaagggc aacggcgctt catgtactac gaattaccag 4140
 cagggagcgt ttacgagtc tgaagccttc ccaagagaag cctgtcctcc aggatattgtg 4200
 tctggagaga accagggtaa cgacactcca tggatatatga gcagcaaaca gatggcgttt 4260
 acgagtccta cgaccaacc tagagtaggt ccgccagggg atggacagga aaaccagggc 4320
 agtcgtgttt cttggaatac gagttatgag cagggaggtt ttatgagttc tacaacatc 4380
 tcaagagaag ccggtccgcc aggttatggg cctggggaaa gtcagggcaa tatcacttca 4440
 tggatttctga cagggcggca gagaatgttt acaggtcctt caactatccc gaacgaagtc 4500
 gggcagcaag gatatggccc ttcaggaatc ccgggaacgg atggctcaag gagtatggac 4560

ES 2 558 745 T3

ttgaaagcct ttccaccagg atatggacct gctggagaaa accagggcgg tgggtgcttca 4620
 gggagcgtca taagcccccga tgctcgggtct gcagagaacg tagagttttc agacgacgct 4680
 agaaaaaagg ctatacagcg ggagctgaga cggcggcaga agttcaggtg a 4731

5 <210> 210
 <211> 1576
 <212> PRT
 <213> *Zea mays*
 <400> 210

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

ES 2 558 745 T3

Gly Ala Val Glu Pro Ala Ala Asn Arg Arg Lys Gly Val Ser Arg Ser
 225 230 235 240
 Lys Glu Gly Ser Pro Asn Ser Ser Val Arg Val Leu Arg Ser Ala Ser
 245 250 255
 Asn Asn Lys Ile Asp Val Ser Ile Glu Pro Leu Asn Asn Ser Thr Thr
 260 265 270
 Gly Gln Pro Ala Ala Lys Arg Arg Lys Asn Gly Ser Ser Phe Lys Ser
 275 280 285
 Ser Ala Arg Val Leu Cys Ser Thr Ser Glu Arg Lys Asn Glu Ala Ser
 290 295 300
 Ser Glu Pro Leu Asn Asp Ser Thr Ala Ala Gln Pro Ala Ala Arg Lys
 305 310 315 320
 Lys Lys Gly Gly Ala Met Pro Lys Thr Asp Asn Pro Lys Ile Gly Val
 325 330 335
 Arg Val Leu Arg Ser Ala Ser Gly Lys Lys Asn Glu Thr Cys Thr Glu
 340 345 350
 Pro Val Asn Arg Ser Thr Ser Arg Glu Pro Thr Val Thr Lys Lys Arg
 355 360 365
 Asn Cys Lys Pro Ser Lys Asp Arg Gly Pro Lys Lys Asp Tyr Leu Lys
 370 375 380
 Ile Cys Gln Arg Ile Lys Tyr Ile Leu Asn Arg Met Lys Tyr Glu Gln
 385 390 395 400
 Ala Phe Ile Gln Ala Tyr Ala Ser Glu Gly Trp Lys Gly Gln Ser Leu
 405 410 415
 Glu Lys Ile Arg Pro Glu Lys Glu Leu Glu Arg Ala Lys Ala Glu Ile
 420 425 430
 Leu Gln Cys Lys Leu Arg Ile Arg Glu Ala Phe Arg Asn Met Asp Ser
 435 440 445
 Leu Leu Ser Lys Gly Lys Leu Glu Glu Ser Leu Phe Asp Ser Ala Gly
 450 455 460
 Gln Ile Ser Ser Glu Asp Ile Phe Cys Ala Ile Cys Gly Ser Lys Asp
 465 470 475 480
 Val Thr Ser Gln Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asp Arg
 485 490 495

ES 2 558 745 T3

Gly Phe His Gln Asn Cys Leu Ser Pro Pro Leu Leu Thr Glu Glu Ile
500 505 510

Pro Pro Gly Asp Glu Gly Trp Leu Cys Pro Ala Cys Val Cys Lys Ala
515 520 525

Asp Tyr Ile Asp Ala Leu Asn Glu Leu Gln Gly Ser Lys Leu Ser Ile
530 535 540

His Asp Ser Trp Val Lys Val Phe Pro Glu Ala Ala Ser Thr Ala Asn
545 550 555 560

Gly Ser Lys Gln Val Asp Ala Ser Asp Leu Leu Pro Asp His Ile Lys
565 570 575

Asp Ser Ala Asn Leu Ala Leu Val Gly Thr His Met Val Asn Glu Ile
580 585 590

Arg Phe Ser Glu Glu Asp Asp Ser Lys Ala Asp Asp Leu Gly Leu Pro
595 600 605

Ser Glu Asp Ser Gly Asp Gly Asp Phe Asp Pro Ala Gly Pro Asp Ser
610 615 620

Ser Glu Asp Gln Asn Asp Gly Leu Asn Ser Glu Glu Ser Asp Phe Thr
625 630 635 640

Ser Asp Ser Asp Tyr Phe Cys Ala Glu Ile Ala Lys Ser Cys Gly Gln
645 650 655

Asp Glu Val Ser Ala Ser Pro Leu Ser Asn Val Ile Asn Arg Thr Asp
660 665 670

Arg Met Lys Leu Arg Ala Tyr Ser Arg Arg Ser Asn Glu Glu Asn His
675 680 685

Asn His Ala Phe Met Asp Met Glu Leu Glu Gln Asp Ile Val Leu Pro
690 695 700

Val Pro Ser Arg Arg Gln Val Glu Arg Leu Asp Tyr Lys Lys Leu Tyr
705 710 715 720

Asp Glu Ala Tyr Glu Glu Glu Ser Ser Asn Ser Ser Asp Glu Glu Glu
725 730 735

Trp Ser Gly Lys Glu Leu Leu Glu Gly Ser Gly Thr Asp Ser Leu Asp
740 745 750

Glu Pro Ser Ser Cys Lys Arg Leu Ser Arg Arg Ala Pro Ala Gly Gln
755 760 765

ES 2 558 745 T3

Gln Asn Asn Glu Leu Thr Pro Gln Ser Arg Pro His Gly Ser Asn Ser
770 775 780

Glu Gln Lys Thr Glu Val Leu Arg Ser Asn Gly Ser Ser Ser Thr Gly
785 790 795 800

Arg Lys Tyr Asp Pro Glu Val Thr Gln Lys Leu Lys Val His Phe Glu
805 810 815

Lys Asp Pro Tyr Pro Ser Arg Glu Thr Lys Glu Asn Leu Ser Glu Glu
820 825 830

Leu Gly Leu Thr Phe Asn Gln Val Ser Lys Trp Phe Ser Ser Thr Arg
835 840 845

His Tyr Ser Arg Val Ala Ser Ala Lys Lys Glu Met His Pro Asp Ser
850 855 860

His Thr Cys Glu Asn Asn Asp Glu Thr Thr Val Asp Ser Met Gln Ala
865 870 875 880

Lys Gln Pro Asn Ala Gly Val Met Glu Lys Leu Thr Gly Asp Arg Asn
885 890 895

Gly Ile Val Pro Glu Lys Pro Met Val Gln Asp Asn Leu Asn Gln Gly
900 905 910

Ala Met Glu Lys Leu Thr Arg Asp Lys Asn Asp Ile Val Pro Glu Lys
915 920 925

Pro Val Met Gln Ser Ser Leu Asn Gln Cys Asn Asn Glu Asp Ile Pro
930 935 940

Leu Ser Gly Thr Glu Ile Glu Met Glu Ser Tyr Glu Gln Glu Ser Ser
945 950 955 960

Asp Ser Ser Asp Glu Glu Trp Ser Ile Leu Ser Thr Pro Arg Lys Thr
965 970 975

Arg Leu Gln Asp His Gly Lys Ser Ser Val Thr Glu Ser Leu Arg Ser
980 985 990

Ala Lys Arg Cys Ser Arg Arg Ala Pro Ala Arg Glu Gln Asn Asn Glu
995 1000 1005

Arg Ile Gln Ser Glu Gln Leu His Gly Ser Ala Ser Glu Gln Gln
1010 1015 1020

Thr Glu Phe Leu Arg Ser Asn Val Ser Ser Gly Lys Val Ser Lys
1025 1030 1035

Tyr His Phe Gly Pro Ile Val Ser Gln Lys Leu Lys Glu His Phe

ES 2 558 745 T3

1040						1045						1050			
Glu	Lys	Asp	Pro	Tyr	Pro	Cys	Arg	Ala	Thr	Lys	Glu	Ser	Leu	Ala	
	1055					1060					1065				
Gln	Glu	Leu	Gly	Leu	Thr	Phe	Asn	Gln	Ile	Ser	Lys	Trp	Phe	Ser	
	1070					1075					1080				
Gly	Thr	Arg	His	Tyr	Ser	Arg	Asp	Ala	Val	Ala	Lys	Asn	Gln	Lys	
	1085					1090					1095				
Arg	Pro	Gly	Glu	Asn	Thr	Thr	Val	Asn	Asn	Asn	Ser	Thr	Thr	Phe	
	1100					1105					1110				
Asp	Asp	Ile	Gln	Val	Ile	Glu	Pro	Asn	Val	Gly	Leu	Ile	Asp	Lys	
	1115					1120					1125				
Pro	Ala	Ala	Asp	Val	Asn	Asp	Met	Ile	Ser	Glu	Lys	Leu	Met	Val	
	1130					1135					1140				
Gln	Ile	Asn	Leu	Asn	Glu	Gly	Ile	Glu	Glu	Asp	Ile	Pro	Pro	Ser	
	1145					1150					1155				
Gln	Tyr	Ala	Arg	Cys	Glu	Glu	Lys	Ile	Thr	Met	Thr	Pro	Ala	Ala	
	1160					1165					1170				
Ile	Ser	Arg	Glu	Ala	Gly	Leu	Pro	Gly	Tyr	Gly	Pro	Gly	Pro	Gly	
	1175					1180					1185				
Glu	Asn	Phe	Leu	Gln	Val	Ser	Ser	Arg	Asn	Thr	Ser	Cys	Glu	Gln	
	1190					1195					1200				
Ser	Val	Ile	Leu	Ala	Pro	Thr	Ala	Ile	Ala	Arg	Glu	Val	Gly	Pro	
	1205					1210					1215				
Pro	Gly	Tyr	Ala	Thr	Gly	Glu	Asn	Gln	Gly	Asn	Gly	Ala	Pro	Trp	
	1220					1225					1230				
Asn	Thr	Ser	Tyr	Glu	Arg	Arg	Leu	Phe	Met	Asn	Pro	Val	Thr	Ser	
	1235					1240					1245				
Ser	Arg	Ala	Val	Ser	Pro	Pro	Gly	Tyr	Gly	Pro	Glu	Glu	Ser	Arg	
	1250					1255					1260				
Gly	Ser	Gly	Ile	Ser	Trp	Asn	Thr	Ser	Ser	Glu	Gln	Ile	Leu	Phe	
	1265					1270					1275				
Met	Ser	Pro	Thr	Thr	Ile	Ser	Arg	Glu	Leu	Gly	Pro	Pro	Gly	Tyr	
	1280					1285					1290				
Gly	Pro	Arg	Glu	Asp	Gln	Cys	Ser	Gly	Thr	Ser	Trp	Asn	Ala	Ser	
	1295					1300					1305				

ES 2 558 745 T3

Cys Glu Leu Gly Val Phe Met Ser Pro Val Thr Thr Ser Arg Glu
 1310 1315 1320
 Asp Ser Pro Pro Gly Tyr Gly Pro Arg Glu Asn Gln Glu Asn Asp
 1325 1330 1335
 Asn Met Arg Cys Glu Leu Arg Met Phe Thr Ser Pro Thr Thr Ile
 1340 1345 1350
 Ser Arg Glu Val Gly Pro Pro Gly Tyr Glu Leu Thr Gly Lys Asn
 1355 1360 1365
 Lys Gly Asn Gly Ala Ser Cys Thr Thr Asn Tyr Gln Gln Gly Ala
 1370 1375 1380
 Phe Thr Ser Pro Glu Ala Phe Pro Arg Glu Ala Cys Pro Pro Gly
 1385 1390 1395
 Phe Val Ser Gly Glu Asn Gln Gly Asn Asp Thr Pro Trp Tyr Met
 1400 1405 1410
 Ser Ser Lys Gln Met Ala Phe Thr Ser Pro Thr Thr Asn Pro Arg
 1415 1420 1425
 Val Gly Pro Pro Gly Tyr Gly Gln Glu Asn Gln Gly Ser Arg Val
 1430 1435 1440
 Ser Trp Asn Thr Ser Tyr Glu Gln Gly Val Phe Met Ser Ser Thr
 1445 1450 1455
 Asn Ile Ser Arg Glu Ala Gly Pro Pro Gly Tyr Gly Pro Gly Glu
 1460 1465 1470
 Ser Gln Gly Asn Ile Thr Ser Trp Ile Ser Thr Gly Gly Gln Arg
 1475 1480 1485
 Met Phe Thr Gly Pro Ser Thr Ile Pro Asn Glu Val Gly Gln Gln
 1490 1495 1500
 Gly Tyr Gly Pro Ser Gly Ile Pro Gly Thr Asp Gly Ser Arg Ser
 1505 1510 1515
 Met Asp Leu Lys Ala Phe Pro Pro Gly Tyr Gly Pro Ala Gly Glu
 1520 1525 1530
 Asn Gln Gly Gly Gly Ala Ser Gly Ser Val Ile Ser Pro His Ala
 1535 1540 1545
 Arg Ser Ala Glu Asn Val Glu Phe Ser Asp Asp Ala Arg Lys Lys
 1550 1555 1560

ES 2 558 745 T3

Ala Ile Gln Arg Glu Leu Arg Arg Arg Gln Lys Phe Arg
 1565 1570 1575

5

<210> 211
 <211> 4620
 <212> ADN
 <213> *Zea mays*
 <400> 211

```

atgggaagga aaggctcaga aatgtcacca agtaaaaggt atcccttgag gtctccacat    60
agtagtggta gagttcttcg ctctgcctta gccaatgata ataaggacag tgagcctctg    120
aatgctgctg ctgctaccca agccgctgtg aaaaaaagaa gtggcagtca gttggatagt    180
cctaatagca gtgtcagggt gcttcaactc acatcaaaaa ataagaacga ggtatgtagc    240
gggcctctaa atgacaggat ccctgataag ccagctgcaa ataaaagaaa acgtgccagt    300
ccttcacagg tgggaagttc cagcaacagt gtcagggtcg ttcgctctgc aataaaaaat    360
aaagatgagg catgcagtaa atctctaagt gacaggaccg ctggtgagcc agctgcaaat    420
aaaagaaaaa acgtcaatcg ctcaaagacg gggagtctga gctgtggtgt cagggtgctt    480
ctatctgctt caaaagctaa ggatgagaca tctaccaagc ctttaaatga tagtgctact    540
gttgaaccag cttcaggtaa aaggaaaggc accagtcctg caaagggtgg aagtgccagc    600
aatagtgtca ggggtgcttc atctacttca aaatataaga atgagacatg tactgagcct    660
ttaaatgata gtggtgctgt tgaaccagct gcaaataaaa ggaaagggtg cagtcgttca    720
aaggagagaa gtctcaatag cagtgtcagg gtcggtgctc tggcctcaaa taataagatt    780
gaggtatcta ttaagcctct aggtaatagt actgctgggg aaccagctgc gaaaagaaga    840
aaaagcggca gttccttcga ggcaggaagc cctgtgagca gtgccagggt gcttcgcctt    900
acctcagaaa ggaagaatga ggcacctaag cttttaaata agagtactgc tgctcaacca    960
gctgcgagga aaaaaaagc tggggtaatt tcaaagaccg acaatccaaa gattggtctc   1020
agggttcttc gctctgcctc aggaaagaag aatgaagcat gtattgggca tgtaaatagat   1080
agtacttctg ctgaaccaac tgtgacaaaa agaaaccgca agccttcaat ggatcgcagt   1140
cccaagaagg actacttaaa aatttgtcaa agagttagat atattttgaa tcgaatgaac   1200
tatcaacaaa cttttattca agcttatgca agtgaagggt ggaaaggcca aagtttgaa   1260
aaaataagac ctgaaaagga gcttgaacga gccaaaggcag aaatcctcca atgcaagtta   1320
agaatccgtg aagcttttctg aaatatggac tctcttttat ccaaggggaa gcttgaagaa   1380
tctttgtttg attccgcagg agaaatttca agcgaagata tattttgtgc tgatgtggc   1440
tcaaaggatg ttactttaca aaatgacatc attctttgtg atggagcctg tgacagagga   1500
ttccaccaga actgtttgaa ccctcctttg ctaactgaag atatacctcc gggggatcaa   1560
agatggcttt gccctgcatg tgtctgcaaa gcagactcta tagatgcaact aaatgaactt   1620
caagggagca agctctccat tcatgactca tgggagaaag ttttcctga ggcagcttca   1680
attgccaacg gttctaaaca agttgatact tctgatctgc tgccagatca tataaagcac   1740
    
```

10

ES 2 558 745 T3

agtgacaacc ctgcattggt tgaagggctg atggtggatg aagtcaggct ttctgcagag 1800
 gatgacagca aagcggatga ccttcgcttg tcttcggagg actcagggga tggtgacttc 1860
 gatccgtcag gtccagattc tagtgaagac caaaaagatg gattgaactc agaagagtca 1920
 gatttcacat cagactctga tgatTTTTgt gctgagattg ccaaactctg tggccaggat 1980
 gaagtctcgg catctccatt atcaaagtgt ataaaccata cttatagaat gaaactcagg 2040
 gccagcaaca accgctctaa tgaagaaaac catgatcatg tatttatgga catggagcta 2100
 gggcaagaca tggtgctacc agtttcaagc aggcggcagg tggaacgggt ggattacaaa 2160
 aaactatatg atgaggcata tgggaaagaa tcatctaatt caagtgatga caaagagtgg 2220
 tctggaaaag aactattaga aggcagtgaa acagattcac ttagtgagcg acctcatcct 2280
 gtaaaaagggt gctctagaag agcacaagca gagcagcaga ataatgaaca tacaccacag 2340
 agggagcgcc ttcattggtc tgaaagtgag cagaaaactg gaattcttcg atccaatggc 2400
 agcagtagca caggacggaa gtttggctct gtagctaccc agaaattgaa ggtacatitt 2460
 gagaaagatc catatcctag ccgtgaaacg aaagagaata taccagaaga actaggcttg 2520
 acatttaatc aggtctccag atggttctct agcactcgtc actattcaag ggttgcttct 2580
 gccagaaaag aatgcatcc tgacggccat actagcgaga ataatgatac caccactgta 2640
 gatagcatgc aagcaaggca acccaacact gtggtgatgg aaaagttgac tgggaataga 2700
 aatgacattg tttctgagaa accgatgggt cagaacaatc ttaaccaagg ggtggcagaa 2760
 aagttgacta gggatagaaa tgacattggt cctgagaaac aggtggtgca gagcaatctt 2820
 aaccaatgta ataatgaaga tatcccgctt tctggaacag aaattgagat ggagtcttac 2880
 gaacaagaat catctgaaag cagtgatgaa gagtgggtcca cacttagcac accacgaaaa 2940
 acaaaattac aaggcaacga aaaagtatca cttactgaat cactacgttc tgcaaaacgg 3000
 tgttctagaa gagcaccagc tagggagcag aataatgaac atactcagag tgagcaactt 3060
 catggttctg taagtaaaca gcaaaactgaa gttcttcgct ccaatgtcag cagtggcgat 3120
 gcctcaaaat gtcattttgg ccctatagtt actcagaagc tgaaggaaca ttttgaaaaa 3180
 gatccatata cttgccgtgc acaaaaagag ggtctggcac aagagctagg actgacattt 3240
 aatcagatca gtaagtgggt ctctgccacc catcattatt caagagatgc tgttgccaag 3300
 aacaaaaaat acccaggaga aaatactact gagaataata gtagcacaat ctttgatggt 3360
 atacaagtaa tagaaccaa ttttggtttg atagataaac cagatgcaga tacaatgac 3420
 atgatttctg aaaagttgat ggtacaaatt aatcttaacg aaggcattga agaggatata 3480
 ccaccgagcc aatatactac caggtgtgaa gagaaattaa ctatgactca agctgccatc 3540
 tcaaggaag ctggacctcc aggttatggt cctggagaaa actttcttca ggtcagttca 3600
 aggattacta gctgtgagca gagtgtgata atggctccta gcgctatcgc aagaggagtt 3660
 ggtccaccag gatatacgc tggagaacac cagggcaatg gggctccatg gaatacaagc 3720
 tatgagcgaa gaccgtttac gagtctgcg actagctcaa gagaagtcgg acctccagga 3780
 tatgggcctg aagaaaacca aggcagtggc atttctgta atacaagcag agagcaaaata 3840

ES 2 558 745 T3

ttgtttatga gccctacaac catctcgaga gaacttggtc caccaggata tttgcccaga 3900
 gaaaaccagc gcagtgcacac ttcattggaat acaagctgcg agtcgggact atttatgagt 3960
 cctgcgacca cctcaataga acatggcca ccaggatatg ggcctggaga aaaccaggag 4020
 aatgacagta caagctgcga gttgagaatg tttaccagtc ctacaatcgt ctccagagaa 4080
 gttggcccac caggatacca gcagggagtg tttatgagtc ctaaaacat ctcaagagaa 4140
 gtctgtcctc caggatttgg gtctggagag aaccagggca acgacactac atggtacaaa 4200
 cagagcgcgt ttacaagtcc tacaaccatc tcgagagaag tcggtccgcc aggttatgga 4260
 caggaagatc agggcaatgg tggttcttgg aacacgagtt gtgagcaggg agtgtttacg 4320
 agtcctacta gaggcactctc aagagaagtc ggtccgccag gatacgggcc cggggaaagt 4380
 cagggcaata ccacttcgtg gatttcgacc tgccagcaga gaatgcttgc aggaatccag 4440
 agaactggtg gctcaaggag tatggacttg gaagcctttc caccgggata cggacctggt 4500
 ggtgcttcag ggagtgtcat aagccccaa gctcggctctg cagagaatgt agaattttcg 4560
 gacgaggcta gaaaaaggc tatacagcgg gagctgagac ggcggcagaa gttcagggtga 4620

<210> 212
 <211> 1539
 <212> PRT
 <213> *Zea mays*

5

<400> 212

ES 2 558 745 T3

Met Gly Arg Lys Gly Ser Glu Met Ser Pro Ser Lys Arg Tyr Pro Leu
1 5 10 15

Arg Ser Pro His Ser Ser Gly Arg Val Leu Arg Ser Ala Leu Ala Asn
20 25 30

Asp Asn Lys Asp Ser Glu Pro Leu Asn Ala Ala Ala Ala Thr Gln Ala
35 40 45

Ala Val Lys Lys Arg Ser Gly Ser Gln Leu Asp Ser Pro Asn Ser Ser
50 55 60

Val Arg Leu Leu His Ser Thr Ser Lys Asn Lys Asn Glu Val Cys Ser
65 70 75 80

Gly Pro Leu Asn Asp Arg Ile Pro Asp Lys Pro Ala Ala Asn Lys Arg
85 90 95

Lys Arg Ala Ser Pro Ser Gln Val Gly Ser Ser Ser Asn Ser Val Arg
100 105 110

Val Val Arg Ser Ala Ile Lys Asn Lys Asp Glu Ala Cys Ser Lys Ser
115 120 125

Leu Ser Asp Arg Thr Ala Gly Glu Pro Ala Ala Asn Lys Arg Lys Asn
130 135 140

ES 2 558 745 T3

Val Asn Arg Ser Lys Thr Gly Ser Leu Ser Cys Gly Val Arg Val Leu
 145 150 155 160
 Leu Ser Ala Ser Lys Ala Lys Asp Glu Thr Ser Thr Lys Pro Leu Asn
 165 170 175
 Asp Ser Ala Thr Val Glu Pro Ala Ser Gly Lys Arg Lys Gly Thr Ser
 180 185 190
 Pro Ser Lys Val Gly Ser Ala Ser Asn Ser Val Arg Val Leu Arg Ser
 195 200 205
 Thr Ser Lys Tyr Lys Asn Glu Thr Cys Thr Glu Pro Leu Asn Asp Ser
 210 215 220
 Gly Ala Val Glu Pro Ala Ala Asn Lys Arg Lys Gly Val Ser Arg Ser
 225 230 235 240
 Lys Glu Arg Ser Leu Asn Ser Ser Val Arg Val Val Ala Leu Ala Ser
 245 250 255
 Asn Asn Lys Ile Glu Val Ser Ile Lys Pro Leu Gly Asn Ser Thr Ala
 260 265 270
 Gly Glu Pro Ala Ala Lys Arg Arg Lys Ser Gly Ser Ser Phe Glu Ala
 275 280 285
 Gly Ser Pro Val Ser Ser Ala Arg Val Leu Arg Pro Thr Ser Glu Arg
 290 295 300
 Lys Asn Glu Ala Ser Lys Pro Leu Asn Glu Ser Thr Ala Ala Gln Pro
 305 310 315
 Ala Ala Arg Lys Lys Lys Ala Gly Val Ile Ser Lys Thr Asp Asn Pro
 325 330 335
 Lys Ile Gly Leu Arg Val Leu Arg Ser Ala Ser Gly Lys Lys Asn Glu
 340 345 350
 Ala Cys Ile Gly His Val Asn Asp Ser Thr Ser Ala Glu Pro Thr Val
 355 360 365
 Thr Lys Arg Asn Arg Lys Pro Ser Met Asp Arg Ser Pro Lys Lys Asp
 370 375 380
 Tyr Leu Lys Ile Cys Gln Arg Val Arg Tyr Ile Leu Asn Arg Met Asn
 385 390 395 400
 Tyr Gln Gln Thr Phe Ile Gln Ala Tyr Ala Ser Glu Gly Trp Lys Gly
 405 410 415

ES 2 558 745 T3

Gln Ser Leu Glu Lys Ile Arg Pro Glu Lys Glu Leu Glu Arg Ala Lys
420 425 430

Ala Glu Ile Leu Gln Cys Lys Leu Arg Ile Arg Glu Ala Phe Arg Asn
435 440 445

Met Asp Ser Leu Leu Ser Lys Gly Lys Leu Glu Glu Ser Leu Phe Asp
450 455 460

Ser Ala Gly Glu Ile Ser Ser Glu Asp Ile Phe Cys Ala Val Cys Gly
465 470 475 480

Ser Lys Asp Val Thr Leu Gln Asn Asp Ile Ile Leu Cys Asp Gly Ala
485 490 495

Cys Asp Arg Gly Phe His Gln Asn Cys Leu Asn Pro Pro Leu Leu Thr
500 505 510

Glu Asp Ile Pro Pro Gly Asp Gln Arg Trp Leu Cys Pro Ala Cys Val
515 520 525

Cys Lys Ala Asp Ser Ile Asp Ala Leu Asn Glu Leu Gln Gly Ser Lys
530 535 540

Leu Ser Ile His Asp Ser Trp Glu Lys Val Phe Pro Glu Ala Ala Ser
545 550 555 560

Ile Ala Asn Gly Ser Lys Gln Val Asp Thr Ser Asp Leu Leu Pro Asp
565 570 575

His Ile Lys His Ser Asp Asn Pro Ala Leu Val Glu Gly Leu Met Val
580 585 590

Asp Glu Val Arg Leu Ser Ala Glu Asp Asp Ser Lys Ala Asp Asp Leu
595 600 605

Arg Leu Ser Ser Glu Asp Ser Gly Asp Gly Asp Phe Asp Pro Ser Gly
610 615 620

Pro Asp Ser Ser Glu Asp Gln Lys Asp Gly Leu Asn Ser Glu Glu Ser
625 630 635 640

Asp Phe Thr Ser Asp Ser Asp Asp Phe Cys Ala Glu Ile Ala Lys Ser
645 650 655

Cys Gly Gln Asp Glu Val Ser Ala Ser Pro Leu Ser Asn Val Ile Asn
660 665 670

His Thr Tyr Arg Met Lys Leu Arg Ala Ser Asn Asn Arg Ser Asn Glu
675 680 685

ES 2 558 745 T3

Glu Asn His Asp His Val Phe Met Asp Met Glu Leu Gly Gln Asp Met
 690 695 700
 Val Leu Pro Val Ser Ser Arg Arg Gln Val Glu Arg Leu Asp Tyr Lys
 705 710 715 720
 Lys Leu Tyr Asp Glu Ala Tyr Gly Lys Glu Ser Ser Asn Ser Ser Asp
 725 730 735
 Asp Lys Glu Trp Ser Gly Lys Glu Leu Leu Glu Gly Ser Glu Thr Asp
 740 745 750
 Ser Leu Ser Glu Arg Pro His Pro Val Lys Arg Cys Ser Arg Arg Ala
 755 760 765
 Gln Ala Glu Gln Gln Asn Asn Glu His Thr Pro Gln Arg Glu Arg Leu
 770 775 780
 His Gly Ser Glu Ser Glu Gln Lys Thr Gly Ile Leu Arg Ser Asn Gly
 785 790 795 800
 Ser Ser Ser Thr Gly Arg Lys Phe Gly Pro Val Ala Thr Gln Lys Leu
 805 810 815
 Lys Val His Phe Glu Lys Asp Pro Tyr Pro Ser Arg Glu Thr Lys Glu
 820 825 830
 Asn Ile Ser Glu Glu Leu Gly Leu Thr Phe Asn Gln Val Ser Arg Trp
 835 840 845
 Phe Ser Ser Thr Arg His Tyr Ser Arg Val Ala Ser Ala Arg Lys Glu
 850 855 860
 Met His Pro Asp Gly His Thr Ser Glu Asn Asn Asp Thr Thr Thr Val
 865 870 875 880
 Asp Ser Met Gln Ala Arg Gln Pro Asn Thr Val Val Met Glu Lys Leu
 885 890 895
 Thr Gly Asn Arg Asn Asp Ile Val Ser Glu Lys Pro Met Val Gln Asn
 900 905 910
 Asn Leu Asn Gln Gly Val Ala Glu Lys Leu Thr Arg Asp Arg Asn Asp
 915 920 925
 Ile Val Pro Glu Lys Gln Val Val Gln Ser Asn Leu Asn Gln Cys Asn
 930 935 940
 Asn Glu Asp Ile Pro Leu Ser Gly Thr Glu Ile Glu Met Glu Ser Tyr
 945 950 955 960
 Glu Gln Glu Ser Ser Glu Ser Ser Asp Glu Glu Trp Ser Thr Leu Ser

ES 2 558 745 T3

			965					970					975		
Thr	Pro	Arg	Lys 980	Thr	Lys	Leu	Gln	Gly 985	Asn	Glu	Lys	Val	Ser 990	Leu	Thr
Glu	Ser	Leu	Arg 995	Ser	Ala	Lys	Arg 1000	Cys	Ser	Arg	Arg	Ala 1005	Pro	Ala	Arg
Glu	Gln 1010	Asn	Asn	Glu	His	Thr 1015	Gln	Ser	Glu	Gln	Leu 1020	His	Gly	Ser	
Val	Ser 1025	Lys	Gln	Gln	Thr	Glu 1030	Val	Leu	Arg	Ser	Asn 1035	Val	Ser	Ser	
Gly	Asp 1040	Ala	Ser	Lys	Cys	His 1045	Phe	Gly	Pro	Ile	Val 1050	Thr	Gln	Lys	
Leu	Lys 1055	Glu	His	Phe	Glu	Lys 1060	Asp	Pro	Tyr	Pro	Cys 1065	Arg	Ala	Thr	
Lys	Glu 1070	Gly	Leu	Ala	Gln	Glu 1075	Leu	Gly	Leu	Thr	Phe 1080	Asn	Gln	Ile	
Ser	Lys 1085	Trp	Phe	Ser	Ala	Thr 1090	His	His	Tyr	Ser	Arg 1095	Asp	Ala	Val	
Ala	Lys 1100	Asn	Gln	Lys	Tyr	Pro 1105	Gly	Glu	Asn	Thr	Thr 1110	Glu	Asn	Asn	
Ser	Ser 1115	Thr	Ile	Phe	Asp	Gly 1120	Ile	Gln	Val	Ile	Glu 1125	Pro	Asn	Phe	
Gly	Leu 1130	Ile	Asp	Lys	Pro	Asp 1135	Ala	Asp	Thr	Asn	Asp 1140	Met	Ile	Ser	
Glu	Lys 1145	Leu	Met	Val	Gln	Ile 1150	Asn	Leu	Asn	Glu	Gly 1155	Ile	Glu	Glu	
Asp	Ile 1160	Pro	Pro	Ser	Gln	Tyr 1165	Thr	Thr	Arg	Cys	Glu 1170	Glu	Lys	Leu	
Thr	Met 1175	Thr	Gln	Ala	Ala	Ile 1180	Ser	Arg	Glu	Ala	Gly 1185	Pro	Pro	Gly	
Tyr	Gly 1190	Pro	Gly	Glu	Asn	Phe 1195	Leu	Gln	Val	Ser	Ser 1200	Arg	Ile	Thr	
Ser	Cys 1205	Glu	Gln	Ser	Val	Ile 1210	Met	Ala	Pro	Ser	Ala 1215	Ile	Ala	Arg	
Gly	Val 1220	Gly	Pro	Pro	Gly	Tyr 1225	Thr	Pro	Gly	Glu	His 1230	Gln	Gly	Asn	

ES 2 558 745 T3

Gly Ala Pro Trp Asn Thr Ser Tyr Glu Arg Arg Pro Phe Thr Ser
1235 1240 1245

Pro Ala Thr Ser Ser Arg Glu Val Gly Pro Pro Gly Tyr Gly Pro
1250 1255 1260

Glu Glu Asn Gln Gly Ser Gly Ile Ser Cys Asn Thr Ser Arg Glu
1265 1270 1275

Gln Ile Leu Phe Met Ser Pro Thr Thr Ile Ser Arg Glu Leu Gly
1280 1285 1290

Pro Pro Gly Tyr Leu Pro Arg Glu Asn Gln Arg Ser Asp Thr Ser
1295 1300 1305

Trp Asn Thr Ser Cys Glu Ser Gly Leu Phe Met Ser Pro Ala Thr
1310 1315 1320

Thr Ser Ile Glu His Gly Pro Pro Gly Tyr Gly Pro Gly Glu Asn
1325 1330 1335

Gln Glu Asn Asp Ser Thr Ser Cys Glu Leu Arg Met Phe Thr Ser
1340 1345 1350

Pro Thr Ile Val Ser Arg Glu Val Gly Pro Pro Gly Tyr Gln Gln
1355 1360 1365

Gly Val Phe Met Ser Pro Lys Thr Ile Ser Arg Glu Val Cys Pro
1370 1375 1380

Pro Gly Phe Gly Ser Gly Glu Asn Gln Gly Asn Asp Thr Thr Trp
1385 1390 1395

Tyr Lys Gln Ser Ala Phe Thr Ser Pro Thr Thr Ile Ser Arg Glu
1400 1405 1410

Val Gly Pro Pro Gly Tyr Gly Gln Glu Asp Gln Gly Asn Gly Gly
1415 1420 1425

Ser Trp Asn Thr Ser Cys Glu Gln Gly Val Phe Thr Ser Pro Thr
1430 1435 1440

Arg Gly Ile Ser Arg Glu Val Gly Pro Pro Gly Tyr Gly Pro Gly
1445 1450 1455

Glu Ser Gln Gly Asn Thr Thr Ser Trp Ile Ser Thr Cys Gln Gln
1460 1465 1470

Arg Met Leu Ala Gly Ile Gln Arg Thr Gly Gly Ser Arg Ser Met
1475 1480 1485

ES 2 558 745 T3

Asp Leu Glu Ala Phe Pro Pro Gly Tyr Gly Pro Gly Gly Ala Ser
 1490 1495 1500
 Gly Ser Val Ile Ser Pro Gln Ala Arg Ser Ala Glu Asn Val Glu
 1505 1510 1515
 Phe Ser Asp Glu Ala Arg Lys Lys Ala Ile Gln Arg Glu Leu Arg
 1520 1525 1530
 Arg Arg Gln Lys Phe Arg
 1535

<210> 213
 <211> 2907
 <212> ADN
 <213> *Vitis vinifera*

5

<400> 213

atggatgctt ccctgccc agagagtaat cggactagga aaagttctag tcctaaacag 60
 aatatttttag aagaggccag aaaactttct gaaagtgtgt gtagcgaatc gtcagagcag 120
 aaacgtccct ctgaaaatgg acagcatgag ccagcagaaa tcagccctgt gctatccaat 180
 tgtattgtta ctgaacagtc agaactgctt ccagaagatg tgggtgatac aatattagga 240
 ttgcctccag cagatgtgac caagaattct cttactgaac atttgggatt gcctccagaa 300
 gatgcaatca agaatgatgg aactgaacaa ttggggtttt ttctgaagt tgtgaccaag 360
 agttctatta ttgaaaatt aggacaatca gagccacctc ctgaaaatgt ggcccgggat 420
 tctggtcttg accagtcagg gtcagcacc aaagatttgg ccaataaaag aactgcaaaa 480
 ctagttaaaa gaaaataata attgagatct tcagtaagcg gttcccgggt tttacgctca 540
 aggtcacaag agaaacctaa agcttcacag ccaagtgata attttgtaaa tgccagtgtc 600
 agcagagaaa ggaaaggaag aaagaagaaa aggatgaata aaacaactgc agatgaattt 660
 gcgagaatca ggaacatct tagatattta ctgaacagaa tgagctatga gcaaaatctg 720
 attgatgctt attctgccga aggttggaag ggacaaagtg tggaaaaatt aaagccagag 780
 aaggaacttc aacgtgcctc atctgaaatt tctcgtcgca aactgmaaat acgtgatcta 840
 tttcaacatc ttgattcatt atgtgctgaa ggaaggtttc cagaatcttt atttgattct 900
 gaagggcaga ttgacagtga ggatatatc tgtgctaaat gtgagtccaa agacatgtct 960
 gctgataatg acataaact ctgtgacggt gcttgtgatc gtggatttca ccagttttgt 1020
 ctggaaccgc cattgttaaa agaagaaatt cctcctgatg acgagggttg gctgtgccct 1080
 gcattgtgatt gcaaagtga ctgcatggac ctgcttaatg actctcaagg aacaaaactt 1140
 tctgtaattg atagctggga gaaggttttt cctgaggcag ctgcagctgg gaataaccag 1200
 gataacaact ctggattttc atcagatgat tctgaggata atgattatga ccccgattgt 1260
 ccagaggttg atgagaaggg tcagggagat aagtcaagtt ctgataaatt tgatgaatct 1320
 gatgaatttg acgaatctga tgaatctgat ttcacttctg catctgatga tatggtggtc 1380

10

ES 2 558 745 T3

tcaccraata atgagcagtg tttggggcctt ccttctgatg attcagagga tgatgatttt 1440
gatcctgatg ctccagaaat tgatgaacag gtaaatcagg ggagttcaag ttctgatttt 1500
acatctgact ccgaggattt tactgccact ttagatcgca gaaacttctc tgacaacgag 1560
gatggtcttg atgagcaaag aagatttggt aggaagaaga aagatacttt aaaggatgag 1620
ctcttatccg tgtagagtc aaattctggt caagataatg cacctctgtc tgcaaagaga 1680
catgtagaaa ggctggatta caaaaagctg catgatgagg catatggaaa tgtttcttct 1740
gattcaagtg atgatgaaga ctggacagag aatgttatac caagaaagag gaagaatctt 1800
agtggtaatg ttgcctcagt gtcaccaaat ggaaacactt caatcactga gaatgggaca 1860
aataccaagg acataaagca cgacttggaa gcagctggat gcactcccaa gcgaagaact 1920
cgtcagaagc tgaacttga aagcacaat aattcacttg ctgagtcaca taaaggttct 1980
cgaagtcttg gttctactgg tgaaaaaagt gggcaatcat catataaaa acttggagaa 2040
gctgtaactg agagactcta caaatccttc caggaaaatc agtaccctga tcgtgctatg 2100
aaagaaaaat tggcagaaga gctgggaata acgagccggc aggttagcaa atggtttgag 2160
aatgcccgtt ggagctttcg ccatcgacca cccaaggaag ccagtgcggg taaaagtgct 2220
gtgaaaaagg atgcatccac gtctcaaaca gatcaaaagc ctgagcaaga agtggttctt 2280
agagaaagtt ctcaaatgg agtgggaaaa aaggagtcgc ccaaagcagg tgcttcgaag 2340
gtagaccgca gcaaggaagc caatgcaggt aaaagtgctg tgaaaaagga tgcattccag 2400
tctcaaacag atcaaaagcc tgagcaagaa gtggttatta aagaaagttc tcacaatgga 2460
gtgggaaaaa aggagtcgac caaagcaggt gcttcgaagg tagaccgctg cagtggagct 2520
aagaggagaa ggaaattggc aactgatgga agccatagac agaaatcttc aactcccaac 2580
tctacaagac aaaagacaaa gtcaaatcat gaagcatctg aggcaaccaa tgggaagcaga 2640
aggcagaatt cttctactcc caagtctaga aggcaaaaga ctaagttagc tgggtgaagcg 2700
tctgaagcaa ctggtggaag cagtagacag aattcgtcaa ccccaaattc taaaaggcga 2760
aaaaccaagt cagatcatga agcatctaat ccgctcctta gcggtagaa gatagcaaag 2820
acggcaaaga gctcatctgg cacacaaaa acaaaggaaa agctaagtga cagaatccaa 2880
acaagaagta ggaaatctat tgcttga 2907

<210> 214
<211> 968
5 <212> PRT
<213> *Vitis vinifera*

<220>
10 <221> INCIERTO
<222> (276)..(276)
<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 214

Met Asp Ala Ser Pro Ala Glu Ser Asn Arg Thr Arg Lys Ser Ser
1 5 10 15

15

ES 2 558 745 T3

Ser Pro Lys Glu Asn Ile Leu Glu Glu Ala Arg Lys Leu Ser Glu Ser
 20 25 30

Val Cys Ser Glu Ser Ser Glu Glu Lys Arg Pro Ser Glu Asn Gly Glu
 35 40 45

His Glu Pro Ala Glu Ile Ser Pro Val Leu Ser Asn Cys Ile Val Thr
 50 55 60

Glu Glu Ser Glu Leu Pro Pro Glu Asp Val Gly Asp Thr Ile Leu Gly
 65 70 75 80

Leu Pro Pro Ala Asp Val Thr Lys Asn Ser Leu Thr Glu His Leu Gly
 85 90 95

Leu Pro Pro Glu Asp Ala Ile Lys Asn Asp Gly Thr Glu Glu Leu Gly
 100 105 110

Phe Phe Pro Glu Val Val Thr Lys Ser Ser Ile Ile Glu Lys Leu Gly
 115 120 125

Glu Ser Glu Pro Pro Pro Glu Asn Val Ala Arg Tyr Ser Gly Leu Asp
 130 135 140

Glu Ser Gly Ser Ala Pro Lys Asp Leu Ala Asn Lys Arg Thr Ala Lys
 145 150 155 160

Leu Val Lys Arg Lys Tyr Lys Leu Arg Ser Ser Val Ser Gly Ser Arg
 165 170 175

Val Leu Arg Ser Arg Ser Glu Glu Lys Pro Lys Ala Ser Glu Pro Ser
 180 185 190

Asp Asn Phe Val Asn Ala Ser Ala Ser Arg Glu Arg Lys Gly Arg Lys
 195 200 205

Lys Lys Arg Met Asn Lys Thr Thr Ala Asp Glu Phe Ala Arg Ile Arg
 210 215 220

Lys His Leu Arg Tyr Leu Leu Asn Arg Met Ser Tyr Glu Glu Asn Leu
 225 230 235 240

Ile Asp Ala Tyr Ser Ala Glu Gly Trp Lys Gly Glu Ser Val Glu Lys
 245 250 255

Leu Lys Pro Glu Lys Glu Leu Glu Arg Ala Ser Ser Glu Ile Ser Arg
 260 265 270

Arg Lys Leu Xaa Ile Arg Asp Leu Phe Glu His Leu Asp Ser Leu Cys
 275 280 285

ES 2 558 745 T3

Ala Glu Gly Arg Phe Pro Glu Ser Leu Phe Asp Ser Glu Gly Gln Ile
 290 295 300

Asp Ser Glu Asp Ile Phe Cys Ala Lys Cys Glu Ser Lys Asp Met Ser
 305 310 315 320

Ala Asp Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asp Arg Gly Phe
 325 330 335

His Gln Phe Cys Leu Glu Pro Pro Leu Leu Lys Glu Glu Ile Pro Pro
 340 345 350

Asp Asp Glu Gly Trp Leu Cys Pro Ala Cys Asp Cys Lys Val Asp Cys
 355 360 365

Met Asp Leu Leu Asn Asp Ser Gln Gly Thr Lys Leu Ser Val Ile Asp
 370 375 380

Ser Trp Glu Lys Val Phe Pro Glu Ala Ala Ala Ala Gly Asn Asn Gln
 385 390 395 400

Asp Asn Asn Ser Gly Phe Ser Ser Asp Asp Ser Glu Asp Asn Asp Tyr
 405 410 415

Asp Pro Asp Cys Pro Glu Val Asp Glu Lys Gly Gln Gly Asp Lys Ser
 420 425 430 435

Ser Ser Asp Lys Phe Asp Glu Ser Asp Glu Phe Asp Glu Ser Asp Glu
 435 440 445

Ser Asp Phe Thr Ser Ala Ser Asp Asp Met Val Val Ser Pro Asn Asn
 450 455 460

Glu Gln Cys Leu Gly Leu Pro Ser Asp Asp Ser Glu Asp Asp Asp Phe
 465 470 475 480

Asp Pro Asp Ala Pro Glu Ile Asp Glu Gln Val Asn Gln Gly Ser Ser
 485 490 495

Ser Ser Asp Phe Thr Ser Asp Ser Glu Asp Phe Thr Ala Thr Leu Asp
 500 505 510

Arg Arg Asn Phe Ser Asp Asn Glu Asp Gly Leu Asp Glu Gln Arg Arg
 515 520 525

Phe Gly Arg Lys Lys Lys Asp Thr Leu Lys Asp Glu Leu Leu Ser Val
 530 535 540

Leu Glu Ser Asn Ser Gly Gln Asp Asn Ala Pro Leu Ser Ala Lys Arg
 545 550 555 560

His Val Glu Arg Leu Asp Tyr Lys Lys Leu His Asp Glu Ala Tyr Gly

ES 2 558 745 T3

565 570 575
 Asn Val Ser Ser Asp Ser Ser Asp Asp Glu Asp Trp Thr Glu Asn Val
 580 585 590
 Ile Pro Arg Lys Arg Lys Asn Leu Ser Gly Asn Val Ala Ser Val Ser
 595 600 605
 Pro Asn Gly Asn Thr Ser Ile Thr Glu Asn Gly Thr Asn Thr Lys Asp
 610 615 620
 Ile Lys His Asp Leu Glu Ala Ala Gly Cys Thr Pro Lys Arg Arg Thr
 625 630 635 640
 Arg Gln Lys Leu Asn Phe Glu Ser Thr Asn Asn Ser Leu Ala Glu Ser
 645 650 655
 His Lys Gly Ser Arg Ser Pro Gly Ser Thr Gly Glu Lys Ser Gly Gln
 660 665 670
 Ser Ser Tyr Lys Lys Leu Gly Glu Ala Val Thr Glu Arg Leu Tyr Lys
 675 680 685
 Ser Phe Gln Glu Asn Gln Tyr Pro Asp Arg Ala Met Lys Glu Lys Leu
 690 695 700
 Ala Glu Glu Leu Gly Ile Thr Ser Arg Gln Val Ser Lys Trp Phe Glu
 705 710 715 720
 Asn Ala Arg Trp Ser Phe Arg His Arg Pro Pro Lys Glu Ala Ser Ala
 725 730 735
 Gly Lys Ser Ala Val Lys Lys Asp Ala Ser Thr Ser Gln Thr Asp Gln
 740 745 750
 Lys Pro Glu Gln Glu Val Val Leu Arg Glu Ser Ser His Asn Gly Val
 755 760 765
 Gly Lys Lys Glu Ser Pro Lys Ala Gly Ala Ser Lys Val Asp Arg Ser
 770 775 780
 Lys Glu Ala Asn Ala Gly Lys Ser Ala Val Lys Lys Asp Ala Ser Thr
 785 790 795 800
 Ser Gln Thr Asp Gln Lys Pro Glu Gln Glu Val Val Ile Lys Glu Ser
 805 810 815
 Ser His Asn Gly Val Gly Lys Lys Glu Ser Thr Lys Ala Gly Ala Ser
 820 825 830
 Lys Val Asp Arg Cys Ser Gly Ala Lys Arg Arg Arg Lys Leu Ala Thr
 835 840 845

ES 2 558 745 T3

Asp Gly Ser His Arg Gln Lys Ser Ser Thr Pro Asn Ser Thr Arg Gln
 850 855 860

Lys Thr Lys Ser Asn His Glu Ala Ser Glu Ala Thr Asn Gly Ser Arg
 865 870 875 880

Arg Gln Asn Ser Ser Thr Pro Lys Ser Arg Arg Gln Lys Thr Lys Leu
 885 890 895

Ala Gly Glu Ala Ser Glu Ala Thr Gly Gly Ser Ser Arg Gln Asn Ser
 900 905 910

Ser Thr Pro Asn Ser Lys Arg Arg Lys Thr Lys Ser Asp His Glu Ala
 915 920 925

Ser Asn Pro Val Leu Ser Gly Lys Lys Ile Ala Lys Thr Ala Lys Ser
 930 935 940

Ser Ser Gly Thr Pro Lys Thr Lys Glu Lys Leu Ser Asp Arg Ile Gln
 945 950 955 960

Thr Arg Ser Arg Lys Ser Ile Ala
 965

<210> 215
 <211> 2805
 <212> ADN
 <213> *Populus tremuloides*

5

<400> 215

atgtctgagg ctgagaacat ggggtgtttct ccatcacaag ttagttctca aactaaaagt 60
 tactcatgcc cttcacaac taaactggaa aatactcatg gatttactgc tgaatacaac 120
 tgtggtggat attcgaaga gaaacataag cttgagtctg aaatcataca aactgaagcc 180
 ggagatagcg gtactgcagt tctccaatct ggtgctgggg aaactgtgga accatctact 240
 gaagatgtaa caaacaattc ttttactgat ttggatccac ctcccgaaga tgccaggggt 300
 gctacttttg atgaaagttc aagacctata ctaactgcta tagatcagaa actcgagcca 360
 ggtgctagaa gtgtgaatac tgcgtgtacg catgggtgaat catcaaaagc aactgactct 420
 ggtattctac aggatgaacc tggaaatacc aatgctgcat catcaagctg cattgctaata 480
 gaaacctcac aagcatcact tgaaaatcta gccaaataatt cttgtactga agatgtgggt 540
 ccgccgtatg gagatgcaag caagggtaac cagattgata aaagttcata cctcaacaa 600
 actatatctg gacatacact tgagttactt tctgacagag cttgtttgga acgatcagaa 660
 gaaagacaga agcctggatc tgaactttca gaaaatgaat caacgggaat tgactgaa 720
 ttatatagtg gcattgctat cgagaattca gagccgctta ctcaactagt gaccaagagt 780
 tctcctatca aacatgtagg cttgcttctt ggtgacagca tcatcattcc tgcaaatgaa 840
 caaacaagac caactcatga tgatgaggac aaaggcccag atcatgaaca tttggaaaca 900

10

ES 2 558 745 T3

ccatctagag ttgcgattgg tattactaga cgcggtagac ctagaggtaa gagtgcttca 960
 agattgtcaa ggaagatata tatgttaaga tctttgagaa gtagtgatag agttctccgg 1020
 tcaagatcac aagagaaacc taaagctcct gaatcaagta ataattcagg taatgttaat 1080
 tccactgggtg ataaaaaagg gaaaagaagg aaaaagagaa gaggaaagaa tattgtggct 1140
 gatgaatatt ccaaaatcag ggcacatctt aggtacttat tgaatcggat gagctatgag 1200
 caaagcttga tcaactgctta ttctggggaa ggctggaaag gacttagcct agaaaagtta 1260
 aagccccgaga aggagttgca gcgagccaca tctgaaatta ccagacgcaa agtgaaaata 1320
 agggatttat ttcaacatat tgattctctg tgcagcgaag gaaggttccc atcatcttta 1380
 tttgattctg aaggacagat tgacagtgag gatgtattct gtgcaaaatg tgggtccaaa 1440
 gatttgaatg ctgataatga tattatacta tgtgatgggtg cttgtgacag aggatttcat 1500
 cagttctgtt tgataccacc attgttaaga gaagacattc ctctgatga tgagggctgg 1560
 ttatgccctg gatgtgattg caaagttgac tgcattgggtt tgcttaacga ttctcaagga 1620
 acaaatatat ctatcagtga tagctgggag aaagttttcc ctgaggctgc cgccacagcg 1680
 tctggacaaa aactggatca caactttgga ccactctctg atgattctga tgataatgat 1740
 tatgaacctg atgggtccaga tattgataag aagagtcagg aggaggaatc gagctctgac 1800
 gaatctgact tcacttctgc atctgatgaa ttcaaggctc ccccgatgg caaagagtac 1860
 ttggggctct cttctgatga ttcagaggat gatgactatg atcctgatgc tccagttctt 1920
 gaggagaagc tgaagcagga aagttcaagt tctgatttta catcagactc tgaggatctt 1980
 gctgcaacta ttaatgggtg tgggttgtct ttagaagatg aatgccacat gcctattgaa 2040
 cctcgtggag tttctaattg gaggaaatct aaatttgatg gaaagaagat gcagcttta 2100
 aacagcgagc ttttatccat gctagaacca gatctgtgcc aagatgagtc agcaactgtt 2160
 tctggaaaaa gaaatgttga cagattggac tacaagaagc tctatgatga aacgtatgga 2220
 aacatttcta cttcaagtga tgatgattac actgatactg ttgggccaag gaagaggagg 2280
 aaaaatactg gagatgttgc tacagtgaca gcaaacggag atgcctctgt taccgagaat 2340
 ggaatgaata gcaagaatat gaaccaggaa ttaaaggaga acaaacgtaa tcctgagaga 2400
 ggaacttgcc aaaactcgag ctttcaagaa acaaatgttt ccccgctaa atcatatgtt 2460
 ggtgcactc tatctggttc tagtggtaaa agcgttaggc cctctgctta taaaaaactt 2520
 ggagaagctg taactcagag actgtacagt tacttcaggg aaaatcagta tccagaccga 2580
 gctgcaaaag caagcttggc agaagagcta ggaattactt ttgagcaggt taacaagtgg 2640
 tttgtgaatg cccgttggag cttcaaccat tcatcateta caggtacaag taaagctgaa 2700
 agtgcttctg gaaaaggtag ctgtgatggc caagtgaggg atagtgaatc gaagaaccga 2760
 aagagcaata aacaaaaaac taacaccccg aaatctagga gataa 2805

<210> 216
 <211> 934
 <212> PRT
 <213> *Populus tremuloides*

5

<400> 216

ES 2 558 745 T3

Met Ser Glu Ala Glu Asn Met Gly Val Ser Pro Ser Gln Val Ser Ser
1 5 10 15

Gln Thr Lys Ser Tyr Ser Cys Pro Ser Gln Thr Lys Leu Glu Asn Thr
20 25 30

His Gly Phe Thr Ala Glu Tyr Asn Cys Gly Gly Tyr Ser Glu Glu Lys
35 40 45

His Lys Leu Glu Ser Glu Ile Ile Gln Thr Glu Ala Gly Asp Ser Gly
50 55 60

Thr Ala Val Leu Gln Ser Gly Ala Gly Glu Thr Val Glu Pro Ser Thr
65 70 75 80

Glu Asp Val Thr Asn Asn Ser Phe Thr Asp Leu Asp Pro Pro Pro Glu
85 90 95

Asp Ala Arg Gly Ala Thr Phe Asp Glu Ser Ser Arg Pro Ile Leu Thr
100 105 110

Ala Ile Asp Gln Lys Leu Glu Pro Gly Ala Arg Ser Val Asn Thr Ala
115 120 125

Cys Thr His Gly Glu Ser Ser Lys Ala Thr Asp Ser Gly Ile Leu Gln
130 135 140

Asp Glu Pro Gly Asn Thr Asn Ala Ala Ser Ser Ser Cys Ile Ala Asn
145 150 155 160

Glu Thr Ser Gln Ala Ser Leu Glu Asn Leu Ala Asn Asn Ser Cys Thr
165 170 175

Glu Asp Val Gly Pro Pro Tyr Gly Asp Ala Ser Lys Gly Asn Gln Ile
180 185 190

Asp Lys Ser Ser Tyr Pro Gln Gln Thr Ile Ser Gly His Thr Leu Glu
195 200 205

Leu Leu Ser Asp Arg Ala Cys Cys Glu Arg Ser Glu Glu Arg Gln Lys
210 215 220

Pro Gly Ser Glu Leu Ser Glu Asn Glu Ser Thr Gly Ile Asp Thr Glu
225 230 235 240

Leu Tyr Ser Gly Ile Ala Ile Glu Asn Ser Glu Pro Leu Thr Gln Leu
245 250 255

Val Thr Lys Ser Ser Pro Ile Lys His Val Gly Leu Leu Pro Gly Asp

ES 2 558 745 T3

260 265 270
 Ser Ile Ile Ile Pro Ala Asn Glu Gln Thr Arg Pro Thr His Asp Asp
 275 280 285
 Glu Asp Lys Gly Pro Asp His Glu His Leu Glu Thr Pro Ser Arg Val
 290 295 300
 Ala Ile Gly Ile Thr Arg Arg Gly Arg Pro Arg Gly Lys Ser Ala Ser
 305 310 315 320
 Arg Leu Ser Arg Lys Ile Tyr Met Leu Arg Ser Leu Arg Ser Ser Asp
 325 330 335
 Arg Val Leu Arg Ser Arg Ser Gln Glu Lys Pro Lys Ala Pro Glu Ser
 340 345 350
 Ser Asn Asn Ser Gly Asn Val Asn Ser Thr Gly Asp Lys Lys Gly Lys
 355 360 365
 Arg Arg Lys Lys Arg Arg Gly Lys Asn Ile Val Ala Asp Glu Tyr Ser
 370 375 380
 Lys Ile Arg Ala His Leu Arg Tyr Leu Leu Asn Arg Met Ser Tyr Glu
 385 390 395 400
 Gln Ser Leu Ile Thr Ala Tyr Ser Gly Glu Gly Trp Lys Gly Leu Ser
 405 410 415
 Leu Glu Lys Leu Lys Pro Glu Lys Glu Leu Gln Arg Ala Thr Ser Glu
 420 425 430
 Ile Thr Arg Arg Lys Val Lys Ile Arg Asp Leu Phe Gln His Ile Asp
 435 440 445
 Ser Leu Cys Ser Glu Gly Arg Phe Pro Ser Ser Leu Phe Asp Ser Glu
 450 455 460
 Gly Gln Ile Asp Ser Glu Asp Val Phe Cys Ala Lys Cys Gly Ser Lys
 465 470 475 480
 Asp Leu Asn Ala Asp Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asp
 485 490 495
 Arg Gly Phe His Gln Phe Cys Leu Ile Pro Pro Leu Leu Arg Glu Asp
 500 505 510
 Ile Pro Pro Asp Asp Glu Gly Trp Leu Cys Pro Gly Cys Asp Cys Lys
 515 520 525
 Val Asp Cys Ile Gly Leu Leu Asn Asp Ser Gln Gly Thr Asn Ile Ser
 530 535 540

ES 2 558 745 T3

Ile Ser Asp Ser Trp Glu Lys Val Phe Pro Glu Ala Ala Ala Thr Ala
545 550 555 560

Ser Gly Gln Lys Leu Asp His Asn Phe Gly Pro Ser Ser Asp Asp Ser
565 570 575

Asp Asp Asn Asp Tyr Glu Pro Asp Gly Pro Asp Ile Asp Lys Lys Ser
580 585 590

Gln Glu Glu Glu Ser Ser Ser Asp Glu Ser Asp Phe Thr Ser Ala Ser
595 600 605

Asp Glu Phe Lys Ala Pro Pro Asp Gly Lys Glu Tyr Leu Gly Leu Ser
610 615 620

Ser Asp Asp Ser Glu Asp Asp Asp Tyr Asp Pro Asp Ala Pro Val Leu
625 630 635 640

Glu Glu Lys Leu Lys Gln Glu Ser Ser Ser Ser Asp Phe Thr Ser Asp
645 650 655

Ser Glu Asp Leu Ala Ala Thr Ile Asn Gly Asp Gly Leu Ser Leu Glu
660 665 670

Asp Glu Cys His Met Pro Ile Glu Pro Arg Gly Val Ser Asn Gly Arg
675 680 685

Lys Ser Lys Phe Asp Gly Lys Lys Met Gln Ser Leu Asn Ser Glu Leu
690 700

Leu Ser Met Leu Glu Pro Asp Leu Cys Gln Asp Glu Ser Ala Thr Val
705 710 715 720

Ser Gly Lys Arg Asn Val Asp Arg Leu Asp Tyr Lys Lys Leu Tyr Asp
725 730 735

Glu Thr Tyr Gly Asn Ile Ser Thr Ser Ser Asp Asp Asp Tyr Thr Asp
740 745 750

Thr Val Gly Pro Arg Lys Arg Arg Lys Asn Thr Gly Asp Val Ala Thr
755 760 765

Val Thr Ala Asn Gly Asp Ala Ser Val Thr Glu Asn Gly Met Asn Ser
770 775 780

Lys Asn Met Asn Gln Glu Leu Lys Glu Asn Lys Arg Asn Pro Glu Arg
785 790 795 800

Gly Thr Cys Gln Asn Ser Ser Phe Gln Glu Thr Asn Val Ser Pro Ala
805 810 815

ES 2 558 745 T3

Lys Ser Tyr Val Gly Ala Ser Leu Ser Gly Ser Ser Gly Lys Ser Val
 820 825 830

Arg Pro Ser Ala Tyr Lys Lys Leu Gly Glu Ala Val Thr Gln Arg Leu
 835 840 845

Tyr Ser Tyr Phe Arg Glu Asn Gln Tyr Pro Asp Arg Ala Ala Lys Ala
 850 855 860

Ser Leu Ala Glu Glu Leu Gly Ile Thr Phe Glu Gln Val Asn Lys Trp
 865 870 875 880

Phe Val Asn Ala Arg Trp Ser Phe Asn His Ser Ser Ser Thr Gly Thr
 885 890 895

Ser Lys Ala Glu Ser Ala Ser Gly Lys Gly Ser Cys Asp Gly Gln Val
 900 905 910

Arg Asp Ser Glu Ser Lys Asn Arg Lys Ser Asn Lys Gln Lys Thr Asn
 915 920 925

Thr Pro Lys Ser Arg Arg
 930

<210> 217
 <211> 2793
 <212> ADN
 <213> *Populus tremuloides*

5

<400> 217

```

atgtccgagg ctgagcacat ggggttttct ccatcaaaag tttctcatac taaaagttat      60
tcatgccctg cacaaactac actggaaaat acgcatgaac ccagtgtgta atacaagttt      120
gggtggatata cggaagagag acataagctt gagtgtgaaa tcatacaaac tgaagctgga      180
gataacaggg ctgcagttct ccaatcttgt tctggtgaag ttgtgcaacc atctactgac      240
gatttaacaa aaagtctctt tattgatttg gaccacctc ctgacgatgc aaggagtgtc      300
ctgtttgata atagtccaag acctatatca accgctatgg atcagaaact tgagccaggt      360
gctacaagtg tgaatactgc ttgtgtgcat agtgaatcat caaaagcaat tgactctagt      420
attctactgg atgaacctag gaatagcaac acagaattgt caagctgcat tgcaaatgaa      480
acctcacaag catcacttga aggtctagcc aatgattctc gtgctgaaga tgcaggactg      540
tcaactgtag aggcaagcaa tagtgacctg attgatgaaa gttcactc tcaacaaact      600
acatctggac agacacgtga gtttcattct gacagagctt gttgtaaacc attggaagaa      660
agacagaagc cgggctctga acttgacaaa aatgaatcaa tgaaattgg cattggatta      720
cccagtggta ttgctattga gaattggag ccgcttactg aacttgtagc caagagttgt      780
cctatcaaac atataggctt gcctcctggg gatgacatca gcattcctgc aatgaacaa      840
ataagaccta ctcatgataa ggagtccaaa taccctgatt gtgaacattt ggaaaaacta      900
    
```

10

ES 2 558 745 T3

tctggaattg tgattggtat tactagtcaa ggtgtaccca gtgtaagag aacttcaaaa 960
 ttgtcagga agaaatatac aagttcttcg agaaaaagtg atagagttct cgggtcaaat 1020
 tcacaagaga aacctaaagc tcccgagccg agtaataatt ctactaatgt taattccact 1080
 ggtgaggaaa aagggaaaag gaggaaaaag agaagagga aaagtatagt ggctgatgaa 1140
 tattccagaa tcagggcacg tcttaggtac ttgttgaatc ggatgagcta tgagcaaagc 1200
 ttgatcactg cttattccgg ggaaggctgg aaaggactta gtctagaaaa gttaaaacct 1260
 gagaaggagt tgcaacgggc cacatctgaa atcatcagac gcaaagtgaa aataagggat 1320
 ttatttcaac atattgattc actatgcggt gaaggaaggt tccagcatc tttgtttgat 1380
 tctgaaggac aaattgacag tgaggatata ttctgtgcaa aatgtggatc caaagatttg 1440
 actgctgata atgatattat actatgtgat ggtgcttgcg accgagggtt tcatcaattc 1500
 tgtttggtac caccattggt aagagaagac attcctcctg gtgatgaggg ctggttatgc 1560
 cctggatgtg attgcaaagt tgactgcatt gacttgctta atgactctca aggaacaaat 1620
 atatctatca gtgacagatg ggataatggt ttccccgagg cagctgcagt agcatctgga 1680
 caaaaactcg actacaactt tggattgtct tctgatgatt ctgatgataa tgattatgat 1740
 cctgatggtc cagatattga cgagaagagt caggaggaat caagctctga cgaatctgac 1800
 ttcagttctg cgtctgatga atttgaggct ccacctgatg ataaacagta cttggggctc 1860
 ctttctgatg attcagagga tgatgactat gatcctgatg ctccagttct cgaggagaag 1920
 ctgaagcagg aaagttccag ttctgatttt acttcggact ctgaggatct tgatgccact 1980
 cttaatggtg atgggttgc tttaggagat gaataccaca tgcctattga acctcatgaa 2040
 gattctaattg ggcgaagatc gagatttggg ggaaagaaga atcattcttt aaacagcaag 2100
 cttttatcca tgctagaacc agattctcac caagaaaagt ctgcacctgt ttctggaaaa 2160
 agaaatattg aaagattgga ctacaagaag ctctatgatg aaacatatgg aaacatttgt 2220
 acttcaagtg atgatgattt cactgatact gttgcaccaa ggaagaggag gaaaaacact 2280
 ggagatgttg ctatggggat agcaaagga gatgcttctg ttactgagaa tggattgaat 2340
 agcaagaata tgaatcagga attaaagaag aatgaacata cttctgggag aactcaccaa 2400
 aactcgagct ttcaagatac aaatgtttcc ccagcaaaaa cacatgttgg cgaatctctc 2460
 tctggttcta gcagtaaaag agttaggccc tctgcttata aaaaacttgg agaagctgta 2520
 actcagaaac tgtacagttt cttcaaggaa aatcggatc ctgaccaagc tgcaaaggca 2580
 agtttggcag aagagctagg aattactttt gagcaggtta acaagtgggt catgaatgct 2640
 cgttggagct tcaaccattc atcacctgaa ggtacaagta aagctgaaag tgcttcagga 2700
 aagggtagct gtgatgggca tgtgagggat agtgaatcga agaaccaaaa gagcaataaa 2760
 caaaaaacta gtaccccaa atctaggaga tag 2793

<210> 218
 <211> 930
 <212> PRT
 <213> *Populus tremuloides*

5

<400> 218

ES 2 558 745 T3

Met Ser Glu Ala Glu His Met Gly Val Ser Pro Ser Lys Val Ser His
1 5 10 15

Thr Lys Ser Tyr Ser Cys Pro Ala Gln Thr Thr Leu Glu Asn Thr His
20 25 30

Glu Pro Ser Ala Glu Tyr Lys Phe Gly Gly Tyr Pro Glu Glu Arg His
35 40 45

Lys Leu Glu Cys Glu Ile Ile Gln Thr Glu Ala Gly Asp Asn Arg Ala
50 55 60

Ala Val Leu Gln Ser Cys Ser Gly Glu Val Val Gln Pro Ser Thr Asp
65 70 75 80

Asp Leu Thr Lys Ser Pro Leu Ile Asp Leu Asp Pro Pro Pro Asp Asp
85 90 95

Ala Arg Ser Ala Leu Phe Asp Asn Ser Pro Arg Pro Ile Ser Thr Ala
100 105 110

Met Asp Gln Lys Leu Glu Pro Gly Ala Thr Ser Val Asn Thr Ala Cys
115 120 125

Val His Ser Glu Ser Ser Lys Ala Ile Asp Ser Ser Ile Leu Leu Asp
130 135 140

Glu Pro Arg Asn Ser Asn Thr Glu Leu Ser Ser Cys Ile Ala Asn Glu
145 150 155 160

Thr Ser Gln Ala Ser Leu Glu Gly Leu Ala Asn Asp Ser Arg Ala Glu
165 170 175

Asp Ala Gly Leu Ser Leu Val Glu Ala Ser Asn Ser Asp Leu Ile Asp
180 185 190

Glu Ser Ser Tyr Ser Gln Gln Thr Thr Ser Gly Gln Thr Arg Glu Phe
195 200 205

His Ser Asp Arg Ala Cys Cys Lys Pro Leu Glu Glu Arg Gln Lys Pro
210 215 220

Gly Ser Glu Leu Ala Glu Asn Glu Ser Met Glu Ile Gly Ile Gly Leu
225 230 235 240

Pro Ser Gly Ile Ala Ile Glu Asn Leu Glu Pro Leu Thr Glu Leu Val
245 250 255

Thr Lys Ser Cys Pro Ile Lys His Ile Gly Leu Pro Pro Gly Asp Asp
260 265 270

ES 2 558 745 T3

Ile Ser Ile Pro Ala Asn Glu Gln Ile Arg Pro Thr His Asp Lys Glu
 275 280 285

Ser Lys Tyr Pro Asp Cys Glu His Leu Glu Lys Leu Ser Gly Ile Val
 290 300

Ile Gly Ile Thr Ser Gln Gly Val Pro Ser Val Lys Arg Thr Ser Lys
 305 310 315 320

Leu Ser Gly Lys Lys Tyr Thr Ser Ser Ser Arg Lys Ser Asp Arg Val
 325 330 335

Leu Arg Ser Asn Ser Gln Glu Lys Pro Lys Ala Pro Glu Pro Ser Asn
 340 345 350

Asn Ser Thr Asn Val Asn Ser Thr Gly Glu Glu Lys Gly Lys Arg Arg
 355 360 365

Lys Lys Arg Arg Gly Lys Ser Ile Val Ala Asp Glu Tyr Ser Arg Ile
 370 375 380

Arg Ala Arg Leu Arg Tyr Leu Leu Asn Arg Met Ser Tyr Glu Gln Ser
 385 390 395 400

Leu Ile Thr Ala Tyr Ser Gly Glu Gly Trp Lys Gly Leu Ser Leu Glu
 405 410 415

Lys Leu Lys Pro Glu Lys Glu Leu Gln Arg Ala Thr Ser Glu Ile Ile
 420 425 430

Arg Arg Lys Val Lys Ile Arg Asp Leu Phe Gln His Ile Asp Ser Leu
 435 440 445

Cys Gly Glu Gly Arg Phe Pro Ala Ser Leu Phe Asp Ser Glu Gly Gln
 450 455 460

Ile Asp Ser Glu Asp Ile Phe Cys Ala Lys Cys Gly Ser Lys Asp Leu
 465 470 475 480

Thr Ala Asp Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asp Arg Gly
 485 490 495

Phe His Gln Phe Cys Leu Val Pro Pro Leu Leu Arg Glu Asp Ile Pro
 500 505 510

Pro Gly Asp Glu Gly Trp Leu Cys Pro Gly Cys Asp Cys Lys Val Asp
 515 520 525

Cys Ile Asp Leu Leu Asn Asp Ser Gln Gly Thr Asn Ile Ser Ile Ser
 530 535 540

ES 2 558 745 T3

Asp Arg Trp Asp Asn Val Phe Pro Glu Ala Ala Val Ala Ser Gly
 545 550 555 560
 Gln Lys Leu Asp Tyr Asn Phe Gly Leu Ser Ser Asp Asp Ser Asp Asp
 565 570 575
 Asn Asp Tyr Asp Pro Asp Gly Pro Asp Ile Asp Glu Lys Ser Gln Glu
 580 585 590
 Glu Ser Ser Ser Asp Glu Ser Asp Phe Ser Ser Ala Ser Asp Glu Phe
 595 600 605
 Glu Ala Pro Pro Asp Asp Lys Gln Tyr Leu Gly Leu Pro Ser Asp Asp
 610 615 620
 Ser Glu Asp Asp Asp Tyr Asp Pro Asp Ala Pro Val Leu Glu Glu Lys
 625 630 635 640
 Leu Lys Gln Glu Ser Ser Ser Ser Asp Phe Thr Ser Asp Ser Glu Asp
 645 650 655
 Leu Asp Ala Thr Leu Asn Gly Asp Gly Leu Ser Leu Gly Asp Glu Tyr
 660 665 670
 His Met Pro Ile Glu Pro His Glu Asp Ser Asn Gly Arg Arg Ser Arg
 675 680 685
 Phe Gly Gly Lys Lys Asn His Ser Leu Asn Ser Lys Leu Leu Ser Met
 690 695 700
 Leu Glu Pro Asp Ser His Gln Glu Lys Ser Ala Pro Val Ser Gly Lys
 705 710 715 720
 Arg Asn Ile Glu Arg Leu Asp Tyr Lys Lys Leu Tyr Asp Glu Thr Tyr
 725 730 735
 Gly Asn Ile Cys Thr Ser Ser Asp Asp Asp Phe Thr Asp Thr Val Ala
 740 745 750
 Pro Arg Lys Arg Arg Lys Asn Thr Gly Asp Val Ala Met Gly Ile Ala
 755 760 765
 Asn Gly Asp Ala Ser Val Thr Glu Asn Gly Leu Asn Ser Lys Asn Met
 770 775 780
 Asn Gln Glu Leu Lys Lys Asn Glu His Thr Ser Gly Arg Thr His Gln
 785 790 795 800
 Asn Ser Ser Phe Gln Asp Thr Asn Val Ser Pro Ala Lys Thr His Val
 805 810 815

ES 2 558 745 T3

Gly Glu Ser Leu Ser Gly Ser Ser Ser Lys Arg Val Arg Pro Ser Ala
 820 825 830

Tyr Lys Lys Leu Gly Glu Ala Val Thr Gln Lys Leu Tyr Ser Phe Phe
 835 840 845

Lys Glu Asn Arg Tyr Pro Asp Gln Ala Ala Lys Ala Ser Leu Ala Glu
 850 855 860

Glu Leu Gly Ile Thr Phe Glu Gln Val Asn Lys Trp Phe Met Asn Ala
 865 870 875 880

Arg Trp Ser Phe Asn His Ser Ser Pro Glu Gly Thr Ser Lys Ala Glu
 885 890 895

Ser Ala Ser Gly Lys Gly Ser Cys Asp Gly His Val Arg Asp Ser Glu
 900 905 910

Ser Lys Asn Gln Lys Ser Asn Lys Gln Lys Thr Ser Thr Pro Lys Ser
 915 920 925

Arg Arg
 930

5

<210> 219
 <211> 2184
 <212> ADN
 <213> *Populus tremuloides*
 <400> 219

ES 2 558 745 T3

atgggtgatt	ctggaaagaa	atcaaagcag	caagacttgc	acgaatcttc	accatctgac	60
acagtgaatg	ggtcgttgct	gattaaatca	ttgaagataa	agaaggggtgg	caaattatct	120
catagaaaaa	gcgaaaaacc	taaaactaaa	cctcatctga	aaactatcat	taattcatct	180
glatcaaaga	aaaaggttac	tcctaagaag	ggcatcagga	atggctctac	cagtagaaga	240
ttgattcaca	ggaaaattct	gcataaagca	cttgataaga	aggcctcaag	aaatggggct	300
tcctcagagc	tccaaggcaa	acagttgtca	actattgatt	ctgaggggaa	tggaaaaaat	360
gccgatgaag	gtgcaattaa	aaaagtcaag	aagaggaagc	ctaagaaaag	gcaaaaggac	420
aaggtaaagc	tagatgaacc	accacgcttg	cagaggagag	caaggtactt	gatgattaa	480
atgaagctgg	agcagaatct	tatagatgct	tactctggag	aaggttggaa	aggtaaaagt	540
cgagagaaga	ttcggccaga	aaaggaacta	ctgagagcca	ggaaacagat	tttgaaatgt	600
aaacttggat	tacgtgaaat	aattcggcag	gtggattctc	taagtacagt	aggatgtatt	660
gaagatgctg	ttatggctcc	agatggatct	gtttcccatg	aacatatatt	ctgtgcgaag	720
tgcaaattga	atgaagtttc	ccaagataat	gatattgtac	tctgtgatgg	gacatgcaac	780
tgtgccttcc	accaaaaatg	ccttgatcct	cctttggata	ctgaaaatat	acctccagga	840
gatcagggat	ggttttgcaa	gttctgtgac	tgtaggatgg	aaattataga	agccatgaat	900
gctcatttgg	ggaccactt	ctcagaggac	agcggttggc	aggatatatt	caaagaagaa	960

ES 2 558 745 T3

gcagcagttc cagatggtgg aaatatgcta ttaaateccag aagaagaatg gccttctgat 1020
gattctgaag atgatgatta tgatccagag aggagagaga acgtcatgag tggagcaggt 1080
actgatgatg atgcatctga tgacaccagc aattcaaccg gcttgagttg gtcttcagat 1140
ggtgaggttt tttcaggatc caggagggtg gaggtggacg gcttggactt cagaaacaat 1200
tctatttata gcagttttaga ttctgatgaa accagtgatg gggaaattat ttgtggccgt 1260
aggcagagaa gagcggttga ttataagaag ttatacaatg agatgtttg aaaggacgct 1320
cctgcacatg agcaaccag cgaagatgaa gattggggtc ctagcaaaag aaagcgacga 1380
gagaaggagt cagatgcagc cagcacccta atgactctat atgaaagtaa aagaagatgt 1440
aaaaatgatg caaccattga aggcattgat aagcttccac gagatcctca aattcgaagg 1500
ccaattttca ggctcccccc tgatgctgtt gagaaacttc gccaaagtatt tgacagagaat 1560
gaacttccat ctagaactgt caaggagaat ctatcaaaag aattgggcct tgaacctggg 1620
aaggtagca aatggttcaa gaattcccgt tatttagctc tgaagtctag aaaggtagag 1680
aagggagaac aagttcatta ttctagtcc aaagtctctg ctgaaccac ttaaatgtc 1740
atgaaggaca aaactgctga tcttctactg ttaaaggata gccaggcaga aactggggtg 1800
tgtaccccag agaatttgaa aaggatcttg cagaggaaga agccgaggtc aataagtaaa 1860
agtttaaga aaaatgaaca gaaaagaggt tccttcgaat cacctactaa aagcaatgag 1920
atgaatggtg agcacaatga tgatctgagc ttgaagaagc tgttaaaagc aaaaacaaag 1980
ggagtaaaga aaaagggtaa tcgtatttca gcagctgctg aatctgatat ggagaaactt 2040
tgtagagcaa agacgagagt agagaatctg aagcagaaac tgggtgaaact acaaactggc 2100
aaagctagga agtcttccaa aatccgtccg ctagacgaat ctgttgctca cgttcctatt 2160
gctgagctaa gagaaaagaa atga 2184

<210> 220
<211> 727
<212> PRT

5

<213> *Populus tremuloides*

<400> 220

10

ES 2 558 745 T3

Met Gly Asp Ser Gly Lys Lys Ser Lys Gln Gln Asp Leu His Glu Ser
1 5 10 15

Ser Pro Ser Asp Thr Val Asn Gly Ser Leu Leu Ile Lys Ser Leu Lys
20 25 30

Ile Lys Lys Gly Gly Lys Leu Ser His Arg Lys Ser Glu Lys Pro Lys
35 40 45

Thr Lys Pro His Leu Lys Thr Ile Ile Asn Ser Ser Val Ser Lys Lys
50 55 60

Lys Val Thr Pro Lys Lys Gly Ile Arg Asn Gly Ser Thr Ser Arg Arg
65 70 75 80

ES 2 558 745 T3

Leu Ile His Arg Lys Ile Leu His Lys Ala Leu Asp Lys Lys Ala Ser
 85 90 95
 Arg Asn Gly Ala Ser Ser Glu Leu Gln Gly Lys Gln Leu Ser Thr Ile
 100 105 110
 Asp Ser Glu Gly Asn Gly Lys Asn Ala Asp Glu Gly Ala Ile Lys Lys
 115 120 125
 Val Lys Lys Arg Lys Pro Lys Lys Arg Gln Lys Asp Lys Val Lys Leu
 130 135 140
 Asp Glu Pro Pro Arg Leu Gln Arg Arg Ala Arg Tyr Leu Met Ile Lys
 145 150 155 160
 Met Lys Leu Glu Gln Asn Leu Ile Asp Ala Tyr Ser Gly Glu Gly Trp
 165 170 175
 Lys Gly Lys Ser Arg Glu Lys Ile Arg Pro Glu Lys Glu Leu Leu Arg
 180 185 190
 Ala Arg Lys Gln Ile Leu Lys Cys Lys Leu Gly Leu Arg Glu Ile Ile
 195 200 205
 Arg Gln Val Asp Ser Leu Ser Thr Val Gly Cys Ile Glu Asp Ala Val
 210 215 220
 Met Ala Pro Asp Gly Ser Val Ser His Glu His Ile Phe Cys Ala Lys
 225 230 235 240
 Cys Lys Leu Asn Glu Val Ser Gln Asp Asn Asp Ile Val Leu Cys Asp
 245 250 255
 Gly Thr Cys Asn Cys Ala Phe His Gln Lys Cys Leu Asp Pro Pro Leu
 260 265 270
 Asp Thr Glu Asn Ile Pro Pro Gly Asp Gln Gly Trp Phe Cys Lys Phe
 275 280 285
 Cys Asp Cys Arg Met Glu Ile Ile Glu Ala Met Asn Ala His Leu Gly
 290 295 300
 Thr His Phe Ser Glu Asp Ser Gly Trp Gln Asp Ile Phe Lys Glu Glu
 305 310 315 320
 Ala Ala Val Pro Asp Gly Gly Asn Met Leu Leu Asn Pro Glu Glu Glu
 325 330 335
 Trp Pro Ser Asp Asp Ser Glu Asp Asp Asp Tyr Asp Pro Glu Arg Arg
 340 345 350

ES 2 558 745 T3

Glu Asn Val Met Ser Gly Ala Gly Thr Asp Asp Asp Ala Ser Asp Asp
 355 360 365
 Thr Ser Asn Ser Thr Arg Leu Ser Trp Ser Ser Asp Gly Glu Val Phe
 370 375 380
 Ser Gly Ser Arg Arg Trp Glu Val Asp Gly Leu Asp Phe Arg Asn Asn
 385 390 395 400
 Ser Ile Tyr Ser Ser Leu Asp Ser Asp Glu Thr Ser Asp Gly Glu Ile
 405 410 415
 Ile Cys Gly Arg Arg Gln Arg Arg Ala Val Asp Tyr Lys Lys Leu Tyr
 420 425
 Asn Glu Met Phe Gly Lys Asp Ala Pro Ala His Glu Gln Pro Ser Glu
 435 440
 Asp Glu Asp Trp Gly Pro Ser Lys Arg Lys Arg Arg Glu Lys Glu Ser
 450 455 460
 Asp Ala Ala Ser Thr Leu Met Thr Leu Tyr Glu Ser Lys Arg Arg Cys
 465 470 475 480
 Lys Asn Asp Ala Thr Ile Glu Gly Met Met Lys Leu Pro Arg Asp Pro
 485 490 495
 Gln Ile Arg Arg Pro Ile Phe Arg Leu Pro Pro Asp Ala Val Glu Lys
 500 505 510
 Leu Arg Gln Val Phe Ala Glu Asn Glu Leu Pro Ser Arg Thr Val Lys
 515 520 525
 Glu Asn Leu Ser Lys Glu Leu Gly Leu Glu Pro Gly Lys Val Ser Lys
 530 535 540
 Trp Phe Lys Asn Ser Arg Tyr Leu Ala Leu Lys Ser Arg Lys Val Glu
 545 550 555 560
 Lys Gly Glu Gln Val His Tyr Ser Ser Ser Lys Val Ser Ala Glu Pro
 565 570 575
 Thr Leu Asn Val Met Lys Asp Lys Thr Ala Asp Leu Ser Leu Leu Lys
 580 585 590
 Asp Ser Gln Ala Glu Thr Gly Val Cys Thr Pro Glu Asn Leu Lys Arg
 595 600 605
 Ile Leu Gln Arg Lys Lys Pro Arg Ser Ile Ser Lys Ser Leu Lys Lys
 610 615 620

ES 2 558 745 T3

Asn Glu Gln Lys Arg Gly Ser Phe Glu Ser Pro Thr Lys Ser Asn Glu
 625 630 635 640
 Met Asn Val Glu His Asn Asp Asp Leu Ser Leu Lys Lys Leu Leu Lys
 645 650 655
 Ala Lys Thr Lys Gly Val Lys Lys Lys Gly Asn Arg Ile Ser Ala Ala
 660 665 670
 Ala Glu Ser Asp Met Glu Lys Leu Cys Arg Ala Lys Thr Arg Val Glu
 675 680 685
 Asn Leu Lys Gln Lys Leu Val Lys Leu Gln Thr Gly Lys Ala Arg Lys
 690 695 700
 Ser Ser Lys Ile Arg Pro Leu Asp Glu Ser Val Val Tyr Val Pro Ile
 705 710 715 720
 Ala Glu Leu Arg Glu Lys Lys
 725

- 5 <210> 221
- <211> 2460
- <212> ADN
- <213> *Ostreococcus tauri*

- 10 <220>
- <221 > misc_feature
- <222> (178)..(178)
- <223> n es a, c, g, o t

- 15 <220>
- <221 > misc_feature
- <222> (214)..(214)
- <223> n es a, c, g, o t

- 20 <220>
- <221 > misc_feature
- <222> (1375)..(1375)
- <223> n es a, c, g, o t

- 25 <220>
- <221 > misc_feature
- <222> (1381)..(1381)
- <223> n es a, c, g, o t

- 30 <220>
- <221 > misc_feature
- <222> (1648)..(1648)
- <223> n es a, c, g, o t

- 35 <220>
- <221 > misc_feature
- <222> (1684)..(1684)
- <223> n es a, c, g, o t

- <400> 221

ES 2 558 745 T3

atgaatgagg tgcgcgacgc ggaggaacga tgtaaactgc gtgtgcgagc ggtcgcggac 60
gcggtgcgga agggaggcgt ggtgtgggcg atcgtggggt ggaaagatag tcaatgcgcg 120

ES 2 558 745 T3

cgcgcgccgt ccgacgcgcg cgttgttggg gattggggcg cgcgcatgtg tgtgaganga 180
 ctgacgatgc atcgatcggg ggttggatgat acgnagtttg cgctcgcgcg atcgcgggata 240
 cagtcgcagt tcgacgat tcgaagacat caggcgttgg tggaaagcgta cgcgagcgcgac 300
 ggctggcgcg gacaggcggc tcaaaagccg aaaccagtca gggagattga aaaggcgcgc 360
 gagaagatat tcgagggtaa attgaagata cgggagtatt ttaaggTTTT ggagttcgat 420
 gagcgcgagc gcgagatcac gacggtggcg gacgagtttg gggagtgcga cgccgcggat 480
 attttttcta gcaagtgcac cctcgcagac gacaggcatg acgacgatat tttgctctgc 540
 gatggcttct gcgatcgtgc gtatcatcaa tcatgcgtgg ctccaccggg tttggcggaa 600
 gacattccgc ccgaggacga aggttggctc tgtccgagat gcgacgcgcg agtggacgtc 660
 atttacgtct tgaacgacga gtacgatcaa aatttgggcc agaggtgtgt ttcagcggat 720
 attttcgttg cagaggcgga catgcgagat aaagggatcg ttccaggaac ggcgcagttc 780
 aaacacgcgc acgaagagga ttggccgagc gatgagagcg atgacgagga cttcgatcag 840
 ggcgacacac gcgacgatgg gcgagacgac gagcacgagg cgctgagcgg aagcgcacag 900
 tcatcgagcg atgaatcgag cagcgaatca gagtctgatc tcatcatcga gggcccacga 960
 aggcgcacga aggttgatta cgtcgccctg aacaacgcca tgttcggcga cagcgaggcg 1020
 tacgagggcg aagctgagga gctcggttgg aaacgcagca acaaaaccaa ggcggagatg 1080
 ctgcgcccgc tcaaggggtga gagcgcgaac gggataacga acggtgccac gaagaaggat 1140
 ccaaaatcaa aacgctcgaa gcgagagcgc gtggatgaga aatctcccgc accgcgcgtc 1200
 aggttcactc ccgatcaaaa acttgagctc gagcgtgtct tctccgagtg cgcgctcgtc 1260
 gacattgacg cgcgcgacgc cctcgcgaag cgactcggga taccggcgcg cgcaaattcg 1320
 attaaaatat ggttcatgaa tcgccgtcgt agggctaaga ttttgggtga gagghaatac 1380
 ngacagcagc ataatccacg tcccgtcgc tcgttcgacg cctcgtcaca aacaacctct 1440
 gaaaaacctt caagtcacct cagaagctca cgcgcgctc aaacatgtcg atcgcacctt 1500
 aacgcgcccc caaactcata tacttctccg cgtggttctg cccgttcgcc catcgagcga 1560
 cgctcgcgtt ggagcattac gcgcacggcg gagagatttc ctacgagtg aaggaatcgc 1620
 tcggtgaggt cacgcttctg tgacattnga cgctcgaatg cgacgaaact tgagcgcctc 1680
 gacngaccgg agacgaatcg ccgcgtagga tgggaaaaac gcaagtcaac ctcgagttag 1740
 gtgaagacga aacacgagaa ttggtaccat tacaagtccc cagagctact gagacacaac 1800
 ccgctgggca tggtgccgac actcgtgacg ccggcgacgt atgaggataa agatggagat 1860
 ccatcgcgga acgtggtgac ggagagcttg gtgtgcgtac agtttatcga tgaactcgtt 1920
 cgggaaacg cgtacgccgg ggagacggcg gcgatcatgc cgagcgacct gtacgcgcgc 1980
 gcgaaggcga gagtggacgc cgactgggtg acaaaaaacg tatgctcgag gtattatcac 2040
 gtctggtgac gacaagaggc gagcgcgagc agagaagcgt ttgaaaagct cgtggagggg 2100
 ctggagacat tcgcccgtg gtgcggcagc ggagaagggg agttctatgg tggacaaacg 2160
 acgcccggat tgggtgacta cgcgctcttt ccatgggctg ggcgcctccc ggtgttcgag 2220

ES 2 558 745 T3

cactaccgcg ggatggattt caagattcct cgcaccgcg cactgaaatc gtatcacgat 2280
 tggatggaag ccatgctcgc gcgagagcaa gttcgcaaga cgcttccacc gtgggacgat 2340
 tacttggaac acattggacg atacgccgac gggagcgcgc gatcgaaagt cgccaacgca 2400
 gtccgcagcg gccgtgcggc gcacgattac gacgacaagg aggacaacac cgatagttga 2460

5 <210> 222
 <211> 819
 <212> PRT
 <213> *Ostreococcus tauri*

10 <220>
 <221 > INCIERTO
 <222> (60)..(60)
 <223> Xaa puede ser cualquier aminoácido de origen natural

15 <220>
 <221 > INCIERTO
 <222> (72)..(72)
 <223> Xaa puede ser cualquier aminoácido de origen natural

20 <220>
 <221 > INCIERTO
 <222> (459)..(459)
 <223> Xaa puede ser cualquier aminoácido de origen natural

25 <220>
 <221 > INCIERTO
 <222> (461)..(461)
 <223> Xaa puede ser cualquier aminoácido de origen natural

30 <220>
 <221 > INCIERTO
 <222> (550)..(550)
 <223> Xaa puede ser cualquier aminoácido de origen natural

35 <220>
 <221 > INCIERTO
 <222> (562)..(562)
 <223> Xaa puede ser cualquier aminoácido de origen natural

<400> 222

ES 2 558 745 T3

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

ES 2 558 745 T3

				85					90				95			
Tyr	Ala	Ser	Asp 100	Gly	Trp	Arg	Gly	Gln 105	Ala	Ala	Gln	Lys	Pro 110	Lys	Pro	
Val	Arg	Glu 115	Ile	Glu	Lys	Ala	Arg 120	Glu	Lys	Ile	Phe	Glu 125	Gly	Lys	Leu	
Lys	Ile 130	Arg	Glu	Tyr	Phe	Lys 135	Val	Leu	Glu	Phe	Asp 140	Glu	Arg	Glu	Arg	
Glu 145	Ile	Thr	Thr	Val	Ala 150	Asp	Glu	Phe	Gly	Glu 155	Cys	Asp	Ala	Ala	Asp 160	
Ile	Phe	Cys	Ser	Lys 165	Cys	Thr	Leu	Ala	Asp 170	Asp	Arg	His	Asp	Asp 175	Asp	
Ile	Leu	Leu	Cys 180	Asp	Gly	Phe	Cys	Asp 185	Arg	Ala	Tyr	His	Gln 190	Ser	Cys	
Val	Ala	Pro 195	Pro	Val	Leu	Ala	Glu 200	Asp	Ile	Pro	Pro	Glu 205	Asp	Glu	Gly	
Trp	Leu 210	Cys	Pro	Arg	Cys	Asp 215	Ala	Arg	Val	Asp	Val 220	Ile	Tyr	Val	Leu	
Asn 225	Asp	Glu	Tyr	Asp	Gln 230	Asn	Leu	Gly	Gln	Arg 235	Cys	Val	Ser	Ala	Asp 240	
Ile	Phe	Val	Ala	Glu 245	Ala	Asp	Met	Arg	Asp 250	Lys	Gly	Ile	Val	Pro 255	Gly	
Thr	Ala	Gln	Phe 260	Lys	His	Ala	His	Glu 265	Glu	Asp	Trp	Pro	Ser 270	Asp	Glu	
Ser	Asp	Asp 275	Glu	Asp	Phe	Asp	Gln 280	Gly	Gly	His	Ser	Asp 285	Asp	Gly	Arg	
Asp 290	Asp	Glu	His	Glu	Ala	Leu 295	Ser	Gly	Ser	Ala	Gln 300	Ser	Ser	Ser	Asp	
Glu 305	Ser	Ser	Ser	Glu	Ser 310	Glu	Ser	Asp	Leu	Ile 315	Ile	Glu	Gly	Pro	Arg 320	
Arg	Arg	Thr	Lys	Val 325	Asp	Tyr	Val	Ala	Leu 330	Asn	Asn	Ala	Met	Phe 335	Gly	
Asp	Ser	Glu	Ala 340	Tyr	Glu	Gly	Glu	Ala 345	Glu	Glu	Leu	Gly	Trp 350	Lys	Arg	
Ser	Asn	Lys 355	Thr	Lys	Ala	Glu	Met 360	Leu	Ala	Ala	Leu	Lys 365	Gly	Glu	Ser	

ES 2 558 745 T3

Ala Asn Gly Ile Thr Asn Gly Ala Thr Lys Lys Asp Pro Lys Ser Lys
 370 375 380

Arg Ser Lys Arg Glu Arg Val Asp Glu Lys Ser Pro Ala Pro Arg Val
 385 390 395 400

Arg Phe Thr Pro Asp Gln Lys Leu Glu Leu Glu Arg Val Phe Ser Glu
 405 410 415

Cys Ala Ser Leu Asp Ile Asp Ala Arg Asp Ala Leu Ala Lys Arg Leu
 420 425 430

Gly Ile Pro Gly Gly Ala Asn Ser Ile Lys Ile Trp Phe Met Asn Arg
 435 440 445

Arg Arg Arg Ala Lys Ile Leu Gly Glu Arg Xaa Tyr Xaa Gln His Asp
 450 455 460

Asn Pro Arg Pro Ala Arg Ser Phe Asp Ala Ser Ser Gln Thr Thr Ser
 465 470 475 480

Glu Lys Pro Ser Ser His Leu Arg Ser Ser Arg Ala Ser Gln Thr Cys
 485 490 495

Arg Ser Pro Leu Asn Ala Pro Pro Asn Ser Tyr Thr Ser Pro Arg Gly
 500 505 510

Ser Ala Arg Ser Pro Ile Glu Arg Arg Ser Arg Trp Ser Ile Thr Arg
 515 520 525

Thr Ala Glu Arg Phe Pro Thr Ser Gly Arg Asn Arg Ser Val Arg Ser
 530 535 540

Arg Phe Val Asp Ile Xaa Arg Ser Asn Ala Thr Lys Leu Glu Arg Leu
 545 550 555 560

Asp Xaa Pro Glu Thr Asn Arg Arg Val Gly Trp Glu Lys Arg Lys Ser
 565 570 575

Thr Ser Ser Glu Val Lys Thr Lys His Glu Asn Trp Tyr His Tyr Lys
 580 585 590

Ser Pro Glu Leu Leu Arg His Asn Pro Leu Gly Met Val Pro Thr Leu
 595 600 605

Val Thr Pro Ala Thr Tyr Glu Asp Lys Asp Gly Asp Pro Ser Arg Asn
 610 615 620

Val Val Thr Glu Ser Leu Val Cys Val Gln Phe Ile Asp Glu Leu Val
 625 630 635 640

ES 2 558 745 T3

Arg Gly Asn Ala Tyr Ala Gly Glu Thr Ala Ala Ile Met Pro Ser Asp
645 650 655

Pro Tyr Ala Arg Ala Lys Ala Arg Val Asp Ala Asp Trp Val Asn Lys
660 665 670

Asn Val Cys Ser Arg Tyr Tyr His Val Leu Val Arg Gln Glu Ala Ser
675 680 685

Glu Gln Arg Glu Ala Phe Glu Lys Leu Val Glu Gly Leu Glu Thr Phe
690 695 700

Ala Ala Trp Cys Gly Ser Gly Glu Gly Lys Phe Tyr Gly Gly Gln Thr
705 710 715 720

Thr Pro Gly Leu Val Asp Tyr Ala Leu Phe Pro Trp Ala Trp Arg Leu
725 730 735

Pro Val Phe Glu His Tyr Arg Gly Met Asp Phe Lys Ile Pro Arg Thr
740 745 750

Ala Ala Leu Lys Ser Tyr His Asp Trp Met Glu Ala Met Leu Ala Arg
755 760 765

Glu Gln Val Arg Lys Thr Leu Pro Pro Trp Asp Asp Tyr Leu Glu His
770 775 780

Ile Gly Arg Tyr Ala Asp Gly Ser Ala Arg Ser Lys Val Ala Asn Ala
785 790 795 800

Val Arg Ser Gly Arg Ala Ala His Asp Tyr Asp Asp Lys Glu Asp Asn
805 810 815

Thr Asp Ser

<210> 223
<211> 2076
5 <212> ADN
<213> *Aquilegia formosa* x *Aquilegia pubescens*

<220>
10 <221 > misc_feature
<222> (692)..(713)
<223> n e s a, c, g, o t

<220>
15 <221 > misc_feature
<222> (1459)..(1462)
<223> n e s a, c, g, o t

<400> 223

atgcaagatg gtgggaacct tggttttcca aagaaacgtg ttggatccca aataaatacc 60
tcatttcaga agaagaatag ttgcaaaaat cgcgatatta ggagatggaa accaaagcca 120

20

ES 2 558 745 T3

agatctcatg tcaagacaat tgcttcgttc ttgtcgaaga gaaagacatc tgagctctca 180
 agcaacagag ccagaaataa atgttccaat ggaaaacctt ttagaagaag agcgtcacag 240
 aaaccaatg ataccagtcc ttctgatagg ccactttcat caatgcttaa agatttcaaa 300
 tcgttaagca ataaatgtca gggacgtgaa aagacaactg acaaggatct tggagtgcaa 360
 aaggctcaca gaaggagaag gagcaagaag aggaagaatg aggagttaga tgaagccgct 420
 cgagtccaaa gaagaacaag gtacctcctg attaaaatga agctggagca gaaccttatt 480
 gaggcttact ctggggaggg ttggaaaggc cagagtcgtg agaaaattaa gccagaaaag 540
 gaactgcaaa gagccaagaa gcagatattg aagtgcacaac ttggaattcg tgatgctatt 600
 cgtcagctgg aatctcttag ttctgagggg agcattgaag attctgtgat aggttcagat 660
 ggatctgttt tccatgaaca tattttctgt tnnnnnnnnn nnnnnnnnnn nnnagcattt 720
 ccagataatg acattatact gtgtgatgga gcatgcaatt gtgctttcca ccagaaatgc 780
 ctggagcctc cattagccac tgaaaatatt cctcctggtg atcaagggtg gtattgcaaa 840
 ttttgtgagt gcaagatggc aatacttgat gcaatcaatg cacatcttgg gactcagttt 900
 tctctagaca gtaaatggca agatattttc atggaagcag ctaatgcacg cgataatgag 960
 gatattctca ttcacctga tggtgactgg ccactctgat actctgaaga tgttgattat 1020
 gaccagaaa aatatgaaat tgattgcagc tatagcagaa tgggctcggg gcacaacgtg 1080
 tctgacgatg caagtagttc cggcaccttg tattggactt ccgaagaaga tgatggtttg 1140
 cattctgagc gattcagaac tgatggtagt ctggaggggg gattttatgg ggataaaact 1200
 aaaatccaga gttctgatag catagattcc aatgacagta ctgatggat ggcatttcat 1260
 cgtaggcagc gaagagatgt tgattataag aactgcatg atgaaatgtt tggtaaagat 1320
 atgcctgaaa gtgaagaaat cagtgaagat gaagattggg gacctagaag aagaaagcgc 1380
 agaagaaatg agtctaagc agccactact ttagtatctt cgtgtggtac tgagaatgaa 1440
 ttttcgaatg tggcactttn nntaccatca agagtggcca gggagagtct ctcgaagcaa 1500
 ttgggtattg caattgagaa gatcaacaag tggttcaaga atgctcggta tacagcttta 1560
 aaaattagaa aggcagaaaa gactgaacaa tcccctggag gtggcaagat gcaaagtgct 1620
 gacgtattaa catcaaagga aaaccattat ttggtctcaa caggaacagc agtccaaaga 1680
 tccaaagttt cgagtgtcca tagaagtaaa aaccgtaaat tgacaaccac accctcaagg 1740
 gaagtacaag agaaagtgtc tgctgttgca ccatccgcta ctgcaaatga ggtgagtgta 1800
 cagttaagaa aagcagtgag cgtagagaag gaaatgggca gtgtaaaaag aaaatctgca 1860
 caagtaaaga ggatcaagcg cacatgctcc actaaggagc agcagctata catgattgag 1920
 ttggagagaa tttgccgtct tgaggagaag cttgaaaaga tgaagaaggt attgcttagt 1980
 atagatgaca atcatagtaa tgtatcagat ggaccccata tgaacgaaca gtctgccgta 2040
 tacattccag ttgcagagct aagagaaaag atgtga 2076

<210> 224
 <211> 691
 <212> PRT
 <213> *Aquilegia formosa* x *Aquilegia pubescens*

5

<220>

ES 2 558 745 T3

<221 > INCIERTO
 <222> (231)..(238)
 <223> Xaa puede ser cualquier aminoácido de origen natural

5 <220>
 <221 > INCIERTO
 <222> (487)..(488)
 <223> Xaa puede ser cualquier aminoácido de origen natural

10 <400> 224

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

ES 2 558 745 T3

Glu Gly Ser Ile Glu Asp Ser Val Ile Gly Ser Asp Gly Ser Val Phe
 210 215 220

His Glu His Ile Phe Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala Phe
 225 230 235

Pro Asp Asn Asp Ile Ile Leu Cys Asp Gly Ala Cys Asn Cys Ala Phe
 245 250 255

His Gln Lys Cys Leu Glu Pro Pro Leu Ala Thr Glu Asn Ile Pro Pro
 260 265 270

Gly Asp Gln Gly Trp Tyr Cys Lys Phe Cys Glu Cys Lys Met Ala Ile
 275 280 285

Leu Asp Ala Ile Asn Ala His Leu Gly Thr Gln Phe Ser Leu Asp Ser
 290 295 300

Lys Trp Gln Asp Ile Phe Met Glu Ala Ala Asn Ala Arg Asp Asn Glu
 305 310 315 320

Asp Ile Ser Ile His Pro Asp Gly Asp Trp Pro Ser Asp Asp Ser Glu
 325 330 335

Asp Val Asp Tyr Asp Pro Glu Lys Tyr Glu Ile Asp Cys Ser Tyr Ser
 340 345 350

Arg Met Gly Ser Glu His Asn Val Ser Asp Asp Ala Ser Ser Ser Gly
 355 360 365

Thr Leu Tyr Trp Thr Ser Glu Glu Asp Asp Gly Leu His Ser Glu Arg
 370 375 380

Phe Arg Thr Asp Gly Ser Leu Glu Gly Gly Phe Tyr Gly Asp Lys Thr
 385 390 395 400

Lys Ile Gln Ser Ser Asp Ser Ile Asp Ser Asn Asp Ser Thr Asp Gly
 405 410 415

Met Ala Phe His Arg Arg Gln Arg Arg Asp Val Asp Tyr Lys Lys Leu
 420 425 430

His Asp Glu Met Phe Gly Lys Asp Met Pro Glu Ser Glu Glu Ile Ser
 435 440 445

Glu Asp Glu Asp Trp Gly Pro Arg Arg Arg Lys Arg Arg Arg Asn Glu
 450 455 460

Ser Asn Ala Ala Thr Thr Leu Val Ser Ser Cys Gly Thr Glu Asn Glu
 465 470 475 480

Phe Ser Asn Val Ala Leu Xaa Xaa Pro Ser Arg Val Val Arg Glu Ser

ES 2 558 745 T3

				485						490						495
	Leu	Ser	Lys	Gln 500	Leu	Gly	Ile	Ala	Ile 505	Glu	Lys	Ile	Asn	Lys 510	Trp	Phe
	Lys	Asn	Ala 515	Arg	Tyr	Thr	Ala	Leu 520	Lys	Ile	Arg	Lys	Ala 525	Glu	Lys	Thr
	Glu	Gln 530	Ser	Pro	Gly	Gly	Gly 535	Lys	Met	Gln	Ser	Ala 540	Asp	Val	Leu	Thr
	Ser 545	Lys	Glu	Asn	His	Tyr 550	Leu	Val	Ser	Thr	Gly 555	Thr	Ala	Val	Gln	Arg 560
	Ser	Lys	Val	Ser	Ser 565	Val	His	Arg	Ser	Lys 570	Asn	Arg	Lys	Leu	Thr 575	Thr
	Thr	Pro	Ser	Arg 580	Glu	Val	Gln	Glu	Lys 585	Val	Ser	Ala	Val	Ala 590	Pro	Ser
	Ala	Thr	Ala 595	Asn	Glu	Val	Ser	Val 600	Gln	Leu	Arg	Lys	Ala 605	Val	Ser	Val
	Glu	Lys 610	Glu	Met	Gly	Ser	Val 615	Lys	Arg	Lys	Ser	Ala 620	Gln	Val	Lys	Arg
	Ile 625	Lys	Arg	Thr	Cys	Ser 630	Thr	Lys	Glu	Gln	Gln 635	Leu	Tyr	Met	Ile	Glu 640
	Leu	Glu	Arg	Ile	Cys 645	Arg	Leu	Glu	Glu	Lys 650	Leu	Glu	Lys	Met	Lys 655	Lys
	Val	Leu	Leu	Ser 660	Ile	Asp	Asp	Asn	His 665	Ser	Asn	Val	Ser	Asp 670	Gly	Pro
	His	Met	Asn 675	Glu	Gln	Ser	Ala	Val 680	Tyr	Ile	Pro	Val	Ala 685	Glu	Leu	Arg
	Glu	Lys	Met													

<210> 225
 <211> 1743
 <212> ADN
 <213> Glycine max

5

<220>
 <221> misc_feature
 <222> (261)..(261)
 <223> n e s a, c, g, o t

10

<400> 225

ES 2 558 745 T3

aagaagaaaa gacaaaggaa taacatagat gttgatgatg cttcacgctt gcgaaggaga 60

acaaggtacc tcttaatcaa aatgaagcta gagcagaacc ttattgatgc ttactctgga 120

gaaggttggg aaggtcaaag tcgggaaaag attaggccag aaaaggagct actacgagca 180

aaaaagcaga ttttaaaatg taaacttagt attcgcgatg ccatacacca gctggattct 240

cttagttctg tgggtagcat ngaagattct gccattgctc cagatggatc tgtttatcat 300

gaaaatata tctgtgccaa ttgcaaaacta catgaagctt tcccagataa tgatattata 360

ctctgtgatg gcacttgcaa tcgtgctttt caccaaagat gtctcaatcc tcctttggat 420

actgaaaata ttctctctgg agatcaaggc tggttttgca agttttgtga atgtaagata 480

gaaatactag aggcaacaaa tgcccatctt gggactcaat tctccctaga cagtaattgg 540

caggatgtat tcaaggaaga ggctgctatg cctgatggtg acattgcatt actaaatcct 600

gaagaagaat ggccatcaga cgatcctgaa gatgatgatt acaatccaga gaggaagaa 660

gacaaccaca actttgacac agaaggagct gatgaaaatg catccaatga ttcaaccagt 720

tgttccagtc tgttgtcttt gaatggcgaa tgtcctccag tagatgaagg catctgtcat 780

gaatattatt ctgttaatag ctgcttagat tctgatgaat ctggggaaat agcatgtggt 840

cctaggcagc gtaaagctgt tgactacaag aaactctatg atgaaatgta tgggaaggat 900

gctccacctt gtgaacaaat gagtgaagat gaagactggg gtcctggcaa aagaaagcga 960

agagaaaagg agtctgatgc tgttgattct cttatgactt tgcatgaaag tgagaatagg 1020

catccaata atgagcacia catgacaagt aaagattctt caagcataaa gatcaaaagg 1080

cattgtttta ggattccacg tgatgcagtt gagaggctt gccaaagttt tgctgagaac 1140

gaacttctc caagatctat aagagagggg ctttcaaaag agctgggcct tgacaccgag 1200

aaggtcagca aatggttcaa aatgcacgt tacttagcac ttaaaaacag aaagcatcaa 1260

gcagaaggag gagcagatca gcttcagagt atcacttcta caaagaacag attacaaaa 1320

caggagaatg tcgatccttt gaaatcaaag accccaaaaa ttactaggac acattcccag 1380

aaggatgtta aaaatgtcaa tgggaagaaag aaaatgaagt catctaaca gaaaagacia 1440

ccagaaattc ctccacctcc aggagaaaat ggcaagaagg actttatgga gatcagtgat 1500

gatgtgagct tgaagaaact gttgaagaaa agaaagaaga ggttgataaa ttttacgttt 1560

ggaggagact ctgagttggc agagttggaa tttgaaagac taagcgaatt gaagactaaa 1620

gtagatagta tgaagcaaaa attaactgca attcaaaatt acagagttaa gggttcaccc 1680

tattcgaatg aaccatccat tgtatatgta cccacagctg tgtaaggga aaaggttgaa 1740

tga 1743

5 <210> 226
 <211> 580
 <212> PRT
 <213> *Glycine max*

10 <220>
 <221> INCIERTO
 <222> (87)..(87)
 <223> Xaa puede ser cualquier aminoácido de origen natural

ES 2 558 745 T3

<400> 226

Lys Lys Lys Arg Gln Arg Asn Asn Ile Asp Val Asp Asp Ala Ser Arg
1 5 10 15

Leu Arg Arg Arg Thr Arg Tyr Leu Leu Ile Lys Met Lys Leu Glu Gln
20 25 30

Asn Leu Ile Asp Ala Tyr Ser Gly Glu Gly Trp Lys Gly Gln Ser Arg
35 40 45

Glu Lys Ile Arg Pro Glu Lys Glu Leu Leu Arg Ala Lys Lys Gln Ile
50 55 60

Leu Lys Cys Lys Leu Ser Ile Arg Asp Ala Ile His Gln Leu Asp Ser
65 70 75 80

Leu Ser Ser Val Gly Ser Xaa Glu Asp Ser Ala Ile Ala Pro Asp Gly
85 90 95

Ser Val Tyr His Glu Asn Ile Phe Cys Ala Asn Cys Lys Leu His Glu
100 105 110

Ala Phe Pro Asp Asn Asp Ile Ile Leu Cys Asp Gly Thr Cys Asn Arg
115 120

Ala Phe His Gln Arg Cys Leu Asn Pro Pro Leu Asp Thr Glu Asn Ile
130 135 140

Pro Pro Gly Asp Gln Gly Trp Phe Cys Lys Phe Cys Glu Cys Lys Ile
145 150 155 160

Glu Ile Leu Glu Ala Thr Asn Ala His Leu Gly Thr Gln Phe Ser Leu
165 170 175

Asp Ser Asn Trp Gln Asp Val Phe Lys Glu Glu Ala Ala Met Pro Asp
180 185 190

Gly Asp Ile Ala Leu Leu Asn Pro Glu Glu Glu Trp Pro Ser Asp Asp
195 200 205

Pro Glu Asp Asp Asp Tyr Asn Pro Glu Arg Lys Glu Asp Asn His Asn
210 215 220

Phe Asp Thr Glu Gly Ala Asp Glu Asn Ala Ser Asn Asp Ser Thr Ser
225 230 235 240

Cys Ser Ser Leu Leu Ser Leu Asn Gly Glu Cys Pro Pro Val Asp Glu
245 250 255

Gly Ile Cys His Glu Tyr Tyr Ser Val Asn Ser Cys Leu Asp Ser Asp
260 265 270

ES 2 558 745 T3

Glu Ser Gly Glu Ile Ala Cys Gly Pro Arg Gln Arg Lys Ala Val Asp
 275 280 285
 Tyr Lys Lys Leu Tyr Asp Glu Met Tyr Gly Lys Asp Ala Pro Pro Cys
 290 295 300
 Glu Gln Met Ser Glu Asp Glu Asp Trp Gly Pro Gly Lys Arg Lys Arg
 305 310 315 320
 Arg Glu Lys Glu Ser Asp Ala Val Asp Ser Leu Met Thr Leu His Glu
 325 330 335
 Ser Glu Asn Arg His Pro Asn Asn Glu His Asn Met Thr Ser Lys Asp
 340 345 350
 Ser Ser Ser Ile Lys Ile Lys Arg His Cys Phe Arg Ile Pro Arg Asp
 355 360 365
 Ala Val Glu Arg Leu Arg Gln Val Phe Ala Glu Asn Glu Leu Pro Pro
 370 375 380
 Arg Ser Ile Arg Glu Gly Leu Ser Lys Glu Leu Gly Leu Asp Thr Glu
 385 390 395 400
 Lys Val Ser Lys Trp Phe Lys Asn Ala Arg Tyr Leu Ala Leu Lys Asn
 405 410 415
 Arg Lys His Gln Ala Glu Gly Gly Ala Asp Gln Leu Gln Ser Ile Thr
 420 425 430
 Ser Thr Lys Asn Arg Leu Gln Lys Gln Glu Asn Val Asp Pro Leu Lys
 435 440 445
 Ser Lys Thr Pro Lys Ile Thr Arg Thr His Ser Gln Lys Asp Val Lys
 450 455 460
 Asn Val Asn Gly Arg Lys Lys Met Lys Ser Ser Asn Lys Lys Arg Gln
 465 470 475 480
 Pro Glu Ile Pro Pro Pro Pro Gly Glu Asn Gly Lys Lys Asp Phe Met
 485 490 495
 Glu Ile Ser Asp Asp Val Ser Leu Lys Lys Leu Leu Lys Lys Arg Lys
 500 505 510
 Lys Arg Leu Ile Asn Phe Thr Phe Gly Gly Asp Ser Gln Leu Ala Glu
 515 520 525
 Leu Glu Phe Glu Arg Leu Ser Glu Leu Lys Thr Lys Val Asp Ser Met
 530 535 540

ES 2 558 745 T3

Lys Gln Lys Leu Thr Ala Ile Gln Asn Tyr Arg Val Lys Gly Ser Pro
545 550 555 560

Tyr Ser Asn Glu Pro Ser Ile Val Tyr Val Pro Thr Ala Val Leu Arg
565 570 575

Glu Lys Val Glu
580

- 5
<210> 227
<211> 1614
<212> ADN
<213> *Lotus corniculatus*

<400> 227

ES 2 558 745 T3

atgcagggta ctgagaagtt aaatggcaca ggatctacaa aatccagtaa ctcagaggaa 60
catgctgagt caaaggtaga ttcattgaga tacagaacaa atataacaaa atgccgaggg 120
aagaaacaaa aactgaaatc aaaatctcat aaactcagcg gttgcttaag tgcacaggg 180
aggacagtta ccaattcttc catcaagaga caagtaaagg attcttctaa taaaaagatt 240
attagacaaa gtctacataa aactgatggc aaatcttcac ggatgctgtc ttcaacaatg 300
atacaaggag gaaaatcttc tcttggttct agaaaggagg gcaaagatgt tgatggggag 360
gtaatgggtc aaaaagttaa aagaaagaga aagaagaaaa ggagacagga tgatgttgat 420
cttgatgatg cttcacgctt actgagaaga acaagggtacc tcttaatcaa aatgaagcta 480
gagcagaatc ttattgatgc ttactctgga gaaggttgga aaggctaaag tcgagagaag 540
attagccag aaaaggagct acaaagagcc aagaagcaga ttttgaattg taaacttagt 600
attcgggatg ccattcgcca gttagattct cttagctcat tgggcagcat tgaagattct 660
gtcattgatc cagatggatc tgtttatcat gaacatatat tctgtgcaaa ttgcaagttg 720
cgtgaagctt tcccagataa tgatataata ctctgtgatg gcacatgcaa tcgtgctttt 780
caccaacggt gtctcaacct tcctttggat actgaaaata ttctctctgg agaccaaggc 840
tggttctgca atttttgtga atgtaagata gaaatattgg aggcaacaaa tgcccatctt 900
gggactcaaa tctccttggga cagtacttgg caggatgtat tcaaggagga ggctgctata 960
cccgatggtg ataatgcatt actgaatcca gaagaagaat ggccatcaga tgatcctgaa 1020
gacgatgatt attatccaga gaggaagaa gacaaccaca gcatcaaagc agaaggaact 1080
gatgataatg catccaatga ttttaagcagt tcttccagtc tagggctctt gaatggtgaa 1140
tgttctccag tagatgaagg cacgggtctt gaatattatt ctgttaattg ctgcatagat 1200
tctgatgatt ctggggaaat agcatgtggc cgtaggcagc gtaaactctgt tgactacaag 1260
aaactatatg atgaaatggt tgggaaggat gtcctccat gtgaaactagt gagtgaagat 1320
gaagactggg gtactggcaa aagaaagcga agagaaaaag agtctgatgc tgtaataact 1380
ctcatgacta tgcataaag tgagaacaaa aatccaata atgagaatag tgatggaata 1440
agagaggggt cttcaggcaa accaataaga aggtcttggt tcaggattcc acatgatgca 1500

cttgagaagc ttcgccaagt ttttgccggag aatgagcttc ctccaagatc tgtcaaagaa 1560
ggctctttcga aagagttggg aattgatgct gcgaagggta tgccttacia cctc 1614

<210> 228
<211> 538
<212> PRT
<213> *Lotus corniculatus*

5

<400> 228

ES 2 558 745 T3

Met Gln Gly Thr Glu Lys Leu Asn Gly Thr Gly Ser Thr Lys Ser Ser
1 5 10 15

Asn Ser Glu Glu His Ala Glu Ser Lys Val Asp Ser Leu Arg Tyr Arg
20 25 30

Thr Asn Ile Thr Lys Cys Arg Gly Lys Lys Gln Lys Leu Lys Ser Lys
35 40 45

Ser His Lys Leu Ser Gly Cys Leu Ser Ala Ser Gly Arg Thr Val Thr
50 55 60

Asn Ser Ser Ile Lys Arg Gln Val Lys Asp Ser Ser Asn Lys Lys Ile
65 70 75 80

Ile Arg Gln Ser Leu His Lys Thr Asp Gly Lys Ser Ser Arg Met Leu
85 90 95

Ser Ser Thr Met Ile Gln Gly Gly Lys Ser Ser Leu Gly Ser Arg Lys
100 105 110

Glu Gly Lys Asp Val Asp Gly Glu Val Met Val Gln Lys Val Lys Arg
115 120 125

Lys Arg Lys Lys Lys Arg Arg Gln Asp Asp Val Asp Leu Asp Asp Ala
130 135 140

Ser Arg Leu Leu Arg Arg Thr Arg Tyr Leu Leu Ile Lys Met Lys Leu
145 150 155 160

Glu Gln Asn Leu Ile Asp Ala Tyr Ser Gly Glu Gly Trp Lys Gly Gln
165 170 175

Ser Arg Glu Lys Ile Arg Pro Glu Lys Glu Leu Gln Arg Ala Lys Lys
180 185 190

Gln Ile Leu Asn Cys Lys Leu Ser Ile Arg Asp Ala Ile Arg Gln Leu
195 200 205

Asp Ser Leu Ser Ser Leu Gly Ser Ile Glu Asp Ser Val Ile Asp Pro
210 215 220

Asp Gly Ser Val Tyr His Glu His Ile Phe Cys Ala Asn Cys Lys Leu

ES 2 558 745 T3

225 230 235 240
 Arg Glu Ala Phe Pro Asp Asn Asp Ile Ile Leu Cys Asp Gly Thr Cys
 245 250 255
 Asn Arg Ala Phe His Gln Arg Cys Leu Asn Pro Pro Leu Asp Thr Glu
 260 265 270
 Asn Ile Pro Pro Gly Asp Gln Gly Trp Phe Cys Asn Phe Cys Glu Cys
 275 280 285
 Lys Ile Glu Ile Leu Glu Ala Thr Asn Ala His Leu Gly Thr Gln Ile
 290 295 300
 Ser Leu Asp Ser Thr Trp Gln Asp Val Phe Lys Glu Glu Ala Ala Ile
 305 310 315 320
 Pro Asp Gly Asp Asn Ala Leu Leu Asn Pro Glu Glu Glu Trp Pro Ser
 325 330
 Asp Asp Pro Glu Asp Asp Asp Tyr Tyr Pro Glu Arg Lys Glu Asp Asn
 340 345
 His Ser Ile Lys Ala Glu Gly Thr Asp Asp Asn Ala Ser Asn Asp Leu
 355 360 365
 Ser Ser Ser Ser Ser Leu Gly Ser Leu Asn Gly Glu Cys Ser Pro Val
 370 375 380
 Asp Glu Gly Thr Gly Leu Glu Tyr Tyr Ser Val Asn Cys Cys Ile Asp
 385 390 395 400
 Ser Asp Asp Ser Gly Glu Ile Ala Cys Gly Arg Arg Gln Arg Lys Ser
 405 410 415
 Val Asp Tyr Lys Lys Leu Tyr Asp Glu Met Phe Gly Lys Asp Ala Pro
 420 425 430
 Pro Cys Glu Leu Val Ser Glu Asp Glu Asp Trp Gly Thr Gly Lys Arg
 435
 Lys Arg Arg Glu Lys Glu Ser Asp Ala Val Asn Thr Leu Met Thr Met
 450 455 460
 His Glu Ser Glu Asn Lys Asn Pro Asn Asn Glu Asn Ser Asp Gly Ile
 465 470 475 480
 Arg Glu Gly Ser Ser Gly Lys Pro Ile Arg Arg Ser Cys Phe Arg Ile
 485 490 495
 Pro His Asp Ala Leu Glu Lys Leu Arg Gln Val Phe Ala Glu Asn Glu
 500 505 510

ES 2 558 745 T3

Leu Pro Pro Arg Ser Val Lys Glu Gly Leu Ser Lys Glu Leu Gly Ile
 515 520 525

Asp Ala Ala Lys Val Met Pro Tyr Asn Leu
 530 535

5 <210> 229
 <211> 573
 <212> ADN
 <213> *Sorghum propinquum*
 <400> 229

aactcgagaa caaatgaaga aaaatctggt atatcagggga agattgattc agtagacaac 60
 tcttgcttgg ttcccttgtc tgaaatcadc aatgtgcata cactgagtgca ccacaatctt 120
 gagatgagga agatggaatc aactagcagt cctgtgtggc tgcacaatga aggaggctgc 180
 ttgtttccaa cactccaagc caaggagagt acattaccta caagcaagcc ttgtttgccg 240
 tcagaaataa cccatccaac aactaatgaa gtgagcactt tagtgcaagc tacttcttgg 300
 atggatgctg gggcgtgcat tgagcagcaa gaaaccactc cttgggtgga taccggggcc 360
 tcaggttacc agcctttcct ggacgtgatt gatgagatgt gcggacttga gtgcaggctg 420
 cagaggctga aggagaacat gttctcatct ggcattggacg gcagaaccgc cgggtgtgagt 480
 gacatgggaa accaggctgt ggtgcttgtg ccgactgctg agctcaagga aaaagcgccg 540
 catggcgggtt tatttgggca ctattgccca tag 573

10
 15 <210> 230
 <211> 190
 <212> PRT
 <213> *Sorghum propinquum*
 <400> 230

ES 2 558 745 T3

Asn Ser Arg Thr Asn Glu Glu Lys Ser Gly Ile Ser Gly Lys Ile Asp
1 5 10 15

Ser Val Asp Asn Ser Cys Leu Val Pro Leu Ser Glu Ile Ile Asn Val
20 25 30

His Thr Arg Val His His Asn Leu Glu Met Arg Lys Met Glu Ser Thr
35 40 45

Ser Ser Pro Val Trp Leu His Asn Glu Gly Gly Cys Leu Phe Pro Thr
50 55 60

Leu Gln Ala Lys Glu Ser Thr Leu Pro Thr Ser Lys Pro Cys Leu Pro
65 70 75 80

Ser Glu Ile Thr His Pro Thr Thr Asn Glu Val Ser Thr Leu Val Gln
85 90 95

Ala Thr Ser Trp Met Asp Ala Gly Ala Cys Ile Glu Gln Gln Glu Thr
100 105 110

Thr Pro Trp Val Asp Thr Gly Ala Ser Gly Tyr Gln Pro Phe Leu Asp
115 120 125

Val Ile Asp Glu Met Cys Gly Leu Glu Cys Arg Leu Gln Arg Leu Lys
130 135 140

Glu Asn Met Phe Ser Ser Gly Met Asp Gly Arg Thr Ala Gly Val Ser
145 150 155 160

Asp Met Gly Asn Gln Ala Val Val Leu Val Pro Thr Ala Glu Leu Lys
165 170 175

Glu Lys Ala Pro His Gly Gly Leu Phe Gly His Tyr Cys Pro
180 185 190

5 <210> 231
<211> 532
<212> ADN
<213> *Sorghum propinquum*

10 <400> 231

ES 2 558 745 T3

aagataaaat tggagcagaa tctactagat gcttactctg gagatggatg gaatgggcaa 60
 agccgggaga aaataaagcc agagaaggaa ctgcagcgtg ccaggaaaca aatcataaaa 120
 tgcaaaattg ctatacgtga tatcattcgc caactttgtt tgtatacttc tactgggagt 180
 gttgatgacc cagcaatgcc tccagatcag tttaccaatc ccgaacatac catgtgctca 240
 acatgcaagt ctcatgaatc atttcccagc aataaattta tcttctgtga agggccctgc 300
 aaaagggcat atcatgagaa atgtttggaa cctcccctaa acaaagggtg acttccaaca 360
 agtagccatg gatggctttg caaattctgt ttgtgcaagg tgaaaatttt agaaactatt 420
 aatgcacatc taggaacaag ttttacagtg atgtgctcct ttgaagatat attcaaggaa 480
 gcaacagaac aaatagactc cgaggatgca ctagatgaag attggctctc tg 532

<210> 232
 <211> 177
 <212> PRT
 <213> *Sorghum propinquum*
 <400> 232

5

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

10

ES 2 558 745 T3

Ala Met Pro Pro Asp Gln Phe Thr Asn Pro Glu His Thr Met Cys Ser
 65 70 75 80

Thr Cys Lys Ser His Glu Ser Phe Pro Ser Asn Lys Phe Ile Phe Cys
 85 90 95

Glu Gly Pro Cys Lys Arg Ala Tyr His Glu Lys Cys Leu Glu Pro Pro
 100 105 110

Leu Asn Lys Gly Val Leu Pro Thr Ser Ser His Gly Trp Leu Cys Lys
 115 120 125

Phe Cys Leu Cys Lys Val Lys Ile Leu Glu Thr Ile Asn Ala His Leu
 130 135 140

Gly Thr Ser Phe Thr Val Met Cys Ser Phe Glu Asp Ile Phe Lys Glu
 145 150 155 160

Ala Thr Glu Gln Ile Asp Ser Glu Asp Ala Leu Asp Glu Asp Trp Leu
 165 170 175

Ser

<210> 233
 <211> 180
 <212> PRT
 <213> *Oryza sativa*
 <400> 233

5

Asp Glu Val Ser Arg Met Glu Lys Arg Ala Arg Tyr Leu Leu Ile Lys
 1 5 10 15

Ile Lys Gln Glu Gln Asn Leu Leu Asp Ala Tyr Ser Gly Asp Gly Trp
 20 25 30

Asn Gly His Ser Arg Glu Lys Ile Lys Pro Glu Lys Glu Leu Gln Arg
 35 40 45

Ala Lys Lys Gln Ile Met Lys Tyr Lys Ile Ala Ile Arg Asp Val Ile
 50 55 60

His Gln Leu Asp Leu Cys Ser Ser Ser Gly Ser Lys Asp Asp Ser Val
 65 70 75 80

Ile Pro Pro Asp Gly Cys His Glu Ser Val Asn Pro Glu His Thr Ile
 85 90 95

Cys Ser Arg Cys Lys Ser His Glu Ser Phe Pro Asp Asn Asn Ile Ile
 100 105 110

10

ES 2 558 745 T3

Phe Cys Glu Gly Gly Cys Lys Leu Ala Cys His Gln Lys Cys Leu Glu
 115 120 125
 Pro Pro Phe Asp Lys Ile Leu Pro Thr Thr Arg His Gly Arg Leu Cys
 130 135 140
 Lys His Cys Ser Ser Lys Met Lys Ile Leu Asp Ala Ile Asn Ala His
 145 150 155 160
 Leu Gly Thr Ser Phe Thr Val Lys Cys Pro Ser Ser Asp Ile Phe Lys
 165 170 175
 Glu Ala Ala Glu
 180

5 <210> 234
 <211> 50
 <212> PRT
 <213> *Oryza sativa*
 <400> 234

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

10 Ala Ala
 50
 15 <210> 235
 <211> 2194
 <212> ADN
 <213> *Oryza sativa*
 <400> 235

ES 2 558 745 T3

aatccgaaaa	gtttctgcac	cgttttcacc	ccctaactaa	caatataggg	aacgtgtgct	60
aaatataaaa	tgagacctta	tatatgtagc	gctgataact	agaactatgc	aagaaaaact	120
catccaccta	ctttagtggc	aatcgggcta	aataaaaaag	agtcgctaca	ctagtttcgt	180
tttccttagt	aattaagtgg	gaaaatgaaa	tcattattgc	ttagaatata	cgttcacatc	240
tctgtcatga	agttaaatta	ttcgaggtag	ccataattgt	catcaaactc	ttcttgaata	300
aaaaaatcct	tctagctgaa	ctcaatgggt	aaagagagag	atTTTTTTta	aaaaaataga	360
atgaagatat	tctgaacgta	ttggcaaaga	tttaaacata	taattatata	atTTTTatagt	420
ttgtgcattc	gtcatatcgc	acatcattaa	ggacatgtct	tactccatcc	caatTTTTat	480
ttagtaatta	aagacaattg	acttattttt	attattttatc	ttttttcgat	tagatgcaag	540
gtacttacgc	acacactttg	tgctcatgtg	catgtgtgag	tgcacctcct	caatacacgt	600

ES 2 558 745 T3

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
 tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaagaa 960
 aaccaagcat cctccttctc ccatctataa attcctcccc ccttttcccc tctctatata 1020
 ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080
 cgaccgcctt ctcgatccat atcttccggt cgagttcttg gtcgatctct tccctcctcc 1140
 acctcctct cacagggtat gtgcctcctc tcggttggtc ttggatttat tgttctaggt 1200
 tgtgtagtag gggcgttgat gttaggaaag gggatctgta tctgtgatga ttctgttct 1260
 tggatttggg atagaggggt tcttgatggt gcatgttatc ggttcggttt gatttagtagt 1320
 atggttttca atcgtctgga gagctctatg gaaatgaaat ggtttaggga tcggaatctt 1380
 gcgattttgt gaggacctt tgtttgaggt aaaatcagag caccggtgat tttgcttggg 1440
 gtaataaagt acggttggtt ggtcctcgat tctggtagtg atgcttctcg atttgacgaa 1500
 gctatccttt gtttattccc tattgaacaa aaataatcca actttgaaga cgggcccggt 1560
 gatgagattg aatgattgat tcttaagcct gtccaaaatt tcgcagctgg cttgtttaga 1620
 tacagtagtc cccatcacga aattcatgga aacagttata atcctcagga acaggggatt 1680
 cctgttctt cggatttgct ttagtcccag aattttttt cccaaatata ttaaaaagtc 1740
 actttctggt tcagttcaat gaattgattg ctacaaataa tgcttttata gcgttatcct 1800
 agctgtagtt cagttaatag gtaatacccc tatagtttag tcaggagaag aacttatccg 1860
 atttctgac tccattttta attatatgaa atgaactgta gcataagcag tattcatttg 1920
 gattatttt tttattagct ctcaccctt cattattctg agctgaaagt ctggcatgaa 1980
 ctgcctcaa ttttgttttc aaattcacat cgattatcta tgcattatcc tcttgtatct 2040
 acctgtagaa gtttctttt ggttattcct tgactgcttg attacagaaa gaaatttatg 2100
 aagctgtaat cgggatagtt atactgcttg ttcttatgat tcatttctt tgtgcagttc 2160
 ttggtgtagc ttgccacttt caccagcaaa gttc 2194

5 <210> 236
 <211> 56
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> cebador: prm09687

<400> 236
 ggggacaagt ttgtacaaaa aagcaggctt aaacaatgaa taccacagaa aagaaa 56

15 <210> 237
 <211> 50
 <212> ADN
 <213> Secuencia artificial

ES 2 558 745 T3

<220>

<223> cebador: prm09688

<400> 237

5 ggggaccact ttgtacaaga aagctgggtg atgcaagggt aagatgcttt 50

<210> 238

<211> 2649

<212> ADN

10 <213> *Physcomitrella patens*

<400> 238

atgatgaaga attaccgtcc tccatcacca cttccaagtg ctcgccaagc tttcttaca 60
gacccggcgt cgtcattcca aatgacagga cgacattggg cagaccaga gggaggtgga 120
gaacggaagc ttcacacaca ctctactca cactcactca ctcacgaagg aaacgaaggg 180
caaaggaagg ggagccccgc cgtgagccga gagcggagag caagcattcc caacggaagt 240
ctccgtcaat gcggtgccgg agccgacgac gacgacgatg atgggctgag cctgtcatcg 300
caatggatgc attgccctgt aagtgggaag gtggaggaa gattggggaa gaggttgagg 360
aagagaaaga gaaagagaaa gagtttgttg ttgttgttg ggaaggggag ggggcgacgg 420
atgatgcccc tgggcgtagt ccgacgacgat gaggcccagg ccgctgctgg ctcgccgcct 480
gcgcccacc ctcccccaat cgccaagccg acgctgacgat cccactcgcg ctccaggggc 540
cacaaggcgc gctccccaca cccgccgctg ggcgatggcg atgcgccctt gccccaccg 600
caaggaacag caacaccagg aggaggagga gcagcagcag tgggcgcacc agggggccatg 660
gatgccgagg ctatggcggc ggggcagcgg cggcgtcgac gacgccgacg acagcggggc 720
aagaactcgg atgtgccgct tgacaatcgc acgcgaatta agaggcgtgt caagtacttg 780
ctcatgaaaa tgcgagtcga ccagaacctg ctggatgcgt actccgggga gggttggaag 840
gggcagagga atttctttt tggcatatcc agtcgggaaa agatcaggcc ggagcaagaa 900
ctacagcgtg ctgaagcgca gatcctgcat tcgaaattgg ccatacggga ggctattcat 960
gaattggata acttgggggtt agaaggctcg ctcgacgaaa atgcatttga tgctgaaggc 1020
cgtgtttacc atgaagagat attttgtgcc aagtgccgat ctcaagatgc tctacctgac 1080
aacgacatta ttttatgtga cggcgcgatgc aacagaggct tccatcagta ctgtttagac 1140
ccacctttag ctacgatcaa cattcctcct ggagacgaag ggtggttgtg tcctgtatgt 1200
gagtgcaaga tggagtgcac tgaagttatt aatgcatact tgggcacaca ctttgaggtg 1260
gagaatagtt gggaggagtt gttttctgaa gaagctctcg tggctgcggg cggtagaact 1320
caaggggtca cagatgggtga ttttccttcc tcagactcag aagacgacga ttataacct 1380
gaaggagcgg agaagcggac agacagtgaag agcgaaggcg aaagtgatga aggtggggga 1440
gatagtgatt caggaagcga ttttagggaa aatgaagcgt ctggtagtga caaagatgag 1500
caagatccaa tgcaggcaac taatttagag agacttgcca gaacagcaaa gcggaccagc 1560
aaggaaggaa gtggaggtcc agaacttca ggcagtgatg actcagaagt ttttcttgtt 1620

ES 2 558 745 T3

gatgggttta cggatggtgg agtcaaatta ggtttaagac gtagagactc acaagcagaa 1680
 gaatcagctt cagacgagga aaccatgggt atagcaggaa aaagacatcg aaaagcagta 1740
 gactacaaga gattacatga tgaaatgttt gggaaacatgg agattgatgt ggacgatatc 1800
 atatccgaag atgaggattg ggggccgaaa cgccgtcggga gaagaacatc tcctgatgac 1860
 ccaagcaggt ctagacctcc tcgagttcga cggagaaagg cgcaaacctc aactggagat 1920
 gctttaaaag aagttgatgc agatggggta ccatcagatt taccctctgg gtcacaaggc 1980
 gagggagcag ccgatactct tgaagctggt agagacaaac gaatgtggag gaagttgcct 2040
 gattctgctg ttgagacggt acgactggct gcggcagtga cgacactgcc ctcaaaatcc 2100
 cgtaaagagg agttggcgac acagcttggc ttgagttttt cacaagtcca tggatggttt 2160
 aaaaatcagc ggtatcttagc tctcagaaat gggttggcga tttcaaagcg accagcagct 2220
 gggcggatgg cggccactgt aaccatacaa tcggaacacg acaacaattc aagtgcctat 2280
 ttaactgaga agttggatga agtccaaatg cgactttttg aactgaagca aactttggaa 2340
 gatttcgcta agaattcatc tggtttaggc ggtggagaat cagataacag ttttgaaaat 2400
 ggaggaattg gacttcctgc aaatgggaaa cggattatct atgtcccagt ggctgaggtt 2460
 agagaacggc cttctctttt ctctgacagc tgtccgcttt cattattagc agtggttaaat 2520
 tgcgaggtga attacaagag gaatctacgt atattcatgt gttcatggcc tagctgctta 2580
 gcgtacgggg ttttgcattg caactgccct tgtcatattg gatcatgttt cgaaggaata 2640
 ggtgggtga 2649

<210> 239
 <211> 882
 <212> PRT
 <213> *Physcomitrella patens*
 <400> 239

5

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

10

ES 2 558 745 T3

Ser Leu Ser Ser Gln Trp Met His Cys Pro Val Ser Gly Lys Val Glu
100 105 110

Glu Gly Leu Gly Lys Arg Leu Arg Lys Arg Lys Arg Lys Arg Lys Ser
115 120 125

Leu Leu Leu Leu Leu Gly Lys Gly Arg Gly Arg Arg Met Met Pro Met
130 135 140

Gly Val Val Arg Asp Asp Glu Ser Gln Ala Ala Ala Gly Ser Pro Pro
145 150 155 160

Ala Pro Thr Pro Pro Ile Ala Lys Pro Thr Leu Arg Ser His Ser
165 170 175

Arg Ser Arg Ala His Lys Ala Arg Ser Gln His Pro Pro Leu Gly Asp
180 185 190

Gly Asp Ala Pro Leu Pro Pro Pro Gln Gly Thr Ala Thr Pro Gly Gly
195 200 205

Gly Gly Ala Ala Ala Val Gly Ala Pro Gly Ala Met Asp Ala Glu Ala
210 215 220

Met Ala Ala Gly Gln Arg Arg Arg Arg Arg Arg Arg Arg Gln Arg Gly
225 230 235 240

Lys Asn Ser Asp Val Pro Leu Asp Asn Arg Thr Arg Ile Lys Arg Arg
245 250 255

Val Lys Tyr Leu Leu Met Lys Met Arg Val Asp Gln Asn Leu Leu Asp
260 265 270

Ala Tyr Ser Gly Glu Gly Trp Lys Gly Gln Arg Asn Ser Leu Phe Gly
275 280 285

Ile Ser Ser Arg Glu Lys Ile Arg Pro Glu Gln Glu Leu Gln Arg Ala
290 295 300

Glu Ala Gln Ile Leu His Ser Lys Leu Ala Ile Arg Glu Ala Ile His
305 310 315 320

Glu Leu Asp Asn Leu Gly Leu Glu Gly Ser Leu Asp Glu Asn Ala Phe
325 330 335

Asp Ala Glu Gly Arg Val Tyr His Glu Glu Ile Phe Cys Ala Lys Cys
340 345 350

Arg Ser Gln Asp Ala Leu Pro Asp Asn Asp Ile Ile Leu Cys Asp Gly
355 360 365

Ala Cys Asn Arg Gly Phe His Gln Tyr Cys Leu Asp Pro Pro Leu Ala

ES 2 558 745 T3

370	375	380																			
Thr 385	Ile	Asn	Ile	Pro	Pro 390	Gly	Asp	Glu	Gly	Trp 395	Leu	Cys	Pro	Val	Cys 400						
Glu	Cys	Lys	Met	Glu 405	Cys	Ile	Glu	Val	Ile 410	Asn	Ala	Tyr	Leu	Gly 415	Thr						
His	Phe	Glu	Val 420	Glu	Asn	Ser	Trp	Glu 425	Glu	Leu	Phe	Ser	Glu 430	Glu	Ala						
Leu	Val	Ala 435	Ala	Gly	Gly	Arg	Thr 440	Gln	Gly	Val	Thr	Asp 445	Gly	Asp	Phe						
Pro	Ser 450	Ser	Asp	Ser	Glu	Asp 455	Asp	Asp	Tyr	Asn	Pro 460	Glu	Gly	Ala	Glu						
Lys 465	Arg	Thr	Asp	Ser	Glu 470	Ser	Glu	Gly	Glu	Ser 475	Asp	Glu	Gly	Gly	Gly 480						
Asp	Ser	Asp	Ser	Gly 485	Ser	Asp	Phe	Arg	Glu 490	Asn	Glu	Ala	Ser	Gly 495	Ser						
Asp	Lys	Asp	Glu 500	Gln	Asp	Pro	Met	Gln 505	Ala	Thr	Asn	Leu	Glu 510	Arg	Leu						
Ala	Arg	Thr 515	Ala	Lys	Arg	Thr	Ser 520	Lys	Glu	Gly	Ser	Gly 525	Gly	Pro	Glu						
Ser	Ser 530	Gly	Ser	Asp	Asp	Ser 535	Glu	Val	Phe	Leu	Val 540	Asp	Gly	Phe	Thr						
Asp 545	Gly	Gly	Val	Lys	Leu 550	Gly	Leu	Arg	Arg	Arg 555	Asp	Ser	Gln	Ala	Glu 560						
Glu	Ser	Ala	Ser	Asp 565	Glu	Glu	Thr	Met	Val 570	Ile	Ala	Gly	Lys	Arg 575	His						
Arg	Lys	Ala	Val 580	Asp	Tyr	Lys	Arg	Leu 585	His	Asp	Glu	Met	Phe 590	Gly	Asn						
Met	Glu	Ile 595	Asp	Val	Asp	Asp	Ile 600	Ile	Ser	Glu	Asp	Glu 605	Asp	Trp	Gly						
Pro 610	Lys	Arg	Arg	Arg	Arg	Arg 615	Thr	Ile	Pro	Asp	Asp 620	Pro	Ser	Arg	Ser						
Arg 625	Pro	Pro	Arg	Val	Arg 630	Arg	Arg	Lys	Ala	Arg 635	Thr	Ser	Thr	Gly	Asp 640						
Ala	Leu	Lys	Glu	Val 645	Asp	Ala	Asp	Gly	Val 650	Pro	Ser	Asp	Leu	Pro 655	Ser						

ES 2 558 745 T3

Gly Ser Gln Gly Glu Gly Ala Ala Asp Thr Leu Glu Ala Val Arg Asp
660 665 670

Lys Arg Met Trp Arg Lys Leu Pro Asp Ser Ala Val Glu Thr Leu Arg
675 680 685

Leu Ala Ala Ala Val Thr Thr Leu Pro Ser Lys Ser Arg Lys Glu Glu
690 695 700

Leu Ala Thr Gln Leu Gly Leu Ser Phe Ser Gln Val His Gly Trp Phe
705 710 715 720

Lys Asn Gln Arg Tyr Leu Ala Leu Arg Asn Gly Leu Ala Ile Ser Lys
725 730 735

Arg Pro Ala Ala Gly Arg Met Ala Ala Thr Val Thr Ile Gln Ser Glu
740 745 750

His Asp Asn Asn Ser Ser Ala Tyr Leu Thr Glu Lys Leu Asp Glu Val
755 760 765

Gln Met Arg Leu Phe Glu Leu Lys Gln Thr Leu Glu Asp Phe Ala Lys
770 775 780

Asn Ser Ser Gly Leu Gly Gly Gly Glu Ser Asp Asn Ser Phe Glu Asn
785 790 795 800

Gly Gly Ile Gly Leu Pro Ala Asn Gly Lys Arg Ile Ile Tyr Val Pro
805 810 815

Val Ala Glu Val Arg Glu Arg Pro Ser Leu Phe Ser Asp Ser Cys Pro
820 825 830

Leu Ser Leu Leu Ala Val Leu Asn Cys Glu Val Asn Tyr Lys Arg Asn
835 840 845

Leu Arg Ile Phe Met Cys Ser Trp Pro Ser Cys Leu Ala Tyr Gly Val
850 855 860

Leu His Ala Asn Cys Pro Cys His Ile Gly Ser Cys Phe Glu Gly Ile
865 870 875 880

Gly Gly

<210> 240
<211> 2379
5 <212> ADN
<213> *Physcomitrella patens*
<400> 240

atggtgccga cgggcgtggt ggagcgggag gaccagcctg agcctgctgc tgcccccca 60

10

ES 2 558 745 T3

gtgccggtga ttacacccaa gacgcttcga tcgactccta tcaggacgca gaccacgcca 120
 ggtgatggcg atgctggtgc ggatcatgct cctggagcag tggcagcagc agcagcatca 180
 gcaggaggag gaggagcagc ggatgtcatg caaggtaacg atgcggggca ggaacggcga 240
 ggacgacggc gcgggcgacg gaagaagtca gacgtgccgc tcgataatcc cacgcgaatc 300
 ctaaagcatg tcaagtactt gctcattaa atgcgagtc accagaacct gctcgatgcg 360
 tacacagggg agggttggaa gggtcagagt cgcgagaaga tcaggccgga gcaagaacta 420
 caacgtgcca aggcgcagat actgcatctt aagttggcca tacgggaggc tattcatgaa 480
 ctcgatgacg tgggtttaga aggttctctc gacaaaaacg cttttgattc tgaaggctgt 540
 atttaccatg aagagatttt ttgtgcaag tgcaaatccc aagaagctct tcttgacaac 600
 gacattatct tgtgacgagc ggcgtgcaac agaggcttcc atcagtattg cttagaccca 660
 cctttagcta cggaagacat tctcctgga gatgaaggat ggttatgccc tgtatgagc 720
 tgcaagatgg agtgtattga agcgataaat tcatattttg gcaccagttt tgaggctgag 780
 aatagttggg agagcttttt ttcgaacgaa gctggtatag ctgcaggcgg tggaacgcaa 840
 gaaggggctg gtggtgattg gccttctca ggcgatgaag atgatgagta tgatccagaa 900
 acagcgggaag ggccttcttc agcttctgaa agtggagctc aagtaatagg gcttctggt 960
 gactggccgt cctcagacga cgaagatcac gactatgatc cggaagcagc ggaagaacct 1020
 cctggagctt ctgggagtgg aactcaaaca gtaggtcttg acgttgattg gccttcaca 1080
 gacgatgaag atgacgactt tgatccagaa ggtatcagaa agcaggcaaa cagtgcagaa 1140
 gaaagtagta gtgatgaaag tgaggaggac ggggaggata gtggttctgt aagcggcttc 1200
 agggaaaatg aggcgtcttc aggtgtaagt gacaaagatg agcaggagcc tagtgaaaaa 1260
 cataatctgg agctaactgc tgtcactaac atcaagaaac ttgcaagaaa aggaaaacgt 1320
 aacataatca gcaacaaagg aaaaagtga tatttagagt cttcagacag tgatgattct 1380
 gatgtctctc ttttgacga gttcaagggg ggtaatgcca atttaggctt aagacgtggg 1440
 gatccacagg cagcagatc agcagcagac gaggaagccg tgattattgc tgggaaaaga 1500
 catcggaag ctgtggacta caaaagacta catgatgaaa tgtatggtaa agcggaggcc 1560
 gataacgatg acgacgttat ctctgaggat gaagattggg gtctgaacg tcgccgaaga 1620
 cgaaccattc ctgatgatcc taactcgcca aggtcccga ttcgaggtag gaaagtgaga 1680
 actccaagtg gagatattc taaaggagcc gatgcagatt taccctctgg attaccagat 1740
 atgccatccc actttccac tggttcaca ggtgaagtag gctcaccaca agacattcca 1800
 gcactagatg gagcagctga cagtcctgga ggagaaaagc ggatgtggcg gcgattgcct 1860
 gattccgcag tcgaggcatt acgggtgatt cttgatgtta ataggctgcc ctcaaaatct 1920
 cggaaggaag aattagcaat taagcttggg ttgagctttt cacaggttca tggatggttt 1980
 aaaaatcagc ggcatacagc tctgaggaaa gggttggcat gtcttaagcg gatacatcca 2040
 gtagcatcaa gctcaagaga tttcaatact gtcgaatctc catcgtctgt ggcgacttta 2100

ES 2 558 745 T3

actgaaagca cattaccgca aaaagagaat gctccaaatt ctattttaac tgagaagttg 2160
gacgaagttc aggtgcgact cctcgaattg aagcaagcgt tggacgattt taccaagaat 2220
gcagcagtggt ttgctgatgg cggaagtatt gaattggata acaacgtgga aaacggagga 2280
attggacttc ctgaaaatgg aaaacgtggt gtctatgtgc cagttgtggc cgacgttagg 2340
gaacaatcac caatgagtac gcatactagc gcatgttga 2379

<210> 241
<211> 792
<212> PRT
<213> *Physcomitrella patens*

5

<400> 241

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

10

ES 2 558 745 T3

Cys Asn Arg Gly Phe His Gln Tyr Cys Leu Asp Pro Pro Leu Ala Thr
 210 215 220
 Glu Asp Ile Pro Pro Gly Asp Glu Gly Trp Leu Cys Pro Val Cys Asp
 225 230 235 240
 Cys Lys Met Glu Cys Ile Glu Ala Ile Asn Ser Tyr Phe Gly Thr Ser
 245 250 255
 Phe Glu Val Glu Asn Ser Trp Glu Ser Phe Phe Ser Asn Glu Ala Gly
 260 265 270
 Ile Ala Ala Gly Gly Gly Thr Gln Glu Gly Ala Gly Gly Asp Trp Pro
 275 280 285
 Ser Ser Gly Asp Glu Asp Asp Glu Tyr Asp Pro Glu Thr Ala Glu Gly
 290 295 300
 Pro Ser Ser Ala Ser Glu Ser Gly Ala Gln Val Ile Gly Leu Ala Val
 305 310 315 320
 Asp Trp Pro Ser Ser Asp Asp Glu Asp His Asp Tyr Asp Pro Glu Ala
 325 330 335
 Ala Glu Glu Pro Pro Gly Ala Ser Gly Ser Gly Thr Gln Thr Val Gly
 340 345 350
 Leu Asp Val Asp Trp Pro Ser Thr Asp Asp Glu Asp Asp Asp Phe Asp
 355 360 365
 Pro Glu Gly Ile Arg Lys Gln Ala Asn Ser Asp Glu Glu Ser Ser Ser
 370 375 380
 Asp Glu Ser Glu Glu Asp Gly Glu Asp Ser Gly Ser Val Ser Gly Phe
 385 390 395 400
 Arg Glu Asn Glu Ala Ser Ser Gly Val Ser Asp Lys Asp Glu Gln Glu
 405 410 415
 Pro Ser Glu Lys His Asn Leu Glu Leu Thr Ala Val Thr Asn Ile Lys
 420 425 430
 Lys Leu Ala Arg Lys Gly Lys Arg Asn Ile Ile Ser Asn Lys Gly Lys
 435 440 445
 Ser Glu Tyr Leu Glu Ser Ser Asp Ser Asp Asp Ser Asp Val Ser Leu
 450 455 460
 Phe Asp Glu Phe Lys Gly Gly Asn Ala Asn Leu Gly Leu Arg Arg Gly
 465 470 475 480

ES 2 558 745 T3

Asp Pro Gln Ala Ala Val Ser Ala Ala Asp Glu Glu Ala Val Ile Ile
 485 490 495
 Ala Gly Lys Arg His Arg Lys Ala Val Asp Tyr Lys Arg Leu His Asp
 500 505 510
 Glu Met Tyr Gly Lys Ala Glu Ala Asp Asn Asp Asp Val Ile Ser
 515 520 525
 Glu Asp Glu Asp Trp Gly Pro Glu Arg Arg Arg Arg Arg Thr Ile Pro
 530 535 540
 Asp Asp Pro Asn Ser Pro Arg Ser Arg Ile Arg Gly Arg Lys Val Arg
 545 550 555 560
 Thr Pro Ser Gly Asp Ile Ser Lys Gly Ala Asp Ala Asp Leu Pro Ser
 565 570 575
 Gly Leu Pro Asp Met Pro Ser His Phe Pro Thr Gly Ser Gln Gly Glu
 580 585 590
 Val Gly Ser Pro Gln Asp Ile Pro Ala Leu Asp Gly Ala Ala Asp Ser
 595 600 605
 Pro Gly Gly Glu Lys Arg Met Trp Arg Arg Leu Pro Asp Ser Ala Val
 610 615 620
 Glu Ala Leu Arg Cys Ile Leu Asp Val Asn Arg Leu Pro Ser Lys Ser
 625 630 635 640
 Arg Lys Glu Glu Leu Ala Ile Lys Leu Gly Leu Ser Phe Ser Gln Val
 645 650 655
 His Gly Trp Phe Lys Asn Gln Arg His Gln Ala Leu Arg Lys Gly Leu
 660 665 670
 Ala Cys Pro Lys Arg Ile His Pro Val Ala Ser Ser Ser Arg Asp Phe
 675 680 685
 Asn Thr Val Glu Ser Pro Ser Ser Val Ala Thr Leu Thr Glu Ser Thr
 690 695 700
 Leu Pro Gln Lys Glu Asn Ala Pro Asn Ser Ile Leu Thr Glu Lys Leu
 705 710 715 720
 Asp Glu Val Gln Val Arg Leu Leu Glu Leu Lys Gln Ala Leu Asp Asp
 725 730 735
 Phe Thr Lys Asn Ala Ala Val Val Ala Asp Gly Gly Ser Ile Glu Leu
 740 745 750
 Asp Asn Asn Val Glu Asn Gly Gly Ile Gly Leu Pro Glu Asn Gly Lys

ES 2 558 745 T3

755 760 765

Arg Val Val Tyr Val Pro Val Val Ala Asp Val Arg Glu Gln Ser Pro
 770 775 780

Met Ser Thr His Thr Ser Ala Cys
 785 790

5 <210> 242
 <211> 2166
 <212> ADN
 <213> *Zea mays*

<400> 242

atgcattcctt cggaaaataa actggggttg aatgagacct ctatgggggtt ggttcctgta 60

ccaaaaagac cagctagacc tgatgcttcc caccaatgca aatctgattc ctttatgaag 120

cgatcaccaa ggaaagttag aaatgctact ctggcaaaaa gcataaagag caaataccac 180

tatagtcccc tcaaacagcg gaagggttca gattctgttc ctgggaaaaat cgtaacagga 240

ctaaccgcaa ggaaaaagaa gaaaaggaaa atccaaatta cagacgaggc aactcgtttg 300

gaacgaagag cgagatattt tctaatacaag ataaaactgg agcagaattt gctagatgct 360

tactctggag atggatggaa tgggcaaagc cgagagaaaa taaagccaga gaaggaactg 420

caacgtgcca ggaaacaaat cataaaatgc aaaattgcta tacgtgatat catccgcaa 480

cttgatttgt atacttctac tgggagtgtt gatgaccac taatgcctac agatcagtcc 540

accaatcccc aacataccat gtgctcaaca tgcaagtctc atgaatcatt tcccagcaat 600

aaaattatct tctgcaaagg gccctgcaaa agggcatgtc atgagaaatg tttggaacct 660

cccttaaaca aaagcgtact tccaacaagt agccatgggt ggctttgcaa attctgtttg 720

tgcaagggtga ggattttaga aactattaat gcacatctag ggacgagttt tacagtgaag 780

tgccactttg aagatatatt caaggaaact actgactaa tagactctga ggatgacta 840

gatgaagatt ggctttctga gtattcaggt gatgaggact atgatcctga tgaaaatgag 900

gccagtggcg actgtatgga cagcggggag aagattatgt ctgatgattc caatgggtca 960

ggaagcccc tttattctcc aaatgacgat attcctgact tcatatcagc agacttaaat 1020

gttgtggaag ggttttgtca taccaattta gatttaggca ttgatgccgt tgaagatgat 1080

tttgcacaga tcctcaccta ccaaaggcca agaagagatg ttgattatag aaggctaac 1140

gaggaaatgt ttgggaaaat aaccgggaat gaagaacaga gcgaggacga agattggggc 1200

catgaaagga gaaagaaaag gacgcattca ggtgttgctg gggataattc tgttggttc 1260

ttgaacgtta tatctgatga aaagagccag aagaagggga gaaaactttt caggattcct 1320

cctgcagctg ttgaggtact tcgcagagct tttgctgaaa atgagcttcc accccgggat 1380

gttaaagaaa atctctcaag agaattggga atttcttttg aaaagattga taaatggttc 1440

aaaaatacac ggtgtgctgc tctcagagat cggaaggctg aaggaaacag tcataataca 1500

gctcccagca aatctcaaaa aaataaagga aaagctggaa tctcaggcaa ggctggaagg 1560

10

ES 2 558 745 T3

aatggtcatg ttactgggtcc ctgcaataat ttgagaataa atgaagaaaa aacaggtata 1620
 tcggggaagt ttgattcagg agacaactct tgtttggttc ctttctctga agtcatcaat 1680
 gtgcctacgc gattgccaca caactttgag atgaggaaga tggaatcaac tagaagtcct 1740
 gcgaggctac acaataaagg agggtgacctg tctgcaacaa tccaaattaa ggagagtaca 1800
 ttgttaccca caggcaagcc ttgttggcag tcagaaatga gccatccaac aactaatgaa 1860
 gtgagcactt cagcgcaagc tactacttgg atcgacgcgg gggcgtgtgc cgaggagcaa 1920
 gaacctactc cttgggtgga catcggggcc tcagactacc agcctttcct ggatgtgatc 1980
 gacgacatgt gcggacttga gtggaggctg cagagactga aggagaacat gctctcatct 2040
 agcatagacg gccaaaccgc cgccagtgag agtgacaaaa gaaaccagac cgtggtgctc 2100
 gtgccaaactg ccgagctcaa ggaaaaactg ccgcatgacg gtttattcgg gcattattgc 2160
 ccatag 2166

<210> 243
 <211> 721
 <212> PRT
 <213> *Zea mays*
 <400> 243

5

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

10

ES 2 558 745 T3

Leu Asp Leu Tyr Thr Ser Thr Gly Ser Val Asp Asp Pro Leu Met Pro
 165 170 175
 Thr Asp Gln Ser Thr Asn Pro Glu His Thr Met Cys Ser Thr Cys Lys
 180 185 190
 Ser His Glu Ser Phe Pro Ser Asn Lys Ile Ile Phe Cys Lys Gly Pro
 195 200 205
 Cys Lys Arg Ala Cys His Glu Lys Cys Leu Glu Pro Pro Leu Asn Lys
 210 220
 Ser Val Leu Pro Thr Ser Ser His Gly Trp Leu Cys Lys Phe Cys Leu
 225 230 235 240
 Cys Lys Val Arg Ile Leu Glu Thr Ile Asn Ala His Leu Gly Thr Ser
 245 250 255
 Phe Thr Val Lys Cys His Phe Glu Asp Ile Phe Lys Glu Thr Thr Glu
 260 265 270
 Leu Ile Asp Ser Glu Asp Ala Leu Asp Glu Asp Trp Leu Ser Glu Tyr
 275 280 285
 Ser Gly Asp Glu Asp Tyr Asp Pro Asp Glu Asn Glu Ala Ser Gly Asp
 290 295 300
 Cys Met Asp Ser Gly Glu Lys Ile Met Ser Asp Asp Ser Asn Gly Ser
 305 310 315 320
 Gly Ser Pro Leu Tyr Ser Pro Asn Asp Asp Ile Pro Asp Phe Ile Ser
 325 330 335
 Ala Asp Leu Asn Val Val Glu Gly Phe Cys His Thr Asn Leu Asp Leu
 340 345 350
 Gly Ile Asp Ala Val Glu Asp Asp Phe Ala Gln Ile Leu Thr Tyr Gln
 355 360 365
 Arg Pro Arg Arg Asp Val Asp Tyr Arg Arg Leu Asn Glu Glu Met Phe
 370 375 380
 Gly Lys Ile Thr Gly Asn Glu Glu Gln Ser Glu Asp Glu Asp Trp Gly
 385 390 395 400
 His Glu Arg Arg Lys Lys Arg Thr His Ser Gly Val Ala Gly Asp Asn
 405 410 415
 Ser Val Gly Phe Leu Asn Val Ile Ser Asp Glu Lys Ser Gln Lys Lys
 420 425 430

ES 2 558 745 T3

Gly Arg Lys₄₃₅ Leu Phe Arg Ile Pro₄₄₀ Pro Ala Ala Val Glu₄₄₅ Val Leu Arg
 Arg Ala Phe Ala Glu Asn Glu₄₅₅ Leu Pro Pro Arg Asp₄₆₀ Val Lys Glu Asn
 Leu Ser Arg Glu Leu Gly₄₇₀ Ile Ser Phe Glu Lys₄₇₅ Ile Asp Lys Trp Phe₄₈₀
 Lys Asn Thr Arg Cys₄₈₅ Ala Ala Leu Arg Asp₄₉₀ Arg Lys Ala Glu Gly₄₉₅ Asn
 Ser His Asn Thr₅₀₀ Ala Pro Ser Lys Ile₅₀₅ Ser Lys Asn Lys Gly₅₁₀ Lys Ala
 Gly Ile Ser₅₁₅ Gly Lys Ala Gly Arg₅₂₀ Asn Gly His Val Thr₅₂₅ Gly Pro Cys
 Asn Asn₅₃₀ Leu Arg Ile Asn Glu₅₃₅ Glu Lys Thr Gly Ile₅₄₀ Ser Gly Lys Phe
 Asp Ser Gly Asp Asn Ser₅₅₀ Cys Leu Val Pro Phe₅₅₅ Ser Glu Val Ile Asn₅₆₀
 Val Pro Thr Arg Leu₅₆₅ Pro His Asn Phe Glu₅₇₀ Met Arg Lys Met Glu₅₇₅ Ser
 Thr Arg Ser Pro₅₈₀ Ala Arg Leu His Asn₅₈₅ Lys Gly Gly Cys Leu₅₉₀ Ser Ala
 Thr Ile Gln Ile Lys Glu Ser Thr₆₀₀ Leu Leu Pro Thr Gly₆₀₅ Lys Pro Cys
 Trp Gln Ser Glu Met Ser His₆₁₅ Pro Thr Thr Asn Glu₆₂₀ Val Ser Thr Ser
 Ala Gln Ala Thr Thr Trp Ile Asp Ala Gly Ala Cys Ala Glu Glu Gln₆₄₀
 Glu Pro Thr Pro Trp₆₄₅ Val Asp Ile Gly Ala Ser Asp Tyr Gln Pro Phe₆₅₅
 Leu Asp Val Ile₆₆₀ Asp Asp Met Cys Gly₆₆₅ Leu Glu Trp Arg Leu Gln Arg
 Leu Lys Glu₆₇₅ Asn Met Leu Ser Ser₆₈₀ Ser Ile Asp Gly Gln₆₈₅ Thr Ala Ala
 Ser Glu Ser Asp Lys Arg Asn Gln Thr Val Val Leu₇₀₀ Val Pro Thr Ala
 Glu Leu Lys Glu Lys Leu Pro His Asp Gly Leu Phe Gly His Tyr Cys

ES 2 558 745 T3

705

710

715

720

Pro

REIVINDICACIONES

1. Un procedimiento para potenciar rasgos relacionados con el rendimiento de las semillas, en plantas con respecto a plantas de control, que comprende aumentar, la expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido dedo de homeodominio de planta-homeodominio (PHDf-HD), cuyo polipéptido PHDf-HD comprende:
- 5
- (i) un dominio que tiene al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más de identidad de secuencia de aminoácidos con un dominio cremallera de leucina/dedo de homeodominio de planta (ZIP/PHDf) como se representa en SEC ID N°: 233; y
- 10 (ii) un dominio que tiene al menos el 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más de identidad de secuencia de aminoácidos con un homeodominio (HD) como se representa en SEC ID N°: 234; y
- 15 (iii) en el que cuando dicho polipéptido PHDf-HD se utiliza en la construcción de un árbol filogenético de HD, tal como el ilustrado en la Figura 1, se agrupa con el grupo de polipéptidos PHDf-HD que comprende la secuencia de polipéptido como se representa en SEC ID N°: 180, en lugar de con cualquier otro grupo HD; y/o en el que dicho polipéptido PHDf-HD tiene en orden creciente de preferencia al menos el 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 %, 99 % o más de identidad de secuencia con el polipéptido PHDf-HD como se representa en SEC ID N°: 180 o con cualquiera de las secuencias de polipéptido proporcionadas en la Tabla D1 en la presente memoria, y opcionalmente la selección de plantas que tienen rasgos relacionados con el rendimiento potenciados.
- 20 2. Procedimiento de acuerdo con la reivindicación 1, en el que dicho polipéptido PHDf-HD comprende: (i) un dominio PHD como se representa en PFAM00628; y (ii) un HD como se representa en PFAM00046.
3. Procedimiento de acuerdo con la reivindicación 1 o 2, en el que dicha secuencia de ácido nucleico que codifica un polipéptido PHDf-HD está representada en una cualquiera de las SEC ID N° de secuencia de ácido nucleico proporcionadas en la Tabla D1 o una porción de las mismas, o una secuencia capaz de hibridar con una cualquiera de las SEC ID N° de secuencias de ácido nucleico proporcionadas en la Tabla D1; y/o en el que dicha secuencia de ácido nucleico codifica un ortólogo o parálogo de cualquiera de las SEC ID N° proporcionadas en la Tabla D1.
- 25 4. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que dicha expresión aumentada se efectúa mediante uno cualquiera o más de: etiquetado en activación de ADN-T, TILLING, o recombinación homóloga.
- 30 5. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que dicha expresión aumentada se efectúa mediante la introducción y expresión en una planta de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD.
6. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que dichos rasgos relacionados con el rendimiento de las semillas, comprenden uno o más de: (i) número de panículas primarias aumentado; (ii) peso total de semillas por planta aumentado; (iii) número de semillas (llenas) aumentado; (iv) PMG aumentado; o (v) índice de cosecha aumentado.
- 35 7. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que dicha secuencia de ácido nucleico está unida operativamente a un promotor constitutivo, preferentemente a un promotor constitutivo de planta, más preferentemente a un promotor GOS2, muy preferentemente a un promotor GOS2 de arroz; y/o en el que dicha secuencia de ácido nucleico que codifica un polipéptido PHDf-HD es de origen vegetal, preferentemente de una planta monocotiledónea, además preferentemente de la familia *Poaceae* más preferentemente del género *Oryza*, muy preferentemente de *Oryza sativa*.
- 40 8. Plantas que tienen rasgos relacionados con el rendimiento de las semillas potenciados o partes de las mismas (incluyendo semillas), o células de planta obtenibles mediante un procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en las que dicha planta, parte o célula de la misma comprende un transgén de ácido nucleico aislado que codifica un polipéptido PHDf-HD unido operativamente a un promotor GOS2.
- 45 9. Construcción que comprende:
- (i) Una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define en una cualquiera de las reivindicaciones 1 a 3;
- 50 (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 terminadora de la transcripción.
- en las que al menos una de las secuencias de control es un promotor GOS2.
- 55 10. Uso de una construcción de acuerdo con la reivindicación 9 en un procedimiento para la producción de plantas que tienen rasgos relacionados con el rendimiento de las semillas potenciados, preferentemente uno o más de: (i) número de panículas primarias aumentado; (ii) peso total de semillas por planta aumentado; (iii) número de semillas

- (llenas) aumentado; (iv) PMG aumentado; o (v) índice de cosecha aumentado; o el uso de una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define en una cualquiera de las reivindicaciones 1 a 4 en potenciar rasgos relacionados con el rendimiento de las semillas, que comprenden uno o más de: (i) número de panículas primarias aumentado; (ii) peso total de semillas por planta aumentado; (iii) número de semillas (llenas) aumentado; (iv) PMG aumentado; o (v) índice de cosecha aumentado.
- 5
11. Planta que tiene rasgos relacionados con el rendimiento de las semillas potenciados con respecto a las plantas de control o parte de planta o célula de planta de la misma transformada con una construcción de acuerdo con la reivindicación 9.
- 10 12. Procedimiento para la producción de plantas transgénicas que tienen rasgos relacionados con el rendimiento de las semillas potenciados con respecto a las plantas de control, que comprende:
- (i) introducir y expresar en una planta, parte de planta, o célula de planta, una secuencia de ácido nucleico que codifica un polipéptido PHDf-HD como se define en una cualquiera de las reivindicaciones 1 a 3, bajo el control de un promotor constitutivo de planta; y
 - (ii) cultivar la célula de planta en condiciones que promueven el crecimiento y el desarrollo de la planta.
- 15 13. Planta transgénica que tiene rasgos relacionados con el rendimiento de las semillas potenciados con respecto a las plantas de control, que se produce por la expresión aumentada de un transgén de ácido nucleico que codifica un polipéptido PHDf-HD como se define en una cualquiera de las reivindicaciones 1 a 3, unido operativamente a un promotor GOS2, o una célula de planta transgénica obtenida de dicha planta transgénica que comprende dicho transgén de ácido nucleico; o partes cosechables de dicha planta que comprenden dicho transgén de ácido nucleico
- 20 que codifica un polipéptido PHDf-HD, en la que dichas partes cosechables son preferentemente semillas.
14. Planta transgénica de acuerdo con la reivindicación 8, 11 o 13, en la que dicha planta es una planta de cultivo o una monocotiledónea o un cereal, tal como arroz, maíz, trigo, cebada, mijo, centeno, triticale, sorgo y avena, o una célula de planta transgénica obtenida de dicha planta transgénica.

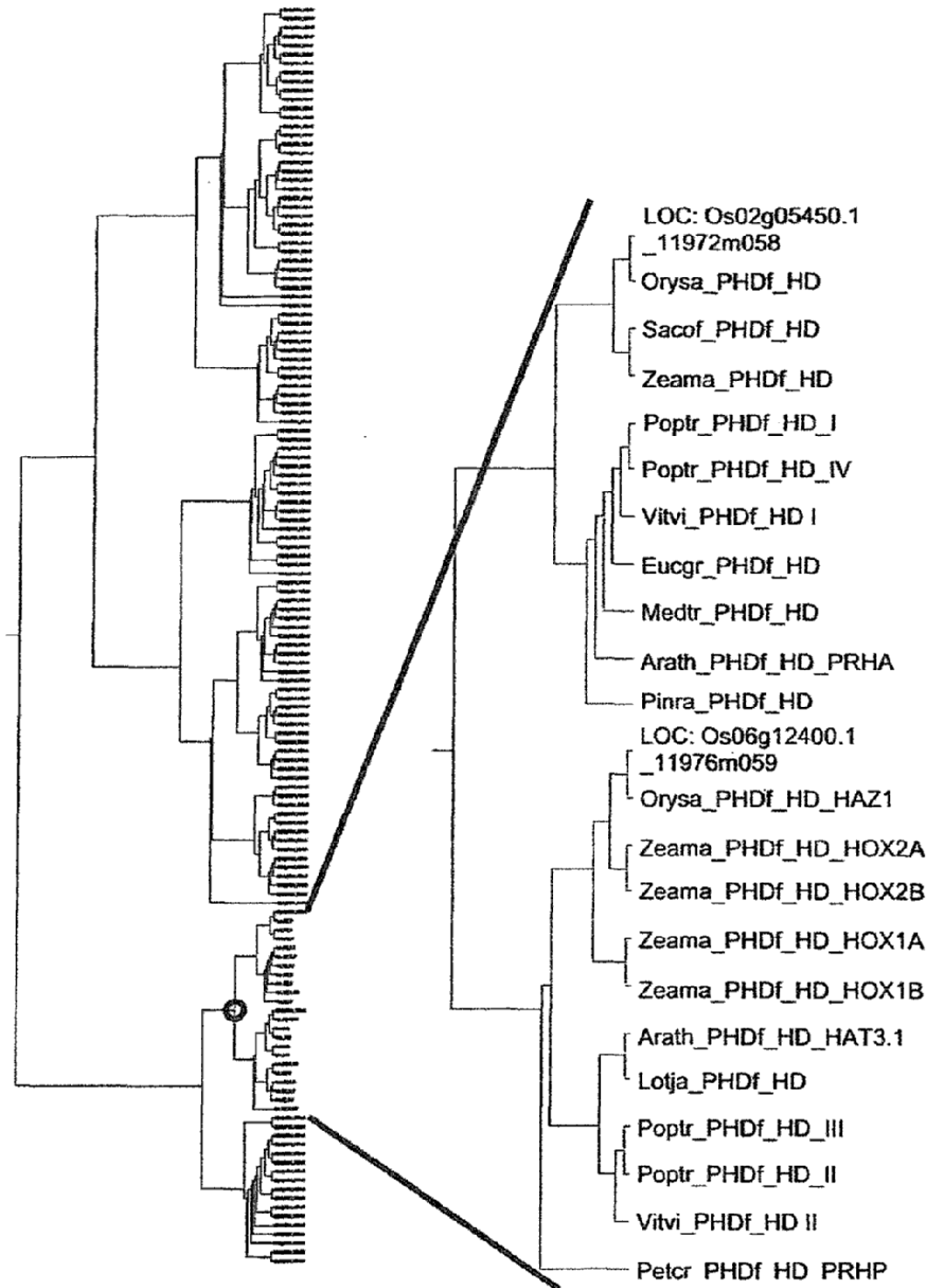


FIGURA 1

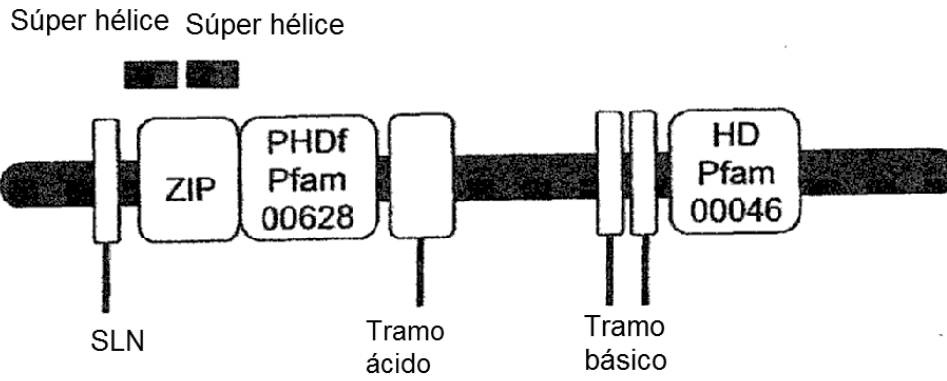


FIGURA 2



FIGURA 3

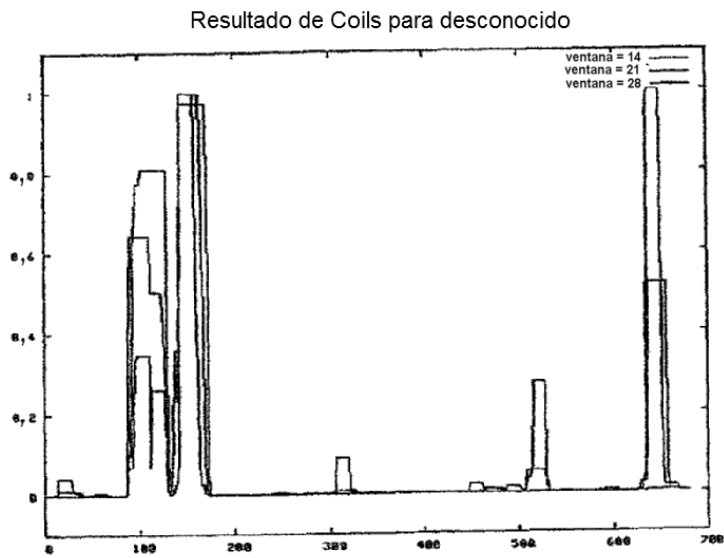


FIGURA 4

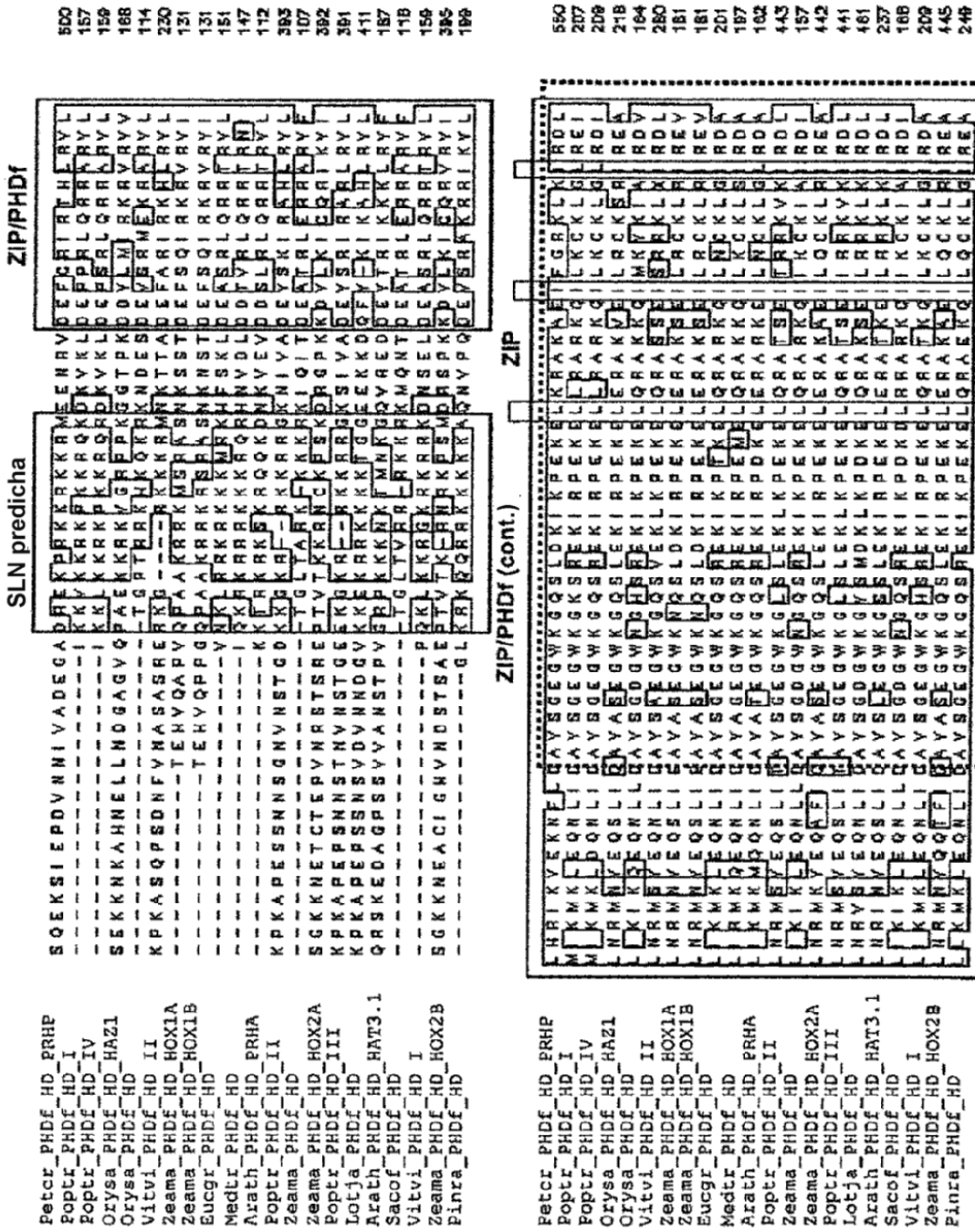


FIGURA 5

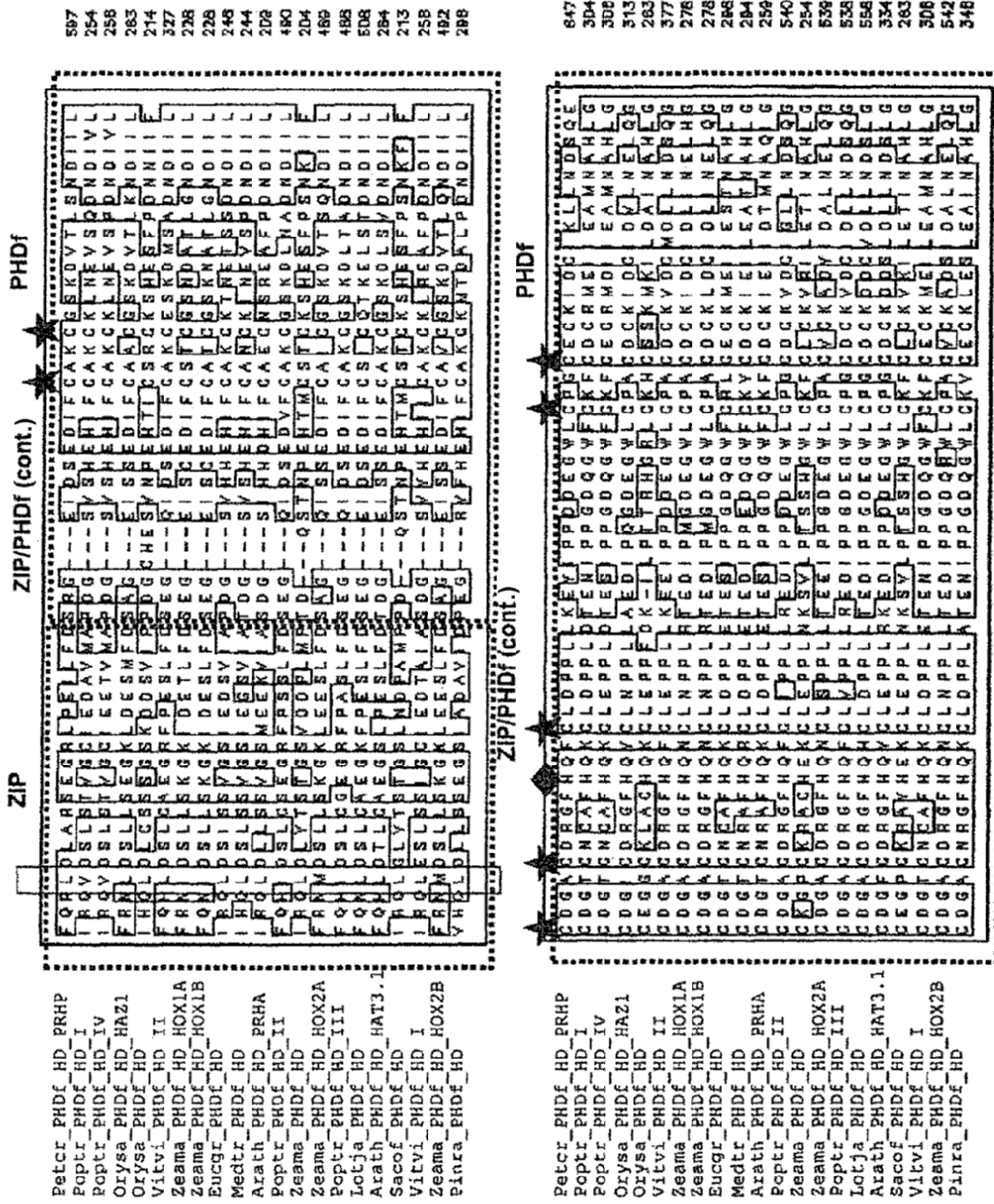


FIGURA 5

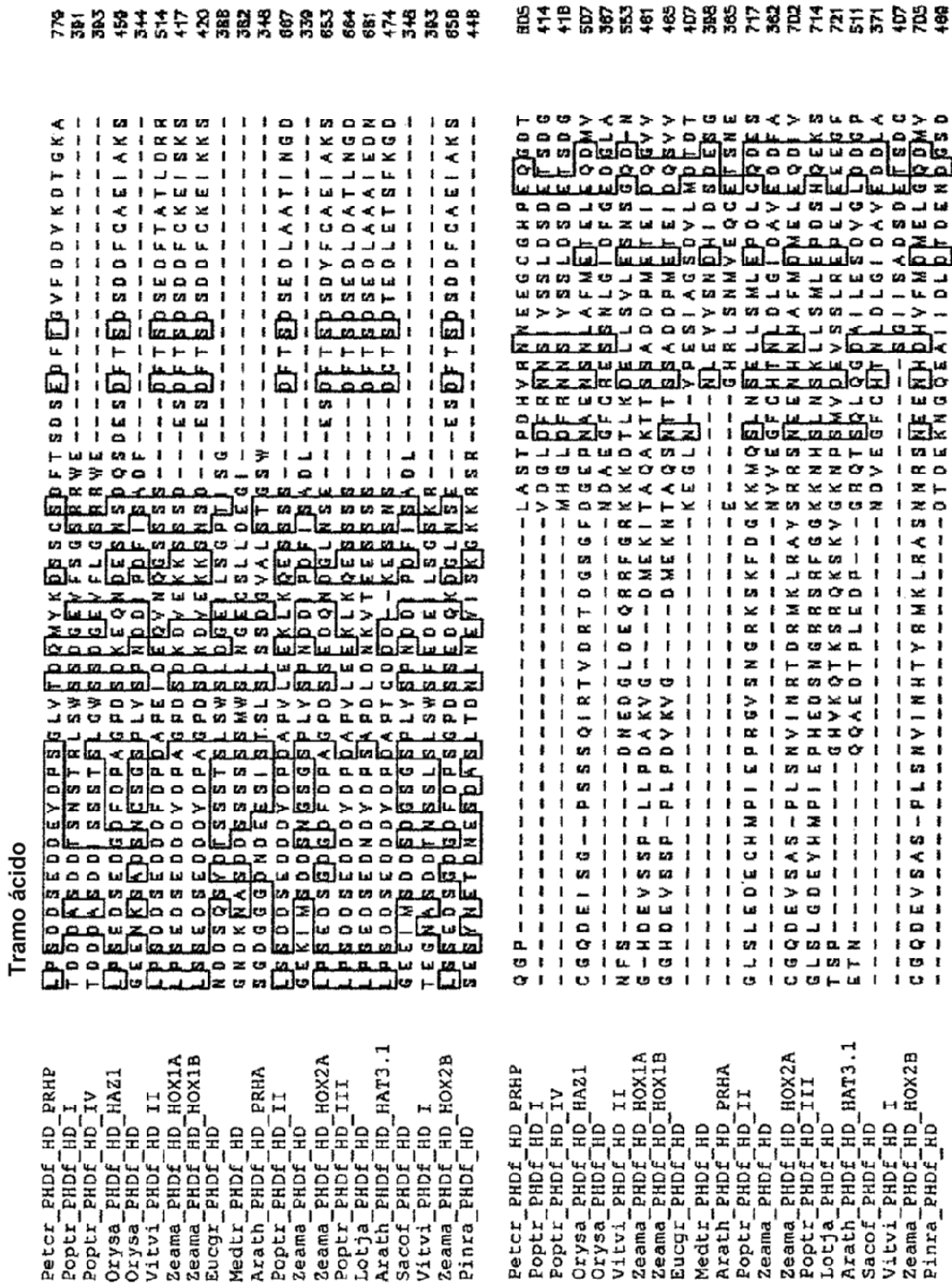


FIGURA 5



FIGURA 5

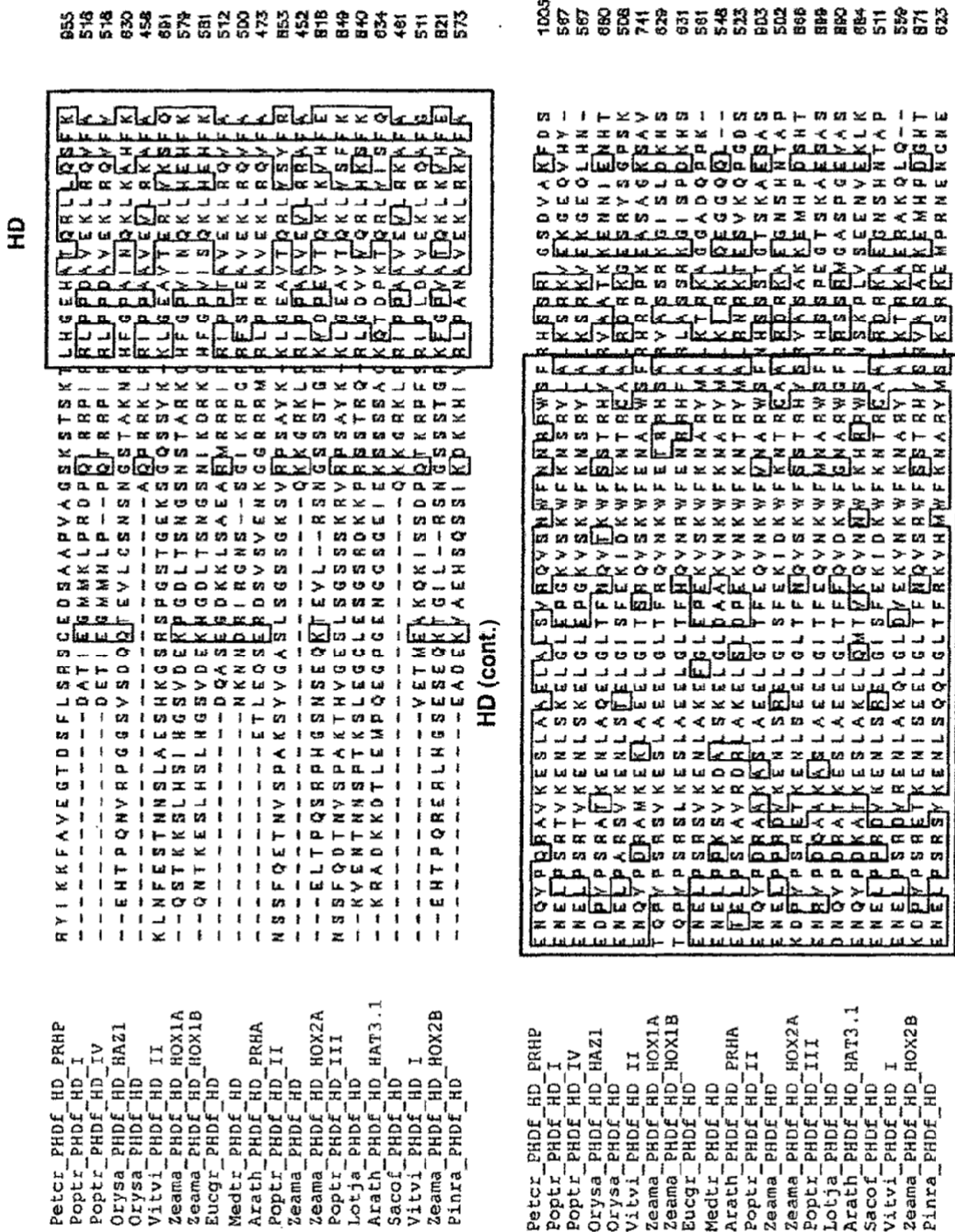


FIGURA 5

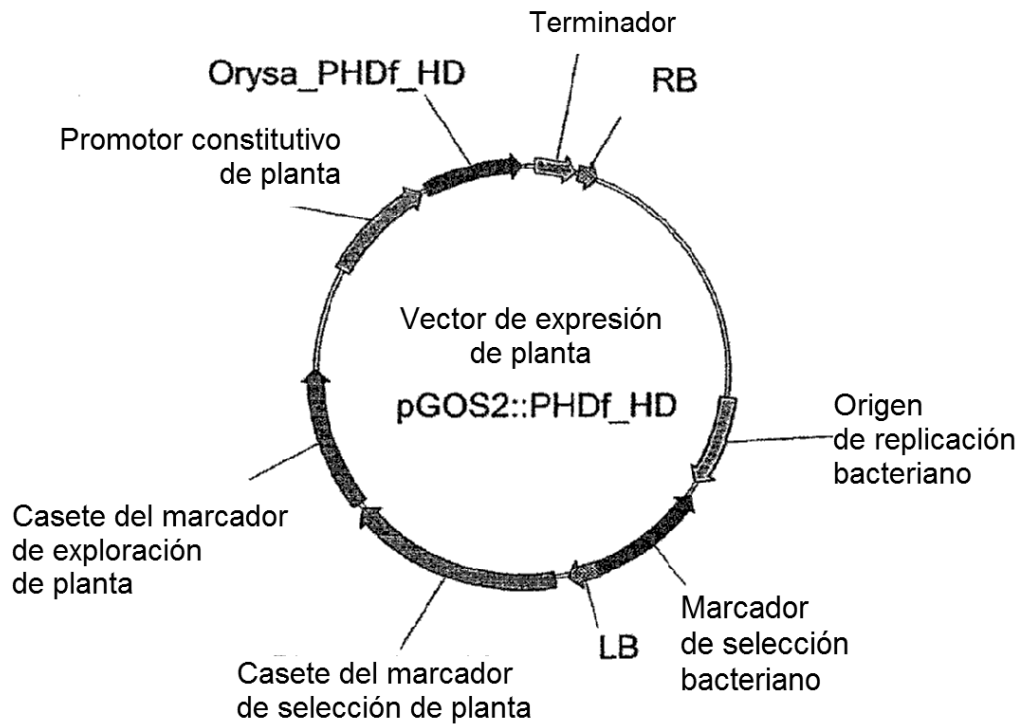


FIGURA 6

SEC ID N°: 179, secuencia de ácido nucleico PHfD_HD de *Oryza sativa* Os02g0147800 NM_001052422.1
 ATGAATACCCCAGAAAAGAAACCACTGTGCTATACTTCTAGAAGAGCACTGCAACAGAGGACAGAA
 TCCTCTTCGGAATTGATCTCTGTATCAAAAAGAGCAACCCGCCAGAATACTCCACGCAAGCCGGAT
 TCTCTCCCAAACGAACAACAAGAAGTAGCGCTAACCTGGCGAAATGTATAGAGAACAAACACCAT
 AGCAGTCCTCTGAAGCGGCGGGCGGGGTTTCCAGATGCGGGTACC GGAAAGTCTGCAACAGGGCCAACA
 AGGAGGAAGCATAAACAGAAAAGGAAAAATGATGAGAGTGATGAGGTCAGTCGCATGGAAAAAAGA
 GCAAGATATTTACTTATCAAAAATAAACAGGAGCAGAATTTGCTAGATGCATACTCTGGAGATGGT
 TGGAAATGGGCACAGTCGAGAGAAAATAAAGCCAGAGAAGGAGCTACAGCGTGCTAAGAAACAGATC
 ATGAAATACAAAATCGCTATCCGTGACGTCATTCACCAACTTGATTTGTGTAGTTCCAGTGGGAGC
 AAGGATGACTCTGTGATTCCACCAGATGGGTGTCACGAGTCTGTCAATCCTGAACATACAATATGT
 TCAAGGTGCAAAATCCCACGAGTCATTTCTGACAATAACATTATATTCTGTGAGGGGGGGCTGCAAA
 CTGGCGTGTATCAGAAATGTTGGAGCCTCCTTTTGATAAAAATCTTCCAAC TACCCGCCATGGG
 CGGCTTTGCAAACTGTTCTTCCAAGATGAAAATTTTAGACGCCATTAATGCCCATCTGGGGACA
 AGTTTTACAGTGAAGTGTCCCTCCAGTGATATTTTCAAGGAAGCAGCTGAACATTTCAACTCTGAT
 GATGGACTGGGCCAAGATTGGCTATCTGAGTATTCAGGTGATGAGGACTATGACCCTGAAGAAAAT
 GAGGCCAGCAGCAGTGGTGAAGAGAATAAATCTGCCGATTCCAA TTGTTGAGGGAGTCCACTTTAT
 TCTCAAATGATGATATTCCAGACTTCATATCAGCAGATTTCAATGATGCTGAAGGATTCTGTTCGT
 GAAAGTTCAAATCTAGGAATCGATTTTGGTGAAGATGGTTTGGCTGAGATTCTTACCCACCAAAGG
 CCAAGAAGAGATGTTGACTATACACA ACTTAATGAGCAAATGTTTGGAGAGCCTATTGGCAATGAC
 GAACAGAGTGAAGATGAAGATTGGGGTCTTAACAAGAGAAAAGAAAAGAAGAACTGGTTCAACTGGT
 GTTGGGACTAATTCCTGAGAGGGTGCATCAGATGTCAAATCCAA TAAGAAAGCCCAACCTCGGAGA
 AAACTTTTAGGATTCTCTCTGCGGCAGTTGAGGTGCTCCGTAAAGCTTTTGTGAAAATGAGCTT
 CCTGCCCGGAGTGTCAAAGAAAATCTCTCAACAGAGCTGGGTATTTCTTTT GAAAAGATTGATAAG
 TGGTTCAAAAATACACGGTGTGCTGCTCTCAGGGATAGGAAGGGT GAAAGTCGTTATTCTGGTCCC
 AGCAAAGGTCAAGAACAAGCATAGAAAAGGCTGAAACTTCAGCTAAAGTGGATCAAATGGACAAC
 TCATGCTTTCTTCCCTTATCTGAAATAATCAACGTGCCACACGCTGCAGAAAAGGTCTTGATAAG
 AAGCCAAAATCAATTAACAGCCCTCCAAGACCTCAGGACAATGAAACTTGCTTGTGCGCCAAC TGAT
 AAAACCAAGGAGGGTACACCGCCTACGATCAAGCCTTCATCACAGATTCTAGTCAACTGATGAAT
 AACGACATCGGCACTGAGGAGACTGCTGTTTCTTGGGTGGACACTTGGGCCTCTGACGCTCTGCAT
 TTCTTGGACGTGAGCGATGATGAACATTTCTTCGATGTGATCGAGAAGGTGTGCGGTCTCGAGAAT
 CGGCTGCAGCGGTTGAAGGAGAACA TGCTATCATCGTCATCGTCAACTGACAACAATGTGGCTGCT
 GAGAGTGGCTTGCAAAACGAAGTTGTGCTTGTACCAGCTGCTGAGCTCAAGGATAAAGCATCTTAA

SEC ID N°: 180, secuencia de polipéptido traducida PHDf_HD de *Oryza sativa*
 MNTPEKKPLCYTSRRALQQRTESSSELI SVSKRATRQNTPRKPDSPPKRTTRSSANLAKCIENKHH
 SSPLKRRRGS DAATGKSATGPTRRKHKQKRKNDESDEVSRMEKRARYLLIKIKQEONLLDAYSGDG
 WNGHSREKIKPEKELQRAKKQIMKYKIAIRDVIHQLDLCS SSGSKDDSVIPPDGCHESVNPEHTIC
 SRCKSHESFPDNNIIFCEGGCKLACHQKCLEPPFDKILP TTRHGRLCKHCSSKMKILDAINAHLGT
 SFTVKCPSSDIFKEAAEHFNSDDGLGQDWLSEYSGDEDYDPEENEASSSGEENKSADSNCSPGLY
 SPND DIPDFISADFNDAEGFCRESSNLGIDFGEDGLAEILTHQRPRRDVDY TQLNEQMFGEPIGND
 EQSEDEDWGLNKRKRRTGSTGVGTNSVEGRSDVKS NKKAQPRRKLFRIPPAAVEVLRKAFAENEL
 PARSVKENLSTELGISFEKIDKWFKNTRCAALRDRKGESRYSGPSKR SRTSIEKAETS AKVDQMDN
 SCFLPLSEI INVPTRLQKGLDKKPKSINSPPRPQDN ETCLSPDKTKEGTPPTIKPSITDSSQLMN
 NDIGTEETA VSWVDTWASDALHFLDVS DDEHFFDVIEKVCLENRLQRLKENMLSSSSSTDNNVAA
 ESGLQNEVVLPAAELKDKAS

FIGURA 7

SEC ID N°: 181, secuencia de ácido nucleico de PHDf_HD_PRHA de *Arabidopsis thaliana* NM_119140

ATGGAGGAGAGTGAACAAAGGGGAGAATCAGTCAAGAACTGATAAGGCTTGTGTTTCAGTTGAG
 AGAATTGGGTCTACTTTACTCTCATCATTTGTGAAGAAAGGTAAAGAAGTATCAAAATAAGAGAAAT
 TCTAAGCAGAATAAGCGAAAAGCGGAAGAAGAACTCTGTTTCGAAATCAAGAACCAAGAAGTACTCT
 CGGGGATGGGTCCGTTGTGAGGAAAATGGAGGAGGAGAAGGTTAAGAAAAACAAGAAAAGAGGAAAAGT
 AAGAGACAGCAAAGGATAATAAGGTTGAGGTTGATGATTCTTTAAGGTTGCAAAGGAGGACAAGG
 TATTTGCTCATTAAGATGAAGATGCAGCAGAATCTTATCGATGCATACGCAACTGAAGGTTGGAAA
 GGTGAGAGCCGGGAAAAGATAAGACCAGACAAGGAGCTTGAGAGAGCCCGGAAAGAGATCTTAAAC
 TGCAAGCTTGGACTACGCGACGCCATTCGCCAACTGGATCTGCTTAGTTCTGTGGGAAGCATGGAA
 GAGAAAGTGATAGCTTCAGATGGATCTATTATCATGATCATATATATTTTGTGCAGAATGCAATTTCT
 CGAGAAGCTTTTCCAGCAATGATATAACTTTTGTGATGGAACGTGTAATCGAGCATTTCCACCAG
 AAATGCCCTTGACCTCCTTTGGAAACAGAAAAGTACCTCCCGGAGATCAGGGCTGGTTTTGCAAA
 TTTTGTGATGCAAAAATTGAAATTATCGACACAATGAATGCACAAAATTGGAACCTCATTTCCCTGTG
 GACAGCAACTGGCAGGATATCTTCAATGAAGAAGCTAGTCTTCCCTATTGGATCTGAAGCTACAGTT
 AACAAATGAAGCAGATTGGCCTTCGGATGACTCCAAGGATGATGATTATGACCCTGAAATGAGGGAA
 AACGGTGGAGGAAACAGCAGCAACGTTAGTGGTGTGTTGGTGGTGGTGGATAACGATGAAGAAAGCATT
 TCAACTAGCTTGAGCTTATCATCTGATGGTGTGCCCCTTCAACGGGATCATGGGAAGGCCATAGA
 CTCAGCAATATGGTGGAGCAGTGCGAAAACAAGTAATGAAGAACTGTATGTGGGCCAAGACAGCGA
 AGGACTGTTGACTATACACAGTTGTACTATGAAATGTTTGGAAAGGATGCAGTTTTGCAAGAGCAA
 GGTAGTGAAGATGAAGACTGGGGTCCGAATGACAGAAGGAAGAGAAAAAGGGAATCTGATGCAGGA
 AGCACACTGTAACATATGTGTGAAAGCAGTAAGAAAGATCAAGATGTTGTAGAAAACGCTAGAACAG
 AGTGAGAGAGATTCTGTCTCTGTGGAAAATAAAGGAGGTGCAAGGCGGATGTTTCAGACTCCCAAGA
 AATGCAGTTGAGAAGCTACGTCAAGTATTTGACAGAGACTGAGCTTCCCTCAAAAGCTGTGAGGGAT
 CGTCTTGCAAAAAGAGCTGAGTCTTGATCCAGAGAAGGTCAATAAATGGTTTTAAAAATACTCGGTAT
 ATGGCACTGCCGAATAGGAAGACTGAGAGCGTGAAACAACCTGGGGATTCTAAGACAGTCTCTGGA
 GGAGATTCTGGACCAGAAGCAGTCATGGAGAACAACCTGAAACAAATGAAGTTCAAGATACTTTG
 GATGACACTGTTCCACCCGGATTTGACGCTACAAATCAAAACATACTCTCCCCATGTAATAACAAC
 CAGGAGGAATTTCAACAGGAAAATGTCTCTTTCCCTTACCCACAGATGAAAGTCAACAGTATCTG
 GAACAAAACGATTTCTTTCTTTCTAGTGCCACATGAAAAGCAAAGCAGCGAAATTAGCTTAAAG
 ACGGCTGTAGAGGAGAATGAGACAGAAAGCAAGATGATGAAAGAGCCACATGAAGAGCTAAGCAGT
 GAAATGAGCCTAAAGACGGCTGCTGAGGAGAAGGAGACAGAGAGCAAGATGATAGAAGAGCCACAT
 GAAGAGCTAAGCAGAGAAATGAGCTTAAAGACGGCTGTAGAGGAGAAGGAGACAGAGAGCAAGATG
 ATGGAAGAGCCACATGATGAGCTAAACAGCGAAATGAGCTTAAGTACAGCGGTAGAGGAGAAGGAG
 ACAGGGAGCAAGATGACGGAAGAGTCACATGAAGAGCTAAGCAACGAAATGAGCTTAGAAGAGAAG
 GAGACGGGGAGAAAAGATGACGGAAGAGGAAGAATTGGAAGCAGTAATGGAGATGCTTTGTAGAACA
 GAAAACAAGTTACTGGACGTGACGCAGAGGCTTGATAGATTTAAAACACCAAAAGGCCGAAAAAAG
 TTAGGCAATTCTTCTTCTCCTTTACTTTGAAGAAGACTCAGTAGTGTATGTTCCAATAGCAGAGATC
 AAGGAGAAAAGGTAA

SEC ID N°: 182, secuencia de polipéptido traducida PHDf_HD_PRHA de *Arabidopsis thaliana*

MEESETKGRI SQETDKACVSVRIGSTLLSSFVKKGKEVSNKRNSKQNKRRKAEELCSKSRTPKYS
 RGWVRCEEMEEKVKTRKRKSKRQKDNKVEVDDSLRLQRRTRYLLIKMKMQQNLIDAYATEGWK
 QQSREKIRPDKELERARKEILNCKLGLRDAIRQLDLLSSVGSMEEKVIASDGSIHHDHIFCAECNS
 REAFPNDIILCDGTCNRAFHQKCLDPPLETESI PPGDQGWFCFKCDCKIEIIDTMNAQIGTHFPV
 DSNWQDIFNEEASLPIGSEATVNNEADWPSDDSKDDYDPEMRENGGNSNVSGDGGDNDEESI
 STLSLSLSSDGVALSTGSWEHRLSNMVEQCETSNEETVCGPRQRRTVDYTYLQYEMFGKDAVLQEQ

FIGURA 7 (continuación)

GSEDEDWGPNDRRKRKRES DAGSTLVTMCES SKKDQDVVETLEQSERDSVSVENKGGRRRMFRLPR
 NAVEKLRQVFAETELPSKAVRDLAKELSLDPEKVNKWFKNTRYMALRNKTESVKQPGDSKTVSG
 GDSGPEAVMENNTETNEVQDTLDDTVPPGFDA TNQNILSPC NNNQEEFQQENVSFPSPTDESQQYL
 EQNDSSFVLPHEKOSSEISLKTAVEENETESKMMKEPHEELSSEMSLKTAAEEKETESKMIEEPH
 EELSREMSLKTAVEEKETESKMMEEPHDELNSEMSLS TAVEEKETGSKMTEESHEELSNEMSLEEK
 ETGRKMTEEEELEAVMEMLCRTENKLLDVTQR LDRFKTPKGRK KLGNSSSPLLEEDSVVYVPIAEI
 KEKR

SEC ID Nº: 183, secuencia de ácido nucleico de PHDf_HD de *Eucalyptus grandis* ADW17964

ATGCATTCCAGAAAGAATCTTGCCCAAAAAGAAGACTTTTTGAAGGCAGAATCTGGATCGGAGGTT
 CTTGCCTCGCTGATATTGAGAAATAGGAGAAGGAAAAGCACAAACAGAAACTACAATCTTCTGT
 AAAGCAAGAGGCTCAACCCCTTGCCAAGAAGAGGGCCAATAACTCCATCAGAAAGAGACTGGTGCCT
 AGGAACAGCTCTGATAAAGACAAACTCATAAATTTAAAGCTTTTCATAATAAGAAAATTGAAATA
 GTTCCCGAGAAGCTTTCATCCCCTCCATCTTCAGAAAGAGAGCTGCCTATTGCTTCTGACGAAAAT
 GTGCAAAACACCGATGAGGATGTCAAAGTTAACAAGCTGAGAAGGAAGAGGAAGAAGAAAATGAGA
 AAGCATTTTTCGAAGTTAGATGAGGCTTCACGATTGCAGAGGAGAACAAGATACCTTCTTATAAAA
 ATGAAGCTGGAACAGAACCCTTATTGATGCTTACTCTGGTGAAGGCTGGAAAGGCCAGAGTCGGGAA
 AAGATCAAGCCAACAAGGAATTACAAAAGGGCCAGGAAGCAGATACTGAATTGCAAACTCGGAATA
 CGTGATGCTATTTCGACAGCTAGACTCCATTAGTTCACTGGGCAGCATTGAAGACTCTGTGATTGCT
 CCTGATGGTTCTGTTCACCATGAACATATATTCTGTGCAAAAATGCAAGACAAATGAAACTTCCTCA
 GATAATGATATCATACTCTGCGATGGGACATGCAATTGTGCTTTTCATCAAAAATGTCTTGATCCT
 CCACTGGAAACTGAAAGCATTCCCCCAGGAGATCAGGGCTGGTTTTGACAGACTCTGCGAGTGTAAG
 ATGGAAATTATAGAGTCAACAAATGCACATCTGGGGACCCACTTTTCTCTTGACAGCAACTGGCAG
 GATATATTTCGAGGAAGAAGCTGCATTGCCTGATGGGGGAATTCAGGTGCTTGACCTAGACGAAAAG
 TGGCCTTCAGATGACTCTGAAGATGATGACTATGACCCAGAAAAGATGACGAACACCGGCATCAAC
 TGTGCCCGCAATGATGATTCTCAATCATAACGACACAAGTAGCTCTACCAGCCTGAGTTGGTCATTA
 GATGGTGAATTTTTGTCAGGACCAACAATATCAGGCAAGGAAGGTTGAACTATCCTGAGAGCATT
 GCAGGGTCTGATGTATTAATGGACACGGATACTGATACTGTATGTGGCCCACGACAGCGAAATGCT
 GTTGATTATAAGCAGCTTTATAATGAAATGTTCCGGTAAGGATACTCCTGCCTGTGAACAAGCCAGT
 GAAGATGAAGACTGGGGTCTTAATAAAAAGAAGAAGAAGAGAAAAGGGAAACAGATGCAGCCAGCACC
 CTCATGACTCTCTGTGAAAGTGAGAAAAACAATCCATTGACCAAGCCTCAGAAGGTGATAAG
 AAACCTCTGCAGAAGCAAGAATGAGAAGGCGGATTTTCAGAATTCCTCCAACCTGCAGTTGAGAAG
 CTTCCGCAAGTATTTGCAGAGAATGAACTTCCCTCAAGATCTGTCAAAGAAAATCTTGCTAAAGAG
 TTTGGTCTTGAACCTGAGAAGGTTAGCAAATGGTTCAAGAATGCACGATACATGGCACTCAAAACA
 AGAAAGGCTGGGGGGGCTGATCAACCTCCAAAATCCCCGCAAAAGCATAAAGAATCTGAACTAGCA
 AAAATGATGAAAGGGGCTGCGCCTATTTTTATCTAAGGATGCCTCAACAGACATTGAATTCATTC
 TCTCTAAATACTGAGAAAGTCATCCAGCAAAAGAGTACGAAGCCAATACACGGGAAGGGTTACCT
 ATGGTACCTGAACATAGCTATATGGAGAATTTTGAGTTTCGATGACAAATATGTCCTTGAAACAAATA
 AAGCTATTGAAAGGAAAACCAAAGGACAGGAAGAGGGTTGACTTCGCAGCCAGTGAGGGATTACGA
 GAAAGAGAAGCTCAGTTGCGGGAGCTATGTAAAATTAGAAGTAGGCTAGATAGTCTGAGGAAGAGA
 ACCCTTACATTTCACTCGAGTGAATCCAAGGCTTCAGACGAATTCAACACTTGTAGATCACGTGTA
 GTATATGTTCCAGTAGCAGAAATAAAGGAGAGAGCTTGA

FIGURA 7 (continuación)

SEC ID N°: 184, secuencia de polipéptido traducida PHDf_HD de *Eucalyptus grandis*

MHSRKNLAQKEDFLKAESGSEVLASLILRNRRRKKKHKQKQLQSSVKARGSTLAKKRANNSIRKRLVR
 RNSDDKDKLINLKAFHNKKIEIVPEKLSPPSSERELPIASDENVQNTDEDVKVNLRRRKRKMMR
 KHFSKLEASRLQRRTRYLLIKMKLEQNLIDAYSSEGWKGQSREKIKPTKELQRARKQILNCKLGI
 RDAIRQLDSISSVGSIEDSVIAPDGSVHHEHIFCAKCKTNETSSDNDIILCDGTCNCAFHQKCLDP
 PLETESI PPGDQGWFCRLCECKMEI IESTNAHLGTHFSLDSNWQDI FEEEAALPDGGIQVLDLDEK
 WPSDDSEDDDYDPEKMTNTGINCAGNDDSQSYDTSSSTLSWSLDGEILSGPTISGKEGLNYPESI
 AGSDVLMDDTDTVCGPRQRNAVVDYKQLYNEMFGKDTPACEQASEDEDWGPNKRRRRERETDAAST
 LMTLCESEKNNPFI DQASEGDKKLSAEARMRRRIFRI PPTAVEKLRQVFAENELPSRSVKENLAKE
 FGLEPEKVSKWFKNARYMALKTRKAGGADQPPKSPQKHKESELAKMMKGAAPILSKDASTDIEFHS
 SLNTEKVIQKSTKPIHGKGSMPVPEHSYMEFEDDNMSLKQIKLLKPKDRKRVDFAASEGLR
 EREAQLRELCKIRSRLDSLKRKRLTFQSSSESKASDEFNTCRSRVVVYVPAEIKERA

SEC ID N°: 185, secuencia de ácido nucleico PHDf_HD de *Medicago truncatula* AC123547

ATGCGGGATACTGAGAGGTTAAATAACCAAGGATCTACAAAACCGAGTAACACAAAGGAACATATC
 AAGTTGAAGGTAGATTCCCCACAGTTCAAACAAGTAGCACAAAACACCAAGGAAGAAACACAGA
 CTGAAATCAAATCTCATAAACTCGGAGGTTGCACAAATGCGTCACTGAGGACAGCTACTGATTCT
 TCCAACAAGGCGTCAACCAAGGATTCTTCAAATAAAACGGATAGGAATTCAACACAGGTTGAGCCT
 TCAAAAAAGATACAAGGAGGAAAAACTTCTCTTAATAATGATAGGAAAGGTGAGAAAGATGTTGTT
 GATCAAGAGGGAAATATTCAAAAAACGTAAGAGAAGGAGAAAGAAGAAAAGACAAAGACATAATGTA
 GATCTTGATGATACTGTACGCCTCCAGAGAAGAACAAGGAATATCTTAATCAGAATGAAGCAAGAA
 CAGAATCTTATCGATGCTTACGCTGGAGAAGGTTGGAAAGGTCAGAGTCGAGAGAAGATTAGGCCA
 GAAATGGAGCTACAAAGAGCAAAAAAGCAGATTTTAAAATGCAAGCTTAGTATCCGTGATGCCATT
 CACCAACTGGATTCTCTTAGTTTCAAGTTGGTAGTATTGAAGGTTCTGTCTATTGCTACAGATGGATCT
 GTTCTCATGAACATATATTCTGTGCAAATTGCAAGATAAACGAAGTTTCTCCAGATAATGATATT
 ATACTCTGTGATGGCACATGCAATCGTGGCTTTCACCAGAGATGCTTGATCCTCCTTTGGAGACT
 GAAGACATTCCTCCTGAAGACCAAGGCTGGTTTTGCAAGTATTGTGACTGTAAGATAGAAATATTG
 GAGGCAACAAACGCGCATCTTGGGACTCGTTTTCCCTTAGACAGTACTTGGCAGGATGTATTCAAG
 GAAGAAGCTGCTATTCTGACGGCGATGCTGCATTACTAAATCAAGAAGAAGAAATGGCCGTCAGAT
 GATCCTGAAGACGATGACTATAATCCAGAGAGGAAAGAAGAAAGCCACGTTGGCTTCAACACAGAA
 GGAAATGATAAAAACGCATCTGATGATTCAGCAGTTCTTCCAGCATGTGGTCTCTAAATGGAGAA
 TGCTCTCTGTTAGATGAAGGCATCAACCTTGAATATTATTCTAACGACCACATAGATTCTGATGAA
 TCCGGGGAAATAGCATGCGGTCGTAGGCAGCGTAGAGCTGTTGACTACAAGAACTTTATGATGAA
 ATGTTTGGAAAGGATGCTCCACCTTGTGAACAAGTGAGTGAAGATGAAGATTGGGGTCTCGAAAA
 AGACGGCGAAGAGAAAAGGAGTCTGAGGCTGTTAATACTCTTATGACTCTACATGAGAGTGAAAAAT
 AAATATCCGAATAATAAGAACAATGATAGAAATAAGAGGGAATTCTTCAGGCATAAAAAGACCTTGT
 TTTAGGTTTTCACATGAGGCAGTTGAGAAGCTTCGCCAAGTTTTTGCAGAGAACGAGCTTCCTCCA
 AAATCCGTGAAGGATGCCCTTTCAAAGAATTGGGACTTGATGCTGCAAAGGTTAACAAATGGTTC
 AAAAATGCACGTTACATGGCACTTAAAATAAGAAAGCTTCAAGAAGGAGGACAACAGCTTCAGAGT
 ATCACTTCTAAGACCTCAAAGGATTCTACATCACAAACATGTGCAGGAAGATGAGGTTTTGAACCCG
 AAGTCCGCGAAAATTACTGTGATTAGTTCCCTGAAGAAGTGTGAAAATGTTACTGGCAAAGAAAAA
 ACAAAGCATCTAGCAATCCTTTGAAGAAAAAGCGACCAGAAATACTCCATGCTTTGGGAGAAAAA
 GGCAACAAGGATTCATGGAGGTCAGTGATGTTAGCTTAATGAAACTATTACAAGACAGGAAA
 AAGAAGGTAAGTTTTGCATTTGAAGGAGACTCTGAGGCAGCAGAGTTGGAATTTGAAAGACTCAGC
 AAAGTGAAGACGAAAATAGATAGATTGAAGCAAAGATTAAGTGGAGTTCAAATTTGCAGATCAAAG
 GGTTCAGATGAAGTTCAATTTGAATGAGCCATCTATTATGTATGTTTCTATAGCTGAGTTAAGGGAA
 AAGGTGAAATAG

FIGURA 7 (continuación)

SEC ID Nº: 186, secuencia de polipéptido traducida PHDf_HD de *Medicago truncatula*

MRDTERLNNQGSTKPSNTKEHIKLVKVDSPQFKTSSTKHQRKKHRLKSKSHKLGCTNASVRTATDS
 SNKASTKDS SNKTDNRNSTQVQPSKKIQGGKTS LNDRKGEKDVVDQEGNIQKRKRKRKRKRQRHNV
 DLDDTVRLQRRTRNLI LIRMKQEQNLIDAYAGEGWKQSQSREKIRPEMELQRAKKQILKCKLSIRDAI
 HQLDSLSSVGSIEGSVIATDGSVSHEHIFCANCKINEVSPDNDIILCDGTCNRAFHQRCCLDPPLET
 EDIPPEDQGWFCYCDCKIEILEATNAHLGTRFPLDSTWQDVFKEEAAIPDGDAALLNQEEWPSD
 DPEDDDYNPERKEESHGGFNTGNDKNASDDSSSSSSMWSLNGECSSLDEGINLEYYSNDHIDSDE
 SGEIACRRRQRRVVDYKPLYDEMFGKDAPPCEQVSEDEWDGPRKRRRREKESAVNTLMTLHESEN
 KYPNNKNNDRIRGNSSGIKRPCFRFSHEAVEKLRQVFAENELPPKSVKDALSKEGLDAAKVNKWF
 KNARYMALKIRKLQEGGQQLQSITSKTSKDSSTSQHVQEDEVLPKSAKITVISSLKCCENVTGKEK
 TKASSNPLKKRPEIPPCLENGENKDSMEVSDVSLMKLLQDRKKVVSFAFEGDSEAAELEFERLS
 KVVKTKIDRLKQRLTGVCNCRSKSGSDEVHLENEPSIMYVPIAELREKVK

SEC ID Nº: 187, secuencia de ácido nucleico PHDf_HD de *Pinus radiata* ADW18458

ATGCCGAAAGCAGGAAAAGGCTCTTCCGGTTCACCTCCGAAAGAAGCAAGCGTTGTCAAATACAGCA
 TCAGAGCTAAGGCCTAAAGCTGCAATTCAGAAGACTAACAATGCCGGGAGTATTTCTGGCAGGATG
 AGGGAAAGGAAACCACGGTTACATGTACAGATGATTGCCTCAATTCTCACGAAAAAAGGACAGGT
 GAAGCATCTGAATCGAGAAATGATTCTGAGCTAGAAGACAAACAATATCTGAGCAATAGTTCAAAA
 ACAAATGCCAATTTTGTAGCGGGGAAGTTAGTTAAAAAGACTTTGAAAAAAGCCCCAAGATTAAC
 TTGATTAAGAATGACAGTCAGAAAAAACCCAGTAAGAGTTCTGCTTCGAGGTCTTATGCTGCACAG
 AAACCTGAACGGCCACAACAGAGAAAAATTAGGAGGACAAAGTGGAAAGGGTATTTAAAAATATA
 ACAAAGGTGAAAACCTGTGAGAACTAGTAATATAGATGGCTTGAAGAGGAAGCAGCAAAGGAGAAAG
 AAGAAAAAGAAGGCGCAGAATGTTCCACAGGATGAAGTATCACGTGCAAAGAGAAGGATCAAATAT
 CTTTTGCTCAAATGAACTTGTAGCAGAATCTTATTGATGCATATTCAGGGGAGGGCTGGAAAGGC
 CAGAGTCGGGAGAAAATAAAGCCAGAGAAGGAGTTACAGAGAGCTGAAAAGCAGATTTCTGCAATGC
 AAACCTTGGAATTCGTGAAGCAGTGCATCAGCTTGATCTTCTAAGTTCCGGAAGGATCTATAGCAGAT
 GCCGTTATGATCCAGAAGGGCGTGTATTTTCATGAGCATATATTTTGTGCAAAATGCAAAATGACA
 GATGCTCTTCCGGATAATGACATAATATTGTGTGATGGAGCTTGCAATCGTGGCTTCCATCAGAAA
 TGTCTGGATCCACCTTTGGCTACTGAAAAATTCCTCCTGGAGATCAGGGTTGGCTCTGTAAAGTT
 TGCGAATGTAAGCTAGAGAGTTTGGAAAGCAATAAATGCTCATTGGGCACCCAGTTTACAGTTGAT
 AATAGCTGGGAGGSAATTTTGTGATGCAGCTAGAATTGCGAATGGTGAAAGTACAGCTACTGGG
 AATGGGGAAGAATGGCCATCAGATGATTGAGAATAATATGAGTGATCTTGGAAAGTCAAGTGGATGGT
 CAGAATCA TATAATGAAACTGACAATGAATCAGATGCTTCCCTAACAGATAATCTAAATGAATACATCTCTAAG
 GGAAAGAAGAGGAGCAGAGACTGATGAAAAAACGGACAAGAAGCTATAATAGATTTAGATACT
 GATGAGAATGATGGCTCAGACATGCCAGTTTTCAGGGAGAAGGCAGCGAAGAGATGTTGATTATAAA
 AAGTTGCATGATGAAATGTTTGGAAAAGATGGACCTTGTGAAGATGAAGTCAGTGAAGATGAAGAC
 TGGGGACCCAATAGGAGGCGTAGAAGAGCGAAAACAGCCGATACAAATACCACTAAGGATTTCTTCA
 AATGAAGAGAAAAGCATGTTAGAAAGGAAGCAGATGAGAAAGTAGCAGAGCATTCACAAAGTAGC
 ATAAAGGATAAGAAACATATTGTCAGACTGCCAGCTAATGCAGTTGAGAAGCTGCGCAAAGTCTTT
 GCTGAAAATGAGCTACCGTCAAGGTCTTACAAGGAAAATCTGTCACAGCAACTGGGCCTTACATTT
 CGAAAGGTCCATATGTGGTTCAAAAATGCCCGTATATGTCTCTGAAGAGCCGAAAGGAAATGCCT
 AGAAATGAGAAGTCAATGAACTAATCTCATGGGAGAAGGTGATATGGGGCCAAAGAAAAAATA
 CCAAAGAGAAAGAAAGCAGAATCAACATGTCTTTTATCTTCTGTTAGAGCAGAAATATAAATCTAGA
 AATACAAAAAGGAAGCATAGGAGAAAAACATCAAAGCACCTTATCTATCTCTCCAGAAGTGAG
 AGAAAAAGAACTATGTGCAGACTGCCTGCCAATGCAGTTGAGAGGTTTCGCCAAGTCTTTGCTGAG

FIGURA 7 (continuación)

AATGAGTTGCCTTCAATGTCCTGTAAAATGGCATTATCAAAGCAGTTGGATCTACCTTATCGGCAG
 GTGCATATGTGGTTCAAAAATGCTCGTTATATGTCTTTGAAGAAGAAAAAGATGAACCGTTTAAAA
 GAATCAGCTCCCGTTATTTCCCTCACATAACCAGAAGACTTGAAGTAGAAAGTAGCAAACCAATCCA
 CAAATCAAAGAGCCAGATGATGCTGGCAATTCATACTTAACAGAGATGGAAAAATGGGTAATATT
 GAGTTGAAACTTGAAAATATGAGAAAGATTTTAGAAACAGTCTGTCTTAAAGAGGTACTGATTTT
 TCCAATGATAAGAAATCAAGAAGGAATCGCCAACCTTCTTAGTGAGGAGCTTCTTGTGTATGTACCA
 GTGGTGGAGTTGAGGGAGAAGAGCTTACAGGTTGCTCCTGTCAGGTGA

SEC ID N°: 188, secuencia de polipéptido traducida PHDf_HD de *Pinus radiata*

MPKAGKGSGLRKKQALSNTASELRPKAAIQKTNNAGSISGRMRERKPRHLVQMIASILTCKRTG
 EASESRNDSELEDKQYLSNSSKTNANFVAGKLVKKTLLKAPRLNLIKNDSSQKKPSKSSASRSYAAQ
 KPERPQQRKIRRTKWKGD SKNITKVKTVRTSNIDGLKRRKQRRKKKKKAQNVPODEVSRARRIKY
 LLLKMKLEQNLI DAYSGEGWKGSREKIKPEKELQRAEKQILQCKLGIREAVHQLDLSSEGSIA
 AVIDPEGRVFEHI FCAKCKMTDALPDNDIILCDGACNRGFHQKCLDPPLATENIPPGDQGWLCKV
 CECKLESLEAINAHLGTQFTVDNSWEGIFADAARIANGESTATGNGEWPSDDSEDDDYDPEKQEI
 RDEGTGYNMMSDSGKSSGWSSESYNETDNESDASLTDNLNEYISKGKKRSRDTDEKNGQEAII DLDT
 DENDGSDMPVSGRRQRD VDYKKLHDEMFGKDGPCEDVSEDEDWGNRRRRRAKQPDNTTKDSS
 NEEKEHVRKEADEKVAEHSQSSIKDKKHI VRLPANAVEKLRKVFAENELPSRSYKENLSQQLGLTF
 RKVHMWFKNARYMSLKSREMPRNENCNETNLMGEGDMGPKKIPKEKKAESTCLLSSVRAEYKSR
 NTKRKHRRKNIKAPLSISSRSERKRTMCRLPANAVERFRQVFAENELPSMSCKMALSQQLDLPYRQ
 VHMWFKNARYMSLKKKKMNRKESAPVISSHTRRLEVESSKPNPQIKEPDDAGNSYLTMEKLGNI
 ELKLENMRKILETVCPKRGTD FSNQKQEGIANLLSEELLVYVPVVELREKSLQVAPVR

SEC ID N°: 189, secuencia ácido nucleico PHDf_HD de *Populus tremuloides*

ATGGGTGATTCTGGAAAGACGTCAAAGCTGCAAGACTCGCACAAAATGTTACCTTCTGACACAGTG
 ACTGGGTCGTTGCTGATTAAATCAATTGAAGATAAAGAAGGACAGCAAATATCTCCTAGAAAAGGC
 CAAAAACGAAAACAAAATCAAACCAAACCAATACCTCATCTAAAAACTATCATCAGTTCAGCT
 GTATCAAAAAGAAAAGTTTCTCCTAAGGGCATTGGGAATGGCTCTACCAGTAGAAAAGTTCAT
 AGGAAAATTTCTGCATAAAGCACTTGATAAGAAGGCCTCAAGGAAGGGGCTTCTCAGGGCTACAA
 TTGTCAACTATTGATTCTAAGGGAAATGGAAAAAATGGTGATGAAGGTGCAATTAAAAAACTCAAG
 AAGAGGAAGCCTAAGAAAAGGCAAAGGGACAAGGTAAAGCTAGATGAACCATCACGATTGCAGAGG
 AGAGCAAGGTACCTGATGATTAAAAATGAAGCTGGACCAGAACCTTATAGATGCTTACTCTGGAGAA
 GGCTGGAAAGGTCAAAGTCGAGAGAAGATCAGGCCAGAAAAGGAAGTACTGCGAGCAAGGAAGCAG
 ATTTTGAAGTGTAAGTGGTTTACGTGATATAATTCGGCAGGTGGATTCTCTTAGTACAGTAGGA
 TGTATTGAAGAGACAGTTATGGCTCCAGATGGATCTGTTTCACATGAACATATATTCTGTGCGAAG
 TGCAAATTGAATGAAGTTTCCCCAGATAATGATAATTGACTCTGTGATGGTACATGCAACTGTGCT
 TTCCACCAAAAATGCCTTGAACCTCCCTTGGATACTGAAAGTATACTCCAGGAGATCAGGGATGG
 TTTTGAAGTTCTGTGAATGCAGGATGGATATTATAGAAGCTATGAATGCTCATTGGGAACCCAT
 TTCTCAGAGGACAGCAGTTGGCAGGATATTTTACAGAGAAGCAGCAATTCCTGATGCTGGAAAT
 GTGCTATTAAACCCCGAAGAAGAATGGCCTTCAGATGATTCTGAAGATGATAATTACGATCCAGAG
 AGGAGAGACAACATCATGAGTGAAGCGGGTACTGATGATGATGCATCTGATGACATCAGCAGTTCA
 ACCAGCTTGGGTTGGTCTTCAGATGGTGAGGTTTTTTGGGATCCAGGAGGTGGGAGATGCATGGC
 TTGGACTTCAGAAACAATTCTATTTATAGCAGTTTAGATTCTGATGAAACCAGTATGGGGAAATT
 GTTTGTGGCCGTAGGCAGAGAAGAGCAATTGACTATAAGAAGTTATATGATGAGGTGTTGGAAAG
 GATGCTCCTGCACACGAGCAAGCCAGTGAAGACGAAGATTGGGGTCTGGCAAAGAAAGCGACGG
 GAGAAGGAGTCAAATGCAGCCAGCACCTTATGACACTGTGTGAAAGTAAAAAAAAGTAAAAGT
 GATGAGACCATTGAAGGTATGATGAACCTTCCCCCTCAAACCTCGAAGGCCAATTTTCAGACTCCCC

FIGURA 7 (continuación)

CCTGATGCAGTTGAGAACTTCGCCAAGTATTTGTAGAAAATGAACTTCCATCTAGAACTGTCAAG
 GAGAATCTGTCAAAAAGAATTGGGCCCTTGAACCTGGGAAGGTTAGCAAATGGTTCAAGAATCCCGT
 TATTTAGCTCTGAAGTCCCGGAAGGTAGAGAAAAGGAGAACAACCTTCATAACTCTAGTTCCAAAGTC
 TCTGCTGAACCCACATTAAATGTCATGGAGGGCAAACTGCTGATCTTTCACAGGATTCCTGGGAA
 GAACTGAGGTATGTATCCCAAAGAATTTGAAAAGGATCCTTCAGAGGAAGAAGCTGAAGTCAATA
 AGCAGAAGTTTAAAGGAAAAATGAACAGAAAAGAGGTTTCCTACGAATCACCTACTAAAAGCAATGAG
 ATGAATGCTGAGTGCAGTGATGATGTGAGCTTGAAGAAGTTGTTAAAAGCAAACCAAAGGGAGTA
 AAGAAAAAGGTTAATCCTATTTTCAGTAGCGGCTGAATATGATATGGAGAGACTCTGTAGAGCGAAG
 GCTATATTAGAGAAGATGAAGCAGAAAGTGGTGAAGCTACAAAATGGCAAAGCTAGGAAGTCATCC
 AAAACCCGTCCTACTTGAAGAATTTATCGTTTACATTCCTATTGCTGAGTTGAGGGGAAAAAATTGA

SEC ID N°: 190, secuencia de polipéptido traducida PHDf_HD de *Populus tremuloides*

MGDSGKTSKLDQSHKCSPTSMTVTGSLLIKSLKIKKDSKISPRKGQKTKTKSKPKPIPHLKTIISSA
 VSKRKVSPKIGINGSTSRKLIHRKILHKALDKKASRKGASSGLQLSTIDSKNGKNGDEGAIKKLLK
 KRKPKRQRDKVKLDEPSRLQRRARYLMIKMKLDQNLIDAYSSEGWKGQSREKIRPEKELLRARKQ
 ILKCKLGLRDIIRQVDSLSTVGCIEETVMAPDGSVSHEHIFCAKCKLNEVSPDNDIVLCDGTCNCA
 FHQKCLEPPLDTEIIPPGDQGWFCFKFCECRMDIIEAMNAHLGTHFSEDSSWQDIFTEEAAIPDAGN
 VLLNPEEWPSSDDSEDDNYDPERRDNIMSEAGTDDASDDISSSTSLGWSSDGEVFLGSRRWEMHG
 LDFRNSIYSSLDSETSDEIVCGRRQRAIDYKLYDEVFGKDAPAHEQASEDEDWPGKRKR
 EKESNAASTLMTLCESKKSDEIIEGMMNLFPQTRRPIFRLPPDAVEKLRQVFVENELPSRTVK
 ENLSKELGLEPGKVKWFKNSRYLALKSRKVEKGEQLHNSSSKVSAPTLNVMEGKTADLSQDSWE
 ETEVCIPKNLKRILQRKLLKISRSRLKNEQKRGSYESPTKSNEMNAECSDVSLKLLKAKPKGV
 KKKVNPISVAAEYDMERLCRAKAIKMKQKVVKLQNGKARKSSKTRPLEEFIVYIPIAELRGKN

SEC ID N°: 191, secuencia de ácido nucleico PHDf_HD de *Saccharum officinarum* cóntigo de CA157855.1, CA261734.1, CA253314.1, CA220753.1, CA201G1958.1

ATGCATTCTTCGGAAAATAAACTGGTATGCTCCAACCTCGGGAAGATCATCAAAGGGAATGAGACC
 TCTATGGAGTTGGTTCCCTGTACCAAAAAGACCAACTAGACATGATGCTTCCCGCCAATGCAAATCT
 GATTCTCCTTTGAAGCGATCACCAAGGAAAGCTAGAAATGCTACTCTGGCAAAAAGCATAAAGAAT
 AAATATCATTGTAGTCCCCTCAAACAGCGGAGGGCTTCAGATTCTGTTTCTGGAAAAGTTGCGACA
 GGACTAACTGTGAGGAGGAGGAAGAAAAGAAAAATGCAAATACTGATGAGGCAACTCGCTTGGAA
 CGAAGAGCGAGATATTTTCTAATCAAGATAAAAATGGAGCAGAATTTGCTAGATGCTTACTCTGGA
 GATGGATGGAATGGGCAAAGCCGGGAAAAAATAAAGCCAGATAAGGATCTGCGGCGTGCCAGGAAA
 CAAATCATAAAATGCAAATGCTATACGTGATATCATTCCGCAACTTGGTTTGTATACTTCTACT
 GGGAGTCTTAATGACCCAGCAATGCCTCCAGATCAGTCTACCAATCCGGAACATAACCATGTGCTCA
 ACATGCAAGTCTCATGAATCATTTCAGCAATAAATTTATCTTCTGTGAAGGGCCCTGCAAAGG
 GCATATCATGAGAAATGTTTGGAACTCCCTTAAACAAAAGCGTACTTCCAACAAGTAGCCATGGA
 TGGCTTTGCAAATCTGTTTGTGCAAAGTAAAATTTTAGAACTATTAATGCACATCTAGGAACA
 AGTTTACAGTGAAGTGCCCTTCGAAGATATATTCAAGGAAGCAACGGAACAAATAGACTCCGAG
 GATGCACTAGATGAAGATTGGCTTCTGAGTATTCAGGTGATGAGGACTATGATCCTGATGAAAAT
 GAGGACAGCGGCAACTGTATGGACAGCGGGGAGGAAATTATGTCTGATGATTCCGATGGTTCAGGA
 AGCCCCTTTATTCTCCAATGACGATATTCTGACTTCATATCAGCAGACTTAAATGACGTGGAA
 GGGTTTGTACACCAATTTAGACTTAGGCATTGATGCCGTTGAAGATGATTTGGCACAGATCCTT
 ACCTACCAAAGGCCAAGAAGAGATGTTGATTATAGAAGCTTAATGAGGAAATGTTTGGAAAAATA
 ATAGGAAATGAAGGACAGAGTGAAGATTGGGGCCATGAAAGGAGAAAGAAAAGGACTTGT
 TCAGGTGGTCTGGGATAATTTCTGTGGTTTCTCAAATGTTATATCTGAGGAAAAGAGCCAAAAG

FIGURA 7 (continuación)

AAGGGGAGAAAACTTTTCAGGATTCTCCTGCAGCTGTGCGAGGTACTTCGCAAAGCTTTTGCTGAA
 AATGAGCTTCCACCCCGGGATGTCAAAGAAAATCTCTCAAGAGAAATGGGAATTTCTTTTGAAAAG
 ATTGATAAATGGTTCAAATAACACGGTGTGCTGCTCTCAGAGATCGCAAGGCTGAAGGGAACAGT
 CATAATACGGCTCCCAGCAAAGCTCAAGAAAATAAAGGAAAAGCTGGAATCTCAAGCCAGGCTGAA
 AGGAATGGTCATGTTACTGGTCCCTGCAATAACTCGAGAACAAATGAAGAAAAATCGTTGACACTA
 GCTCCTAATGCCAGGGAGAGAAGAATGGGATTGCAGCAGATTATGTCTTACAAAAATCTGGTATA
 TCGGGGAAGGTTGATTCAGTAGACAACCTCTTGCTTGGTTCCCTTGTCTGAAAGCATCAATGTGCAT
 ACACGATTGCAACAGAACTTTGAGATGAGGAAGATGGAATCAACTAGCAGTCCTGTGTGGCTGCAC
 AATAAAGGAGGTTGCTTGTTCACACAGGCCAAGCCAAGGAGAGTACATTACCTACAGGCAAGCCT
 TGTTTGCAGTCAGAAATAACCCATTCAACAATTAATGAAGTGAGCACTTTAGTGCAAGCTACTTCT
 TGGATGGATGCTGGGTGCTGCGCTGAGGTGCAAGAAGCCACTCCTGGGTGGACATCGGGCCCTCA
 GATTACCAGCCTTCTGACGTGATCGATGAGATGTGCGGACTTGAGTGACAGGCTGCAGAGACTG
 AAGGAGAACATGCTCTCATCTGGCATAAGACGGCAAACCACAGGTGAGAGTGACATGGGAAACCAG
 GCTGTGGTGTGCTGCGGACCGCCGAGCTCAAGGAAAAGCGCCGCATGGCAGTTTTTTTGGGCAT
 TATTGCCCATAG

SEC ID N°: 192, secuencia de polipéptido traducida PHDf_HD de *Saccharum officinarum*
 MHSSENKLVCSNSGRSSKGNETSMEVLPVPRPTRHDASRQCKSDSPLKRSRPRKARNATLAKSIKN
 KYHCSPLKQRRASDSVSGKVATGLTVRRRKKRKMONTDEATRLERRARYFLIKIKLEQNLDDAYS
 DGWNGQSREKIKPKDLRRARKQIIKCKIAIRDIIRQLGLYTSSTGSLNDPAMPDQSTNPEHTMCS
 TCKSHESFPNKFIFCEGPCKRAYHEKCLEPPLNKSVLPTSSHWLCKFCLCKVKIILETINAHLGT
 SFTVKCPFEDIFKEATEQIDSEALDEDWLSEYSGDEYDPDENEDSGNCMDSGEEIMSDSDGSG
 SPLYSNDDIPDFISADLNDVEGFCHTNLDLGDIDAVEDDLAQILTYQRPRRDVDYRRLNEEMFGKI
 IGNEGQSEDEDWGHERRKRTCSGGAGDNSVGFNSVISEEKSQKKGRKLFRIPPAAVEVLRKAFAE
 NELPPRDVKENLSRELGISFEKIDKWFKNTRCAALRDRKAEGNSHNTAPSKSSRNKGKAGISSQAE
 RNGHVTGPCNNSRTNEEKSLTLAPNAOGEKNGIAADYVLQKSGISGKVDSDNSCLVPLSESINVH
 TRLQONLEMRKMESTSSPVWLHNKGGCLFPTGQAKESTLPTGKPCLOSEITHSTINEVSTLVQATS
 WMDAGSCAEVQEATPWVDIGASDYQPFLLVIDEMCGLECRLORLKENMLSSGIDGKTTGESDMGNQ
 AVVLVPTAELKEKAPHGSFFGHYCP

SEC ID N°: 193, secuencia de ácido nucleico PHDf_HD de *Vitis vinifera* AM477372.2, AM488059.1
 ATGCGTGGAAGTGGAAAGAAAGCAGGACATCAGGAATCTGGAAAGTCTTGCTTTCTAAAAGAAAT
 ATCGGGCCTAAGCTGAATGCAGCATTGCAGATTAATAATGGTAGTAAAATATCTCAGACCAGGAAA
 TGCAAGCCAAAATCAAATCTCATGCAAAGACAATTGGTGCAATCCTGTCTAAGAGAACAATACT
 GACTCTCCAAGCAAGGGAAGCAGGAGTGGATCCACAACCAGAAAAGTATTACAAAGAAAACCCTG
 CATAAAGCCATTGATACTGAGTCTTCCAAGAAGGAGTCTTCATCAAAGCTTAAAGGTGAGAAGCCC
 CCACAAATTAGCACAAAACAAAATGGGGAAACTGTGGACAAGAAATGTTAAACCTCAAAAATTAAG
 AAAAGAGGAAAGCGGAAGCGGCGAAAGGATAATTCAGAGCTAGATGAAGCTTCTCGCTTGCAAAGA
 AGAACCAGGTACTCTTGATCAAATGAAACTGGAGCAGAATCTTATTGATGCTTACTCTGGGGAA
 GGTGGAAGGTCATAGTCGAGAAAAGATCAGGCCAGAGAAGGAACTACAAAGAGCTACGAAGCAA
 ATATTGAAATGTAAACTTGGGATCCGTGATGCAATTCGCCAGCTGGAGTCACTGAGCTCCATAGGA
 TGTATTGAAGATACTGCAATTGCTTCAGATGGATCTGTTTATCATGAACACATAATCTGTGCAAAG
 TGCAAGTTGCGGGAGGCTTTCCAGATAATGATATTACTCTGTGATGGGACGTGCAATTGTGCT
 TTCCACCAGAAATGTCTTGATCCTCCATTGGAAGTGAATAATTCCTCCGGGAGACCAAGGTTGG
 TTTTGCAAGTTCTGCGAGTGTAAGATGGAAATACTAGAAGCAATGAATGCACATCTTGAACCCGC
 TTCTCTGTGGACAGTACTTGGCAGGATATTTTCAAAGAAGAAGCTGCTTTGCCTGATGGTGGGAGT

FIGURA 7 (continuación)

GCACTGCCCTATCCAGAGGAAAGATTGGCCTTCTGATGATTCACAAGATCATGATTATGATCCTGAA
 AGAAATGAAAATAGCTGCAGCATCAGCACTGCTGGCACTGAAGGAAATGCTTCAGATGATACAAAT
 AGTTCTTTAAGCTTGAGTTGGTCGTTTGAAGACGAAATCCTTTCTGGATCCAAAAGATCAGGGATT
 ATCAGTGCAGATTCTGATGAAACATCTGATTGTGAAATCATAAGTGGTCGTAGGCAACGAAGAGCT
 GTTGATTATAGAAAGTTATATGATGAAATGTTTGGAAAGGATGCTCATGCAAAATGAACAAGTTAGT
 GAAGATGAAGACTGGGGTCTGCTAACAAAAGGCGGAGAGAAAAGGAGTCTGATGCAGCCAGCACC
 CTCATAACTCTATATGAAGGTGAGAAAAGTTACCTAATGTGGAAACCATGGAAGCCAAACAAAAA
 ATTTCTTCAGATCCACAAACTAAAAGGCCATTTTCCAGAATTCCACTTGATGCGGTTGAGAAGCTT
 CGCCAAGCATTTGGGGAGAATGAACCTCCCTCTAGAGATGTGAGGGAGAATCTTGCAAAGCAGTTG
 GGTCTTGATTATGAGAAGGTGAACAAATGGTTCAAAAATGCACGATATATAGCGCTAAAAACAAGA
 AAGGCAGAAAGAGCAAAACAACCTCAGACTTCTCCAGAATCTCCAAGGAATCCAGATCAGAAATC
 GTGAAGGACAAAACCTGTTGATCTAGTGGCATCAAGGGATAACTCATCAGCATCACTGGTCCGTGCA
 CTAAAGAAATTTGAAAAGGTGCGGAGGAGAAAAAATCCAAAGCCAATAATGACCAGCCCTGTAAAG
 AAAAAGCATCATAGAAGGGCTTTACTTGAGTCACCTACAAATGACAAGGTTACTATGGAGTTTGAC
 GATGATGTGAGCTTGAAGAAAACAACCTGAAGCTTTTGAAAGAAAAAAGTAAGAGGGACAAGCAAAGG
 GTCGACTTCAAGGAGGGAACCTGGTGTACAAGATGCAGAGAAGGAGATGGAGAGACTCTGCCAAAT
 AAGGATAAGATAGAGAAGTTGAAGCAGGTAATACTGAGACTTCAATGTGACAAAACATAATCAATGG
 CAAGACCAGTCAGTGATCTATGTTCTGTGGCAGAGCTAAGGGAGAAAGGTCGATTTTGCATCACA
 TCT

SEC ID N°: 194, secuencia de polipéptido traducida PHDf_HD de *Vitis vinifera*

MRGTGKKAGHQESGKSCFPKRNI GPKLNAAALQIKNGSKISQTRKCKPKSKSHAKTIGAILSKR'TTT
 DSPSKGSRSGSTTRKLIHKKTLHKAIDTESSKKESSSKLKGEKPPQISTNKNGETVDKNVKPQKPK
 KRGRKRKRKDNSELDEASRLQRRTRYLLIKMKLEQNLIDAYS GEGWKGHSREKIRPEKELQ'RA TKQ
 ILKCKLGIRD'AI RQLESLS SIGIEDTAIASDGSVYHEHI ICAKCKLREAFPDNDI ILC'DGTCNCA
 FHQKCLDPPLETENIPPGDQGWFC'KFCECKMEILEAMNAHLGTRF'SVDSTWQDIFKEEAALPDGGS
 ALPYPEEDWPSDDSDHDYDPERNENSCSISTAGTEGNASDDTNS SLSLSWSFEDEILSGSKRSGI
 ISADSDETS'DCEIISGRRQRRAVDYR'KLYDEMFGKDAHANEQVSEDEDWGPANKRRRREKESDAAST
 LITLYE'GEEKLPNVETMEAKQKISSDPQTKR'PFSRIPLDAVEKLRQAFGENELPSRDVREN'LAKQL
 GLDYEKVNKWFKNARYIALKTRKAERAKQLQTS'PRI SKESRSEIVKDKTVDLVA SRDNSSASLVRA
 LKNLKKVRRRKNPKPIMTSPVKKKHRRALLES'PTNDKVTMEFDDDVSLKKQLKLLKEKSKRDKQR
 VDFKEGTGVQDAEKEMERLCQIKDKIEK'LKQVILRLQCDKTNQWQDQSVIYVPAELREKGRFCIT
 S

SEC ID N°: 195, secuencia de ácido nucleico PHDf_HD de *Zea mays* cóntigo de EE162310, DN204182, CF057937

ATGCATTCTTCGGAAAAATAA'ACTGGGTTGGAAATGAGACCTCTATGGGGTTGGTTCCTGTACCAAAA
 AGACCAGCTAGACCTGATGCTTCC'CAACCAATGCAAATCTGATTCTTTATGAAGCGATCACCAAGG
 AAAGTTAGAAATGCTACTCTGG'CAAAAAGCATAAAGAGCAAATACCACTATAGTCCCTCAAACAG
 CGGAAGGGTT' CAGATTCTGTTCTGGGAAAA'TCGTAACAGGACTAACC'GCAAGGAAAAAGAGAAA
 AGGAAAATCCAAAT' TACAGACGAGGCAACTCGTTT'GGAACGAAGAGCGAGATATTTCTAATCAAG
 ATAAA'ACTGGAGCAGAATTTGCTAGATGCTTACTCTGGAGATGGATGGAA'ATGGGCAAAGCCGAGAG
 AAAATAAAGCCAGAGAAGGAACTGCAACGTGCCAGGAAACAAATCATAAAATGCAAAAT'TGCTATA
 CGTGATATCATCCGCCAACTTGATTTGTATACTTCTACTGGGAGTGTGATGACCCACTAATGCCT
 ACAGATCAGTCCACCAATCCC'GAACATAACCATGTGCTCAACATGCAAGTCTCATGAATCATTTCCC
 AGCAATAAAAT'TATCTTCTGCAAAGGGCCCTGCAA'AAAGGGCATGTGATGAGAAATGTTTGGAACTT
 CCCTTAAACAAAAGCGTACTTCCAACAAGTAGCCATGGGTGGCTTTGCAAAT'TCTGTTTGTGCAAG

FIGURA 7 (continuación)

GTGAGGATTTTAGAAACTATTAATGCACATCTAGGGACGAGTTTTACAGTGAAGTGCCACTTTGAA
 GATATATTCAAGGAACTACTGAACTAATAGACTCTGAGGATGCACTAGATGAAGATTGGCTTTCT
 GAGTATTCAGGTGATGAGGACTATGATCCTGATGAAAATGAGGCCAGTGGCGACTGTATGGACAGC
 GGGGAGAAGATTATGTCTGATGATTCCAATGGTTTCAGGAAGCCCCCTTTATTCTCCAAATGACGAT
 ATTCCTGACTTCATATCAGCAGACTTAAATGTTGTGGAAGGGTTTTGTGCATACCAATTTAGATTTA
 GGCATTGATGCCGTTGAAGATGATTTTGCACAGATCCTCACCTACCAAAGGCCAAGAAGAGATGTT
 GATTATAGAAGGCTTAACGAGGAAATGTTTGGGAAAATAACCGGGAATGAAGAACAGAGCGAGGAC
 GAAGATTGGGGCCATGAAAGGAGAAAAGAAAAGGACGCATTTCAGGTGTTGCTGGGGATAATTCTGTT
 GGTTTCTTGAACGTTATATCTGATGAAAAGAGCCAGAAGAAGGGGAGAAAACTTTTCAGGATTCCT
 CCTGCAGCTGTTGAGGTACTTCGCAGAGCTTTTGTGAAAATGAGCTTCCACCCCGGGATGTTAAA
 GAAAATCTCTCAAGAGAAATGGGAATTTCTTTGAAAAGATTGATAAATGGTTCAAAAAATACACGG
 TGTGCTGCTCTCAGAGATCGGAAGGCTGAAGGAAACAGTCATAATACAGCTCCCAGCAAAGCTCA
 AAAAAATAAGGAAAAGCTGGAATCTCAGGCAAGACTGGAAGGAATGGTCATGTTACTGGTCCCTGC
 AATAATTTGAGAACAAATGAAGAAAAACAGGTATATCGGGGAAGTTTGATTTCAGGAGACAACCTCT
 TGTTTGGTTCCTTCTCTGAAGTCATCAATGTGCCTACGCGATTGCCACACAACCTTTGAGATGAGG
 AAGATGGAATCAACTAGAAGTCCCTGCGAGGCTACACAATAAAGGAGGGTTCCTGTCTGCAACAGTC
 CAAATTAAGGAGAGTACGTTGTTACCCACAGGCAAGCCTTGTGGCAGTCAGAAATGAGCCATCCA
 ACAACTAATGAAGTGAGCACTTCAGCGCAAGCTACTACTTGGATCGACCGGGGGCGTGTGCCGAG
 GAGCAAGAACCTACTCCTTGGGTGGACATCGGGTCTCAGACTACCAGCCTTTTCTGGATGTGATC
 GACGACATGTGCGGACTTGAGTGGAGGCTGCAGAGACTGAAGGAGAACATGCTCTCATCTAGCATA
 GACGGCCAAACCGCCGAGTGAGAGTGACAAAAGAAACCAGACCGTGGTGTAGTGCCAACCGCC
 GAGCTCAAGGAAAAGCTGCCGCATGACGGTTTTATTCGGGCATTATTGCCCATAG

SEC ID N°: 196, secuencia de polipéptido traducida PHDf_HD de *Zea mays*

MHSSENKLGWNETSMGLVFPVKRPARPDASHQCKSDSFMKRSRPRKVRNATLAKSIKSKYHYSPLKQ
 RKGSDSVPVKIVTGLTARKKKRRIQITDEATRLERRARYFLIKIKLEQNLLDAYSGDGWNGQSRE
 KIKPEKELQRRKQIIKCKIAIRDIIRQLDLYTSTGSVDDPLMPTDQSTNPEHTMCSTCKSHESFP
 SNKIIFCCKGPKRACHEKCLEPPLNKSVLPTSSHWLCKFCLCKVRILETINAHLGTSFTVKCHFE
 DIFKETTELIDSEDALDEDWLSEYSGDEYDPDENEASGDCMDSGEKIMSDSNGSGSPLYSPNDD
 IPDFISADLNVVEGFCHTNLDLGI DAVEDDFAQILTYQRPRRDVYRRLNEEMFGKITGNEEQSED
 EDWGHERRKTRTHSGVAGDNSVGFNLVI SDEKSQKKGKLFRI PPAAVEVLRRAFAENELPPRDVK
 ENLSRELGISFEKIDKWFKNTRCAALRDRKAEGNSHNTAPSKSSKNKGKAGISGKTGRNGHVTGPC
 NNLRTNEEKTGISGKFDSDNSCLVPFSEVINVPTRLPHNFEMRKMESTRSPARLHNKGGFLSATV
 QIKESTLLPTGKPCWQSEMSHPTTNEVSTSAQATTWIDAGACAEQEPTPWVDIGVSDYQPFLDVI
 DDMCGLEWRLQRLKENMLSSSIDGQTAASESDKRNQTVVLPVPTAELKEKLPDGLFGHYCP

SEC ID N°: 197, secuencia de ácido nucleico PHDf_HD_HAT3.1 de *Arabidopsis thaliana* NM_112838 At3g19510

ATGTATAAAGCAGTAAGTAAGCGGGTAACTAGAAGTAGTGGATCTGGTTTGAAGCAGACAAATGTA
 GATAATGGAGGAGAGATTAGTCCAACCTGTTGACAGAGTATCTGAACAAGGCAAATCATCAGAGGCA
 GGGAGTCACATGCCAACTGATGCAAACGGAAAATGGACATTTGCATCATGAGATTATGGATCATGGA
 AAGGGGAATGAAGAACAGAAACCAACCCCTCAAACCGTGAAGAAGGATTCTAATACCAATACCAA
 TTTTCAGGGTCTCATCGTGAGCTTGTTCATCGGCTTTCCTTGCCGTGGACAATTTGAAATCCACAAC
 CGGTCCAGGGCATCTACGAGTTCGAAGAGGTTAGGAGGAGGTGGAGAAAGGAATGTGTGTTTGCA
 AGTCACAAGAGAGCTCAGAGATCTAAGGAAGATGCTGGACCATCTTCGGTTGTTGCTAATTCACA
 CCTGTAGGTCGACCAAAGAAAAAGAATAAGACTATGAACAAAGGACAAGTTAGAGAAGATGATGAG
 TACACAAGAATCAAAAAGAACTTCGATACTTCTAAACCGGATTAATGAGCAAAGCCTAATT
 GACGCCATTTCGTTAGAAGCTGGAAAGGTTCAAGCCTGGAAAAGATAAGGCCAGAGAAAAGAGCTT

FIGURA 7 (continuación)

GAACGAGCCACAAAAGGAAATTTTGC GGCGCAAGCTTAAAATTAGAGATCTTTTTTCAGCACCTTGAC
 AACTCTGTGCTGAAGGAAGCCTTCCAGAGTCTTTATTTGATACTGATGGAGAGATTTCTAGCGAG
 GATATCTTCTGTGCAAAATGTGGTTCAAAGGACTTGTCTGTTGATAATGATATCATACTATGCGAT
 GGGTTTTGCGATCGAGGCTTTCACCAATACGTCTTGAACCACCTCTGCGGAAAGAAGATATTCCT
 CCGGATGATGAGGGTTGGCTATGCCCCGGATGTGATTGCAAAGATGACAGCTTGGACTTGCTTAAT
 GATTCTTTAGGGACAAAATTTTCGGTCAGCGACAGTTGGGAGAAAATATTTTCTGAGGCAGCAGCA
 GCGCTGGTTGGTGGAGGTCAAAATCTTGACTGTGATCTTCCCTCAGATGATTCTGATGATGAAGAG
 TATGATCCAGATTGTTTGAATGATAATGAAAATGATGAAGATGGCAGTGATGACAATGAAGAATCT
 GAGAACGAAGATGGTAGTTCTGATGAAACTGAATTCACATCTGCATCTGACGAAATGATTGAGTCA
 TTTAAAGAAGGGAAAGATATAATGAAAGATGTAATGGCCCTTCCATCTGATGATTCTGAGGATGAT
 GACTATGATCCTGATGCCCGACTTGTGATGATGATAAAGAAAGTTCAAATTCTGATTGTACATCA
 GACACTGAAGATCTTGAAACTTCTTCAAAGGAGATGAGACTAATCAGCAAGCTGAAGATACACCA
 CTTGAGGATCCTGGCAGACAGCTCTCAGTCCAAGGTGATGCAATTTTAGAGTCAGATGTTGGG
 CTTGATGATGGTCCCGGTTGATCCAGAAGGAAATGTTGAAAGGCTGGATTACAAAAGTTG
 TATGATGAGGAATATGACAATGTTCGACTTCGTCAAGTATGATGACGATTGGGATAAAAAGTCA
 AGAATGGGAAAAGAAGATTCCGAGTCGGAAGATGAAGGGGACACTGTACCTCTGAAACAGTCTTCA
 AATGCAGAAGATCATACTTCTAAGAACTTATACGGAAGTCGAAAAGGGCAGATAAAAAGGATAACC
 TTAGAAATGCCACAAGAAGGTCCTGGAGAAAATGGTGGTCTGGTCAAATTGAAAAAGCAGCTCT
 TCAGCATGTAAACAGACAGATCCCAAACTCAGAGACTGTACATATCTTTTCAAGAGAATCAATAT
 CCAGACAAAGCCACGAAGGAGAGCTTAGCCAAAGAGCTACAAATGACAGTTAAGCAAGTTAATAAC
 TGGTTTTAAGCATAGGCGTTGGTCCGATAAACAGCAAACCTTTAGTTTCAGAGGAAAATGTCGAGAAG
 TTAAAACAGGCAAGGAAGGTGAATGTGAGACATCTGTTGCTGGAAGTAGCAAACAACTATGGAA
 ACCGAGTCAGTTGCAGAGAAACCAACCAATACTGGATCACGGAAAAGAAGAAGAAAGTAG

SEC ID N°: 198, secuencia de polipéptido traducida PHDf_HD_HAT3.1 de *Arabidopsis thaliana*
 MYKAVSKRVTRSSGSLKQTNVDNGGEISPTVDRVSEQKSSSEAGSHMPTDANGNHLHHEIMDHG
 KGNEEQKPTPQTVKKSNTNFKSGSHRELVIGLPCRGQFEIHNRSRSTSSKRLGGGGERNVLF
 SHKRAQRSKEDAGPSSVVANSTPVGPRPKKNKTMNKGQVREDDEYTRIKKKLRYFLNRIYEQSLI
 DAYSLEGWKGSLEKIRPEKELERATKEILRRKLIKIRDLFQHLDTLCAEGSLPESLFDTDGEISSE
 DIFCAKCGSKDLSVDNDIILCDGFCDRGFHQYCLEPPLRKEDI PPDDEGWLCPGCDCKDDSLDLLN
 DSLGTFKFSVSDSWEKIFPEAAAALVGGGQNLDCDLPSDDSDDEEYDPDCLNDNENDEDEGSDDNEES
 ENEDGSSDETEFTSASDEMIESFKEGKDIMKDVMLPSDDSEDDDYDPDAPTCDDDKESSNSDCTS
 DTEDLETSFKGDETNQQAEDTPLEDPGRQTSQLOQDAILES DVGLDDGPAGVSRRRNVERLDYKKL
 YDEEYDNVPTSSSDDDDWDKTARMGKEDSESEDEGDTVPLKQSSNAEDHTSKKLIKRSKRADKKDT
 LEMPQEGPGENGGSGEIEKSSSACKQTDPKTQRLYISFQENQYDPKATKESLAKELQMTVKQVNN
 WFKHRRWSINSKPLVSEENVKLTGKEGECETSVAGSSKQTMETESVAEKPTNTGSRKRRRK

SEC ID N°: 199, secuencia de ácido nucleico PHDf_HD de *Lotus japonicus* AP006117.1
 ATGATTAGTTCAACTGGAGAATTGACAAGCCATAACAACAGTAGCATTGAGCACATGGAAAAGTGG
 CAACCTGAATTGCTGAAAAACACCTCAGATAGGCTCTGAATGTCTGGACGGTGAACAAAGGGTG
 CTTGGTACTGTGTCAACCAGTCTGTTAATAATGAAAATTCAAATCAAGTTTCTGCCAATGTGACC
 GAGAATTCTGTTTTCAACTACCAGCAACACCTCAACATTGCTTAGAAAAGAGTGGTCAGACTGTA
 ACAAGTCCACCCCTGGAAGAAAGCACCTTGAACAAGTTGCTACTAATGTATCTAATGGTAAATCA
 GAAAACAAATGTCAGACATCCTCTCAAGATGTTCAAACGAAGTGTGGAAGCAAGTGTGTAATA
 ACTGGCTCTTTTGTGCGGACCAACACAATCCATTCTGCTCAAGTGAATACGAGTTCTGTGAAT
 GAACTATTGGATATCCCTCCGGAGATGTTGCAGTAGAAGTTGGTTCTGATAGTTTAGAAAAAAG

FIGURA 7 (continuación)

TCTAAAGGGTGGCTTGGTAGACAGCGAACAAAAGGGTACTTGGTACTGTGTTAACCAATTCTATTAAT
 GATGGAAATFCAAATCAAGTTTCTGCCAATGTGACCGAGAATTCTGTTATTCAACTGCCAGCACCA
 CCTCAACATTCGCTTTGAAAAGAGTGGACAGACTGTAAAAAGTTCATCTCTGGAAGAAAGTACCTTA
 AAACAAGTCTCCACTAATTTATCTAACAGTAAATCAGAAAACATATGTCAGACATCCTCTCAAGAT
 GTTCAAAACGAACTTGTGGAACAAGTGTGAGTAACTGGCTCTGTGTTGCAGACAAAACACAA
 TCCATTCTGCTCAAGTGAATACAAGTTCTGTGAATGAACTATTGGATAATCCTTCCAGAGATGTT
 GCAGAAGAAGTGGTTCGATAGTTCAGAAAAGGAAGTCTAAAAGCACAACTCCTTCCACAATTGAGA
 CATATAGGTAAAAGTAACTCAAAGTTATCGAAAAAGAAAGTATATCCTAAGGTCGTTAGGAAGCAGT
 GACAGAGCTTTGCGGTCCAGGACAAAAGAAAAACCTAAAAGCGCCTGAACCAAGTAGTAACTCAGTT
 GACGTTAATAATGATGGAGTGAAGAAGGAAAAAGGAGGAAGAAGAAAAAACAGGAGGGGAGGAA
 AAGAAGGATCAATTTTATAAAATCAAAGCTCACCTAAGATATTTATTAAACAGAGTAAGCTATGAG
 CAAAACCTAATGATGCTTACTCTGGTGCAGGCTGGAAAGGATACAGTATGGACAAGTTAAAGCCT
 GATAAGGAGATTCAACGGGCTAAATCTGAAATTCCTCGACGTAAATGAAAAATTAGGGATCTATTT
 CAAAACCTGGATTCTTGTGTGCTGAAGGAAAGTTCCAGAATCTTGTGTTGATTCTGAGGGAGAG
 ATTGACAGTGAAGTATATTCTGTTCAATTTGCCAGACCAAGAGATTGAGCACTGATAATGATATA
 ATTCTCTGTGATGGTGCCTGTGACCGTGGATTTACCAGCACTGCTTGGATCCTCCATTGTTAACT
 GAAGACATTCCACCTGGCGATGAGGGTTGGTTATGCCCGGATGTGATTGCAAAGATGACTGCGTT
 GACCTTCTTAAACGACTCCTTAGGAACACGTCTCTCTTCTTAGTGACACCTGGGAGAAGGTTTTTCT
 GAAGCAGCAGCTGTTACTGGAAATAATATGGATCAGGGACTTCTTCTGATGATTCTGATGATGAT
 GATTATAACCCATAATGGTAAAGAGGATGTGGAGGTTGAAGGAGGTGAATCAAGTCTGATGAATCT
 GAATATGCTTCTGCTTCTGAGAAATTGGAAGATTCACACCATGAGGATCCATACTTAGGCCTTCT
 TCTGAAGATTGAGGACAATGATTATGATCCTAGTGCTCCAGATCTTGACAATAAAGTCACAGAA
 GAAAGTTCAAGCTCTGATTTACATCTGACTCTGAGGATCTTGCTGCTGCTATTGAGGATAACACG
 TCCCCTGGACATGTTAAACAGACTAAAAGCAGACAAAAAAGTAAAGTTGGTAAGAACCATCAATG
 GTTGATGAAGTGTCACTTTACGAGAGCCAGAACTTGAGGAAGAGGGTTTCACTCCAGTTTCTGGG
 AAAAGAAATGTGGAAAGGCTGGACTACAAAAGCTGTATGATGAGACATACAAAAGTGATACTAGT
 GAAGATGAAGAATGGACTGCCAGTGCCACTCCAAGTAGAAAAAGAACTCTGTGGCAAATGACT
 CCAGTGTCAATGAGTGGAAAGGCTCAAATAATTCTAGACATACTCCTGAAAGGAACACCCAACAA
 GATAAAGTTGAAAAACAAAATAATTCACCCACTAAATCACTTGAAGGCTGTCTGGAATCTGGCTCA
 AGGATAAAAAAGCCTAGGCTTCAACACGTCAAAGCTTGGAGACGTTGTAGTGCAGAGACTTAC
 AAATCTTTTTAAAGACAATCAGTATCCGGATCGAGGACAAAAGAAAGCTTGGCAGAAGAACTGGGG
 CTCACTTTTTTCCAGGTGACAAAATGGTTTGGCAATGCTCGATGGGGCTTTCGGCGTTCATCGCGT
 ATGGGAGCTAGTCCAGGCGAATATGCTTCGCCACAGGCCACTGGCAGCGGACCTGAAAATACTGGG
 GAGAGGGAACCGAATTAGCTTCCCAGGAAGTCGGTGAAGAAAAATTGAAAACCCCCAGTCTTAGA
 AAAAGGAAGCATTGTCTGAACCGCAGGCATCTGAAGCACCGCAGATTATTGTTCTTGGCTTAGCA
 GCATCTCCTGGCTCACCACGGGCTCATGCAGGCAATAAAAAGAAGACAGTGAAGAAGGAAATGA

SEC ID Nº: 200, secuencia de polipéptido traducida PHDf_HD de *Lotus japonicus*
 MISSTGELTSHNNSIEHMETEQPELSEKTPQIGSECLDGEQRLVLTSTSSVNNNSNQVSANVT
 ENSVFLPATPQHCHLEKSGQTVTSPPLEESTLQVATNVSNGKSENKQTSQDVQNELVEASDVI
 TGSFVADQTSIPAQVNTSSVNELLDIPSGDVAVEVGSDSLEKSKGCLVDSEQRLVLTNSIN
 DGNSNQVSANVTENSVIQLPAPPQHCFEKSGQTVKSSSLEESTLQVSTNLSNSKSENIQTSQD
 VQNELVETSDAVTGSVVADKTSIPAQVNTSSVNELLDIPSRDVAEEVGSDSSEKSKSTTPSQR
 HIGKSNKLSKKYILRSLGSSDRALRSRTKEKPKAPESSNSVDVNDGVKKEKRRKKKKTGGEE
 KKDQFYKIKAHLRYLLNRVSYEQNLI DAYSGDGWKGYSMDKLPDKEIQRAKSEILRRKLIKIRDLF
 QNLDSLCAEGKFPESLFDSEGEIDSEDI FCSI QTKELSTDNDI ILCDGACDRGFHQHCLDPPLLT
 EDIPPGDEGWLCPGCDCKDDCVDLLNDSLGLRSLSDTWEKVFPEAAAVTGNMMDQGLPSDDSDDD

FIGURA 7 (continuación)

DYNPNGKEDVEVEGGESSSESEYASASEKLEDSHHEDPYLGLPSEDSNDYDPSAPDLDNKVTE
 ESSSSDFTSDSEDLAAAIEDNTSPGHVKQTKSRQKSKVGNPSMVDEVSSLREPELEEEGFTPVSG
 KRNVERLDYKKLYDETYKSDTSEDEEWTASATPSRKKKLCGKMPVSSDGKASNNSRHTPERNTQQ
 DKVENTNNSPTKSLEGCSGRDKKPRSSTRQRLGDVVVQRLHKSFKDNQYPDRATKESLAEELG
 LTFQVQDKWFGNARWGFRRSSRMGASPGEYASPOATGSGPENTGERERELASQEVGEEKLKTSPSR
 KRKHLSEPOASEAPQIIIVLGLAASPGSPRAHAGNKKKTVKRK

SEC ID N°: 201, secuencia de ácido nucleico PHDf_HD_HAZ1 de *Oryza sativa* AB081340

ATGGATAAGACAACACTACTTCTGATCTTGT'TTTGGACAACGACAACATTGGGAGTAATGCTGGTCT
 GCACAGGAGCCTCTTACTACCAATGGTAAGACTAGTGGGGTCAGAAATAGATACAAACAACTGTT
 AAAAGGGGAAGGAAAGGCTCCCAAATATCACCAAGTAAAAATATCCCCTGAGATCTTCACATAGT
 AATGTCAGGGTGCTTCGCTCTGCCTCAAAAAAGAAGAATGAGACACCTATTGTGCCACAAAATGAT
 AATACTGCTGTTCAACGAGTTGCGAAGAAAAGGAAAAGAAGCAAACCCTTGAGACCTGCACCTAGC
 AGGGTGCTTCGCTCTACCTCGGAAAAGAAGAATAAGGCACATAATGAGCTTCTAAATGATGGTGCT
 GGTGTTCAACCAGCCGAAAAGAAAAGAAAGGTAGGCAGACCCCAAAAGGAGGAACTCCCAAGGAC
 GATTACCTCATGATCCGAAAGCGAGTTAGATATGTTTTGAATCGAATGAACTATGAACAAAGTTTA
 ATCAAGCCTATGCCAGTGAAGGTTGGAAAAGGTCAAAGCTTGGAAAAATAAGACCTGAGAAGGAG
 CTTGAACGAGCCAAGGTGGAAATTCTGCGATGCAAGTCAAGAATACGGGAAGCTTTTCGGAATTTG
 GATTCCCTCCTATCTGAGGGAAAGCTGGATGAATCTATGTTTGATTCTGCTGGAGAAATATCTAGT
 GAAGATATTTCTGTGCTGCATGTGGTTCAAAGGATGTTACATTAATAAATGACATAATTCTTTGT
 GATGGAATCTGCGACAGAGGGTTCACCAGTATTGTTTTAAACCCTCCTTTGCTTGCTGAAGATATT
 CCACAGGGGGATGAAGGATGGCTTTGCCCCTGCATGTGATTGCAAAATTGACTGCATAGATGTATTA
 AATGAACTCCAAGGGGTCAAACCTCTATCCATGACTCTTGGGAGAAAGTTTTCTCTGAGGCAGCA
 TCCTTTTTGAATGGTTCTAAGCAAATTGATGCGTCCGATCTTCCATCAGATGATTCAGCAGACAAT
 GATTATGACCCTACTTTGGCTCAAGGGCACAAAGGTGGATGAAGAAAATCTTCTGGAGAAGATGGT
 GGTGAAGGATTAGATTCTGATGACTCATCTTCTGAGGATTCTGAATCTTCTGAGAAGGAAAAGTCT
 AAAACTTCACAGAATGGAAGGACAGTTGATGATCTTGGTTTGCCTTCTGAGGATTCAGAGGATGGT
 GACTTTGATCCAGCAGGTCCAGATTCCGACAAAGAACAAAATGATGAATCAAACTCAGATCAGTCA
 GATGAATCAGATTTTACATCTGACTCTGATGATTTTTGTGCTGAGATTGCTAAATCCTGTGGTCAG
 GATGAAATCTCAGGCCCTTCATCATCAAAAACAGAACAGTCGATCGCACCCGATGGAAGTGGTTTT
 GATGGTGAGCCTAATGCAGAAAATTCGAATCTTGCAATTTATGGAGACAGAGCTAGAGCAAGACATG
 GTGCTACCAATTTCTAGTAAACGACAAGTTGAACGTTTGGATTACAAAAACTGTATAATGAGGCT
 TATGGTAAAGCATCATCTGATTCAAGTGACGATGAAGAGTGGTATGGAAATAGTACACCCGAAAAA
 GGTAATTTAGAAAGACAGTGAAAACAGATTTCGCTTGCTGAATCACCTCAGGGTGGAAAAGGATTCTCT
 AGAAGAGCACCAGTTAGGTACCATAATAATGAACATACTCCACAGAATGTAAGGCCCGGTGGTTCA
 GTAAGCGACCAGCAAACCTGAAGTACTTTGCTCCAATAGCAATGGTAGCACAGCCAAAAACAGACAT
 TTTGGTCTGCAATTAATCAGAACTGAAGGCACATTTCAAGGAAGATCCATACCCTAGTCGTGCA
 ACAAAAGAAAATCTGGCACAAGAGTTAGGCCTCACATTTAATCAGGTCATAAATGGTTTTCCAGT
 ACTCGTCATTACGCAAGAGTGGCAGCCACAAAAAGGAGAACAACATTGAAAATCATACGGCTGAG
 AATAATAACAATACCAACTGTTGATAGCATACACTGAGAGGATCCAATGATATTTGTTAGTGTA
 GATAGAAATGATATGGTCTCCGAGGAAAGGACAGGACAAAGTAATCTCAATGAAGGTACTCCATTA
 AGATCAGATACCAGCTGTGGGCAGAGTGTGGCTGTGACTCCTATGGTACACCCAGAAAACCAGGGC
 AATGATTCCCTCAAGTAATGTTAGAACACCAAATGCTAAGTCTGCGGAGAAATTTGATCCCTGGACTG
 GAAAATTCAGATGAGGCTAGGCCGAAAGCGGTACAACGGGAACTGAGAAAGATGAAGACTGGCAGG
 TGA

FIGURA 7 (continuación)

SEC ID N°: 202, secuencia de polipéptido traducida PHDf_HD_HAZ1 de *Oryza sativa*

MDKTTTSDLVLDNDNIGSNAGSAQEPLTTNGKTSGVRNRYKQTVKRGRKGSQISPSKTYPLRSSHS
 NVRVLRASAKKNETPIVPTNDNTAVQORVAKKRKRKPLRPAPSRVLRSTSEKKNKAHNE LLNDGA
 GVQPAEKRRKVRPFPKGGTPKDDYLMIRKRVRYVLRNRMNYESLIQAYASEGWKQSLEKIRPEKE
 LERAKVEILRCKSRIREAFRNLD SLLSEGKLDSEMFDSAGEISSEDI FCAACGSKDVT LKNDI ILC
 DGICDRGFHQYCLNPPLLAEDI PQGDEGWLC PACDCKIDCIDVLNELQGVKLSIHDSWEKVFPEAA
 SFLNGSKQIDASDLPSDDADNDYDPTLAQGHKVDEEKSSGEDGGEGLDSDSSSEDESSEKEKS
 KTSQNGRTVDDLGLPSEDES DGFDPAGPDS DKEQND ESNSDQSD ESDFTSDSDDFCAEIAKSCGQ
 DEI SGPSSSQIRTVDRTD GSGFDGEPNAENSNLAFMETELEQDMVLP ISSKRQVERLDYKKLYNEA
 YGKASSDSSDEEWYGNSTPEKGNLEDSETSLAES PQGGKGF SRRAPVRYHNNEHTPQNVRP GGS
 VSDQQTEVLC SNSNGSTAKNRHFGPA INOKLKAHF KEDPYP SRATKENLAQELGLTFNQVTKWFSS
 TRHYARVAATKKENNIENHTAENNNNTNTVDSIQLRGSNDI VSVDRNDMVSEERTGQSNLNEGTP L
 RSDTSCGQSVAVTPMVHPENQGNDS SSVRTPNAKSAEKLI PGLENSDEARRKAVQRELRKMK TGR

SEC ID N°: 203, secuencia de ácido nucleico PHDf_HD_PRHP de *Petroselinum crispum* L21975

ATGGAAGAGATCTCAGACCCTAAACCAAATGCATTAGAGCAAGTACTTCCTACTGTCCCAAATGGT
 AAATGTACTGCACCGGTGCAAATGGAAAGTTTGGCTGTTGATGTTCAAAAAGTATCAGGTGAAGCA
 AAGGTGAGGATATGTTCCCTGTTGGTGCAGATAGTACGTTCTCCAGAAGATTTAACTAACTTGTG
 CCTTGTAATGATTTTGCAGAAGATATAAAAATTGTTTGATTCTGATCCCATGCAACAAGAAGCTGAA
 AGCTCCATTGGAATACCATTGATCCCTAAACAAGTTACAATGTCCATAACCATGACCATGAATCT
 GGTCTGAAATGGTCAGTAACGAAGTAATGCAAGAAAACCATGTAATTGCCACAGAAAATACTTAC
 CAGAAATCTGATTTTGACCGTATAAATATGGGACAAAAAGAAACCATGCCAGAAGAAGTGATCCAT
 AAGTCTTTCTTGGAATCCTCAACTTCCTCGATAGATATACTATTAAATAATCATAACAGTTATCAA
 TCAGGACTGCCACCAGAGAATGCAGTAACAGATTGCAAGCAAGTCCAATTGGGACACCCGAAGTGAT
 GATGCTATAAAGAATTTCTGGTCTTGTGGAATTGGTAATAGGTCAGAAAAATGTTGCCAAGAGTCCCT
 AGCCAGTTGGTAGAAACGGGGAAAAGAGGTCGTGGTCCGCCAGGAAAGTACAACTGGCCTTGAG
 CAATTGGTAATAGGTCAGAAAAGTCTGCCAAGAGCTCTAGCCAGTTAGGTGACACAGGGAAAAGA
 TCTCGTGGTCCGGCCAGAAAAGTACAAAATTTCTCAACTTCTTTCTTGGAGAATATAAATATGGAA
 CAAAAGAAACAATACCAGAACAAGTGACCCAAAATTTCTATCTTGGAGTCTTGGACTATCCCGACA
 GACAATCAGTCAAGAACTTATAACAGTGATCAATCAGAAGTCCACCAGAGAATGCAGCAAAAAAT
 TGCAATCAGCCCAATTTGGACATCAAAGTGATGATACTACAAAGATTTCTGGATTCAAGGAATTG
 GTAATAGGTCAGGAAACAGTTGCAAAGAGTCCTAGCCAGTTGGTAGACGCAGGAAAAAGAGGTCGT
 GGTCCGCCAGAAAAGTACAACTGGTCTTGAGCAATTGGTACCAGTTCAGGAAACTGCGGCCAAG
 AGCTCTAGCCAGTTGGGAGACACAGGGAAAAGATCTCGCGGCCGCCAGGAAAGTACAAGATTCT
 CCAACTTCTTTAGGGGGTAATGTTAAAGTGGTACCAGAGAAAGGAAAAGATTCTCAGGAAATGTCG
 GTAAACAGCAGCAGATCATTGCGCTCAAGGTCACAAGAGAAATCTATCGAACCCGATGTCAATAAT
 ATTGTGGCAGATGAAGGTGCTGACAGAGAAAAACCAAGGAAGAAGAGGAAAAGCGAATGGAGGAA
 AATAGGTTGATGAATTTCTGTAGAATAAGGACACACCTAAGGTA CTACTACACAGAATAAAATAC
 GAGAAGAATTTCTTGATGCTTACTCTGGGGAGGGATGGAAGGGACAAAGCCTAGACAAAAATAAAA
 CCTGAGAAGGAGCTTAAACGAGCAAAAAGCTGAAATTTTTGGACGCAAATTGAAAATTAGAGATCTA
 TTTACAGCGCTTAGATTTAGCACGTTCCGAGGGAAAGGCTTCCAGAAATTTCTATTTGATTCTCGAGGA
 GAAATTGACAGCGAGGATATATTTCTGCGCGAAATGTGGATCCAAGGATGTGACACTCAGTAATGAT
 ATCATCCTTTGCGATGGTGATGTGATCGTGGATTTACCAATTTGTTTGGACCCACCATTGCTG
 AAAGAATATATTTCTCCTGACGACGAGGGTTGGCTTTGCCCGGGATGCGAGTGCAAAATTGACTGC
 ATCAAGTTACTTAATGATTCCCAAGAGACAAATATTTTACTCGGTGACAGCTGGGAGAAAAGTTTTT
 GCTGAGGAGGCAGCAGCAGCAGCTTCCGGAAAAAATCTAGATGATAATTCTGGACTGCCTTCAGAT

FIGURA 7 (continuación)

GATTCCGAGGATGATGATTATGATCCTGGGGGCCCTGATTTGGATGAGAAGGTGCAGGGAGATGAT
 TCGAGTACGGATGAATCTGATTATCAGTCGGAATCAGATGATATGCAAGTTATACGTCAAAAAAAT
 TCGCGCGGGCTTCCTTCAGATGATTCCGAAGATGATGAATACGATCCTAGTGGCTTAGTTACTGAT
 CAGATGTATAAGGACAGCTCATGTTCTGATTTTACATCTGATTCTGAGGATTTTACAGGTGTGTTT
 GACGATTATAAGGATACTGGTAAAGCTCAAGGGCCACTGGCATCGACCCAGATCACGTCAGGAAC
 AATGAGGAAGGGTGTGGGCATCCTGAGCAAGGTGATACTGCCCCCTGTATCCGAGAAGACAAGTA
 GAGAGTTTGGACTACAAAAAGCTAAATGATATTGAGTTCTCTAAAATGTGTGACATCCTTGATATT
 CTTTCAAGTCAACTTGATGTAATTATTTGTTACTGGTAACCAGGAAGAGTATGGGAACACATCTTCT
 GATTCAAGTGATGAAGATTATATGGTTACTTCTTCACCAGATAAAAATAATTCTGATAAAGAAGCT
 ACAGCAATGGAGCGTGGAAAGAGAGCGGTGATTTGGAGCTAGACC AAAAAGCGAGAGAATCCACT
 CATAATAGGAGATATATAAAGAAGTTTGGCTGTTGAAGGCACAGATTCTTTTCTTTCTAGATCGTGT
 GAAGACTCTGCAGCACCTGTTGCTGGTAGTAAAAGTACTTCAAAAACATTGCACGGGGAACATGCA
 ACTCAGAGACTCCTCCAATCCTTCAAGGAAAATCAATACCCTCAACGAGCTGTGAAAGAGAGTTTG
 GCTGCAGAATTAGCACTTTCAGTCCGACAGGTTAGCAACTGGTTTAATAATAGACGTTGGAGCTTT
 CGCCACTCATCACGTATTGGATCGGATGTAGCTAAATTTGATTCAAACGATACTCCCGCTCAGAAA
 AGCATTGACATGTCTGGACCCAGCTTGAAATCGGTCTTGGACAGTGCTACTTACAGTGA AATTGAA
 AAGAAGGAACAAGATACGGCAAGCCTTGGTTTTGACAGAAGGTTGTGACAGATACATGACGCTGAAT
 ATGGTTGCAGATGAAGGCAATGTGCACACTCCTTGTATTGCAGAACTAGGGGAAGAAAAACCGAA
 GTCGGTATAAAGCCACAACAGAACCCTATAA

SEC ID Nº: 204, secuencia de polipéptido traducida PHDf_HD_PRHP de *Petroselinum crispum*
 MEEISDPKPNALEQVLPVFNPKCTAPVQMESLAVDVQKVSGEAKVTRICSCWCEIVRSPEDLTKLV
 PCNDFAEEDIKLFDS DPMQOEAESSIGIPLIPKQVTM SHNH DHESGSEMVSNEVMQENHVIATENTY
 QKSDFDRIINMGQKEMPEEVIHKSFLSSTSSIDILLNNHNSYQSGLPPENAVTDCKQVQLGHRSD
 DAIKNSGLVELVIGQKNVAKSPSQLVETGKRGRGRPRKVQTGLEQLVIGQKTAAKSSSQLGDTGKR
 SRGRPRKVQNSPTSFLENINMEQKETIPEQVTQNSILES LTIPTDNQSRTYNSDQSELPPENA AKN
 CNHAQFGHQSDDTTKISGFKELVIGQETVAKSPSQLVDAGKRGRGRPRKVQTGLEQLVPVQETA AAK
 SSSQLGDTGKR SRGRPRKVQDSPTS LGGNVQVPEKGKDSQELSVNSSRSLRSRSQEKSI EPDVNN
 IVADEGADREKPRKKRKRMEENRVDEF CRIRTHLRYLLHRIKYEKNFLDAYSGEGWKGQSLDKIK
 PEKELKRAKAEIFGRKLRDLFQRLDLARSEGRLPEILFDSRGEIDSEDFCAKCGSKDVTLSND
 IILCDGACDRGFHQFCLDPPLLKEYIPPDDEGWLCPGCECKIDCIKLLNDSQETNILLGDSWEKVF
 AEEAAAAASGKNLDDNSGLPSDDSEDDDYDPGGPDLDEKVQGDDSDTDES DYQSESDDMQVIRQKN
 SRGLPSDDSEDEYDPSGLVTDQMYKSSCSDFTSDSEDFTG VFDYKDTGKAQGPLASTPDHVRN
 NEEGCGHPEQGD TAPLYPRRQVESLDYKLLNDIEFSKMCDILDILSSQLDVIICTGNQE EYGNTSS
 DSSDEDYMTSSPDKNNSDKEATAMERGESGDL ELDQKARESTHNRRYIKKFAVEGTD SFLSRSC
 EDSAAPVAGSKSTSKTLHGEHATQRL LQSFKENQYPQRAVKESLAEELALSVRQVSNWFNNRRWSF
 RHSSRIGSDVAKFDSNDTPRQKSIDMSGPSLKSVLDSATYSEIEKKEQDTASLGLTEGCDRYMTLN
 MVADEGNVHTPCIAETREEKTEVGIK PQQNPL

SEC ID Nº: 205, secuencia de ácido nucleico PHDf_HD_HOX1a de *Zea mays* X67561
 ATGGAAAAAATATAGCTCACTGTCCTGTTGAGGGCAATGGTGAGATTGAAAATGGTGCGAGCTCT
 AGTCAGAACTCTGAGTCTCTTGAGCATTGAGTTTTGCTGTCTACATCTCAAACGATGCCAAATAAC
 TTGGGGATAAGGAAAACTATAAAAAGGGCAGCAAACAGAGGTAAAAAAGGTTCCCAAGGACTAACA
 GGCCAAGCATATACTTTGATGTCATCTAATAGTGATGTCAGGGTGCTTCGCTCCACATCAAGTTCA
 AAGACAACATCTACTGAGCATGTT CAGGCTCCAGTACAACCGGCTGCCAAGAGAAGAAAAATGAGC
 AGAGCCTCAAACAAAAGTTGACTGATGAGTTTTCACAAATTCGTAAACGAGTTAGATATATTTTG

FIGURA 7 (continuación)

AACCGAATGAATTATGAGCAAAGCCTAATTGAAGCTTATGCTAGCGAAGGCTGGAAAAATCAAAGT
 TTGGATAAGATAAGACCTGAGAAGGAGCTTGAACGAGCCAAATCAGAAATCTTACGATGCAAATTG
 AGAATACGAGAAGTCTTCCGGAATATTGATTCTCTTCTCTCCAAGGGGAAAAATTGACGAAACTTTA
 TTTGATTCTGAAGGAGAAATATCTTTCGGAAGATATCTTTTGTTCACCTTGTGGTTCGAATGACGCT
 ACATTGGGTAATGATATCATTCTCTGCGATGGAGCTTGTGACAGAGGGTTCACCAGAACTGTTTA
 AATCCTCCCCTGCGTACTGAAGATATCCCCATGGGAGATGAAGGATGGCTCTGCCCAGCATGTGAT
 TGCAAGATAGATTGTATAGATCTAATAAATGAACTCCATGGGAGTAACATCTCCATTGAGGACTCT
 TGGGAGAAAGTTTTTCCAGATGCAGCTGCTATGGCAAATGACTCTAAACAAGATGACGCATTTGAT
 CTCCATCAGATGACTCAGACGACAATGACTTTGACCCCAATATGCCTGAAGAGCATGTGGTTGGT
 AAGGATGAAGAATCGTCTGAAGAGGATGAGGATGGAGGATCAGACTCTGATGACTCAGACTTTTTG
 ACATGTTCTGATGATTCGGAACCTTTGATAGACAAGAAGGTTGATGACCTCAGATTACCTTCTGAA
 GATTCAGAGGATGATGACTATGATCCAGCAGGTCTGATTCAGATAAAGATGTCGAAAAGAAGTCA
 AGTCTGATGAATCTGACTTCACATCTGACTCAGATGATTTTTGTAAGGAGATTTCAAATCTGGA
 CATGATGAAGTTTCATCACCTCTATTACCCGATGCCAAGGTGGGTGATATGGAAAAGATTACTGCT
 CAAGCTAAAACAACAAGTTCTGCTGACGACCCTATGGAGACTGAAATAGACCAGGGTGTGGTTTTA
 CCAGATTCAGGAGACGGCAAGCTGAACGGTTGGACTACAAAAGTTGTATGATGAGGCATATGGT
 GAAGCATCATCTGATTCAGTGATGATGAAGAATGGTCTGGGAAGAACACGCCAATAATAAAAAGC
 AATGAAGAGGGTGAAGCCAATTCTCTGCGGGAAAAGGCTCCAGAGTTGTGCACCATAATGATGAA
 CTGACTACACAAAGCACTAAAAAAGCTTGCATTCTATTCATGGTTCAGTAGATGAGAAACCTGGA
 GACCTTACTTCTAATGGCAGCAACAGCACAGCCAGGAAAGGACATTTTTGGCCAGTAATTAATCAG
 AAGCTACATGAACATTTAAGACACAACCATACCCTAGTCGTTCTGTTAAAGAAAGTCTGGCAGAA
 GAACTAGGGTTAACATTTTCGTCAGGTTAACAAATGGTTTGAAGTGGCGTCATTCGCCAAGGGTA
 GCTTCTTCCAGGAAAGGCATCAGTCTAGATAAGCATAGCCCCAAAATACTAATAGTCAAGTGACA
 GCTAGTATGGAACCTAAAGAACCTGAGGGGACTGTGGTGGGAAGAATCTAATGTATGCCTAAATGGG
 GGTACTACCATCTCTAAAGAAGCAGTGTCTTCAAAGTTGGATCAAGAACCCTGGAAGTGATGTG
 GGAGGGTCAAAGTTGATTCTGCTGAGGATCAGAATCCGGGGCCTGATCTTGCAGAGAAGGCAAGG
 CAAAAGGCAATACAGCAGGAGTTGAGGAAGAAAAAATGGGACGATGA

SEC ID N°: 206, secuencia de polipéptido traducida PHDf_HD_HOX1a de *Zea mays*

MEKNIAHCPVEGNGEIENGASSSQNPESLEHSVLLSTSQTMPNNLGIRKKNYKRAANRGRKKSQGLT
 GQAYTLMSSNSDVRVLRSTSSSKTTSTEHVQAPVQPAKRKRKMSRASNKSSTDEFSQIRKRVRYIL
 NRMNYEQSLIEAYASEGWKNQSLDKIRPEKELERAKSEILRCKLRIREVFRNIDSLLSKGKIDETL
 FDSEGEISCEDIFCSTCGSN DATLGNDI ILCDGACDRGFHQNCLNPPLRTEDI PMGDEGWLCPACD
 CKIDCIDLINELHGSNISIEDSWEKVFPDAAAMANDSKQDDAFDLPSSDSDNDNDFPNMPEEHVVG
 KDEESSEDEDGGSDSDDSDFLTCSDDEPLIDKKVDDLRLPSEDSDDDYDPAGPDSDKDVEKKS
 SSDESDFTSDDDFCKEISKSGHDEVSSPLLPDAKVGDMEKITAQAKTSSADDPMETEIDQGVVL
 PDSRRRQAEERLDYKKLYDEAYGEASSDSSDDEEWSGKNTPIIKSNEEGEANS PAGKGSRVVHND
 LTTQSTKKSLSHSIHGSVDEKPGDLTSNGSNSTARKGHFGPVINQKLHEHFKTQPYPSRSVKESLAE
 ELGLTFRQVKNWFETRHSARVASSRKGISLDKHSQNTNSQVTASMEPKPEGTVVVEESNVCLNG
 GTTISKEAVSSKVGSRTPGSDVGGSKVDSAEDQNPGLAELAKARQKAIQQELRKKMGR

SEC ID N°: 207, secuencia de ácido nucleico de PHDf_HD_HOX1b de *Zea mays* X92428

ATGGAAAAAAGACAGTTCACGTTCTGGCTGAGGGTAATGGTGAGATTGAAAATGGTACGAGCTCT
 AGTCAGAACTCTGAGACCTTGAGCATCCAGTTTTGCTGCTGCATCTCAAACGTGCCAAATAAC
 TTGGGGATAAGGAAAACTATAAAAGGGCAGCAAACAGAGGTAAAAGGGTCCCAAGGACTAACA
 GGTCAGGCATATACATTGAGGTCATCTGATATTGATGTCAGGGTGCTTCGCTCTACATCAGTTCA

FIGURA 7 (continuación)

AAGACAACATCTACTGAGCATGTGCAGCCTCCAGGACAACCAGCTGCCAAGAGAAGAAAAAGGAGC
 AGAGCCTCAAACAAAAATTTCGACAGATGAGTTCTCACAAATTCGTAAACGGGTTAGATACATTTTG
 AACCGAATGAATTATGAGCAAAGCCTAATTGAAGCTTATGCTAGTGAAGGCTGGAAAAATCAAAGT
 TTGGATAAGATAAGACCTGAGAAGGAGCTTGAACGAGCCAAATCAGAAATCTTACGATGCAAATTG
 AGAATACGGGAAGTCTTCCAAAATATTGATTCTCTTCTGTCCAAAGGGAAAATTGACGAATCTTTA
 TTTGATTCTGAAGGAGAAATATCTTGTGAAGATATCTTTTGTGCTACTTGTGGTTTCGAAAAATGCT
 ACATTGGGTAATGATATCATTCTCTGTGATGGAGCTTGTGACAGAGGGTTCCACCAGAACTGTCTA
 AATCCTCCTCTGCGTACTGAAGATATCCCATGGGAGATGAAGGATGGCTCTGCCAGCATGTGAT
 TGCAAGTTAGACTGTATAGATCTAATAAATGAACTCCAGGGCAGTGACATCTCCATTGAGGACCCT
 TGGGAGAAAGTTTTTCCAGAGGCAGCTGCTATGACCAATGGCTCTAAGCAAGATGAACGATTTGAT
 CTTCCATCTGATGATTCGGATGACAATGATTTTGACCCCAATATGCCTGAAGAGCATGTGGCTAGT
 AAGGAAGAAGGATCTTCTGAAGAGGAAGAGGATGATGACGGAGGATCAGACTCTGACGACTCTGAC
 TTCTTAACTGTTCTGATGATTTGGAACCTTTGATAGATAAAAAGAAGGTTGATGACCTGGGATTA
 TCTTCTGAAGATTCAGAGGATGATGACTATGATCCAGCAGGTCTGATTCGGATAAAGATGTCGAA
 AAGAAGTCAAATCTGATGAGTCTGATTTACGCTGACTCGGATGATTTTTGTAAGGAGATTA
 AAATCTGGTGGACATGATGAAGTTTCATCGCCTCCATTACCCGATGTCAAGGTGGGTGATATGGAA
 AAGAACACTGCTCAGTCTAACACAACAAGTTCTGCTGACGACCCTATGGAGACTGAAATAGACCAG
 AGTGTGGTTTTTACCAGTATCAAGGAGACGTCAAGCTGAACGGTTGGACTACAAAAGGCTGTATGAT
 GAGGCATATGGTGAAGCATCATCTGATTCTAGTGTGAAGAAGAATGGTCTGGGAAGAACACACCA
 ATAAAAAGCAATGAAGAGGGTGAAGTTGGTTCTCTGCAGGAAAAGGCTCTAGAGTTGCGCACCAT
 AATGAACTGACTACACAAAACACTAAAGAAAGCTTGCATTCTCTTCATGGTTCAGTAGATGAGAAA
 CATGGAGATCTTACTTCTAATGGCAGCAATATCAAAGATAGGAAAGGACATTTTGGTCCAGTGATT
 AGTCAGAAGCTACATGAACATTTTAAAGACACAACCATACCCTAGTCGTTCTCTTAAAGAAAGCTTG
 GCAGAAGAAGTGGGCTAACATTTTCAATCAGGTCAACAGATGGTTTTGAAAATAGGCGTCATTTTGCA
 AGGTTAGCTTCTTCCAGGAAAGGCATCAGTCCAGATAAGCATAGCCCCAAAATACTAATAGCCCA
 GTGACACCTAGTATGCAACCTAAAGAACCAGAGGGGACTGTGATGGAAGAATCTAATGTATGCATA
 AACAGGGATGCTACCATCTCTAAAAAAGCGGTTTTCTTCAAAGTCCGGATCAAGAAAGAACCTTAGC
 AAGAGTTCCCTAAAAGCCGGGTCAAAGTTGA

SEC ID N°: 208, secuencia de polipéptido traducida PHDf_HD_HOX1b de *Zea mays*

MEKKTVHVLAEENGIEIENGTSSSQNPETLEHPVLLSASQTVPNNLGIRKKNYKRAANRGKKSQGLT
 GQAYTLRSSDIDVRVLRSTSGSKTSTEHVQPPGQPAAKRRKRSRASNKNSTDEFSQIRKRVRYIL
 NRMNYEQSLIEAYASEGWKNQSLDKIRPEKELERAKSEILRCKLRIREVFQNIIDSLLSKGIKIDESL
 FDSEGEISCEDIFCFCATCGSKNATLGNDIILCDGACDRGFHQNCLNPPLRTEDI PMGDEGWLCPACD
 CKLDCIDLINELQGSISIEDPWKEKVFPEAAAMTNGSKQDERFDLPSDDSDNDNFENMPEEHVAS
 KEEGSSEEEEDDDGGSDDSDDFLTCSDDLEPLIDKKKVDLGLSSEDEDDDYDPAGPDSKDKVE
 KKSNSDESDFTSDDDFCKEIKKSGGHDEVSSPPLPDVKVGDMEKNTAQSNNTSSADDPMETEIDQ
 SVVLPVSRRRQAERLDYKRLYDEAYGEASSDSSDEEWSGKNTPIKSNEEGEVGSPAGKGSRVAAH
 NELTTQNTKESLHSLHGSVDEKHGDLTNGSNIKDRKGFHGPVISOQLHEHFKTQPYPSRSLKESL
 AEELGLTFHQVNRWFENRRHFARLASSRKGISPDKHSQNTNSPVTSPMQPKEPEGTVMESNVC I
 NRDATISKKAVSSKVGSRKNLSKSSPKSRVKS

SEC ID N°: 209, secuencia de ácido nucleico PHDf_HD_HOX2a de *Zea mays* X89760

ATGGGAAGGAAAGGGTCAGAAGTATCACCAGCAAAGGATCCCTTGAGGTCTGCACACAGTAGT
 GGTAGAGTTCTTCGCTCTGCCTTAGCCAATAATAAAGGACAGTGAGCCTCTGAATGCGGCTGCT
 GCTACCCAAGCAGCTGTGAAAAAAGAAGTGGAACCAGCTAGATAGTCCTAATAGCACTGTCAGG

FIGURA 7 (continuación)

C T A C T T C G T T C T A C G T C A A A A A T A A G G A C G A G G C A C A T A G C G G G C C T C T T A A T G A C A G G A T C A T T
 G A T A A G C C A G C T G C A A A T A A A A G A A A A C A T G C C A A T C C C T C A C A G G T T G G A A G T C C C A G C A G C A G T
 G T C A G G G T G C T T C G C T C T G C A A T A A A A A A T A A A G A T A T G G C G C G C A G T G A G C C T C A A A G T G A C A G G
 A C C G C T G G T G A G G C A G C T A C A A A T A A A A G A A A A A C G C C A A T T C C T C A A A G G T G G G G A G T C T G A G C
 T C T G G T C T C A G T G T G C T T C G G T C T G C T T C A A A A G C T A A G G A T G A G A C A T G T A T C A A A C C T T T A A A T
 G A T A G T C C T G C T G T T G A A C C A G C T T C A A G T A A A A G A A A A T G C A C C A G T C C G T C A A A G G T G G G A A G T
 C C C A G C A A C A G T G T C A G G G T G C T T C G A T C T A C T T C A A A A T A T A G A A T G A G A C A T G T A C C G A G C C T
 T T A A A T G A T A G T G G T G C T G T T G A A C C A G C T G C A A A T A G A A G G A A A G G T G T C A G T C G T T C A A A G G A G
 G G A A G T C C C A A T A G C A G T G T C A G G G T G C T T C G C T C T G C C T C A A A T A A T A A G A T T G A T G T A T C T A T T
 G A G C C A C T G A A T A A T A G T A C T A C T G G T C A A C C G G C T G C G A A A A G A A G A A A A A T G G C A G T T C C T T T
 A A G A G C A G T G C C A G G G T G C T T T G C T C T A C C T C A G A A A G G A A A A T G A G G C A T C T A G T G A G C C T T T A
 A A T G A T A G T A C T G C T G C T C A G C C A G C T G C G A G G A A A A G A A A A G G T G G T G C A A T G C C A A A G A C T G A T
 A A C C C A A A G A T T G G T G T C A G G G T T C T C G C T C T G C C T C A G G A A A G A A G A A T G A G A C A T G T A C T G A G
 C C T G T A A A T C G G A G T A C T T C T C G T G A C C A A C T G T G A C A A A A A A A G A A A T T G C A A G C C T T C A A A G
 G A T C G C G G T C C C A A G A A G G A C T A C T T A A A A T T T G T C A A A G A A T T A A A T A T A T T T T G A A T C G A A T G
 A A G T A T G A A C A A G C C T T T A T T C A G G C T T A T G C A A G T G A A G G T T G G A A A G G C C A A G T T T G G A A A A
 A T A A G A C C C G A G A A G G A G C T T G A A C G A G C C A A G G C A G A A A T C C T T C A A T G C A A G T T A A G A A T C C G G
 G A A G C T T T T C G A A A T A T G G A C T C T C T T T A T C G A A G G G G A A G C T T G A A G A A T C T T T G T T T G A T T C A
 G C A G G A C A A A T T T C A A G C G A A G A T A T A T T T T G T G C C A T A T G T G G C T C A A A G G A T G T T A C T T C A C A A
 A A T G A C A T C A T T C T T T G T G A T G G A G C C T G T G A C A G A G G A T T C C A C C A G A A T T G T T T G A G C C C T C C T
 T T G C T A A C T G A A G A A A T T C C T C C G G G G G A T G A A G G A T G G C T T T G C C C T G C A T G T G T G T G C A A A G C A
 G A C T A T A T A G A T G C G T T A A A T G A A C T T C A A G G A A G C A A A C T C T C C A T T C A T G A C T C A T G G G T G A A A
 G T T T T C C C T G A G G C A G C T T C A A C T G C C A A T G G T T C T A A A C A A G T T G A T G C T T C T G A T C T G C T G C C A
 G A T C A T A T A A A G G A C A G T G C T A A C C T T G C A T T G G T T G G A A C A C A C A T G G T G A A T G A A A T C A G G T T T
 T C T G A A G A G G A T G A T A G C A A A G C G G A T G A C C T C G G C T T G C C T T C G G A G G A C T C A G G G G A T G G T G A C
 T T C G A T C C A G C A G G T C C A G A T T C T A G T G A A G A C C A A A A T G A T G G A T T G A A C T C A G A A G A G T C A G A T
 T T C A C A T C T G A C T C T G A T T A T T T T T G T G C T G A G A T T G C T A A A T C T T G C G G C C A G G A T G A A G T C T C G
 G C A T C T C C A T T A T C A A A T G T G A T A A A C C G T A C T G A T A G A A T G A A A C T C A G G G C C T A C A G T A G A C G C
 T C T A A T G A A G A A A C C A C A A T C A T G C A T T T A T G G A C A T G G A G C T A G A G C A A G A C A T A G T G C T A C C A
 G T T C C A A G C A G G C G G C A G G T T G A A C G G T T G G A T T A C A A A A A C T A T A C G A T G A G G C A T A T G A G G A A
 G A A T C A T C T A A T T C A A G T G A T G A A G A A G A G T G G T C T G G A A A A G A A C T C T T A G A A G G C A G T G G A A C A
 G A T T C A C T T G A T G A G C C T T C G T C C T G T A A A A G G C T C T C T A G A A G A G C A C C A G C A G G G C A G C A G A A T
 A A T G A A C T T A C C A C A G A G C C G G C C T C A T G G T T C T A A C A G T G A G C A G A A A A C T G A A G T T C T T C G C
 T C C A A T G G C A G T A G T A G C A C A G G A C G G A A G T A T G A T C C T G A A G T T A C C C A G A A A T T G A A G G T G C A T
 T T T G A G A A A G A T C C A T A T C C T A G C C G T G A A A C A A A G A A A A T C T A T C A G A A G A A C T A G G C T T G A C A
 T T T A A T C A G G T C T C C A A A T G G T T C T A G T A C T C G T C A C T A T T C A A G G G T T G C T T C T G C C A A A A A A
 G A A A T G C A T C C T G A C A G C C A T A C T T G T G A G A A T A A T G A T G A G A C A A C T G T A G A C A G C A T G C A A G C A
 A A A C A A C C C A A T G C T G G G G T G A T G G A A A A G T T G A C G G G G G A T A G A A A T G G C A T T G T T C C T G A G A A A
 C C G A T G G T G C A G G A C A A T C T T A A C C A A G G A G C G A T G G A A A A G T T A A C T A G G G A T A A A A A T G A C A T T
 G T T C C T G A G A A A C C G G T G A T G C A G A G C A G T C T T A A C C A A T G T A A T A A T G A A G A T A T A C C G C T T T C T
 G G A A C A G A A A T T G A G A T G G A G T C T T A C G A A C A A G A A T C A T C T G A C T C A A G T G A T G A A G A G T G G T C C
 A T A C T T A G C A C A C C A C G A A A A C A A G A T T A C A A G A C C A T G G A A A A T C A T C A G T T A C T G A A T C A C T T
 C G T T C T G C A A A A C G G T G T T C C A G A A G A G C A C C A G C T A G G G A G C A G A A T A A T G A A C G T A T T C A G A G T
 G A G C A G C T T C A T G G T T C T G C A A G T G A G C A G C A A A C T G A A T T T C T T C G C T C C A A T G T C A G C A G T G G C
 A A G G T C T C A A A A T A T C A T T T T G G C C C T A T A G T T A G T C A G A A G C T T A A G G A A C A T T T T G A A A A G A T
 C C A T A T C C T T G C C G T G C A A C A A A A G A G A G T C T G G C A C A A G A G C T A G G A C T G A C A T T T A A T C A G A T C
 A G T A A G T G G T T C T C T G G C A C T C G T C A T T A T T C A A G A G A T G C T G T T G C C A A G A A T C A G A A A C G C C C G
 G G A G A A A A T A C T A C T G T G A A T A A T A T A G C A C A A C C T T T G A T G A C A T A C A G T A A T A G A C C C A A T

FIGURA 7 (continuación)

GTTGGGTTGATAGATAAACAGCTGCAGATGTAATGACATGATTTCTGAAAAGTTGATGGTCCAA
 ATAAATCTTAACGAAGGCATTGAAGAAGATATCCACCCAGCCAATATGCCAGGTGTGAAGAGAAA
 ATAACTATGACTCCAGCTGCCATCTCAAGGGAAGCTGGTCTTCCAGGATATGGTCCTGGTCTGGA
 GAAAACCTTCTCAGGTCAGTTCAAGGAATACTAGCTGTGAGCAGAGTGTGATACTGGCTCCTACC
 GCTATCGCAAGAGAAGTTGGTCCACCAGGATATGCGACTGGAGAAAACCAGGGCAATGGGGCTCCA
 TGGAATACAAGTTATGAGCGAAGACTATTTATGAATCCTGTGACCAGCTCAAGAGCAGTCAGTCCT
 CCAGGATATGGGCCTGAAGAAAGCCGAGGCAGTGGCATTTCCTGGAATACAAGCAGTGAACAAATA
 TTGTTTATGAGTCCTACAACCATCTCAAGAGAAGTTGGTCCACCAGGATATGGGCCTAGAGAAGAC
 CAGTGCAGTGGCACTTCATGGAATGCAAGCTGCGAGTTGGGAGTATTTATGAGTCCTGTGACCACC
 TCAAGAGAAGACAGTCCGCCAGGATATGGGCCTCGAGAAAACCAGGAGAATGACAATATGAGATGC
 GAGTTGAGAAATGTTTACAAGTCTACAACCATCTCCAGAGAAGTTGGCCCACCAGGATATGAGCTT
 ACTGGA AAAACAAGGGCAACGGCGCTTCATGTACTACGAATTACCAGCAGGGAGCGTTTACGAGT
 CCTGAAGCCTTCCCAAGAGAAGCCTGTCTCCAGGATTTGTGTCTGGAGAGAACCAGGGTAACGAC
 ACTCCATGGTATATGAGCAGCAAAACAGATGGCGTTTACGAGTCTACGACCAACCCTAGAGTAGGT
 CCGCCAGGGTATGGACAGGAAAACCAGGGCAGTCGTGTTTCTTGGAAATACGAGTTATGAGCAGGGA
 GTGTTTATGAGTCTACAACATCTCAAGAGAAGCCGGTCCGCCAGGTTATGGGCCTGGGGAAAGT
 CAGGGCAATATCACTTCATGGATTTGACAGGCGGGCAGAGAATGTTTACAGGTCCTTCAACTATC
 CCGAACGAAGTCGGGCAGCAAGGATATGGCCCTCAGGAATCCCGGGAACGGATGGCTCAAGGAGT
 ATGGACTTGAAAGCCTTCCACCAGGATATGGACCTGCTGGAGAAAACCAGGGCGGTGGTGCTTCA
 GGGAGCGTCATAAGCCCCATGCTCGGTCTGCAGAGAACGTAGAGTTTTCAGACGACGCTAGAAAA
 AAGGCTATACAGCGGGAGCTGAGACGGCGGCAAGTTCAGGTGA

SEC ID Nº: 210, secuencia de polipéptido traducida PHDf_HD_HOX2a de *Zea mays*

MGRKGVSEVSPSKRYPLRSAHSSGRVLRSAALANNNDSEPLNAAAAATQAAVKKRSGNQLDSPNSTVR
 LLRSTSKNKDEAHSGPLNDRIIDKPAANKRKHANPSQVGS PSSSVRVLRSAIKNKDMARSEPQSDR
 TAGEAATNKRKNANSSKVGSLSSGLSVLRSASKAKDETCIKPLNDSPAVEPASSKRKCTSPSKVGS
 PSNSVRVLRSTSKYKNETCTEPLNDGAVEPAANRRKGVSRKEGSPNSSVRVLRASANNKIDVSI
 EPLNNSTGQPAARRRKNSSFKSSARVLCSTSERKNEASSEPLNDSTAAQPAARKKKGGAMPKTD
 NPKIGVRVLRASAGKKNCTEFPVNRSTSREPTVTKRCKPSKDRGPKDYKICQRIKYILNRM
 KYEQAFIQAYASEGWKQSLEKIRPEKELERAKAEILQCKLRIREAFRNMDSLLSKGKLEESLFD
 AGQISSEDI FCAICGSKDVTSONDIILCDGACDRGFHQNCLSPPLLTEEIPPGDEGLCPACVCKA
 DYIDALNELQGSKLSIHDSWVKVPEAASTANGSKQVDASDLLPDHIKDSANLALVGTHMVNEIRF
 SEEDDSKADDLGLPSEDSGDGDFDPAGPDSSDQNDGLNSEESDFTSDSDYFCAEIAKSCGQDEV
 ASPLSNVINRTRDRMKLRAYSRRSNEENHNHAFMDMELEQDIVLPVPSRRQVERLDYKKLYDEAYEE
 ESSNSSDEEWSGKELLESGTDSLDEPSSCKRLSRRAPAGQQNNELTPQSRPHGSNSEQKTEVLR
 SNGSSSTGRKYDPEVTQKLKVHFEKDPYPSRETENLSEELGLTFNQVSKWFSSTRHYSRVASAKK
 EMHPDSHTCENNETTVDSMQAKQPNAGVMEKLTGDRNGIVPEKPMVQDNLNQGAMEKLTRDKNDI
 VPEKPMQSSLNQCNNEDIPLSGTEIEMESYEQESSDSSDEEWSILSTPRKTRLQDHGKSSVTESL
 RSAKRCRRAPAREQNNERIQSEQLHGSASEQQTEFLRSNVSSGKVSKYHFGPIVSQKLKEHFEKD
 PYPCRATKESLAQELGLTFNQISKWFSGTRHYSRDAVAKNQKRPGENTTVNNNSTTFDDIQVIEPN
 VGLIDKPAADVNDMISEKLMVQINLNEGIEEDI PPSQYARCEEKITMTPAAISREAGLPGYGP
 ENFLOVSSRNTSCEQSVILAPTAIAREVGGPYATGENQNGAPWNTSYERRLFMNPVTSSRAVSP
 PGYGPESRSGSISWNTSSEQILFMSPTTISRELGPYGPREDQCSTSWNASCELGVFMSPVTT
 SREDSPPGYGPENQENDNMRCELRMFTSPTTISREVGPYELTGKNKNGASCTTNYQQGAFST
 PEAFPREACPPGFVSGENQNDTPWYMSSQMAFTSPTTNPVGVPPGYQENQGSRVSWNTSYEQG
 VFMSSTNISREAGPPGYGPESQGNITSWISTGGQRMFTGPSTIPNEVGQQGYGPSGIPGTDGSR
 MDLKAFFPGYGPAGENQGGGASGSVISPHARSAENVEFSDARKKAIQRELRRRQKFR

FIGURA 7 (continuación)

SEC ID Nº: 211, secuencia de ácido nucleico PHDf_HD_HOX2b de *Zea mays* X89761

ATGGGAAGGAAAGGCTCAGAAATGTCACCAAGTAAAAGGTATCCCTTGAGGTCTCCACATAGTAGT
 GGTAGAGTTCTTCGCTCTGCCTTAGCCAATGATAATAAGGACAGTGAGCCTCTGAATGCTGCTGCT
 GCTACCCAAGCCGCTGTGAAAAAAGAAGTGGCAGTCAGTTGGATAGTCCTAATAGCAGTGTCAGG
 TTGCTTCACTCTACATCAAAAAATAAGAACGAGGTATGTAGCGGGCCTCTAAATGACAGGATCCCT
 GATAAGCCAGCTGCAAAATAAAAGAAAACGTGCCAGTCCCTCACAGGTGGGAAGTTCCAGCAACAGT
 GTCAGGGFCGTTTCGCTCTGCAATAAAAAATAAAGATGAGGCATGCAGTAAATCTCTAAGTGACAGG
 ACCGCTGGTGAGCCAGCTGCAAAATAAAAGAAAAACGTCAATCGCTCAAAGACGGGGAGTCTGAGC
 TGTGGTGTCAGGGTGTCTTCTATCTGCTTCAAAGCTAAGGATGAGACATCTACCAAGCCTTTAAAT
 GATAGTGCTACTGTTGAACCAGCTTCAGGTAAGGAAAGGCACCAGTCCGTCAAAGGTGGGAAGT
 GCCAGCAATAGTGTCAGGGTGTTCGATCTACTTCAAATAATAAGAATGAGACATGTACTGAGCCT
 TAAATGATAGTGGTGTGTTGAACCAGCTGCAAAATAAAAGGAAAGGTGTCAGTCGTTCAAAGGAG
 AGAAGTCTCAATAGCAGTGTCAGGGTCGTTGCTCTGGCCTCAAATAATAAGATTGAGGTATCTATT
 AAGCCTCTAGGTAATAGTACTGCTGGGGAACCAGCTGCGAAAAGAAGAAAAGCGGCAGTTCCTTC
 GAGGCAGGAAGCCCTGTGAGCAGTGCCAGGGTGTTCGCCCTACCTCAGAAAGGAAGAATGAGGCA
 TCTAAGCCTTTAAATGAAAGTACTGCTGCTCAACCAGCTGCGAGGAAAAAAAAGCTGGGGTAATT
 TCAAAGACCGACAATCCAAAGATTGGTCTCAGGGTCTTCGCTCTGCCTCAGGAAAGAAGAATGAA
 GCATGTATTGGGCATGTAATGATAGTACTTCTGCTGAACCAACTGTGACAAAAAGAAACCGCAAG
 CTTTCAATGGATCGCAGTCCCAAGAAGGACTACTTAAAAATTTGTCAAAGAGTTAGATAATTTTTG
 AATCGAATGAACATCAACAAACCTTTATTCAAGCTTATGCAAGTGAAGGTTGGAAAGGCCAAAGT
 TTGGAAAAAATAAGACCTGAAAAGGAGCTTGAACGAGCCAAGGCAGAAATCCTCCAATGCAAGTTA
 AGAATCCGTGAAGCTTTTCGAAATATGGACTCTCTTTTATCCAAGGGGAAGCTTGAAGAATCTTTG
 TTTGATTCCGCAGGAGAAATTTCAAGCGAAGATATATTTTTGTGCTGTATGTTGGCTCAAAGGATGTT
 ACTTTACAAAATGACATCATTCTTTGTGATGGAGCCTGTGACAGAGGATTCACCAGAACTGTTTTG
 AACCTCCTTTGCTAACTGAAGATATACCTCCGGGGGATCAAAGATGGCTTTGCCCTGCATGTGTC
 TGCAAAGCAGACTCTATAGATGCACTAAATGAACTTCAAAGGAGCAAGCTCTCCATTCATGACTCA
 TGGGAGAAAAGTTTTCCCTGAGGCAGCTTCAATTGCCAACGGTCTTAAACAAGTTGATACTTCTGAT
 CTGCTGCCAGATCATATAAAGCACAGTGACAACCCTGCATTGGTTGAAGGGCTGATGGTGGATGAA
 GTCAGGCTTCTGCAGAGGATGACAGCAAAGCGGATGACCTTCGCTTGTCTTCGGAGGACTCAGGG
 GATGGTGACTTCGATCCGTCAGGTCCAGATTCTAGTGAAGACCAAAAAGATGGATTGAACTCAGAA
 GAGTCAGATTTACATCAGACTCTGATGATTTTTGTGCTGAGATTGCCAAATCTTTGTGGCCAGGAT
 GAAGTCTCGGCATCTCCATTATCAAATGTGATAAACCATACTTATAGAATGAAACTCAGGGCCAGC
 ACAACCCTCTAATGAAGAAAACCATGATCATGATTTATGGACATGGAGCTAGGGCAAGACATG
 GTGCTACCAGTTTCAAGCAGGCGGCAGGTGGAACGGTTGGATTACAAAAACTATATGATGAGGCA
 TATGGGAAAAGAATCATCTAATTCAAGTGTGACAAAGAGTGGTCTGGAAAAGAACTATTAGAAGGC
 AGTGAACAGATTCACTTAGTGAGCGACCTCATCCTGTAAAACGGTCTCTAGAAGAGCACAAGCA
 GAGCAGCAGAATAATGAACATACACCACAGAGGGAGCGCCTTCATGGTCTGAAAGTGAGCAGAAA
 ACTGGAATTCCTTCGATCCAATGGCAGCAGTAGCACAGGACGGAAGTTTGGTCTGTAGCTACCCAG
 AAATTGAAGGTACATTTTGAAGAAAGATCCATATCCTAGCCGTGAAACGAAAGAGAATATATCAGAA
 GAACTAGGCTTGACATTTAATCAGGTCTCCAGATGGTCTCTAGCACTCGTCACTATTCAAGGGTT
 GCTTCTGCCAGAAAAGAAATGCATCCTGACGGCCATACTAGCGAGAATAATGATACCACGACTGTA
 GATAGCATGCAAGCAAGGCAACCAACTGTGGTGTGATGGAAAAGTTGACTGGGAATAGAAAATGAC
 ATTGTTTCTGAGAAACCGATGGTGCAGAACAATCTTAACCAAGGGGTGGCAGAAAAGTTGACTAGG
 GATAGAAATGACATTTGTTCTGAGAAACAGGTGGTGCAGAGCAATCTTAACCAATGTAATAATGAA
 GATATCCCGCTTCTTGGAACAGAAATTTGAGATGGAGTCTTACGAACAAGAATCATCTGAAAGCAGT
 GATGAAGAGTGGTCCACACTTAGCACACCACGAAAAACAAAATTACAAGGCAACGAAAAAGTATCA
 CTTACTGAATCACTACGTTCTGCAAAACGGTGTCTAGAAGAGCACCAGCTAGGGAGCAGAATAAT

FIGURA 7 (continuación)

GAACATACTCAGAGTGAGCAACTTCATGGTTCTGTAAGTAAACAGCAAACCTGAAGTTCTTCGCTCC
 AATGTCAGCAGTGGCGATGCCTCAAAAATGTCATTTTGGCCCTATAGTTACTCAGAAGCTGAAGGAA
 CATTTTGAAAAAGATCCATATCCTTGCCGTGCAACAAAAGAGGGTCTGGCACAAGAGCTAGGACTG
 ACATTTAATCAGATCAGTAAGTGGTTCTCTGCCACCCATCATTATTCAAGAGATGCTGTTGCCAAG
 AACCAAAAATACCCAGGAGAAAATACTACTGAGAATAATAGTAGCACAATCTTTGATGGTATACAA
 GTAATAGAACCCAAATTTTGGTTTGGATAGATAAACAGATGCAGATACAAAATGACATGATTTCTGAA
 AAGTTGATGGTACAAATTAATCTTAACGAAGGCATTGAAGAGGATATCCCACCGAGCCAATATACT
 ACCAGGTGTGAAGAGAAATTAATACTATGACTCAAGCTGCCATCTCAAGGGAAAGCTGGACCTCCAGGT
 TATGGTCTCGGAGAAAACCTTTCTTCAGGTGAGTTCAAGGATTACTAGCTGTGAGCAGAGTGTGATA
 ATGGCTCCTAGCGCTATCGCAAGAGGAGTTGGTCCACCAGGATATACGCTGGAGAACACCAGGGC
 AATGGGGCTCCATGGAATACAAGCTATGAGCGAAGACCGTTTACGAGTCTCGGACTAGCTCAAGA
 GAAGTCGGACCTCCAGGATATGGGCTGAAGAAAACCAAGGCAGTGGCATTCTCTGTAATACAAGC
 AGAGAGCAAATATTGTTTATGAGCCCTACAACCATCTCGAGAGAACTTGGTCCACCAGGATATTTG
 CCCAGAGAAAACAGCGCAGTGACACTTCATGGAATACAAGCTGCGAGTCCGGGACTATTTATGAGT
 CCTGCGACCACCTCAATAGAACATGGTCCACCAGGATATGGGCTGGAGAAAACAGGAGAATGAC
 AGTACAAGCTGCGAGTTGAGAATGTTTACCAGTCTACAATCGTCTCCAGAGAAGTTGGCCACCA
 GGATACCAGCAGGGAGTGTATGAGTCTAAAACCATCTCAAGAGAAGTCTGTCCTCCAGGATTT
 GGGTCTGGAGAGAACCAGGGCAACGACACTACATGGTACAAACAGAGCGCGTTTACAAGTCTTACA
 ACCATCTCGAGAGAAGTCCGGTCCGCCAGGGTATGGACAGGAAGATCAGGGCAATGGTGGTTCTTG
 AACACGAGTTGTGAGCAGGGAGTGTATACGAGTCTACTAGAGGCATCTCAAGAGAAGTCCGGTCCG
 CCAGGATACGGGCCCGGGAAAGTCAGGGCAATACCACCTTCGTGGATTTGACCTGCCAGCAGAGA
 ATGCTTGCAGGAATCCAGAGAAGTGGTGGCTCAAGGAGTATGGACTTGGAAAGCCTTCCACCGGGA
 TACGGACCTGGTGGTGTTCAGGGAGTGTACATAAGCCCCAAGCTCGGTCTGCAGAGAATGTAGAA
 TTTTCGGACGAGGCTAGAAAAAAGGCTATACAGCGGGAGCTGAGACGGCGGCAGAAGTTCAGGTGA

SEC ID N°: 212, secuencia de polipéptido traducida PHDf_HD_HOX2b de *Zea mays*

MGRKGSEMSPSKRYPLRSPHSSGRVLRSLANDNKDSEPLNAAAATQAAVKKRSGSQLDSPNSSVR
 LLHSTSKNKNEVCSGPLNDRI PDKPAANKRKRASPSQVGSSSNSVRVVRSAIKNKDEACSKSLSDR
 TAGEPAANKRKNVNRSKTGSLSGVRVLLSASKAKDETSTKPLNDSATVEPASGKRKGTSPSKVGS
 ASNSVRVLRSTSKYKNETCTEPLNDGAVEPAANKRKGVSRSKERSLNSSVRVVALASNNKIEVSI
 KPLGNSTAGEPAAKRRKSGSSFEAGSPVSSARVLRPTSERKNEASKPLNESTAAQPAARKKKAGVI
 SKTDNPKIGLRVLRASGKNEACIGHVNDSTSAEPTVTKRNRKPSMDRSPRKDYLIKQVRVYIL
 NRMNYQQTFIQAYASEGWKQSLEKIRPEKELERAKAEILOCKLRIREAFRNMDSLLSKGKLEESL
 FDSAGEISSEDI FCAVCGSKDVTLQNDIILCDGACDRGFHQNCLNPPLLTEDI PPGDQRWLCPACV
 CKADSIDALNELQGSKLSIHDSWEKVFPEAASIANGSKQVDTSDLLPDHIKHSNDPALVEGLMVDE
 VRLSAEDDSKADDLRLSSEDSGDGDFDPSGPDSESDQKDLNSEESDFTSDSDDFCAEIAKSCGQD
 EVSASPLSNVINHTYRMKLRASNNRSNEENHDHVFMDMELGQDMVLPVSSRRQVERLDYKKLYDEA
 YGKESNSSDDKEWSGKELLEGETDSLSEPHVVKRCSRRAQAEQONNEHTPQRERLHGSESEQK
 TGI LRNSGSSSTGRKFGPVATQKLVHFEKDPYPSRETKENISEELGLTFNQVSRWFSSTRHYSRV
 ASARKEMHPDGHTSENNDTTTVDSMQARQPNTVVMEKLTGNRNDIVSEKPMVQNNLNQVAAEKLTR
 DRNDIVPEKQVVQSNLNQCNNEDIPLSGTEIEMESYEQESSESSDEEWSTLSTPRKTKLOGNEKVS
 LTESLRSKRCSRRAQAEQONNEHTQSEQLHGSVSKQOTEVLRNSVSSGDASKCHFGPIVTVQKLKE
 HFEKDPYPCRATKEGLAQELGLTFNQISKWFSATHHYSRDAVAKNQKYPGENTTENNSSTIFDGIQ
 VIEPNFGLIDKPDADTNDMISEKLMVQINLNEGIEEDI PPSQYTTTRCEEKLTMTQAAISREAGPPG
 YGPGENFLQVSSRITSCEQSVIMAPSAIARGVGGPYTPGEHQNGAPWNTSYERRPFTSPATSSR
 EVGPPGYGPEENQGSGISCNTSREQILFMSPTTISRELGPPLYLPRENQRSDTSWNTSCESGLFMS

FIGURA 7 (continuación)

PATTSIEHGPPGYGPGENQENDSTSCELRMFTSPTIVSREVGPPGYQQGVFMSPKTISREVCPPGF
 GSGENQGNDDTTWYKQSAFTSPTTISREVGPPGYQEDQNGGWSWNTSCEQGVFTSPTRGI SREVG
 PGYGPGESQGNNTSWISTCQQRMLAGIQRGGSRSMLEAFPPGYGPGGASGSVISPQARSAENVE
 FSDEARKKAIQRELRRRQKFR

SEC ID N°: 213, secuencia de ácido nucleico PHDf_HD_II de *Vitis vinifera* cóntigo de AM464161.2,
 AM478203.2

ATGGATGCTTCCCCTGCCAAGAGAGTAATCGGACTAGGAAAAGTTCTAGTCCTAACAGAATATT
 TTAGAAGAGGCCAGAAAACCTTCTGAAAGTGTGTGTAGCGAATCGTCAGAGCAGAAAACGTCCCCTCT
 GAAAATGGACAGCATGAGCCAGCAGAAATCAGCCCTGTGCTATCCAATTGTATTGTTACTGAACAG
 TCAGAATGCCTCCAGAAGATGTGGGTGATACAATATTAGGATTGCCTCCAGCAGATGTGACCAAG
 AATTCTCTTACTGAACATTTGGGATTGCCCTCAGAAGATGCAATCAAGAATGATGGAACCTGAACAA
 TTGGGGTTTTTCTGAAAGTGTGACCAAGAGTTCTATTATTGAAAAATTAGGACAATCAGAGCCA
 CCTCCTGAAAATGTGGCCCGGTATTCTGGTCTTGACCAGTCAGGGTCAGCACCCAAAGATTTGGCC
 AATAAAAAGAACTGCAAAACTAGTTAAAAGAAAATATAAATGAGATCTTCAGTAAGCGGTTCCCGG
 GTTTTACGCTCAAGGTCACAAGAGAAAACCTAAAGCTTACAGCCAAGTGATAATTTTGTAATGCC
 AGTGCTAGCAGAGAAAGGAAAGGAAAGAAAAGGATGAATAAAAACAACCTGCAGATGAATTT
 GCGAGAATCAGGAAACATCTTAGATATTTACTGAACAGAATGAGCTATGAGCAAAATCTGATTGAT
 GCTTATTCTGCCGAAGGTTGGAAAGGACAAAGTGTGGAAAAATTAAAGCCAGAGAAGGAACTTCAA
 CGTGCCTCATCTGAAATTTCTCGTCGCAAACTGMAAATACGTGATCTATTTCAACATCTTGATTCA
 TTATGTGCTGAAGGAAGGTTCCAGAATCTTTATTTGATTCTGAAGGGCAGATTGACAGTGAGGAT
 ATATTCTGTGCTAAATGTGAGTCCAAAGACATGTCTGCTGATAATGACATAATACTCTGTGACGGT
 GCTTGTGATCGTGGATTTACCAGTTTTGTCTGGAACCGCCATTGTTAAAAGAAGAAATTCCTCCT
 GATGACGAGGGTTGGCTGTGCCCTGCATGTGATTGCAAAGTTGACTGCATGGACCTGCTTAATGAC
 TCTCAAGGAACAAAACCTTCTGTAATTGATAGCTGGGAGAAGGTTTTTCTGAGGCAGCTGCAGCT
 GGGAAATAACCAGGATAACAACCTCTGGATTTTCATCAGATGATTCTGAGGATAATGATTATGACCC
 GATTGTCCAGAGTTGATGAGAAGGGTCAGGGAGATAAGTCAAGTTCTGATAAATTTGATGAATCT
 GATGAATTTGACGAATCTGATGAATCTGATTTCACTTCTGCATCTGATGATATGGTGGTCTCACC
 AATAATGAGCAGTGTGGGGCTTCTCTGATGATTGAGAGGATGATGATTTGATCCTGATGCT
 CCAGAAATGATGAACAGGTTAATCAGGGGAGTTCAAGTTCTGATTTTACATCTGACTCCGAGGAT
 TTTACTGCCACTTTAGATCGCAGAACTTCTCTGACAACGAGGATGGTCTTGATGAGCAAAGAAGA
 TTTGGTAGGAAGAAGAAAGATACTTTAAAGGATGAGCTCTTATCCGTGTTAGAGTCAAATTTCTGGT
 CAAGATAATGCACCTCTGTCTGCAAAGAGACATGTAGAAAGGCTGGATTACAAAAGCTGCATGAT
 GAGGCATATGGAAATGTTTCTCTGATTCAAGTGATGATGAAGACTGGACAGAGAATGTTATACCA
 AGAAAGAGGAAGAATCTTAGTGGTAATGTTGCCCTCAGTGTACCAAATGGAAACACTTCAATCACT
 GAGAATGGGACAAATACCAAGGACATAAAGCAGACTTGGAAAGCAGCTGGATGCACTCCCAAGCGA
 AGAACTCGTCAGAAGCTGAACTTTGAAAGCACAAATAATTCACTTGCTGAGTCACATAAAGGTTCT
 CGAAGTCTGGTTCTACTGGTGAAAAAGTGGGCAATCATCATATAAAAAACTTGGAGAAGCTGTA
 ACTGAGAGACTCTACAAATCCTTCCAGGAAAAATCAGTACCCTGATCGTGCTATGAAAGAAAAATTG
 GCAGAAGAGCTGGGAATAACGAGCCGGCAGGTTAGCAAATGGTTTGAGAATGCCCGTTGGAGCTTT
 CGCCATCGACCACCAAGGAAGCCAGTGCAGGTAAGTGTGTAAGGATGCATCCACGTCT
 CAAACAGATCAAAGCCTGAGCAAGAAGTGGTCTTAGAGAAAAGTTCTCACAATGGAGTGGGAAAA
 AAGGAGTCGCCCAAAGCAGGTGCTTCGAAGGTAGACCGCAGCAAGGAAGCCAATGCAGGTAAGG
 GCTGTGAAAAAGGATGCATCCACGTCTCAAACAGATCAAAGCCTGAGCAAGAAGTGGTTATTTAA
 GAAAGTTCTCACAATGGAGTGGGAAAAAGGAGTCGACCAAAGCAGGTGCTTCGAAGGTAGACCGC
 TGCAGTGGAGCTAAGAGGAGAAGGAAATTTGGCAACTGATGGAAGCCATAGACAGAAATCTTCAACT
 CCCAACTCTACAAGACAAAAGACAAAGTCAAATCATGAAGCATCTGAGGCAACCAATGGAAGCAGA
 AGGCAGAATTTCTACTCCCAAGTCTAGAAGGCAAAAGACTAAGTTAGCTGGTGAAGCGTCTGAA

FIGURA 7 (continuación)

GCAACTGGTGGGAAGCAGTAGACAGAATTCGTCAACCCCAAATCTAAAAGGCGAAAAACCAAGTCA
 GATCATGAAGCATCTAATCCCGTCCTTAGCGGTAAGAAGATAGCAAAGACGGCAAAGAGCTCATCT
 GGCACACCAAAAACAAAGGAAAAGCTAAGTGACAGAATCCAAACAAGAAGTAGGAAATCTATTGCT
 TGA

SEC ID N°: 214, secuencia de polipéptido traducida PHDf_HD-II de *Vitis vinifera*

MDASPAQESNRTRKSSSPKQNIILEARKLSESVCSSESSEQKRPSENGQHEPAEISPVLNSNCIVTEQ
 SELPPEDVGDITILGLPPADVTKNSLTEHLGLPPEDA IKN DGTEQLGFFPEVVTKSSIIIEKLGQSEP
 PPENVARYSGLDQSGSAPKDLANKRTAKLVKRKYKLRSSVSGSRVLRSRVSRQEKPKASQPSDNFVNA
 SASRERKGRKKRKNKTADDEFARIRKHLRYLLNRMSYEQNLIDAYS AEGWKGQSVEK LKPEKELQ
 RASSEISRRLXIRDLFQHLDSLCAEGRFPESLFDSEGQIDSEDFCAKCESKDM SADNDIILCDG
 ACDRGFHQFCLEPPLLKEEIPDDEGWLCPACDCKVDCMDLLNDSQGT KLSVIDSWEKVFPEAAAA
 GNNQDNNSGFSSDDSEDNDYDPDCPEVDEKQGQDKSSSDKFDESDEFDESDESDFTSASDDMVVSP
 NNEQCLGLPSDDSEDDDFDPAPEIDEQVNQGSSSSDFTSDEDF TATLDRRNFSDNEDGLDEQRR
 FGRKKKDTLKDILLSVLESNSGQDNAPLSAKRHVERLDYK K LHDEAYGNVSSDSSDDEDWTE NVIP
 RKRKNLSGNVASVSPNGNTSITENGTNTKDIKHDLEAAGCTPKRRTRQKLNFE STNNSLAESHKGS
 RSPGSTGEKSGQSSYKKLGEAVTERLYKSFQENQYPDRAMKEKLA EELGITSRQVSKWFENARWSF
 RHRPPKEASAGKSAVKKDASTSQTDQKPEQEVVLRSSHNGVGK KESPKAGASKVDRSKEANAGKS
 AVK KDASTSQTDQKPEQEVVIKES SHNGVGKKESTKAGASKVDRCSGAKRRRKLATDGS HRQKSS T
 PNSTRQKTKSNHEASEATNGSRRQNSSTPKSRRQKTKLAGEASEATGGSSRQNSSTPNSKRKRTKS
 DHEASNPVLSGKKIAKTAKSSSGTPKTKEKLSDRIQTRSRKSIA

SEC ID N°: 215, secuencia de ácido nucleico PDHf_HD_II de *Populus tremuloides* scaff_IX.730

ATGTCTGAGGCTGAGAACATGGGTGTTTCTCCATCACAAGTTAGTCTCAAAC TAAAAGTTACTCA
 TGCCCTTCACAAACTAAACTGGAAAATACTCATGGATTTACTGCTGAATACA ACTGTGGTGGATAT
 TCGGAAGAGAAACATAAGCTTGAGTCTGAAATCATACAA ACTGAAGCCGGAGATAGCGGTACTGCA
 GTTCTCCAATCTGGTGCTGGGGAAACTGTGGAACCATCTACTGAAGATGTAACAAACAATTCTTTT
 ACTGATTTGGATCCACCTCCCGAAGATGCCAGGGGTGCTACTTTTGATGAAAGTTCAAGACCTATA
 CTAACTGCTATAGATCAGAACTCGAGCCAGGTGCTAGAAGTGTGAATACTGCGTGTACGCATGGT
 GAATCATCAAAGCAACTGACTCTGGTATTCTACAGGATGAACCTGGA AATACCAATGCTGCATCA
 TCAAGCTGCATTGCTAATGAAACCTCACAAGCATCAC TTGAAAATCTAGCCAATAATTCTTGTACT
 GAAGATGTGGGTCCGCCGTATGGAGATGCAAGCAAGGTAACCAGATTGATAAAAAGTTCATAC CCT
 CAACAACTATATCTGGACATACTTGGAGTTACTTTCTGACAGAGCTTGTTGTGAACGATCAGAA
 GAAAGACAGAAGCCTGGATCTGAACTTTCAGAAAATGAATCAACGGGAATTGACACTGAATTATAT
 AGTGGCATTGCTATCGAGAATTCAGAGCCGCTTACTCAACTAGTGACCAAGAGTTCTCCTATCAAA
 CATGTAGGCTTGCTTCCTGGTGACAGCATCATCATTCCTGCAAATGAACAAACAAGACCAACTCAT
 GATGATGAGGACAAAGGCCAGATCATGAACATTTGGAACACCATCTAGAGTTGCGATTGGTATT
 ACTAGACGCGGTAGACCTAGAGGTAAGAGTGCTTCAAGATTGTCAAGGAAGATATATATGTTAAGA
 TCTTTGAGAAGTAGTGATAGAGTCTCCGGTCAAGATCACAAGAGAAACCTAAAGCTCCTGAATCA
 AGTAATAATTCAGGTAATGTTAATTCCTACTGGTGATAAAAAGGGAAAAGAAGGAAAAAGAGAAGA
 GGAAAGAATATTGTGGCTGATGAATATCCAAAATCAGGGCACATCTTAGG TACTTATTGAATCGG
 ATGAGCTATGAGCAAAGCTTGATCACTGCTTATTCTGGGGAAGGCTGGAAAGGACTTAGCCTAGAA
 AAGTTAAAGCCCGAGAAGGAGTTGCAGCGAGCCACATCTGAAATTACCAGACGCAAAGTGAAAATA
 AGGGATTTATTTCAACATATTGATTCTCTGTGCAGCGAAGGAAGGTTCCCATCATCTTTATTTGAT
 TCTGAAGGACAGATTGACAGTGAGGATGTATTCTGTGCAAATGTGGGTCCAAAGATTTGAATGCT
 GATAATGATATTATACTATGTGATGGTGCTTGTGACAGAGGATTT CATCAGTTCTGTTTGATACCA

FIGURA 7 (continuación)

CCATTGTTAAGAGAAGACATTCCCTCCTGATGATGAGGGCTGGTTATGCCCTGGATGTGATTGCAAA
 GTTACTGCATTGGTTTGGCTTAAACGATTCTCAAGGAACAAATATATCTATCAGTGATAGCTGGGAG
 AAAGTTTTCCCTGAGGCTGCCGCCACAGCGTCTGGACAAAACTGGATCACAACCTTTGGACCATCT
 TCTGATGATTCTGATGATAATGATTATGAACCTGATGGTCCAGATATTGATAAGAAGAGTCAGGAG
 GAGGAATCGAGCTCTGACGAATCTGACTTCACTTCTGCATCTGATGAATTCAAGGCTCCCCGGAT
 GGCAAAGAGTACTTGGGGCTCTCTTCTGATGATTGAGGATGATGACTATGATCCTGATGCTCCA
 GTTCTTGAGGAGAAGCTGAAGCAGGAAAGTTCAAGTTCTGATTTTACATCAGACTCTGAGGATCTT
 GCTGCAACTATTAATGGTGATGGGTGTCTTTAGAAAGATGAATGCCACATGCCTATTGAACCTCGT
 GGAGTTTTCTAATGGGAGGAAATCTAAATTTGATGGAAAGAAGATGCAGTCTTTAAACAGCGAGCTT
 TTATCCATGCTAGAACAGATCTGTGCCAAGATGAGTCAGCAACTGTTTCTGGAAAAGAAATGTT
 GACAGATTGGACTACAAGAAGCTCTATGATGAAACGTATGGAAACATTTCTACTTCAAGTGATGAT
 GATTACACTGATACTGTTGGGCCAAGGAAGAGGAGGAAAAATACTGGAGATGTTGCTACAGTGACA
 GCAAACGGAGATGCCCTGTACCAGAAATGGAATGAATAGCAAGAATATGAACCAGGAATTAAG
 GAGAACAAACGTAATCCTGAGAGAGGAACTTGCCAAAACTCGAGCTTCAAGAAACAAATGTTTCC
 CCGGCTAAATCATATGTTGGTGCATCTCTATCTGGTCTAGTGGTAAAAGCGTTAGGCCCTCTGCT
 TATAAAAACTTGAGAGAAGCTGTAACCTCAGAGACTGTACAGTTACTTCAAGGAAAATCAGTATCCA
 GACCGAGCTGCAAAAAGCAAGCTTGGCAGAAGAGCTAGGAATTACTTTTGAGCAGGTTAAACAAGTGG
 TTTGTGAATGCCCGTTGGAGCTTCAACCATTATCATCTACAGGTACAAGTAAAGCTGAAAGTGCT
 TCTGGAAAAGGTAGCTGTGATGGCCAAGTGAGGGATAGTGAATCGAAGAACCAGAAAGAGCAATAAA
 CAAAAACTAACACCCCGAAATCTAGGAGATAA

SEC ID N°: 216, secuencia de polipéptido traducida PHDf_HD_II de *Populus tremuloides*

MSEAENMGVSPSQVSSQTKSYSCPSQTKLENTGHFTAENYCGGYSEEKHKLESEI IQTEAGDSGTA
 VLQSGAGETVEPSTEDVTNNSFTDLDPPEPARGATFDESSRPILTAIDQKLEPGARSVNTACTHG
 ESSKATDSGILQDEPGNTNAASSSCIANETSQASLENLANNSCTEDVGPYPYGDASKGNQIDKSSYP
 QQTISGHTLELLSDRACCERSEERQKPGSELSENESTGIDTELYSGIAIENSEPLTQLVTKSSPIK
 HVGLLPGDSIIIPANEQTRPTHDEDEKGPDHEHLETSPRVAIGITRRGRPRGKSASRLSRKIYMLR
 SLRSSDRVLRSSQEKPKAPSSNNSGNVNSTGDKKGRKRRKRRGKNIVADEYSKIRAHLYLLNR
 MSYEQSLITAYSGEGWGLSLEKLPKPEKELQRATSEITRRKVKIRDLFQHIDSLCSEGRFPSSLFD
 SEGQIDSEDFVCAKCGSKDLNADNDIILCDGACDRGFHQFCLIPLLREDIPDDEGWLCPGCDCK
 VDCIGLLNDSQGTNISISDSWEKVFPEAAATASGQKLDHNFGPSDDSDNDYEPDGPDI DKKSQE
 EESSDESDFTSASDEFKAPPDGKEYLGLSSDDSEDDYDPDAPVLEEKLKQESSSDFTSDSEDL
 AATINGDGLSLEDECHMPIEPRGVSNRKS KFDGKKMQLNSELLSMLEPDLQCDESATVSGKRV
 DRLDYKKLYDETYGNI STSSDDDYDTVGP RRKNTGDVATVTANGDASVTENGMNSKNMNQELK
 ENKRNPERGTCQNSSFQETNVPKASYVGASLSGSSGKSVRPSAYKKLGEAVTQRLYSYFRENQYP
 DRAAKASLAEELGITFEQVNKWFVNARWSFNHSSSTGTSKAESASGKGS CDGQVRDSESKNRKSNK
 QKTNTPKSRR

SEC ID N°: 217, secuencia de ácido nucleico PHDf_HD_III de *Populus tremuloides* LG_1002624

ATGTCCGAGGCTGAGCACATGGGTGTTTCTCCATCAAAAGTTTCTCATACTAAAAGTTATTCATGC
 CCTGCACAAACTACACTGGAAAATACGCATGAACCCAGTGCTGAATACAAGTTTGGTGGATATCCG
 GAAGAGAGACATAAGCTTGAGTGTGAAATCATACAAACTGAAGCTGGAGATAACAGGGCTGCAGTT
 CTCCAATCTTGTTCTGGTGAAGTTGTGCAACCATCTACTGACGATTTAACAAAAAGTCTCTTATT
 GATTTGGACCCACCTCCTGACGATGCAAGGAGTGCTCTGTTTGATAATAGTCCAAGACCTATATCA
 ACCGCTATGGATCAGAACTTGAGCCAGGTGCTACAAGTGTGAATACTGCTTGTGTGCATAGTGAA
 TCATCAAAAGCAATTGACTCTAGTATTTACTGATGAACCTAGGAATAGCAACACAGAATTGTCA

FIGURA 7 (continuación)

AGCTGCATTGCAAATGAAACCTCACAAAGCATCACTTGAAGGTCTAGCCAATGATTCTCGTGCTGAA
 GATGCAGGACTGTCACCTGTAGAGGCAAGCAATAGTGACCTGATTGATGAAAGTTCATACTCTCAA
 CAAACTACATCTGGACAGACACGTGAGTTTCATTCTGACAGAGCTTGTGTAAACCATTGGAAGAA
 AGACAGAAGCCCGGCTCTGAACTTGCAGAAAATGAATCAATGGAAATTGGCATTGGATTACCCAGT
 GGTATTGCTATTGAGAATTTGGAGCCGCTTACTGAACTTGTGACCAAGAGTTGTCTATCAAACAT
 ATAGGCTTGCCTCCTGGTGATGACATCAGCATTCTGCAAATGAACAAATAAGACCTACTCATGAT
 AAGGAGTCCAAATACCCTGATTGTGAACATTTGGAAAACTATCTGGAATTGTGATTGGTATTACT
 AGTCAAGGTGTACCCAGTGTTAAGAGAACTTCAAATTTGTCAGGGAAGAAATATAACAAGTTCTTCG
 AGAAAAAGTGATAGAGTTCTCCGGTCAAATTCACAAGAGAAACCTAAAGCTCCCGAGCCGAGTAAT
 AATTCTACTAATGTTAATCCACTGGTGAGGAAAAAGGGAAAAAGGAGGAAAAAGAGAAGAGGGAAA
 AGTATAGTGGCTGATGAATATTCAGAATCAGGGCACGTCTTAGGTACTTGTGAATCGGATGAGC
 TATGAGCAAAGCTTGATCACTGCTTATTCGGGGAAAGGCTGGAAAGGACTTAGTCTAGAAAAGTTA
 AAACCTGAGAAGGAGTTGCAACGGGCCACATCTGAAATCATCAGACGCAAAGTGAATAAAGGGAT
 TTATTTCAACATATTGATTCACTATGCGGTGAAGGAAGGTTCCAGCATCTTTGTTGATTCTGAA
 GGACAAATTGACAGTGAGGATATATTCTGTGCAAATGTGGATCCAAAGATTTGACTGCTGATAAT
 GATATTATACTATGTGATGGTGCTTGCAGCCGAGGGTTTCATCAATTCCTGTTTGGTACCACCATTG
 TTAAGAGAAGACATTCCTCCTGGTGATGAGGGCTGGTTATGCCCTGGATGTGATTGCAAAGTTGAC
 TGCATTGACTTGCTTAATGACTCTCAAGGAACAAATATATCTATCAGTGACAGATGGGATAATGTT
 TTCCCCGAGGCAGCTGCAGTAGCATCTGGACAAAACCTCGACTACAACCTTTGGATTGTCTTCTGAT
 GATTCTGATGATAATGATTATGATCCTGATGGTCCAGATATTGACGAGAAGAGTCAGGAGGAATCA
 AGCTCTGACGAATCTGACTTCAGTTCTGCGTCTGATGAATTTGAGGCTCCACCTGATGATAAACAG
 TACTTGGGGCTCCCTTCTGATGATTGAGGATGATGACTATGATCCTGATGCTCCAGTCTCGAG
 GAGAAGCTGAAGCAGGAAAGTTCCAGTTCTGATTTTACTTCCGACTCTGAGGATCTTGATGCCACT
 CTTAATGGTGATGGGTTGTCTTTAGGAGATGAATACCACATGCCTATGAACTCATGAAGATTCT
 AATGGGCGAAGATCGAGATTTGGTGGAAAGAAGAATCATTCTTTAAACAGCAAGCTTTTATCCATG
 CTAGAACCAGATTTCTACCAAGAAAAGTCTGCACCTGTTTCTGGAAAAAGAAATATTGAAAGATTG
 GACTACAAGAAGCTCTATGATGAAACATATGGAAACATTTGTACTTCAAGTGATGATGATTTCACT
 GATACTGTTGCACCAAGGAAGAGGAGGAAAAACACTGGAGATGTTGCTATGGGGATAGCAAATGGA
 GATGCTTCTGTTACTGAGAATGGATTGAATAGCAAGAATATGAATCAGGAATTAAGAAGAATGAA
 CATACTTCTGGGAGAACTCACAAAACCTCGAGCTTTCAGATACAAATGTTTCCCAGCAAAAACA
 CATGTTGGCGAATCTCTCTCTGTTTCTAGCAGTAAAAGAGTTAGGCCCTCTGCTTATAAAAAACTT
 GGAGAAGCTGTAACCTCAGAAACTGTACAGTTTCTTCAAGGAAAATCGGTATCTGACCAAGCTGCA
 AAGGCAAGTTTGGCAGAAGAGCTAGGAATTACTTTTGGAGCAGGTTAACAAGTGGTTTCATGAATGCT
 CGTTGGAGCTTCAACCATTCATCACCTGAAGGTACAAGTAAAGCTGAAAGTGCTTCAGGAAAGGGT
 AGCTGTGATGGGCATGTGAGGGATAGTGAATCGAAGAACC AAAAGAGCAATAAACAAAAAACTAGT
 ACCCCAAAATCTAGGAGATAG

SEC ID N°: 218, secuencia de polipéptido traducida PHDf_HD_III de *Populus tremuloides*
 MSEAEHMGVSPSKVSHTKSYSCPAQTTLENTHEPSAEYKFGGYPEERHKLCEEIIQTEAGDNRAAV
 LQSCSGEVVQPSTDDLTKSPLIDLPPPPDDARSALFDNSPRPISTAMDQKLEPGATSVNTACVHSE
 SSKAIDSSILLDEPRNSNTELSSCIANETSQASLEGLANDSRAEDAGLSLVEASNSDLIDESSYSQ
 QTTSGQTRFHS DRACCKPLEERQKPGSELAENESMEIGIGLPSGIAIENLEPLTELVTKSCPIKH
 IGLPPGDDISIPANEQIRPHDKESKYPDCEHLEKLSGIVIGITSQGVPSVKRTSKLSGKKYPTSS
 RKSDRVLRNSQEKPKAPEPSNNSTNVNSTGEEKGKRRKRRGKSIVADEYSRIRARLRYLLNRMS
 YEQSLITAYSGEGWKGLSLEKLPKPEKELQRATSEIRRKVKIRDLFQHI DSLCGEGRFPASLFDSE
 GQIDSEDI FCAKCGSKDLTADNDIILCDGACDRGFHQFCLVPPLLREDIPPGDEGWLCPGCDCKVD
 CIDLLNDSQGTNISISDRWDNVFPEAAAVASGQKLDYNFGLSSDSDSDNDYDPDGPDI DEKSQEE

FIGURA 7 (continuación)

SSDESDFSSASDEFEAPPDDKQYLGLPSDDSEDDDDYDPDAPVLEEKLKQESSSSDFTSDSEDLDAT
 LNGDGLSLGDEYHMP IEPHEDSNRRSRFRGGKKNHSLNSKLLSMLEPDSHQEK SAPVSGKRNIERL
 DYKKLYDETYGNICTSSDDDFDTVAPRKRKNTGDVAMGIANGDASVTENGLNSKNMNQELKNE
 HTSGRTHQNSSFQDTNVS PAKTHVGESLSGSSSKRVRPSAYKKLGEAVTQKLYSFFKENRYPDQAA
 KASLAEELGITFEQVNWFMNARWSFNHSSPEGTSKAESASGKGS CDGHVRDSESKNQKSNKQKTS
 TPKSRR

SEC ID Nº: 219, secuencia de ácido nucleico PHDf_HD_IV de *Populus tremuloides* LG_XVIII1192

ATGGGTGATTCTGGAAAGAAATCAAAGCAGCAAGACTTGCACGAATCTTCACCATCTGACACAGTG
 AATGGGTGCGTTGCTGATTAAATCATTGAAGATAAAGAAGGGTGGCAAATTATCTCATAGAAAAAGC
 GAAAAACCTAAACTAAACCTCATCTGAAAACCTATCATTAATTCATCTGTATCAAAGAAAAAGGTT
 ACTCCTAAGAAGGCATCAGGAATGGCTCTACCAGTAGAAGATTGATTCACAGGAAAAATCTGCAT
 AAAGCACTTGATAAGAAGGCCTCAAGAAATGGGGCTTCCTCAGAGCTCCAAGGCAAACAGTTGTCA
 ACTATTGATTCTGAGGGGAATGGAAAAAATGCCGATGAAGGTGCAATTAAGTCAAGAAGAGG
 AAGCCTAAGAAAAGGCAAAGGACAAGGTAAAGCTAGATGAACCACCACGCTTGCAGAGGAGAGCA
 AGGTACTTGATGATTAAAATGAAGCTGGAGCAGAATCTTATAGATGCTTACTCTGGAGAAGGTTGG
 AAAGGTAAAAGTCGAGAGAAGATTCGGCCAGAAAAGGAAGTACTGAGAGCCAGGAAACAGATTTTG
 AAATGTAAACTTGGATTACGTGAAATAATTCGGCAGGTGGATTCTCTAAGTACAGTAGGATGTATT
 GAAGATGCTGTTATGGCTCCAGATGGATCTGTTCCCATGAACATATATCTGTGCGAAGTGCAAA
 TTGAATGAAGTTTCCCAAGATAATGATATTGACTCTGTGATGGGACATGCAACTGTGCCTTCCAC
 CAAAATGCCTTGATCCTCTTTGGATACTGAAAATATACCTCCAGGAGATCAGGGATGGTTTTGC
 AAGTTCTGTGACTGTAGGATGGAAATTATAGAAGCCATGAATGCTCATTGGGGACCCACTTCTCA
 GAGGACAGCGGTGGCAGGATATTTTCAAAGAAGAAGCAGCAGTTCCAGATGGTGGAAATATGCTA
 TTAAATCCAGAAGAAGAATGGCCTTCTGATGATTCTGAAGATGATGATTATGATCCAGAGAGGAGA
 GAGAACGTCATGAGTGGAGCAGGTACTGATGATGATGCATCTGATGACACCAGCAATTC AACCCGC
 TTGAGTTGGTCTTCAGATGGTGAGGTTTTTTTCAGGATCCAGGAGGTGGGAGGTGGACGGCTTGGAC
 TTCAGAAACAATCTATTTATAGCAGTTTAGATTCTGATGAAACCAGTGATGGGGAAATTTATTTGT
 GGCCGTAGGCAGAGAAGAGCGGTTGATTATAAGAAGTTATACAATGAGATGTTTGGAAAGGACGCT
 CCTGCACATGAGCAACCAGCGAAGATGAAGATTGGGGTCTTAGCAAAAAGAAAGCGACGAGAGAAG
 GAGTCAGATGCAGCCAGCACCTAATGACTCTATATGAAAGTAAAAGAAGATGTA AAAATGATGCA
 ACCATTGAAGGCATGATGAAGCTTCCACGAGATCCTCAAATTCGAAGGCCAATTTTCAGGCTCCCC
 CCTGATGCTGTTGAGAACTTCGCCAAGTATTTGCAGAGAATGAACTTCCATCTAGA ACTGTCAAG
 GAGAATCTATCAAAAAGAAATGGGCCTTGAACCTGGGAAGGTTAGCAAATGGTTCAAGAATTC CCGT
 TATTTAGCTCTGAAGTCTAGAAAGGTAGAGAAGGGAGAACAAGTTCATTATTCTAGTTCCAAAAGTC
 TCTGCTGAACCCACTTTAAATGTCATGAAGGACAAAACCTGCTGATCTTTCACTGTTAAAGGATAGC
 CAGGCAGAAACTGGGGTGTGTACCCAGAGAATTTGAAAAGGATCTTGCAGAGGAAGAAGCCGAGG
 TCAATAAGTAAAAGTTTAAAGAAAAATGAACAGAAAAGAGGTTCTT CGAATCACCTACTAAAAGC
 AATGAGATGAATGTTGAGCAATGATGATCTGAGCTTGAAGAAGCTGTTAAAAGCAAAAACAAG
 GGAGTAAAGAAAAGGGTAATCGTATTT CAGCAGCTGCTGAATCTGATATGGAGAACTTTGTAGA
 GCAAAGACGAGAGTAGAGAATCTGAAGCAGAACTGGTGAAACTACAACTGGCAAAGCTAGGAAG
 TCTTCCAAAATCCGTCCGCTAGACGAATCTGTTGTCTACGTTCTTATTGCTGAGCTAAGAGAAAAG
 AAATGA

SEC ID Nº: 220, secuencia de polipéptido traducida PHDf_HD_IV de *Populus tremuloides*
 MGDGKSKQDLHESPSTVNGSLLIKSLKIKKGGKLSHRKSEKPKTKPHLKT I INSSVSKKKV
 TPKKGI RNGSTSRRLIHRKILHKALDKKASRNGASSELQKQLSTIDSENGKNADEGA I KKVKKR

FIGURA 7 (continuación)

KPKKRQKDKVKLDEPPRLQRRARYLMIKMKLEQNLI DAYS GEGWKGSREKIRPEKELLRARKQIL
 KCKLGLREIIRQVDSLSTVGCIEDAVMAPDGSVSHEHIFCAKCKLNEVSQDNDIVLCDGTNCFAFH
 QKCLDPPLDTENIPPGDQGWFCFKDCRMEIEAMNAHLGTHFSEDSGWQDIFKEEAAVPDGGNML
 LNPEEEWPSDDSEDDYDPERRENVMSGAGTDDDDASDDTSNSTRLSWSSDGEVFGSRRWEVDGLD
 FRNSIYSSLDSDETS DGEIICGRRQRRAVDYKKLYNEMFGKDAPAHEQPSEDEDWGPSKRKRREK
 ESDAASLTMLTYESKRRCNDATIEGMMKLPRDPQIRRPFRFPDAVEKLRQVFAENELPSRTVK
 ENLSKELGLEPGKVKWFKNSRYLALKSRKVEKGEQVHYSSSKVSAEPTLNVMKDKTADLSLLKDS
 QAETGVCTPENLKRILQRKKPRSISKSLKKNEQKRGSFESPTKSNEMNVEHNDL SLLKLLKAKTK
 GVKKKGNRISAAAESDMEKLCRAKTRVENLKQKLVKLQTKARKSSKIRPLDES VVYVPIAELREK
 K

SEC ID N°: 221, secuencia de ácido nucleico PHDf_HD de *Ostreococcus tauri* CR954214.4

ATGAATGAGGTGCGGACGCGGAGGAACGATGTAAACGTCGTGTGCGAGCGGTGCGGGACGCGGTG
 CGGAAGGGAGGCGTGGTGTGGGCGATCGTGGGGTGGAAAGATAGTCAATGCGCGCGCGCGCGTCC
 GACGCGCGCGTGTGGAGATTGGGCGCGCGGATGTGTGTGAGANGACTGACGATGCATCGATCG
 GTGGTTGATGATACGNAGTTTGCCTCGCGGATCGCGGATACAGTTCGAGTTCGCGACGATTGCA
 AGACATCAGGCGTTGGTGGAAAGCGTACGCGAGCGACGGCTGGCGCGGACAGGCGGCTCAAAGCCG
 AAACCAGTCAGGGAGATTGAAAAGGCGCGGAGAAGATATTCGAGGGTAAATGAAAGATACGGGAG
 TATTTTAAGGTTTTGGAGTTCGATGAGCGCGAGCGGAGATCACGACGGTGGCGGACGAGTTTGGG
 GAGTGCAGCGCCGCGGATATTTTTGTAGCAAGTGCACCCTCGCAGACGACAGGCATGACGACGAT
 ATTTTGTCTGCGATGGCTTCTGCGATCGTGCATCATCAATCATGCGTGGCTCCACCGTTTTG
 GCGGAAGACATTCCGCCCGAGGACGAAGGTTGGCTCTGTCCGAGATGCGACGCGGAGTGGACGTC
 ATTTACGTCTTGAACGACGAGTACGATCAAAATTTGGGCCAGAGGTGTGTTTCAGCGGATATTTTC
 GTTGCAGAGGCGGACATGCGAGATAAAGGATCGTTCAGGAACGGCGCAGTTCAAACACGCGCAC
 GAAGAGGATTGGCCGAGCGATGAGAGCGATGACGAGGACTTCGATCAGGGCGGACACAGCGACGAT
 GGGCGAGACGACGAGCAGGCGCTGAGCGGAAGCGCACAGTCATCGAGCGATGAATCGAGCAGC
 GAATCAGAGTCTGATCTCATCATCGAGGGCCCACGAAGGCGCACGAAGGTTGATTACGTCGCCCTG
 AACAAACGCCATGTTCCGGCGACAGCGAGGCGTACGAGGGCGAAGCTGAGGAGCTCGGTTGGAAACGC
 AGCAACAAAACCAAGGCGGAGATGCTCGCGGCGCTCAAGGGTGAGAGCGCGAACGGGATAACGAAC
 GGTGCCACGAAGAAGGATCCAAAATCAAAACGCTCGAAGCGAGAGCGCGTGGATGAGAAATCTCCC
 GCACCGCGGTCAGGTTCACTCCCGATCAAAAACCTTGAGCTCGAGCGTGTCTTCTCCGAGTGC
 TCGCTCGACATTGACGCGCGGACGCGCCCTCGCGAAGCGACTCGGGATACCCGGCGGCGCAAATTCG
 ATTAATAATGTTTCATGAATCGCCGTCGTAGGGCTAAGATTTGGGTGAGAGGNAATACNGACAG
 CACGATAATCCACGTCCCGCTCGCTCGTTCGACGCTCGTCACAAACAACCTCTGAAAAACCTTCA
 AGTCACCTCAGAAGCTCACGCGGCTCTCAAACATGTCGATCGCCCCATAACGCGCCCCCAAACCTCA
 TATACTTCTCCGCGTGGTTCTGCCCGTTCGCCCATCGAGCGACGCTCGCGTTGGAGCATTACGCGC
 ACGGCGGAGAGATTTCTACGAGTGGAAGGAATCGCTCGGTGAGGTCACGCTTCGTTGACATTNGA
 CGCTCGAATGCGACGAACTTGAGCGCCTCGACNGACCGGAGACGAATCGCCGCGTAGGATGGGAA
 AAACGCAAGTCAACCTCGAGTGAGGTGAAGACGAAACACGAGAATTGGTACCATTACAAGTCCCCA
 GAGCTACTGAGACACAACCCGCTGGGCATGGTGCCGACACTCGTGACGCCGGGACGSTATGAGGAT
 AAAGATGGAGATCCATCGCGGAACGTGGTGACGGGAGCGGAGCTTGGTGTGCGTACAGTTTATCGATGAA
 CTCGTTCCGGGAAACGCGTACGCCGGGAGACGGCGGATCATGCCGAGCGACCCGTTACGCGCGC
 GCGAAGGCGAGAGTGGACCGCGACTGGGTGAACAAAACGTTATGCTCGAGGTATTATCACGTCCTG
 GTGCGACAAGAGGCGAGCGAGCAGAGAAGCGTTTTGAAAAGCTCGTGGAGGGCTGGAGACATTC
 GCGGCGTGGTGGCGAGCGGAGAAGGGAAGTTCTATGGTGGACAAACGACGCCCGGATTGGTGGAC
 TACGCGCTCTTCCATGGGCGTGGCGCTCCCGGTGTTGAGCACTACCGCGGGATGGATTTCAAG

FIGURA 7 (continuación)

ATTCCCTCGCACCGCGGCACTGAAATCGTATCACGATTGGATGGAAGCCATGCTCGCGCGAGAGCAA
 GTTCGCAAGACGCTTCCACCGTGGGACGATTACTTGGAACACATTGGACGATACGCCGACGGGAGC
 GCGCGATCGAAAGTCGCCAACGCAGTCCGCAGCGGCCGTGCGGCGCACGATTACGACGACAAGGAG
 GACAACACCGATAGTTGA

SEC ID N°: 222, secuencia de polipéptido traducida PHDf_HD de *Ostreococcus tauri*

MNEVRDAEERCKRRVRVAVADAVRKGGVWVAIVGWKDSQCARAASDARVVGDWARAMCVRXLTMHRS
 VVDDTXFALARSRIQSQFATIRRHQALVEAYASDGWRGQAAQKPKPVREIEKAREKIFEGKLIKIRE
 YFKVLEFDEREREITTVDEFGECDAAIDFC SKCTLADDRHDDDI LLCDGFCDRAYHQSCVAPPVL
 AEDI PPEDEGWLCPRCDARVDVIYVLNDEYDQNLGQRCVSADI FVAEADM RDKGIVPGTAQFKHAH
 EEDWPSDESDEDFDQGGHSDDRDDEHEALSGSAQSSSDESSSESDLIIEGPRRRRTKVDYVAL
 NNAMEFGDSEAYEGEAEELGWKRSNKTKAEMLAALKGESANGITNGATKKDKPKSKRSKRERVDEKSP
 APRVRFPTDQKLELERVFSECASLDIDARDALAKRLGIPGGANSIKIWFMNRRRRRAKILGERXYXQ
 HDNPRPARSFDASSQTTSEKPSHLLRSSRASQTCRSPLNAPPNSYTS PRGSARSPIERRSRWSITR
 TAERFPTSGRNRSVRSRFDIXRSNATKLERLDXPETNRRVGVWEKRKSTSSEVKT KHENWYHYKSP
 ELLRHNP LGMVPTLVTPATYEDKDGDP SRNVVTE SLVCVQF IDELV RGNAYAGETAAIMPSDPYAR
 AKARVDADWVNKNVCSRYHYVLRQEASEQREAF EKLV EGLETF AAWCGSGEGKFYGGQTPGLVD
 YALFPWAWRLPVFEHYRGMDFKI PRTAALKSYHDWMEAMLAREQVRKTLPPWDDYLEHIGRYADGS
 ARSKVANAVRSGRAAHDYDDKEDNTDS

SEC ID N°: 223, secuencia de ácido nucleico parcial PHDf_HD de *Aquilegia formosa* x *Aquilegia pubescens* cóntigo de DR914726.1, DR941696.1, DR943570.1

ATGCAAGATGGTGGGAACCTTGGTTTTCCAAAGAAACGTGTTGGATCCCAAATAAATACCTCATTT
 CAGAAGAAGAATAGTTGCAAAAATCCGCATATTAGGAGATGGAAACCAAAGCCAAGATCTCATGTC
 AAGACAATTGCTTCGTTCTGTGCGAAGAGAAAGACATCTGAGCTCTCAAGCAACAGAGCCAGAAAT
 AAATGTTCCAATGGAAAACCTTTAGAGAAGAGCGTCACAGAAACCAATGATACCAGTCCTTCT
 GATAGGCCACTTTCATCAATGCTTAAAGATTTCAAATCGTTAAGCAATAAATGTCAGGGACGTGAA
 AAGACAACCTGACAAGGATCTTGGAAAGTCAAAGGCTCACAGAAGGAGAAGGAGCAAGAAGAGGAAG
 AATGAGGAGTTAGATGAAGCCGCTCGAGTCAAAGAAGAACAAGGTACCTCCTGATTAAAATGAAG
 CTGGAGCAGAACCTTATTGAGGCTTACTCTGGGGAGGGTTGGAAAGGCCAGAGTCGTGAGAAAATT
 AAGCCAGAAAAGGAACTGCAAAGAGCCAAGAAGCAGATATTGAAGTGCAAACCTTGGAAATTCGTGAT
 GCTATTCGTGTCAGCTGGAATCTCTTAGTTCTGAGGGGAAGCATTGAAGATTCTGTGATAGGTTCCAGAT
 GGATCTGTTTTCCATGAACATATTTCTGTNNNNNNNNNNNNNNNNNNNNNNNAGCATTTCAGAT
 AATGACATTATACTGTGTGATGGAGCATGCAATTGTGCTTTCCACCAGAAATGCCTGGAGCCTCCA
 TTAGCCACTGAAAATATTCCTCCTGGTGATCAAGGGTGGTATTGCAAATTTTGTGAGTGCAAGATG
 GCAATACTTGATGCAATCAATGCACATCTTGGGACTCAGTTTTCTCTAGACAGTAAATGGCAAGAT
 ATTTTCATGGAAGCAGCTAATGCACGGATAATGAGGATATCTCAATTCATCCTGATGGTGACTGG
 CCATCTGATGACTCTGAAGATGTTGATTATGACCCAGAAAAATATGAAATTGATTGCAGCTATAGC
 AGAATGGGCTCGGAGCACAACGTGTCTGACGATGCAAGTAGTCCGGCACCTTGATTTGGACTTCC
 GAAGAAGATGATGGTTTGCATTCTGAGCGATTCCAGAACTGATGGTAGTCTGGAGGGGGGATTTTAT
 GGGGATAAACTAAAATCCAGAGTTCTGATAGCATAGATTCCAATGACAGTACTGATGGTATGGCA
 TTTTCATCGTAGGCAGCGAAGAGATGTTGATTATAAGAACTGCATGATGAAATGTTTGGTAAAGAT
 ATGCCTGAAAGTGAAGAAATCAGTGAAGATGAAGATTGGGGACCTAGAAGAAGAAAGCCAGAAAGA
 AATGAGTCTAATGCAGCCACTACTTTAGTATCTTCGTGTGGTACTGAGAATGAATTTTCGAATGTG
 GCACTTNNNTACCATCAAGAGTGGTCAGGGAGAGTCTCTCGAAGCAATTGGGTATTGCAATTGAG
 AAGATCAACAAGTGGTTCAAGAATGCTCGGTATACAGCTTTAAAAATTAGAAAGGCAGAAAAGACT

FIGURA 7 (continuación)

GAACAATCCCCTGGAGGTGGCAAGATGCAAAGTGCTGACGTATTAACATCAAAGGAAAACCATTAT
 TTGGTCTCAACAGGAACAGCAGTCCAAAAGTTCGAGTGCCATAGAAGTAAAAACCGT
 AAATTGACAACCACACCCTCAAGGGAAAGTACAAGAGAAAAGTGTCTGCTGTTGCACCATCCGCTACT
 GCAAATGAGGTGAGTGTACAGTTAAGAAAAGCAGTGAGCGTAGAGAAGGAAATGGGCAGTGTAAAA
 AGAAAATCTGCACAAGTAAAGAGGATCAAGCGCACATGCTCCACTAAGGAGCAGCAGCTATACATG
 ATTGAGTTGGAGAGAATTTGCCGCTTTGAGGAGAAGCTTGAAAAGATGAAGAAGGTATTGCTTAGT
 ATAGATGACAATCATAGTAATGTATCAGATGGACCCCATATGAACGAACAGTCTGCCGTATACATT
 CCAGTTGCAGAGCTAAGAGAAAAGATGTGA

SEC ID N°: 224, secuencia de polipéptido parcial traducida PHD_HD de *Aquilegia formosa* x *Aquilegia pubescens*

MQDGGNLFPPKRRVGSQINTSFQKKNCKNPHIRRWPKPRSHVKTIASFLSKRKTSELSSNRARN
 KCSNGKPFRRRASQKPNDTSPSDRPLSSMLKDFKLSNKCQGREKTTDKDLGSQKAHRRRRRSKKRK
 NEELDEAARVQRRTRYLLIKMKLEQNLIEAYS GEGWKGQSREKIKPEKELQRAKKQILKCKLGIRD
 AIRQLESLSSEGSIEDSVIGSDGSVFHEHIFCXXXXXXXXXAFPNDIILCDGACNCAFHQKCLEPP
 LATENIPPGDQGWYCKFCECKMAILDAINAHLGTQFSLDSKWQDIFMEANARDNEDISIHDPGDW
 PSDDSEVDVYDPEKYEIDCSYSRMGSEHNVSDDASSSGTLYWTSEEDDGLHSEFRFDGSLGGFY
 GDKTKIQSSDSIDSNDSIDGMAFHRRQRDQVYKLLHDEMFGKDMPESEEISEDEDWGP RRKRKR
 NESNAATTLVSSCGTENEF SNVALXXPSRVVRESLSKQLGIAIEKINKWFKNARYTALKIRKAECT
 EQSPGGGKMQSADVLTSKENHYLVSTGTAVQRSKVS SVHRSKNRKLTTTPSREVQEKVSAVAPSAT
 ANEVSVQLRKAVSVEKEMGSV KRKSAQVKRIRKTCSTKEQQLYMIELERICRLEEKLEKMKVLLS
 IDDNHNSVSDGPHMNEQSAVYIPVAELREKM

SEC ID N°: 225, secuencia de ácido nucleico parcial PHD_HD de *Glycine max* cóntigo de GM06LC25006

AAGAAGAAAAGACAAAAGGAATAACATAGATGTTGATGATGCTTCACGCTTGCGAAGGAGAACAAGG
 TACCTCTTAATCAAAAATGAAGCTAGAGCAGAACCCTTATTGATGCTTACTCTGGAGAAGGTTGGAAA
 GGTCAAAGTCGGGAAAAAGATTAGGCCAGAAAAGGAGCTACTACGAGCAAAAAGCAGATTTTAAAA
 TGTAACCTTAGTATTCGCGATGCCATACACCAGCTGGATTCTCTTAGTTCTGTGGGTAGCATNGAA
 GATTCTGCCATTGCTCCAGATGGATCTGTTTATCATGAAAATATATTCGTGCCAATTGCAACTA
 CATGAAGCTTTCCAGATAATGATATTACTCTGTGATGGCACTTGCAATCGTGCTTTTCACCAA
 AGATGTCTCAATCCTCCTTTGGATACTGAAAATATTCCTCCTGGAGATCAAGGCTGGTTTTGCAAG
 TTTTGTGAATGTAAGATAGAAATACTAGAGGCAACAAATGCCCATCTTGGGACTCAATTCTCCCTA
 GACAGTAATTGGCAGGATGTATTCAAGGAAGAGGCTGCTATGCCTGATGGTGACATTGCATTA
 AATCCTGAAGAAGAATGGCCATCAGACGATCCTGAAGATGATGATTACAATCCAGAGAGGAAAGAA
 GACAACCACAACCTTTGACACAGAAGGAGCTGATGAAAATGCATCCAATGATTCAACCAGTTGTTCC
 AGTCTGTTGTCTTTGAATGGCGAATGTCCTCCAGTAGATGAAGGCATCTGTCATGAATATTATTCT
 GTTAATAGCTGCTTAGATTCTGATGAATCTGGGAAAATAGCATGTGGTCCTAGGCAGCGTAAAGCT
 GTTGACTACAAGAACTCTATGATGAAATGTATGGGAAGGATGCTCCACCTTGTGAACAAAATGAGT
 GAAGATGAAGACTGGGGTCTTGCCAAAAGAAAGCGAAGAGAAAAGGAGTCTGATGCTGTTGATTCT
 CTTATGACTTTGATGAAAAGTGAGAAATAGGCATCCCAATAATGAGCACAACATGACAAGTAAAGAT
 TCTTCAAGCATAAAGATCAAAAAGGCATTTGTTTFFAGGATTCCACGTGATGCAGTTGAGAGGCTTCGC
 CAAGTTTTTGTGATGAGAACGAACCTTCTCCAAGATCTATAAGAGAGGGTCTTTCAAAGAGCTGGGC
 CTTGACACCGAGAAGGTCAGCAAATGGTTCAAATAATGCACGTTACTTAGCACTTAAAAACAGAAAG
 CATCAAGCAGAAGGAGGAGCAGATCAGCTTCAGAGTATCACTTCTACAAAGAACAGATTACAAAA
 CAGGAGAATGTGCATCCTTTGAAATCAAAGACCCCAAAAATTAAGTAGGACACATCCCAGAAGGAT
 GTTAAAAATGTCAATGGAAGAAAAGAAAATGAAGTCATCTAACAAAGAAAAGACAACCAGAAATTCCT
 CCACCTCCAGGAGAAAATGGCAAGAAGGACTTTATGGAGATCAGTATGATGTGAGCTTGAAGAAA

FIGURA 7 (continuación)

CTGTTGAAGAAAAGAAAGAAGAGGTTGATAAATTTTACGTTTGGAGGAGACTCTCAGTTGGCAGAG
 TTGGAATTTGAAAGACTAAGCGAATTGAAGACTAAAGTAGATAGTATGAAGCAAAAATTAAGTCA
 ATTCAAATTAACAGAGTTAAGGGTTCACCCTATTCGAATGAACCATCCATTGTATATGTACCCACA
 GCTGTGTTAAGGGAAAAGGTTGAATGA

SEC ID N°: 226, secuencia de polipéptido parcial traducida PHD_HD de *Glycine max*

KKKRQRNNIDVDDASRLRRRTRYLLIKMKLEQNLI DAYSGEGWKGQSREKIRPEKELLRAKKQILK
 CKLSIRDAIHQLDSLSSVGSXEDSAIAPDGSVYHENIFCANCCKLHEAFPNDIILCDGTCNRAFHQ
 RCLNPPLDTENIPPGDQGWFCCKFCECKIEILEATNAHLGTQFSLDSNWQDVFKEEAAMPDGDIAL
 NPEEEWPSDDPEDDDYNPERKEDNHNFDTEGADENASNDSTSCSSLLSLNGECPVDEGICHEYY
 VNSCLDSDESGEIACGPRQRKAVDYKKLYDEMYGKDAPPCEQMSSEDEWGPGRKRREKESDAVDS
 LMTLHESENHRPNNEHNMTSKDSSSIKIKRHC FRI PRDAVERLRQVFAENELPPRSIREGLSKELG
 LDTEKVS KWFKNARYLALKNRKHQAEGGADQLQSITSTKNRLQKQENVDP LKSKTPK ITRTHSQKD
 VKNVNGRKKMKSSNKKRQPEIPPPGGENGKKDFMEISDDVSLKLLKRRKRLINFTFGGDSQLAE
 LEFERLSELKTKVDSMKQKLTAIQNYRVKGPSYSNEPSIVYVPTAVLREKVE

SEC ID N°: 227, secuencia de ácido nucleico parcial PHD_HD de *Lotus corniculatus* AP004517

ATGCAGGGTACTGAGAAGTTAAATGGCACAGGATCTACAAAATCCAGTAACTCAGAGGAACATGCT
 GAGTCAAAGGTAGATTCATTGAGATACAGAACAAATATAACAAAATGCCGAGGGAAGAAACAAAA
 CTGAAATCAAATCTCATAAACTCAGCGGTTGCTTAAGTGCATCAGGGAGGACAGTTACCAATTCT
 TCCATCAAGAGACAAGTAAAGGATTCTTCTAATAAAAAGATTATTAGACAAAGTCTACATAAACT
 GATGGCAAATCTTCACGGATGCTGTCTTCAACAATGATACAAGGAGGAAAATCTTCTCTTGGTCT
 AGAAAGGAGGGCAAAGATGTTGATGGGGAGGTAATGGTTCAAAAAGTTAAAAGAAAGAGAAAGAAG
 AAAAGGAGACAGGATGATGTTGATCTTGATGATGCTTCACGCTTACTGAGAAGAACAAGGTACCTC
 TTAATCAAATGAAGCTAGAGCAGAATCTTATTGATGCTTACTCTGGAGAAGGTTGGAAAGGTCAA
 AGTCGAGAGAAGATTAGGCCAGAAAAGGAGCTACAAAGAGCCAAGAAGCAGATTTTGAATTGTA
 AACTTAGTATTCGGGATGCCATTGCGCAGTTAGATTCTCTTAGCTCATTGGGCAGCATTGAAGATTCT
 GTCATTGATCCAGATGGATCTGTTTATCATGAACATATATTCTGTGCAAATTGCAAGTTGCGTGAA
 GCTTCCCAGATAATGATATAATACTCTGTGATGGCACATGCAATCGTGCTTTTACCAACGGTGT
 CTCAACCCTCCTTTGGATACTGAAAATATTCTCTGGAGACCAAGGCTGGTCTGCAATTTTTGT
 GAATGTAAGATAGAAATATTGGAGGCAACAAATGCCCATCTTGGGACTCAAATCTCCTTGGACAGT
 ACTTGGCAGGATGTATTCAAGGAGGAGGCTGCTATACCCGATGGTGATAATGCATTACTGAATCCA
 GAAGAAGAATGGCCATCAGATGATCCTGAAGACGATGATTATTATCCAGAGAGGAAAGAAGACAAC
 CACAGCATCAAAGCAGAAGGAAGTATGATAATGCATCCAATGATTTAAGCAGTTCTTCCAGTCTA
 GGGTCTTTGAATGGTGAATGTTCTCCAGTAGATGAAGGCACGGGTCTTGAATATTATTCTGTTAAT
 TGCTGCATAGATTCTGATGATTCTGGGGAAATAGCATGTGGCCGTAGGCAGCGTAAATCTGTTGAC
 TACAAGAACTATATGATGAAATGTTGGGAAGGATGCTCCTCCATGTGAACTAGTGAGTGAAGAT
 GAAGACTGGGGTACTGGCAAAGAAAGCGAAGAGAAAAAGAGTCTGATGCTGTTAATACTCTCATG
 ACTATGCATGAAAGTGAGAACAAAAATCCCAATAATGAGAATAGTGATGGAATAAGAGAGGGTCT
 TCAGGCAAACCAATAAGAAGGCTTGTTCAGGATCCACATGATGCACTTGAGAAGCTTCGCCAA
 GTTTTTGGCGAGAATGAGCTTCTCCAAGATCTGTCAAAGAAGGCTTTCGAAAGAGTTGGGAATT
 GATGCTGCGAAGGTTATGCCTTACAACCTC

FIGURA 7 (continuación)

SEC ID N°: 228, secuencia de polipéptido parcial traducida PHD_HD de *Lotus corniculatus*
 MQGTEKLNLTGSGTKSSNSEEHAESKVDLSLRRTNITKCRGKKQKLKSKSHKLSGCLSASGRVTNS
 SIKRQVKDSSNKKIIRQSLHKTGKSSRMLSSTMIQGGKSSLGSRKEGKDVDGEMVQVKRKRKK
 KRRQDDVLDLDDASRLLRTRYLLIKMKLEQNLI DAYSGEGWKGQSREKIRPEKELQRAKQIILNCK
 LSIRDALRQLDSLSSLSGSIEDSVIDPDGSVYHEHIFCANCKLREAFPDNDIILCDGTCNRAFQRC
 LNPPLDTENIPPGDQGWFCNFCECKIEILEATNAHLGTQISLDSTWQDVFKEEAAIPDGDNALLNP
 EEEWPSDDPEDDDYYPERKEDNHSIKAEGTDDNASNDLSSSSSLGSLNGECSPVDEGTGLEYYSVN
 CCIDSDDSGEIACGRRQRKSVDYKKLYDEMFGKDAPPCELVSEDEDWGTGKRKRREKESDAVNTLM
 TMHESENKNPNNSDGIREGSSGKPIRRSCFRI PHDALEKLRQVFAENELPPRSVKEGLSKELGI
 DAAKVMPLYNL

SEC ID N°: 229, secuencia de ácido nucleico parcial PHD_HD de *Sorghum propinquum* BF656332
 AACTCGAGAACAATGAAGAAAAATCTGGTATATCAGGGAAGATTGATTGAGTAGACAACCTTTGC
 TTGGTTCCCTTGTCTGAAATCATCAATGTGCATACACGAGTGCACCACAATCTTGAGATGAGGAAG
 ATGGAATCAACTAGCAGTCTGTGTGGCTGCACAATGAAGGAGGCTGCTTGTTCACACTCCAA
 GCCAAGGAGAGTACATTACCTACAAGCAAGCCTTGTTCGCGTCAGAAATAACCCATCCAACAAC
 AATGAAGTGAGCACTTTAGTGCAAGCTACTTCTTGATGGATGCTGGGGCGTGCATTGAGCAGCAA
 GAAACCACCTCTGGGTGGATACCGGGCCCTCAGGTTACCAGCCTTCTCCTGGACGTGATTGATGAG
 ATGTGCGGACTTGAGTGCAGGCTGCAGAGGCTGAAGGAGAACATGTTCTCATCTGGCATGGACGGC
 AGAACCGCCGGTGTGAGTGCATGGGAAACCAGGCTGTGGTGCTTGTGCCGACTGCTGAGCTCAAG
 GAAAAAGCGCCGATGGCGGTTTATTTGGGCACACTATTGCCCATAG

SEC ID N°: 230, secuencia de polipéptido parcial traducida PHD_HD de *Sorghum propinquum*
 NSRTNEEKSGISGKIDSVDNSCLVPLSEIINVHTRVHHNLEMRKMESTSSPVWLHNEGGLFPTLQ
 AKESTLPTSKPCLPSEITHPTTNEVSTLVQATSWMDAGACIEQQETTPWVDTGASGYQPFLDVIDE
 MCGLECRQLQRLKENMFSSGMDGRTAGVSDMGNQAVVLPVPTAELKEKAPHGGLEFGHYCP

SEC ID N°: 231, secuencia de ácido nucleico parcial PHD_HD de *Sorghum propinquum* BF704605
 AAGATAAAATTGGAGCAGAATCTACTAGATGCTTACTCTGGAGATGGATGGAATGGGCAAAGCCGG
 GAGAAAATAAAGCCAGAGAAGGAACCTGCAGCGTGCCAGGAAACAAATCATAAAATGCAAATTGCT
 ATACGTGATATCATTCGCCAACTTTGTTTGTATACTTCTACTGGGAGTGTGATGACCCAGCAATG
 CCTCCAGATCAGTTTACCAATCCCGAACATACCATGTGCTCAACATGCAAGTCTCATGAATCATTT
 CCCAGCAATAAAATTTATCTTCTGTGAAGGGCCCTGCAAAGGGCATATCATGAGAAATGTTTGGAA
 CCTCCCTTAAACAAAGGTGTACTTCCAACAAGTAGCCATGGATGGCTTTGCAAATCTGTTTGTGC
 AAGGTGAAAATTTAGAAACTATTAATGCACATCTAGGAACAAGTTTTACAGTGATGTGCTCCTTT
 GAAGATATATTCAAGGAAGCAACAGAACAATAGACTCCGAGGATGCACTAGATGAAGATTGGCTC
 TCTG

SEC ID N°: 232, secuencia de polipéptido parcial traducida PHD_HD de *Sorghum propinquum*
 KIKLEQNLLDAYSGDGWNGQSREKIKPEKELQRAKQIIRQLCLYTSSTGSVDDPAM
 PPDQFTNPEHTMCSTCKSHESFPSNKFIFCEGPCKRAYHEKCLEPPLNKGVLPTSSHGWLCKFCLC
 KVKILETINAHLGSFTVMCSFEDIFKEATEQIDSEDALEDDWLS

FIGURA 7 (continuación)

SEC ID N°: 233, dominio cremallera de leucina/dedo de homeodominio de planta (ZIP/PHDf), comprendido en SEC ID N°: 180

DEVSRMEKRARYLLIKIKQEQLLDAYS GDGWNGHSREKIKPEKELQRAKKQIMKYKIAIRDVIHQ
LDLCS SSGSKDDSVIPP DGCHE SVNPEHTICSRCKSHESFPDNNII FCEGGCKLACHQKCLEPPFD
KILPTTRHGR LCKHCSSKMKILDAINAHLGTSFTVKCPSSDIFKEAAE

SEC ID N°: 234, homeodominio, comprendido en SEC ID N°: 180

RIPPA AVEVLRKAF AENELPARSVKENLSTELGISFEKIDKWFKNTRCAA

SEC ID N°: 235, Promotor GOS2 de *Oryza sativa*

AATCCGAAAAGTTTCTGCACCGTTTTACCCCCTAACTAACAATATAGGGAACGTGTGCTAAATAT
AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAACTCATCCACCTACT
TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCTTAGTAATTAAGT
GGGAAAATGAAATCATTATTGCTTAGAATATACGTTACATCTCTGTGTCATGAAGTTAAATTATTCG
AGGTAGCCATAATTGTCATCAAACCTCTCTTGAATAAAAAAATCTTCTAGCTGAACTCAATGGGT
AAAGAGAGAGATTTTTTTTAAAAAATAGAATGAAGATATTCTGAACGTATTGGCAAAGATTTAAA
CATATAATTATATAATTTTATAGTTTGTGCATTTCGTCATATCGCACATCATTAAAGGACATGTCTTA
CTCCATCCCAATTTTTATTTAGTAATTAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG
ATTAGATGCAAGGTA CTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT
ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATACATTTAGGTAGCAATATC
TGAATTCAAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT
ACAGAATAGCATGAAAAGTATGAAACGAACTATTTAGGTTTTTCACATACAAAAAAGAAATT
TTGCTCGTGC GCGAGCGCCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCACA
GAACAACCCACAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG
GCTTTGCGGCCAGGAGAGAGGAGGAGAGGCAAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA
ATTCCTCCCCCTTTTTCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACCAG
GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTTCGAGTTCTTGGT
CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCCTTCGGTTGTTCTTGGATTT
ATTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATTCC
TGTTCTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTTCGGTTTGATTAGTAGT
ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTAGGGATCGGAATCTTGCAT
TTGTGAGTACCTTTTGTGTTGAGGTAAAATCAGAGCACCGGTGATTTTGCTTGGTGTAAATAAGTAC
GGTTGTTTGGTCCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC
CCTATTGAACAAAAATAATCCAACCTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCTT
AAGCCTGTCCAAAATTTTCGAGCTGGCTTGTTTAGATACAGTAGTCCCATCACGAAATTCATGGA
AACAGTTATAATCCTCAGGAACAGGGGATCCCTGTTCTTCCGATTTGCTTTAGTCCCAGAATTTT
TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTTCAGTTCAATGAAATTGATTGCTACAAATAATG
CTTTTATAGCGTTATCCTAGCTGTAGTTTCAATAGGTAATACCCCTATAGTTTAGTCAGGAGA
AGA ACTTATCCGATTTCTGATCTCCATTTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT
CATTTGGATTATTTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA
CTGTCCCTCAATTTTGTGTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT
AGAAGTTCTTTTTGGTTATTCTTACTGCTTGTGATTACAGAAAGAAATTTATGAAGCTGTAATCG
GGATAGTTATACTGCTTGTCTTATGATTCAATTCCTTTGTGCAGTTCTTGGTGTAGCTTGCCACT
TTCACCAGCAAAGTTC

SEC ID N°: 236, prm09687

GGGGACAAGTTTGTACAAAAAGCAGGCTTAAACAATGAATACCCAGAAAAGAAA

FIGURA 7 (continuación)

SEC ID N°: 237, prn09688

GGGGACCACTTTGTACAAGAAAGCTGGGTGATGCAAGGTTAAGATGCTTT

SEC ID N°: 238, secuencia de ácido nucleico Phypa_PHD-HD de *Physcomitrella patens*

XM_001762483.1

ATGATGAAGAATTACCGTCCCTCCATCACCCTTCCAAGTGCCTCGCCAAGCTTTCTTACAAGACCCG
 GCGTCGTCAATCCAAATGACAGGACGACATTTGGGCAGACCCAGAGGGAGGTGGAGAACGGAAGCTT
 CACACACACTCTCACTCACACTCACTCACTCACGAAGGAAACGAAGGGCAAAGGAAGGGGAGCCCC
 GCCGTGAGCCGAGAGCGGAGAGCAAGCATTTCCCAACGGAAGTCTCCGTCAATGCGGTGCCGGAGCC
 GACGACGACGACGATGATGGGCTGAGCCTGTATCGCAATGGATGCATTTGCCCTGTAAGTGGGAAG
 GTGGAGGAAGGATTGGGGAAGAGGTTGAGGAAGAGAAAGAGAAAGAGAAAGAGTTTGTGTTGTTG
 TTGGGGAAGGGGAGGGGCGACGGATGATGCCCATGGGCGTAGTCCGCGACGATGAGTCCCAGGCC
 GCTGCTGGCTCGCCGCTGCGCCACCCCTCCCCAATCGCCAAGCCGACGCTGCGATCCCCTCG
 CGCTCCAGGGCGCACAAAGGCGCGCTCCCAACACCCGCGCTGGGCGATGGCGATGCGCCCTTGCCC
 CCACCGCAAGGAACAGCAACACCAGGAGGAGGAGGAGCAGCAGCAGTGGGCGCACAGGGGCCATG
 GATGCCGAGGCTATGGCGGCGGGGAGCGGGCGGCTGACGACGCGGACGACAGCGGGCAAGAAC
 TCGGATGTGCCGCTTGACAATCGCACGCGAATTAAGAGGCGTGTCAAGTACTTGCTCATGAAAATG
 CGAGTCGACCAGAACCTGCTGGATGCGTACTCCGGGGAGGGTTGGAAGGGGCAGAGGAATTTCTCTT
 TTTGGCATATCCAGTCGGGAAAAGATCAGGCCGGAGCAAGAACTACAGCGTGCTGAAGCGCAGATC
 CTGCATTCGAAATTTGCCATACGGGAGGCTATTCATGAATTGGATAACTTGGGGTTAGAAGGCTCG
 CTCGACGAAAATGCATTTGATGCTGAAGGCCGTGTTTACCATGAAGAGATATTTTGTGCCAAGTGC
 CGATCTCAAGATGCTCTACCTGACAACGACATTATTTTATGTGACGGCGCATGCAACAGAGGCTTC
 CATCAGTACTGTTTAGACCCACCTTTAGCTACGATCAACATTCCTCCTGGAGACGAAGGGTGGTTG
 TGTCTGTATGTGAGTGCAAGATGGAGTGCATGAAGTTATTAATGCATACTTGGGCACACACTTT
 GAGGTGGAGAATAGTTGGGAGGAGTTGTTTTCTGAAGAAGCTCTCGTGGCTGCGGGCGGTAGA
 CAAGGGGTACAGATGGTGATTTTCTTCTCAGACTCAGAAGACGACGATTATAACCCTGAAGGA
 GCGGAGAAGCGGACAGACAGTGAAGCGAAGGCGAAAGTGAAGGTTGGGGGAGATAGTGATTCA
 GGAAGCGATTTTAGGAAAATGAAGCGTCTGCTAGTGACAAAGATGAGCAAGATCCAATGCAGGCA
 ACTAATTTAGAGAGACTTGCAGAACAGCAAAGCGGACCAGCAAGGAAGGAAGTGGAGGTCCAGAA
 TCTTCAGGCAGTGATGACTCAGAAGTTTTTCTGTTGATGGGTTTACGGATGGTGGAGTCAAATTA
 GGTTFAGACGTAGAGACTCACAAAGCAGAAGAATCAGCTTCAGACGAGGAAACCATGGTTATAGCA
 GGAAAAAGACATCGAAAAGCAGTAGACTACAAGAGATTACATGATGAAATGTTTGGGAACATGGAG
 ATTGATGTGGACGATATCATATCCGAAGATGAGGATTTGGGGTCCGAAACGCCGTGCGGAGAAGAAC
 ATTCTGATGACCCAAGCAGGCTTAGACCTCCTCGAGTTTCGACGGAGAAAGGCGCGAACCTCAACT
 GGAGATGCTTTAAAAGAAAGTTGATGCAGATGGGGTACCATCAGATTTACCCCTCTGGGTACAAAGG
 GAGGGAGCAGCCGATACTCTTGAAGCTGTTAGAGACAAACGAATGTGGAGGAAGTTGCCTGATTCT
 GCTGTTGAGACGTTACGACTGGCTGCGGCAGTGACGACACTGCCCTCAAAATCCCGTAAAGAGGAG
 TTGGCGACACAGCTTGGCTTGGTTTTTACAAGTCCATGGATGGTTTTAAAATCAGCGGTATTTA
 GCTCTCAGAAATGGGTTGGCGATTTCAAAGCGACCAGCAGCTGGGCGGATGGCGGCCACTGTAACC
 ATACAATCGGAACACGACAACAATTCAAGTGCCTATTTAACTGAGAAGTTGGATGAAGTCCAAATG
 CGACTTTTGAAGTGAAGCAAACTTTGAAGATTTTCGCTAAGAATTCATCTGGTTTAGGCGGTGGA
 GAATCAGATAACAGTTTTGAAAATGGAGGAATTGGACTTCCTGCAAATGGGAAACGGATTATCTAT
 GTCCAGTGGCTGAGGTTAGAGAACGGCCTTCTCTTTCTCTGACAGCTGTCCGCTTTTATTATTA
 GCAGTGTAAATTGCGAGGTGAATTACAAGAGGAATCTACGTATATTTCATGTGTTTCATGGCCTAGC
 TGCTTAGCGTACGGGTTTTGCATGCCAACTGCCCTGTTCATATTGGATCATGTTTCAAGGAATA
 GGTGGGTGA

FIGURA 7 (continuación)

SEC ID N°: 239, secuencia de polipéptido traducida Phypa_PHD-HD de *Physcomitrella patens*
 MMKNYRPPSPPLPSARQAFLODPASSFQMTGRHWADPEGGGERKLHTHSHSHSLTHEGNEGQRKGGSP
 AVSRERRASIPNGSLRQCGAGADDDDDGLSLSSQWMHCPVSGKVEEGLGKRLRKRKRKRSLLLLL
 LGKGRRRMMPMGVVRDESQAAAGSPPAPTPPPIAKPTLRSHSRRAHKARSQHPPLGDGDAPLP
 PPQGTATPGGGGAAAVGAPGAMDAEAMAAGQRRRRRRRQRGKNSDVPLDNRTRIKRRVKYLLMKM
 RVDQNLDDAYSSEGWKGQRNSLFGISSREKIRPEQELQRAEAQILHSKLAIREAHELNDNLGLEGS
 LDENAFDAEGRVYHEEIFCAKCRSQDALPDNDIILCDGACNRGFHOYCLDPPLATINI PPGDEGWL
 CPVCECKMECIEVINAYLGFHFEVENSWEELFSEEALVAAGGRTQGVTDGDFPSSDSEDDDDYNPEG
 AEKRTDSESEGESDEGGGDSDSGSDFRENEASGSDKDEQDPMQATNLERLARTAKRTSKEGSGGPE
 SSGSDDSEVFLVDGFTDGGVKLGLRRRDSQAEESASDEETMVIAGKRHRKAVDYKRLHDEMFGNME
 IDVDDII SEDEDWGPKRRTTIPDDPSRSRPPRVRRRKARTSTGDALKEVDADGVPSDLPSGSQG
 EGAADTLEAVRDKRMWRKLPDSAVETLRLAAAVTTLPSKSRKEELATQLGLSFSQVHGWFKNQRYL
 ALRNGLAISKRPAAGRMAATVTIQSEHDNNSAYLTEKLDEVQMRLFELKQTLDFAKNSSGLGGG
 ESDNSFENGGIGLPANGKRIIYVPVAEVRERPSLFSDSCLPLSLLAVLNCEVNYKRNLRIFMCSWPS
 CLAYGVLHANCPCHIGSCFEGIGG

SEC ID N°: 240, secuencia de ácido nucleico Phypa_PHD_HD II de *Physcomitrella patens*
XM_001779822.1
 ATGGTGCCGACGGGCGTGGTGGAGCGGGAGGACCAGCCTGAGCCTGCTGCTGCCCCCAGTGCCG
 GTGATTACACCCAAGACGCTTCGATCGCACTCTATCAGGACGCAGACCACGCCAGGTGATGGCGAT
 GCTGGTGCAGTATGCTCCTGGAGCAGTGGCAGCAGCAGCAGCATCAGCAGGAGGAGGAGGAGCA
 CGGATGTGATGCAAGGTAACGATGCGGGGCAGGAACGGCGAGGACGACGGCGCGGGCGACGGAAG
 AAGTCAGACGTGCCGCTCGATAATCCCACGCGAATCCTAAAGCATGTCAAGTACTTGCTCATTAAA
 ATGCGAGTCCACCAGAACCTGCTCGATGCGTACACAGGGGAGGGTTGGAAGGGTCAGAGTCGCGAG
 AAGATCAGGCCCGGAGCAAGAACAACGTCGAAGGCGCAGATACTGCGATCTAAGTTGGCCATA
 CGGGAGGCTATTCATGAACTCGATGACGTGGGTTAGAAGGTTCTCTCGACAAAAACGCTTTTGAT
 TCTGAAGGTCGTATTTACCATGAAGAGATTTTTTGTGCGAAGTGCAAATCCCAAGAAGCTCTTCT
 GACAACGACATTATCTTGTGCGACGGGGCGTGCAACAGAGGCTTCCATCAGTATTGCTTAGACCCA
 CCTTTAGCTACGGAAGACATTCTCTGGAGATGAAGSATGGTTATGCCCTGTATGCGACTGCAAG
 ATGGAGTGTATTGAAGCGATAAATTCATATTTTTGGCACCAGTTTTGAGGTCGAGAATAGTTGGGAG
 AGCTTTTTTTGCAACGAAGCTGGTATAGCTGCAGGCGGTGGAACGCAAGAAGGGGCTGGTGGTGAT
 TGGCCTTCTCAGGCGATGAAGATGATGAGTATGATCCAGAAACAGCGGAAGGGCCTCTTCAGCT
 TCTGAAAGTGGAGCTCAAGTAATAGGGCTTGCTGTTGACTGGCCGCTCAGACGACGAAGATCAC
 GACTATGATCCGGAAGCAGCGGAAGAACCTCCTGGAGCTTCTGGGAGTGGAACCAACAGTAGGT
 CTTGACGTTGATTGGCCTTCCACAGACGATGAAGATGACGACTTTGATCCAGAAGGTATCAGAAAG
 CAGGCAACAGTGACGAAGAAAGTAGTAGTGATGAAAGTGAGGAGGACGGGGAGGATAGTGGTTCT
 GTAAGCGGCTTCAGGAAAATGAGGCGTCTTCAGGTGTAAGTGACAAAGATGAGCAGGAGCCTAGT
 GAAAAACATAATCTGGAGCTAACTGCTGTCACTAACATCAAGAAACTTGCAAGAAAAGGAAAACGT
 AACATAATCAGCAACAAAGGAAAAAGTGAATATTTAGAGTCTTCAGACAGTGATGATTCTGATGTC
 TCTCTTTTTGACGAGTTCAAGGGTGGTAATGCCAATTTAGGCTTAAGACGTGGGGATCCACAGGCA
 GCAGTATCAGCAGCAGCAGGGAAGCCGTGATTATTGCTGGGAAAAGACATCGGAAAAGCTGTGGAC
 TACAAAAGACTACATGATGAAATGTATGGTAAAGCGGAGGCCGATAACGATGACGACGTTATCTCT
 GAGGATGAAGATTGGGGTCTGAACTGCGCCGAAGACGAACCATTCCTGATGATCCTAATCGCCA
 AGGTCCCGAATTCGAGGTAGGAAAAGTGAGAACCTCAAGTGGAGATATTTCTAAAGGAGCCGATGCA
 GATTTACCCTCTGGATTACCAGATATGCCATCCCCTTTCCACTGGTTTACAAGGTGAAGTAGGC
 TCACCACAAGACATTCAGCACTAGATGGAGCAGCTGACAGTCTGGAGGAGAAAAGCGGATGTGG
 CGGCGATTGCCTGATTCCGCAGTCGAGGCATTACGGTGTATTCTTGATGTTAATAGGCTGCCCTCA

FIGURA 7 (continuación)

AAATCTCGGAAGGAAGAATTAGCAATTAAGCTTGGGTTGAGCTTTTCACAGGTTTCATGGATGGTTT
 AAAAAATCAGCGGCATCAAGCTCTGAGGAAAGGGTTGGCATGTCTTAAGCGGATACATCCAGTAGCA
 TCAAGCTCAAGAGATTTCAATACTGTGCAATCTCCATCGTCTGTGGCGACTTTAACTGAAAGCACA
 TTACCGCAAAAAGAGAATGCTCCAAATCTATTTAACTGAGAAGTTGGACGAAGTTCAGGTGCGA
 CTCTCGAATTGAAGCAAGCGTTGGACGATTTTACCAAGAATGCAGCAGTGGTTGCGGATGGCGGA
 AGTATTGAATTGGATAACAACGTGGAAAACGGAGGAATTGGACTTCTGAAAATGGAAAACGTGTT
 GTCTATGTGCCAGTTGTGGCCGACGTTAGGGAACAATCACCAATGAGTACGCATACTAGCGCATGT
 TGA

SEC ID N°: 241, secuencia de polipéptido traducida Phypa_PHD_HD II de *Physcomitrella patens*
 MVPTGVVEREDQPEPAAAPPVPIVITPKTLRSHSIRTQTPGDGDAGADHAPGAVAAAAASAGGGGA
 ADVMOGNDAGQERRRRRRRRKKSVDVPLDNPTRI LKHVKYLLIKMRVHQNLLDAYTGEGWKQSRE
 KIRPEQELQRAKAQILRSKLAIREAHELDDVGLGSLDKNAFDSEGRIYHEEIFCAKCKSQEALP
 DNDIILCDGACNRGFHQYCLDPPLATEDIPPGDEGWLCPVCDCKMECIEAINSYFGTSFEVENSWE
 SFFSNEAGIAAGGGTQEGAGGDWPSGDEDEYDPETAEGPSSASESGAQVIGLAVDWPSDDDEDH
 DYDPEAAEPPGASGSGTQTVGLDWDVWSTDEDDDFDPEGIRKQANSDEESSDESEEDGEDSGS
 VSGFRENEASSGVSDKDEQEPSEKHNELELTAVTNIKKLARKGKRNIISNKGKSEYLESSDSDSDV
 SLDFEFKGGNANLGLRRGDPQAAVSAADEEAVI IAGKRHRKAVDYKRLHDEMYGKAEADNDDVSI
 EDEDWGPERRRRRTIPDDPNSPRSRIRGRKVRTPSGDISKGADADLPSGLPDMPSHFPTGSQGEVG
 SPQDI PALDGAADSPGGEKRMWRRLPDSAVEALRCLDVRNRLPSKSRKEELAIKLGLSFSQVHGF
 KNQRHQALRKLACPKRIHPVASSSRDFNTVESPSVATLTESTLPQKENAPNSILTEKLEDEVQVR
 LLELQALDDFTKNAVVADGGSIELDNNVENGIGLPENGKRVVYVPVADVREQSPMSTHTSAC

SEC ID N°: 242, secuencia de ácido nucleico patentada Zeama_PHD_HD de *Zea mays*
 ATGCATTCTTCGGAATAAACTGGGTTGGAATGAGACCTCTATGGGGTTGGTTCTGTACCAAAA
 AGACCAGCTAGACCTGATGCTTCCCACCAATGCAAACTGATTCCTTTATGAAGCGATCACCAAGG
 AAAGTTAGAAATGCTACTCTGGCAAAAAGCATAAAGAGCAAAATACCCTATAGTCCCCTCAAACAG
 CGGAAGGGTTTCAGATTCTGTTCC TGGGAAAATCGTAAACAGGACTAACCAGCAAGGAAAAAGAA
 AGGAAAATCCAAATTACAGACGAGGCAACTCGTTTGGAAACGAAGAGCGAGATATTTTCTAATCAAG
 ATAAACTGGAGCAGAATTTGCTAGATGCTTACTCTGGAGATGGATGGAATGGGCAAAGCCGAGAG
 AAAATAAAGCCAGAGAAGGAAGTGCACGTCAGGAAACAAATCATAAAATGCAAAATGCTATA
 CGTGATATCATCCGCAACTTGATTTGTATACTTCTACTGGGAGTGTGATGACCCACTAATGCCT
 ACAGATCAGTCCACCAATCCCGAACATACCATGTGCTCAACATGCAAGTCTCATGAATCATTTCCC
 AGCAATAAAATTATCTTCTGCAAAGGGCCCTGCAAAGGGCATGTCATGAGAAATGTTTGGAACT
 CCTTAAACAAAAGCGTACTTCCAACAAGTAGCCATGGGTGGCTTTGCAAATCTGTTTGTGCAAG
 GTGAGGATTTTAGAACTATTAATGCACATCTAGGGACGAGTTTACAGTGAAGTGCCACTTTGAA
 GATATATTCAAGGAACTACTGAACTAATAGACTCTGAGGATGCACTAGATGAAGATTGGCTTTCT
 GAGTATTCAGGTGATGAGGACTATGATCCTGATGAAAATGAGGCCAGTGGCGACTGTATGGACAGC
 GGGGAGAAGATTATGTCTGATGATTCCAATGGTTCAGGAAGCCCCCTTTATTCTCAAATGACGAT
 ATTCTGACTTCATATCAGCAGACTTAAATGTTGTGGAAGGGTTTGTGCATACCAATTTAGATTTA
 GGCATTGATGCCGTTGAAGATGATTTTGCACAGATCCTCACCTACCAAAGGCCAAGAAGAGATGTT
 GATTATAGAAGGCTTAACGAGGAAAATGTTTGGGAAAATAACCGGGAATGAAGAACAGAGCGAGGAC
 GAAGATTGGGGCCATGAAAGGAGAAAAGAAAGGACGCATTTCAGGTGTTGCTGGGGATAATTCTGTT
 GGTTCCTGAACGTTATATCTGATGAAAAGAGCCAGAAGAAGGGGAGAAAACTTTTAGGATTCTCT
 CCTGCAGCTGTTGAGGTAATTCGCAGAGCTTTTGTGAAAATGAGCTTCCACCCCGGGATGTTAAA
 GAAAATCTCTCAAGAGAATTGGGAATTTCTTTGAAAAGATTGATAAATGGTTCAAAAATACACGG

FIGURA 7 (continuación)

TGTGCTGCTCTCAGAGATCGGAAGGCTGAAGGAAACAGTCATAATACAGCTCCCAGCAAAATCTCA
 AAAAAATAAGGAAAAGCTGGAATCTCAGGCAAGGCTGGAAGGAATGGTCATGTTACTGGTCCCTGC
 AATAATTTGAGAATAAATGAAGAAAAACAGGTATATCGGGGAAGTTTGATTCAGGAGACAACCTCT
 TGTTTGGTTCCTTCTCTGAAGTCATCAATGTGCCTACGCGATTGCCACACAACCTTGAGATGAGG
 AAGATGGAATCAACTAGAAAGTCCTGCGAGGCTACACAATAAAGGAGGGTGCCTGTCTGCAACAATC
 CAAATTAAGGAGAGTACATTGTTACCCACAGGCAAGCCTTGTGGCAGTCAGAAATGAGCCATCCA
 ACAACTAATGAAGTGAGCACTTCAGCGCAAGCTACTACTTGGATCGACGCGGGGGCGTGTGCCGAG
 GAGCAAGAACCTACTCCTTGGGTGGACATCGGGGCCTCAGACTACCAGCCTTTCCTGGATGTGATC
 GACGACATGTGCGGACTTGAGTGGAGGCTGCAGAGACTGAAGGAGAACATGCTCTCATCTAGCATA
 GACGGCCAAACCGCCGCCAGTGAGAGTGACAAAAGAAACCAGACCGTGGTGCTCGTGCCAACTGCC
 GAGCTCAAGGAAAAACTGCCGCATGACGGTTTATTTCGGGCATTATTGCCCATAG

SEC ID N°: 243, secuencia de polipéptido traducida patentada Zeama_PHD_HD de Zea mays
 MHSENKLGWNETSMGLVPVPKRPARPDASHQCKSDSFMKRSRPRKVRNATLAKSIKSKYHYSPLKQ
 RKGSDSVPVKIVTGLTARKKKRRIQITDEATRLERRARYFLIKIKLEQNLLDAYSGDGWNGQSRE
 KIKPEKELQARKQI IKCKIAIRDI IRQLDLYTSTGSDVDDPLMPTDQSTNPEHTMCSTCKSHESFP
 SNKIIFCKGPKRACHEKCLEPPLNKSVLPTSSHWLCKFCLCKVRILETINAHLGTSFTVKCHFE
 DIFKETTELIDSEDALDEDWLSEYSGDEDYDPDENEASGDCMDSGEKIMSDDSNNGSGSPLYSPNDD
 IPDFISADLNVVEGFCHTNLDLGDIDAVEDDFAQILTYQRPRRDVDYRRLNEEMFGKITGNEEQSED
 EDWGHERRKTRTHSGVAGDNSVGFLNVI SDEKSQKKGRKLFRI PPAAVEVLRRAFAENELPPRDVK
 ENLSRELGISFEKIDKWFKNTRCAALRDRKAEGNSHNTAPSKI SKNKGKAGISGKAGRNGHVTGPC
 NNLRINEEKTGISGKFDSDGNSCLVPPFSEVINVPTRLPHNFEMRKMESTRSPARLHNKGGCLSATI
 QIKESTLLPTGKPCWQSEMHPPTNEVSTSAQATTWIDAGACAEQEPTPWVDIGASDYQPFLLDVI
 DDMCGLEWRLQRLKENMLSSSIDGQTAASESDKRNQTVVLVPTAELKEKLPDGLFGHYCP

FIGURA 7 (continuación)