

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 381 031**

51 Int. Cl.:
C12N 5/04 (2006.01)
C12N 15/29 (2006.01)
C12N 15/52 (2006.01)
C12N 15/09 (2006.01)
C12N 15/82 (2006.01)
C12N 15/90 (2006.01)
A01H 5/00 (2006.01)
A01H 5/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **02797860 .0**
96 Fecha de presentación: **05.09.2002**
97 Número de publicación de la solicitud: **1423508**
97 Fecha de publicación de la solicitud: **02.06.2004**

54 Título: **Polipéptidos relacionados con el estrés regulado por proteína fosfatasa y métodos de uso en las plantas**

30 Prioridad:
05.09.2001 US 317305 P

45 Fecha de publicación de la mención BOPI:
22.05.2012

45 Fecha de la publicación del folleto de la patente:
22.05.2012

73 Titular/es:
**BASF PLANT SCIENCE GMBH
CARL-BOSCH-STRASSE 38
67056 LUDWIGSHAFEN, RHEINLAND-, DE**

72 Inventor/es:
**BOHNERT, Hans, J. ;
CHEN, Ruoying ;
ISHITANI, Manabu ;
VAN THIELEN, Nocha y
DA COSTA E SILVA, Oswaldo**

74 Agente/Representante:
Carvajal y Urquijo, Isabel

ES 2 381 031 T3

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

DESCRIPCIÓN

Polipéptidos relacionados con el estrés regulado por proteína fosfatasa y métodos de uso en las plantas.

Antecedentes de la invención

Campo de la invención

- 5 Esta invención se refiere en general a las secuencias de ácidos nucleicos que codifican polipéptidos que están asociados con respuestas de estrés abiótico y con la tolerancia al estrés abiótico en plantas. En particular, la presente invención se relaciona con secuencias de ácido nucleico que codifican polipéptidos que confieren tolerancia a la sequía a las plantas.

Antecedentes de la técnica

- 10 Los estreses abióticos del medio ambiente, tales como el estrés por sequía, estrés por salinidad, estrés por calor y estrés por frío, son los principales factores limitantes del crecimiento y productividad vegetal. Las pérdidas de cosechas y las pérdidas de rendimiento de las cosechas de los principales cultivos tales como la soja, el arroz, el maíz, el algodón y el trigo causadas por estas tensiones representan un importante factor económico y político y contribuyen a la escasez de alimentos en muchos países subdesarrollados.
- 15 Las plantas suelen estar expuestas durante su ciclo de vida a condiciones de reducción del contenido de agua del medio ambiente. La mayoría de las plantas han desarrollado estrategias para protegerse contra estas condiciones de desecación. Sin embargo, si la severidad y la duración de las condiciones de sequía son demasiado grandes, los efectos sobre el desarrollo, el crecimiento y el rendimiento de la mayoría de plantas de cultivo son profundos. La exposición continua a condiciones de sequía provoca grandes alteraciones en el metabolismo de la planta, que en
- 20 última instancia conduce a muerte celular y por consiguiente a pérdidas de rendimiento.

- El desarrollo de plantas tolerantes al estrés es una estrategia que tiene el potencial para resolver o mediar por lo menos algunos de estos problemas. Sin embargo, las estrategias tradicionales de fitomejoramiento para desarrollar nuevas líneas de plantas que muestran una gran resistencia (tolerancia) a este tipo de tensiones son relativamente
- 25 lentas y requieren determinadas líneas resistentes para el cruce con la línea deseada. La escasez de recursos de germoplasma para la tolerancia al estrés y la incompatibilidad en los cruces entre especies de plantas alejadas representan problemas significativos encontrados en el fitomejoramiento convencional. Además, los procesos celulares que conducen a la tolerancia a la sequía, el frío y la salinidad en el modelo, plantas tolerantes a la sequía y/o a la salinidad son de naturaleza compleja e involucran múltiples mecanismos de adaptación celular y numerosas rutas metabólicas. Esta naturaleza de múltiples componentes de la tolerancia al estrés no sólo ha hecho muy poco
- 30 exitoso el fitomejoramiento para tolerancia, sino que también ha limitado la capacidad para modificar genéticamente las plantas para tolerar al estrés utilizando métodos biotecnológicos.

- Los estreses por sequía y por frío, así como los estreses por salinidad, tienen un factor común importante para el crecimiento de las plantas, y es la disponibilidad de agua. Las plantas están expuestas durante todo su ciclo de vida a condiciones de un contenido reducido de agua en el medio ambiente, y la mayoría de las plantas han desarrollado
- 35 estrategias para protegerse contra estas condiciones de desecación. Sin embargo, si la severidad y la duración de las condiciones de sequía son demasiado grandes, los efectos sobre el desarrollo, crecimiento y rendimiento de las plantas para la mayoría de las plantas de cultivo son profundos. Además, la mayoría de las plantas de cultivo son muy susceptibles a las concentraciones más altas de sal en el suelo. Debido a que un alto contenido de sal en algunos suelos trae como resultado una menor disponibilidad de agua para consumo por parte de las células, una
- 40 alta concentración de sal tiene un efecto sobre las plantas similar al efecto de sequía sobre las plantas. Además, bajo temperaturas de congelación, las células de las plantas pierden agua como resultado de la formación de hielo que se inicia en el apoplasto y, retira el agua del simplasto. Los mecanismos moleculares de respuesta de una planta a cada una de estas condiciones de estrés son comunes, y las proteínas fosfatasas juegan un papel esencial en estos mecanismos moleculares.

- 45 Es bien sabido que la fosforilación reversible de proteínas controla muchos procesos celulares en plantas y animales. El estado de fosforilación de las proteínas está regulado por las actividades opuestas de las proteínas quinasas y las proteínas fosfatasas. La fosforilación de las proteínas eucariotas se produce predominantemente sobre residuos de serina y de treonina, y en menor medida, sobre residuos de tirosina. En los animales, la fosforilación de proteínas desempeña un papel bien conocido en diversos procesos celulares tales como el
- 50 metabolismo del glicógeno, el control del ciclo celular, y la transducción de señales (Smith, RD y Walker, JC, 1996, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47: 101 - 125).

- Las actividades de la proteína fosfatasa han sido reportadas en la mayoría de los compartimentos subcelulares de las plantas, incluidas las mitocondrias, los cloroplastos, los núcleos y el citosol, y están asociadas con diferentes fracciones de partículas y de la membrana. Algunas proteínas fosfatasas están pobremente caracterizadas y pueden
- 55 representar nuevas enzimas que son únicas para las plantas. Otras tienen propiedades bioquímicas que son muy similares a las de las bien conocidas proteínas fosfatasas de mamíferos, tales como las proteínas citosólicas serina /

treonina fosfatasa (Mackintosh C. Cohen y P. 1989 Biochem. J. 262: 335 - 339). Se ha identificado que dos de tales serina / treonina fosfatasa de las plantas actúan en forma similar a las proteínas serina / treonina fosfatasa de tipo 1 (PP1) y de tipo 2 (PP2) de mamífero. Los estudios bioquímicos y genéticos en plantas implican actividad de PP1 y / o PP2 en la transducción de señales, la regulación hormonal, la mitosis, y el control del metabolismo del carbono y el nitrógeno (Smith, R. D. y Walker, J. C., 1996, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47: 101 - 125).

La evidencia experimental ha implicado la participación de las proteínas fosfatasa en la cascada de señalización de estrés de las plantas y, más concretamente, en la percepción del estrés y la transducción de señales relacionados con los mecanismos fisiológicos de adaptación en las plantas. Por ejemplo, se ha demostrado que la proteína fosfatasa 2C (PP2C) está involucrada en las respuestas al estrés en las plantas (Sheen, J. 1998 Proc. Natl. Acad. EE.UU. 95: 975 - 980). También se ha demostrado que, en la levadura, la calcineurina fosfatasa PP2B (CaN) es un componente central de una ruta de transducción de señal que depende de Ca^{2+} que media la tolerancia de Na^+ , Li^- , y Mn^{2+} de *Saccharomyces cerevisiae* (Cunningham, K. W. y Fink G. R. (1996) Mol. Cell. Biol. 16: 2226 - 2237). CaN actúa para limitar la acumulación intracelular de Na^+ por medio de procesos que regulan esa restricción de entrada y mejoran la salida de este catión a través de la membrana plasmática. CaN también participa en la homeostasis citosólica de Ca^{2+} a través de la regulación positiva del aparato de Golgi y el bombeo del ion tipo P localizado en la membrana vacuolar y el control negativo de un intercambiador de H^+ / Ca^{2+} vacuolar. Curiosamente, la sobreexpresión de CaN de levadura confiere tolerancia a la salinidad en las plantas, lo que indica firmemente que la modulación de las rutas de señalización del estrés mediante la expresión de una proteína fosfatasa activada mejora sustancialmente la tolerancia de las plantas al estrés (Pardo, I. M. et al. (1998), Proc. Natl. Acad. Sci. EE.UU. 95: 9681 - 9686).

A pesar de que algunos genes que están involucrados en respuestas al estrés en las plantas han sido caracterizados, la caracterización y clonación de genes de las plantas que confieren tolerancia al estrés sigue siendo muy incompleta y fragmentada. Por ejemplo, algunos estudios han indicado que el estrés por sequía y salinidad en algunas plantas puede ser debido a efectos aditivos de genes, en contraste con otra investigación que indica que genes específicos son activados transcripcionalmente en el tejido vegetativo de las plantas bajo condiciones de estrés osmótico. Aunque en general se supone que las proteínas inducidas por estrés juegan un papel en la tolerancia, se carece aún de evidencia directa, y las funciones de muchos genes de respuesta al estrés son desconocidas.

Es necesario, por lo tanto, identificar los genes expresados en plantas tolerantes al estrés que tienen la capacidad de conferir resistencia al estrés a su planta huésped y a otras especies de plantas. Las plantas tolerantes al estrés recientemente generadas tendrán muchas ventajas, tales como el aumento en el rango en el cual las plantas de cultivo pueden ser cultivadas, por ejemplo, disminuyendo las necesidades de agua de una especie vegetal.

Resumen de la invención

Esta invención cumple en parte la necesidad de identificar nuevas proteínas fosfatasa únicas capaces de conferir tolerancia al estrés a las plantas por sobreexpresión. La presente invención describe un ácido nucleico que codifica al Polipéptido Relacionado con el Estrés de la Proteína Fosfatasa (PPSRP por sus siglas en inglés) y que es importante para la modulación de una respuesta de la planta a un estrés ambiental, que comprende al ácido nucleico que codifica al PPSRP, en donde el ácido nucleico que codifica al PPSRP comprende un polinucleótido seleccionado del grupo que consiste de a) el polinucleótido como se define en la SEQ ID NO: 2, b) un polinucleótido que codifica para un PPSRP, en donde el PPSRP es una Proteína Fosfatasa 2A-1 (PP2A-1) como se define en la SEQ ID NO: 3, c) un polinucleótido complementario a un polinucleótido de cualquiera de las secuencias de ácido nucleico de a) o b), y c) al menos 80 - 85% idénticas a la secuencia de la SEQ ID NO: 2 sobre la longitud completa de la región de codificación de la secuencia de la SEQ ID NO: 2. Más particularmente, la sobreexpresión de este ácido nucleico que codifica PPSRP en una planta resulta en la mayor tolerancia de la planta al estrés por sequía.

Por lo tanto, la presente invención incluye una célula vegetal aislada que comprende al ácido nucleico que codifica PPSRP, en donde la expresión de la secuencia de ácido nucleico en la células de la planta resulta en una mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la célula de la planta. Es decir, en este documento se describe la Proteína Fosfatasa 2A-1 (PP2A-1), un PPSRP seleccionado de entre el grupo que consiste de una secuencia como la definida en la SEQ ID NO: 3 y secuencias que son al menos 90 - 95% idénticas a la secuencia completa de aminoácidos mostrada en la SEQ ID NO: 3.

Los PPSRP y los ácidos nucleicos que los codifican pueden ser encontrados en miembros de los generos *Physcomitrella*, *Brassica*, *Glycine*, u *Oryza*. La invención establece que el estrés ambiental puede ser sequía.

La invención proporciona además una semilla producida por una planta transgénica transformada por el ácido nucleico que codifica PPSRP de la invención, en donde la planta es una línea genéticamente pura para mayor tolerancia al estrés por sequía comparada con una variedad de sipo silvestre de la planta. La invención provee además una semilla producida por una planta transgénica que expresa el PPSRP, en donde la planta es una línea genéticamente pura para mayor tolerancia al estrés por sequía comparada con una variedad de sipo silvestre de la planta.

También se divulga el uso de dicha planta o semilla de la invención para la producción de un producto agrícola. La invención proporciona además al PPSRP aislado como se describe a continuación. La invención proporciona además al ácido nucleico aislado que codifica PPSRP, en donde el ácido nucleico que codifica PPSRP codifica para el PPSRP como se describe a continuación.

5 La invención proporciona además un vector aislado de expresión recombinante que comprende un ácido nucleico que codifica PPSRP como se describe a continuación, en donde la expresión del vector en una célula huésped resulta en una mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la célula huésped. La invención proporciona además una célula huésped que contiene al vector y una planta que contiene la célula huésped.

10 La invención proporciona además un método para producir una planta transgénica con el ácido nucleico que codifica PPSRP, en donde la expresión del ácido nucleico en la plantas resulta en una mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la planta, que comprende: (a) transformar una célula vegetal con un vector de expresión que comprende al ácido nucleico que codifica PPSRP, y (b) la generación a partir de la célula de la planta de una planta transgénica con una mayor tolerancia al estrés por sequía en comparación con una
15 variedad de tipo silvestre de la planta. En modalidades preferidas, el PPSRP y el ácido nucleico que codifica PPSRP son como se describe más adelante.

También se describe aquí, un método para la identificación de un nuevo PPSRP, que comprende (a) elevar la respuesta de un anticuerpo específico para un PPSRP, o un fragmento del mismo, como se describe más adelante; (b) seleccionar el material PPSRP putativo con el anticuerpo, en donde el enlazamiento específico del anticuerpo
20 con el material indica la presencia de un PPSRP potencialmente novedoso; y (c) la identificación del material enlazado a un nuevo PPSRP en comparación con PPSRP conocido. Alternativamente, la hibridación con sondas de ácido nucleico como se describe más adelante puede utilizarse para identificar nuevos ácidos nucleicos para PPSRP.

25 La presente invención también proporciona métodos para aumentar la tolerancia a la sequía de una planta que comprende, aumentar la expresión del ácido nucleico para PPSRP en la planta, en donde el PPSRP es como se describe más adelante.

Breve descripción de los dibujos

30 La Figura 1 muestra un diagrama del vector de expresión de la planta pBPSJHOOI que contiene al súper promotor que dirige la expresión de un ácido nucleico que codifica PPSRP ("Gen de interés"). Los componentes son: el gen aacCI de resistencia a la gentamicina (Hajdukiewicz et al., Plant Molec. Biol. 25: 989 - 94, 1994), el promotor NOS (Becker et al (1992), Plant Molec. Biol. 20: 1195 - 1197), el terminador g7T (Becker et al., 1992), el terminador NOSpA (Jefferson et al. (1987), EMBO J. 6:3901 - 7).

35 La Figura 2 muestra una alineación de la secuencia de aminoácidos deducida de PP2A-1 (SEQ ID NO: 3) de *Physcomitrella patens* con las secuencias de aminoácidos deducidas de OsPP2A-1 (SEQ ID NO: 24), OsPP2A-2 (SEQ ID NO: 26), OsPP2A-3 (SEQ ID NO: 28), OsPP2A-4 (SEQ ID NO: 30), y OsPP2A-5 (SEQ ID NO: 32). de *Oryza sativa*.

40 La Figura 3 muestra una alineación de la secuencia de aminoácidos deducida de PP2A-1 (SEQ ID NO: 3) de *Physcomitrella patens* con las secuencias de aminoácidos deducidas de GmPP2A-1 (SEQ ID NO: 14), GmPP2A-2 (SEQ ID NO: 16), GmPP2A-3 (SEQ ID NO: 18), GmPP2A-4 (SEQ ID NO: 20), y GmPP2A-5 (SEQ ID NO: 22) de *Glycine max*, y BnPP2A-1 (SEQ ID NO: 8), BnPP2A-2 (SEQ ID NO: 10), y BnPP2A-3 (SEQ ID NO: 12). de *Brassica napus*.

45 La Figura 4 muestra una alineación de la secuencia de aminoácidos deducida de PP2A-1 (SEQ ID NO: 3) de *Physcomitrella patens* con secuencias de aminoácidos deducidas de OsPP2A-1 (SEQ ID NO 24), OsPP2A-2 (SEQ ID NO: 26), OsPP2A-3 (SEQ ID NO: 28), OsPP2A-4 (SEQ ID NO: 30), y OsPP2A-5 (SEQ ID NO: 32) de *Oryza sativa*. La alineación también incluye GmPP2A-1 (SEQ ID NO: 14), GmPP2A-2 (SEQ ID NO: 16), GmPP2A-3 (SEQ ID NO: 18), GmPP2A-4 (SEQ ID NO: 20), y GmPP2A-5 (SEQ ID NO: 22) de *Glycine max*, y BnPP2A-1 (SEQ ID NO: 8), BnPP2A-2 (SEQ ID NO: 10), y BnPP2A-3 (SEQ ID NO: 12) de *Brassica napus*.

Descripción detallada de la invención

50 La presente invención se puede entender más fácilmente por referencia a la siguiente descripción detallada de las modalidades preferidas de la invención y los Ejemplos incluidos en este documento. Sin embargo, antes de que los presentes compuestos, composiciones y métodos sean divulgados y descritos, se debe entender que esta invención no se limita a ácidos nucleicos específicos, polipéptidos específicos, tipos específicos de células, células huésped específicas, condiciones específicas, o métodos específicos, etc., como tales, pueden, por supuesto, variar, y las numerosas modificaciones y variaciones de ellos serán evidentes para aquellos capacitados en la técnica. También
55 se debe entender que la terminología usada aquí tiene el propósito de describir las modalidades específicas únicamente y no se pretende que sean limitantes. En particular, la designación de las secuencias de aminoácidos

como polipéptido "Polipéptidos Relacionados con el Estrés de la Proteína Fosfatasa" (PPSRP), no limita de ninguna manera la funcionalidad de esas secuencias.

La presente invención describe al ácido nucleico que codifica al Polipéptido Relacionado con el Estrés de la Proteína Fosfatasa (PPSRP), en donde el ácido nucleico que codifica PPSRP comprende un polinucleótido seleccionado de entre del grupo que consiste de a) el polinucleótido como el definido en la SEQ ID NO: 2, b) un polinucleótido que codifica para un PPSRP, en donde el PPSRP es proteína fosfatasa 2A-1 (PP2A-1) como se define en la SEQ ID NO: 3, c) un polinucleótido complementario a un polinucleótido de cualquiera de las secuencias de ácido nucleico de a) o b), y d) al menos 80 - 85% idénticas a la secuencia de la SEQ ID NO: 2 sobre la longitud completa de la región de codificación de la secuencia de la SEQ ID NO: 2. Más particularmente, la sobreexpresión de este ácido nucleico que codifica PPSRP en una planta resulta en la mayor tolerancia de la planta al estrés por sequía. Miembros representativos del género PPSRP incluyen, pero no se limitan a, PP2A-1, PP2A-2, PP2A-3, PP2A-4, PP-1, BnPP2A-1, BnPP2A-2, BnPP2A-3, GmPP2A-1, GmPP2A-2, GmPP2A-3, GmPP2A-4, GmPP2A-5, OsPP2A-1, OsPP2A-2, OsPP2A-3, OsPP2A-4, y OsPP2A-5. En una modalidad preferida, PP2A-1 es una proteína fosfatasa biológicamente activa. Los polipéptidos PP2A-2, PP2A-3, y PP2A 4 se describen en la solicitud de patente de los Estados Unidos N° 09/828.302, mientras que PP2A-1, PP-1, BnPP2A-1, BnPP2A-2, BnPP2A-3, GmPP2A-1, GmPP2A-2, GmPP2A-3, GmPP2A-4, GmPP2A-5, OsPP2A-1, OsPP2A-2, OsPP2A-3, OsPP2A-4, y OsPP2A-5 se describen en este documento.

La presente invención proporciona una célula de una planta transgénica transformada por el ácido nucleico que codifica PPSRP de la invención, en donde la expresión de la secuencia de ácido nucleico en la célula de la planta resulta en una mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la célula de la planta. La invención proporciona además partes de plantas transgénicas y plantas transgénicas que contienen las células de las plantas descritas en este documento. Las partes de la planta incluyen, pero no se limitan a, tallos, raíces, óvulos, estambres, hojas, embriones, regiones meristemáticas, tejido caloso, gametófitos, esporofitos, polen, microesporas y similares. También se proporciona una semilla de planta producida por una planta transgénica transformada por el ácido nucleico que codifica PPSRP de la invención, en donde la semilla contiene al ácido nucleico que codifica PPSRP de la invención, y en donde la planta es una línea genéticamente pura para mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la planta. La invención proporciona además una semilla producida por una planta transgénica que expresa al PPSRP de la invención, en donde la semilla contiene al PPSRP de la invención, y en donde la planta es una línea genéticamente pura para mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la planta. También se divulga el uso de dicha planta o semilla de la invención para la producción de un producto agrícola. Los productos agrícolas incluyen, pero no se limitan a, extractos de la planta, proteínas, aminoácidos, carbohidratos, grasas, aceites, polímeros, vitaminas y similares.

Tal como se usa aquí, el término "variedad" se refiere a un grupo de plantas dentro de una especie que comparten caracteres constantes que las separan de la forma típica y de otras variedades posibles dentro de esa especie. Aunque poseen al menos un rasgo distintivo, una variedad se caracteriza también por una cierta variación entre individuos dentro de la variedad, con base principalmente en la segregación Mendeliana de rasgos entre la progenie de generaciones sucesivas. Una variedad es considerada "una línea genéticamente pura" para un rasgo particular si es genéticamente homocigota para ese rasgo en la medida en que, cuando la variedad de la línea genéticamente pura es autógama, no se observa una cantidad significativa de segregación independiente del rasgo entre la progenie. En la presente invención, el rasgo surge de la expresión transgénica de una o más secuencias de ADN introducidas en una variedad de planta.

La presente invención describe por primera vez que la PPSRP de *Physcomitrella patens*, PP2A-1 es útil para incrementar la tolerancia de una planta al estrés por sequía. Tal como se usa aquí, el término polipéptido se refiere a una cadena de al menos cuatro aminoácidos unidos por enlaces peptídicos. La cadena puede ser lineal, ramificada circular, o combinaciones de los mismos. En consecuencia, la presente invención proporciona al PPSRP aislado PP2A-1 como se define en la SEQ ID NO: 3 y secuencias que son al menos 90 - 95% idénticas a la secuencia completa de aminoácidos mostrada en la SEQ ID NO: 3. Otras fosfatasas son: 1) el polipéptido de la proteína fosfatasa 1 de *Physcomitrella patens* (PP-1) como se define en SEQ ID NO: 6; 2) el polipéptido de la proteína fosfatasa PP2A-1 (BnPP2A-1) de *Brassica napus* como se define en la SEQ ID NO : 8; 3) el polipéptido de BnPP2A-2 como se define en SEQ ID NO: 10; 4) el polipéptido de BnPP2A-3 como se define en la SEQ ID NO: 12; 5) el polipéptido de la proteína fosfatasa PP2A-1 de *Glycine max* (GmPP2A-1) como se define en SEQ ID NO: 14; 6) el polipéptido de GmPP2A-2 como se define en SEQ ID NO: 16; 7) el polipéptido de GmPP2A-3 como se define en la SEQ ID NO: 18; 8) el polipéptido de GmPP2A-4 como se define en la SEQ ID NO: 20; 9) el polipéptido de GmPP2A-5 como se define en la SEQ ID NO: 22; 10) el polipéptido de la proteína fosfatasa 2A-1 de *Oryza sativa* (OsPP2A-1) como se define en SEQ ID NO: 24; 11) el polipéptido de OsPP2A-2 como se define en la SEQ ID NO: 26; 12) el polipéptido de OsPP2A-3 como se define en la SEQ ID NO: 28; 13) el polipéptido de OsPP2A-4 como se define en la SEQ ID NO: 30; 14) el polipéptido de OsPP2A-5 como se define en la SEQ ID NO: 32; y homólogos y ortólogos de los mismos. Los homólogos y ortólogos de las secuencias de aminoácidos se definen más adelante.

El PPSRP de la presente invención es producido preferiblemente por medio de técnicas de ADN recombinante. Por ejemplo, una molécula de ácido nucleico que codifica al polipéptido se clona en un vector de expresión (como se

describe más adelante), se introduce el vector de expresión en una célula huésped (como se describe más adelante) y el PPSRP se expresa en la célula huésped. El PPSRP puede ser aislado entonces de las células mediante un esquema de purificación adecuado utilizando técnicas estándar de purificación de polipéptidos. Para los fines de la invención, el término "polinucleótido recombinante" se refiere a un polinucleótido que ha sido alterado, reorganizado o modificado por ingeniería genética. Los ejemplos incluyen a cualquier polinucleótido clonado, y a polinucleótidos que están enlazados o unidos a secuencias heterólogas. El término "recombinante" no se refiere a las alteraciones con polinucleótidos que resultan de eventos que ocurren naturalmente, tales como mutaciones espontáneas. Como alternativa a la expresión recombinante, el PPSRP de la invención, o el péptido pueden ser sintetizados químicamente usando técnicas estándar de síntesis de péptidos. Además, el PPSRP nativo de la invención puede ser aislado de las células, por ejemplo, de *Physcomitrella patens*, por ejemplo utilizando un anticuerpo anti-PPSRP, que pueden ser producido por medio de técnicas estándar que utilizan el PPSRP de la invención o un fragmento del mismo.

La invención proporciona además el ácido nucleico aislado que codifica PPSRP de la invención. La presente invención incluye al ácido nucleico para codificación de PPSRP que codifica al PPSRP como se describe aquí. El ácido nucleico que codifica PPSRP se selecciona de entre el grupo que consiste de a) el ácido nucleico para la proteína fosfatasa 2A-1 (PP2A-1) como se define en SEQ ID NO: 2, b) un polinucleótido que codifica para una Proteína Relacionada con Estrés de Proteína Fosfatasa (PPSRP), en donde el PPSRP es proteína fosfatasa 2A-1 (PP2A-1) como define en la SEQ ID NO: 3, c) un polinucleótido complementario a un polinucleótido de cualquiera de a) o b) anteriores; y d) secuencias de ácido nucleico al menos 80 - 85% idénticas a la secuencia de la SEQ ID NO: 2 sobre la longitud completa de la región de codificación de la secuencia de la SEQ ID NO: 2. Se describen adicionalmente 1) el ácido nucleico para la proteína fosfatasa 1 (PP-1) de *Physcomitrella patens* como se define en la SEQ ID NO: 5; 2) el ácido nucleico para proteína fosfatasa PP2A-1 (BnPP2A-1) de *Brassica napus* como se define en la SEQ ID NO: 9; 4) ácido nucleico BnPP2A-3 como se define en la SEQ ID NO: 11; 5) el ácido nucleico para la proteína fosfatasa PP2A-1 de *Glycine max* (GmPP2A-1) como se define en la SEQ ID NO: 13; 6) el ácido nucleico GmPP2A-2 como se define en la SEQ ID NO: 15; 7) el ácido nucleico GmPP2A-3 como se define en la SEQ ID NO: 17; 8) el ácido nucleico GmPP2A-4 como se define en la SEQ ID NO: 19; 9) el ácido nucleico GmPP2A-5 como se define en la SEQ ID NO: 21; 10) el ácido nucleico para la proteína fosfatasa 2A-1 de *Oryza sativa* (OsPP2A-1) como se define en la SEQ ID NO: 23; 11) el ácido nucleico OsPP2A-2 como se define en la SEQ ID NO: 25; 12) el ácido nucleico OsPP2A-3 como se define en la SEQ ID NO: 27; 13) el ácido nucleico OsPP2A-4 como se define en la SEQ ID NO: 29; 14) el ácido nucleico OsPP2A-5 como se define en la SEQ ID NO: 31; y homólogos y ortólogos de los mismos. Los homólogos y ortólogos de las secuencias de nucleótidos como se define más adelante. De acuerdo con la invención, el ácido nucleico y polipéptido son de una planta de *Physcomitrella patens* (P. patens).

Tal como se usa aquí, el término "estrés ambiental" se refiere a cualquier condición de crecimiento por debajo del óptimo e incluye, pero no se limita a, condiciones por debajo del óptimo asociadas con estreses por salinidad, sequía, temperatura, metales, compuestos químicos, patógenos y oxidativos, o combinaciones de los mismos. Preferiblemente, el estrés ambiental puede ser sequía. Debe entenderse también que tal como se utiliza en la especificación y en las reivindicaciones, "un" o "unos" puede significar uno o más, dependiendo del contexto en el que se utiliza. Así, por ejemplo, la referencia a "una célula" puede significar que al menos una célula puede ser utilizada.

Como también se usa aquí, el término "ácido nucleico" y el término "polinucleótido" se refieren a ARN o ADN que es lineal o ramificado, monocatenario o bicatenario, o un híbrido de los mismos. El término también abarca híbridos de ARN / ADN. Estos términos incluyen también la secuencia no traducida localizada tanto en el extremo 3' como en el 5' de la región de codificación del gen: al menos aproximadamente 1000 nucleótidos de la secuencia hacia arriba del extremo 5' de la región de codificación y al menos aproximadamente 200 nucleótidos de la secuencia hacia abajo del extremo 3' de la región de codificación del gen. Las bases menos comunes, tales como inosina, 5-metilcitosina, 6-metiladenina, hipoxantina y otras también pueden utilizarse para apareamiento antisentido, de ARNbc y de ribozima. Por ejemplo, los polinucleótidos que contienen análogos de propino C-5 de uridina y citidina han demostrado que enlazan ARN con una alta afinidad y que son inhibidores antisentido potentes de la expresión génica. También puede hacerse otras modificaciones, tales como la modificación de la columna vertebral fosfodiéster, o la 2'-hidroxi en el grupo del azúcar ribosa del ARN. Los polinucleótidos antisentido y los ribozimas pueden consistir enteramente de ribonucleótidos, o puede contener ribonucleótidos y desoxirribonucleótidos mezclados. Los polinucleótidos de la invención pueden ser producidos por cualquier medio, incluidas preparaciones genómicas, preparaciones de ADNc, síntesis *in vitro*, RT-PCR y transcripción *in vitro* o *in vivo*.

Una molécula "aislada" de ácido nucleico es una que está sustancialmente separada de otras moléculas de ácido nucleico que están presentes en la fuente natural del ácido nucleico (es decir, secuencias que codifican otros polipéptidos). Preferiblemente, un ácido nucleico "aislado" está libre de algunas de las secuencias que flanquean naturalmente al ácido nucleico (es decir, las secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en su replicón de origen natural. Por ejemplo, un ácido nucleico clonado se considera aislado. La molécula aislada de ácido nucleico para PPSRP puede contener aproximadamente menos de 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb o 0,1 kb de secuencias de nucleótidos que flanquean naturalmente la molécula de ácido nucleico en ADN genómico de la célula a partir de la cual se deriva el ácido nucleico (por ejemplo, una célula de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u *Oryza sativa*). Un ácido nucleico también se considera aislado si ha sido alterado por medio

de intervención humana, o colocado en un sitio o ubicación que no es su sitio natural, o si se introduce en una célula por medio de agroinfección. Además, una molécula "aislada" de ácido nucleico, tal como una molécula de ADNc, puede estar libre de una parte de otro material celular con el que está asociado de forma natural, o medio de cultivo cuando es producida por medio de técnicas recombinantes, o precursores químicos u otros productos químicos cuando es sintetizada químicamente.

Se excluyen específicamente de la definición de "ácidos nucleicos aislados": los cromosomas de origen natural (por ejemplo, propagaciones de cromosomas), bibliotecas de cromosomas artificiales, bibliotecas genómicas y bibliotecas de ADNc que existen, ya sea como en preparaciones *in vitro* de ácido nucleico o como preparaciones de células huésped transfectadas / transformadas, en el que las células huésped son ya sea una preparación heterogénea *in vitro* o sembradas en placa como una población heterogénea de colonias individuales. También se excluyen específicamente las bibliotecas anteriores en las que un ácido nucleico específico representa menos del 5% del número de insertos de ácido nucleico en las moléculas del vector. Se excluyen específicamente además ADN genómico de células enteras o preparaciones de ARN de células enteras (incluidas las preparaciones de células enteras que son cortadas mecánicamente o digeridas enzimáticamente). Aún más específicamente se excluyen las preparaciones de células enteras que se encuentran ya sea como una preparación *in vitro* o como una mezcla heterogénea separada por electroforesis en donde el ácido nucleico de la invención no ha sido separado adicionalmente de los ácidos nucleicos heterólogos en el medio de electroforesis (por ejemplo, separar adicionalmente por escisión de una sola banda a partir de una población heterogénea de bandas en un gel de agarosa o transferencia a una membrana de nylon).

La molécula de ácido nucleico de la presente invención, por ejemplo, una molécula de ácido nucleico que tiene una secuencia de nucleótidos de la SEQ ID NO: 2, o una porción de la misma, puede ser aislada utilizando técnicas estándar de biología molecular y la información de la secuencia suministrada aquí. Por ejemplo, un ADNc para PPSRP de *P. patens* puede ser aislado a partir de una biblioteca de *P. patens* utilizando la totalidad o una porción de las secuencias de la SEQ ID NO: 1. Además, una molécula de ácido nucleico que comprende la totalidad o una porción de la secuencia SEQ ID NO: 2, se puede aislar por medio de la reacción en cadena de la polimerasa utilizando iniciadores oligonucleótidos diseñados con base en esta secuencia. Por ejemplo, el ARNm puede ser aislado de las células de las plantas (por ejemplo, por medio del procedimiento de extracción de guanidinio-tiocianato de Chirgwin et al., 1979 Biochemistry 18: 5294 - 5299) y se puede preparar ADNc utilizando transcriptasa inversa (por ejemplo, transcriptasa inversa MLV de Moloney, de Gibco / BRL, Bethesda, MD; o la transcriptasa inversa AMV, de Seikagaku America, Inc., St. Petersburg, FL). Los iniciadores oligonucleótidos sintéticos para la amplificación por medio de la reacción en cadena de la polimerasa pueden ser diseñados con base en las secuencias de nucleótidos mostradas en la SEQ ID NO: 2. Una molécula de ácido nucleico de la invención puede ser amplificada utilizando ADNc o, alternativamente ADN genómico, como una plantilla e iniciadores oligonucleótidos apropiados de acuerdo con técnicas estándar de amplificación por PCR. La molécula de ácido nucleico así amplificada puede ser clonada en un vector apropiado y caracterizada por medio de análisis de secuencia de ADN. Además, los oligonucleótidos correspondientes a una secuencia de nucleótidos para PPSRP pueden ser preparados por medio de técnicas estándar de síntesis, por ejemplo, utilizando un sintetizador automatizado de ADN.

Una molécula aislada de ácido nucleico de acuerdo con la invención comprende la secuencia de nucleótidos mostrada en la SEQ ID NO: 2. Este ADNc comprende una secuencia que codifican el PPSRP (es decir, la "región de codificación"), así como secuencias 5' no traducidas y secuencias 3' no traducidas. Se debe entender que la SEQ ID NO: 2 puede contener tanto regiones de codificación como regiones 5' y 3' no traducidas. Alternativamente, las moléculas de ácido nucleico de la presente invención pueden contener únicamente la región de codificación de las secuencias en la SEQ ID NO: 2, o puede contener fragmentos genómicos completos aislados a partir de ADN genómico. La presente invención también incluye al ácido nucleico codificador de PPSRP que codifica al PPSRP como se describe aquí. Se prefiere un ácido nucleico codificador de PPSRP que codifica al PPSRP PP2A-1 (SEQ ID NO: 3).

Además, la molécula de ácido nucleico de la divulgación puede comprender únicamente una porción de la región de codificación de la secuencia en la SEQ ID NO: 2, por ejemplo, un fragmento que puede ser utilizado como una sonda o iniciador o un fragmento que codifica una porción biológicamente activa de un PPSRP. Las secuencias de nucleótidos determinadas a partir de la clonación de los genes para PPSRP de *P. patens*, *B. napus*, *G. max*, y *O. sativa* permitir la generación de sondas y de iniciadores diseñados para su uso en la identificación y / o clonación de homólogos de PPSRP en otros tipos de células y organismos, así como homólogos de PPSRP de otros musgos y especies relacionadas. La porción de la región de codificación también puede codificar un fragmento biológicamente activo de un PPSRP.

Tal como se usa aquí, el término "porción biológicamente activa de" un PPSRP está destinado a incluir una porción, por ejemplo, un dominio / motivo, de un PPSRP que participa en la modulación de la tolerancia al estrés en una planta, y más preferiblemente, tolerancia a la sequía o tolerancia a la sal. Para los propósitos de la presente invención, la modulación de la tolerancia al estrés se refiere a por lo menos a un aumento del 10% en la tolerancia al estrés por sequía de una planta transgénica que contiene un casete de expresión del PPSRP (o un vector de expresión), en comparación con la tolerancia al estrés por sequía de una planta de control no transgénica. Los métodos para cuantificación de la tolerancia al estrés se proporcionan por lo menos en el Ejemplo 7 más adelante.

En una modalidad preferida, la porción biológicamente activa de un PPSRP aumenta la tolerancia de una planta al estrés por sequía.

Las porciones biológicamente activas de un PPSRP que incluye péptidos que contienen secuencias de aminoácidos derivadas de la secuencia de aminoácidos de un PPSRP (por ejemplo, una secuencia de aminoácidos de la SEQ ID NO: 3, SEC ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30 y SEQ ID NO: 32) o la secuencia de aminoácidos de un polipéptido idéntico a un PPSRP, que incluye menos aminoácidos que un PPSRP de longitud completa o el polipéptido de longitud completa que es idéntico a un PPSRP, y exhibe al menos una actividad de un PPSRP. Típicamente, las porciones biológicamente activas (por ejemplo, los péptidos que son, por ejemplo, de 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 o más aminoácidos de longitud) contienen un dominio o motivo con al menos una actividad de un PPSRP. Además, se pueden preparar otras porciones biológicamente activas en las cuales se suprimen otras regiones del polipéptido, por medio de técnicas recombinantes y se evalúan para una o más de las actividades descritas en este documento. Preferiblemente, las porciones biológicamente activas de un PPSRP incluyen uno o más dominios / motivos seleccionados o porciones de los mismos que tienen actividad biológica.

También se describen polipéptidos quiméricos o de fusión de PPSRP. Como se usa aquí, un "polipéptido quimérico" o "polipéptido de fusión" de PPSRP comprende un PPSRP operativamente enlazado a uno que no es PPSRP. Un PPSRP se refiere a un polipéptido que tiene una secuencia de aminoácidos correspondiente a un PPSRP, mientras que uno que no es PPSRP se refiere a un polipéptido que tiene una secuencia de aminoácidos correspondiente a un polipéptido que no es sustancialmente idéntico al PPSRP, por ejemplo, un polipéptido que es diferente del PPSRP y se deriva del mismo o de un organismo diferente. Dentro del polipéptido de fusión, el término "operativamente enlazado" pretende indicar que el PPSRP y el que no es PPSRP se fusionan entre sí de manera que ambas secuencias cumplen la función propuesta atribuida a la secuencia utilizada. El que no es PPSRP se puede fusionar con el terminal N o el terminal C del PPSRP. Por ejemplo, el polipéptido de fusión es un polipéptido de fusión GST-PPSRP en el cual las secuencias de PPSRP se fusionan con el terminal C de las secuencias de GST. Tales polipéptidos de fusión pueden facilitar la purificación de los PPSRP recombinantes. Alternativamente, el polipéptido de fusión es un PPSRP que contiene una secuencia señal heteróloga en su terminal N. En ciertas células huésped (por ejemplo, células huésped de mamífero), se puede incrementar la expresión y / o la secreción de un PPSRP mediante el uso de una secuencia señal heteróloga.

Un polipéptido quimérico o de fusión de PPSRP puede ser producido por medio de técnicas estándar de ADN recombinante. Por ejemplo, los fragmentos de ADN que codifican para las diferentes secuencias de polipéptidos se ligan entre sí dentro del marco de acuerdo con técnicas convencionales, por ejemplo mediante el empleo de terminales romos o escalonados para la ligación, digestión con enzimas de restricción para proporcionar terminales apropiados, rellenando de extremos cohesivos según sea conveniente, tratamiento con fosfatasa alcalina para evitar uniones indeseables y ligación enzimática. El gen de fusión puede ser sintetizado por medio de técnicas convencionales que incluyen sintetizadores automatizados de ADN. Alternativamente, la amplificación por PCR de fragmentos de genes puede llevarse a cabo utilizando iniciadores de anclaje que dan lugar a salientes complementarios entre dos fragmentos de genes consecutivos que posteriormente pueden ser apareados y amplificados nuevamente para generar una secuencia del gen quimérico (véase, por ejemplo, Current Protocols in Molecular Biology, Eds Ausubel et al. John Wiley & Sons: 1992). Además, muchos vectores de expresión se encuentran disponibles comercialmente que ya codifican una unidad estructural de fusión (por ejemplo, un polipéptido GST). Un ácido nucleico que codifica PPSRP puede ser clonado dentro de tal vector de expresión de tal manera que la unidad estructural de fusión de fusión se enlace en el marco con el PPSRP.

Además de los fragmentos y de los polipéptidos de fusión de los PPSRP aquí descritos, la presente divulgación incluye homólogos y análogos de los PPSRP de origen natural y los ácidos nucleicos que codifican PPSRP en una planta. "Homólogos" se definen aquí como dos ácidos nucleicos o polipéptidos que tienen secuencias similares, o "idénticas", de nucleótidos o aminoácidos, respectivamente. Los homólogos incluyen variantes alélicas, ortólogos, parálogos, agonistas y antagonistas de los PPSRP como se definen aquí a continuación. El término "homólogo" abarca además moléculas de ácido nucleico que difieren de una de las secuencias de nucleótidos mostrada en la SEQ ID NO: 2, SEC ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29 y SEQ ID NO: 31 (y porciones de los mismos) debido a la degeneración del código genético y, por tanto codifican al mismo PPSRP que aquel codificado por las secuencias de nucleótidos mostradas en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29 y SEQ ID NO: 31. Como se usa aquí, un PPSRP de "origen natural" se refiere a una secuencia de aminoácidos de PPSRP que se presenta en la naturaleza. Preferiblemente, un PPSRP de origen natural comprende una secuencia de aminoácidos seleccionada de entre el grupo que consiste de la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30 y SEQ ID NO: 32.

Un agonista del PPSRP puede retener sustancialmente la misma, o un subconjunto, de las actividades biológicas del PPSRP. Un antagonista del PPSRP puede inhibir una o más de las actividades de la forma natural del PPSRP. Por ejemplo, el antagonista de PPSRP puede unirse competitivamente con un miembro secuencia abajo o secuencia arriba de la cascada metabólica de componentes de la membrana celular que incluye al PPSRP, o se une con un PPSRP que media el transporte de los compuestos a través de dichas membranas, lo que impide que se lleve a cabo la translocación.

Las moléculas de ácido nucleico que corresponden a variantes alélicas naturales y análogos, ortólogos y parálogos de un ADNc para PPSRP se puede aislar con base en su identidad de ácidos nucleicos para PPSRP de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u *Oryza sativa* descritos aquí utilizando los ADNc para PPSRP, o una porción de los mismos, como una sonda de hibridación de acuerdo con técnicas estándar de hibridación bajo condiciones estrictas de hibridación. Alternativamente, se pueden identificar homólogos del PPSRP por medio de la selección de bibliotecas combinatorias de mutantes, por ejemplo, mutantes de truncamiento del PPSRP de la actividad agonista o antagonista de PPSRP. Se puede generar una biblioteca abigarrada de variantes de PPSRP por medio de mutagénesis combinatoria a nivel del ácido nucleico y es codificada por una biblioteca abigarrada de genes. Se puede producir una biblioteca abigarrada de variantes de PPSRP, por ejemplo, ligando enzimáticamente una mezcla de oligonucleótidos sintéticos en las secuencias de genes de tal manera que un conjunto degenerado de secuencias potenciales de PPSRP es expresable como polipéptidos individuales, o alternativamente, como un conjunto de polipéptidos de fusión más grandes (por ejemplo, para despliegue en fagos) que contiene allí el conjunto de secuencias de PPSRP. Existen una variedad de métodos que pueden ser utilizados para producir las bibliotecas de homólogos potenciales de PPSRP a partir de una secuencia degenerada de oligonucleótidos. La síntesis química de una secuencia degenerada de genes puede realizarse en un sintetizador automático de ADN, y el gen sintético se liga a continuación en un vector de expresión apropiado. El uso de un conjunto degenerado de genes permite proporcionar, en una mezcla, todas las secuencias que codifican el conjunto deseado de secuencias potenciales de PPSRP. Los métodos para sintetizar oligonucleótidos degenerados son conocidos en el arte (véase, por ejemplo, Narang, S. A. (1983), *Tetrahedron* 39:3; Itakura et al., (1984) *Ann. Rev. Biochem* 53: 323; Itakura et al. (1984) *Science* 198: 1056; Ike et al. (1983), *Nucleic Acid Res.* 11: 477).

Además, las bibliotecas de fragmentos de las regiones codificadoras de PPSRP pueden ser utilizadas para generar una población abigarrada de fragmentos de PPSRP para el cribado y selección posterior de homólogos de un PPSRP. Se puede generar una biblioteca de fragmentos de la secuencia de codificación por tratamiento de un fragmento bicatenario de la PCR de una secuencia de codificación de PPSRP con una nucleasa bajo condiciones en donde se produce amellamiento aproximadamente sólo una vez por molécula, la desnaturalización del ADN bicatenario, renaturalización del ADN para formar ADN bicatenario, que puede incluir pares sentido / antisentido de diferentes productos amellados, la remoción de porciones monocatenarias de dúplex reformados por tratamiento con nucleasa S1, y ligando la biblioteca de fragmentos resultante en un vector de expresión. Por medio de este método, se puede derivar una biblioteca de expresión que codifica fragmentos del terminal N, del terminal C e internos de diversos tamaños del PPSRP.

Se conocen varias técnicas en el arte para la selección de productos génicos de bibliotecas de combinación realizadas por mutaciones puntuales o de truncamiento, y para la selección de bibliotecas de ADNc para los productos génicos que tienen una propiedad seleccionada. Tales técnicas se pueden adaptar para la selección rápida de las bibliotecas de genes generadas por la mutagénesis combinatoria de homólogos de PPSRP. Las técnicas más ampliamente utilizadas, que son susceptibles a análisis de alto rendimiento, para selección de grandes bibliotecas de genes típicamente incluyen la clonación de la biblioteca de genes en vectores de expresión replicables, transformación de células apropiadas con la biblioteca resultante de vectores, y expresión de los genes de combinación bajo condiciones en las cuales la detección de una actividad deseada facilita el aislamiento del vector que codifica al gen cuyo producto se detectó. La mutagénesis recursiva de conjunto (REM), una nueva técnica que aumenta la frecuencia de mutantes funcionales en las bibliotecas, puede ser utilizada en combinación con los ensayos de selección para identificar homólogos de PPSRP (Arkin y Yourvan (1992), *PNAS* 89: 7811 - 7815; Delgrave et al. (1993), *Polypeptide Engineering* 6(3): 327 - 331). También, se pueden explotar los ensayos basados en las células para analizar una biblioteca abigarrada de PPSRP, utilizando métodos bien conocidos en el arte. Adicionalmente se divulga un método para identificación de un nuevo PPSRP, que comprende (a) elevar una respuesta de anticuerpos específicos con un PPSRP, o un fragmento del mismo, como se describe aquí; (b) seleccionar el material de PPSRP putativo con el anticuerpo, en donde el enlazamiento específico del anticuerpo con el material indica la presencia de un PPSRP potencialmente nuevo; y (c) analizar el material unido en comparación con un PPSRP conocido, para determinar su novedad.

Como se estableció anteriormente, la presente invención incluye al PPSRP como se define en la SEQ ID NO: 3 así como secuencias que sean al menos 90 a 95% idénticas a la secuencia entera de aminoácidos mostrada en la SEQ ID NO: 3. Para determinar el porcentaje de identidad de la secuencia de dos secuencias de aminoácidos (por ejemplo la SEC ID NO: 3 y una forma mutante de la misma), las secuencias se alinean con el fin de una comparación óptima (por ejemplo, se pueden introducir espacios en la secuencia de un polipéptido para la alineación óptima con el otro polipéptido o ácido nucleico). Se comparan entonces los residuos de aminoácidos en las correspondientes posiciones de los aminoácidos. Cuando una posición en una secuencia (SEQ ID NO: 3) está ocupada por el mismo residuo aminoácido que en la posición correspondiente en la otra secuencia (una forma

mutante de la secuencia seleccionada del polipéptido de la SEQ ID NO: 3), entonces las moléculas son idénticas en esa posición. Se puede hacer el mismo tipo de comparación entre dos secuencias de ácido nucleico.

5 El porcentaje de identidad de secuencia entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias (es decir, el porcentaje de identidad de secuencia = número de posiciones idénticas / número total de posiciones x 100). Preferiblemente, los homólogos aislados de aminoácidos incluidos en la presente invención son al menos aproximadamente 90 - 95 %, y lo más preferible al menos aproximadamente 96%, 97%, 98%, 99% o más, idénticos a una secuencia entera de aminoácidos mostrada en la SEQ ID NO: 3. En aún otra modalidad, los homólogos aislados de aminoácidos incluidos en la presente invención son al menos aproximadamente 90 - 95%, y lo más preferible al menos aproximadamente 96%, 97%, 98%, 99% o más, idénticos a una secuencia entera de aminoácidos codificada por una secuencia de ácido nucleico mostrada en SEQ ID NO: 2.

10 En otra realización preferida, un homólogo aislado de ácido nucleico de la invención comprende una secuencia de nucleótidos que es al menos aproximadamente 80 - 85%, preferiblemente al menos 85 - 90% ó 90 - 95%, e incluso más preferiblemente aproximadamente al menos 95%, 96%, 97%, 98%, 99% o más, idéntica a una secuencia de nucleótidos mostrada en la SEQ ID NO: 2. La longitud preferible de comparación de la secuencia para ácidos nucleicos es la longitud total de la región de codificación.

15 Se prefiere adicionalmente que el homólogo aislado de ácido nucleico de la invención codifique un PPSRP, o una porción del mismo, que sea al menos 90% idéntica a una secuencia de aminoácidos de la SEQ ID NO: 3, y que funciona como un modulador de una respuesta al estrés ambiental en una planta. En una modalidad más preferida, la sobreexpresión del homólogo de ácido nucleico en una planta aumenta la tolerancia de la planta a un estrés ambiental. En una modalidad preferida adicional, el homólogo de ácido nucleico codifica un PPSRP que funciona como una proteína fosfatasa.

20 Para los propósitos de la invención, el porcentaje de identidad de secuencia entre dos secuencias de ácido nucleico o de polipéptido se determina utilizando el paquete de software Vector NTI 6,0 (PC) (Informax, 7600 Wisconsin Ave., Bethesda, MD 20814). Se utiliza una penalización por abertura del espacio de 15 y una penalización por extensión del espacio de 6,66 para determinar el porcentaje de identidad de dos ácidos nucleicos. Se utiliza una penalización por abertura del espacio de 10 y una penalización por extensión del espacio de 0,1 para determinar el porcentaje de identidad de dos polipéptidos. Todos los demás parámetros se fijan en los ajustes predeterminados. Para el propósito de un alineamiento múltiple (algoritmo de Clustal W), la penalización por abertura del espacio es de 10, y la penalización por extensión del espacio es de 0,05, con la matriz blosum62 matriz. Se debe entender que para los fines de determinar la identidad de secuencia cuando se compara una secuencia de ADN con una secuencia de ARN, un nucleótido de timidina es equivalente a un nucleótido de uracilo.

25 También se describe un ácido nucleico aislado que comprende un polinucleótido que se hibrida con el polinucleótido de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31 bajo condiciones estrictas. Más particularmente, una molécula aislada de ácido nucleico es al menos de 15 nucleótidos de longitud e hibrida bajo condiciones estrictas con la molécula de ácido nucleico que comprende una secuencia de nucleótidos de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31. Por ejemplo, el ácido nucleico es de al menos 30, 50, 100, 250 o más nucleótidos de longitud. Preferiblemente, un homólogo aislado de ácido nucleico comprende una secuencia de nucleótidos que hibrida en condiciones altamente rigurosas con la secuencia de nucleótidos mostrada en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEC ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31 y funciona como un modulador de la tolerancia de la planta a un estrés ambiental. Se prefiere, el homólogo aislado de ácido nucleico codifique un PPSRP que funciona como una proteína fosfatasa.

35 Como se usa aquí con respecto a la hibridación para ADN con una transferencia de ADN, el término "condiciones estrictas" se refiere a la hibridación durante la noche a 60° C en solución de Denhart 10X, SSC 6X, SDS al 0,5% y 100 ug / ml de ADN de esperma de salmón desnaturalizado. Las transferencias se lavan secuencialmente a 62° C durante 30 minutos cada vez en SSC 3X / SDS al 0,1%, seguido por SSC 1X / SDS al 0,1% y, finalmente, SSC 0,1X / SDS al 0,1%. Como también se utiliza aquí, "condiciones altamente rigurosas" se refiere a la hibridación durante la noche a 65° C en solución de Denhart 10X, SSC 6X, SDS al 0,5% y 100 ug / ml de ADN de esperma de salmón desnaturalizado. Las transferencias se lavaron secuencialmente a 65° C durante 30 minutos cada vez en SSC 3X / SDS al 0,1%, seguido por SSC 1X / SDS al 0,1% y, finalmente, SSC 0,1X / SDS al 0,1%. Los métodos para hibridaciones de ácido nucleico se describen en Meinkoth y Wahl (1984) Anal. Biochem. 138: 267 - 284; Current Protocols in Molecular Biology, Capítulo 2, Ausubel et al. Eds., Greene Publishing and Wiley-Interscience, Nueva York, 1995; y Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization with Nucleic Acid Probes, Parte I, Capítulo 2, Elsevier, Nueva York, 1993. Una molécula aislada de ácido nucleico de la invención que hibrida bajo condiciones rigurosas o altamente estrictas con una secuencia de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEC ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31 corresponde a una

molécula de ácido nucleico de origen natural. Como se usa aquí, una molécula de ácido nucleico de "origen natural" se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que se presenta en la naturaleza (por ejemplo, que codifica un polipéptido natural). Los ácidos nucleicos de la divulgación codifican un PPSRP de origen natural de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u *Oryza sativa*.

5 Utilizando los métodos descritos anteriormente, y otros conocidos por aquellos capacitados en la técnica, alguien ordinariamente experimentado en el arte puede aislar homólogos de los PPSRP que comprenden las secuencias de aminoácidos mostradas en la SEQ ID NO: 3. Un subconjunto de estos homólogos es de variantes alélicas. Tal como se usa aquí, el término "variante alélica" se refiere a una secuencia de nucleótidos que contiene polimorfismos que conducen a cambios en las secuencias de aminoácidos de un PPSRP y que existen dentro de una población natural
10 (por ejemplo, una especie de planta o variedad). Tales variaciones alélicas naturales típicamente puede resultar en una varianza del 1 - 5% en un ácido nucleico para PPSRP. Las variantes alélicas pueden ser identificadas por medio de la secuenciación de la secuencia de ácido nucleico de interés en una cantidad de plantas diferentes, que puede ser fácilmente llevada a cabo mediante el uso de sondas de hibridación para identificar el mismo locus genético para PPSRP en esas plantas. Todas y cada una de las variaciones de ácidos nucleicos y los polimorfismos de los
15 aminoácidos resultantes o las variaciones en un PPSRP que son el resultado de la variación alélica natural y que no alteran la actividad funcional de un PPSRP, con tal de que estos ácidos nucleicos que codifican para PPSRP y las secuencias de proteína que ellos codifican cumplan con la definición en las reivindicaciones 1 y 4 de la invención, se pretende que estén dentro del alcance de la invención.

Además, se pretende que la molécula de ácido nucleico que codifica PPSRP de la invención de la misma o de otras especies tales como análogos, ortólogos y parálogos de PPSRP, estén dentro del alcance de la presente invención. Como se usa aquí, el término "análogos" se refiere a dos ácidos nucleicos que tienen la misma función o una función similar, pero que han evolucionado por separado en organismos no relacionados. Tal como se usa aquí, el término "ortólogos" se refiere a dos ácidos nucleicos de diferentes especies, pero que han evolucionado a partir de un gen ancestral común por especiación. Normalmente, los ortólogos codifican polipéptidos que tienen las mismas funciones o similares. Como también se usa aquí, el término "parálogos" se refiere a dos ácidos nucleicos que están relacionados por la duplicación dentro de un genoma. Los parálogos usualmente tienen funciones diferentes, pero estas funciones puede estar relacionadas (Tatusov, R. L. et al. 1997 Ciencia 278 (5338): 631 - 637). Los análogos, ortólogos y parálogos de un PPSRP de origen natural pueden diferir del PPSRP de origen natural por modificaciones posteriores a la traducción, por diferencias en la secuencia de aminoácidos, o por ambos. La modificaciones posteriores a la traducción incluyen formación de derivados químicos *in vivo* e *in vitro* de polipéptidos, por ejemplo, acetilación, carboxilación, fosforilación o glicosilación, y tales modificaciones pueden presentarse durante la síntesis o el procesamiento del polipéptido o después del tratamiento con enzimas modificadoras aisladas. En particular, los ortólogos de la invención generalmente exhibirán al menos 90 - 95%, y más preferiblemente 95%, 96%, 97%, 98% o incluso 99% de identidad o de identidad de secuencia con toda o parte de la secuencia de aminoácidos del PPSRP de origen natural y exhibirá una función similar al PPSRP. Preferiblemente, el ortólogo de PPSRP de la presente invención actúa como un modulador de la respuesta al estrés por sequía en una planta y / o actúa como una proteína fosfatasa. Más preferiblemente, los ortólogos de PPSRP incrementan la tolerancia al estrés de una planta. En una modalidad, los ortólogos de PPSRP mantienen la capacidad de participar en el metabolismo de los compuestos necesarios para la construcción de las membranas celulares en una planta, o en el transporte de moléculas a través de estas membranas.
20
25
30
35
40

Además de las variantes de origen natural de una secuencia de PPSRP que pueda existir en la población, el experto en la materia se dará cuenta además que se pueden introducir cambios por mutación en una secuencia de nucleótidos de la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, o SEQ ID NO: 32, lo que conduce por lo tanto a cambios en la secuencia de aminoácidos del PPSRP codificado, sin alterar la actividad funcional del PPSRP. Por ejemplo, pueden hacerse sustituciones de nucleótidos que conducen a sustituciones de aminoácidos en residuos de aminoácidos "no esenciales" en una secuencia de la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, o SEQ ID NO: 32. Un residuo de aminoácido "no esencial" es un residuo que puede ser alterado a partir de la secuencia de tipo silvestre de uno de los PPSRP sin alterar la actividad de dicho PPSRP, mientras que se requiere de un residuo de aminoácido "esencial" para la actividad del PPSRP. Otros residuos de aminoácidos, sin embargo, (por ejemplo, aquellos que no se han conservado o únicamente conservados en forma parcial en el dominio que tiene actividad de PPSRP) pueden no ser esenciales por la actividad y por lo tanto son probablemente susceptibles a modificaciones sin alterar la actividad del PPSRP.
45
50
55

En consecuencia, otro aspecto de la invención se refiere a la molécula de ácido nucleico que codifica el PPSRP de la invención que contiene cambios en los residuos de aminoácidos que no son esenciales para la actividad de PPSRP. De acuerdo con la definición de la Reivindicación 4, tal PPSRP difiere en la secuencia de aminoácidos de una secuencia contenida en la SEQ ID NO: 3, sin embargo, mantiene al menos una de las actividades de PPSRP descritas aquí. En una modalidad, la molécula aislada de ácido nucleico comprende una secuencia de nucleótidos que codifica un polipéptido, en donde el polipéptido comprende una secuencia de aminoácidos al menos aproximadamente 90 - 95% idéntica a las secuencias de la SEQ ID NO: 3, y lo más preferible aproximadamente al
60

menos 96%, 97%, 98%, o 99% idéntica a las secuencias de la SEQ ID NO: 3. Los homólogos preferidos de PPSRP de la presente invención preferiblemente participan en la respuesta de tolerancia al estrés por sequía en una planta, o más particularmente en la transcripción de un polipéptido involucrado en la respuesta de tolerancia al estrés por sequía en una planta, y / o actúan como una proteína fosfatasa.

5 Una molécula aislada de ácido nucleico que codifica un PPSRP que tiene por ejemplo 90 a 95% de identidad de secuencia comparado con la secuencia entera de aminoácidos con una secuencia del polipéptido de la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, o SEQ ID NO: 32 puede ser creada mediante la introducción de una o más sustituciones, adiciones o supresiones de nucleótidos en una secuencia de nucleótidos de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31, respectivamente, de tal manera que una o más sustituciones, adiciones o supresiones de aminoácidos son introducidas en el polipéptido codificado. Las mutaciones pueden ser introducidas en una de las secuencias de SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31 mediante técnicas estándar, tales como mutagénesis dirigida al sitio y mutagénesis mediada por PCR. Preferiblemente, las sustituciones conservadoras de aminoácidos se realizan en uno o más residuos de los aminoácidos no esenciales predichos. Una "sustitución conservadora de aminoácidos" es una en la que se sustituye el residuo de aminoácido con un residuo de aminoácido que tiene una cadena lateral similar.

Las familias de residuos de aminoácidos que tienen cadenas laterales similares han sido definidas en el arte. Estas familias incluyen aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares sin carga (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas en beta (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). De este modo, un residuo predicho de aminoácido no esencial en un PPSRP preferiblemente se sustituye con otro residuo de aminoácido de la misma familia de cadena lateral. Alternativamente, en otra modalidad, se pueden introducir mutaciones al azar a lo largo de toda o parte de una secuencia de codificación de PPSRP, por ejemplo mediante mutagénesis de saturación, y los mutantes resultantes pueden ser seleccionados por una actividad de PPSRP descrita aquí para identificar mutantes que retienen la actividad de PPSRP. Después de la mutagénesis de una de las secuencias de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, y SEQ ID NO: 31, el polipéptido codificado puede ser expresado en forma recombinante y se puede determinar la actividad del polipéptido mediante el análisis de la tolerancia al estrés de una planta que expresa al polipéptido como se describe en el Ejemplo 7.

Adicionalmente, se pueden crear los ácidos nucleicos optimizados para PPSRP. Preferiblemente, un ácido nucleico optimizado para PPSRP que codifica un PPSRP que se enlaza con un grupo fosfato y / o modula la tolerancia de una planta a un estrés ambiental, y más preferiblemente aumenta la tolerancia de una planta a un estrés ambiental por su sobreexpresión en la planta. Como se usa aquí, "optimizado" se refiere a un ácido nucleico que se manipula genéticamente para aumentar su expresión en una planta o animal dado. Para proporcionar ácidos nucleicos optimizados para PPSRP de una planta, se puede modificar la secuencia de ADN del gen para 1) incluir codones preferidos por los genes de plantas altamente expresados, 2) incluir un contenido de A + T en la composición base de nucleótidos con aquella sustancialmente encontrada en las plantas; 3) formar una secuencia de inicio de la planta, 4) eliminar las secuencias que causan desestabilización, poliadenilación inadecuada, degradación y terminación del ARN, o que forma horquillas de estructura secundaria de ARN o sitios de empalme del ARN. El aumento de expresión de ácido nucleicos para PPSRP en las plantas puede lograrse mediante la utilización de la frecuencia de distribución del uso de codones en las plantas en general o en una planta particular. Los métodos para la optimización de la expresión de ácido nucleico en las plantas se pueden encontrar en EPA 0359472; EPA 0385962, WO 91/16432; patente de los Estados Unidos No. 5.380.831; Patente de los Estados Unidos No. 5.436.391; Perlack et al. (1991) Proc. Natl. Acad. Sci. EE.UU. 88: 3324 - 3328; y Murray et al. (1989) Nucleic Acids Res. 17: 477 - 498.

Como se usa aquí, "frecuencia de uso del codón preferido" se refiere a la preferencia mostrada por una célula huésped específica en el uso de codones de nucleótidos para especificar un aminoácido dado. Para determinar la frecuencia de uso de un codón particular en un gen, se divide el número de apariciones de ese codón en el gen por el número total de apariciones de todos los codones que especifican el mismo aminoácido en el gen. De manera similar, se puede calcular la frecuencia de uso del codón preferido exhibida por una célula huésped promediando la frecuencia de uso del codón preferido en un gran número de genes expresados por la célula huésped. Es preferible que este análisis se limite a los genes que son altamente expresados por la célula huésped. El porcentaje de desviación de la frecuencia de uso del codón preferido para un gen sintético a partir de aquella empleada por una célula huésped se calcula primero mediante la determinación del porcentaje de desviación de la frecuencia de uso de un solo codón a partir de aquella de la célula huésped seguido por la obtención de la desviación promedio sobre

5 todos los codones. Tal como se define en este documento, este cálculo incluye codones únicos (es decir, ATG y TGG). En términos generales, la desviación media total de la utilización de codones de un gen optimizado a partir de aquella de una célula huésped se calcula utilizando la ecuación $1A = n = 1 ZX_n - Y_n X_n$ veces 100 Z, donde X_n = frecuencia de uso para el codón n en la célula huésped; Y_n = frecuencia de uso para el codón n en el gen sintético, n representa un codón individual que especifica un aminoácido y el número total de codones es Z. La desviación total de la frecuencia de uso del codón, A, para todos los aminoácidos debe ser preferiblemente aproximadamente menor al 25%, y más preferiblemente aproximadamente menor al 10%.

10 Por lo tanto, un ácido nucleico para PPSRP se puede optimizar de manera que su frecuencia de distribución de uso del codón se desvíe, preferiblemente, no más del 25% a partir de aquella de los genes de plantas altamente expresados y, más preferiblemente, aproximadamente no más del 10%. Además, se tiene en cuenta el contenido en porcentaje de G + C de la tercera base degenerada (las monocotiledóneas parecen favorecer G + C en esta posición, mientras que las dicotiledóneas no lo hacen). También se reconoce que el nucleótido XCG (donde X es A, T, C o G) es el codón menos preferido en dicotiledóneas mientras que el codón XTA es evitado tanto en monocotiledóneas como en dicotiledóneas. Los ácidos nucleicos optimizados para PPSRP de esta invención también tienen preferiblemente indicios de evitar el doblete CG y TA estrechamente cercanos a aquellos de la planta huésped escogida (es decir, *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u *Oryza sativa*). Más preferiblemente, estos índices se desvían de aquel del huésped aproximadamente no más del 10 - 15%.

20 Además de las moléculas de ácido nucleico que codifican los PPSRP descritas anteriormente, se conocen moléculas aisladas de ácido nucleico que son antisentido de las mismas. Se cree que los polinucleótidos antisentido inhiben la expresión génica de un polinucleótido objetivo enlazando específicamente al polinucleótido objetivo e interfiriendo con la transcripción, empalme, transporte, traducción y / o estabilidad del polinucleótido objetivo. En el estado del arte se describen métodos para dirigir el polinucleótido antisentido hacia el ADN cromosómico, hacia un transcrito primario de ARN o hacia un ARNm procesado. Preferiblemente, las regiones objetivo incluyen sitios de empalme, codones de iniciación de la traducción, codones de terminación de la traducción, y otras secuencias dentro del marco de lectura abierto.

30 El término "antisentido", para los fines de la invención, se refiere a un ácido nucleico que comprende un polinucleótido que es suficientemente complementario a todo o a una porción de un gen, transcrito primario o ARNm procesado, para interferir con la expresión del gen endógeno. Polinucleótidos "complementarios" son aquellos que son capaces de apareamiento de bases de acuerdo con las normas de complementación estándar de Watson-Crick. Específicamente, las purinas se aparearán con pirimidinas para formar una combinación de guanina apareada con citosina (G:C) y adenina apareada con cualquier timina (A:T) en el caso de ADN, o adenina apareada con uracilo (A:U) en el caso de ARN. Se entiende que dos polinucleótidos pueden hibridarse entre sí, incluso si no son totalmente complementarios entre sí, siempre y cuando cada uno tenga al menos una región que sea sustancialmente complementaria con la otra. El término "ácido nucleico antisentido" incluye casetes de expresión de ARN monocatenario, así como de ADN bicatenario que pueden ser transcritos para producir un ARN antisentido. Ácidos nucleicos antisentido "activos" son moléculas de ARN antisentido que son capaces de hibridarse selectivamente con un transcrito primario o ARNm que codifica un polipéptido que tiene al menos 80% de identidad de secuencia con el polipéptido de la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, y SEQ ID NO: 32.

45 El ácido nucleico antisentido puede ser complementario a una hebra completa de codificación de PPSRP, o solamente a una parte de la misma. Por ejemplo, una molécula de ácido nucleico antisentido es antisentido a una "región de codificación" de la hebra de codificación de una secuencia de nucleótidos que codifica una PPSRP. El término "región de codificación" se refiere a la región de la secuencia de nucleótidos que comprende codones que se traducen en residuos de aminoácidos. En otro ejemplo, la molécula de ácido nucleico antisentido es antisentido a una "región no codificante" de la hebra de codificación de una secuencia de nucleótidos que codifica un PPSRP. El término "región no codificante" se refiere a secuencias 5' y 3' que flanquean la región de codificación que no se traducen en aminoácidos (es decir, también denominadas como regiones 5' y 3' no traducidas). La molécula de ácido nucleico antisentido puede ser complementaria a la región de codificación completa del ARNm para PPSRP, pero más preferiblemente es un oligonucleótido que es antisentido sólo a una porción de la región de codificación o no codificadora del ARNm para PPSRP. Por ejemplo, el oligonucleótido antisentido puede ser complementario a la región que rodea al sitio de inicio de la traducción del ARNm para PPSRP. Un oligonucleótido antisentido puede ser, por ejemplo, aproximadamente de 5, 10, 15, 20, 25, 30, 35, 40, 45 ó 50 nucleótidos de longitud.

55 Se puede construir un ácido nucleico antisentido utilizando síntesis química y reacciones enzimáticas de ligación utilizando procedimientos conocidos en el arte. Por ejemplo, un ácido nucleico antisentido (por ejemplo, un oligonucleótido antisentido) puede ser sintetizado químicamente utilizando nucleótidos de origen natural o nucleótidos modificados de diferente forma diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre los ácidos nucleicos sentido y antisentido, por ejemplo, se pueden utilizar derivados de fosforotioato y nucleótidos sustituidos con acridina. Los ejemplos de nucleótidos modificados que pueden ser utilizados para generar el ácido nucleico antisentido incluyen 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetil citosina, 5-(carboxihidroximetil)uracilo, 5-

5 carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N-6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N-6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metil-N-6-isopenteniladenina, ácido uracil-5-oxiacético (v), wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico del ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w, y 2,6-diaminopurina. Alternativamente, el ácido nucleico antisentido puede ser producido biológicamente usando un vector de expresión dentro del cual se ha subclonado un ácido nucleico en una orientación antisentido (es decir, el ARN transcrito a partir del ácido nucleico insertado será de una orientación antisentido para un ácido nucleico objetivo de interés, descrito más detalladamente en la subsección siguiente).

15 Por ejemplo, la molécula de ácido nucleico antisentido de la invención es una molécula de ácido nucleico α -anomérico. Una molécula de ácido nucleico α -anomérico específicas forma híbridos bicatenarios específicos con ARN complementario en los cuales, contrariamente a las unidades β habituales, las hebras corren paralelas entre sí (Gaultier et al. (1987) *Nucleic Acids Res.* 15: 6625 - 6641). La molécula de ácido nucleico antisentido puede comprender también un 2'-O-metilribonucleótido (Inoue et al. (1987) *Nucleic Acids. Res.* 15: 6131 - 6148) o un análogo químico de ARN - ADN (Inoue et al. (1987), *FEBS Lett.* 215: 327 - 330).

20 Las moléculas de ácido nucleico antisentido se administran típicamente a una célula o se generan *in situ* de tal manera que se hibridan con o se unen al ARNm celular y / o al ADN genómico que codifica un PPSRP para inhibir así la expresión del polipéptido, por ejemplo, por inhibición de la transcripción y / o de la traducción. La hibridación puede ser por medio de la complementariedad convencional de nucleótidos para formar un dúplex estable, o, por ejemplo, en el caso de una molécula de ácido nucleico antisentido que se une a dúplex de ADN, a través de interacciones específicas en el surco mayor de la doble hélice. La molécula antisentido puede ser modificada de tal forma que se une específicamente a un receptor o a un antígeno expresado sobre una superficie de la célula seleccionada, por ejemplo, mediante el enlazamiento de la molécula de ácido nucleico antisentido con un péptido o un anticuerpo que se une a un receptor de la superficie de la célula o antígeno. También se puede suministrar la molécula de ácido nucleico antisentido a las células utilizando los vectores descritos aquí. Para lograr suficientes concentraciones intracelulares de las moléculas antisentido, se prefieren las construcciones de vectores en las que se coloca la molécula de ácido nucleico antisentido bajo el control de un promotor procarionta, viral o eucariota fuerte (incluidas las plantas).

35 Como una alternativa a los polinucleótidos antisentido, se pueden utilizar ribozimas, polinucleótidos sentido o ARN bicatenario (ARNbc) para reducir la expresión de un polipéptido PPSRP. Por "ribozima" se entiende una enzima catalítica con base en ARN con actividad de ribonucleasa, que es capaz de escindir un ácido nucleico monocatenario, tal como un ARNm, con el cual tiene una región complementaria. Se pueden utilizar las ribozimas (por ejemplo, ribozimas cabeza de martillo descritos en Haselhoff y Gerlach (1988), *Nature* 334: 585 - 591) para escindir catalíticamente transcritos de ARNm para PPSRP para inhibir así la traducción del ARNm para PPSRP. Se puede diseñar una ribozima que tiene especificidad por un ácido nucleico que codifica PPSRP con base en la secuencia de nucleótidos de un ADNc para PPSRP, como se describe aquí (es decir, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31), o sobre la base de una secuencia heteróloga que se aísla de acuerdo con los métodos que se enseñan en esta invención. Por ejemplo, se puede construir un derivado de un L-19 IVS del ARN de *Tetrahymena* en el que la secuencia de nucleótidos del sitio activo es complementaria a la secuencia de nucleótidos que se escinde en un ARNm que codifica PPSRP. Véase, por ejemplo, Cech et al., patente de los Estados Unidos No. 4.987.071 y Cech et al., patente de los Estados Unidos No. 5.116.742. Alternativamente, se puede utilizar ARNm para PPSRP para seleccionar un ARN catalítico que tenga una actividad específica de ribonucleasa a partir de una reserva de moléculas de ARN. Véase, por ejemplo, Bartel, D. y Szostak, J. W., 1993 *Science* 261: 1411 - 1418. Por ejemplo, la ribozima contendrá una porción que tiene al menos 7, 8, 9, 10, 12, 14, 16, 18 ó 20 nucleótidos, y más preferiblemente 7 u 8 nucleótidos, que tiene 100% de complementariedad con una porción del ARN objetivo. Los métodos para preparar las ribozimas son conocidos por aquellos capacitados en el arte. Véase, por ejemplo, las patentes de los Estados Unidos Nos. 6.025.167, 5.773.260 y 5.496.698.

55 El término "ARNbc" (ARN bicatenario), tal como se utiliza aquí, se refiere a híbridos de ARN que comprenden dos cadenas de ARN. Los ARN bicatenarios puede ser de estructura lineal o circular. Por ejemplo, el ARNbc es específico para un polinucleótido que codifica ya sea al polipéptido de la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, o SEQ ID NO: 32; o un polipéptido que tiene al menos 80% de identidad de secuencia con la SEQ ID NO: 3, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, y SEQ ID NO: 32. Los ARN de hibridación puede ser sustancial o completamente complementarios. Por "sustancialmente complementarios", se entiende que cuando los dos ARN de hibridación están alineados de forma óptima utilizando el programa BLAST como se describió anteriormente, las porciones de hibridación son al menos el 95% complementarias. Preferiblemente, el ARNbc tendrá por lo menos 100

pares de bases de longitud. Típicamente, los ARN de hibridación serán de longitud idéntica sin extremos 5' o 3' que sobresalgan y sin espacios. Sin embargo, los ARNbc que tienen salientes 5' o 3' de hasta 100 nucleótidos pueden ser utilizados en los métodos de la invención.

5 El ARNbc puede comprender ribonucleótidos o análogos de ribonucleótidos, por ejemplo residuos de 2'-O-metil ribosilo o combinaciones de los mismos. Véase las patentes de los Estados Unidos Nos. 4.130.641 y 4.024.222. Un ácido polirribonucleosínico ARNbc: el ácido polirribocitidílico está descrito en la patente de los Estados Unidos No. 4.283.393. Los métodos para elaborar y utilizar ARNbc son conocidos en el arte. Un método comprende la transcripción simultánea de dos hebras complementarias de ADN, ya sea *in vivo*, o en una sola mezcla de reacción *in vitro*. Véase, por ejemplo, la Patente de los Estados Unidos No. 5.795.715. El ARNbc puede ser introducido en
10 una planta o en una célula de una planta directamente por medio de procedimientos estándar de transformación. Alternativamente, el ARNbc puede ser expresado en una célula de una planta transcribiendo dos ARN complementarios.

15 Otros métodos para la inhibición de la expresión de genes endógenos, tales como la formación de una hélice triple (Moser et al. (1987), *Science* 235: 645 - 650 y Cooney et al. (1988), *Science* 241: 456 - 459) y cosupresión (Napoli et al. (1990), *The Plant Cell* 2: 279 - 289) son conocidos en el arte. Los ADNc parciales y de longitud completa han sido utilizados para la cosupresión de genes de plantas endógenas. Véase, por ejemplo, las patentes de los Estados Unidos Nos. 4.801.340, 5.034.323, 5.231.020 y 5.283.184, Van der Kroll et al. (1990), *The Plant Cell* 2: 291 - 299, Smith et al. (1990), *Mol Gen Genetics* 224: 477 - 481 y Napoli et al. (1990), *The Plant Cell* 2: 279 - 289.

20 Para la supresión sentido, se cree que la introducción de un polinucleótido sentido bloquea la transcripción del gen objetivo correspondiente. El polinucleótido sentido tendrá al menos 65% de identidad de secuencia con el gen objetivo de la planta o ARN. Preferentemente, el porcentaje de identidad es al menos del 80%, 90%, 95% o más. El polinucleótido sentido introducido no necesita ser de longitud completa con respecto al gen objetivo o transcripto. Preferiblemente, el polinucleótido sentido tendrá al menos 65% de identidad de secuencia con al menos 100 nucleótidos consecutivos de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31. Las regiones de identidad puede comprender intrones y / o exones y regiones no traducidas. El polinucleótido sentido introducido puede estar presente en la célula de la planta en forma transitoria, o puede ser integrado en forma estable en un cromosoma de la planta o replicón extracromosómico.

30 Alternativamente, la expresión del gen para PPSRP puede ser inhibida dirigiendo las secuencias de nucleótidos complementarias a la región reguladora de una secuencia de nucleótidos para PPSRP (por ejemplo, un promotor y / o potenciador de PPSRP) para formar estructuras helicoidales triples que impiden la transcripción de un gen para PPSRP en las células objetivo. Véase, en general, Helene, C. (1991) *Anticancer Drug Des.* 6(6): 569 - 84; Helene, C. et al. (1992), *Ann. N. Y. Acad. Sci.* 660: 27 - 36; y Maher, L. J. (1992), *Bioassays* 14(12): 807 - 15.

35 Además de los ácidos nucleicos para PPSRP y de los polipéptidos descritos anteriormente, estos ácidos nucleicos y polipéptidos pueden ser unidos a una unidad estructural. Estas unidades estructurales incluyen, pero no se limitan a, las unidades estructurales de detección, unidades estructurales de hibridación, unidades estructurales de purificación, unidades estructurales de suministro, unidades estructurales de reacción, unidades estructurales de enlazamiento, y similares. Un grupo típico de ácidos nucleicos que tienen unidades estructurales enlazadas son sondas e iniciadores. Las sondas y los iniciadores comprenden típicamente un oligonucleótido sustancialmente aislado. El oligonucleótido comprende típicamente una región de la secuencia de nucleótidos que hibrida bajo condiciones rigurosas con al menos aproximadamente 12, preferiblemente alrededor de 25, más preferiblemente alrededor de 40, 50 ó 75 nucleótidos consecutivos de una cadena de sentido de una de las secuencias expuestas en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31; una secuencia antisentido de una de las secuencias expuestas en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31; o mutantes de origen natural de los mismos. Los iniciadores basados en una secuencia de nucleótidos de la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, o SEQ ID NO: 31 se pueden utilizar en reacciones PCR para clonar homólogos de PPSRP. Las sondas basadas en las secuencias de nucleótidos para PPSRP se puede utilizar para detectar transcritos o secuencias genómicas que codifican los mismos polipéptidos o sustancialmente idénticos. Por ejemplo, la sonda comprende además un grupo de marcación unido al mismo, por ejemplo, el grupo de marcación puede ser un isótopo radioactivo, un compuesto fluorescente,
55 una enzima, o un cofactor de la enzima. Tales sondas pueden ser utilizados como parte de un kit de prueba marcador genómico para la identificación de las células que expresan un PPSRP, por ejemplo mediante la medición del nivel de ácido nucleico que codifica PPSRP, en una muestra de células, por ejemplo, detectando los niveles de ARNm para PPSRP o determinando si un gen genómico para PPSRP ha sido mutado o suprimido.

60 En particular, un método útil para comprobar el nivel de transcripción del gen (un indicador de la cantidad de ARNm disponible para la traducción hasta el producto génico) es realizar una transferencia tipo Northern (para referencia

véase, por ejemplo, Ausubel et al. (1988), Current Protocols in Molecular Biology, Wiley: Nueva York). Esta información demuestra al menos parcialmente el grado de transcripción del gen transformado. El ARN celular total se puede preparar a partir de células, tejidos u órganos por diferentes métodos, todos ellos bien conocidos en el arte, tal como el descrito en Bormann, E. R et al. (1992), Mol. Microbiol. 6: 317 - 326. Para evaluar la presencia o la cantidad relativa de polipéptido traducido a partir de este ARNm, se pueden emplear técnicas estándar, tales como una transferencia tipo Western. Estas técnicas son bien conocidas para una persona ordinariamente capacitada en el arte. (Véase, por ejemplo, Ausubel et al. (1988), Current Protocols in Molecular Biology, Wiley: Nueva York).

La invención proporciona además un vector de expresión recombinante aislado que comprende al ácido nucleico para PPSRP de la invención como se describió anteriormente, en donde la expresión del vector en una célula huésped resulta en una mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la célula huésped. Como se usa aquí, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico con el cual se ha enlazado. Un tipo de vector es un "plásmido", que se refiere a un bucle circular de ADN bicatenario dentro del cual se pueden ligar segmentos adicionales de ADN. Otro tipo de vector es un vector viral, en donde se pueden ligar segmentos adicionales de ADN en el genoma viral. Algunos vectores son capaces de replicación autónoma en una célula huésped dentro de la cual se introducen (por ejemplo, vectores bacterianos que tiene un origen de replicación bacteriano y vectores episomales de mamíferos). Otros vectores (por ejemplo, vectores no episomales de mamífero) se integran en el genoma de una célula huésped después de la introducción en la célula huésped, y con ello se replican junto con el genoma del huésped. Además, ciertos vectores son capaces de dirigir la expresión de genes a los que están operativamente enlazados. Tales vectores se denominan aquí como "vectores de expresión". En general, los vectores de expresión de utilidad en las técnicas de ADN recombinante están a menudo en forma de plásmidos. En la presente especificación, "plásmido" y "vector" se puede utilizar indistintamente ya que el plásmido es la forma más usual de vector. Sin embargo, la invención está destinada a incluir tales otras formas de vectores de expresión, tales como vectores virales (por ejemplo, los retrovirus de replicación defectuosa, adenovirus y virus adeno-asociados), que realizan funciones equivalentes.

Los vectores de expresión recombinante de la invención comprenden un ácido nucleico de la invención en una forma adecuada para la expresión del ácido nucleico en una célula huésped, lo que significa que los vectores de expresión recombinante incluyen una o más secuencias reguladoras, seleccionadas sobre con base en las células huésped que se utilizan para la expresión, que están operativamente enlazadas a la secuencia de ácido nucleico que va a ser expresada. Dentro de un vector de expresión recombinante, "operativamente enlazada" pretende significar que la secuencia de nucleótidos de interés está enlazada a la(s) secuencia(s) reguladora(s) de tal manera que permite la expresión de la secuencia de nucleótidos (por ejemplo, en un sistema de transcripción / traducción *in vitro* o en una célula huésped cuando se introduce el vector en la célula huésped). El término "secuencia reguladora" está destinado a incluir promotores, potenciadores y otros elementos de control de la expresión (por ejemplo, señales de poliadenilación). Tales secuencias reguladoras se describen, por ejemplo, en Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990) o véase: Gruber y Crosby, en: Methods in Plant Molecular Biology and Biotechnology, eds. Glick y Thompson, Capítulo 7, 89 - 108, CRC Press: Boca Raton, Florida, incluyendo las referencias citadas allí. Las secuencias reguladoras incluyen aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células huésped y aquellas que dirigen la expresión de la secuencia de nucleótidos sólo en ciertas células huésped o bajo ciertas condiciones. Se apreciará por parte de aquellos capacitados en el arte que el diseño del vector de expresión puede depender de factores tales como la elección de la célula huésped a transformar, el nivel de expresión del polipéptido deseado, etc. Los vectores de expresión de la invención puede ser introducidos en células huésped para producir con ello polipéptidos o péptidos, incluyendo polipéptidos o péptidos de fusión, codificados por ácidos nucleicos como los aquí descritos (por ejemplo, los PPSRP, formas mutantes de los PPSRP, polipéptidos de fusión, etc.).

Los vectores de expresión recombinante de la invención pueden ser diseñados para la expresión del PPSRP de la invención en células procariotas o eucariotas. Por ejemplo, los genes para PPSRP pueden ser expresados en células bacterianas tales como *C. glutamicum*, células de insectos (usando vectores de expresión de baculovirus), levaduras y otras células de hongos (véase Romanos, M. A. et al. (1992), Foreign gene expression in yeast: una revisión, Yeast 8: 423 - 488; van den Hondel, C. A. M. J. J. et al. (1991), Heterologous gene expression in filamentous fungi, en: More Gene Manipulations in Fungi, J. W. Bennet & L. L. Lasure, eds., p. 396 - 428: Academic Press: San Diego; y van den Hondel, C. A. M. J. J. & Punt, P. J. (1991), Gene transfer systems and vector development for filamentous fungi, en: Applied Molecular Genetics of Fungi, Peberdy, J. F. et al., eds., p. 1 - 28, Cambridge University Press: Cambridge), algas (Falciatore et al. (1999), Marine Biotechnology 1(3): 239 - 251), ciliados de los tipos: *Holotrichia*, *Peritrichia*, *Spirotrichia*, *Suctorina*, *Tetrahymena*, *Paramecium*, *Colpidium*, *Glaucoma*, *Platyophrya*, *Potomacus*, *Pseudocohnilembus*, *Euplotes*, *Engelmanniella*, y *Stylonychia*, especialmente del género *Stylonychia lemnae* con vectores de acuerdo con el método de transformación descrito en la solicitud PCT No. WO 98/01572 y células de plantas multicelulares (véase Schmidt, R. y Willmitzer, L. (1988), High efficiency Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana leaf and cotyledon explants, Plant Cell Rep. 583 - 586; Plant Molecular Biology and Biotechnology, C Press, Boca Raton, Florida, capítulo 6/7, p. 71 - 119 (1993); F. F. White, B. Jenes et al., Techniques for Gene Transfer, en: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. Kung y R. Wu, 128 - 43, Academic Press: 1993; Potrykus (1991), Annu. Rev. Plant Physiol. Plant Molec. Biol. 42: 205 - 225 y las referencias citadas allí) o células de mamífero. Células huésped adecuadas se discuten adicionalmente en Goeddel (1990), Gene Expression Technology: Methods in Enzymology 185, Academic

Press: San Diego, CA. Alternativamente, el vector de expresión recombinante puede ser transcrito y traducido *in vitro*, por ejemplo utilizando secuencias reguladoras del promotor T7 y polimerasa T7.

5 La expresión de polipéptidos en procariotas se lleva a cabo con mayor frecuencia con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión ya sea de polipéptidos de fusión o que no son de fusión. Los vectores de fusión añaden una cantidad de aminoácidos a un polipéptido codificado allí, por lo general al terminal amino del polipéptido recombinante, pero también al terminal C o se fusionan dentro de las regiones adecuadas en los polipéptidos. Tales vectores de fusión típicamente sirven para tres propósitos: 1) para aumentar la expresión de un polipéptido recombinante, 2) para aumentar la solubilidad de un polipéptido recombinante; y 3) para ayudar en la purificación de un polipéptido recombinante, al actuar como un ligando en la purificación por afinidad. A 10 menudo, en vectores de expresión de fusión, se introduce un sitio de escisión proteolítico en la unión de la unidad estructural de fusión y el polipéptido recombinante para permitir la separación del polipéptido recombinante a partir de la unidad estructural de fusión posteriormente a la purificación del polipéptido de fusión. Estas enzimas, y sus secuencias de reconocimiento de cognatos, incluyen al factor Xa, trombina y enteroquinasa.

15 Los vectores típicos de expresión por fusión incluyen pGEX (Pharmacia Biotech Inc.; Smith, D. B. y Johnson, K. S. (1988), *Gene* 67: 31 - 40), pMAL (New England Biolabs, Beverly, MA) y pRIT5 (Pharmacia, Piscataway, NJ), que fusionan glutatión S-transferasa (GST), polipéptido de enlazamiento de maltosa E, o polipéptido A, respectivamente, con el polipéptido recombinante objetivo. En una modalidad, la secuencia de codificación del PPSRP se clona en un vector de expresión pGEX para crear un vector que codifica un polipéptido de fusión que comprende, desde el terminal N hasta el terminal C, polipéptido X del sitio de escisión de GST-trombina. El polipéptido de fusión puede 20 ser purificado por cromatografía de afinidad utilizando resina de glutatión-agarosa. El PPSRP recombinante no fusionado a GST puede ser recuperado por la escisión del polipéptido de fusión con la trombina.

25 Los ejemplos de vectores de expresión apropiados inducibles no fusionados de *E. coli* incluyen pTrc (Amann et al. (1988), *Gene* 69: 301 - 315) y pET 11d (Studier et al. (1990), *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California 60 - 89). La expresión del gen objetivo a partir del vector pTrc se basa en la transcripción de la ARN polimerasa del huésped de un promotor de fusión trp-lac híbrido. La expresión del gen objetivo a partir del vector pET 11d se basa en la transcripción de un promotor de fusión T7 gn10-lac mediada por una ARN polimerasa viral coexpresada (T7 gn1). Esta polimerasa viral es suministrada por cepas huésped BL21 DE3) o HMS174 (DE3) de un profago A residente que alberga un gen T7 gn1 bajo el control transcripcional del promotor lacUV 5.

30 Una estrategia para maximizar la expresión del polipéptido recombinante es expresar el polipéptido en una bacteria huésped, con una capacidad mermada para escindir proteolíticamente al polipéptido recombinante (Gottesman, S. (1990), *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California 119 - 128). Otra estrategia es alterar la secuencia del ácido nucleico que va a ser insertado en un vector de expresión de modo que los codones individuales para cada aminoácido son aquellos preferentemente utilizados en la bacteria escogida para expresión, tales como *C. glutamicum* (Wada et al. (1992), *Nucleic Acids Res.* 20: 2111 - 2118). Tal 35 alteración de la secuencia de ácido nucleico de la invención puede llevarse a cabo mediante técnicas estándar de síntesis de ADN.

40 Un PPSRP de la invención puede ser expresado en un vector de expresión de levadura. Los ejemplos de vectores para expresión en levadura *S. cerevisiae* incluyen pYepSecl (Baldari, et al. (1987), *EMBO J.* 6: 229 - 234), pMFa (Kurjan y Herskowitz (1982), *Cell* 30: 933 - 943), pJRY88 (Schultz et al. (1987), *Gene* 54: 113 - 123), y pYES2 (hivitrogen Corporation, San Diego, CA). Los vectores y métodos para la construcción de vectores adecuados para su uso en otros hongos, tales como los hongos filamentosos, incluyen aquellos los detallados en: van den Hondel, C. A. M. J. J. & Punt, P. J. (1991), "Gene transfer systems and vector development for filamentous fungi", en: *Applied Molecular Genetics of Fungi*, J. F. Peberdy, et al, eds., p. 1 - 28, Cambridge University Press, Cambridge.

45 Alternativamente, un PPSRP de la invención se puede expresar en células de insecto utilizando vectores de expresión de baculovirus. Los vectores baculovirus disponibles para la expresión de polipéptidos en células de insectos cultivadas (por ejemplo, células Sf 9) incluyen la serie pAc (Smith et al. (1983), *Mol. Cell Biol.* 3: 2156 - 2165) y la serie pVL (Lucklow y Summers (1989), *Virology* 170: 31 - 39).

50 Además, un ácido nucleico para PPSRP de la invención se puede expresar en células de mamífero utilizando un vector de expresión de mamíferos. Los ejemplos de vectores de expresión de mamífero incluyen pCDM8 (Seed, B. (1987), *Nature* 329: 840) y pMT2PC (Kaufman et al. (1987), *EMBO J.* 6: 187 - 195). Cuando se utiliza en células de mamífero, las funciones de control del vector de expresión a menudo son proveídas por elementos reguladores virales. Por ejemplo, los promotores comúnmente utilizados se derivan de polioma, Adenovirus 2, citomegalovirus y virus de simio 40. Para otros sistemas de expresión adecuados tanto para células procariotas como eucariotas véanse los capítulos 16 y 17 de Sambrook, J., Fritsh, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, la 55 última edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

El vector recombinante de expresión de mamífero es capaz de dirigir la expresión del ácido nucleico preferentemente en un tipo de célula particular (por ejemplo, se utilizan elementos reguladores específicos del tejido para expresar el ácido nucleico). Los elementos reguladores específicos del tejido son conocidos en el arte. Los

ejemplos no limitantes de promotores adecuados específicos del tejido incluyen al promotor de albúmina (específico del hígado; Pinkert et al. (1987), Genes Dev. 1: 268 - 277), a los promotores específicos linfocitos (Calame y Eaton (1988), Adv Immunol 43: 235 - 275), en particular promotores de receptores de células T (Winoto y Baltimore (1989), EMBO J. 8: 729 - 733) e inmunoglobulinas (Banerji et al. (1983), Cell 33: 729 - 740; Queen y Baltimore (1983), Cell 33: 741 - 748), promotores específicos de neuronas (por ejemplo, el promotor de neurofilamentos; Byrne y Ruddle (1989), PNAS 86: 5473 - 5477), promotores específicos del páncreas (Edlund et al. (1985), Science 230: 912 - 916), y promotores específicos de la glándula mamaria (por ejemplo, el promotor del suero de la leche; patente de los Estados Unidos No. 4.873.316 y La publicación de la solicitud europea No. 264.166). Los promotores regulados por desarrollo también están incluidos, por ejemplo, los promotores del múrido Hox (Kessel y Grass (1990), Science 249: 374 - 379) y el promotor del polipéptido fetal (Campes y Tilghman (1989), Genes Dev. 3: 537 - 546).

Para la transfección estable de células de mamíferos, se sabe que, dependiendo del vector de expresión y de la técnica de transfección utilizada, sólo una pequeña fracción de las células puede integrar el ADN foráneo en su genoma. Con el fin de identificar y seleccionar estos integrantes, se introduce generalmente un gen que codifica un marcador seleccionable (por ejemplo, de resistencia a los antibióticos) en las células huésped junto con el gen de interés. Los marcadores seleccionables preferidos incluyen aquellos que confieren resistencia a fármacos, tales como G418, higromicina y metotrexato o en las plantas que confieren resistencia a un herbicida como el glifosato o al glufosinato. Las moléculas de ácido nucleico que codifican un marcador seleccionable pueden ser introducidas en una célula huésped en el mismo vector que aquel que codifica un PPSRP o pueden ser introducidas en un vector separado. Las células transfectadas de manera estable con la molécula introducida de ácido nucleico pueden ser identificadas, por ejemplo, por selección de fármacos (por ejemplo, las células que han incorporado al gen marcador seleccionable sobrevivirán, mientras que las otras células morirán).

En una modalidad preferida de la presente invención, el PPSRP de la invención se expresa en plantas y en células de plantas, tal como células de plantas unicelulares (tal como algas) (véase Falciatore et al. (1999), Marine Biotechnology 1 (3): 239 - 251 y las referencias citadas allí) y células de plantas superiores (por ejemplo, las espermatofitas, tales como las plantas de cultivo). El PPSRP de la invención puede ser "introducido" en una célula de una planta por cualquier medio, incluyendo transfección, transformación o transducción, electroporación, bombardeo de partículas, agroinfección y similares.

Los métodos adecuados para transformación o transfección de células huésped incluyendo células vegetales se pueden encontrar en Sambrook, et al. (Molecular Cloning: A Laboratory Manual, la última edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, 1989) y otros manuales de laboratorio, tales como Methods in Molecular Biology, 1995, vol. 44, Agrobacterium protocolos, ed.: Gartland y Davey, Humana Press, Totowa, Nueva Jersey. Como la tolerancia al estrés biótico y abiótico es un rasgo general que se desea que sea heredado en una gran variedad de plantas como el maíz, trigo, centeno, avena, triticale, arroz, cebada, soja, maní, algodón, colza y canola, mandioca, pimiento, girasol y tagetes, plantas solanáceas como la patata, el tabaco, la berenjena y el tomate, especies Vicia, arveja, alfalfa, plantas arbustivas (café, cacao, té), especies Salix, árboles (palma, cocotero), pastos perennes y cultivos forrajeros, estas plantas de cultivo también son plantas objetivo preferidas para modificación por ingeniería genética como una modalidad adicional de la presente invención. Los cultivos forrajeros incluyen, pero no se limitan a, heno de trigo, alpiste, bromo, heno de centeno silvestre, heno azul, dátilo aglomerado, alfalfa, salfoin, loto, trébol híbrido, trébol rojo y trébol de olor.

La transfección de un PPSRP de la invención en una planta se consigue mediante la transferencia génica mediada por *Agrobacterium*. Transformación de una planta mediada con *Agrobacterium* se puede realizar utilizando por ejemplo las cepas GV3101 (pMP90) (Koncz y Schell (1986), Mol. Gen. Genet. 204: 383 - 396) o LBA4404 (Clontech) de *Agrobacterium tumefaciens*. La transformación puede llevarse a cabo por medio de técnicas estándar de transformación y de regeneración (Deblaere et al. (1994), Nucl. Acids. Res. 13: 4777 - 4788; Gelvin, Stanton B. y Schilperoort, Robert A, Plant Molecular Biology Manual, 2. Ed. - Dordrecht: Kluwer Academic Publ, 1995 - en Sect., Ringbuc Centrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R.; Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, 1993 360 S., ISBN 0-8493-5164-2). Por ejemplo, la colza puede ser transformada a través de transformación del cotiledón o del hipocotilo (Moloney et al. (1989), Plant cell Report 8: 238 - 242; De Block et al. (1989), Plant Physiol. 91: 694 - 701). El uso de Antibióticos para *Agrobacterium* y la selección de plantas depende del vector binario y de la cepa de *Agrobacterium* utilizada para la transformación. La selección de la colza se realiza normalmente utilizando kanamicina como marcador seleccionable de la planta. La transferencia génica mediada por *Agrobacterium* para lino se puede realizar utilizando, por ejemplo, una técnica descrita por Mlynarova et al. (1994) Plant Cell Report 13: 282 - 285. Adicionalmente, la transformación de la soja se puede realizar utilizando por ejemplo una técnica descrita en la Patente Europea No. 0424 047, en la patente de los Estados Unidos No. 5.322.783, en la Patente Europea No. 0 397 687, en la patente de los Estados Unidos No. 5.376.543 o en la patente de los Estados Unidos No. 5.169.770. La transformación de maíz se puede lograr mediante bombardeo de partículas, captación de ADN mediada por polietilén glicol o mediante la técnica de fibra de carburo de silicio. (Véase, por ejemplo, Freeling y Walbot "The maize handbook" Springer Verlag: Nueva York (1993) ISBN 3-540-97826-7). Un ejemplo específico de transformación del maíz se encuentra en la patente de los Estados Unidos No. 5.990.387 y un ejemplo específico de la transformación de trigo se puede encontrar en la solicitud PCT No. WO 93/07256.

El PPSRP introducido puede ser mantenido en la célula de la planta de manera estable si se incorpora en un replicón autónomo no cromosómico o se integra en los cromosomas de la planta. Alternativamente, el PPSRP introducido puede estar presente en un vector no replicante extracromosómico y ser expresado de forma transitoria o activa transitoriamente.

5 Por ejemplo, se puede crear un microorganismo recombinante homólogo en el que el PPSRP está integrado en un cromosoma, se prepara un vector que contiene al menos una porción de un gen para PPSRP en el que se ha introducido una supresión, adición o sustitución para alterar con ello, por ejemplo, interrumpir funcionalmente, el gen para PPSRP. Preferiblemente, el gen para PPSRP es un gen para PPSRP de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u *Oryza sativa*, pero puede ser un homólogo de una planta relacionada o incluso de origen mamífero, de levadura o de insecto. Se divulga que el vector está diseñado de tal manera que, después de la recombinación homóloga, el gen endógeno para PPSRP está funcionalmente interrumpido (es decir, ya no codifica un polipéptido funcional; también denominado como un vector de desactivación). Alternativamente, el vector puede ser diseñado de tal manera que, tras la recombinación homóloga, el gen endógeno para PPSRP está mutado o bien alterado, pero aún codifica un polipéptido funcional (por ejemplo, la región reguladora secuencia arriba puede ser alterada para alterar con ello la expresión del PPSRP endógeno). Para crear un punto de mutación mediante recombinación homóloga, se pueden utilizar híbridos de ADN - ARN en una técnica conocida como quimeroplastia (Cole-Strauss et al. (1999), *Nucleic Acids Research* 27(5): 1323 - 1330 y Kmiec (1999), *Gen Therapy American Scientist*. 87(3): 240 - 247). Los procedimientos de recombinación homóloga en *Physcomitrella patens* son también bien conocidos en el arte y se contemplan para ser utilizados aquí.

20 Mientras que en el vector de recombinación homóloga, la porción alterada del gen para PPSRP está flanqueada en sus extremos 5' y 3' por una molécula adicional de ácido nucleico del gen para PPSRP para permitir que se produzca la recombinación homóloga entre el gen exógeno para PPSRP transportado por el vector y un gen endógeno para PPSRP, en un microorganismo o una planta. El molécula adicional de ácido nucleico que flanquea al PPSRP es de longitud suficiente para recombinación homóloga exitosa con el gen endógeno. Típicamente, desde varios cientos de pares de bases hasta kilobases de ADN que flanquea (tanto en el extremo 5' como en el extremo 3') están incluidos en el vector (véase, por ejemplo, Thomas, K. R., y Capecchi, M. R. (1987), *Cell* 51: 503 para una descripción de los vectores de recombinación homóloga o Strepp et al. (1998), *PNAS*, 95 (8): 4368 - 4373 para la recombinación con base en ADNc en *Physcomitrella patens*). El vector se introduce en un microorganismo o en una célula de una planta (por ejemplo, a través del ADN mediado por polietilén glicol), y se seleccionan células en las que el gen introducido para PPSRP se ha recombinado en forma homóloga con el gen endógeno para PPSRP usando técnicas conocidas en el arte.

Alternativamente, se pueden producir microorganismos recombinantes que contienen sistemas seleccionados que permiten la expresión regulada del gen introducido. Por ejemplo, la inclusión de un gen para PPSRP sobre un vector colocándolo bajo el control del operón lac permite la expresión del gen para PPSRP sólo en presencia de IPTG. 35 Tales sistemas de regulación son bien conocidos en el arte.

Ya sea que se encuentre presente en un vector no replicante extracromosómico o en un vector que está integrado en un cromosoma, el polinucleótido PPSRP preferiblemente reside en un casete de expresión de la planta. Un casete de expresión de la planta contiene preferiblemente secuencias reguladoras capaces de dirigir la expresión génica en células de la planta que están operativamente enlazadas de manera que cada secuencia pueda cumplir su función, por ejemplo, terminación de la transcripción por medio de señales de poliadenilación. Las señales de poliadenilación preferidas son aquellas que se originan a partir de t-ADN de *Agrobacterium tumefaciens*, tales como el gen 3 conocido como octopina sintasa del plásmido Ti pTiACH5 (Gielen et al. (1984), *EMBO J.* 3:835) o equivalentes funcionales del mismo, sino también son adecuados todos los otros terminadores funcionalmente activos en plantas. Como la expresión génica de la planta muy a menudo no está limitada a niveles transcripcionales, un casete de expresión de la planta preferiblemente contiene otras secuencias operativamente enlazadas como reforzadores de la traducción tales como la secuencia de sobreexcitación que contiene la secuencia líder 5' no traducida del virus del mosaico del tabaco que mejora la relación de polipéptido por ARN (Gallie et al. (1987), *Nucl. Acids Research* 15: 8693 - 8711). Los ejemplos de vectores de expresión de plantas incluyen a aquellos detallados en: Becker, D., Kemper, E., Schell, J. y Masterson, R. (1992), *New plant binary vectors with selectable markers located proximal to the left border*, *Plant Mol. Biol.* 20: 1195 - 1197, y Bevan, M. V. (1984), *Binary Agrobacterium vectors for plant transformation*, *Nucl. Acid. Res.* 12: 8711 - 8721; *Vectors for Gene Transfer in Higher Plants*; en: *Transgenic Plants, Vol. 1, Engineering and Utilization*, eds. Kung y R. Wu, Academic Press, 1993, p. 15 - 38.

La expresión génica de la planta debe estar operativamente enlazada a un promotor apropiado que confiere expresión génica en una forma oportuna, específica de la célula o del tejido. Los promotores útiles en los casetes de expresión de la invención incluyen cualquier promotor que es capaz de iniciar la transcripción en una célula de una planta. Tales promotores incluyen, pero no se limitan a, aquellos que pueden obtenerse a partir de plantas, virus de plantas, y bacterias que contienen los genes que se expresan en las plantas, tales como *Agrobacterium* y *Rhizobium*.

El promotor puede ser constitutivo, inducible, preferiblemente de etapa de desarrollo, preferiblemente de tipo celular, preferido del tejido o preferido de un órgano. Los promotores constitutivos son activos bajo la mayoría de las condiciones. Los ejemplos de promotores constitutivos incluyen a los promotores 19S y 35 de CaMV (Odell et al. (1985), Nature 373: 810 - 812), al promotor 35S de CaMV sX (Ray et al. (1987), Science 236: 1299 - 1302), al promotor Sep1, al promotor de actina de arroz (McElroy et al. (1990), Plant Cell 2: 163 - 171), al promotor de actina de *Arabidopsis*, al promotor de ubiquitina (Christensen et al. (1989), Plant Molec Biol. 18: 675 - 689.); pEmu (Last et al. (1991), Theor Appl Genet 81: 581 - 588), al promotor 35S del virus del mosaico de la escrofularia, al promotor Smas (Velten et al. (1984), EMBO J. 3: 2723 - 2730), al promotor GRPI-8, al promotor de cinamil alcohol deshidrogenasa (patente de los Estados Unidos No. 5.683.439), a los promotores del T-ADN de *Agrobacterium*, tales como manopina sintasa, nopalina sintasa, y octopina sintasa, la pequeña subunidad del promotor de ribulosa difosfato carboxilasa (ssuRUBISCO), y similares.

Los promotores inducibles son activos bajo ciertas condiciones ambientales, tales como la presencia o ausencia de un nutriente o metabolito, frío o calor, luz, ataque de patógenos, condiciones anaeróbicas, y similares. Por ejemplo, el promotor hsp80 de *Brassica* es inducido por choque térmico, el promotor PPKK es inducido por la luz, el promotor RP-1 de tabaco, *Arabidopsis* y maíz son inducibles por infección con un patógeno, y el promotor Adhl es inducido por la hipoxia y estrés por frío. La expresión del gen de la planta también puede ser facilitada a través de un promotor inducible (para revisión véase Gatz (1997), Annu. Rev. Plant Physiol. Plant Mol. Biol. 48: 89 - 108). Los promotores químicamente inducibles son especialmente adecuados si se desea que ocurra la expresión del gen de una manera específica en el tiempo. Los ejemplos de tales promotores son un promotor inducible por ácido salicílico (solicitud PCT No. WO 95/19443), un promotor inducible por tetraciclina (Gatz et al. (1992), Plant J. 2: 397 - 404) y un promotor inducible por etanol (solicitud PCT No. WO 93/21334). Además, promotores adecuados que responden a condiciones de estrés biótico o abiótico son aquellos tales como el promotor del gen para PRP1 inducible por patógeno (Ward et al. (1993), Plant. Mol. Biol. 22: 361 - 366), el promotor hsp80 inducible por calor de tomate (patente de los Estados Unidos No. 5.187.267), el promotor alfa-amilasa inducible por frío de patata (solicitud PCT No. WO 96/12814) o el promotor pinll inducible por una lesión (Patente Europea No. 375091). Para otros ejemplos de promotores inducibles por sequía, frío, y sal, tales como el promotor RD29A, consultar Yamaguchi-Shinozaki et al. (1993 Mol. Gen. Genet. 236: 331 - 340).

Los promotores preferidos del estado de desarrollo se expresan preferentemente en ciertas etapas de desarrollo. Los promotores preferidos de tejidos y órganos incluyen a aquellos que se expresan preferentemente en determinados tejidos u órganos, tales como hojas, raíces, semillas o xilema. Los ejemplos de promotores preferidos de tejidos y preferidos de órganos incluyen, pero no se limitan a promotores preferidos de la fruta, preferidos de óvulos, preferidos de tejido macho, preferido de semillas, preferidos de tegumento, preferidos de tubérculos, preferidos del tallo, preferidos del pericarpio, y preferidos de la hoja y preferidos del estigma, preferidos del polen, preferidos de las anteras, preferidos de los pétalos, preferidos del sépalo, preferidos del pedicelo, preferidos de la silicua, preferidos del tronco, preferidos de la raíz, etc. Los promotores preferidos de la semillas se expresan preferentemente durante el desarrollo y / o germinación de las semillas. Por ejemplo, los promotores preferidos de las semillas pueden ser preferidos del embrión, preferidos del endospermo y preferidos del recubrimiento de la semilla. Véase Thompson et al. (1989), BioEssays 10: 108. Los ejemplos de promotores preferidos de la semilla incluyen, pero no se limitan a la celulosa sintasa (celA), Cim1, gamma-zeína, globulina-1, zeína del maíz de 19 kDa (cZ19BI) y similares.

Otros promotores adecuados preferidos de tejidos y preferidos de órganos incluyen el promotor del gen de la napina de colza (patente de los Estados Unidos No. 5.608.152), al promotor USP de Vicia faba (Baeumlein et al. (1991), Mol Gen Genet. 225 (3): 459 - 67), al promotor de oleosina de *Arabidopsis* (solicitud PCT No. WO 98/45461), al promotor de faseolina de *Phaseolus vulgaris* (patente de los Estados Unidos No. 5.504.200), al promotor Bce4 de *Brassica* (solicitud PCT No. WO 91/13980) o al promotor B4 de legumina (LeB4; Baeumlein et al. (1992), Plant Journal, 2 (2): 233 - 9), así como promotores que confieren expresión específica de semillas en plantas monocotiledóneas como el maíz, cebada, trigo, centeno, arroz, etc. Los promotores adecuados para señalar son el promotor del gen lpt2 o del gen lpt1 de la cebada (solicitud PCT WO 95/15389 y la solicitud PCT WO 95/23230), o aquellos descritos en la solicitud PCT WO 99/16890 (promotores del gen de hordelina de la cebada, el gen de la glutelina del arroz, el gen de la orizina del arroz, el gen de la prolamina del arroz, el gen de la gliadina del trigo, el gen de la glutelina del trigo, el gen de la glutelina de la avena, el gen de la kasirina del sorgo y el gen de la secalina del centeno secalina).

Otros promotores útiles en los casetes de expresión de la invención incluyen, pero no se limitan al promotor principal de proteína de enlazamiento de clorofila alba, a promotores de histona, al promotor Ap3, al promotor de β -conglucina, al promotor de napina, al promotor de la lectina de soja, al promotor de zeína de maíz de 15 kD, al promotor de zeína de 22 kD, al promotor de zeína de 27 kD, al promotor de g-zeína, los promotores de cera, de apergaminado 1, apergaminado 2 y de bronceado, al promotor Zm13 (patente de los Estados Unidos No. 5.086.169), los promotores de la poligalacturonasa del maíz (PG) (patentes de los Estados Unidos Nos. 5.412.085 y 5.545.546) y el promotor SGB6 (patente de los Estados Unidos No. 5.470.359), así como promotores sintéticos y otros naturales.

Se puede obtener flexibilidad adicional en el control de la expresión del gen heterólogo en plantas mediante el uso de dominios de enlazamiento de ADN y elementos de respuesta a partir de fuentes heterólogas (es decir, los dominios de enlazamiento de ADN procedentes de fuentes no vegetales). Algunos ejemplos de tales dominios de

enlazamiento de ADN heterólogo incluyen los dominios de enlazamiento de ADN LexA y GAL4, Schwechheimer et al. (1998), Plant Mol Biol 36: 195 - 204. El dominio de enlazamiento de ADN LexA es parte de la proteína represora LexA de *Escherichia coli* (*E. coli*) (Brent y Ptashne (1985), Cell 43: 129 - 136 (1985)).

5 La invención proporciona además un vector de expresión recombinante que comprende una molécula de ADN para PPSRP de la invención clonada en el vector de expresión en una orientación antisentido. Es decir, la molécula de ADN está ligada operativamente a una secuencia reguladora de una manera que permite la expresión (por transcripción de la molécula de ADN) de una molécula de ARN que es antisentido a un ARNm para PPSRP. Las secuencias reguladoras operativamente enlazadas a una molécula de ácido nucleico clonada en la orientación antisentido se puede escoger para que dirijan la expresión continua de la molécula de ARN antisentido en una
10 variedad de tipos de células. Por ejemplo, se pueden escoger promotores virales y / o potenciadores, o secuencias reguladoras que dirijan la expresión constitutiva, específica del tejido o específica del tipo de célula de ARN antisentido. El vector de expresión antisentido puede estar en la forma de un plásmido recombinante, fagémido o virus atenuado en donde los ácidos nucleicos antisentido se producen bajo el control de una región reguladora de alta eficiencia. La actividad de la región reguladora puede ser determinada por el tipo de célula dentro de la cual se
15 introduce el vector. Para una discusión sobre la regulación de la expresión génica utilizando genes antisentido, véase Weintraub, H. et al. (1986), Antisense RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, Vol. 1 (1) y Mol et al. (1990), FEBS Letters 268: 427 - 430.

Un vector de expresión recombinante de la invención puede ser introducido en células huésped. Los términos "célula huésped" y "célula huésped recombinante" se utilizan aquí indistintamente. Se entiende que tales términos se refieren no sólo a la célula objetivo particular, sino que también se aplican a la progenie o a la progenie potencial de dicha célula. Debido a que ciertas modificaciones pueden ocurrir en las generaciones sucesivas, debido a mutaciones o influencias ambientales, dicha progenie no puede, de hecho, ser idéntica a la célula progenitora, pero está incluida dentro del alcance del término como se usa aquí. Una célula huésped puede ser cualquier célula procarionota o eucariota. Por ejemplo, un PPSRP puede ser expresado en células bacterianas tales como *C. glutamicum*, células de insectos, células de hongos o células de mamíferos (por ejemplo, células de ovario de hámster chino (CHO) o células COS), algas, ciliados, células de plantas, hongos u otros microorganismos como *C. glutamicum*. Otras células huésped adecuadas son conocidas por las personas capacitadas en el arte.

Una célula huésped de la invención, tal como una célula huésped procarionota o eucariota en cultivo, puede ser utilizada para producir al PPSRP (es decir, expresa al ácido nucleico que codifica PPSRP) de la invención. En consecuencia, la invención proporciona además métodos para producir al PPSRP de la invención utilizando células huésped. En primer lugar, el método comprende cultivar la célula huésped de la invención (dentro de la cual se ha introducido un vector de expresión recombinante que codifica un PPSRP, o en cuyo genoma se ha introducido un gen que codifica un PPSRP de tipo silvestre o alterado) en un medio adecuado hasta que se produce PPSRP. En segundo lugar, el método comprende además el aislamiento de los PPSRP del medio o de la célula huésped.

35 Otro aspecto de la invención se refiere al PPSRP aislado de la invención. Un polipéptido "aislado" o "purificado" o una porción biológicamente activa del mismo está libre de una parte del material celular cuando es producido mediante técnicas de ADN recombinante, o precursores químicos u otros productos químicos cuando es sintetizado químicamente. La expresión "sustancialmente libre de material celular" incluye preparaciones de PPSRP en las cuales el polipéptido es separado de algunos de los componentes celulares de las células en las que es producido
40 en forma natural o recombinantemente. La expresión "sustancialmente libre de material celular" incluye preparaciones de un PPSRP que tiene aproximadamente menos de 30% (en peso seco) de material que no es PPSRP (también denominado aquí como un "polipéptido contaminante"), más preferiblemente aproximadamente menos de 20% de material que no es PPSRP, aún más preferiblemente aproximadamente menos de 10% de material que no es PPSRP, y lo más preferible aproximadamente menos de 5% de material que no es PPSRP.

45 Cuando el PPSRP o una porción biológicamente activa del mismo es producido en forma recombinante, también está preferiblemente sustancialmente libre de medio de cultivo, es decir, que el medio de cultivo representa aproximadamente menos de 20%, más preferiblemente aproximadamente menos de 10%, y lo más preferible aproximadamente menos de 5% del volumen de la preparación del polipéptido. La expresión "sustancialmente libre de precursores químicos u otros productos químicos" incluye preparaciones de PPSRP en las que el polipéptido es
50 separado de otros precursores químicos u otros compuestos químicos que están implicados en la síntesis del polipéptido. La expresión "sustancialmente libre de precursores químicos o de otros compuestos químicos" incluye preparaciones de un PPSRP que tienen aproximadamente menos de 30% (en peso seco) de precursores químicos o productos químicos que no son PPSRP, más preferiblemente aproximadamente menos de 20% de precursores químicos o compuestos químicos que no son PPSRP, aún más preferiblemente aproximadamente menos de 10% de precursores químicos o compuestos químicos que no son PPSRP, y lo más preferible aproximadamente menos de
55 5% de precursores químicos o compuestos químicos que no son PPSRP. Los polipéptidos aislados, o porciones biológicamente activas de los mismos, carecen de polipéptidos contaminantes del mismo organismo del que se deriva el PPSRP. Típicamente, tales polipéptidos son producidos por expresión recombinante, por ejemplo, de un PPSRP de *Physcomitrella patens* en plantas diferentes a *Physcomitrella patens* o microorganismos tales como *C. glutamicum*, ciliados, algas u hongos.
60

Las moléculas de ácido nucleico, polipéptidos, homólogos de polipéptidos , polipéptidos de fusión, iniciadores, vectores, y células huésped descritas aquí pueden ser utilizados en uno o más de los métodos siguientes: identificación de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, *Oryza sativa*, y organismos relacionados; mapeo de los genomas de organismos relacionados con *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u

5 *Oryza sativa*; identificación y localización de secuencias de interés de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, u *Oryza sativa*; estudios evolutivos; determinación de las regiones de PPSRP requeridas para la función; modulación de una actividad de PPSRP; modulación del metabolismo de una o más funciones de las células; modulación del transporte transmembrana de uno o más compuestos; y modulación de la resistencia al estrés.

10 El musgo *Physcomitrella patens* representa un miembro de los musgos. Se relaciona con otros musgos tales como *Ceratodon purpureus*, que es capaz de crecer en ausencia de luz. Musgos como *Ceratodon* y *Physcomitrella* comparten un alto grado de identidad de secuencia sobre la secuencia de ADN y el nivel de polipéptido que permite el uso de selección heteróloga de moléculas de ADN con sondas que evolucionan a partir de otros musgos u organismos, permitiendo así la derivación de una secuencia de consenso adecuada para selección heteróloga o

15 anotación funcional y predicción de funciones de los genes en otras especies. La capacidad para identificar dichas funciones puede tener por lo tanto relevancia significativa, por ejemplo, la predicción de la especificidad del sustrato de las enzimas. Además, estas moléculas de ácido nucleico puede servir como puntos de referencia para el mapeo de genomas de musgo, o de genomas de organismos relacionados.

20 Las moléculas de ácido nucleico para PPSRP de la invención tienen una variedad de usos. Lo más importante, las secuencias de ácido nucleico y de aminoácidos de la presente invención pueden ser utilizadas para transformar plantas, induciendo así tolerancia a estreses tales como sequía, alta salinidad y frío. La presente invención proporciona por tanto una planta transgénica transformada por el ácido nucleico para PPSRP de la invención, en donde la expresión de la secuencia de ácido nucleico en la planta resulta en una mayor tolerancia al estrés por sequía en comparación con una variedad de tipo silvestre de la planta. La planta transgénica puede ser una

25 monocotiledónea o una dicotiledónea. La invención permite además, que la planta transgénica pueda ser seleccionada a partir de maíz, trigo, centeno, avena, triticale, arroz, cebada, soja, maní, algodón, colza, canola, mandioca, pimiento, girasol, tagetes, plantas solanáceas, papa, tabaco, berenjena, tomate, especies Vicia, arveja, alfalfa, café, cacao, té, especies *Salix*, palma oleaginosa, cocotero, hierba perenne y cultivos forrajeros, por ejemplo.

30 En particular, la presente invención describe el uso de la expresión de PP2A-1 de *Physcomitrella patens*; otras fosfatasa son PP-1 de *Physcomitrella patens*, BnPP2A-1, BnPP2A-2, y BnPP2A-3 de *Brassica napus*; GmPP2A-1, GmPP2A-2, GmPP2A -3, GmPP2A-4, y GmPP2A-5 de *Glycine max*, y OsPP2A-1, OsPP2A-2, OsPP2A-3, OsPP2A-4, y OsPP2A-5 de *Oryza sativa* para diseñar plantas tolerantes a la sequía, tolerante a la sal y / o tolerantes al frío. Esta estrategia ha sido demostrada aquí para *Arabidopsis thaliana*, colza / canola, sojas, maíz y trigo, pero su aplicación no se limita a estas plantas. En consecuencia, la invención proporciona una planta transgénica que

35 contiene PP2A-1 (SEQ ID NO: 3), en donde la planta tiene una mayor tolerancia a estrés por la sequía.

Por lo tanto, la invención provee un método para producir una planta transgénica con el ácido nucleico que codifica PPSRP de la invención, en donde la expresión del ácido nucleico en la planta resulta en una mayor tolerancia al estrés por sequía comparado con una variedad de tipo silvestre de la planta que comprende: (a) introducir en una célula de una planta un vector de expresión que comprende al ácido nucleico para PPSRP de la invención, y (b)

40 generar a partir de la célula de la planta una planta transgénica con mayor tolerancia al estrés por sequía comparado con una variedad de tipo silvestre de la planta. La célula de la planta incluye, pero no se limita a, un protoplasto, una célula que produce gametos, y una célula que regenera una planta completa. Como se utiliza aquí, el término "transgénico" se refiere a cualquier planta, célula de una planta, callo, tejido de una planta parte de una planta que contiene todo o parte de al menos un polinucleótido recombinante. En muchos casos, todo o parte de polinucleótido recombinante está integrado en forma estable en un cromosoma o elemento extracromosómico estable, de modo que es pasado a generaciones sucesivas.

La presente invención también proporciona un método para incrementar la tolerancia de una planta al estrés por sequía que comprende, incrementar la expresión del ácido nucleico que codifica PPSRP de la invención en la planta. Se puede incrementar la tolerancia de la planta al estrés por sequía como se logra por medio del incremento de la

50 expresión del PPSRP. Preferiblemente, la tolerancia de la planta al estrés por sequía se incrementa por medio del aumento de la expresión del PPSRP. La expresión del PPSRP se puede incrementar por medio de cualquier método conocido por aquellos capacitados en el arte. Los métodos para incrementar la expresión de los PPSRP se pueden utilizar en donde la planta sea transgénica o no transgénica. En los casos cuando la planta es transgénica, la planta puede ser transformada con un vector que contiene cualquiera de los ácidos nucleicos que codifican PPSRP

55 descritos anteriormente, o la planta puede ser transformada con un promotor que dirige la expresión de PPSRP nativo en la planta, por ejemplo. Tal promotor puede ser específico del tejido. Además, tal promotor puede ser regulado por desarrollo. Alternativamente, las plantas no transgénicas puede tener expresión de PPSRP nativo incrementada por inducción de un promotor nativo. La expresión de PP2A-1 (SEQ ID NO: 2) en plantas objetivo puede ser lograda, pero no se limita a, uno de los siguientes ejemplos: (a) un promotor constitutivo, (b) un promotor inducible por estrés, (c) un promotor inducido químicamente, y (d) sobreexpresión de un promotor modificado por

60

ingeniería genética por ejemplo con factores de transcripción derivados de un dedo de zinc (Greisman y Pabo (1997), Science 275: 657).

5 La transcripción del PPSRP se puede incrementar utilizando factores de transcripción derivados de un dedo de zinc (ZFP) como se describe en Greisman y Pabo, 1997 Science 275: 657 y fabricados por Sangamo Biosciences, Inc. Estos ZFP incluyen tanto un dominio de reconocimiento de ADN como un dominio funcional que causa activación o represión de un ácido nucleico objetivo tal como un ácido nucleico para PPSRP. Por lo tanto, se puede crear la activación de los ZFP para que reconozca específicamente a los promotores de PPSRP descritos anteriormente y utilizarla para incrementar la expresión de PPSRP en una planta, incrementando así la tolerancia a la sequía de la planta. La presente invención también incluye la identificación de los homólogos de PP2A-1 (SEQ ID NO: 2) en una planta objetivo así como el promotor del homólogo. La invención también provee un método para incrementar la expresión de un gen de interés dentro de una célula huésped comparado con una variedad de tipo silvestre de la célula huésped, en donde el gen de interés es transcrito en respuesta al PPSRP de la invención, que comprende: (a) transformar la célula huésped con un vector de expresión que comprende al ácido nucleico que codifica PPSRP de la invención, y (b) expresar el PPSRP de la invención dentro de la célula huésped, incrementando así la expresión del gen transcrito en respuesta al PPSRP, comparado con una variedad de tipo silvestre de la célula huésped.

10 Además de introducir las secuencias de ácido nucleico para PPSRP en plantas transgénicas, estas secuencias pueden utilizarse también para identificar un organismo como perteneciente a la especie *Physcomitrella patens*, *Brassica napus*, *Glycine max*, *Oryza sativa* o una especie próxima de la misma. Asimismo, aquéllas pueden utilizarse para identificar la presencia de *Physcomitrella patens*, *Brassica napus*, *Glycine max*, *Oryza sativa* o una especie próxima de la misma en una población mixta de microorganismos. La invención proporciona las secuencias de ácido nucleico del genes de *Physcomitrella patens*; por medio del sondeo del ADN genómico extraído de un cultivo de una población única o mixta de microorganismos bajo condiciones estrictas con una sonda que abarca una región de un gen, que es exclusiva de este organismo, se puede comprobar si está presente este organismo.

25 Adicionalmente, las moléculas de ácido nucleico y del polipéptido de la invención pueden servir como marcadores para regiones específicas del genoma. Esto tiene utilidad no sólo en el mapeo del genoma, sino también en estudios funcionales de polipéptidos de *Physcomitrella patens*. Por ejemplo, para identificar la región del genoma a la que se enlaza un polipéptido particular para enlazamiento de ADN de *Physcomitrella patens*, se podría digerir el genoma de *Physcomitrella patens*, e incubarse los fragmentos con el polipéptido de enlazamiento de ADN. Aquellos fragmentos que enlazan al polipéptido pueden sondarse adicionalmente con la molécula de ácido nucleico de la invención, preferiblemente con marcadores fácilmente detectables. El enlazamiento de tal molécula de ácido nucleico al fragmento de genoma permite la localización del fragmento con el mapa del genoma de *Physcomitrella patens*, y, cuando se realiza varias veces con enzimas diferentes, facilita una determinación rápida de la secuencia de ácido nucleico a la que se enlaza el polipéptido. Adicionalmente, la molécula de ácido nucleico de la invención pueden ser suficientemente idéntica a las secuencias de especies relacionadas de tal modo que esta molécula de ácido nucleico pueda servir como marcador para la construcción de un mapa genómico en musgos relacionados.

35 La molécula de ácido nucleico para el PPSRP de la invención es también útil para estudios evolutivos y estructurales del polipéptido. Los procesos metabólicos y de transporte en los cuales participa la molécula de la invención son utilizados por una diversidad de células procariontas y eucariontas; por comparación de la secuencia de la molécula de ácido nucleico de la presente invención con aquéllas que codifican enzimas similares de otros organismos, puede evaluarse la afinidad evolutiva de los organismos. Análogamente, tal comparación permite una evaluación de cuáles son las regiones de la secuencia que se conservan y cuáles no, lo que puede ayudar a determinar aquellas regiones del polipéptido que son esenciales para el funcionamiento de la enzima. Este tipo de determinación es valiosa para estudios de modificación del polipéptido por ingeniería genética y puede proporcionar una indicación de lo que puede tolerar el polipéptido en términos de mutagénesis sin pérdida de función.

45 La manipulación de las moléculas de ácido nucleico para la PPSRP de la invención puede dar como resultado la producción de los PPSRP que tengan diferencias funcionales con respecto a los PPSRP de tipo silvestre. Estos polipéptidos pueden mejorarse en eficiencia o actividad, pueden estar presentes en mayor número en la célula de lo que es usual, o pueden tener menor eficiencia o actividad.

50 Existen varios mecanismos por medio de los cuales la alteración del PPSRP de la invención puede afectar directamente la respuesta al estrés y/o la tolerancia al estrés. En el caso de plantas que expresan los PPSRP, el mayor transporte puede conducir a un mejor reparto de sal y/o de soluto en el tejido y los órganos de la planta. Ya sea incrementado el número o la actividad de las moléculas transportadoras que exportan moléculas iónicas desde la célula, puede ser posible afectar a la tolerancia de la célula a la sal.

55 El efecto de la modificación genética en plantas, *C. glutamicum*, hongos, algas, o ciliados sobre la tolerancia al estrés puede evaluarse cultivando el microorganismo o la planta modificados bajo condiciones inferiores a las adecuadas y analizando posteriormente las características de crecimiento y/o el metabolismo de la planta. Tales técnicas de análisis son bien conocidos por una persona capacitada en el arte, e incluyen peso seco, peso húmedo, síntesis de polipéptidos, síntesis de carbohidratos, síntesis de lípidos, tasas de evapotranspiración, rendimiento general de la planta y/o la cosecha, floración, reproducción, producción de semillas, crecimiento de raíces, tasas de

respiración, tasas de fotosíntesis, etc. (Aplicaciones de HPLC en Bioquímica en: *Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17; Rehm et al. (1993), *Biotechnology*, vol. 3, capítulo III: Recuperación y purificación de productos, páginas 469 - 714, VCH: Weinheim; Belter, P. A. et al. (1988), *Bioseparations: downstream processing for biotechnology*, John Wiley and Sons; Kennedy, J. F. y Cabral, J. M. S. (1992), *Recovery processes for biological materials*, John Wiley and Sons; Shaeiwitz, J. A. y Henry, J. D. (1988), *Biochemical Separations*, en: *Ullmann's Encyclopedia of Industrial Chemistry*, vol. B3, capítulo 11, páginas 1 - 27, VCH: Weinheim y Dechow, F. J. (1989), *Separation and purification techniques in biotechnology*, Noyes Publications).

Por ejemplo, pueden construirse y transformarse vectores de expresión de levadura que comprenden a los ácidos nucleicos divulgados aquí, o fragmentos de los mismos, en *Saccharomyces cerevisiae* utilizando protocolos estándar. Las células transgénicas resultantes pueden analizarse luego en cuanto al decaimiento o alteración al estrés por sequía, salinidad, y temperatura. Análogamente, se pueden construir y transformar vectores de expresión de plantas que comprenden a los ácidos nucleicos divulgados aquí, o fragmentos de los mismos, en una célula de una planta apropiada tal como *Arabidopsis*, soja, colza, maíz, trigo, *Medicago truncatula*, etc., utilizando protocolos estándar. Las células transgénicas y/o plantas resultantes derivadas de ellas pueden analizarse luego en cuanto al decaimiento o alteración al estrés por sequía, salinidad, y temperatura.

La modificación por ingeniería genética del gen para PPSRP de la invención puede dar también como resultado PPSRP que tengan actividades alteradas que afectan indirectamente la respuesta al estrés y/o la tolerancia al estrés de algas, plantas, ciliados u hongos u otros microorganismos tales como *C. glutamicum*. Por ejemplo, los procesos bioquímicos normales del metabolismo dan como resultado la producción de una diversidad de productos (por ejemplo, peróxido de hidrógeno y otras especies oxigenadas reactivas) que pueden interferir activamente con estos mismos procesos metabólicos (por ejemplo, se sabe que el peroxinitrito nitra las cadenas laterales de tirosina, inactivando con ello algunas enzimas que tienen tirosina en el sitio activo (Groves, J. T. (1999), *Curr. Opin. Chem. Biol.* 3(2): 226 - 235). Si bien estos productos son típicamente excretados, las células pueden ser genéticamente alteradas para transportar más productos de lo que es típico para una célula de tipo silvestre. Por medio de la optimización de la actividad del PPSRP de la invención que está involucrada en la exportación de moléculas específicas, tales como moléculas de sal, puede ser posible mejorar la tolerancia de la célula al estrés.

Adicionalmente, las secuencias divulgadas aquí, o fragmentos de las mismas, pueden utilizarse para generar mutaciones que desactivan un gen determinado en los genomas de diversos organismos, tales como bacterias, células de mamífero, células de levadura, y células de plantas (Girke, T. (1998), *The Plant Journal* 15: 39 - 48). Las células desactivadas resultantes pueden evaluarse luego respecto a su habilidad o capacidad para tolerar diversas condiciones de estrés, su respuesta a diversas condiciones de estrés, y el efecto sobre el fenotipo y/o el genotipo de la mutación. Para otros métodos de inactivación de genes véase la patente de los Estados Unidos No. 6.004.804 "Non-Chimeric Mutational Vectors" y Puttaraju et al. (1999), *Spliceosome-mediated RNA trans-splicing as a tool for gene therapy*, *Nature Biotechnology* 17: 246 - 252.

Las estrategias de mutagénesis mencionadas anteriormente para los PPSRP que dan como resultado mayor resistencia al estrés no deben considerarse limitantes; variaciones en estas estrategias serán fácilmente evidentes para un experto en la técnica. Utilizando tales estrategias, e incorporando los mecanismos divulgados en esta memoria, las moléculas de ácido nucleico y polipéptidos de la invención pueden utilizarse para generar algas, ciliados, plantas, hongos u otros microorganismos tales como *C. glutamicum* que expresan moléculas mutadas de polipéptido y de ácido nucleico para PPSRP de tal manera que la tolerancia al estrés se mejora.

También se describen anticuerpos que enlazan específicamente a un PPSRP, o una porción del mismo, tal como es codificado por un ácido nucleico descrito aquí. Los anticuerpos pueden elaborarse por muchos métodos bien conocidos (véase, por ejemplo Harlow y Lane, "Antibodies; A Laboratory Manual" Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, (1988)). En resumen, puede inyectarse antígeno purificado en un animal en una cantidad y en intervalos suficientes para provocar una respuesta inmunitaria. Los anticuerpos pueden purificarse ya sea directamente, o pueden obtenerse a partir de células del bazo del animal. Las células pueden fusionarse luego con una línea de células inmortales y escrutarse respecto a la secreción de anticuerpos. Los anticuerpos pueden utilizarse para seleccionar bibliotecas de clones de ácido nucleico en cuanto a células secretoras del antígeno. Aquellos clones positivos pueden secuenciarse posteriormente. (Véase, por ejemplo, Kelly et al. (1992), *Bio/Technology* 10: 163 - 167; Bebbington et al. (1992), *Bio/Technology* 10: 169 - 175).

Las expresiones se "enlaza selectivamente" y se "enlaza específicamente" con el polipéptido hacen referencia a una reacción de enlazamiento que es determinante de la presencia del polipéptido en una población heterogénea de polipéptidos y otros compuestos biológicos. Así, bajos las condiciones diseñadas del inmunoensayo, los anticuerpos especificados enlazados a un polipéptido particular no se enlazan en una cantidad significativa a otros polipéptidos presentes en la muestra. El enlazamiento selectivo de un anticuerpo bajo tales condiciones puede requerir de un anticuerpo que se selecciona por su especificidad por un polipéptido particular. Pueden utilizarse una diversidad de formatos de inmunoensayo para seleccionar anticuerpos que se enlazan selectivamente a un polipéptido particular. Por ejemplo, se utilizan rutinariamente inmunoensayos ELISA en fase sólida para seleccionar anticuerpos que son selectivamente inmunorreactivos con un polipéptido. Véase Harlow y Lane, "Antibodies; A Laboratory Manual" Cold Spring Harbor Publications, Nueva York, (1988), para una descripción de formatos y condiciones de inmunoensayo

que podrían utilizarse para determinar el enlazamiento selectivo.

En algunos casos, es deseable preparar anticuerpos monoclonales a partir de diversos huéspedes. Una descripción de técnicas para preparar tales anticuerpos monoclonales puede encontrarse en Stites et al. editores, "Basic and Clinical Immunology" (Lange Medical Publications, Los Altos, Calif., Cuarta Edición) y las referencias citadas allí, y en Harlow y Lane, ("Antibodies; A Laboratory Manual" Cold Spring Harbor Publications, Nueva York, (1988).

A lo largo de esta solicitud se han citado diversas publicaciones. Las divulgaciones de la totalidad de estas publicaciones y las referencias citadas en dichas publicaciones se incorporan aquí en su totalidad como referencia en esta solicitud a fin de describir más plenamente el estado de la técnica a la que pertenece esta invención.

Debe entenderse también que lo anterior se relaciona con modalidades preferidas de la presente invención y que pueden hacerse numerosos cambios en ella sin apartarse del alcance de la invención. La invención se ilustra adicionalmente por medio de los siguientes ejemplos, que no deben interpretarse en modo alguno como imposición de limitaciones en cuanto al alcance de la misma. Por el contrario, debe entenderse claramente que se puede recurrir a muchas otras modalidades, modificaciones, y equivalentes de la misma, que, después de la lectura de la presente descripción, pueden sugerirse en sí mismas a aquellos capacitados en el arte sin apartarse del espíritu de la presente invención y/o del alcance de las reivindicaciones adjuntas.

EJEMPLOS

Ejemplo 1

Crecimiento de cultivos de *Physcomitrella patens*

Para este estudio se utilizaron plantas de la especie *Physcomitrella patens* (Hedw.) B.S.G. de la colección de la sección de estudios genéticos de la Universidad de Hamburgo. Dichas plantas proceden de la cepa 16/14 recolectada por H. L. K. Whitehouse en Gransden Wood, Huntingdonshire (Inglaterra), que fue subcultivada a partir de una spora por Engel (1968, Am. J. Bot. 55, 438 - 446). La proliferación de las plantas se llevó a cabo por medio de esporas y por medio de regeneración de los gametofitos. El protonema se desarrolló a partir de la spora haploide como un cloronema rico en cloroplastos y un caulonema pobre en cloroplastos, sobre los cuales se formaron brotes al cabo de aproximadamente 12 días. Éstos crecieron para producir gametóforos que tenían anteridios y arquegonios. Después de la fertilización, se obtuvo el esporofito diploide con una pequeña seta y la cápsula de esporas, en la cual maduraron las meiosporas.

El cultivo se llevó a cabo en una cámara climatizada a una temperatura del aire de 25°C y con una intensidad de luz de 55 micromoles⁻¹m² (luz blanca; tubo fluorescente Philips TL de 65W/25) y un cambio luz/oscuridad de 16/8 horas. El musgo se modificó ya sea en cultivo líquido utilizando medio de Knop de acuerdo con Reski y Abel (1985, Planta 165: 354 - 358) o se cultivó en medio sólido de Knop utilizando agar oxoide al 1% (Unipath, Basingstoke, Inglaterra). Los protonemas utilizados para el aislamiento del ARN y del ADN se cultivaron en medios de cultivo líquidos aireados. Los protonemas fueron triturados cada 9 días y transferidos a medio de cultivo fresco.

Ejemplo 2

Aislamiento del ADN total de las plantas

Los detalles para el aislamiento del ADN total se refieren al tratamiento de un gramo de peso fresco de material de planta. Los materiales utilizados incluyen los amortiguadores siguientes: amortiguador CTAB: bromuro de N-cetil-N,N,N-trimetilamonio al 2% (p/v) (CTAB); Tris HCl 100 mM de pH 8,0; NaCl 1,4 M; EDTA 20 mM; amortiguador de N-laurilsarcosina: N-laurilsarcosina al 10% (p/v); Tris HCl 100 mM de pH 8,0; EDTA 20 mM.

Se trituró el material de la planta bajo nitrógeno líquido en un mortero para obtener un polvo fino y se lo transfirió a matraces Eppendorf de 2 ml. El material congelado de la planta se cubrió luego con una capa de 1 ml de amortiguador de descomposición (1 ml de amortiguador CTAB, 100 µl de amortiguador de N-laurilsarcosina, 20 µl de β-mercaptoetanol y 10 µl de solución de proteinasa K, 10 mg/ml) y se incubó a 60°C durante una hora con agitación continua. El homogeneizado obtenido se distribuyó en dos matraces Eppendorf (de 2 ml) y se extrajo dos veces mediante agitación con el mismo volumen de cloroformo/alcohol isoamílico (24:1). Para la separación de fases, se llevó a cabo una centrifugación a 8000 x g y a temperatura ambiente durante 15 minutos en cada caso. Se precipitó luego el ADN a -70°C durante 30 minutos utilizando isopropanol enfriado en hielo. Se sedimentó el ADN precipitado a 4°C y 10.000 g durante 30 minutos y se resuspendió en 180 µl de amortiguador TE (Sambrook et al., 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6). Para purificación adicional, se trató el ADN con NaCl (concentración final 1,2 M) y se precipitó de nuevo a -70°C durante 30 minutos utilizando dos veces el volumen de etanol absoluto. Después de una etapa de lavado con etanol al 70%, se secó el ADN y se lo recogió posteriormente en 50 µl de H₂O + ARNasa (concentración final 50 mg/ml). Se disolvió el ADN durante una noche a 4°C y se llevó a cabo posteriormente la digestión con ARNasa a 37°C durante una hora. El almacenamiento del ADN se hizo a 4°C.

Ejemplo 3

Aislamiento de ARN total y ARN poli(A)+ y construcción de una biblioteca de ADNc de *Physcomitrella patens*

Para la investigación de las transcripciones, se aislaron tanto ARN total como ARN poli(A)+. El ARN total se obtuvo de protonemas de tipo silvestre de 9 días siguiendo el método GTC (Reski et al. 1994, Mol. Gen. Genet., 244: 352 - 359). El ARN poli(A)+ se aisló utilizando Dyna Beads® (Dyna, Oslo, Noruega) siguiendo las instrucciones del protocolo del fabricante. Después de la determinación de la concentración del ARN o del ARN poli(A)+, se precipitó el ARN por medio de la adición de volúmenes de 1/10 de acetato de sodio 3 M de pH 4,6 y 2 volúmenes de etanol y se almacenó a -70°C.

Para la construcción de la biblioteca de ADNc, se logró la síntesis de la primera cadena utilizando transcriptasa inversa del Virus de la Leucemia de Múrido (Roche, Mannheim, Alemania) e iniciadores de oligo-d(T), y la síntesis de la segunda cadena por incubación con ADN polimerasa I, enzima Klenow y digestión con ARNasaH a 12°C (2 horas), 16°C (1 hora) y 22°C (1 hora). Se detuvo la reacción por medio de incubación a 65°C (10 minutos) y se la transfirió posteriormente a hielo. Se volvieron romas las moléculas de ADN bicatenario con ADN polimerasa T4 (Roche, Mannheim) a 37°C (30 minutos). Se removieron los nucleótidos por extracción con fenol/cloroformo y columnas centrífugas de Sefadex G50. Se ligaron adaptadores EcoRI (Pharmacia, Friburgo, Alemania) a los extremos del ADNc con ADN ligasa T4 (Roche, 12°C, una noche) y se fosforilaron por medio de incubación con polinucleótido quinasa (Roche, 37°C, 30 minutos). Esta mezcla fue sometida a separación sobre un gel de agarosa de bajo punto de fusión. Se eluyeron las moléculas de ADN mayores a 300 pares de bases del gel, se las extrajo con fenol, se las concentró sobre columnas Elutip-D (Schleicher y Schuell, Dassel, Alemania), se las ligó a los brazos del vector y empacó en fagos lambda ZAPII o fagos lambda ZAP-Express utilizando el Kit Gigapack Gold (Stratagene, Ámsterdam, Países Bajos) utilizando material y siguiendo las instrucciones del fabricante.

Ejemplo 4Secuenciación y anotación de la función de las EST de *Physcomitrella patens*

Se utilizaron bibliotecas de ADNc como se describe en el Ejemplo 3 para secuenciación de ADN de acuerdo con métodos estándar, y en particular, por medio del método de terminación de cadena utilizando el Kit ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer, Weiterstadt, Alemania). Se realizó una secuenciación aleatoria posteriormente a la recuperación preparativa del plásmido a partir de bibliotecas de ADNc por escisión de masa *in vivo*, nueva transformación, y posterior siembra en placa de DH10B en placas de agar (detalles del material y el protocolo de Stratagene, Ámsterdam, Países Bajos). Se preparó ADN plasmídico a partir de cultivos de *E. coli* desarrollados durante una noche en medio de Caldo de Luria que contenía ampicilina (véase Sambrook et al. 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) en un robot de preparación de ADN Qiagene (Qiagen, Hilden), de acuerdo con los protocolos del fabricante. Se utilizaron iniciadores de secuenciación con las siguientes secuencias de nucleótidos:

5' -CAGGAAACAGCTATGACC-3' SEQ DD NO: 33

5' -CTAAAGGGAACAAAAGCTG-3' SEQ DD NO: 34

5' -TGTAACGACGGCCAGT-3' SEQ DD NO: 35

Las secuencias se procesaron y anotaron utilizando el paquete de software EST-MAX proporcionado comercialmente por Bio-Max (Múnich, Alemania). El programa incorpora prácticamente todos los métodos bioinformáticos importantes para la caracterización funcional y estructural de secuencias de proteína. Los algoritmos más importantes incorporados en EST-MAX son: FASTA: Búsquedas muy sensibles en bases de datos de secuencias con estimaciones de significación estadística; Pearson W. R. (1990) Comparación rápida y sensible de secuencias con FASTP y FASTA. Methods Enzymol. 183: 63 - 98; BLAST: Búsquedas muy sensibles en bases de datos de secuencias con estimaciones de significación estadística. Altschul S. F., Gish W., Miller W., Myers E. W., y Lipman D. J. Herramienta de búsqueda de alineación local básica. Journal of Molecular Biology 215: 403 - 10; PREDATOR: Predicción de gran exactitud de la estructura secundaria de secuencias individuales y múltiples. Frishman, D. y Argos, P. (1997) 75% de exactitud en la predicción de la estructura secundaria de proteínas. Proteins, 27: 329 - 335; CLUSTAL W: Alineación de secuencias múltiples. Thompson, J. D., Higgins, D. G. y Gibson, T. J. (1994) CLUSTAL W: Mejora de la sensibilidad de la alineación de secuencias progresivas múltiples por ponderación de la secuencia, penalización por espacios en posiciones específicas y elección de la matriz por peso. Nucleic Acids Research, 22: 4673 - 4680; TMAP: Predicción de la región transmembrana a partir de secuencias alineadas en forma múltiple. Persson, B. y Argos, P. (1994) Predicción de segmentos transmembrana en proteínas utilizando alineaciones de secuencias múltiples. J. Mol. Biol. 237: 182 - 192; ALOM2: Predicción de la región transmembrana de secuencias individuales. Klein, P., Kanehisa, M., and DeLisi, C. Predicción de la función de proteínas a partir de las propiedades de las secuencias: Un análisis discriminado de una base de datos. Biochem. Biophys. Acta 787: 221 - 226 (1984). Versión 2 por Dr. K. Nakai; PROSEARCH: Detección de patrones de secuencia de proteínas PROSITE. Kolakowski L. F. Jr., Leunissen J. A. M., Smith J. E. (1992) ProSearch: búsqueda rápida de secuencias de proteínas con patrones de expresión regular relacionados con la estructura y función de las proteínas. Biotechniques 13, 919 - 921; BLIMPS: Búsquedas de similitudes contra una base de datos de bloques sin espaciamentos. J. C. Wallace y Henikoff S., (1992); PATMAT: Un programa de búsqueda y extracción para

consultas y bases de datos de secuencias, patrones y bloques, CABIOS 8: 249 - 254. Escrito por Bill Alford.

Ejemplo 5

Identificación de los ORF de *Physcomitrella patens* correspondientes a PpPP2A-1 y PpPP-1

5 Los ADNc parciales de *Physcomitrella patens* (las EST) para PpPP2A-1 (SEQ ID NO: 1) y PpPP-1 (SEQ ID NO: 4) fueron identificados en el programa de secuenciación de EST de *Physcomitrella patens* utilizando el programa EST-MAX a través del análisis BLAST. Estos clones particulares, que se encontró que codifican para proteína fosfatasa, fueron escogidos para análisis adicionales (véanse las Tablas 1 - 3 a continuación).

Tabla 1

Identificación de Marcos de Lectura Abiertos				
Nombre	Vector	Nucleótidos totales en el clon	Posición del ORF	Aminoácidos totales en el ORF
PP2A-1	PCR2.1	1279	123 - 1057	311
PP-1	PCR2.1	1014	39 - 950	303

10

Tabla 2

Grado de identidad y similitud de aminoácidos de PpPP2A-1 y otras proteínas homólogas (se utilizó el programa GCG Gap: penalización por espacio: 10; penalización por extensión del espacio: 0,1; matriz de puntuación: blosum62)					
Swiss-Prot #	Q9XGH7	Q07100	Q9SBW3	P48578	Q9XF94
Nombre de la proteína	Subunidad catalítica de la proteína serina/treonina fosfatasa PP2A	Subunidad catalítica de la proteína serina/treonina fosfatasa PP2A-3	Subunidad catalítica de la proteína serina/treonina fosfatasa PP2A-4	Subunidad catalítica de la proteína serina/treonina fosfatasa PP2A-4	Subunidad catalítica de la proteína serina/treonina fosfatasa PP2A-2
especie	Nicotiana Tabacum (tabaco común)	Arabidopsis thaliana (berro oreja de ratón)	Oryza sativa (arroz)	Arabidopsis thaliana (berro oreja de ratón)	Oryza sativa (arroz)
% de identidad	89%	88%	88%	88%	88%
% de similitud	92%	91%	89%	91%	90%

Tabla 3

Grado de identidad y similitud de aminoácidos de PpPP-1 y otras proteínas homólogas (se utilizó el programa GCG Gap: penalización por espacio: 10; penalización por extensión del espacio: 0,1; matriz de puntuación: blosum62)					
Swiss-Prot #	Q42912	Q9LHE7	Q9SX52	Q9U9A3	000743
Nombre de la proteína	Proteína serina/treonina fosfatasa	Fosfo proteína fosfatasa	Proteína F1413.5	Subunidad catalítica de la proteína fosfatasa 6	Proteína serina/treonina fosfatasa 6
especie	Malus domestica (manzana) (Malus sylvestris)	Arabidopsis thaliana (berro oreja de ratón)	Arabidopsis thaliana (berro oreja de ratón)	Dictyostelium discoideum (moho mucilaginoso)	Homo sapiens (humano)
% de identidad	91%	90%	90%	72%	68%
% de similitud	95%	95%	95%	81%	80%

Ejemplo 6

Clonación del ADNc de longitud completa de *Physcomitrella patens* que codifica para PpPP2A-1 y PpPP-1

Para aislar PpPP2A-1 (SEQ ID NO: 2) y PpPP-1 (SEQ ID NO: 5) de longitud completa de *Physcomitrella patens*, se realizó una PCR (como se describe más abajo en la Amplificación de Longitud Completa) utilizando las EST originales descritas en el Ejemplo 5 como plantillas ya que ellas eran de longitud completa. Los iniciadores utilizados para amplificación se enlistan más abajo en la Tabla 4.

Amplificación de longitud completa

Se obtuvieron clones de longitud completa correspondientes a PpPP2A-1 (SEQ ID NO: 2), y PpPP-1 (SEQ ID NO: 5) llevando a cabo la reacción en cadena de la polimerasa (PCR) con iniciadores específicos del gen (véase la Tabla 4) y la EST original como plantilla. Las condiciones para la reacción eran condiciones estándar con ADN polimerasa PWO (Roche). Se llevó a cabo la PCR de acuerdo con condiciones estándar y con los protocolos del fabricante (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2ª Edición. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY, Termociclador Biometra T3). Los parámetros para la reacción eran: cinco minutos a 94°C seguidos por cinco ciclos de un minuto a 94°C, un minuto a 50°C y 1,5 minutos a 72°C. Esto fue seguido por 25 ciclos de un minuto a 94°C, un minuto a 65°C y 1,5 minutos a 72°C.

Se extrajeron los fragmentos amplificados del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) y se los ligó dentro del vector TOPO pCR 2.1 (Invitrogen) siguiendo las instrucciones del fabricante. Se transformaron los vectores recombinantes en células Top10 (Invitrogen) utilizando condiciones estándar (Sambrook et al. 1989. Molecular Cloning, A Laboratory Manual, 2ª Edición. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY). Se seleccionaron las células transformadas en agar LB que contenía 100 µg/ml de carbenicilina, 0,8 mg de X-gal (5-bromo-4-cloro-3-indolil-β-D-galactosido) y 0,8 mg de IPTG (isopropiltio-β-D-galactosido) que crecieron durante la noche a 37°C. Se seleccionaron y utilizaron las colonias blancas para inocular 3 ml de LB líquido que contenía 100 µg/ml de ampicilina y se las dejó crecer durante una noche a 37°C. Se extrajo el ADN plasmático utilizando el kit QIAprep Spin Miniprep (Qiagen) siguiendo las instrucciones del fabricante. Los análisis de los clones subsiguientes y el mapeado de restricción se llevaron a cabo de acuerdo con técnicas estándar de biología molecular (Sambrook et al., 1989. Molecular Cloning, A Laboratory Manual, 2ª Edición. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY).

Tabla 4

Esquema e iniciadores utilizados para la clonación de clones de longitud completa				
Nombre del gen	Sitio de clonación	Método de clonación	Iniciador 5' RACE	Iniciadores de longitud completa
PpPP2A-1	XmaI/SacI	PCR del clon EST original		<p>RC001: 5'ATCCCGGGACGA CATGAGTGTGCCT CCGATATC 3' (SEQ ID NO:36)</p> <p>RC002: 5'CTGAGCTCAAGT CCCACTATAAGAA GTAGTCT 3' (SEQ ID NO:37)</p>

(continuación)

Esquema e iniciadores utilizados para la clonación de clones de longitud completa				
Nombre del gen	Sitio de clonación	Método de clonación	Iniciador 5' RACE	Iniciadores de longitud completa
PpPP-1	Xmal/Hpal	PCR del clon EST original		RC381: 5'ATCCCGGGAGGA AGGGGACTGGACA CAACGTGATG3' (SEQ ID NO:38)
				RC382: 5'GCGTTAACGCAC CATATGATGCTTTC CGGTCGTC3' (SEQ ID NO:39)

Ejemplo 7

- 5 Modificación por ingeniería genética de plantas de *Arabidopsis* tolerantes al estrés por sobreexpresión del gen PpPP2A-1 y represión del gen PpPP-1

Construcción del vector binario: pBPSJH001

- 10 El vector pLMNC53 (Mankin, 2000, Tesis de PhD, University of North Carolina) fue digerido con HindIII (Roche) y se relleno el extremo romo con enzima Klenow y dNTP 0,1 mM (Roche) de acuerdo con las instrucciones del fabricante. Se extrajo este fragmento de un gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. Se digirió luego el fragmento purificado con EcoRI (Roche) y se extrajo del gel de agarosa con un Kit de Extracción de Gel QIAquick de acuerdo con las instrucciones del fabricante. El fragmento resultante de 1,4 kilobases, el casete de gentamicina, incluía al promotor nos, al gen aacCI y el terminador g7.

- 15 Se digirió el vector pBlueScript con EcoRI y SmaI (Roche) de acuerdo con las instrucciones del fabricante. Se extrajo el fragmento resultante del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. El vector pBlueScript digerido y los fragmentos del casete de gentamicina fueron ligados con ADN Ligasa T4 (Roche) de acuerdo con las instrucciones del fabricante, uniendo los dos sitios EcoRI respectivos y uniendo el sitio de extremo romo de HindIII con el sitio SmaI.

- 20 El vector recombinante (pGMBS) fue transformado en células Top10 (Invitrogen) utilizando condiciones estándar. Se seleccionaron las células transformadas sobre agar LB que contenía 100 µg/ml de carbenicilina, 0,8 mg de X-gal (5-bromo-4-cloro-3-indolil-β-D-galactosido) y 0,8 mg de IPTG (isopropiltio-β-D-galactosido), y se las dejó crecer durante una noche a 37°C. Se seleccionaron las colonias blancas y se las utilizó para inocular 3 ml de LB líquido que contenía 100 µg/ml de ampicilina y se dejaron crecer durante una noche a 37°C. Se extrajo el ADN plasmídico utilizando el Kit Spin Miniprep QIAprep (Qiagen) siguiendo las instrucciones del fabricante. Se llevaron a cabo los análisis de los clones subsiguientes y el mapeado de restricción de acuerdo con técnicas estándar de biología molecular (Sambrook et al., 1989. Molecular Cloning, A Laboratory Manual, 2ª edición, Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.).

- 30 Tanto el vector pGMBS como el vector p1bxSuperGUS fueron digeridos con XbaI y KpnI (Roche) de acuerdo con las instrucciones del fabricante, cortando el casete de gentamicina de pGMBS y produciendo la columna vertebral del vector p1bxSuperGUS. Se extrajeron los fragmentos resultantes del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. Se ligaron estos dos fragmentos se ligaron con ADN Ligasa T4 (Roche) de acuerdo con las instrucciones del fabricante.

- 35 El vector recombinante resultante (pBPSJH001) fue transformado en células Top10 (Invitrogen) utilizando condiciones estándar. Se seleccionaron las células transformadas sobre agar LB que contenía 100 µg/ml de carbenicilina, 0,8 mg de X-gal (5-bromo-4-cloro-3-indolil-β-D-galactosido) y 0,8 mg de IPTG (isopropiltio-β-D-galactosido), y se las dejó crecer durante una noche a 37°C. Se seleccionaron las colonias blancas y se las utilizó para inocular 3 ml de LB líquido que contenía 100 µg/ml de ampicilina y se dejaron crecer durante una noche a 37°C. Se extrajo el ADN plasmídico utilizando el Kit Spin Miniprep QIAprep (Qiagen) siguiendo las instrucciones del

fabricante. Se llevaron a cabo los análisis de los clones subsiguientes y el mapeado de restricción de acuerdo con técnicas estándar de biología molecular (Sambrook et al., 1989. Molecular Cloning, A Laboratory Manual, 2ª edición, Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.).

Subclonación de PpPP2a-1 y PnPP-1 en el vector binario

- 5 Se cortaron los fragmentos que contenían las diferentes proteína fosfatasa de *Physcomitrella patens* a partir de los vectores recombinantes TOPO PCR2.1 por medio de digestión doble con enzimas de restricción (véase la Tabla 5) de acuerdo con las instrucciones del fabricante. Se cortó el fragmento posterior del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante, y se lo ligó dentro del vector binario pBPSJH001, se lo escindió con las enzimas apropiadas (véase la Tabla 5) y se desfosforiló antes de la ligación. El vector recombinante resultante contenía la fosfatasa correspondiente en la orientación sentido bajo el control del súper promotor constitutivo.
- 10

Tabla 5

Se enlistan los nombres de los dos constructos de las fosfatasas de <i>Physcomitrella patens</i> usadas para la transformación de la planta			
Gen	Enzimas utilizadas para generar un fragmento de gen	Enzimas utilizadas para restringir pBPSJH001	Constructo de vector binario
PpPP2A-1	XmaI/SacI	XmaI/SacI	pBPSSH004
PpPP-1	XmaI/HpaI	XmaI/Ec1136	pBPSSLVM018

Transformación en *Agrobacterium*

- 15 Los vectores recombinantes fueron transformados en *Agrobacterium tumefaciens* C58C1 y PMP90 de acuerdo con condiciones estándar (Hoefgen y Willmitzer, 1990).

Transformación de plantas

Se cultivaron y transformaron *Arabidopsis thaliana* del ecotipo C24 de acuerdo con condiciones estándar (Bechtold 1993, Acad. Sci. Paris. 316: 1194 - 1199; Bent et al. 1994, Science 265: 1856 - 1860).

- 20 Selección de las plantas transformadas

- Se esterilizaron semillas T1 de acuerdo con protocolos estándar (Xiong et al. 1999, Plant Molecular Biology Reporter 17: 159 - 170). Se sembraron en placa las semillas sobre medio ½ Murashige y Skoog (MS) (Sigma-Aldrich), agar al 0,6% complementado con sacarosa al 1%, 150 µg/ml de gentamicina (Sigma-Aldrich) y 2 µg/ml de benomilo (Sigma-Aldrich). Se vernalizaron las semillas en las placas durante 4 días a 4°C. Las semillas germinaron en una cámara climática a temperatura ambiente de 22°C y con intensidad de luz de 40 micromoles⁻¹m² (luz blanca; tubo fluorescente Philips TL 65W/25) y 16 horas de luz y 8 horas de oscuridad como ciclo diario. Se seleccionaron las plántulas al cabo de 14 días y se las transfirió a medio ½ MS complementado con agar al 0,6%, sacarosa al 1%, y se dejó que se recuperaran durante 5 - 7 días.
- 25

Selección por tolerancia a la sequía

- 30 Se transfirieron plántulas T1 a papel de filtro estéril seco en una caja de Petri y se dejaron desecar durante 2 horas a con una RH de 80% (humedad relativa) en una Cabina de Crecimiento Sanyo MLR-350H, micromoles⁻¹m² (luz blanca; tubo fluorescente Philips TL 65W/25). Se redujo luego la RH a 60% y se desecaron adicionalmente las plántulas durante 8 horas. Se removieron luego las plántulas y colocaron sobre placas de ½ MS con agar al 0,6% complementadas con 2 µg/ml de benomilo (Sigma-Aldrich) y se evaluaron después de 5 días. Se seleccionaron las plantas transgénicas por su mayor tolerancia a la sequía.
- 35

- Bajo condiciones de estrés por sequía, las plantas de *Arabidopsis thaliana* que sobreexpresan PpPP2A-1 mostraron una tasa de supervivencia del 62% (8 sobrevivientes de 13 plantas estresadas) para la selección por estrés; mientras que el control no transformado mostró una tasa de supervivencia menor al 6%. Las líneas transgénicas sobrevivieron al tratamiento mientras que las plantas de tipo silvestre estaban casi muertas o no sobrevivieron. Es notable que los análisis de estas líneas transgénicas fueran realizadas con plantas T1, y por lo tanto, los resultados serán mejores cuando se encuentra un expresador homocigoto fuerte.
- 40

Selección por tolerancia a la congelación

Se transportaron las plántulas a cajas de Petri que contenían MS con agar al 0,6% complementado con sacarosa al

2% y 2 µg/ml de benomilo. Después de 4 días, se incubaron las plántulas a 4°C durante 1 hora y luego se las cubrió con hielo raspado. Se colocaron luego las plántulas en una Cámara Ambiental Environmental Specialist ES2000 y se las incubó durante 3,5 horas comenzando a -1,0°C y descendiendo -1°C cada hora. Se incubaron luego las plántulas a -5,0°C durante 24 horas y se permitió que se descongelaran a 5°C durante 12 horas. Se retiró el agua y se hizo un registro de las plántulas al cabo de 5 días.

Se seleccionan las plantas transgénicas luego por su tolerancia mejorada al frío, demostrando que la expresión del transgén (o la represión en el caso de PP-1) confiere tolerancia al frío. Es notable que los análisis de estas líneas transgénicas fueran realizadas con plantas T1, y por lo tanto, los resultados serán mejores cuando se encuentra un expresador homocigoto fuerte.

10 Selección por tolerancia a la sal

Se transfirieron las plántulas a papel de filtro remojado en ½ MS y se las colocó sobre ½ MS con agar al 0,6% complementado con 2 µg/ml de benomilo la noche anterior a la selección por tolerancia a la sal. Para la selección por tolerancia a la sal, se trasladó el papel de filtro con las plántulas a pilas de papel de filtro estéril, impregnado en NaCl 50 mM, en una caja de Petri. Después de dos horas, se trasladó el papel de filtro con las plántulas a pilas de papel filtro estéril, impregnado con NaCl 200 mM, en una caja de Petri. Después de 2 horas, se trasladó el papel de filtro con las plántulas a pilas de papel de filtro estéril, impregnadas en NaCl 600 mM, en una caja de Petri. Después de 10 horas, se trasladaron las plántulas a cajas de Petri que contenían ½ MS con agar al 0,6% complementado con 2 µg/ml de benomilo. Se hizo un registro de las plántulas al cabo de 5 días. Se seleccionaron luego las plantas transgénicas por su tolerancia mejorada a la sal demostrando que la expresión del transgén (o represión en el caso de PP-1) confiere tolerancia a la sal.

Ejemplo 8

Detección del transgén PP2A-1 en las líneas transgénicas de *Arabidopsis*

Para comprobar la presencia del transgén PP2A-1 en líneas transgénicas de *Arabidopsis*, se realizó una PCR sobre ADN genómico que contamina las muestras de ARN tomadas como se describe en el Ejemplo 9 más adelante. Se utilizaron 2,5 microlitros de la muestra de ARN en una reacción PCR de 50 µl utilizando Taq ADN polimerasa (Roche Molecular Biochemicals) de acuerdo con las instrucciones del fabricante.

Se utilizó el plásmido del vector binario con cada gen clonado como control positivo, y como control negativo se utilizó ADN genómico C24 de tipo silvestre en las reacciones PCR. Se analizó la reacción PCR de 10 µl sobre gel de agarosa al 0,8%/bromuro de etidio.

30 PP2A-1: Los iniciadores utilizados en las reacciones son:

5' CCGTTATCCGAGGTCGAGGTCAGAG 3' (SEQ ID NO: 40)

5' CCAGGTCCCGAATGTGGTCCAAGGA 3' (SEQ ID NO: 41)

El programa para la PCR fue el siguiente: 35 ciclos de 1 minuto a 94°C, 30 segundos a 62°C y 1 minuto a 72°C, seguido por 5 minutos a 72°C. Se produjo un fragmento de 0,45 kilobases a partir del control positivo y las plantas transgénicas.

Se amplificaron exitosamente los transgenes a partir de líneas transgénicas T1, pero no a partir de C24 de tipo silvestre. Este resultado indica que las plantas transgénicas T1 contienen al menos una copia de los transgenes. No hubo indicios de la existencia ya sea de genes idénticos o muy similares en el control no transformado de *Arabidopsis thaliana* que podría ser amplificado por este método a partir de las plantas de tipo silvestre.

40 Ejemplo 9

Detección del ARNm del transgén PP2A-1 en líneas transgénicas de *Arabidopsis*

Se detectó la expresión del transgén utilizando RT-PCR. Se aisló el ARN total a partir de plantas tratadas por estrés utilizando un procedimiento adaptado de Verwoerd et al. (1989 NAR 17: 2362). Se recogieron muestras de hojas (50 - 100 mg) y trituraron hasta un polvo fino en nitrógeno líquido. Se resuspendió el tejido triturado en 500 µl de una mezcla 1:1 a 80°C de fenol con amortiguador de extracción (LiCl 100 mM, Tris 100 mM de pH 8, EDTA 10 mM, SDS al 1%), seguido por una breve agitación tipo vórtice para mezclar. Después de la adición de 250 µl de cloroformo, se agitó brevemente cada muestra con agitación tipo vórtice. Se centrifugaron luego las muestras durante 5 minutos a 12.000 x g. Se llevó la fase acuosa superior a un tubo Eppendorf nuevo. Se precipitó el ARN por medio de la adición de una décima parte del volumen con acetato de sodio 3 M y dos volúmenes de etanol al 95%. Se mezclaron las muestras por inversión y se dejaron en hielo durante 30 minutos. Se precipitó el ARN por centrifugación a 12.000 x g durante 10 minutos. Se retiró el sobrenadante y se secó brevemente el precipitado al aire, se resuspendió la muestra de ARN en 10 µl de agua tratada con DEPC. Para eliminar el ADN contaminante de las muestras, se trató cada una con ADNasa exenta de ARNasa (Roche) de acuerdo con las recomendaciones del fabricante. Se sintetizó ADNc a

partir del ARN total utilizando el Sistema de Síntesis Superscript para la Primera Hebra para RT-PCR (Gibco-BRL) siguiendo las recomendaciones del fabricante.

5 Se realizó la amplificación por PCR de un fragmento específico del gen a partir del ADNc sintetizado utilizando Taq ADN polimerasa (Roche) y los iniciadores específicos del gen (véase el Ejemplo 8 para los iniciadores) en la reacción siguiente: amortiguador PCR 1X, MgCl₂ 1,5 mM, 0,2 μM de cada iniciador, los dNTP 0,2 μM, 1 unidad de polimerasa, y 5 μl de ADNc de la reacción de síntesis. La amplificación se realizó bajo las condiciones siguientes: Desnaturalización, 95°C, 1 minuto; apareamiento, 62°C, 30 segundos; extensión, 72°C, 1 minuto, 35 ciclos; extensión, 72°C, 5 minutos; mantenimiento, 4°C, hasta terminar. Se corrieron los productos PCR sobre un gel de agarosa al 1%, se tiñeron con bromuro de etidio, y se visualizaron bajo luz UV utilizando el sistema de documentación en gel Quantity-One (Bio-Rad).

10 Se detectó la expresión de los transgenes en la línea transgénica T1. Este resultado indicó que los transgenes se expresan en las líneas transgénicas y demostró que su producto génico mejoró la tolerancia de la planta al estrés en la línea transgénica. De acuerdo con la exposición anterior, no se pudo detectar por este método expresión alguna de genes endógenos idénticos o muy similares. Estos resultados están de acuerdo con los datos del Ejemplo 8.

15 **Ejemplo 10**

Modificación por ingeniería genética de plantas de soja tolerantes al estrés por sobreexpresión del gen PP2A-1 y represión del gen PP-1

Se utilizan los constructos pBPSSH004 y pBPSLVM018 para transformar soja como se describe a continuación.

20 Se esteriliza la superficie de semillas de soja con etanol al 70% durante 4 minutos a la temperatura ambiente con agitación continua, seguido por Clorox al 20% (v/v) complementado con Tween al 0,05% (v/v) durante 20 minutos con agitación continua. Luego, se lavaron las semillas 4 veces con agua destilada y se las colocó sobre papel de filtro estéril humedecido en una caja de Petri a temperatura ambiente durante 6 a 39 horas. Se pelaron las cubiertas de las semillas, y se desprendieron los cotiledones del eje del embrión. Se examinó el eje del embrión para asegurarse de que no se había deteriorado la región meristemática. Se recolectaron los ejes de embrión cortados en una caja de Petri estéril medio abierta y se secaron al aire hasta un contenido de humedad menor al 20% (peso fresco) en una caja de Petri sellada hasta su uso posterior.

30 Se preparó un cultivo de *Agrobacterium tumefaciens* a partir de una sola colonia en medio LB sólido más los antibióticos apropiados (por ejemplo 100 mg/l de estreptomina, 50 mg/l de kanamicina) seguido por crecimiento de la única colonia en medio LB líquido hasta una densidad óptica a 600 nm de 0,8. A continuación, se centrifuga el cultivo de bacterias a 7000 rpm durante 7 minutos a temperatura ambiente, y se resuspende en medio MS (Murashige y Skoog, 1962) complementado con acetosiringona 100 μM. Se incuban los cultivos bacterianos en este medio de preinducción durante 2 horas a temperatura ambiente antes de su utilización. Los ejes de los embriones de las semillas cigóticas de soja con un contenido aproximado de 15% de humedad se embeben durante 2 horas a temperatura ambiente con el cultivo preinducido en suspensión de *Agrobacterium*. Se remueven los embriones del cultivo en imbibición y se transfieren a cajas de Petri que contienen medio MS sólido complementado con sacarosa al 2% y se incuba durante 2 días en la oscuridad, a temperatura ambiente.

40 Alternativamente, se colocan los embriones sobre la parte superior de papel de filtro estéril húmedo (medio MS líquido) en una caja de Petri y se incuba bajo las mismas condiciones descritas anteriormente. Después de este periodo, se transfieren los embriones ya sea a medio MS sólido o líquido complementado con 500 mg/L de carbenicilina o 300 mg/l de cefotaxima para matar las agrobacterias. Se utiliza el medio líquido para humedecer el papel de filtro estéril. Se incuban los embriones durante 4 semanas a 25°C, bajo 150 μmol m⁻²s⁻¹ y un periodo de luz de 12 horas. Una vez que las plántulas de semillero han producido raíces, se las transfiere a suelo Metromix estéril. Se lava el medio de las plantas *in vitro* antes de transferir las plantas al suelo. Se mantienen las plantas bajo una cubierta de plástico durante 1 semana para favorecer el proceso de aclimatación. Se transfieren luego las plantas a una sala de crecimiento donde se las incuba a 25°C, bajo 150 μmol m⁻²s⁻¹ de intensidad lumínica y un periodo de luz de 12 horas aproximadamente durante 80 días.

45 Se seleccionan luego las plantas transgénicas en cuanto a su tolerancia mejorada a la sequía, la sal y/o el frío, de acuerdo con el método de selección descrito en el Ejemplo 7, demostrando que la expresión del transgén confiere tolerancia al estrés.

50 **Ejemplo 11**

Modificación por ingeniería genética de plantas de Colza/Canola tolerantes al estrés por sobreexpresión del gen PP2A-1 y represión del gen PP-1

Se utilizan los constructos pBPSSH004 y pBPSLVM018 para transformar colza/canola como se describe a continuación.

55 El método de transformación de plantas descrito en esta memoria es también aplicable a *Brassica* y otros cultivos.

Se esteriliza la superficie de semillas de canola con etanol al 70% durante 4 minutos a temperatura ambiente con agitación continua, seguido por Clorox al 20% (v/v) complementado con Tween al 0,05% (v/v) durante 20 minutos, a temperatura ambiente con agitación continua. A continuación, se lavan 4 veces las semillas con agua destilada y se dejan sobre papel de filtro estéril húmedo en una caja de Petri a temperatura ambiente durante 18 horas. Luego se remueven las cubiertas de las semillas y se secan las semillas al aire durante la noche en una caja de Petri estéril medio abierta. Durante este período, las semillas pierden aproximadamente un 85% de su contenido de agua. Se almacenan luego las semillas a temperatura ambiente en una caja de Petri sellada hasta su uso posterior. Los constructos de ADN y la imbibición de los embriones son como se describe en el Ejemplo 10. Se analizan muestras de las plantas transgénicas primarias (T0) por PCR para confirmar la presencia de T-ADN. Estos resultados se confirman por hibridación tipo Southern en la cual se somete a electroforesis al ADN sobre un gel de agarosa al 1% y se lo transfiere a una membrana de nylon cargada positivamente (Roche Diagnostics). Se utiliza el Kit de Síntesis de Sonda DIG para PCR (Roche Diagnostics) para preparar una sonda marcada con digoxigenina por PCR, y se la utiliza de acuerdo con las recomendaciones del fabricante.

Se seleccionan luego las plantas transgénicas por su tolerancia mejorada al estrés de acuerdo con el método de selección descrito en el Ejemplo 7, demostrando que la expresión del transgén confiere tolerancia a la sequía.

Ejemplo 12

Modificación por ingeniería genética de plantas de maíz tolerantes al estrés por sobreexpresión del gen PP2A-1 y represión del gen PP-1

Se utilizan los constructos pBPSSH004 y pBPSLVM018 para transformar maíz como se describe a continuación.

La transformación de maíz (*Zea Mays L.*) se realiza con el método descrito por Ishida et al. (1996, Nature Biotech. 14745 - 50). Se cultivan conjuntamente los embriones inmaduros con *Agrobacterium tumefaciens* que transportan vectores "súper binarios", y se recuperan las plantas transgénicas a través de organogénesis. Este procedimiento proporciona una eficiencia de transformación comprendida entre 2,5% y 20%. Se seleccionan luego las plantas transgénicas por su tolerancia mejorada a la sequía, la sal y/o el frío de acuerdo con el método de selección descrito en el Ejemplo 7, demostrando que la expresión del transgén confiere tolerancia al estrés.

Ejemplo 13

Modificación por ingeniería genética de plantas de trigo tolerantes al estrés por sobreexpresión del gen PP2A-1 y represión del gen PP-1

Se utilizan los constructos pBPSSH004 y pBPSLVM018 para transformar trigo como se describe a continuación.

La transformación de trigo se realiza con el método descrito por Ishida et al. (1996, Nature Biotech. 14745 - 50). Se cultivan conjuntamente los embriones inmaduros con *Agrobacterium tumefaciens* que transportan vectores "súper binarios", y se recuperan las plantas transgénicas a través de organogénesis. Este procedimiento proporciona una eficiencia de transformación comprendida entre 2,5% y 20%. Se seleccionan luego las plantas transgénicas por su tolerancia mejorada al estrés de acuerdo con el método de selección descrito en el Ejemplo 7, demostrando que la expresión del transgén confiere tolerancia al estrés.

Ejemplo 14

Monitoreo de los cambios en la concentración de ARNm de PpPP-1 en cultivos de *Physcomitrella patens* tratados contra el frío

Preparación en portaobjetos de un microarreglo de ADN

Se realizó la amplificación por PCR en placas de 96 pozos a partir de los EST seleccionados de *Physcomitrella patens* clonados en el vector pBluescript. Se empleó el juego de amortiguadores para PCR (Boehringer Mannheim) para la reacción PCR. Cada mezcla de reacción PCR contiene 10 µl de amortiguador para PCR sin MgCl₂, 10 µl de MgSO₄, 3 µl del iniciador hacia delante SK (MWG-Biotech, secuencia: 5'-CGCCAAGCGCGCAATTAACCCTCACT-3', SEQ ID NO: 42), 3 µl del iniciador inverso SK (MWG-Biotech, secuencia: 5-GCGTAATACGACTCACTATAGGG CGA-3', SEQ ID NO: 43), 2 µl de dNTP, 1 µl de Taq ADN polimerasa (Roche), 72 µl de agua y 1 µl de plantilla de ADN. Después de desnaturalización a 95°C durante 3 minutos, se realizaron las reacciones PCR con 35 ciclos de 3 etapas consecutivas que incluían desnaturalización a 95°C durante 45 segundos, apareamiento a 63°C durante 45 segundos, y alargamiento a 72°C durante 60 segundos. El último alargamiento se realizó a 72°C durante 10 minutos. Se purificaron luego los productos PCR con el Kit de purificación de PCR QIAquick (Qiagen, Inc.), se eluyeron con agua y se midió la concentración de ADN a 260 nm en un espectrofotómetro.

Se secaron de 2 a 5 µg de cada producto PCR y se disolvieron en 50 µl de DMSO. Los productos PCR fueron luego formateados desde placas de 96 pozos hasta placas de 384 pozos para la impresión. Se empleó el formador de arreglos Microarray GenIII (Molecular Dynamics) para imprimir los productos PCR en portaobjetos para microarreglos (Molecular Dynamics) con el formato recomendado por el fabricante. Los puntos impresos tenían un

diámetro de aproximadamente 290 μm y estaban espaciados aproximadamente 320 μm de centro a centro. Después de la impresión, se dejó el portaobjetos en la cámara exenta de polvo durante 1 hora hasta secarlo. Se llevó a cabo el entrecruzamiento con UV con 600 $\mu\text{J}/\text{mm}$. Los portaobjetos entrecruzados estaban listos para hibridación y se almacenaron en cámaras oscuras y secas.

5 Síntesis de la sonda de microarreglo

Se extrajo el ARN total de cultivos de *Physcomitrella patens* tratados contra el frío (12 horas a 4°C en la oscuridad) siguiendo el método de extracción de ARN descrito en Ausubel et al. (Curr. Prot. in Mol. Biol. 1987, J. Wiley and Sons, Nueva York). Se aplicó el kit midi de ARNm Oligotex (Qiagen Inc.) para aislar el ARNm del ARN total con un enfoque que combina tanto un protocolo de lotes como uno estándar recomendados por el fabricante. Después de enlazar el ARN total con Oligotex, se centrifugó la muestra a 14000 x g para separar la mezcla Oligotex:ARNm con la fase líquida en lugar de pasarla a través de una columna. Después de cuatro lavados con amortiguador OW2 como se describe en el protocolo de lotes, se resuspendió el Oligotex:ARNm en 400 μl de OW2 y se recogió luego por medio de la columna como en el protocolo estándar. Se eluyó el ARNm de acuerdo con el protocolo estándar.

Se sintetizaron sondas de ADNc marcadas con Cy3 y Cy5 a partir de ARNm con el Sistema de Escogencia Superscript para síntesis de ADNc (Gibco BRL). Se mezclaron tanto el iniciador oligo-(dT)₂₅ (Genosys Biotechnologies) como el iniciador Nonamer (Amersham Pharmacia Biotech) con ARNm para alcanzar un volumen total de 20 μl . Se calentó primero la mezcla a 70°C durante 10 minutos y se dejó luego a temperatura ambiente durante 15 minutos antes de transferirla sobre hielo. Una vez que la muestra está sobre hielo, se añaden 8 μl de Amortiguador de Síntesis de la Primera Hebra, 4 μl de DTT 0,1 M, 2 μl de dNTP (Amersham Pharmacia Biotech), 2 μg de Cy3- o Cy5-dCTP (Amersham Pharmacia Biotech), 2 μl de Inhibidor de ARNasa (Gibco BRL) y 2 μl de Transcriptasa Inversa Superscript II. Se llevó a cabo la síntesis de la primera hebra a 42°C durante 8 horas y se calentó luego la mezcla a 94°C durante tres minutos después de la reacción.

Después de la síntesis de la primera hebra, se añadieron 4 μl de hidróxido de sodio 2,5 M a la reacción y se incubó la mezcla a 37°C durante diez minutos. Se añadieron luego a cada reacción 20 μl de MOPS 2M (pH 5,0) y 500 μl de amortiguador PB (Qiagen Inc.). Se purificó luego la sonda por medio del Kit de Purificación para PCR QIAquick (Qiagen Inc.) con el protocolo proporcionado por el fabricante.

Hibridación y lavados de microarreglos de ADNc

Se mezclaron las sondas purificadas marcadas con Cy3 y Cy5 y se secaron al vacío para obtener un volumen final de 9 μl . Se añadieron luego a las sondas de ADNc 9 μl de Solución de Hibridación de Microarreglos (Amersham Pharmacia Biotech) y 18 μl de Formamida (Sigma) hasta obtener un volumen final de 36 μl . Se aplicó la muestra al portaobjetos impreso de microarreglos que se cubrió luego con un cubreobjetos limpio y libre de polvo sin aire atrapado. Se llevó a cabo la hibridación en una cámara de hibridación a 42°C durante 16 a 20 horas. Después de la hibridación, se lavaron 2 veces los portaobjetos con 0,5 X SSC, SDS al 0,2% a temperatura ambiente durante 5 minutos y 15 minutos. Se realizaron 2 lavados bajo condiciones estrictas con 0,25 X SSC, SDS al 0,1% a 55°C durante 10 y 30 minutos respectivamente. Después de los lavados, se enjuagaron brevemente los portaobjetos con agua Millipore y se secaron bajo nitrógeno comprimido.

Escaneo y análisis de los datos del microarreglo

Se escanearon los microarreglos de ADNc utilizando el Escáner de Microarreglos GenIII (Molecular Dynamics) equipado con dos canales láser. Las imágenes escaneadas fueron vistas y ajustadas primeramente en el software ImageQuant (Molecular Dynamics) y luego analizadas por medio del software ArrayVision (Molecular Dynamics). Se normalizó la intensidad de señal para cada punto por medio del software ArrayVision (Molecular Dynamics) y se la transfirió a Excel (Microsoft). Los datos obtenidos se normalizaron dividiendo la diferencia del valor de la intensidad y la señal de fondo y la diferencia del valor de control y la señal de fondo. Se obtuvo luego la relación por medio de la división de los datos normalizados. El nivel de transcripto de PpPP-1 disminuyó 2 veces comparado con el control no tratado.

Ejemplo 15

Identificación de genes idénticos y heterólogos

Pueden utilizarse secuencias génicas para identificar genes idénticos o heterólogos a partir de ADNc o de bibliotecas genómicas. Se pueden aislar genes idénticos (por ejemplo, clones de ADNc de longitud completa) a través de hibridación de ácido nucleico utilizando por ejemplo bibliotecas de ADNc. Dependiendo de la abundancia del gen de interés, se siembran en placas 100.000 hasta 1.000.000 de bacteriófagos recombinantes y se los transfiere a membranas de nylon. Después de desnaturalización con álcali, se inmoviliza el ADN sobre la membrana mediante, por ejemplo, entrecruzamiento con UV. La hibridación se lleva a cabo en condiciones de alta rigurosidad. En solución acuosa, se lleva a cabo la hibridación y el lavado con una fuerza iónica de NaCl 1M y una temperatura de 68°C. Se generan sondas de hibridación mediante, por ejemplo, marcación de la transcripción de la muesca radiactiva (³²P) (High Prime, Roche, Mannheim, Alemania). Las señales se detectan por medio de autorradiografía.

Los genes parcialmente idénticos o heterólogos que están relacionados pero no son idénticos pueden ser identificados de una forma análoga al procedimiento anteriormente descrito utilizando condiciones de hibridación y de lavado de baja rigurosidad. Para hibridación acuosa, la fuerza iónica se mantiene normalmente con NaCl 1 M mientras se reduce progresivamente la temperatura desde 68 a 42°C.

- 5 El aislamiento de las secuencias génicas con homologías (o identidad/ semejanza de secuencia) solamente en un dominio distinto de (por ejemplo 10 - 20 aminoácidos) puede llevarse a cabo por medio del uso de sondas sintéticas de oligonucleótidos marcadas en forma radioactiva. Los oligonucleótidos marcados en forma radioactiva se preparan por fosforilación del extremo 5' de dos oligonucleótidos complementarios con polinucleótido quinasa T4. Los oligonucleótidos complementarios se aparean y se ligan para formar concatémeros. Los concatémeros bicatenarios se someten luego a marcación radioactiva mediante, por ejemplo, transcripción de la muesca. La hibridación se lleva a cabo normalmente en condiciones de baja rigurosidad utilizando concentraciones altas de oligonucleótidos.

Solución para hibridación de oligonucleótidos:

6 X SSC

fosfato de sodio 0,01 M

- 15 EDTA 1 mM (pH 8)

SDS al 0,5%

100 µg/ml de ADN desnaturalizado de esperma de salmón

0,1% de leche desengrasada seca

- 20 Durante la hibridación, la temperatura se reduce gradualmente hasta 5 - 10°C por debajo del valor T_m estimado de los oligonucleótidos o hasta temperatura ambiente seguido por etapas de lavado y autorradiografía. El lavado se realiza con baja rigurosidad tal como tres etapas de lavado utilizando 4 X SSC. Los detalles adicionales son descritos descritos por Sambrook, J. et al. (1989), "Molecular Cloning: A Laboratory Manual" Cold Spring Harbor Laboratory Press, o Ausubel, F. M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.

Ejemplo 16

- 25 Identificación de genes idénticos por medio de selección de bibliotecas de expresión con anticuerpos

Pueden utilizarse clones de ADNc para producir polipéptido recombinante, por ejemplo en *E. coli* (por ejemplo el sistema QIAexpress pQE de Qiagen). Los polipéptidos recombinantes son luego normalmente purificados por afinidad mediante cromatografía de afinidad Ni-NTA (Qiagen). Se utilizan luego polipéptidos recombinantes para producir anticuerpos específicos, por ejemplo utilizando técnicas estándar para inmunización de conejos. Los anticuerpos se purifican por afinidad utilizando una columna Ni-NTA saturada con el antígeno recombinante como ha sido descrito por Gu et al., 1994 BioTechniques 17: 257 - 262. El anticuerpo puede utilizarse luego para seleccionar bibliotecas de expresión de ADNc para identificar genes idénticos o heterólogos a través de selección inmunológica (Sambrook, J. et al. (1989), "Molecular Cloning: A Laboratory Manual" Cold Spring Harbor Laboratory Press, o Ausubel, F. M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.

- 35 **Ejemplo 17**

Mutagénesis *in vivo*

- 40 La mutagénesis *in vivo* de microorganismos puede realizarse mediante el paso de ADN plasmídico (u otro vector) a través de *E. coli* u otros microorganismos (por ejemplo, *Bacillus* spp. o levaduras tales como *Saccharomyces cerevisiae*) que están afectados en su capacidad para mantener la integridad de su información genética. Las cepas mutantes típicas tienen mutaciones en los genes para el sistema de reparación del ADN (por ejemplo, mutHLS, mutD, mutT, etc.; para referencia, véase Rupp W. D. (1996) Mecanismos de reparación de ADN, en: *Escherichia coli* y *Salmonella*, p. 2277 - 2294, ASM: Washington). Tales cepas son bien conocidas por los expertos en la técnica. El uso de dichas cepas se ilustra, por ejemplo, en Greener, A. y Callahan, M. (1994) Strategies 7: 32 - 34. La transferencia de moléculas de ADN mutadas en plantas se lleva a cabo preferiblemente después de selección y ensayo en microorganismos. Las plantas transgénicas se generan de acuerdo con diversos ejemplos dentro de la sección de ejemplos de este documento.

Ejemplo 18

Análisis *in vitro* de la función de los genes de *Physcomitrella* en organismos transgénicos

- 50 La determinación de las actividades y de los parámetros cinéticos de las enzimas está bien establecida en el arte. Los experimentos para determinar la actividad de cualquier enzima alterada dada debe adaptarse a la actividad específica de la enzima de tipo silvestre, lo cual se encuentra dentro de las capacidades de un experto en el arte.

- Revisiones acerca de enzimas en general, así como detalles específicos concernientes a estructura, cinética, principios, métodos, aplicaciones y ejemplos para la determinación de muchas actividades de las enzimas pueden encontrarse, por ejemplo, en las referencias siguientes: Dixon, M., y Webb, E. C., (1979) *Enzymes*. Longmans: London; Fersht, (1985) *Enzyme Structure and Mechanism*. Freeman: Nueva York; Walsh, (1979) *Enzymatic Reaction Mechanisms*. Freeman: San Francisco; Price, N. C., Stevens, L. (1982) *Fundamentals of Enzymology*. Oxford Univ. Press: Oxford; Boyer, P. D., ed. (1983) *The Enzymes*, 3rd ed. Academic Press: Nueva York; Bisswanger, H, (1994) *Enzymkinetik*, 2nd ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H. U., Bergmeyer, J., Graßl, M., eds. (1983 - 1986) *Methods of Enzymatic Analysis*, 3rd ed., vol. I - XII, Verlag Chemie: Weinheim; y Ullmann's Encyclopedia of Industrial Chemistry (1987) vol. A9, *Enzymes*. VCH: Weinheim, p. 352 - 363.
- 10 La actividad de los polipéptidos que se enlazan a ADN puede medirse por medio de diferentes métodos bien establecidos, tales como ensayos de desplazamiento de bandas de ADN (denominados también ensayos de retardo en gel). El efecto de tales polipéptidos sobre la expresión de otras moléculas puede medirse utilizando ensayos de genes reporteros (tales como aquel descrito en Kolmar, H. et al . (1995) *EMBO J* . 14: 3895 - 3904 y las referencias citadas allí). Los sistemas de ensayo de genes reporteros son bien conocidos y están bien establecidos para
- 15 aplicaciones tanto en células procariotas como eucariotas, utilizando enzimas tales como β -galactosidasa, polipéptido fluorescente verde, y varias otras.

La determinación de la actividad de los polipéptidos de transporte de membrana puede realizarse de acuerdo con técnicas tales como las descritas en Gennis, R. B. Pores, Channels and Transporters, en *Biomembranes, Molecular Structure and Function*, páginas 85 - 137, 199 - 234 y 270 - 322, Springer, Heidelberg (1989).

20 Ejemplo 19

Purificación del producto deseado a partir de organismos transformados

- La recuperación del producto deseado a partir de material de la planta (es decir, *Physcomitrella patens* o *Arabidopsis thaliana*), hongos, algas, ciliados, células de *C. glutamicum*, u otras células bacterianas transformadas con las secuencias de ácido nucleico descritas aquí, o el sobrenadante de los cultivos descritos más arriba puede realizarse por diversos métodos bien conocidos en el arte. Si el producto deseado no es secretado por las células, puede ser recolectado a partir del cultivo por medio de centrifugación a baja velocidad, las células pueden ser lisadas por medio de técnicas estándar, tales como fuerza mecánica o sonicación. Los órganos de las plantas pueden separarse mecánicamente de otros tejidos u órganos. Después de la homogeneización, se separan los residuos celulares por centrifugación, y la fracción sobrenadante que contiene las proteínas solubles se retiene para
- 25 purificación adicional del compuesto deseado. Si el producto es secretado por las células deseadas, entonces se separan las células del cultivo por centrifugación a baja velocidad, y se retiene la fracción sobrenadante para purificación adicional.

- La fracción sobrenadante de cualquier método de purificación se somete a cromatografía con una resina adecuada, en la cual la molécula deseada es retenida ya sea en una resina cromatográfica en tanto que muchas de las impurezas de la muestra no son retenidas o, por el contrario, las impurezas son retenidas por la resina mientras que la muestra no lo es. Tales etapas de cromatografía pueden repetirse en caso necesario, utilizando la misma o diferentes resinas cromatográficas. Un experto en la técnica estaría plenamente capacitado para la selección de resinas cromatográficas apropiadas y en su aplicación más eficaz para una molécula particular que va a ser purificada. El producto purificado puede concentrarse por medio de filtración o ultrafiltración, y almacenado a una
- 35 temperatura a la cual se maximiza la estabilidad del producto.

- Existe una amplia gama de métodos de purificación conocidos en el arte y el método de purificación anterior no pretende constituirse en limitante. Dichas técnicas de purificación están descritas, por ejemplo, en Bailey J. E. & Ollis, D. F. *Biochemical Engineering Fundamentals*, McGraw-Hill: Nueva York (1986). Adicionalmente, la identidad y pureza de los compuestos aislados puede evaluarse por medio de técnicas estándar en el arte. Éstas incluyen cromatografía líquida de alta resolución (HPLC), métodos espectroscópicos, métodos de coloración, cromatografía en capa delgada, NIRS, ensayo enzimático, o microbiológicamente. Tales métodos de análisis se revisan en: Patek et al., 1994 *Appl. Environ. Microbiol.* 60: 133 - 140; Malakhova et al., 1996 *Biotekhnologiya* 11: 27 - 32; y Schmidt et al., 1998 *Bioprocess Engineer.* 19: 67 - 70. Ullmann's Encyclopedia of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89 - 90, p. 521 - 540, p. 540 - 547, p. 559 - 566, 575 - 581 y p. 581 - 587; Michal, G. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A. et al. (1987) *Aplicaciones de HPLC en Bioquímica en: Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17.
- 45
- 50

Ejemplo 20

Identificación de ortólogos de PpPP₂A-1

- Se identificaron los ADNc parciales de *Brassica napus*, *Glycine max*, y *Oryza sativa* (las EST) para BnPP2A-1, BnPP2A-2, BnPP2A-3, GmPP2A-1, GmPP2A-2, GmPP2A-3, GmPP2A-4, GmPP2A-5, OsPP2A-1, OsPP2A-2, OsPP2A-3, OsPP2A-4, y OsPP2A-5 en una base de datos licenciada en forma privada por medio de la búsqueda para las EST con similitud con la secuencia de nucleótidos de PP2A-1 de *Physcomitrella patens*, y se secuenciaron
- 55

ES 2 381 031 T3

los ADNc de longitud completa de las EST identificadas. Estos clones particular clones fueron escogidos para análisis adicionales (véase la Tabla 6 a continuación).

Tabla 6

Nombre del gen	ID del clon de Hyseq	Fuente de Hyseq	posición para la traducción	vs. Identidad/ similitud con PpPP2A-1	vs. identidad/ similitud pública
BnPP2A-1	BN48706417	canola	150 - 1067	78% / 86%	AAG52565:96%/97%
BnPP2A-2	BN51288093	canola	90 - 1046	42% / 54%	AAC39460: 90%/95%
BnPP2A-3	BN51387173	canola	155 - 1036	40% / 54%	BAB09762: 92%/92%
OsPP2A-1	OS41502678	arroz	140 - 1081	88% / 90%	AAD048068: 99%/99%
OsPP2A-2	OS32806943	arroz	173 - 1093	79% / 86%	AAC72838: 96%/98%
OsPP2A-3	OS35083313	arroz	208 - 1155	42% / 57%	AAA33545: 94%/98%
OsPP2A-4	OS33003814	arroz	58-978	88% / 90%	AAD22116: 100%/100%
OsPP2A-5	OS34738749	arroz	104 - 1021	78% / 87%	BAA92697: 78%/87%

(continuación)

Nombre del gen	ID del clon de Hyseq	Fuente de Hyseq	posición para la traducción	vs. Identidad/ similitud con PpPP2A-1	vs. identidad/ similitud pública
GmPP2A-1	GM50770660	soja	97 - 1005	55% / 67%	CAA87385: 55%/67%
GmPP2A-2	GM48922444	soja	28 - 975	41% / 57%	BAA92244: 41%/57%
GmPP2A-3	GM50131069	soja	206 - 1144	88% / 91%	BAA92699: 88%/91%
GmPP2A-4	GM47171610	soja	114 - 1082	43% / 56%	CAA05491: 43%/56%
GmPP2A-5	GM49671923	soja	91 - 1008	78% / 87%	BAA92697: 78%/87%

Apéndice

5 Secuencia de nucleótidos del PP2A-1 parcial de *Physcomitrella patens* (SEQ ID NO: 1)
 GAAGTCTTCCATACGTGGCATTTCGGTACTTTTGAAGACACTCGTCATAGAAACC
 ATACACTTGGGTGATCTGCCGACTCTCATGATTGCCACGCAGAATGGTAATACGA
 TCAGGATATCGTACTTTTTAAGCCACTAGCAGTGTAGCAGTTTTCAACTGAATAAT
 AACCACGATCCACATAATCTCCCATAAATAAGTAATTGGTGTCTGGACACATTCC
 10 TCCAATTCGGAAGAGTTCAGCAAGATCATGAAACTGGCCATGAATATCACCACA
 AATTGTGACTGGACACTTCACTGGCTGAACATTATTTCCCGCATCAATATCTCCT
 TCGCCTTCTCACATAGTCTCTGACCTCGACCTCGGATAACGGTTTGCATTGTATG
 AGCTGAGCAATCTGTGTGTCCAGCTGCCATTAGAGGATATCGGAGGCACACTC
 ATGTCGTCCCTCCTGTTGCGCTTCCCCTTCTCCACGCTATCGCCCTACCCTCCGTT
 15 TCCGCTGATCTCCGCCTCAAAAACCAACTCCGACACTCTCGAAACGCAATCTGCA
 ACACCGACAACAAAAAGAAAATCACGTGACGAAAGAAAGGGTGAGAAGCAACA
 GGGCGAAAAAGAAAATCACGGACGAAAGAAAGGGTGAGATGCAACAGGGCGAG
 AGGGGAACGCAAGAGGAACGACAGAGGAGCGACCTACGGTGAGCTGGTGC
 Secuencia de nucleótidos del PpPP2A-1 de longitud completa de *Physcomitrella patens* (SEQ ID NO: 2)
 20 CTTGTTCTAGGGTTGCGATTGCGTTTCGAGAGTGTGGAGTTGGTTTTGAGGCG
 GAGATCAGCGGAAACGGAGGGTAGGGCGATAGCGTGGAGAAGGGGAAGCGCAA
 CAGGAGGGACGACATGAGTGTGCCTCCGATATCCTCTAATGGGCAGCTGGACAC
 ACAGATTGCTCAGTCCATAAATGCAAACCGTTATCCGAGGTCGAGGTCAGAGG
 ACTATGTGAGAAGGCGAAGGAGATATTGATGCGGGAAAATAATGTTTCAGCCAGT
 25 GAAGTGCCAGTCACAATTTGTGGTGTATTCATGGCCAGTTTCATGATCTTGCT
 GAACTCTTCCGAATTGGAGGAATGTGTCCAGACACCAATTACTTATTTATGGGAG
 ATTATGTGGATCGTGGTTATTATTCAGTTGAAACTGCTACACTGCTAGTGGCTTTA
 AAAGTACGATATCCTGATCGTATTACCATTCTGCGTGGCAATCATGAGAGTCGGC
 AGATCACCCAAGTGTATGGTTTCTATGACGAGTGTCTTCGAAAGTACGGAAATGC

CAACGTATGGAAGATCTTCACTGACCTGTTTGATTATTTTCCTTTAACAGCACTCG
 TAGAGTCGGAGATTTTTTGTTTACATGGAGGGCTTTTCGCCAAGCATCGATTCCCT
 GGACCACATTCGGGACCTGGATCGAGTTCAAGAGGTTCCCTCATGAAGGTCCGAT
 GTGTGATCTACTTTGGTCTGACCCCGATGACCGTTGTGGTTGGGGCATTCTCCCC
 5 GTGGTGCTGGCTACACATTTGGCCAGGATATATCTGAGCAGTTCAATCACAACAA
 CAATCTGAAGTTGGTCGCAAGGGCACATCAATTAGTTATGGAGGGCTACAATTG
 GGGACATGAACACAAGGTGGTCACTATTTTCAGCGCACCTAATTATTGCTATCGC
 TGTGGAACATGGCTTCTATATTGGAAGTGGATGACAATATGGGCCACACTTTCA
 TTCAGTTTGAACCAGCCCCGAGACGAGGTGAACCAGATGTGACAAGGCGCACGC
 10 CAGACTACTTCTTATAGTGGGACTTTCTGATAGTAGTTTTTAAAGTATGCTTTGAG
 CTATTTTGGATCGTCTGTAGTCCATGCATTCAATGATGTAGATTTTCCTCAGGTTA
 GCATGGTGTTACCAAGCGATAGCAGCCTGAATGCTGTCATAACCGCCACACCAT
 CATATGATATGTATTTCAATTGAGCGGGCATGCTACTCTGCGCTTGAGATGTAAGC
 GAGTCTCTATTTGGAGTG

15 Secuencia deducida de aminoácidos de PpPP2A-1 de *Physcomitrella patens* (SEQ ID NO: 3)

MSWPISSNGQLDTQIAQLIQ LSEVEVRGLCEKAKELLMRENNVQPVKCPVTICG
 DEHQFHDLAELFRIGGMCPDTNYLFMGDYVDRGYYSVETATLLVALKVRYPDRTI
 LRGNHESRQITQVYGFYDECLRKYGNANVWKIFTDLFDYFPLTALVESEIFCLHGGLS
 PSDDSLDHD DLDRVQEVPHGPMCDLLWSDPDDRCGWGISPRGAGYTFGQDISEQF
 20 NHNNNLKLVARAHQLVMEGYNWGHEHRVVTIFSAPNYCYRCGNMASEVDDNM
 GHTFIQFEPAPRRGEPDVTRRTPDYFL*

Secuencia de nucleótidos del PpPP-1 parcial de *Physcomitrella patens* (SEQ ID NO: 4)

GGCACGAGGTTTTAGCTGCGCGCGGAAGAAGCAGCGTGCGCGGCGGTGGTTGTT
 TGGTTTTTGTTCCTGTGTTGCTGTTAGCTGCGCAAAGGAAGGGGACTGGACACA
 25 ACGTGATGGACTTAGATCAGTGGCTTGAGAAAGTGAAGAGCGGCAACTACCTCT
 TGAAGACGAGCTCAAGCAACTATGTGAATATGTGAAAGAAATATTGGTGGAGG
 AATCCAATGTTTCAGCCTGTCAACAGTCCCGTTACTGTTTGTGGCGATATCCATGG
 CCAGTTTCATGACTTGATGAAGCTTTTTTCAGACTGGAGGACACGTCCCCAGCACA
 AACTACATCTTCATGGGTGATTTTGTGGATCGAGGTTACAACAGTTTGAAGTAT
 30 TTACAATACTTTTGTGCTGCTGAAAGCAAGATACCCTGCTCATATGACGTTGTTGAG
 GGGTAACCATGAGAGTAGACAGATAACTCAGGTATATGGATTTTATGACGAATG
 CCAGCGGAAGT

Secuencia de nucleótidos de PpPP-1 de longitud completa de *Physcomitrella patens* (SEQ ID NO: 5)

GCCCTTATCCCGGGAGGAAGGGGACTGGACACAACGTGATGGACTTAGATCAGT
 35 GGCTTGAGAAAGTGAAGAGCGGCAACTACCTCTTGAAGACGAGCTCAAGCAAC
 TATGTGAATATGTGAAAGAAATATTGGTGGAGGAATCCAATGTTTCAGCCTGTCA

ACAGTCCCGTTACTGTTTGTGGCGATATCCATGGCCAGTTTCATGACTTGATGAA
 GCTTTTTCAGACTGGAGGACACGTCCCCAGCACAACTACATCTTCATGGGTGAT
 TTTGTGGATCGAGGTTACAACAGTTTGAAGTATTTACAATACTTTTGCTGCTGA
 AAGCAAGATACCCTGCTCATATGACGTTGTTGAGGGGTAACCATGAGAGTAGAC
 5 AGATAACTCAGGTATATGGATTTTATGACGAATGCCAGCGGAAGTATGGAAACC
 CAAATGCTTGGCGGTTACTGCACTGATGTTTTTACTACCTTACACTCTCAGCCAT
 AATAGATGGAAGGGTGTGTGTGTTTATGGAGGTCTGTCTCCAGACATTCGGACA
 ATTGATCAGATTAGGGTGATAGAGAGGCAGTGTGAGATTCCTCATGAAGGGCCA
 TTCTGTGACTTGATGTGGAGTGATCCTGAGGATATCGAACTTGGGCTGTTAGCC
 10 CACGAGGTGCTGGGTGGCTTTTTGGTGCACGCGTTACCTCTGAGTTCAATCACAT
 AAACGGATTGGAGCTTGTATGCCGTGCGCATCAATTAGTTCAAGAGGGATTGAA
 GTACATGTTTTCTGACAAAGGACTTGTACCGGTGTGGTCCGCTCCAACTATTGC
 TACAGATGTGAAAATGTTGCTTCAATCTTAAGCTTCAACGAAAATATGGAGAGA
 GATGTGAAATTTTTACTGAGACCGAGGAGAACCAGGCTATGATGGCACCTCGA
 15 GCAGGAGTTCCTTACTTCTGTAGAGATATTTGTCGCAGATACCACATGACGACC
 GGAAAGCATCATATGGTGCCTAACGCAAGGGC

Secuencia deducida de aminoácidos de PpPP-1 de *Physcomitrella patens* (SEQ ID NO: 6)

MDLDQWLEKVKSGNYLLEDELKQLCEYVKEILVEESNVQPVNSPVTVCGDIHGQFHDLMLK
 FQTGGHVPSTNYIFMGDFVDRGYNLSLEVFTILLKARYPAHMTLLRGNHESRQITQVYGFY
 20 DECQRKYGNPNAWRYCTDVFYDLTLSAIDGRVLCVHGGLSPDIRTIDQIRVIERQCEIPHE
 GPFCDLMWSDPEDIETWAVSPRGAGWLFGARVTSEFNHINGLELVCRAHQLVQEGLKYM
 FPKGLVTVWVWAPNYCYRCGNVASILSFNENMERDVKFFTEETEENQAMMAPRAGVPYFL*

Secuencia de nucleótidos de BnPP2A-1 de *Brassica napus* (SEQ ID NO: 7)

TTGAAATTGAAATTCGCATTTTTGCTATGGGGAAGTACTGAGTGATGGTAGATAAATTCG
 AATCCAAATCCGCATCCGGATCCAATACTATCCGAATCCGGATTTTGGTTTTTGG
 TCAGATCGGGGATCGGATCTGAGGGAAGGAGAAGACGATGCCGGAGACGGGAGA
 CATCGATCGTCAGATCGAGCAGCTGATGGAGTGTAAGCGTTGTCCGAGGCGGAG
 GTGAAGACGCTGTGCGAGCAAGCGAGGGCGATTCTGGTGGAGGAGTGGAATGTT
 30 CAGCCGGTTAAGTGTCCGGTCACCGTCTGCGGCGACATCCACGGCCAGTTTTACG
 ATCTGATTGAGCTTTTTAAGATCGGTGGTTCTTCGCCTGACACCAATTATCTCTTC
 ATGGGCGATTACGTAGATCGAGGGTATTATTCTGTGGAGACAGTCTCGCTCTTGG
 TAGCACTCAAAGTTCGCTACAGAGATAGGCTTACCATCTTAAGAGGGAATCACGA
 AAGCCGCCAAATACTCAAGTGTATGGATTTTATGATGAGTGCTTGAGAAAATAT
 35 GGAAATGCTAATGTCTGAAACACTTCACTGACCTTTTTGATTATCTTCCTCTTAC
 AGCTCTCATCGAGAGTCAGTTTTCTGTTTACATGGAGGGCTCTCACCTTCTTTAG
 ATACACTTGACAACATCCGTTCTCTAGATCGAATCCAAGAGGTTCCACATGAAGG

ACCTATGTGTGATCTGTTATGGTCCGATCCAGATGATCGATGCGGTTGGGAATA
 TCTCCTCGTGGCGCAGGCTACACGTTCCGACAAGATATCGCTACTCAGTTTAACC
 ACACCAATGGACTCAGTCTGATCTCAAGAGCACACCAACTTGTCATGGAAGGTTA
 TAATTGGTGCCAAGAAAAGAACGTTGTGACTGTGTTTAGCGCCCCAAACTATTGC
 5 TACCGATGCGGCAACATGGCTGCTATTCTAGAGATAGACGAGAACATGGACCAG
 AACTTCCTCAGTTCGATCCAGCCCCACGTCAAGTAGAACCCGAAACTACACGCA
 AACTCCAGATTACTTTTTGTAAGTACCCAAAAAGAAAAAACATCCTTAACCTT
 GTTCTGTAATTTCAATTCCTGTTCTGTTAACTCGTAGTTGTCTTTTTGGTTTTAGTT
 AAGAATGTGTAACCTTTTAACTGATACAAAGCGTTACAAAAGATTCTGGTCCATA
 10 TGAATAAGGCAATTGTTGTTGAGAGCTA
 Secuencia deducida de aminoácidos de BnPP2A-1 de *Brassica napus* (SEQ ID NO: 8)
 MPETGDIDRQIEQLMECKALSEAIEVKTLCQARAILVEEWNVQPVKCPVTVCGDIHG
 QFYDLIELFKIGGSSPDNYLFMGDYVDRGYYSVETVSLLVALKVRYRDRLTILRGNH
 ESRQITQVYGFYDECLRKYGNANVWKHFTDLFDYLPALTALIESQVFCLHGGGLSPSLDT
 15 LDNLRSLDRIQVPEHGPMCDLLWSDPDRCGWGISPRGAGYTFGQDIATQFNHTNGL
 SLISRAHQLVMEGYNWCQEKNVTSAPNYCYRCGNMAAILEIDENMDQNFLQFD
 PAPERQVEPETTRKTPDYFL*
 Secuencia de nucleótidos de BnPP2A-2 de *Brassica napus* (SEQ ID NO: 9)
 ATTTTATTCACCTACTTAACTCTCTCCGTCGTCGGATTTGTATTCGTTCCGAGAAA
 20 GGGGGTTAAGGTTTTAGGCTTGTCCGTACGATGGACGAGAATTTGCTGGACGA
 TATAATACGGCGGCTGTTGGAGACTAACAACGGGAAGCAGGTGAAGCTACTCGA
 GGCTGAGATACGCCAGCTCTGCTCTGCTTCCAAAGAGGTTTTTCTCAGCCAGCCTA
 ATCTCCTCGAGCTCGAGGCTCCTATCAAGATTTGCGGTGATGTTTCATGGTCAGTTT
 CCAGACCTCTTGGGTTGTTTGAGTACGGCGGCTACCCTCCAGCTGCAAATACTT
 25 GTTCCTCGGAGACTACGTTGATCGTGGTAAGCGGAGCATAGAGACCATATGCCTT
 CTCCTTGCCCTACAAGCTCAAATACAAGCTCAACTTCTTTCTCCTCAGAGGCAATCA
 CGAATGCGCTTCTATCAACCGTGTTCACGGCTTCTACGATGAGTGCAAGAGAAGA
 TACAACGTGCGCCTGTGGAAGAGTTTCACCGACTGTTTCAACTGCCTCCCCGTTGC
 TGCTCTCATCGACGACAAGATCCTCTGTATGCACGGTGGACTTTCTCCTGATCTCA
 30 AGACCTTGGATGATATCAGGCGGATTCCTCGTCCTGTTGATGTTCTGATCAGGGC
 GTCCTTTGTGATTTGTTATGGGCTGATCCTGACAAAGAAATCCAAGGCTGGGGGG
 AGAATGACAGAGGTGTGTCTTATACATTTGGTCCCGACAAAGTGGCTGAGTTCCT
 TCAGACTCATGACCTTGATCTTGTGGCCGAGCTCATCAGGTTGTAGAAGATGGAT
 ATGAGTTCTTTGCAAAGAGACAACCTGGTGACAATATTCTCTGCACCCAACTACTG
 35 TGGTGAGTTTGACAATGCTGGCGCAATGATGAGTGTTGATGATAGTTTGACATGTT
 CTTTCAAATCCTCAAGTCAACTGAGAAGAAAGGAAGATTTGGATACAACAACAA

CGTTCATAGGCCAGGAACCCACCTCATAAGGGGGGAAAAGGTGGTTGAGATGG
 GGAATCAAGAGAAGAGTGAAGCCGAAGGGTTCGAACTTTATGGTCAATGTAAT
 GTAGGTGATTTGAGGCAATACCGTTTGTGTTTTGTTTGATTGATGCAAAGATTTT
 GGTTTTGTTAGATTGTTTTGTAAGTACGACGATTTTCAACTTAAGAAAGTTGG
 5 GTTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de BnPP2A-2 de *Brassica napus* (SEQ ID NO: 10)

MDENLLDDPIRRRLEETNNGKQVKLLEAEERQLCSASKEVFLSQPNLLELEAPIKICGDVH
 GQFPDLLRLFYGGYPPAANYLFLGDYVDRGKRSIETICLLLAYKLRKLNFFLLRGN
 HECASENRVYGFYDECKRRYNVRLWKSFTDCFNCLPVAALDDDKILCMHGGLSPDLK
 10 TLDDIRRLPRVDVDPDQGVLCDLLWADPDKEIQGWGENDRGVSYTFGPDKVAEFLQT
 HDLDLVCRAHQWEDGYEFFAKRQLVTIFSAPNYCGEFDNAGAMMSVDDSLTCSFQI
 LKSTEEKGRFGYNNNVHRPGTPPHRGKGG*

Secuencia de nucleótidos de BnPP2A-3 de *Brassica napus* (SEQ ID NO: 11)

GGCGAAAACCTTTTTGGTCTGAAGATCGAGAAAAGATTTGAAAATCAAATTTGG
 15 GTTCCTGGGCAATCGATTTATGTCCGAAAGATTGAAGCTTTGTAAGATAAAAAG
 ATCGATTGGTGTGTTAAGTTTCGATCGCGTGGGGTCTTTGTAATGACGCAGCAA
 GGGCAGGGAAGCATGGACCCTGCCGTTCTCGACGACATCATTGTCGTTTGTGG
 ATTACAGAAACCCTAAGCCTGGAACCAAACAGGTCATGCTCAACGAGTCTGAGAT
 CCGACAGCTTTCAGCGTGTCTAGAGAGATTTTCCTCAGCAGCCTAACCTCCTTG
 20 AGCTCGAGGCTCCAATTAAGATCTGTGGTGATTCATGGACAGTACTCAGATCT
 ACTGAGGCTATTTGAGTACGGAGGCTTACCTCCTGCAGCTAACTATCTATTCCTAG
 GAGATTACGTGGATCGCGGGAAGCAGAGCCTAGAAACAATCTGCCTTCTCCTTGC
 CTACAAGATCAAATACCCTGAGAACTTCTCCTCCTAAGAGGCAACCACGAATGC
 GCTTCCATCAACCGAATCTACGGATTCTACGATGAACGTAACGCAGGTTGAGTG
 25 TCAGACTCTGGAAAGTGTTCACAGATTCTTCAACTGCCTCCCTGTAGCTGCTGTA
 ATAGACGATAAGATATTGTGTATGCACGGTGGCCTTTCTCCTGATTTGACCAGCGT
 GGAACAGATTAAGAACATTAAGCGACCTACCGATGTTCCGGACTCCGTTTTGTTA
 TGTGATCTGCTTTGGTCTGATCCGAGCAAAGATGTGAAAGGCTGGGGGATGAATG
 ACCGTGGAGTTTCTTACACGTTTGGGCCTGATAAAGTTGCTGAGTTTTTGATAAAG
 30 AATGATATGGATCTCATCTGTCGTGCTCACCAGGTTGTAGAGGATGGTTATGAGTT
 CTTTGC GGATAGACAGCTTGTTACTATATTTTCAGCTCCTAATTACTGTGGTGAAT
 TCGATAATGCTGGTGCTATGATGAGTGTTGATGAGAGTTAATGTGCTCTTTTCAA
 TTCTTAAGCCTGCGGATCGGAGGCCTCGTTCTTATGAGTTAGAGCCTCACTGGA
 AAGAAGACGAAATTGGCGAGATGAAAACGGGAGAGAGAGAGAGAGACATTTGA
 35 AACTCCCGGAGACTTTGTCCTGAGGCCTTTGCAAGAAGGCAGGAAAAAAAAACA
 CAGTGTTACATGTTATATCATATAATCTTATTTGAACTTTTGAATTTCTTTTCTCA

AAACTTTTATGTTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAA

Secuencia deducida de aminoácidos de BnPP2A-3 de *Brassica napus* (SEQ ID NO: 12)

MTQQGQGSMDPAVLDDPIRRLLDYRNPKEGKQVMLNESEIRQLCSVSREIFLQQPNL
5 LELEAPACGDIHGQYSDLLRLFYGGGLPPAANYLFLGDYVDRGKQSLETICLLLAYKI
KYPENFLLRGNHECASD RIYGFYDERKRRFSVRLWKVFTDSFNCLPVAVIDDKILC
MHGGLSPDLTSVEQIKNP RPTDVPDGLLWSDPSKDVKGWGMNDRGVSYTFG
PDKVAEFLIKNDMDLICRAHQVVEDGYEFFADRQLVTIFAPNYCGEFDNAGAMMSV
DES*

10 Secuencia de nucleótidos de GmPP2A-1 de *Glycine Max* (SEQ ID NO: 13)

CCTAACCCATCGTTCGCACGGAGCCGTGGTGGTCTGGCTCTGCCTTTTTTCCTTT
CTCTCGTTACCAGAACCAAGAACCATTTGTGGCCAACATGGATTTGGACCAGT
GGATCTCCAAGGTCAAAGAGGGCCAGCATCTTCTTGAAGACGAGCTTCAACTTCT
CTGCGAATATGTAAAAGAGATCCTTATCGAGGAGTCCAATGTGCAGCCTGTCAAT
15 AGTCCAGTGACTGTTTGTGGTGATATTCATGGTCAATTCATGATCTAATGAAACT
TTCCAGACTGGGGGTCATGTGCCTGAGACAAATTATTTTTATGGGAGACTTTG
TTGATCGAGGTTACAATAGTCTTGAAGTTTTTACCATCCTTTTACTTCTAAAAGCT
AGATACCCAGCTAATATACTCTTTTACGTGGAAATCATGAAAGTAGACAATTAA
CCCAGTCTATGGATTTTATGATGAATGCCAGAGGAAGTATGGCAATGCCAATGC
20 TTGGCGGTATTGTACAGATGTGTTTACTATTTAACACTTTCTGCAATTATTGATG
GAACTGTGCTTTGTGTTTATGGGGGCTTTCTCCTGACATTCGAACAATTGATCAG
ATAAGGGTCATTGACCGGAAGTGTGAAATTCCTCATGAGGGTCTTTTCTGTGATCT
AATGTGGAGTGATCCTGAAGATATTGAAACATGGGCAGTCAGTCCCGTGGAGCA
GGTTGGCTTTTTGGATCCAGGGTCACTTCGGAGTTCAATCACATAAATAACCTTGA
25 TCTTGTGTTGTCGAGCGCACCAACTCGTTCAGGAAGGCCTTAAGTATATGTTCCAAG
ATAAAGGCCTTGTAAGTGTATGGTCTGCACCTAATACTGTTACCGATGTGGAAT
GTAGCTTCTATTCTGAGTTTCAATGAAAATATGGAAAGAGAAGTTAAATTTTTCAC
CGAAACAGAGGAGAACAATCAGATGAGAGGGCCAGGACAGGCGTTCCATATTT
CTTATAAGTTGGTGCAAAATTTGTTTTGAATTTATTGAAAATTAGACTCATGT
30 ATTTATGCTTTGCCTTTTAAAGGTGGATTTTATTGGTCACAAGATTACCAATCAA
CTATATCTTAGCTCTGGGTCGCACAGATAATTTTATGTTTAAATTTTATTGAAAA
AAAAAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de GmPP2A-1 de *Glycine Max* (SEQ ID NO: 14)

MDLDQWISKVKEGQHLLLEDELQLLCEYVKEILIEESNVQPVNSPVTVCGDPIGQFHDL
35 MKLFQTGGHVPEYNYIFMGDFVDRGYNLEVFILLLLKARYPANITLLRGNHESRQL
TQVYGFYDECQRKYGNANAWRYCTDVFYDLTSLAIDGTVLCVHGGLSPDIRTEDQR

ES 2 381 031 T3

VDDRNCETHEGPFCDLMWSDPEDIETWAVSPRGAGWLFGRVTSEFNHINNDLVCR
AHQLVQEGLKYMFDKGLVTVWSAPNYCYRCGNVASE SFNENMEREVKFFTETEE
NNQMRGPRTGVVYFL*

Secuencia de nucleótidos de GmPP2A-2 de *Glycine Max* (SEQ ID NO: 15)

5 AAGAAGAAGAGGTTTTGATCGGATGCGATGAGCACACAGGGGCAAGTGATTATT
GATGAGGCGGTTCTGGATGACATAATCCGGCGCTTAACGGAGGTCCGACTGGCCC
GACCCGGCAAGCAGGTTACAGCTCTCCGAGTCTGAGATCAAGCAACTCTGCGTCGC
TTCCAGAGACATCTTCATTAACCAGCCCAATTTGCTTGAAGTGAAGCCCCCATCA
AGATTTGTGGTGACATTCATGGGCAGTACAGTGATTTGTTAAGGCTATTTGAGTAT
10 GGGGTTTGCCTCCTACTGCGAATTATCTCTTTTTGGGGGAATACGTGGACCGTGG
GAAGCAGAGCTTAGAAACCATATGTCTTTTGGCTTATGCGTATAAAATCAAATATCCA
GAAAACTTTTCTGTTAAGGGGAATCATGAGTGTGCTTCCATTAATAGGATTTA
TGGGTTTTATGATGAATGTAAGCGAAGGTTTAACGTGAGGCTTTGGAAAGCCTTT
ACCGACTGTTTTAACTTCCTTCTGTGGCAGCCCTTATAGATGATAAAATATTGTG
15 CATGCATGGTGGTCTTCCCCTGAACTCACAACTTGGATGAAATCAGGAATCTA
CCTCGTCCTACTGCGATTCCCAGACACCGGCTTGGCTTTGTGATTTGCTTTGGTCTGAT
CCTGGTAGGGATGTGAAGGGTGGGGTATGAATGACAGAGGAGTGTCTACACCT
TTGGCCCTGATAAGGTCGCTGAGTCTTGACAAAGCATGACTTGGACCTCATTGT
CGTGCTCATCAGGTTGTAGAGGATGGGTATGAATCTTTGCTGATAGGCAACTTGT
20 TACGATATTTTACAGCTCCAACTATTGTGGTGAATTTGACAATGCTGGTGGATGA
TGAGTGTGGACGAAAACCTGATGTGCTCATTTCAGATTCTTAAGCCTGCAGAGAA
AAAATCAAAGTTTGTGATGTCAAACAAGATGTGATGGTTGGCACATCACTGTCAA
GTAATTAACCAAGATGTATTCGTGGAGCTAAATTAATCCTGAAGATTTAGATTG
CATGGTCTTAAGTTCTATCTATTCTGAGGTGATGATGATGAACAAACAAGTTTACT
25 GCTATAACATCCAGTCAGGCAGTGAGCATGAGGTACTACAAGAGATATTAAGCAC
TGTTGGATGGCCATAAAAGCAAAGGCTTTTATCTTTTTTTTTCTTTTCTTTTAT
AATTATTCTGCAACACAATATGTACATATATGTGTTGTAGATGCTCTGGAATGAC
CTTCTTTGCTCTGAAAGGTCCTCTTAGACTATCGATTTACACTGATAGAGCAGTTT
GTGTTGATGATTGTGGCCAATTTTATCCAGTTAGTAAAGGTGCAATTGATGGATTT
30 TATGGTTTAAAAAAAAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de GmPP2A-2 de *Glycine Max* (SEQ ID NO: 16)

MSTQGQVELDBAVLDDIIRRLTEVRLARPGRQVQLSESEKQLCVASRDFFLNQPNLLELE
APIKICGDPIGQYSDLLRLFYEGGLPPTANYLFLGEYVDRGKQSLETICLLLAYKIRYPE
NFFLLRGNHECAS1NRIYGFYDECKRRFNRLWKAFTDCFNFLPVAALDDDKLLCMHG
35 GLSPELTNLDEIRJNLPRPTAIPDTGLLCDLLWSDPGRDVKGWGMNDRGVSYTFGPK
VAEFLTRHDLDLICRAHQWEDGYEFFADRQLVTIFSAPNYCGEFDNAGAMMSVDEN

ES 2 381 031 T3

NHTNSLKLIAHQLVMDGFINWAHEQKVVTLSAPNYCYRCGNMASHLEVDCKGH
TFIQFEPAPRRGEPDVTRRTPDYFL*

Secuencia de nucleótidos de GmPP2A-4 de *Glycine Max* (SEQ ED NO: 19)

GCTCACTCTTCCAACACTACTACTGTTGTTCTTCTCGTCGTCTTGGCCTTCGCATCTT
5 CACAATCACATCCAATAGAGACACGTTGACTTTGCTGGAAGAAGAAGAAGA
GAATGGAACAATCGCTTTTGGATGACATAATCAATCGCCTCCTCGAAGTTCCTAC
CCTACCGGCTAAGCAGGTTGAGCTATCCGAGTCCGAGATCCGTCAACTCTGCGTA
GTTTCCAGAGAAATTTCTTGAACAACCTAATTTATTGGAGCTCGAAGCACCTAT
TAAGATTTGTGGTGATGTACATGGGCAATATTCTGATCTTTAAGGCTTTTTGAGT
10 ACGGTGGATTACCTCCTGAAGCCAACCTATTTGTTTTGGGGATTATGTTGATCGA
GGGAAGCAGAGTTTAGAAACAATTTGCCTCCTCCTTGCTTATAAAATAAAATATC
CTGAGAACTTTTTCTTGTTAAGGGGAAACCATGAATGTGCTTCTATAAACCGGAT
ATATGGATTTTATGATGAGTGCAAGAGAAGGTTCAATGTAAGGTTATGGAAGACA
TTACAGACTGCTTCAATTGCCTGCCTGTGGCAGCCCTTGTGATGAAAAGATTTT
15 GTGTATGCATGGGGGACTTTCTCCCGACTTAAATAATTTGGACCAAATTAGAAAT
TTACAGCGGCCACAGATGTTCTGATACAGGTTTGCTTTGTGATCTGCTTTGGTC
TGACCCGAGCAAAGATGTTCAAGGATGGGAATGAATGACAGAGGAGTTTCATA
CACATTTGGTGCTGATAAGGTCTCACAATTTCTCAGAAACATGATCTTGATCTTG
TTTGTGCTGCTCATCAGTTGTGGAAGATGGATACGAGTTCTTTGCTAATCGACAA
20 CTTGTAACAATATTTTCAGCACCTAATTATTGTGGGGAGTTTGACAATGCTGGTGC
TATGATGAGTGTTGATGAGACGCTAATGTGCTCTTTCCAAATATTAAGCCAGCT
GATAAAAAGCAAAGCTCAATTTGGAAGTACAACCACTGCTAAGCCTGGAACT
CTCCAGCAGGTGTAAGGTTGGAAGATATTAGTCCTCCTGGATGCGAAAGTGTG
AAATTAATTTGGCTAAAAGATTGCTACTACTACGGATCAGCTTGGGCTTGA
25 CCTAATGGTTGCAAGAAGGGGAAAATCAAGTTCATTTGCCTACTATGATATTTT
GGAATTGTA
AATCAAAGAGAACCATTATGAAGTTTGTAACCATTGTTTATT
ATTGGTACAAATTTGCATTTCAAGATGGAGAGCCATAATCTCCTTGCTTCTTGT
ACACTAATAACTGGTATATTTTCTTAAGCTTCAAGCTTCAAGCGTAGATGGATAC
ATCCGAATCTGTTGCTGAGAACCATTTAAAATGCTTATACGATTTGGCATATATG
30 GATGGCAGTTGAGGCTGGTGA
AAAAAAAAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de GmPP2A-4 de *Glycine Max* (SEQ ID NO: 20)

MEQSLDDDDNRLLVPTLPKQVQLSESEPIQLCVSREIFLQQPNLLELEAPIKICGDV
HGQYSDLLRFLFEYGGPLPEANYLFLGDYVDRGKQSLETICLLLAYKLYPENFFLLRGN
HECASINR1YGFYDECKRRFNVRLWRTFTDCFNCLPVAALVDEKILCMHGGLSPDLNN
35 LDQD NLQRPTDVPDGLLCDLLWSDPSKDVQGWGMNDRGVSYTFGADRVSQFLQK
HDLDLVCRAHQVVEDGYEFFANRQLVTFFSAPNYCGEFDNAGAMMSVDETLMCSFQI

ES 2 381 031 T3

LKPADKKAKLNFSTTTAKPGNSPAGVKVGRY*

Secuencia de nucleótidos de GmPP2A-5 de *Glycine Max* (SEQ ID NO: 21)

ACCGTCCCGGTTCGCGCCAACCGCCGCAACCCGAAGAAACCGAATCGATCTGAGA
GAAGGTGCGATCTCGGAGGTGGGAGCCAAACGAAACGATGCCGTCTCACGCGGA
5 TCTGGAGCGACAGATCGAGCAGCTGATGGACTGCAAGCCTCTGTTCGGAGTCGGAG
GTGAAGGCGCTGTGCGATCAAGCGAGGACGATTCTTGTGGAGGAGTGAACGTG
CAACCGGTTAAGTCCCCGTCACCGTCTGCGGCGATATTCACGGCCAGTTCTACG
ATCTCATCGAGCTGTTTCGGATTGGAGGGAACGCTCCCGATACCAATTATCTCTTC
ATGGGTGATTATGTAGATCGTGGATACTATTCAGTGGAGACTGTTACACTTTTGGT
10 GGCTTTGAAAAGTCCGTTATAGAGATAGAATCACAATTCTCAGGGGAAATCATGAA
AGCCGTCAAATTAAGTCAAGTGTATGGCTTCTATGATGAATGCTTGAGAAAATATG
GAAATGCGAATGTCTGGAATACTTTACAGACTTGTTCGATTATTTGCCTCTGACT
GCCCTCATTGAGAGTCAGATTTTCTGCTTGCATGGAGGTCTCTCACCTTCTTTGGA
TACACTGGATAACATCAGAGCATTGGATCGTATTCAAGAGGTTCCACATGAAGGA
15 CCAATGTGTGATCTCTTGTGGTCTGACCCTGATGATCGCTGTGGATGGGGAATATC
TCCACGTGGTGCAGGATACACATTTGGGCAGGATATAGCTGCTCAGTTTAATCAT
ACCAATGGCCTCTCCCTGATATCGAGAGCACATCAGCTTGTTATGGAAGGATTCA
ATTGGTGCCAGGACAAGAATGTGGTGACTGTATTTAGTGCTCCAAATTAAGTGTAT
CGATGTGGGAATATGGCTGCCATACTAGAAAATAGGAGAGAATATGGATCAGAATT
20 TTCTTCAGTTTGATCCAGCTCCAGGCAAATTGAGCCTGACACCACACGCAAGAC
TCCAGATTATTTTTGTAAGTTCATTTATCTGCCTGTTTGTAGTTACTGCTTTCTGC
CATTACTGTAGATGTGTCTTTAAGGAAAGGAGTTTTACTGTGTAAGTGGAGGGTG
GTCATCAACATAATTCTTTCTTTTGGAGTTTACCTGTTGCTGCTGCCGCTGCCTTAT
CTGTACAAGAAACCAATAGAAGTACACATGACACCAATTGGGGTTGTTGTATAT
25 TTTTGGGAGGAAGCAGCATAACATGGTATATCTTTTCTGTAATCCTTTTTCTTTTCT
TTAAATTAATCTCAAGTTAAAGAGCAGATTTTTGAGTCTGACAATGATGTCCTT
TTGAGACTTTTGTATGATGCCAAATGAAATTGCAGGTTTTCAAAAAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de GmPP2A-5 de *Glycine Max* (SEQ ID NO: 22)

MPSHADLERQIEQLMDCKPLSESEVKALCDQARTILVEEWNVQPVKCPVTVCEDIHG
30 QFYDLD3LFRIGGNAPDTNYLFMGDYVDRGYYSVETVLLVALKVRYRDRITILRGNH
ESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLPALTALIESQIFCLHGGLSPSLDTL
DNIRALDRIQEVPEHGPMDLLWSDPDDRCGWGISPRGAGYTFGQDIAAQFNHTNGL
SLISRAHQLVMEGFNWCQDKNVVTSAPNYCYRCGNMAAILEIGENMDQNFLQFDP
APRQIEPDTRKTPDYFL*

35 Secuencia de nucleótidos de OsPP2A-1 de *Oryza sativa* (SEQ ID NO: 23)

GCGCCCCCAACCGCATCTCAATCTCTCTCCTCTCTCTCTCTCTCCAGCTCGCCTCC

CCCTCGCCGGCGACGAGCTCTCCCCGGCTAGGGTTAGACCAGGTGCCCTCTTG
TCTGCGGCGGAGGAGGTGGGAGGAAGGATGGAGGAGTCCGTGGGCTCGCGCGGC
GGCGGCGGCGGCGGCTGGACGCCAGATCGAGCAGCTCATGGAGTGCCGCCCG
CTCTCCGAGCCCGAGGTCAAGACGTTGTGCGAGAAGGCGAAGGAGATATTGATG
5 GAAGAAAGCAACGTTTCAGCCAGTTAAGAGCCCAGTCACAATTTGTGGTGATATCC
ATGGGCAATTCCATGATCTAGTAGAGCTCTTTCGGATTGGTGGGAAGTGTCCAGA
TACAAATTATTTGTTTATGGGAGATTATGTAGATCGTGGCTACTATTCTGTTGAGA
CTGTTACACTTTTGGTTGCACTGAAGGTGCGCTACCCACAGCGGATTACAATCCTT
CGTGGAACCATGAGAGTCGGCAGATCACACAGGTGTATGGATTCTACGACGAAT
10 GCCTACGAAAGTATGGAAGTGCAATGTCTGGAAGATCTTCACCGATCTTTTTGA
CTATTTCCATTGACAGCATTGTTGAATCAGAGATTTTCTGCCTCCATGGTGGTT
TATCGCCATCAATCGACAATCTTGATAGTGTTCGCAGCTTAGATCGTGTTCAAGAG
GTCCCTCATGAGGGACCAATGTGTGATCTTCTATGGTCTGACCCGGATGATCGAT
GCGGTTGGGGCATATCTCCTCGTGGTGCTGGCTACACTTTTGGCCAGGACATATCG
15 GAGCAGTTTAACCATAACCAATAATCTCAAACCTGTAGCCCGGGCTCATCAATTAG
TTATGGAAGGATATAACTGGGCGCACGAACAAAAGGTCGTGACCATATTCAGTGC
ACCTAATTATTGTTATCGCTGTGGCAACATGGCATCCATCCTGGAGGTTGATGACT
GCAGGAATCACACATTTATTGAGTTTGAACCAGCTCCTAGGAGAGGTGAACCAGA
TGTGACACGGAGAACACCTGATTATTTCTTTAAATTATCTGTTGTAATTTGTATT
20 GTTTTGTTTCTTTTGTCTCTAAGACCGCAATAGTGAGTGCTGGTCAGTAAAATT
TTGTTGGATCCCTTTGGTAACTAACTGGCCAGCGATAGCATGAGAATGCCGATG
CCCAAAAAAATGTGAACTTATGCCCTCATTGATCATTGTGAGAATGGTGCTG
TCATCCAGGATGCAACGCATTGCATACGATTGAGTCTCTTACCCACCCTTCCCAAG
CCATGTTTAGGTGGCATTGTGTTGACAGATATCAAATTCATTTTGGTATAAGCT
25 GCTTGAGTTATGTATTGGCTGGTTTTGTAAGTGTGCTTGGACCTTCTATCATT
AATGACAGACAAGCTGATCTCTCGGTTGCGAAAAAAAAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de OsPP2A-1 de *Oryza sativa* (SEQ ID NO: 24)

MEESVGSRRGGGGGLDAQDEQLMECRPLSEPEVKTLCEAKEILMEESNVQPVKSPVT
ICGDIHGQFHDLVELFRIGGKCPDTNYLFMGDYVDRGYYSVETVLLVALRVRYPRQRI
30 TILRGNHESRQITQVYGFYDECLRKYGSANVWKIFTDLFDYFPLTALVESEIFCLHGGL
SPSDDNLDVRSRLDRVQVEVPHEGPMCDLLWSDPDDRCGWGISPRGAGYTFGQDISEQF
NHTNNLKLVARAHQLVMEGYNWAHEQKVVT SAPNYCYRCGNLVLASLLEVDDCRN
HTFIQFEPAPRRGEPDVTRRTPDYFL*

Secuencia de nucleótidos de OsPP2A-2 de *Oryza sativa* (SEQ ID NO: 25)

35 GAGGCTTGAGCTCCACCTCCACCTCCTCCACCTCCAACCCCGATCCCCCGAAA
CCCTAGCCCTCTCCCCACCTCCTCGCCGGCGGCGAGCGGCGGCGGCGCGCGGC

GGGACCCGGAGCCCCAGTAGCGCCTCCTCCGTCCTCCCCTCCCTGAGGTGCGGG
 GGAGAGGATGCCGTCGTCGCACGGGGATCTGGACCGGCAGATCGCGCAGCTGCG
 GGAGTGCAAGCACCTGGCGGAGGGGAGGTGAGGGCGCTGTGCGAGCAGGCGAA
 GGCCATCCTCATGGAGGAGTGGAACGTGCAGCCGGTGCGGTGCCCCGTACGGTC
 5 TCGGGCGACATCCACGGCCAGTTCTACGACCTCATCGAGCTCTCCGCATCGGCG
 GCGAGGCGCCCGACACCAACTACCTCTTCATGGGCGACTACGTCGACCGTGGCTA
 CTA CTACTCAGTGGAGACTGTTTCGTTGTTGGTGGCTTTGAAAGTACGCTACAGAGATC
 GAATTACAATATTGAGAGGAAATCATGAGAGCAGACAAATCACTCAAGTGTACG
 GCTTCTACGATGAATGCTTGAGAAAGTATGGAAATGCAAATGTATGGAAATACTT
 10 TACAGACTTGTGTTGATTATTTGCCTCTCACAGCTCTTATAGAAAACCAGGTGTTCT
 GCCTTCACGGTGGTCTCTCTCCATCATTGGATACTTTAGATAACATCCGTGCTCTT
 GATCGTATAACAAGAGTTCCCTCATGAAGGACCCATGTGTGATCTTTTGTGGTCTGA
 CCCAGATGACAGATGCGGGTGGGAATTTACCGAGAGGAGCAGGTTATACATTT
 GGGCAAGATATCGCTCAACAGTTTAACCATACAAATGGTCTATCTCTCATCTCAA
 15 GGGCACATCAACTTGAATGGAAGGATTTAATTGGTGTGTCAGGACAAGAATGTTGT
 GACGGTCTTCAGTGCACCAAACACTACTGTTATCGCTGTGGTAACATGGCTGCAATTC
 TTGAGATTGGCGAAAACATGGATCAGAACTTCCCTCAATTTGATCCAGCTCCTCG
 GCAAATTGAACCAGACACAACACGCAAGACTCCCGACTACTTTTTGTAATTTGTG
 GTGTTGACAATTTTAACTCACCTGTGTTGATGCTCCTCTCCTCCGCGGTGTCGGGG
 20 TCTGTAGATCTTCTGTCTTAGATACGGGTTCCACGAGCCCGGCTGTATGTCTCTC
 AATTCTTTTGTGTTGGAGATTTTGTGCTGCTTCTCAACCTTTATACAAGACGTTAAA
 AGTTACATGCACTGGATTTTTTTCTCC

Secuencia deducida de aminoácidos de OsPP2A-2 de *Oryza sativa* (SEQ ID NO: 26)

MPSSHGDLDRQIAQLRECKHLAEGEVRALCEQAKAL MEEWNVQPVRCPVTVCGDIH
 25 GQFYDLIELFRIGGEAPDTNYLFMGDYVDRGYYSVETVSLLVALKVRYRDRITD RGN
 HESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLPLTALIENQVFCLHGGLSPSLD
 TLDNIRALDRIQEVPHGPMCDLLWSDPDDRCGWGISPRGAGYTFGQDIAQQFNHTN
 GLSLISRAHQLVMEGFNWCQDKNVVTSAPNYCYRCGNMAALLEIGENMDQNFLQ
 FDPAPRQD3PDTRKTPDYFL*

30 Secuencia de nucleótidos de OsPP2A-3 de *Oryza sativa* (SEQ ID NO: 27)

CCAATGCCAAGAACAAGAAATCAGAGAGCCCAAGCAAAAAATCCAATCCGAATC
 CCCTCCAAAAACCCACCAACAATCTCCTCCATCTCCAAGAACAAGAACAAGAG
 AGCATCCTAGAACATTAGGGGATAAAGGAGGAGGAGAAGAAGACCAAGAAAAG
 TTTCGAGAGGAGAGGGAGGAGAGTTCAAGAACTTGGAGAGAAGGGGATGGATCC
 35 GGTGTTGCTGGACGACATCATCCGGAGGCTTATCGAGGTGAAGAATCTGAAGCCG
 GGAAGAACGCGCAGCTTTCGGAGTCGGAGATTAAGCAGCTCTGCGCAACCTCCA

AGGAGATCTTCCTGAATCAGCCCAACCTGCTCGAGCTCGAGGCCCCCATCAA
 CTGCGGTGATGTTTCATGGACAGTATTCTGATCTCCTGAGGCTGTTTGATTATGGTG
 GGTATCCACCTCAGTCCAACCTATCTCTTCTGGGCGATTATGTGGACCGGGGAAA
 GCAAAGCCTTGAGACGATATGCCTTCTTTTGGCTTATAAGATCAAGTACCCTGAA
 5 AACTTCTTCTACTCAGAGGCAACCATGAATGTGCATCGGTCAACCGCATCTATG
 GATTTTATGACGAGTGCAAGCGCAGATTGAGTAAAACCTCTGGAAGACTTTTAC
 TGACTGTTTTAACTGCTTACCAGTGGCAGCATTGATAGATGAAAAGATTCTTTGTA
 TGCACGGAGGTCTTTCTCCAGAGTTGAATAAGCTGGATCAAATACTCAACCTCAA
 CCGCCCCACGGATGTGCCTGATACTGGGTTACTTTGTGATCTCCTTTGGTCCGATC
 10 CATCCAATGACGCACAAGGGTGGGCTATGAATGATCGAGGTGTCTCATATACATT
 CGGGCCAGACAAAGTGTCTGAATTTCTTGAGAAGCATGATTTAGACCTCATCTGT
 CGAGCCCATCAGGTTGTGGAAGATGGGTACGAGTTCTTTGCTAACCGCCAACTTG
 TAACAATATTCTCGGCCCTAATTACTGTGGAGAATTTGATAATGCTGGTGCCATG
 ATGAGTGTAGATGATACACTGATGTGCTCTTTTCAAATACTAAAACCAGCGAGGA
 15 AAATGTTGGGTGGTTCACGAATTCAAATCCGGCTTCAAGTCACTGAGAGGGTG
 GTGACGATGAGCAAAGCTGTGATCTGATCTGCTGGCGCATGTCTTCTACAGCGGC
 TGCGACTAACCGGCATTTTCGCCTACAGCTCGGGTCCATAAACAGCGAAGCAGAT
 AGAAATGTGTACAACCTTTCCAGCCGATGGAAGTGTACATCATCGTTCATGTTGGA
 TTAACACTTGTGTAATGTATTATTGGTTTTACCATGCGGATCTCTTATCATATGA
 20 GAGGATGTGAATGAAAACCTGTTCTCCCGTCTCCCCCTAAATTCAGAAAAGTTC
 AGACAGAAGGACTCCAATAAAAATAGCTAGAATCGAATGCTTTTGAACCAAAAA
 AAAAAAAAAAAAAA

Secuencia deducida de aminoácidos de OsPP2A-3 de *Oryza sativa* (SEQ ID NO: 28)

MDPVLLDDIIRLFFIVKLNKPKGKNAQLSESEIKQLCATSKEFLNQPNNLELEAPACGD
 25 VHGQYSDLLRLFDYGGYPPQSNYFLGDYVDRGKQSLETICLLLAYKFKYPENFFLLR
 GNHECASVNRIFYGYDECKRRFSVKLWKFTDFCNCLPVAALLDEKXLCMHGGLSPEL
 NKLDQELNLRPTDVPDGLLCDLLWSDPSNDAQGWAMNDRGVSYTFGPKVSEFL
 EKHDLDLICRAHQVVEDGYEFFANRQLVTFDFAPNYCGEFDNAGAMMSVDDTLMCSF
 QE KPARKMLGGSTNSRSGFKSLRGW*

30 Secuencia de nucleótidos de OsPP2A-4 de *Oryza sativa* (SEQ ID NO: 29)

GCCGCGCCGAGATCTAGGGTTGGGCGCGCGGACGCCCCCCCCCGCGGCGAGG
 AGGATGAGCAGCCCCATGGCGGCCTCGACGACCAGATCGAGCGCCTCATGCAGT
 GCAAGCCCCTCCCCGAGCCCAGGTCAGAGCACTTTGCGAGAAGGCAAAAGAGA
 TATTGATGGAGGAGAGCAACGTTCAACCTGTAAAGAGTCCTGTTACAATATGTGG
 35 TGATATTCATGGGCAGTTTCATGACCTTGCAAGACTGTTCCGAATCGGTGGAAG
 TGCCCAGATACAACTACTTGTATTATGGGAGATTACGTGGATCGTGGTTATTATTC

TGTTGAAACTGTCACGCTTTTGGTGGCTTTAAAGGTTTCGTTATCCTCAGCGAATTA
 CTATTCTCAGAGGAAACCACGAAAGTCGACAGATCACTCAAGTTTATGGATTCTA
 TGACGAGTGCTTAAGGAAGTACGGGAATGCAAATGTGTGGAAAACCTTTACAGAT
 CTCTTCGATTACTTCCCCTTGACAGCATTGGTTGAGTCAGAAATATTTTGCCTGCA
 5 TGGTGGATTATCGCCATCCATTGAGACACTTGATAACATACGTAACCTTCGATCGTG
 TCCAAGAAGTTCCTCATGAAGGGCCCATGTGTGATCTTCTGTGGTCTGATCCAGA
 CGATCGATGTGGTTGGGGTATTTCTCCTCGAGGTGCTGGATACACCTTCGGGCAG
 GATATATCAGAGCAGTTCACCATAACCAATAATTTAAGACTTATTGCTAGAGCTC
 ACCAGTTGGTCATGGAGGGATTCAATTGGGCTCATGAGCAAAAAGTTGTTACCAT
 10 ATTTAGTGACCTAATTATTGCTATCGCTGTGGGAACATGGCATCAATCTTGAAG
 TTGATGATTGCAGGGAGCATACATTCATCCAGTTTGGAGCCAGCCCCAAGAAGGGG
 AGAGCCAGATGTAACCTCGTAGAACACCTGACTATTTCTGTGATGTAAAAGTGGT
 GGACTGTCTCTGCAGCAAATGTTTGATAGCTAGCTGGGAGGATTCATCGTGTCTC
 ACTTATCTCTAATTGGCTGATGCTTGGCTTGGGGCTGCAGTGGTACTCGAAGC
 15 ATCAAGTAGCAAATTTGTATTATGAAAGGAAAACCTATTCTCTTTGTATTCAATTTG
 TTCGCCTTTCTTCCCACAAATTTACCTAATTTCTTTTTCTTTTTTCATGATCCT
 TGTAGAGATGAACAATGTAGTTGTATGGCTCCCTGTTGAGCCGGTAGGTCTTTCTG
 AGTACATCTTGATTTGCCGTACATAATTGCTTGAAAAACAAGTATTAGAATCTTT
 GTGACCAAAAAAAAAAAAAAAAAAAAA

20 Secuencia deducida de aminoácidos de OsPP2A-4 de *Oryza sativa* (SEQ ID NO: 30)
 MSSPHGGLDDQIERLMQCKPLPEPEVRALCEKAREILMEESNVQPVRSPVTICGDIHG
 QFHDLAELFRIGGRCPDNYLFMGDYVDRGYYSVETVLLVALRVRYRQRITILRGNH
 ESRQITQVYGFYDECLRKYGNANVWKTFTDLFDYFPLTALVESEDFCLHGGLSPSIETL
 DNRNFRVDRVQVEPHEGPMCDLLWSDPDDRCGWGISPRGAGYTFGQDISEQFNHTNNL
 25 RLIARAHQLVMEGFNWAHEQKVVTIFSAPNYCYRCGNMASILEVDDCREHTFIQFEPA
 PRRGEPDVTRRTPDYFL*

Secuencia de nucleótidos de OsPP2A-5 de *Oryza sativa* (SEQ ID NO: 31)
 GCCGTCACCGTCGCGCCAACTGCCGCAAACCGAATAAACCGAATCGATCTGAGAG
 AAGAAGAAGAAGAAGACGCGATCTCGGAGGTGGGAGCGAAACGAAACGATGCC
 30 GTCTCACGCGGATCTGGAACGACAGATCGAGCAGCTGATGGAGTGCAAGCCTCTG
 TCGGAGTCGGAGGTGAAGGCGCTGTGTGATCAAGCGAGGGCGATTCTCGTGGAG
 GAATGGAACGTGCAACCGGTGAAGTGCCCCGTCACCGTCTGCGGCGATATTCAG
 GCCAGTTTTACGATCTCATCGAGCTGTTTCGGATTGGAGGGAACGCACCCGATAC
 CAATTATCTCTTCATGGGTGATTATGTAGATCGTGGATACTATTCAGTGGAGACTG
 35 TTACACTTTTGGTGGCTTTGAAAGTCCGTTACAGAGATAGAATCACAATTCTCAGG
 GGAAATCATGAAAGTCGTCAAATACTCAAGTGTATGGCTTCTATGATGAATGCT

ES 2 381 031 T3

TGAGAAAATATGGAAATGCCAATGTCTGGAAATACTTTACAGACTTGTTTGATTA
TTTACCTCTGACTGCCCTCATTGAGAGTCAGATTTTCTGCTTGCATGGAGGTCTCT
CACCTTCTTTGGATACACTGGATAACATCAGAGCATTGGATCGTATACAAGAGGT
TCCACATGAAGGACCAATGTGTGATCTCTTGTGGTCTGACCCTGATGATCGCTGTG
5 GATGGGGAATATCTCCACGTGGTGCAGGATACACATTTGGACAGGATATAGCTGC
TCAGTTTAATCATACCAATGGTCTCTCCCTGATATCGAGAGCTCATCAGCTTGTTA
TGGAAGGATTCAATTGGTGCCAGGACAAAAATGTGGTGACTGTATTTAGTGCACC
AAATTACTGTTACCGATGTGGGAATATGGCTGCTATACTAGAAATAGGAGAGAAT
ATGGATCAGAATTTCTTCAGTTTGATCCAGCGCCCAGGCAAATTGAGCCTGACA
10 CCACACGCAAGACTCCAGATTATTTTTATAATTTTCATTTATCTGCCTGTTTGTAGT
TACTGCTCTCTGCCATTACTGTAGATGTGTCTTTAAGGAAAGGAGTTTTGCTGTTT
AAGTGGAGGGTGGTCATCAACATAATTCTTTCTTTTGGAGTTTACCTCCTGCTGCT
GCCGCTGCCGCTGCCTTATTTGTACAAGAAACCAATAGAAGTACACAAGCCACC
AATTGGGGTTGTATATTTTTGGGAGGAAGCGGTAATAACATGGTATATCTTGTCT
15 GTAATCCTTTTTCTTTAAATTGAATCTCAAGTTAGAGAGCAAAAAAAAAAAAAAAAA
AAA

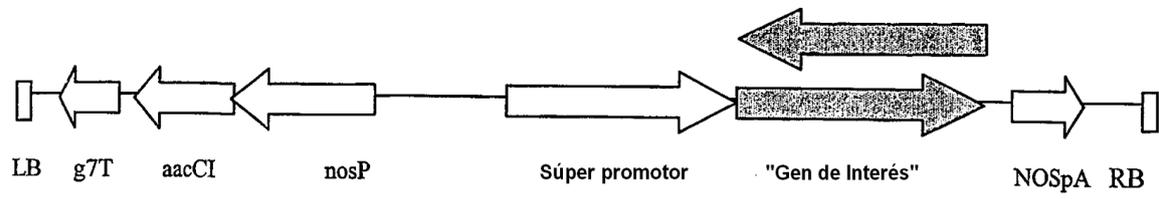
Secuencia deducida de aminoácidos de OsPP2A-5 de *Oryza sativa* (SEQ ID NO: 32)

MPSHADLERQDSQLMECRPLSESEVKALCDQARAILVEEWNVQPVKCPVTVCGDFFLG
QFYDLIELFRIGGNAPDTNYLFMGDYVDRGYYSVETVLLVALKVRYRDRITPI RGNH
20 ESRQITQVYGFYDECLRKYGNANVWRYFTDLFDYLPLTALFFISQFFCLHGGLSPSLDTL
DNIRALDRIQEVPHGPMCDLLWSDPDDRCGWGISPRGAGYTFGQDIAAQFNHTNGL
SLISRAHQLVMEGFNWCQDKNVVTVFSAPNYCYRCGNMAAILEIGENMDQNFLQFDP
APRQLEPDTRKTPDYFL*

REIVINDICACIONES

1. Un ácido nucleico que codifica para la Proteína Relacionada con el Estrés de la Proteína Fosfatasa en donde el ácido nucleico que codifica para PPSRP comprende un polinucleótido seleccionado de entre el grupo que consiste de
- 5 a) el polinucleótido como se define en la SEQ ID NO: 2;
- b) un polinucleótido que codifica para una Proteína Relacionada con el Estrés de la Proteína Fosfatasa (PPSRP), en donde la PPSRP es Proteína Fosfatasa 2A-1 (PP2A-1) como se define en la SEQ ID NO: 3;
- c) un polinucleótido complementario a un polinucleótido de cualquiera de a) o b) anteriores; y
- 10 d) secuencias de ácido nucleico al menos 80 - 85% idénticas a la secuencia de la SEQ ID NO: 2 sobre toda la longitud de la región de codificación de a secuencia de la SEQ ID NO: 2.
2. Un vector de expresión recombinante que comprende el ácido nucleico de la reivindicación 1.
3. Un casete de expresión de la planta, en donde el ácido nucleico que codifica para PPSRP de la reivindicación 1 reside en el casete de expresión de la planta.
- 15 4. Una Proteína Relacionada con el Estrés de la Proteína Fosfatasa (PPSRP), en donde la PPSRP es una Proteína Fosfatasa 2A-1 (PP2A-1) seleccionada de entre el grupo que consiste de la secuencia como la definida en la SEQ ID NO: 3.
- 20 5. Una célula de una planta transgénica transformada por y que comprende al ácido nucleico que codifica para la Proteína Relacionada con el Estrés de la Proteína Fosfatasa (PPSRP) de la reivindicación 1 o el casete de expresión de la reivindicación 3, en donde la expresión de la secuencia de ácido nucleico en la célula de la planta trae como resultado una mayor tolerancia al estrés por sequía comparado con una variedad de tipo silvestre de la célula de la planta.
6. Una planta transgénica que comprende una célula de una planta de acuerdo con la reivindicación 5.
- 25 7. Un método para producir una planta transgénica que contiene un ácido nucleico que codifica para la Proteína Relacionada con el Estrés de la Proteína Fosfatasa (PPSRP) en donde la expresión del ácido nucleico en la planta resulta en mayor tolerancia al estrés por sequía comparado con una variedad de tipo silvestre de la planta que comprende: (a) transformar una célula de una planta con el vector de expresión de la reivindicación 2 o el casete de expresión de a reivindicación 3 que comprende un ácido nucleico que codifica para PPSRP, y (b) generar a partir de la célula de la planta una planta transgénica con una mayor tolerancia al estrés por sequía comparado con una variedad de tipo silvestre de la planta en donde el PPSRP es el PPSRP de acuerdo con a reivindicación 4 o el ácido
- 30 nucleico que codifica para PPSRP es el ácido nucleico que codifica para PPSRP de acuerdo con la reivindicación 1.
8. El método de la Reivindicación 7 o la planta de la reivindicación 6, en donde la planta es una planta monocotiledónea o una dicotiledónea.
- 35 9. El método de la Reivindicación 7 u 8 o la planta de la reivindicación 6 u 8 en donde la planta se selecciona de entre el grupo que consiste de maíz, trigo, centeno, avena, triticale, arroz, cebada, soja, maní, algodón, colza, canola, mandioca, pimiento, girasol y tagetes, plantas solanáceas, patata, tabaco, berenjena, tomate, especies Vicia, arveja, alfalfa, café, cacao, té, especies Salix, palma oleaginosa, cocotero, pastos perennes y cultivos forrajeros.
- 40 10. Semilla de una planta transgénica producida por la planta de cualquiera de las reivindicaciones 6, 8 ó 9, en donde la semilla comprende al ácido nucleico que codifica para PPSRP de la reivindicación 1, o la proteína PPSRP de la reivindicación 4 o el casete de expresión de la reivindicación 3 y en donde la semilla es de una línea genéticamente pura para una mayor tolerancia al estrés por sequía comparado con una variedad de tipo silvestre de la semilla.
11. Un método para incrementar la tolerancia al estrés por sequía de una planta que comprende incrementar la expresión de la Proteína Relacionada con el Estrés de la Proteína Fosfatasa (PPSRP) de la reivindicación 4.

Figura 1



pBPSJH001

Figura 2

		1		40
PpPP2A-1	(1)	--MSVPP	ISS-NGOLDTQIAQLI	QCKPLS-----EVE
OsPP2A-5	(1)	-----MPSHADLERQIEQLMECKPLS-----ESE		
OsPP2A-4	(1)	-----MSSPHGGLDDQIERLMQCKPLP-----EPE		
OsPP2A-3	(1)	-----MDPVLLDDIIRRLIEVKNLKPGKNAQLSESE		
OsPP2A-2	(1)	-----MPSSHGDLDRQIAQLRECKHLA-----EGE		
OsPP2A-1	(1)	MEEVSGSRGGGGGLDAQIEQLMECRPLS-----EPE		
Consenso	(1)	M S H G L D Q I E Q L M E C K P L S		E E
		41		80
PpPP2A-1	(30)	VRGLCEKAKEILMRENNVQPVKCPVTICGDIHGQFHDIAE		
OsPP2A-5	(25)	VKALCDQARATLMEEWNVQPVKCPVTICGDIHGQFYDLIE		
OsPP2A-4	(26)	VRALCEKAKEILMEESNVQPVKSPVTICGDIHGQFHDIAE		
OsPP2A-3	(32)	IKQLCATSKETFLNQPNLELEAPTKICGDVHGQYSDLER		
OsPP2A-2	(26)	VRALGEQAKATLMEEWNVQPVRCPTVTCGDIHGQFYDLIE		
OsPP2A-1	(33)	VKTLCEKAKEILMEESNVQPVKSPVTICGDIHGQFHDLVE		
Consenso	(41)	VKALCEKAKEILMEE NVQPVKCPVTICGDIHGQFHDLIE		
		81		120
PpPP2A-1	(70)	LFRIGCMCPDTNYLFMGDYVDRGYYSVETATLLVALKVRV		
OsPP2A-5	(65)	LFRIGGNAPDTNYLFMGDYVDRGYYSVETVTLIIVALKVRV		
OsPP2A-4	(66)	LFRIGGKCPDTNYLFMGDYVDRGYYSVETVTLIIVALKVRV		
OsPP2A-3	(72)	LFDYGGYPPOSNYLFLGDYVDRGKOSLETICLLAYKIKY		
OsPP2A-2	(66)	LFRIGGEAPDTNYLFMGDYVDRGYYSVETVSLIIVALKVRV		
OsPP2A-1	(73)	LFRIGGKCPDTNYLFMGDYVDRGYYSVETVTLIIVALKVRV		
Consenso	(81)	LFRIGG CPDTNYLFMGDYVDRGYYSVETVTLIIVALKVRV		
		121		160
PpPP2A-1	(110)	PDRITILRGNHESRQITQVYGFYDECLRKYGNANVWKKIFT		
OsPP2A-5	(105)	RDRITILRGNHESRQITQVYGFYDECLRKYGNANVWKKYFT		
OsPP2A-4	(106)	PQRITILRGNHESRQITQVYGFYDECLRKYGNANVWKTFT		
OsPP2A-3	(112)	PENFFLIRGNHECASVNRITYGFYDECLRRRES-VKLEWKTFT		
OsPP2A-2	(106)	RDRITILRGNHESRQITQVYGFYDECLRKYGNANVWKKYFT		
OsPP2A-1	(113)	PQRITILRGNHESRQITQVYGFYDECLRKYGSANVWKKIFT		
Consenso	(121)	PDRITILRGNHESRQITQVYGFYDECLRKYGNANVWK FT		
		161		200
PpPP2A-1	(150)	DLEDYFPLTALVESEIFCLHGGLSPSIDSLDHIRDLDRVQ		
OsPP2A-5	(145)	DLEDYLPLTALTESQIFCLHGGLSPSIDTLDNIRALDRIQ		
OsPP2A-4	(146)	DLEDYFPLTALVESEIFCLHGGLSPSLETLDNIRNFDRVQ		
OsPP2A-3	(151)	DCFNCLPVAALIDEKILCMHGGLSPELNKLDQILNLNRPT		
OsPP2A-2	(146)	DLEDYLPLTALIEENQVFCFHGGLSPSIDTLDNIRALDRIQ		
OsPP2A-1	(153)	DLEDYFPLTALVESEIFCLHGGLSPSIDNLDVRSFLDRVQ		
Consenso	(161)	DLEDYLPLTALIESEIFCLHGGLSPSIDTLDNIRALDRVQ		
		201		240
PpPP2A-1	(190)	EVPHEGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDISE		
OsPP2A-5	(185)	EVPHEGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIAA		
OsPP2A-4	(186)	EVPHEGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDISE		
OsPP2A-3	(191)	DVPDTGLECDLLWSDPSNDAQWAMNDRGVSYTFGPDKVS		
OsPP2A-2	(186)	EVPHEGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIAQ		
OsPP2A-1	(193)	EVPHEGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDISE		
Consenso	(201)	EVPHEGPMCDLLWSDP DDRCGWGISPRGAGYTFGQDISE		

Figura 2 continuación

		241		280
PpPP2A-1	(229)	QFNHNNLKLVARAHQLVMEGYNWCHEHKVVTIIFSA	NYC	
OsPP2A-5	(224)	QFNHTNGLSLISRAHQLVMEGENWCODKNVTVFSA	NYC	
OsPP2A-4	(225)	QFNHTNRLRIARAHQLVMEGENWAHEQKVVTIIFSA	NYC	
OsPP2A-3	(231)	EFLEKHDLDLICRAHOVVEDGYEFFANRQLVTIIFSA	NYC	
OsPP2A-2	(225)	QFNHTNGLSLISRAHQLVMEGENWCODKNVTVFSA	NYC	
OsPP2A-1	(232)	QFNHTNNLKLVARAHQLVMEGYNWAHEQKVVTIIFSA	NYC	
Consenso	(241)	QFNHTNNLKLVARAHQLVMEGENWAHEKVVTVIIFSA	NYC	
		281		320
PpPP2A-1	(269)	YRCGNMASILEVDDNMGHTFIQFEPAPRRGEPDVTRRTPD		
OsPP2A-5	(264)	YRCGNMAAILEIGENMDONFLQFDPAAPRQIEPDTTRKTPD		
OsPP2A-4	(265)	YRCGNMASILEVDDCREHTFIQFEPAPRRGEPDVTRRTPD		
OsPP2A-3	(271)	GEFDNAGAMMSVDDTLMCSFQILKPAKMLGGSTNSKSGF		
OsPP2A-2	(265)	YRCGNMAAILEIGENMDONFLQFDPAAPRQIEPDTTRKTPD		
OsPP2A-1	(272)	YRCGNMASILEVDDCRNHTFIQFEPAPRRGEPDVTRRTPD		
Consenso	(281)	YRCGNMAAILEVDDNMDHTFIQFEPAPRRGEPDVTRKTPD		
		321		
PpPP2A-1	(309)	YFL----		
OsPP2A-5	(304)	YFL----		
OsPP2A-4	(305)	YFL----		
OsPP2A-3	(311)	KSLRGW-		
OsPP2A-2	(305)	YFL----		
OsPP2A-1	(312)	YFL----		
Consenso	(321)	YFL		

Figura 3

		1		40
PpPP2A-1	(1)	---MSVPP	SSNG---OLDTQIAQT	IQCK-----PLSEVE
GmPP2A-5	(1)	-----	MPSHADLERQIEQL	MDCK-----PLSESE
GmPP2A-4	(1)	-----	MEQSLDDI	INRLL
GmPP2A-3	(1)	---MGANSML	SESSHLD	DDQISQLM
GmPP2A-2	(1)	MSTQGQVIL	DEAVLDDI	IRRLTEVRLARPG-KQV
GmPP2A-1	(1)	-----	MDLDQWISKV	KEGQ-----HLL
BnPP2A-3	(1)	MTQQGQGS	MDPAVLDDI	IRRLDYRNP
BnPP2A-2	(1)	-----	MDENLDDI	IRRLLETNNGK---
BnPP2A-1	(1)	-----	MPETGDI	DRQIEQLMECK-----
Consenso	(1)		M DIDR I L K	LSESE
		41		80
PpPP2A-1	(30)	VRGLCEKAKE	EILMRENNVQPVKCP	VTVICGDIHGQFHD
GmPP2A-5	(25)	VKALCDOART	ILVEEWNVQPVKCP	VTVICGDIHGQFYD
GmPP2A-4	(31)	IRQLCVVSREI	FLOQPNLLELEAP	TKICGDVHGQYSD
GmPP2A-3	(32)	VRGLCEKAKE	EILMDES	NVQPVKSPVTICGDIHGQFHD
GmPP2A-2	(40)	IKQLCVASRDI	FINQPNLLELEAP	TKICGDIHGQYSD
GmPP2A-1	(21)	LQLLCEYVKEI	LTEESNVQPVNSP	VTVICGDIHGQFHD
BnPP2A-3	(41)	IRQLCVVSREI	FLOQPNLLELEAP	TKICGDIHGQYSD
BnPP2A-2	(29)	IRQLCVSASKE	VFLSQPNLLELEAP	TKICGDVHGQFPD
BnPP2A-1	(25)	VKTLCEQARA	ILVEEWNVQPVKCP	VTVICGDIHGQFYD
Consenso	(41)	IR LCE ARE	ILL E NVQPV	APVTICGDIHGQFHD
		81		120
PpPP2A-1	(70)	LEFRIGGMC	PDTNYL	FMGDYVDRGYYSVET
GmPP2A-5	(65)	LEFRIGGNAP	PDTNYL	FMGDYVDRGYYSVET
GmPP2A-4	(71)	LFEYCGLP	PPAANYL	FLGDYVDRGKQSL
GmPP2A-3	(72)	LEFRIGGKC	PDTNYL	FMGDYVDRGYYSVET
GmPP2A-2	(80)	LFEYCGLP	PPAANYL	FLGDYVDRGKQSL
GmPP2A-1	(61)	LFQTCGHV	PEITNYL	FMGDFVDRGYNS
BnPP2A-3	(81)	LFEYCGLP	PPAANYL	FLGDYVDRGKQSL
BnPP2A-2	(69)	LFEYCGYP	PPAANYL	FLGDYVDRGKRST
BnPP2A-1	(65)	LEFRIGGSS	PDTNYL	FMGDYVDRGYYSVET
Consenso	(81)	LF GG	PDTNYL	FMGDYVDRGY SLETITLL
		121		160
PpPP2A-1	(110)	PDRIITLL	LRGNHESRQITQVYGFY	DECLRKYGNANVVK
GmPP2A-5	(105)	RDRIITLL	LRGNHESRQITQVYGFY	DECLRKYGNANVVK
GmPP2A-4	(111)	PENFFLL	LRGNHECASINRIYGFY	DECLRKYGNANVVK
GmPP2A-3	(112)	PQRIITLL	LRGNHESRQITQVYGFY	DECLRKYGNANVVK
GmPP2A-2	(120)	PENFFLL	LRGNHECASINRIYGFY	DECLRKYGNANVVK
GmPP2A-1	(101)	PANITLL	LRGNHESRQITQVYGFY	DECLRKYGNANVVK
BnPP2A-3	(121)	PENFFLL	LRGNHECASINRIYGFY	DECLRKYGNANVVK
BnPP2A-2	(109)	KLNFFLL	LRGNHECASINRVYGFY	DECLRKYGNANVVK
BnPP2A-1	(105)	RDRIITLL	LRGNHESRQITQVYGFY	DECLRKYGNANVVK
Consenso	(121)	PDNITLL	LRGNHESRQITQVYGFY	DECLRKYGNANVVK

Figura 3 Continuación

		161	200
PpPP2A-1	(150)	DLFDYFPLTALVESEIFCLHGGLSPSLD	SLDHIIRDLDLRVQ
GmPP2A-5	(145)	DLFDYLPLTALIESQIFCLHGGLSPSLD	TLDNIRALDRIQ
GmPP2A-4	(150)	DCFNCLPVAALVDEKILCMHGGLSPDL	NNLDQIRNLRPT
GmPP2A-3	(152)	DLFDFPPLTALVESEIFCLHGGLSPS	LETLDNIRNFDLRVQ
GmPP2A-2	(159)	DCFNCLPVAALIDDKILCMHGGLSPEL	TNLDEIRNLRPT
GmPP2A-1	(141)	DVFDYLTLSALIDGTVLGVHGGLSPDI	RTIDQIRVLDNRN
BnPP2A-3	(160)	DSENCLPVAAVIDDKILCMHGGLSPDL	TVEQIKNIKRP
BnPP2A-2	(148)	DCFNCLPVAALIDDKILCMHGGLSPDL	KTLDDIRRI
BnPP2A-1	(145)	DLFDYLPLTALIESQIFCLHGGLSPSLD	TLDNIRSLDRIQ
Consenso	(161)	DLFDYLPLTALID ILCLHGGLSPDL	TLDNIR LDR
		201	240
PpPP2A-1	(190)	EVPHEGPMCDLLWSDP-DDRCGWGISPRG	AGYTFEQDISE
GmPP2A-5	(185)	EVPHEGPMCDLLWSDP-DDRCGWGISPRG	AGYTFEQDIAA
GmPP2A-4	(190)	DVPDTGLLCDLLWSDPSKDVQGWGMNDR	GVSYTFEGADKVS
GmPP2A-3	(192)	EVPHEGPMCDLLWSDP-DDRCGWGISPRG	AGYTFEQDISE
GmPP2A-2	(199)	AIPDTGLLCDLLWSDPGRDVKGWGMNDR	GVSYTFEGPDKVA
GmPP2A-1	(181)	EIPHEGPFCDLMWSDP-EDLETWAVSPRG	AGWLFGRVITS
BnPP2A-3	(200)	DVPDSGLLCDLLWSDPSKDVKGWGMNDR	GVSYTFEGPDKVA
BnPP2A-2	(188)	DVPDQGVLCDLLWADPKETQGWGENDR	GVSYTFEGPDKVA
BnPP2A-1	(185)	EVPHEGPMCDLLWSDP-DDRCGWGISPRG	AGYTFEQDIAT
Consenso	(201)	EVPHEGPLCDLLWSDP DDV GWGISPRG	AGYTFG DI A
		241	280
PpPP2A-1	(229)	QFNHNNNLKLVARAHQLVMEGYNWGHEHK-	VVTIFSAAPNY
GmPP2A-5	(224)	QFNHTNGLSLISRAHQLVMEGENWCODKN-	VVTVFSAPNY
GmPP2A-4	(230)	QFLQKHDLDLVCRAHOVVEDGYEFFANRO-	LVTIFSAAPNY
GmPP2A-3	(231)	QFNHTNSLKLIVARAHQLVMDGFNWAHEOK-	VVTIFSAAPNY
GmPP2A-2	(239)	EFLTKHDLDLIVCRAHOVVEDGYEFFADRO-	LVTIFSAAPNY
GmPP2A-1	(220)	EFNHINNLDLVCRAHQLVQEGCLKYMFQDKGL	VTVWSAPNY
BnPP2A-3	(240)	EFLIKNDMDLIVCRAHOVVEDGYEFFADRO-	LVTIFSAAPNY
BnPP2A-2	(228)	EFLQTHDLDLVCRAHOVVEDGYEFFAKRO-	LVTIFSAAPNY
BnPP2A-1	(224)	QFNHTNGLSLISRAHQLVMEGYNWCQEK-	VVTVFSAPNY
Consenso	(241)	QFNH N LDLICRAHQLV DGY F DRQ	LVTIFSAAPNY
		281	320
PpPP2A-1	(268)	CYRCGNMASILEVDDNMGHTFIQFEPAPRR	GEPDVTRRTP
GmPP2A-5	(263)	CYRCGNMAAILETGENMDQNFLOFDPAPRQ	IEPDTTRKTP
GmPP2A-4	(269)	CGEFDNAGAMMSVDETLMCSFQILKPADKKA	KLNFSTTT
GmPP2A-3	(270)	CYRCGNMASILEVDDCKGHTFIQFEPAPRR	GEPDVTRRTP
GmPP2A-2	(278)	CGEFDNAGAMMSVDENLMCSFQILKPAEKKS	KFVMSNKM-
GmPP2A-1	(260)	CYRCGNVASILSFNENMEREVKFFTEETEEN	NQMRGPRTGV
BnPP2A-3	(279)	CGEFDNAGAMMSVDES-	
BnPP2A-2	(267)	CGEFDNAGAMMSVDDSLTCSFQILKSTEKKG	RFGYNNNVH
BnPP2A-1	(263)	CYRCGNMAAILETIDENMDQNFLOFDPAPRQ	VEPETTRKTP
Consenso	(281)	CYRCGNMAAILSVDENM SF F PA RK	TRKT

Figura 3 Continuación

		321	336
PpPP2A-1	(308)	DYFL-----	
GmPP2A-5	(303)	DYFL-----	
GmPP2A-4	(309)	AKPGNSPAGVKVGRY-	
GmPP2A-3	(310)	DYFL-----	
GmPP2A-2	(317)	-----	
GmPP2A-1	(300)	PYFL-----	
BnPP2A-3	(295)	-----	
BnPP2A-2	(307)	RPGTPPHKGGKGG---	
BnPP2A-1	(303)	DYFL-----	
Consenso	(321)	YFL	

Figura 4

	1	40
PpPP2A-1	(1) ---MSVPPISSNGOLDTOIAQLTQCKPLS-----	EV
OsPP2A-5	(1) -----MPSHADLERQIEQLMECKPLS-----	ES
OsPP2A-4	(1) -----MSSPHGGLDDQIERLMQCKPLP-----	EP
OsPP2A-3	(1) -----MDPVLLDDIIRRLIEVKNLKPG-KNAQLSES	ES
OsPP2A-2	(1) -----MPSSHGGLDROIAQLRECKHLA-----	EG
OsPP2A-1	(1) MEESVGSRRGGGGGLDAQIEQLMECRPLS-----	EP
GmPP2A-5	(1) -----MPSHADLERQIEQLMDCKPLS-----	ES
GmPP2A-4	(1) -----MEQSLDDIIRRLIEVPTL-PA-KQVQLSES	ES
GmPP2A-3	(1) -MGANSMLSESSHLDLDDQISQLMQCKPLS-----	EQ
GmPP2A-2	(1) -MSTQGQVIIDEAVLDDIIRRLTEVRLARPG-KQVQLSES	ES
GmPP2A-1	(1) -----MDLDQWISKVKEGQHLL-----	ED
BnPP2A-3	(1) -MTQQGQGSMDPAVLDDIIRRLIDYRNPKPGTKQVMLNES	ES
BnPP2A-2	(1) -----MDENLDDIIRRLIETNNG-----KQVKLL	EA
BnPP2A-1	(1) -----MPETGDIIDROIEQLMECKALS-----	EA
Consenso	(1) G LDDQI QLMECK LS	ES
	41	80
PpPP2A-1	(29) EVRGLCEKAKEILMRENNVQPVKCPVTICGDIHGQFHDLA	
OsPP2A-5	(24) EVKALCDQARAILMEEWNVQPVKCPVTICGDIHGQFYDLI	
OsPP2A-4	(25) EVRALCEKAKEILMEESNVQPVKSPVTICGDIHGQFHDLA	
OsPP2A-3	(31) ETKOLGATSKIEIFLNQPNLLELEAPTKICGDVHGQYSDLL	
OsPP2A-2	(25) EVRALCEQAKAILMEEWNVQPVRCPTVTCGDIHGQFYDLI	
OsPP2A-1	(32) EVKTLCEKAKEILMEESNVQPVKSPVTICGDIHGQFHDLM	
GmPP2A-5	(24) EVKALCDQARTILMEEWNVQPVKCPVTICGDIHGQFYDLI	
GmPP2A-4	(30) ETRQLCVVSREIFLQQPNLLELEAPTKICGDVHGQYSDLL	
GmPP2A-3	(31) QVRGLCEKAKEILMDES NVQPVKSPVTICGDIHGQFHDLA	
GmPP2A-2	(39) ETKQLGVA SRDI FLNQPNLLELEAPTKICGDVHGQYSDLL	
GmPP2A-1	(20) ELQQLCEYVKEILTEESNVQPVNSPVTICGDIHGQFHDLM	
BnPP2A-3	(40) ETRQLCSVSREIFLQQPNLLELEAPTKICGDVHGQYSDLL	
BnPP2A-2	(28) ETRQLCSASKEVFLSQPNLLELEAPTKICGDVHGQFPDLL	
BnPP2A-1	(24) EVKTLCEQARAILMEEWNVQPVKCPVTICGDIHGQFYDLI	
Consenso	(41) EVR LCE AKEILMEE NVQPVKAPVTICGDIHGQFHDLL	
	81	120
PpPP2A-1	(69) ELFRIGGMCPDTNYLFMGDYVDRGYYSVETATLLVALKVR	
OsPP2A-5	(64) ELFRICGNAPDTNYLFMGDYVDRGYYSVETVILLVALKVR	
OsPP2A-4	(65) ELFRIGGKCPDTNYLFMGDYVDRGYYSVETVILLVALKVR	
OsPP2A-3	(71) RLFDYGGYPPOSNYLFLGDYVDRGKQSLETICLLLAYKIK	
OsPP2A-2	(65) ELFRIGGEAPDTNYLFMGDYVDRGYYSVETVSLVALKVR	
OsPP2A-1	(72) ELFRIGGKCPDTNYLFMGDYVDRGYYSVETVILLVALKVR	
GmPP2A-5	(64) ELFRICGNAPDTNYLFMGDYVDRGYYSVETVILLVALKVR	
GmPP2A-4	(70) RLFEYGGLPPEANYLFLGDYVDRGKQSLETICLLLAYKIK	
GmPP2A-3	(71) ELFRIGGKCPDTNYLFMGDYVDRGYYSVETVSLVALKVR	
GmPP2A-2	(79) RLFEYGGLPPTANYLFLGEYVDRGKQSLETICLLLAYKIK	
GmPP2A-1	(60) KLFQTCGHVPEPTNYLFMGDFVDRGYNSLEVFILLLLKAR	
BnPP2A-3	(80) RLFEYGGLPPEANYLFLGDYVDRGKQSLETICLLLAYKIK	
BnPP2A-2	(68) RLFEYGGYPPEANYLFLGDYVDRGKRSIETICLLLAYKIK	
BnPP2A-1	(64) ELFKIGGSSPDTNYLFMGDYVDRGYYSVETVSLVALKVR	
Consenso	(81) ELFRIGG PDTNYLFMGDYVDRGYYSVETVILLVALKVR	

Figura 4 Continuación

		121	160
PpPP2A-1	(109)	YPRDRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKIF	
OsPP2A-5	(104)	YRDRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKYF	
OsPP2A-4	(105)	YPQRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKTF	
OsPP2A-3	(111)	YPENFFLLRGNHESRQITQVYGFYDECKRRFS-VKLWKTF	
OsPP2A-2	(105)	YRDRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKYF	
OsPP2A-1	(112)	YPQRIITILRGNHESRQITQVYGFYDECLRKYGSANVWKIF	
GmPP2A-5	(104)	YRDRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKYF	
GmPP2A-4	(110)	YPENFFLLRGNHESRQITQVYGFYDECKRRFN-VRLWKTF	
GmPP2A-3	(111)	YPQRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKTF	
GmPP2A-2	(119)	YPENFFLLRGNHESRQITQVYGFYDECKRRFN-VRLWKAF	
GmPP2A-1	(100)	YPANITILRGNHESRQITQVYGFYDECORRYGNANAWRYC	
BnPP2A-3	(120)	YPENFFLLRGNHESRQITQVYGFYDECKRRFS-VRLWKVF	
BnPP2A-2	(108)	YKLNFFLLRGNHESRQITQVYGFYDECKRRYN-VRLWKSF	
BnPP2A-1	(104)	YRDRIITILRGNHESRQITQVYGFYDECLRKYGNANVWKHF	
Consenso	(121)	YPRDRIITILRGNHESRQITQVYGFYDECLRKYGNANVWK F	
		161	200
PpPP2A-1	(149)	TDLFDYFPLTALVESEIFCLHGGLSPSIDSLDHIRDLDRV	
OsPP2A-5	(144)	TDLFDYLPLTALIESQIFCLHGGLSPSLDTLDNIRALDRI	
OsPP2A-4	(145)	TDLFDYFPLTALVESEIFCLHGGLSPSIETLDNIRNFDRV	
OsPP2A-3	(150)	TDCFNCLPVAALIDEKILCMHGGLSPELNKLDIILNLNRP	
OsPP2A-2	(145)	TDLFDYLPLTALIEENQVIFCLHGGLSPSLDTLDNIRALDRI	
OsPP2A-1	(152)	TDLFDYFPLTALVESEIFCLHGGLSPSIDNLDVRSLSDRV	
GmPP2A-5	(144)	TDLFDYLPLTALIESQIFCLHGGLSPSLDTLDNIRALDRI	
GmPP2A-4	(149)	TDCFNCLPVAALVDEKILCMHGGLSPDLNLDIIRNLQRP	
GmPP2A-3	(151)	TDLFDYFPLTALVESEIFCLHGGLSPSIETLDNIRNFDRV	
GmPP2A-2	(158)	TDCFNCLPVAALIDDKILCMHGGLSPELTNLDEIRNLPRP	
GmPP2A-1	(140)	TDVFDYLTLSAIDGTMLCVHGGLSPDIRTIDQIRVIDRN	
BnPP2A-3	(159)	TDCFNCLPVAAVIIDDKILCMHGGLSPDLTSVEQIKNTKRP	
BnPP2A-2	(147)	TDCFNCLPVAALIDDKILCMHGGLSPDLKTLDDIRRIIPRP	
BnPP2A-1	(144)	TDLFDYLPLTALIESQVIFCLHGGLSPSLDTLDNIRSLDRI	
Consenso	(161)	TDLFDYLPLTALIES IFCLHGGLSPSLDTLDNIR LDRI	
		201	240
PpPP2A-1	(189)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIS	
OsPP2A-5	(184)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIA	
OsPP2A-4	(185)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIS	
OsPP2A-3	(190)	TDVFDTGLLCDLLWSDPSNDAQWAMNDRGVSYTFGPDKV	
OsPP2A-2	(185)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIA	
OsPP2A-1	(192)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIS	
GmPP2A-5	(184)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIA	
GmPP2A-4	(189)	TDVFDTGLLCDLLWSDPSKDVQGWGMNDRGVSYTFGADKV	
GmPP2A-3	(191)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIS	
GmPP2A-2	(198)	TALPDTGLLCDLLWSDPGRDVKGWGMNDRGVSYTFGPDKV	
GmPP2A-1	(180)	CEIPHEGPFCDLMWSDP-EDIETWAVSPRGAGWLFGRMT	
BnPP2A-3	(199)	TDVFDSSGLLCDLLWSDPSKDVKGWGMNDRGVSYTFGPDKV	
BnPP2A-2	(187)	VDVFDQGVLCDLLWADPKETIQGWGENDRGVSYTFGPDKV	
BnPP2A-1	(184)	QEVPHGPMCDLLWSDP-DDRCGWGISPRGAGYTFGQDIA	
Consenso	(201)	QEVPHGPMCDLLWSDP DDRCGWGISPRGAGYTFGQDIA	

Figura 4 Continuación

		241	280
PpPP2A-1	(228)	EQFNHNNLKLVARAHQLVMEGYNWCHEHK-VVTIFSAPN	
OsPP2A-5	(223)	AQFNHTNGLSLISRAHQLVMEGFNWCQDKN-VVTVFSAPN	
OsPP2A-4	(224)	EQFNHTNRLRLIARAHQLVMEGFNWAHEQK-VVTIFSAPN	
OsPP2A-3	(230)	SEFLEKHDLDLICRAHQVVEDGYEFFANRQ-LVTIFSAPN	
OsPP2A-2	(224)	QQFNHTNGLSLISRAHQLVMEGFNWCQDKN-VVTVFSAPN	
OsPP2A-1	(231)	EQFNHTNKLKLVARAHQLVMEGFNWAHEQK-VVTIFSAPN	
GmPP2A-5	(223)	AQFNHTNGLSLISRAHQLVMEGFNWCQDKN-VVTVFSAPN	
GmPP2A-4	(229)	SQFLQKHDLDLVCRAHQVVEDGYEFFANRQ-LVTIFSAPN	
GmPP2A-3	(230)	EQFNHTNSLKLVARAHQLVMEGFNWAHEQK-VVTIFSAPN	
GmPP2A-2	(238)	AEFLTKHDLDLICRAHQVVEDGYEFFADRQ-LVTIFSAPN	
GmPP2A-1	(219)	SEFNHINNLDLVCRAHQVVEGLKYMFDKGLVTVWSAPN	
BnPP2A-3	(239)	AEFLIKNDMDLICRAHQVVEDGYEFFADRQ-LVTIFSAPN	
BnPP2A-2	(227)	AEFLQTHDLDLVCRAHQVVEDGYEFFAKRQ-LVTIFSAPN	
BnPP2A-1	(223)	TQFNHTNGLSLISRAHQLVMEGFNWCQEKV-VVTVFSAPN	
Consenso	(241)	AQFNHTN L LIARAHQLVMEGFNWC DRQ VVTIFSAPN	
		281	320
PpPP2A-1	(267)	YCYRCGNMAAILEVDDNMGHTFTQFEPAPRRGEPDVTRRT	
OsPP2A-5	(262)	YCYRCGNMAAILEIGENMDQNFLOFDPAPRQIEPDTTRKT	
OsPP2A-4	(263)	YCYRCGNMAAILEVDDCREHTFTQFEPAPRRGEPDVTRRT	
OsPP2A-3	(269)	YCGEFDNAGAMMSVDDTLMCSFQILKPAKMLGGSTNSKS	
OsPP2A-2	(263)	YCYRCGNMAAILEIGENMDQNFLOFDPAPRQIEPDTTRKT	
OsPP2A-1	(270)	YCYRCGNMAAILEVDDCRNHTFTQFEPAPRRGEPDVTRRT	
GmPP2A-5	(262)	YCYRCGNMAAILEIGENMDQNFLOFDPAPRQIEPDTTRKT	
GmPP2A-4	(268)	YCGEFDNAGAMMSVDETLMCSFQILKPADKAKLNFVSTT	
GmPP2A-3	(269)	YCYRCGNMAAILEVDDCKGHTFTQFEPAPRRGEPDVTRRT	
GmPP2A-2	(277)	YCGEFDNAGAMMSVDENLMCSFQILKPAEKKSDFVMSNKM	
GmPP2A-1	(259)	YCYRCGNMAAILEFNENMEREVKFFTEENNQMRGPRTG	
BnPP2A-3	(278)	YCGEFDNAGAMMSVDES-----	
BnPP2A-2	(266)	YCGEFDNAGAMMSVDDSLTCSFQILKSTEEKGRFGYNNNV	
BnPP2A-1	(262)	YCYRCGNMAAILEIDENMDQNFLOFDPAPRQVEPETTRKT	
Consenso	(281)	YCYRCGNMAAILEVDENM SFIQFDPAPRR EPD TRKT	
		321	337
PpPP2A-1	(307)	PDYFL-----	
OsPP2A-5	(302)	PDYFL-----	
OsPP2A-4	(303)	PDYFL-----	
OsPP2A-3	(309)	GFKSLRGW-----	
OsPP2A-2	(303)	PDYFL-----	
OsPP2A-1	(310)	PDYFL-----	
GmPP2A-5	(302)	PDYFL-----	
GmPP2A-4	(308)	TAKPGNSPAGVKVGRY-	
GmPP2A-3	(309)	PDYFL-----	
GmPP2A-2	(317)	-----	
GmPP2A-1	(299)	VPYFL-----	
BnPP2A-3	(295)	-----	
BnPP2A-2	(306)	HRPGTTPPHKGGKGG---	
BnPP2A-1	(302)	PDYFL-----	
Consenso	(321)	PDYFL	