

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 451 669**

51 Int. Cl.:

C12N 15/82 (2006.01)

A01H 1/00 (2006.01)

C07K 14/415 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.08.2007 E 07788181 (1)**

97 Fecha y número de publicación de la concesión europea: **08.01.2014 EP 2054516**

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

30 Prioridad:

02.08.2006 EP 06118347 10.08.2006 US 836804 P
20.09.2006 EP 06120994 05.10.2006 EP 06121856
06.10.2006 EP 06121928 12.10.2006 US 851265 P
12.10.2006 US 851258 P 12.10.2006 US 851250 P
27.10.2006 EP 06123066 31.10.2006 EP 06123237
17.11.2006 US 859717 P 01.12.2006 US 868095 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.03.2014

73 Titular/es:

CROPDESIGN N.V. (100.0%)
TECHNOLOGIEPARK 3
9052 ZWIJNAARDE, BE

72 Inventor/es:

FRANKARD, VALERIE;
MIRONOV, VLADIMIR;
REUZEAU, CHRISTOPHE y
SANZ MOLINERO, ANA ISABEL

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 451 669 T3

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

DESCRIPCIÓN

Plantas que tienen características mejoradas y un procedimiento de fabricación de las mismas

5 La presente invención se refiere, en general, al campo de la biología molecular y se refiere a un procedimiento para mejorar diversas características de las plantas modulando en una planta la expresión de un ácido nucleico que codifica una PRC (Proteína Relacionada con el Crecimiento). La presente invención también se refiere a plantas que tienen expresión modulada de un ácido nucleico que codifica un polipéptido de PRC cuyas plantas tienen características mejoradas con respecto a las plantas de tipo silvestre correspondientes u otras plantas de control. La invención también proporciona construcciones útiles en los procedimientos de la invención.

10 La constante cada vez mayor 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 selectivas para identificar plantas que tienen características deseables. Sin embargo, dichas técnicas de siembra selectivas tienen diversos inconvenientes, concretamente que estas técnicas son típicamente laboriosas y producen plantas que a menudo contienen componentes genéticos heterogéneos que no siempre pueden producir el rasgo deseable que se transmite desde la planta progenitora. Los avances en biología molecular han permitido a la humanidad modificar el plasma germinal de animales y plantas. La ingeniería genética de las plantas implica el aislamiento y manipulación del 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.

20 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, número y tamaño de los órganos, arquitectura de la planta (por ejemplo, del número de ramas), rendimiento de semillas, senescencia de las hojas y 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 las semillas 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 los 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 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, se asocian brotes con mayor vigor. Cuando se practica siembra con perforación, para la buena emergencia de las plántulas, son importantes mesocótilos y coleótilos más grandes. La capacidad de modificar genéticamente el vigor temprano en plantas sería de gran importancia en la agricultura. Por ejemplo, el pobre vigor temprano ha sido una limitación para la introducción de híbridos de maíz (*Zea mays* L.) basados en plasma germinal del Cinturón de Maíz en el Atlántico Europeo.

50 Un rasgo importante adicional es el de tolerancia mejorada al estrés abiótico. El estrés abiótico es la causa principal de pérdidas de cultivo en todo el mundo, reduciendo las producciones promedio para la mayoría de las plantas de cultivo principales en más de un 50 % (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 al optimizar uno de los factores mencionados anteriormente.

55 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 semillas, ya sea que esté en la forma de tamaño de semilla aumentado o número de semilla aumentado.

5 Una estrategia para aumentar el rendimiento (rendimiento de semillas y/o biomasa) en 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.

10 Ahora se ha descubierto que en las plantas pueden mejorarse diversas características modulando, en una planta, la expresión de un ácido nucleico que codifica un polipéptido PRC (Proteína Relacionada con el Crecimiento) en una planta. El polipéptido PRC, como se describe en el presente documento, es un polipéptido de Anquirina con dedo de Cinc (AZ). Las características mejoradas comprenden rasgos relacionados con el rendimiento, tales como rendimiento aumentado y/o crecimiento aumentado, o contenido de compuestos de almacenamiento modificados.

Antecedentes

Polipéptido de Anquirina con dedo de Cinc

15 La biomasa vegetal se produce para cultivos de forraje como alfalfa, ensilaje de maíz y heno. En los cultivos de grano se han utilizado muchos indicadores para el rendimiento. El principal entre ellos son las estimaciones del tamaño de la planta. El tamaño de la planta puede medirse de diversas maneras dependiendo de la especie y de la fase de desarrollo, aunque incluye el peso seco total de la planta, el peso seco sobre la superficie, el peso fresco sobre la superficie, el área foliar, el volumen del tallo, la altura de la planta, el diámetro de roseta, la longitud foliar, la longitud radicular, la masa radicular, el número de retoños y número de hojas. Muchas especies mantienen una proporción conservativa entre el tamaño de las diferentes partes de la planta en una fase de desarrollo determinada. Estas relaciones alométricas se usan para extrapolar desde una de estas medidas de tamaño a otras (por ejemplo, Tiftonell y col. 2005 Agric Ecosys & Environ 105: 213). El tamaño de la planta en una fase de desarrollo temprana se correlacionará típicamente con el tamaño de la planta más tarde en el desarrollo. Una planta más grande con mayor área foliar típicamente puede absorber más luz y dióxido de carbono que una planta más pequeña y por tanto probablemente obtendrá mayor peso durante el mismo período (Fasoula & Tollenaar 2005 Maydica 50: 39). Esto es por lo que, además de la posible continuación de la ventaja microambiental o genética, la planta ha de alcanzar inicialmente el tamaño más grande. Existe un fuerte componente genético con respecto al tamaño de la planta y la tasa de crecimiento (por ejemplo, ter Steege et al 2005 Plant Physiology 139: 1078) y por eso para un intervalo de tamaño de planta de genotipos diversos en una condición ambiental se correlaciona posiblemente con el tamaño en otra (Hittalmani y col. 2003 Theoretical Applied Genetics 107: 679). De este modo se usa un entorno estándar como un indicador para los entornos diversos y dinámicos encontrados en diferentes localizaciones y tiempos por los cultivos en el campo.

35 El índice de cosecha, la proporción de rendimiento de semilla con respecto al peso seco sobre la superficie, es relativamente estable en muchas condiciones ambientales y por tanto a menudo puede obtenerse una fuerte correlación entre el tamaño de la planta y el rendimiento de grano (por ejemplo Rebetzke y col. 2002 Crop Science 42: 739). Estos procesos están intrínsecamente relacionados porque la mayoría de la biomasa en grano es dependiente de la productividad fotosintética normal o almacenada por las hojas y los tallos de la planta (Gardener y col. 1985 Physiology of Crop Plants. Iowa State University Press, págs. 68-73) Por lo tanto, la selección del tamaño de la planta, incluso en fases tempranas del desarrollo, se ha usado como un indicador para futuras posibles producciones (por ejemplo Tiftonell y col. 2005 Agric Ecosys & Environ 105: 213). Cuando se analiza el impacto de las diferencias genéticas sobre la tolerancia al estrés, la capacidad de normalizar las propiedades del suelo, temperatura, disponibilidad de agua y nutrientes e intensidad de luz es una ventaja intrínseca de los entornos de invernaderos o cámaras de cultivo de plantas en comparación con el campo. Sin embargo, limitaciones artificiales sobre producciones debidas a una mala polinización por ausencia de viento o de insectos, o espacio insuficiente para la maduración de las raíces o crecimiento del dosel arbóreo, pueden restringir el uso de estos entornos controlados para analizar diferencias de rendimiento. Por lo tanto, las mediciones del tamaño de las plantas en el desarrollo temprano, en condiciones normalizadas en una cámara de cultivo o en un invernadero, son prácticas convencionales para proporcionar indicaciones de las posibles ventajas de rendimiento genético.

50 La transcripción se realiza mediante ARN polimerasas. Estas polimerasas están normalmente asociadas con otras proteínas (factores de transcripción) que determinan la especificidad de los procesos de transcripción. Los factores de transcripción se unen a elementos reguladores en cis del gen y también pueden mediar en la unión de otras proteínas reguladoras. Stegmaier y col. propusieron una clasificación de factores de transcripción basándose en sus dominios de unión a ADN. Se diferenciaron 5 superclases, basándose en la presencia de: 1) dominios básicos, 2) dominios de coordinación de cinc, 3) dominios hélice-giro-hélice, 4) dominios con estructura beta con contactos en surco menor y 0) otros dominios (Stegmaier y col., Genome informatics 15, 276-286, 2004). El grupo de factores de transcripción que comprenden dominios de coordinación de cinc es muy diverso y adicionalmente puede clasificarse de acuerdo con sus restos de cisteína e histidina conservados, incluyendo los dominios WRKY, grupos de cinc C6, dominios DM y GCM.

Además de motivos de unión a ADN, los factores de transcripción también pueden comprender motivos de

interacción proteína-proteína. Un motivo de este tipo es el motivo de anquirina. Este está presente en muchas familias de proteínas diversas, normalmente como una repetición de 2 a más de 20 unidades. Cada unidad contiene dos hélices antiparalelas y una horquilla beta.

5 Aunque muchas proteínas vegetales con dominios de dedo de cinc están bien caracterizadas, se sabe poco sobre proteínas vegetales que comprendan el motivo de dedo de cinc C3H1. PEI1, un factor de transcripción que según se informa desempeña una función en el desarrollo embrionario, tiene un motivo de dedo de cinc que se asemeja al motivo C3H1 pero carece de un motivo de anquirina (Li y Thomas, Plant Cell 10, 383-398, 1998). El documento WO 02/44389 describe AtSIZ, un factor de transcripción aislado de *Arabidopsis*. Se descubrió que la expresión de AtSIZ bajo el control del promotor CaMV35S promovía la transcripción de genes inducidos por estrés y, según se informa, plantas con expresión de AtSIZ aumentada tenían una tasa de supervivencia más alta bajo estrés salino en comparación con las plantas de control, pero no se proporcionó ningún análisis con respecto al rendimiento de semillas. Se postuló que AtSIZ podía usarse en plantas para aumentar la resistencia al estrés osmótico.

Sumario

15 Sorprendentemente, ahora se ha descubierto que la modulación de la expresión de un ácido nucleico que codifica un polipéptido PRC proporciona plantas que tienen características mejoradas con respecto a plantas de control.

La presente memoria descriptiva describe que la modulación de la expresión en una planta de un ácido nucleico que codifica el polipéptido AZ de *Arabidopsis* (AtAZ) o un homólogo del mismo proporciona plantas que tienen rendimiento aumentado con respecto a plantas de control.

20 La memoria descriptiva describe un procedimiento para aumentar el rendimiento de las plantas, que comprende modular en una planta la expresión de un ácido nucleico que codifica el polipéptido AZ o un homólogo del mismo. Ventajosamente, la realización de los procedimientos de acuerdo con la presente invención produce plantas con rendimiento aumentado, particularmente rendimiento de semillas, con respecto a plantas de tipo silvestre correspondientes. La presente memoria descriptiva también describe secuencias de ácido nucleico y construcciones útiles en la realización de dichos procedimientos.

25 La presente invención proporciona la materia objeto como se expone en uno cualquiera y en todos los puntos (1) (22) indicados a continuación:

1. Un procedimiento para aumentar el rendimiento de semillas de plantas con respecto a plantas de tipo silvestre correspondientes, que comprende introducir y expresar en una planta un ácido nucleico AZ o una variante del mismo, en el que dicho ácido nucleico AZ codifica un polipéptido AZ o un homólogo del mismo, en el que el homólogo proporciona plantas que tienen rendimiento aumentado, y en el que dicho polipéptido AZ o el homólogo del mismo comprende dos repeticiones de anquirina y dos dominios C3H1 de dedo de cinc, y en el que dichas repeticiones de anquirina se localizan aguas arriba de los dominios C3H1 de dedo de cinc.

30 2. Un procedimiento de acuerdo con el punto 1, en el que polipéptido AZ comprende al menos uno de los siguientes motivos:

35 (P/A)CSRAY(S/T)HDWTEC (motivo 1, SEC ID N°: 3)
 HPGENARRRDPR (motivo 2, SEC ID N°: 4)
 HG(V/I)FE(C/S)WLHP(A/S)QY(R/K)TRLCK (motivo 3, SEC ID N°: 5)
 CFFAH (motivo 4, SEC ID N°: 6).

40 3. El procedimiento de acuerdo con el punto 1 o 2, en el dicho ácido nucleico AZ codifica un polipéptido AZ de SEC ID N°: 2: o un homólogo del mismo.

4. El procedimiento de acuerdo con uno cualquiera de los puntos 1 a 3, en el que dicha variante es una parte de un ácido nucleico AZ o una secuencia capaz de hibridarse con un ácido nucleico AZ, cuya parte o secuencia de hibridación codifica un polipéptido AZ que comprende dos repeticiones de anquirina y dos dominios de dedo de cinc C3H1, y en el que dichas repeticiones de anquirina se localizan aguas arriba de los dominios de dedo de cinc C3H1.

45 5. El procedimiento de acuerdo con uno cualquiera de los puntos 1 a 4, en el que dicho ácido nucleico AZ o variante del mismo se sobreexpresa en una planta.

6. El procedimiento de acuerdo con uno cualquiera de los puntos 1 a 5, en el que dicho ácido nucleico AZ o variante del mismo es de origen vegetal, preferentemente de una planta dicotiledónea, más preferentemente de la familia *Brassicaceae*, más preferentemente el ácido nucleico es de *Arabidopsis thaliana*.

50 7. El procedimiento de acuerdo con uno cualquiera de los puntos 1 a 6, en el que dicho ácido nucleico AZ o variante del mismo está unido operativamente a un promotor específico de semilla.

8. El procedimiento de acuerdo con el punto 7, en el que dicho promotor específico de semilla es un promotor WSI18.

55 9. El procedimiento de acuerdo con uno cualquiera de los puntos 1 a 8, en el que dicho rendimiento de semilla aumentada comprende peso de mil granos aumentado.

10. Una construcción que comprende:

(i) un ácido nucleico AZ o una variante del mismo como se define en cualquiera de los puntos 1 a 4 y 6;

(ii) un promotor WSI18 específico de semilla o un promotor GOS2 constitutivo unido operativamente a la secuencia de ácido nucleico de (i).

11. Una construcción de acuerdo con el punto 10, en el que dicho promotor WSI18 es como se representa en la SEC ID N° 55 o SEC ID N° 9.

12. Una construcción de acuerdo con el punto 10, en el que dicho promotor GOS2 es como se representa en la SEC ID N° 56 o SEC ID N° 54.

13. Una planta transformada con una construcción de acuerdo con uno cualquiera de los puntos 10 a 12.

14. Un procedimiento para la producción de una planta transgénica que tiene rendimiento de semilla aumentado en comparación con plantas de tipo silvestre correspondientes, cuyo procedimiento comprende:

(i) introducir y expresar en una planta o en una célula de planta un ácido nucleico AZ o variante del mismo como se define en cualquiera de los puntos 1 a 4 y 6; y

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

15. Una planta transgénica de acuerdo con el punto 13, en la que dicha planta es una planta monocotiledónea tal como caña de azúcar o en la que la planta es un cereal, tal como arroz, maíz, trigo, cebada, mijo, centeno, avena o sorgo.

16. Partes cosechables de una planta de acuerdo con cualquiera de los puntos 13 o 15, en las que dichas partes cosechables comprenden una construcción de acuerdo con uno cualquiera de los puntos 10 a 12.

17. Partes cosechables de una planta de acuerdo con el punto 16 en las que dichas partes cosechables son semillas.

18. Productos directamente derivados de una planta de acuerdo con el punto 13 o 15 y/o de partes cosechables de una planta de acuerdo con los puntos 16 o 17, en los que dichos productos comprenden una construcción de acuerdo con uno cualquiera de los puntos 10 a 12.

19. El uso de un ácido nucleico AZ o variante del mismo como se define en cualquiera de los puntos 1 a 4 y 6, o el uso de un polipéptido AZ o un homólogo del mismo, en el rendimiento de semillas aumentado, con respecto a plantas de tipo silvestre correspondientes.

20. El uso de acuerdo con el punto 19, en el que dicho rendimiento de semillas aumentada comprende el peso de mil granos aumentado.

21. El uso de un ácido nucleico AZ o variante del mismo como se define en cualquiera de los puntos 1 a 4 y 6, o el uso de un polipéptido AZ o un homólogo del mismo, como un marcador molecular.

22. El uso de una construcción de acuerdo con cualquiera de los puntos 10 a 12 en un procedimiento para fabricar plantas que tengan rendimiento de semillas aumentado con respecto a plantas de control.

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

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

Planta (plantas) 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 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 nulizigotos 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, que incluye semillas y partes de semilla.

Homólogo/ (homólogos)

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

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. De manera general, las inserciones dentro de la secuencia de aminoácidos serán más pequeñas que las fusiones de N o C terminal, del orden de aproximadamente 1 a 10 restos. Ejemplos de las 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 utiliza 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, epítoto Tag•100, epítoto c-myc, epítoto FLAG[®], lacZ, CMP (péptido de unión a calmodulina), epítoto HA, epítoto de proteína C y epítoto VSV.

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

Tabla 1: Ejemplos de sustituciones de aminoácido conservativas

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

Las sustituciones, deleciones y/o inserciones de aminoácido pueden realizarse fácilmente utilizando técnicas sintéticas de péptidos bien conocidas en la técnica, tal como síntesis peptídica de fase sólida y similares, o por manipulación de ADN recombinante. Se conocen bien en la técnica los procedimientos para la manipulación de secuencias de ADN para producir variantes de sustitución, inserción o deleción de una proteína. Por ejemplo, técnicas para realizar mutaciones de sustitución en sitios predeterminados en ADN son bien conocidas por aquellos expertos en la técnica e incluyen mutagénesis M13, mutagénesis *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.

25 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 no origen no natural en comparación con la secuencia de aminoácidos de forma no natural del polipéptido. Un derivado también puede comprender uno o más sustituyentes o adiciones no aminoácido en comparación con la secuencia de aminoácidos de la que deriva, por ejemplo, una molécula indicadora u otro ligando, unido covalente o no covalentemente a la secuencia de aminoácidos, tal como una molécula indicadora que se une para facilitar su detección, y restos de aminoácido que no son de origen natural con respecto a la secuencia de aminoácidos de una proteína de origen natural. Además, los "derivados" también incluyen fusiones de la forma de origen natural de la proteína con péptidos etiquetadores tal como FLAG, HIS6 o tiorredoxina (para una revisión de péptidos etiquetadores, véase Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

Ortólogo(s)/Parólogo(s)

Los ortólogos y parálogos incluyen conceptos evolutivos utilizados 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 especiación y también derivan de un gen antecesor común.

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 son altamente 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 alienadas de una familia de homólogos de proteína, éstas pueden utilizarse como identificadores para determinar si cualquier polipéptido en cuestión pertenece a una familia polipeptídica previamente identificada.

Motivo/Secuencia consenso/Firma

La expresión "motivo" o "secuencia consenso" o "firma" se refiere 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 se ubican fuera del dominio conservado (si todos los aminoácidos del motivo caen fuera de un dominio definido).

Hibridación

El término "hibridación" como se define en el presente documento es un proceso en el que la secuencia de nucleótidos complementaria sustancialmente homóloga hibrida entre sí. El proceso de hibridación puede producirse completamente en la solución, es decir, ambos ácidos nucleicos complementarios están en la 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 cuando uno de los ácidos nucleicos complementarios inmovilizados en un soporte sólido, tal como una membrana de nitrocelulosa o de nailon o se inmoviliza mediante, por ejemplo, fotolitografía a, por ejemplo, un soporte vítreo silíceo (el último conocido como matrices o micromatrices de ácido nucleico o como microplacas de ácido nucleico). Con el fin de permitir que ocurra la hibridación, las moléculas de ácido nucleico de manera general se desnaturalizan térmica o químicamente para fundir una doble cadena en dos cadenas únicas y/o para retirar horquillas u otras estructuras secundarias de ácidos nucleicos de cadena sencilla.

El término "rigurosidad" se refiere a las condiciones bajo las cuales se produce la hibridación. La rigurosidad de la hibridación está influenciada por condiciones tales como temperatura, concentración de sal, 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_m) 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_m y las condiciones de rigurosidad alta son cuando la temperatura es 10 °C por debajo de la T_m . Las condiciones de hibridación de rigurosidad alta se usan típicamente para aislar las secuencias de hibridación que tienen alta similitud de secuencia con la secuencia de ácido nucleico diana. Sin embargo, los ácidos nucleicos pueden desviarse en cuanto a la secuencia y aún codificar un polipéptido sustancialmente idéntico, debido a la degeneración del código genético. Por lo tanto, algunas veces pueden necesitarse condiciones de hibridación de rigurosidad media para identificar dichas moléculas de ácidos nucleicos.

La T_m 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_m 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. El índice máximo de hibridación se obtiene de aproximadamente 16 °C hasta 32 °C por debajo de la T_m . La presencia de cationes monovalentes en la solución de hibridación reduce la repulsión electrostática entre las dos cadenas de ácido nucleico que promueven por lo tanto la formación del híbrido; este efecto es visible para concentraciones de sodio de hasta 0,4 M (para concentraciones más elevadas, este efecto puede ignorarse). La formamida reduce la temperatura de fusión de dúplex ADN-ADN y ADN-ARN con 0,6 a 0,7 °C para cada porcentaje de formamida y la adición de formamida al 50 % permite que se realice la hibridación de 30 a 45 °C, aunque se disminuirá el índice de hibridación. Los emparejamientos erróneos de los pares de bases reducen el índice de hibridación y la estabilidad térmica de los dúplex. En promedio y para sondas grandes, la T_m disminuye aproximadamente 1 °C por % de emparejamiento erróneo de bases. La T_m puede calcularse utilizando las siguientes ecuaciones, dependiendo de los tipos de híbridos:

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

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

2) Híbridos de ADN-ARN o ARN-ARN:

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

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

Para <20 nucleótidos: $T_m = 2 (I_n)$

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

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

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

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

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

10 La unión no específica puede controlarse usando uno cualquiera de un número de técnicas conocidas tal como, por ejemplo, bloqueo de la membrana con soluciones que contienen la proteína, adiciones del ARN heterólogo, ADN y SDS heterólogo al tampón de hibridación y tratamiento con Rnasa. 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 de 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.

20 Además de las condiciones de hibridación, la especificidad de hibridación también depende típicamente de la función de 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: la menor concentración de sal y la mayor temperatura de lavado, la mayor rigurosidad de lavado. Las condiciones de lavado se realizan típicamente a o por debajo de la rigurosidad de hibridación. Una hibridación positiva proporciona una señal que es al menos dos veces la del fondo. Generalmente, las condiciones de rigurosidad adecuadas para los ensayos de hibridación de ácido nucleico o los procedimientos de detección de amplificación de genes son como se establecieron anteriormente. También pueden seleccionarse condiciones más o menos rigurosas. El experto en la técnica es consciente de diversos parámetros que pueden alterarse durante el lavado y que mantendrán o cambiarán las condiciones de rigurosidad.

30 Por ejemplo, las condiciones de hibridación de alta rigurosidad típicas para híbridos de ADN mayores de 50 nucleótidos incluyen 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 0,3 x SSC. Ejemplos de condiciones de hibridación de rigurosidad media para híbridos de ADN mayores de 50 nucleótidos incluyen 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 la secuencia conocida se hibridan, la longitud del híbrido puede determinarse alineando las secuencias e identificando las regiones conservadas descritas en el presente documento. 1xSSC es NaCl 0,15M y citrato sódico 15 mM; la solución de hibridación y las soluciones de lavado pueden incluir adicionalmente reactivo de Denhard 5x, SDS 0,5-1,0 %, 100 µg/ml de ADN de espera de salmón fragmentado, desnaturalizado, pirofosfato sódico al 0,5 %.

35 Con objeto 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).

40 Variante de corte y empalme

45 El término "variante de corte y empalme" como se usa en el presente documento incluye variantes de una secuencia de ácido nucleico en la que los intrones y/o exones seleccionados se han escindido, reemplazado, desplazado o añadido, o en la que los intrones se han acortado o alargado. Dichas variantes serán unas en las que la actividad biológica de la proteína se conserva sustancialmente; esto puede conseguirse conservando selectivamente los segmentos funcionales de la proteína. Dichas variantes de corte y empalme pueden encontrarse en la naturaleza o pueden ser fabricadas por el hombre. Los procedimientos para predecir y aislar dichas variantes de corte y empalme son bien conocidos en la técnica (véase por ejemplo, Foissac y Schiex (2005) BMC Bioinformatics 6: 25).

Variante alélica

50 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), así como Polimorfismos de Inserción/Delección pequeños (INDEL). El tamaño de INDEL es normalmente menor de 100 pb. Los SNP e INDEL forman el grupo 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 consta de iteraciones de combinación de ADN seguido de detección 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 N° 5.811.238 y 6.395.547).

Elemento regulador/Secuencia de control/Promotor

Las expresiones “elemento regulador”, “secuencia de control” y “promotor” se utilizan 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” típicamente se refiere a una secuencia de control de ácidos nucleicos localizada en dirección 5' del inicio transcripcional de un gen y que está implicado en el reconocimiento y unión de la ARN polimerasa y otras proteínas, dirigiendo de este modo la transcripción de un ácido nucleico unido operativamente. Incluidas en los términos anteriormente mencionados se encuentran las secuencias reguladoras transcripcionales derivadas de un gen genómico eucariota clásico (que incluye la caja TATA que es necesaria para el inicio de transcripción adecuado, con o sin secuencia de caja CCAAT) y elementos reguladores adicionales (es decir, secuencias activadoras en dirección 5', potenciadores y silenciadores) que alteran la expresión génica en respuesta al desarrollo y/o estímulo externo, o en una forma específica de tejido. También se incluye dentro del término una secuencia reguladora transcripcional de un gen procarionota clásico, en cuyo caso puede incluir una secuencia de caja -35 y/o secuencias reguladoras transcripcionales de caja -10. La expresión “elemento regulador” también incluye una molécula de fusión sintética o un derivado que confiere, activa o potencia la expresión de una molécula de ácido nucleico en una célula, tejido u órgano.

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 necesita ser de origen vegetal, pero puede originarse a partir 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 proceso de la invención y que se describe en el presente documento. Esto también se aplica a las otras señales reguladoras de “planta”, tales como terminadores de “planta”. Los promotores en la dirección 5' de la secuencia 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 e 3' tal como terminadores u otras regiones reguladoras 3' que se ubican lejos de la ORF. Además es posible que la actividad de los promotores aumente por modificación de su secuencia o que se reemplacen completamente por 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 patrón del gen indicador en diversos tejidos de la planta. Los genes indicadores bien conocidos adecuados incluyen, por ejemplo, beta-glucuronidasa o beta-galactosidasa. La actividad del promotor se evalúa midiendo la actividad enzimática de la beta-glucuronidasa o beta-galactosidasa. El patrón de expresión y/o la fuerza del promotor pueden después compararse con aquel de un promotor de referencia (tal como el utilizado en los procedimientos de la presente invención). Como alternativa, la fuerza del promotor puede evaluarse cuantificando los niveles de ARNm o comparando niveles de ARNm del ácido nucleico utilizado en el procedimientos de la presente invención, con niveles de ARNm de genes constitutivos tal como ARNr 18S, utilizando los procedimientos conocidos en la técnica, tal 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 1/500.000 transcritos por célula. Por el contrario, un “promotor fuerte” dirige la expresión de una secuencia codificante de alto nivel, o aproximadamente 1/10 transcritos a aproximadamente 1/100 transcritos a aproximadamente 1/1000 transcritos por célula.

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 es capaz de 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, pero no necesariamente todas, las fases de crecimiento y desarrollo y en condiciones más ambientales, en al menos una célula, tejido u órgano. La siguiente Tabla 2a proporciona ejemplos de promotores constitutivos.

5

Tabla 2a: Ejemplos de promotores constitutivos

Fuente del gen	Referencia
Actina	McElroy y col, Plant Cell, 2: 163-171, 1990
HMGP	Documento WO 2004/070039
CAMV 35S	Odell y col, Nature, 313: 810-812, 1985
CaMV 19S	Nilsson y col., Physiol. Plant. 100: 456-462, 1997
GOS2	de Pater y col, Plant J Nov; 2(6): 837-44, 1992, documento WO 2004/065596
Ubiquitina	Christensen y col, Plant Mol. Biol. 18: 675-689, 1992
Ciclofilina de arroz	Buchholz y col, Plant Mol Biol. 25(5): 837-43, 1994
Histona H3 de maíz	Lepetit y col, Mol. Gen. Genet. 231:276-285, 1992
Histona H3 de alfalfa	Wu y col. Plant Mol. Biol. 11:641-649, 1988
Actina 2	An y col, Plant J. 10(1); 107-121, 1996
34S FMV	Sanger y col., Plant. Mol. Biol., 14, 1990: 433-443
Subunidad pequeña Rubisco	Documento 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	Documento WO 01/14572
Súper promotor	Documento WO 95/14098
Proteínas de caja G	Documento WO 94/12015

Promotor ubicuo

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

Promotor regulado evolutivamente

10 Un promotor regulador evolutivamente es activo durante ciertas etapas de desarrollo o en partes de la planta que experimentan cambios evolutivos.

Promotor inducible

15 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 de diversos patógenos.

Promotor específico de órgano/Específico de tejido

20 Un promotor específico de órgano o específico de tejido es uno que puede iniciar preferentemente la transcripción en determinados órganos o tejidos, tales como hojas, raíces, tejido de semilla, etc. Por ejemplo, un “promotor específico de raíz” es un promotor que es transcripcionalmente activo predominantemente en las raíces de las plantas, sustancialmente con la exclusión de cualesquiera otras partes de una planta, mientras que aún permite

cualquier expresión parcial en estas otras partes de la planta. En el presente documento a los promotores que pueden iniciar la transcripción solo en determinadas células se les denomina “específicos de célula”.

5 Un promotor específico de semilla es transcripcionalmente activo predominantemente en el tejido de semilla, pero no necesariamente exclusivamente en tejidos de semilla (en casos de expresión parcial). El promotor específico de semilla puede ser activo durante el desarrollo y/o durante la germinación de las semillas. El promotor específico de semilla puede ser específico de endospermo/aleurona/embrión. Ejemplos de promotores específicos de semillas se muestran en las Tablas 2b a 2e a continuación. Otros ejemplos de promotores específicos de semilla se proporcionan en Qing Qu y Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuya descripción se incorpora por referencia en el presente documento en su totalidad.

10

Tabla 2b: 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 LMW y HMW 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 , γ -gliadinas de trigo	EMBO J. 3: 1409-15, 1984
Promotor It1 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 250:7 50-60, 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
a-globulina Glb-1 de arroz	Wu y col, Plant Cell Physiology 39(8) 885-889, 1998
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
Pirofosforilasa ADP glucosa de arroz	Trans Res 6: 157-68, 1997
Familia del gen ESR de maíz	Plant J 12:235-46, 1997
α -kafirina de sorgo	DeRose 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

(continuación)

Fuente del gen	Referencia
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
PRO0095	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 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-Peru de maíz	Selinger y col., Genetics 149:1125-38,1998

Tabla 2c: 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 LMW y HMW 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 ltr1 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
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
Pirofosforilasa ADP glucosa de arroz	Russell y col. (1997) Trans Res 6: 157-68

(continuación)

Fuente del gen	Referencia
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 2d: 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

5

Tabla 2e: 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

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 cualesquiera otras partes de una planta, mientras que aún permanece cualquier expresión parcial en estas otras partes de la planta.

- 10 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 cualesquiera otras partes de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de las plantas.

Terminador

- 15 El término "terminador" incluye una secuencia de control que es una secuencia de ADN en el extremo de una unidad que señala el procesamiento y la poliadenilación en 3' de un transcrito primario y la terminación de la transcripción. El terminador puede derivar del gen natural, de una diversidad de otros genes de planta, o de ADN-T. El terminador a añadir puede derivar de, por ejemplo, los genes de nopalina sintasa u octopina sintasa, o como alternativa, de otro gen de planta, o menos preferentemente de cualquier otro gen eucariota.

20 Modulación

- El término "modulación" significa en relación a la expresión o expresión de gen, un proceso en el que 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 no modulada, original, 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 ácidos nucleicos de la invención o las proteínas codificadas, que conduce a la rendimiento aumentado y/o a crecimiento aumentado de las plantas.
- 25

Expresión

El término “expresión” o “expresión de gen” significa la transcripción de un gen específico o genes específicos o construcción genética específica. El término “expresión” o “expresión de gen” 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 proceso incluye la transcripción de ADN y el procesamiento del producto de ARNm resultante.

Expresión aumentada/sobreexpresión

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

Los procedimientos para aumentar la expresión de los genes o productos génicos se documentan bien en la técnica e incluyen, por ejemplo, la sobreexpresión conducida mediante promotores adecuados, el uso de potenciadores de la transcripción o potenciadores de la traducción. Los ácidos nucleicos aislados que sirven como promotor o elementos potenciadores pueden introducirse en una posición adecuada (típicamente en la dirección 5') de una forma no heteróloga de un polinucleótido para regular por aumento la expresión de un ácido nucleico que codifica el polipéptido de interés. Por ejemplo, los promotores endógenos pueden alterarse *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 como se describe en el presente documento de tal manera que se controle la expresión del gen.

Si se desea la expresión del polipéptido, es generalmente deseable incluir una región de poliadenilación en el extremo 3' de una región codificante de polinucleótido. La región de poliadenilación puede derivar del gen natural, de una diversidad de otros genes de planta, o de ADN-T. La secuencia de extremo 3' a añadir puede derivar de, por ejemplo, los genes 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 información general véase: The Maize Handbook, capítulo 116, Freeling y Walbot, Eds., Springer, N.Y. (1994).

Gen endógeno

La referencia en el presente documento 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 en hombre, por ejemplo, mediante síntesis química.

Expresión reducida

En el presente documento la referencia a “expresión reducida” o “reducción o eliminación sustancial” de expresión significa una reducción en la expresión del gen endógeno y/o los niveles de polipeptido y/o actividad de polipéptido con relación a las plantas de control. La reducción o eliminación sustancial está en orden aumentado de preferencia por lo menos de 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, este 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 derivar del ácido nucleico que codifica la proteína de interés (gen diana), o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u 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 aumentado 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

requerimiento para los diversos procedimientos analizados en el presente documento para la reducción o eliminación sustancial de la expresión de un gen endógeno.

Esta reducción o eliminación sustancial de la expresión puede realizarse usando herramientas y técnicas habituales. Un procedimiento para la reducción o eliminación sustancial de la expresión del gen endógeno es al introducir y expresar en una planta una construcción genética en la que el ácido nucleico (en este caso un tramo de nucleótidos sustancialmente contiguos derivados del gen de interés, o de cualquier nucleótido capaz de codificar un ortólogo, parálogo u 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, la expresión del gen endógeno se reduce o se elimina sustancialmente a través del silenciamiento mediado por ARN utilizando una repetición invertida de un ácido nucleico o una parte del mismo (en este caso un tramo de nucleótidos sustancialmente contiguos derivados del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), preferentemente capaz de formar una estructura en horquilla. La repetición invertida se clona en un vector de expresión que comprende las 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), 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 (ARNhp). El ARNhp se procesa en la planta en los ARNip que se incorporan en un complejo de silenciamiento inducido por ARN (RISC, RNA). El RISC escinde adicionalmente los transcritos de ARNm, reduciendo por lo tanto sustancialmente el número de transcritos de ARNm que van a traducirse en los polipéptidos. Para detalles generales adicionales véase, por ejemplo, Grierson y col. (1998) WO 98/53083; Waterhouse y col. (1999) WO 99/53050).

La realización de los procedimientos como se describe en el presente documento no se basa en introducir y expresar en una planta una construcción genética en la que el ácido nucleico se clona como una repetición invertida, sino que pueden utilizarse cualquiera de uno o más de los diversos procedimientos de "silenciamiento génico" bien conocidos para lograr 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 del gen (regulación por disminución). El silenciamiento en este caso se activa en una planta mediante una secuencia de ARN bicatenario (ARNbc) que es sustancialmente similar al gen endógeno diana. Este ARNbc se 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 derivados del gen de interés, o del 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 reduciría la expresión del gen endógeno, dando lugar a un fenómeno conocido como co-supresión. La reducción de la expresión del gen sería más pronunciada si en una planta se introducen diversas copias adicionales de una secuencia de ácidos nucleicos, cuando hay una correlación positiva entre altos niveles de transcrito y la activación de la co-supresión.

Otro ejemplo de un procedimiento de silenciamiento de ARN implica el uso de las 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 codificante de una molécula de ADNc bicatenario o complementario a una secuencia de transcrito de ARNm. La secuencia de ácidos nucleicos antisentido es preferentemente complementaria al 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 aminoácido. 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 (denominado también regiones no traducidas 5' y 3').

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 derivados del gen de interés, o de cualquier ácido nucleico capaz de codificar un ortólogo, parálogo u homólogo de la proteína de interés), pero también puede ser un oligonucleótido que es antisentido para solo a 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 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 iniciar de aproximadamente 50, 45, 40, 35, 30, 25, 20, 15 o 10 nucleótidos de longitud o menos. Una secuencia de ácidos nucleicos antisentido como se describe en el presente documento puede construirse usando síntesis química y las reacciones de ligamiento enzimático utilizando los procedimientos conocidos en la técnica. Por ejemplo, una

5 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 el aumento de la estabilidad física del dúplex formado entre las secuencias de ácidos nucleicos antisentido y sentido, por ejemplo, pueden utilizarse derivados fosforotioato y nucleótidos sustituidos por acridina. En la técnica se conocen bien ejemplos de nucleótidos modificados que pueden

10 utilizarse para generar las secuencias de ácidos nucleicos antisentido. Las modificaciones de nucleótido conocidos 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.

La secuencia de ácidos nucleicos antisentido puede producirse biológicamente utilizando un vector de expresión en el que se ha subclonado una secuencia de ácidos nucleicos en una orientación antisentido (es decir, el ARN 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.

15

Las moléculas de ácido nucleico utilizadas para el silenciamiento en los procedimientos descritos en el presente documento (tanto si se introduce en una planta como se genera *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 cualquier complementariedad convencional de nucleótidos 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 doble hélice. Las secuencias de ácidos nucleicos antisentido pueden introducirse en una planta mediante 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 administración sistémica, las secuencias de ácidos nucleicos antisentido pueden modificarse de tal manera que se unan específicamente a los receptores o antígenos expresados sobre una superficie celular seleccionada, por ejemplo, al ligar la secuencia de ácidos nucleicos antisentido a los péptidos o anticuerpos que se unen a los receptores o antígenos de superficie celular. Las secuencias de ácidos nucleicos antisentido también pueden administrarse a células usando los vectores descritos en el presente documento.

20

25

30

Como se describe en el presente documento, la secuencia de ácidos nucleicos antisentido es una secuencia de ácidos nucleicos anomérica a. Una secuencia de ácidos nucleicos anomérica a forma híbridos bicatenarios específicos con ARN de complementariedad en el que, contrario a las unidades b habituales, las cadenas corren paralelas entre sí (Gaultier y col. (1987) Nucl Ac Res 15: 6625-6641). La secuencia de ácidos nucleicos antisentido también puede comprender un 2'-o-metilribonucleótido (Inoue y col. (1987) Nucl Ac Res 15, 6131-6148) o un análogo de ARN-ADN quimérico (Inoue y col. (1987) FEBS Lett. 215, 327-330).

35

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 un ARNm, en el que tienen una región complementaria. Por tanto, las ribozimas (por ejemplo, ribozimas cabeza de martillo (descritas en Haselhoff y Gerlach (1988) Nature 334, 585-591) pueden utilizarse para escindir catalíticamente los transcritos de ARNm que codifican un polipéptido, reduciendo sustancialmente por tanto el número de transcritos de ARNm que van a traducirse en un polipéptido. Se puede diseñar una ribozima que tenga especificidad para una secuencia de ácidos nucleicos (véase, por ejemplo: Cech y col. Patente de Estados Unidos N° 4.987.071; y Cech y col. Patente de Estados Unidos N° 5.116.742). Como alternativa, los transcritos de ARNm correspondientes a una secuencia de ácidos nucleicos pueden utilizarse para seleccionar un ARN catalítico que tenga una actividad específica de ribonucleasa 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) WO 94/00012; Lenne y col. (1995) WO 95/03404; Lutziger y col. (2000) WO 00/00619; Prinsen y col. (1997) WO 97/13865 y Scott y col. (1997) WO 97/38116).

40

45

50

El silenciamiento génico también puede producirse por mutagénesis de inserción (por ejemplo, inserción de ADN-T o inserción de transposón) o mediante estrategias como describen, entre otros, Angell y Baulcombe ((1999) Plant J 20(3): 357-62), (Amplicon VIGS 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 o eliminación sustancial puede estar provocada por un polipéptido no funcional. Por ejemplo, el polipéptido puede unirse a diversas proteínas que interaccionan; por lo tanto pueden proporcionarse una o más mutaciones y/o truncamientos para un polipéptido que es aún capaz de unir las proteínas que interaccionan (tal como proteínas receptoras) pero que no pueden exhibir su función normal (tal como ligando de señalización).

55

60

Una estrategia adicional para el 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.

- 5 Un experto conocerá otros procedimientos, tales como el uso de anticuerpos dirigidos a un polipéptido endógeno para inhibir su función en la planta, o la interferencia en la ruta de señalización en la que un polipéptido está implicado. En particular, se puede prever que las moléculas fabricadas por el hombre pueden 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.
- 10 Como alternativa, puede establecerse un programa de detección hasta 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 utilizarse, por ejemplo, para realizar recombinación homóloga.

Para desactivar la expresión génica y/o traducción de ARNm pueden utilizarse microARN (miARN) artificiales y/o naturales. Los miARN endógenos son ARN pequeños monocatenarios típicamente con una longitud de 19-24 nucleótidos. Principalmente funcionan para regular la expresión génica y/o la traducción de ARNm. La mayor parte de los microARN (miARN) de planta tienen complementariedad perfecta o casi perfecta con sus secuencias diana. Sin embargo, hay dianas naturales con hasta cinco emparejamiento 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) al unirse a su componente principal, una proteína Argonauta. Los miARN sirven como componentes de especificidad del RISC, debido al emparejamiento de bases con 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 se reflejan frecuentemente en niveles de ARNm reducidos de los genes diana.

- 25 Los microARN artificiales (amiARN), que típicamente tienen 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 utilizarse para ayudar en el diseño de los amiARN específicos (Schwab y col., *Dev. Cell* 8, 517-527, 2005). Están disponibles para el público las herramientas convenientes para el diseño y la generación de los amiARN y sus precursores (Schwab y col., *Plant Cell* 18, 1121-1133, 2006).

Para un rendimiento óptimo, las técnicas de silenciamiento génico utilizadas 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 dada 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 exista una homología sustancial entre el gen diana endógeno y el ácido nucleico a introducir.

- 40 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 será capaz de 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 a través del uso de un promotor apropiado, por ejemplo.

45 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 el que se expresa para facilitar la identificación y/o selección de las células que se transfieren o se transforman con una construcción de ácido nucleico como se describe en el presente documento. Estos genes marcadores permiten la identificación de una transferencia exitosa de las moléculas de ácido nucleico mediante una serie de diferentes principios. Los marcadores adecuados pueden seleccionarse de marcadores que confieren resistencia herbicida o antibiótica, que introducen un nuevo rasgo metabólico o que permiten selección visual. Como ejemplos de genes marcadores de selección se incluyen genes que confieren resistencia a antibióticos (tales como nptII que fosforila neomicina y canamicina, o hpt, que higromicina fosforilante, o genes que confieren resistencia a, por ejemplo, bleomicina, estreptomocina, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomicina o blastidina), a los herbicidas (por ejemplo bar que proporciona resistencia a Basta[®], aroA o gox que proporciona resistencia contra el glifosato, o los genes que confieren resistencia a, por ejemplo, imidazolinona, fosfinotricina o sulfonilurea) o genes que proporcionan un rasgo metabólico (tal como manA que permite a las plantas utilizar manosa como única fuente de carbono o 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/lucefarasa) 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. Dependiendo del organismo y del procedimiento de selección se prefieren diferentes marcadores.

Se sabe que, tras la integración estable o transitoria de los ácidos nucleicos en las 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 utilizado y de la técnica de transfección utilizada. 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, por ejemplo, utilizarse 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 como se describe en el presente documento o utilizarse en los procedimientos como se describe en el presente documento, o incluso en un vector distinto. Las células que se han transfectado de manera estable con el ácido nucleico introducido pueden identificarse por ejemplo mediante selección (por ejemplo, células que tienen integrado el marcador de selección sobreviven mientras que las otras células mueren).

Dado que los genes marcadores, particularmente genes que confieren resistencia a antibióticos y herbicidas, ya no requieren o no se desean en la célula huésped transgénica, una vez que los ácidos nucleicos se han introducido satisfactoriamente, el proceso como se describe en el presente documento para introducir los ácidos nucleicos emplea ventajosamente técnicas que permiten la eliminación o escisión de estos genes marcadores. Un procedimiento de este tipo es el conocido como co-transformación. El procedimiento de co-transformación emplea dos vectores simultáneamente para la transformación, un vector que lleva el ácido nucleico como se describe en el presente documento y un segundo vector que lleva el gen (o genes) marcador. Una proporción grande de transformantes recibe o, en el caso de plantas, comprende (hasta 40 % o más de los transformantes), ambos vectores. En el caso de transformación con Agrobacteria, los transformantes normalmente reciben solo una parte del vector, es decir, la secuencia flanqueada por el ADN-T, que normalmente representa el casete de expresión. Los genes marcadores pueden retirarse posteriormente de la planta transformada realizando cruzamientos. En otro procedimiento, se utilizan genes marcadores integrados en un transposón para la transformación junto con 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 se ha producido la transformación exitosamente y se pierde. En un número adicional de casos, el transposón salta a una localización diferente. En estos casos el gen marcador debe eliminarse realizando cruzamientos. En microbiología, se desarrollaron 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 el sistema Cre/lox. Cre1 es una recombinasa que elimina 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 de manera satisfactoria, 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). Como se describe en el presente documento 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.

45 Transgénico/Transgén/Recombinante

Para el propósito 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, como se describe en el presente documento, todas estas construcciones llevadas a cabo por procedimientos recombinantes en los que

- (a) las secuencias de ácidos nucleicos que codifican las proteínas útiles en los procedimientos como se describe en el presente documento, o
- (b) la secuencia (o secuencias) de control genético que se une operativamente con la secuencia de ácidos nucleicos como se describe en el presente documento, 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 de una biblioteca genómica. En el caso de una biblioteca genómica, 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, especialmente 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 como se describe en el presente documento, como se define anteriormente – llega a ser un casete de expresión transgénica cuando este 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 tanto, una planta transgénica, como se describe en el presente documento, se entiende que significa, como se ha indicado anteriormente, que los ácidos nucleicos utilizados en el procedimiento, como se describe en el presente documento, 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, como se describe en el presente documento o se usan 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, como se describe en el presente documento, 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 utilizado para la transferencia. El tejido de planta capaz de propagación clonal posterior, mediante organogénesis o embriogénesis, puede transformarse con una construcción genética, como se describe en el presente documento, 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, megagametófitos, 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 transitoria o establemente en una célula huésped y puede mantenerse no integrado, por ejemplo, como un plásmido. Como alternativa, puede integrarse en el genoma huésped. La célula de planta transformada resultante puede después utilizarse 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 de plantas es actualmente una técnica bastante habitual. Ventajosamente, puede usarse cualquiera de los diversos procedimientos de transformación para introducir el gen de interés en una célula antecesora adecuada. Los procedimientos 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 utilizando virus o polen y microinyección. Pueden seleccionarse procedimientos entre, el procedimiento de calcio/poli-etilenglicol para protoplastos (Krens, F.A. y col., (1982) Nature 296, 72-74; Negrutiu I y col. (1987) Plant Mol Biol 8: 363-373); electroporación de protoplastos (Shillito R.D. y col. (1985) Bio/Technol 3, 1099-1102); microinyección en el material de planta (Crossway A y col., (1986) Mol. Gen Genet 202: 179-185); bombardeo de partículas revestido con ADN o ARN (Klein TM y col., (1987) Nature 327: 70) infección con virus (no integrantes) y similares. Las plantas transgénicas, 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 las agrobacterias actúen en las semillas de la planta o inocular las agrobacterias en el meristemo de la planta. Se ha demostrado de un modo particularmente conveniente, como se describe en el presente documento, 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 mediada por *Agrobacterium* de arroz incluyen procedimientos bien conocidos para la transformación del arroz, tales como los descritos en cualquiera de las siguientes referencias bibliográficas: la 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), cuyas descripciones se incorporan por referencia en el presente documento en su totalidad. 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), cuyas descripciones se incorporan por referencia como si se expusieran en el presente documento. 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 la transformar *Agrobacterium tumefaciens*, por ejemplo, pBin19 (Bevan y col., Nucl. Acids Res. 12 (1984) 8711). La agrobacteria transformada por dicho vector puede después utilizarse de una manera conocida para la transformación de plantas, tales como las plantas utilizadas como un modelo, como *Arabidopsis* (*Arabidopsis thaliana* no se considera como una planta de cultivo) o plantas de cultivo tales como, por ejemplo, plantas de tabaco, por ejemplo sumergiendo hojas magulladas u hojas 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; en 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 meristemas 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, págs. 274-289]. Los procedimientos alternativos se basan en la eliminación repetida de las influorescencias y la 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 final de tiempo (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 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 incuba 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 de cloroplasto se realiza generalmente mediante un proceso 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 de cloroplasto. Estas secuencias flanqueantes homólogas dirigen la integración específica de sitio en el plástoma. 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):v425-38 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 co-integrado transitorio (Klaus y col., 2004, Nature Biotechnology 22(2), 225-229).

40 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 en la dirección 3' o 5' 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 inducido. El promotor se incorpora típicamente en un ADN-T. Este ADN-T se inserta al azar en 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 debido a expresión modificada de los genes cerca al promotor introducido.

50 TILLING

El término "TILLING" es una abreviatura de "Targeted Induced Local Lesions In Genomes" (Inducción Dirigida de Lesiones Locales en el Genoma) y se refiere a una tecnología de mutagénesis que es 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 (por ejemplo, si las mutaciones afectan 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 que típicamente se siguen en el TILLING son: (a) mutagénesis EMS (Redei GP y Koncz C (1992) En Methods in Arabidopsis Research, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, págs. 16-82; Feldmann y col., (1994) En Meyerowitz EM, Somerville CR, eds, Arabidopsis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; Lightner J y Caspar T (1998) En J Martinez- Zapater, J Salinas, eds, Methods on Molecular Biology, Vol. 82. Humana Press, Totowa, NJ, págs. 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, donde 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. En la técnica se conocen bien procedimientos de TILLING (McCallum y col., (2000) Nat Biotechnol 18: 455-457; revisado por Stemple (2004) Nat Rev Genet 5(2): 145-50).

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).

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 en base a su número, tamaño y/o peso, o al rendimiento real es el rendimiento por metro cuadrado de un cultivo y año, que se determina dividiendo el rendimiento total (incluyendo la producción cosechada y valorada) por metro cuadrado plantado. El término “rendimiento” de una planta puede relacionarse 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 (con el crecimiento de 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 la 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 un 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 misma 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) índice 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 semillas aumentado, y también puede ser el resultado de un aumento en el tamaño de embriones y/o endospermo.

También puede manifestarse un aumento en el rendimiento de 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 semilla en sí mismo como un aumento en el área de semilla y/o longitud de semilla y/o anchura de semilla y/o 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 pixel 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 pixeles para el que la proporción de 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 reducida, 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.

5 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, donde 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 donde 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 excelsea*, *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*, *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., *Triticosecale rimpaui*, *Triticum* spp. (por ejemplo *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum* 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.

45 Descripción detallada

Descripción detalla para el polipéptido de Anquirina – dedo de Zn

Sorprendentemente, la memoria descriptiva describe que la modulación de la expresión en una planta de un ácido nucleico que codifica el polipéptido AZ de *Arabidopsis* (AtAZ) o un homólogo del mismo proporciona plantas que tienen un rendimiento aumentado con respecto a plantas de control. De acuerdo con una realización de la presente invención, la memoria descriptiva describe un procedimiento para aumentar el rendimiento de plantas, que comprende modular la expresión en una planta de un ácido nucleico que codifica el polipéptido AZ o un homólogo del mismo. Ventajosamente, la realización de los procedimientos de acuerdo con la presente memoria descriptiva da como resultado plantas que tienen rendimiento aumentado, particularmente rendimiento de semillas, con respecto a plantas de tipo silvestre correspondientes.

Una “referencia”, “planta de referencia”, “control”, “planta de control”, “tipo silvestre” o “planta de tipo silvestre” es en particular una célula, un tejido, un órgano, una planta, o una parte de la misma, que no se produjo de acuerdo con el procedimiento descrito en el presente documento. Por consiguiente, las expresiones “tipo silvestre”, “control” o “referencia” son intercambiables y pueden ser una célula o una parte de la planta tal como un orgánulo o tejido, o una planta, que no se ha modificado o tratado de acuerdo con el procedimiento descrito en el presente documento. Por consiguiente, la célula o una parte de la planta tal como un orgánulo o una planta utilizada como tipo silvestre,

control o referencia corresponde a la célula, planta o parte de la misma tanto como sea posible y es en cualquier otra propiedad aunque en el resultado del proceso como se describe en el presente documento sea idéntica a la materia objeto como se describe en el presente documento tanto como sea posible. Por lo tanto, el tipo silvestre, control o referencia se trata de modo idéntico o tan idéntico como sea posible, diciendo únicamente que las condiciones o propiedades deben ser diferentes, que no influyen en la calidad de la propiedad sometida a ensayo. Esto significa en otras palabras que el tipo silvestre indica (1) una planta, que lleva la forma no alterada o no modulada de un gen o alelo o (2) el material/planta de partida a partir del cual se producen las plantas mediante el proceso o procedimiento descrito en el presente documento.

Preferentemente, cualquier comparación entre las plantas de tipo silvestre y las plantas producidas por el procedimiento de la invención se realiza en condiciones análogas. La expresión "condiciones análogas" significa que todas las condiciones tales como, por ejemplo, cultivo o condiciones de cultivo, condiciones de ensayo (tal como composición del tampón, temperatura, sustratos, cepa patógena, concentración y similar) se mantienen idénticas entre los experimentos a comparar.

La "referencia", "control" o "tipo silvestre" es preferentemente un sujeto, por ejemplo, un orgánulo, una célula, una tejido, una planta, que no se ha modulado, modificado o tratado de acuerdo con el procedimiento descrito en el presente documento como se describe en el presente documento y es en cualquier otra característica tan similar a la materia objeto como se describe en el presente documento como sea posible. La referencia, control o tipo silvestre es en su genoma, transcriptoma, proteoma o metaboloma tan similar como sea posible al sujeto como se describe en el presente documento. Preferentemente, la expresión orgánulo, célula, tejido o planta de "referencia", "control" o "tipo silvestre" se refiere a un orgánulo, célula, tejido o planta, que es casi genéticamente idéntico al orgánulo, célula, tejido o planta, como se describe en el presente documento o una parte de la misma, preferentemente 95 %, más preferentemente 98 %, incluso más preferentemente 99,00 %, en particular 99,10 %, 99,30 %, 99,50 %, 99,70 %, 99,90 %, 99,99 %, 99,999 % o más. Más preferentemente, la "referencia", "control" o "tipo silvestre" es un sujeto, por ejemplo, un orgánulo, una célula, un tejido, una planta, que es genéticamente idéntico a la planta, orgánulo, célula usado de acuerdo con el procedimiento como se describe en el presente documento, excepto que las moléculas de ácido nucleico o el producto génico codificado por las mismas se cambia, modula o modifica de acuerdo con el procedimiento de la invención.

El término "expresión" o "expresión génica" es como se define en el presente documento, preferentemente resulta en la aparición de un rasgo fenotípico como una consecuencia de la transcripción de un gen o genes específicos.

El aumento se refiere a la actividad de las cantidades del polipéptido en una célula, un tejido, un orgánulo, un órgano o un organismo o una parte de los mismos, preferentemente a al menos 5 %, preferentemente al menos 10 % o al menos 15 %, especialmente preferentemente al menos 20 %, 25 %, 30 % o más, muy especialmente preferentemente es al menos 40 %, 50 % o 60 %, más preferentemente es al menos 70 % o más en comparación con el control, referencia o tipo silvestre.

La expresión "rendimiento aumentado" como se define en el presente documento se considera que significa un aumento en biomasa (peso) de una o más partes de una planta, que puede incluir las partes sobre la superficie (cosechables) y/o partes (cosechables) por debajo de la superficie. En una realización preferida, el rendimiento aumentado es rendimiento de semilla aumentado.

Por lo tanto, dichas partes cosechables son preferentemente 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.

Un aumento en el rendimiento de semilla también puede manifestarse como un aumento del tamaño de la semilla y/o del volumen de la semilla, que también puede influir en la composición de las semillas (incluyendo contenido y composición total de aceite, proteína y carbohidrato).

Tomando el maíz como un ejemplo, un aumento en el rendimiento puede manifestarse como uno o más de los siguientes: aumento en el número de plantas por metro cuadrado, un aumento en el número de espigas por planta, un aumento en el número de filas, número de granos por fila, peso de grano, peso de mil granos, longitud/diámetro de espiga, aumento en el índice 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 un ejemplo, un aumento en el rendimiento puede manifestarse como un aumento en uno o más de los siguientes: número de plantas por metro cuadrado, número de panículas por planta, número de espiguillas por panícula, número de flores (floretes) por panícula (que se expresa como una proporción del número de semillas llenas dividido entre el número de panículas primarias), aumento en el índice de llenado de semilla (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), aumento en peso de mil granos, entre otros. Un aumento en el rendimiento también puede ser resultado de arquitectura modificada o puede ocurrir como un resultado de arquitectura modificada.

De acuerdo con la memoria descriptiva, la realización de los procedimientos de la invención puede dar como resultado plantas que tienen un rendimiento aumentado, particularmente producción de la semilla. Por lo tanto, de

acuerdo con la presente memoria descriptiva, se proporciona un procedimiento para aumentar el rendimiento de una planta, cuyo procedimiento comprende modular la expresión en una planta de un ácido nucleico que codifica el polipéptido AZ o un homólogo del mismo.

5 Dado que las plantas transgénicas como se describe en el presente documento tienen rendimiento aumentado, 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 vigor temprano, tasa de crecimiento, índice de verdor, tiempo de floración y velocidad de la maduración de 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 durante sustancialmente el ciclo de vida completo de la planta. Una tasa de crecimiento aumentada durante las etapas tempranas en el ciclo de vida de una planta puede reflejar un vigor potenciado. El aumento en la tasa de crecimiento puede alterar el ciclo de la cosecha de una planta lo que permite que las plantas se siembren más tarde y/o se cosechen más pronto de lo que de otra manera sería posible (puede obtenerse un efecto similar con el 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 posible. Alterar el ciclo de cosecha de una planta puede conducir a un aumento en producción de biomasa anual por metro cuadrado (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 las limitaciones territoriales para cosechar un cultivo a menudo se determinan por condiciones ambientales adversas en el momento de plantar (estación temprana) o en el momento de cosechar (estación tardía). Dichas condiciones adversas pueden evitarse si se acorta el ciclo de cosecha. La tasa de crecimiento puede determinarse derivando diversos parámetros de curvas de crecimiento, dichos parámetros pueden ser: T-Medio (el tiempo que tarda una planta en alcanzar el 50 % de su tamaño máximo) y T-90 (el tiempo que tarda una planta en alcanzar el 90 % de su tamaño de máximo), entre otros.

35 De acuerdo con una característica preferida de la presente memoria descriptiva, la realización de los procedimientos de la invención proporciona plantas que tienen una tasa de crecimiento aumentada o rendimiento aumentado en comparación con plantas de control. Por lo tanto, de acuerdo con la presente memoria descriptiva, se proporciona un procedimiento para aumentar el rendimiento y/o la tasa de crecimiento en plantas, cuyo procedimiento comprende modular la expresión en una planta de un ácido nucleico que codifica el polipéptido AZ o un homólogo del mismo.

40 Un aumento en el rendimiento y/o tasa de crecimiento se produce cuando la planta está en condiciones de cero estrés o cuando la planta está expuesta a diversos tipos de estrés en comparación con las plantas de control. Típicamente las plantas responden a exposición frente al estrés creciendo más lentamente. En condiciones de estrés intenso, la planta puede incluso detener por completo el crecimiento. Por otro lado, estrés leve se define en el presente documento como cualquier estrés al que se expone una planta que no da como resultado el cese total del crecimiento de la planta sin la capacidad de reanudar el crecimiento. El estrés leve, como se describe en el presente documento, 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 del 14 %, 13 %, 12 %, 11 % o 10% o menos en comparación con la planta de control en condiciones de cero estrés. Debido a los avances en las prácticas agrícolas (riego, fertilización, tratamientos con pesticidas) las formas de estrés intenso no se encuentran frecuentemente en plantas cultivadas. Como una consecuencia, el crecimiento comprometido inducido por estrés leve es a menudo una característica indeseable para la agricultura. El estrés leve es un estrés biótico y/o abiótico (ambiental) al cual está expuesta la planta. El estrés abiótico puede deberse a sequía o a exceso de agua, estrés anaeróbico, estrés salino, toxicidad química, estrés oxidativo y temperaturas altas, frías o heladas. El estrés abiótico puede ser un estrés osmótico provocado por estrés al agua (particularmente debido a sequía), estrés salino, estrés oxidativo y un estrés iónico. El estrés biótico es típicamente aquel estrés provocado por patógenos, tales como bacterias, virus, hongos e insectos.

55 En particular, los procedimientos descritos en el presente documento pueden realizarse en condiciones de cero estrés o en condiciones de sequía leve para dar plantas con un rendimiento aumentado con relación a plantas de control. Como describen Wang y col. (Planta (2003) 218: 1-14), el estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos o moleculares que afectan adversamente al crecimiento de las plantas y a su productividad. Se sabe que la sequía, salinidad, temperaturas extremas y el estrés oxidativo se interconectan y que pueden inducir daño celular y en el crecimiento a través de mecanismos similares. Rabbani y col. (Plant Physiol (2003) 133: 1755-1767) describen un grado particularmente alto de "interferencia" entre estrés por sequía y estrés por alta salinidad. Por ejemplo, la sequía y/o salinización se manifiestan principalmente como estrés osmótico, que

produce la interrupción de homeostasis y distribución de iones en la célula. El estrés oxidativo, que frecuentemente acompañada alta o baja temperatura, estrés por salinidad o sequía, puede producir desnaturalización de las proteínas funcionales y estructurales. Como una consecuencia, este diverso estrés ambiental a menudo activa rutas de señalización celular y respuestas celulares similares, tales como la producción de proteínas de estrés, regulación por aumento de antioxidantes, acumulación de solutos compatibles y paralización del crecimiento. La expresión condiciones de “no estrés”, como se usa en el presente documento, son aquellas condiciones ambientales que no imponen estrés a las plantas, tales como diversos tipos de estrés descritos anteriormente. Las condiciones de no estrés permiten el crecimiento óptimo de las plantas. Los expertos en la técnica son conscientes de las condiciones normales en el suelo y de las condiciones climáticas en una localización determinada.

La realización de los procedimientos como se describe en el presente documento da lugar a plantas cultivadas en condiciones de cero estrés o en condiciones de sequía leve de rendimiento aumentado con respecto a plantas de control que crecen en condiciones comparables. Por lo tanto, de acuerdo con la presente memoria descriptiva, se proporciona un procedimiento para aumentar el rendimiento en plantas que crecen en condiciones de cero estrés o en condiciones de sequía leve, cuyo procedimiento comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido AZ.

Como se describe en el presente documento, el aumento del rendimiento y/o tasa de crecimiento se produce de acuerdo con los procedimientos de la presente memoria descriptiva en condiciones de cero estrés.

La realización de los procedimientos como se describe en el presente documento da lugar a plantas que crecen en condiciones de déficit de nutrientes, particularmente en condiciones de déficit de nitrógeno, rendimiento aumentado con respecto a plantas de control que se cultivan en condiciones comparables. Por lo tanto, de acuerdo con la presente memoria descriptiva, se proporciona un procedimiento para aumentar el rendimiento en plantas que se cultivan en condiciones de déficit de nutrientes, cuyo procedimiento comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido AZ. El déficit de nutrientes puede resultar de una carencia de nutrientes, tales como nitrógeno, fosfatos y otros compuestos que contienen fósforo, potasio, calcio, cadmio, magnesio, manganeso, hierro y boro, entre otros.

Los procedimientos como se describe en el presente documento son ventajosamente aplicables a cualquier planta. Las características de cultivo mencionadas anteriormente pueden modificarse ventajosamente en cualquier planta. Las plantas que son particularmente útiles en los procedimientos descritos en el presente documento incluyen todas las plantas que pertenecen a la superfamilia *Viridiplantae*, en particular plantas monocotiledóneas y dicotiledóneas que incluyen forraje o leguminosas forrajeras, plantas ornamentales, cultivos alimenticios, árboles o arbustos. Como se describe en el presente documento, la planta es una planta de cultivo. 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.

Otras plantas ventajosas se seleccionan del grupo que consiste en Asteraceae, tal como los géneros *Helianthus*, *Tagetes*, por el ejemplo las especies *Helianthus annuus* [girasol], *Tagetes lucida*, *Tagetes erecta* o *Tagetes tenuifolia* [caléndula]; *Brassicaceae* tal como el género *Brassica*, por ejemplo las especies *Brassica napus*, *Brassica rapa* ssp. [canola, colza, nabo]; *Fabaceae* tal como el género *Glycine* por ejemplo las especies *Glycine max*, *Soja hispida* o *Soja max* [soja]; *Linaceae* tal como el género *Linum*, por ejemplo las especies *Linum usitatissimum*, [lino, linaza]; *Poaceae* tal como los géneros *Hordeum*, *Secale*, *Avena*, *Sorghum*, *Oryza*, *Zea*, *Triticum* por ejemplo las especies *Hordeum vulgare* [cebada], *Secale cereale* [centeno], *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida* [avena], *Sorghum bicolor* [sorgo, mijo], *Oryza sativa*, *Oryza latifolia* [arroz], *Zea mays* [mazorca, maíz] *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernum*, *Triticum macha*, *Triticum sativum* o *Triticum vulgare* [trigo, trigo harinero, trigo común]; *Solanaceae* tal como los géneros *Solanum*, *Lycopersicon* por ejemplo las especies *Solanum tuberosum* [patata], *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon pyriforme*, *Solanum integrifolium* o *Solanum lycopersicum* [tomate].

La expresión “polipéptido AtAZ o un homólogo del mismo” como se define en el presente documento se refiere a proteínas que comprenden al menos una repetición de anquirina y al menos un dominio C3H1 de dedo de Cinc, cuya repetición de anquirina se localiza cadena arriba del dominio C3H1. Preferentemente, la proteína AZ comprende dos repeticiones de anquirina y dos dominios C3H1 de dedo de Cinc, tal como en la proteína representada en la SEC ID N°: 2. Además preferentemente, las dos repeticiones de anquirina se localizan próximas entre sí. Preferentemente de manera adicional, los dominios C3H1 de dedo de Cinc también se localizan próximos entre sí y en el extremo C de las repeticiones de anquirina. En la SEC ID N°: 2, las repeticiones de anquirina se localizan en las posiciones D90 a R120 y D125 a L157, los dos dominios de dedo de Cn se localizan en las posiciones H301 a V327 y Q336 a P359 (Figura 1).

También preferentemente, la proteína AtAZ comprende al menos una de las siguientes secuencias consenso:

(P/A)CSRAY(S/T)HDWTEC (motivo 1, SEC ID N°: 3)
HPGENARRRDPR (motivo 2, SEC ID N°: 4)

HG(V/I)FE(C/S)WLHP(A/S)QY(R/K)TRLCK (motivo 3, SEC ID N°: 5)

CFFAH (motivo 4, SEC ID N°: 6)

Preferentemente el motivo 1 es PCSRAYSHDWTEC, y el motivo 3 es preferentemente HGVFECWLHPAQYRTRLCK.

- 5 Más preferentemente, la proteína AtAZ comprende dos de los motivos mencionados anteriormente, especialmente de modo preferente 3 de los motivos mencionados anteriormente, más preferentemente los cuatro motivos.

La repetición de anquirina (SMART SM00248, Interpro IPR002110), como se describe en la base de datos Interpro, es uno de los motivos de interacción proteína-proteína más habituales en la naturaleza. Las repeticiones de anquirina son moléculas repetidas (normalmente en tándem) normalmente de aproximadamente 33 aminoácidos. Se producen en una gran cantidad de proteínas funcionalmente diversas principalmente de eucariotas. Los pocos ejemplos conocidos de procariontas y virus pueden ser el resultado de transferencias génicas horizontales. La repetición se ha encontrado en proteínas de diversa función tales como iniciadores transcripcionales, reguladores del ciclo celular, transportadores citoesqueléticos, iónicos y transductores de señales. El pliegue de anquirina parece definirse por su estructura en lugar de por su función ya que no reconoce universalmente ninguna secuencia o estructura específica. El plegamiento conservado de la unidad de repetición de anquirina se conoce a partir de diversas estructuras cristalinas y en solución. Cada repetición se pliega en una estructura de hélice-bucle-hélice con una región beta-horquilla/bucle que se proyecta fuera de las hélices a un ángulo de 90°. Las repeticiones se apilan a la vez para formar una estructura en forma de L.

Se piensa que el dominio de dedo de Cinc ZnF_C3H1 (también conocido como Znf_CCCH, SMART SM00356; Interpro IPR000571), como se describe en la base de datos Interpro, está implicado en la unión a ADN. Los dedos de Cinc existen como tipos diferentes, dependiendo de las posiciones de los restos de cisteína. Las proteínas que contienen dominios de dedo de Cinc del tipo C-x8-C-x5-C-x3-H (representando x en el presente documento cualquier aminoácido y los dígitos 8, 5 y 3 el número de aminoácidos entre los restos C o H conservados) incluyen proteínas de dedo de cinc de eucariotas implicadas en el ciclo celular o regulación relacionada con la fase de crecimiento, por ejemplo TIS11B humano (factor 1 de respuesta a butirato), una probable proteína reguladora implicada en la regulación de la respuesta a factores de crecimiento, y la proteína nuclear inducible por el factor de crecimiento TTP de ratón, que tiene la misma función. La proteína TTP de ratón está inducida por factores de crecimiento. Otra proteína que contiene este dominio es la subunidad de 35 Kd del factor de corte y empalme humano U2AF, que desempeña una función crítica en el corte y empalme dependiente de potenciador y constitutivo mediando interacciones proteína-proteína esenciales e interacciones proteína-ARN necesarias para la selección en el sitio de corte y empalme 3'. Se ha observado que diferentes proteínas de dedo de Cinc CCCH interaccionan con la región no traducida 3' de diversos ARNm. Es muy frecuente que este tipo de dedo cinc esté presente en dos copias.

La Figura 2 describe las secuencias consenso como se definen en la base de datos SMART para el dominio de anquirina y el dominio de dedo de cinc; sin embargo debe observarse que estas secuencias consenso podrían estar sesgadas hacia secuencias de proteínas animales.

Los términos "domino" y "motivo" se definen en la sección "definiciones" del presente documento. Existen bases de datos especialistas para la identificación de dominios. El dominio C3H1 o de anquirina en una proteína AZ puede identificarse usando, 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 sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., págs. 53-61, AAAIPress, Menlo Park; Hulo y col., Nucl. Acids. Res. 32: D134-D137, (2004)) or Pfam (Bateman y col., Nucleic Acids Research 30(1): 276-280 (2002)) o Pfam (Bateman y col., Nucleic Acids Research 30(1): 276-280 (2002)). En el servidor proteómico ExpASY se encuentra disponible un conjunto de herramientas para análisis por ordenador de las secuencias de proteína (proporcionado por el Swiss Institute of Bioinformatics (Gasteiger y col., ExpASY: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31: 3784-3788(2003)). La secuencia de la proteína AZ se analizó con la herramienta SMART (versión 4.1; Schultz y col. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic y col. (2002) Nucleic Acids Res 30, 242-244) y se usó para seleccionar la base de datos Pfam (Versión 17.0, marzo de 2005; Bateman y col. (2004) Nucl. Acids Res. 32, D138-141) e InterPro (Release 11.0, 26 de julio de 2005; Mulder y col. (2005) Nucl. Acids. Res. 33, D201-205).

Alineando otras secuencias de proteína con la SEC ID N°: 2, las secuencias consenso correspondientes, el dominio C3H1, el dominio de anquirina u otros motivos de secuencia pueden identificarse fácilmente. De este modo, los polipéptidos AZ u homólogos de los mismos (que incluyen ortólogos y parálogos) pueden identificarse fácilmente, utilizando técnicas rutinarias bien conocidas en la técnica, tal como mediante alineamiento de secuencias. Los procedimientos para el alineamiento de las secuencias para comparación son bien conocidos en la técnica, dichos procedimientos incluyen GAP, BESTFIT, BLAST, FASTA y TFASTA. GAP utiliza el algoritmo de Needleman y Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar el alineamiento de las dos secuencias que maximiza 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 desarrolla 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. Los homólogos pueden identificarse fácilmente utilizando, por ejemplo, el algoritmo de alineamiento de secuencia múltiple ClustalW (versión 1.83), con los parámetros de alineamiento por pares predeterminados, y un procedimiento de 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. 10 de julio, 2003; 4: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando las secuencias de ADN o proteína.). Puede realizarse edición manual menor para optimizar el alineamiento entre los motivos conservados, como sería obvio para un experto en la técnica. Adicionalmente, en lugar de utilizar las secuencias de longitud completa para la identificación de homólogos, también pueden utilizarse dominios específicos (tal como el dominio C3H1 o de anquirina o uno de los motivos definidos anteriormente). Los valores de identidad de secuencia, que se indican adelante como un porcentaje, se determinaron sobre el dominio conservado o sobre la secuencia de ácidos nucleicos o aminoácidos completa utilizando los programas mencionados anteriormente utilizando parámetros predeterminados.

En la Tabla 1 del Ejemplo 1 se proporcionan ejemplos de proteínas AZ u homólogos de las mismas.

Las secuencias que se encuentran bajo la definición de "polipéptido AZ u homólogo del mismo" como se define en la reivindicación 1 y preferentemente comprenden también al menos una de las secuencias consenso de las SEC ID Nos: 3, 4, 5 o 6 como se define anteriormente, pueden ser adecuadas para su uso en los procedimientos de la invención. Preferentemente, el polipéptido es un polipéptido de *Arabidopsis thaliana*.

En el término "homólogos" se incluyen las secuencias ortólogas y las secuencias parálogas. Ortólogos y parálogos pueden encontrarse realizando una búsqueda BLAST denominada recíproca. Esta puede realizarse mediante un primer BLAST que implica una secuencia de consulta BLASTing (por ejemplo, SEC ID N°: 1 o SEC ID N°: 2) contra cualquier base de datos de secuencia, tal como la base de datos NCBI públicamente disponible. BLASTN o TBLASTX (utilizando valores predeterminados convencionales) pueden utilizarse 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 deriva la secuencia de consulta (donde la secuencia de consulta es la SEC ID N°: 1 o SEC ID N°: 2, por lo que el segundo BLAST se realizaría contra las secuencias de *Arabidopsis*). Después se comparan los resultados del primer y segundo BLAST. Se identifica un parólogo si un acierto de alta clasificación del segundo BLAST es de la misma especie de la que deriva la secuencia de consulta; un ortólogo se identifica si un acierto de alta clasificación no es de la misma especie del que deriva la de secuencia de consulta. Los ortólogos preferidos son los ortólogos de la SEC ID N°: 1 o SEC ID N°: 2. 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 ácidos nucleicos (o polipéptido) sobre una longitud particular. Preferentemente, la puntuación es mayor de 50, más preferentemente mayor de 100; y preferentemente el valor E es menor de e^{-5} , más preferentemente menor de e^{-6} . 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. Ejemplos de secuencias ortólogas a la SEC ID N°: 2 incluyen la SEC ID N°: 11, SEC ID N°: 15 y SEC ID N°: 17. La SEC ID N°: 19 es un ejemplo de un parólogo de la SEC ID N°: 2.

La memoria descriptiva describe proteínas de AZ que tienen, además de al menos una repetición de anquirina y al menos un dominio C3H1, en orden creciente de preferencia, una identidad de secuencia de al menos 26 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % con la proteína de SEC ID N°: 2. La matriz mostrada en la Figura 3 (Matriz A) muestra similitudes e identidades (en negrita) sobre la longitud completa de diversas proteínas AZ. En caso de comparar solo dominios específicos, la identidad o similitud puede ser mayor entre las proteínas diferentes (Matriz B: comparación de una secuencia con dominio de dedo de Cn).

Puede realizarse un ensayo para determinar la actividad AZ. La actividad de unión a ADN e interacciones proteína-proteína pueden determinarse fácilmente *in vitro* o *in vivo* usando técnicas bien conocidas en el campo. Ejemplos de ensayos *in vitro* para la actividad de unión a ADN incluyen: análisis de retraso en gel o análisis de un híbrido en levadura. Un ejemplo de un ensayo *in vitro* para interacciones proteína-proteína es el análisis de doble híbrido en levadura (Fields y Song (1989) Nature 340: 245-6).

Además, como se describe en el presente documento, la expresión en plantas de la proteína AZ o de un homólogo de la misma, y en particular en arroz, tiene el efecto de aumentar el rendimiento de la planta transgénica cuando se compara con plantas de tipo silvestre correspondientes, donde el rendimiento aumentado comprende al menos uno de: peso total de semillas, número total de semillas y número de semillas llenas.

Un polipéptido de AZ u homólogo del mismo, como se define en el presente documento, está codificado por un ácido

nucleico/gen AZ. Por lo tanto, la expresión “ácido nucleico/gen AZ” como se define en el presente documento, es cualquier ácido nucleico/gen que codifica un polipéptido AZ o un homólogo del mismo como se define anteriormente. Los ejemplos de ácidos nucleicos AZ incluyen, pero sin limitación, los representados en la Tabla A del Ejemplo 1. Los ácidos nucleicos/genos AZ y variantes de los mismos pueden ser adecuados en la realización práctica de los procedimientos descritos en el presente documento. Preferentemente, las variantes de un gen AZ originado de *Arabidopsis thaliana*. Los ácidos nucleicos/genos de AZ variantes incluyen partes de un ácido nucleico/gen AZ, variantes de cortes y empalme, variantes alélicas y/o ácidos nucleicos que pueden hibridarse con un ácido nucleico/gen AZ.

En el presente documento la referencia a una “secuencia de ácidos nucleicos” se considera que significa una forma polimérica de un polímero de desoxirribonucleótido o ribonucleótido de cualquier longitud, mono o bicatenario o análogos del mismo, que tiene la característica esencial de un ribonucleótido natural en que puede hibridarse con secuencias de ácidos nucleicos de una manera similar a polinucleótidos de origen natural.

El término porción, como se describe en el presente documento, se refiere a un segmento de ADN que codifica un polipéptido que comprende al menos una repetición de anquirina y al menos un dominio de dedo Zn C3H1. Una porción puede prepararse, por ejemplo, realizando una o más deleciones en un ácido nucleico AZ. Las porciones pueden usarse en forma aislada o pueden fusionarse con otras secuencias codificantes (o no codificantes) para, por ejemplo, producir una proteína que combine varias actividades. Cuando se fusiona con otras secuencias codificantes, el polipéptido resultante producido después de la traducción puede ser más grande que el esperado para el fragmento AZ. La porción tiene típicamente al menos 500, 700 o 900 nucleótidos de longitud, preferentemente al menos 1100, 1300 o 1500 nucleótidos de longitud, más preferentemente al menos 1700, 1900 o 2100 nucleótidos de longitud y más preferentemente al menos 2300 o 2400 nucleótidos de longitud. Preferentemente, la porción es una porción de un ácido nucleico como se representa en la Tabla A del Ejemplo 1. Más preferentemente, la porción de un ácido nucleico AZ es como se representa en la SEC ID N°: 1 o SEC ID N°: 53.

Las expresiones “fragmento”, “fragmento de una secuencia” o “parte de una secuencia”, “porción” o “porción de la misma” significa una secuencia truncada de la secuencia original a la que se hace referencia. La secuencia truncada (secuencia de ácidos nucleicos o proteínas) puede variar ampliamente en longitud; siendo el tamaño mínimo una secuencia de suficiente tamaño para proporcionar una secuencia con al menos una función y/o actividad comparable de la secuencia original a la que se hace referencia o que se hibrida, en condiciones rigurosas, con la molécula de ácido nucleico como se describe en el presente documento o se usa en el procedimiento como se describe en el presente documento, aunque que el tamaño máximo no es crítico. En algunas aplicaciones, normalmente el tamaño máximo no es sustancialmente mayor que el necesario para proporcionar la actividad y/o función (o funciones) deseadas de la secuencia original. Una función comparable significa al menos 40 %, 45 % o 50 %, preferentemente al menos 60 %, 70 %, 80 % o 90 % o más de la función de la secuencia original.

Otra variante de un ácido nucleico/gen AZ como se define en el presente documento es un ácido nucleico que puede hibridarse en condiciones de rigurosidad reducida, preferentemente en condiciones rigurosas, con un ácido nucleico/gen AZ como se ha definido en el presente documento anteriormente o con una porción como se ha definido en el presente documento anteriormente. La secuencia de hibridación tiene típicamente al menos 300 nucleótidos de longitud, preferentemente al menos 400 nucleótidos de longitud, más preferentemente al menos 500 nucleótidos de longitud y más preferentemente al menos 600 nucleótidos de longitud.

Preferentemente, la secuencia de hibridación es una que puede hibridarse con un ácido nucleico como se representa por la SEC ID N°: 1, SEC ID N°: 10, SEC ID N°: 12, SEC ID N°: 14, SEC ID N°: 16, SEC ID N°: 18, SEC ID N°: 20, SEC ID N°: 22, SEC ID N°: 24, SEC ID N°: 26, SEC ID N°: 28, SEC ID N°: 30, SEC ID N°: 32, SEC ID N°: 34, SEC ID N°: 36, SEC ID N°: 38, SEC ID N°: 40, SEC ID N°: 42 o SEC ID N°: 53 o con una porción de cualquiera de las secuencias mencionadas anteriormente, una porción como se ha definido anteriormente. Más preferentemente, la secuencia de hibridación puede hibridarse con la SEC ID N°: 1, con la SEC ID N°: 53 o con porciones (o sondas) de las mismas. En la técnica se conocen bien procedimientos para diseñar sondas. Las sondas tienen generalmente una longitud menor de 1000 pb, 900 pb, 800 pb, 700 pb, 600 pb, preferentemente menor de 500 pb, 400 pb, 300 pb 200 pb o 100 pb. Normalmente, las longitudes de sonda para hibridaciones de ADN-ADN, tal como transferencia de Southern, varían entre 100 y 500 pb, mientras que la región de hibridación en sondas para hibridaciones de ADN-ADN, tales como en amplificación por PCR, generalmente son más cortas que 50 pero más largas que 10 nucleótidos, preferentemente tienen una longitud de 15, 20, 25, 30, 35, 40, 45 o 50 pb.

Son también útiles en los procedimientos de la invención, ácidos nucleicos que codifican homólogos (incluyendo ortólogos o parálogos) de la secuencia de aminoácidos representada por la SEC ID N°: 2 o derivados de la misma.

Otra variante de ácido nucleico útil en los procedimientos descritos en el presente documento es una variante de corte y empalme que codifica un polipéptido AZ como se define anteriormente. Como se describe en el presente documento las variantes de corte y empalme son variantes de corte y empalme del ácido nucleico que codifica un polipéptido que comprende al menos una repetición de anquirina y al menos un dominio C3H1. Preferentemente, el polipéptido AZ o el homólogo del mismo codificado por la variante de corte y empalme tiene una identidad de secuencia al menos 26 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 60

96 %, 97 %, 98 % o 99 % con la SEC ID N°: 2. Adicionalmente, la memoria descriptiva describe variantes de corte y empalme representadas por los ácidos nucleicos representados en la Tabla A del Ejemplo 1. Por ejemplo, la SEC ID N°: 25 y SEC ID N°: 47 están codificadas por variantes de corte y empalme del mismo gen. Por ejemplo la variante de corte y empalme representada por la SEC ID N°: 1 o SEC ID N°: 53.

5 Otra variante de ácido nucleico útil en los procedimientos descritos en el presente documento es una variante alélica de un ácido nucleico que codifica un polipéptido AZ como se define anteriormente. Preferentemente, el polipéptido codificado por la variante alélica está representado por las secuencias polipeptídicas indicadas en la Tabla A del Ejemplo 1. Más preferentemente, la variante alélica que codifica AZ se representa por la SEC ID N°: 1.

10 Una variante de ácido nucleico adicional útil en los procedimientos descritos en el presente documento es una variante de ácido nucleico obtenida por combinación de genes. Además, puede usarse mutagénesis dirigida a sitio para generar variantes de ácidos nucleicos AZ. Se dispone de diversos procedimientos para realizar mutagénesis dirigida a sitio; siendo los más habituales los procedimientos basados en PCR (Current Protocols in Molecular Biology. Wiley Eds.).

15 Por lo tanto, la memoria descriptiva describe un procedimiento para aumentar el rendimiento y/o la tasa de crecimiento de una planta, que comprende modular la expresión en una planta de una variante de un ácido nucleico AZ seleccionado entre:

- (i) una porción de un ácido nucleico AZ;
- (ii) un ácido nucleico que se hibrida con un ácido nucleico AZ;
- (iii) una variante de corte y empalme de un ácido nucleico que codifica un polipéptido AZ; y
- (iv) una variante alélica de un ácido nucleico que codifica un polipéptido AZ; y
- (v) una variante de ácido nucleico que codifica un polipéptido AZ obtenida por combinación de genes o mutagénesis dirigida a sitio.

25 El ácido nucleico AZ o variante del mismo, como se define en el presente documento, puede derivar de cualquier fuente natural o artificial. Este ácido nucleico puede modificarse de su forma nativa en composición y/o en medio genómico a través de manipulación humana deliberada. El ácido nucleico es preferentemente de origen vegetal, adicionalmente preferentemente de una especie dicotiledónea, más preferentemente de la familia Brassicaceae, más preferentemente de *Arabidopsis thaliana*. Más preferentemente, el ácido nucleico AZ es la secuencia de *Arabidopsis thaliana* representada por la SEC ID N°: 1, y la secuencia de aminoácidos AZ es como se representa en la SEC ID N° 47. Como alternativa, el ácido nucleico AZ representado por la SEC ID N°: 53 o la secuencia de aminoácidos AZ como se representa por la SEC ID N°: 47 también puede ser útil en los procedimientos de la presente invención.

Por lo tanto, cualquier referencia en el presente documento a un polipéptido AZ se entiende que se refiere a una proteína AZ como se define anteriormente. Cualquier ácido nucleico que codifique dicha proteína AZ es adecuado para su uso en los procedimientos como se describe en el presente documento.

35 De acuerdo con la presente memoria descriptiva, se contempla la expresión modulada, preferentemente aumentada, del ácido nucleico AZ o variante del mismo. Procedimientos para aumentar la expresión de genes o productos génicos están bien documentados en la técnica. La expresión de una secuencia de ácido nucleico que codifica un polipéptido de AZ puede modularse introduciendo una modificación genética, que puede introducirse, por ejemplo, mediante cualquiera de uno (o más) de los siguientes procedimientos: activación de ADN-T, TILLING, mutagénesis dirigida a sitio, evolución dirigida y recombinación homóloga o introduciendo y expresando en una planta un ácido nucleico que codifique un polipéptido AZ o un homólogo del mismo. Después de la introducción de la modificación genética, viene una etapa de selección para la expresión modificada de un ácido nucleico que codifica un polipéptido AZ o un homólogo del mismo, cuya modificación en la expresión da lugar a plantas con rendimiento aumentado. En la técnica también se conocen procedimientos para disminuir la expresión de genes o productos génicos.

45 La activación de ADN-T se describe anteriormente. En el locus de un gen AZ también puede introducirse una modificación genética usando la técnica de TILLING (Inducción Dirigida de Lesiones Locales en el Genoma). El locus de un gen, como se define en el presente documento, se entiende que se refiere a una región genómica, que incluye el gen de interés y 10 kb cadena arriba o cadena abajo de la región codificante. Puede usarse mutagénesis dirigida a sitio para generar variantes de ácidos nucleicos AZ. Se dispone de diversos procedimientos para realizar mutagénesis dirigida a sitio; siendo los más comunes los procedimientos basados en PCR (Current Protocols in Molecular Biology. Wiley Eds.). De acuerdo con la presente memoria descriptiva puede aplicarse recombinación homóloga.

55 La memoria descriptiva describe un procedimiento para introducir una modificación genética que consiste en introducir y expresar, en una planta, un ácido nucleico que codifica un polipéptido AZ o un homólogo del mismo como se define anteriormente. El ácido nucleico a introducir en una planta puede ser un ácido nucleico de longitud completa o puede ser una porción o una secuencia de hibridación como se ha definido anteriormente en el presente documento.

Adicionalmente, la memoria descriptiva describe construcciones genéticas y vectores para facilitar la introducción y/o

expresión de las secuencias de nucleótidos útiles en los procedimientos como se describe en el presente documento.

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

- 5 (i) un ácido nucleico AZ o variante del mismo como se define anteriormente en el presente documento;
 (ii) una o más secuencias de control unidas operativamente a la secuencia de ácidos nucleicos de (i);

10 En los procedimientos, como se describe en el presente documento, pueden crearse construcciones útiles utilizando tecnología de ADN recombinante bien conocida por los expertos en la técnica. Las construcciones génicas pueden insertarse en vectores, que pueden adquirirse en el comercio, adecuados para la transformación en plantas y adecuados para la expresión del gen de interés en las células transformadas. Por lo tanto, la memoria descriptiva describe el uso de una construcción génica, como se ha definido anteriormente en el presente documento, en los procedimientos como se describe en el presente documento.

15 Las plantas se transforman con un vector que comprende la secuencia de interés (es decir, un ácido nucleico que codifica un polipéptido AZ u homólogo del mismo). El experto es consciente de los elementos genéticos que deben estar presentes en el vector para realizar satisfactoriamente la transformación, selección y propagación de células hospedadoras que contienen la secuencia de interés. La secuencia de interés está unida operativamente a una o más secuencias de control (al menos a un promotor). Las expresiones "elemento regulador", "secuencia de control" y "promotor" se han definido anteriormente. Ventajosamente, puede utilizarse cualquier tipo de promotor para dirigir la expresión de la secuencia de ácidos nucleicos.

20 Generalmente se conocen promotores adecuados, que son funcionales en plantas. Estos pueden tener forma de promotores constitutivos o inducibles. Promotores adecuados pueden permitir la expresión específica de desarrollo y/o tejido en eucariotas multicelulares; por tanto, promotores específicos de hoja, raíz, flor, semilla, estroma, tubérculo o fruta pueden usarse ventajosamente en plantas.

Diferentes promotores de plantas que pueden usarse en plantas son promotores tales como, por ejemplo, el promotor USP, LegB4, CD3 o el promotor de ubiquitina de perejil.

25 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 necesariamente es de origen vegetal, sino que puede originarse de virus o microorganismos, en particular, por ejemplo, de virus que atacan a células de plantas.

30 El promotor "de planta" también puede originarse una célula de planta, por ejemplo, de la planta que se transforma con la secuencia de ácido nucleico a expresar en el proceso de la invención y como se describe en el presente documento. Esto también se aplica a las otras señales reguladoras "de plantas", por ejemplo en terminadores "de plantas".

35 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 en el momento temporal correcto y de una manera específica de célula o de tejido. Son promotores útiles los promotores constitutivos (Benfey y col., EMBO J. 8 (1989) 2195-2202), tales como los que se originan de virus de plantas, tal como 35S CAMV (Franck y col., Cell 21 (1980) 285-294), 19S CaMV (véanse también los documentos US 5352605 y WO 84/02913), 34S FMV (Sanger y col., Plant. Mol. Biol., 14, 1990: 433-443), el promotor de ubiquitina de perejil, o promotores de plantas tales como el promotor de la subunidad pequeña de Rubisco descrito en el documento US 4.962.028 o los promotores PRP1 de plantas [Ward y col., Plant. Mol. Biol. 22 (1993)], SSU, PGEL1, OCS [Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553-2557], lib4, usp, mas [Comai (1990) Plant Mol Biol 15 (3): 373-381], STLS1, ScBV (Schenk (1999) Plant Mol Biol 39(6): 1221-1230), B33, SAD1 o SAD2 (promotores de lino, Jain y col., Crop Science, 39 (6), 1999: 1696-1701) o nos [Shaw y col. (1984) Nucleic Acids Res. 12(20): 7831-7846]. Ejemplos adicionales de promotores constitutivos de plantas son los promotores ATPasa-V de la remolacha azucarera (documento WO 01/14572). Ejemplos de promotores constitutivos sintéticos son el promotor Super (documento WO 95/14098) y promotores derivados de cajas G (documento WO 94/12015). Si fuera apropiado, adicionalmente también se usan promotores químicamente inducibles, compárense los documentos EP-A 388186, EP-A 335528, WO 97/06268. La expresión estable, constitutiva de las proteínas como se describe en el presente documento de una planta puede ser ventajosa. Sin embargo, la expresión inducible del polipéptido como se describe en el presente documento es ventajosa, si, por ejemplo, la expresión tardía antes de la cosecha es de ventaja, ya que la manipulación metabólica puede conducir a retrasar el crecimiento de la planta.

45 La expresión de genes de plantas también puede facilitarse mediante un promotor químicamente inducible. Los promotores químicamente inducibles son particularmente adecuados cuando se desea expresar el gen de una manera específica del tiempo. Son ejemplos de dichos promotores un promotor inducible por ácido salicílico (documento WO 95/19443) y un promotor inducible por ácido abscísico (documento EP 335 528), un promotor inducible por tetraciclina (Gatz y col. (1992) Plant J. 2, 397-404), un promotor inducible por ciclohexanol o etanol (documento WO 93/21334) u otros como se describe en el presente documento.

Otros promotores adecuados son aquellos que reaccionan frente a condiciones de estrés biótico o abiótico, por ejemplo el promotor del gen PRP1 inducido por patógenos (Ward y col., Plant. Mol. Biol. 22 (1993) 361-366), el promotor hsp80 termoinducible del tomate (documento US 5.187.267), el promotor alfa-amilasa inducible por frío de la patata (documento WO 96/12814) o el promotor pinll inducible por daños (documento EP-A-0 375 091) u otros como se describe en el presente documento.

En el presente documento se describen promotores en particular aquellos que efectúan expresión de genes en tejidos y órganos, en células de semilla, tales como células de endospermo y células del embrión en desarrollo. Son promotores adecuados el promotor del gen napina de colza (documento US 5.608.152), el promotor USP de *Vicia faba* (Baeumlein y col., Mol Gen Genet, 1991, 225 (3): 459-67), el promotor de olesina de *Arabidopsis* (documento WO 98/45461), el promotor de faseolina de *Phaseolus vulgaris* (documento US 5.504.200), el promotor Bce4 de *Brassica* (documento WO 91/13980), el promotor arc5 de alubia, el promotor DcG3 de zanahoria, o el promotor B4 de Legumina (LeB4; Baeumlein y col., 1992, Plant Journal, 2 (2): 233-9) y promotores que efectúan la expresión específica de semilla en plantas monocotiledóneas tales como maíz, cebada, trigo, centeno, arroz y similares. Promotores específicos de semilla ventajosos son el promotor de proteína de unión a sacarosa (documento WO 00/26388), el promotor de faseolina y el promotor de napina. Promotores adecuados que deben tenerse en cuenta son el promotor del gen lpt2 o lpt1 de cebada (documentos WO 95/15389 y WO 95/23230) y los promotores descritos en el documento WO 99/16890 (promotores del gen de hordeína de cebada, del gen de glutelina de arroz, del gen de orizina de arroz, del gen de prolamina de arroz, del gen de gliadina de trigo, del gen de glutelina de trigo, del gen de zeína de maíz, del gen de glutelina de cebada, del gen de kasirina de sorgo y del gen de secalina de centeno). Otros promotores adecuados son Amy32b, Amy 6-6 y Aleuraina [documento US 5.677.474], Bce4 (colza) [documento US 5.530.149], glicinina (soja) [documento EP 571 741], fosfoenolpiruvato carboxilasa (soja) [documento JP 06/62870], ADR12-2 (soja) [documento WO 98/08962], isocitrato liasa (colza) [documento US 5.689.040] o α -amilasa (cebada) [documento EP 781 849]. Otros promotores que están disponibles para la expresión de genes en plantas son los promotores específicos de hoja tales como los descritos en el documento DE-A 19644478 o promotores regulados por luz tales como, por ejemplo, el promotor petE del guisante.

Otros promotores de plantas adecuados son el promotor de FBPasa citosólico o el promotor de ST-LSI de la patata Stockhaus y col., EMBO J. 8, 1989, 2445), el promotor de fosforribosil pirofosfato aminotransferasa de *Glycine max* (Nº de entrada GenBank U87999) o el promotor específico de nodo descrito en el documento EP-A-0 249 676.

De acuerdo con una característica preferida de la invención, el ácido nucleico AZ o variante del mismo como se define en el presente documento está unido operativamente a un promotor específico de semilla. El promotor específico de semilla puede ser activo durante el desarrollo de la semilla y/o durante la germinación. Los promotores específicos de semilla son bien conocidos en la técnica. Preferentemente, el promotor específico de semilla es un promotor específico de embrión/específico de endospermo/específico de aleurona. Preferentemente además, el promotor específico de semilla conduce la expresión en al menos uno de: embrión, endospermo y aleurona. Más preferentemente, el promotor es un promotor WSI18 o un promotor funcionalmente equivalente. Más preferentemente, la secuencia promotora es como se representa en la SEC ID Nº: 9 o SEC ID Nº: 55. Anteriormente se han proporcionado ejemplos de otros promotores específicos de semilla (promotores específicos de embrión/específicos de endospermo/específicos de aleurona) que también pueden usarse para dirigir la expresión de un ácido nucleico AZ.

De acuerdo con la memoria descriptiva, el ácido nucleico AZ o variante del mismo como se define en el presente documento está unido operativamente a un promotor constitutivo. Un promotor constitutivo es un promotor constitutivo que también se expresa sustancialmente de manera ubicua. Adicionalmente se prefiere que el promotor derive de una planta, más preferentemente de una planta monocotiledónea. Un ejemplo de dicho promotor es el promotor GOS2 del arroz (SEC ID Nº: 54 o SEC ID Nº: 56). Anteriormente se han ofrecido ejemplos de otros promotores constitutivos que también pueden usarse para dirigir la expresión de un ácido nucleico AZ como se define anteriormente.

Opcionalmente, en la construcción introducida en una planta pueden usarse 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 secuencias terminadoras y potenciadoras. 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 UTR 3' y/o UTR 5') pueden ser elementos estabilizadores de proteína y/o ARN. Un experto en la materia conocería dichas secuencias o podría obtenerlas fácilmente.

Las construcciones genéticas, como se describe en el presente documento, 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.

Para la detección de la transferencia satisfactoria de las secuencias de ácidos nucleicos, como se usa en los procedimientos como se describe en el presente documento y/o para la selección de plantas transgénicas que comprenden estos ácidos nucleicos, es ventajoso usar genes marcadores (o genes indicadores). Por lo tanto, la construcción genética puede comprender opcionalmente un gen marcador de selección. Los marcadores de selección se describen con más detalle en la sección de "definiciones" del presente documento. Los genes marcadores pueden retirarse o escindirse de la célula transgénica cuando ya no se necesiten. En la técnica se conocen técnicas para retirar marcadores, técnicas útiles como las descritas anteriormente en la sección de definiciones.

La presente memoria descriptiva también describe plantas, partes de planta o células de planta que pueden obtenerse mediante los procedimientos descritos en el presente documento. Por lo tanto, la presente memoria descriptiva describe plantas que pueden obtenerse mediante el procedimiento como se describe en el presente documento, cuyas plantas tienen en su interior un ácido nucleico AZ o una variante del mismo introducido, como se ha definido anteriormente.

La memoria descriptiva también describe un procedimiento para la producción de plantas transgénicas que tienen rendimiento aumentado, que comprende la introducción y expresión en una planta de un ácido nucleico AZ o una variante del mismo, como se ha definido anteriormente.

Las plantas huésped para los ácidos nucleicos o para el vector utilizado en el procedimiento como se describe en el presente documento, el casete de expresión o construcción o vector son, en principio, ventajosamente todas las plantas, que son capaces de sintetizar los polipéptidos utilizados en el procedimiento como se describe en el presente documento.

Más específicamente, la presente memoria descriptiva describe un procedimiento para la producción de plantas transgénicas que tienen rendimiento aumentado, cuyo procedimiento comprende:

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

El ácido nucleico puede introducirse directamente en una célula de planta o en la propia planta (incluyendo la introducción en un tejido, órgano o en cualquier otra parte de una planta). De acuerdo con una característica de la presente memoria descriptiva, el ácido nucleico se introduce en una planta por transformación. Los términos "introducción" o "transformación" se describen con más detalle en la sección definiciones.

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

Como se ha mencionado, también puede utilizarse *Agrobacteria* transformada con un vector de expresión como se describe en el presente documento, de una manera de por sí conocida, para la transformación de plantas tales como plantas utilizadas como un modelo, como *Arabidopsis*, o plantas de cultivo, tales como cereales, maíz, avena, centeno, cebada, trigo, soja, arroz, algodón, remolacha azucarera, canola, girasol, lino, cáñamo, patata, tabaco, tomate, zanahoria, pimientos dulce, colza, tapioca, yuca, arruruz, tagetes, alfalfa, lechuga y las diversas especies de árboles, nueces y vid, en particular plantas de cultivo que contienen aceite, tales como soja, cacahuete, ricino, girasol, maíz, algodón, lino, colza, coco, palma de aceite, cártamo (*Carthamus tinctorius*) o grano de cacao, por ejemplo introduciendo hojas laceradas o segmentos foliares en una solución agrobacteriana y posteriormente cultivándolos en medios adecuados.

Las células de planta modificadas genéticamente pueden regenerarse mediante cualquiera de los procedimientos que son familiares para el experto en la técnica. Procedimientos adecuados pueden encontrarse en las publicaciones anteriormente mencionadas de S.D. Kung y R. Wu, Potrykus o Höfgen y Willmitzer.

Generalmente, después de la transformación, las células o los grupos de células de planta se seleccionan para detectar la presencia de uno o más marcadores que están codificados por genes que pueden expresarse en plantas co-transfectados 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 semillas 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.

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.

5 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.

10 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).

15 La presente memoria descriptiva se describe para cualquier célula de planta o planta producida mediante cualquiera de los procedimientos descritos en el presente documento, y todas las partes de plantas y propágulos de las mismas. La presente memoria descriptiva se describe 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 genotípicas y/o fenotípicas que las producidas por el precursor en los procedimientos descritos en el presente documento. Se describen células huésped que contienen un ácido nucleico AZ aislado o variante del mismo. Las células huésped preferidas son células de plantas. También se describen partes cosechables de una planta, tales como, pero sin limitación, semillas, hojas, frutos, flores, tallos, raíces, rizomas, tubérculos y bulbos. 20 Adicionalmente se describen productos directamente derivados de una parte cosechable de dicha planta, tales como gránulos deshidratados o polvos, aceite, grasa y ácidos grasos, almidón o proteínas.

La presente memoria descriptiva describe el uso de ácidos nucleicos AZ o variantes de los mismos y el uso de polipéptidos AZ u homólogos de los mismos.

25 Un uso de este tipo se refiere a mejorar las características de cultivo de las plantas en particular la mejora del rendimiento, especialmente el rendimiento de semillas. El aumento de rendimiento de semillas comprende preferentemente el aumento del peso de mil granos.

30 Los ácidos nucleicos que codifican polipéptidos AZ como se define en el presente documento pueden encontrar uso en programas de reproducción en los que se identifica un marcador de ADN que puede estar unido genéticamente a un gen AZ. Los ácidos nucleicos/genes pueden usarse para definir un marcador molecular. Este marcador de ADN o de proteína puede después usarse en programas de reproducción para seleccionar plantas con mayor rendimiento.

35 Las variantes alélicas de un ácido nucleico/gen de AZ como se define en el presente documento también pueden encontrar uso en programas de reproducción asistida mediante marcador. Dichos programas de reproducción a veces requieren la introducción de variación alélica por tratamiento mutagénico de las plantas, usando, por ejemplo, mutagénesis con EMS; como alternativa, el programa puede comenzar con una colección de variantes alélicas de origen denominado "natural" producida involuntariamente. Después, la identificación de variantes alélicas se realiza, por ejemplo, por PCR. A esto le sigue una etapa de selección de variantes alélicas superiores de la secuencia en cuestión y que dará un rendimiento aumentado. La selección se realiza típicamente supervisando el rendimiento del crecimiento de las plantas que contienen diferentes variantes alélicas de la secuencia en cuestión, por ejemplo, 40 diferentes variantes alélicas de uno cualquiera de los ácidos nucleicos indicados en la Tabla A del Ejemplo 1. El rendimiento del crecimiento puede supervisarse en un invernadero o en el campo. Etapas opcionales adicionales incluyen cruzamiento de plantas, en el que se identifica la variante alélica superior, con otra planta. Esto podría usarse, por ejemplo, para realizar una combinación de características fenotípicas interesantes.

45 Un ácido nucleico AZ como se define en el presente documento también puede usarse como una sonda para mapear, genética y físicamente, genes que forman parte de, y como marcadores para rasgos ligados a estos genes. Dicha información puede ser útil en reproducción de plantas para desarrollar líneas con fenotipos deseados. Dicho uso de ácidos nucleicos AZ o variantes de los mismos requiere solo una secuencia de ácido nucleico de al menos 15 nucleótidos de longitud. Los ácidos nucleicos AZ o variantes de los mismos pueden usarse como marcadores de polimorfismos de longitud de fragmentos de restricción (PLFR). Con los ácidos nucleicos AZ pueden explorarse transferencias de Southern (Sambrook J, Fritsch EF y Maniatis T (1989) Molecular Cloning, A Laboratory Manual) de 50 ADN genómico de planta digerido por enzimas de restricción. Los patrones de banda resultantes pueden después someterse a análisis genético usando programas informáticos tales como MapMaker (Lander y col. (1987) Genomics 1: 174-181) para construir un mapa genético. Además, los ácidos nucleicos pueden usarse para explorar transferencias de Southern que contienen ADN genómicos tratados con endonucleasas de restricción de un conjunto de individuos que representan a los padres y a la progenie de un cruce genético definido. La segregación de 55 polimorfismos de ADN se observa y se usa para calcular la posición del ácido nucleico AZ en el mapa genético previamente obtenido usando esta población (Botstein y col. (1980) Am. J. Hum. Genet. 32: 314-331).

La producción y la utilización 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.

- 5 Las sondas de 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, págs. 319-346, y referencias citadas en su interior).

De acuerdo con la memoria descriptiva pueden utilizarse sondas de ácidos nucleicos en mapeo directo de hibridación *in situ* con fluorescencia (HISF) (Trask (1991) *Trends Genet.* 7: 149-154). Aunque los procedimientos actuales de mapeo HISF favorecen el uso de clones grandes (de varias 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.

Usando los ácidos nucleicos pueden realizarse diversos procedimientos basados en amplificación de ácidos nucleicos para el mapeo genético y físico. Como ejemplos se incluyen amplificación específica de alelo (Kazazian (1989) *J. Lab. Clin. Med.* 11: 95-96), polimorfismo de fragmentos amplificados por PCR (CAPS; Sheffield y col. (1993) *Genomics* 16: 325-332), ligamiento específico de alelo (Landegren y col. (1988) *Science* 241: 1077-1080), reacciones de extensión de nucleótidos (Sokolov (1990) *Nucleic Acid Res.* 18: 3671), Mapeo Híbrido por Radiación (Walter y col. (1997) *Nat. Genet.* 7: 22-28) y Mapeo Happy (Dear y Cook (1989) *Nucleic Acid Res.* 17: 6795-6807). Para estos procedimientos, se utiliza la secuencia de un ácido nucleico para diseñar y producir pares de cebadores para su uso en la reacción de amplificación o en reacciones de extensión con cebador. El diseño de dichos cebadores es bien conocido por los expertos en la técnica. En los procedimientos que emplean mapeo genético basado en PCR, puede ser necesario identificar diferencias de secuencias de ADN entre los padres del cruce del mapeo en la región correspondiente a la secuencia de ácido nucleico instantánea. Sin embargo, generalmente esto no es necesario para procedimientos de mapeo.

25 Los procedimientos de acuerdo con la presente memoria descriptiva dan como resultado plantas que tienen rendimiento aumentado, como se ha descrito anteriormente en el presente documento. Estas características de cultivo ventajosas también pueden combinarse con otros rasgos económicamente ventajosos, tales como rasgos que potencian adicionalmente el rendimiento, tolerancia a diversos estreses, rasgos que modifican diversas características arquitectónicas y/o características bioquímicas y/o fisiológicas.

30 **Descripción de las figuras**

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

35 La Figura 1 muestra un ejemplo de la estructura del dominio de un polipéptido AZ. La proteína codificada por la SEC ID N°: 2 comprende dos repeticiones de anquirina (en negrita, subrayadas) y dos dominios C3H1 (en cursiva, subrayados).

La Figura 2 muestra las secuencias consenso de dedo de Cinc de C3H1 (C3H1) y Anquirina (ANQ) de acuerdo con la base datos SMART, los símbolos de los diversos grupos de aminoácidos se indican en la leyenda.

40 La Figura 3 representa una matriz de identidad/similitud de secuencia preparada por MATGAT (BLOSUM62, penalización por apertura de hueco 11, penalización por extensión de hueco 1). Por encima de la diagonal, en negrita, se muestran las identidades de secuencia, por debajo de la diagonal se muestran las similitudes de secuencia para: A) secuencias de proteína de longitud completa, B) secuencias de proteína parcial que comprenden el supuesto dominio de dedo de Cinc que está más en el extremo C.

45 La Figura 4 muestra el vector binario p056, para la expresión en *Oryza sativa* de una secuencia codificante AZ de *Arabidopsis thaliana* bajo el control de un promotor WSI18 (referencia interna PRO0151).

Ejemplos

La presente invención se describirá ahora con referencia a los siguientes ejemplos.

50 Manipulación de ADN: a menos que se indique otra cosa, se realizaron 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 ácidos nucleicos usado en los procedimientos descritos en el presente documento

Las secuencias (ADNc de longitud completa, EST o genómicas) relacionadas con la secuencia de ácidos nucleicos utilizadas en los procedimientos como se describe en el presente documento se identificaron entre las conservadas en la base de datos de Nucleótidos Entrez en el National Center for Biotechnology Information (NCBI) utilizando las herramientas de búsqueda de secuencia de bases de datos, tal como la Herramienta de Alineamiento Local Básica (BLAST) (Altschul y col. (1990) J. Mol. Biol. 215:403-410; y Altschul y col. (1997) Nucleic Acids Res. 25: 3389-3402). El programa se utiliza para encontrar regiones de similitud local entre las secuencias comparando las secuencias de ácidos nucleicos o polipéptidos con las bases de datos de secuencia y calculando el significado estadístico de los emparejamientos. Por ejemplo, el polipéptido codificado por el ácido nucleico como se define en el presente documento se utilizó para el algoritmo TBLASTN, con configuraciones predeterminadas y el filtro para ignorar las secuencias de baja complejidad desencadenadas. El resultado de los análisis se observó por comparación en forma de pares, y se clasificó de acuerdo con la clasificación de probabilidad (valor E), en el que la clasificación refleja la probabilidad de que se produzca un alineamiento particular por casualidad (cuanto menor sea el valor E, más significativo el acierto). Además de valores E, también se clasificaron comparaciones por porcentaje de entidad. Porcentaje de identidad se refiere al número de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de ácidos nucleicos (o polipéptido) comparadas sobre una longitud particular. En algunos casos, los parámetros predeterminados pueden ajustarse para modificar la exigencia de la búsqueda. Por ejemplo, el valor E puede aumentarse para mostrar menos emparejamientos exigentes. De esta manera, pueden identificarse emparejamientos cortos casi exactos.

La Tabla A proporciona una lista de secuencias de ácidos nucleicos relacionadas con la secuencia de ácidos nucleicos utilizada en los procedimientos como se describe en el presente documento.

Tabla A: Ejemplos de polipéptidos AZ:

Fuente de planta	Ácido nucleico SEC ID N°:	Proteína SEC ID N°: Proteína
<i>Arabidopsis thaliana</i>	1	2
<i>Eucalyptus grandis</i>	10	11
<i>Oryza sativa</i>	12	13
<i>Medicago truncatula</i>	14	15
<i>Oryza sativa</i>	16	17
<i>Arabidopsis thaliana</i>	18	9
<i>Oryza sativa</i>	20	21
<i>Arabidopsis thaliana</i>	22	23
<i>Arabidopsis thaliana</i>	24	25
<i>Eucalyptus grandis</i>	26	27
<i>Eucalyptus grandis</i>	28	29
<i>Glycine max</i>	30	31
<i>Eucalyptus grandis</i>	32	33
<i>Arabidopsis thaliana</i>	34	35
<i>Oryza sativa</i>	36	37
<i>Hordeum_vulgare</i>	38	39
<i>Pinus radiata</i>	40	41
<i>Pinus radiata</i>	42	43
<i>Glycine max</i>		44
<i>Glycine max</i>		45

(continuación)

Fuente de planta	Acido nucleico SEC ID N°:	Proteína SEC ID N°: Proteína
<i>Glycine max</i>		46
<i>Arabidopsis thaliana</i>		47
<i>Arabidopsis thaliana</i>		48
<i>Arabidopsis thaliana</i>		49
<i>Arabidopsis thaliana</i>		50
<i>Arabidopsis thaliana</i>		51
<i>Arabidopsis thaliana</i>		52
<i>Arabidopsis thaliana</i>	53	

En algunos casos, las secuencias relacionadas se han ensamblado provisionalmente y descrito públicamente por instituciones de investigación, tal como el Instituto para la Investigación Genómica (TIGR). La base de datos de Ortólogos de Gen Eucariota (EGO) puede utilizarse para identificar dichas secuencias relacionadas, mediante búsqueda de palabra clave o utilizando el algoritmo BLAST con la secuencia de ácidos nucleicos o polipéptidos de interés.

Ejemplo 2: Clonación de genes AZ

El gen que codifica AZ de *Arabidopsis* (CDS3104) se amplificó por PCR usando como molde una biblioteca de ADNc de plántulas de *Arabidopsis thaliana* (Invitrogen, Paisley, RU). Después de transcripción inversa de ARN extraído de las plántulas, los ADNc se clonaron en pCMV Sport 6.0. El tamaño medio del inserto del banco fue de 1,5 kb, y el número original de clones fue de $1,59 \times 10^7$ ufc. El título original se determinó que era de $9,6 \times 10^5$ ufc/ml, después de una primera amplificación de 6×10^{11} ufc/ml. Después de la extracción del plásmido, se utilizaron 200 ng de molde en una mezcla PCR de 50 μ l. Se utilizaron los cebadores ceb06717 (sitio AttB1, directo, en cursiva, codón de inicio en negrita: 5'-**ggggacaagttgtacaaaaagcaggcttaacaat**gtgctgtggatcagacc-3') (SEC ID N°: 7) y ceb06718 (sitio AttB2, inverso, complementario, en cursiva: 5'-185 **ggggaccacttgtacaagaaagctgggtggttagtctctcaattctgc**-3') (SEC ID N°: 8), que incluían los sitios AttB para la recombinación Gateway, para la amplificación por PCR. La PCR se realizó usando Taq ADN polimerasa Hifi en condiciones convencionales. Se amplificó un fragmento de PCR del tamaño esperado y se purificó también utilizando procedimientos convencionales. Después, se realizó la primera etapa del procedimiento Gateway, la reacción BP, durante la cual el fragmento de PCR se combina *in vivo* con el plásmido pDONR201 para producir, de acuerdo con la terminología Gateway, un "clon de entrada", p07. El plásmido pDONR201 se adquirió en Invitrogen, como parte de la tecnología Gateway®.

Ejemplo 3: Construcción de vectores

El clon de entrada p07 se usó posteriormente en una reacción LR con p02417, un vector destinatario utilizado para transformación de plantas (*Oryza sativa*). Este vector contiene, como elementos funcionales dentro de los límites de ADN-T: un marcador de selección de plantas; un casete de expresión marcador detectable; y un casete Gateway diseñado para recombinación LR *in vivo* con la secuencia de interés ya clonada en el clon de entrada. Un promotor WSI18 de arroz (SEC ID N°: 9) para expresión específica de semilla (PRO0151) se localizaba cadena arriba de este casete Gateway (p056, Figura 4).

Se han descrito muchos sistemas vectoriales binarios diferentes (y súper binarios) para la transformación de plantas (por ejemplo, An, G. en *Agrobacterium* Protocols. Methods in Molecular Biology vol 44, págs. 47-62, Gartland KMA y MR Davey eds. Humana Press, Totowa, New Jersey). Muchos se basan en el vector pBIN19 descrito por Bevan (Nucleic Acid Research. 1984. 12: 8711-8721) que incluye un casete de expresión de genes de plantas flanqueado por las secuencias límite izquierda y derecha del plásmido Ti de *Agrobacterium tumefaciens*. Un casete de expresión de genes de plantas consiste en al menos dos genes – un gen marcador de selección y un promotor de planta que regula la transcripción del ADNc o ADN genómico del gen atributo. Pueden utilizarse diversos genes marcadores de selección incluyendo el gen de *Arabidopsis* que codifica una enzima ácido acetohidroxisintasa (AHAS) (Patentes de Estados Unidos 57673666 y 6225105). De manera similar, pueden utilizarse varios promotores para regular el gen atributo para proporcionar regulación constitutiva, evolutiva, tisular o ambiental de la transcripción génica.

Después de la etapa de recombinación LR, el vector de expresión resultante, p056 (Figura 4) se transformó en la cepa de *Agrobacterium* LBA4044 usando protocolos de choque térmico o electroporación. Las colonias transformadas se cultivaron en medio YEP y se seleccionaron mediante antibióticos respectivos durante dos días a 28 °C. Estos cultivos de *Agrobacterium* se utilizaron para la transformación de plantas.

Pueden usarse otras cepas de *Agrobacterium tumefaciens* para la transformación de plantas y son bien conocidas en la técnica. Son ejemplos de dichas cepas C58C1 o EHA105.

Ejemplo 4. Transformación de plantas

Transformación de arroz

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

15 La cepa LBA4404 de *Agrobacterium* que contenía el vector de expresión se utilizó para el co-cultivo. Se inoculó *Agrobacterium* en medio AB con los antibióticos apropiados y se cultivó durante 3 días a 28 °C. Después las bacterias se recogieron y se suspendieron en medio de co-cultivo 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 a un papel de filtro y se transfirieron a medio co-cultivo solidificado y se incubaron durante 3 días en la oscuridad a 25 °C. Los callos co-cultivados 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
25 suelo. Los brotes endurecidos se cultivaron a alta humedad y días cortos en un invernadero.

Se generaron aproximadamente 35 transformantes de arroz T0 independientes para una construcción. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero. Después de un análisis PCR cuantitativo para verificar el número de copias del inserto de ADN-T, solo una única copia de plantas transgénicas que presentaban tolerancia al agente de selección se mantuvieron para cosecha de la semilla T1.
30 Después las semillas se cosecharon de tres a cinco meses después del trasplante. El procedimiento produjo transformantes de sitio único a una tasa de alrededor de 50 % (Aldemita 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 son genotipos específicos susceptibles a transformación y regeneración. Como progenitores, la línea endogámica A188 (Universidad de Minnesota) o híbridos con A188, son buenas fuentes de material donante para la transformación, pero también pueden usarse otros genotipos satisfactoriamente. Se cosecharon espigas de plantas de maíz aproximadamente 11 días después de la polinización (DDP) cuando la longitud del embrión inmaduro era de
40 aproximadamente 1 a 1,2 mm. Los embriones inmaduros se co-cultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través organogénesis. Los embriones cortados se cultivaron en medio de inducción de callo y después en medio de regeneración de maíz que contenía el agente de selección (por ejemplo imidazolinona, pero pueden usarse diversos marcadores de selección). Las placas de Petri se incubaron a la luz a 25 °C durante 2-3 semanas, o hasta que se desarrollasen los brotes. Los brotes
45 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. La variedad de cultivo Bobwhite (disponible de CIMMYT, México) se utiliza normalmente para transformación. Embriones inmaduros se co-cultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través de organogénesis. Después de la incubación con *Agrobacterium*, los embriones se cultivaron *in vitro* en medio de inducción de callo, después el medio de
55 regeneración que contenía 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 las plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

Transformación de soja

5 La transformación de soja se realizó de acuerdo con una modificación del procedimiento descrito en el Texas A&M patente de Estados Unidos 5.164.310. Diversas variedades comerciales de soja se encuentran son susceptibles a transformación mediante este procedimiento. La variedad de cultivo Jack (disponible de la fundación Illinois Seed) se utiliza normalmente para transformación. Semillas de soja se esterilizaron para 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 restante 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 de tratamiento de co-cultivo, los explantes se lavaron y se transfirieron al medio 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 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 colza/canola

Peciolos cotiledonarios e hipocotiledóneos de plántulas jóvenes de 5-6 días de vida se utilizaron como explantes para el cultivo tisular y transformaron de acuerdo con Babic y col. (1998, Plant Cell Rep 17: 183-188). La variedad de cultivo comercial Westar (Agriculture Canada) es la variedad convencional usada para la transformación, pero 20 también pueden usarse otras variedades. Semillas de canola se esterilizaron en la superficie para la siembra *in vitro*. Explantes de peciolos cotiledonarios con el cotiledón unido se extirparon de 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 co-cultivo con *Agrobacterium*, 25 los explantes de peciolo 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 cefotaxime, 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 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 de 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, éstas pueden seleccionarse de variedades de cultivo Rangelander (Agriculture Canada) o de cualquier otra variedad comercial de alfalfa como describen Brown DCW y A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Como alternativa, la variedad RA3 (Universidad de Wisconsin) se ha seleccionado para su uso en cultivo tisular (Walker y col., 1978 Am J Bot 65: 654-659). Explantes de peciolo se co-cultivaron con un 40 cultivo durante una noche de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie y col., 1999 Plant Physiol 119: 839-847) o LBA4404 que contiene el vector de expresión. Los explantes se co-cultivaron durante 3 d 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 las plantas que presentaban tolerancia al agente de selección y que contenían una sola copia 50 de inserto de ADN-T.

Ejemplo 5: Configuración de evaluación de expresión de AZ en arroz bajo el control del promotor de arroz WSI18

Se generaron aproximadamente de 15 a 20 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 55 T1. Se conservaron siete acontecimientos, de los que la progenie T1 segregó 3:1 para la presencia/ausencia del transgén. Para cada uno de estos acontecimientos, aproximadamente 10 plántulas T1 que contenían el transgén (hetero- y homo-zigotos) y aproximadamente 10 plántulas T1 que carecían del transgén (nulicigotos) se seleccionaron supervisando la expresión visual del marcador. Las plantas T1 seleccionadas se transfirieron a un invernadero. Cada planta recibió una sola etiqueta de código de barras para relacionar con claridad los datos de

fenotipado con la planta correspondiente. Las plantas T1 seleccionadas se cultivaron en suelo en macetas especialmente diseñadas de un diámetro de 10 cm con fondo transparente para permitir la visualización de las raíces, con la siguiente configuración ambiental: fotoperiodo = 11,5 h, intensidad de luz diurna = 30.000 lux o más, temperatura diurna = 28 °C, temperatura nocturna = 22 °C, humedad relativa = 60-70 %. Las plantas transgénicas y los correspondientes nulicigotos se cultivaron a cada lado en posiciones aleatorias. Se tomaron medidas para que las plantas no tuviesen ningún estrés. 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 pixeles, 16 millones de colores) de cada planta desde al menos 6 ángulos diferentes. Una cámara digital también registró imágenes desde el fondo de la maceta durante el crecimiento de la planta.

10 *Detección de sequía*

Se cultivaron plantas de semillas T2 en suelo de maceta en condiciones normales hasta que alcanzaron la etapa de partida. Después se transfirieron a una sección "seca" donde se mantiene el riego. Sondeas de humedad se insertaron en macetas seleccionadas al azar para supervisar el contenido de agua en el suelo (CAS). Cuando el CAS está por debajo de determinados umbrales, las plantas vuelven a regarse automáticamente de manera continua hasta que de nuevo se alcanza 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 para las plantas que no se cultivan en condiciones de estrés abiótico. Los parámetros de cultivo y rendimiento se registraron como se detalla para el cultivo en condiciones normales.

Detección de la eficiencia del uso de nitrógeno

20 Se cultivaron plantas de arroz de semillas T2 en suelo de maceta en condiciones normales excepto para la solución nutrientes. Las macetas se regaron desde el trasplante hasta la maduración con una solución de nutrientes específica que contenía un contenido de nitrógeno N reducido (N), normalmente entre 7 a 8 veces menos. El resto del cultivo (maduración de planta, cosecha de semilla) es igual para las plantas que no se cultivan bajo estrés abiótico. Los parámetros de cultivo y rendimiento se registraron como se detalla para el cultivo en condiciones normales.

El área por encima de la superficie de la planta (o biomasa foliar) se determinó contando el número total de pixeles en las imágenes digitales a partir de las partes de la planta por encima de la superficie discriminadas del fondo. Este valor se promedió para las imágenes tomadas en el mismo momento de los diferentes ángulos y se convirtió a un valor de superficie físico expresado en mm al cuadrado por calibración. Los experimentos mostraron que el área de la planta por encima de la superficie medida de esta forma se correlacionaba con la biomasa de las partes de la planta por encima de la superficie. El área por encima de la superficie es el momento en el que la planta ha alcanzado su máxima biomasa foliar. Las características radiculares tales como el área total proyectada (que puede correlacionarse con volumen radicular total), diámetro promedio y longitud de las raíces por encima de un determinado umbral de espesor (longitud de raíces gruesas, o longitud de raíces finas) se dedujo a partir de la imagen generada usando un programa informático apropiado.

Las panículas primarias maduras se cosecharon, se embolsaron, se etiquetaron con códigos de barra y después se secaron durante tres días en el horno a 37 °C. Después las panículas se trillaron y se recogieron todas las semillas. Las cáscaras llenas se separaron de las vacías usando un dispositivo de soplado con aire. Después de la separación, ambos lotes de semillas se contaron usando una contadora disponible en el comercio. Las cáscaras vacías se descartaron. Las cáscaras llenas se pesaron en una balanza analítica y el área transversal de las semillas se midió usando formación de imágenes digitales. Este procedimiento da lugar al conjunto de los siguientes parámetros relacionados con las semillas:

El número de semillas llenas se determinó contando el número de cáscaras llenas que quedó después de la etapa de separación. El rendimiento de semilla total (peso de semilla total) se midió pesando todas las cáscaras llenas cosechadas de una planta. El número total de semillas por planta se midió contando el número de cáscaras cosechadas de una planta. El índice de cosecha (IC) en la presente invención se define como la proporción entre el rendimiento de semilla total y el área por encima de la superficie (mm²), multiplicado por un factor de 10⁶. El Peso de Mil Granos (PMG) se extrapoló a partir del número de semillas llenas contadas y su peso total. El índice de carga de semilla, como se define en la presente invención, es la proporción (expresada como un %) del número de semillas llenas sobre el número total de semillas (o flósculos). Estos parámetros derivaron de una forma automática de las imágenes digitales usando un programa informático de análisis de formación de imágenes y se analizaron estadísticamente. Los parámetros de semilla individuales (incluyendo anchura, longitud, área, peso) se midieron usando un dispositivo fabricado a medida que consistía en dos componentes principales, un dispositivo de pesaje y de formación de imágenes, acoplado a un programa informático para análisis de imágenes.

Se utilizó un ANOVA (análisis de varianza) de dos factores corregido para el diseño desequilibrado 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 en todos los acontecimientos transformados con ese gen. El ensayo F se realizó para verificar un efecto del gen sobre todos acontecimientos de transformación y para verificar un efecto

global del gen, denominado también en el presente documento “efecto génico global”. Si el valor del ensayo F muestra que los datos son significativos, es decir se llega a la conclusión de que es un efecto “génico”, significa que no solo la presencia o la posición del gen ocasionan el efecto. El umbral de significado de un efecto génico global verdadero se establece al 5 % del nivel de probabilidad para el ensayo F.

- 5 Para verificar un efecto sobre los genes dentro de un acontecimiento, es decir, un efecto específico de línea, se realizó un ensayo T dentro de cada acontecimiento utilizando conjuntos de datos de las plantas transgénicas y de plantas nulas correspondientes. “Plantas nulas” o “segregantes nulos” o “nulicigotos” son las plantas tratadas de la misma manera que la planta transgénica, pero a partir de las cuales se ha segregado el transgén. Las plantas nulas también pueden describirse como plantas transformadas negativas homocigotas. El umbral de significado del ensayo T se establece a un nivel de probabilidad del 10 %. Los resultados de algunos acontecimientos pueden estar por encima o por debajo de este umbral. Esto se basa en la hipótesis de que un gen podría tener solo un efecto en determinadas posiciones en el genoma y de que la aparición de este efecto dependiente de posición no es atípico. Este tipo de efecto génico también se denomina en el presente documento un “efecto de línea del gen”. El valor p se obtiene comparando el valor t con la distribución t o como alternativa, comparando el valor F con la distribución F. El valor p entonces proporciona la probabilidad de que la hipótesis nula (es decir, que no haya efecto del transgén) sea correcta.

- 20 Los datos obtenidos para AZ en el primer experimento se confirmaron en un segundo experimento con plantas T2. Se seleccionaron cuatro líneas que tenían el patrón de expresión correcto para análisis posterior. Se exploraron lotes de semillas de las plantas positivas (tanto hetero como homocigotas) en T1, verificando la expresión de marcadores. Para cada acontecimiento seleccionado, los lotes de semillas heterocigotas se conservaron después para evaluación T2. Dentro de cada lote de semillas se cultivó el mismo número de plantas positivas y negativas en el invernadero para la evaluación.

Se evaluó un número total de 120 plantas AZ transformadas en la generación T2, es decir 30 plantas por acontecimiento de las cuales 15 fueron positivas del transgén y 15 negativas.

- 25 Dado que se habían realizado dos experimentos con acontecimientos solapantes, se realizó un análisis combinado. Esto es útil para verificar la coherencia de los efectos sobre los dos experimentos, y si fuera éste el caso, acumular pruebas de ambos experimentos para aumentar la confianza en la conclusión. El procedimiento utilizado fue una estrategia de modelo mixta que considera la estructura multinivel de los datos (es decir, experimento – acontecimiento – segregantes). Los valores P se obtuvieron comparando el ensayo de relación de probabilidad con distribuciones de chi cuadrado.

Ejemplo 6: Evaluación de transformantes AZ: medición de parámetros relacionados con el rendimiento

- 35 Después del análisis de las semillas, como se ha descrito anteriormente, los autores de la invención descubrieron que las plantas transformadas con la construcción génica AZ tenían una mayor rendimiento de semillas, expresada como peso de mil granos, en comparación con plantas que carecían del transgén AZ. Además, se observó vigor de emergencia aumentado e índice de verdor aumentado en plantas que llevaban el transgén en comparación con las plantas de control.

- 40 En una de las construcciones, en la generación T1 aumentó el peso de mil granos 2,7 %. Estos resultados positivos se obtuvieron de nuevo en la generación T2 (aumento del 2,1 %). Los datos T2 volvieron a evaluarse en un análisis combinado con los resultados de la generación T1, y los valores de p obtenidos mostraron que los efectos observados eran muy significativos.

LISTADO DE SECUENCIAS

- 45 <110> CropDesign N.V.
 <120> Plantas que tienen rendimiento aumentado y procedimiento para producirlas
 <130> PF58401-prio
 50 <160> 347
 <170> PatentIn versión 3.3
 55 <210> 1
 <211> 2151
 <212> ADN
 <213> *Arabidopsis thaliana*

ES 2 451 669 T3

<400> 1

atgtgctgtg	gatcagaccg	attaaaccag	atcgtgtcat	caagatcttc	gttgccaatt	60
tctttcgagg	aagataacaa	tcttgttacc	aacacagaca	tgaatcactt	aacagtcgaa	120
acagaggata	cgtttgcgag	cttgcttgag	cttgacagcta	acaacgatgt	tgaaggtgta	180
aggctatcta	tcgagagaga	cccttcttgt	gtagacgaag	ctggtctctg	gtacggtcgt	240
caaaaagggt	ctaaagctat	ggtcaacgat	tacaggactc	cgttgatggt	tgctgctact	300
tacggaagca	ttgatgtgat	caagcttatt	gtttctttga	ctgatgctga	cgtgaaccgt	360
gcttgcgga	atgatcagac	cactgcgtta	cactgcgctg	cttctggagg	agctgtgaat	420
gctatccaag	ttgttaagct	gcttcttgca	gctggagctg	atttgaatct	gttggatgct	480
gaaggtcaac	gagctggtga	tgttattggt	gttcctccta	agcttgaagg	cgtgaagctg	540
atgcttcagg	agcttctttc	tgctgatgga	tcatctactg	cggagcggaa	tctacggggt	600
gtgacaaatg	ttccgaatag	aagctcatct	ccgtgtcatt	ctcctactgg	agagaatggg	660
ggatcagggg	ctgggtcacc	gctcggctct	ccttttaagc	tgaaatctac	tgaattcaag	720
aaagagtatc	cggttgatcc	gtctttgcca	gatatcaaga	acagtatcta	cgcgactgat	780
gagtttagaa	tgtattcctt	caaggtccgg	ccttgctctc	gtgcttattc	acatgattgg	840
actgagtgtc	cttttgttca	cccgggtgaa	aacgcgagga	ggagagaccc	gaggaagttc	900
cattacagct	gcgttccttg	cccggattht	aggaaggag	cttgtaggag	aggagatatg	960
tgtgagtatg	cgcaacggtg	gtttgaatgc	tggttctatc	cggctcagta	caggaccctg	1020
ctttgcaaag	atggaacagg	ctgtgctcgg	cgggtttggt	tctttgcgca	tacacccgag	1080
gagcttcgac	ctttgtaogc	atcaactggt	tcagcgggtc	cttcgcctag	atcgaatgct	1140
gattatgcag	ctgctttgag	tctccttctt	ggttctccat	caggagtctc	tgtcatgtcc	1200
ccgctttccc	catcagcagc	ggggaacgga	atgtctcatt	cgaatatggc	ttggccacaa	1260
ccaaatgtcc	ctgcgttgca	cttaccagga	agcaatctac	agtcaagcag	gctaaggtct	1320
tctctcaatg	caagggatat	cccgacggat	gagttcaata	tgtagcggga	ttacgagcag	1380
cagcaactcc	tcaacgagta	ttccaatgct	ctgagccggt	ctggtcggat	gaaatcaatg	1440
cctccttcga	atcttgaaga	tcttttctca	gcagaaggct	cttcatctcc	ccggttcaact	1500
gattccgctt	tagcttccgc	ggtgttctcg	cctacacaca	agtcagctgt	cttcaaccag	1560
ttccaacaac	agcaacagca	gcagcagagc	atggtgtctc	caatcaacac	aagcttttct	1620
tcaccaaaaga	gcggtgacca	ctcattgttt	tcaggtggag	gaagaatgtc	tcctcggaat	1680
gttggtgaac	caatatcacc	catgagtgtc	cgggtttcca	tgtaggctca	gtgctggaag	1740
caacaacaac	agcaacagca	gcagcagcag	cagcaacatc	agttccgtag	ccttagctcc	1800
agagagctca	gaacaaactc	tagcccaatc	gttggttcac	cggtaacaa	caacacatgg	1860
tcatcaaaaat	ggggatcttc	aaatggtcaa	ccgattggg	gaatgagctc	agaagcactt	1920
ggtaagttga	gatcttcgtc	atcgtttgat	ggtgatgagc	ctgatgtgtc	atgggtccag	1980
tactggtga	aggagactcc	agcagaagcc	aaagagaaag	cagcaacatc	ttcctcaggg	2040
gaacacgtga	tgaagcagcc	aaatccggtt	gaaccggtaa	tggatcatgc	tgggctagaa	2100
gcttggtattg	agcaaatgca	gctcagatcag	cttgtggctc	agcagaattg	a	2151

5 <210> 2
 <211> 716
 <212> PRT
 <213> *Arabidopsis thaliana*

10 <400> 2

ES 2 451 669 T3

Met Cys Cys Gly Ser Asp Arg Leu Asn Gln Ile Val Ser Ser Arg Ser
1 5 10 15
Ser Leu Pro Ile Ser Phe Glu Glu Asp Asn Asn Leu Val Thr Asn Thr
20 25 30
Asp Met Asn His Leu Thr Val Glu Thr Glu Asp Thr Phe Ala Ser Leu
35 40 45
Leu Glu Leu Ala Ala Asn Asn Asp Val Glu Gly Val Arg Leu Ser Ile
50 55 60
Glu Arg Asp Pro Ser Cys Val Asp Glu Ala Gly Leu Trp Tyr Gly Arg
65 70 75 80
Gln Lys Gly Ser Lys Ala Met Val Asn Asp Tyr Arg Thr Pro Leu Met
85 90 95
Val Ala Ala Thr Tyr Gly Ser Ile Asp Val Ile Lys Leu Ile Val Ser
100 105 110
Leu Thr Asp Ala Asp Val Asn Arg Ala Cys Gly Asn Asp Gln Thr Thr
115 120 125
Ala Leu His Cys Ala Ala Ser Gly Gly Ala Val Asn Ala Ile Gln Val
130 135 140
Val Lys Leu Leu Leu Ala Ala Gly Ala Asp Leu Asn Leu Leu Asp Ala
145 150 155 160
Glu Gly Gln Arg Ala Gly Asp Val Ile Val Val Pro Pro Lys Leu Glu
165 170 175
Gly Val Lys Leu Met Leu Gln Glu Leu Leu Ser Ala Asp Gly Ser Ser
180 185 190
Thr Ala Glu Arg Asn Leu Arg Val Val Thr Asn Val Pro Asn Arg Ser
195 200 205
Ser Ser Pro Cys His Ser Pro Thr Gly Glu Asn Gly Gly Ser Gly Ser
210 215 220
Gly Ser Pro Leu Gly Ser Pro Phe Lys Leu Lys Ser Thr Glu Phe Lys
225 230 235 240
Lys Glu Tyr Pro Val Asp Pro Ser Leu Pro Asp Ile Lys Asn Ser Ile
245 250 255
Tyr Ala Thr Asp Glu Phe Arg Met Tyr Ser Phe Lys Val Arg Pro Cys
260 265 270
Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro
275 280 285
Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys Phe His Tyr Ser Cys
290 295 300
Val Pro Cys Pro Asp Phe Arg Lys Gly Ala Cys Arg Arg Gly Asp Met
305 310 315 320
Cys Glu Tyr Ala His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln
325 330 335
Tyr Arg Thr Arg Leu Cys Lys Asp Gly Thr Gly Cys Ala Arg Arg Val
340 345 350
Cys Phe Phe Ala His Thr Pro Glu Glu Leu Arg Pro Leu Tyr Ala Ser
355 360 365
Thr Gly Ser Ala Val Pro Ser Pro Arg Ser Asn Ala Asp Tyr Ala Ala
370 375 380
Ala Leu Ser Leu Leu Pro Gly Ser Pro Ser Gly Val Ser Val Met Ser
385 390 395 400
Pro Leu Ser Pro Ser Ala Ala Gly Asn Gly Met Ser His Ser Asn Met
405 410 415
Ala Trp Pro Gln Pro Asn Val Pro Ala Leu His Leu Pro Gly Ser Asn
420 425 430
Leu Gln Ser Ser Arg Leu Arg Ser Ser Leu Asn Ala Arg Asp Ile Pro
435 440 445
Thr Asp Glu Phe Asn Met Leu Ala Asp Tyr Glu Gln Gln Gln Leu Leu

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

210> 3
 <211> 13
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> motivo 1

10

<220>
 <221> VARIANTE
 <222> (1) .. (1)
 <223> \reemplazo= "Ala"

15

<220>
 <221> VARIANTE
 <222> (7) .. (7)
 <223> \reemplazo= "Thr"

20

<400> 3

Pro Cys Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys
 1 5 10

25

<210> 4
 <211> 12
 <212> PRT
 <213> Secuencia artificial

ES 2 451 669 T3

```

<220>
<223> motivo 2

<400> 4
5
      His Pro Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg
      1          5          10

<210> 5
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> motivo 3
15

<220>
<221> VARIANTE
<222> (3)..(3)
<223> \reemplazo= "Ile"
20

<220>
<221> VARIANTE
<222> (6)..(6)
<223> \reemplazo= "Ser"
25

<220>
<221> VARIANTE
<222> (11)..(11)
<223> \reemplazo= "Ser"
30

<220>
<221> VARIANTE
<222> (14)..(14)
<223> \reemplazo= "Lys"
35

<400> 5

      His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln Tyr Arg Thr Arg
      1          5          10          15
      Leu Cys Lys

<210> 6
<211> 5
<212> PRT
<213> Secuencia artificial

<220>
<223> motivo 4
45

<400> 6

      Cys Phe Phe Ala His
      1          5

<210> 7
<211> 54
<212> ADN
<213> Secuencia artificial
55

<220>
<223> cebador: prm06717

<400> 7
60

```

ES 2 451 669 T3

ggggacaagt ttgtacaaaa aagcaggctt aaacaatgtg ctgtggatca gacc 54

5 <210> 8
 <211> 50
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> cebador: prm06718

<400> 8
 ggggaccact ttgtacaaga aagctgggtg gttaggtctc tcaattctgc 50

15 <210> 9
 <211> 981
 <212> ADN
 <213> *Oryza sativa*

20 <400> 9

aagaggcaag	agcatccgta	ttaaccagcc	ttttgagact	tgagagtgtg	tgtgactcga	60
tccagcgtag	tttcagttcg	tgtgttggtg	agtgattcca	gccaagtttg	cgatggcttc	120
tcagcaggaa	cgggctagct	accacgccgg	cgagaccaag	gcccgcgccg	aggagaagac	180
ggggcgcatg	atgggcacgg	cgcaggagaa	ggcgcgggag	gccaaggaca	cggcgtccga	240
cgcccgggg	cgcgcgatgg	gcaggggaca	cggcgccaag	gaggcgacca	aggagaaggc	300
gtacgagacc	aaggacgcga	ccaaggagaa	ggcgtacgag	gcaaaggacg	cggcctccga	360
cgccaccggc	cgcgccatgg	acaagggccg	cggcgcgccg	ggcgccacga	gggacaaggc	420
gtacgatgcc	aaggacaggg	cggctgacac	ggcgcagtcc	gccgcccagc	gcgcccgcga	480
cggcgcggg	cagaccggga	gctacattgg	acagaccgcc	gaggccgcca	agcagaaagc	540
ggccggcggc	gcgcagtagc	ccaaggagac	cgcgatcgcc	ggcaaggaca	agaccggcgc	600
cgtgctccag	caggcagggg	agcaggtgaa	gagcgtggcg	gtgggggcca	aggacgcggt	660
gatgtacacg	ctcgggatgt	caggcgataa	caagaacaac	gccgctgccg	gcaaggacac	720
cagcacctac	aagcctggaa	ctgggagtga	ctaccagtaa	tacggtagaa	gaagcatgtg	780
tcgtctttgg	caactgatgcc	aaagtgtacg	tgttgtatcc	tcttttttaa	gtttcagctc	840
gacttcgacg	tgttcgggtg	cacactttgg	tttttcagtt	gtgctcaact	gttcatgttt	900
ctggttccat	ggagggccag	tgtggaggtc	aatgtttaag	ctttcgtttt	aaaatctgat	960
aataaagttg	gttaagacct	g				981

25 <210> 10
 <211> 3372
 <212> ADN
 <213> *Eucalyptus grandis*

<400> 10

ES 2 451 669 T3

ggaagacgaa	gagcaacaaa	ataggtctca	ctccctctc	tctctctct	cctctctctt	60
ctttctctct	ctttctgctc	taacaaagtc	tcttctctga	gagacagggc	tgcgtcgtcg	120
tctttctctc	tcctcgtctc	gagcttctga	gaaagttcaa	ttctttttct	cttgttctct	180
ctctctaccc	ttctgggtac	cactgtgaag	ctccggtctt	ttctatTTTT	TTTTTTTTtg	240
ggctatctgg	gtctgggcaa	atccatcgcg	cgctctgctc	tggactgaga	ggcgcgcagt	300
ggctttagat	ctgcgacgcc	tcttgcttgc	tcagtgaagc	gggctagttc	aaatcgacga	360
agaaagcatg	cgctagtgat	tgggtgtggg	aatacactgc	attcgatctc	tactaagtat	420
ccccaaagtat	actaagatcc	cgttctcagc	catgagtcaa	ctgaccattc	agactgagga	480
cacttttgcc	agcttgcttg	agcttgetgc	taacaacgac	acagaatctt	tcggacgggtg	540
tgtggaacgt	gatccttcga	gcatagatga	aattggatat	tgggatggtc	gccccaaagg	600
ttcgaagcag	gtgggtcaata	tgcaaaagac	tcctcttatg	gtggctgcta	catatggtag	660

tgttgatgta	atgagactca	ttctttgcct	atctgatgct	gatgtgaatc	gaacctgcag	720
cacagacaag	agcacagccc	ttcactgtgc	tgectctggt	ggtgctgtga	atgctgtaga	780
tgctgtgagg	ctactcctgt	cagctgggtc	tgaccaagt	ttagcagatg	ctaaccggtca	840
gcggcctgtg	gatgttattg	ttgttctctc	aaagctcctt	tcaataaagt	ttgctcttga	900
agagctcttg	tcgaccgaag	gatctgtaaa	tgaacacaat	ctgagagtg	ccgtagccac	960
ttccaattca	acctctcccc	cactttcctc	ttccccggat	aatgggtccc	cagcatctgc	1020
taattgttct	ttccccaaaga	actcaaagtt	aagtgatgcc	cctgttcttt	atgcatcaga	1080
aaagaaggaa	taccgggtgg	atccatctct	tccagatctc	aagaatagca	tttactcaac	1140
agatgaattc	cgaatgtatt	cttttaaagt	gcggccttgt	tcacgagcgt	actcgcatag	1200
ttggacggag	tgcccttttg	ttcatccagg	ggagaatgcc	cgtagaagg	atccaaggaa	1260
gttccactac	agctgtgtcc	cttgccctga	tttccggaag	ggtgcttgta	gacgtggaga	1320
tatgtgtgaa	tatgctcatg	gtgtttttga	gtgctggctc	catcctgctc	agtatcggac	1380
tcgattatgc	aaggatggta	caagttgtgc	tcggagagtg	tgcttctttg	cccacacgga	1440
gcaagagctg	cgtccattgt	acgtctccac	tggttctgct	gttccgctc	ctcgtcggag	1500
tacctctgga	gctgctgcca	tggattttgc	tcgagccatg	agcctcttac	ctggttcccc	1560
atcatcagta	tccatcatgt	ccccctcacc	cttcaactct	cccattgtct	catctgctaa	1620
ttgtatttct	cacccatctg	ttgcctggcc	ccagcaaaat	gtaccaactt	tgcactctcc	1680
cggaagcaat	cttcagtcca	gccgcttgag	atcttctctt	aatgcaagag	atattcctca	1740
ggaggatttt	gacttgctgt	cagattatga	tgtgcaacag	cagcagctcc	taaataagtt	1800
ttccatcctt	tcacaacaat	cgatgggtgc	taattccttg	aaccgttctg	gtcggctgaa	1860
aactttgacc	ccctcaaacc	ttgatgatct	cttctctgct	gagagctcat	cccctcgtca	1920
cgctgatcaa	gccctggctt	ctgctgtttt	ttcaccaacg	cacaaatctg	cagtaatcaa	1980
tcaatttcag	cagcagcagc	agagcatggt	atcacccatc	aacacaacct	tctctcctaa	2040
gagtgtcgag	caccctttgt	tgcaagcgtc	ttctcgtggt	caatctgggc	gaaatgtccc	2100
tcgtaaacatg	gatcccatct	ctcctataag	ttctcgtgtg	tcgatgttgg	ccccacgaga	2160
gaaacagcaa	cagcaattac	gcagcctaag	ctctcgtgaa	ctcggttcca	attcagccgc	2220
cattgtgggt	ttccccgtgg	gttcttggtc	gaaatgggga	gctacaaatg	ggaaaccaga	2280
ctgggctggt	agtgcagatg	aactaggtaa	gcttcgcagg	tctaattcat	ttgagcttgg	2340
gaacaatggt	gaggagccag	atctttcatg	ggttcaatcc	ctcgttaaag	aatctcctac	2400
cgagatgaaa	gaaaagcttt	cgtcaactct	ctctggtggt	ccagcccccg	ctacatccag	2460
tgaggttccg	agtatcagct	cgcagatgga	atcggttgat	cacgaagtgc	taggagcatg	2520
gctccagcag	atgcagctcg	atccgctcgt	ggctcagcaa	aactaggttg	ttttttttcc	2580
tacatggcct	tgaggaagta	gacagcggaa	agtttttttt	ggtaataact	atgttttttc	2640
tggaaatttt	tgatgctggg	ggtggggtct	ggaagaagat	aacaaggcag	gaaaggggtc	2700
agtgaagtca	ctggagaaaa	ggaattcatt	tttaaccatt	ttatcattct	attacaacag	2760
aaagtaggga	aaaaaaagga	agaccctctg	ggttatgaag	agaaattaa	cccaggctag	2820
gcgttctcct	ttctaataat	tccaatttta	ggtccatatt	actgtcattt	cctttttgct	2880
gtcttatcat	atttcatcaa	aatggaactg	gggactaatg	tttgttccat	tctttcgtct	2940
ttctgattta	tttgcaccct	tggggttaaga	tcaaaagaga	aattatgatc	attttctttt	3000
gaggatattt	ttttttccca	atattttgta	gaatgaaagt	taagagggga	tatgatgtgt	3060
ctggtgtttg	agtatgaaa	accaataacc	gagttcaact	gttgctgctg	gtggtagaag	3120
aagtggagaa	gaagctatga	tcctttgatg	taacagtcaa	tcaaacattt	taatccttt	3180
attttttggt	tcctcatgta	atccatcctt	tgtgattgtc	ctctctctct	ctctctctct	3240
ctctctctcc	ctccccgtgt	tctttcttca	taagcgtctt	gcttgtcgat	ctgtaaatata	3300
ttgaaaagg	tcatggaaag	ccgtgccggt	gtggattctc	atttttgcaa	aaaaaaaaaa	3360
aaaaaaaaaa	aa					3372

<210> 11
 <211> 704
 <212> PRT
 <213> *Eucalyptus grandis*

ES 2 451 669 T3

<400> 11

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

ES 2 451 669 T3

Lys Gly Ser Lys Gln Val Val Asn Met Gln Arg Thr Pro Leu Met Val
 50 55 60
 Ala Ala Thr Tyr Gly Ser Val Asp Val Met Arg Leu Ile Leu Cys Leu
 65 70 75 80
 Ser Asp Ala Asp Val Asn Arg Thr Cys Ser Thr Asp Lys Ser Thr Ala
 85 90 95
 Leu His Cys Ala Ala Ser Gly Gly Ala Val Asn Ala Val Asp Ala Val
 100 105 110
 Arg Leu Leu Leu Ser Ala Gly Ala Asp Pro Ser Leu Ala Asp Ala Asn
 115 120 125
 Gly Gln Arg Pro Val Asp Val Ile Val Val Pro Pro Lys Leu Leu Ser
 130 135 140
 Ile Lys Phe Ala Leu Glu Glu Leu Leu Ser Thr Glu Gly Ser Val Asn
 145 150 155 160
 Glu His Asn Leu Arg Val Ser Val Ala Thr Ser Asn Ser Thr Ser Pro
 165 170 175
 Pro Leu Ser Ser Ser Pro Asp Asn Gly Ser Pro Ala Ser Ala Asn Cys
 180 185 190
 Ser Ser Pro Lys Asn Ser Lys Leu Ser Asp Ala Pro Val Leu Tyr Ala
 195 200 205
 Ser Glu Lys Lys Glu Tyr Pro Val Asp Pro Ser Leu Pro Asp Ile Lys
 210 215 220
 Asn Ser Ile Tyr Ser Thr Asp Glu Phe Arg Met Tyr Ser Phe Lys Val
 225 230 235 240
 Arg Pro Cys Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe
 245 250 255
 Val His Pro Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys Phe His
 260 265 270
 Tyr Ser Cys Val Pro Cys Pro Asp Phe Arg Lys Gly Ala Cys Arg Arg
 275 280 285
 Gly Asp Met Cys Glu Tyr Ala His Gly Val Phe Glu Cys Trp Leu His
 290 295 300
 Pro Ala Gln Tyr Arg Thr Arg Leu Cys Lys Asp Gly Thr Ser Cys Ala
 305 310 315 320
 Arg Arg Val Cys Phe Phe Ala His Thr Glu Gln Glu Leu Arg Pro Leu
 325 330 335
 Tyr Val Ser Thr Gly Ser Ala Val Pro Ser Pro Arg Ser Ser Thr Ser
 340 345 350
 Gly Ala Ala Ala Met Asp Phe Ala Ala Ala Met Ser Leu Leu Pro Gly
 355 360 365
 Ser Pro Ser Ser Val Ser Ile Met Ser Pro Ser Pro Phe Thr Pro Pro
 370 375 380
 Met Ser Pro Ser Ala Asn Gly Ile Ser His Pro Ser Val Ala Trp Pro
 385 390 395 400
 Gln Gln Asn Val Pro Thr Leu His Leu Pro Gly Ser Asn Leu Gln Ser
 405 410 415
 Ser Arg Leu Arg Ser Ser Leu Asn Ala Arg Asp Ile Pro Gln Glu Asp
 420 425 430
 Phe Asp Leu Leu Ser Asp Tyr Asp Val Gln Gln Gln Gln Leu Leu Asn
 435 440 445
 Glu Phe Ser Ile Leu Ser Gln Gln Ser Met Gly Ala Asn Ser Leu Asn
 450 455 460
 Arg Ser Gly Arg Leu Lys Thr Leu Thr Pro Ser Asn Leu Asp Asp Leu
 465 470 475 480
 Phe Ser Ala Glu Ser Ser Ser Pro Arg Tyr Ala Asp Gln Ala Leu Ala
 485 490 495
 Ser Ala Val Phe Ser Pro Thr His Lys Ser Ala Val Ile Asn Gln Phe
 500 505 510
 Gln Gln Gln Gln Gln Ser Met Leu Ser Pro Ile Asn Thr Thr Phe Ser

ES 2 451 669 T3

		515					520					525			
Pro	Lys	Ser	Val	Asp	His	Pro	Leu	Leu	Gln	Ala	Ser	Phe	Gly	Val	Gln
	530					535					540				
Ser	Gly	Arg	Met	Ser	Pro	Arg	Asn	Met	Asp	Pro	Ile	Ser	Pro	Ile	Ser
545					550					555					560
Ser	Arg	Val	Ser	Met	Leu	Ala	Gln	Arg	Glu	Lys	Gln	Gln	Gln	Gln	Leu
				565					570					575	
Arg	Ser	Leu	Ser	Ser	Arg	Glu	Leu	Gly	Ser	Asn	Ser	Ala	Ala	Ile	Val
				580				585					590		
Gly	Ser	Pro	Val	Gly	Ser	Trp	Ser	Lys	Trp	Gly	Ala	Thr	Asn	Gly	Lys
		595					600					605			
Pro	Asp	Trp	Ala	Val	Ser	Ala	Asp	Glu	Leu	Gly	Lys	Leu	Arg	Arg	Ser
	610					615					620				
Asn	Ser	Phe	Glu	Leu	Gly	Asn	Asn	Gly	Glu	Glu	Pro	Asp	Leu	Ser	Trp
625					630					635					640
Val	Gln	Ser	Leu	Val	Lys	Glu	Ser	Pro	Thr	Glu	Met	Lys	Glu	Lys	Leu
				645					650					655	
Ser	Ser	Thr	Leu	Ser	Gly	Val	Pro	Ala	Pro	Ala	Thr	Ser	Ser	Glu	Val
				660				665						670	
Pro	Ser	Ile	Ser	Ser	Gln	Met	Glu	Ser	Val	Asp	His	Glu	Val	Leu	Gly
		675					680					685			
Ala	Trp	Leu	Gln	Gln	Met	Gln	Leu	Asp	Pro	Leu	Val	Ala	Gln	Gln	Asn
	690					695					700				

<210> 12
 <211> 1860
 <212> ADN
 <213> *Oryza sativa*

 <400> 12

5

ES 2 451 669 T3

atgggggagc	ctgggggagc	cgaggcgcc	gtctccgca	ggctgctcga	gctggcgcc	60
gacgacaacg	cggcgggct	cgggagctc	ctcgcggct	ggccctccct	cgccgacgag	120
cccgcgccgt	ggtacacccc	ggcgcgggc	gcgagcgc	tgaccccgct	catggctgcc	180
gccgtgtaag	gctcgggtgg	ctgcctcgac	gcgctcctct	cgccgcccta	cctcgtggac	240
cccaaccgag	cctcgggtgc	gtcgtctctc	accccgctcc	acctcgcgc	cgcgggcggg	300
tccgcctccg	ccccgcggc	ggtctccgc	ctcctcgcgc	ccggcgccga	cccggccctc	360
ctcgaccacc	tccagcgccg	ggcgtccgac	ctcgtcgcgc	tcccgcccaa	ctcgtcccg	420
ctcaagaacc	acctcctctc	cctcctcggc	gcccgcgaag	agtggcctcc	cgacccctcc	480
ctccccgaca	tcaagaacgg	cgctacgcc	tccgacgact	tcaggatgta	ctcgttcaag	540
gtgcgcgct	gctcgcgggc	ctactccat	gactggacgg	agtggccctt	cgtccacccc	600
ggcgagaacg	cgcgcgccg	cgacccgag	aagtaccact	acagctgcgt	gocgtgccc	660
gagttcaaga	agggggccg	gtgcagaga	ggggacatgt	gcgagtacgc	gcacggggtg	720
ttcgagagct	ggctccaccc	ggcgagtac	cggacgcgc	tctgcaagga	cgcgctcggc	780
tgcgcccgc	gcgtctgctt	cttcgcccac	acgcccgcgc	agctccgccc	gctctacgtc	840
tccacgggct	ccgcccgtgc	gtcgcgcgc	ggggcggttg	agatggcggc	ggcgggcgcg	900
gcgatgggga	tgggctgtc	gtcgcgggg	togtctgct	tcacgcccgc	gctatcgcgc	960
tggccggcg	ggggcgggg	cggggcggg	ggcagcgcg	gcggcgggc	gtggccgcag	1020
cagccgagcg	tgccggcgct	ctgcctgccc	gggagcgccg	ggaacctcca	cctgagccgg	1080
ctgcgcacgt	cgctgagcgc	gcgcgacatg	gcccgtcgc	agctgctcgc	cgcgggcgcg	1140
gcggcgcgcg	actacgacgg	cctcgtcgc	tccccgcct	ccatccggtc	cgcgaggggg	1200
aaggcgcttg	tgccgtcaaa	tctcagcag	ctctctccg	ctgagctcgc	cgccgcccgc	1260
gcgtcgcgt	cgccgcgcta	cgccgaccaa	ggcgcgccg	cgttctcccc	gacccgcaag	1320
gccaccgtgc	tcaaccaatt	ccagctgcag	cagcagcata	gcttgctctc	gcggggcgcg	1380
gccgcggtga	caccagagcc	ggtctcccca	atgagctccc	gcctcctcgc	cgcgctggcg	1440
cagcgggaga	agatgcagca	gcagacgctg	cggagcatga	gctcacggga	cctcggcaac	1500
gccgcgtcgc	tgctggtcgg	ctcgcgggtg	agctcagca	tgtccaaatg	ggggttcccc	1560
tccggcaacc	cggactgggg	cgccgacgac	gaggagctcg	gcccctcaa	gcggtgctcc	1620
tcgttcgagc	tccggtcggg	agccgccaat	ggcaaccatg	agcctgacct	ctcatgggtc	1680
aacaccctag	tgaaggagcc	gacaccggag	aagatgatga	cgacgacatc	ggcaatggat	1740
tccattggca	tcttgggaca	gaacacaagc	cgtgatcaca	tcgtcggagg	cgaggatgac	1800
actgccggag	tcacagcag	ctggcttgaa	cagctccagc	tcgatgagat	ggttgctctag	1860

<210> 13
 <211> 619
 <212> PRT
 <213> *Oryza sativa*

<400> 13

ES 2 451 669 T3

Met Gly Glu Pro Gly Gly Ala Glu Ala Ala Val Ser Ala Arg Leu Leu
1 5 10 15
Glu Leu Ala Ala Asp Asp Asn Ala Ala Gly Leu Gly Glu Leu Leu Ala
20 25 30
Ala Trp Pro Ser Leu Ala Asp Glu Pro Ala Pro Trp Tyr Thr Pro Ala
35 40 45
Arg Gly Ala Glu Pro Leu Thr Pro Leu Met Val Ala Ala Val Tyr Gly
50 55 60
Ser Val Gly Cys Leu Asp Ala Leu Leu Ser Pro Pro Tyr Leu Val Asp
65 70 75 80
Pro Asn Arg Ala Ser Ala Ser Ser Leu Ser Thr Pro Leu His Leu Ala
85 90 95
Ala Ala Gly Gly Ser Ala Ser Ala Pro Ala Ala Val Ser Arg Leu Leu
100 105 110
Ala Ala Gly Ala Asp Pro Ala Leu Leu Asp His Leu Gln Arg Arg Ala
115 120 125
Ser Asp Leu Val Ala Leu Pro Pro Asn Ser Leu Pro Leu Lys Asn His
130 135 140
Leu Leu Ser Leu Leu Gly Ala Arg Lys Glu Trp Pro Pro Asp Pro Ser
145 150 155 160
Leu Pro Asp Ile Lys Asn Gly Ala Tyr Ala Ser Asp Asp Phe Arg Met
165 170 175
Tyr Ser Phe Lys Val Arg Ala Cys Ser Arg Ala Tyr Ser His Asp Trp
180 185 190
Thr Glu Cys Pro Phe Val His Pro Gly Glu Asn Ala Arg Arg Arg Asp
195 200 205
Pro Arg Lys Tyr His Tyr Ser Cys Val Pro Cys Pro Glu Phe Lys Lys
210 215 220
Gly Ala Gly Cys Arg Arg Gly Asp Met Cys Glu Tyr Ala His Gly Val
225 230 235 240
Phe Glu Ser Trp Leu His Pro Ala Gln Tyr Arg Thr Arg Leu Cys Lys
245 250 255
Asp Gly Val Gly Cys Ala Arg Arg Val Cys Phe Phe Ala His Thr Pro
260 265 270
Asp Glu Leu Arg Pro Leu Tyr Val Ser Thr Gly Ser Ala Val Pro Ser
275 280 285
Pro Arg Gly Ala Leu Glu Met Ala Ala Ala Ala Ala Met Gly Met
290 295 300
Gly Leu Ser Ser Pro Gly Ser Ser Ser Phe Thr Pro Pro Leu Ser Pro
305 310 315 320
Ser Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly
325 330 335
Ala Trp Pro Gln Gln Pro Ser Val Pro Ala Leu Cys Leu Pro Gly Ser
340 345 350
Ala Gly Asn Leu His Leu Ser Arg Leu Arg Thr Ser Leu Ser Ala Arg
355 360 365
Asp Met Ala Val Asp Glu Leu Leu Ala Ala Ala Ala Ala Ala Asp
370 375 380
Tyr Asp Gly Leu Val Ala Ser Pro Ala Ser Ile Arg Ser Ala Arg Gly
385 390 395 400

ES 2 451 669 T3

Lys Ala Leu Val Pro Ser Asn Leu Asp Glu Leu Phe Ser Ala Glu Leu
 405 410 415
 Ala Ala Ala Ala Ala Ser Arg Ser Pro Arg Tyr Ala Asp Gln Gly Gly
 420 425 430
 Ala Ala Phe Ser Pro Thr Arg Lys Ala Thr Val Leu Asn Gln Phe Gln
 435 440 445
 Leu Gln Gln Gln His Ser Leu Leu Ser Pro Arg Ala Ala Ala Val Thr
 450 455 460
 Pro Glu Pro Val Ser Pro Met Ser Ser Arg Leu Leu Ala Ala Leu Ala
 465 470 475 480
 Gln Arg Glu Lys Met Gln Gln Gln Thr Leu Arg Ser Met Ser Ser Arg
 485 490 495
 Asp Leu Gly Asn Ala Ala Ser Leu Leu Val Gly Ser Pro Val Ser Ser
 500 505 510
 Ser Met Ser Lys Trp Gly Phe Pro Ser Gly Asn Pro Asp Trp Gly Ala
 515 520 525
 Asp Asp Glu Glu Leu Gly Arg Leu Lys Arg Cys Ser Ser Phe Glu Leu
 530 535 540
 Arg Ser Gly Ala Ala Asn Gly Asn His Glu Pro Asp Leu Ser Trp Val
 545 550 555 560
 Asn Thr Leu Val Lys Glu Pro Thr Pro Glu Lys Met Met Thr Thr Thr
 565 570 575
 Ser Ala Met Asp Ser Ile Gly Ile Leu Gly Gln Asn Thr Ser Arg Asp
 580 585 590
 His Ile Val Gly Gly Glu Asp Asp Thr Ala Gly Val Ile Ser Ser Trp
 595 600 605
 Leu Glu Gln Leu Gln Leu Asp Glu Met Val Val
 610 615

<210> 14
 <211> 2106
 <212> ADN
 <213> *Medicago truncatula*

5

<400> 14

ES 2 451 669 T3

atgaaaaatc	taactgttccg	tactgatgat	tctttttcca	gcttacttga	acatgcttct	60
aacaatgatt	ttgaagattt	caaggtagct	ctagatagtg	atgcttcact	tattaatgaa	120
gttggcttct	ggtatgtccg	tcaaaagga	tctaaccaaa	ttgttcttga	gcaccgaacc	180
cctttaatgg	tggctgcttc	ctatgggagt	attgatattc	taaagcttat	actctcatat	240
cccgaggctg	atgttaattt	ctcctgtgga	actgataaaa	gcactgctct	tcactgtgct	300
gcctcaagtg	gttcagttaa	tgctgttgat	gctataaaaat	tgcttttacc	agctgggtget	360
gatatcaatt	ctgtggatgc	taatgggaaa	cgccctgtgg	atgttatcgt	tgttccctatt	420
gttgttcctc	ataagctcga	agggtgttaa	acaattcttg	aagaacttct	ctcagacagt	480
gcttctgaag	gatctgtgga	tgattgctct	cttcccctgt	ctcttatttc	atcgagtcct	540
ggttcactctg	cccctttatc	atctgctgaa	aatggatctc	catcctctcc	tggtggctccc	600
aagtttacag	atacagctgt	taattctaca	tcagaaaaga	aagagtatcc	agttgaccca	660
tctcttcctg	acataaaaaa	cagcatgtat	gccacagatg	aattccgcat	gtattcattc	720
aaggttcgct	cttgttctcg	tgcatactct	catgattgga	ctgagtgtcc	ttttgtgcat	780
cctggagaga	atgctcgaag	gagagaccct	agaaagtttc	actacagctg	tggtccatgc	840
cctgatthta	ggaaaggggc	ttgccgacgt	tcggatatgt	gtgaatatgc	tcattggagta	900
ttcgagtgtc	ggctacaccc	agctcagtat	cggacaaggc	tggtcaaaga	cggtatgggt	960
tgtaaccgaa	gggtgtgctt	cttcgctcac	taacctgaag	agctgcgtcc	gctgtatgtg	1020
tcactgggtt	ctgctgttcc	ttcaccccga	tcagctgctt	ctactgctaa	tgctcatggac	1080
atggctgctg	ctatgagcct	tttcccctgg	tcaccatcat	caatctcttt	gatgtctcaa	1140
tcaccctttg	cacagcctcc	tctatctcca	tctgcaaagt	gcaataatgc	ttggccacag	1200
cccaatgtgc	cagctcttca	tttaccagga	agcattaatc	aaactagctg	tttgagatct	1260
tctcttagtg	cccgtgatat	gccacacgac	gacttcaaca	atatgttgca	agactttgat	1320
gggcagcagc	agatactaaa	tgacttgagc	tgtttctcac	agccccgtcc	tggtgctatt	1380
tcagttggtc	gatctggccg	ccctaaaaca	ctaactccct	caaatctgga	tgatcttttt	1440
tggtctgaga	ttgcttcatc	tcctaggtat	tccgaccccg	ctgcggttc	tgtatthtcc	1500
ccaacacaca	aatctgctgt	cttcaaccag	ttcaacagc	ttcaaagctc	cttatcacc	1560
atcaacacaa	atgtcatgtc	tcctacaaac	gtagagcatc	ccctgttcca	ccaggcttca	1620
tatggtctct	cttctcctgg	aaggatgtca	ccaagaagta	tggaagccct	atctccaatg	1680
agttctcggc	tgtcagcttt	tgctcagcgt	gagaaacaac	agcagcagca	gcaacagctg	1740
cgtagcctca	gctcaagaga	actcgggtgct	aacaatcctc	tctcagctgt	tggttcccct	1800
gttaactcct	ggtccaagtg	gggatcatcc	cctattggaa	aagctgattg	gtcggtaaat	1860
ccaaatgact	tcgggtcaa	acagagatca	acttcttttg	agcatggaaa	caatggagaa	1920
gagcctgatg	taggttgggt	ccattccctt	gtcaaggatc	ccacacctga	gaagaaagag	1980
aagcttgtag	gttccggccc	aattccatcc	gttgaaaaga	atcccaatcc	tcaagcggac	2040
ggcattgata	actctgtttt	gggagcttgg	ctcagagcaac	tgtagctgga	tcaacttgta	2100
gtctag						2106

<210> 15
 <211> 701
 <212> PRT
 <213> *Medicago truncatula*

<400> 15

ES 2 451 669 T3

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

Cys Asn Arg Arg Val Cys Phe Phe Ala His Ser Pro Glu Glu Leu Arg
 325 330 335
 Pro Leu Tyr Val Ser Thr Gly Ser Ala Val Pro Ser Pro Arg Ser Ala
 340 345 350
 Ala Ser Thr Ala Asn Val Met Asp Met Ala Ala Ala Met Ser Leu Phe
 355 360 365
 Pro Gly Ser Pro Ser Ser Ile Ser Leu Met Ser Gln Ser Pro Phe Ala
 370 375 380
 Gln Pro Pro Leu Ser Pro Ser Ala Asn Gly Asn Asn Ala Trp Pro Gln
 385 390 395 400
 Pro Asn Val Pro Ala Leu His Leu Pro Gly Ser Ile Asn Gln Thr Ser
 405 410 415
 Arg Leu Arg Ser Ser Leu Ser Ala Arg Asp Met Pro His Asp Asp Phe
 420 425 430
 Asn Asn Met Leu Gln Asp Phe Asp Gly Gln Gln Gln Ile Leu Asn Asp
 435 440 445
 Leu Ser Cys Phe Ser Gln Pro Arg Pro Gly Ala Ile Ser Val Gly Arg
 450 455 460
 Ser Gly Arg Pro Lys Thr Leu Thr Pro Ser Asn Leu Asp Asp Leu Phe
 465 470 475 480
 Cys Ala Glu Ile Ala Ser Ser Pro Arg Tyr Ser Asp Pro Ala Ala Ala
 485 490 495
 Ser Val Phe Ser Pro Thr His Lys Ser Ala Val Phe Asn Gln Phe Gln
 500 505 510
 Gln Leu Gln Ser Ser Leu Ser Pro Ile Asn Thr Asn Val Met Ser Pro
 515 520 525
 Thr Asn Val Glu His Pro Leu Phe His Gln Ala Ser Tyr Gly Leu Ser
 530 535 540
 Ser Pro Gly Arg Met Ser Pro Arg Ser Met Glu Ala Leu Ser Pro Met
 545 550 555 560
 Ser Ser Arg Leu Ser Ala Phe Ala Gln Arg Glu Lys Gln Gln Gln Gln
 565 570 575
 Gln Gln Gln Leu Arg Ser Leu Ser Ser Arg Glu Leu Gly Ala Asn Asn
 580 585 590
 Pro Leu Ser Ala Val Gly Ser Pro Val Asn Ser Trp Ser Lys Trp Gly
 595 600 605
 Ser Ser Pro Ile Gly Lys Ala Asp Trp Ser Val Asn Pro Asn Asp Phe
 610 615 620
 Gly Gln Thr Gln Arg Ser Thr Ser Phe Glu His Gly Asn Asn Gly Glu
 625 630 635 640
 Glu Pro Asp Val Gly Trp Val His Ser Leu Val Lys Asp Pro Thr Pro
 645 650 655
 Glu Lys Lys Glu Lys Leu Ala Gly Ser Gly Pro Ile Pro Ser Val Glu
 660 665 670
 Lys Asn Pro Asn Pro Gln Ala Asp Gly Ile Asp His Ser Val Leu Gly
 675 680 685
 Ala Trp Leu Glu Gln Leu Gln Leu Asp Gln Leu Val Val
 690 695 700

<210> 16
 <211> 2841
 <212> ADN
 <213> *Oryza sativa*
 <400> 16

ES 2 451 669 T3

atgaacggca	cgccgatctc	cgcgtccgcc	gcgcccgccg	tgcacggagt	cgccgcggcg	60
gtggcgctgg	cggccgcgac	caagaagagt	gcccgcgcgg	cggccgcggt	cgccgagatg	120
gcgaaaacc	tcaccgtcga	cacggacgac	gccttcgcgg	ggctcctcga	gctcgcgcgg	180
gacgacgacg	cggagggcct	gcgccgcgcg	ctggagcgcg	ccccgcccgc	cgccgcggac	240
gaggcgggcc	tctggtacgg	cgcgcgcaag	gtcctcagac	accgcacgcc	gctgatggtc	300
gcggccacct	atggcagcct	cgcggtgctt	cgctgctgc	tgtccctccc	gtccgtcgat	360
gtcaatcgcc	gctgtggctc	cgacggcacc	accgccctcc	actgtgcggc	gtctggtggc	420
tcgccgtctt	gtgtggaggc	cgtcaagctg	ctgcttgcgt	ctggggctga	tgctgatgcc	480
acggatgctt	ccgatatcgc	tccagctgat	gtgatctctg	ttcctccaaa	gatgtttgac	540
gccaagattg	ccctccaaga	tcttcttggg	tgcccaaagg	ctgggcatgg	cgttctccgg	600
gtggtgacaa	ggccgcgcaa	ctctatggtg	tcacctgat	catcccctac	agcagaagat	660
gcacgatctc	catcagctgc	tgtgatgatg	acgacaaaagt	ttgcagatct	tccaaggggt	720
gtgacatcgg	aaaagaaaga	atatccagtg	gatccgtccc	ttcccgatat	caagaacagc	780
atctatgctt	ccgatgagtt	ccgcatgtac	tcatthaaga	tcaggccatg	ctcgcggggc	840
tactcacatg	attggactga	gtgcccgttt	gttcacccag	gggagaacgc	acggcgtcgg	900
gaccctcgca	agtatcacta	cagctgtgtg	ccatgccccg	actttagaaa	gggagtttgc	960
cggcgtggtg	acatgtgtga	atatgctcat	ggcgtgttcg	agtgttggct	ccatccagca	1020
cagtaccgta	ctcgcctttg	caaggatggc	acaagctgta	atcgcctgtg	ctgtttcttt	1080
gcgcatacaa	ctgatgagct	ccgaccacta	tatgtttcca	ctggatctgc	agtaccatcc	1140
ccaagagcct	cggcaacagc	tacaatggag	atggctgcag	caatgggctt	gatgcctggg	1200
tctccatcat	cagtttcagc	agtcatgtcc	ccatttacac	caccaatgtc	cccttcaggc	1260
aatgggatgc	ccccttcatt	gggctggcag	cagccaaatg	ttccgacact	acaccttcca	1320
ggcagcagcc	ttcagtcgag	ccggtccgtg	acctcactta	gtgcaagggg	tatgcctgct	1380
gatgattact	ccctgatgca	ggatattgat	tcacagctta	taaatgattt	gtgcctttca	1440
cgtattgggt	catcaacagg	aaaccacacg	tctcggacca	agtccctaaa	tcctgcaaac	1500
ttggatgata	tcttctctgc	tgagatggtc	tcttccccga	ggatatagtaa	tgctgatcag	1560
ggtggtatgt	tttaccatc	tcacaaggct	gctttcctta	atcagttcca	gcaacagcag	1620
caggcacttc	tttaccaat	caacacagtc	ttctccccga	agtctgtgga	caaccagcag	1680
ttgccttcac	actcatctct	gttgcaagca	tcacttggtg	tatcctcccc	tggccgcgatg	1740
tctcctcgat	gtgttgaatc	tgggtcccct	atgaactctc	atcttgctgc	tgctcttgct	1800
cagcgtgaga	agcaacagca	gacaatgaga	agtctcagtt	ctcgtgatct	tgggcccaggt	1860
gctgcaagag	catcaggtgt	tgttggctcc	cctctaagct	catcatggtc	aaagtgggga	1920
tcaccttcag	ggacacctga	ctgggggtgt	aatggtgaag	aattgggcaa	gcttcgcccg	1980
tcacatcagt	ttgagctgag	atctggtggt	gatgatccag	atctctcttg	ggtacacaca	2040
ctggttaagg	aatctccacc	agagaagcaa	gtcactactg	atgaatccat	aaactctggt	2100
ggaccctcac	cactgatgcc	tcccagtgta	agcaacggtg	aaggctcctag	tctgaatgcc	2160
ccgctggatg	ggcatgacca	agctgctggt	attggagcat	tgcttgaaca	gatgcagcct	2220
gatcagcata	ttggtagtct	agcaacataa	gcgctgaatg	agcctggaaa	gtgcaaggag	2280
ttattattct	tagttaatga	atttggagta	atTTTTTtcc	tgttcattaa	gatggctcagc	2340
aagcaaaagg	atggatagct	gatggtgggtg	attcagagat	tggTTTTctt	tactttattg	2400
aggtaaatca	tatacattat	tgaggttcca	gtaggttgaa	agattgaagt	accttgattg	2460
gggtcgtttc	aagaccgacc	caggtagaat	cgcacccccg	cagcttcaat	tcacgggtca	2520
aaaatatctt	cctgttttgt	taattaaccc	cgttaaaaaa	gaagactcgt	ttggtgtttc	2580
ggaattcttt	tctttacctt	agcgggtggtt	atTTTgttta	ttatgatatt	gatacttgat	2640
gtactgatgg	gtataagggt	ggttaccagc	catgctatag	tggtatatca	agtcccaaag	2700
tattcttttt	ctccctttca	ccatttgtcg	aggatcatac	tatggccttg	ttttggtcag	2760
atcttgaggc	ctgtataatc	cttggatttg	taataatgta	atattgtcat	tgaacttaca	2820
ttgctattgt	tttgcaatcg	c				2841

<210> 17
 <211> 749
 <212> PRT
 <213> *Oryza sativa*

<400> 17

ES 2 451 669 T3

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

ES 2 451 669 T3

Glu Gly Leu Arg Arg Ala Leu Glu Arg Ala Pro Pro Ala Ala Ala Asp
 65 70 75 80
 Glu Ala Gly Leu Trp Tyr Gly Arg Arg Lys Val Leu Glu His Arg Thr
 85 90 95
 Pro Leu Met Val Ala Ala Thr Tyr Gly Ser Leu Ala Val Leu Arg Leu
 100 105 110
 Leu Leu Ser Leu Pro Ser Val Asp Val Asn Arg Arg Cys Gly Ser Asp
 115 120 125
 Gly Thr Thr Ala Leu His Cys Ala Ala Ser Gly Gly Ser Pro Ser Cys
 130 135 140
 Val Glu Ala Val Lys Leu Leu Leu Ala Ala Gly Ala Asp Ala Asp Ala
 145 150 155 160
 Thr Asp Ala Ser Gly Tyr Arg Pro Ala Asp Val Ile Ser Val Pro Pro
 165 170 175
 Lys Met Phe Asp Ala Lys Ile Ala Leu Gln Asp Leu Leu Gly Cys Pro
 180 185 190
 Lys Ala Gly His Gly Val Leu Arg Val Val Thr Arg Ala Ala Asn Ser
 195 200 205
 Met Leu Ser Pro Val Ser Ser Pro Thr Ala Glu Asp Ala Arg Ser Pro
 210 215 220
 Ser Ala Ala Val Met Met Thr Thr Lys Phe Ala Asp Leu Pro Arg Val
 225 230 235 240
 Val Thr Ser Glu Lys Lys Glu Tyr Pro Val Asp Pro Ser Leu Pro Asp
 245 250 255
 Ile Lys Asn Ser Ile Tyr Ala Ser Asp Glu Phe Arg Met Tyr Ser Phe
 260 265 270
 Lys Ile Arg Pro Cys Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys
 275 280 285
 Pro Phe Val His Pro Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys
 290 295 300
 Tyr His Tyr Ser Cys Val Pro Cys Pro Asp Phe Arg Lys Gly Val Cys
 305 310 315 320
 Arg Arg Gly Asp Met Cys Glu Tyr Ala His Gly Val Phe Glu Cys Trp
 325 330 335
 Leu His Pro Ala Gln Tyr Arg Thr Arg Leu Cys Lys Asp Gly Thr Ser
 340 345 350
 Cys Asn Arg Arg Val Cys Phe Phe Ala His Thr Thr Asp Glu Leu Arg
 355 360 365
 Pro Leu Tyr Val Ser Thr Gly Ser Ala Val Pro Ser Pro Arg Ala Ser
 370 375 380
 Ala Thr Ala Thr Met Glu Met Ala Ala Ala Met Gly Leu Met Pro Gly
 385 390 395 400
 Ser Pro Ser Ser Val Ser Ala Val Met Ser Pro Phe Thr Pro Pro Met
 405 410 415
 Ser Pro Ser Gly Asn Gly Met Pro Pro Ser Leu Gly Trp Gln Gln Pro
 420 425 430
 Asn Val Pro Thr Leu His Leu Pro Gly Ser Ser Leu Gln Ser Ser Arg
 435 440 445
 Leu Arg Thr Ser Leu Ser Ala Arg Asp Met Pro Ala Asp Asp Tyr Ser
 450 455 460
 Leu Met Gln Asp Ile Asp Ser Gln Leu Ile Asn Asp Leu Cys Tyr Ser
 465 470 475 480
 Arg Ile Gly Ser Ser Thr Gly Asn His Thr Ser Arg Thr Lys Ser Leu
 485 490 495
 Asn Pro Ser Asn Leu Asp Asp Leu Phe Ser Ala Glu Met Val Ser Ser
 500 505 510
 Pro Arg Tyr Ser Asn Ala Asp Gln Gly Gly Met Phe Ser Pro Ser His
 515 520 525
 Lys Ala Ala Phe Leu Asn Gln Phe Gln Gln Gln Gln Gln Ala Leu Leu

ES 2 451 669 T3

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

<210> 18
 <211> 2769
 <212> ADN
 <213> *Arabidopsis thaliana*

5

<400> 18

ES 2 451 669 T3

acaccagtta	ccctctcatc	cgttttcggt	ttttttttct	ctctttcaaa	aatctctcag	60
ctgagggtga	tcgatcttct	tcttcttctt	cctcactctt	tagatttggt	ccttttgcat	120
tttacacttt	tggatctgaa	aatgtgggtc	tctgtttcgc	ccgttacogt	ttagattcag	180
ttctgttttt	ttcttaccog	atcgcttgat	tcggactgtg	atctttgatc	ttttttcttc	240
tccagtgcog	tgaagatgt	gtggctctgc	taagaagctg	gatatagagg	atactttgac	300
atcactgtca	gaccaagaga	atgaatcttt	ggccaaacct	atgaatgatg	ctgctgaatg	360
ggaacattcg	ttttctgcct	tgcttgagtt	tgctgcagac	aacgatgtgg	aggggttag	420
gcggaactc	tctgatgtgt	cttgtatcaa	ccagatgggt	ctttggtaca	gacggcagag	480
gtttgttaga	agaatgggtc	ttgagcaaag	aaccccgctg	atggttgctt	cgttatatgg	540
gagtttagat	gltgtgaagt	ttattcttct	tttcccggaa	gctgagttga	atctgtcttg	600
tggtcctgat	aaaagtactg	ctcttcattg	cgctgcttct	ggtgcttctg	tgaattcctt	660
ggatgttgtc	aagttgcttt	tgagtgtagg	agcagatcct	aatatccctg	atgctcatgg	720
aaatcgctct	gttgatgttc	ttgttggtgc	tccacacgct	cctgggttga	gaaccatcct	780
tgaagagatc	ttgaagaaag	acgagattat	atctgaagat	ctgcatgcct	cgctcatctag	840
cttgggatca	agttccgggt	ctctctcatc	atcccctgat	aatgggtcct	cgttactctc	900
cttagattca	gtatcctctc	cgactaagcc	acacggtaact	gatgtaactt	tcgcatcaga	960
gaagaaagag	taccaaatg	atccatcatt	gcctgatatc	aaaagcggga	tttattcaac	1020
cgatgagttt	cgtatgttct	cgttcaagat	ccgccatgt	tctcgagcat	attcccatga	1080
ctggactgaa	tgtccatttg	cacaccagg	tgagaatgca	aggagaagag	acccgaggaa	1140
gtttcactat	acgtgtgttc	catgcccogga	ttttaagaaa	ggatcctgta	agcaaggtga	1200
tatgtgtgaa	tatgctcatg	gggtttttga	atgctggcta	caccctgctc	agtacagaac	1260
acgattgtgc	aaggacggaa	tgggttgcaa	ccgaagggtt	tgcttctttg	ctcacgcaaa	1320
tgaggagttg	cgcccttctg	acccttccac	aggatctgga	ttgccatctc	ctcgggcttc	1380
gtctgctgtt	tccgctctca	ctatggacat	ggcgtcagtt	ttgaacatgt	taccaggctc	1440
accatctgct	gctcaacatt	cgttcaccct	accaatatct	ccttctggaa	atggtagtat	1500

gccccattca	tcgatgggtt	ggcctcagca	gaacataccg	gcgttgaatc	ttcttggaag	1560
caatatccag	ttgagtcgtc	tgagatcttc	tcttaacogt	agagatattc	cttctgagca	1620
gctttagatg	ctgcatgagt	ttgaaatgca	acgtcagctt	gctggcgata	tgcacagtcc	1680
acgctttatg	aatcattccg	ctcgtcctaa	gacactgaac	ccttcaaatc	tggaggaact	1740
cttctcagct	gaggttgcat	ctcctcgttt	ctctgatcaa	cttgctgttt	catctgttct	1800
atcgccttcc	cacaagtccg	cgcttcttaa	tcagctgcag	aataataagc	agagcatgct	1860
ttctctatc	aagacaaatc	taatgtcttc	tccaaagaat	gtggagcaac	attctcttct	1920
gcagcaagcc	tcgtcaccct	gaggcggaga	gcctatttcc	ccaatgaatg	ctcgaatgaa	1980
acagcagcta	cattcacgca	gcctaagctc	ccgtgatctt	ggatctagtc	tgccccgtga	2040
tttaatgccc	actgattctg	gttcgccatt	aagtcctatg	tcaagttggg	accagacca	2100
tggaaagcaag	gtggattggt	cagtccaatc	agatgagttg	ggtcgggtga	gaaaatctca	2160
ttcttggct	aataaccctaa	acagggaagc	agatgtttca	tgggctcagc	agatgttaaa	2220
agactcttca	tcacctagga	acggaaaccg	tgttgtgaac	atgaatgggtg	caaggccatt	2280
gactcaaggt	ggttcgagtg	tgaatcctca	caacagtgac	actcgtgaga	gcgacattct	2340
tgatgcgtgg	cttgaacagc	tgcacctaga	tcgctgagcc	tcagctgcca	gagagaggtt	2400
cacatttctg	tgaagctgtg	aaactgatga	ttcgtttatt	tattattcaa	gaaagcaaac	2460
ggaacaaaaa	gcaaactccg	ggtaagcttt	tttcgattct	aataacccta	aaaggctcag	2520
ttttttcagg	cttctttctg	aaatttcttt	actttcttat	ttttatcacc	tcattaaatt	2580
aattattgta	tcactctctg	tgtaacaatg	gccaaagtgc	gcctctatta	cttcccgat	2640
ttctgattta	cattttttgt	atcctctcag	tttgtcaatt	gtttctaata	tctccttcat	2700
atthgtcaaa	gaacactgta	tgagaaataa	taacatattg	tttcagctaa	taagattcat	2760
tcatttct						2769

<210> 19
 <211> 706
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 19

ES 2 451 669 T3

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

ES 2 451 669 T3

Ile Asp Pro Ser Leu Pro Asp Ile Lys Ser Gly Ile Tyr Ser Thr Asp
 245 250 255
 Glu Phe Arg Met Phe Ser Phe Lys Ile Arg Pro Cys Ser Arg Ala Tyr
 260 265 270
 Ser His Asp Trp Thr Glu Cys Pro Phe Ala His Pro Gly Glu Asn Ala
 275 280 285
 Arg Arg Arg Asp Pro Arg Lys Phe His Tyr Thr Cys Val Pro Cys Pro
 290 295 300
 Asp Phe Lys Lys Gly Ser Cys Lys Gln Gly Asp Met Cys Glu Tyr Ala
 305 310 315 320
 His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln Tyr Arg Thr Arg
 325 330 335
 Leu Cys Lys Asp Gly Met Gly Cys Asn Arg Arg Val Cys Phe Phe Ala
 340 345 350
 His Ala Asn Glu Glu Leu Arg Pro Leu Tyr Pro Ser Thr Gly Ser Gly
 355 360 365
 Leu Pro Ser Pro Arg Ala Ser Ser Ala Val Ser Ala Ser Thr Met Asp
 370 375 380
 Met Ala Ser Val Leu Asn Met Leu Pro Gly Ser Pro Ser Ala Ala Gln
 385 390 395 400
 His Ser Phe Thr Pro Pro Ile Ser Pro Ser Gly Asn Gly Ser Met Pro
 405 410 415
 His Ser Ser Met Gly Trp Pro Gln Gln Asn Ile Pro Ala Leu Asn Leu
 420 425 430
 Pro Gly Ser Asn Ile Gln Leu Ser Arg Leu Arg Ser Ser Leu Asn Ala
 435 440 445
 Arg Asp Ile Pro Ser Glu Gln Leu Ser Met Leu His Glu Phe Glu Met
 450 455 460
 Gln Arg Gln Leu Ala Gly Asp Met His Ser Pro Arg Phe Met Asn His
 465 470 475 480
 Ser Ala Arg Pro Lys Thr Leu Asn Pro Ser Asn Leu Glu Glu Leu Phe
 485 490 495
 Ser Ala Glu Val Ala Ser Pro Arg Phe Ser Asp Gln Leu Ala Val Ser
 500 505 510
 Ser Val Leu Ser Pro Ser His Lys Ser Ala Leu Leu Asn Gln Leu Gln
 515 520 525
 Asn Asn Lys Gln Ser Met Leu Ser Pro Ile Lys Thr Asn Leu Met Ser
 530 535 540
 Ser Pro Lys Asn Val Glu Gln His Ser Leu Leu Gln Gln Ala Ser Ser
 545 550 555 560
 Pro Arg Gly Gly Glu Pro Ile Ser Pro Met Asn Ala Arg Met Lys Gln
 565 570 575
 Gln Leu His Ser Arg Ser Leu Ser Ser Arg Asp Phe Gly Ser Ser Leu
 580 585 590
 Pro Arg Asp Leu Met Pro Thr Asp Ser Gly Ser Pro Leu Ser Pro Trp
 595 600 605
 Ser Ser Trp Asp Gln Thr His Gly Ser Lys Val Asp Trp Ser Val Gln
 610 615 620
 Ser Asp Glu Leu Gly Arg Leu Arg Lys Ser His Ser Leu Ala Asn Asn
 625 630 635 640
 Pro Asn Arg Glu Ala Asp Val Ser Trp Ala Gln Gln Met Leu Lys Asp
 645 650 655
 Ser Ser Ser Pro Arg Asn Gly Asn Arg Val Val Asn Met Asn Gly Ala
 660 665 670
 Arg Pro Leu Thr Gln Gly Gly Ser Ser Val Asn Pro His Asn Ser Asp
 675 680 685
 Thr Arg Glu Ser Asp Ile Leu Asp Ala Trp Leu Glu Gln Leu His Leu
 690 695 700
 Asp Arg

5
 <210> 20
 <211> 2674
 <212> ADN
 <213> *Oryza sativa*
 <400> 20

```

tgcgagtcct cctcctcttc tegtgcgct gtaactctcg ctttctctct ctctctctct      60
cacttgttcc ccaaggcgag aagcagccgc cgccggcgag cgtcgcgggg gaggggaggg      120
aagggagggg ggagcggtgg atccgggctt gattggatg ggtcggattc gattttggat      180
caaccccgga gggcgggagc ggttgctaca gatgcgttga gctttggtta atctatccgg      240
cgagagataa tggcgagct tgctgatctc gttgtcgtgc cgtcgcagcc gccgctcgcc      300
ggcggccggc gggacaggct ggcggcgtg ctggagctcg cggcggcgga tgatggtgat      360
gggctcaggg gggcgctcgc ggagggagge gaggagcgg cggagttggc tgatggggtc      420
gggctgtggt atggtcggag caaggcgtac gaggcgcga cggcgtgat ggtggcggcg      480
acgtacggca gcgccggggt ggtctcgtg ctggtgggcc tcggcggttg cgtcgacgtc      540
aaccgtcgcc ctggagccga cggcgccacc gcgctccact gcgccgcctc cggtggtctg      600
cgcaacgccg tcgctggtgt caagctgctt ttggccgctg gcgccgatcc ggccaccccc      660
gattccgccc gccgcttccc cgcgacgtc atcctagctc ctccggcttc gccagatgcc      720
cttggcgatc tcgaggtgct cctcggcgcg cgcgagcac tcctcaccag tcgccgtgg      780
gcttcaggtt cgtcatcccc tccgctctcg tccctaccag atgagggcaa caggtogccc      840
tegtcgcggt cgtcgtgct gtctcccatc actgtggatc gtgggaagaa ggagtatccg      900
gtggatccaa ctctgccgga catcaagagc agcgtgtatg cttcggatga gttccgcatg      960
tttgcgttca aggtccggcc ctgctcccgt gcctactcac acgactggac tgagtggccc      1020
tttgtcacc ccggcgagaa cgcgcgcgc cgtgatcccc gcaagcacc atactactgct      1080
gtgccttgcc ccaactttcg cggcctggt ggctgcccta gcggcgatag ctgtgagttc      1140
tcgcatggcg tgtttgagag ctggctacac ccatcacagt atcgacacaag gctctgcaag      1200
gagggagcag ctgcgcccc tcgcatgtgc ttctttgccc atgatgagga tgagctccgc      1260
catgtgctc acaacagtgg tgccggcctg ctgtctcccc gcgcttcttc atccattgat      1320
atgctgtgt cagctgcgct cgggcttctc ccaggttctc ctaccagaca ctttgaccg      1380
ccgctgtgt caccatctgc tgggagcaat ggagagctg ctgctgcgca ttggctccaa      1440
ggcagtaggc tgcgttcttc tttcaatgca agggagctg ctggtgatga ccttggcatg      1500
ctcctcgaat ggaatcaca atacctggg gcactctgccc tgccaccag cagccgcccc      1560
caaccacgcc tttcagctgg tctgagtatc agccaacaa ttgctccatc caatcttga      1620
gacatgtatg cttcagacat ggcaatgtct ccgaggttcc ctaatgacca aggtcactca      1680
gtctactcac cagcccacaa atcagccctc ctcaacaagc tcatcaaca gaagggcctc      1740
ttatcacctg ttaacaccaa cagaatgtac tccccagggt ctcttgatcc gtcactttg      1800
gcacattctc catttgggtg catgtctccc cggctcccccc gtaccatgga acctacatca      1860
ccctaagtg ctcgtgtagg agcccctgcc acacagcggc cttctggtgg ttaccacgg      1920
aattccagtg cttggggcac cgtggggtcc ccgatgggta aggttgactg ggggtctgat      1980
agcgagagc tagtccgct gagacgccct gcacaaccag ggtttggaga agatgagaca      2040
gatgtatcat gggtcagtc actggtaagc aatgctgagc ttaatggcaa gagggcgaa      2100
gtacaaggca tgcctggtac ttctgcattg atgaacaggc ctgacctgaa caatcagggt      2160
gacttggtg accagacggt gatcgggtgct tggcttgagc agatgcacct ggatcagaag      2220
tgatttccaa ggaagccat gaagtccaa agtggatgaa gcctttattt tgccaagggt      2280
atttaccaaa gaatagtgt tggctcctagt aaataataat ttattctttt taattcttga      2340
aatttttggg gggcaaagtc agagatgggtg gtcaagttca acaaacatt tggtcacaga      2400
ttggtagctg aatcagttc cagagattgg taacaacct cattacttgg ggtcctaact      2460
agtattcttt tgattagctc agatgagtct ttattttagt gggttaaaat tcatatgttc      2520
cccattggtt ttatgtccat gatctcttcc taacaaaaga gagattataa ttgtccattt      2580
ttcatttatc aatgaatgat tttgttaaaa caatgtaagt tacattctta atttttctc      2640
tgttcaatgg aattaccttc cttgggttagt cctc      2674
    
```

10
 <210> 21
 <211> 657
 <212> PRT
 <213> *Oryza sativa*
 15
 <400> 21

ES 2 451 669 T3

Met Gly Glu Leu Ala Asp Leu Val Val Val Pro Ser Gln Pro Pro Leu
1 5 10 15
Ala Gly Gly Arg Arg Asp Arg Leu Ala Ala Leu Leu Glu Leu Ala Ala
20 25 30
Ala Asp Asp Val Asp Gly Leu Arg Gly Ala Leu Ala Glu Gly Gly Glu
35 40 45
Glu Ala Ala Glu Leu Ala Asp Gly Val Gly Leu Trp Tyr Gly Arg Ser
50 55 60
Lys Ala Tyr Glu Ala Arg Thr Pro Leu Met Val Ala Ala Thr Tyr Gly
65 70 75 80
Ser Ala Gly Val Val Ser Leu Leu Val Gly Leu Gly Gly Cys Val Asp
85 90 95
Val Asn Arg Arg Pro Gly Ala Asp Gly Ala Thr Ala Leu His Cys Ala
100 105 110
Ala Ser Gly Gly Ser Arg Asn Ala Val Ala Val Val Lys Leu Leu Leu
115 120 125
Ala Ala Gly Ala Asp Pro Ala Thr Pro Asp Ser Ala Gly Arg Phe Pro
130 135 140
Ala Asp Val Ile Leu Ala Pro Pro Ala Ser Pro Asp Ala Leu Gly Asp
145 150 155 160
Leu Glu Val Leu Leu Gly Arg Arg Arg Ala Leu Ala Val Ala Thr Ser
165 170 175
Val Ala Ser Gly Ser Ser Ser Pro Pro Leu Ser Ser Ser Pro Asp Glu
180 185 190
Gly Asn Arg Ser Pro Ser Ser Arg Ser Ser Ser Leu Ser Pro Ile Thr
195 200 205
Val Asp Arg Gly Lys Lys Glu Tyr Pro Val Asp Pro Thr Leu Pro Asp
210 215 220
Ile Lys Ser Ser Val Tyr Ala Ser Asp Glu Phe Arg Met Phe Ala Phe
225 230 235 240
Lys Val Arg Pro Cys Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys
245 250 255
Pro Phe Val His Pro Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys
260 265 270
His Pro Tyr Thr Ala Val Pro Cys Pro Asn Phe Arg Arg Pro Gly Gly
275 280 285
Cys Pro Ser Gly Asp Ser Cys Glu Phe Ser His Gly Val Phe Glu Ser
290 295 300
Trp Leu His Pro Ser Gln Tyr Arg Thr Arg Leu Cys Lys Glu Gly Ala
305 310 315 320
Ala Cys Ala Arg Arg Ile Cys Phe Phe Ala His Asp Glu Asp Glu Leu
325 330 335
Arg His Val Pro His Asn Ser Gly Ala Gly Leu Leu Ser Pro Arg Ala
340 345 350
Ser Ser Ser Ile Asp Met Thr Ala Ala Ala Ala Leu Gly Leu Leu Pro
355 360 365
Gly Ser Pro Thr Arg His Phe Ala Pro Pro Pro Val Ser Pro Ser Ala
370 375 380
Gly Ser Asn Gly Gly Ala Ala Ala Ala His Trp Leu Gln Gly Ser Arg
385 390 395 400
Leu Arg Ser Ser Phe Asn Ala Arg Asp Ala Ala Val Asp Asp Leu Gly
405 410 415
Met Leu Leu Glu Trp Glu Ser Gln Tyr Leu Gly Ala Leu Cys Leu Pro
420 425 430
Pro Ser Ser Arg Pro Gln Pro Arg Leu Ser Ala Gly Leu Ser Ile Arg
435 440 445
Pro Thr Ile Ala Pro Ser Asn Leu Glu Asp Met Tyr Ala Ser Asp Met
450 455 460

ES 2 451 669 T3

Ala Met Ser Pro Arg Phe Pro Asn Asp Gln Gly His Ser Val Tyr Ser
 465 470 475 480
 Pro Ala His Lys Ser Ala Leu Leu Asn Lys Leu His Gln Gln Lys Gly
 485 490 495
 Leu Leu Ser Pro Val Asn Thr Asn Arg Met Tyr Ser Pro Arg Ala Leu
 500 505 510
 Asp Pro Ser Ser Leu Ala His Ser Pro Phe Gly Gly Met Ser Pro Arg
 515 520 525
 Ser Pro Arg Thr Met Glu Pro Thr Ser Pro Leu Ser Ala Arg Val Gly
 530 535 540
 Ala Pro Ala Thr Gln Arg Pro Ser Val Gly Ser Pro Arg Asn Ser Ser
 545 550 555 560
 Ala Trp Gly Thr Val Gly Ser Pro Met Gly Lys Val Asp Trp Gly Val
 565 570 575
 Asp Ser Glu Glu Leu Val Arg Leu Arg Arg Pro Ala Gln Pro Gly Phe
 580 585 590
 Gly Glu Asp Glu Thr Asp Val Ser Trp Val Gln Ser Leu Val Ser Asn
 595 600 605
 Ala Glu Leu Asn Gly Lys Arg Gly Glu Val Gln Gly Met Pro Gly Thr
 610 615 620
 Ser Ala Leu Met Asn Arg Pro Asp Leu Asn Asn Gln Gly Asp Leu Leu
 625 630 635 640
 Asp Gln Thr Val Ile Gly Ala Trp Leu Glu Gln Met His Leu Asp Gln
 645 650 655
 Lys

- <210> 22
- <211> 2223
- <212> ADN
- <213> *Arabidopsis thaliana*
- <400> 22

5

ES 2 451 669 T3

ttcttcaaaa	accccaacca	cttcttctcc	ccaaaaacct	ccaaagtffc	aatctttact	60
tctctctttt	tctccaagtt	atcttctttt	ctaggaagag	atatgtgctg	tgcaaagagc	120
aacctttgct	catctaaaac	cctaacagaa	gtcgaattca	tgaggcagaa	atcagaagac	180
ggagcttccg	ccacgtgtct	cctcgaattc	gccgcctgtg	atgatctttc	atcgtttaag	240
agagagatcg	aagagaatcc	atcgggtggag	attgatgagt	caggggtttg	gtattgcaga	300
cgggtcgggt	ctaagaagat	gggttttgaa	gaaagaacac	cacttatggg	tgctgctatg	360
tatggaagca	tggaagtgtt	gaattacata	attgccacag	gaagatccga	tgtgaacaga	420
gtttgacgtg	acgagaaagt	cactgctctt	cactgtgcag	tttctggctg	ttctgtttct	480
atcgttgaga	tcatcaagat	cttgcttgat	gcttctgctt	cacctaattg	tgttgacgct	540
aatgggaaca	aaccggttga	ttgttggtc	aaagattctc	ggtttgttcc	taaccagagt	600
agaaagccg	ttgaggtttt	actgaccggg	attcatggtt	cggttatgga	agaagaggag	660
gaggaactga	agagtgttgt	gactaagtat	ccagctgatg	catcacttcc	tgatattaac	720
gaaggtgttt	atggaactga	tgattttagg	atgttttagc	ttaaggttaa	gccatgttct	780
agggcttatt	cacatgattg	gactgaatgt	ccttttgttc	atcctgggtg	gaatgcaagg	840
aggagagatc	ctaggaagta	tccttacact	tgtgtgcctt	gtcccagagt	tcgtaaaggg	900
tcttgcctta	aaggagattc	gtgtgagtac	gcgcacgggt	tttccagagc	ttggcttcac	960
ccggcgagct	ataggacacg	gctttgcaaa	gatgagactg	gttgtgctag	gagagtttgt	1020
ttctttgctc	atagacggga	tgagttaaga	ccggttaatg	cttctactgg	ttctgcaatg	1080
gtttcaccaa	ggtcgtctaa	tcagtcctct	gagatgtctg	ttatgtctcc	tttgacgctg	1140
ggatcatcgc	caatgaactc	tcctatggct	aatgggttcc	ctttgtctcc	aagaaatggt	1200
ggtttatggc	agaacagagt	taatagcctt	acaccaccac	cgttgcagct	taatggtagc	1260
agattgaagt	cgactttgag	tgctagagat	atggatatgg	agatggaact	taggtttcgc	1320
ggtttgata	accgggagact	tggtgatctc	aagccatcca	acctcgaaga	gactttcggg	1380
tcatatgact	cagcttctgt	gatgcaactt	caatcaccaa	gcaggcattc	tcagatgaac	1440
cactatccgt	cttcacctgt	gaggcagcct	cctcctcatg	gattcgaatc	ttcagcagcc	1500
atggcagctg	cagtgatgaa	tgcaagatcc	tcagcgtttg	cgaaacgcag	cttgagtttc	1560
aaaccagctc	cagtagcttc	taatgtctcc	gattggggat	caccaaattg	gaagcttgag	1620
tggggaatgc	aaagagatga	gctgaacaag	ttgaggagaa	gtgcctcctt	cggcattcat	1680
ggaaacaaca	acaacagtgt	gtcacgccct	gctagagact	acagtgacga	gccagatgtg	1740
tcgtgggtga	actcactggt	gaaagagaat	gcaccagaga	gagtgaatga	gagggttggg	1800
aatacgggtg	atgggtgcagc	gagtagagac	aagtttaagc	tgccgtcgtg	ggcagagcaa	1860
atgtatatag	accatgagca	gcagattgtg	gcataagaag	cagaaagaaa	gatgtgggat	1920
ttatattgct	tttgtcttct	gggcctctct	acacagaatc	taacaaatct	ggcaataatt	1980
ctttgatattg	tgtttgacct	atagtttgg	tactagtata	tgttttttta	tgttcttttt	2040
ttctttgtca	ttctcttgtc	cttcgtgaca	ctatgtaatg	attaaaagca	aataattgat	2100
gcatgagttc	aaatgttctt	tgaaggatcc	atcttattag	ctttgttaatt	gttgtgatat	2160
cttaacttta	ttggttacgt	atttcaagtg	ctttagaaaa	aatgggccta	agagattttg	2220
ggg						2223

<210> 23

<211> 597

<212> PRT

<213> *Arabidopsis thaliana*

<400> 23

ES 2 451 669 T3

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

ES 2 451 669 T3

Ser Thr Gly Ser Ala Met Val Ser Pro Arg Ser Ser Asn Gln Ser Pro
 325 330 335
 Glu Met Ser Val Met Ser Pro Leu Thr Leu Gly Ser Ser Pro Met Asn
 340 345 350
 Ser Pro Met Ala Asn Gly Val Pro Leu Ser Pro Arg Asn Gly Gly Leu
 355 360 365
 Trp Gln Asn Arg Val Asn Ser Leu Thr Pro Pro Pro Leu Gln Leu Asn
 370 375 380
 Gly Ser Arg Leu Lys Ser Thr Leu Ser Ala Arg Asp Met Asp Met Glu
 385 390 395 400
 Met Glu Leu Arg Phe Arg Gly Leu Asp Asn Arg Arg Leu Gly Asp Leu
 405 410 415
 Lys Pro Ser Asn Leu Glu Glu Thr Phe Gly Ser Tyr Asp Ser Ala Ser
 420 425 430
 Val Met Gln Leu Gln Ser Pro Ser Arg His Ser Gln Met Asn His Tyr
 435 440 445
 Pro Ser Ser Pro Val Arg Gln Pro Pro Pro His Gly Phe Glu Ser Ser
 450 455 460
 Ala Ala Met Ala Ala Ala Val Met Asn Ala Arg Ser Ser Ala Phe Ala
 465 470 475 480
 Lys Arg Ser Leu Ser Phe Lys Pro Ala Pro Val Ala Ser Asn Val Ser
 485 490 495
 Asp Trp Gly Ser Pro Asn Gly Lys Leu Glu Trp Gly Met Gln Arg Asp
 500 505 510
 Glu Leu Asn Lys Leu Arg Arg Ser Ala Ser Phe Gly Ile His Gly Asn
 515 520 525
 Asn Asn Asn Ser Val Ser Arg Pro Ala Arg Asp Tyr Ser Asp Glu Pro
 530 535 540
 Asp Val Ser Trp Val Asn Ser Leu Val Lys Glu Asn Ala Pro Glu Arg
 545 550 555 560
 Val Asn Glu Arg Val Gly Asn Thr Val Asn Gly Ala Ala Ser Arg Asp
 565 570 575
 Lys Phe Lys Leu Pro Ser Trp Ala Glu Gln Met Tyr Ile Asp His Glu
 580 585 590
 Gln Gln Ile Val Ala
 595

<210> 24
 <211> 1761
 <212> ADN
 <213> *Arabidopsis thaliana*

<400> 24

ES 2 451 669 T3

atggaaaaag	atagtattat	gtgcagtgga	ccaaagagca	atctctgctc	ttcaagaacc	60
ttaacagaaa	tcgaatcaag	gcaaaaggaa	gaagaaacaa	tgcttctcct	cgaattcgct	120
gcttgatgat	atcttgactc	gttcaagaga	gaggttgaag	agaaagggct	tgatttgat	180
gagtcagggg	tatggatttg	cagacgtgtc	ggttctaaga	agatgggtct	tgaagaaaga	240
acacctttaa	tggttgcagc	tatgtatgga	agcataaagg	ttttgacttt	catcgtttcc	300
actggaaaa	ctgatgtgaa	cagagcttgt	ggtgaagaga	gagttactcc	gcttcaactgt	360
gctggtgctg	gctggttctg	gaatatgatt	gaagtcatca	atgtcttgct	tgatgcttct	420
gctttgggta	actctggtga	tgctaagggg	aatcaacctt	tggatgtggt	tgttcgagtt	480
tcgaggtttg	tggctagtcc	gaggaggaaa	gcggttgagt	tgttgctgag	aggaggaggt	540
gttggaggat	tgatcgatga	ggcggttgaa	gaagagatca	agattgtctc	taagtatcca	600
gctgatgctt	ctttaccgga	tataaacgaa	ggggtttatg	gaagtgatga	gtttaggatg	660
tatagcttta	aggttaagcc	atggtctagg	gcttattctc	atgattggac	cgagtgtgct	720
tttgttcatc	cgggagaaaa	tgcgaggagg	agagatccga	ggaagtatcc	ttacacttgt	780
gtccccgtgc	ccgagttccg	taaaggatca	tgcccgaag	gagattcttg	cgagtatgct	840
cacgggggtt	tcgagtcgtg	gcttcacccc	gcgcagtata	aaacccggct	ttgtaaagat	900
gaaacggggt	gtgcaaggaa	agtttgtttc	tttgctcata	aacgcgaaga	gatgagacct	960

gttaatgctt	caactggctc	tgccgtggct	cagtctccgt	ttagcagctt	ggagatgatg	1020
ccaggggttg	ctcctcttgc	ttattcttca	ggagtttoga	ctcctccggt	ttctccaatg	1080
gctaattggg	ttccttctc	tccaagaaac	ggcggatcat	ggcagaacag	agtcaatacc	1140
cttactccac	cggctttgca	gctcaatggt	ggaagcagat	tgaagtccac	actgagcgct	1200
agagatatcg	atatggagat	ggagatggaa	ttgagactcc	gcggttttgg	caacaatgtg	1260
gaagagacgt	tcgggtctta	tgtttctct	ccaagtagga	attctcaaat	gggtcaaac	1320
atgaaccaac	attatccatc	ttccccggtg	agacaaccgc	catctcaaca	cgggttcgaa	1380
tcttcagcag	ctgcagcggg	tgcaagtgat	aaagcgagat	caaccgcctt	tgcgaaaacgt	1440
agcttgagct	tcaaaccagc	tactcaagca	gcaccacagt	cgaatctctc	ggattgggga	1500
tctccaaaac	ggaagctgga	atggggaatg	aaaggagaag	agctgaataa	gatgagaaga	1560
agtgtttcct	ttggaatcca	tggaaacaac	aacaataacg	cagctagaga	ctacagggac	1620
gagccagatg	tgatcatggg	taactcttta	gttaaagaca	gtactgtggt	gtctgagaga	1680
agctttggaa	tgaatgagag	ggttcggata	atgtcgtggg	ctgagcaaat	gtacagagag	1740
aaggagcaga	ctgtggtgta	a				1761

<210> 25
 <211> 586
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 25

5

ES 2 451 669 T3

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

ES 2 451 669 T3

Ala Arg Lys Val Cys Phe Phe Ala His Lys Arg Glu Glu Met Arg Pro
 305 310 315 320
 Val Asn Ala Ser Thr Gly Ser Ala Val Ala Gln Ser Pro Phe Ser Ser
 325 330 335
 Leu Glu Met Met Pro Gly Leu Ser Pro Leu Ala Tyr Ser Ser Gly Val
 340 345 350
 Ser Thr Pro Pro Val Ser Pro Met Ala Asn Gly Val Pro Ser Ser Pro
 355 360 365
 Arg Asn Gly Gly Ser Trp Gln Asn Arg Val Asn Thr Leu Thr Pro Pro
 370 375 380
 Ala Leu Gln Leu Asn Gly Gly Ser Arg Leu Lys Ser Thr Leu Ser Ala
 385 390 395 400
 Arg Asp Ile Asp Met Glu Met Glu Met Glu Leu Arg Leu Arg Gly Phe
 405 410 415
 Gly Asn Asn Val Glu Glu Thr Phe Gly Ser Tyr Val Ser Ser Pro Ser
 420 425 430
 Arg Asn Ser Gln Met Gly Gln Asn Met Asn Gln His Tyr Pro Ser Ser
 435 440 445
 Pro Val Arg Gln Pro Pro Ser Gln His Gly Phe Glu Ser Ser Ala Ala
 450 455 460
 Ala Ala Val Ala Val Met Lys Ala Arg Ser Thr Ala Phe Ala Lys Arg
 465 470 475 480
 Ser Leu Ser Phe Lys Pro Ala Thr Gln Ala Ala Pro Gln Ser Asn Leu
 485 490 495
 Ser Asp Trp Gly Ser Pro Asn Gly Lys Leu Glu Trp Gly Met Lys Gly
 500 505 510
 Glu Glu Leu Asn Lys Met Arg Arg Ser Val Ser Phe Gly Ile His Gly
 515 520 525
 Asn Asn Asn Asn Asn Ala Ala Arg Asp Tyr Arg Asp Glu Pro Asp Val
 530 535 540
 Ser Trp Val Asn Ser Leu Val Lys Asp Ser Thr Val Val Ser Glu Arg
 545 550 555 560
 Ser Phe Gly Met Asn Glu Arg Val Arg Ile Met Ser Trp Ala Glu Gln
 565 570 575
 Met Tyr Arg Glu Lys Glu Gln Thr Val Val
 580 585

<210> 26
 <211> 2709
 <212> ADN
 <213> *Eucalyptus grandis*

<400> 26

ES 2 451 669 T3

cttctgaaag	ctttttgact	taagacgaga	gagaaggaga	gaaggteccc	ctcctcgtcc	60
tcgtccccc	gtggattttg	aagaagaaaa	gtcgcacott	cotttctcctt	teccactcct	120
ccctctgctc	gaagcttttc	tcttccgcag	aattacataa	aaacctcgac	tttgcgcatic	180
attccgattc	acctcacacc	ttcactttcc	cactcgaggt	ctccccctc	ttttcctagc	240
tctttccott	tcctccctc	tctctcgaga	atcgccgcat	ttggaggagc	tccaatctgc	300
tttgctttgc	tttgctctct	tcttgctcgg	ttccccctca	taaggagtgc	attatgtgca	360
acggttcttc	gaagggtaaa	cttttcccc	cgagtatggg	catggagggc	gaattccaca	420
acaaggatgg	cgaagcacc	cgtaaagtct	ctgccttget	tgaattggca	gcctcggacg	480
atctctcgtc	gttcaaaaagt	gaagtggaag	agaagggctg	cgacgttgat	gaggccagct	540
tttggatgg	taggagaatc	gggtcgaaga	agatgggttt	tgaagagagg	actccattga	600
tgatctctgc	tttgtttggg	agcaccaagg	tcttgaaata	cataatcgag	accgccagag	660
ctgatgtcaa	caggtcttgt	gggtccgaca	aggtggccgc	cctccattgc	gcagcccgcg	720
gtgggtccag	ttcttcactt	gaaattgtga	agctcttgat	tgaggcctca	gcggatatta	780
attctgtaga	tggcaatgga	aataggcca	tcgacgtgct	tgccccggca	gggaagtctc	840
gctgcaatc	cagaaataag	tttgtagat	cgttgctgaa	aggtgaaaac	tatgtcgtgg	900
aaggtgacca	atcctttgac	atagaaggag	aggagaagct	agtcgctctt	ccaaaggagg	960
gaggcgagaa	gaaagagtat	cctggtgatg	tctctctacc	tgacataaac	aatgggttct	1020
acagtaccga	tgagttccgg	atgtatgctt	tcaaggtgaa	gccttgctcg	agggcttact	1080
cccacgactg	gaccgagtgc	cggtttgtgc	accctgggga	gaacgcgagg	aggagggacc	1140
cacgcaagta	cccttacagc	tgtgtccott	gtcctgagtt	tcgcaaggtt	tcgtgcgtaa	1200
ggggggatgc	ttgtgagtat	gctcatggag	tctttgagtc	gtggcttcac	ccagcgcaat	1260
accgaaccgg	gctgtgcaag	gatgaaactg	gttgtactcg	caaagtttgc	ttctttgctc	1320
acaagtccga	agaattgcgt	cccgtgatg	cttccacagg	ttctgctatg	ccctcacca	1380
agtccttttc	agctaattgcc	ctagacatga	caacctgag	ccccttatcc	cttaattcac	1440
catctctgcc	tttgctgct	acttccacgc	cccccatgtc	acctttggct	gcctcatctt	1500
caccaagggg	catgaacttg	tggcataaca	aaattaacct	gacccacca	agcctgcagc	1560
ttcctggcag	ccggctgaag	acggctatga	gtgcgcggga	cttcgatttt	gagttggaat	1620
ttcttgggct	ggaaaagcaa	gcttctcagc	ggcagcaact	gatagaagag	atctctcgtc	1680
tctcatcgcc	ctctcatatg	tggaaactcg	aatttggcag	aaccgcagag	ctgaagccca	1740
ctaaccttga	tgatgcgttt	ggatctcttg	acacttctct	tttgtctccg	ttgcaggggt	1800
cgctgatgaa	aacatcgact	cctaccagt	tgcaatcccc	cacagggctt	aaaatttcga	1860
atttgaacca	actccgtgcg	agctaccctg	ctagcagctt	gtcgtcctct	cctgtgagga	1920
agacctcttc	ttttgggttc	gactcatcca	gtgcagttgc	tgcagcagtc	atgaactcac	1980
ggctctgctgc	tatgacgaag	cggagccaga	gcttcattga	ccgtggagca	gtgggtcaac	2040
ggctctggact	cattggacct	gctaattctg	ctcctaggat	gtccaacctt	tcggactggg	2100
gctcgcctga	tgggaagtgg	gattggggtg	ttcaagggga	cgagctcaac	aagcttagga	2160
agtcogcttc	cttcggcttt	agaaacaaca	gtatggcgaa	cccaaacaac	gtggcgtctc	2220
ccagtgctga	tgagccggac	gtgtcgtggg	ttggttcatt	ggtgaaggat	gtggctccgc	2280
ccgaagggtg	tccacagtat	ctgtacatag	aacaggagca	gatggtggca	taactaaagc	2340
gaagagcacc	acacgaactc	tctcctgatg	gcttaagatg	acttgtttga	cattctttat	2400
attcttacia	acagcgcggt	cttaggagtt	agctggagga	agaaggaag	cggtattgag	2460
tttgagattc	aggctcttag	ctggacagcg	aaaatttggg	gaaggaagag	aatttggttt	2520
cttgcccaac	ttagataatg	atgcttttga	aggcttaaaa	gaaagatgaa	ggcaaacatt	2580
cttttgttag	tattgtatta	ttgtttta	ttttcatccc	ctctgtcggg	gtgtggtggg	2640
tgctgatggt	tctttcatca	gtaaaatata	taatgaggtt	tactcatcta	tttttacta	2700
aaaaaaaa						2709

- <210> 27
- <211> 659
- <212> PRT
- <213> *Eucalyptus grandis*
- <400> 27

5

ES 2 451 669 T3

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

Gly Glu Asn Tyr Val Val Glu Gly Asp Gln Ser Phe Asp Ile Glu Gly
 180 185 190
 Glu Glu Lys Leu Val Ala Leu Pro Lys Glu Gly Gly Glu Lys Lys Glu
 195 200 205
 Tyr Pro Val Asp Val Ser Leu Pro Asp Ile Asn Asn Gly Phe Tyr Ser
 210 215 220
 Thr Asp Glu Phe Arg Met Tyr Ala Phe Lys Val Lys Pro Cys Ser Arg
 225 230 235 240
 Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro Gly Glu
 245 250 255
 Asn Ala Arg Arg Arg Asp Pro Arg Lys Tyr Pro Tyr Ser Cys Val Pro
 260 265 270
 Cys Pro Glu Phe Arg Lys Gly Ser Cys Val Arg Gly Asp Ala Cys Glu
 275 280 285
 Tyr Ala His Gly Val Phe Glu Ser Trp Leu His Pro Ala Gln Tyr Arg
 290 295 300
 Thr Arg Leu Cys Lys Asp Glu Thr Gly Cys Thr Arg Lys Val Cys Phe
 305 310 315 320
 Phe Ala His Lys Ser Glu Glu Leu Arg Pro Val Tyr Ala Ser Thr Gly
 325 330 335
 Ser Ala Met Pro Ser Pro Lys Ser Phe Ser Ala Asn Ala Leu Asp Met
 340 345 350
 Thr Thr Leu Ser Pro Leu Ser Leu Asn Ser Pro Ser Leu Pro Leu Pro
 355 360 365
 Ala Thr Ser Thr Pro Pro Met Ser Pro Leu Ala Ala Ser Ser Ser Pro
 370 375 380
 Lys Gly Met Asn Leu Trp His Asn Lys Ile Asn Leu Thr Pro Pro Ser
 385 390 395 400
 Leu Gln Leu Pro Gly Ser Arg Leu Lys Thr Ala Met Ser Ala Arg Asp
 405 410 415
 Phe Asp Phe Glu Leu Glu Phe Leu Gly Leu Glu Lys Gln Ala Ser Gln
 420 425 430
 Arg Gln Gln Leu Ile Glu Glu Ile Ser Arg Leu Ser Ser Pro Ser His
 435 440 445
 Met Trp Asn Ser Glu Phe Gly Arg Thr Ala Glu Leu Lys Pro Thr Asn
 450 455 460
 Leu Asp Asp Ala Phe Gly Ser Leu Asp Thr Ser Leu Leu Ser Pro Leu
 465 470 475 480
 Gln Gly Ser Ser Met Lys Thr Ser Thr Pro Thr Gln Leu Gln Ser Pro
 485 490 495
 Thr Gly Leu Lys Ile Ser Asn Leu Asn Gln Leu Arg Ala Ser Tyr Pro
 500 505 510
 Ser Ser Ser Leu Ser Ser Ser Pro Val Arg Lys Thr Ser Ser Phe Gly
 515 520 525
 Phe Asp Ser Ser Ser Ala Val Ala Ala Val Met Asn Ser Arg Ser
 530 535 540
 Ala Ala Met Thr Lys Arg Ser Gln Ser Phe Ile Asp Arg Gly Ala Val
 545 550 555 560
 Gly Gln Arg Ser Gly Leu Ile Gly Pro Ala Asn Ser Ala Pro Arg Met
 565 570 575
 Ser Asn Leu Ser Asp Trp Gly Ser Pro Asp Gly Lys Leu Asp Trp Gly
 580 585 590
 Val Gln Gly Asp Glu Leu Asn Lys Leu Arg Lys Ser Ala Ser Phe Gly
 595 600 605
 Phe Arg Asn Asn Ser Met Ala Asn Pro Asn Asn Val Ala Ser Pro Ser
 610 615 620
 Ala Asp Glu Pro Asp Val Ser Trp Val Gly Ser Leu Val Lys Asp Val
 625 630 635 640
 Ala Pro Pro Glu Gly Tyr Pro Gln Tyr Leu Tyr Ile Glu Gln Glu Gln

ES 2 451 669 T3

		645		650		655	
	Met Val Ala						
5	<210> 28						
	<211> 2518						
	<212> ADN						
	<213> <i>Eucalyptus grandis</i>						
	<400> 28						
	tctccttcga gtttctttct tcactagaat tcgctcccga gtctgttggt gctcgtgtag 60						
	ttttgcttac tccgtccttc gtttagctcg ctgaccagcg cggagctagg agcggtcgct 120						
	aaaggattac tcgtacaaaa cgtaaaactca gctctgcca ttttcccatg gagggggaat 180						
	cttacttcga gaaagatgaa aaatattcta attgctcaat ctgctcgaa ttatctgctt 240						
	cggacgatct cccagctttt gaaaggaaag cgaaaagagaa gggctgtaac attgatggtg 300						
	ctagcttctg gtacggtaga agaattggct caaggaagat gggctctgaa gagaggactc 360						
	ctctcatggt ggcttctctg tttggaagct ctagggttgt gaagtacatt ctcgaaatctg 420						
	gcaaagtcca tgtaaatagg gcttgtgggt cgacaaggt cactgccctt cactgtgctg 480						
	ttgccagtgg ctctgcttct gcggtggagg ttgtcaagct ctgcttcac gcactctgccg 540						
	atgctaattg cattgatggc aatggaaaga agccaattga tgtgatagcc cttccattaa 600						
	agtcaacgag cgattcaagg aggaagctga tggagctggt gctgaaaggc gataattctg 660						
	atgggggaatt tgaatcccac gaggagaagc cgattgccgc accgcaagca tccaaagagg 720						
	gaagcgaaaa gaaagatgat caatttcctg ttgatatctc tctgctgac ataaatggtg 780						
	ggatttacag tactgatgag ttcagaatgt atgctttcaa agtaaagcct tgctcgcggg 840						
	catactccca tgactggaca gagtgccat ttgttcatcc tggcgagaat gcgaggaggc 900						
	gggaccctcg caagtacccc tacagctgag tcccttgccc tgaatttcgg aagggatctt 960						
	gccaaaaggg tgactcctgt gagtacgagc acggcgtatt tgagtcgtgg cttcatcctg 1020						
	cacagtatag aacaagactg tgcaaggatg agactggatg tgctcgcaaa gtttgtttct 1080						
	ttgctcacia gcccgaaaga ttaaggcctg tctatgcttc gacgggatca gctatgcctt 1140						
	ccccaaaatc ctactcatca agtgggctgg acatgtccac attgagtcct ctctcaatca 1200						
	gttctccgtc agcatcgttg cctgttactt caacagcacc catgtctcct cttgcagcct 1260						
	cgctcatctc gatgtctgtg aacatgtggc agagcaagc taacaagctc tccccgcaa 1320						
	tgctgcagct ctcagtagt aggtgaaga gctgcttgag tgctagggag ttggacctgg 1380						
	agatggaatt gcgtggtcta gagagtcaga tggccactca acagcatcag ttgatggaag 1440						
	agatatctcg tctctcctca ccatcatcct gctttagtag taggattggg gaagtgaaac 1500						
	ccactaacct cgatgacgtt tttgggtctc cggatcctgc tttgctgctt caattgcagg 1560						
	ggctgtcaag acctcaaca ccaagccagt tgcaatctcc aactgggctt cagatgcgcc 1620						
	agaatgcaac ccagttctgt ggggcgtacc agagcaatgc aaatgcattg tcatctccag 1680						
	caatgaagca ggcacctctt tatgggtttg actcatctag tgcagttgca gcagcggatg 1740						
	tgaattcgag gtcagccgct tttgcgaagc ggagtcagag ttttatcgac aggggaatgg 1800						
	cgtgccctgg aattgccaat tcttccccta tgatgtcttc agctatgtcg agctggagct 1860						
	cacctcatgg gaaattggat tggggcgtcc aaggagatga gttgaatagg ctgaggaaag 1920						
	ctgcttctct taagatgaga agcagcacc gacaggtgc taatactgtc tcggcagcag 1980						
	ccatggctga tgagccagat atttcttggg tcagttcatt ggttaaggac gtgccttctg 2040						
	cggaggacgc gatgttctgt gcagagaaag gacagcgcac ttatgggaaa gacatccgag 2100						
	aaaggattac cccatgggtg gagcagctgt acagagaagt gccacggatg gcgatgtaag 2160						
	attgccactg caagtccgat gccttagtat gctgactaat tgatattctt tgcatttgtt 2220						
	ttgaggcatt tggtagccat tagatacgag aaaaggccaa gcagcagggtg gtgtcttggc 2280						
	aaggaatagg atgcacatag tctgttatcg agtagaatag acttggaac aatggttata 2340						
	gccaaatggt aaaagttag atattctttt ccaattcttt ctcttctca tagtaggttt 2400						
	ctaccaagt ctttttagtga gagcctgcgg gatgtactat atgtttccct tatgtaacgt 2460						
10	ctcttcggtg aaagaaatgg ctttataata taaagcatca agttttttaa aaaaaaaa 2518						
	<210> 29						
	<211> 663						
	<212> PRT						
15	<213> <i>Eucalyptus grandis</i>						
	<400> 29						

ES 2 451 669 T3

Met Glu Gly Glu Ser Tyr Phe Glu Lys Asp Glu Lys Tyr Ser Asn Cys
1 5 10 15
Ser Ile Leu Leu Glu Leu Ser Ala Ser Asp Asp Leu Pro Ala Phe Glu
20 25 30
Arg Lys Ala Lys Glu Lys Gly Cys Asn Ile Asp Gly Ala Ser Phe Trp
35 40 45
Tyr Gly Arg Arg Ile Gly Ser Arg Lys Met Gly Leu Glu Glu Arg Thr
50 55 60
Pro Leu Met Val Ala Ser Leu Phe Gly Ser Ser Arg Val Val Lys Tyr
65 70 75 80
Ile Leu Glu Ser Gly Lys Val Asp Val Asn Arg Ala Cys Gly Ser Asp
85 90 95
Lys Val Thr Ala Leu His Cys Ala Val Ala Ser Gly Ser Ala Ser Ala
100 105 110
Val Glu Val Val Lys Leu Leu Leu His Ala Ser Ala Asp Ala Asn Cys
115 120 125
Ile Asp Gly Asn Gly Lys Lys Pro Ile Asp Val Ile Ala Leu Pro Leu
130 135 140
Lys Ser Arg Gly Asp Ser Arg Arg Lys Leu Met Glu Leu Leu Leu Lys
145 150 155 160
Gly Asp Asn Ser Asp Gly Glu Phe Glu Ser His Glu Glu Lys Pro Ile
165 170 175
Ala Ala Pro Gln Ala Ser Lys Glu Gly Ser Glu Lys Lys Glu Tyr Gln
180 185 190
Phe Pro Val Asp Ile Ser Leu Pro Asp Ile Asn Val Gly Ile Tyr Ser
195 200 205
Thr Asp Glu Phe Arg Met Tyr Ala Phe Lys Val Lys Pro Cys Ser Arg
210 215 220
Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro Gly Glu
225 230 235 240
Asn Ala Arg Arg Arg Asp Pro Arg Lys Tyr Pro Tyr Ser Cys Val Pro
245 250 255
Cys Pro Glu Phe Arg Lys Gly Ser Cys Gln Lys Gly Asp Ser Cys Glu
260 265 270
Tyr Ala His Gly Val Phe Glu Ser Trp Leu His Pro Ala Gln Tyr Arg
275 280 285
Thr Arg Leu Cys Lys Asp Glu Thr Gly Cys Ala Arg Lys Val Cys Phe
290 295 300
Phe Ala His Lys Pro Glu Glu Leu Arg Pro Val Tyr Ala Ser Thr Gly
305 310 315 320
Ser Ala Met Pro Ser Pro Lys Ser Tyr Ser Ser Ser Gly Leu Asp Met
325 330 335
Ser Thr Leu Ser Pro Leu Ser Ile Ser Ser Pro Ser Ala Ser Leu Pro
340 345 350
Val Thr Ser Thr Ala Pro Met Ser Pro Leu Ala Ala Ser Ser Ser Pro
355 360 365
Met Ser Val Asn Met Trp Gln Ser Lys Ala Asn Lys Leu Ser Pro Pro
370 375 380
Met Leu Gln Leu Ser Gly Ser Arg Leu Lys Thr Ala Leu Ser Ala Arg
385 390 395 400
Asp Leu Asp Leu Glu Met Glu Leu Arg Gly Leu Glu Ser Gln Met Ala
405 410 415
Thr Gln Gln His Gln Leu Met Glu Glu Ile Ser Arg Leu Ser Ser Pro
420 425 430
Ser Ser Cys Phe Ser Ser Arg Ile Gly Glu Val Lys Pro Thr Asn Leu
435 440 445
Asp Asp Val Phe Gly Ser Pro Asp Pro Ala Leu Leu Pro Gln Leu Gln
450 455 460
Gly Leu Ser Arg Pro Ser Thr Pro Ser Gln Leu Gln Ser Pro Thr Gly

ES 2 451 669 T3

```

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

```

```

<210> 30
<211> 2001
5 <212> ADN
<213> Triticum aestivum

```

```

10 <220>
<221> misc_feature
<222> (481)..(530)
<223> n i s a , c , g o t

```

```

<400> 30

```

ES 2 451 669 T3

cgaattccgg	tgcacgattt	ctcgatttcc	ttctctataa	cacaacgctc	tcttctcttg	60
caaccaaagt	acttgttcca	gtgtctactc	tactcaaaaa	ggatttggga	catcatgtgc	120
agtgattcga	aaagtaaact	ttcttcccca	accctcgtcg	tcatggagaa	tagtaacatt	180
cagaagcaga	atctggatgg	tctctacaac	tgggttttgc	ttgaattgtc	tgcatctgat	240
gattatgaag	ctttcaaaag	agaggtggag	gaaaaaggct	tagatgtgaa	cgaggcaggc	300
ttttggtacg	gtagaagaat	tgggtcaaaag	aagatgggat	ctgaaacgag	gaccctctg	360
atgattgctt	ctttgtttgg	aagcgccaag	gtgctcaatt	atattcttct	tcagaaagga	420
ggaggtggtg	atgtgaacag	ggtctgtggt	tctgataggg	ccactgctct	ccattgtgct	480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	tcccgggtcg	540
cgatttcgta	ggagattcac	agagagaaaa	gaagcaatta	gtgataataa	gaaagaatac	600
cctgttgata	tactactgcc	agacataaac	aacggtgtat	atggaacaga	tgattttagg	660
atgtacaact	tcaaggtgaa	gccttgctca	agggcttact	cccattgactg	gaccgagtgt	720
ccattcgttc	accagggga	gaacgctagg	aggagagacc	cacggaaata	cccttacagc	780
tgtgttcctt	gccctgagtt	cgcgaaaggg	acctgccaga	aggggtgattc	ctgtgagtat	840
gctcatgggtg	tttttgagtc	ctggctgcat	cctgcccaat	accggacaag	gctttgcaag	900
gatgagactg	gctgcgctag	aaaagtctgc	ttctttgccc	acaaacctga	agagctacgc	960
cctgtgtatg	cttccactgg	gtcggctatg	ccatcaccaa	aatcatatctc	agctagtgga	1020
cttgacatga	cagcgatgag	tccattggct	ctaagtcca	catctttgcc	taatgcccc	1080
ccgtttccag	cctcacccta	tcgtgcgccc	tcgttcttct	ctcagagtga	agctgtgcag	1140
aacaaaataa	accttactcc	accatcgttg	cagctccctg	gtagccgact	gaaggctgct	1200
ttgagtgcc	gggatctgga	gatggagatg	gaactgctcg	gtctagaaag	ccctgctcgc	1260
caacaacagc	agcagcagca	acaattgatc	gaagagattg	ccaggatctc	ttccccatct	1320
ttccggagca	aggaattcaa	taggattggt	gatttgaatc	ctactaacct	tgatgacctg	1380
ttagcatctg	ctgacccttc	tgtatcttct	caactacatg	gactttctgt	gcaaccttca	1440
acaccacac	aaagtgggct	tcagatgctc	caaacatga	accacctcgc	tgcgagttat	1500
ccatccaaca	tccttctctc	tcctgtgagg	aagccctcag	cttttggggt	tgactcatca	1560
gctgctgtgg	caactgcagt	gatgaattct	aggtctgctg	ccttcgcaaa	gcgaagccaa	1620
agtttcattg	atcgtggagc	tgcaaccac	catcttgggc	tgtcttcagc	ttccaactct	1680
tottgcaggg	tatcctctac	cctttcagat	tggagtccc	ctaccgggaa	actggattgg	1740
ggtgtaaacy	gagacaagct	gaacaagctg	aggaaatcta	cttcctttgg	attcagaaac	1800
agtggggtaa	ctgcatcccc	catagcacag	cctgaatttg	gtgctgagcc	ggatgtctca	1860
tgggttcatt	cattggttaa	agatgttccc	tccgagaggt	ctgagatatt	tgggtgctgag	1920
aagcaacaat	atgatctcag	taaagagatg	cttcaccat	ggatggagca	gctgtatata	1980
gagcaggagc	agatggtagc	a				2001

<210> 31
 <211> 667
 <212> PRT
 <213> *Triticum aestivum*

5

<220>
 <221> UNSURE
 <222> (161)..(177)
 <223> Xaa puede ser cualquier aminoácido de origen natural

10

<400> 31

ES 2 451 669 T3

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

ES 2 451 669 T3

	275					280					285				
Leu	His	Pro	Ala	Gln	Tyr	Arg	Thr	Arg	Leu	Cys	Lys	Asp	Glu	Thr	Gly
	290					295					300				
Cys	Ala	Arg	Lys	Val	Cys	Phe	Phe	Ala	His	Lys	Pro	Glu	Glu	Leu	Arg
305					310					315					320
Pro	Val	Tyr	Ala	Ser	Thr	Gly	Ser	Ala	Met	Pro	Ser	Pro	Lys	Ser	Tyr
				325					330						335
Ser	Ala	Ser	Gly	Leu	Asp	Met	Thr	Ala	Met	Ser	Pro	Leu	Ala	Leu	Ser
			340					345					350		
Ser	Thr	Ser	Leu	Pro	Asn	Ala	Pro	Pro	Phe	Pro	Ala	Ser	Pro	Tyr	Arg
		355				360						365			
Ala	Pro	Ser	Phe	Phe	Ser	Gln	Ser	Glu	Ala	Val	Gln	Asn	Lys	Ile	Asn
	370					375						380			
Leu	Thr	Pro	Pro	Ser	Leu	Gln	Leu	Pro	Gly	Ser	Arg	Leu	Lys	Ala	Ala
385					390					395					400
Leu	Ser	Ala	Arg	Asp	Leu	Glu	Met	Glu	Met	Glu	Leu	Leu	Gly	Leu	Glu
				405				410						415	
Ser	Pro	Ala	Arg	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Leu	Ile	Glu	Glu
			420				425							430	
Ile	Ala	Arg	Ile	Ser	Ser	Pro	Ser	Phe	Arg	Ser	Lys	Glu	Phe	Asn	Arg
		435					440					445			
Ile	Val	Asp	Leu	Asn	Pro	Thr	Asn	Leu	Asp	Asp	Leu	Leu	Ala	Ser	Ala
	450					455					460				
Asp	Pro	Ser	Val	Phe	Ser	Gln	Leu	His	Gly	Leu	Ser	Val	Gln	Pro	Ser
465					470					475					480
Thr	Pro	Thr	Gln	Ser	Gly	Leu	Gln	Met	Arg	Gln	Asn	Met	Asn	His	Leu
				485					490					495	
Arg	Ala	Ser	Tyr	Pro	Ser	Asn	Ile	Pro	Ser	Ser	Pro	Val	Arg	Lys	Pro
			500				505						510		
Ser	Ala	Phe	Gly	Phe	Asp	Ser	Ser	Ala	Ala	Val	Ala	Thr	Ala	Val	Met
		515				520						525			
Asn	Ser	Arg	Ser	Ala	Ala	Phe	Ala	Lys	Arg	Ser	Gln	Ser	Phe	Ile	Asp
	530					535					540				
Arg	Gly	Ala	Ala	Thr	His	His	Leu	Gly	Leu	Ser	Ser	Ala	Ser	Asn	Ser
545					550					555					560
Ser	Cys	Arg	Val	Ser	Ser	Thr	Leu	Ser	Asp	Trp	Ser	Ser	Pro	Thr	Gly
				565					570					575	
Lys	Leu	Asp	Trp	Gly	Val	Asn	Gly	Asp	Lys	Leu	Asn	Lys	Leu	Arg	Lys
			580					585					590		
Ser	Thr	Ser	Phe	Gly	Phe	Arg	Asn	Ser	Gly	Val	Thr	Ala	Ser	Pro	Ile
		595					600					605			
Ala	Gln	Pro	Glu	Phe	Gly	Ala	Glu	Pro	Asp	Val	Ser	Trp	Val	His	Ser
	610					615					620				
Leu	Val	Lys	Asp	Val	Pro	Ser	Glu	Arg	Ser	Glu	Ile	Phe	Gly	Ala	Glu
625					630					635					640
Lys	Gln	Gln	Tyr	Asp	Leu	Ser	Lys	Glu	Met	Leu	Pro	Pro	Trp	Met	Glu
			645						650					655	
Gln	Leu	Tyr	Ile	Glu	Gln	Glu	Gln	Met	Val	Ala					
			660					665							

<210> 32
 <211> 2683
 <212> ADN
 <213> *Eucalyptus grandis*

5

<400> 32

ES 2 451 669 T3

gcaaaggctg	atcacttctt	ccctagaaag	cgagtgtgga	gttgaagctt	gataaccaga	60
ggccgctctt	cgtctcgtct	cgcccgcttg	cgcttgcctt	gctctccgcg	tgccaagggg	120
gtgttctctg	gtgctgaatc	tttccatgtg	tagcggttca	aaaggaaggg	gagagtgaat	180
tcgagaagca	gaggatgtcg	gcccgtcagt	tctcgatcct	gctcgagtta	tctgctgctg	240
atgatctgac	gaactttaag	aaagcagttg	aggaagacgg	ctacgatatt	gatgagtcga	300
gcttgtggta	tggtaggagg	atcgggtcga	agaagattgg	gcttgaagag	agaactcccc	360
tcatgattgc	cgcgatgttc	ggcagtatgt	ccgtgctgga	ttatattatc	aagtctggcc	420
gggccaatgt	aaacaaggcg	tgtggttcag	atggtgctac	cgcgcttcac	tgtgctgctg	480
ctggtggctc	ggtacaatct	cctgaggtgg	tcaagctggt	gcttgattct	tcagcgaatg	540
ctaactccat	tgatgcgaat	gggaaacgag	cgggagactt	gatttctgag	gtctctggtt	600
cgcccttcaa	ttcgagaagg	aagactttgg	atgtcatggt	gactggaggt	gggactgttg	660
agtttgttga	ggaaacttac	aatctgcctg	agaatctggg	tagtcaaatt	gaaggaaacg	720
aacaagaga	gagtccaacg	gcccgcgctt	ccaaggatgg	ttctgaaaag	aaagagtatc	780
ctgtcgacct	ttctcttccg	gacatcaaca	atggaatata	tagcacagat	gagtttagga	840
tgtattcttt	caaagtgaag	ccttgctcga	gagcttactc	tcatgactgg	actgagtgtc	900
catttgttca	ccctggggag	aatgcaagac	ggcgtgacct	acggaaatat	cactacagct	960
gtgtgccttg	ccctgagttc	cgcaaggggt	catgcaggca	aggggatggc	tgcgagtatg	1020
ctcatggtat	atgtgagtgc	tggcttcacc	cagctcaata	tgcaccccg	ctctgtaagg	1080
atgagattgg	atgcaccaga	aaagtctggt	tctttgccca	caaacatgaa	gagcttctgc	1140
cattgtatcg	atcaactggt	tcggcgcttc	tctctctcgg	atcattttcg	cccgttgcctg	1200
cttctctaga	catgggatca	ctgagccctc	tctctctcgg	ttcttcttca	gtccggatac	1260
cgccaacttc	aacaccacct	atgactccat	cagggggctc	ttctcccctt	gggtgggtcga	1320
tgtggaaaag	ccaaattaat	agcactccgc	ctggcttgca	gcttccaggt	agcaggttga	1380
gaagcgcatt	gagtgcctaga	gacatggatt	tagatgttga	cttgatcgat	ctagaaaata	1440
attatcgttt	gcagaagcag	ttgctcgaac	actttcctga	tctgtcctct	cctcgtgggt	1500
ggaacaactc	ttcatccacc	acgtcggctt	tccctgagta	ttcaggtgac	atgactggag	1560
aaataagtag	gttaggagta	aaaccaata	atctcgagga	tagtttcagg	tcattggacc	1620
tgacctctt	gtctcagtta	caagggctgt	cacttgatgg	tgcaatatcc	cagctgcaat	1680
ctcctactgg	aatgaagatt	cggcagaaca	tgaccagca	gctctactca	aactatactg	1740
acaagctttc	ctcgtcacct	agggcaatgc	catcatttgg	aaccgatcct	tccagagctt	1800
cagcagcagc	cactctgagt	tccaggtcat	tggcatttgc	aaaaaggagc	cacagcttca	1860
ttgagcggag	tacagtgaac	agtcagtctg	gatattcagc	aggtgctgct	tctccaactg	1920
caaggatgtc	ttcccagaat	gactggggct	cgcccgatgg	caaactagac	tggggcattc	1980
aaggggagga	gctgaacaag	ctgaggaat	ctgcatcatt	cgggctcagg	agcagcagca	2040
accgcttcca	tgcgtctgca	gattctgcga	cagcaactgt	aggggaccca	gacatgccct	2100
ggattcagtc	cttggcaaa	gaagcccgt	cacaaaaccc	tggcaattht	ggagcagagc	2160
atcagcagca	gcagcagcag	cagcagcagc	agcagtatca	tcttaattct	ggaggtactg	2220
agctgcttcc	agcttgggtg	gagcagttgt	acgcgatca	ggagcagatg	gtcgcctgag	2280
atcaacattg	gcttcttctc	taaccactat	tagtcatctc	gttattgctt	taattttttt	2340
tcttctgagt	ctagtattaa	tgtctaggat	tcgaacgaac	tggaaaatta	aatctagagg	2400
gaagatggga	agaaaagagc	aggatggaag	gtttctgctc	ggtccgagat	ttctcatagt	2460
ctattataga	ctatcgtatt	tctcgttctt	ttccgtccca	atgttcttga	tttggttctc	2520
agcatgtttt	ctggatgagg	cttacaact	atgtaatctt	gtcttgctaa	aagaatcaga	2580
gctgcacctg	caccaaaagg	tgtgatacta	ccgcttattg	atgatgatga	taataataat	2640
aattcggaca	tttagtacca	agtccgatgt	ctcaaaaaaa	aaa		2683

<210> 33
 <211> 694
 <212> PRT
 <213> *Eucalyptus grandis*

<400> 33

ES 2 451 669 T3

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

Lys Ala Cys Gly Ser Asp Gly Ala Thr Ala Leu His Cys Ala Ala Ala
 85 90 95
 Gly Gly Ser Val Gln Ser Pro Glu Val Val Lys Leu Leu Leu Asp Ser
 100 105 110
 Ser Ala Asn Ala Asn Ser Ile Asp Ala Asn Gly Lys Arg Ala Gly Asp
 115 120 125
 Leu Ile Ser Glu Val Ser Gly Ser Pro Phe Asn Ser Arg Arg Lys Thr
 130 135 140
 Leu Asp Val Met Leu Thr Gly Gly Gly Thr Val Glu Phe Val Glu Glu
 145 150 155 160
 Thr Tyr Asn Leu Pro Glu Asn Leu Gly Ser Gln Ile Glu Gly Asn Glu
 165 170 175
 Gln Arg Glu Ser Pro Thr Ala Arg Ala Ser Lys Asp Gly Ser Glu Lys
 180 185 190
 Lys Glu Tyr Pro Val Asp Leu Ser Leu Pro Asp Ile Asn Asn Gly Ile
 195 200 205
 Tyr Ser Thr Asp Glu Phe Arg Met Tyr Ser Phe Lys Val Lys Pro Cys
 210 215 220
 Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro
 225 230 235 240
 Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys Tyr His Tyr Ser Cys
 245 250 255
 Val Pro Cys Pro Glu Phe Arg Lys Gly Ser Cys Arg Gln Gly Asp Gly
 260 265 270
 Cys Glu Tyr Ala His Gly Ile Phe Glu Cys Trp Leu His Pro Ala Gln
 275 280 285
 Tyr Arg Thr Arg Leu Cys Lys Asp Glu Ile Gly Cys Thr Arg Lys Val
 290 295 300
 Cys Phe Phe Ala His Lys His Glu Glu Leu Arg Pro Leu Tyr Ala Ser
 305 310 315 320
 Thr Gly Ser Ala Leu Pro Ser Pro Arg Ser Phe Ser Pro Val Ala Ala
 325 330 335
 Ser Leu Asp Met Gly Ser Leu Ser Pro Leu Ser Leu Gly Ser Ser Ser
 340 345 350
 Val Arg Ile Pro Pro Thr Ser Thr Pro Pro Met Thr Pro Ser Gly Ala
 355 360 365
 Ser Ser Pro Leu Gly Gly Ser Met Trp Lys Ser Gln Ile Asn Ser Thr
 370 375 380
 Pro Pro Gly Leu Gln Leu Pro Gly Ser Arg Leu Arg Ser Ala Leu Ser
 385 390 395 400
 Ala Arg Asp Met Asp Leu Asp Val Asp Leu Ile Asp Leu Glu Asn Asn
 405 410 415
 Tyr Arg Leu Gln Lys Gln Leu Leu Glu His Phe Pro Asp Leu Ser Ser
 420 425 430
 Pro Arg Gly Trp Asn Asn Ser Ser Ser Thr Thr Ser Ala Phe Pro Glu
 435 440 445
 Tyr Ser Gly Asp Met Thr Gly Glu Ile Ser Arg Leu Gly Val Lys Pro
 450 455 460
 Asn Asn Leu Glu Asp Ser Phe Arg Ser Leu Asp Leu Thr Leu Leu Ser
 465 470 475 480
 Gln Leu Gln Gly Leu Ser Leu Asp Gly Ala Ile Ser Gln Leu Gln Ser
 485 490 495
 Pro Thr Gly Met Lys Ile Arg Gln Asn Met Thr Gln Gln Leu Tyr Ser
 500 505 510
 Asn Tyr Thr Asp Lys Leu Ser Ser Ser Pro Arg Ala Met Pro Ser Phe
 515 520 525
 Gly Thr Asp Pro Ser Arg Ala Ser Ala Ala Ala Thr Leu Ser Ser Arg
 530 535 540
 Ser Leu Ala Phe Ala Lys Arg Ser His Ser Phe Ile Glu Arg Ser Thr

ES 2 451 669 T3

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

<210> 34
 <211> 2499
 <212> ADN
 <213> *Arabidopsis thaliana*

5

<400> 34

ES 2 451 669 T3

atthttgacct	taagaagaaa	gtgacaagga	gaggaagaag	aagaaaaaaa	acataatttg	.60
aggaagaaga	aaaaaaattc	ggatttgtht	tttcaataaa	ttgactaatt	gagtactcgt	120
ttaaaggaag	tgaagagcgg	ttttttggta	gtggtggtcg	agaaaagaga	gagtttgct	180
ctgtgactca	gagtgaaatc	aatagagcgg	gaaaagattg	ttgctttttt	ttgccatggg	240
agttgatgag	ctgtctcacc	tcaaattctc	tcttctgcta	gaatcatcag	cctgcaatga	300
tttgtccgg	tttaagtctc	tagttgaaga	agaaggtctt	gagagcattg	atggctctgg	360
tttgtggtat	gggaggagat	taggatcaaa	gaagatgggt	tttgaggaga	ggacgcctct	420
tatgattgct	gccttgthttg	gaagcaaa	ggttgthttg	tacatcatta	gtactggtht	480
tgthtgactg	aaccgctctt	gtggctctga	tggtgccacg	gctcttcact	gtgaggthct	540
tggtctgtct	gccaatagcc	ttgagattgt	tactcttctg	ctgaagggtc	ctgogaatcc	600
ggattcttgt	gatgcttatg	gtaacaagcc	tggagatgtg	atthtccctt	gthttgagthc	660
ggtthtttagc	gcgaggatga	aggtthttgga	gcgtthttg	aaaggaaatg	atgatttgaa	720
tgaagthaat	gggcaagaag	aaagcgagcc	agaggttgag	gthtgagtht	aggtthctgcc	780
tcctcggggg	tctgagagga	aggagthtcc	ggttgatcca	acgcttctct	atatcaagaa	840
cggthtatat	gggacggatg	agthtccgat	gtatgctthc	aagatcaagc	cgtgctctag	900
agcatactct	cacgactgga	cggaatgtcc	ctthgttcat	ccgggtgaga	acgcaaggag	960
gcgtgatccg	aggaagthacc	attatagtht	tgthccctgt	cctgaatthc	ggaagggthc	1020
thgtthccaga	ggtgatactt	gcgagthtgc	tcatgthtgc	thtgagthgt	ggctthcacc	1080
ggctcagthc	cgactcgtc	tctgcaagga	cgagacgaat	tgctcgagaa	gagthttgtht	1140
ctthgcccac	aaacccgagg	agctgctgthc	thtgthaccct	tcaactggat	caggtgthtcc	1200
gtccccgctg	tcttctctct	catcttgcaa	thctctgacc	gctthctgaca	tgggaccgat	1260
tagthcctgt	cctatcggag	caacaaccac	acctctthtt	agthctaaag	gtgthtctct	1320
thccaatagth	ggaggaaaaa	cgtggatgaa	ctggctaac	ataaccctct	ctgcattgca	1380
gctthcagg	agcagatthg	aatctgcat	gaatgcaaga	gaaatcgatt	tctctgaaga	1440
gatgcaaa	cttactthct	caactacatg	gaacaacacg	ccaatgthcat	ctccattctc	1500
cggaaaggcc	atgaacaggc	thgcaaggag	agcaatgagc	ccggthgaata	gtctcagthg	1560
tatgthttgg	acagaggata	atacatcggg	thtgcaatc	cgacgcagcg	thcattaacc	1620
gcagctgcat	thccaacagthc	thtctthcat	acctgthggg	gccaattctc	gtthttctgat	1680
ggattctctc	gcagthcttg	ctthcaagagc	gctgthaat	gctaaacagc	gaagccaaag	1740
cttcatagaa	cgcaacaacg	gactgthtca	ccatcccgca	atctctthcca	tgactacaac	1800
thgtthtaaac	gattggggct	cattgthtgg	gaagctthgac	thggagcgtc	aaggagacga	1860
gctacagaag	ctcagaaaa	ccactthctt	ccgtctcaga	gccggthggca	thggaatcaag	1920
actgcctaac	gaaggthctg	ggctcgaaga	gccagatgthc	tcatggthtgg	agccgctggt	1980
gaaagagcca	caggagacaa	gactagctcc	ggtthtgatg	gagcaatcat	acatggagac	2040
agaacagacc	gtggcttgaa	tcaaaagtht	tgaactthca	thtaaccgthc	cacaagaagc	2100
aaagthcagaa	agattccgag	aggtcagthc	thaatctatth	catthttatth	gththtaagct	2160
thgtthattth	tctthtagaat	aaaaagaaaa	aattctthagg	ggacaaaaaga	gagthtctgtht	2220
gtctctctct	ctgtctccaa	agaaaaacag	aggtgaaaaa	aggtthtcaaa	acctaaagaaa	2280
cctthgaa	cctcacctca	cttctctgat	tctthtactat	thcacaatgag	thaatcgatth	2340
thttthttct	thgthaacact	ctcacgctga	ataatathgt	thttthtagth	thaatataatt	2400
ggaatacaga	aatgthttht	cactthtgth	gthtaggaaa	gtgthtgaat	gtthtctctct	2460
aaagthttht	ctaaagatgth	thgagctata	tctthcgtct			2499

<210> 35

<211> 607

<212> PRT

<213> *Arabidopsis thaliana*

<400> 35

ES 2 451 669 T3

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

ES 2 451 669 T3

			340					345				350			
Pro	Pro	Leu	Ser	Pro	Asn	Gly	Val	Ser	Ser	Pro	Ile	Gly	Gly	Gly	Lys
		355					360					365			
Thr	Trp	Met	Asn	Trp	Pro	Asn	Ile	Thr	Pro	Pro	Ala	Leu	Gln	Leu	Pro
	370					375				380					
Gly	Ser	Arg	Leu	Lys	Ser	Ala	Leu	Asn	Ala	Arg	Glu	Ile	Asp	Phe	Ser
385					390					395					400
Glu	Glu	Met	Gln	Ser	Leu	Thr	Ser	Pro	Thr	Thr	Trp	Asn	Asn	Thr	Pro
			405						410					415	
Met	Ser	Ser	Pro	Phe	Ser	Gly	Lys	Gly	Met	Asn	Arg	Leu	Ala	Gly	Gly
			420					425					430		
Ala	Met	Ser	Pro	Val	Asn	Ser	Leu	Ser	Asp	Met	Phe	Gly	Thr	Glu	Asp
		435					440					445			
Asn	Thr	Ser	Gly	Leu	Gln	Ile	Arg	Arg	Ser	Val	Ile	Asn	Pro	Gln	Leu
	450					455					460				
His	Ser	Asn	Ser	Leu	Ser	Ser	Ser	Pro	Val	Gly	Ala	Asn	Ser	Leu	Phe
465					470					475					480
Ser	Met	Asp	Ser	Ser	Ala	Val	Leu	Ala	Ser	Arg	Ala	Ala	Glu	Phe	Ala
			485					490						495	
Lys	Gln	Arg	Ser	Gln	Ser	Phe	Ile	Glu	Arg	Asn	Asn	Gly	Leu	Asn	His
			500					505					510		
His	Pro	Ala	Ile	Ser	Ser	Met	Thr	Thr	Thr	Cys	Leu	Asn	Asp	Trp	Gly
	515						520					525			
Ser	Leu	Asp	Gly	Lys	Leu	Asp	Trp	Ser	Val	Gln	Gly	Asp	Glu	Leu	Gln
	530					535					540				
Lys	Leu	Arg	Lys	Ser	Thr	Ser	Phe	Arg	Leu	Arg	Ala	Gly	Gly	Met	Glu
545					550					555					560
Ser	Arg	Leu	Pro	Asn	Glu	Gly	Thr	Gly	Leu	Glu	Glu	Pro	Asp	Val	Ser
			565						570					575	
Trp	Val	Glu	Pro	Leu	Val	Lys	Glu	Pro	Gln	Glu	Thr	Arg	Leu	Ala	Pro
			580					585					590		
Val	Trp	Met	Glu	Gln	Ser	Tyr	Met	Glu	Thr	Glu	Gln	Thr	Val	Ala	
		595					600						605		

<210> 36
 <211> 1806
 <212> ADN
 <213> *Oryza sativa*

5

<400> 36

ES 2 451 669 T3

atgtgctctg	ggccgcgcaa	gccgtccaca	ccgcccgtgc	cgcagcagca	gaaggaggcg	60
acggatgatg	cggcgtcctt	gcttcttgag	ctggcggcag	cggacgacgt	ggcggcgggtg	120
aggagggtcg	tggaggagga	gaaggtgtct	cttggcgtgg	ctgggttggtg	gtatgggcct	180
tcggcgagcg	gcgtggcgag	gctcgggatg	gagcggagga	cggcggcgat	ggtggcggcg	240
ctgtacggga	gcacgggggt	gcttgggtat	gtcgtggcgg	cagcgcgggc	ggaggccgcg	300
cgcgcgtcgg	agacgggatg	ggccacgccg	ctgcacatgg	cggctgccgg	tggcgcggcg	360
aacgcggtcg	cggccacgcg	cctgttgctc	gccgcggggg	cgtcggtcga	cgcgctctcg	420
gcttcggggc	tccgcgcggg	tgacctcctc	ccgcgcgcca	ccgcggcggga	gaaggccatc	480
cggctgctgc	tcaagtgcgc	ggcctgtctg	ccgtcgtcgt	cgcggaagaa	gtcggcctcg	540
ccgcgcgtgc	cgccgcggcc	gcaggaggcg	aagaaggagt	accgcctga	cctgacgctg	600
cccgaacctca	agagcggact	gttcagcacc	gacgagttcc	gcatgtacag	cttcaagggtg	660
aagccgtgct	cccgcgccta	ctccatgac	tggaccagat	gccccttcgt	ccaccccggc	720
gagaacgcgc	gccgcgcgca	ccctgcggc	tactccataca	gctgcgtgcc	ttgcccggag	780
ttccgcaagg	gcggctcgtg	ccgcaagggc	gacgcgtgcg	agtacgccca	tggcgtgttc	840
gagtgcgtgg	tccacccggc	gcagtacagg	acgcgcctct	gcaaggacga	ggtcggctgc	900
gcgcgcgcga	tctgcttctt	cgcccacaag	cccgaacgagc	tccgcgcctg	caacccctcc	960
gccgtgtccg	tcggcatgca	gcccaccgta	togtcgcggc	gctcctcgcc	gcccacggg	1020
ctcgacatgg	cggcggcggc	ggcggcgatg	atgagccccg	cctggccgctc	gtcccagcg	1080
agccgcctca	agacggcgct	cggcgcgcgg	gagctcgact	tcgacctcga	gatgctcggc	1140
ctggaccagt	accagcagaa	gctgttcgac	aaggtgtccg	gcgcgcggtc	gccgagggcg	1200
agctggggcg	ccgcggcgaa	cggcctcgcc	accgcgtcgc	cggcggaggc	cgtgccggac	1260
tacaccgacc	tgctcggctc	cgtcgaccgg	gcatgctgt	cccagctcca	cgcgctgtcc	1320
ctcaagcagg	ccggcgacat	gcccgcgtac	agctccatgg	cggacaccac	gcagatgcac	1380
atgccgacct	cgccgatggt	ggcggcgcg	aacaccgcgt	tggggtgga	ccactccatg	1440
gcgaaggcga	tcatgagctc	ccgcgcctcg	gcgttcgcca	agcgcagcca	gagcttcatc	1500
gaccgcggag	gccgcgcccc	ggcggcgcgt	tcgctcatgt	cgcgcggcag	gaccggcgcg	1560
ccgtccattc	tctcggactg	ggctcggccg	gacggcaagc	tggactgggg	cgtccagggc	1620
gacgagctgc	acaagctccg	caagtccggc	togttcggct	tccgcggcca	atccgccatg	1680
ccggtggcga	cgcacgcggc	ggcggcggag	ccggacgtgt	catgggtgaa	ctctcttctc	1740
aaggacggcc	acgccgcggg	cgacatattc	gcgcagtgcc	cggagcagga	gcagatggtg	1800
gcatga						1806

<210> 37
 <211> 601
 <212> PRT
 <213> *Oryza sativa*
 <400> 37

5

ES 2 451 669 T3

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

ES 2 451 669 T3

Ala Val Ser Val Gly Met Gln Pro Thr Val Ser Ser Pro Arg Ser Ser
 325 330 335
 Pro Pro Asn Gly Leu Asp Met Ala Ala Ala Ala Ala Met Met Ser
 340 345 350
 Pro Ala Trp Pro Ser Ser Pro Ala Ser Arg Leu Lys Thr Ala Leu Gly
 355 360 365
 Ala Arg Glu Leu Asp Phe Asp Leu Glu Met Leu Ala Leu Asp Gln Tyr
 370 375 380
 Gln Gln Lys Leu Phe Asp Lys Val Ser Gly Ala Pro Ser Pro Arg Ala
 385 390 395 400
 Ser Trp Gly Ala Ala Ala Asn Gly Leu Ala Thr Ala Ser Pro Ala Arg
 405 410 415
 Ala Val Pro Asp Tyr Thr Asp Leu Leu Gly Ser Val Asp Pro Ala Met
 420 425 430
 Leu Ser Gln Leu His Ala Leu Ser Leu Lys Gln Ala Gly Asp Met Pro
 435 440 445
 Ala Tyr Ser Ser Met Ala Asp Thr Thr Gln Met His Met Pro Thr Ser
 450 455 460
 Pro Met Val Gly Gly Ala Asn Thr Ala Phe Gly Leu Asp His Ser Met
 465 470 475 480
 Ala Lys Ala Ile Met Ser Ser Arg Ala Ser Ala Phe Ala Lys Arg Ser
 485 490 495
 Gln Ser Phe Ile Asp Arg Gly Gly Arg Ala Pro Ala Ala Arg Ser Leu
 500 505 510
 Met Ser Pro Ala Thr Thr Gly Ala Pro Ser Ile Leu Ser Asp Trp Gly
 515 520 525
 Ser Pro Asp Gly Lys Leu Asp Trp Gly Val Gln Gly Asp Glu Leu His
 530 535 540
 Lys Leu Arg Lys Ser Ala Ser Phe Ala Phe Arg Gly Gln Ser Ala Met
 545 550 555 560
 Pro Val Ala Thr His Ala Ala Ala Ala Glu Pro Asp Val Ser Trp Val
 565 570 575
 Asn Ser Leu Val Lys Asp Gly His Ala Ala Gly Asp Ile Phe Ala Gln
 580 585 590
 Trp Pro Glu Gln Glu Gln Met Val Ala
 595 600

<210> 38
 <211> 1692
 <212> ADN
 <213> *Hordeum vulgare*

<400> 38

5

ES 2 451 669 T3

cggcagcagg	cacatccatc	atctaacctc	acctctcctc	tcctcccctc	tcctcctacc	60
aaacccaaaa	ccaagcagag	caagagcaag	agcaagagca	agagcaagca	agcatgtgcc	120
ctggcctgcg	caacctcgcc	gccgccatgc	cacctcgcg	ccaagaccac	ccctcctcct	180
acctgctcga	gctcgccgcc	gacgacgacc	tccccgcctt	ccgcccgcgc	gtccaggagg	240
acaacctctc	cctcgacgcc	gcatccccga	ggtacgagcc	atcccccaaa	tcagaccaac	300
aacaacaaca	cgccccagct	cgcgctccac	ctgcgcaccc	ccgccatggg	cgccgcgctc	360
tacggcagca	ccaccgtcct	ctcctacgtc	ctctccatcg	ccccctccga	ggccgcccgc	420
gcctccgcat	ccgacggcgc	caccccgcct	ctcctcgccc	accagggccg	cgcgccatcc	480
gcgccccacg	ccgcaagcct	cctcctcacc	gacggcgcct	catcgtcctc	cctactcgcg	540
ccccaaagtc	accctctcaa	ccacaaaaac	aaaaccaaa	acagccccac	caagaaagac	600
tcgcccgcgg	actccagcag	gaccaccacc	aagaaggact	actcctccgc	ctccgactcc	660
cagacggagg	acatcaacgc	ggcgctcttc	gccaccgacg	acttccggat	gtacagcttc	720
aagggtgaacc	cggtctcccg	cgcctacacg	cacgactgga	ccgagtgcc	cttcgcccac	780
cccggcgaga	acgcccgcgg	ccgcccaccc	cgcccgcgtg	catactcgtg	cgccccatgc	840
ccggacttcc	gccgcgaccc	ggcccgcctg	cgcaagggcg	acgcctgcga	gtacgcgcac	900
ggcgtcttcg	agtcatggct	ccaccccgcg	cagtaccgca	ccaggctctg	caaggacgag	960

gtcggatgcc	cgcccgcgat	ctgcttcttc	gcgcaaggcg	cccagacagct	acgcccgcgtc	1020
aacccctccg	ccgcatccat	ggactcgcca	tccccaaactt	cctcttcgcc	gccgccaacc	1080
tccaggccgg	ccgcccctcac	cgcgtcgcct	agctcgcggg	acctcgactt	ggacgccgac	1140
aaccaggccc	agtacgcgcg	caggatgatg	atggccaggg	ccaactcccc	gccggactac	1200
tgcgccgacc	tcgtcgcggc	ctacgtacag	gcgctctcct	ccctgcaaca	gcagcagcat	1260
cagcagaacc	agcaacagca	gcatcagcag	cagaaccagc	accagcagca	acatcagcag	1320
aaccagcacc	agcagcatca	gcagcaacat	cagcagagca	tggggatggg	ggggctgagc	1380
gcccgcgccc	ccgccttcac	caaccgcagc	cagaccttcg	tgcaccgctc	tccgtccccg	1440
gtcccgggc	ggtcgttcaa	gtctccggcg	ccgtcgtcca	tgctcgcgga	ctgggggctg	1500
ccggacggga	agctggactg	ggcggtgcag	gccgcggagc	tgcgcaagtc	cacgtctttc	1560
ggagtcaaaa	gcagcagcag	gccgcctcat	gagacgacga	gggcccggga	caacatgtac	1620
ccgtcgtgga	tgaaggacgg	cagcgatatg	ctgctggcgg	cgccgtggtc	ggacctggag	1680
cagatggctg	cc					1692

<210> 39
 <211> 564
 <212> PRT
 <213> *Hordeum vulgare*

5

<400> 39

ES 2 451 669 T3

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

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

<210> 40
 <211> 3610
 <212> ADN
 <213> *Pinus radiata*

5

<400> 40

tgtttccagg	cgggcactaa	agcaagggag	ggggtaggct	ttactttctg	ctctgcgcaa	60
agaacgttga	aatcaatcgc	cctggctggt	ctggcgtgac	tactagattc	aatttcttca	120
tggccgtctt	cacataccca	ttctttaccg	gttcagagct	gtgatcttta	tttttaacag	180
ccacaatcat	ggtttgtggt	tccagtgtta	tgatctgagt	gaagttcggt	ctttttctcg	240
tgaccaggc	ttgatactag	gccggacctt	tctgaggtgg	aagagatcta	tacatttgag	300
gcctatthtg	tgtagccatg	tgtggaggcc	cagaacattt	gaagcctgcc	agcccacacg	360
aaggagaaga	taaagtcaaa	atggccgaga	atcagtctat	caaagtgaag	gaattgtctg	420
aatcttgttc	aagtctacat	gaactagctg	ctaataatga	ccttattggc	tttaagaaag	480
caatggagga	agaaggggtca	aagatagatg	aggttaactt	ttggtacggg	aggcagaatg	540
gttctaataca	gatggctctg	gagcaaagga	ctccattgat	ggttgctgca	ctttatggca	600
gtgtagatgc	gctgagttac	atcttatcca	tttatgtaac	ttgtggagca	gatgttaacc	660
aagcctgtgg	gtcagataac	tccactgcct	tgcattgtgc	ggctgtggga	gggtctgcct	720
gtgcagttga	aactgtaaaa	ttgttacttc	atgcagccag	tgatgtgaat	cgcttggatg	780
cttatggcag	aagaccagca	gatgtgatta	tggtttctcc	taagctaacc	gaaatcaagg	840
ccaagctaga	agaaatgtta	aacgcagctg	gttcatgtca	aacttctccg	gcaaagttgc	900
ctaacatagt	ttcagggcca	cctgggtttg	agtcaaaggg	gatggagtcc	atgtccccat	960
tgccattggt	gcctctttca	ttgtctttag	aagcatccaa	taatagatca	ggttgtgtga	1020
attctccaac	atcttcgcca	aagtccatgg	aagcattaaa	gggtttcggg	gatgttaatg	1080
agaagaagga	atatcctgtg	gacccttctt	ttccagacat	aaagaatagc	atctatacta	1140

10

cagatgaatt tcggatgttt tccttcaagg tgcggccatg ttcacgggca tattctcatg 1200
 attggactga atgcccattt gtgcatcctg gtgaaaatgc cagaaggcgg gatccaagaa 1260
 ggtatcatta tagctgtgtt ccttgcccag attttcggaa agggacttgt aggcgcagtg 1320
 atgtttgtga atatgcacac ggtgtttttg agtgctggtt acatcctgct caatatagga 1380
 cacggttggtg caaagatggg actaattgtt cacgtagagt ttgcttcttt gctcacacat 1440
 ctgaggaact acgccctctc attgtctcta ctgggtctgc tgttccatcc ccaagggcat 1500
 catcatctct ggacatgaca tctgtcatga gtctctctgc ccttggttct cctctctcag 1560
 tttcaatgat gtcacccttc ctatcaaadc ctgagcaagg cagtgtgctt actccgccta 1620
 tgtctccatc agcgtcctct gtaaatggat atggaggctg gccacagcct aatgtaccaa 1680
 ccttacacct tcctggtagc aatgttcaaa ccagccgtct tagagcggaa cttaatgcca 1740
 gagacatgcc tgttgaggat tctcctcgaa tttcagacta tgaagggcag caactcctga 1800
 atgatttttc tccactgtcc acacaagcca ggctgaatgc tgctgctgct gttatatctg 1860
 gtggcgggaa caccacaaca aggtctggaa aatacaagag tcacgggatc aatactgttg 1920
 ctccaacgaa tcttgaagac ttgtttgcct ccgaggtaac atctcctaga gttagcagttc 1980
 ttgaaccttc catcttttct cagatgagtc cccaaatgca agctcataag actgcccagg 2040
 catatatgca gattcaaaac cagatgctgc ctctataaaa tacacaggca ttttcgcagg 2100
 gaattacaca gatgcagcag gctgcaatag agcctcagag ccctggacat tctttgatgc 2160
 aatcaccttt ccaatcttcc tcgtatgggt tgggatcccc tggtagaatg tcacctcgtt 2220
 gtgtggatgt ggaacgtcat aatacatgtg ggtctcctt atcaccggct atggctgcaa 2280
 cgataaattc aagaatggct atggctgctt ttgttcagag ggaaaaacgg agccatagtt 2340
 cccgtgactt gggagctaat gtgaatccca gttcatggtc tgattggggc tcgcctacag 2400
 gtaaagttga ctgggggggt caaggagaag agttgagcaa attaagaaag tcggcttcat 2460
 ttggtccccg cagttatgaa gaaccggatt tgtcttgggt tcaaacactg gtaaaggaaa 2520
 ctacaccaga gggtaaagat ggaggaaatg taagctgttc tggggaaact ccacacaagg 2580
 ggcaaataga aaatgttgat cattcagttt tgggtgcctg gattgaacag atgcagcttg 2640
 atcagattgt agcttgagat taggattatt tatttggagt ggtggtaggg ataggctcat 2700
 ttaaaattca atttctcatt ttttactatt tctttataa aaattcccca ttatagtta 2760
 ggaaatagtc tggttttcta cctattatca gaattacacc tgcaggaaat tttggaggaa 2820
 agcatgcaaa aagtagatag ggatgttatt cctatcagca ggttgacaag ctgaaaatca 2880
 ctgggggtgt agaccagaga atgacactat tttttgttga catggcaact gaagatgctg 2940
 ttttctttac ttatcattaa caaccctata tatatttgtt ttgaaagaac tgagcggaga 3000
 aatgtttgca gttggttact ctgocgaagg ccttggaga aatccaagat gtggcatctt 3060
 ggtgcatttt taatttatca agtgtgaaat ccataacagg tttcagtgag tgacttctga 3120
 ggttgtatat ggaaaaacct atgatgttg ctgtctactg ctatttttct gtgcctaaac 3180
 tgtcaactaa agtttgcagg tggcaatttt gtggcagcat atttgcacat tgaagcggat 3240
 ggtctgcacc tgctatagaa gttttcgagt ctgtagaatt tgatggtgca agatgatatt 3300
 ctagttgata tatttgggaag gctttgcca agtagtggca tgtacatttt gcaaaaattt 3360
 aaaggatggc aatccattgt tttgccatgt agcttactt tattgattag gtggaaagga 3420
 attttgagac acttcaattt gtgcatact ttgttctgaa ctgcaaaatc agtctcttgt 3480
 gatgtcctca agctattat gctcaggat ttgcctaaaa ccataagtg ccttagataa 3540
 ggtaccattg tattaccttt tattgtttgg atattttatt tatgaaagtg aattttattt 3600
 aaaaaaaaaa 3610

<210> 41
 <211> 779
 <212> PRT
 <213> *Pinus radiata*

5

<400> 41

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

10

Leu Glu Gln Arg Thr Pro Leu Met Val Ala Ala Leu Tyr Gly Ser Val
 85 90 95
 Asp Ala Leu Ser Tyr Ile Leu Ser Ile Tyr Val Thr Cys Gly Ala Asp
 100 105 110
 Val Asn Gln Ala Cys Gly Ser Asp Asn Ser Thr Ala Leu His Cys Ala
 115 120 125
 Ala Val Gly Gly Ser Ala Cys Ala Val Glu Thr Val Lys Leu Leu Leu
 130 135 140
 His Ala Gly Ser Asp Val Asn Arg Leu Asp Ala Tyr Gly Arg Arg Pro
 145 150 155 160
 Ala Asp Val Ile Met Val Ser Pro Lys Leu Thr Glu Ile Lys Ala Lys
 165 170 175
 Leu Glu Glu Met Leu Asn Ala Ala Gly Ser Cys Gln Thr Ser Pro Ala
 180 185 190
 Lys Leu Pro Asn Ile Val Ser Gly Pro Pro Gly Phe Glu Ser Lys Gly
 195 200 205
 Met Glu Ser Met Ser Pro Leu Pro Leu Leu Pro Leu Ser Leu Ser Leu
 210 215 220
 Glu Ala Ser Asn Asn Arg Ser Gly Cys Val Asn Ser Pro Thr Ser Ser
 225 230 235 240
 Pro Lys Ser Met Glu Ala Leu Lys Gly Phe Gly Asp Val Asn Glu Lys
 245 250 255
 Lys Glu Tyr Pro Val Asp Pro Ser Phe Pro Asp Ile Lys Asn Ser Ile
 260 265 270
 Tyr Thr Thr Asp Glu Phe Arg Met Phe Ser Phe Lys Val Arg Pro Cys
 275 280 285
 Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro
 290 295 300
 Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Arg Tyr His Tyr Ser Cys
 305 310 315 320
 Val Pro Cys Pro Asp Phe Arg Lys Gly Thr Cys Arg Arg Ser Asp Val
 325 330 335
 Cys Glu Tyr Ala His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln
 340 345 350
 Tyr Arg Thr Arg Leu Cys Lys Asp Gly Thr Asn Cys Ser Arg Arg Val
 355 360 365
 Cys Phe Phe Ala His Thr Ser Glu Glu Leu Arg Pro Leu Ile Val Ser
 370 375 380
 Thr Gly Ser Ala Val Pro Ser Pro Arg Ala Ser Ser Ser Leu Asp Met
 385 390 395 400
 Thr Ser Val Met Ser Pro Leu Ala Pro Gly Ser Pro Ser Ser Val Ser
 405 410 415
 Met Met Ser Pro Phe Leu Ser Asn Pro Gln Gln Gly Ser Val Leu Thr
 420 425 430
 Pro Pro Met Ser Pro Ser Ala Ser Ser Val Asn Gly Tyr Gly Gly Trp
 435 440 445
 Pro Gln Pro Asn Val Pro Thr Leu His Leu Pro Gly Ser Asn Val Gln
 450 455 460
 Thr Ser Arg Leu Arg Ala Glu Leu Asn Ala Arg Asp Met Pro Val Glu
 465 470 475 480
 Asp Ser Pro Arg Ile Ser Asp Tyr Glu Gly Gln Gln Leu Leu Asn Asp
 485 490 495
 Phe Ser Pro Leu Ser Thr Gln Ala Arg Leu Asn Ala Ala Ala Val
 500 505 510
 Ile Ser Gly Gly Gly Asn Thr Thr Thr Arg Ser Gly Lys Tyr Lys Ser
 515 520 525
 His Gly Ile Asn Thr Val Ala Pro Thr Asn Leu Glu Asp Leu Phe Ala
 530 535 540
 Ser Glu Val Thr Ser Pro Arg Val Ala Val Leu Glu Pro Ser Ile Phe

ES 2 451 669 T3

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

<210> 42
 <211> 3610
 <212> ADN
 <213> *Pinus radiata*

5

<400> 42

tgtttccagg	cgggcactaa	agcaagggag	ggggtaggct	ttactttctg	ctotgocgaa	60
agaacgttga	aatcaatcgc	cctggctggg	ctggcgtgac	tactagattc	aatttcttca	120
tgcccgctct	ccatacccca	ttctttaccg	gttcagagct	gtgatcttta	tttttaacag	180
ccacaatcat	ggtttggtgt	tccagtgtaa	tgatctgagt	gaagttcggt	ctttttctcg	240
tgaccacagg	ttgatactag	gccggacctt	tctgaggtgg	aagagatcta	tacatttgag	300
gcctattttg	tgtagccatg	tgtggaggcc	cagaacattt	gaagcctgcc	agccacacag	360
aaggagaaga	taaagtcaaa	atggccgaga	atcagtctat	caaagtgaag	gaattgtctg	420
aatcttggtc	aagtctacat	gaactagctg	ctaataatga	ccttattggc	tttaagaaag	480
caatggagga	agaagggcca	aagatagatg	aggtttaact	ttgggtacggg	aggcagaatg	540
gttctaataca	gatggctctg	gagcaaagga	ctccattgat	ggttgctgca	ctttatggca	600
gtgtagatgc	gctgagttac	atcttatcca	tttatgtaac	ttgtggagca	gatgttaacc	660
aagcctgtgg	gtcagataac	tccactgcct	tgattgtgct	ggctgtggga	gggtctgcct	720
gtgcagttga	aactgtaaaa	ttgttacttc	atgcaggcag	tgatgtgaat	cgcttgatg	780
cttatggcag	aagaccagca	gatgtgatta	tggtttctcc	taagctaacc	gaaatcaagg	840
ccaagctaga	agaaatgtta	aacgcagctg	gttcatgtca	aacttctccg	gcaaagttgc	900
ctaacaatagt	ttcagggcca	cctggggttg	agtcaaaggg	gatggagtcc	atgtccccat	960
tgccattggt	gcctctttca	ttgtctttag	aagcatccaa	taatagatca	ggttgtgtga	1020
attctccaac	atcttcgcca	aagtccatgg	aagcattaaa	gggtttcggt	gatgttaatg	1080
agaagaagga	atatacctgtg	gacccttctt	ttccagacat	aaagaatagc	atctatacta	1140
cagatgaatt	tcggatgttt	tccttcaagg	tgccggccatg	ttcacgggca	tattctcatg	1200
attggactga	atgcccattt	gtgcatcctg	gtgaaaatgc	cagaaggcgg	gatccaagaa	1260
ggatcatta	tagctgtggt	ccttgcccag	attttcggaa	agggacttgt	aggcgcagtg	1320
atgtttgtga	atatgcacac	ggtgtttttg	agtgctggtt	acatctctgct	caatatagga	1380

10

ES 2 451 669 T3

```

cacggttgtg caaagatggg actaattggt cacgtagagt ttgcttcttt gctcacacat 1440
ctgaggaact acgcctctc attgtctcta ctgggtctgc tgttccatcc ccaagggcat 1500
catcatctct ggacatgaca tctgtcatga gtctcttgc ccctggttct ccctcttcag 1560
tttcaatgat gtcacccttc ctatcaaadc ctcagcaagg cagtgtgctt actccgccta 1620
tgtctccatc agcgtcctct gtaaattgat atggaggctg gccacagcct aatgtaccaa 1680
ccttacacct tcttggtagc aatgttcaaa ccagccgtct tagagcggaa cttaatgcca 1740
gagacatgcc tgttgaggat tctcctcgaa tttcagacta tgaagggcag caactcctga 1800
atgatttttc tccactgtcc acacaagcca ggctgaatgc tgctgctgct gttatatctg 1860
gtggcgggaa caccacaaca aggtctggaa aatacaagag tcacgggatc aatactgttg 1920
ctccaacgaa tcttgaagac ttgtttgctt cccaaatgca atctcctaga gttagcagttc 1980
ttgaaccttc catcttttct cagatgagtc cccaaatgca agctcataag actgcccagg 2040
catatatgca gattcaaaac cagatgctgc ctctataaaa tacacaggca ttttcgcagg 2100
gaattacaca gatgcagcag gctgcaatag agcctcagag ccctggacat tctttgatgc 2160
aatcaccttt ccaatcttcc tcgtatgggt tgggatcccc tggtagaatg tcacctcgtt 2220
gtgtggatgt ggaacgtcat aatacatgtg ggtctccctt atcaccggct atggctgcaa 2280
cgataaattc aagaatggct atggctgctt ttgttcagag ggaaaaacgg agccatagtt 2340
cccgtagctt gggagctaata gtgaatccca gttcatggct tgattggggc tcgcctacag 2400
gtaaagtcca ctgggggggt caaggagaag agttgagcaa attaagaaag tcggcttcat 2460
ttggtccccg cagttatgaa gaaccggatt tgccttgggt tcaaacactg gtaaaaggaaa 2520
ctacaccaga gggtaaagat ggaggaaatg taagctgctt tggggaaact ccacacaagg 2580
ggcaaataga aaatgttgat cattcagttt tgggtgcctg gattgaacag atgcagcttg 2640
atcagattgt agcttgagat taggattatt tatttggagt ggtggtaggg ataggctcat 2700
ttaaatttca atttctcatt ttttactatt tcttttataa aaattcccca ttatagttta 2760
ggaaatagtc tggtttttcta cctattatca gaattacacc tgcaggaaat tttggaggaa 2820
agcatgcaaaa aagtagatag ggatgttatt cctatcagca ggttgacaag ctgaaaaatca 2880
cttgggtggg agaccagaga atgacactat tttttgttga catggcaact gaagatgctg 2940
tttcttttac ttatcattaa caaccctata tatatttgtt ttgaaagaac tgagcggaga 3000
aatgtttgca gttggttact ctgcgcaagg ccttgggaaga aatccaagat gtggcatctt 3060
ggtgcatttt taatttatca agtgtgaaat ccataacagg tttcagttag tgacttctga 3120
ggtgtatat ggaaaaacct atgatgttgg ctgtctactg ctatttttct gtgcctaaac 3180
tgtcaactaa agtttgcagg tggcaatttt gtggcagcat atttgcacat tgaagcggat 3240
ggtctgcacc tgctatagaa gttttcgagt ctgtagaatt tgatggtgca agatgatttt 3300
ctagttgata tatttggaa gctttgccaa agtagtggca tgtacatttt gcaaaaaattt 3360
aaaggatggc aatccattgt tttgccatgt agcttcactt tattgattag gtggaaagga 3420
atttgagac acttcaattt gtgcatactt ttgttctgaa ctgcaaaatc agtctcttgt 3480
gatgtcctca aggtattat gctcagggat ttgcctaaaa ccataagtgg ccttagataa 3540
ggtaccattg tattaccttt tattgtttgg atattttatt tatgaaagtg aatttatttt 3600
aaaaaaaaaa 3610

```

<210> 43
 <211> 749
 <212> PRT
 <213> *Pinus radiata*

5

<400> 43

```

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

```

10

ES 2 451 669 T3

Leu Ile Lys Ala Gly Ala Asp Val Asn Phe Met Asn Ala Gly Gly Arg
 115 120 125
 Lys Pro Ala Asp Val Ile Met Val Ser Pro Lys Leu Ala His Phe Lys
 130 135 140
 Asn Val Leu Glu Asp Leu Leu Ile Met Gly Ser Asn Ser Pro Met Lys
 145 150 155 160
 Ile Pro Cys Arg Val Ser Gly Ser Gly Phe Tyr Leu Pro Glu Gly Gly
 165 170 175
 Gly Cys Phe Phe Asp Glu His Gly Cys Val Val Ser Val Pro Thr Ser
 180 185 190
 Ser Pro Leu Phe Ser Ser Pro Asp Ala Thr Ser Pro Ala Thr Val Asn
 195 200 205
 Ser Pro Leu Ser Ser Pro Pro Thr Ser Leu Asp Thr Pro Lys Asn Leu
 210 215 220
 Cys Asp Cys Gly Gln Lys Lys Glu Phe Ala Val Asp Ser Ser Leu Pro
 225 230 235 240
 Asp Ile Lys Asn Ser Ile Tyr Ser Thr Asp Glu Phe Arg Met Tyr Ser
 245 250 255
 Phe Lys Val Arg Pro Cys Ser Arg Ala Tyr Ser His Asp Trp Thr Glu
 260 265 270
 Cys Pro Phe Val His Pro Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg
 275 280 285
 Lys Tyr His Tyr Ser Cys Val Pro Cys Pro Asp Phe Arg Lys Gly Ala
 290 295 300
 Cys Arg Arg Gly Asp Val Cys Glu Tyr Ala His Gly Val Phe Glu Cys
 305 310 315 320
 Trp Leu His Pro Ala Gln Tyr Arg Thr Arg Leu Cys Lys Asp Gly Thr
 325 330 335
 Asn Cys Ser Arg Arg Val Cys Phe Phe Ala His Thr Pro Glu Glu Leu
 340 345 350
 Arg Pro Leu Tyr Pro Pro Ala Cys Ser Ser Met Leu Ser Gln Arg Thr
 355 360 365
 Thr Met Thr Ser Ser Asp Lys Met Ala Val Met His Pro Leu Ala Pro
 370 375 380
 Gly Ser Ala Ser Ser Val Leu Met Met Ser Ser Ser Asn Ser Ser Gln
 385 390 395 400
 Ser Ser Phe Pro Asn Ser Pro Val Ser Pro Leu Ser Ser Ala Asn Thr
 405 410 415
 Ser Ser His Ser Ser Phe Gly Gly Gly Ser Trp Ala His Pro Asn Leu
 420 425 430
 Pro Thr Leu His Leu Ser Asn Gly Ala Leu Gln Ala Ser Arg Leu Arg
 435 440 445
 Thr Ala Val Asn Ala Arg Asp Met His Pro Asp Cys Ser Ile Glu Ser
 450 455 460
 Gly Asp Tyr Glu Gly Gln Leu Leu Asn Glu Phe Ala Tyr Leu Ser Thr
 465 470 475 480
 Gln Ala Arg Gly Asn Gly Pro Met Ala Thr Val Ser Ser Ser Gly Asn
 485 490 495
 Thr Pro Cys Arg Pro Arg Lys Phe Arg Ala His Asn Val Ala Pro Thr
 500 505 510
 Asn Leu Glu Asp Leu Phe Ala Ser Glu Val Phe Ser Pro Lys Met Thr
 515 520 525
 Ala Ser Glu Ser Ala Phe Leu Ser Glu Ile Gln Ser His Lys Ser Ala
 530 535 540
 Gln Leu Ser Pro Gln Leu Gln Ser Gln Met Leu Ser Ser Phe Asn Thr
 545 550 555 560
 Gln Val Tyr Pro Gln Gly Ser Thr Gln Gly Gln Met His Met Gln His
 565 570 575
 Gly Gly Val Asp Cys Gln Ser Pro Ser Val Phe Leu Ser Pro Pro Pro

ES 2 451 669 T3

```

                    580                                585                                590
Val Gln Leu Ala Ser Tyr Ser Leu Ser Ser Leu Gly Pro Leu Ser Ser
                    595                                600                                605
Leu Thr Gly Glu Leu Glu Arg Gln Asn Ser Asn Gly Ser Pro Leu Ser
    610                                615                                620
Pro Ile Met Ser Thr Ala Ala Asp Ser Arg Ala Val Ala Phe Ser Gln
    625                                630                                635
Arg Asp Lys Gly Ser Ser Arg Ser Gly Asp Leu Gly Gly Ala Thr Thr
                    645                                650                                655
Trp Ser Glu Trp Gly Ser Pro Thr Gly Lys Val Asn Trp Gly Ile Arg
    660                                665                                670
Gly Glu Glu Leu Gln Lys Phe Arg Lys Ser Ala Ser Phe Gly Ile Arg
    675                                680                                685
Ser Ser Asp Glu Pro Asp Leu Ser Trp Val Gln Lys Leu Phe Lys Glu
    690                                695                                700
Ala Pro Met Glu Ser Met Asp Arg Gly Thr Met Gly Arg Ser Met Asp
    705                                710                                715
Ile Ala Asn Ser Val Gln Met Glu Ala Thr Asp Leu Gly Gly Trp Ile
                    725                                730                                735
Ser Gln Ile Asn Pro Asp Gln Val Ala Pro Leu Thr Leu
                    740                                745

```

<210> 44
 <211> 711
 <212> PRT
 <213> *Glycine max*

5

<400> 44

```

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

```

10

ES 2 451 669 T3

				245					250				255		
Ser	His	Asp	Trp	Thr	Glu	Cys	Pro	Phe	Val	His	Pro	Gly	Glu	Asn	Ala
			260					265					270		
Arg	Arg	Arg	Asp	Pro	Arg	Lys	Phe	His	Tyr	Ser	Cys	Val	Pro	Cys	Pro
		275					280					285			
Asp	Phe	Arg	Lys	Gly	Ala	Cys	Arg	Arg	Gly	Asp	Met	Cys	Glu	Tyr	Ala
	290					295					300				
His	Gly	Val	Phe	Glu	Cys	Trp	Leu	His	Pro	Ala	Gln	Tyr	Arg	Thr	Arg
305					310					315					320
Leu	Cys	Lys	Asp	Gly	Thr	Ser	Cys	Asn	Arg	Arg	Val	Cys	Phe	Phe	Ala
				325					330					335	
His	Thr	Ala	Glu	Glu	Leu	Arg	Pro	Leu	Tyr	Val	Ser	Thr	Gly	Ser	Ala
			340					345					350		
Val	Pro	Ser	Pro	Arg	Ser	Ser	Ala	Ser	Ala	Pro	Asn	Val	Met	Asp	Met
		355					360						365		
Ala	Ala	Ala	Met	Ser	Leu	Leu	Pro	Gly	Ser	Pro	Ser	Ser	Val	Ser	Ser
	370					375						380			
Met	Ser	Pro	Ser	His	Phe	Gly	Gln	Pro	Met	Ser	Pro	Ser	Ala	Asn	Gly
385					390					395					400
Met	Ser	Leu	Ser	Ser	Ala	Trp	Ala	Gln	Pro	Asn	Val	Ser	Ala	Leu	His
				405					410					415	
Leu	Pro	Gly	Ser	Asn	Leu	Gln	Ser	Ser	Arg	Leu	Arg	Ser	Ser	Leu	Ser
			420					425					430		
Ala	Arg	Asp	Met	Pro	Pro	Asp	Asp	Leu	Asn	Met	Met	Ser	Asp	Leu	Asp
		435					440					445			
Gly	Gln	Gln	Gln	His	Pro	Leu	Asn	Asp	Leu	Ser	Cys	Tyr	Leu	Gln	Pro
	450					455					460				
Arg	Pro	Gly	Ala	Gly	Ser	Val	Ser	Arg	Ser	Gly	Arg	Ser	Lys	Ile	Leu
465					470					475					480
Thr	Pro	Ser	Asn	Leu	Glu	Asp	Leu	Phe	Ser	Ala	Glu	Ile	Ser	Ser	Ser
				485					490					495	
Pro	Arg	Tyr	Ser	Asp	Pro	Ala	Ala	Gly	Ser	Val	Phe	Ser	Pro	Thr	His
			500					505					510		
Lys	Ser	Ala	Val	Leu	Asn	Gln	Phe	Gln	Gln	Leu	Gln	Ser	Met	Leu	Ser
		515					520					525			
Pro	Ile	Asn	Thr	Asn	Leu	Leu	Ser	Pro	Lys	Asn	Val	Glu	His	Pro	Leu
	530					535					540				
Leu	Gln	Ala	Ser	Phe	Gly	Val	Ser	Pro	Ser	Gly	Arg	Met	Ser	Pro	Arg
545					550					555					560
Ser	Val	Glu	Pro	Ile	Ser	Pro	Met	Ser	Ser	Arg	Ile	Ser	Ala	Phe	Ala
				565					570					575	
Gln	Arg	Glu	Lys	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Leu	Arg	Ser	Leu	Ser
			580					585					590		
Ser	Arg	Asp	Leu	Gly	Ala	Asn	Ser	Pro	Ala	Ser	Leu	Val	Gly	Ser	Pro
		595					600					605			
Ala	Asn	Pro	Trp	Ser	Lys	Trp	Gly	Ser	Pro	Asn	Gly	Lys	Ala	Asp	Trp
	610					615					620				
Ser	Val	Asn	Gly	Asp	Thr	Leu	Gly	Arg	Gln	Met	Arg	Arg	Ser	Ser	Ser
625					630					635					640
Phe	Glu	Leu	Lys	Asn	Asn	Gly	Glu	Glu	Pro	Asp	Leu	Ser	Trp	Val	Gln
			645						650					655	
Ser	Leu	Val	Lys	Glu	Ser	Pro	Pro	Glu	Met	Ile	Lys	Glu	Lys	Phe	Ala
			660					665					670		
Ser	Pro	Met	Pro	Thr	Ala	Ser	Ala	Asp	Gly	Pro	Asn	Ser	Asn	Ser	Gln
	675						680					685			
Ile	Glu	Ser	Ile	Asp	His	Ser	Val	Leu	Gly	Ala	Trp	Leu	Glu	Gln	Met
	690					695					700				
Gln	Leu	Asp	Gln	Leu	Val	Val									
705					710										

ES 2 451 669 T3

<210> 45
 <211> 643
 <212> PRT
 <213> *Glycine max*

5

<400> 45

Arg Gly Ser Gly Phe Pro Gly Arg Pro Thr Arg Pro Arg Ser Gly Arg
 1 5 10 15
 Thr Arg Gly Arg Thr Arg Gly Val Asn Phe Ala Cys Gly Ala Asn Lys
 20 25 30
 Thr Thr Ala Leu His Cys Ala Ala Ser Gly Ala Ser Thr Lys Ala Val
 35 40 45
 Asp Ala Val Lys Leu Leu Leu Ser Ala Gly Ala Asp Val Asn Cys Val
 50 55 60
 Asp Ala Asn Gly Asn Arg Pro Ile Asp Val Ile Ala Val Pro Pro Lys
 65 70 75 80
 Leu Gln Gly Ala Lys Ala Val Leu Glu Glu Leu Leu Ser Asp Asn Ala
 85 90 95
 Ser Asp Val Ser Val Gly Glu Phe Ser Val Pro Val Ser Val Asn Ser
 100 105 110
 Ser Ser Pro Gly Ser Pro Ala His Ser Ser Asn Gly Met Pro Tyr Thr
 115 120 125
 Pro Ser Val Ser Pro Pro Ser Pro Val Ala Ala Lys Phe Thr Asp Ala
 130 135 140
 Ala Ile Cys Ser Leu Ser Glu Lys Ala Arg Glu Tyr Pro Ile Asp Pro
 145 150 155 160
 Ser Leu Pro Asp Ile Lys Asn Ser Ile Tyr Ala Thr Asp Glu Phe Arg
 165 170 175
 Met Phe Ser Phe Lys Val Arg Pro Cys Ser Arg Ala Tyr Ser His Asp
 180 185 190
 Trp Thr Glu Cys Pro Phe Val His Pro Gly Glu Asn Ala Arg Arg Arg
 195 200 205
 Asp Pro Arg Lys Phe His Tyr Ser Cys Val Pro Cys Pro Asp Phe Arg
 210 215 220
 Lys Gly Ala Cys Arg Arg Gly Asp Met Cys Glu Tyr Ala His Gly Val
 225 230 235 240
 Phe Glu Cys Trp Leu His Pro Ala Gln Tyr Arg Thr Arg Leu Cys Lys
 245 250 255
 Asp Gly Thr Ser Cys Asn Arg Arg Val Cys Phe Phe Ala His Thr Ala
 260 265 270
 Glu Glu Leu Arg Pro Leu Tyr Val Ser Thr Gly Ser Ala Ala Pro Ser
 275 280 285
 Pro Arg Ser Ser Ala Ser Gly Pro Asn Val Met Asp Met Ala Ala Ala
 290 295 300
 Met Ser Leu Phe Pro Gly Ser Pro Ser Ser Gly Ser Ser Ile Ser Leu
 305 310 315 320
 Ser Ile Ser Phe Ser Leu Asp Pro Met Ser Pro Ser Ala Asn Gly Met
 325 330 335
 Pro Leu Ser Ser Ala Trp Ala Gln Pro Asn Val Pro Ala Leu His Leu
 340 345 350
 Pro Gly Ser Asn Leu Gln Ser Ser Arg Leu Arg Ser Ser Leu Ser Ala
 355 360 365
 Arg Asp Ile Pro Pro Glu Asp Leu Asn Met Met Ser Asp Leu Asp Gly
 370 375 380
 Gln Gln Gln His His Leu Asn Asp Leu Ser Cys Tyr Ile Gln Pro Arg
 385 390 395 400
 Pro Gly Ala Ser Ser Val Ser Arg Ser Gly Arg Ser Lys Thr Leu Thr
 405 410 415

ES 2 451 669 T3

Pro Ser Asn Leu Glu Glu Leu Phe Ser Ala Glu Ile Ser Leu Ser Pro
 420 425 430
 Arg Tyr Ser Asp Pro Ala Ala Gly Ser Val Phe Ser Pro Thr His Lys
 435 440 445
 Ser Ala Val Leu Asn Gln Phe Gln Gln Leu Gln Ser Met Leu Ser Pro
 450 455 460
 Ile Asn Thr Asn Leu Leu Ser Pro Lys Asn Val Glu His Pro Leu Phe
 465 470 475 480
 Gln Ala Ser Phe Gly Val Ser Pro Ser Gly Arg Met Ser Pro Arg Ser
 485 490 495
 Val Glu Pro Ile Ser Pro Met Ser Ala Arg Leu Ser Ala Phe Ala Gln
 500 505 510
 Arg Glu Lys Gln Gln Gln Gln Leu Arg Ser Val Ser Ser Arg Asp Leu
 515 520 525
 Gly Ala Asn Ser Pro Ala Ser Leu Val Gly Ser Pro Ala Asn Pro Trp
 530 535 540
 Ser Lys Trp Gly Ser Pro Ile Gly Lys Ala Asp Trp Ser Val Asn Gly
 545 550 555 560
 Asp Ser Leu Gly Arg Gln Met Arg Arg Ser Ser Ser Phe Glu Arg Lys
 565 570 575
 Asn Asn Gly Glu Glu Pro Asp Leu Ser Trp Val Gln Ser Leu Val Lys
 580 585 590
 Glu Ser Pro Pro Glu Met Ile Lys Glu Lys Phe Ala Ser Pro Met Pro
 595 600 605
 Thr Ala Ser Ala Asp Gly Pro Asn Ser Asn Ser Gln Ile Glu Ser Ile
 610 615 620
 Asp His Ser Val Leu Gly Ala Trp Leu Glu Gln Met Gln Leu Asp Gln
 625 630 635 640
 Leu Val Val

<210> 46
 <211> 669
 <212> PRT
 <213> *Glycine max*

5

<400> 46

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

10

ES 2 451 669 T3

Ala Glu Asp Leu Lys Met Lys Thr Asn Glu Val Ser Glu Lys Lys Glu
180 185 190
Tyr Pro Val Asp Leu Ser Leu Pro Asp Ile Lys Asn Ser Ile Tyr Ser
195 200 205
Ser Asp Glu Phe Arg Met Tyr Ser Phe Lys Val Arg Pro Cys Ser Arg
210 215 220
Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro Gly Glu
225 230 235 240
Asn Ala Arg Arg Arg Asp Pro Arg Lys Phe His Tyr Ser Cys Val Pro
245 250 255
Cys Pro Glu Phe Arg Lys Gly Ala Cys Arg Arg Gly Asp Met Cys Glu
260 265 270
Tyr Ala His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln Tyr Arg
275 280 285
Thr Arg Leu Cys Lys Asp Gly Thr Asn Cys Ala Arg Arg Val Cys Phe
290 295 300
Phe Ala His Thr Asn Glu Glu Leu Arg Pro Leu Tyr Val Ser Thr Gly
305 310 315 320
Ser Ala Val Pro Ser Pro Arg Ser Ser Ala Ser Ser Ala Met Asp Phe
325 330 335
Val Ala Ala Ile Ser Pro Ser Ser Met Ser Val Met Ser Pro Ser Pro
340 345 350
Phe Thr Pro Pro Met Ser Pro Ser Ser Ala Ser Ile Ala Trp Pro Gln
355 360 365
Pro Asn Ile Pro Ala Leu His Leu Pro Gly Ser Asn Phe His Ser Ser
370 375 380
Arg Leu Arg Ser Ser Leu Asn Ala Arg Asp Phe Ser Val Asp Asp Phe
385 390 395 400
Asp Leu Leu Leu Pro Asp Tyr Asp His His His Gln Gln Gln Gln
405 410 415
Gln Gln Phe Leu Asn Glu Leu Ser Cys Leu Ser Pro His Ala Met Asn
420 425 430
Cys Asn Thr Met Asn Arg Ser Gly Arg Met Lys Pro Leu Thr Pro Ser
435 440 445
Asn Leu Asp Asp Leu Phe Ser Ala Glu Ser Ser Ser Pro Arg Tyr Ala
450 455 460
Asp Pro Ala Leu Ala Ser Ala Val Phe Ser Pro Thr His Lys Ser Ala
465 470 475 480
Val Phe Asn Gln Phe Gln His Gln Gln Ser Met Leu Ala Pro Leu Asn
485 490 495
Thr Asn Phe Ala Ser Lys Asn Phe Glu His Pro Leu Leu Gln Ala Ser
500 505 510
Leu Gly Met Ser Pro Arg Asn Val Glu Pro Ile Ser Pro Met Gly Ser
515 520 525
Arg Ile Ser Met Leu Ala Gln Arg Glu Lys Gln Gln Phe Arg Ser Leu
530 535 540
Ser Phe Arg Glu Leu Gly Ser Asn Ser Ala Ala Ala Ser Ala Asp Ser
545 550 555 560
Trp Ser Lys Trp Gly Ser Pro Asn Val Lys Leu Asp Trp Pro Val Gly
565 570 575
Ala Gly Glu Val Gly Lys Leu Arg Arg Ser Ser Ser Phe Glu Leu Gly
580 585 590
Asn Asn Gly Glu Glu Pro Asp Leu Ser Trp Val Gln Ser Leu Val Lys
595 600 605
Glu Ser Pro Ala Glu Val Lys Asp Lys Leu Ala Thr Thr Val Ser Tyr
610 615 620
Val Ala Ala Ala Ala Ala Gly Ser Ser Ser Glu Gly Ser Asn Ile Ser
625 630 635 640
Thr Gln Met Glu Ser Val Val Asp His Ala Val Leu Gly Ala Trp Leu

ES 2 451 669 T3

645 650 655
 Glu Gln Met Gln Leu Asp His Leu Val Ala Gln Gln Asn
 660 665

<210> 47
 <211> 580
 <212> PRT
 <213> *Arabidopsis thaliana*
 <400> 47

5

Met Cys Ser Gly Pro Lys Ser Asn Leu Cys Ser Ser Arg Thr Leu Thr
 1 5 10 15
 Glu Ile Glu Ser Arg Gln Lys Glu Glu Glu Thr Met Leu Leu Leu Glu
 20 25 30
 Phe Ala Ala Cys Asp Asp Leu Asp Ser Phe Lys Arg Glu Val Glu Glu
 35 40 45
 Lys Gly Leu Asp Leu Asp Glu Ser Gly Leu Trp Tyr Cys Arg Arg Val
 50 55 60
 Gly Ser Lys Lys Met Gly Leu Glu Glu Arg Thr Pro Leu Met Val Ala
 65 70 75 80
 Ala Met Tyr Gly Ser Ile Lys Val Leu Thr Phe Ile Val Ser Thr Gly
 85 90 95
 Lys Ser Asp Val Asn Arg Ala Cys Gly Glu Glu Arg Val Thr Pro Leu
 100 105 110
 His Cys Ala Val Ala Gly Cys Ser Val Asn Met Ile Glu Val Ile Asn
 115 120 125
 Val Leu Leu Asp Ala Ser Ala Leu Val Asn Ser Val Asp Ala Asn Gly
 130 135 140
 Asn Gln Pro Leu Asp Val Phe Val Arg Val Ser Arg Phe Val Ala Ser
 145 150 155 160
 Pro Arg Arg Lys Ala Val Glu Leu Leu Leu Arg Gly Gly Gly Val Gly
 165 170 175
 Gly Leu Ile Asp Glu Ala Val Glu Glu Glu Ile Lys Ile Val Ser Lys
 180 185 190
 Tyr Pro Ala Asp Ala Ser Leu Pro Asp Ile Asn Glu Gly Val Tyr Gly
 195 200 205
 Ser Asp Glu Phe Arg Met Tyr Ser Phe Lys Val Lys Pro Cys Ser Arg
 210 215 220
 Ala Tyr Ser His Asp Trp Thr Glu Cys Ala Phe Val His Pro Gly Glu
 225 230 235 240
 Asn Ala Arg Arg Arg Asp Pro Arg Lys Tyr Pro Tyr Thr Cys Val Pro
 245 250 255
 Cys Pro Glu Phe Arg Lys Gly Ser Cys Pro Lys Gly Asp Ser Cys Glu
 260 265 270
 Tyr Ala His Gly Val Phe Glu Ser Trp Leu His Pro Ala Gln Tyr Lys
 275 280 285
 Thr Arg Leu Cys Lys Asp Glu Thr Gly Cys Ala Arg Lys Val Cys Phe
 290 295 300
 Phe Ala His Lys Arg Glu Glu Met Arg Pro Val Asn Ala Ser Thr Gly
 305 310 315 320
 Ser Ala Val Ala Gln Ser Pro Phe Ser Ser Leu Glu Met Met Pro Gly
 325 330 335
 Leu Ser Pro Leu Ala Tyr Ser Ser Gly Val Ser Thr Pro Pro Val Ser
 340 345 350
 Pro Met Ala Asn Gly Val Pro Ser Ser Pro Arg Asn Gly Gly Ser Trp
 355 360 365
 Gln Asn Arg Val Asn Thr Leu Thr Pro Pro Ala Leu Gln Leu Asn Gly
 370 375 380
 Gly Ser Arg Leu Lys Ser Thr Leu Ser Ala Arg Asp Ile Asp Met Glu

10

ES 2 451 669 T3

```

385                390                395                400
Met Glu Met Glu Leu Arg Leu Arg Gly Phe Gly Asn Asn Val Glu Glu
                405                410                415
Thr Phe Gly Ser Tyr Val Ser Ser Pro Ser Arg Asn Ser Gln Met Gly
                420                425                430
Gln Asn Met Asn Gln His Tyr Pro Ser Ser Pro Val Arg Gln Pro Pro
                435                440                445
Ser Gln His Gly Phe Glu Ser Ser Ala Ala Ala Val Ala Val Met
                450                455                460
Lys Ala Arg Ser Thr Ala Phe Ala Lys Arg Ser Leu Ser Phe Lys Pro
465                470                475                480
Ala Thr Gln Ala Ala Pro Gln Ser Asn Leu Ser Asp Trp Gly Ser Pro
                485                490                495
Asn Gly Lys Leu Glu Trp Gly Met Lys Gly Glu Glu Leu Asn Lys Met
                500                505                510
Arg Arg Ser Val Ser Phe Gly Ile His Gly Asn Asn Asn Asn Ala
                515                520                525
Ala Arg Asp Tyr Arg Asp Glu Pro Asp Val Ser Trp Val Asn Ser Leu
530                535                540
Val Lys Asp Ser Thr Val Val Ser Glu Arg Ser Phe Gly Met Asn Glu
545                550                555                560
Arg Val Arg Ile Met Ser Trp Ala Glu Gln Met Tyr Arg Glu Lys Glu
                565                570                575
Gln Thr Val Val
                580

```

<210> 48
 <211> 719
 <212> PRT
 <213> *Arabidopsis thaliana*

5

<400> 48

```

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

```

10

ES 2 451 669 T3

210						215						220			
Gly	Ser	Pro	Leu	Gly	Ser	Pro	Phe	Lys	Leu	Lys	Ser	Thr	Glu	Phe	Lys
225						230				235					240
Lys	Glu	Tyr	Pro	Val	Asp	Pro	Ser	Leu	Pro	Asp	Ile	Lys	Asn	Ser	Ile
				245						250					255
Tyr	Ala	Thr	Asp	Glu	Phe	Arg	Met	Tyr	Ser	Phe	Lys	Val	Arg	Pro	Cys
			260					265					270		
Ser	Arg	Ala	Tyr	Ser	His	Asp	Trp	Thr	Glu	Cys	Pro	Phe	Val	His	Pro
		275					280					285			
Gly	Glu	Asn	Ala	Arg	Arg	Arg	Asp	Pro	Arg	Lys	Phe	His	Tyr	Ser	Cys
	290					295					300				
Val	Pro	Cys	Pro	Asp	Phe	Arg	Lys	Gly	Ala	Cys	Arg	Arg	Gly	Asp	Met
305					310					315					320
Cys	Glu	Tyr	Ala	His	Gly	Val	Phe	Glu	Cys	Trp	Leu	His	Pro	Ala	Gln
				325						330					335
Tyr	Arg	Thr	Arg	Leu	Cys	Lys	Asp	Gly	Thr	Gly	Cys	Ala	Arg	Arg	Val
			340					345						350	
Cys	Phe	Phe	Ala	His	Thr	Pro	Glu	Glu	Leu	Arg	Pro	Leu	Tyr	Ala	Ser
		355					360						365		
Thr	Gly	Ser	Ala	Val	Pro	Ser	Pro	Arg	Ser	Asn	Ala	Asp	Tyr	Ala	Ala
	370					375					380				
Ala	Leu	Ser	Leu	Leu	Pro	Gly	Ser	Pro	Ser	Gly	Val	Ser	Val	Met	Ser
385					390					395					400
Pro	Leu	Ser	Pro	Ser	Ala	Ala	Gly	Asn	Gly	Met	Ser	His	Ser	Asn	Met
				405					410					415	
Ala	Trp	Pro	Gln	Pro	Asn	Val	Pro	Ala	Leu	His	Leu	Pro	Gly	Ser	Asn
			420					425					430		
Leu	Gln	Ser	Ser	Arg	Leu	Arg	Ser	Ser	Leu	Asn	Ala	Arg	Asp	Ile	Pro
		435					440						445		
Thr	Asp	Glu	Phe	Asn	Met	Leu	Ala	Asp	Tyr	Glu	Gln	Gln	Gln	Leu	Leu
	450					455					460				
Asn	Glu	Tyr	Ser	Asn	Ala	Leu	Ser	Arg	Ser	Gly	Arg	Met	Lys	Ser	Met
465					470					475					480
Pro	Pro	Ser	Asn	Leu	Glu	Asp	Leu	Phe	Ser	Ala	Glu	Gly	Ser	Ser	Ser
			485						490						495
Pro	Arg	Phe	Thr	Asp	Ser	Ala	Leu	Ala	Ser	Ala	Val	Phe	Ser	Pro	Thr
			500					505					510		
His	Lys	Ser	Ala	Val	Phe	Asn	Gln	Phe	Gln	Gln	Gln	Gln	Gln	Gln	Gln
		515				520							525		
Gln	Gln	Gln	Ser	Met	Leu	Ser	Pro	Ile	Asn	Thr	Ser	Phe	Ser	Ser	Pro
	530					535					540				
Lys	Ser	Val	Asp	His	Ser	Leu	Phe	Ser	Gly	Gly	Gly	Arg	Met	Ser	Pro
545					550					555					560
Arg	Asn	Val	Val	Glu	Pro	Ile	Ser	Pro	Met	Ser	Ala	Arg	Val	Ser	Met
			565						570					575	
Leu	Ala	Gln	Cys	Val	Lys	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln
			580					585						590	
Gln	Gln	Gln	His	Gln	Phe	Arg	Ser	Leu	Ser	Ser	Arg	Glu	Leu	Arg	Thr
		595					600					605			
Asn	Ser	Ser	Pro	Ile	Val	Gly	Ser	Pro	Val	Asn	Asn	Asn	Thr	Trp	Ser
	610					615						620			
Ser	Lys	Trp	Gly	Ser	Ser	Asn	Gly	Gln	Pro	Asp	Trp	Gly	Met	Ser	Ser
625					630					635					640
Glu	Ala	Leu	Gly	Lys	Leu	Arg	Ser	Ser	Ser	Ser	Phe	Asp	Gly	Asp	Glu
			645						650					655	
Pro	Asp	Val	Ser	Trp	Val	Gln	Ser	Leu	Val	Lys	Glu	Thr	Pro	Ala	Glu
			660					665					670		
Ala	Lys	Glu	Lys	Ala	Ala	Thr	Ser	Ser	Ser	Gly	Glu	His	Val	Met	Lys
		675					680						685		

ES 2 451 669 T3

Gln Pro Asn Pro Val Glu Pro Val Met Asp His Ala Gly Leu Glu Ala
 690 695 700
 Trp Ile Glu Gln Met Gln Leu Asp Gln Leu Val Ala Gln Gln Asn
 705 710 715

<210> 49
 <211> 686
 <212> PRT
 <213> *Arabidopsis thaliana*
 <400> 49

5

Met Asn His Ile Thr Val Glu Thr Glu Asp Thr Phe Ala Ser Leu Leu
 1 5 10 15
 Glu Leu Ala Ala Asn Asn Asp Val Glu Gly Val Arg Leu Ser Ile Glu
 20 25 30
 Arg Asp Pro Ser Cys Val Asp Glu Ala Gly Leu Trp Tyr Gly Arg Gln
 35 40 45
 Lys Gly Ser Lys Ala Met Val Asn Asp Tyr Arg Thr Pro Leu Met Val
 50 55 60
 Ala Ala Thr Tyr Gly Ser Ile Asp Val Ile Lys Leu Ile Val Ser Leu
 65 70 75 80
 Thr Asp Ala Asp Val Asn Arg Ala Cys Gly Asn Asp Gln Thr Thr Ala
 85 90 95
 Leu His Cys Ala Ala Ser Gly Gly Ala Val Asn Ala Ile Gln Val Val
 100 105 110
 Lys Leu Leu Leu Ala Ala Gly Ala Asp Leu Asn Leu Leu Asp Ala Glu
 115 120 125
 Gly Gln Arg Ala Gly Asp Val Ile Val Val Pro Pro Lys Leu Glu Gly
 130 135 140
 Val Lys Leu Met Leu Gln Glu Leu Leu Ser Ala Asp Gly Ser Ser Thr
 145 150 155 160
 Ala Glu Arg Asn Leu Arg Val Val Thr Asn Val Pro Asn Arg Ser Ser
 165 170 175
 Ser Pro Cys His Ser Pro Thr Gly Glu Asn Gly Gly Ser Gly Ser Gly
 180 185 190
 Ser Pro Leu Gly Ser Pro Phe Lys Leu Lys Ser Thr Glu Phe Lys Lys
 195 200 205
 Glu Tyr Pro Val Asp Pro Ser Leu Pro Asp Ile Lys Asn Ser Ile Tyr
 210 215 220
 Ala Thr Asp Glu Phe Arg Met Tyr Ser Phe Lys Val Arg Pro Cys Ser
 225 230 235 240
 Arg Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Val His Pro Gly
 245 250 255
 Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys Phe His Tyr Ser Cys Val
 260 265 270
 Pro Cys Pro Asp Phe Arg Lys Gly Ala Cys Arg Arg Gly Asp Met Cys
 275 280 285
 Glu Tyr Ala His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln Tyr
 290 295 300
 Arg Thr Arg Leu Cys Lys Asp Gly Thr Gly Cys Ala Arg Arg Val Cys
 305 310 315 320
 Phe Phe Ala His Thr Pro Glu Glu Leu Arg Pro Leu Tyr Ala Ser Thr
 325 330 335
 Gly Ser Ala Val Pro Ser Pro Arg Ser Asn Ala Asp Tyr Ala Ala Ala
 340 345 350
 Leu Ser Leu Leu Pro Gly Ser Pro Ser Gly Val Ser Val Met Ser Pro
 355 360 365
 Leu Ser Pro Ser Ala Ala Gly Asn Gly Met Ser His Ser Asn Met Ala
 370 375 380

10

ES 2 451 669 T3

Trp Pro Gln Pro Asn Val Pro Ala Leu His Leu Pro Gly Ser Asn Leu
 385 390 395 400
 Gln Ser Ser Arg Leu Arg Ser Ser Leu Asn Ala Arg Asp Ile Pro Thr
 405 410 415
 Asp Glu Phe Asn Met Leu Ala Asp Tyr Glu Gln Gln Gln Leu Leu Asn
 420 425 430
 Glu Tyr Ser Asn Ala Leu Ser Arg Ser Gly Arg Met Lys Ser Met Pro
 435 440 445
 Pro Ser Asn Leu Glu Asp Leu Phe Ser Ala Glu Gly Ser Ser Ser Pro
 450 455 460
 Arg Phe Thr Asp Ser Ala Leu Ala Ser Ala Val Phe Ser Pro Thr His
 465 470 475 480
 Lys Ser Ala Val Phe Asn Gln Phe Gln Gln Gln Gln Gln Gln Gln Gln
 485 490 495
 Gln Gln Ser Met Leu Ser Pro Ile Asn Thr Ser Phe Ser Ser Pro Lys
 500 505 510
 Ser Val Asp His Ser Leu Phe Ser Gly Gly Gly Arg Met Ser Pro Arg
 515 520 525
 Asn Val Val Glu Pro Ile Ser Pro Met Ser Ala Arg Val Ser Met Leu
 530 535 540
 Ala Gln Cys Val Lys Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
 545 550 555 560
 Gln Gln His Gln Phe Arg Ser Leu Ser Ser Arg Glu Leu Arg Thr Asn
 565 570 575
 Ser Ser Pro Ile Val Gly Ser Pro Val Asn Asn Asn Thr Trp Ser Ser
 580 585 590
 Lys Trp Gly Ser Ser Asn Gly Gln Pro Asp Trp Gly Met Ser Ser Glu
 595 600 605
 Ala Leu Gly Lys Leu Arg Ser Ser Ser Ser Phe Asp Gly Asp Glu Pro
 610 615 620
 Asp Val Ser Trp Val Gln Ser Leu Val Lys Glu Thr Pro Ala Glu Ala
 625 630 635 640
 Lys Glu Lys Ala Ala Thr Ser Ser Ser Gly Glu His Val Met Lys Gln
 645 650 655
 Pro Asn Pro Val Glu Pro Val Met Asp His Ala Gly Leu Glu Ala Trp
 660 665 670
 Ile Glu Gln Met Gln Leu Asp Gln Leu Val Ala Gln Gln Asn
 675 680 685

<210> 50
 <211> 633
 <212> PRT
 <213> *Arabidopsis thaliana*

 <400> 50

5

ES 2 451 669 T3

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

Arg Val Val Thr Asn Val Pro Asn Arg Ser Ser Ser Pro Cys His Ser
 115 120 125
 Pro Thr Gly Glu Asn Gly Gly Ser Gly Ser Gly Ser Pro Leu Gly Ser
 130 135 140
 Pro Phe Lys Leu Lys Ser Thr Glu Phe Lys Lys Glu Tyr Pro Val Asp
 145 150 155 160
 Pro Ser Leu Pro Asp Ile Lys Asn Ser Ile Tyr Ala Thr Asp Glu Phe
 165 170 175
 Arg Met Tyr Ser Phe Lys Val Arg Pro Cys Ser Arg Ala Tyr Ser His
 180 185 190
 Asp Trp Thr Glu Cys Pro Phe Val His Pro Gly Glu Asn Ala Arg Arg
 195 200 205
 Arg Asp Pro Arg Lys Phe His Tyr Ser Cys Val Pro Cys Pro Asp Phe
 210 215 220
 Arg Lys Gly Ala Cys Arg Arg Gly Asp Met Cys Glu Tyr Ala His Gly
 225 230 235 240
 Val Phe Glu Cys Trp Leu His Pro Ala Gln Tyr Arg Thr Arg Leu Cys
 245 250 255
 Lys Asp Gly Thr Gly Cys Ala Arg Arg Val Cys Phe Phe Ala His Thr
 260 265 270
 Pro Glu Glu Leu Arg Pro Leu Tyr Ala Ser Thr Gly Ser Ala Val Pro
 275 280 285
 Ser Pro Arg Ser Asn Ala Asp Tyr Ala Ala Ala Leu Ser Leu Leu Pro
 290 295 300
 Gly Ser Pro Ser Gly Val Ser Val Met Ser Pro Leu Ser Pro Ser Ala
 305 310 315 320
 Ala Gly Asn Gly Met Ser His Ser Asn Met Ala Trp Pro Gln Pro Asn
 325 330 335
 Val Pro Ala Leu His Leu Pro Gly Ser Asn Leu Gln Ser Ser Arg Leu
 340 345 350
 Arg Ser Ser Leu Asn Ala Arg Asp Ile Pro Thr Asp Glu Phe Asn Met
 355 360 365
 Leu Ala Asp Tyr Glu Gln Gln Gln Leu Leu Asn Glu Tyr Ser Asn Ala
 370 375 380
 Leu Ser Arg Ser Gly Arg Met Lys Ser Met Pro Pro Ser Asn Leu Glu
 385 390 395 400
 Asp Leu Phe Ser Ala Glu Gly Ser Ser Ser Pro Arg Phe Thr Asp Ser
 405 410 415
 Ala Leu Ala Ser Ala Val Phe Ser Pro Thr His Lys Ser Ala Val Phe
 420 425 430
 Asn Gln Phe Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Ser Met Leu
 435 440 445
 Ser Pro Ile Asn Thr Ser Phe Ser Ser Pro Lys Ser Val Asp His Ser
 450 455 460
 Leu Phe Ser Gly Gly Gly Arg Met Ser Pro Arg Asn Val Val Glu Pro
 465 470 475 480
 Ile Ser Pro Met Ser Ala Arg Val Ser Met Leu Ala Gln Cys Val Lys
 485 490 495
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln His Gln Phe
 500 505 510
 Arg Ser Leu Ser Ser Arg Glu Leu Arg Thr Asn Ser Ser Pro Ile Val
 515 520 525
 Gly Ser Pro Val Asn Asn Asn Thr Trp Ser Ser Lys Trp Gly Ser Ser
 530 535 540
 Asn Gly Gln Pro Asp Trp Gly Met Ser Ser Glu Ala Leu Gly Lys Leu
 545 550 555 560
 Arg Ser Ser Ser Ser Phe Asp Gly Asp Glu Pro Asp Val Ser Trp Val
 565 570 575
 Gln Ser Leu Val Lys Glu Thr Pro Ala Glu Ala Lys Glu Lys Ala Ala

ES 2 451 669 T3

			580					585					590		
Thr	Ser	Ser	Ser	Gly	Glu	His	Val	Met	Lys	Gln	Pro	Asn	Pro	Val	Glu
			595					600				605			
Pro	Val	Met	Asp	His	Ala	Gly	Leu	Glu	Ala	Trp	Ile	Glu	Gln	Met	Gln
	610					615						620			
Leu	Asp	Gln	Leu	Val	Ala	Gln	Gln	Asn							
625					630										

<210> 51
 <211> 678
 <212> PRT
 <213> *Arabidopsis thaliana*

 <400> 51

5

ES 2 451 669 T3

Met Asn Asp Ala Ala Glu Trp Glu His Ser Phe Ser Ala Leu Leu Glu
 1 5 10 15
 Phe Ala Ala Asp Asn Asp Val Glu Gly Phe Arg Arg Gln Leu Ser Asp
 20 25 30
 Val Ser Cys Ile Asn Gln Met Gly Leu Trp Tyr Arg Arg Gln Arg Phe
 35 40 45
 Val Arg Arg Met Val Leu Glu Gln Arg Thr Pro Leu Met Val Ala Ser
 50 55 60
 Leu Tyr Gly Ser Leu Asp Val Val Lys Phe Ile Leu Ser Phe Pro Glu
 65 70 75 80
 Ala Glu Leu Asn Leu Ser Cys Gly Pro Asp Lys Ser Thr Ala Leu His
 85 90 95
 Cys Ala Ala Ser Gly Ala Ser Val Asn Ser Leu Asp Val Val Lys Leu
 100 105 110
 Leu Leu Ser Val Gly Ala Asp Pro Asn Ile Pro Asp Ala His Gly Asn
 115 120 125
 Arg Pro Val Asp Val Leu Val Val Ser Pro His Ala Pro Gly Leu Arg
 130 135 140
 Thr Ile Leu Glu Glu Ile Leu Lys Lys Asp Glu Ile Ile Ser Glu Asp
 145 150 155 160
 Leu His Ala Ser Ser Ser Ser Leu Gly Ser Ser Phe Arg Ser Leu Ser
 165 170 175
 Ser Ser Pro Asp Asn Gly Ser Ser Leu Ser Leu Asp Ser Val Ser
 180 185 190
 Ser Pro Thr Lys Pro His Gly Thr Asp Val Thr Phe Ala Ser Glu Lys
 195 200 205
 Lys Glu Tyr Pro Ile Asp Pro Ser Leu Pro Asp Ile Lys Ser Gly Ile
 210 215 220
 Tyr Ser Thr Asp Glu Phe Arg Met Phe Ser Phe Lys Ile Arg Pro Cys
 225 230 235 240
 Ser Arg Ala Tyr Ser His Asp Trp Thr Glu Cys Pro Phe Ala His Pro
 245 250 255
 Gly Glu Asn Ala Arg Arg Arg Asp Pro Arg Lys Phe His Tyr Thr Cys
 260 265 270
 Val Pro Cys Pro Asp Phe Lys Lys Gly Ser Cys Lys Gln Gly Asp Met
 275 280 285
 Cys Glu Tyr Ala His Gly Val Phe Glu Cys Trp Leu His Pro Ala Gln
 290 295 300
 Tyr Arg Thr Arg Leu Cys Lys Asp Gly Met Gly Cys Asn Arg Arg Val
 305 310 315 320
 Cys Phe Phe Ala His Ala Asn Glu Glu Leu Arg Pro Leu Tyr Pro Ser
 325 330 335
 Thr Gly Ser Gly Leu Pro Ser Pro Arg Ala Ser Ser Ala Val Ser Ala
 340 345 350
 Ser Thr Met Asp Met Ala Ser Val Leu Asn Met Leu Pro Gly Ser Pro

ES 2 451 669 T3

		355					360					365			
Ser	Ala	Ala	Gln	His	Ser	Phe	Thr	Pro	Pro	Ile	Ser	Pro	Ser	Gly	Asn
	370					375					380				
Gly	Ser	Met	Pro	His	Ser	Ser	Met	Gly	Trp	Pro	Gln	Gln	Asn	Ile	Pro
385					390					395					400
Ala	Leu	Asn	Leu	Pro	Gly	Ser	Asn	Ile	Gln	Leu	Ser	Arg	Leu	Arg	Ser
				405					410					415	
Ser	Leu	Asn	Ala	Arg	Asp	Ile	Pro	Ser	Glu	Gln	Leu	Ser	Met	Leu	His
			420						425				430		
Glu	Phe	Glu	Met	Gln	Arg	Gln	Leu	Ala	Gly	Asp	Met	His	Ser	Pro	Arg
		435					440					445			
Phe	Met	Asn	His	Ser	Ala	Arg	Pro	Lys	Thr	Leu	Asn	Pro	Ser	Asn	Leu
	450					455					460				
Glu	Glu	Leu	Phe	Ser	Ala	Glu	Val	Ala	Ser	Pro	Arg	Phe	Ser	Asp	Gln
465					470					475					480
Leu	Ala	Val	Ser	Ser	Val	Leu	Ser	Pro	Ser	His	Lys	Ser	Ala	Leu	Leu
				485					490					495	
Asn	Gln	Leu	Gln	Asn	Asn	Lys	Gln	Ser	Met	Leu	Ser	Pro	Ile	Lys	Thr
			500					505					510		
Asn	Leu	Met	Ser	Ser	Pro	Lys	Asn	Val	Glu	Gln	His	Ser	Leu	Leu	Gln
		515					520					525			
Gln	Ala	Ser	Ser	Pro	Arg	Gly	Gly	Glu	Pro	Ile	Ser	Pro	Met	Asn	Ala
	530					535					540				
Arg	Met	Lys	Gln	Gln	Leu	His	Ser	Arg	Ser	Leu	Ser	Ser	Arg	Asp	Phe
545					550					555					560
Gly	Ser	Ser	Leu	Pro	Arg	Asp	Leu	Met	Pro	Thr	Asp	Ser	Gly	Ser	Pro
				565					570					575	
Leu	Ser	Pro	Trp	Ser	Ser	Trp	Asp	Gln	Thr	His	Gly	Ser	Lys	Val	Asp
			580					585					590		
Trp	Ser	Val	Gln	Ser	Asp	Glu	Leu	Gly	Arg	Leu	Arg	Lys	Ser	His	Ser
		595				600						605			
Leu	Ala	Asn	Asn	Pro	Asn	Arg	Glu	Ala	Asp	Val	Ser	Trp	Ala	Gln	Gln
	610					615					620				
Met	Leu	Lys	Asp	Ser	Ser	Ser	Pro	Arg	Asn	Gly	Asn	Arg	Val	Val	Asn
625					630					635					640
Met	Asn	Gly	Ala	Arg	Pro	Leu	Thr	Gln	Gly	Gly	Ser	Ser	Val	Asn	Pro
				645					650					655	
His	Asn	Ser	Asp	Thr	Arg	Glu	Ser	Asp	Ile	Leu	Asp	Ala	Trp	Leu	Glu
			660					665					670		
Gln	Leu	His	Leu	Asp	Arg										
		675													

<210> 52
 <211> 640
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 52

ES 2 451 669 T3

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

ES 2 451 669 T3

				85					90					95	
Val	Val	Ser	Pro	His	Ala	Pro	Gly	Leu	Arg	Thr	Ile	Leu	Glu	Glu	Ile
			100					105					110		
Leu	Lys	Lys	Asp	Glu	Ile	Ile	Ser	Glu	Asp	Leu	His	Ala	Ser	Ser	Ser
		115					120					125			
Ser	Leu	Gly	Ser	Ser	Phe	Arg	Ser	Leu	Ser	Ser	Ser	Pro	Asp	Asn	Gly
	130					135					140				
Ser	Ser	Leu	Leu	Ser	Leu	Asp	Ser	Val	Ser	Ser	Pro	Thr	Lys	Pro	His
145						150					155				160
Gly	Thr	Asp	Val	Thr	Phe	Ala	Ser	Glu	Lys	Lys	Glu	Tyr	Pro	Ile	Asp
				165				170						175	
Pro	Ser	Leu	Pro	Asp	Ile	Lys	Ser	Gly	Ile	Tyr	Ser	Thr	Asp	Glu	Phe
			180					185					190		
Arg	Met	Phe	Ser	Phe	Lys	Ile	Arg	Pro	Cys	Ser	Arg	Ala	Tyr	Ser	His
		195					200					205			
Asp	Trp	Thr	Glu	Cys	Pro	Phe	Ala	His	Pro	Gly	Glu	Asn	Ala	Arg	Arg
	210					215					220				
Arg	Asp	Pro	Arg	Lys	Phe	His	Tyr	Thr	Cys	Val	Pro	Cys	Pro	Asp	Phe
225						230				235					240
Lys	Lys	Gly	Ser	Cys	Lys	Gln	Gly	Asp	Met	Cys	Glu	Tyr	Ala	His	Gly
				245					250					255	
Val	Phe	Glu	Cys	Trp	Leu	His	Pro	Ala	Gln	Tyr	Arg	Thr	Arg	Leu	Cys
			260					265					270		
Lys	Asp	Gly	Met	Gly	Cys	Asn	Arg	Arg	Val	Cys	Phe	Phe	Ala	His	Ala
		275					280					285			
Asn	Glu	Glu	Leu	Arg	Pro	Leu	Tyr	Pro	Ser	Thr	Gly	Ser	Gly	Leu	Pro
	290					295					300				
Ser	Pro	Arg	Ala	Ser	Ser	Ala	Val	Ser	Ala	Ser	Thr	Met	Asp	Met	Ala
305						310					315				320
Ser	Val	Leu	Asn	Met	Leu	Pro	Gly	Ser	Pro	Ser	Ala	Ala	Gln	His	Ser
				325					330					335	
Phe	Thr	Pro	Pro	Ile	Ser	Pro	Ser	Gly	Asn	Gly	Ser	Met	Pro	His	Ser
		340						345					350		
Ser	Met	Gly	Trp	Pro	Gln	Gln	Asn	Ile	Pro	Ala	Leu	Asn	Leu	Pro	Gly
		355					360					365			
Ser	Asn	Ile	Gln	Leu	Ser	Arg	Leu	Arg	Ser	Ser	Leu	Asn	Ala	Arg	Asp
	370					375					380				
Ile	Pro	Ser	Glu	Gln	Leu	Ser	Met	Leu	His	Glu	Phe	Glu	Met	Gln	Arg
385						390				395					400
Gln	Leu	Ala	Gly	Asp	Met	His	Ser	Pro	Arg	Phe	Met	Asn	His	Ser	Ala
			405						410					415	
Arg	Pro	Lys	Thr	Leu	Asn	Pro	Ser	Asn	Leu	Glu	Glu	Leu	Phe	Ser	Ala
			420					425					430		
Glu	Val	Ala	Ser	Pro	Arg	Phe	Ser	Asp	Gln	Leu	Ala	Val	Ser	Ser	Val
		435					440					445			
Leu	Ser	Pro	Ser	His	Lys	Ser	Ala	Leu	Leu	Asn	Gln	Leu	Gln	Asn	Asn
	450					455					460				
Lys	Gln	Ser	Met	Leu	Ser	Pro	Ile	Lys	Thr	Asn	Leu	Met	Ser	Ser	Pro
465						470				475					480
Lys	Asn	Val	Glu	Gln	His	Ser	Leu	Leu	Gln	Ala	Ser	Ser	Pro	Arg	
				485					490					495	
Gly	Gly	Glu	Pro	Ile	Ser	Pro	Met	Asn	Ala	Arg	Met	Lys	Gln	Gln	Leu
			500					505					510		
His	Ser	Arg	Ser	Leu	Ser	Ser	Arg	Asp	Phe	Gly	Ser	Ser	Leu	Pro	Arg
		515					520					525			
Asp	Leu	Met	Pro	Thr	Asp	Ser	Gly	Ser	Pro	Leu	Ser	Pro	Trp	Ser	Ser
	530					535					540				
Trp	Asp	Gln	Thr	His	Gly	Ser	Lys	Val	Asp	Trp	Ser	Val	Gln	Ser	Asp
545					550					555					560

ES 2 451 669 T3

Glu Leu Gly Arg Leu Arg Lys Ser His Ser Leu Ala Asn Asn Pro Asn
 565 570 575
 Arg Glu Ala Asp Val Ser Trp Ala Gln Met Leu Lys Asp Ser Ser
 580 585 590
 Ser Pro Arg Asn Gly Asn Arg Val Val Asn Met Asn Gly Ala Arg Pro
 595 600 605
 Leu Thr Gln Gly Gly Ser Ser Val Asn Pro His Asn Ser Asp Thr Arg
 610 615 620
 Glu Ser Asp Ile Leu Asp Ala Trp Leu Glu Gln Leu His Leu Asp Arg
 625 630 635 640

<210> 53
 <211> 2158
 <212> ADN
 <213> *Oryza sativa*

5

<400> 53

```

tgctccttct tcattgcaaa gaaagcacca ctttttaaag aatcctcctc acactccatt      60
cttcttaaaa aacccaccac aacacaattc ccacttgttt cttcatcctc acctacttca      120
atcaaaaaac attccaactt tcttctcaat ttcattccag gatagtatta tgtgcagtgg      180
accaaagagc aatctctgct cttcaagaac cttaacagaa atcgaatcaa ggcaaaagga      240
agaagaaaca atgcttctcc tcgaattcgc tgcttggat gatcttgact cgttcaagag      300
agaggttgaa gagaaagggc ttgatttgga tgagtcaggg ttatggatt gcagacgtgt      360
cggttctaag aagatgggtc ttgaagaaag aacacctta atggttgacg ctatgtatgg      420
aagcataaag gtttgactt tcatcgtttc cactggaaaa tctgatgtga acagagcttg      480
tgggtgaagag agagttactc cgcttcactg tgcgttgct ggctgttctg tgaatatgat      540
tgaagtcatc aatgtcttgc ttgatgcttc tgctttggtt aactctgttg atgctaattg      600
gaatcaacct ttggatgtgt ttggtcagat ttcgaggttt gtggctagtc cgaggaggaa      660
agcggttgag ttgttgctga gaggaggagg tggtggagga ttgatcgatg aggcggttga      720
agaagagatc aagattgtct ctaagtatcc agctgatgct tctttaccgg atataaacga      780
aggggtttat ggaagtgatg agtttaggat gtatagcttt aaggttaagc catgttctag      840
ggcttattct catgattgga ccgagtggtc tttgttcat ccgggagaaa atgcgaggag      900
gagagatccg aggaagtatc cttacacttg tgcacctgt cccgagttcc gtaaaggatc      960
atgcccgaaa ggagattctt gcgagtatgc tcacggggtt ttcgagtcgt ggcttcacc      1020
cgcgcagtat aaaaccggc tttgtaaaga tgaacgggt tgtgcaagga aagtttgttt      1080
ctttgctcat aaacgcgaag agatgagacc tgtaaatgct tcaactggct ctgccgtggc      1140
tcagtctccg tttagcagct tggagatgat gccagggttg tctcctcttg cttattcttc      1200
aggagtctcg actcctccgg tttctccaat ggctaattgg gttccttctc ctccaagaaa      1260
cggcggatca tggcagaaca gagtcaatac cttactcca ccggctttgc agctcaatgg      1320
tggagcaga ttgaagtcca cactgagcgc tagagatata gatatggaga tggagatgga      1380
attgagactc cgcggttttg gcaacaatgt ggaagagacg ttcgggtctt atgttctctc      1440
tccaagttag aattctcaaa tgggtcaaaa catgaaccaa cattatccat cttccccggt      1500
gagacaaccg ccatctcaac acgggttoga atcttcagca gotgcagcgg ttgcagtgat      1560
gaaagcgaga tcaaccgctt ttgcgaaacg tagcttgagc ttcaaaccag ctactcaagc      1620
agcaccacag tcgaatctct cggattgggg atctccaaac ggggaagctgg aatggggaat      1680
gaaaggagaa gagctgaata agatgagaag aagtgtttcc tttggaatcc atggaacaaa      1740
caacaataac gcagctagag actacagggg cgagccagat gtgtcatggg ttaactcttt      1800
agttaaagac agtactgtgg tgtctgagag aagctttgga atgaatgaga gggttcggat      1860
aatgtcgtgg gctgagcaaa tgtacagaga gaaggagcag actgtggtgt aaacacacac      1920
aaagatggtt tcttatatat attgcttttg gccatctct gcaaatttga ttctttaatt      1980
tttgtgactt tctttagttg ttactgttat tagtagtata tggtttgttg tcaactacgag      2040
tctacgtgat gaaaagatag aagttaattg cattagtttc tatattcgtt tctcatcctc      2100
ttgtaattta tcaaacctatg aaatggctaa gcaatccaaa ccgaaaaaaa aaaaaaaa      2158
    
```

10

<210> 54
 <211> 2193
 <212> ADN
 <213> *Oryza sativa*

15

<400> 54

ES 2 451 669 T3

aatccgaaaa	gtttctgcac	cgttttcacc	ccctaactaa	caatataggg	aacgtgtgct	60
aaatataaaa	tgagacctta	tatatgtagc	gctgataact	agaactatgc	aagaaaaact	120
catccacctt	cttttagtggc	aatcgggcta	aataaaaaag	agtcgctaca	ctagtttctgt	180
tttccttagt	aattaagtgg	gaaaaatgaaa	tcattattgc	ttagaatata	cgttcacatc	240
tctgtcatga	agttaaatta	ttcgaggtag	ccataattgt	catcaaactc	ttcttgaata	300
aaaaaatctt	tctagctgaa	ctcaatgggt	aaagagagag	atTTTTTTT	aaaaaataga	360
atgaagatat	tctgaacgta	ttggcaaaga	tttaaacata	taattatata	atTTTatagt	420
ttgtgcattc	gtcatatcgc	acatcattaa	ggacatgtct	tactccatcc	caatTTTTat	480
ttagtaatta	aagacaattg	acttattttt	attattttatc	TTTTTctgat	tagatgcaag	540
gtacttacgc	acacactttg	tgctcatgtg	catgtgtgag	tgcacctcct	caatacacgt	600
tcaactagca	acacatctct	aatatcactc	gcctatttaa	tacatttagg	tagcaatatc	660
tgaattcaag	cactccacca	tcaccagacc	actTTTtaata	atatctaaaa	tacaaaaaat	720
aatTTTtacag	aatagcatga	aaagtatgaa	acgaactatt	taggtTTTTc	acatacaaaa	780
aaaaaaagaa	TTTTgctcgt	gcgcgagcgc	caatctccca	tattgggcac	acaggcaaca	840
acagagtggc	tgcccacaga	acaaccacaca	aaaaacgatg	atctaacgga	ggacagcaag	900
tccgcaacaa	cTTTTtaaca	gcaggctttg	cggccaggag	agaggaggag	aggcaaaaga	960
aaccaagcat	cctcctcctc	ccatctataa	attcctcccc	cTTTTcccc	tctctatata	1020
ggagggcatcc	aagccaagaa	gagggagagc	accaaggaca	cgcgactagc	agaagccgag	1080
cgaccgcctt	cttcgatcca	tatcttccgg	tcgagttctt	ggtcgatctc	ttccctcctc	1140
cacctcctcc	tcacagggta	tgtgcccttc	ggttgttctt	ggatttattg	ttctaggttg	1200
tgtagtacgg	gcgttgatgt	taggaaaggg	gatctgtatc	tgtgatgatt	cctgttcttg	1260
gatttgggat	agaggggttc	ttgatgttgc	atgttatcgg	ttcggtttga	ttagtagtat	1320
ggttttcaat	cgtctggaga	gctctatgga	aatgaaatgg	tttaggttac	ggaatcttgc	1380
gattttgtga	gtaccttttg	tttgaggtaa	aatcagagca	ccggtgattt	tgcttgggtg	1440
aataaaaagta	cggttgtttg	gtcctcgatt	ctggtagtga	tgcttctcga	tttgacgaag	1500
ctatcctttg	tttattccct	attgaacaaa	aataatccaa	ctttgaagac	ggtcccgttg	1560
atgagattga	atgattgatt	cttaagcctg	tccaaaattt	cgcagctggc	ttgttttagat	1620
acagtagtcc	ccatcacgaa	attcatggaa	acagttataa	tcctcaggaa	caggggattc	1680
cctgttcttc	cgatttgctt	tagtcccaga	atTTTTtttc	ccaaatatct	taaaaagtca	1740
ctttctgggt	cagttcaatg	aattgattgc	tacaaataat	gcttttatag	cgttatccta	1800
gctgtagttc	agttaatagg	taatacccct	atagtttagt	caggagaaga	acttatccga	1860
tttctgatct	ccatTTTTaa	ttatatgaaa	tgaactgtag	cataagcagt	atTCatttgg	1920
attatTTTTt	ttattagctc	tcaccccttc	attattctga	gctgaaagtc	tggcatgaac	1980
tgtcctcaat	tttgTTTTca	aattcacatc	gattatctat	gcattatcct	cttgtatcta	2040
cctgtagaag	tttctTTTTg	gttattcctt	gactgcttga	ttacagaaag	aaatTTtatga	2100
agctgtaatc	gggatagtta	tactgcttgt	tcttatgatt	catTTccttt	gtgcagttct	2160
tgggtgtagct	tgccactttc	accagcaaaag	ttc			2193

<210> 55
 <211> 1827
 <212> ADN
 <213> *Oryza sativa*

5

<400> 55

ES 2 451 669 T3

gcttgagtca	tagggagaaa	acaaatcgat	catatttgac	tcttttccct	ccatctctct	60
taccggcaaa	aaaagtagta	ctggtttata	tgtaaagtaa	gattctttaa	ttatgtgaga	120
tccggcttaa	tgccttttctt	ttgtcacata	tactgcattg	caacaattgc	catatattca	180
cttctgccat	cccattatat	agcaactcaa	gaatggattg	atatatcccc	tattactaat	240
ctagacatgt	taaggctgag	ttgggcagtc	catcttccca	accaccacc	ttcgtttttc	300
gcgcacatac	ttttcaaact	actaaatggg	gtgtttttta	aaaatatttt	caatacaaaa	360
gttgctttaa	aaaattatat	tgatccattt	ttttaaaaaa	aatagcta	acttaattaa	420
tcacgtgta	aaagaccgct	ccgttttgcg	tgcaggaggg	ataggttcac	atcctgcatt	480
accgaacaca	gcctaaatct	tgttgtctag	attcgtagta	ctggatata	taaatacatgt	540
tctaagttac	tataactga	gatgaataga	ataagtaaaa	ttagaccac	cttaagtctt	600
gatgaagtta	ctactagctg	cgtttgggag	gacttcccaa	aaaaaaaaagt	attagccatt	660
agcacgtgat	taattaagta	ctagtttaa	aaacttaaaa	aataaattaa	tatgattctc	720
ttaagtaact	ctcctataga	aaacttttac	aaaattacac	cgtttaatag	tttggaatat	780
atgtcagtaa	aaaataagag	agtagaagtt	atgaaagtta	gaaaaagaat	tgtttttagta	840

gtatacagtt	ataaactatt	ccctctgttc	taaaacataa	gggattatgg	atggattcga	900
catgtaccag	taccatgaat	cgaatccaga	caagtttttt	atgcatattt	atttactat	960
aatatatcac	atctgctcta	aatatcttat	atttcgaggt	ggagactgtc	gctatgtttt	1020
tctgcccgtt	gctaagcaca	cgccaccccc	gatgcgggga	cgctctggc	cttcttgcca	1080
cgataattga	atggaacttc	cacattcaga	ttcgataggt	gaccgtcgac	tccaagtgtc	1140
ttgcacaaaa	caactccggc	ctcccggcca	ccagtcacac	gactcacggc	actaccacc	1200
ctgactccct	gaggcggacc	tgccactgtt	ctgcatgcga	agctatctaa	aattctgaag	1260
caaagaaagc	acagcacatg	ctccgggaca	cgcgccacc	ggcggaaaag	ggctcgggtg	1320
ggcgatctca	cagccgcata	tcgcatttca	caagccgccc	atctccaccg	gcttcacgag	1380
gctcatcgcg	gcacgaccgc	gcacggaacg	cacgcggcgg	accgcgcgc	ctcgatgcgc	1440
gagcccaccc	gcccgcgtct	ccctttgcct	ttgcccgtat	cctctcggtc	gtatcccgtt	1500
tctctgtctt	ttgctccccg	gcgcgcgcca	gttcggagta	ccagcgaac	ccggacacct	1560
ggtacacctc	cgccggccac	aacgcgtgtc	cccctacgtg	gccgcgcagc	acatgcccac	1620
gcgcgacaog	tgacctctct	catccaaact	ctcaagtctc	aacggtccta	taaatagcag	1680
gatagcctca	agctgctcgt	cacaaggcaa	gaggcaagag	gcaagagcat	ccgtattaac	1740
cagccttttg	agacttgaga	gtgtgtgtga	ctcgatccag	cgtagtttca	gttcgtgtgt	1800
tggtgagtga	ttccagccaa	gtttgcg				1827

<210> 56
 <211> 2194
 <212> ADN
 <213> *Oryza sativa*
 <400> 56

5

ES 2 451 669 T3

aatccgaaaa	gtttctgcac	cgttttcacc	ccctaactaa	caatataggg	aacgtgtgct	60
aaatataaaa	tgagacctta	tatatgtagc	gctgataact	agaactatgc	aagaaaaact	120
catccaccta	ctttagtggc	aatcgggcta	aataaaaaag	agtcgctaca	ctagtttctg	180
tttcttagt	aattaagtgg	gaaaatgaaa	tcattattgc	ttagaatata	cgttcacatc	240
tctgtcatga	agttaaatta	ttcgaggtag	ccataattgt	catcaaactc	ttcttgaata	300
aaaaaatctt	tctagctgaa	ctcaatgggt	aaagagagag	atTTTTTTT	aaaaaataga	360
atgaagatat	tctgaacgta	ttggcaaaga	tttaaacata	taattatata	atTTTtatag	420
ttgtgcattc	gtcatatcgc	acatcattaa	ggacatgtct	tactccatcc	caatttttat	480
ttagtaatta	aagacaattg	acttattttt	attattttatc	ttttttcgat	tagatgcaag	540
gtacttacgc	acacactttg	tgctcatgtg	catgtgtgag	tgcacctcct	caatacacgt	600
tcaactagca	acacatctct	aatatcactc	gcctatttaa	tacatttagg	tagcaatata	660
tgaattcaag	cactccacca	tcaccagacc	acttttaata	atatctaaaa	tacaaaaaat	720
aatTTTtacag	aatagcatga	aaagtatgaa	acgaactatt	taggtTTTT	acatacaaaa	780
aaaaaaagaa	ttttgctcgt	gcgcgagcgc	caatctccca	tattgggcac	acaggcaaca	840
acagagtggc	tgcccacaga	acaaccaca	aaaaacgatg	atctaacgga	ggacagcaag	900
tccgcaacaa	cctTTTtaaca	gcaggettgg	cggccaggag	agaggaggag	aggcaagaa	960
aaccaagcat	cctccttctc	ccatctataa	attcctcccc	cctTTTcccc	tctctatata	1020
ggaggcatcc	aagccaagaa	gagggagagc	accaaggaca	cgcgactagc	agaagccgag	1080
cgaccgcctt	ctcgatccat	atcttccggt	cgagttcttg	gtcgatctct	tcctcctcc	1140
acctcctcct	cacaggggat	gtgcctcctt	tcggttggtc	ttggatttat	tgttctaggt	1200
tgtgtagtac	gggctgtgat	gttaggaaag	gggatctgta	tctgtgatga	ttcctgttct	1260
tggatttggg	atagaggggt	tcttgatggt	gcatgttatc	ggttcgggtt	gattagtagt	1320
atggTTTTca	atogtctgga	gagctctatg	gaaatgaaat	ggTTtaggga	tcggaatctt	1380
gcgattttgt	gagtaccttt	tgTTtgaggt	aaaatcagag	caccggtgat	tttgettgg	1440
gtaataaagt	acggttgTTT	ggtcctcgat	tctggtagtg	atgcttctcg	atttgacgaa	1500
gctatccttt	gTTTattccc	tattgaacaa	aaataatcca	actTTgaaga	cggTcccgtt	1560
gatgagattg	aatgattgat	tcttaagcct	gtccaaaatt	tcgcagctgg	cttgtttaga	1620
tacagtagtc	cccatcacga	aattcatgga	aacagttata	atcctcagga	acaggggatt	1680
ccctgttctt	ccgatttgct	ttagtcccag	aattTTTTTT	cccaaatac	ttaaaaagtc	1740
actttctggt	tcagttcaat	gaattgattg	ctacaaataa	tgcttttata	gcgttatcct	1800
agctgtagtt	cagttaatag	gtaatacccc	tatagtttag	tcaggagaag	aacttatccg	1860
atTTctgatc	tccattTTT	attatatgaa	atgaactgta	gcataagcag	tattcatttg	1920
gattattTTT	tttatttagct	ctcaccctt	cattattctg	agctgaaagt	ctggcatgaa	1980
ctgtcctcaa	TTTTgtTTT	aaattcacat	cgattatcta	tgcatatcc	tcttgtatct	2040
acctgtagaa	gTTTctTTTT	ggttattcct	tgactgcttg	attacagaaa	gaaatttatg	2100
aagctgtaat	egggatagtt	atactgcttg	ttcttatgat	tcatttcctt	tgtgcagttc	2160
ttggtgtagc	ttgccacttt	caccagcaaa	gttc			2194

REIVINDICACIONES

1. Procedimiento para aumentar el rendimiento de semillas de plantas con respecto a plantas de tipo silvestre correspondientes, que comprende introducir y expresar en una planta un ácido nucleico AZ o una variante del mismo, en el que dicho ácido nucleico AZ codifica un polipéptido AZ o un homólogo del mismo, en el que el homólogo proporciona plantas que tienen un rendimiento aumentado y en el que dicho polipéptido AZ o el homólogo del mismo comprende dos repeticiones de anquirina y dos dominios C3H1 de dedo de Cinc, y en el que dichas repeticiones de anquirina se localizan cadena arriba de los dominios C3H1 de dedo de Cinc.
2. El procedimiento de acuerdo con la reivindicación 1, en el que polipéptido AZ comprende al menos uno de los siguientes motivos:
 10 (P/A)CSRAY(S/T)HDWTEC (motivo 1, SEC ID N°: 3)
 HPGENARRRDPR (motivo 2, SEC ID N°: 4)
 HG(V/I)FE(C/S)WLHP(A/S)QY(R/K)TRLCK (motivo 3, SEC ID N°: 5)
 CFFAH (motivo 4, SEC ID N°: 6)
3. El procedimiento de acuerdo con la reivindicación 1 o 2, en el que dicho ácido nucleico AZ codifica un polipéptido AZ de SEC ID N°: 2 o un homólogo del mismo.
4. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que dicha variante es una porción de un ácido nucleico AZ o una secuencia capaz de hibridarse con un ácido nucleico AZ, cuya porción o secuencia de hibridación codifica un polipéptido AZ que comprende dos repeticiones de anquirina y dos dominios C3H1 de dedo de Cinc, y en el que dichas repeticiones de anquirina se localizan cadena arriba de los dominios C3H1 de dedo de Cinc.
5. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que dicho ácido nucleico AZ o variante del mismo se sobreexpresa en una planta.
6. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que dicho ácido nucleico AZ o variante del mismo es de origen vegetal, preferentemente de una planta dicotiledónea, más preferentemente de la familia *Brassicaceae*, más preferentemente el ácido nucleico es de *Arabidopsis thaliana*.
7. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que dicho ácido nucleico AZ o variante del mismo está unido operativamente a un promotor específico de semilla.
8. El procedimiento de acuerdo con la reivindicación 7, en el que dicho promotor específico de semilla es un promotor WSI18.
9. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 8, en el que dicho rendimiento de semilla aumentado comprende un peso de mil granos aumentado.
10. Construcción que comprende:
 35 (i) un ácido nucleico AZ o una variante del mismo como se define en cualquiera de las reivindicaciones 1 a 4 y 6;
 (ii) un promotor WSI18 específico de semilla o un promotor GOS2 constitutivo unido operativamente a la secuencia de ácidos nucleicos de (i)
11. La construcción de acuerdo con la reivindicación 10, en la que dicho promotor WSI18 es como se representa por la SEC ID N°: 55 o SEC ID N°: 9.
12. La construcción de acuerdo con la reivindicación 10, en la que dicho promotor GOS2 es como se representa por la SEC ID N°: 56 o SEC ID N°: 54.
13. Planta transformada con una construcción de acuerdo con una cualquiera de las reivindicaciones 10 a 12.
14. Procedimiento para la producción de una planta transgénica que tiene rendimiento de semilla aumentado en comparación con plantas de tipo silvestre correspondientes, cuyo procedimiento comprende:
 45 (i) introducir y expresar en una planta o célula de planta un ácido nucleico AZ o variante del mismo como se define en cualquiera de las reivindicaciones 1 a 4 y 6; y
 (ii) cultivar la célula de planta en condiciones que promuevan el crecimiento y el desarrollo de las plantas.
15. Planta transgénica de acuerdo con la reivindicación 13, en la que dicha planta es una planta monocotiledónea, tal como caña de azúcar o en la que la planta es un cereal, tal como arroz, maíz, trigo, cebada, mijo, centeno, avena o sorgo.
16. Partes cosechables de una planta de acuerdo con una cualquiera de las reivindicaciones 13 o 15, en la que dichas partes cosechables comprenden una construcción de acuerdo con una cualquiera de las reivindicaciones 10 a 12.

17. Partes cosechables de una planta de acuerdo con la reivindicación 16 en la que dichas partes cosechables son semillas.
- 5 18. Productos directamente derivados de una planta de acuerdo con la reivindicación 13 o 15 y/o de partes cosechables de una planta de acuerdo con las reivindicaciones 16 o 17, en la que dichos productos comprenden una construcción de acuerdo con una cualquiera de las reivindicaciones 10 a 12.
19. El uso de un ácido nucleico AZ o de una variante del mismo como se define en cualquiera de las reivindicaciones 1 a 4 y 6, o el uso de un polipéptido AZ o de un homólogo del mismo, en el aumento de rendimiento de semilla, con respecto a plantas de tipo silvestre correspondientes.
- 10 20. El uso de acuerdo con la reivindicación 19, en el que dicho aumento de rendimiento de semilla comprende un peso de mil granos aumentado.
21. El uso de un ácido nucleico AZ o de una variante del mismo, como se define en cualquiera de las reivindicaciones 1 a 4 y 6, o el uso de un polipéptido AZ o de un homólogo del mismo como un marcador molecular.
22. El uso de una construcción, de acuerdo con cualquiera de las reivindicaciones 10 a 12, en un procedimiento para producir plantas que tienen un rendimiento de semilla aumentado con respecto a plantas de control.

Organización de dominios en la SEC ID N°: 2

MCCGSDRLNQIVSSRSLPI SFEEENNLTNTDMNHLTVETEDTFASLLELAANNDVEGVRLS IER
 DP SCVDEAGLWYGRQKGSKAMVNDYRTPLMVAATYGSIDVIKLI VSLTDADVNRACGNDQTALHC
AASGGAVNATQVVKLLLAAGADLNLLLDAEGQRAGDVIVVPPKLEGVKLMLQELLSADGSSTAERNL
 RVVTNVPNRSSSPCHSPTGENGGSGSGSPLGSPFKLKSTEFKKEYPVDPSLPDIKNSIYATDEFM
 YSFKVRPCSRAYSHDWTECPFVHPGENARRRDPKRFHYSCVPCPDRKGA CRRGDMCEYAHGVFEC
 WLHPAQYRTRLCKDGTGCARRVCFFAHTPEELRPLYASTGSAVPSPRSADYAAA LSLPGSPSGV
 SVMSP LSPSAAGNGMSHSNMAWPQPNVPALHLPGSNLQSSRLRSSLNARDIPTDEFNMLADYEQQQ
 LLNEYSNALSRSGRMKSMPPSNLEDLFSAEGSSSPRFTDSALASAVFSPTHKSAVFNQFQQQQQQQ
 QSM LSPINTSFSSPKSVDHSLFSGGGRMSPRNVVEPI SPMSARVSM LAQCVKQQQQQQQQQQHQ
 FRSLSSRELRTNSSPIVGS P VNNNTWSSKWGSSNGQPDWGMSSSEALGKLRSSSSFDGDEPDVSWVO
 SLVKETPAEAKKAAATSSSGEHVMKQPNPVEPVM DHAGLEAWIEQMQLDQLVAQQN

FIGURA 1

ANQ:

O04242/1-30 NGHTALHIAASK-----GDEQCVKLLLEHGA-----DPNA
 CONSENSO/80% .t.sshhsh.t.....tp.phhphlp.t.....pht.
 CONSENSO/65% pstosLphAstp.....sphphlpLlptss.....shsh
 CONSENSO/50% sGpTsLHhAsps.....sshcllchLlspus.....slst

C3H1:

TTP_BOVIN/13-4 RYKTELCRTF---SESG-----RCRYGA-KCQFAHGL
 CONSENSO/80% t.p...C..a.....tG.....C.hu..pC.a.H..
 CONSENSO/65% p.+..hCpha....tpG.....hCthGs.pCpahHs.
 CONSENSO/50% ph+.slCctF....ppG.....pCshGs.pCpFtHst

Leyenda:

Clase	Clave	Restos
alcohol	o	S, T
alifático	l	I, L, V
ninguna	.	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
aromático	a	F, H, W, Y
cargado	c	D, E, H, K, R
hidrófobo	h	A, C, F, G, H, I, K, L, M, R, T, V, W, Y
negativo	-	D, E
polar	p	C, D, E, H, K, N, Q, R, S, T
positivo	+	H, K, R
pequeño	s	A, C, D, G, N, P, S, T, V
diminuto	u	A, G, S
de tipo giro	t	A, C, D, E, G, H, K, N, Q, R, S, T

FIGURA 2

Tabla MATCAT: A) Secuencias de longitud completa

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.SEC ID N°2		63.8	44	58.5	54.6	49.5	43.8	38.2	36.9	38.6	37.9	35.7	38.6	38.5	35.9	27.8	48.8	43.1
2.SEC ID N°11	78.1		47.2	65.7	57	51.5	44	39	35.8	39.6	40.8	34.4	41.5	39.7	37	27.9	51.2	45.1
3.SEC ID N°13	57.5	60.4		45.4	48.2	39.1	39.1	37.7	37	37.2	38.1	33.5	37.1	37.1	38.2	30.2	37.9	35.2
4.SEC ID N°15	74.3	80.1	58.3		54.6	51.5	42.8	37.5	35.8	39.1	39.4	34.5	40.9	39.4	36.4	27	50.2	43.5
5.SEC ID N°17	69.6	72.5	57.7	68.6		47	45.5	36.3	35.2	39.2	38.1	33.9	38.9	38.5	34.9	29.5	46.4	41.7
6.SEC ID N°19	68.7	70.5	54.7	69.7	62.9		43.3	37.4	35.3	36.9	38.9	33.8	38.3	38.9	35	27.5	42.6	38.7
7.SEC ID N°21	57.4	60.8	54.5	59.3	57.9	58.9		36.6	34.9	36.3	36.7	33.6	35.6	36.4	37.5	28.1	39.7	37.3
8.SEC ID N°23	52.1	54.3	54.3	52.1	49.3	53.7	52.5		68.3	50.4	48.1	45	44.5	46.8	39.9	32.1	36.7	34.9
9.SEC ID N°25	49.9	50.9	53	51.6	48.2	52.1	50.7	81.7		49.6	47	45.3	43.2	46.6	40.1	30.9	35.8	34
10.SEC ID N°27	54.9	55.5	52.8	54.6	52.9	55	53.9	65.6	63		63.3	54.7	53.8	51	45.8	31.5	40.7	38.5
11.SEC ID N°29	54.5	57.7	53.2	58.2	52.5	55.5	54	65.9	62.7	78		53.7	54.1	47.3	41.7	31	40.3	37.6
12.SEC ID N°31	52.2	51.7	49.8	53.8	49.1	51.7	52.8	60	60	68.4	69		44.7	42.2	39.3	30.9	34.4	32.3
13.SEC ID N°33	56.1	58.8	52.4	59.1	53.8	57.9	53.2	61	59.2	69.5	69.9	62.7		48.7	38.3	29.8	40.8	39.2
14.SEC ID N°35	53.4	54.3	55.4	53.4	52.1	55.5	54.9	63.4	63.9	63.6	62.9	58	62.1		42.4	31.1	38	38.1
15.SEC ID N°37	50.3	50.9	55.3	51.1	47.5	49.9	53.3	58.7	57.2	60.1	57.3	55.8	55	60		39.8	34.5	33.9
16.SEC ID N°39	41.8	40.6	43.8	41.1	39.5	41.2	43.1	49.4	49	46.6	46	43.8	44.1	47.1	53.7		28.1	26.7
17.SEC ID N°41	63.9	65.5	50.6	64.8	64.1	58.5	53.3	50.2	48.1	54.8	54.4	51.3	56	50.2	48.5	38.1		54.8
18.SEC ID N°43	58.3	61.5	50.3	59.9	60.5	57	52.7	50.3	49.5	53.3	54.7	48.3	56.6	53.4	48.9	38.1	67.3	

FIGURA 3A

B) Secuencias de dominio de dedo de Cn

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.SEC ID N°2		92.0	96.0	88.0	88.0	84.0	72.0	88.0	80.0	80.0	88.0	88.0	76.0	84.0	84.0	76.0	88.0	92.0
2.SEC ID N°11	92.0		88.0	80.0	92.0	80.0	76.0	84.0	76.0	76.0	80.0	80.0	72.0	80.0	76.0	72.0	88.0	88.0
3.SEC ID N°13	96.0	88.0		88.0	84.0	84.0	72.0	84.0	76.0	76.0	84.0	84.0	76.0	80.0	88.0	80.0	84.0	88.0
4.SEC ID N°15	92.0	84.0	96.0		84.0	92.0	68.0	80.0	72.0	76.0	80.0	80.0	76.0	80.0	80.0	76.0	80.0	84.0
5.SEC ID N°17	88.0	92.0	84.0	88.0		84.0	68.0	80.0	72.0	76.0	76.0	76.0	72.0	80.0	72.0	72.0	88.0	88.0
6.SEC ID N°19	84.0	80.0	88.0	96.0	84.0		68.0	80.0	72.0	76.0	76.0	76.0	76.0	76.0	76.0	76.0	80.0	80.0
7.SEC ID N°21	84.0	92.0	84.0	80.0	84.0	80.0		68.0	60.0	60.0	64.0	64.0	60.0	64.0	72.0	68.0	68.0	68.0
8.SEC ID N°23	88.0	84.0	84.0	80.0	80.0	80.0	80.0		88.0	84.0	88.0	88.0	80.0	84.0	84.0	80.0	80.0	80.0
9.SEC ID N°25	88.0	84.0	84.0	80.0	80.0	80.0	80.0	100.0		88.0	92.0	92.0	84.0	80.0	80.0	72.0	72.0	72.0
10.SEC ID N°27	84.0	80.0	80.0	80.0	84.0	84.0	76.0	92.0	92.0		92.0	92.0	92.0	84.0	80.0	76.0	80.0	76.0
11.SEC ID N°29	92.0	84.0	88.0	84.0	80.0	80.0	80.0	96.0	96.0	92.0		100.0	88.0	88.0	88.0	76.0	76.0	80.0
12.SEC ID N°31	92.0	84.0	88.0	84.0	80.0	80.0	80.0	96.0	96.0	92.0	100.0		88.0	88.0	88.0	76.0	76.0	80.0
13.SEC ID N°33	80.0	76.0	84.0	84.0	76.0	88.0	76.0	88.0	88.0	92.0	88.0	88.0		80.0	80.0	76.0	72.0	72.0
14.SEC ID N°35	88.0	88.0	84.0	84.0	88.0	80.0	80.0	92.0	92.0	92.0	96.0	96.0	88.0		84.0	76.0	88.0	92.0
15.SEC ID N°37	88.0	80.0	92.0	88.0	76.0	84.0	80.0	92.0	92.0	88.0	96.0	96.0	92.0	92.0		88.0	72.0	76.0
16.SEC ID N°39	80.0	76.0	84.0	84.0	76.0	84.0	76.0	84.0	84.0	88.0	84.0	84.0	88.0	80.0	88.0		72.0	72.0
17.SEC ID N°41	92.0	96.0	88.0	88.0	100.0	88.0	84.0	84.0	84.0	88.0	84.0	84.0	80.0	88.0	80.0	80.0		96.0
18.SEC ID N°43	96.0	96.0	92.0	92.0	96.0	84.0	84.0	84.0	84.0	84.0	88.0	88.0	80.0	92.0	84.0	76.0	96.0	

FIGURA 3B

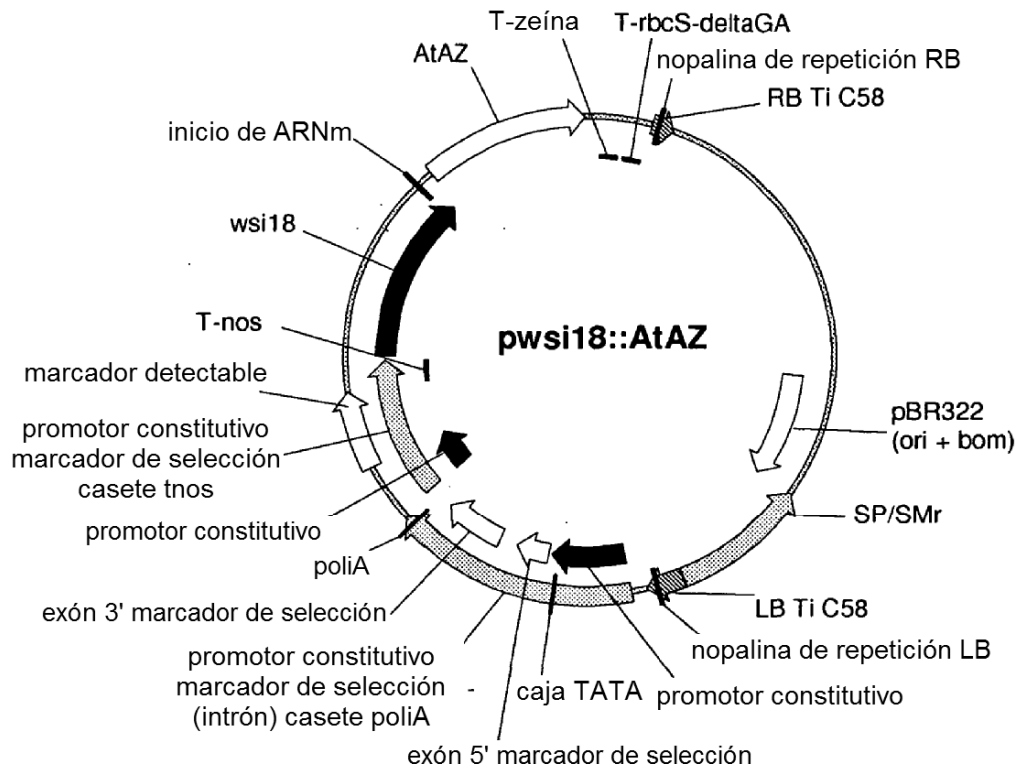


FIGURA 4