



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 362 716**

(51) Int. Cl.:

C12N 15/82 (2006.01)

C12N 15/11 (2006.01)

C07K 14/415 (2006.01)

C07K 14/47 (2006.01)

C07K 16/16 (2006.01)

C12Q 1/68 (2006.01)

G01N 33/50 (2006.01)

A01H 5/00 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **01943269 .9**

(96) Fecha de presentación : **19.04.2001**

(97) Número de publicación de la solicitud: **1276887**

(97) Fecha de publicación de la solicitud: **22.01.2003**

(54) Título: **Protección contra la toxicidad ambiental a través de la manipulación del procesamiento de precursores de ARN mensajero.**

(30) Prioridad: **19.04.2000 ES 200001102**

(45) Fecha de publicación de la mención BOPI:
12.07.2011

(45) Fecha de la publicación del folleto de la patente:
12.07.2011

(73) Titular/es: **Universidad Politécnica de Valencia
Centro de Transferencia de Tecnología
CTT-Edif. 6G - Camino de Vera, s/n
46022 Valencia, ES**

(72) Inventor/es: **Vicente Meana, Óscar;
Roldán Medina, Marta;
Serrano Salom, Ramón;
Forment Millet, José Javier;
Naranjo Olivero, Miguel Ángel;
Ros Palau, Roc y
Kahnonou, Rodolphe, Arthur**

(74) Agente: **Carvajal y Urquijo, Isabel**

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

Protección contra la toxicidad ambiental a través de la manipulación del procesamiento de precursores de ARN mensajero

Campo de la invención

Esta invención se refiere al uso de ácidos nucleicos y de proteínas involucrados en el procesamiento de precursores de ARN mensajero para mejorar la tolerancia al estrés ambiental tal como la toxicidad por sales minerales en células y organismos eucariotas.

Antecedentes de la invención

La naturaleza de los objetivos celulares sensibles a la toxicidad por sodio y litio representa un importante vacío en nuestro conocimiento sobre la fisiología de la homeostasis iónica en células eucariotas. La caracterización de estos objetivos es esencial para la comprensión de problemas clínicos tales como los efectos del litio sobre la terapia para el trastorno dipolar [Schou (1997) Arch. Gen. Psychiatry 54, 9] o altos niveles de sodio asociados con hipertensión [Lifton (1996) Science 272, 676]. Otro problema, completamente diferente pero también relacionado con homeostasis iónica, es la salinización progresiva de tierras cultivadas sometidas a intensa irrigación, que ha convertido al fitomejoramiento de las plantas de cultivo para tolerancia a la salinidad en una urgente necesidad para el desarrollo de una agricultura sostenible en las regiones [Serrano (1996) Int. Rev. Cytol. 165, 1; Yeo (1998) J. Exp. Bot. 49, 915; Holmberg & Bülow (1998) Trends Plant Sci. 3, 61].

Aparte del transporte de iones [Haro et al. (1991) FEBS Lett. 291, 189; Gaxiola et al. (1999) Proc. natl. Acad. Sci. USA 96, 1480; Apse et al. (1999) Science 285, 1256] y de la síntesis de osmolitos [Tarczynski et al. (1993) Science 259, 508; Kishor et al. (1995) Plant Physiol. 108, 1387; Alia et al. (1998) Plant J. 16, 155], la manipulación de sistemas celulares más sensibles a altas concentraciones de iones y a estrés por agua ofrece rutas alternativas para mejorar la tolerancia a la sal por parte de las plantas de cultivo [Serrano (1996) Int. Rev. Cytol. 165, 1; Tezara et al. (1999) Nature 401, 914].

Los análisis genéticos y bioquímicos han permitido identificar el producto del gen *HAL2* de lavadura como un importante objetivo fisiológico de la toxicidad por sal [Gläser et al. (1993) EMBO J. 12, 3105; Dichtl et al. (1997) EMBO J. 16, 7184]. *HAL2* codifica una 3',5'-bifosfato nucleotidasa, que es muy sensible a la inhibición por litio y sodio [Murguía et al. (1995) Science 267, 232]. La inhibición de sal de Hal2p resulta en una acumulación intracelular de 3'-fosfoadenosina 5'-fosfato (pAp) [Murguía et al. (1996) J. Biol. Chem. 271, 29029], un compuesto tóxico que a su vez inhibe las reacciones de reducción y transferencia de grupos sulfato, así como algunas exorribonucleasas [Dichtl et al. (1997) EMBO J. 16, 7184; Gil-Mascarell et al. (1999) Plant J. 17, 373]. Existen genes homólogos a *HAL2* en plantas [Gil-Mascarell et al. (1999) Plant J. 17, 373] y en mamíferos [López-Coronado et al. (1999) J. Biol. Chem. 274, 16043] aunque en el último caso la enzima codificada es inhibida por el litio pero no por el sodio.

La tolerancia a la sal conferida por la sobreexpresión de Hal2p, la proteína de tipo silvestre así como versiones mutadas resistentes al litio y al sodio, es relativamente modesta [Albert et al. (2000) J. Mol. Biol. 295, 927]. Esto sugiere la existencia de objetivos adicionales de toxicidad por sal, que se convierte en limitante una vez que se supera el cuello de botella de *HAL2*, pero aún no se conoce la naturaleza de estos importantes procesos sensibles a la sal.

Dos solicitudes de patente que se relacionan con estrés osmótico pero que describen mecanismos de protección diferentes al mecanismo general descrito en la presente invención son las siguientes: ES2110918A (1998-02-16) que se relaciona con la producción de plantas tolerantes a estrés osmótico a través de la manipulación del metabolismo de carbohidratos y ES2134155A1 (1998-10-07) que se relaciona con un método para conferir tolerancia a estrés osmótico, por agua y sal en plantas glicofílicas, a través del uso de genes que codifican proteínas con actividad de peroxidasa.

Resumen de la invención

El problema técnico fundamental de la presente invención es el de proporcionar un método que pueda ser utilizado para mejorar la tolerancia al estrés de células y de organismos que sufran de condiciones de estrés tales como estrés osmótico, causado por salinidad, sequía o estrés por frío y congelación.

5 Una solución a este problema técnico se logra suministrando una protección contra estrés osmótico en células y en organismos a través de la manipulación del procesamiento del ARN mensajero. La presente invención proporciona un conjunto de genes aislados que son capaces de conferirle a una célula huésped heteróloga u organismo huésped tolerancia a las condiciones de estrés. Estos genes están todos involucrados en el procesamiento de precursores de ARNm (tal como síntesis, empalme, modificación del extremo 5' y 3', movimiento, transporte hasta el citoplasma, metabolismo, etc.) y todos mostraron un fenotipo de resistencia a la salinidad cuando transformaron separadamente a un mutante de levadura sensible a la sal. Haciendo esto, cada gen actúa como un eficiente reforzador de la tolerancia al estrés, sin la ayuda de factores adicionales.

15 Este grupo de genes, que incluye proteínas del tipo SR, factores de enlazamiento de ARN (nuclear), componentes de complejos de ribonucleoproteína, factores de transcripción, y proteínas de movimiento nuclear, permiten que la persona capacitada en el arte altere genéticamente al organismo de interés con el propósito de hacerlo tolerante a situaciones de estrés tales como situaciones de estrés osmótico, más particularmente toxicidad por Na⁺ o Li⁺ o por una sal mineral. Cada uno de los genes divulgados le permite a la persona capacitada en el arte modificar el destino de la célula y/o el desarrollo de la planta y/o la bioquímica y/o la fisiología por medio de la introducción de al menos uno de estos genes dentro de la célula. Para el cultivo de plantas de cultivo por ejemplo, muchas de la cuales son sensibles a condiciones de estrés como salinidad, sequía o frío, los genes divulgados ofrecen la posibilidad de resolver el problema de un menor rendimiento y menor beneficio económico.

25 Esta invención ofrece una solución para la toxicidad celular provocada por estrés ambiental a través de la manipulación del procesamiento de precursores de ARN mensajero y las modalidades de la invención incluyen métodos, ácidos nucleicos, polipéptidos, vectores, células huésped y organismos transgénicos tales como plantas transgénicas.

30 Descripción detallada de la invención

La salinidad del suelo es uno de los estreses abióticos o ambientales más significativos para la agricultura de plantas. Aparte del éxito práctico del mejoramiento genético de la tolerancia a la sal de las plantas de cultivo, la investigación de tolerancia a la sal representa una parte importante de la biología básica de la planta. También la investigación sobre otros dos estreses abióticos principales, sequía y frío, está íntimamente relacionada con el trabajo sobre estrés por salinidad. Por ejemplo, muchos genes que están regulados por el estrés por salinidad son también responsables por el estrés por sequía o por frío (Zhu J. K., 1997, Molecular aspects of osmotic stress in plants, CRC Crit. Rev. Plant Sci. 16 253 - 277). Una persona capacitada en el arte puede asumir por lo tanto que cuando un gen aislado confiere tolerancia a la sal a un organismo huésped cuando se lo transfecta aquí, podría conferir también tolerancia al estrés por sequía y/o frío. La salinidad, la sequía y el frío se considera que son las tres formas más importantes de estrés osmótico.

Con el propósito de identificar nuevos objetivos de toxicidad ambiental en células eucariotas, los inventores caracterizaron genes que confieren un incremento en la tolerancia a sales minerales cuando se expresan en células de levadura. Por lo tanto, se buscaron bibliotecas de ADNc de *Arabidopsis thaliana* [Minet et al. (1992) Plant J. 2, 417] y *Beta vulgaris* que pudieran conferir tolerancia al estrés por salinidad por medio de sobreexpresión en células de levadura, en las cuales la acumulación de pAp, y su efecto tóxico, fueron evitados por medio de la adición de metionina al medio de cultivo [Gläser et al. (1993) EMBO J. 12, 3105; Dichtl et al. (1997) EMB J. 16, 7184; Murguía et al. (1995) Science 267, 232; Murguía et al. (1996) J. Biol. Chem. 271, 29029]. Los resultados de esta estrategia han dado lugar a la presente invención. Esta es una aproximación funcional para la identificación de genes y proteínas que están involucrados en la respuesta de las plantas al estrés por salinidad. Para este propósito se construyeron bibliotecas de expresión de ADNc como se describe en el ejemplo 1 y en el ejemplo 3 y se utilizaron cepas competentes de levadura (ver el ejemplo 2) para seleccionar los ADNc que incrementaron la tolerancia de la

levadura a la salinidad por sobreexpresión. El crecimiento de estas cepas de levadura es normalmente inhibida a altas concentraciones de NaCl o LiCl de (150 mM) en forma similar a aquellas que perjudican el desarrollo de la mayoría de las especies de cultivo. Después de transformar las células de levadura con la biblioteca de ADNc, se reunieron las colonias y se hizo una selección por su habilidad para desarrollarse en presencia de NaCl 150 mM. Este procedimiento de selección es descrito más adelante en el ejemplo 4.

Aislamiento de genes de *Arabidopsis thaliana* que mejoran la tolerancia a la sal

Se seleccionó una biblioteca de ADNc de la planta *Arabidopsis thaliana* de aproximadamente $7,5 \times 10^5$ transformantes, y se aislaron tres clones independientes que fueron capaces de conferir tolerancia a altas concentraciones de sal en levadura (Fig. 1). Sorprendentemente, estos tres ADNc codifican proteínas implicadas en el procesamiento de precursores de ARN mensajero. Estas tres proteínas fueron llamadas Ct-SAL1, RCY1 y U1A.

Dos de estas proteínas pertenecen a la familia de los así llamados factores “tipo SR” o “ricos en arginina alternante”, definidos por tener un dominio con un alto contenido en residuos de Arg alternando con Ser, Asp, y/o Glu (dominio RS) (Fig. 2). Los miembros de esta familia han sido involucrados en el empalme constitutivo y/o alternativo, y en el acoplamiento de diferentes etapas durante el procesamiento y metabolismo del ARN mensajero (transcripción, modificaciones en los extremos 5' y 3', y empalme de pre-ARNm, transporte de ARN maduro al citoplasma, etc.). La proteína Ct-SRL1, con una composición de aminoácidos como la expuesta en la SEQ ID NO. 3, es codificada por el ADNc identificado aquí como la SEQ ID NO. 1, que puede ser encontrada también sobre la región genómica del cromosoma 5 de *Arabidopsis thaliana*, clon MNJ8 de P1 (número de acceso del GenBank AB017069). También está presente la secuencia de la proteína de la SEQ ID NO. 3 en la base de datos pública bajo el número de acceso BAB09109 pero no se le asignó función a esta secuencia. El ADNc de Ct-SRL1 no es de longitud completa y codifica al extremo del terminal carboxilo de una proteína tipo SR, que incluye al dominio RS. Su expresión en levadura confiere tolerancia al litio y al sodio (Fig. 1).

El segundo clon codifica una proteína putativa con un dominio de ciclina del terminal N relacionado con aquellas ciclinas tipo K y tipo T y con un dominio rico en arginina en el terminal C (Fig. 2A). Esta proteína tipo SR fue llamada RCY1 por alterar a la ciclina 1 rica en arginina. Su expresión en levadura le confirió tolerancia al sodio y al litio (Fig. 1). Para determinar el papel del dominio RS en esta proteína tipo SR para el mejoramiento de la tolerancia a la sal en levadura, el dominio de ciclina y el dominio RS de RCY1 fueron clonados separadamente y transformados en células de levadura. La división de ambos dominios está marcada en la figura 2A por medio del residuo subrayado de metionina y la parte del terminal C, que contiene al dominio RS se conoce como la SEQ ID NO. 21 (Figura 5). La expresión únicamente del dominio RS confiere el mismo fenotipo que la expresión de la proteína de longitud completa (Fig. 3). Estos resultados demuestran que, en el caso de factores “tipo SR”, es la expresión del dominio en sí mismo, y no de una proteína específica, lo que confiere tolerancia al estrés por salinidad. La composición de aminoácidos de la proteína RCY1 está representada aquí como la SEQ ID NO. 4. El ADNc de RCY1 se identifica aquí como la SEQ ID NO. 2 y esta secuencia puede ser encontrada sobre la región genómica de la secuencia del cromosoma II de *Arabidopsis thaliana* del clon T9J22. La secuencia de nucleótidos de este clon genómico de *Arabidopsis thaliana*, que corresponde a una parte de la secuencia de nucleótidos del gen para RCY1, está anotada como una ciclina putativa (número de acceso AAC14513.1). Esta predicción de la proteína a partir de la base de datos pública, carece de los primeros 55 aminoácidos de la proteína RCY1. Como tal, el fragmento de la proteína RCY1 correspondiente a la región del terminal amino está representado en la SEQ ID NO. 22.

El tercer clon de *Arabidopsis* codifica una proteína U1A de *Arabidopsis*, un componente de la U1-snRNP, el complejo de ribonucleoproteína que reconoce al sitio de empalme 5' en una de las primeras etapas en el procesamiento de intrones del pre-ARN mensajero. La proteína U1a de *Arabidopsis* fue previamente descrita por Simpson et al (1995, EMBO J. 14: 4540 - 4550) como un componente específico del U1-snRNP spliceosomal. La expresión de U1A en levadura confirió una tolerancia más débil al LiCl y ninguna tolerancia al NaCl (Fig. 1). La proteína, con una composición de aminoácidos como la expuesta en la SEQ ID NO. 6, es codificada por el ADNc identificado aquí como la SEQ ID NO. 5. La secuencia de proteína así como la secuencia de ácido nucleico pueden encontrarse en las bases de datos públicas bajo los números de acceso CAA90283.1 y Z49991 respectivamente y están completamente anotadas. Las secuencias descritas anteriormente se presentan en la figura 5.

Fenotipos como los descritos anteriormente fueron observados en presencia y en ausencia de metionina, y en diferentes ambientes genéticos.

La mejora de la tolerancia a la sal por medio de la expresión de los clones de Arabidopsis no fue debida a la estimulación del transporte de iones en levadura, ya que no estaba asociada a cambios en las concentraciones intracelulares de litio. Esto se determinó como se describe en el ejemplo 5. También se mantuvieron los fenotipos en una cepa de levadura deficiente en el transporte vacuolar.

Los datos anteriormente mencionados sugieren que un impacto sobre otro proceso celular fue responsable por la tolerancia al estrés observado. Los inventores creyeron que el procesamiento de precursores de ARN mensajero podría ser un objetivo de toxicidad iónica en células eucariotas. Esto no ha sido descrito previamente.

De acuerdo con lo anterior, los inventores han confirmado que el procesamiento de intrones de los pre-ARNm está inhibido en levadura en la presencia, por ejemplo, de cloruro de litio. Dos ensayos independientes para empalme de pre-ARNm *in vivo* dieron soporte a esta invención. Primero, los inventores han medido la actividad específica de la enzima β -galactosidasa sintetizada en células de levadura a partir de un plásmido que contiene al gen *lacZ* de *E. coli* interrumpido en forma artificial por un intrón (ejemplo 6). Ellos detectaron una disminución en la acumulación de esta enzima, comparada con aquella producida a partir de la construcción de control sin intrón, cuando se añade LiCl al medio de cultivo. La expresión simultánea del dominio RS de SRL1 de Arabidopsis en estas células de levadura bloqueó parcialmente la inhibición observada (datos no mostrados). Estos resultados han sido confirmados por el segundo ensayo, en el cual los inventores determinaron directamente, por medio de la técnica de RT-PCR como se describe en el ejemplo 7, la inhibición del empalme en presencia de litio. Se demostró la acumulación de precursores de ARN mensajero de levadura endógena en presencia de LiCl, por ejemplo el pre-ARNm correspondiente al gen *SAR1*. Debido a que una inhibición general del empalme afectaría primero la remoción de aquellos intrones normalmente procesados con menor eficiencia, los inventores escogieron para estos experimentos al pre-ARNm de *SAR1*, que contiene tal intrón (Kao y Siliciano, 1996, Mol. Cell. Biol. 16: 960 - 967). Aquí nuevamente, los inventores observaron la acumulación de pre-ARNm de *SAR1* por medio de la incubación de células de levadura bajo condiciones de estrés por salinidad, y cómo fue parcialmente revertida por medio de co-expresión simultánea de Ct-SRL1. De este modo los inventores demostraron que la inhibición del procesamiento del ARNm precursor en presencia de sal es parcialmente revertida por medio de la expresión de uno de los clones de Arabidopsis mencionado antes (por ejemplo, Ct-SRL1).

La importancia general de la presente invención fue corroborada por medio del fenotipo de plantas transgénicas de Arabidopsis que sobreexpresaron al ADNc de Ct-SRL1. Soportando al carácter general del mecanismo, la expresión del mismo ADNc de Ct-SRL1 bajo el control del promotor 35S de CaMV en plantas transgénicas de Arabidopsis, incrementa su tolerancia al NaCl y al LiCl en una forma similar que en la levadura. Esto puede ser observado, por ejemplo, por medio de germinación de semillas transgénicas en placas de agar que contienen concentraciones de LiCl que son tóxicas para las semillas de control de tipo silvestre (Fig. 4). Las tres líneas transgénicas independientes fueron capaces de crecer indicando la eficiencia del método de la presente invención.

A partir de estos resultados y de la propia naturaleza de los clones aislados de Arabidopsis, se puede deducir que cualquier estimulación del procesamiento de los precursores del ADN mensajero, independientemente del mecanismo involucrado, contrarresta el efecto tóxico de la sal, y que este efecto protector contra el estrés por salinidad es general en todas las células y organismos eucariotas. El carácter universal de la invención, es decir que el efecto protector contra el estrés por salinidad no depende de la especie, fue confirmado por el aislamiento de genes de la remolacha azucarera y genes eucariotas que también confieren tolerancia a la sal y que fueron también relacionados con el procesamiento de precursores de ARN mensajero.

Aislamiento de genes de *Beta vulgaris* que mejoran la tolerancia a la sal

Otro aspecto de la presente invención es el procedimiento de selección de una biblioteca de ADNc a partir de hojas de remolacha azucarera inducidas por NaCl y posterior aislamiento de los siete genes de la remolacha azucarera que confieren tolerancia al estrés a células de levadura. Se siguió una aproximación funcional para identificar genes de la remolacha azucarera y proteínas que están involucradas en la respuesta de plantas al estrés por salinidad. Para este propósito se construyó una biblioteca de expresión de ADNc inducida por NaCl a partir de hojas de la

remolacha azucarera como se describe en el ejemplo 1 y en el ejemplo 3 y se utilizó la cepa JM26 mutante de levadura sensible al Na^+ (ver el ejemplo 2) para seleccionar los ADNc de remolacha azucarera que incrementaron la tolerancia a la salinidad de la levadura por sobreexpresión. El crecimiento de este mutante de levadura está normalmente inhibido con concentraciones de NaCl (150 mM) similares a aquellas que perjudican el crecimiento de la mayoría de las especies de cultivo. Después de transformar las células de levadura con la biblioteca de ADNc, se reunieron y seleccionaron las colonias por su habilidad para crecer en presencia de NaCl 150 mM. Este procedimiento de selección es descrito adicionalmente en el ejemplo 4. Seis clones positivos que sobrevivieron a las altas concentraciones de sal, fueron caracterizadas posteriormente y contenían un gen con una secuencia como en las SEQ ID NO. 7, SEQ ID NO 9, SEQ ID NO 11, SEQ ID NO. 13, SEQ ID NO. 15 o SEQ ID NO. 17. Las secuencias de la proteína correspondiente codificadas por estos genes tienen una composición de aminoácidos como la expuesta en las SEQ ID NO 8, SEQ ID NO. 10, SEQ ID NO 12, SEQ ID NO. 14, SEQ ID NO. 16 o SEQ ID NO. 18 respectivamente. Todas estas secuencias son presentadas en la figura 5. Sorprendentemente, cada uno de los seis clones de levadura seleccionados son transformados por un gen que codifica una proteína que está implicada en el procesamiento del ARNm: una proteína es una proteína putativa de enlazamiento de ARN de arginina-aspartato, dos son una proteína putativa de enlazamiento de ARN, dos son un factor putativo de transcripción y uno se presenta en forma similar a una proteína de movimiento. Por lo tanto estas proteínas (y los genes que las codifican) son adecuadas para ser utilizadas como reforzadoras de la tolerancia al estrés a través de la manipulación del procesamiento de pre-ARNm, como se describe para los genes de *Arabidopsis thaliana* de la presente invención y como se describió más arriba.

Los genes seleccionados y sus proteínas codificadas son descritos adicionalmente en los siguientes párrafos.

Por lo tanto, la invención se relaciona con un nuevo ácido nucleico aislado de remolacha roja como se expone en la SEQ ID NO. 7, que codifica una proteína putativa de enlazamiento de ARN de arginina-aspartato y capaz de reforzar la tolerancia a la salinidad en células de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 51 y que termina en la posición 1079 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 8. Este polipéptido tiene 76% de identidad y 86% de similitud con la proteína putativa de enlazamiento de ARN de arginina-aspartato de *Arabidopsis thaliana* (número de acceso del Swiss-Prot AAB68037). La comparación de aminoácidos se hizo con el programa GAP (Tabla de comparación de símbolos: blosum62.cmp, CompCheck: 6430, matriz de sustitución de aminoácidos BLOSUM62, Referencia: Henikoff, S. y Henikoff, J. G. (1992). Matrices de sustitución de aminoácidos de bloques de proteína. Proc. Natl. Acad. Sci. USA 89: 10915 - 10919.). Este programa también fue utilizado para comparar las secuencias de aminoácidos de los siguientes polipéptidos de la remolacha roja.

La invención también se relaciona con un nuevo ácido nucleico aislado de remolacha roja como el expuesto en la SEQ ID NO. 9, que codifica una proteína putativa de enlazamiento de ARN y capaz de mejorar la tolerancia a la sal en células de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 14 y termina en la posición 625 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 10. Este polipéptido tiene 68% de identidad y 78% de similitud con la proteína putativa de enlazamiento de ARN de *Arabidopsis thaliana* (número del acceso del Swiss-Prot AAG52616.1).

La invención también se relaciona con un nuevo ácido nucleico aislado de la remolacha roja como el expuesto en la SEQ ID NO. 11, más conocido como clon o secuencia número 10, que codifica un factor de transcripción putativo y capaz de mejorar la tolerancia a la sal en células de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 51 y termina en la posición 1119 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 12. Este polipéptido tiene 81% de identidad y 86% de similitud con la proteína de enlazamiento de ARN nuclear de *Spinacia oleracea* (número del acceso del Swiss-Prot AAF14145.1).

Los nucleótidos 1 a 475 de la SEQ ID NO. 11 fueron publicados en la base de datos EST bajo el número de acceso BF011019 con la descripción de fragmentos de ADN de una biblioteca de ADNc de germinación de remolacha azucarera y que es similar a un factor de transcripción putativo.

La invención también se relaciona con un nuevo ácido nucleico aislado de remolacha roja como el expuesto en la SEQ ID NO. 13, que codifica un factor de transcripción putativo y capaz de mejorar la tolerancia a la sal en células

de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 51 y termina en la posición 1121 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 14. Este polipéptido tiene 81% de identidad y 87% de similitud con la proteína de enlazamiento de ARN nuclear de *Spinacia oleracea* (número de acceso del Swiss-Prot AAF14144.1).

La invención también se relaciona con un nuevo ácido nucleico aislado de remolacha roja como el expuesto en la SEQ ID NO. 15, que codifica una proteína putativa de enlazamiento de ARN y capaz de mejorar la tolerancia a la sal en células de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 2 y termina en la posición 970 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 16. Este polipéptido tiene 68% de identidad y 74% de similitud con la proteína putativa de enlazamiento de ARN de *Oryza sativa* (número de acceso del Swiss-Prot AAG59664.1). La invención también se relaciona con un nuevo ácido nucleico aislado de remolacha roja como se expone en la SEQ ID NO. 17, que codifica un tipo desconocido de proteína capaz de mejorar la tolerancia a la sal en células de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 35 y termina en la posición 922 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 18. Este polipéptido tiene 61% de identidad y 69% de similitud con la proteína de *Arabidopsis thaliana* con similitud con una proteína de movimiento nuclear (número de acceso del Swiss-Prot BAAF97317.1). Por lo tanto, una modalidad preferida de la presente invención se relaciona con un método para la inducción de tolerancia al estrés para un organismo que comprende la expresión de un (o al menos un) gen de *Beta vulgaris*, que está involucrado en el procesamiento de precursores de ARN mensajero.

También se puede utilizar el procedimiento de detección y selección como se describió anteriormente para seleccionar genes de organismos diferentes a las plantas. A manera de ejemplo, se seleccionó una secuencia de *Mus musculus* que mejora la tolerancia a la sal en células de levadura. Esta secuencia está expuesta en la SEQ ID NO. 19 y codifica una subunidad putativa pequeña de una proteína del factor auxiliar U2 snRNP y es capaz de mejorar la tolerancia a la sal en células de levadura. El marco de lectura abierto, que comienza en la posición del nucleótido 37 y termina en la posición 969 codifica la secuencia de aminoácidos como la expuesta en la SEQ ID NO. 20. Este polipéptido tiene 78% de identidad y 84% de similitud con la subunidad pequeña, del factor auxiliar U2 snRNP de *Arabidopsis thaliana* (número de acceso del Swiss-Prot BABF10638.1). Esta secuencia está presentada en la figura 5.

Este resultado ilustra que también pueden utilizarse genes de mamífero en el método de la presente invención. El método de la presente invención es por lo tanto generalmente aplicable para conferir tolerancia al estrés a una célula huésped u organismo a través de la manipulación de precursores de ARN mensajero.

El fenotipo sorprendentemente fuerte de algunos de los clones de levadura seleccionados como se describió anteriormente y el hecho de que estos genes en una posición aislada y en un ambiente heterólogo actuaran como mejoradores de la tolerancia al estrés, convierte a estos genes en herramientas muy atractivas para inducir tolerancia al estrés en cualquier organismo de interés, sin la necesidad de compuestos accesorios. La habilidad de estos genes para mejorar la tolerancia al estrés osmótico, particularmente al estrés por sales minerales tal como el estrés por Na^+ y Li^+ en células de levadura cuando se las aísla y transfecta aquí, demuestra claramente su potencial para conferir por su propia cuenta tolerancia al estrés osmótico a cualquier organismo huésped heterólogo. Alternativamente, cada uno de los genes de tolerancia al estrés de la presente invención puede ser combinado con otro gen, con el propósito de alterar el destino de la célula o la morfología de la planta, el desarrollo de la planta, la bioquímica de la planta o la fisiología de la planta.

De acuerdo con una primera modalidad, la presente invención se relaciona con un método para mejorar la tolerancia al estrés en células y organismos que comprende la manipulación del proceso de procesamiento de precursores de ARN mensajero.

De acuerdo con una modalidad preferida, dicho estrés podría ser estrés ambiental, tal como, pero sin limitarse a, estrés osmótico, estrés por salinidad, estrés por sequía, estrés por frío o congelación. De acuerdo con una modalidad preferida, los métodos de la invención se relacionan con el mejoramiento de la tolerancia a la sal de células y organismos.

Otra modalidad de la invención se relaciona con un método para proteger células y organismos contra la toxicidad por sal que comprende la manipulación del proceso de procesamiento de precursores de ARN mensajero (ARNm).

De acuerdo con una modalidad, una forma de manipulación del proceso de procesamiento de precursores de ARN mensajero es por medio de manipulación genética o bioquímica de al menos una molécula que está involucrada en o que interfiere con el proceso de procesamiento de precursores de ARN mensajero o con una de las rutas de procesamiento de precursores de ARNm. El término "células" se relaciona con cualquier célula procariota o eucariota. El término "organismo" se relaciona con cualquier organismo mono o multicelular de origen procariota o eucariota.

La presente invención claramente describe varios genes y proteínas que pertenecen a diferentes clases de genes y proteínas que pueden ser utilizados para mejorar la tolerancia al estrés. Estos genes y proteínas de la invención se ha demostrado que tienen un efecto sobre el proceso de procesamiento del ARNm.

Por lo tanto, de acuerdo con aún modalidades preferidas, la invención se relaciona con cualquiera de los métodos anteriormente mencionados en donde está involucrada la manipulación genética o bioquímica de una proteína que posee un dominio con un alto contenido de dipéptidos Arg-Ser, Arg-Glu y Arg-Asp (dominio RS), una proteína de enlazamiento de ARN, un componente del complejo U1-snRNP o de U2-snRNP, un factor de transcripción, o una proteína de movimiento nuclear.

La invención también se relaciona con el uso de un ácido nucleico aislado que comprende una secuencia de ácido nucleico como la representada en la SEQ ID NO. 1 con una secuencia de aminoácidos como la expuesta en la SEQ ID NO. 3, o un ácido nucleico como el representado en la SEQ ID NO. 2 con una secuencia de aminoácidos como la expuesta en la SEQ ID NO. 4 o la SEQ ID NO. 21, o un ácido nucleico que codifica un polipéptido que comprende la secuencia de aminoácidos representada en la SEQ ID NO. 22, para al menos uno, o en al menos uno, de los métodos anteriormente descritos. Las SEQ ID NOs. 1 y 2 comparten una homología sustancial con genes que codifican proteínas tipo SR.

La invención también se relaciona con un ácido nucleico aislado que comprende una secuencia de ácido nucleico como la representada en la SEQ ID NO. 5 con una secuencia de aminoácidos como la expuesta en la SEQ ID NO. 6 para cualquiera de los métodos anteriormente descritos. La SEQ ID NO. 5 comparte homología sustancial con los genes que codifican un componente del complejo I1-snRNP o U2-snRP.

La invención también se relaciona con un ácido nucleico aislado que comprende una secuencia de ácido nucleico como la representada en cualquiera de las SEQ ID NOs. 7, 9, 11, 13, 15, 17 ó 19, con una secuencia de aminoácidos como la expuesta en cualquiera de las SEQ ID NOs. 8, 10, 12, 14, 16, 18 ó 20 para cualquiera de los métodos anteriormente descritos. Las SEQ ID NOs. 7, 9, 13 y 15 comparten homología sustancial con los genes que codifican proteínas de enlazamiento de ARN. La SEQ ID NO. 19 comparte homología sustancial con los genes que codifican un componente del complejo I1-snRNP o U2-snRP. Las SEQ ID NOs 11 y 17 comparten homología sustancial con los genes que codifican factores transcripcionales.

La invención se relaciona además con un ácido nucleico aislado que codifica una proteína o un fragmento inmunológicamente activo y/o funcional de tal proteína seleccionada del grupo que consiste de:

a) un ácido nucleico que comprende una secuencia de ADN como figura en cualquiera de las SEQ ID NOs 1, 2, 5, 7, 9, 11, 13, 15, 17 ó 19 ó el complemento de las mismas,

b) un ácido nucleico que comprende la secuencia de ARN correspondiente a cualquiera de las SEQ ID NOs 1, 2, 5, 7, 9, 11, 13, 15, 17 ó 19 como en (a) ó el complemento de las mismas,

c) un ácido nucleico que hibrida específicamente a la secuencia de nucleótidos como la definida en (a) ó (b),

d) un ácido nucleico que codifica un polipéptido ó proteína con una secuencia de aminoácidos que es al menos 50 %, preferiblemente al menos 60 %, 70 % ó 80%, más preferiblemente al menos 85% ó 90%, lo más preferible 95 % idéntica al polipéptido representado en cualquiera de las SEQ ID NOs 3, 4, 6, 8, 10, 12, 14, 16, 18, 20 ó 21,

e) un ácido nucleico que codifica un polipéptido ó proteína que comprende la secuencia de aminoácidos como figura en cualquiera de las SEQ ID NOs, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 21 ó 22,

f) un ácido nucleico que está degenerado como resultado del código genético con una secuencia de nucleótidos de un ácido nucleico como figura en cualquiera de las SEQ ID NOs 1, 2, 5, 7, 9, 11, 13, 15, 17 ó 19 ó como se define en (a) hasta (e),

g) un ácido nucleico que es divergente debido a las diferencias en el uso del codón entre los organismos con una secuencia de nucleótidos que codifica un polipéptido ó proteína tal como figura en cualquiera de las SEQ ID NOs 3, 4, 6, 8, 10, 12, 14, 16, 18, 20 ó 21 o como se define en (a) hasta (e),

h) un ácido nucleico que es divergente debido a las diferencias en alelos que codifican un polipéptido ó proteína como figura en cualquiera de las SEQ ID NOs, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20 ó 21 o como se define en (a) hasta (e),

i) un ácido nucleico que codifica un fragmento inmunológicamente activo y/o funcional de un polipéptido ó proteína codificado por una secuencia de ADN como figura en cualquiera de las SEQ ID NOs 1, 2, 5, 7, 9, 11, 13, 15, 17 ó 19 o como se define en (a) hasta (e),

j) un ácido nucleico que codifica una proteína o polipéptido como se define en las SEQ ID Nos. 3, 4, 6, 8, 10, 12, 14, 16, 18, 20 ó 21 o como se define en (a) hasta (i) caracterizado porque dicha secuencia es de ADN, ADNc, ADN genómico ó ADN sintético.

Debe entenderse que la presente invención también se relaciona con cualquiera de los ácidos nucleicos definidos en a) hasta j) para uso en cualquiera de los métodos descritos anteriormente.

La invención se relaciona además con una molécula de ácido nucleico de al menos 15 nucleótidos de longitud que hibrida específicamente con, o que amplifica específicamente uno de los ácidos nucleicos de la invención, preferiblemente aquellos ácidos nucleicos como los definidos en a) hasta j).

La invención también se relaciona con un vector que comprende un ácido nucleico de la invención, en donde dicho vector preferiblemente es un vector de expresión en donde el ácido nucleico está operativamente enlazado a una o más secuencias de control permitiendo la expresión de dicha secuencia en células procariotas y/o eucariotas. Como tal, la invención también se relaciona con una célula huésped que contiene un ácido nucleico o un vector de la invención. Dicha célula huésped puede ser escogida de una célula bacteriana, de insecto, de hongo, de levadura, de planta o de animal.

De acuerdo con aún otra modalidad, la invención se relaciona con el uso de un ácido nucleico natural o sintético que codifica una proteína que contiene un "dominio RS" como se definió anteriormente para uso en cualquiera de los métodos descritos aquí.

La invención también se relaciona con el uso de un ácido nucleico natural o sintético que codifica una proteína involucrada en el proceso del procesamiento de precursores de ARN mensajero en células eucariotas y en cualquiera de los métodos de la invención.

La invención se relaciona además con el uso de un ácido nucleico con al menos 50% de identidad con al menos una de las secuencias descritas aquí en un método para mejorar la tolerancia al estrés que comprende la manipulación del proceso o ruta de procesamiento de precursores de ARNm. Preferiblemente dicho ácido nucleico se origina a partir de una célula u organismo eucariota.

También están incluidas dentro de la invención moléculas antisentido correspondientes a al menos uno de los ácidos nucleicos de la invención y su uso en cualquiera de los métodos de la invención.

De acuerdo con aún otra modalidad, la presente invención se relaciona con un polipéptido codificable por al menos uno de los ácidos nucleicos de la invención, o un homólogo de los mismos o un derivado de los mismos, o un fragmento inmunológicamente activo y/o funcional de los mismos, siendo dicho polipéptido natural, sintético, enriquecido, aislado, libre de la célula y/o recombinante. Cada uno de estos polipéptidos puede ser utilizado en cualquiera de los métodos de la invención.

Los polipéptidos preferidos son aquellos que comprenden una secuencia de aminoácidos tal como figura en cualquiera de las SEQ ID NOs 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 21 ó 22 un homólogo de las mismas o un derivado de las mismas, o un fragmento inmunológicamente activo y/o funcional de las mismas.

- 5 La invención también se relaciona con un método para producir un polipéptido de la invención que comprende el cultivo de una célula huésped como se describió anteriormente bajo las condiciones que permitan la expresión del polipéptido y la recuperación del polipéptido producido del cultivo. La invención también se relaciona con un método para la producción de plantas transgénicas, células de la planta o tejidos de la planta que comprende la introducción de un ácido nucleico de la invención en un formato expresable o un vector de la invención en dicha planta, célula de la planta o tejido de la planta.

10 La invención también se relaciona con un método para la producción de plantas alteradas, células de planta o tejidos de planta que comprende la introducción de un polipéptido de la invención directamente en una célula; un tejido o un órgano de dicha planta.

- 15 La invención se relaciona además con un método para efectuar la expresión de un polipéptido de la invención que comprende la introducción de un ácido nucleico de la invención operativamente enlazado a una o más secuencias de control o a un vector de la invención en forma estable dentro del genoma de una célula de la planta.

- 20 La invención también se relaciona con dichos métodos para producir plantas transgénicas, que comprende además la regeneración de una planta a partir de dicha célula de la planta.

25 La invención se relaciona además con una célula de una planta transgénica que puede ser obtenida por uno de los métodos anteriormente mencionados en donde dicho ácido nucleico está integrado en forma estable dentro del genoma de dicha célula de la planta.

- 30 De acuerdo con la invención, las plantas transgénicas tolerantes al estrés por salinidad pueden ser producidas como resultado de la expresión de al menos uno de los ácidos nucleicos de la reivindicación 13 o al menos uno de los polipéptidos de la reivindicación 25 ó 26 o una molécula antisentido de la reivindicación 23. Dichas plantas transgénicas también hacen parte de la invención, así como las plantas transgénicas que como resultado de la expresión de al menos uno de los ácidos nucleicos o moléculas antisentido de la invención o al menos uno de los polipéptidos de la invención muestran una alteración de su fenotipo. La invención se relaciona con cualquier parte cosechable de una planta de la invención que se selecciona preferiblemente del grupo que consiste de semillas, hojas, frutos, cultivos de tallo, rizomas, raíces, tubérculos y bulbos. También la progenie derivada de cualquiera de las plantas o partes de la planta de la invención hace parte de la presente invención.

40 De acuerdo con otra modalidad, la invención se relaciona con un método para mejorar la tolerancia al estrés en (a) un(a) planta(s) que comprende la expresión de al menos uno de los ácidos nucleicos o al menos uno de los polipéptidos en una molécula antisentido de la invención en células, tejidos o partes de dicha(s) planta(s).

De acuerdo con otra modalidad, la invención se relaciona con un método para mejorar la tolerancia al estrés en (a) un(a) planta(s) que comprende la expresión de al menos uno de los ácidos nucleicos o al menos uno de los polipéptidos en una molécula antisentido de la invención en células, tejidos o partes de dicha(s) planta(s).

- 45 Debe ser claro que la tolerancia al estrés en los métodos anteriores puede significar cualquier estrés provocado por el medio ambiente tal como, pero sin limitarse a estrés osmótico, estrés por salinidad, estrés por sequía, estrés por congelación o estrés por frío.

50 Además, cualquiera de los métodos, los ácidos nucleicos o moléculas antisentido o los polipéptidos de la invención pueden ser utilizados (en un método) para incrementar el rendimiento, para estimular el crecimiento que puede ser el cualquier parte de esa planta, tal como raíz, hoja, semilla.

La invención también se relaciona con una planta que puede ser obtenida por medio de cualquiera de los métodos descritos anteriormente para cultivos sobre suelo con altas concentraciones de sal, preferiblemente suelos con un contenido de sal de más de 1 mM de iones de sal.

También forman parte de la invención nuevas cepas de levadura u otras eucariotas unicelulares más tolerantes al estrés por salinidad como resultado de la expresión de cualquiera de los ácidos nucleicos y/o proteínas de la invención.

5 La invención se relaciona además con un sistema de cultivo de células *in vitro* que comprende células huésped, de plantas o de animales como se definió anteriormente, tolerantes a la sal, obtenidas como resultado de la expresión de al menos uno de los ácidos nucleicos, vectores, polipéptidos, o molécula antisentido de la reivindicación 23. La invención se relaciona además con al menos una aplicación terapéutica en humanos derivada de los métodos descritos aquí.

10 La invención también se relaciona con un anticuerpo que reconoce específicamente un polipéptido de la invención o un epítipo específico del mismo.

15 La invención se relaciona además con una composición para diagnóstico que comprende al menos un ácido nucleico, un vector, una molécula antisentido, un polipéptido o un anticuerpo de la invención.

Definiciones y elaboración de las modalidades

Definiciones específicas

20 “Manipulación de un proceso” significa aquí la interferencia con, o la modulación de ese proceso, preferiblemente el mejoramiento, la catálisis, el cambio o la alteración de ese proceso. Esta interferencia puede tener un impacto sobre cada etapa o cada componente, o cada producto o cada resultado de ese proceso. También, esta interferencia puede tener un impacto sobre la eficiencia, la velocidad o el rendimiento de ese proceso.

25 “Manipulación genética de un proceso” se refiere aquí a la manipulación de un proceso por medio de cualquier clase de interferencia con las secuencias genéticas (por ejemplo, secuencias de nucleótidos, ARN, ADN) que están involucradas en ese proceso. Después de la actividad genética natural de la célula tal como replicación, transcripción, traducción, y el procesamiento de diferentes ácidos nucleicos, también están incluidas dentro del
30 término “manipulación genética de un proceso”, técnicas de biología molecular y técnicas de tecnología génica. Estas técnicas genéticas artificiales son conocidas por las personas capacitadas en el arte y son por ejemplo clonación, transformación, recombinación, expresión, sobreexpresión, silenciamiento, etc. Como ejemplo de “manipulación genética de un proceso” se puede interferir con la secuencia genética que codifica una proteína, que está involucrada en el proceso del procesamiento de precursores de ARN mensajero. Alternativamente, se puede
35 interferir con una molécula de ARN (tal como un ARN nucleolar pequeño), que está directamente involucrada en el proceso del procesamiento de precursores de ARN mensajero.

En el caso en que dicha secuencia genética codifique una proteína y la expresión, la constitución, la estructura o la localización de esa secuencia de codificación sean alteradas, se utiliza el término “manipulación genética de una
40 proteína”. Más particularmente se puede alterar la composición o la expresión de dichas secuencias de codificación o se puede introducir o suprimir dicha secuencia de codificación en la célula huésped, que lleva a cabo dicho proceso. Más específicamente, dicha proteína puede ser cualquier componente que esté involucrado en el procesamiento de pre-ARNm, tal como, pero sin limitarse a un componente del complejo U1-snRNP o U2-snRNP, un factor de transcripción, una proteína de enlazamiento de ARN o una proteína de movimiento nuclear.

45 “Manipulación bioquímica de un proceso” se refiere aquí a la manipulación de dicho proceso por medio del uso de métodos bioquímicos que interfieren con dicho proceso. Con métodos bioquímicos significa el uso de cualquier sustancia (compuestos químicos, péptidos y moléculas) que tienen un impacto sobre los procesos biológicos. Por ejemplo se puede introducir un péptido, o una proteína, o una sustancia bioquímica o química en la célula que lleva
50 a cabo ese proceso, con el propósito de interferir con ese proceso. Más particularmente, se puede introducir una proteína o un péptido derivado de los genes de la presente invención directamente dentro de la célula, con el propósito de mejorar el proceso del procesamiento de precursores de ARN mensajero.

En el caso en que dicho método bioquímico involucre el uso de una proteína o de un péptido, o en el caso en que dicho método bioquímico resulte en la modificación de una proteína, utilizamos el término “manipulación bioquímica de una proteína”.

- 5 “Células” es tomado aquí en su contexto más amplio e incluye a cada célula viviente, tal como células procariotas y eucariotas.

10 “Precursor de ARN mensajero” se refiere aquí a cualquier molécula de ARN, que aún no es operacional como ARN mensajero maduro, a partir de la cual se pueden transcribir los polipéptidos, debido a que su composición, su estructura o su localización tienen que ser cambiadas.

15 “Procesamiento de precursores de ARN mensajero” se refiere aquí a cualquier proceso, que cambia la composición, la estructura o la localización de un precursor de ARN mensajero. Los ejemplos de tales procesos en una célula eucariota son la síntesis del precursor de ARN mensajero durante la transcripción, la modificación de los extremos 5’ y/o 3’ del precursor tal como una poliadenilación, el empalme de intrones a partir del precursor como el empalme constitutivo o el alternativo, el metabolismo del precursor o el traslado del precursor hacia la envoltura nuclear y el transporte del ARN maduro a los citoplasmas, etc. También el acoplamiento de las diferentes etapas de este procesamiento o metabolismo del ADN precursor hacen parte del término “procesamiento de precursores de ARN mensajero” como se lo utiliza en esta descripción. El procesamiento post-transcripcional del ARN precursor
20 comprende una ruta compleja de escisiones endonucleolíticas, digestión exonucleolítica y modificaciones covalentes. El orden general de las diferentes etapas del procesamiento está bien conservado en células eucariotas, pero los mecanismos subyacentes son en gran medida desconocidos. El procesamiento pre-ARNm es una actividad celular importante y ha sido estudiado en las células de levadura de *Saccharomyces cerevisiae* por Venema y Tollervey (Yeast, (1995) 11 (16): 1629 - 1650). Las etapas del procesamiento involucran una variedad de interacciones proteína - proteína así como interacciones proteína - ARN y ARN - ARN. El papel preciso de los
25 diferentes factores de transcripción, proteínas de enlazamiento de ARN, y otras proteínas nucleares tales como las proteínas de movimiento nuclear y del ARN nuclearen el procesamiento del pre-ARN permanece en gran medida desconocido. Por lo tanto, una proteína que interfiere con el proceso del procesamiento de precursores de ARN mensajero, puede ser una proteína de muchas clases diferentes. El transporte nuclear de pre-ARN involucra
30 diferentes factores que pueden ser utilizados en el método de la presente invención, ya que están involucrados en el procesamiento de pre-ARN. El pre-ARN se transcribe principalmente desde los genes localizados en la interfaz entre los dominios de cromatina y el espacio intercromatínico. Después del procesamiento parcial o completo y de la formación del complejo con proteínas nucleares, los transcriptos dejan su sitio de síntesis y viajan a través del espacio intercromatínico a los poros nucleares, donde son capturados por la maquinaria de exportación para
35 exportarlos hacia el citoplasma. Los ARNm competentes para el transporte forman un complejo con el complemento correcto de proteínas nucleares (revisado en Politz y Pederson, J. struct. Biol 2000, 129 (2 - 3): 252 - 257). El papel de la U1-snRNP y de la U2-snRNP está principalmente situado en el empalme de pre-ARNm. El producto de la particular de ribonucleoproteína nuclear pequeña U1 o del complejo del gen U1-snRNP de 70K (U1-70K), una proteína específica de U1-sn-RNP, ha sido implicada en el empalme básico así como alternativo o pre-ARNm tanto
40 en animales como en plantas (Golovkin y Reddy, Plant Cell, 1996, (8): 1421 - 1435). En otros reportes se describen diferentes proteínas de interacción de la proteína U1-70K, tal como por ejemplo una proteína del tipo SC35 y proteína rica en serina/arginina (Golovkin y Reddy, J. Biol Chem 1999, 245(51): 36428 - 36438). Para la ribonucleoproteína nuclear pequeña U2 se han descrito factores auxiliares tales como U2A (Domon et al, J. Biol Bhem 1998: 273(51): 34603 - 34610). Dentro del alcance de la presente invención están cualquiera de las proteínas,
45 que están involucradas en el empalme de los precursores de ARN mensajero. Se ha descrito el papel de los ARN nucleolar pequeño en el procesamiento de pre-ARNm en plantas (Brown y Shaw, the Plant Cell, 1998, 10: 649 - 657). Por lo tanto la manipulación de los procesos del procesamiento de pre-ARNm puede ser establecida también interfiriendo con moléculas de ARN. Jarrous et al (J. Cell Biol. 1999, 146(3): 559 - 572) mostraron que Rpp29 y Rpp38 (subunidades de proteína de ARNasaP humana) se encuentran en los nucleolos y que ellos residen en
50 cuerpos enrollados en espiral, organelos que están implicados en la biogénesis de varias otras ribonucleoproteínas nucleares pequeñas requeridas para el procesamiento de ARNm precursor. Estos tipos de proteínas que hacen parte de un complejo de Ribonucleoproteína-Ribonucleasa están involucradas en el procesamiento de pre-ARNm pueden ser también utilizadas por lo tanto en el método de la presente invención. El papel de las proteínas de

movimiento nuclear no está bien establecido en el estado del arte. Incluso una cantidad de moléculas involucradas en el procesamiento de precursores de ARN mensajero deben ser dilucidadas.

“Proteínas tipo SR” como se las utiliza aquí fueron descritas previamente en Blencowe et al. 1999, 77(4): 277 - 291:

“El procesamiento de precursores de ARN mensajero (pre-ARNm) con ARNm requiere de un grana gran cantidad de proteínas que contienen dominios ricos en residuos alternantes de arginina y serina (dominios RS). Estos incluyen a miembros de la familia SR de factores de empalme y de proteínas que son estructural y funcionalmente distintos de la familia SR colectivamente denominados más abajo como proteínas relacionadas con SR. Ambos grupos de proteínas del dominio RS actúan en el empalme de pre-ARNm constitutivo y regulado. Recientemente, se han identificado diferentes proteínas relacionadas con SR que están asociadas con la maquinaria transcripcional. Otras proteínas relacionadas con SR están asociadas con la formación del extremo 3' de ARNm y han sido implicadas en la exportación”. La evidencia de que las proteínas que contienen dominios RS pueden jugar un papel fundamental en la coordinación de diferentes etapas en la síntesis y procesamiento de pre-ARNm es revisada adicionalmente en Blencowe et al. 1999, 77(4): 277 - 291.

Definiciones generales

Los términos “proteína(s)”, “péptido(s)” u “oligopéptido(s)” o polipéptido, cuando se los utiliza aquí se refieren a aminoácidos en una forma polimérica de cualquier longitud. Dichos términos también incluyen modificaciones conocidas de aminoácidos tales como la formación de un enlace disulfuro, cisteinilación, oxidación, glutationilación, metilación, acetilación, farnesilación, biotinilación, estearoilación, formilación, adición lipóica ácida, fosforilación, sulfatación, ubiquitanición, miristoilación, palmitoilación, geranilgeranilación, ciclización (por ejemplo formación de ácido piroglutámico), oxidación, desamidación, deshidratación, glicosilación (por ejemplo pentosas, hexosaminas, N-acetilhexosaminas, desoxihexosas, hexosas, ácido siálico, etc.) acilación y marcadores radioactivos (por ejemplo, ^{125}I , ^{131}I , ^{35}S , ^{14}C , ^{32}P , ^{33}P , ^3H) así como residuos aminoácidos de ocurrencia no natural, residuos de L-aminoácido y residuos de D-aminoácidos.

“Homólogos” de una proteína de la invención son aquellos péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que contienen sustituciones, supresiones y/o adiciones de aminoácidos con relación a dicha proteína con respecto a la cual son homólogas sin alterar una o más de sus propiedades funcionales, en particular sin reducir la actividad de las resultantes. Por ejemplo, un homólogo de dicha proteína consistirá de una variante de una secuencia bioactiva de aminoácidos de dicha proteína. Para producir tales homólogos, se pueden reemplazar aminoácidos presentes en dicha proteína por otros aminoácidos que tengan propiedades similares, por ejemplo Hidrofobicidad, hidrofiliidad, momento hidrófobo, antigenicidad, propensión a formar o a romper estructuras de hélice α o estructuras de lámina β , etc.

Las variantes de sustitución de una proteína de la invención son aquellas en las cuales se ha removido al menos un residuo en dicha secuencia de aminoácidos de la proteína y se ha insertado un residuo diferente en su lugar. Las sustituciones de aminoácidos son típicamente de residuos individuales, pero pueden agruparse dependiendo de las limitaciones funcionales impuestas por el polipéptido; las inserciones serán usualmente del orden de aproximadamente 1 - 10 residuos aminoácidos y las supresiones estarán en el rango aproximadamente de 1 - 20 residuos. Preferiblemente, las sustituciones de aminoácidos incluirán sustituciones conservadoras de aminoácidos, tales como aquellas descritas más arriba.

Las variantes de inserción de la secuencia de aminoácidos de una proteína de la invención son aquellas en las cuales se introducen uno o más residuos aminoácidos en un sitio predeterminado en dicha proteína. Las inserciones pueden incluir fusiones en el terminal carboxilo y/o en el terminal amino así como inserciones dentro de la secuencia de múltiples aminoácidos o de aminoácidos individuales. Generalmente, las inserciones dentro de la secuencia de aminoácidos serán menores que las fusiones en los terminales amino o carboxilo, aproximadamente del orden de 1 a 10 residuos. Los ejemplos de proteínas o péptidos de fusión en el terminal amino o carboxilo incluyen al dominio de enlazamiento o al dominio de activación de un activador transcripcional como el utilizado en el sistema híbrido doble de levadura, proteínas de recubrimiento de fago, etiqueta de (histidina) $_6$, glutatona S-transferasa, proteína A, proteína de enlazamiento de maltosa, dihidrofolato reductasa, epítipo de Tag•100 (EETARFQPGYRS), epítipo de

c-myc (EQKLISEEDL), epítipo de FLAG® (DYKDDDK), lacZ, CMP (péptido de enlazamiento de calmodulina), epítipo de HA (YPYDVPDYA), epítipo de proteína C (EDQVDPRLIDGK) y epítipo de VSV (YTDIEMNRLGK).

Las variantes de supresión de una proteína de la invención se caracterizan por la remoción de uno o más aminoácidos de la secuencia de aminoácidos de dicha proteína.

5 Las variantes de aminoácidos de una proteína de la invención pueden ser elaboradas fácilmente utilizando técnicas de síntesis de péptidos conocidas en el arte, tales como síntesis de péptidos en fase sólida y similares, o por medio de manipulaciones de ADN recombinante. La manipulación de secuencias de ADN para producir proteínas
10 variantes, que se manifiestan como variantes de sustitución, de inserción o de supresión conocidas en el arte. Por ejemplo, las técnicas para elaborar mutaciones de sustitución en sitios predeterminados en el ADN que tienen secuencias conocidas son conocidas por aquellos capacitados en el arte, tales como por medio de mutagénesis de M13, un kit de mutagénesis *in vitro* del Gen de T7 (USB, Cleveland, OH), un kit de mutagénesis dirigida al sitio QuickChange (Stratagene, San Diego, CA), mutagénesis dirigida al sitio mediada por PCR u otros protocolos de mutagénesis dirigida al sitio. Otra alternativa para manipular secuencias de ADN para producir proteínas variantes,
15 que se manifiestan como variantes de sustitución, de inserción o de supresión comprende una modificación dirigida del gen *in vivo* que puede lograrse por medio de oligonucleótidos quiméricos de ARN/ADN como describen por ejemplo (Palmgren 1997; Yoon *et al.* 1996).

20 “Derivados” de una proteína de la invención son aquellos péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que comprenden al menos aproximadamente cinco residuos aminoácidos contiguos de dicho polipéptido pero que contienen la actividad biológica de dicha proteína. Un “derivado” puede incluir además residuos aminoácidos adicionales de ocurrencia natural, glicosilados alterados, acilados o no naturales comparado con la secuencia de aminoácidos de una forma de ocurrencia natural de dicho polipéptido. Alternativamente o adicionalmente, un derivado puede incluir uno o más sustituyentes no aminoácidos, comparado con la secuencia de aminoácidos de
25 una forma de ocurrencia natural de dicho polipéptido, por ejemplo una molécula reportera u otro ligando enlazado en forma covalente o no covalente a la secuencia de aminoácidos tal como, por ejemplo, una molécula reportera que está enlaza allí para facilitar su detección.

30 Con “inmunológicamente activo” se entiende que una molécula o fragmentos específicos de la misma tales como epítopos o haptenos son reconocidos por, es decir enlazados a anticuerpos.

En el contexto de la presente invención se incluyen también homólogos, derivados y/o fragmentos inmunológicamente activos de cualquiera de los polipéptidos de la invención.

35 “Anticuerpos” incluye a los anticuerpos de camello monoclonales, policlonales, sintéticos o de cadena pesada así como fragmentos de anticuerpos tales como Fab, Fv o fragmentos de scFv. Los anticuerpos monoclonales pueden ser preparados por medio de técnicas como las descritas previamente por ejemplo (Liddle & Cryer 1991) que comprende la fusión de células de mieloma de ratón con células de bazo derivadas de animales inmunizados. El término “anticuerpos” incluye además derivados de los mismos tales como anticuerpos etiquetados. Las etiquetas de anticuerpo incluyen fosfatasa alcalina, PKH2, PKH26, PKH67, fluoresceína (FITC), Hoechst 33258, R-ficoeritrina (PE), rodamina (TRITC), Quantum Red, Texas Red, Cy3, biotina, agarosa, peroxidasa, esferas doradas y marcadores radioactivos (por ejemplo ¹²⁵I, ¹³¹I, ³⁵S, ¹⁴C, ³²P, ³³P, ³H). Las herramientas en biología molecular que cuentan con anticuerpos contra una proteína incluyen análisis de transferencias en gel de proteína, selección de bibliotecas de expresión que permiten identificación de los genes, métodos cuantitativos de proteína que incluyen
40 ELISA y RIA, purificación de inmutafinidad de proteínas, inmunoprecipitación de proteínas por ejemplo (Magyar *et al.* 1997) e inmunolocalización. Otros usos de los anticuerpos y especialmente de anticuerpos de proteína incluyen el estudio de procesamiento proteolítico (Loffler *et al.* 1994; Woulfe *et al.* 1994), determinación de sitios activos de proteína (Lerner 1982), el estudio de procesamiento del precursor y post-traducción (Baron & Baltimore 1982; Lerner *et al.* 1981; Semler *et al.* 1982), identificación de dominios de proteína involucrados en las interacciones proteína-proteína (Murakami *et al.* 1992) y el estudio del uso de exones en expresión génica (Tamura *et al.* 1991).
50

Dentro del alcance de la actual invención están también anticuerpos que reconocen a las proteínas de la presente invención o un homólogo, derivado o fragmento de los mismos como se definió más arriba.

Los términos “gen(es)”, “polinucleótido(s)”, “secuencia(s) de ácido nucleico”, “secuencia(s) de nucleótido(s)”, “secuencia(s) de ADN” o “molécula(s) de ácido nucleico”, cuando se los utiliza aquí se refieren a nucleótidos, ya sea ribonucleótidos o desoxirribonucleótidos o una combinación de ambos, en una forma polimérica de cualquier longitud. Dichos términos incluyen además ADN monocatenario y bicatenario y ARN. Dichos términos también incluyen modificaciones conocidas de nucleótidos tales como metilación, ciclización y ‘taponamientos’ y la sustitución de uno o más de los nucleótidos de ocurrencia natural con un análogo tal como inosina. Dichos términos también abarcan ácidos nucleicos del péptido (PNA), un análogo de ADN en el cual la columna vertebral es un pseudopéptido que consiste de unidades de N-(2-aminoetil)-glicina en vez de un azúcar. Los PNA imitan el comportamiento del ADN y enlazan cadenas complementarias de ácido nucleico. La columna vertebral neutra de PNA resulta en un enlazamiento más fuerte y en una especificidad mayor que lo logrado normalmente. Además, las propiedades químicas, físicas y biológicas únicas de PNA han sido explotadas para producir herramientas biomoleculares poderosas, agentes antigénicos y antisentido, sondas moleculares y biosensores.

Con “molécula de ADN recombinante” o “gen quimérico” se entiende un ADN híbrido producido por medio de la unión de piezas de ADN de diferentes fuentes. Con “secuencia heteróloga de nucleótidos” se entiende una secuencia que no se presenta en forma natural con la secuencia promotora. Aunque esta secuencia de nucleótidos es heteróloga a la secuencia promotora, puede ser homóloga, o nativa, o heteróloga, o extraña, al huésped de la planta. “Cadena sentido” se refiere a la cadena de una molécula bicatenaria de ADN que es homóloga a un transcripto de ARNm de la misma. La “cadena antisentido” contiene una secuencia invertida, que es complementaria a aquella de la “cadena sentido”.

Una “secuencia de codificación” o un “marco abierto de lectura” o “ORF” se define como una secuencia de nucleótidos que puede ser transcrita en ARNm y/o traducida en un polipéptido cuando se lo coloca bajo el control de secuencias reguladoras apropiadas, es decir, cuando dicha secuencia de codificación u ORF está presente en un formato expresable. Dicha secuencia de codificación de ORF es enlazada por un codón de inicio de la traducción 5’ y un codón de detención de la traducción 3’. Una secuencia de codificación u ORF puede incluir, pero no se limita a ARN, ARNm, ADNc, secuencias recombinantes de nucleótido, secuencias de nucleótidos elaboradas en forma sintética o ADN genómico. Dicha secuencia de codificación u ORF puede ser interrumpida por medio de la intervención de secuencias de ácido nucleico. Los genes y secuencias de codificación que codifican esencialmente la misma proteína pero aislados de diferentes fuentes pueden consistir sustancialmente de secuencias divergentes de ácido nucleico. Recíprocamente, se pueden diseñar secuencias de ácido nucleico sustancialmente divergentes para llevar a cabo la expresión esencialmente de la misma proteína. Dichas secuencias de ácido nucleico son el resultado por ejemplo de la existencia de alelos diferentes de un gen dado, o de la degeneración del código genético o de diferencias en el uso del codón. De este modo aminoácidos tales como la metionina y el triptófano son codificados por un codón único mientras que otros aminoácidos tales como arginina, leucina y serina pueden ser traducidos cada uno a partir de hasta seis codones diferentes. Las diferencias en el uso de un codón preferidos se ilustran más abajo para *Agrobacterium tumefaciens* (una bacteria), *A. thaliana*, *M. sativa* (dos plantas dicotiledóneas) y *Oryza sativa* (un planta monocotiledónea). Estos ejemplos fueron extraídos de (<http://www.kazusa.or.jp/codon>). Para dar un ejemplo, el codón GGC (para glicina) es el codón más frecuentemente utilizado en *A. tumefaciens* (36.2 ‰), es el segundo codón más frecuentemente utilizado en *O. sativa* pero se utiliza con mucha menos frecuencia en *A. thaliana* y *M. sativa* (9 ‰ y 8.4 ‰, respectivamente). De los cuatro codones posibles que codifican glicina (ver la Tabla 2), se utiliza más preferiblemente dicho codón GGC en *A. tumefaciens* y *O. sativa*. Sin embargo, en *A. thaliana* este es el codón GAGA (y GGU) mientras que en *M. sativa* este es el codón GGU (y GGA). Las variantes alélicas están definidas adicionalmente ya que comprenden polimorfismos únicos de nucleótidos (SNP) así como polimorfismos pequeños de inserción/supresión (los INDEL; el tamaño de los INDEL es usualmente menos a 100 b pb). Los SNP y los INDEL forman el conjunto más grande de variantes de secuencias en cadenas polimórficas de ocurrencia natural de la mayoría de los organismos. Ellos son útiles en el mapeo de genes y en el descubrimiento de genes y de funciones de los genes. Son además útiles en la identificación de loci genéticos, por ejemplo genes de plantas involucrados en procesos de determinación tales como velocidad de crecimiento, tamaño de la planta y rendimiento de la planta, vigor de la planta, resistencia a las enfermedades, tolerancia al estrés, etc. Hoy en día se encuentran disponibles muchas técnicas para identificar los SNP y/o los INDEL incluyendo (i) PCR seguida por cromatografía líquida desnaturalizante de alto rendimiento (DHPLC; por ejemplo (Cho *et al.* 1999)); (ii) electroforesis capilar desnaturalizante constante (CDCE) combinada con PCR de alta fidelidad (por ejemplo (Li-Sucholeiki *et al.* 1999)); (iii) electroforesis desnaturalizante en gradiente de gel (por ejemplo Fischer y Lerman 1983); (iv)

espectrometría de masas con detector de iones de tiempo de vuelo de desorción/ionización láser asistida por matriz (MALDI-TOF MS; por ejemplo (Ross *et al.* 2000)); (v) ensayos de PCR para monitoreo de fluorescencia en tiempo real (por ejemplo Tapp *et al.* 2000); (vi) tecnología en gel de Acrydite™ (por ejemplo Kenney *et al.* 1998); (vii) ciclo de formación de huellas digitales con didesoxi (CddF; por ejemplo (Langemeier *et al.* 1994); (viii) análisis de polimorfismo de conformación monocatenaria (SSCP) (por ejemplo (Vidal-Puig & Moller 1994)) y (ix) reacción de extensión del iniciador por mini-secuenciación (por ejemplo Syvanen 1999). La técnica de Detección de Lesiones Locales Inducidas en Genomas (TILLING; (McCallum *et al.* 2000a; McCallum *et al.* 2000b)), que es una variante de (i) ver más arriba, puede ser aplicada también para identificar rápidamente un gen alterado por ejemplo en plantas individuales mutadas químicamente que muestran fenotipos interesantes.

“Hibridación” es el proceso en donde se hibridan entre sí secuencias complementarias de nucleótidos sustancialmente homólogas. El proceso de hibridación puede presentarse completamente en solución, es decir ambos ácidos nucleicos complementarios están en solución. Las herramientas en biología molecular que cuentan con tal proceso incluyen la reacción en cadena de la polimerasa (PCR; y todos los métodos basados en la misma), hibridación sustractiva, extensión aleatoria del iniciador, mapeo de la nucleasa S1, extensión del iniciador, transcripción inversa, síntesis de ADNc, expresión diferencial de los ARN, y determinación de la secuencia de ADN. El proceso de hibridación puede presentarse también con uno de los ácidos nucleicos complementarios inmovilizados a una matriz tal como cuentas magnéticas, cuentas de Sefarosa o cualquier otra resina. Las herramientas en biología molecular que cuentan con tal proceso incluyen el aislamiento de poli(A+) ARNm. El proceso de hibridación puede presentarse además con uno de los ácidos nucleicos complementarios inmovilizados a un soporte sólido tal como una membrana de nylon o de nitrocelulosa o inmovilizados por ejemplo por medio de fotolitografía por ejemplo a un soporte de vidrio síliceo (este último conocido como arreglos o microarreglos de ácido nucleico o como chips de ácido nucleico. Las herramientas en biología molecular que cuentan con tal proceso incluyen análisis de transferencias en gel de ARN y ADN, hibridación de colonias, hibridación de placa, hibridación *in situ* e hibridación de microarreglos. Con el propósito de permitir que ocurra la hibridación, generalmente se desnaturalizan térmica o químicamente las moléculas de ácido nucleico para fundir una cadena doble en dos cadenas sencillas y/o para remover horquillas u otras estructuras secundarias de los ácidos nucleicos monocatenarios. La rigurosidad de la hibridación está influenciada por condiciones tales como la temperatura, la concentración de sal y la composición del amortiguador de hibridación. Las condiciones de alta rigurosidad para la hibridación incluyen alta temperatura y/o baja concentración de sal (las sales incluyen NaCl y citrato de Na₃) y/o la inclusión de formamida en el amortiguador de hibridación y/o la disminución de la concentración de compuestos tales como SDS (detergente) en el amortiguador de hibridación y/o la exclusión de compuestos tales como sulfato de dextrano o polietilén glicol (promoción del amontonamiento molecular) del amortiguador de hibridación. Las condiciones convencionales de hibridación están descritas por ejemplo en (Sambrook *et al.* 1989) pero la persona capacitada se dará cuenta que se pueden diseñar numerosas condiciones diferentes de hibridación en función de la homología conocida o la esperada y/o la longitud de la secuencia de ácido nucleico. Con hibridación específica se entiende una hibridación bajo condiciones rigurosas. Se prefieren particularmente las condiciones de hibridación de rigurosidad suficientemente baja para aislar heterólogos de ácidos nucleicos con las secuencias de ADN de la invención definidas más arriba. Los elementos que contribuyen a dicha heterología incluyen alelismo, degeneración del código genético y diferencias en el uso de un codón preferido como se discutió más arriba.

Por lo tanto, el alcance de la presente invención está también relacionado con el uso de las secuencias de ADN de la invención que codifican los polipéptidos de la presente invención, homólogos, derivados y/o fragmentos inmunológicamente activos de los mismos como se define más arriba en cualquier método de hibridación. La invención actual se relaciona también además con secuencias de ADN que hibridan a dichas secuencias de ADN de la invención.

“Amplificación específica” aquí utilizando un método de amplificación que es selectivo y únicamente amplifica una secuencia de ácido nucleico con una composición específica de pares de bases (por ejemplo, reacción en cadena de la polimerasa).

Las secuencias de ADN como se definen en la presente invención pueden ser interrumpidas por medio de secuencias de intervención. Con “secuencias de intervención” se entiende cualquier secuencia de ácido nucleico que interrumpa una secuencia de codificación que contenga dicha secuencia de ADN de la invención o que

interrumpa el formato expresable de una secuencia de ADN que contenga dicha secuencia de ADN de la invención. La remoción de la secuencia de intervención restablece dicha secuencia de codificación o dicho formato expresable. Los ejemplos de dichas secuencias de intervención incluyen intrones, secuencias de ADN que pueden ser movilizadas tales como transposones y etiquetas de ADN tales como por ejemplo un T-ADN. Con "secuencia de ADN que puede ser movilizada" se entiende cualquier secuencia de ADN que pueda ser movilizada como resultado de un evento de recombinación.

Para llevar a cabo la expresión de una proteína en una célula, tejido u órgano, preferiblemente de origen vegetal, cualquier proteína puede ser introducida directamente a dicha célula, por medio de microinyección o de un medio balístico o alternativamente, se puede introducir una molécula aislada de ácido nucleico que codifica dicha proteína dentro de dicha célula, tejido u órgano en un formato expresable.

Preferiblemente, la secuencia de ADN de la invención comprende una secuencia de codificación o un marco abierto de lectura (ORF) que codifica una proteína de la presente invención o un homólogo o derivado de la misma o un fragmento inmunológicamente activo de la misma como se define más arriba. Las proteínas preferidas de la invención incluyen la secuencia de aminoácidos como la presentada en las SEQ ID NOs. 3, 4, 6, 8, 10, 12, 14, 16, 18, 20 y 21.

Con "vector" o "secuencia del vector" se entiende una secuencia de ADN, que puede ser introducida en un organismo por medio de transformación y puede ser mantenida en forma estable en dicho organismo. El mantenimiento del vector es posible por ejemplo en cultivos de *Escherichia coli*, *A. tumefaciens*, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*. Otros vectores tales como fagémidos y vectores cósmidos pueden ser mantenidos y multiplicados en bacterias y/o virus. Las secuencias de los vectores generalmente comprenden un conjunto de sitios únicos reconocidos por enzimas de restricción, el sitio de clonación múltiple (MCS), en donde uno o más secuencias que no son del vector pueden ser insertadas.

Con "secuencias que no son del vector" se entiende por lo tanto una secuencia de ADN que está integrada en uno o más de los sitios del MCS incluido dentro de un vector.

Los "vectores de expresión" forman un subgrupo de vectores que, en virtud de contener las secuencias reguladoras apropiadas que permiten la creación de un formato expresable para la(s) secuencia(s) que no son del vector, permitiendo así la expresión de la proteína codificada por dicha(s) que no son del vector. Se sabe en el estado del arte que los vectores de expresión permiten la expresión de proteína en organismos incluidas bacterias (por ejemplo *E. coli*), hongos (por ejemplo *S. cerevisiae*, *S. pombe*, *Pichia pastoris*), células de insecto (por ejemplo vectores de expresión baculoviral), células animales (por ejemplo células COS o CHO) y células vegetales (por ejemplo vectores de expresión con base en el virus X de la patata, ver, por ejemplo Vance et al. 1998 - WO9844097). Ver también además en esta especificación, vectores de expresión típicos de la planta.

La presente invención claramente incluye cualquier vector o vector de expresión que contenga una secuencia de ADN que no sea del vector que comprende las secuencias de nucleótidos de acuerdo con la presente invención o una secuencia que no es del vector que codifica las proteínas de la presente invención, o el homólogo, derivado y/o fragmento inmunológicamente activo de las mismas como se define más arriba. Como una alternativa a la producción de proteína mediada por el vector de expresión en sistemas biológicos, se puede aplicar síntesis química de proteína. Se pueden fabricar péptidos sintéticos en fase de solución o en fase sólida. La síntesis de péptidos en fase sólida (Merrifield 1963) es, sin embargo, la vía más común e involucra la adición secuencial de aminoácidos para crear una cadena lineal de péptidos.

Por "formato expresable" o "bajo el control de secuencias de control de expresión" se entiende que la molécula aislada de ácido nucleico está en una forma adecuada para ser transcrita en ARNm y/o traducida para producir una proteína, ya sea constitutivamente o después de la inducción por medio de una señal intracelular o extracelular, tal como un estímulo o estrés ambiental (mitógenos, anoxia, hipoxia, temperatura, sal, luz, deshidratación, etc.) o un compuesto químico tal como IPTG (isopropil- β -D-tiogalactopiranosido) o tal como un antibiótico (tetraciclina, ampicilina, rifampicina, kanamicina), hormona (por ejemplo, giberelina, auxina, citoquinina, glucocorticoide, brasinosteroides, etileno, ácido abscísico, etc.), análogo de hormona (ácido iodoacético (IAA), 2,4-D, etc.), metal

(zinc, cobre, hierro, etc.), o dexametasona, entre otros. Como lo saben aquellos capacitados en el arte, la expresión de una proteína funcional puede requerir también una o más modificaciones post-traducción, tales como glicosilación, fosforilación, desfosforilación, o una o más interacciones proteína - proteína, entre otras. Todos estos procesos están incluidos dentro del alcance del término “formato expresable”.

- 5 Preferiblemente, expresión de una proteína en una célula específica, tejido u órgano, preferiblemente de origen vegetal, se efectúa por medio de la introducción y expresión de una molécula aislada de ácido nucleico que codifica dicha proteína, tal como una molécula de ADNc, gen genómico, molécula sintética de oligonucleótido, molécula de ARNm o marco abierto de lectura, a dicha célula, tejido u órgano, en donde dicha molécula de ácido nucleico se coloca operativamente en conexión con secuencias reguladoras apropiadas incluido un promotor, preferiblemente un
- 10 promotor expresable de una planta, y una secuencia de terminación.

- “Secuencia reguladora” se refiere a secuencias de control de ADN, que son necesarias para afectar la expresión de las secuencias de codificación a las cuales están ligadas. La naturaleza de tales secuencias de control difiere dependiendo del organismo huésped. En procariotas, las secuencias de control generalmente incluyen promotores,
- 15 sitios de enlazamiento ribosomal, y terminadores. En eucariotas generalmente las secuencias de control incluyen promotores, terminadores y potenciadores o silenciadores.

- Dentro del alcance de la invención están también las secuencias de nucleótidos como se define en la presente invención, fusionadas a cualquier secuencia reguladora. El término “secuencia de control” pretende incluir, como
- 20 mínimo, todos los componentes cuya presencia es necesaria para expresión, y puede incluir también componentes convenientes adicionales y que determinan cuándo, cuánto y dónde se expresa un gen específico.

- La referencia aquí a un “promotor” se debe tomar en su contexto más amplio e incluye las secuencias reguladoras de la transcripción derivadas de un gen genómico eucariota clásico, incluida la caja TATA que es requerida para la iniciación precisa de la transcripción, con o sin una secuencia de caja CCAAT y elementos reguladores adicionales
- 25 (es decir, secuencias de activación secuencia arriba, potenciadores y silenciadores) que alteran la expresión génica en respuesta a estímulos externos y/o de desarrollo, o en una forma específica del tejido.

- El término “promotor” también incluye a las secuencias reguladoras de la transcripción de un gen procariota clásico, en cuyo caso puede incluir una secuencia de caja -35 y/o secuencias reguladoras de la transcripción de caja -10.
- 30

El término “promotor” es utilizado también para describir una molécula sintética o de fusión o un derivado, que confiere, activa o potencia la expresión de una molécula de ácido nucleico en una célula, tejido u órgano.

- 35 Los promotores pueden contener copias adicionales de uno o más elementos reguladores específicos, para potenciar adicionalmente la expresión y/o para alterar la expresión espacial y/o la expresión temporal de una molécula de ácido nucleico a la cual está operativamente conectada. Tales elementos reguladores pueden estar colocados en forma adyacente a una secuencia promotora heteróloga para dirigir la expresión de una molécula de ácido nucleico en respuesta por ejemplo a cobre, glucocorticoides, dexametasona, tetraciclina, giberelina, cAMP,
- 40 ácido abscísico, auxina, lesiones, etileno, jasmonato o ácido salicílico o para conferir expresión de una molécula de ácido nucleico a células específicas, tejidos u órganos tales como meristemas, hojas, raíces, embriones, flores, semillas o frutos. En el contexto de la presente invención, el promotor es preferiblemente una secuencia promotora expresable por la planta. Los promotores, sin embargo, que también funcionan o únicamente funcionan en células que no son de plantas, tales como bacterias, células de levadura, células de insecto y células animales no están
- 45 excluidas de la invención. Por “expresable por la planta” se entiende que la secuencia promotora incluido cualquiera de los elementos reguladores adicionales añadidos a la misma o contenidos allí, es al menos capaz de inducir, conferir, activar o reforzar la expresión de una célula tejido u órgano de la planta, preferiblemente una célula, tejido u órgano de una planta monocotiledónea o dicotiledónea.

- 50 Los términos “operable por la planta” y “operable en una planta” cuando se los utiliza aquí, con respecto a una secuencia promotora, se entenderá que es equivalente a una secuencia promotora expresable de una planta.

En el presente contexto, un “promotor regulado” o “secuencia promotora regulable” es un promotor que es capaz de conferir expresión sobre un gen estructural en una célula, tejido u órgano particular o grupo de células, tejidos u

órganos de una planta, opcionalmente bajo condiciones específicas, sin embargo generalmente no confiere expresión a través de la planta bajo todas las condiciones. Por lo tanto, una secuencia promotora regulable puede ser una secuencia promotora que confiera expresión sobre un gen al cual está operativamente conectado en una ubicación particular dentro de la planta o alternativamente, a través de la planta bajo un conjunto específico de condiciones, tales como después de la inducción de expresión del gen por medio de un compuesto químico u otro desencadenante. Preferiblemente, el promotor regulable utilizado en la realización de la presente invención confiere expresión en una ubicación específica dentro de la planta, ya sea constitutivamente o después de la inducción, Sin embargo no en la planta completa bajo ninguna circunstancia. Incluido dentro del alcance de tales promotores están las secuencias promotoras específicas de la célula, secuencias promotoras específicas del tejido, secuencias promotoras específicas del órgano, secuencias promotoras génicas específicas del ciclo celular, secuencias promotoras inducibles y secuencias promotoras constitutivas que han sido modificadas para conferir expresión en una parte particular de la planta en un momento dado, tal como por medio de la integración de dicho promotor constitutivo dentro de un elemento genético transponible (*Ac*, *Ds*, *Spm*, *En* u otro transposón). Aquellos capacitados en el arte serán consientes de que un “promotor inducible” es un promotor cuya actividad transcripcional se incrementa o se induce en respuesta a estímulos de desarrollo, químicos, ambientales o físicos. Dentro del alcance de la presente invención están las secuencias de nucleótidos como se define en las reivindicaciones, fusionadas a un promotor inducible por estrés. En forma similar, la persona capacitada comprenderá que un “promotor constitutivo” es un promotor que está transcripcionalmente activo a través de la mayoría, pero no necesariamente de todas las partes de un organismo, preferiblemente una planta, durante la mayoría, pero no necesariamente todas las fases de su crecimiento y desarrollo. Contrariamente el término “promotor ubicuo” es utilizado para indicar un promotor que es transcripcionalmente activo a través de la mayoría, pero no necesariamente todas las partes de un organismo, preferiblemente una planta.

Aquellos capacitados en el arte serán capaces de seleccionar fácilmente secuencias promotoras apropiadas para uso en la regulación de la expresión apropiada de las proteínas de la presente invención como se describió más arriba a partir de fuentes fácilmente disponibles o públicamente disponibles, sin una experimentación excesiva.

Colocando una molécula de ácido nucleico bajo el control regulatorio de una secuencia promotora, o en conexión operable con una secuencia promotora significa posicionar dicha molécula de ácido nucleico de tal manera que la expresión sea controlada por la secuencia promotora. Un promotor está usualmente, pero no necesariamente, posicionado secuencia arriba, o en el extremo 5', y dentro de 2 kb del sitio de iniciación de la transcripción, de la molécula de ácido nucleico que regula. En la construcción de promotores heterólogos, combinaciones genéticas estructurales se prefiere generalmente posicionar al promotor a una distancia del sitio de inicio de la transcripción génica que sea aproximadamente la misma que la distancia entre ese promotor y el gen que controla en su entorno natural (es decir, el gen a partir del cual se deriva el promotor).

“Expresión” significa la producción de una secuencia de proteína o nucleótidos en la célula misma o en un sistema libre de células. Incluye la transcripción en un producto de ARN, una modificación post-transcripcional y/o de traducción a un producto de proteína o polipéptido a partir de un ADN que codifica ese producto, así como posibles modificaciones posteriores a la traducción.

“Operativamente enlazado” se refiere a una yuxtaposición en donde los componentes así descritos están en una relación que les permite funcionar en la forma prevista. Una secuencia de control “operativamente enlazada” a una secuencia de codificación está ligada de tal forma que la expresión de la secuencia de codificación se logra bajo condiciones compatibles con las secuencias de control. En caso de que la secuencia de control sea un promotor, es obvio para una persona capacitada que se utiliza preferiblemente ácido nucleico bicatenario.

El término “terminador” se refiere a una secuencia de ADN en el extremo de una unidad transcripcional que señala la terminación de la transcripción. Los terminadores son secuencias de ADN 3' no traducidas que contienen una señal de poliadenilación, que facilita la adición de secuencias de poliadenilato al extremo 3' de un transcrito primario. Los terminadores activos en células derivadas de virus, levaduras, mohos, bacterias, insectos, aves, mamíferos y plantas son conocidos y están descritos en la literatura. Pueden ser aislados de bacterias, hongos, virus, animales y/o plantas.

Los ejemplos de terminadores particularmente adecuados para uso en las construcciones génicas de la presente invención incluyen al terminador del gen de la nopalina sintasa (NOS) de *Agrobacterium tumefaciens*, la secuencia terminadora del gen de la octopina sintasa (OCS) de *Agrobacterium tumefaciens*, la secuencia terminadora del gen 35S del virus del mosaico de la coliflor (CaMV), la secuencia terminadora de ADP-glucosa pirofosforilasa de *Oryza sativa* (t3'Bt2), la secuencia terminadora del gen de zeína de *Zea mays*, el terminador del gen *rbcs-1A*, y las secuencias terminadoras del gen *rbcs-3A*, entre otros.

Aquellos capacitados en el arte serán consientes de secuencias promotoras adecuadas y secuencias terminadoras que pueden ser adecuadas para uso en la realización de la invención. Tales secuencias pueden ser fácilmente utilizadas sin ninguna experimentación excesiva. Alterando la expresión de un gen se puede reducir la expresión. En el contexto de la actual expresión se prevé la reducción de la expresión de algunas proteínas que están involucradas en el proceso del procesamiento de precursores de ARN mensajero. Esto puede resultar por ejemplo en un ejemplo benéfico sobre el procesamiento del ARN mensajero y de tolerancia al estrés. Cuando dicha proteína era un regulador negativo del proceso. La "reducción de la expresión" como se lo utiliza aquí significa la disminución de los niveles de expresión génica y/o de los niveles de producto génico activo y/o de los niveles de actividad del producto génico. La disminución en la expresión se puede lograr por ejemplo por medio de la adición de las secuencias de codificación o de partes de las mismas en una orientación sentido (si resulta en una co-supresión) o en una orientación antisentido con relación a la secuencia promotora y además por ejemplo por medio de mutagénesis de inserción (por ejemplo, inserción de T-ADN o inserción de un transposón) o por medio de estrategias de silenciamiento de genes como lo describen por ejemplo Angell y Baulcombe 1998 (WO9836083), Lowe et al. 1989 (WO9853083), Lederer et al. 1999 (WO9915682) o Wang et al. 1999 (WO9953050). Las construcciones genéticas que tienen el propósito de silenciar la expresión génica pueden tener la secuencia de nucleótidos de dicho gen (o una o más partes del mismo) contenidas allí en una orientación sentido y/o antisentido con relación a la secuencia promotora. Otro método para reducir la expresión génica incluye el uso de ribozimas, por ejemplo como se describe en Atkins et al. 1994 (WO9400012) Lenée et al. 1995 (WO9503404), Lutziger et al. 2000 (WO0000619), Prinsen et al. 1997 (WO9713865) y Scott et al. 1997 (WO9738116).

La modulación, incluida la disminución en el nivel de productos génicos activos o de la actividad del producto génico puede ser lograda por medio de la administración o exposición de las células, tejidos, órganos u organismos a dicho producto génico, un homólogo, análogo, derivado y/o fragmento inmunológicamente activo de los mismos. La inmunomodulación es otro ejemplo de una técnica capaz de reducir la expresión de los niveles de producto génico activo y/o de actividad del producto génico y comprende la administración de o la exposición a o la expresión de anticuerpos a dicho producto génico o en células, tejidos, órganos u organismos en donde los niveles de dicho producto génico y/o actividad del producto génico han de ser modulados. Tales anticuerpos incluyen "cuerpos de la planta", anticuerpos de cadena sencilla, anticuerpos IgG y anticuerpos de camello de cadena pesada así como fragmentos de los mismos. Dentro del alcance de la presente invención están anticuerpos, que reconocen las proteínas de la presente invención y que pueden ser utilizados para dicha inmunomodulación.

La modulación, incluida la disminución del nivel de productos génicos activos o de la actividad del producto génico se puede lograr adicionalmente por medio de la administración o exposición de células, tejidos, órganos u organismos a un inhibidor o activador de dicho producto génico o la actividad del mismo. Tales inhibidores o activadores incluyen proteínas (que comprenden por ejemplo proteinasas y quinasas) y compuestos químicos identificados de acuerdo con la presente invención.

Por "destino de la célula y/o desarrollo de la planta y/o morfología de la planta y/o bioquímica y/o fisiología" se entiende que una o más características de desarrollo y/o morfológicas y/o bioquímicas y/o fisiológicas de una planta se altera por medio de la realización de una o más etapas relacionadas con la invención descrita aquí.

"Destino celular" se refiere al tipo de célula o características celulares de una célula particular que son producidas durante el desarrollo de la planta o n proceso celular de la misma, en particular durante el ciclo celular o como consecuencia de un proceso del ciclo celular.

"Desarrollo de la planta" o el término "característica del desarrollo de la planta" o un término similar se tomarán, cuando se los utiliza aquí, para significar cualquier proceso celular de una planta que está involucrado en la

determinación del destino del desarrollo de una célula vegetal, en particular del tejido específico o tipo de órgano dentro del cual se desarrollará la célula progenitora. Los procesos celulares relevantes para el desarrollo de una planta serán conocidos por aquellos capacitados en el arte. Tales procesos incluyen, por ejemplo, morfogénesis, fotomorfogénesis, desarrollo de un brote, desarrollo de la raíz, desarrollo vegetativo, desarrollo reproductivo, alargamiento del tallo, floración, y mecanismos reguladores involucrados en la determinación del destino de la célula, en particular un proceso o proceso regulador que involucra el ciclo celular.

“Morfología de la planta” o el término “característica morfológica de la planta” o un término similar cuando se lo utiliza aquí, será entendido por aquellos capacitados en el arte que se refiere a la apariencia externa de una planta, incluida una cualquiera o más características estructurales o combinación de características estructurales de la misma. Tales características estructurales incluyen la forma, el tamaño, la cantidad, la posición, el color, la textura, la disposición, y la formación de un patrón de cualquier célula, tejido u órgano o grupos de células, tejidos u órganos de una planta, incluida la raíz, tallo, hojas, brotes, pecíolos, tricoma, flores, pétalos, estigma, estilo, estamen, polen, óvulos, semillas, embriones, endospermo, recubrimiento de las semillas, aleurona, fibra, frutos, cambium, madera, corazón de la madera, parénquima, aerénquima, elementos de criba, floema o tejido vascular, entre otros.

“Bioquímica de la planta” o el término “característica bioquímica de la planta” o un término similar, cuando se los utiliza aquí, será entendido por aquellos capacitados en el arte que se refiere a los procesos metabólicos y catalíticos de una planta, incluido el metabolismo primario y secundario y los productos del mismo, incluida cualquiera de las moléculas pequeñas, macromoléculas o compuestos químicos, tales como, pero sin limitarse a, almidones, azúcares, proteínas, péptidos, enzimas, hormonas, factores de crecimiento, moléculas de ácido nucleico, celulosas, hemicelulosas, callos, lectinas, fibras, pigmentos tales como antocianinas, vitaminas, minerales, micronutrientes, o macronutrientes, que son producidos por las plantas.

“Fisiología de la planta” o el término “característica fisiológica de la planta” o un término similar, cuando se lo utiliza aquí, se entiende que se refiere a los procesos funcionales de una planta, incluidos los procesos de desarrollo tales como crecimiento, expansión y diferenciación, desarrollo sexual, reproducción sexual, conjunto de semillas, desarrollo de las semillas, llenado del grano, reproducción asexual, división celular, inactividad, germinación, adaptación a la luz, fotosíntesis, expansión de la hoja, producción de fibra, crecimiento secundario o producción de madera, entre otros; respuestas de una planta a factores aplicados en forma externa tales como metales, compuestos químicos, hormonas, factores de crecimiento, factores de estrés del ambiente y del medio ambiente (por ejemplo, anoxia, hipoxia, alta temperatura, baja temperatura, deshidratación, luz, duración del día, inundaciones, sal, metales pesados, entre otros) incluyendo respuesta adaptativas de las plantas a dichos factores aplicados externamente. El término “estrés ambiental” ha sido definido en diferentes formas en el estado del arte y en gran medida se superpone con el término “estrés osmótico”. (Holmberg & Bülow, 1998, Trends plant sci. 3, 61 - 66), define por ejemplo diferentes factores de estrés ambiental que resultan en estrés abiótico. El término estrés osmótico como se lo utiliza aquí se entiende como una situación de estrés inducida por condiciones como salinidad, sequía, calor, enfriamiento (o frío) y congelación. Con el término “estrés ambiental” como se lo utiliza en la presente invención se refiere a cualquier efecto adverso sobre el metabolismo, crecimiento o viabilidad de la célula, tejido, semilla, órgano o la planta completa que es producido por un estrés ambiental no viviente o no biológico. Más particularmente, abarca también factores ambientales tales como estrés por agua (inundación, anegamiento, sequía, deshidratación), estrés anaeróbico (bajo nivel de oxígeno, CO₂, etc.), estrés aeróbico, estrés osmótico, estrés por salinidad, estrés por temperatura (calentamiento/calor, frío, congelación, escarcha) o falta de nutrientes, estrés por contaminantes (metales pesados, compuestos químicos tóxicos), ozono, luz elevada, patógenos (incluidos virus, bacterias, hongos, insectos y nematodos) y combinaciones de estos.

El término “estrés anaeróbico” significa cualquier reducción en los niveles de oxígeno suficiente para producir un estrés como se lo definió aquí anteriormente, incluida hipoxia y anoxia.

El término “estrés por inundación” se refiere a cualquier estrés que esté asociado con o inducido por inmersión prolongada o transitoria de una planta, parte de una planta, tejido o célula aislada en un medio líquido tal como ocurre durante el monzón, temporada lluviosa, inundaciones repentinas o irrigación excesiva de las plantas, etc.

“Estrés por frío” o “estrés por congelación” y “estrés por calor” son estreses inducidos por temperaturas que están respectivamente, por debajo o por encima del rango óptimo de temperaturas de desarrollo para una especie

particular de planta. Tales rangos de temperatura óptimos para el crecimiento se determinan fácilmente o son conocidos por aquellos capacitados en el arte.

“Estrés por deshidratación” es cualquier estrés que está asociado con o inducido por la pérdida de humedad, turgencia reducida o contenido reducido de humedad de una célula, tejido, órgano o planta completa.

“Estrés por sequía” se refiere a cualquier estrés que es inducido por o asociado con la privación de agua o suministro reducido de agua a una célula, tejido, órgano u organismo. “Estrés oxidativo” se refiere a cualquier estrés que incrementa el nivel intracelular de especies reactivas de oxígeno.

Los términos “estrés inducido por salinidad”, “estrés por sal” o “toxicidad iónica por sal” o toxicidad por sal mineral o términos similares se refieren a cualquier estrés que esté asociado con o inducido por concentraciones elevadas de sal y que resulten en una perturbación en el potencial osmótico del ambiente intracelular o extracelular de una célula. Los ejemplos más conocidos de sales minerales que inducen dicho estrés por sal son Na^+ y Li^+ .

Las plantas transgénicas obtenidas de acuerdo con el método de la presente invención, por la presencia del ácido polinucleico y/o de la secuencia reguladora introducida dentro de dicha planta, eran susceptibles a alcanzar una resistencia, tolerancia o una tolerancia o resistencia mejoradas contra estrés ambiental que la correspondiente planta de tipo silvestre.

Los términos “tolerancia”, “resistencia” cubren el rango de protección de una demora para completar la inhibición de la alteración en el metabolismo celular, crecimiento reducido de la célula y/o muerte celular provocada por las condiciones de estrés ambiental definidas aquí anteriormente. Preferiblemente, la planta transgénica obtenida de acuerdo con el método de la presente invención es tolerante o resistente a las condiciones de estrés ambiental en el sentido en que dicha planta es capaz de crecer en forma sustancialmente normal bajo condiciones ambientales donde la correspondiente planta de tipo silvestre muestra un crecimiento, metabolismo, viabilidad, productividad y/o esterilidad de machos y hembras reducidos. Como se lo utiliza aquí, “tolerancia al estrés” se refiere a la capacidad para crecer y producir biomasa durante el estrés, la capacidad para reiniciar el crecimiento y la producción de biomasa después del estrés, y la capacidad para sobrevivir al estrés. El término “tolerancia al estrés” también cubre la capacidad de la planta para someter su programa de desarrollo durante un estrés similar a las condiciones que bajo condiciones no estresadas, por ejemplo para cambiar de inactividad a germinación y de fase vegetativa a reproductiva bajo condiciones estresadas en forma similar que bajo condiciones no estresadas. Las metodologías para determinar el crecimiento en la planta o la respuesta al estrés incluyen, pero no se limitan a, mediciones de altura, área de la hoja, relaciones de agua de la planta, habilidad para florecer, habilidad para generar progenie y rendimiento o cualquier otra metodología conocida por aquellos capacitados en el arte.

“Crecimiento” se refiere a la capacidad de la planta o de partes de la planta para crecer e incrementar su biomasa mientras que “rendimiento” se refiere a la biomasa cosechable de las plantas o de partes de la planta, particularmente aquellas partes de valor comercial. “Crecimiento y/o rendimiento bajo condiciones estresadas y no estresadas” se refiere al hecho de que las plantas que crecieron en el campo casi siempre experimentarán alguna forma de estrés, aunque suaves. Por lo tanto se prefiere no distinguir condiciones no estresadas de condiciones con estrés suave. Como ciertos efectos benéficos de la invención sobre el crecimiento y el rendimiento se espera que ocurran tanto bajo condiciones de estrés severo como de estrés moderado, son descritos por lo tanto como mayor crecimiento y/o rendimiento bajo condiciones estresadas y no estresadas.

Los medios para introducir ADN recombinante en un tejido o en células de una planta incluyen, pero no se limitan a, transformación utilizando CaCl_2 y variaciones de los mismos, en particular el método descrito previamente (Hanahan 1983), incorporación directa de ADN en los protoplastos (Krens *et al.* 1982; Paszkowski *et al.* 1984), captación mediada por PEG a los protoplastos (Armstrong *et al.* 1990), bombardeo de micropartículas, electroporación (Fromm *et al.* 1985), microinyección de ADN (Crossway *et al.* 1986; Fromm *et al.* 1985), bombardeo con micropartículas de explantes de tejido o de células (Christou *et al.* 1988), infiltración al vacío de tejidos con ácido nucleico, o en el caso de plantas, transferencia mediada por T-ADN de *Agrobacterium* al tejido de la planta como se describe esencialmente en (An *et al.* 1985; Dodds 1985; Herrera-Estrella *et al.* 1983a; Herrera-Estrella *et al.* 1983b). Los métodos para transformación de plantas monocotiledóneas son bien conocidos en el arte e incluyen transformación

mediada por *Agrobacterium* (Cheng et al. 1997 - W09748814; Hansen 1998 - WO9854961, Hiei et al. 1994 - WO9400977; Hiei et al. 1998 - WO9817813; Rikiishi et al. 1999 - W09904618; Saito et al. 1995 - WO9506722), bombardeo de microproyectiles (Adams et al. 1999 - US5969213; Bowen et al. 1998 - US5736369; Change et al. 1994 - WO9413822; Lundquist et al. 1999 - US5874265/US5990390; Vasil y Vasil 1995 -US5405765; Walker et al. 1999 - US5955362), incorporación de ADN (Eyal et al. 1993 - WO9318168), microinyección de células de *Agrobacterium* (von Holt 1994 - DE4309203) y sonicación (Finer et al. 1997 - US5693512).

Una planta completa puede ser regenerada a partir de la célula transformada o transfectada, de acuerdo con procedimientos conocidos en el arte. El tejido de la planta capaz de una propagación clonal posterior, ya sea por medio de organogénesis o de embriogénesis, puede ser transformado con una construcción génica de la presente invención y se puede regenerar una planta completa a partir de allí. El tejido particular escogido variará dependiendo de: los sistemas clonales de propagación disponibles para, y más adecuados para, la especie particular que está siendo transformada. Los ejemplos de tejidos objetivo incluyen discos de hojas, polen, embriones, cotiledones, hipocotiledones, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, yemas axilares, y meristemos de raíz), y tejido inducido de meristemo (por ejemplo, meristemo de cotiledón y meristemo de hipocotiledón).

Preferiblemente, la planta producida de acuerdo con el método de la invención es transfectada o transformada con una secuencia genética o sensible a la introducción de una proteína, por cualquier medio reconocido en el estado del arte, tal como bombardeo de microproyectiles, microinyección, transformación mediada por *Agrobacterium* (incluido el método de transformación de "inmersión de la flor"; (Bechtold & Pelletier 1998; Trieu et al. 2000)), fusión de protoplastos, o electroporación, entre otros. Más preferiblemente dicha planta es producida por transformación mediada por *Agrobacterium*.

Con "vector de transformación binaria" se entiende un vector de transformación de T-ADN que comprende: una región de T-ADN que contiene al menos un gen de interés y/o al menos un marcador seleccionable activo en la célula eucariota que va a ser transformada; y una región de la columna vertebral del vector que contiene al menos orígenes de replicación activos en *E. coli* y *Agrobacterium* y marcadores para selección en *E. coli* y *Agrobacterium*. Alternativamente, la replicación del vector de transformación binario en *Agrobacterium* depende de la presencia de un plásmido auxiliar separado. El vector binario pGreen y el plásmido auxiliar pSoup forman un ejemplo de tal sistema como se describe por ejemplo en (Hellens et al. 2000) o como se encuentra disponible en el sitio de Internet <http://www.pgreen.ac.uk>.

Los límites del T-ADN de un vector de transformación binario se pueden derivar de plásmidos Ti del tipo octopina o del tipo nopalina o de ambos. El T-ADN de un vector binario es transfectado únicamente a una célula eucariota junto con un plásmido auxiliar. También se conocen en el arte múltiples cepas del vector binario de *Agrobacterium* para una co-transformación eficiente de plantas (Bidney y Scelonge 2000 - WO018939).

"Huésped" o "célula huésped" u "organismo huésped" es aquí cualquier célula procariota o eucariota u organismo que pueda ser receptor de las secuencias de la presente invención. Un "huésped", como se utiliza aquí el término, incluye organismos procariotas o eucariotas que pueden ser modificados por medio de ingeniería genética. Para los ejemplos de tales huéspedes, ver Maniatis et al., Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982). "Célula vegetal" comprende cualquier célula derivada de cualquier planta y que existe en cultivo como una célula individual, un grupo de células o un callo. Una célula vegetal puede ser también cualquier célula en una planta madura o en desarrollo en cultivo o creciendo en la naturaleza.

"Planta" o "Plantas" incluyen todas las especies de plantas que pertenecen a la súper familia Viridiplantae. La presente invención es aplicable a cualquier planta, en particular a plantas monocotiledóneas y plantas dicotiledóneas incluida una forrajera o legumbre forrajera, planta ornamental, planta alimenticia, árbol, o arbusto seleccionado de la lista que comprende *Acacia spp.*, *Acer spp.*, *Actinidia spp.*, *Aesculus spp.*, *Agathis australis*, *Albizia amara*, *Alsophila tricolor*, *Andropogon spp.*, *Arachis spp.*, *Areca catechu*, *Astelia fragrans*, *Astragalus cicer*, *Baikiaea plurijuga*, *Betula spp.*, *Brassica spp.*, *Bruguiera gymnorrhiza*, *Burkea africana*, *Butea frondosa*, *Cadaba farinosa*, *Calliandra spp.*, *Camellia sinensis*, *Canna indica*, *Capsicum spp.*, *Cassia spp.*, *Centroema pubescens*, *Chaenomeles spp.*, *Cinnamomum cassia*, *Coffea arabica*, *Colophospermum mopane*, *Coronillia varia*, *Cotoneaster serotina*,

- Crataegus* spp., *Cucumis* spp., *Cupressus* spp., *Cyathea dealbata*, *Cydonia oblonga*, *Cryptomeria japonica*,
Cymbopogon spp., *Cynthea dealbata*, *Cydonia oblonga*, *Dalbergia monetaria*, *Davallia divaricata*, *Desmodium* spp.,
Dicksonia squarosa, *Diheteropogon amplexans*, *Dioclea* spp., *Dolichos* spp., *Dorycnium rectum*, *Echinochloa*
5 *pyramidalis*, *Ehrhartia* spp., *Eleusine coracana*, *Eragrostis* spp., *Erythrina* spp., *Eucalyptus* spp., *Euclea schimperii*,
Eulalia villosa, *Fagopyrum* spp., *Feijoa sellowiana*, *Fragaria* spp., *Flemingia* spp., *Freycinetia banksii*, *Geranium*
thunbergii, *Ginkgo biloba*, *Glycine javanica*, *Gliricidia* spp., *Gossypium hirsutum*, *Grevillea* spp., *Guibourtia*
coleosperma, *Hedysarum* spp., *Hemarthria altissima*, *Heteropogon contortus*, *Hordeum vulgare*, *Hyparrhenia rufa*,
Hypericum erectum, *Hyperthelia dissoluta*, *Indigo incarnata*, *Iris* spp., *Leptarrhena pyrolifolia*, *Lespedeza* spp., *Lettuca*
10 *spp.*, *Leucaena leucocephala*, *Loudetia simplex*, *Lotonus bainesii*, *Lotus* spp., *Macrotyloma axillare*, *Malus* spp.,
Manihot esculenta, *Medicago sativa*, *Metasequoia glyptostroboides*, *Musa sapientum*, *Nicotianum* spp., *Onobrychis*
spp., *Omithopus* spp., *Oryza* spp., *Peltophorum africanum*, *Pennisetum* spp., *Persea gratissima*, *Petunia* spp.,
Phaseolus spp., *Phoenix canariensis*, *Phormium cookianum*, *Photinia* spp., *Picea glauca*, *Pinus* spp., *Pisum sativum*,
Podocarpus totara, *Pogonarthria fleckii*, *Pogonarthria squarrosa*, *Populus* spp., *Prosopis cineraria*, *Pseudotsuga*
15 *menziesii*, *Pterolobium stellatum*, *Pyrus communis*, *Quercus* spp., *Rhaphiolepis umbellata*, *Rhopalostylis sapida*,
Rhus natalensis, *Ribes grossularia*, *Ribes* spp., *Robinia pseudoacacia*, *Rosa* spp., *Rubus* spp., *Salix* spp.,
Schyzachyrium sanguineum, *Sciadopitys verticillata*, *Sequoia sempervirens*, *Sequoiadendron giganteum*, *Sorghum*
bicolor, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus alopecuroides*, *Stylosanthos humilis*, *Tadehagi* spp., *Taxodium*
distichum, *Themeda triandra*, *Trifolium* spp., *Triticum* spp., *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vitis*
20 *vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*, *Zea mays*, amaranto, alcachofa, espárrago, brócoli, col de
Bruselas, col, canola, zanahoria, coliflor, apio, col rizada, lino, kale, lenteja, colza de semilla oleaginosa, oca,
cebolla, patata, arroz, soja, paja, remolacha azucarera, caña de azúcar, girasol, tomate, calabaza, y té, entre otras, o
las semillas de cualquiera de las plantas específicamente mencionadas anteriormente o un cultivo de tejido, célula u
órgano de cualquiera de las especies anteriores. La presente invención es aplicable a cualquier planta, en particular
a plantas monocotiledóneas y a plantas dicotiledóneas incluido un arbusto o leguminosa forrajera, planta
25 ornamental, cultivo alimenticio, árbol, o arbusto seleccionado de la lista que comprende *Acacia* spp., *Acer* spp.,
Actinidia spp., *Aesculus* spp., *Agathis australis*, *Albizia amara*, *Alsophila tricolor*, *Andropogon* spp., *Arachis* spp.,
Areca catechu, *Astelia fragrans*, *Astragalus cicer*, *Baikiaea plurijuga*, *Betula* spp., *Brassica* spp., *Bruguiera*
gymnorhiza, *Burkea africana*, *Butea frondosa*, *Cadaba farinosa*, *Calliandra* spp., *Camellia sinensis*, *Canna indica*,
Capsicum spp., *Cassia* spp., *Centroema pubescens*, *Chaenomeles* spp., *Cinnamomum cassia*, *Coffea arabica*,
30 *Colophospermum mopane*, *Coronilla varia*, *Cotoneaster serotina*, *Crataegus* spp., *Cucumis* spp., *Cupressus* spp.,
Cyathea dealbata, *Cydonia oblonga*, *Cryptomeria japonica*, *Cymbopogon* spp., *Cynthea dealbata*, *Cydonia oblonga*,
Dalbergia monetaria, *Davallia divaricata*, *Desmodium* spp., *Dicksonia squarosa*, *Diheteropogon amplexans*, *Dioclea*
spp., *Dolichos* spp., *Dorycnium rectum*, *Echinochloa pyramidalis*, *Ehrhartia* spp., *Eleusine coracana*, *Eragrostis* spp.,
Erythrina spp., *Eucalyptus* spp., *Euclea schimperii*, *Eulalia villosa*, *Fagopyrum* spp., *Feijoa sellowiana*, *Fragaria* spp.,
35 *Flemingia* spp., *Freycinetia banksii*, *Geranium thunbergii*, *Ginkgo biloba*, *Glycine javanica*, *Gliricidia* spp., *Gossypium*
hirsutum, *Grevillea* spp., *Guibourtia coleosperma*, *Hedysarum* spp., *Hemarthria altissima*, *Heteropogon contortus*,
Hordeum vulgare, *Hyparrhenia rufa*, *Hypericum erectum*, *Hyperthelia dissoluta*, *Indigo incarnata*, *Iris* spp.,
Leptarrhena pyrolifolia, *Lespedeza* spp., *Lettuca* spp., *Leucaena leucocephala*, *Loudetia simplex*, *Lotonus bainesii*,
Lotus spp., *Macrotyloma axillare*, *Malus* spp., *Manihot esculenta*, *Medicago sativa*, *Metasequoia glyptostroboides*,
40 *Musa sapientum*, *Nicotianum* spp., *Onobrychis* spp., *Omithopus* spp., *Oryza* spp., *eltophorum africanum*, *Pennisetum*
spp., *Persea gratissima*, *Petunia* spp., *Phaseolus* spp., *Phoenix canariensis*, *Phormium cookianum*, *Photinia* spp.,
Picea glauca, *Pinus* spp., *Pisum sativum*, *Podocarpus totara*, *Pogonarthria fleckii*, *Pogonarthria squarrosa*, *Populus*
spp., *Prosopis cineraria*, *Pseudotsuga menziesii*, *Pterolobium stellatum*, *Pyrus communis*, *Quercus* spp.,
Rhaphiolepis umbellata, *Rhopalostylis sapida*, *Rhus natalensis*, *Ribes grossularia*, *Ribes* spp., *Robinia*
45 *pseudoacacia*, *Rosasp.*, *Rubusspp.*, *Salix* spp., *Schyzachyrium sanguineum*, *Sciadopitys varticillata*, *Sequoia*
sempervirens, *Sequoiadendron giganteum*, *Sorghum bicolor*, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus*
alopecuroides, *Stylosanthos humilis*, *Tadehagi* spp., *Taxodium distichum*, *Themeda triandra*, *Trifolium* spp., *Triticum*
spp., *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vitis vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*,
50 *Zea mays*, amaranto, alcachofa, espárrago, brócoli, col de Bruselas, col, canola, zanahoria, coliflor, apio, col rizada,
lino, kale, lenteja, colza de semilla oleaginosa, oca, cebolla, patata, arroz, soja, paja, remolacha azucarera, caña de
azúcar, girasol, tomate, calabaza, y té, entre otras, o las semillas de cualquiera de las plantas específicamente
mencionadas anteriormente o un cultivo de tejido, célula u órgano de cualquiera de las especies anteriores.

“Cereal” comprende plantas de cultivo con grano comestible por ejemplo plantas que pertenecen a la familia de las gramíneas que se cultiva por sus granos nutritivos tales como avena, cebada, centeno, trigo, arroz y maíz, etc.

Dentro del alcance de la presente invención está también la aplicación del sistema Doble híbrido, en donde cualquiera de las secuencias de la presente invención es utilizada para estudiar interacciones con otros factores tales como proteínas. Con “ensayo doble híbrido de levadura” se entiende un ensayo que está basado en la observación de que muchos factores de transcripción eucariotas contienen dos dominios, un dominio de enlazamiento de ADN (DB) y un dominio de activación (AD) que, cuando se separan físicamente (es decir ruptura del enlace covalente) no efectúa la expresión génica objetivo. Dos proteínas capaces de interactuar físicamente con una de dichas proteínas fusionadas a DB y la otra de dichas proteínas fusionadas a AD reunirán nuevamente a los dominios DB y AD del factor de transcripción resultando en expresión génica objetivo. El gen objetivo en el ensayo doble híbrido de levadura es usualmente un gen reportero tal como el gen de la β -galactosidasa. La interacción entre compañeros de la proteína en el ensayo doble híbrido de la levadura puede por lo tanto ser cuantificada por medio de la medición de la actividad del producto del gen reportero (Bartel & Fields 1997). Alternativamente, se puede utilizar un sistema doble híbrido de mamífero que incluye por ejemplo una proteína fluorescente verde quimérica que codifica un gen reportero (Shioda *et al.* 2000). Incluso otra alternativa consiste de un sistema doble híbrido bacteriano utilizando por ejemplo *HIS* como gen reportero (Joung *et al.* 2000). Una persona capacitada en el arte reconocerá también que en versiones adaptadas del sistema doble híbrido y también otras técnicas (por ejemplo, retardo de gel, inmunoprecipitación, inhibición competitiva), es posible para estudiar interacciones proteína - oligonucleótido, y por lo tanto estas técnicas que utilizan cualquiera de la secuencias de la presente invención están también dentro del alcance de la presente invención.

El término “fragmento de una secuencia” o “parte de una secuencia” significa una secuencia truncada de la secuencia original a la cual se hace referencia. La secuencia truncada (secuencia de ácido nucleico o de proteína) pueden variar ampliamente en longitud; el tamaño mínimo que es una secuencia de tamaño suficiente para proporcionar una secuencia con al menos una función comparable y/o actividad o la secuencia original a la cual se hace referencia, mientras que el tamaño máximo no es crítico. En algunas aplicaciones, el tamaño máximo usualmente no es sustancialmente mayor que aquel requerido para proporcionar la actividad deseada y/o la(s) función(es) de la secuencia original. Típicamente, la secuencia truncada de aminoácidos o de nucleótidos estará en el rango aproximadamente desde 5 hasta aproximadamente 60 aminoácidos de longitud. Más típicamente, sin embargo, la secuencia tendrá un máximo de aproximadamente 50 aminoácidos de longitud, preferiblemente un máximo de aproximadamente 60 aminoácidos. Usualmente es deseable seleccionar secuencias de al menos aproximadamente 10, 12 ó 15 aminoácidos o nucleótidos hasta un máximo de aproximadamente 20 ó 25 aminoácidos o nucleótidos.

La invención también incluye métodos para selección de compuestos de alto rendimiento. La persona capacitada en el arte puede diseñar fácilmente tales métodos utilizando uno o varios elementos de la invención.

Los compuestos que aún no se han obtenido o identificado pueden ser compuestos que son capaces de enlazarse con cualquiera de los ácido nucleicos, péptidos o proteínas involucrados en el proceso del procesamiento del ARN mensajero precursor y son por lo tanto útiles en el método de la presente invención. Dicho compuesto o pluralidad de compuestos pueden estar incluidos por ejemplo, en muestras, por ejemplo, extractos celulares por ejemplo de plantas, animales o microorganismos. Además, dicho(s) compuesto(s) puede(n) ser conocido(s) en el arte pero hasta ahora no se sabe que sean capaces de suprimir o activar proteínas involucradas en el procesamiento del ARN mensajero precursor. En el alcance de la presente invención está también incluida la introducción en una célula vegetal de una o más moléculas recombinantes de ácido nucleico, tales como una molécula de ADN que codifica una proteína que cuando se expresa en dicha célula vegetal en una cantidad efectiva incrementa o induce la expresión de un ácido polinucleótido endógeno de acuerdo con la presente invención o como se define en las reivindicaciones o incrementa o induce la actividad de un polipéptido de la reivindicación.

La presente invención está descrita adicionalmente haciendo referencia a las siguientes figuras y ejemplos no limitantes.

Breve descripción de las figuras

Figura 1

Fenotipos de tolerancia a la sal de cepas de levadura que expresan los clones aislados de ADNc de *Arabidopsis*, U1A, Ct-SRL1 o RCY1, como se determinó por medio de los "ensayos de caída"; en los cuales se analizan sobre placas diluciones seriales de cultivos saturados de las diferentes cepas, y de la cepa de control que contiene al vector de expresión vacío, sin sal o conteniendo las concentraciones indicadas de LiCl o NaCl.

Figura 2

Secuencias de aminoácidos derivadas de las secuencias de nucleótidos de los clones RCY1 (A: SEQ ID NO 4) y Ct-SRL1 (B: SEQ ID NO 3). Los dipéptidos Arg-Ser, Arg-Glu y Arg-Asp, que definen al dominio RS, están escritos en negrilla. El M subrayado en (A) marca al residuo de metionina utilizado como iniciador durante la expresión en levadura del dominio RS de RCY1 (SEQ ID NO 21).

Figura 3

Fenotipo de tolerancia a la sal de cepas de levadura que expresan al dominio del terminal amino (ciclina) de RCY1, su dominio RS del terminal carboxilo, o la proteína de longitud completa. Se llevaron a cabo los ensayos de caída como en el experimento de la Fig. 1.

Figura 4

Fenotipo de tolerancia a la sal de plantas transgénicas de *Arabidopsis*. Las semillas T2 de tres líneas transgénicas independientes (L1, L3 y L5, como ejemplo de las 12 líneas obtenidas), transformadas con el ADNc de Ct-SRL1 bajo el control del promotor 35S del CaMV, fueron germinadas sobre placas de agar sin (A) o con (B) LiCl 20 mM. Se utilizaron semillas de plantas silvestres (Wt) como control. Todas las líneas transgénicas mostraron fenotipos similares.

Figura 5

Secuencias de los genes y proteínas de la presente invención.

EJEMPLOS

Ejemplo 1: Material de la planta

Se cultivaron semillas de remolacha roja (variedad DITA de *Beta vulgaris*, también denominada aquí como "remolacha azucarera"), sobre macetas que contienen una mezcla de arena y vermiculita (1:1 p/p). Se cultivaron las plantas bajo condiciones de invernadero (8 horas a 20°C, 16 horas a 25°C con iluminación adicional para estimular un fotoperíodo mínimo de 12 horas). Fueron periódicamente irrigadas con una solución nutritiva que contenía 2,4 g/l de $\text{Ca}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$, 1 g/l de KNO_3 , 1 g/l de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,3 g/l de KH_2PO_4 , 5,6 mg/l de quelato de Fe (Kelantren, Bayer), 1,1 mg/l de $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 3,3 mg/l de $\text{MnO}_4 \cdot \text{H}_2\text{O}$, 0,3 mg/l de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 3,8 mg/l de H_3BO_3 , 0,18 mg/l de $(\text{NH}_4)_6\text{MO}_7 \cdot 4\text{H}_2\text{O}$. Para la construcción de la biblioteca de ADNc, se irrigaron plantas de tres semanas de edad con NaCl 200 mM durante 24 horas antes de la cosecha.

Ejemplo 2: Cepas de levadura y condiciones de cultivo

Se transformaron las células competentes de *Saccharomyces cerevisiae* W303-1 A (*MATa ura3, leu2, his3, trp1, ade2, ena 1-4::HIS3*) con la biblioteca de ADNc de *Arabidopsis*.

Se utilizó la cepa JM26 de *Saccharomyces cerevisiae* (*MATa leu 2-3,112 ura 3-1 trp1-1, ade 2-1 his3-11,15 can 1-100, ena 1-4::HIS3, nha1::TRP1*) suministrada por J. M. Mulet (Universidad Politécnica de Valencia, Instituto de Biología Molecular y Celular de Plantas) para la selección de la biblioteca de ADNc de remolacha roja. La cepa JM26 es derivada de W303.1 A (Wallis et al., 1989, Cell 58: 409 - 419) con mutaciones nulas de los genes *ENA1-4* y

NHA1, que codifican una ATPasa que bombea Na^+ y un antiportador de Na^+/H^+ , respectivamente, responsable por la mayor parte de la extrusión del sodio de la levadura (Garcia-deblas et al. 1993, Mol. Gen. Genet. 236, 363 - 368), (Bañuelos et al. 1998, Microbiology 144: 2749 - 2758).

- 5 Se cultivaron las células de levadura ya sea en medio de glucosa sintética mínima (SD) o en medio rico (YPD). El medio SD contenía glucosa al 2%, base nitrogenada de levadura al 0,7% sin aminoácidos y ácido succínico 50 mM, ajustado a pH 5 con Tris, más los aminoácidos requeridos [100 mg/ml de leucina, 30 mg/ml de adenina, 100 mg/ml de metionina] como se indicó. El medio YPD contenía extracto de levadura al 1%, peptona Bacto al 2% y glucosa al 2%. Se suplementó el medio con NaCl y LiCl como se indica en las figures y en los ejemplos. El medio sólido
10 contenía agar grado bacteriológico al 2%.

Ejemplo 3: Construcción de las bibliotecas de ADNc

- 15 La construcción de la biblioteca de ADNc de *Arabidopsis thaliana* en el vector pFL61 está descrita en Minet et al. (1992) Plant J. 2(3): 417 - 422.

Para la construcción de una biblioteca de ADNc de remolacha roja inducida por estrés por salinidad, se utilizó el material de la planta como se describe en el ejemplo 1. Se sintetizaron los ADNc direccionales (kit de síntesis de ADNc, Stratagene) utilizando ARN poli(A)⁺ preparado a partir de las hojas de plantas de remolacha roja tratadas con
20 sal. Se ligaron los ADNc dentro del vector fago λ PG15 y se los empacó utilizando extracto de empaquetamiento Gigapack III dorado (Stratagene). Este fago tiene insertado al plásmido de expresión pYPGE15 que puede ser extirpado (*URA3* como marcador de selección) que puede ser utilizado directamente tanto para complementación de *E. coli* como de levadura (Brunelli y Pall, 1993, Yeast 9: 1309 - 1318). Se recuperó una biblioteca de ADNc del plásmido λ PG15 por medio del sistema cre-lox recombinasa (Brunelli y Pall, 1993, Yeast 9: 1309 - 1318).

25

Ejemplo 4: Selección y asilamiento de clones de ADNc que confieren tolerancia a la sal a la levadura

Para seleccionar los ADNc de *Arabidopsis thaliana* que incrementan la tolerancia a la sal en levadura, se utilizó la biblioteca de ADNc construida en pFL61 para transformar la cepa de levadura W303-1 A por medio del método del
30 LiCl (Gietz et al. 1992, Nucleic Acids Res. 20: 1425). Se seleccionaron los transformantes para tolerancia a los alógenos en placas con medio mínimo más LiCl 25 mM y 50 mM, o como se indica en las figuras, y que contenían metionina 400 mM. Se sometieron clones resistentes a la pérdida de plásmido inducida por ácido fluoroorótico (Boeke et al (1984) Mol. Gen. Genet. 197: 354 - 346) para seleccionar aquellos clones que demuestran tolerancia al LiCl que depende del plásmido. Se confirmaron los resultados en una cepa de tipo silvestre y en el mutante doble
35 *ena1-4::HIS3 ttp1::LEU2*, deficiente en el transporte vacuolar.

Para seleccionar los ADNc de remolacha roja que incrementan la tolerancia a la sal en levadura, se utilizó la biblioteca de ADNc construida en pYPGE15 para transformar la cepa mutante de levadura JM26. Se reunieron los transformantes seleccionados sobre placas de SD con leucina y adenina por medio de prototrofia de uracilo y se los sembró en placa nuevamente sobre medio de selección (SD con leucina, adenina y metionina suplementada con
40 NaCl 0,15 M) una densidad de 2×10^5 células por placa (12 x 12 cm). Se añadió metionina al medio selectivo para evitar la selección de los homólogos tipo HAL2 ya encontrados en *Arabidopsis* (Quintero et al. 1996, Plant Cell 8: 529 - 537) (Gil-Mascarell et al. 1999, Plant J. 17(4): 373 - 383). Alternativamente, para la selección de células de levadura resistentes a Li^+ , se sembraron nuevamente en placa los transformantes sobre medio de selección (SD con
45 leucina y adenina suplementada con LiCl 20 mM). Se seleccionaron nuevamente los clones putativos positivos sobre el mismo medio de NaCl o sobre medio de LiCl.

Ejemplo 5: Determinación del contenido intracelular de litio

- 50 Se cultivaron células de levadura que expresan los clones de ADNc de la presente invención o la cepa de control transformada con el vector vacío, hasta fase exponencial y se suplementó luego el medio con LiCl hasta una concentración final 30 mM. Se tomaron muestras de 10 ml en diferentes momentos y se determinó el contenido intracelular de litio como se describió previamente (Murguía et al. 1996, J. Biol. Chem. 271: 29029 - 29033).

Ejemplo 6: Ensayo de beta galactosidasa

Se transformaron células de levadura que contenían al gen *LacZ* de *E. Coli*, interrumpido o no con un intrón (Legrain y Rosbash (1989), cell, 75: 573 - 583), bajo el control de un promotor inducible por galactosa, con el ADNc de Ct-SRL1 en un vector de expresión de levadura o, como control, con el plásmido vacío. Se desarrollaron los cultivo hasta fase exponencial en medio mínimo de glucosa, con o sin LiCl 35 mM, y luego se cambió a medio de galactosa, manteniendo las mismas condiciones de salinidad, para inducir la expresión de beta-galactosidasa. Se recolectaron las muestras 0 (valor de fondo) y 4 horas después de la inducción, y se midió la actividad de la beta-galactosidasa en células permeabilizadas como lo describen (Serrano et al, 1973: Eur. J. Biochem. 34: 479 - 482).

Ejemplo 7: Ensayo RT-PCR

Se cultivaron células de levadura que sobreexpresan al ADNc de Ct-SRL o transformadas con el vector vacío hasta fase exponencial; se suplementó luego el medio con LiCl hasta una concentración final de 150 mM y se purificó el ARN total en diferentes momentos. Se digirieron cantidades igual del ARN total con ADNasa1 libre de ARNasa, y se transcribió en forma inversa con RTasa de M-MuLV (Roche Molecular Biochemicals) utilizando al iniciador *preSAR1rp* (5' CATCAAATCTTTTCAGGG-3': SEQ ID NO. 23), que hibrida a los segundos 277 nucleótidos del exón SAR1 secuencia abajo del sitio de empalme de 3'. Esto fue seguido por 15 ciclos de amplificación por PCR con la ADN polimerasa Netzyme (Molecular Netline Bioproducts), y el iniciador *preSAR1fp*, (5'-CTTTATTTTACTGTACAG-3': SEQ ID NO. 24), que hibrida al extremo 3' del intrón SAR1. Se añadió [α -³²P]dCTP (740kBq, 110 TBq/mmol, Amersham) después del 5° ciclo y se resolvieron los productos en geles de urea de PA al 6% y visualizados por medio de autorradiografía. Se determinó la intensidad relativa de las bandas amplificadas utilizando un dispositivo formador de imágenes fosforescentes PhosphorImager de FUJI (Fujifilm BAS-1500). Se utilizaron las muestras que carecían de transcriptasa inversa como controles.

Ejemplo 8: Transformación de plantas

Los clones de la presente invención, por ejemplo el ADNc de Ct-SRL1 de *Arabidopsis*, se subclonan en pBI121 (Clontech), en lugar del gen GUS, y se introduce el vector binario resultante en la cepa C58C1 de *Agrobacterium tumefaciens* por medio de electroporación. Se obtuvieron plantas transgénicas del tipo *Arabidopsis* por medio de infiltración *in vivo* como se describe en http://www.arabidopsis.org/protocols_Mundy2.html#trans.inf.

Alternativamente, se pueden transformar los genes de la presente invención con otras plantas monocotiledóneas o dicotiledóneas. Por lo tanto, se clonan en un vector adecuado de transformación de la planta, tal como un vector binario, bajo el control de secuencias reguladoras operables de la planta, tal como un promotor operable de la planta y un terminador operable de la planta. Estos vectores que contienen un gen de la presente invención pueden ser transformados en plantas (tal como una planta de cultivo) utilizando técnicas estándar conocidas por las personas capacitadas en el arte, tales como transferencia de genes mediada por *Agrobacterium*. La planta transgénica que expresa al gen de la presente invención se espera que muestre una mayor tolerancia al estrés ambiental, tal como por ejemplo una mayor tolerancia a la toxicidad por sales minerales. Esta mayor tolerancia puede ser observada por ejemplo por la habilidad de las plantas transgénicas para germinar en un medio que contiene sal.

Ejemplo 9: Transformación de arroz con los genes de la presente invención

La expresión de genes de remolacha roja o de *Arabidopsis thaliana* que están involucrados en la tolerancia a la sal en levadura en monocotiledóneas, tal como por ejemplo Arroz, pueden conferir tolerancia al estrés a esa planta monocotiledónea.

Para transferir la actividad de tolerancia al estrés de los genes de la presente invención a monocotiledóneas, los genes anteriormente mencionados (SEQ ID NO. 1, 2, 5, 7, 9, 11, 13, 15, 17 y 19), operativamente enlazados a un promotor, son transformados cada uno en arroz utilizando los procedimientos estándar de transformación conocidos por las personas capacitadas en el arte y esquematizados en el siguiente párrafo. Después de diferentes períodos de tiempo en el rango de 1 día hasta 1 o más semanas, se revisan las plantas de semillero por la expresión del gen transformado. Esto se hace cultivando las paltas de semillero en medio para organogénesis, y revisando la presencia del ADN o de ARNm por medio de PCR o PCR inversa. Después de la confirmación de la expresión del

gen se revisan las plantas de arroz transformadas con relaciona a la tolerancia mejorada a situaciones de estrés incluida salinidad, sequía y frio (ver WO97/13843). Esto se hace cultivando las plantas de arroz transformadas en medio que contiene mayores cantidades de NaCl o LiCl. También se analizó la mayor resistencia al frio o a la sequía cultivando las plantas transformadas en temperaturas de cultivo por debajo del óptimo y niveles de humedad por debajo del óptimo, respectivamente (WO97/13843).

Transformación de arroz mediada por *Agrobacterium*

Los genes de la presente invención están operativamente enlazados a un promotor y clonados en un vector. Estos vectores son transformados con la cepa LBA4404 o C58 de *Agrobacterium tumefaciens* por medio de electroporación y posteriormente se seleccionaron células bacterianas transformadas sobre un medio de agar sólido que contiene los antibióticos apropiados.

Para demostración de la expresión de los genes de la presente invención en arroz, se les quitó la cascarilla a 309 semillas secas maduras de las variedades cultivadas japonesas de arroz Nipponbare o Taipei, se esterilizaron y germinaron sobre un medio que contenía ácido 2,4-diclorofenoxiacético (2,4-D). Después de la incubación en la oscuridad durante cuatro semanas, se cortan los callos embrionarios derivados de escutelo y se los propaga sobre el mismo medio. El callo embrionario seleccionado es co-cultivado luego con *Agrobacterium*. El callo co-cultivado se cultiva sobre medio que contiene 2,4-D durante 4 a 5 semanas en la oscuridad en la presencia de una concentración adecuada del agente selectivo apropiado. Durante éste período, se desarrollan islas de callo resistente de crecimiento rápido. Después de la transferencia de este material a un medio con una concentración reducida de 2,4-D e incubación a la luz, se libera el potencial embriogénico y se desarrollan brotes en las siguientes a cuatro a cinco semanas. Se extirpan los brotes del callo y se incuban durante una semana sobre un medio que contiene auxina a partir del cual pueden ser transferidas al suelo. Se cultivan los brotes endurecidos bajo gran humedad y días cortos en un fitotrón. Las semillas pueden ser cosechadas de tres a cinco meses después del trasplante. El método produce transformantes de un solo locus a una tasa superior al 50 % (Chan et al. 1993, Plant Mol. Biol. 22: 491 - 506) (Hiei et al. 1994, Plant J. 6: 271 - 282).

REFERENCIAS

- An G., Watson B. D., Stachel S., Gordon M.P., & Nester E. W. (1985) New cloning vehicles for transformation of higher plants. EMBO J. 4, 277 - 284.
- Armstrong C. L., Petersen W. P., Buchholz W. G., Bowen B. A., & Sulc S. L. (1990) Factors affecting PEG-mediated stable transformation of maize protoplasts. Plant Cell Reports 9, 335 - 339.
- Baron M. H. & Baltimore D. (1982) Antibodies against the chemically synthesized genome-linked protein of poliovirus react with native virus-specific proteins. Cell 28, 395 - 404. Bartel P. L. & Fields S. (1997) The Yeast Two-Hybrid System. Oxford University Press. Bechtold N. & Pelletier G. (1998) In planta *Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. Methods Mol.Biol. 82, 259 - 266.
- Cho R. J., Mindrinos M., Richards D. R., Sapolsky R. J., Anderson M., Drenkard E., Dewdney J., Reuber T. L., Stammers M., Federspiel N., Theologis A., Yang W. H., Hubbell E., Au M., Chung E. Y., Lashkari D., Lemieux B., Dean C., Lipshutz R. J., Ausubel F. M., Davis R. W., & Oefner P. J. (1999) Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. Nat.Genet. 23, 203 - 207.
- Christou P., McCabe D. E., & Swain W. F. (1988) Stable transformation of soybean callus by DNA-coated gold particles. Plant Physiol. 87, 671 - 674.
- Crossway A., Oakes J. V., Irvine J. M., Ward B., Knauf V. C., & Shewmaker C. K. (1986) Integration of foreign DNA following microinjection of tobacco mesophyll protoplasts. Mol. Gen. Genet. 202, 179 - 185.
- Dodds J. H. (1985) Plant genetic engineering. Cambridge University Press.
- Fromm M., Taylor L.P., & Walbot V. (1985) Expression of genes transferred into monocot and dicot plant cells by electroporation. Proc.Natl.Acad.Sci.U.S.A 82, 5824 - 5828.
- Hanahan D. (1983) Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol 166,557 - 580.
- Hellens R. P., Edwards E. A., Leyland N. R., Bean S., & Mullineaux P.M. (2000) pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*- mediated plant transformation. Plant Mol. Biol. 42, 819 - 832.
- Herrera-Estrella L., De Block M., Messens E.H.J.P., Van Montagu M., & Schell J. (1983a) Chimeric genes as dominant selectable markers in plant cells. EMBO J. 2, 987 - 995.

- Herrera-Estrella L., Depicker A., Van Montagu M., & Schell J. (1983b) Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature* 303, 209 - 213.
- Joung J. K., Ramm E. I., & Pabo C.O. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A* 97, 7382 - 7387.
- 5 Krens F. A., Molendijk L., Wullems G. J., & Schilperoort R. A. (1982) In vitro transformation of plant protoplasts with Ti-plasmid DNA. *Nature* 296, 72 - 74.
- Langemeier J. L., Cook R. F., Issel C. J., & Montelaro R. C. (1994) Application of cycle dideoxy fingerprinting to screening heterogeneous populations of the equine infectious anemia virus. *Biotechniques* 17, 484 - 6, 488, 490.
- Lerner R. A. (1982) Tapping the immunological repertoire to produce antibodies of predetermined specificity. *Nature* 10 299, 593 - 596.
- Lerner R. A., Green N., Alexander H., Liu F. T., Sutcliffe J. G., & Shinnick T. M. (1981) Chemically synthesized peptides predicted from the nucleotide sequence of the hepatitis B virus genome elicit antibodies reactive with the native envelope protein of Dane particles. *Proc. Natl. Acad. Sci. U.S.A* 78, 3403 - 3407.
- Li-Sucholeiki X. C., Khrapko K., Andre P. C., Marcelino L. A., Karger B. L., & Thilly W. G. (1999) Applications of constant denaturant capillary electrophoresis/high-fidelity polymerase chain reaction to human genetic analysis. *Electrophoresis* 20, 1224 - 1232.
- 15 Liddle J. E. & Cryer A. (1991) *A Practical Guide to Monoclonal Antibodies*. Wiley New York.
- Loffler J., Langui D., Probst A., & Huber G. (1994) Accumulation of a 50 kDa N-terminal fragment of beta-APP695 in Alzheimer's disease hippocampus and neocortex. *Neurochem. Int* 24, 281 - 288.
- 20 Magyar Z., Meszaros T., Miskolczi P., Deak M., Feher A., Brown S., Kondorosi E., Athanasiadis A., Pongor S., Bilgin M., Bako L., Koncz C., & Dudits D. (1997) Cell cycle phase specificity of putative cyclin-dependent kinase variants in synchronized alfalfa cells. *Plant Cell* 9, 223 - 235.
- McCallum C. M., Comai L., Greene E. A., & Henikoff S. (2000b) Targeted screening for induced mutations. *Nat. Biotechnol.* 18, 455 - 457.
- 25 McCallum C. M., Comai L., Greene E. A., & Henikoff S. (2000a) Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol* 123, 439 - 442.
- Merrifield R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Amer. Chem. Soc.* 85, 2149 - 2154.
- Murakami T., Simonds W. F., & Spiegel A. M. (1992) Site-specific antibodies directed against G protein beta and gamma subunits: effects on alpha and beta gamma subunit interaction. *Biochemistry* 31, 2905 - 2911.
- 30 Palmgren G. (1997) Transgenic plants: environmentally safe factories of the future. *Trends Genet.* 13, 348.
- Paszkowski J., Shillito R. D., Saul M., Mandak V., & Hohn T.H.B.P.I. (1984) Direct gene transfer to plants. *EMBO J.* 3, 2717 - 2722.
- Ross P., Hall L., & Haff L. A. (2000) Quantitative approach to single-nucleotide polymorphism analysis using MALDITOF mass spectrometry [In Process Citation]. *Biotechniques* 29, 620 - 629.
- 35 Sambrook J., Fritsch E. F., & Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Semler B. L., Anderson C. W., Hanecak R., Dorner L. F., & Wimmer E. (1982) A membrane-associated precursor to poliovirus VPg identified by immunoprecipitation with antibodies directed against a synthetic heptapeptide. *Cell* 28, 405 - 412.
- 40 Shioda T., Andriole S., Yahata T., & Isselbacher K. J. (2000) A green fluorescent protein-reporter mammalian two hybrid system with extrachromosomal maintenance of a prey expression plasmid: Application to interaction screening. *Proc. Natl. Acad. Sci. U.S.A* 97, 5220 - 5224.
- Tamura R. N., Cooper H. M., Collo G., & Quaranta V. (1991) Cell type-specific integrin variants with alternative alpha chain cytoplasmic domains. *Proc. Natl. Acad. Sci. U.S.A* 88, 10183 - 10187.
- 45 Trieu A. T., Burleigh S. H., Kardailsky I. V., Maldonado-Mendoza I. E., Versaw W. K., Blaylock L. A., Shin H., Chiou T. J., Katagi H., Dewbre G. R., Weigel D., & Harrison M. J. (2000) Technical Advance: Transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium*. *Plant J.* 22, 531 - 541.
- Vidal-Puig A. & Moller D. E. (1994) Comparative sensitivity of alternative single-strand conformation polymorphism (SSCP) methods. *Biotechniques* 17, 490 - 2, 494, 496.
- 50 Woulfe J., Lafortune L., de Nadai F., Kitabgi P., & Beaudet A. (1994) Post-translational processing of the neurotensin/neuromedin N precursor in the central nervous system of the rat-II. Immunohistochemical localization of maturation products. *Neuroscience* 60, 167 - 181.

Yoon K., Cole-Strauss A., & Kmiec E. B. (1996) Targeted gene correction of episomal DNA in mammalian cells mediated by a chimeric RNA. DNA oligonucleotide. Proc. Natl. Acad. Sci. U.S.A 93, 2071 - 2076.

LISTADO DE SECUENCIAS

5

<110> Universidad Politecnica de Valencia

<120> Protección contra la toxicidad ambiental a través de la manipulación del procesamiento de precursores de ARN mensajero

<130> CROP-016-PCT

10

<140>

<141>

<150> P200001102

<151> 2000-04-19

<160> 24

15

<170> PatentIn Ver. 2.1

<210> 1

<211> 840

<212> ADN

<213> Arabidopsis thaliana

20

<400> 1

```

caaaccttga gaagatgaat ttaccaacta aaccttcttg ttcaaccgga gacatgaccc 60
gtggctcaga agacactgcc cgtcgtccac catcagtaaa agcatctctc tctgtttcat 120
ttggtcagcg tgcacctcat cgtgcttcca ccagaggctc ttctcctggt cgccgtcctc 180
caccgactgg ttatgacaga aatggaggcg atgaagtaca acagcgggcc ccacgtagaa 240
gccagagccg agactattat tctgacagag actcagatag acaacgggaa agagagaggg 300
agaaagaccg cgaaagagag agggggaggg atagatacag agaaaggag agggattatg 360
gtaatgatag gagatcaagg cgcgactatg atagtagaag caggcgcaat gattatgagg 420
acgacagaag tagacatgac cggagaagca ggagcagaag cagaagtagg agcaggagtg 480
tgcagattga gcgtaaccg actcctaaaa gagatagtag caacaaagag aaatcggcgg 540
tgacagtga cagcaatctc gcaaagctaa aagatttgta tggagacgca agtagtcaga 600
aaagggatga aggatttgga acaaggaaag attcaagttc agaagaagtg ataaagcttg 660
gtggttcctc ttggagggtga aaaaacaaac aaaacaaaac caaactgttg atttaaaatg 720
cttcttctat ttagcaggat gatgcatggt gtttaactat actgttttga tttccagcaa 780
actatttgtc atatgcttta tattacagtt taagagttga tctttatctt gaaaaaaaaa 840

```

<210> 2

25

<211> 1433

<212> ADN

<213> Arabidopsis thaliana

<400> 2

```

gtttagtgaa tcgatgattt acactgctat cgacaatttt tacctaaccg acgagcagct 60
gaaggcttca ccttcgagga aagatgggat agatgaaaca actgaaatct ctcttagaat 120
ctatggatgt gatctcatcc aagaggggtg aatattgtct aaactaccac aggcagttat 180
ggctactggg caggttctgt ttcagcgatt ctattgcaag aagtcttttg ctaaatttga 240
tgtcaagata gttgctgcca gctgtgtatg gcttgcata aaactggaag aaaaccctaa 300
aaaagctaga caggatcatca tcgtattcca caggatggag tgcgcaggag agaacttgcc 360
attagaacat ctggatatgt atgccaagaa gttctctgag ttgaaagttg aattaagcag 420
aactgagaga catatactga aagagatggg ttttgtttgt catgttgaac atcctcacia 480
gttcatatca aactaccttg ccacattaga aacacctcca gaattgaggc aagaagcttg 540
gaatttggcc aatgatagtc tgcgtacaac cctctgtgta aggttcagaa gtgaggttgt 600
ggcttgtggg gtagtgatg ctgctgccc taggtttcaa gtaccactcc ctgagaatcc 660
gccgtgggtg aaagcatttg atgcagataa atctagtatt gacgaagtgt gtagagttct 720
tgctcattta tacagtcttc caaaggctca gtatatctct gtttgcaagg atgggaagcc 780
atttacattt tctagcagat ccgggaattc tcaaggctca tcagcgacaa aggatctgtt 840
gccgggagca ggcgaggctg ttgatactaa atgtactgca ggatcagcta ataacgactt 900
gaaggatgga atggttacta caccacacga aaaggctaca gattccaaga aaagtggtag 960
cgagtcaaac tctcagccaa ttgtaggaga ctcaagctat gaaagaagta aagtaggaga 1020
tagagaaaga gagagtgata gagagaagga acgaggtaga gagagggaca ggggtaggtc 1080
tcacagaggc agagattctg acagagacag tgataggag agagacaaac tcaagatcg 1140
aagtcattcat cggatcaagag acagattgaa ggattcaggt ggacattcag ataaatcaag 1200
gcatcattct tctcgggacc gtgactaccg cgactcatcg aaagaccgtc gtaggcacca 1260
ttaagccaat cttcttgtca tctacatccc cttgagccta cttgatgtta agacagtata 1320
gtgttgatt gtgttaagag tcaaaaccca tgtgtactta atcacatgct aagatcacgt 1380
tggttcgaca tataaatcga gaaagtctga tatgtttcta aaaaaaaaaa aaa 1433

```

<210> 3

<211> 221

5 <212> PRT

<213> Arabidopsis thaliana

<400> 3

```

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

```

```

Gly Ser Glu Asp Thr Ala Arg Arg Pro Pro Ser Val Lys Ala Ser Leu
          20                      25                      30

```

```

Ser Val Ser Phe Gly Gln Arg Ala Pro His Arg Ala Ser Thr Arg Gly
          35                      40                      45

```

```

Ser Ser Pro Val Arg Arg Pro Pro Pro Thr Gly Tyr Asp Arg Asn Gly
          50                      55                      60

```

```

Gly Asp Glu Val Gln Gln Arg Ser Pro Arg Arg Ser Gln Ser Arg Asp
          65                      70                      75                      80

```

```

Tyr Tyr Ser Asp Arg Asp Ser Asp Arg Gln Arg Glu Arg Glu Arg Glu

```


				85						90					95				
Lys	Asp	Arg	Glu	Arg	Glu	Arg	Gly	Arg	Asp	Arg	Tyr	Arg	Glu	Arg	Glu				
			100					105					110						
Arg	Asp	Tyr	Gly	Asn	Asp	Arg	Arg	Ser	Arg	Arg	Asp	Tyr	Asp	Ser	Arg				
		115					120					125							
Ser	Arg	Arg	Asn	Asp	Tyr	Glu	Asp	Asp	Arg	Ser	Arg	His	Asp	Arg	Arg				
	130					135					140								
Ser	Arg	Ser	Arg	Ser	Arg	Ser	Arg	Ser	Arg	Ser	Val	Gln	Ile	Glu	Arg				
145					150					155					160				
Glu	Pro	Thr	Pro	Lys	Arg	Asp	Ser	Ser	Asn	Lys	Glu	Lys	Ser	Ala	Val				
				165					170					175					
Thr	Val	Asn	Ser	Asn	Leu	Ala	Lys	Leu	Lys	Asp	Leu	Tyr	Gly	Asp	Ala				
		180						185					190						
Ser	Ser	Gln	Lys	Arg	Asp	Glu	Gly	Phe	Gly	Thr	Arg	Lys	Asp	Ser	Ser				
	195						200					205							
Ser	Glu	Glu	Val	Ile	Lys	Leu	Gly	Gly	Ser	Ser	Trp	Arg							
	210					215					220								

<210> 4

<211> 416

5 <212> PRT

<213> Arabidopsis thaliana

<400> 4

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

Ala Ala Ser Cys Val Trp Leu Ala Ser Lys Leu Glu Glu Asn Pro Lys
 85 90 95

Lys Ala Arg Gln Val Ile Ile Val Phe His Arg Met Glu Cys Arg Arg
 100 105 110

Glu Asn Leu Pro Leu Glu His Leu Asp Met Tyr Ala Lys Lys Phe Ser
 115 120 125

Glu Leu Lys Val Glu Leu Ser Arg Thr Glu Arg His Ile Leu Lys Glu
 130 135 140

Met Gly Phe Val Cys His Val Glu His Pro His Lys Phe Ile Ser Asn
 145 150 155 160

Tyr Leu Ala Thr Leu Glu Thr Pro Pro Glu Leu Arg Gln Glu Ala Trp
 165 170 175

Asn Leu Ala Asn Asp Ser Leu Arg Thr Thr Leu Cys Val Arg Phe Arg
 180 185 190

Ser Glu Val Val Ala Cys Gly Val Val Tyr Ala Ala Ala Arg Arg Phe
 195 200 205

Gln Val Pro Leu Pro Glu Asn Pro Pro Trp Trp Lys Ala Phe Asp Ala
 210 215 220

Asp Lys Ser Ser Ile Asp Glu Val Cys Arg Val Leu Ala His Leu Tyr
 225 230 235 240

Ser Leu Pro Lys Ala Gln Tyr Ile Ser Val Cys Lys Asp Gly Lys Pro
 245 250 255

Phe Thr Phe Ser Ser Arg Ser Gly Asn Ser Gln Gly Gln Ser Ala Thr
 260 265 270

Lys Asp Leu Leu Pro Gly Ala Gly Glu Ala Val Asp Thr Lys Cys Thr
 275 280 285

Ala Gly Ser Ala Asn Asn Asp Leu Lys Asp Gly Met Val Thr Thr Pro
 290 295 300

His Glu Lys Ala Thr Asp Ser Lys Lys Ser Gly Thr Glu Ser Asn Ser
 305 310 315 320

Gln Pro Ile Val Gly Asp Ser Ser Tyr Glu Arg Ser Lys Val Gly Asp
 325 330 335

Arg Glu Arg Glu Ser Asp Arg Glu Lys Glu Arg Gly Arg Glu Arg Asp
 340 345 350
 Arg Gly Arg Ser His Arg Gly Arg Asp Ser Asp Arg Asp Ser Asp Arg
 355 360 365
 Glu Arg Asp Lys Leu Lys Asp Arg Ser His His Arg Ser Arg Asp Arg
 370 375 380
 Leu Lys Asp Ser Gly Gly His Ser Asp Lys Ser Arg His His Ser Ser
 385 390 395 400
 Arg Asp Arg Asp Tyr Arg Asp Ser Ser Lys Asp Arg Arg Arg His His
 405 410 415

<210> 5

<211> 920

5 <212> ADN

<213> Arabidopsis thaliana

<400> 5

ccacgcgtcc gagagatttg gtgaagacga tggagatgca agaggctaata caaggaggag 60
 gatcggagggt ttctccgaat cagacgattt acatcaacaa tctcaacgaa aaagtgaagc 120
 ttgatgagct gaagaaatcg ctgaatgcag tgttctctca gttcgggaag atactggaga 180
 tattggcgtt taagaccttt aagcacaaag gacaagcttg ggtagtcttc gacaacaccg 240
 agtctgcttc cactgctatt gctaaaatga ataattttcc tttctacgac aaggagatga 300
 gaatacaata tgccaaaaca aaatcagatg ttgttgccaa ggccgatggg acatttggtc 360
 ctgcgcgagaa gagaaaagaga catgaggaga aaggaggcgg caagaaaaag aaagaccagc 420
 accatgattc tacacagatg ggcatgccca tgaactcagc atatccagggt gtctatggag 480
 ctgcacctcc tctatcgcaa gtaccatacc ctgggtggcat gaaacceaat atgcccgagg 540
 caccagctcc gccaaataat attctctttg tccaaaacct tcctcacgag acaactccaa 600
 tgggtgcttca gatgttggtc tgccagtacc aaggatttaa ggaagttaga atgattgaag 660
 ccaaacccggg aatcgctttt gtggagtttg ctgatgagat gcagtcgacg gtcgcaatgc 720
 agggacttca aggtttcaag attcagcaaa accagatgct catcacgtat gccaaagaaat 780
 agacaatttc gttttatttg tgtttcgatg agatatgttt gtatctgtca atgttacttc 840
 ttgccatggg ggctgtcttc tgggttggtg gatgctagat atccctctct acttacattt 900
 tttcatcaaa aaaaaaaaaa 920

10

<210> 6

<211> 250

<212> PRT

<213> Arabidopsis thaliana

15 <400> 6

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

Asn Gln Met Leu Ile Thr Tyr Ala Lys Lys
245 250

<210> 7
<211> 1492
5 <212> ADN
<213> Beta vulgaris
<400> 7

```

tcctctcttt taacctaat tgagctctca ttatcctgat tttttcaacc atggatgctc 60
agagagctct tctcgatgaa ttaatgggcg cagctcgaaa tctgactgat gaagaaaaga 120
aaggttatag agagataaag tgggatgaca aggaagtttg tgcgccgtat atgattcgat 180
tttgccctca cgatctcttc gtcaatactc gaagtgatct tggaccatgt ccaagagttc 240
atgacacaaa gctgaaagag agctttgaga actctccaag gcatgactca tatgtcccac 300
gttttgaagc agagcttgcc caattttggt agaagctggt ggcagatttg gataggaaag 360
taagacgtgg gagagagcgg ctggaccagg aggttgaacc tccacctccc cctcctattt 420
ctgcagaaaa agctgagcag ctatctgtac ttgaagagaa aataaaaaat ttgcttgaac 480
aagtagagtc actgggagaa gctggcaaaag tcatgaagc agaagcactc atgcgaaagg 540
tggaagtct taatttagag aaagctgcat taactcaaca gcccagaat gcagcaacaa 600
tgcttaccga agagaaaaag atggcactat gtgaaatttg cggttccttc ctggtagcca 660
atgatgctgt ggaaagaact caatctcata taactggcaa gcagcatatt ggctatggca 720
tggtccgtga ttaccttgct gagtataagg aggctaagga gaaggcaaga gaagaggaaa 780
gattagcaag ggagaaagaa gcagaagaac gtcggaagca gagggaaaag gaaaatgaga 840
gtaaaaacag aagaagcatc tccagtgaga gggaccgtca tcgtgatagg gattatggcc 900
gagatcgtga aagatcacga gaatggaaca atagggggaa tcgagacgag ggaagaggaa 960
tggtatcggag aaggcaatat gatcgcaatg gaagggatgg agggaggaat acgtatcatg 1020
gtcgtgaacg tgaaaggagc aggtcacggt cccctgttag gcatggccac cggaggtgat 1080
ctaagagtgc tggttgccga tattagtagg cagtgggttg ttagataaaa cgatgatctt 1140
aaacctactg aggtagatgc tttatatctc aagatgtttt gtgtctgttt tcgaggtggt 1200
gcattgcagt cttattgggg gttaaacttt tctttattgt cccacagtgt tgagactata 1260
ctgtctcttc tcatcaatct tgtagaggt caaagagatt gaggtaggta aaacttcac 1320
gttgtaatct tacctatagt caacttgagt tttgtccaat tatagcacat ggtctttgaa 1380
acatttttta atcatgcggg ggtacgcaag aaaatatgca actatgctgc atggcttgtg 1440
tgcaaaaaaa aaaaaaaaaa aaaaaactcg agggggggcc cggtagcaag at 1492

```

10 <210> 8
<211> 342
<212> PRT
15 <213> Beta vulgaris
<400> 8

Met Asp Ala Gln Arg Ala Leu Leu Asp Glu Leu, Met Gly Ala Ala Arg
1 5 10 15
Asn Leu Thr Asp Glu Glu Lys Lys Gly Tyr Arg Glu Ile Lys Trp Asp

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

275 280 285
 Ser Arg Glu Trp Asn Asn Arg Gly Asn Arg Asp Glu Gly Arg Gly Met
 290 295 300
 Asp Arg Arg Arg Gln Tyr Asp Arg Asn Gly Arg Asp Gly Gly Arg Asn
 305 310 315 320
 Thr Tyr His Gly Arg Glu Arg Glu Arg Ser Arg Ser Arg Ser Pro Val
 325 330 335
 Arg His Gly His Arg Arg
 340

<210> 9
 <211> 650
 5 <212> ADN
 <213> Beta vulgaris
 <400> 9

acgaacaaca aaaatggcgg cagcagatgt tgaagcagta gacttcgaac ctgaagaaga 60
 tgatctcatg gacgaagatg gcggtgcggc tgaagctgac ggctctcctc gagctcctca 120
 ccctaagatt aaatcagcca ttactggcgc cggagctcca tcttctggcg gcttcggagc 180
 taagaaaact aaaggtcgcg gcttcctgta agacgccgat gctgagcgta acagccgtat 240
 gactgctcgt gaatttgatt ctcttgactc cgatgggtgga cctggctctg ctcgatcaat 300
 tgagggctgg attatacttg tcacgggagt gcatgaagag gctcaagaag aggatctcct 360
 taatgtcttt ggagagtttg gccagcttaa gaatttgcatt ttgaatctgg atcgctgtac 420
 tgggtttgtc aagggttatg cattgatcga gtatgagaag tttgaagaag cacaagctgc 480
 aataaaggag atgaatggtg ccaaaatgct tgagcagccg ataaatggtg attgggcatt 540
 ctgcaatggt ccttacagga ggaggggcaa ccgaagaaga tccccacgtg gtcaccgatc 600
 aaggagtctt agaagaagat attaaatctg tttgctgcat gtggaagttg 650

10 <210> 10
 <211> 203
 <212> PRT
 <213> Beta vulgaris
 15 <400> 10

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

35	40	45
Pro Ser Ser Gly Gly Phe Gly Ala Lys Lys Thr Lys Gly Arg Gly Phe		
50	55	60
Arg Glu Asp Ala Asp Ala Glu Arg Asn Ser Arg Met Thr Ala Arg Glu		
65	70	75
Phe Asp Ser Leu Asp Ser Asp Gly Gly Pro Gly Pro Ala Arg Ser Ile		
	85	90
Glu Gly Trp Ile Ile Leu Val Thr Gly Val His Glu Glu Ala Gln Glu		
	100	105
Glu Asp Leu Leu Asn Val Phe Gly Glu Phe Gly Gln Leu Lys Asn Leu		
	115	120
His Leu Asn Leu Asp Arg Arg Thr Gly Phe Val Lys Gly Tyr Ala Leu		
	130	135
Ile Glu Tyr Glu Lys Phe Glu Glu Ala Gln Ala Ala Ile Lys Glu Met		
145	150	155
Asn Gly Ala Lys Met Leu Glu Gln Pro Ile Asn Val Asp Trp Ala Phe		
	165	170
Cys Asn Gly Pro Tyr Arg Arg Arg Gly Asn Arg Arg Arg Ser Pro Arg		
	180	185
Gly His Arg Ser Arg Ser Pro Arg Arg Arg Tyr		
195	200	

<210> 11

<211> 1464

5 <212> ADN

<213> Beta vulgaris

<400> 11

```

ttttgtagca tttgattttt gctgaaaaac ccaattcata ttttgaagaa atgacaacaa 60
tgaacccttt tgatttggtg ggtgacaatg ataacgatga cccatctcag cttttagagt 120
ctgcaactgc tcagttgcag aaaattgctg ttaaaaaaac cccaactcag gttgctcaac 180
aacctcagca acagaaagct gcaaagttac ccaccaaac tcttctcca actcaagctg 240
tccgggaggc aaagaatgat tcccagcgcg gagggggcg tggaggaggt cgcggtagt 300
gccgtgggcg tgggtggatac aatagggact actcaaacaa tgaaaatgct tttaacagca 360
ctggagtaac tggcagtcaa ggggatgatg gggaaagggg aaggcgacct tatgcgggac 420
ctcgtggccc ttatcgtggt ggtcgccgag atgggttcaa caatgaggag ggaagagacg 480

```



```

gggaacgccc gcgtagaacc tatgagcgac gaagtgggac tgggcgtgga agtgagatca 540
aacgtgaggg agcaggacgt ggaaactggg gtgctgaatc agatgaagtt gcaccggtta 600
ctgaggaagc tggagaacaa aatgagaaga agttgaaccc tgagaatctt ccagctgtag 660
aagatgctgc tgatggcatc aaggagggcc agccagatga gactgaagaa aaggaaccag 720
aggaaaagga gatgacactt gaagagtatg agaagttgct ggaagagaag aggaaggctt 780
tatcagcact caaggctgag gaacgcaagg tggagggtga caaagatttc gagtccatgc 840
aacagcttat aaacaaaaaa aaggatgaag actcagtttt catcaaattg ggttctgaca 900
aggataagaa gaaggaagca gctgaaaagg agaaagtga gaagtctgtc agcattaatg 960
aatcttctgaa gcctgctgaa ggggatagat atggtggtcg tggcagggga cgtggctgtg 1020
gcccaagagg tggtggtatg ggtggaggta ataggatgtt tagtacgtct gctccagcaa 1080
tcgaagatcc aggggagttc ccaaccctag gtggcaagtg aggccacatc tttgaacttt 1140
ggtctctatt tggggtttta cttgaccccc tctgatttta agtcatttga gtgacaggaa 1200
tggaactcca gctgtggggt tectgtacca aatccacttt taagaaaatt tttatgcttt 1260
ttaaatattgt atatttatc tgttaaaaaa aaaaaaaaaa actcgtgccg aattcgatat 1320
caagcttatc gataccgtcg acctcgaggg ggggcccggg accaagatgg cctttggtgg 1380
gttgaagaag gaaaaagaca gaaacgactt aattacctac ttgaaaaaag cctgtgagta 1440
aacaggcccc ttttcctttg tcga                                     1464

```

<210> 12

<211> 356

5 <212> PRT

<213> Beta vulgaris

<400> 12

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

Arg	Gly	Gly	Arg	Arg	Asp	Gly	Phe	Asn	Asn	Glu	Glu	Gly	Arg	Asp	Gly	130	135	140	
Glu	Arg	Pro	Arg	Arg	Thr	Tyr	Glu	Arg	Arg	Ser	Gly	Thr	Gly	Arg	Gly	145	150	155	160
Ser	Glu	Ile	Lys	Arg	Glu	Gly	Ala	Gly	Arg	Gly	Asn	Trp	Gly	Ala	Glu	165	170	175	
Ser	Asp	Glu	Val	Ala	Pro	Val	Thr	Glu	Glu	Ala	Gly	Glu	Gln	Asn	Glu	180	185	190	
Lys	Lys	Leu	Asn	Pro	Glu	Asn	Leu	Pro	Ala	Val	Glu	Asp	Ala	Ala	Asp	195	200	205	
Gly	Ile	Lys	Glu	Gly	Gln	Pro	Asp	Glu	Thr	Glu	Glu	Lys	Glu	Pro	Glu	210	215	220	
Glu	Lys	Glu	Met	Thr	Leu	Glu	Glu	Tyr	Glu	Lys	Leu	Leu	Glu	Glu	Lys	225	230	235	240
Arg	Lys	Ala	Leu	Ser	Ala	Leu	Lys	Ala	Glu	Glu	Arg	Lys	Val	Glu	Val	245	250	255	
Asp	Lys	Asp	Phe	Glu	Ser	Met	Gln	Gln	Leu	Ile	Asn	Lys	Lys	Lys	Asp	260	265	270	
Glu	Asp	Ser	Val	Phe	Ile	Lys	Leu	Gly	Ser	Asp	Lys	Asp	Lys	Lys	Lys	275	280	285	
Glu	Ala	Ala	Glu	Lys	Glu	Lys	Val	Lys	Lys	Ser	Val	Ser	Ile	Asn	Glu	290	295	300	
Phe	Leu	Lys	Pro	Ala	Glu	Gly	Asp	Arg	Tyr	Gly	Gly	Arg	Gly	Arg	Gly	305	310	315	320
Arg	Gly	Arg	Gly	Pro	Arg	Gly	Gly	Gly	Tyr	Gly	Gly	Gly	Asn	Arg	Met	325	330	335	
Phe	Ser	Thr	Ser	Ala	Pro	Ala	Ile	Glu	Asp	Pro	Gly	Glu	Phe	Pro	Thr	340	345	350	
Leu	Gly	Gly	Lys													355			

<210> 13
 <211> 1310
 <212> ADN
 <213> Beta vulgaris
 <400> 13

5

```

ccgacaagtt aggggttttc caagggtttt ctgaaaaaca gcgaataaca atggcaacca 60
ctaacccttt cgacttgctc gacgacgatg ctgaggaccc agccctcctt attgctgcgc 120
aggagcagaa ggtttccgcc gtcgttgccg gagataagaa aactccggca gtcgctgcta 180
agcctgctaa actccctact aagcctcttc ctcttctca agctgtgaga gaggcaagga 240
atgatggtgg tcgtggtgga ggtggccgcg ggggccgtgg ttatgggcgg ggacgtggtc 300
caggtggacc taatagagat tcaacaaata atgatgaaat atatcccaac gagaatgggg 360
gttctatggg atataggagg gacagagata agccatctga aagacgtgga ggatatggcg 420
gtcctcgtgg tggttatcgt ggaggacgac gtggagggtta tgataatgga gaagctgctg 480
aaggagaacg tcctaggagg atgtatgaac gccgtagtgg cactggacga ggaggtgaga 540
ttaaacgtga gggttctggt cgtggaaact ggggaŕctcc tactgatgag atagctccgg 600
agactgaaga acctgttgtg gaaaatgaag cagctgttgc agctgataag ccagcaggag 660
agggagaaaa tgttgatgct gaaaaggaga gtcaagagaa ggaagttgta gaagcagagc 720
ctgaagaaaa ggaaatgact cttgaggagt atgagaaggt attggaggag aagaggaagg 780
ccttgctatc attgaaaggg gaggaagaa aggtggattt ggacaaggag tttgaatcta 840
tgcagctggt ttcaaagaag aagaatgatg atgaggtttt cataaagctg ggttctgata 900
aggacaagag aaaggaggct gcagaaagag aagaaaggtc caagaagtct gtgagcatca 960
atgaatttct taagcctgcc gagggtgacg gataccacag gcgtggaaga ggaagaggcc 1020
gtggtggtag gggaggctat ggtggaggat acggcatgaa caatgcatct gtccttctta 1080
ttgaggatcc caatcaattc ccattcttgg gtgcgaactg agtttttgtc cgttggtgtc 1140
ttagttattt ttgggtcttt cttatatatt gagacttatt tatgatgttc aggagcctca 1200
tcaattacaa aaaaagatat ttgacaggaa taatgtgttt ttcctgtgtt aagagtgtaa 1260
atcttagatg tttcatcttt caaaaaaaa aaaaaaaact cgagggggggg 1310

```

<210> 14
 <211> 356
 <212> PRT
 <213> Beta vulgaris
 <400> 14

10

```

Met Ala Thr Thr Asn Pro Phe Asp Leu Leu Asp Asp Asp Ala Glu Asp
  1                      5                      10                      15

Pro Ala Leu Leu Ile Ala Ala Gln Glu Gln Lys Val Ser Ala Val Val
          20                      25                      30

Ala Gly Asp Lys Lys Thr Pro Ala Val Ala Ala Lys Pro Ala Lys Leu
          35                      40                      45

Pro Thr Lys Pro Leu Pro Pro Ser Gln Ala Val Arg Glu Ala Arg Asn
          50                      55                      60

```

15

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

Gly Arg Gly Arg Gly Gly Arg Gly Gly Tyr Gly Gly Gly Tyr Gly Met
325 330 335

Asn Asn Ala Ser Ala Pro Ser Ile Glu Asp Pro Asn Gln Phe Pro Ser
340 345 350

Leu Gly Ala Asn
355

<210> 15
<211> 1155
<212> ADN
<213> Beta vulgaris
<400> 15

gatggacgag	aattatatc	gtacttgttt	cgctcaatcc	ggcgagcttg	ttaatgttaa	60
aatcatccgt	aataagcaaa	ccatgcagtc	agagtgcctat	ggatttattg	agttttccac	120
ccatgctgct	gctgaaagga	ttttgcagac	ttacaataac	accttgatgc	caaagtgtga	180
gcaaaactac	agactgaatt	gggctttcta	tggatctggt	gagaagcgtg	gagaggatgc	240
ttctgattat	acaatttttg	ttggggattt	agctccagat	gttactgatt	acacattgca	300
agagacattt	agagttcgct	atccatctgt	aaaagggtgct	aagggttgta	tagatagact	360
gacaagtaga	tcaaagggtt	atggatttgt	tcgtttcgga	gatgaaagtg	aacaagcacg	420
tgccatgtca	gagatgaatg	gaatgatgtg	cttaggccgt	gcaatgcgta	ttggagcagc	480
tgcaaacaag	aaaagtgttg	gcggaacagc	ttcatatcag	aataatcagg	gaactccaaa	540
tgacagtgat	ccgagtaaca	ctactatatt	tgttggcaat	ttggattcta	atgtgactga	600
tgaacatttg	agacaaacat	ttagccctta	cggagaattg	gtccatgtaa	aaattcctgc	660
gggcaaacag	tgcggggttg	ttcaatttac	taacagaagt	agtgtgagg	aagcattgag	720
ggtattgaac	ggaatgcaat	taggcggacg	aaatgttaga	ctttcgtggg	gccgtagtc	780
taacaacaga	cagtctcaac	ctgaccagaa	ccagtggaac	aatgctgctt	attatggtta	840
tcctcaagga	tacgactctt	atggatatgt	atctgctcct	caagacccaa	acatgtacta	900
tggtggctac	cctggttatg	gtggttacgc	gatgcctcag	caggctcaga	tgccattgca	960
acaacagtga	tctaccttat	gccaagcagg	agaggtcggt	tgccagggag	ctgtcattgt	1020
acttgagggc	tgagcttctg	gagttggatg	attcctccca	gagatggcag	aatgtagtat	1080
aacttggtca	ttgtgctggt	cgaattttat	ttactgtctt	gggtttttgc	tctgtgctgc	1140
ttttttgtag	cttgc					1155

<210> 16
<211> 322
<212> PRT
<213> Beta vulgaris
<400> 16

Met Asp Glu Asn Tyr Ile Arg Thr Cys Phe Ala Gln Ser Gly Glu Leu
1 5 10 15

Val Asn Val Lys Ile Ile Arg Asn Lys Gln Thr Met Gln Ser Glu Cys
 20 25 30

Tyr Gly Phe Ile Glu Phe Ser Thr His Ala Ala Ala Glu Arg Ile Leu
 35 40 45

Gln Thr Tyr Asn Asn Thr Leu Met Pro Asn Val Glu Gln Asn Tyr Arg
 50 55 60

Leu Asn Trp Ala Phe Tyr Gly Ser Gly Glu Lys Arg Gly Glu Asp Ala
 65 70 75 80

Ser Asp Tyr Thr Ile Phe Val Gly Asp Leu Ala Pro Asp Val Thr Asp
 85 90 95

Tyr Thr Leu Gln Glu Thr Phe Arg Val Arg Tyr Pro Ser Val Lys Gly
 100 105 110

Ala Lys Val Val Ile Asp Arg Leu Thr Ser Arg Ser Lys Gly Tyr Gly
 115 120 125

Phe Val Arg Phe Gly Asp Glu Ser Glu Gln Ala Arg Ala Met Ser Glu
 130 135 140

Met Asn Gly Met Met Cys Leu Gly Arg Ala Met Arg Ile Gly Ala Ala
 145 150 155 160

Ala Asn Lys Lys Ser Val Gly Gly Thr Ala Ser Tyr Gln Asn Asn Gln
 165 170 175

Gly Thr Pro Asn Asp Ser Asp Pro Ser Asn Thr Thr Ile Phe Val Gly
 180 185 190

Asn Leu Asp Ser Asn Val Thr Asp Glu His Leu Arg Gln Thr Phe Ser
 195 200 205

Pro Tyr Gly Glu Leu Val His Val Lys Ile Pro Ala Gly Lys Gln Cys
 210 215 220

Gly Phe Val Gln Phe Thr Asn Arg Ser Ser Ala Glu Glu Ala Leu Arg
 225 230 235 240

Val Leu Asn Gly Met Gln Leu Gly Gly Arg Asn Val Arg Leu Ser Trp
 245 250 255

Gly Arg Ser Pro Asn Asn Arg Gln Ser Gln Pro Asp Gln Asn Gln Trp
 260 265 270

Asn Asn Ala Ala Tyr Tyr Gly Tyr Pro Gln Gly Tyr Asp Ser Tyr Gly
275 280 285

Tyr Val Ser Ala Pro Gln Asp Pro Asn Met Tyr Tyr Gly Gly Tyr Pro
290 295 300

Gly Tyr Gly Gly Tyr Ala Met Pro Gln Gln Ala Gln Met Pro Leu Gln
305 310 315 320

Gln Gln

<210> 17
<211> 1200
5 <212> ADN
<213> Beta vulgaris
<400> 17

```

aaaaacctct ttttctctct cctaaatcac aacaatggcg atactctcag attacgagga 60
agaagaacac caaccacaac cagaaaagaa gcaaccttca aagaaatddd cagcaacttt 120
cgatccttcg aatccgctag ggtttcttca atctactctc gaattcgtct caaaagagtc 180
cgattttttc gctaaggaat catctgcgaa agatgttggt tctctgggtc agaaagtga 240
ggagaagtac attgaagaag tagagaataa gaagaagaag cttctagatg aatctgccgc 300
tgccgcgcgc gccgcgcgtg ctgctgctgc gtcgtcgtct tcatctgatt tggagaagaa 360
ggttgatgat aatgagagtg cggaagagac agagaaatct aagtacaaag ctccaaacag 420
tgggaatggg caagatctcg agaactactc atggatacag tccttgcaag aagttactgt 480
taatgttcct gttccacctg gaacaaagtc taggtttatc gattgtcaga taaagaagaa 540
tcatctgaaa gttggcctca agggtcagcc tcccatcatc gatgggtgaac tgttcaagcc 600
tgttaagcca gatgattggt tttggagttt ggaggatcaa aagtcaatct ctatgctgct 660
aacaagcat gatcaaatgg agtgggtggag aagtctggtc aaagggtgaac ctgaaatcga 720
cactcagaag gttgaacctg agagcagtaa gctgtctgac ttggaccctg aaacaaggtc 780
aactgttgag aagatgatgt ttgaccaaag gcaaaaatcc atgggcttgc ccacaagtga 840
tgatatgcag aagcaagaca tgctgaagaa gttcatgtcc gagcatccgg aaatggactt 900
ttctaacgcg aagtttaact agatatcgat gtcggtgatg gactatgatt ttttgggtgg 960
caaattctcg aaacaggaac tgaagaaagc ttttggtatg tctaatactg agcttggtca 1020
tagtagttac agtctctagg gtagatgtct catgaagagg ggaacattgc tttttgttta 1080
actcttattt atatgcaagt gatattcggg ttgctaagca gtacattcgt gcacccctgcg 1140
cttgattcgg gtctgtttca atcatatatg taatgttata gctgcaaaaa aaaaaaaaaa 1200

```

10 <210> 18
<211> 295
<212> PRT
<213> Beta vulgaris
15 <400> 18

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

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

Lys Gln Asp Met Leu Lys Lys Phe Met Ser Glu His Pro Glu Met Asp
 275 280 285

Phe Ser Asn Ala Lys Phe Asn
 290 295

<210> 19

<211> 1050

5 <212> ADN

<213> Mus musculus

<400> 19

```

gaaaagagag gagagagaaa accaaaatca acaaaaatgg cggaacatct agcatcgata 60
ttcgggacag agaaagacag agtgaactgt ccattctact tcaagatcgg agcttgtaga 120
catggagatc gttgctcaag gcttcatact aagcctagta ttagccctac tttgttgctt 180
gctaatatgt atcaacgccc tgatatgatt actcctggtg ttgatcctca aggacagcct 240
cttgatcctc gcaaaattca acaacatttt gaggattttt atgaggattt atttgaggaa 300
ctaagcaagt atggggagat tgaaagtctc aacatctgtg acaatttggc tgaccacatg 360
gttgggaaatg tttatgtgca gttcagagag gaagaacatg ctggcgaggc actacgaaac 420
ttgagtggaa gatttttatgc cggtcgtcca atcattgttg atttttctcc tgtaacggac 480
ttcagagaag caacctgcag acagtatgag gaaaatgtgt gcaatcgtgg aggttactgc 540
aactttatgc atttgaaaaa aattagcagg gagcttaggc gacagttgtt tggaaggtag 600
agaaggaggc atagccgtag tagaagtcgc agtcctcaag cacatcgggg gcatggagat 660
cgtccacatg gtggccgtgg ttatggtaga agagatgatg atagaaatca gcggtaccat 720
gacaagggaa gaaggcctag aagccgtagc cctgggcata gaggacgaag cagaagccct 780
cccggcagga gggataggag tccagtgagg gagaatagtg aggagagaag agcaaagatt 840
gcacaatgga acagggaaaa ggaacaggca gacactggta ataacgatgt taatcatgat 900
gtcactgaca accatgcaaa tggatttcag gacaatgggg aggattacta tgaccatcct 960
cagcagtaac tggatgaagt gcacaagcag gctttattca ctacttctgg tttgctgtta 1020
tcagagtctg ctcgtttgca ggatttttcg 1050

```

10

<210> 20

<211> 310

<212> PRT

<213> Mus musculus

15 <400> 20

Met Ala Glu His Leu Ala Ser Ile Phe Gly Thr Glu Lys Asp Arg Val
 1 5 10 15

Asn Cys Pro Phe Tyr Phe Lys Ile Gly Ala Cys Arg His Gly Asp Arg
 20 25 30

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

Val Thr Asp Asn His Ala Asn Gly Phe Gln Asp Asn Gly Glu Asp Tyr
290 295 300

Tyr Asp His Pro Gln Gln
305 310

<210> 21

<211> 117

5 <212> PRT

<213> Arabidopsis thaliana

<400> 21

Met Val Thr Thr Pro His Glu Lys Ala Thr Asp Ser Lys Lys Ser Gly
1 5 10 15

Thr Glu Ser Asn Ser Gln Pro Ile Val Gly Asp Ser Ser Tyr Glu Arg
20 25 30

Ser Lys Val Gly Asp Arg Glu Arg Glu Ser Asp Arg Glu Lys Glu Arg
35 40 45

Gly Arg Glu Arg Asp Arg Gly Arg Ser His Arg Gly Arg Asp Ser Asp
50 55 60

Arg Asp Ser Asp Arg Glu Arg Asp Lys Leu Lys Asp Arg Ser His His
65 70 75 80

Arg Ser Arg Asp Arg Leu Lys Asp Ser Gly Gly His Ser Asp Lys Ser
85 90 95

Arg His His Ser Ser Arg Asp Arg Asp Tyr Arg Asp Ser Ser Lys Asp
100 105 110

Arg Arg Arg His His
115

10 <210> 22

<211> 55

<212> PRT

<213> Arabidopsis thaliana

<400> 22

15

Met Ile Tyr Thr Ala Ile Asp Asn Phe Tyr Leu Thr Asp Glu Gln Leu

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

<210> 23

<211> 17

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Sonda o Iniciador

<400> 23

10 catcaaatct ttcaggg 17

<210> 24

<211> 18

<212> ADN

<213> Secuencia Artificial

15 <220>

<223> Descripción de la Secuencia Artificial: Sonda o Iniciador

<400> 24

ctttatttta ctgtacag 18

20

25

30

35

40

REIVINDICACIONES

1. Un método para mejorar la tolerancia al estrés en células y organismos que comprende la manipulación genética y/o bioquímica en dichas células y organismos de:

(A) al menos un polipéptido con una secuencia de aminoácidos que es al menos 50% idéntica al polipéptido representado en la SEQ ID NO: 3 y/o un fragmento funcional del mismo en donde dicho fragmento funcional se caracteriza por:

(i) contener un dominio RS; y
(ii) mejorar la tolerancia al estrés en una planta transgénica cuando se expresa transgénicamente en células, tejidos o partes de dicha planta; o

(B) un ácido nucleico que codifica al polipéptido o fragmento funcional del mismo como se define en (A).

2. El método de acuerdo con la reivindicación 1 que comprende la manipulación genética y/o bioquímica en las células y organismos de:

(A) al menos un polipéptido con la secuencia de aminoácidos representada en la SEQ ID NO: 3 y/o un fragmento funcional de la misma en donde dicho fragmento funcional se caracteriza por:

(i) contener un dominio RS; y
(ii) mejorar la tolerancia al estrés en una planta transgénica cuando se expresa transgénicamente en células, tejidos o partes de dicha planta; o

(B) un ácido nucleico que codifica al polipéptido o fragmento funcional del mismo como se define en (A).

3. El método de acuerdo cualquiera de las reivindicaciones 1 ó 2, en donde el ácido nucleico como se define en la reivindicación 1(B) o en la reivindicación 2(B) se escoge del grupo que consiste de:

(a) un ácido nucleico que comprende una secuencia de ADN como la dada en la SEQ ID NO 1 o el complemento de la misma,

(b) un ácido nucleico que comprende la secuencia de ARN correspondiente a la SEQ ID NO 1 o el complemento de la misma,

(c) un ácido nucleico que se degenera como resultado del código genético hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),

(d) un ácido nucleico que diverge debido a las diferencias en el uso del codón entre los organismos hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),

(e) un ácido nucleico que diverge debido a las diferencias en alelos hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),

(f) un ácido nucleico que codifica un fragmento funcional de la proteína codificada por la secuencia de ADN como la dada en la SEQ ID NO 1, en donde dicho fragmento funcional es como se define en cualquiera de las reivindicaciones 1 ó 2.

4. El método de acuerdo con cualquiera de las reivindicaciones 1, 2, ó 3(a) y (c) hasta (f), caracterizado porque dicho ácido nucleico es ADN, ADNc, ADN genómico o ADN sintético.

5. El método de acuerdo con cualquiera de las reivindicaciones 1 a 4, en donde dicho ácido nucleico se deriva de una célula eucariota.

6. El método para mejorar o alterar la tolerancia al estrés en una planta de acuerdo con cualquiera de las reivindicaciones 1 a 5, que comprende la expresión en células, tejidos o partes de dicha planta de al menos un ácido nucleico como se define en cualquiera de las reivindicaciones 1 a 5; o de al menos un polipéptido y/o fragmento

funcional del mismo como se define en cualquiera de las reivindicaciones 1 a 5; o al menos una molécula antisentido que contiene una secuencia complementaria a la de dicho ácido nucleico.

5 7. El método de acuerdo con cualquiera de las reivindicaciones 1 a 6, en donde dicho estrés es estrés osmótico, estrés por salinidad, estrés por sequía, estrés por congelación o estrés por frío.

10 8. Un método para la producción de plantas transgénicas, células vegetales o tejidos vegetales que tienen tolerancia mejorada al estrés, que comprende la introducción de un ácido nucleico en un formato expresable, o un vector que contiene dicho ácido nucleico, dentro de dicha planta, célula vegetal o tejido vegetal, en donde dicho ácido nucleico codifica un polipéptido con una secuencia de aminoácido que es al menos 50% idéntica a la del polipéptido representado en la SEQ ID NO: 3 o fragmento funcional del mismo, en donde dicho fragmento funcional se caracteriza por:

15 (i) contener un dominio RS; y
(ii) mejorar la tolerancia al estrés en una planta transgénica cuando se expresa transgénicamente en células, tejidos o partes de dicha planta.

20 9. El método de acuerdo con la reivindicación 8, que comprende la introducción de un ácido nucleico en un formato expresable, o un vector que contiene dicho ácido nucleico, dentro de dicha planta, célula vegetal o tejido vegetal, en donde dicho ácido nucleico codifica un polipéptido con la secuencia de aminoácidos representada en la SEQ ID NO: 3 o fragmento funcional de la misma, en donde dicho fragmento funcional se caracteriza por:

25 (i) contener un dominio RS; y
(ii) mejorar la tolerancia al estrés en una planta transgénica cuando se expresa transgénicamente en células, tejidos o partes de dicha planta.

10. El método de acuerdo con cualquiera de las reivindicaciones 8 ó 9, en donde el ácido nucleico se escoge del grupo que consiste de:

30 (a) un ácido nucleico que comprende una secuencia de ADN como la dada en la SEQ ID NO 1 o el complemento de la misma,
(b) un ácido nucleico que comprende la secuencia de ARN correspondiente a la SEQ ID NO 1 o el complemento de la misma,
(c) un ácido nucleico que se degenera como resultado del código genético hasta una secuencia de nucleótidos com
35 la dada en la SEQ ID NO 1 o como se define en (a) o (b),
(d) un ácido nucleico que diverge debido a las diferencias en el uso del codón entre los organismos hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),
(e) un ácido nucleico que diverge debido a las diferencias en alelos hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),
40 (f) un ácido nucleico que codifica un fragmento funcional de la proteína codificada por la secuencia de ADN como la dada en la SEQ ID NO 1, en donde dicho fragmento funcional es como se define en cualquiera de las reivindicaciones 8 ó 9.

45 11. El método de acuerdo con cualquiera de las reivindicaciones 8, 9, ó 10(a) y (c) hasta (f), caracterizado porque dicho ácido nucleico es ADN, ADNc, ADN genómico o ADN sintético.

12. El método de acuerdo con cualquiera de las reivindicaciones 8 a 11, en donde dicho ácido nucleico se deriva de una célula eucariota.

50 13. El método de acuerdo con cualquiera de las reivindicaciones 8 a 12, en donde el ácido nucleico o un vector que contiene dicho ácido nucleico es introducido en forma estable en el genoma de la célula vegetal.

14. El método de cualquiera de las reivindicaciones 8 a 13, que comprende además la regeneración de una planta a partir de dicha célula vegetal.

- 5 15. Un método para la producción de plantas alteradas, células vegetales o tejidos vegetales con tolerancia mejorada al estrés, que comprende la introducción de un polipéptido con una secuencia de aminoácido que es al menos 50% idéntica a la del polipéptido representado en la SEQ ID NO: 3 y/o fragmento funcional del mismo directamente dentro de una célula, un tejido o un órgano de dicha planta, en donde dicho fragmento funcional se caracteriza por:
- 10 (i) contener un dominio RS; y
(ii) mejorar la tolerancia al estrés en una planta transgénica cuando se expresa transgénicamente en células, tejidos o partes de dicha planta.
- 15 16. El método de la reivindicación 15 que comprende la introducción de un polipéptido con la secuencia de aminoácidos como la representada en la SEQ ID NO 3 y/o fragmento funcional de la misma directamente en una célula, un tejido o un órgano de dicha planta, en donde dicho fragmento funcional se caracteriza por:
- (i) contener un dominio RS; y
(ii) mejorar la tolerancia al estrés en una planta transgénica cuando se expresa transgénicamente en células, tejidos o partes de dicha planta.
- 20 17. El método de cualquiera de las reivindicaciones 8 a 16, en donde dicho estrés es estrés osmótico, estrés por salinidad, estrés por sequía, estrés por congelación o estrés por frío.
18. El método de cualquiera de las reivindicaciones 8 a 16 para incrementar el rendimiento.
- 25 19. El uso del ácido nucleico como se define en cualquiera de las reivindicaciones 8 a 13, un vector que contiene dicho ácido nucleico, o el polipéptido y/o fragmento funcional del mismo como se define en cualquiera de las reivindicaciones 15 ó 16 para incrementar el rendimiento.
- 30 20. El uso del ácido nucleico como se define en cualquiera de las reivindicaciones 8 a 13, un vector que contiene dicho ácido nucleico, o el polipéptido y/o fragmento funcional del mismo como se define en cualquiera de las reivindicaciones 15 ó 16 para estimular el crecimiento de la planta, caracterizado porque dicho crecimiento de la planta puede ser en cualquier parte de esa planta, tal como raíz, hoja, semilla.
- 35 21. El uso de una planta que puede ser obtenida por cualquiera de los métodos de las reivindicaciones 8 a 16 para cultivar sobre un suelo con un contenido de sal de más de 1 mM de iones de sal.
22. Un polipéptido de la secuencia de aminoácidos como la representada en la SEQ ID NO: 3.
- 40 23. Un ácido nucleico que codifica al polipéptido como se define en la reivindicación 22.
24. El ácido nucleico de acuerdo con la reivindicación 23 escogido del grupo que consiste de:
- a) un ácido nucleico que comprende una secuencia de ADN como la dada en la SEQ ID NO 1 o el complemento de la misma,
- 45 b) un ácido nucleico que comprende la secuencia de ARN correspondiente a la SEQ ID NO 1 o el complemento de la misma,
- c) un ácido nucleico que se degenera como resultado del código genético hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),
- 50 d) un ácido nucleico que diverge debido a las diferencias en el uso del codón entre los organismos hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b),
- e) un ácido nucleico que diverge debido a las diferencias en alelos hasta una secuencia de nucleótidos como la dada en la SEQ ID NO 1 o como se define en (a) o (b).

25. El ácido nucleico de acuerdo con cualquiera de las reivindicaciones 23, 24(a) y (c) hasta (e), caracterizado porque dicho ácido nucleico es ADN, ADNc, ADN genómico o ADN sintético.
- 5 26. El ácido nucleico de acuerdo con cualquiera de las reivindicaciones 23 a 25, en donde dicho ácido nucleico se deriva de una célula eucariota.
27. Un vector que contiene al ácido nucleico como se define en cualquiera de las reivindicaciones 23 a 26.
- 10 28. El vector de acuerdo a la reivindicación 27, en donde dicho ácido nucleico está operativamente enlazado a una o más secuencias de control permitiendo la expresión de dicho ácido nucleico en células huésped procariontas y/o eucariotas.
- 15 29. El vector de acuerdo a la reivindicación 28, en donde dicha secuencia de control es un promotor constitutivo.
- 30 30. Un método para la producción de plantas transgénicas, células vegetales o tejidos vegetales, que comprende la introducción del ácido nucleico como se define en cualquiera de las reivindicaciones 23 a 26 en un formato expresable, o un vector que contiene dicho ácido nucleico, dentro de dicha planta, célula vegetal o tejido vegetal.
- 20 31. Un método para la producción de plantas alteradas, células vegetales o tejidos vegetales, que comprende la introducción del polipéptido como se define en la reivindicación 22 directamente en una célula, un tejido o un órgano de dicha planta.
- 25 32. Un método para efectuar la expresión del polipéptido como se define en la reivindicación 22, que comprende la introducción de un ácido nucleico que codifica a dicho polipéptido operativamente enlazado a una o más secuencias de control o un vector que contiene dicho ácido nucleico en forma estable dentro del genoma de una célula vegetal.
- 30 33. El método de la reivindicación 30 ó 32, que comprende además la regeneración de una planta a partir de dicha célula vegetal.
- 35 34. Una célula de una planta transgénica que puede ser obtenida por medio del método de la reivindicación 30, toda célula vegetal conteniendo un transgén como se define en cualquiera de las reivindicaciones 23 - 26.
- 40 35. Una planta transgénica tolerante al estrés por salinidad que expresa en forma transgénica al menos un ácido nucleico como se define en cualquiera de las reivindicaciones 23 a 26, o una molécula antisentido que contiene una secuencia complementaria a la de dicho ácido nucleico, o al menos un polipéptido como se define en la reivindicación 22.
- 45 36. Una planta transgénica que expresa transgénicamente al menos un ácido nucleico como se define en cualquiera de las reivindicaciones 23 a 26, o una molécula antisentido que contienen una secuencia complementaria a la de dicho ácido nucleico, o al menos un polipéptido expresado transgénicamente como se define en la reivindicación 22, en donde la planta muestra una alteración de su fenotipo.
- 50 37. Una parte cosechable de una planta de la reivindicación 35 ó 36, en donde dicha parte contiene al menos un ácido nucleico expresado transgénicamente como se define en cualquiera de las reivindicaciones 23 a 26, o una molécula antisentido que contiene una secuencia complementaria a dicho ácido nucleico, o al menos un polipéptido expresado transgénicamente como se define en la reivindicación 22.
38. Una parte cosechable de acuerdo a la reivindicación 37 que se selecciona del grupo que consiste de semillas, hojas, frutos, cultivos de tallo, rizomas, raíces, tubérculos y bulbos.
39. La progenie transgénica derivada de cualquiera de las plantas o partes de la planta de cualquiera de las reivindicaciones 35 a 38, en donde dicha progenie contiene al menos un ácido nucleico expresado transgénicamente como se define en cualquiera de las reivindicaciones 23 a 26, o una molécula antisentido que contiene una

secuencia complementaria a dicho ácido nucleico, o al menos un polipéptido expresado transgénicamente como se define en la reivindicación 22.

5 40. Nuevas cepas de levadura u otras eucariotas unicelulares más tolerantes al estrés por salinidad que expresan cualquiera de los ácidos nucleicos definidos en cualquiera de las reivindicaciones 23 a 26 y/o del polipéptido como se define en la reivindicación 22.

10 41. Un sistema de cultivo de células *in vitro* que comprende células huésped animales o de planta tolerantes al estrés por salinidad obtenido como resultado de la expresión transgénica de al menos un ácido nucleico como se define en cualquiera de las reivindicaciones 23 a 26, o de al menos un polipéptido expresado transgénicamente como se define en la reivindicación 22, o una molécula antisentido que contiene una secuencia complementaria a la de dicho ácido nucleico.

15 42. Una célula huésped que contiene al ácido nucleico expresado transgénicamente de acuerdo con cualquiera de las reivindicaciones 23 a 26, o el vector de acuerdo con cualquiera de las reivindicaciones 27 a 29.

43. Una célula huésped de acuerdo a la reivindicación 42, escogida del grupo que consiste de una célula bacteriana, de insecto, de hongo, de levadura, de planta y de animal.

20 44. Un método para producir un polipéptido de la reivindicación 22 que comprende el cultivo de una célula huésped de la reivindicación 42 ó 43 bajo las condiciones que permitan la expresión del polipéptido y recuperar el polipéptido producido del cultivo.

25

30

35

40

45

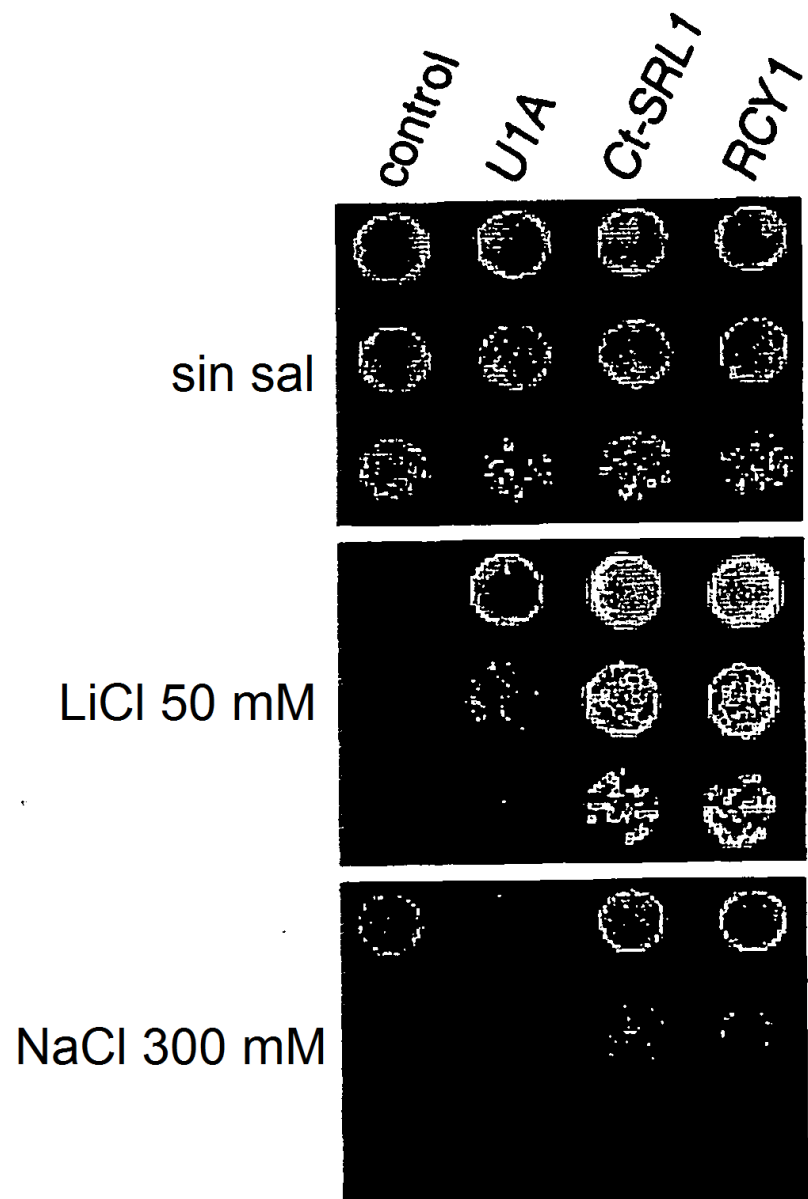


FIGURA 1

A

MIYTAIDNFYLTDEQLKASPSRKDGIDETTEISLRIYGCDLI
 QEGGILLKL PQAVMATGQVLFQRFYCKKSLAKFDVKIVA
 ASCVWLASKLEENPKKARQVIIVFHRMECRRENLPLEHL
 DMYAKKFSELKVELSRTERHILKEMGFVCHVEHPHKFIS
 NYLATLETPPELRQEAWNLANDSLRTTLCVRF RSEVVAC
 GVVYAAARRFQVPLPENPPWWKAFDADKSSIDEVCRVL
 AHL YSLPKAQYISVCKDGKPFTFSSRSGNSQGGQSATKDL
 LPGAGEAVDTKCTAGSANN DLKDGMVTTPHEKATDSKK
 SGTESNSQPIVGDSSYERSKVGDREERESDREKERGRER
DRGRSHRGRDSDRDSDRERDKLKDRSHHRSRDRLKDS
 GGHS DKSRHHSSRDRDYRDSSKDRRRHH

B

MNLPTKPSGSTGDMTRGSEDTARRPPSVKASLSVSFGQ
 RAPHRASTRGSSPVRRPPPTGYDRNGGDEVQQRSPRR
SQSRDYYS DRDSDRQRE REREKDRERERGRDRYRERE
RDYGNDRRSRRDYDSRSRRNDYEDDRSRHDRRSRSRS
RSRSRSVQIEREPTPKRDSSNKEKSAVTVNSNLAKLKD L
 YGDASSQKRDEGFGTRKDS SSEEVIKLGGSWR

FIGURA 2

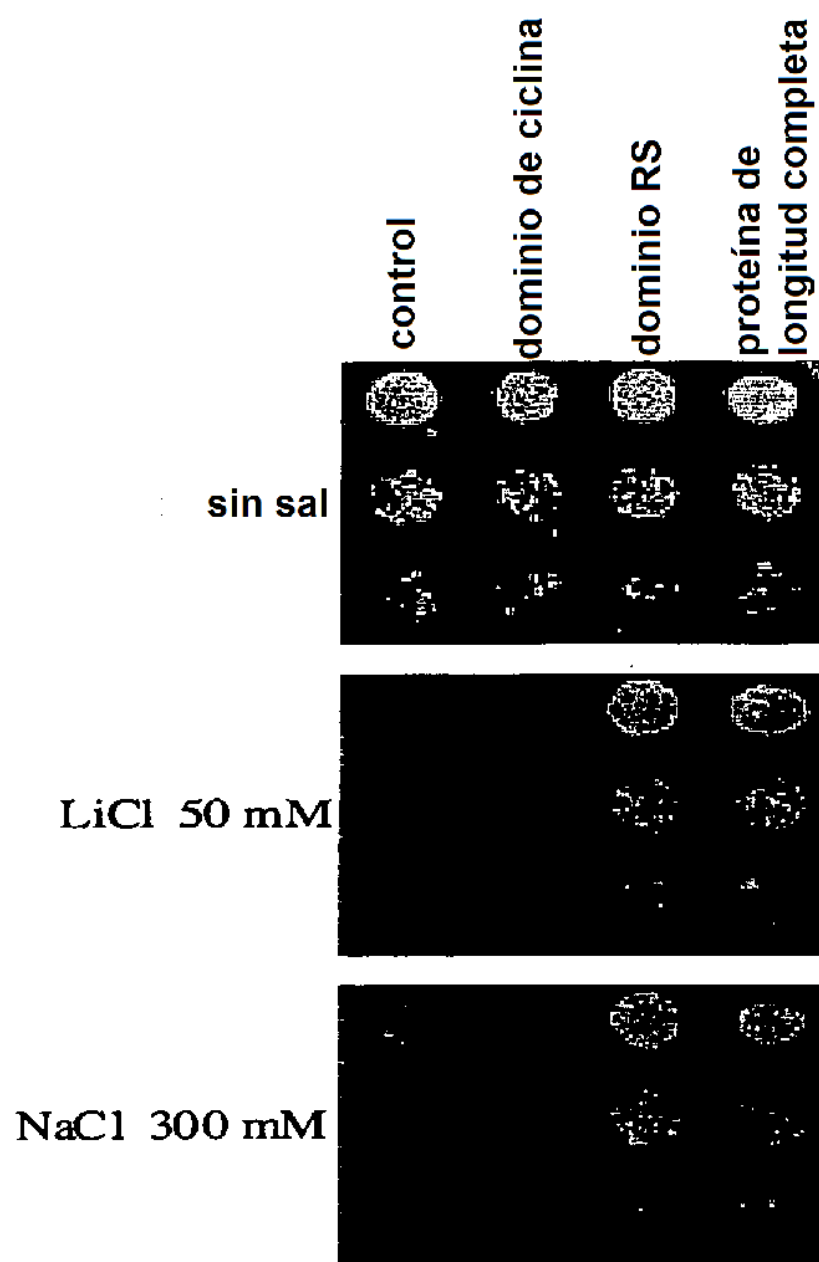


FIGURA 3

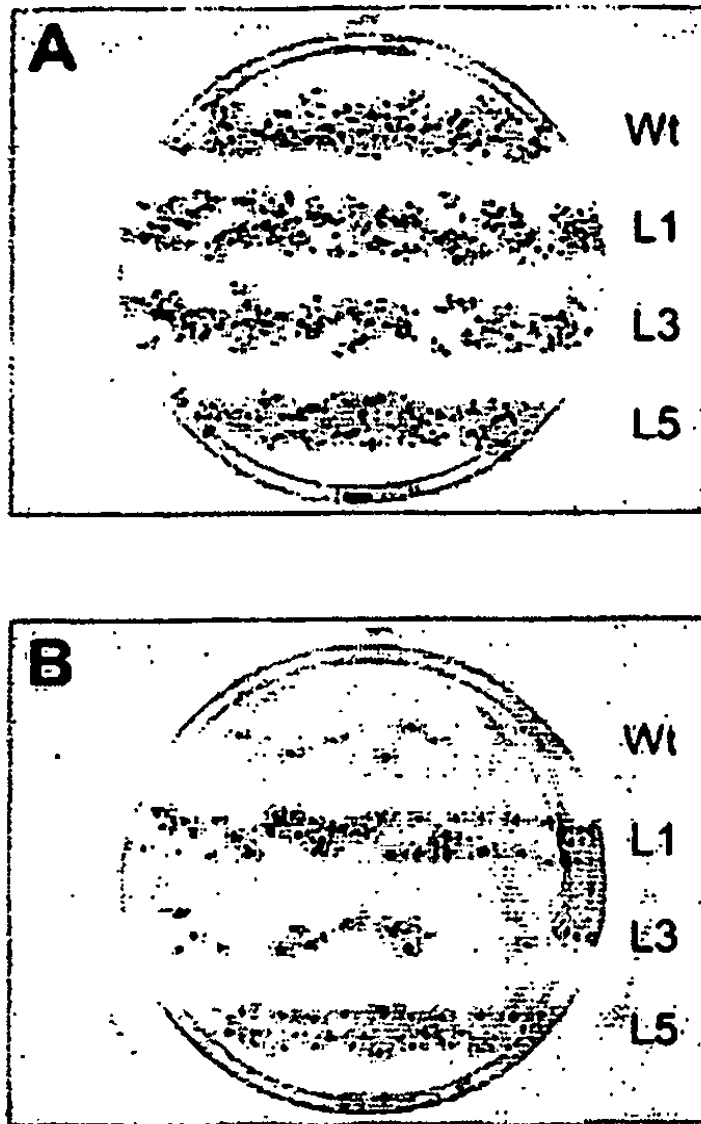


FIGURA 4

Figura 5

SEQ ID NO. 1: secuencia de nucleótidos de Ct-SRL1 (840 nucleótidos). La región que codifica a Ct-SRL1 va desde la posición 15 hasta la 678 (el codón de inicio y el de detención están subrayados)

```

CAAACCTTGA GAAGATGAAT TTACCAACTA AACCTTCTGG TTCAACCGGA
GACATGACCC GTGGCTCAGA AGACACTGCC CGTCGTCCAC CATCAGTAAA
AGCATCTCTC TCTGTTTCAT TTGGTCAGCG TGCACCTCAT CGTGCTTCCA
CCAGAGGCTC TTCTCCTGTT CGCCGTCCTC CACCGACTGG TTATGACAGA
AATGGAGGCG ATGAAGTACA ACAGCGGTCC CCACGTAGAA GCCAGAGCCG
AGACTATTAT TCTGACAGAG ACTCAGATAG ACAACGGGAA AGAGAGAGGG
AGAAAGACCG CGAAAGAGAG AGGGGGAGGG ATAGATACAG AGAAAGGGAG
AGGGATTATG GTAATGATAG GAGATCAAGG CGCGACTATG ATAGTAGAAG
CAGGCGCAAT GATTATGAGG ACGACAGAAG TAGACATGAC CGGAGAAGCA
GGAGCAGAAG CAGAAGTAGG AGCAGGAGTG TGCAGATTGA GCGTGAACCG
ACTCCTAAAA GAGATAGTAG CAACAAAGAG AAATCGGCGG TGACAGTGAA
CAGCAATCTC GCAAAGCTAA AAGATTTGTA TGGAGACGCA AGTAGTCAGA
AAAGGGATGA AGGATTTGGA ACAAGGAAAG ATTCAAGTTC AGAAGAAGTG
ATAAAGCTTG GTGGTTCCTC TTGGAGGTGA AAAAACAAAC AAAACAAAAC
CAAACGTGGG ATTTAAATG CTTCTTCTAT TTAGCAGGAT GATGCATGTT
GTTTAACTAT ACTGTTTGA TTTCCAGCAA ACTATTTGTC ATATGCTTTA
TATTACAGTT TAAGAGTTGA TCTTTATCTT GAAAAAAAAA

```

SEQ ID NO. 2: secuencia de nucleótidos de RCY1 (1433 nucleótidos). La región que codifica a RCY1 putativa va desde la posición 14 hasta la 1262 (los codones putativos de inicio y de detención están subrayados)

```

GTTTAGTGAA TCGATGATTT ACACTGCTAT CGACAATTTT TACCTAACCG
ACGAGCAGCT GAAGGCTTCA CCTTCGAGGA AAGATGGGAT AGATGAAACA
ACTGAAATCT CTCTTAGAAT CTATGGATGT GATCTCATCC AAGAGGGTGG
AATATTGCTC AAACCTACCAC AGGCAGTTAT GGCTACTGGG CAGGTTCTGT
TTCAGCGATT CTATTGCAAG AAGTCTTTGG CTAAATTTGA TGTCAAGATA
GTTGCTGCCA GCTGTGTATG GCTTGCATCA AAACCTGGAAG AAAACCCCTAA
AAAAGCTAGA CAGGTCATCA TCGTATTCCA CAGGATGGAG TGTGCGAGGG
AGAACTTGCC ATTAGAACAT CTGGATATGT ATGCCAAGAA GTTCTCTGAG
TTGAAAGTTG AATTAAGCAG AACTGAGAGA CATATACTGA AAGAGATGGG
TTTTGTTTGT CATGTTGAAC ATCCTCACAA GTTCATATCA AACTACCTTG
CCACATTAGA AACACCTCCA GAATTGAGGC AAGAAGCTTG GAATTTGGCC
AATGATAGTC TGCGTACAAC CCTCTGTGTA AGGTCAGAA GTGAGGTTGT
GGCTTGTTGG GTAGTGTATG CTGCTGCCCC TAGGTTTCAA GTACCACTCC
CTGAGAATCC GCCGTGGTGG AAAGCATTTG ATGCAGATAA ATCTAGTATT
GACGAAGTGT GTAGAGTTCT TGCTCATTTA TACAGTCTTC CAAAGGCTCA
GTATATCTCT GTTTGCAAGG ATGGGAAGCC ATTTACATTT TCTAGCAGAT
CCGGGAATTC TCAAGGTCAA TCAGCGACAA AGGATCTGTT GCCGGGAGCA
GGCGAGGCTG TTGATACTAA ATGTACTGCA GGATCAGCTA ATAACGACTT
GAAGGATGGA ATGGTTACTA CACCACACGA AAAGGCTACA GATTCCAAGA
AAAGTGGTAC CGAGTCAAAC TCTCAGCCAA TTGTAGGAGA CTCAAGCTAT
GAAAGAAGTA AAGTAGGAGA TAGAGAAAGA GAGAGTGATA GAGAGAAGGA
ACGAGGTAGA GAGAGGGACA GGGGTAGGTC TCACAGAGGC AGAGATCTCTG
ACAGAGACAG TGATAGGGAG AGAGACAAAC TCAAAGATCG AAGTCATCAT

```

Figura 5 (Continuación)

```

CGGTCAAGAG ACAGATTGAA GGATTCAGGT GGACATTCAG ATAAATCAAG
GCATCATTCT TCTCGGGACC GTGACTACCG CGACTCATCG AAAGACCGTC
GTAGGCACCA TTAAGCCAAT CTTCTTGTC TCTACATCCC CTTGAGCCTA
CTTGATGTTA AGACAGTATA GTGTTGTATT GTGTTAAGAG TCAAAACCCA
TGTGTA CTTA ATCACATGCT AAGATCACGT TGGTTCGACA TATAAATCGA
GAAAGTCTGA TATGTTTCTA AAAAAAAAAA AAA

```

SEQ ID NO. 3: secuencia de la proteína de Ct-SRL1 (221 aminoácidos)

```

MNLPTKPSGSTGDMTRGSEDARRPPSVKASLSVSFGQRASTRGSSPVRRPPPTGYDR
NGGDEVQQRS PRRSQSRDYYS DRDS DRQRE REKDR ERERGR DRYRE REDYGN DRSR RD
YDS SRSR RNDYED DRSR HRDR SRSR SRSR SRSR SVQIER EPTPKR DSSN KEKSA VTVNS NLAK LK
DLYGDASSQK RDE GFG TRK DSS SSE VIK LGG SSW

```

SEQ ID NO. 4: secuencia de la proteína putativa de RCY1 (416 aminoácidos). La caja cuadrada corresponde a la SEQ ID NO. 21. La secuencia subrayada corresponde a la SEQ ID NO. 22.

MIYTAIDNFYLTDEOLKASPSRKDGIDETTEISLRIYGC DLIOEGGILLKLPOAV

SEQ ID NO. 5: secuencia de nucleótidos de U1A (920 nucleótidos). La región que codifica a U1A va desde la base 30 hasta la base 782 (los codones putativos de inicio y de detención están subrayados)

```

ccacgcgtcc gagagatttg gtgaagacga tggagatgca agaggctaata
caaggaggag gatcggagggt ttctccgaat cagacgattt acatcaacaa
tctcaacgaa aaagtgaagc ttgatgagct gaagaaatcg ctgaatgcag
tggttctctca gttcgggaag atactggaga tattggcgtt taagaccttt
aagcacaag gacaagcttg ggtagtcttc gacaacaccg agtctgcttc
cactgctatt gctaaaatga ataattttcc tttctacgac aaggagatga
gaatacaata tgccaaaaca aatcagatg ttggttgccaa ggccgatggt
acatttggtc ctgcgcgagaa gagaaagaga catgaggaga aaggaggcgg
caagaaaaag aaagaccagc accatgattc tacacagatg ggcattgccca
tgaactcagc atatccagggt gtctatggag ctgcacctcc tctatcgcaa
gtaccatacc ctggtggcat gaaacccaat atgcccaggg caccagctcc
gccaaataat attctctttg tccaaaacct tcctcacgag acaactccaa
tggtgcttca gatgttggtc tgccagtagc aaggatttaa ggaagttaga
atgattgaag ccaaaccggg aatcgccttt gtggagtttg ctgatgagat
gcagtcgacg gtcgcaatgc agggacttca aggtttcaag attcagcaaa
accagatgct catcacgtat gccaagaaat agacaatttc gttttatttg
tggttcgatg agatatgttt gtatctgtca atgttacttc ttgccatggg
ggctgtcttc tgggttggtg gatgctagat atccctctct acttacattt
tttcatcaaa aaaaaaaaaa

```


Figura 5 (Continuación)

SEQ ID NO. 6: secuencia de la proteína putativa de U1A (250 aminoácidos)

MEMQEANQGGGSEVSPNQTIYINNLEKVKLDELKKSLSNAVFSQFGKILEILAFKTFKHKGQ
AWVVDNTESASTAIKMNFPFYDKEMRIQYAKTKSDVVAKADGTFVPREKRKRHEEKGGG
KKKKDQHHDSTQMGMPMNSAYPGVYGAAPPLSQVPYPGGMKPNMPEAPAPPNNILFVQNLPH
ETTPMVLQMLFCQYQGFKEVRMIEAKPGIAFVEFADEMQSTVAMQGLQGFKIQQNQMLITYA
KK

SEQ ID NO. 7: secuencia de proteína de Beta vulgaris (1492 nucleótidos). La secuencia putativa de codificación va desde la posición 51 hasta la 1079 (los codones putativos de inicio y de detención están subrayados)

TCCTCTCTTTTAACCTAATTTGAGCTCTCATTATCCTGATTTTTTCAACCATGGATGCTCAG
AGAGCTCTTCTCGATGAATTAATGGGCGCAGCTCGAAATCTGACTGATGAAGAAAAGAAAGG
TTATAGAGAGATAAAGTGGGATGACAAGGAAGTTTGTGCGCCGTATATGATTTCGATTTTGCC
CTCACGATCTCTTCGTCAATACTCGAAGTGATCTTGGACCATGTCCAAGAGTTCATGACCAA
AAGCTGAAAGAGAGCTTTGAGAACTCTCCAAGGCATGACTCATATGTCCCACGTTTGAAGC
AGAGCTTGCCCAATTTTGTGAGAAGCTGGTGGCAGATTTGGATAGGAAAGTAAGACGTGGGA
GAGAGCGGCTGGACCAGGAGGTTGAACCTCCACCTCCCCCTCCTATTTCTGCAGAAAAAGCT
GAGCAGCTATCTGTACTTGAAGAGAAAAATAAAAAATTTGCTTGAACAAGTAGAGTCACTGGG
AGAAGCTGGCAAAGTCGATGAAGCAGAAGCACTCATGCGAAAGGTGGAAAGTCTTAATTTAG
AGAAAGCTGCATTAACCAACAGCCCCAGAATGCAGCAACAATGCTTACCCAAGAGAAAAAG
ATGGCACTATGTGAAATTTGCGGTTCTTCTTGGTAGCCAATGATGCTGTGGAAAGAACTCA
ATCTCATATAACTGGCAAGCAGCATATTGGCTATGGCATGGTCCGTGATTACCTTGCTGAGT
ATAAGGAGGCTAAGGAGAAGGCAAGAGAAGAGGAAAGATTAGCAAGGGAGAAAGAAGCAGAA
GAACGTCGGAAGCAGAGGGGAAAAGGAAAATGAGAGTAAAAACAGAAGAAGCATCTCCAGTGA
GAGGGACCGTCATCGTGATAGGGATTATGGCCGAGATCGTGAAAGATCACGAGAATGGAACA
ATAGGGGGAATCGAGACGAGGGAAGAGGAATGGATCGGAGAAGGCAATATGATCGCAATGGA
AGGGATGGAGGGAGGAATACGTATCATGGTCGTGAACGTGAAAGGAGCAGGTCACGGTCCCC
TGTTAGGCATGGCCACCGGAGGTGATCTAAGAGTGCTGGTTGCCGATATTAGTAGGCAGTGG
GTTGTGTAGATAAACGATGATCTTAAACCTACTGAGGTAGATGCTTTATATCTCAAGATGTT
TTGTGTCTGTTTTTCGAGGTGTTGCATTGCAGTCTTATTGGGGGTAAACTTTTCTTTATTGT
CCCACAGTGTGAGACTATACTGTCTCCTCTCATCAATCTTGTTAGAGGTCAAAGAGATTGA
GGTAGGTAAACTTCATCGTTGTAATCTTACCTATAGTCAACTTGAGTTTGTCCAATTATA
GCACATGGTCTTTGAAACATTTTTTAATCATGCGGGGTACGCAAGAAAATATGCAACTATG
CTGCATGGCTTGTGTGCAAAAAAAAAAAAAAAAAAAAAAACTCGAGGGGGGGCCCGGTACCA
AGAT

SEQ ID NO. 8: secuencia de la proteína putativa de Beta vulgaris (342 aminoácidos)

MDAQRALLDELMGAARNLTDEEKKGyreIKWDDKEVCAPYMIRFCPHDLFVNTRSDLGPCPR
VHDQKLKESFENSPRHDSYVPRFEAELAQFCEKLVLADLDRKVRGRERLDQEVPPPPPPIS
AEKAEQLSVLEEKIKNLLEQVESLGEAGKVDEAEALMRKVESLNLEKAALTQQPQNAATMLT
QEKKMALCEICGSFLVANDAVERTOQSHITGKQHIGYGMVRDYLAEYKEAKEKAREEERLARE
KEAEERRKQREKENESKNRRSISSEDRDRDRDYGRDRERSREWNNRGNRDEGRGMDRRRQY
DRNGRDGGRNTYHGRERERSRSRSPVRHGHRR

Figura 5 (Continuación)

SEQ ID NO. 9: secuencia de nucleótidos de Beta vulgaris (650 nucleótidos). La secuencia putativa de codificación va desde la posición 14 hasta la 625 (los codones putativos de inicio y de detención están subrayados)

ACGAACAACAAAAATGGCGGCAGCAGATGTTGAAGCAGTAGACTTCGAACCTGAAGAAGATG
ATCTCATGGACGAAGATGGCGGTGCGGCTGAAGCTGACGGCTCTCCTCGAGCTCCTCACCCT
AAGATTAAATCAGCCATTACTGGCGCCGGAGCTCCATCTTCTGGCGGCTTCGGAGCTAAGAA
AACTAAAGSTCGCGGCTTCCGTGAAGACGCCGATgCTGAGCGTAACAGCCGTATGACTGCTC
GTGAATTTGATTCTCTTGACTCCGATgGTGGACCTGGTCCTGCTCGATCAATTGAGGGCTGG
ATTATACTTGTCACGGGAGTGCATGAAGAGGCTCAAGAAGAGGATCTCCTTAATGTCTTTGG
AGAGTTTGGCCAGCTTAAGAATTTGCATTTGAATCTGGATCGTCGTACTGGGTTTGTCAAGG
GTTATGCATTGATCGAGTATGAGAAGTTTGAAGAAGCACAAAGCTGCAATAAAGGAGATGAAT
GGTGCCAAAATGCTTGAGCAGCCGATAAATGTTGATTGGGCATTCTGCAATGGTCCTTACAG
GAGGAGGGGCAACCGAAGAAGATCCCCACGTGGTCACCGATCAAGGAGTCCTAGAAGAAGAT
ATTAAATCTGTTTGCTGCATGTGGAAGTTG

SEQ ID NO. 10: secuencia de la proteína putativa de Beta vulgaris (203 aminoácidos)

MAAADVEAVDFEPEEDDLMDDEDGGAAEADGSPRAPHPKIKSAITGAGAPSSGGFGAKKTKGRGFRED
ADAERNRMTAREFDSLSDGGPGPARSIEGWIIILVTGVHEEAQEEDLLNVFGEFGQLKNLHLNLDR
RTGFVKGYALIEYEKFEEAQAAIKEMNGAKMLEQPINVDWAFcNGPYRRRGNNRRRSRGHRSRSPRR
RY

SEQ ID NO. 11: secuencia de nucleótidos de Beta vulgaris (1464 nucleótidos). La secuencia putativa de codificación va desde la posición 51 hasta la 1119 (los codones putativos de inicio y de detención están subrayados)

TTTTGTAGCATTGATTGCTGAAAAACCCAATTCATATTTTGAAGAAATGACAACAATG
AACCCTTTTGATTGTTGGGTGACAAATGATAACGATGACCCATCTCAGCTTTTAGAGTCTGC
AACTGCTCAGTTGCAGAAAATTGCTGTAAAAAAACCCCAACTCAGGTTGCTCAACAACCTC
AGCAACAGAAAGCTGCAAAGTTACCCACCAAACCTCTTCCTCCAACCTCAAGCTGTCCGGGAG
GCAAAGAATGATTCCCAGCGCGGAGGGGGGCGTGGAGGAGGTCGCGGTAGTGGCCGTGGGCG
TGGTGGATACAATAGGGACTACTCAAACAATGAAAATGCTTTTAAACAGCACTGGAGTAACTG
GCAGTCAAGGGGATGATGGGGAAAGGGAAAGGCGACCTTATGCGGGACCTCGTGGCCCTTAT
CGTGGTGGTCGCCGAGATGGGTTCAACAATGAGGAGGGAAGAGACGGGGAAACGCCCGCGTAG
AACCTATGAGCGACGAAGTGGGACTGGGCGTGGAAGTGAGATCAAACGTGAGGGAGCAGGAC
GTGGAACCTGGGGTGTGAATCAGATGAAGTTGCACCGGTTACTGAGGAAGCTGGAGAACAA
AATGAGAAGAAGTTGAACCCTGAGAATCTTCAGCTGTAGAAGATGCTGCTGATGGCATCAA
GGAGGGCCAGCCAGAtgAGACTGAAGAAAAGGAACCAGAGGAAAAGGAGATGACACTTGAAG
AGTATGAGAAGTTGCTGGAAGAGAAGAGGAAGGCTTTATCAGCACTCAAGGCTGAGGAACGC
AAGTGAGAGTTGACAAAGATTTGAGTCCATGCAACAGCTTATAAACAAAAAAGGATGA
AGACTCAGTTTTCATCAAATTTGGGTTCTGCAAGGATAAGAAGAAGGAAGCAGCTGAAAAGG
AGAAAGTGAAGAAGTCTGTGACGATTAATGAATTTCTGAAGCCTGCTGAAGGGGATAGATAT
GGTGGTCGTGGCAGGGGACGTGGTCGTGGCCCAAGAGGTGGTGGATATGGTGGAGTAAATAG
GATGTTTAGTACGTCTGCTCCAGCAATCGAAGATCCAGGGGAGTCCCAACCCTAGGTGGCA
AGTGAGGCCACATCTTTGAACTTTGGtCTCTATTTGGGGTTTACTTGACCCCTCTGATTT

Figura 5 (Continuación)

TAAGTCATTTGAGTGACAGGAATGGACTTCCAGCTGTGGGTTTCCTGTACCAAATCCACTTT
 TAAGAAAATTTTATGCTTTTAAATTTGTATATTTATTCTGTTAAAAAATAAAAAAAC
 TCGTGCCGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCA
 AGATGGCCTTTGGTGGGTTGAAGAAGGAAAAAGACAGAAACGACTTAATTACCTACTTGAAA
 AAAGCCTGTGAGTAAACAGGCCCTTTTCCTTTGTCGA

SEQ ID NO. 12: secuencia de la proteína putativa de Beta vulgaris (356 aminoácidos)

MTTMNPFDLLGDNDNDPSQLLESATAQLQKIAVKKTPPTQVAQQPQQQKAAKLPTKPLPPTQ
 AVREAKNDSQRGGGRGGGRGSGRGRGGYNRDYSNNENAFNSTGVTGSQDDGERERRPYAGP
 RGPYRGGRDGFNNEEGRDGERPRRTYERRSGTGRGSEIKREGAGRGNWGAESDEVAPVTEE
 AGEQNEKKLNPNENLPAVEDAADGIKEGQPDTEEKEPEEKEMTLEEYEKLLLEEKRLSALK
 AEERKVEVDKDFESMQQLINKKKDEDSVFIKLGSDDKDKKEAAEKEKVKKSVSINEFLKPAE
 GDRYGGRRGRGRGRPRGGGYGGNRMFSTSAPIEDPGEFPTLGGK

SEQ ID NO. 13: secuencia de nucleótidos de Beta vulgaris (1310 nucleótidos). La secuencia putativa de codificación va desde la posición 51 hasta la 1121 (los codones putativos de inicio y de detención están subrayados)

CCGACAAGTTAGGGTTTTTCCAAGGGTTTTCTGAAAAACAGCGAATAACAATGGCAACCACT
 AACCTTTTCGACTTGCTCGACGACGATgCTGAGGACCCAGCCCTCCTTATTGCTGCGCAGGA
 GCAGAAGGTTTCCGCCGTCGTTGCCGGAGATAAGAAAACCTCCGGCAGTCGCTGCTAAGCCTG
 CTAACCTCCCTACTAAGCCTCTTCCTCCTTCTCAAGCTGTGAGAGAGGCAAGGAATGATGGT
 GGTCGTGGTGGAGGTGGCCGCGGGGGCCGTGGTTATGGGCGGGGACGTGGTCCAGGTGGACC
 TAATAGAGATTCAACAAATAATGATGAAATATATCCCAACGAGAATGGGGGTCTATGGGAT
 ATAGGAGAGACAGAGATAAGCCATCTGAAAGACGTGGAGGATATGGCGGTCCTCGTGGTGGT
 TATCGTGGAGGACGACGTGGAGGTTATGATAATGGAGAAGCTGCTGAAGGAGAACGTCCTAG
 GAGGATGTATGAACGCCGTAGTGGCACTGGACGAGGAGGTGAGATTAAACGTGAGGGTTCTG
 GTCGTGGAACCTGGGGATCTCCTACTGATGAGATAGCTCCGGAGACTGAAGAACCTGTTGTG
 GAAAATGAAGCAGCTGTTGCAGCTGATAAGCCAGCAGGAGAGGGAGAAAATGTTGATGCTGA
 AAAGGAGAGTCAAGAGAAGGAAGTTGTAGAAGCAGAGCCTGAAGAAAAGGAAATGACTCTTG
 AGGAGTATGAGAAGGTATTGGAGGAGAAGAGGAAGGCCTTGCTATCATTTGAAAGGGGAGGAA
 AGAAAGGTGGATTTGGACAAGGAGTTTGAATCTATGCAGCTGGTTTCAAAGAAGAAGAATGA
 TGATGAGGTTTTCATAAAGCTGGGTCTGATAAGGACAAGAGAAAGGAGGCTGCAGAAAGAG
 AAGAAAGGTCCAAGAAGTCTGTGAGCATCAATGAATTTCTTAAGCCTGCCGAGGGTGACGGA
 TACCACAGGCGTGGAAGAGGAAGAGGCCGTGGTGGTAGGGGAGGCTATGGTGGAGGATACGG
 CATGAACAATGCATCTGCTCCTTCTATTGAGGATCCCAATCAATTCCCATCTTTGGGTGCGA
 ACTGAGTTTTTTGTCCGTTGTTGTCTTAGTTATTTTGGGTCTTTCTTATATTTTGAGACTTA
 TTTATGATGTTTCAAGAGCCTCATCAATTACAAAAAAGATATTTGACAGGAATAATGTGTTT
 TTCCTGTGTTAAGAGTGTAATCTTAGATGTTTCATCTTTCAAAAAAaAAAAAAACTCG
 AGGGGGGG

SEQ ID NO. 14: secuencia de la proteína putativa de Beta vulgaris (356 aminoácidos)

MATTNPFDLLDDDAEDPALLIAAQEQKVS AVVAGDKKTPAVAAPAKLPTKPLPPSQA VREA
 RNDGGRGGGGRGGRGYGRGRPGGPNRDSTNNDEIYPNENGSMGYREDRDKPSERRGGYGG
 PRGGYRGGRRGYDNGEAAEGERPRRMYERRSGTGRGGEIKREGSGRGNWGSPTDEIAPETE
 EPVVENEA AVAADKPAGEGENVDAEKESQEKEVVEAEPEEKEMTLEEYEKLVLEEKRLALLSL
 KGEERKVDLDKEFESMQLVSKKKNDDEVFIKLGSDKDKRKEAEREERSKKSVSINEFLKPA
 EGDGYHRRGRGRGRGGYGGGYGMNNASAPSIEDPNQFPSLGAN

Figura 5 (Continuación)

SEQ ID NO. 15: secuencia de nucleótidos de Beta vulgaris (1155 nucleótidos). La secuencia putativa de codificación va desde la posición 2 hasta la 970 (los codones putativos de inicio y de detención están subrayados)

GATGGACGAGAATTATATTTCGTACTTGTTTCGCTCAATCCGGCGAGCTTGTTAATGTTAAAA
TCATCCGTAATAAGCAAACCATGCAGTCAGAGTGCTATGGATTTATTGAGTTTTCCACCCAT
GCTGCTGCTGAAAGGATTTTGCAGACTTACAATAACACCTTGATGCCAAATGTTGAGCAAAA
CTACAGACTGAATTGGGCTTTCTATGGATCTGGTGAGAAGCGTGAGAGGATGCTTCTGATT
ATACAATTTTTGTTGGGGATTTAGCTCCAGATGTTACTGATTACACATTGCAAGAGACATTT
AGAGTTCGCTATCCATCTGTAAAAGGTGCTAAGGTTGTGATAGATAGACTGACAAGTAGATC
AAAGGGTTATGGATTTGTTTCGTTTCGGAGATGAAAAGTGAACAAGCACGTGCCATGTCAGAGA
TGAATGGAATGATGTGCTTAGGCCGTGCAATGCGTATTGGAGCAGCTGCAAACAAGAAAAGT
GTTGGCGGAACAGCTTCATATCAGAATAATCAGGGAACCCAAATGACAGTGATCCGAGTAA
CACTACTATATTTGTTGGCAATTTGGATTCTAATGTGACTGATGAACATTTGAGACAAACAT
TTAGCCCTTACGGAGAATTGGTCCATGTAAAAATTCCTGCGGGCAAACAGTGCGGGTTTGT
CAATTTACTAACAGAAGTAGTGCTGAGGAAGCATTGAGGGTATTGAACGGAATGCAATTAGG
CGGACGAAATGTTAGACTTTCGTGGGGCCGTAGTCCTAACACAGACAGTCTCAACCTGACC
AGAACCAGTGGAACAATGCTGCTTATTATGTTATCCTCAAGGATACGACTCTTATGGATAT
GTATCTGCTCCTCAAGACCCTCAACATGTACTATGGTGGCTACCCCTGGTTATGGTGGTTACGC
GATGCTTCAGCAGGCTCAGATGCCATTGCAACAACAGTGATCTACCTTATGCCAAGCAGGAG
AGGTTCGGTTGCCAGGGAGCTGTCATTGTACTTTGGAGGCTGAGCTTCTGGAGTTGGATGATTC
CTCCCAGAGATGGCAGAATGTAGTATAACTTGGTCATTGTGCTGGTCGAATTTTATTTACTG
TCTTGGGTTTTTTGCTCTGTGCTGCTTTTTTTGTAGCTTGC

SEQ ID NO. 16: secuencia de la proteína putativa de Beta vulgaris (322 aminoácidos)

MDENYIRTCFAQSGELVN VKIIRNKQTMQSECYGFIEFSTHAAAERILQTYNNTLMPNVEQN
YRLNWAFYGSGEKRGEDASDYTIFVGDLAPDVTDTYTLQETFRVRYPSVKGAKVVIDRLTSRS
KGYGFVRFGDESEQARAMSEMNGMMCLGRAMRIGAAANKKSVGGTASYQNNQGTPNDS DPSN
TTIFVGNLDSNVTDEHLRQTFSPYGELVHV KIPAGKQCGFVQFTNRSSAEELRVLNGMQLG
GRNVRLSWGRSPNNRQSQPDQNQWNNAAYYGYPQGYDSYGYVSAPQDPNMYYGYPGYGGYA
MPQQAQMPLQQQ

SEQ ID NO. 17: secuencia de nucleótidos de Beta vulgaris (1200 nucleótidos). La secuencia putativa de codificación va desde la posición 35 hasta la 922 (los codones putativos de inicio y de detención están subrayados)

AAAAACCTCTTTTTCTCTCTCCTAAATCACAACAATGGCGATACTCTCAGATTACGAGGAAG
AAGAACACCAACCACAACCAGAAAAGAAGCAACCTTCAAAGAAATTTTCAGCAACTTTTCGAT
CCTTCGAATCCGCTAGGGTTTCTTCAATCTACTCTCGAATTCGTCTCAAAGAGTCCGATTT
TTTCGCTAAGGAATCATCTGCGAAAAGATGTTGTTTCTCTGGTTCAGAAAGTGAAGGAGAAGT
ACATTGAAGAAGTAGAGAATAAGAAAGAAGCTTCTAGATGAATCTGCCGCTGCCGCCGCC
GCCGCCGCTGCTGCTGCTGCGTCGTCGTCTTCATCTGATTTGGAGAAGAAGGTTGATGATAA
TGAGAGTGCGGAAGAGACAGAGAAATCTAAGTACAAAGCTCCAAACAGTGGGAATGGTCAAG
ATCTCGAGAACTACTCATGGATACAGTCCTTGCAAGAAGTTACTGTTAATGTTCCGTGTTCCA
CCTGGAACAAAGTCTAGGTTTATCGATTGTCAGATAAAGAAGAATCATCTGAAAGTTGGCCT
CAAGGGTCAGCCTCCCATCATCGATGGTGAAGCTGTTCAAGCCTGTTAAGCCAGATGATTGTT

Figura 5 (Continuación)

TTTGGAGTTTGGAGGATCAAAAGTCAATCTCTATGCTGCTAACAAAGCATGATCAAATGGAG
 TGGTGGAGAAGTCTGGTCAAAGGTGAACCTGAAATCGACACTCAGAAGGTTGAACCTGAGAG
 CAGTAAGCTGTCTGACTTGGACCCTGAAACAAGGTCAACTGTTGAGAAGATGATGTTTGACC
 AAAGGCAAAAATCCATGGGCTTGCCCAAGTGATGATATGCAGAAGCAAGACATGCTGAAG
 AAGTTCATGTCCGAGCATCCGGAATGGACTTTTCTAACGCGAAGTTTAACTAGATATCGAT
 GTCGGTGATGGACTATGATTTTTTGGGTGGCAAATTTCTCGAAACAGGAACTGAAGAAAGCTT
 TTGTTATGTCTAATACTGAGCTTGTTTCATAGTAGTTACAGTCTCTAGGGTAGATGTCTCATG
 AAGAGGGGAACATTGCTTTTTGTTTAACTCTTATTTATATGCAAGTGATATTCGGTTTGCTA
 AGCAGTACATTCTGTCATCCTGCGCTTGATTTCGGTCTCTGTTCAATCATATATGTAATGTTA
 TAGCTGCAAAAAAAAAAAAAA

SEQ ID NO. 18: secuencia de la proteína putativa de Beta vulgaris (295 aminoácidos)

MAILSDYEEEEHQPQPEKKQPSKKFSATFDPSNPLGFLQSTLEFVSKESDFFAKESSAKDVV
 SLVQKVKEKYIEEVENKKKKLLDESAAAAAAAAAAAASSSSDLEKKVDDNESAEETESKY
 KAPNSGNGQDLENYSWIQSLQEVTVNVPVPPGTKSRFIDCQIKKNHLKVGLKGQPPIIDGEL
 FKPVKPDDCFWSLEDQKSI SMLLTKHDQMEWWSLVKGEPEIDTQKVEPESSKLSLDLPETR
 STVEKMMFDQRQKSMGLPTSDDMQKQDMLKKFMSEHPMDFSNAKFN

**SEQ ID NO. 19: secuencia de nucleótidos de mus musculus (1050 nucleótidos).
 La secuencia putativa de codificación va desde la posición 37 hasta la 969 (los
 codones putativos de inicio y de detención están subrayados)**

GAAAAGAGAGGAGAGAGAGAAAACCAAATCAACAAAATGGCGGAACATCTAGCATCGATATT
 CGGGACAGAGAAAGACAGAGTGAACCTGTCCATTCTACTTCAAGATCGGAGCTTGTAGACATG
 GAGATCGTTGCTCAAGGCTTCATACTAAGCCTAGTATTAGCCCTACTTTGTTGCTTGCTAAT
 ATGTATCAACGCCCTGATATGATTACTCCTGGTGTTGATCCTCAAGGACAGCCTCTTGATCC
 TCGCAAAATTCAACAACATTTTGAGGATTTTATGAGGATTTATTTGAGGAACTAAGCAAGT
 ATGGGGAGATTGAAAGTCTCAACATCTGTGACAATTTGGCTGACCACATGGTTGGGAATGTT
 TATGTGCAGTTTCAGAGAGGAAGAACATGCTGGCGAGGCACTACGAACTTGAGTGGAAGATT
 TTATGCCGGTTCGTTCCAATCATTTGTTGATTTTTCTCCTGTAACGGACTTCAGAGAAGCAACCT
 GCAGACAGTATGAGGAAATGTGTGCAATCGTGAGGTTACTGCAACTTTATGCATTTGAAA
 AAAATTAGCAGGGAGCTTAGGCGACAGTTGTTTGGAAAGGTACAGAAGGAGGCATAGCCGTAG
 TAGAAGTCGCAGTCTTCAAGCACATCGGGGGCATGGAGATCGTCCACATGGTGGCCGTGGTT
 ATGGTAGAAGAGATGATGATAGAAATCAGCGGTACCATGACAAGGGAAGAAGGCCCTAGAAGC
 CGTAGCCCTGGGCATAGAGGACGAAGCAGAAGCCCTCCCGGCAGGAGGGATAGGAGTCCAGT
 GAGGGAGAATAGTGAGGAGAGAAGAGCAAAGATTGCACAATGGAACAGGGAAAAGGAACAGG
 CAGACACTGGTAATAACGATGTTAATCATGATGTCACTGACAACCATGCAATGGATTTTCAG
 GACAATGGGGAGGATTACTATGACCATCCTCAGCAGTAAGTGGATGAAGTGCACAAGCAGGC
 TTTATTTACTACTTCTGGTTTGCTGTTATCAGAGTCTGCTCGTTTGCAGGATTTTTCG

**SEQ ID NO. 20: secuencia de la proteína putativa de la secuencia de mus
 musculus (310 aminoácidos)**

MAEHLASIFGTEKDRVNCPFYFKIGACRHGDRCSRLHTKPSISPTLLLANMYQRPDMITPGV
 DPQGQPLDPRKIQQHFEDFYEDLFEELSKYGEIESLNICDNLADHVMGVNVYVQFREEEHAGE
 ALRNLSGRFYAGRPIIVDFSPVTDFREACRQYEENVNCRGGYCNFMHLKKISRELRRQLFG
 RYRRRHSRSRSPQAHRGHGDRPHGGRGYGRRDDDRNQRYHDKGRPRSRSPGHRGRSRSP
 PGRDRSPVRENSEERRAKIAQWNREKEQADTGNNDVNHVDVDNHNANGFQDNGEDYDDHPQQ

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Este listado de referencias citado por el solicitante es únicamente para conveniencia del lector. No forma parte del documento europeo de la patente. Aunque se ha tenido gran cuidado en la recopilación, no se pueden excluir los errores o las omisiones y la OEP rechaza toda responsabilidad en este sentido.

Documentos de patente citados en la descripción

- 10 • ES 2110918 A [0006]
- ES 2134155 A1 [0006]
- WO 9844097 A, Vance [0099]
- WO 9836083 A, Angell and Baulcombe [0117]
- WO 9853083 A, Lowe [0117]
- WO 9915682 A, Lederer [0117]
- 15 • WO 9953050 A, Wang [0117]
- WO 9400012 A, Atkins [0117]
- WO 9503404 A, Lenee [0117]
- WO 0000619 A, Lutziger [0117]
- WO 9713865 A, Prinsen [0117]
- 20 • WO 9738116 A, Scott [0117]
- WO 9748814 A, Cheng [0135]
- WO 9854961 A, Hansen [0135]
- WO 9400977 A, Hiei [0135]
- WO 9817813 A, Hiei [0135]
- WO 9904618 A, Rikiishi [0135]
- WO 9506722 A, Saito [0135]
- US 5969213 A, Adams [0135]
- US 5736369 A, Bowen [0135]
- WO 9413822 A, Change [0135]
- US 5874265 A, Lundquist [0135]
- US 5990390 A [0135]
- US 5405765 A, Vasil and Vasil [0135]
- US 5955362 A, Walker [0135]
- WO 9318168 A, Eyal [0135]
- DE 4309203, von Holt [0135]
- US 5693512 A, Finer [0135]
- WO 0018939 A, Bidney and Scelonge [0139]
- WO 9713843 A [0167]

25 Literatura citada en la descripción que no es de patente:

- Schou. *Arch. Gen. Psychiatry*, 1997, vol. 54, 9 [0002]
- Lifton. *Science*, 1996, vol. 272, 676 [0002]
- Serrano. *Int. Rev. Cytol.*, 1996, vol. 165, 1 [0002] [0003]
- 30 • Yeo. *J. Exp. Bot.*, 1998, vol. 49, 915 [0002]
- Holmberg ; Bülow. *Trends Plant Sci.*, 1998, vol. 3, 61 [0002]
- Haro et al. *FEBS Lett.*, 1991, vol. 291, 189 [0003]
- Gaxiola et al. *Proc. natl. Acad. Sci. USA*, 1999, vol. 96, 1480 [0003]
- Apse et al. *Science*, 1999, vol. 285, 1256 [0003]
- 35 • Tarczynski et al. *Science*, 1993, vol. 259, 508 [0003]
- Kishor et al. *Plant Physiol.*, 1995, vol. 108, 1387 [0003]
- Alia et al. *Plant J.*, 1998, vol. 16, 155 [0003]
- Tezara et al. *Nature*, 1999, vol. 401, 914 [0003]
- Gläser et al. *EMBO J.*, 1993, vol. 12, 3105 [0004] [0012]
- 40 • Dichtl et al. *EMBO J.*, 1997, vol. 16, 7184 [0004]
- Murguía et al. *Science*, 1995, vol. 267, 232 [0004]
- Murguía et al. *J. Biol. Chem.*, 1996, vol. 271, 29029 [0004]
- Gil-Mascarell et al. *Plant J.*, vol. 17, 373 [0004]
- Gil-Mascarell et al. *Plant J.*, 1999, vol. 17, 373 [0004]
- 45 • López-Coronado et al. *J. Biol. Chem.*, 1999, vol. 274, 16043 [0004]
- Albert et al. *J. Mol. Biol.*, 2000, vol. 295, 927 [0005]
- Zhu J. K. Molecular aspects of osmotic stress in plants. *CRC Crit. Rev. Plant Sci.*, 1997, vol. 16, 253 - 277 [0011]
- Minet et al. *Plant J.*, 1992, vol. 2, 417 [0012]
- Dichtl et al. *EMB J.*, 1995, vol. 16, 7184 [0012]
- 50 • Murguía et al. *Science*, 1995, vol. 267, 232 [0012]
- Murguía et al. *J. Biol. Chem.*, 1996, vol. 271, 29029 [0012]
- Simpson et al. *EMBO J.*, 1995, vol. 14, 4540 - 4550 [0016]
- Kao ; Siliciano. *Mol. Cell. Biol.*, 1996, vol. 16, 960 - 967 [0020]

- **Henikoff, S. ; Henikoff, J. G.** Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA*, 1992, vol. 89, 10915 - 10919 [0025]
- **Yeast.** *Yeast*, 1995, vol. 11 (16), 1629 - 1650 [0075]
- **Politz ; Pederson.** *J. struct. Biol.*, 2000, vol. 129 (2-3), 252 - 257 [0075]
- 5 • **Golovkin ; Reddy.** *Plant Cell*, 1996, 1421 - 1435 [0075]
- **Golovkin ; Reddy.** *J. Biol Chem*, 1999, vol. 245 (51), 36428 - 36438 [0075]
- **Domon et al.** *J. Biol Bhem*, 1998, vol. 273 (51), 34603 - 34610 [0075]
- **Brown ; Shaw.** *Plant Cell*, 1998, vol. 10, 649 - 657 [0075]
- **Jarrous et al.** *J. Cell Biol.*, 1999, vol. 146 (3), 559 - 572 [0075]
- 10 • **Holmberg ; Bülow.** *Trends plant sci.*, 1998, vol. 3, 61 - 66 [0125]
- **Maniatis et al.** *Molecular Cloning.ALaboratory Manual.* Cold Spring Harbor Laboratory, 1982 [0140]
- **Garciadeblas et al.** *Mol. Gen. Genet.*, 1993, vol. 236, 363 - 368 [0155]
- **Ban ~ uelos et al.** *Microbiology*, 1998, vol. 144, 2749 - 2758 [0155]
- **Minet et al.** *Plant J.*, 1992, vol. 2 (3), 417 - 422 [0157]
- 15 • **Brunelli ; Pall.** *Yeast*, 1993, vol. 9, 1309 - 1318 [0158]
- **Gietz et al.** *Nucleic Acids Res.*, 1992, vol. 20, 1425 [0159]
- **Boeke et al.** *Mol. Gen. Genet.*, 1984, vol. 197, 354 - 346 [0159]
- **Quintero et al.** *Plant Cell*, 1996, vol. 8, 529 - 537 [0160]
- **Gil-Mascarell et al.** *Plant J.*, 1999, vol. 17 (4), 373 - 383 [0160]
- 20 • **Murgüia et al.** *J. Biol. Chem.*, 1996, vol. 271, 29029 - 29033 [0161]
- **Legrain ; Rosbash.** *cell*, 1989, vol. 75, 573 - 583 [0162]
- **Serrano et al.** *Eur. J. Biochem.*, 1973, vol. 34, 479 - 482 [0162]
- **Chan et al.** *Plant Mol. Biol.*, 1993, vol. 22, 491 - 506 [0169]
- **Hiei et al.** *Plant J.*, 1994, vol. 6, 271 - 282 [0169]
- 25 • **An G. ; Watson B. D. ; Stachel S. ; Gordon M. P. ; Nester E. W.** New cloning vehicles for transformation of higher plants. *EMBO J.*, 1985, vol. 4, 277 - 284 [0170]
- **Armstrong C. L. ; Petersen W. P. ; Buchholz W. G. ; Bowen B. A. ; Sulc S. L.** Factors affecting PEG-mediated stable transformation of maize protoplasts. *Plant Cell Reports*, 1990, vol. 9, 335 - 339 [0170]
- **Baron M.H. ; Baltimore D.** Antibodies against the chemically synthesized genome-linked protein of poliovirus react with native virus-specific proteins. *Cell*, 1982, vol. 28, 395 - 404 [0170]
- 30 • **Bartel P. L. ; Fields S.** The Yeast Two-Hybrid System. Oxford University Press [0170]
- **Bechtold N. ; Pelletier G.** In planta Agrobacterium-mediated transformation of adult Arabidopsis thaliana plants by vacuum infiltration. *Methods Mol. Biol.*, 1998, vol. 82, 259 - 266 [0170]
- **Cho R. J. ; Mindrinos M. ; Richards D. R. ; Sapolsky R. J. ; Anderson M. ; Drenkard E. ; Dewdney J. ; Reuber T. L. ; Stammers M. ; Federspiel N.** Genome-wide mapping with biallelic markers in Arabidopsis thaliana. *Nat.Genet.*, 1999, vol. 23, 203 - 207 [0170]
- 35 • **Christou P. ; McCabe D. E. ; Swain W. F.** Stable transformation of soybean callus by DNA-coated gold particles. *Plant Physiol.*, 1988, vol. 87, 671 - 674 [0170]
- **Crossway A. ; Oakes J. V. ; Irvine J. M. ; Ward B. ; Knauf V. C. ; Shewmaker C. K.** Integration of foreign DNA following microinjection of tobacco mesophyll protoplasts. *Mol. Gen. Genet.*, 1986, vol. 202, 179 - 185 [0170]
- 40 • **Dodds J. H.** Plant genetic engineering. Cambridge University Press, 1985 [0170]
- **Fromm M. ; Taylor L. P. ; Walbot V.** Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proc. Natl. Acad. Sci. U.S.A.*, 1985, vol. 82, 5824 - 5828 [0170]
- **Hanahan D.** Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol*, 1983, vol. 166, 557 - 580 [0170]
- 45 • **Hellens R. P. ; Edwards E. A. ; Leyland N. R. ; Bean S. ; Mullineaux P. M.** pGreen: a versatile and flexible binary Ti vector for Agrobacterium- mediated plant transformation. *Plant Mol. Biol.*, 2000, vol. 42, 819 - 832 [0170]
- **Herrera-Estrella L. ; De Block M. ; Messens E. H. J. P. ; Van Montagu M. ; Schell J.** Chimeric genes as dominant selectable markers in plant cells. *EMBO J.*, 1983, vol. 2, 987 - 995 [0170]
- 50 • **Herrera-Estrella L. ; Depicker A. ; Van Montagu M. ; Schell J.** Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature*, 1983, vol. 303, 209 - 213 [0170]
- **Joung J. K. ; Ramm E. I. ; Pabo C. O.** A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.*, 2000, vol. 97, 7382 - 7387 [0170]

- **Krens F. A. ; Molendijk L. ; Wullems G. J. ; Schilperoort R. A.** In vitro transformation of plant protoplasts with Ti-plasmid DNA. *Nature*, 1982, vol. 296, 72 - 74 [0170]
- **Langemeier J. L. ; Cook R. F. ; Issel C. J. ; Montelaro R. C.** Application of cycle dideoxy fingerprinting to screening heterogeneous populations of the equine infectious anemia virus. *Biotechniques*, 1994, vol. 17, 484 - 6488, 490 [0170]
- **Lerner R. A.** Tapping the immunological repertoire to produce antibodies of predetermined specificity. *Nature*, 1982, vol. 299, 593 - 596 [0170]
- **Lerner R. A. ; Green N. ; Alexander H. ; Liu F. T. ; Sutcliffe J. G. ; Shinnick T. M.** Chemically synthesized peptides predicted from the nucleotide sequence of the hepatitis B virus genome elicit antibodies reactive with the native envelope protein of Dane particles. *Proc. Natl. Acad. Sci. U.S.A.*, 1981, vol. 78, 3403 - 3407 [0170]
- **Li-Sucholeiki X. C. ; Khrapko K. ; Andre P. C. ; Marcelino L. A. ; Karger B. L. ; Thilly W. G.** Applications of constant denaturant capillary electrophoresis/high-fidelity polymerase chain reaction to human genetic analysis. *Electrophoresis*, 1999, vol. 20, 1224 - 1232 [0170]
- **Liddle J. E. ; Cryer A.** A Practical Guide to Monoclonal Antibodies. *A Practical Guide to Monoclonal Antibodies*, 1991 [0170]
- **Loffler J. ; Langui D. ; Probst A. ; Huber G.** Accumulation of a 50 kDa N-terminal fragment of beta-APP695 in Alzheimer's disease hippocampus and neocortex. *Neurochem.Int.*, vol. 24, 281 - 288 [0170]
- **Magyar Z. ; Meszaros T. ; Miskolczi P. ; Deak M. ; Feher A. ; Brown S. ; Kondorosi E. ; Athanasiadis A. ; Pongor S. ; Bilgin M.** Cell cycle phase specificity of putative cyclin-dependent kinase variants in synchronized alfalfa cells. *Plant Cell*, vol. 9, 223 - 235 [0170]
- **McCallum C. M. ; Comai L. ; Greene E. A. ; Henikoff S.** Targeted screening for induced mutations. *Nat.Biotechnol.*, vol. 18, 455 - 457 [0170]
- **McCallum C. M. ; Comai L. ; Greene E. A. ; Henikoff S.** Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol*, 2000, vol. 123, 439 - 442 [0170]
- **Merrifield R.B.** Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Amer. Chem. Soc.*, 1963, vol. 85, 2149 - 2154 [0170]
- **Murakami T. ; Simonds W. F. ; Spiegel A. M.** Site-specific antibodies directed against G protein beta and gamma subunits: effects on alpha and beta gamma subunit interaction. *Biochemistry*, 1992, vol. 31, 2905 - 2911 [0170]
- **Palmgren G.** Transgenic plants: environmentally safe factories of the future. *Trends Genet.*, 1997, vol. 13, 348 [0170]
- **Paszkowski J. ; Shillito R. D. ; Saul M. ; Mandak V. ; Hohn T. H. B. P. I.** Direct gene transfer to plants. *EMBO J.*, 1984, vol. 3, 2717 - 2722 [0170]
- **Ross P. ; Hall L. ; Haff L. A.** Quantitative approach to single-nucleotide polymorphism analysis using MALDI-TOF mass spectrometry [In Process Citation]. *Biotechniques*, 2000, vol. 29, 620 - 629 [0170]
- **Sambrook J. ; Fritsch E. F. ; Maniatis T.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0170]
- **Semler B. L. ; Anderson C. W. ; Hanecak R. ; Dorner L. F. ; Wimmer E.** A membrane-associated precursor to poliovirus VPg identified by immunoprecipitation with antibodies directed against a synthetic heptapeptide. *Cell*, 1982, vol. 28, 405 - 412 [0170]
- **Shioda T. ; Andriole S. ; Yahata T. ; Isselbacher K. J.** A green fluorescent protein-reporter mammalian two-hybrid system with extrachromosomal maintenance of a prey expression plasmid: Application to interaction screening. *Proc. Natl. Acad. Sci. U.S.A.*, 2000, vol. 97, 5220 - 5224 [0170]
- **Tamura R. N. ; Cooper H. M. ; Collo G. ; Quaranta V.** Cell type-specific integrin variants with alternative alpha chain cytoplasmic domains. *Proc. Natl. Acad. Sci. U.S.A.*, 1991, vol. 88, 10183 - 10187 [0170]
- **Trieu A. T. ; Burleigh S. H. ; Kardailsky I. V. ; Maldonado-Mendoza I. E. ; Versaw W. K. ; Blaylock L. A. ; Shin H. ; Chiou T. J. ; Katagi H. ; Dewbre G. R.** Technical Advance: Transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium*. *Plant J.*, 2000, vol. 22, 531 - 541 [0170]
- **Vidal-Puig A. ; Moller D. E.** Comparative sensitivity of alternative single-strand conformation polymorphism (SSCP) methods. *Biotechniques*, 1994, vol. 17, 490 - 2494, 496 [0170]
- **Woulfe J. ; Lafortune L. ; de Nadai F. ; Kitabgi P. ; Beaudet A.** Post-translational processing of the neurotensin/neuromedin N precursor in the central nervous system of the rat-II. Immunohistochemical localization of maturation products. *Neuroscience*, 1994, vol. 60, 167 - 181 [0170]
- **Yoon K. ; Cole-Strauss A. ; Kmiec E. B.** Targeted gene correction of episomal DNA in mammalian cells mediated by a chimeric RNA-DNA oligonucleotide. *Proc. Natl. Acad. Sci. U.S.A.*, 1996, vol. 93, 2071 - 2076 [0170]