



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 279 777**

51 Int. Cl.:

**A01H 5/00** (2006.01)

**A01H 5/10** (2006.01)

**C12N 15/29** (2006.01)

**C07K 14/415** (2006.01)

**C12N 5/10** (2006.01)

**C12N 15/82** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **00990297 .4**

86 Fecha de presentación : **22.12.2000**

87 Número de publicación de la solicitud: **1280397**

87 Fecha de publicación de la solicitud: **05.02.2003**

54

Título: **Proteínas relacionadas con el estrés por factores de transcripción y métodos de utilización en las plantas.**

30

Prioridad: **22.12.1999 US 171745 P**

45

Fecha de publicación de la mención BOPI:  
**01.09.2007**

45

Fecha de la publicación del folleto de la patente:  
**01.09.2007**

73

Titular/es: **BASF Plant Science GmbH  
67056 Ludwigshafen, DE**

72

Inventor/es: **Costa e Silva, Oswaldo da;  
Van Thielen, Nocha y  
Chen, Ruoying**

74

Agente: **Lehmann Novo, María Isabel**

ES 2 279 777 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

Proteínas relacionadas con el estrés por factores de transcripción y métodos de utilización en las plantas.

5 **Antecedentes de la invención****Campo de la invención**

Esta invención se refiere en general a secuencias de ácido nucleico que codifican proteínas que están asociadas con respuestas de estrés abiótico y tolerancia al estrés abiótico en las plantas. En particular, esta invención se refiere a secuencias de ácido nucleico que codifican proteínas que confieren a las plantas tolerancia a la sequía.

**Antecedentes de la técnica**

Los estados de estrés ambiental abiótico, tales como estrés por sequía, estrés por salinidad, estrés por calor, y estrés por frío, son factores limitantes importantes del crecimiento y la productividad de las plantas. Las pérdidas de cosecha y pérdidas de rendimiento de cosecha en el caso de cosechas importantes tales como arroz, maíz (grano) y trigo causados por estos estados de estrés representan un factor económico y político importante y contribuyen a la escasez de alimentos en muchos países subdesarrollados.

Las plantas se ven expuestas típicamente durante su ciclo vital a condiciones de contenido reducido de agua en el ambiente. La mayoría de las plantas han desarrollado estrategias para protegerse a sí mismas contra estas condiciones de desecación. No obstante, si la severidad y duración de las condiciones de sequía son demasiado grandes, los efectos sobre el desarrollo, el crecimiento y el rendimiento de la mayoría de las plantas de cosecha son profundos. Adicionalmente, la mayoría de las plantas de cosecha son muy sensibles a concentraciones elevadas de sal en el suelo. La exposición continua a sequía y contenido elevado de sal causa alteraciones importantes en el metabolismo de las plantas. Estos grandes cambios en el metabolismo conducen finalmente a la muerte celular y por consiguiente a pérdidas de rendimiento.

El desarrollo de plantas tolerantes al estrés es una estrategia que tiene potencial para resolver o mediar al menos alguno de estos problemas. Sin embargo, las estrategias tradicionales de reproducción de plantas para desarrollar nuevas líneas de plantas que exhiban resistencia (tolerancia) a estos tipos de estrés son relativamente lentas y requieren líneas específicas resistentes para cruzamiento con la línea deseada. Los recursos limitados de germoplasma para tolerancia al estrés y la incompatibilidad en los cruzamientos entre especies de plantas remotamente afines representan problemas importantes encontrados en la reproducción convencional. Adicionalmente, los procesos celulares que conducen a tolerancia a la sequía, el frío y la sal en plantas modelo, tolerantes a la sequía y/o la sal son de naturaleza compleja e implican mecanismos múltiples de adaptación celular y numerosos caminos metabólicos. Esta naturaleza multi-componente de la tolerancia al estrés no sólo ha hecho infructuosa en gran parte la reproducción para tolerancia, sino que ha limitado también la posibilidad de elaborar por ingeniería genética plantas tolerantes al estrés utilizando métodos biotecnológicos.

Así pues, es necesaria la identificación de los genes y proteínas implicados en estos procesos multi-componente que conducen a la tolerancia al estrés. La elucidación de la función de los genes expresados en las plantas tolerantes al estrés no sólo hará avanzar la comprensión actual de la adaptación y la tolerancia de las plantas a los estados de estrés ambiental, sino que puede proporcionar también información importante para el diseño de nuevas estrategias para mejora de las cosechas.

Una planta modelo utilizada en el estudio de la tolerancia al estrés es *Arabidopsis thaliana*. Existen al menos cuatro caminos diferentes de transducción de señales que conducen a la tolerancia al estrés en la planta modelo *Arabidopsis thaliana*. Estos caminos se encuentran bajo el control de factores de transcripción distintos (Shinozaki *et al.*, 2000 Curr. Op. Pl. Biol. 3: 217-23). Reguladores de genes, especialmente factores de transcripción, implicados en estos caminos de tolerancia son particularmente adecuados para modificar por ingeniería genética la tolerancia en las plantas dado que un solo gen puede activar una cascada completa de genes que conducen al fenotipo tolerante. Por consiguiente, los factores de transcripción son dianas importantes en la búsqueda con objeto de identificar los genes que confieren a las plantas tolerancia al estrés.

Un factor de transcripción que ha sido identificado en la técnica anterior es el factor de transcripción CBF de *Arabidopsis thaliana* (Jaglo-Ottosen *et al.*, 1998 Science 280: 104-6). La sobre-expresión de este gen en *Arabidopsis* confería a esta planta tolerancia a la sequía (Kasuga *et al.*, 1999 Nature Biotech. 17: 287-91). Sin embargo, CBF es el único ejemplo hasta la fecha de un factor de transcripción capaz de conferir a las plantas tolerancia a la sequía por sobre-expresión.

Aunque han sido caracterizados algunos genes que están implicados en respuestas al estrés en plantas, la caracterización y clonación de genes de plantas que confieren tolerancia al estrés sigue siendo en gran parte incompleta y fragmentada. Por ejemplo, ciertos estudios han indicado que el estrés por sequía y sal en algunas plantas puede ser debido a efectos aditivos de genes, en contraste con otras investigaciones que indican que genes específicos se activan por transcripción, lo que conduce a la acumulación de nuevas proteínas en el tejido vegetativo de las plantas en condiciones de estrés osmótico. Aunque está admitido generalmente que las proteínas inducidas por estrés tienen

cierto papel en la tolerancia, falta todavía la evidencia directa, y las funciones de muchos genes sensibles al estrés son desconocidas.

5 Por consiguiente, existe necesidad de identificar los genes expresados en las plantas tolerantes al estrés que tengan la capacidad de conferir resistencia al estrés a su planta hospedadora y a otras especies de plantas. Las plantas tolerantes al estrés de nueva generación tendrán muchas ventajas, tales como el aumento de la extensión en que pueden cultivarse las plantas de cosecha, por ejemplo por disminución de los requerimientos de agua de una especie vegetal.

## 10 Sumario de la invención

Esta invención satisface en parte la necesidad de identificar nuevos factores de transcripción singulares capaces de conferir a las plantas tolerancia al estrés por sobre-expresión. Concretamente, se describe en esta memoria el factor de transcripción: CRT/DRE Binding Factor (CBF-1) de *Physcomitrella patens*.

15 La presente invención proporciona una planta transgénica transformada por un ácido nucleico codificante de una proteína relacionada con el estrés por factores de transcripción (TFSRP), en donde la expresión de la secuencia de ácido nucleico en la planta da como resultado una tolerancia incrementada al estrés ambiental en comparación con una variedad de tipo salvaje de la planta y en donde la TFSRP es un Factor de Fijación-1 CRT/DRE (CBF-1) como se define en SEQ ID NO: 17 y homólogos del mismo que tienen al menos 50-60% de homología con la proteína como se define en SEQ ID NO: 17.

20 La invención estipula en algunas realizaciones que la TFSRP y el ácido nucleico codificante son los que se encuentran en miembros del género *Physcomitrella*. En otra realización preferida, el ácido nucleico y la proteína son de una planta *Physcomitrella patens*. La invención estipula que el estrés ambiental puede ser estrés por sequía.

30 La invención proporciona adicionalmente una semilla producida por una planta transgénica transformada por un ácido nucleico codificante de TFSRP, en donde la planta es progenie auténtica para tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta. La invención proporciona adicionalmente una semilla producida por una planta transgénica que expresa una TFSRP en donde la planta es progenie auténtica para tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta.

35 La invención proporciona adicionalmente un producto agrícola producido por cualquiera de las plantas transgénicas arriba descritas. La invención proporciona adicionalmente una TFSRP aislada, en donde la TFSRP es como se describe más adelante. La invención proporciona adicionalmente un ácido nucleico codificante de TFSRP, en donde el ácido nucleico codificante de TFSRP codifica una TFSRP como se describe más adelante.

40 La invención proporciona adicionalmente un vector de expresión recombinante aislado que comprende un ácido nucleico codificante de TFSRP como se describe más adelante, en donde la expresión del vector en una célula hospedadora da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la célula hospedadora. La invención proporciona adicionalmente una célula hospedadora que contiene el vector y una planta que contiene la célula hospedadora.

45 La invención proporciona adicionalmente un método de producción de una planta transgénica con un ácido nucleico codificante de TFSRP, en donde la expresión del ácido nucleico en la planta da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta que comprende: (a) transformar una célula de planta con un vector de expresión que comprende un ácido nucleico codificante de TFSRP, y (b) generar a partir de la célula de la planta una planta transgénica con una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta. En realizaciones preferidas, la TFSRP es como se describe más adelante. En realizaciones preferidas, el ácido nucleico codificante de TFSRP es como se describe más adelante.

55 La invención proporciona adicionalmente un método de aumentar la expresión de un gen de interés en una célula hospedadora en comparación con una variedad de tipo salvaje de la célula hospedadora, en donde el gen de interés se transcribe en respuesta a una TFSRP, que comprende: (a) transformar la célula hospedadora con un vector de expresión que comprende un ácido nucleico codificante de TFSRP, y (b) expresar la TFSRP en la célula hospedadora, aumentando con ello la expresión del gen transcrito en respuesta a la TFSRP en comparación con una variedad de tipo salvaje de la célula hospedadora. En realizaciones preferidas, la TFSRP es como se describe más adelante. En realizaciones preferidas, el ácido nucleico codificante de TFSRP es como se describe más adelante.

60 La presente invención proporciona adicionalmente un método de identificación de una nueva TFSRP, que comprende (a) generar una respuesta específica de anticuerpos para una TFSRP, o fragmento de la misma, como se ha descrito arriba; (b) escrutar material supuesto de TFSRP con el anticuerpo, en donde la fijación específica del anticuerpo al material indica la presencia de una TFSRP potencialmente nueva; y (c) analizar el material fijado en comparación con TFSRP conocidas para determinar su novedad.

**Breve descripción de los dibujos**

Las Figuras 1 (A-G) muestran las secuencias parciales de cDNA de CABF-1 (SEQ ID NO: 1) (Ejemplo Comparativo), DBF-1 (SEQ ID NO: 2) (Ejemplo Comparativo), CBF-1 (SEQ ID NO: 3) de acuerdo con la invención, HDZ-1 (SEQ ID NO: 4) (Ejemplo Comparativo), ZF-1 (SEQ ID NO: 5) (Ejemplo Comparativo), LZ-1 (SEQ ID NO: 6) (Ejemplo Comparativo) y CABF-2 (SEQ ID NO: 7) (Ejemplo Comparativo) de *Physcomitrella Patens*.

Las Figuras 2 (A-H) muestran las secuencias de cDNA de longitud total de CABF-1 (SEQ ID NO: 8) (Ejemplo Comparativo), DBF-1 (SEQ ID NO: 9) (Ejemplo Comparativo), DBF-1 variante (SEQ ID NO: 22) (Ejemplo Comparativo), CBF-1 (SEQ ID NO: 10) de acuerdo con la invención, HDZ-1 (SEQ ID NO: 11) (Ejemplo Comparativo), ZF-1 (SEQ ID NO: 12) (Ejemplo Comparativo), LZ-1 (SEQ ID NO: 13) (Ejemplo Comparativo) y CABF-2 (SEQ ID NO: 14) (Ejemplo Comparativo) de *Physcomitrella patens*.

Las Figuras 3 (A-H) muestran las secuencias de aminoácidos deducidas de CABF-1 (SEQ ID NO: 15) (Ejemplo Comparativo), DBF-1 (SEQ ID NO: 16) (Ejemplo Comparativo), DBF-1 variante (SEQ ID NO: 23) (Ejemplo Comparativo), CBF-1 (SEQ ID NO: 17) de acuerdo con la invención, HDZ-1 (SEQ ID NO: 18) (Ejemplo Comparativo), ZF-1 2 (SEQ ID NO: 19) (Ejemplo Comparativo), LZ-1 (SEQ ID NO: 20) (Ejemplo Comparativo) y CABF-2 (SEQ ID NO: 21) (Ejemplo Comparativo) de *Physcomitrella patens*.

La Figura 4 muestra un diagrama del vector de expresión de plantas pGMSG que contiene el super-promotor que dirige la expresión de SEQ ID NOs: 8, 9, 10, 11, 12, 13, y 14 (“gen deseado”). Los componentes son: gen de resistencia a la gentamicina aacCI (Hajdukiewicz *et al.*, 1994 Plant Molecular Biology 25: 989-94), promotor NOS (Becker *et al.*, 1992 Plant Molecular Biology 20: 1195-7), terminador g7T (Becker *et al.*, 1992), terminador NOSpA (Jefferson *et al.*, 1987 EMBO J. 6: 3901-7).

La Figura 5 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de HDZ-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 6 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de ZF-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 7 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de CABF-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 8 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de DBF-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas individuales transformantes.

La Figura 9 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de CABF-2 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 10 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de LZ-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 11 muestra los resultados de un ensayo de estrés por sequía con sobre-expresión de CBF-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (de acuerdo con la presente invención). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 12 muestra los resultados de un ensayo de estrés por sal con sobre-expresión de ZF-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 13 muestra los resultados de un ensayo de estrés por sal con sobre-expresión de CABF-2 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

La Figura 14 muestra los resultados de un ensayo de estrés por sal con sobre-expresión de LZ-1 de *Physcomitrella patens* en plantas transgénicas y líneas de tipo salvaje de *Arabidopsis* (Ejemplo Comparativo). Las líneas transgénicas exhiben un fenotipo tolerante. Se muestran líneas transformantes individuales.

### Descripción detallada de la invención

La presente invención puede comprenderse más fácilmente haciendo referencia a la descripción detallada que sigue de las realizaciones preferidas de la invención y los ejemplos incluidos en esta memoria. Sin embargo, antes que se expongan y describan los presentes compuestos, composiciones, y métodos, debe entenderse que esta invención no está limitada a ácidos nucleicos específicos, polipéptidos específicos, tipos de células específicos, células hospedadoras específicas, condiciones específicas, o métodos específicos, etc., dado que los mismos pueden, por supuesto, variar, y las numerosas modificaciones y variaciones en los mismos serán evidentes para los expertos en la técnica. Debe entenderse también que la terminología utilizada en esta memoria tiene únicamente por objeto describir realizaciones específicas y no debe entenderse como limitante. En particular, la designación de las secuencias de aminoácidos como "Proteínas relacionadas con el estrés por factores de transcripción" (TFSRPs), no limita en modo alguno la funcionalidad de tales secuencias.

La presente invención proporciona una planta transgénica transformada por un ácido nucleico codificante de TFSRP como se define en la reivindicación 1, en donde la expresión de la secuencia de ácido nucleico en la planta da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta. La invención proporciona adicionalmente una semilla producida por una planta transgénica transformada por un ácido nucleico codificante de TFSRP, en donde la semilla contiene el ácido nucleico codificante de TFSRP, y en donde la planta es progenie auténtica para tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta.

La invención proporciona adicionalmente una semilla producida por una planta transgénica que expresa una TFSRP, en donde la semilla contiene la TFSRP, y en donde la planta es progenie auténtica para tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta. La invención proporciona adicionalmente un producto agrícola producido por cualquiera de las plantas transgénicas descritas anteriormente o más adelante. Tal como se utiliza en esta memoria, el término "variedad" hace referencia a un grupo de plantas dentro de una especie que comparten caracteres constantes que los separan de la forma típica y de otras variedades posibles dentro de dicha especie. Si bien posee al menos un rasgo distintivo, una variedad se caracteriza también por cierta variación entre los individuos dentro de la variedad, basada fundamentalmente en la segregación mendeliana de rasgos entre la progenie de generaciones sucesivas. Una variedad se considera "progenie auténtica" para un rasgo particular si la misma es genéticamente homocigótica para dicho rasgo en tal grado que, cuando la variedad de progenie auténtica se auto-poliniza, no se observa un grado significativo de segregación independiente del rasgo entre la progenie. En la presente invención, el rasgo procede de la expresión transgénica de una secuencia simple de DNA introducida en una variedad de la planta.

La invención proporciona adicionalmente una TFSRP aislada. La invención estipula que la TFSRP es un Factor de Fijación CRT/DRE (CBF-1) como se define en SEQ ID NO: 17 y homólogos del mismo. Los homólogos de las secuencias de aminoácidos se describen más adelante. Es un nuevo descubrimiento de la presente invención que esta clase de factores de transcripción está involucrada en la tolerancia al estrés en las plantas y que la expresión de un miembro de esta clase de proteínas en una planta puede aumentar la tolerancia de dicha planta al estrés por sequía.

La invención proporciona adicionalmente un ácido nucleico aislado codificante de TFSRP. La presente invención incluye ácidos nucleicos codificantes de TFSRP que codifican TFSRPs como se describe en esta memoria. En realizaciones preferidas, el ácido nucleico codificante de TFSRP es un Factor de Fijación CRT/DRE (CBF-1) como se define en SEQ ID NO: 3 y homólogos del mismo. Homólogos de las secuencias de nucleótidos se definen más adelante. En una realización preferida, el ácido nucleico y la proteína se aíslan del género de plantas *Physcomitrella*. En otra realización preferida, el ácido nucleico y la proteína son de una planta *Physcomitrella patens* (*P. patens*).

Tal como se utiliza en esta memoria, el término "estrés ambiental" se refiere a cualquier condición de crecimiento sub-óptima e incluye, aunque sin carácter limitante, condiciones sub-óptimas asociadas con sequía. En realizaciones preferidas, el estrés ambiental puede ser sequía, y en particular, puede ser contenido de agua bajo. Debe entenderse también que, tal como se utiliza en la memoria descriptiva y en las reivindicaciones, "un" o "uno(a)" puede significar uno o más, dependiendo del contexto en el que se utilice. Así, por ejemplo, la referencia a "una célula" puede significar que puede utilizarse al menos una célula.

De acuerdo con los propósitos de esta invención, tal como se materializa y describe en términos generales en esta memoria, esta invención, en un aspecto, proporciona un ácido nucleico aislado de un musgo que codifica una Proteína Relacionada con el Estrés (SRP), o una porción de la misma. En particular, la presente invención proporciona ácidos nucleicos que codifican TFSRPs que incluyen la secuencia de ácido nucleico representada en SEQ ID NO: 10. La presente invención proporciona también secuencias de aminoácidos de TFSRPs que incluyen la secuencia de aminoácidos que se muestra en SEQ ID NO: 17.

La presente invención proporciona una nueva proteína predicha de *P. patens* CBF-1 (SEQ ID NO: 17), que es homóloga del factor de transcripción CBF-1 de *Arabidopsis thaliana*. (La homología con otras proteínas se muestra en la Tabla 8). Como se ha mencionado anteriormente, la expresión de CBF-1 conduce a plantas tolerantes al estrés. Dado que CBF-1 (SEQ ID NO: 17) procede de una planta tolerante al estrés, *Physcomitrella patens*, es concebible que este gen confiera un nivel más alto de tolerancia al estrés a otras plantas distintas del homólogo de *Arabidopsis*.

Un aspecto de la invención se refiere a moléculas de ácido nucleico aisladas que codifican polipéptidos de TFSRP o porciones biológicamente activas de los mismos, así como fragmentos de ácido nucleico suficientes para uso como sondas de hibridación o iniciadores para la identificación o amplificación de ácido nucleico codificante de TFSRP (v.g., DNA de TFSRP). Tal como se utiliza en esta memoria, debe entenderse que la expresión “molécula de ácido nucleico” incluye moléculas de DNA (v.g., cDNA o DNA genómico) y moléculas de RNA (v.g., mRNA) y análogos del DNA o RNA generados utilizando análogos de nucleótidos. Este término abarca también secuencias no traducidas localizadas en ambos extremos 3' y 5' de la región codificante del gen: al menos aproximadamente 1000 nucleótidos de secuencia aguas arriba del extremo 5' de la región codificante y al menos aproximadamente 200 nucleótidos de secuencia situados aguas abajo del extremo 3' de la región codificante del gen. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria, pero preferiblemente es DNA bicatenario. Una molécula de ácido nucleico “aislada” es una que está separada sustancialmente de otras moléculas de ácido nucleico que están presentes en la fuente natural del ácido nucleico. Preferiblemente, un ácido nucleico “aislado” está exento de algunas de las secuencias que flanquean naturalmente el ácido nucleico (es decir, secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en el DNA genómico del organismo del que se deriva el ácido nucleico. Por ejemplo, en diversas realizaciones, la molécula de ácido nucleico TFSRP aislada puede contener menos de aproximadamente 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 el DNA genómico de la célula de la que deriva el ácido nucleico (v.g., una célula de *Physcomitrella patens*). Además, una molécula de ácido nucleico “aislada” tal como una molécula de cDNA, puede estar exenta de algo del otro material celular con el cual está asociada aquella naturalmente, o medio de cultivo cuando se produce por técnicas recombinantes, o precursores químicos u otros productos químicos cuando se sintetiza químicamente.

Una molécula de ácido nucleico de la presente invención, v.g., una molécula de ácido nucleico que tiene una secuencia de nucleótidos de SEQ ID NO: 10, o una porción de la misma, puede aislarse utilizando técnicas estándar de biología molecular y la información de secuencias proporcionada en esta memoria. Por ejemplo, un cDNA de TFSRP de *P. patens* puede aislarse a partir de una genoteca de *P. patens* utilizando la totalidad o una parte de una de las secuencias SEQ ID NO: 3, como una sonda de hibridación y técnicas de hibridación estándar (v.g., como se describe en Sambrook *et al.*, 1989 *Molecular Cloning: A Laboratory Manual*, 2ª edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). Además, una molécula de ácido nucleico que abarca la totalidad o una parte de la secuencia de SEQ ID NO: 3, puede aislarse por la reacción en cadena de la polimerasa utilizando iniciadores oligonucleotídicos diseñados sobre la base de esta secuencia (v.g., una molécula de ácido nucleico que abarca la totalidad o una parte de la secuencia de SEQ ID NO: 3 puede aislarse por la reacción en cadena de la polimerasa utilizando iniciadores oligonucleotídicos diseñados basándose en esta misma secuencia de SEQ ID NO: 3.

Por ejemplo, es posible aislar mRNA a partir de células de plantas (v.g., por el procedimiento de extracción con tiocianato de guanidinio de Chirgwin *et al.*, 1979 *Biochemistry* 18: 5294-5299) y puede prepararse cDNA utilizando transcriptasa inversa (v.g., transcriptasa inversa de MLV de Moloney, disponible de Gibco/BRL, Bethesda, MD; o transcriptasa inversa de AMV, disponible de Seikagaku America Inc., St. Petersburg, FL). Iniciadores oligonucleotídicos sintéticos para amplificación de la reacción en cadena de la polimerasa pueden diseñarse basándose en las secuencias de nucleótidos que se muestran en SEQ ID NO: 3.

Una molécula de ácido nucleico de la invención puede amplificarse utilizando cDNA o, alternativamente, DNA genómico, como molde e iniciadores oligonucleotídicos apropiados de acuerdo con las técnicas estándar de amplificación PCR. La molécula de ácido nucleico así amplificada puede clonarse en un vector apropiado y caracterizarse por análisis de la secuencia de DNA. Adicionalmente, pueden prepararse oligonucleótidos correspondientes a una secuencia de nucleótidos de TFSRP por técnicas de síntesis estándar, v.g., utilizando un sintetizador automático de DNA.

En una realización preferida, una molécula de ácido nucleico aislada de la invención comprende la secuencia de nucleótidos que se muestra en SEQ ID NO: 10.

La secuencia de SEQ ID NO: 10 corresponde a los cDNAs de TFSRP de *Physcomitrella patens* de la invención. Este cDNA comprende secuencias que codifican TFSRPs (es decir, la “región codificante”, indicada en la Tabla 1), así como secuencias no traducidas 5' y secuencias no traducidas 3'. Debe entenderse que SEQ ID NO: 10 comprende a la vez regiones codificantes y regiones 5' y 3' no traducidas. Alternativamente, la molécula de ácido nucleico puede comprender únicamente la región codificante de la secuencia en SEQ ID NO: 10, o puede contener fragmentos genómicos enteros aislados de DNA genómico. Una región codificante de esta secuencia se indica como “posición ORF”.

En otra realización preferida, una molécula de ácido nucleico aislada de la invención comprende una molécula de ácido nucleico que es un complemento de la secuencia de nucleótidos que se muestra en SEQ ID NO: 10, o una porción de la misma. Una molécula de ácido nucleico que es complementaria a la secuencia de nucleótidos que se muestra en SEQ ID NO: 10, es una que es suficientemente complementaria a la secuencia de nucleótidos que se muestra en SEQ ID NO: 10, tal que la misma puede hibridarse a la secuencia de nucleótidos que se muestra en SEQ ID NO: 10, formando de este modo un dúplex estable.

En otra realización preferida adicional, una molécula de ácido nucleico aislada de la invención comprende una secuencia de nucleótidos que es al menos aproximadamente 50-60%, con preferencia al menos aproximadamente 60-70%, de modo más preferible al menos aproximadamente 70-80%, 80-90%, o 90-95%, y de modo aún más preferible al

menos aproximadamente 95%, 96%, 97%, 98%, 99% o más homóloga a una secuencia de nucleótidos que se muestra en SEQ ID NO: 10, o una porción de la misma. En una realización preferida adicional, una molécula de ácido nucleico aislada de la invención comprende una secuencia de nucleótidos que se hibrida, v.g., se hibrida en condiciones severas, a la secuencia de nucleótidos que se muestra en SEQ ID NO: 10, o una porción de la misma. Estas condiciones de hibridación incluyen lavado con una solución que tiene una concentración de sal aproximadamente 0,02 molar a pH 7 a aproximadamente 60°C.

Además, la molécula de ácido nucleico de la invención puede comprender sólo una porción de la región codificante de la secuencia en SEQ ID NO: 10, por ejemplo un fragmento que puede utilizarse como sonda o iniciador o un fragmento que codifica una porción biológicamente activa de una TFSRP. Las secuencias de nucleótidos determinadas por la clonación de los genes TFSRP de *P. patens* permiten la generación de sondas e iniciadores diseñados para uso en la identificación y/o clonación de homólogos de TFSRP en otros tipos de células y organismos, así como homólogos de TFSRP de otros musgos o especies afines. Por consiguiente, esta invención proporciona también compuestos que comprenden las moléculas de ácido nucleico descritas en esta memoria, o fragmentos de las mismas. Estos compuestos incluyen las moléculas de ácido nucleico fijadas a un resto. Estos restos incluyen, pero sin carácter limitante, restos de detección, restos de hibridación, restos de purificación, restos de suministro, restos de reacción, restos de fijación, y análogos. La sonda/iniciador comprende por regla general oligonucleótidos sustancialmente aislados. El oligonucleótido comprende típicamente una región de secuencia de nucleótidos que se hibrida en condiciones severas a al menos aproximadamente 12, con preferencia aproximadamente 25, de modo más preferible aproximadamente 40, 50 ó 75 nucleótidos consecutivos de una cadena de sentido de las secuencias expuestas en SEQ ID NO: 10, o una secuencia antisentido de la secuencia indicada en SEQ ID NO: 10, o mutantes de la misma existentes naturalmente. Los iniciadores basados en una secuencia de nucleótidos de SEQ ID NO: 10 pueden utilizarse en reacciones PCR para clonar homólogos de TFSRP. Pueden utilizarse sondas basadas en las secuencias de nucleótidos TFSRP para detectar transcritos o secuencias genómicas que codifican la misma proteína o proteínas homólogas. En realizaciones preferidas, la sonda comprende adicionalmente un grupo marcador fijado a ella, v.g. el grupo marcador puede ser un radioisótopo, un compuesto fluorescente, una enzima, o un co-factor enzimático. Tales sondas pueden utilizarse como parte de un kit de ensayo de marcadores genómicos para identificar células que expresan una TFSRP, por ejemplo por medida de un nivel de un ácido nucleico codificante de TFSRP en una muestra de células, v.g., detección de niveles de mRNA de TFSRP o determinación de si un gen de TFSRP genómico ha sido mutado o deletado.

En particular, un método útil para determinar el nivel de transcripción del gen (un indicador de la cantidad de mRNA disponible para traducción al producto génico) consiste en realizar una transferencia Northern (para referencia véase, por ejemplo, Ausubel *et al.*, 1988 Current Protocols in Molecular Biology, Wiley; Nueva York), en la cual un iniciador diseñado para fijarse al gen de interés está marcado con un marcador detectable (usualmente radioactivo o quimioluminiscente), tal que cuando se extrae el RNA total de un cultivo del organismo, se pasa sobre un gel, se transfiere a una matriz estable y se incuba con esta sonda, la fijación y cantidad de fijación de la sonda indica la presencia y también la cantidad de mRNA para este gen. Esta información demuestra al menos parcialmente el grado de transcripción del gen transformado. Puede prepararse RNA celular total a partir de células, tejidos u órganos por varios métodos, todos ellos bien conocidos en la técnica, tales como el descrito en Bormann, E.R. *et al.*, 1992 Mol. Microbiol. 6: 317-326).

Para evaluar la presencia o la cantidad relativa de proteína traducida a partir de este mRNA, pueden emplearse técnicas estándar (tales como una transferencia Western (véase, por ejemplo, Ausubel *et al.*, 1988 Current Protocols in Molecular Biology, Wiley, Nueva York). En este proceso, se extraen las proteínas celulares totales, se separan por electroforesis en gel, se transfieren a una matriz tal como nitrocelulosa, y se incuban con una sonda, tal como un anticuerpo, que se fija específicamente a la proteína deseada. Esta sonda está marcada generalmente con un marcador quimioluminiscente o colorimétrico que puede detectarse fácilmente. La presencia y cantidad de marcador observado indica la presencia y cantidad de la proteína mutante deseada presente en la célula.

En una realización, la molécula de ácido nucleico de la invención codifica una proteína o porción de la misma que incluye una secuencia de aminoácidos que es suficientemente homóloga a una secuencia de aminoácidos de SEQ ID NO: 17, de tal modo que la proteína o porción de la misma mantiene la misma función o una función similar que la secuencia de aminoácidos con la que se compara. Tal como se utiliza en esta memoria, la expresión lingüística “suficientemente homólogo” hace referencia a proteínas o porciones de las mismas que tienen secuencias de aminoácidos que incluyen un número mínimo de residuos de aminoácidos idénticos o equivalentes (v.g., un residuo de aminoácido que tiene una cadena lateral similar que un residuo de aminoácido en los ORFs de la secuencia de SEQ ID NO: 17), a una secuencia de aminoácidos de TFSRP tal que la proteína o porción de la misma es capaz de participar en una respuesta de tolerancia al estrés en una planta, o más particularmente puede participar en la transcripción de una proteína implicada en una respuesta de tolerancia a la sequía en una planta de *Physcomitrella patens*. Ejemplos de tales actividades se describen también en esta memoria. Ejemplos de actividades de TFSRP se indican en la Tabla 1.

En otra realización, la proteína es al menos aproximadamente 50-60%, con preferencia al menos aproximadamente 60-70%, y de modo más preferible al menos aproximadamente 70-80%, 80-90%, 90-95%, y de modo muy preferible al menos aproximadamente 96%, 97%, 98%, 99%, o más homóloga a una secuencia entera de aminoácidos que se muestra en SEQ ID NO: 17. En otra realización adicional, al menos aproximadamente 50-60%, con preferencia al menos aproximadamente 60-70%, de modo más preferible al menos aproximadamente 70-80%, 80-90%, 90-95%, y de modo muy preferible al menos aproximadamente 96%, 97%, 98%, 99% o más homóloga de una secuencia entera de aminoácidos codificada por una secuencia de ácido nucleico representada en SEQ ID NO: 10.

Porciones de proteínas codificadas por las moléculas de ácido nucleico TFSRP de la invención son con preferencia porciones biológicamente activas de una de las TFSRPs. Tal como se utiliza en esta memoria, la expresión “porción biológicamente activa de una TFSRP” tiene por objeto incluir una porción, v.g., un dominio/motivo, de una TFSRP que participa en una respuesta de tolerancia al estrés en una planta, o participa más particularmente en la transcripción de una proteína implicada en una respuesta de tolerancia al estrés en una planta, o tiene una actividad como se indica en la Tabla 1. Para determinar si una TFSRP o una porción biológicamente activa de la misma puede participar en la transcripción de una proteína implicada en una respuesta de tolerancia al estrés en una planta, puede realizarse un análisis de estrés de una planta que expresa la TFSRP. Tales métodos de análisis son bien conocidos por los expertos en la técnica, como se detalla en el Ejemplo 7.

Fragmentos adicionales de ácido nucleico que codifican porciones biológicamente activas de una TFSRP pueden prepararse por aislamiento de una porción de las secuencias en SEQ ID NO: 17, que expresan la porción codificada de la TFSRP o el péptido (v.g., por expresión recombinante *in vitro*) y evaluación de la actividad de la porción codificada de la TFSRP o péptido.

La invención abarca adicionalmente moléculas de ácido nucleico que difieren de la secuencia de nucleótidos que se muestra en SEQ ID NO: 10 (y porciones de la misma) debido a la degeneración del código genético y codifican por tanto la misma TFSRP que la codificada por la secuencia de nucleótidos que se muestra en SEQ ID NO: 10.

En una realización adicional, la molécula de ácido nucleico de la invención codifica una proteína de *Physcomitrella patens* de longitud total que es sustancialmente homóloga a una secuencia de aminoácidos de un polipéptido representado en SEQ ID NO: 17.

Además de las secuencias de nucleótidos de TFSRP de *Physcomitrella patens* representadas en SEQ ID NO: 10, será apreciado por los expertos en la técnica que pueden existir polimorfismos de secuencias de DNA que conducen a cambios en las secuencias de aminoácidos de TFSRPs en una población (v.g., la población de *Physcomitrella patens*). Dicho polimorfismo genético en el gen de TFSRP puede existir entre individuos de una misma población debido a la variación natural. Tal como se utiliza en esta memoria, los términos o expresiones “gen” y “gen recombinante” hacen referencia a moléculas de ácido nucleico que comprenden un marco de lectura abierto que codifica una TFSRP, preferiblemente una TFSRP de *Physcomitrella patens*. Tales variaciones naturales pueden dar como resultado típicamente 1-5% de variación en la secuencia de nucleótidos del gen de TFSRP. Debe entenderse que cualquiera y la totalidad de dichas variaciones de nucleótidos y polimorfismos de aminoácidos resultantes en una TFSRP que son resultado de la variación natural y que no alteran la actividad funcional de las TFSRPs están comprendidas dentro del alcance de la invención.

Moléculas de ácido nucleico correspondientes a variantes naturales y homólogos no pertenecientes a la especie *Physcomitrella patens* del cDNA de TFSRP de *Physcomitrella patens* de la invención pueden aislarse sobre la base de su homología con el ácido nucleico de TFSRP de *Physcomitrella patens* descrito en esta memoria utilizando el cDNA de *Physcomitrella patens*, o una porción del mismo, como sonda de hibridación de acuerdo con técnicas de hibridación estándar en condiciones de hibridación severas. De acuerdo con ello, en otra realización, una molécula de ácido nucleico aislada de la invención tiene una longitud de al menos 15 nucleótidos y se hibrida en condiciones severas a la molécula de ácido nucleico que comprende una secuencia de nucleótidos de: SEQ ID NO: 10. En otras realizaciones, el ácido nucleico tiene al menos una longitud de 30, 50, 100, 250 o más nucleótidos. Tal como se utiliza en esta memoria, la expresión “se hibrida en condiciones severas” tiene por objeto describir condiciones para hibridación y lavado en las cuales las secuencias de nucleótidos que tienen una homología de al menos 60% unas con otras se mantienen típicamente hibridadas unas a otras. Preferiblemente, las condiciones son tales que las secuencias que son al menos aproximadamente 65%, de modo más preferible al menos aproximadamente 70%, y de modo todavía más preferible al menos aproximadamente 75% o más homólogas una a otras se mantienen típicamente hibridadas unas a otras. Tales condiciones severas son conocidas por los expertos en la técnica y pueden encontrarse en *Current Protocols in Molecular Biology*, 6.3.1-6.3.6, John Wiley & Sons, N.Y. (1989). Un ejemplo preferido y no limitante de condiciones de hibridación severas son la hibridación en cloruro de sodio/citrato de sodio (SSC) 6X a aproximadamente 45°C, seguida por uno o más lavados en 0,2 X SSC, 0,1% de SDS a 50-65°C. Preferiblemente, una molécula de ácido nucleico aislada de la invención que se hibrida en condiciones severas a una secuencia de SEQ ID NO: 10, corresponde a una molécula de ácido nucleico existente naturalmente. Tal como se utiliza en esta memoria, una molécula de ácido nucleico “existente naturalmente” hace referencia a una molécula de RNA o DNA que tiene una secuencia de nucleótidos que existe en la naturaleza (v.g., codifica una proteína natural). En una realización, el ácido nucleico codifica una TFSRP natural de *Physcomitrella patens*.

Además de variantes existentes naturalmente de la secuencia de TFSRP que puede existir en la población, el experto en la técnica apreciará adicionalmente que pueden introducirse cambios por mutación en una secuencia de nucleótidos de SEQ ID NO: 10, conduciendo con ello a cambios en la secuencia de aminoácidos de la TFSRP codificada, sin alterar la capacidad funcional de la TFSRP. Por ejemplo, pueden hacerse sustituciones de nucleótidos que conducen a sustituciones de aminoácidos en residuos de aminoácidos “no esenciales” en una secuencia de SEQ ID NO: 10. Un residuo de aminoácido “no esencial” es un residuo que puede alterarse con respecto a la secuencia de tipo salvaje de una de las TFSRPs (SEQ ID NO: 17) sin alterar la actividad de dicha TFSRP, mientras que un residuo de aminoácido “esencial” es necesario para la actividad de TFSRP. Otros residuos de aminoácidos, sin embargo, (v.g., aquéllos que no están conservados o están solo semi-conservados en el dominio que tiene actividad de TFSRP) pueden no ser esenciales



para la actividad y por consiguiente son probablemente susceptibles de alteración sin que se altere la actividad de TFSRP.

De acuerdo con ello, otro aspecto de la invención se refiere a moléculas de ácido nucleico que codifican TFSRPs que contienen cambios en residuos de aminoácidos que no son esenciales para la actividad de TFSRP. Tales TFSRPs difieren en secuencia de aminoácidos de una secuencia contenida en SEQ ID NO: 17, pero retienen al menos una de las actividades de TFSRP descritas en esta memoria. En una realización, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos que codifica una proteína, en donde la proteína comprende una secuencia de aminoácidos que es al menos 50% homóloga a una secuencia de aminoácidos de SEQ ID NO: 17, y es capaz de participar en la respuesta de tolerancia al estrés en una planta, o más particularmente participa en la transcripción de una proteína implicada en una respuesta de tolerancia al estrés en una planta *Physcomitrella patens*, o tiene una o más de las actividades indicadas en la Tabla 1. Preferiblemente, la proteína codificada por la molécula de ácido nucleico es al menos aproximadamente 50-60% homóloga a la secuencia de SEQ ID NO: 17, de modo más preferible al menos aproximadamente 60-70% homóloga a una de las secuencias de SEQ ID NO: 17, de modo todavía más preferible al menos aproximadamente 70-80, 80-90, 90-95% homóloga a la secuencia de SEQ ID NO: 17, y de modo muy preferible al menos aproximadamente 96%, 97%, 98%, o 99% homóloga a la secuencia de SEQ ID NO: 17.

Para determinar el porcentaje de homología de dos secuencias de aminoácidos (v.g., la secuencia de SEQ ID NO: 17, y una forma mutante de la misma) o de dos ácidos nucleicos, las secuencias se alinean para propósitos de comparación óptima (v.g., pueden introducirse lagunas en las secuencias de una proteína o de un ácido nucleico para alineación óptima con la otra proteína o el otro ácido nucleico). Se comparan luego los residuos de aminoácidos o nucleótidos para posiciones de aminoácidos o posiciones de nucleótidos correspondientes. Cuando una posición en una secuencia (v.g., la secuencia de SEQ ID NO: 17) está ocupada por el mismo residuo de aminoácido o nucleótido que la posición correspondiente en la otra secuencia (v.g., una forma mutante de la secuencia del polipéptido de SEQ ID NO: 17), entonces las moléculas son homólogas en dicha posición (es decir, tal como se utiliza en esta memoria, "homología" de aminoácido o ácido nucleico es equivalente a "identidad" de aminoácido o ácido nucleico). El porcentaje de homología entre las dos secuencias es función del número de posiciones idénticas compartidas por las secuencias (es decir, % de homología = número de posiciones idénticas/número total de posiciones x 100). Preferiblemente, la longitud de comparación de la secuencia es al menos 15 residuos de aminoácidos, más preferiblemente al menos 25 residuos de aminoácidos, y muy preferiblemente al menos 35 residuos de aminoácidos.

Alternativamente, una determinación de la homología porcentual entre dos secuencias puede realizarse utilizando un algoritmo matemático. Un ejemplo preferido y no limitante de un algoritmo matemático utilizado para la comparación de dos secuencias es el algoritmo de Karlin y Altschul (1990 Proc. Natl. Acad. Sci. USA 90: 5873-5877). Un algoritmo de este tipo está incorporado en los programas NBLAST y XBLAST de Altschul, *et al.* (1990, J. Mol. Biol. 215: 403-410). Las búsquedas de ácidos nucleicos BLAST pueden realizarse con el programa NBLAST, registro = 100, longitud de palabra = 12 para obtener secuencias de ácido nucleico homólogas a las moléculas de ácido nucleico TFSRP de la invención. Pueden realizarse búsquedas de proteínas BLAST con el programa XBLAST, registro = 50, longitud de palabra = 3 para obtener secuencias de aminoácidos homólogas a las TFSRPs de la presente invención. Para obtener alineaciones con lagunas para propósitos de comparación, puede utilizarse Gapped BLAST como se describe en Altschul *et al.* (1997) Nucleic Acids Res. 25: 3389-3402). Cuando se utilizan los programas BLAST y Gapped BLAST, pueden utilizarse los parámetros por defecto de los programas respectivos (v.g., XBLAST y NBLAST). Otro ejemplo no limitante preferido de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo de Myers y Miller (CABIOS 1989). Dicho algoritmo está incorporado en el programa ALIGN (versión 2.0) que forma parte del paquete de soporte lógico de alineación de secuencias GCG. Cuando se utiliza el programa ALIGN para comparación de secuencias de aminoácidos, puede utilizarse una tabla de residuos de peso PAM120, una penalidad por longitud de laguna de 12 y una penalidad por laguna de 4 a fin de obtener secuencias de aminoácidos homólogas a las TFSRPs de la presente invención. Para obtener alineaciones con lagunas para propósitos de comparación, puede utilizarse Gapped BLAST como se describe en Altschul *et al.* (1997 Nucleic Acids Res. 25: 3389-3402). Cuando se utilizan los programas BLAST y Gapped BLAST, pueden utilizarse los parámetros por defecto de los programas respectivos (v.g., XBLAST y NBLAST). Otro ejemplo no limitante preferido de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo de Myers y Miller (CABIOS 1989). Dicho algoritmo está incorporado en el programa ALIGN (versión 2.0) que forma parte del paquete de soporte lógico de alineación de secuencias GCG. Cuando se utiliza el programa ALIGN para comparación de secuencias de aminoácidos, puede utilizarse una tabla de residuos de peso PAM120, una penalidad por longitud de laguna de 12 y una penalidad por laguna de 4.

Una molécula de ácido nucleico aislada que codifica una TFSRP homóloga a una secuencia de proteínas de SEQ ID NO: 17 puede crearse por introducción de una más sustituciones, adiciones o deleciones de nucleótidos en una secuencia de nucleótidos de SEQ ID NO: 10, de tal modo que se introducen en la proteína codificada una o más sustituciones, adiciones o deleciones de aminoácidos. Pueden introducirse mutaciones en la secuencia de SEQ ID NO: 10 por técnicas estándar, tales como mutagénesis orientada y mutagénesis mediada por PCR. Preferiblemente, se hacen sustituciones conservadoras de aminoácidos en uno o más residuos de aminoácidos no esenciales predichos. Una "sustitución conservadora de aminoácidos" es una en la cual el residuo de aminoácido se reemplaza con un residuo de aminoácido que tiene una cadena lateral similar. Familias de residuos de aminoácidos que tienen cadenas laterales similares han sido definidas en la técnica. Estas familias incluyen aminoácidos con cadenas laterales básicas (v.g., lisina, arginina, histidina), cadenas laterales ácidas (v.g., ácido aspártico, ácido glutámico), cadenas laterales polares desprovistas de carga (v.g., glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (v.g., alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas en

posición beta (v.g., treonina, valina, isoleucina) y cadenas laterales aromáticas (v.g., tirosina, fenilalanina, triptófano, histidina). Así, un residuo de aminoácido no esencial predicho en una TFSRP se reemplaza preferiblemente con otro residuo de aminoácido de la misma familia de cadenas laterales. Alternativamente, en otra realización, pueden introducirse mutaciones aleatoriamente a lo largo de toda o parte de una secuencia codificante de TFSRP, por ejemplo por mutagénesis de saturación, y los mutantes resultantes pueden escrutarse respecto a una actividad de TFSRP descrita en esta memoria a fin de identificar mutantes que retienen la actividad de TFSRP. Después de la mutagénesis de la secuencia de SEQ ID NO: 6, la proteína codificada puede expresarse recombinantemente y la actividad de la proteína puede determinarse por análisis de la tolerancia al estrés de una planta que expresa la proteína que se describe en el Ejemplo 7.

Además de las moléculas de ácido nucleico que codifican las TFSRPs descritas anteriormente, otro aspecto de la misma se refiere a moléculas de ácido nucleico aisladas que son antisentido respecto a aquéllas. Un ácido nucleico “antisentido” comprende una secuencia de nucleótidos que es complementaria a un ácido nucleico “sentido” que codifica una proteína, v.g., complementaria a la cadena codificante de una molécula de cDNA bicatenaria o complementaria a una secuencia de mRNA. De acuerdo con ello, un ácido nucleico antisentido puede estar unido por hidrógeno a un ácido nucleico sentido. El ácido nucleico antisentido puede ser complementario a una cadena codificante entera de TFSRP; o sólo a una porción de la misma. En una realización, una molécula de ácido nucleico antisentido es antisentido respecto a una “región codificante” de la cadena codificante de la cadena codificante de una secuencia de nucleótidos que codifica una TFSRP. La expresión “región codificante” hace referencia a la región de la secuencia de nucleótidos que comprende codones que se traducen en residuos de aminoácidos (v.g., la región codificante entera de ,,,, , comprende los nucleótidos 1 a ....). En otra realización, la molécula de ácido nucleico antisentido es antisentido respecto a una “región no codificante” de la cadena codificante de una secuencia de nucleótidos que codifica una TFSRP. La expresión “región no codificante” se refiere a secuencias 5' y 3' que flanquean la región codificante que no se traducen en aminoácidos (es decir, a las que se hace referencia también como regiones 5' y 3' no traducidas).

Dadas las secuencias de la cadena codificante que codifican TFSRP descritas en esta memoria, (v.g., la secuencia indicada en SEQ ID NO: 10), pueden diseñarse ácidos nucleicos antisentido de la invención de acuerdo con las reglas de apareamiento de bases de Watson y Crick. La molécula de ácido nucleico antisentido puede ser complementaria a la región codificante entera de mRNA de TFSRP, pero más preferiblemente es un oligonucleótido que es antisentido respecto a solamente una porción de la región codificante o no codificante de mRNA de TFSRP. Por ejemplo, el oligonucleótido antisentido puede ser complementario a la región que circunda el sitio de comienzo de la traducción de mRNA de TFSRP. Un oligonucleótido antisentido puede tener, por ejemplo, una longitud de aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45 ó 50 nucleótidos. Un ácido nucleico antisentido de la invención puede construirse utilizando reacciones de síntesis química y de ligación enzimática que emplean procedimientos conocidos en la técnica. Por ejemplo, un ácido nucleico antisentido (v.g. un oligonucleótido antisentido) puede sintetizarse químicamente utilizando nucleótidos existentes naturalmente o nucleótidos modificados de diversas maneras, diseñados para aumentar la estabilidad biológica de las moléculas o para aumentar la estabilidad física del dúplex formado entre los ácidos nucleicos antisentido y sentido, v.g., pueden utilizarse derivados fosforotioato y nucleótidos sustituidos con acridina. Ejemplos de nucleótidos modificados que pueden utilizarse para generar el ácido nucleico antisentido incluyen 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroximetil)-uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metil-N6-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, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxi-propil)-uracilo, (acp3)w, y 2,6-diaminopurina. Alternativamente, el ácido nucleico antisentido puede producirse biológicamente utilizando un vector de expresión en el cual un ácido nucleico ha sido subclonado en una orientación antisentido (es decir, el RNA transcrito a partir del ácido nucleico insertado tendrá una orientación antisentido respecto a un ácido nucleico diana de interés, descrito con mayor detalle en la subsección que sigue).

Las moléculas de ácido nucleico antisentido de la invención se administran típicamente a una célula o se generan *in situ* de tal manera que las mismas se hibridan con o se fijan a mRNA celular y/o DNA genómico que codifica una TFSRP para inhibir con ello la expresión de la proteína, v.g., por inhibición de la transcripción y/o la traducción. La hibridación puede realizarse por complementariedad convencional de nucleótidos para formar un dúplex estable, o, por ejemplo, en el caso de una molécula de ácido nucleico antisentido que se fija a dúplex de DNA, por interacciones específicas en la hendidura mayor de la doble hélice. La molécula antisentido puede modificarse de tal manera que la misma se fije específicamente a un receptor o un antígeno expresado en una superficie de la célula seleccionada, v.g., por enlace de la molécula de ácido nucleico antisentido a un péptido o un anticuerpo que se fija a un receptor o antígeno de la superficie celular. La molécula de ácido nucleico antisentido puede suministrarse también a las células utilizando los vectores descritos en esta memoria. Para alcanzar concentraciones intracelulares suficientes de las moléculas antisentido, se prefieren constructos de vectores en los cuales la molécula de ácido nucleico antisentido se halla bajo el control de un promotor procarionta, viral, o eucariota (con inclusión de plantas) fuerte.

En otra realización adicional, la molécula de ácido nucleico antisentido es una molécula de ácido nucleico  $\alpha$ -anomérica. Una molécula de ácido nucleico  $\alpha$ -anomérica forma híbridos bicatenarios específicos con RNA complementario en los cuales, contrariamente a las unidades  $\beta$  usuales, las cadenas avanzan paralelas una a otra (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 quimérico RNA-DNA (Inoue *et al.* 1987, FEBS Lett. 215: 327-330).

5 En otra realización adicional, un ácido nucleico antisentido es una ribozima. Las ribozimas son moléculas catalíticas de RNA con actividad de ribonucleasa que son capaces de escindir un ácido nucleico monocatenario, tal como un mRNA, para el cual tienen las mismas una región complementaria. Así, pueden utilizarse ribozimas (v.g., las ribozimas de cabeza de martillo descritas en Haselhoff y Gerlach, 1988 Nature 334: 585-591)) para escindir catalíticamente transcritos de mRNA de TFSRP a fin de inhibir con ello la traducción de mRNA de TFSRP. Una ribozima que tiene  
10 especificidad para un ácido nucleico codificante de TFSRP puede diseñarse basándose en la secuencia de nucleótidos de un cDNA de TFSRP descrito en esta memoria (a saber, SEQ ID NO: 10) o sobre la base de una secuencia heteróloga a aislar de acuerdo con los métodos propuestos en esta invención. Por ejemplo, puede construirse un derivado de un RNA de IVS de *Tetrahymena* L-19 en el cual la secuencia de nucleótidos del sitio activo es complementaria a la secuencia de nucleótidos a escindir en un mRNA codificante de TFSRP. Véase, v.g., Cech *et al.* Patente U.S. No. 4.987.071 y Cech *et al.* Patente U.S. No. 5.116.742. Alternativamente, puede utilizarse mRNA de TFSRP para seleccionar un RNA catalítico que tenga una actividad específica de ribonucleasa a partir de una agrupación de moléculas de RNA. Véase, v.g., Bartel, D. y Szostak, J.W., 1993 Science 261: 1411-1418.

Alternativamente, la expresión de los genes de TFSRP puede inhibirse por direccionamiento de secuencias de nucleótidos complementarias a la región reguladora de una secuencia nucleotídica de TFSRP (v.g., un promotor y/o intensificador de TFSRP) para formar estructuras de triple hélice que impiden la transcripción de un gen de TFSRP en las células diana. Véase en líneas generales, 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.

25 La invención proporciona adicionalmente un vector de expresión recombinante aislado que comprende un ácido nucleico como se ha descrito arriba, en donde la expresión del vector en una célula hospedadora da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la célula hospedadora. Como se utiliza en esta memoria, el término "vector" hace referencia a una molécula de ácido nucleico capaz de transportar otro ácido nucleico con el que se ha enlazado. Otro tipo de vector es un "plásmido", término que hace  
30 referencia a un bucle circular de DNA bicatenario al cual pueden estar ligados segmentos de DNA adicionales. Otro tipo de vector es un vector viral, en el cual segmentos adicionales de DNA pueden ligarse al genoma viral. Ciertos vectores son capaces de replicación autónoma en una célula hospedadora en la cual se han introducido (v.g., vectores bacterianos que tienen un origen de replicación bacteriano y vectores episómicos de mamífero). Otros vectores (v.g., vectores no episómicos de mamífero) se integran en el genoma de una célula hospedadora después de introducción en la célula hospedadora, y con ello se replican junto con el genoma del hospedador. Además, ciertos vectores son capaces de dirigir la expresión de genes a los cuales están enlazados operativamente. En esta memoria se hace referencia a vectores de este tipo como "vectores de expresión". En general, los vectores de expresión de utilidad en las técnicas de DNA recombinante se encuentran a menudo en la forma de plásmidos. En la presente memoria descriptiva, "plásmido" y "vector" pueden utilizarse intercambiamente dado que el plásmido es la forma de vector más corrientemente  
40 utilizada. Sin embargo, la invención tiene por objeto incluir otras formas de vectores de expresión, tales como vectores virales (v.g., retrovirus deficientes en replicación, adenovirus y virus adeno-asociados), que desempeñan funciones equivalentes.

Los vectores de expresión recombinantes de la invención comprenden un ácido nucleico de la invención en una  
45 forma adecuada para la expresión del ácido nucleico en una célula hospedadora, lo que significa que los vectores de expresión recombinantes incluyen una o más secuencias reguladoras, seleccionadas sobre la base de las células hospedadoras a utilizar para la expresión, que está enlazada operativamente a la secuencia de ácido nucleico a expresar. En un vector de expresión recombinante, la expresión "enlazado operativamente" tiene por objeto significar que la secuencia de nucleótidos de interés está enlazada a la o las secuencias reguladoras de una manera que permite la  
50 expresión de la secuencia de nucleótidos (v.g., en un sistema de transcripción/traducción *in vitro* o en una célula hospedadora cuando el vector se introduce en la célula hospedadora). La expresión "secuencia reguladora" tiene por objeto incluir promotores, intensificadores y otros elementos de control de la expresión (v.g., señales de poliadenilación). Secuencias reguladoras de este tipo 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  
55 Biology and Biotechnology, compiladores Glick y Thompson, capítulo 7, 89-108, CRC Press: Boca Raton, Florida, con inclusión de las referencias que aparecen en dicho lugar. Secuencias reguladoras incluyen aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células hospedadoras y aquellas que dirigen la expresión de la secuencia de nucleótidos solamente en ciertas células hospedadoras o en ciertas condiciones. Será apreciado por los expertos en la técnica que el diseño del vector de expresión puede depender de factores tales como  
60 la elección de la célula hospedadora a transformar, el nivel de expresión de proteína deseado, etc. Los vectores de expresión de la invención pueden introducirse en células hospedadoras para producir con ello proteínas o péptidos, con inclusión de proteínas o péptidos de fusión, codificados por ácidos nucleicos como se describe en esta memoria (v.g., TFSRPs, formas mutantes de TFSRPs, proteínas de fusión, etc.).

65 Los vectores de expresión recombinantes de la invención pueden diseñarse para expresión de TFSRPs en células procariontas o eucariontas. Por ejemplo, los genes TFSRP pueden expresarse en células bacterianas tales como *C. glutamicum*, células de insecto (utilizando vectores de expresión de baculovirus), células de levadura y otras células fúngicas (véase Romanos, M.A. *et al.* 1992 Foreign gene expression in yeast: a review, 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, compiladores, 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.* compiladores, p. 1-28, Cambridge University Press: Cambridge), algas (Falcitore *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 siguiendo un método de transformación como el descrito en WO 98/01572 y células de plantas multicelulares (véase Schmidt, R. and 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, chapter 6/7, S. 71-119 (1993); F.F. White, B. Jenes *et al.* Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. Kung und R. Wu, 128-43, Academic Press: 1993; Potrykus, 1991 Annu. Rev. Plant Physiol. Plant Molec. Biol. 42:205-225 y las referencias citadas en dichos lugares) o células de mamífero. Células hospedadoras adecuadas se exponen adicionalmente en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press: San Diego, CA (1990). Alternativamente, el vector de expresión recombinante puede transcribirse y traducirse *in vitro*, utilizando por ejemplo secuencias reguladoras del promotor T7 y la polimerasa T7.

La expresión de proteínas en procariotas se realiza en muchos casos con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas, sean o no de fusión. Los vectores de fusión añaden cierto número de aminoácidos a una proteína codificada en ellos, usualmente al término amino de la proteína recombinante, pero también al término C, o se fusionan dentro de regiones adecuadas en las proteínas. Tales vectores de fusión desempeñan típicamente tres finalidades: 1) aumentar la expresión de una proteína recombinante; 2) aumentar la solubilidad de la proteína recombinante; y 3) ayudar a la purificación de la proteína recombinante por actuar como un ligando en la purificación por afinidad. A menudo, en los vectores de expresión de fusión, se introduce un sitio de escisión proteolítica en la unión del resto de fusión y la proteína recombinante para permitir la separación de la proteína recombinante del resto de fusión subsiguientemente a la purificación de la proteína de fusión. Enzimas de este tipo, y sus secuencias de reconocimiento cognadas, incluyen Factor Xa, trombina y enteroquinasa.

Vectores de expresión de fusión típicos 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), proteína E de fijación de maltosa, o proteína A, respectivamente, a la proteína recombinante diana. En una realización, la secuencia codificante de la TFSRP se somete a clonación en un vector de expresión pGEX para crear un vector que codifica una proteína de fusión que comprende, desde el término N al término C, GST-sitio de escisión de trombina-proteína X. La proteína de fusión puede purificarse por cromatografía de afinidad utilizando glutatión-resina de agarosa. La TFSRP recombinante no fusionada a GST puede recuperarse por escisión de la proteína de fusión con trombina.

Ejemplos de vectores de expresión de *E. coli* inducibles adecuados que no son de fusión incluyen pTrc (Amann *et al.* 1988 Gene 69:301-315) y pET 11d (Studier *et al.* *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). La expresión de genes diana por el vector pTrc está basada en la transcripción por la RNA-polimerasa del hospedador a partir de un promotor de fusión *trp-lac* híbrido. La expresión de genes diana por el vector pET 11d está basada en la transcripción a partir de un promotor de fusión *gn10-lac* T7 mediado por una RNA-polimerasa viral co-expresada (T7 *gn1*). Esta polimerasa viral es suministrada por las cepas hospedadoras BL21(DE3) o HMS174(DE3) a partir de un profago  $\lambda$  residente que alberga un gen T7 *gn1* bajo el control de transcripción del promotor *lacUV 5*.

Una estrategia para maximizar la expresión de proteínas recombinantes consiste en expresar la proteína en una bacteria hospedadora que tiene una capacidad deteriorada para escindir proteolíticamente la proteína recombinante (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Otra estrategia consiste en alterar la secuencia del ácido nucleico a insertar en un vector de expresión de tal manera que los codones individuales para cada aminoácido sean los utilizados preferentemente en la bacteria seleccionada para expresión, tal como *C. glutamicum* (Wada *et al.* 1992 Nucleic Acids Res. 20:2111-2118). Dicha alteración de secuencias de ácido nucleico de la invención puede realizarse por técnicas estándar de síntesis de DNA.

En otra realización, el vector de expresión de TFSRP es un vector de expresión de levadura. Ejemplos de vectores para expresión en la levadura *S. cerevisiae* incluyen pYepSec1 (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 (Invitrogen Corporation, San Diego, CA). Vectores y métodos para la construcción de vectores apropiados para uso en otros hongos, tales como los hongos filamentosos, incluyen 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.* compiladores, p. 1-28 Cambridge University Press: Cambridge.

Alternativamente, las TFSRPs de la invención pueden expresarse en células de insecto utilizando vectores de expresión de baculovirus. Vectores de expresión de baculovirus disponibles para expresión de proteínas en células de insecto cultivadas (v.g., células Sf9) 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).

En otra realización adicional, un ácido nucleico de TFSRP de la invención se expresa en células de mamífero utilizando un vector de expresión de mamífero. 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 utilizan en células de mamífero, las funciones de control del vector de expresión están proporcionadas a menudo por elementos reguladores virales. Por ejemplo, promotores utilizados comúnmente se derivan de poliovirus, Adenovirus 2, citomegalovirus y el Virus 40 de los Simios. Para otros sistemas de expresión adecuados para células tanto procariotas como eucariotas véanse los capítulos 16 y 17 de Sambrook, J. Fritsch, E.F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2ª Edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

En otra realización, el vector de expresión recombinante de mamífero es capaz de dirigir la expresión del ácido nucleico preferentemente en un tipo de célula particular (v.g., se utilizan elementos reguladores específicos de tejidos para expresar el ácido nucleico). Elementos reguladores específicos de tejidos se conocen la técnica. Ejemplos no limitantes de promotores específicos de tejidos adecuados incluyen el promotor albúmina (específico de hígado; Pinkert *et al.* 1987 Genes Dev. 1:268-277), promotores específicos linfoides (Calame y Eaton, 1988 Adv. Immunol. 43:235-275), en particular promotores de receptores de las 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 (v.g., el promotor de neurofilamentos; Byrne y Ruddle, 1989 PNAS 86:5473-5477), promotores específicos de páncreas (Edlund *et al.* 1985 Science 230:912-916), y promotores específicos de glándula mamaria (v.g., el promotor de lactosuero; patente U.S. No. 4.873.316 y la Publicación de Solicitud de Patente Europea No. 264.166). Están comprendidos también promotores regulados por desarrollo, por ejemplo, los promotores hox murinos (Kessel y Gruss, 1990 Science, 249:374-379) y el promotor de fetoproteína (Campes y Tilghman, 1989 Genes Dev. 3:537-546).

En otra realización, las TFSRPs de la invención pueden expresarse en células de plantas unicelulares (tales como algas) (véase Falciatore *et al.* 1999 Marine Biotechnology 1 (3):239-251 y las referencias citadas en dicho lugar) y células de plantas procedentes de plantas superiores (v.g., las espermatofitas, tales como plantas de cosecha). Ejemplos de vectores de expresión de plantas incluyen los 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.W., 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, compiladores: Kung y R. Wu, Academic Press, 1993; p. 15-38.

Una casete de expresión de plantas contiene preferiblemente secuencias reguladoras capaces de impulsar la expresión génica en células de plantas y que están enlazadas operativamente de tal manera que cada secuencia puede desempeñar su función, por ejemplo, la terminación de la transcripción por señales de poliadenilación. Señales de poliadenilación preferidas son las originarias de t-DNA 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 de los mismos, pero también son adecuados cualesquiera otros terminadores funcionalmente activos en plantas.

Dado que la expresión de genes de plantas no está limitada en muchos casos por niveles de transcripción, una casete de expresión de plantas contiene preferiblemente otras secuencias enlazadas operativamente tales como intensificadores de la traducción tales como la secuencia superimpulsora que contiene la secuencia conductora no traducida 5' del virus del mosaico del tabaco, que mejora la relación de proteína por RNA (Gallie *et al.* 1987 Nucl. Acids Research 15:8693-8711).

La expresión de genes de plantas debe enlazarse operativamente a un promotor apropiado que confiere la expresión de los genes de una manera oportuna, específica de la célula o del tejido. Se prefieren promotores que impulsan expresión constitutiva (Benfey *et al.* 1989 EMBO J. 8:2195-2202) como los derivados de virus de plantas tales como el CAMV 35S (Franck *et al.* 1980 Cell 21:285-294), el CaMV 19S (véase también la patente U.S. No. 5352605 y el documento WO 8402913) o promotores de plantas como los de la subunidad pequeña Rubisco descrita en la patente U.S. No. 4.962.028.

Otras secuencias preferidas para uso en casetes de expresión de genes de plantas son secuencias de direccionamiento necesarias para dirigir el producto génico a su compartimiento celular apropiado (para revisión, véase Kermode, 1996 Crit. Rev. Plant Sci. 15(4):285-423 y las referencias citadas en dicho lugar) tales como la vacuola, el núcleo, todos los tipos de plástidos tales como amiloplastos, cloroplastos, cromoplastos, el espacio extracelular, la mitocondria, el retículo endoplasmático, los cuerpos de aceite, peroxisomas y otros compartimientos de las células de las plantas.

La expresión de genes de plantas puede ser facilitada también por un promotor inducible (para revisión, véase Gatz, 1997 Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:89-108). Promotores químicamente inducibles son especialmente adecuados si se desea que ocurra la expresión génica de una manera específica en el tiempo. Ejemplos de tales promotores son un promotor inducible por ácido salicílico (documento WO 95/19443), un promotor inducible por tetraciclina (Gatz *et al.* 1992 Plant J. 2:397-404) y un promotor inducible por etanol (documento WO 93/21334).

Asimismo, promotores adecuados que responden a condiciones de estrés biótico o abiótico son aquellos tales como el promotor del gen PRP1 inducible por patógenos (Ward *et al.* 1993 Plant Mol. Biol. 22:361-366), el promotor hsp80 del tomate inducible por calor (Patente U.S. No. 5187267), el promotor alfa-amilasa de la patata inducible por frío (documento WO 96/12814) o el promotor pinII inducible por lesiones (Patente Europea No. 375091). Para otros

ejemplos de promotores inducibles por sequía, frío y sal, tales como el promotor RD29A, véase Yamaguchi-Shinozaki *et al.* (1993 Mol. Gen. Genet. 236:331-340).

Se prefieren especialmente aquellos promotores que confieren expresión génica en tejidos y órganos específicos, tales como células de guarda y las células de los pelos radicales. Promotores adecuados incluyen el promotor del gen napin de la colza (Patente U.S. No. 5.608.152), el promotor USP de *Vicia faba* (Baeumlein *et al.* 1991 Mol. Gen Genet. 225(3):459-67), el promotor oleosina de *Arabidopsis* (documento WO 98/45461), el promotor faseolina de *Phaseolus vulgaris* (Patente U.S. No. 5504200), el promotor Bce4 de Brassica (documento WO 91/13980) o el 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 tales como maíz, cebada, trigo, centeno, arroz, etc. Promotores adecuados que deben indicarse son el promotor del gen 1pt2 o 1pt1 de la cebada (documento WO 95/15389 y el documento WO 95/23230) o los descritos en el documento WO 99/16890 (promotores del gen hordeína de la cebada, el gen glutelina del arroz, el gen oricina del arroz, el gen prolamina del arroz, el gen gliadina del trigo, el gen glutelina del trigo, el gen zeína del maíz, el gen glutelina de la avena, el gen kasirina del sorgo y el gen secalina del centeno).

Son también especialmente adecuados promotores que confieren expresión de genes específicos de plástidos, dado que los plástidos son el compartimiento en el que se sintetizan los precursores y algunos productos finales de la biosíntesis de los lípidos. Promotores adecuados son el promotor de RNA-polimerasa viral descrito en el documento WO 95/16783 y en el documento WO 97/06250 y el promotor clpP de *Arabidopsis* descrito en el documento WO 99/43394.

La invención proporciona adicionalmente un vector de expresión recombinante que comprende una molécula de DNA de la invención clonada en el vector de expresión en orientación antisentido. Es decir, la molécula de DNA está enlazada operativamente a una secuencia reguladora de una manera que permite la expresión (por transcripción de la molécula de DNA) de una molécula de RNA que es antisentido respecto al mRNA de TFSRP. Pueden seleccionarse secuencias reguladoras enlazadas operativamente a una molécula de ácido nucleico clonada en orientación antisentido que dirigen la expresión continua de la molécula de RNA antisentido en una diversidad de tipos de células. Por ejemplo, pueden seleccionarse promotores y/o intensificadores virales, o secuencias reguladoras que dirigen la expresión constitutiva, específica de tejido o específica del tipo de células de RNA antisentido. El vector de expresión antisentido puede encontrarse en la forma de un plásmido recombinante, fagémido o virus atenuado en el cual los ácidos nucleicos antisentido se producen bajo el control de una región reguladora de alta eficiencia, cuya actividad puede estar determinada por el tipo de célula en el que se introduce el vector. Para una discusión de la regulación de la expresión génica utilizando genes antisentido véase Weintraub, H. *et al.* Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986 y Mol *et al.* 1990 FEBS Letters 268:427-430).

Otro aspecto de la invención se refiere a células hospedadoras en las cuales se ha introducido un vector de expresión recombinante de la invención. Las expresiones "célula hospedadora" y "célula hospedadora recombinante" se utilizan intercambiabilmente en esta memoria. Debe entenderse que dichas expresiones se refieren no sólo a la célula objeto particular sino que se aplican también a la progenie o progenie potencial de dicha célula. Dado que pueden producirse ciertas modificaciones en generaciones sucesivas debido a mutación o influencias ambientales, dicha progenie puede de hecho no ser idéntica a la célula originaria, pero se incluye todavía dentro del alcance de la expresión tal como se utiliza en esta memoria.

Una célula hospedadora puede ser cualquier célula procariota o eucariota. Por ejemplo, una TFSRP puede expresarse en células bacterianas tales como *C. glutamicum*, células de insecto, células fúngicas o células de mamífero (tales como células de ovario de hámster chino (CHO) o células COS), algas, ciliados, células de plantas, hongos u otros microorganismos como *C. glutamicum*. Otras células hospedadoras adecuadas son conocidas por los expertos en la técnica.

Puede introducirse DNA vector en células procariotas o eucariotas por técnicas convencionales de transformación o transfección. Tal como se utilizan en esta memoria, los términos "transformación", "transfección", "conjugación" y "transducción" tienen por objeto hacer referencia a una diversidad de métodos reconocidos en la técnica para la introducción de ácido nucleico extraño (v.g., DNA) en una célula hospedadora, con inclusión de co-precipitación con fosfato de calcio o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección, competencia natural, transferencia mediada por productos químicos y electroporación. Métodos adecuados para transformación o transfección de células hospedadoras con inclusión de células de plantas pueden encontrarse en Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, 2ª edición, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) y otros manuales de laboratorio tales como *Methods in Molecular Biology*, 1995, vol. 44, *Agrobacterium* protocols, compiladores: Gartland y Davey, Humana Press, Totowa, New Jersey. Dado que la tolerancia al estrés biótico o abiótico es un rasgo general que se desea heredar en una diversidad de plantas tales como maíz, trigo, centeno, avena, trigo híbrido, arroz, cebada, soja, cacahuete, algodón, colza y canola, mandioca, pimienta, girasol y tagetes, plantas solanáceas tales como patata, tabaco, berenjena y tomate, especies de *Vicia*, guisante, alfalfa, plantas arbustivas (café, cacao, té), especies de *Sáliz*, árboles (palma de aceite, cocotero), céspedes perennes y cosechas forrajeras, estas plantas de cosecha son también plantas diana preferidas para la ingeniería genética como una realización adicional de la presente invención.

En particular, la invención proporciona un método de producción de una planta transgénica con un ácido nucleico codificante de TFSRP, en donde la expresión del ácido nucleico en la planta da como resultado una tolerancia

incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta, que comprende: (a) transformar una célula de planta con un vector de expresión que comprende un ácido nucleico de TFSRP, y (b) generar a partir de la célula de la planta una planta transgénica con una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta. En realizaciones preferidas, la TFSRP es como se ha descrito arriba. En realizaciones preferidas, el ácido nucleico codificante de TFSRP es como se ha descrito arriba. La invención proporciona también un método de aumentar la expresión de un gen de interés en una célula hospedadora en comparación con una variedad de tipo salvaje de la célula hospedadora, en donde el gen de interés se transcribe en respuesta a una TFSRP, que comprende: (a) transformar la célula hospedadora con un vector de expresión que comprende un ácido nucleico codificante de TFSRP, y (b) expresar la TFSRP en la célula hospedadora, incrementando con ello la expresión del gen transcrito en respuesta a la TFSRP en comparación con una variedad de tipo salvaje de la célula hospedadora. En realizaciones preferidas, la TFSRP es como se ha descrito arriba. En realizaciones preferidas, el ácido nucleico codificante de TFSRP es como se ha descrito arriba.

Para dicha transformación de plantas, pueden utilizarse vectores binarios tales como pBinAR (Höfgen y Willmitzer, 1990 *Plant Science* 66:221-230). La construcción de los vectores binarios puede realizarse por ligación del cDNA en orientación de sentido o antisentido en el T-DNA. En la posición 5-prima respecto al cDNA, un promotor de planta activa la transcripción del cDNA. Una secuencia de poliadenilación está localizada en posición 3-prima respecto al cDNA. La expresión específica de tejido puede lograrse utilizando un promotor específico de tejido. Por ejemplo, la expresión específica de semilla puede lograrse por clonación del promotor *napin* o *LeB4* o *USP* en posición 5-prima respecto al cDNA. Asimismo, puede utilizarse cualquier otro elemento promotor específico de semilla. Para expresión constitutiva en la planta entera, puede utilizarse el promotor 35S de CaMV. La proteína expresada puede estar direccionada a un compartimiento celular utilizando un péptido señal, por ejemplo para plástidos, mitocondria o retículo endoplasmático (Kermode, 1996 (*Crit. Rev. Plant Sci.*, 1996 4(15):285-423). El péptido señal se somete a clonación en posición 5-prima en marco respecto al cDNA para conseguir la localización subcelular de la proteína de fusión.

La transformación de plantas mediada por *Agrobacterium* puede realizarse utilizando por ejemplo la cepa de *Agrobacterium tumefaciens* GV3101 (pMP90) (Koncz y Schell) 1986 *Mol. Gen. Genet.* 204:383-396) o LBA4404 (Clontech) de *Agrobacterium tumefaciens*. La transformación puede realizarse por técnicas estándar de transformación (Deblaere *et al.* 1994 *Nucl. Acids Res.* 13:4777-4788). En una realización, pueden utilizarse promotores que son sensibles a estados de estrés abiótico, tales como el promotor RD29A de *Arabidopsis*, con las secuencias de ácido nucleico descritas en esta memoria. Un experto en la técnica reconocerá que el promotor utilizado debería estar enlazado operativamente al ácido nucleico de tal modo que el promotor cause la transcripción del ácido nucleico, lo cual da como resultado la síntesis de un mRNA que codifica un polipéptido. Alternativamente, el RNA puede ser un RNA antisentido para uso con objeto de producir la expresión subsiguiente del mismo u otro gen o genes.

La transformación de plantas mediada por *Agrobacterium* puede realizarse utilizando técnicas estándar de transformación y regeneración (Gelvin, Stanton B. y Schilperoort, Robert A, *Plant Molecular Biology Manual*, 2ª Edición - Dordrecht: Kluwer Academic Publ., 1995. - in Sect., Ringbuc Zentrale 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 transformarse por transformación del cotiledón o el hipocótilo (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 la selección de *Agrobacterium* y plantas depende del vector binario y de la cepa de *Agrobacterium* utilizada para la transformación. La selección de colza se realiza normalmente utilizando kanamicina como marcador de plantas seleccionable. La transferencia de genes mediada por *Agrobacterium* al lino puede realizarse utilizando, por ejemplo, una técnica descrita por Mlynarova *et al.* 1994 *Plant Cell Report* 13:282-285. Adicionalmente, la transformación de la soja puede realizarse utilizando por ejemplo una técnica descrita en la Patente Europea No. 0424 047, la Patente U.S. No. 5.322.783 (Pioneer Hi-Bred International) o en la Patente Europea No. 0397 687, la Patente U.S. No. 5.376.543 y la Patente U.S. No. 5.169.770 (Universidad de Toledo).

La transformación de plantas puede realizarse por bombardeo con partículas, absorción del DNA mediada por polietileno-glicol o por la técnica de las fibras de carburo de silicio se describe, por ejemplo, por 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 U.S. No. 5.990.387 y un ejemplo específico de transformación del trigo puede encontrarse en el documento WO 93/07256.

Para la transfección estable de células de mamífero, se sabe que, dependiendo del vector de expresión y de la técnica de transfección utilizada, solamente una pequeña fracción de células puede integrar el DNA extraño en su genoma. A fin de identificar y seleccionar estos integrantes, se introduce generalmente un gen que codifica un marcador seleccionable (v.g., resistencia a antibióticos) en las células hospedadoras junto con el gen de interés. Marcadores seleccionables preferidos incluyen aquéllos que confieren resistencia a fármacos, tales como G418, higromicina y metotrexato o en plantas que confieren resistencia a un herbicida tal como glifosato o glufosinato. Las moléculas de ácido nucleico que codifican un marcador seleccionable pueden introducirse en una célula hospedadora en el mismo vector que el que codifica una TFSRP o pueden introducirse en un vector separado. Las células transfectadas de manera estable con la molécula de ácido nucleico introducida pueden identificarse, por ejemplo, mediante selección de fármacos (v.g., las células que han incorporado el gen marcador seleccionable sobrevivirán, mientras que las otras células morirán).

Para crear un microorganismo recombinante homólogo, se prepara un vector que contiene al menos una porción de un gen de TFSRP en el cual se ha introducido una delección, adición o sustitución para alterar con ello, v.g., desorganizar funcionalmente, el gen de TFSRP. Preferiblemente, este gen de TFSRP es un gen TFSRP de *Physcomitrella patens*, pero puede ser un homólogo de una planta afín o incluso de una fuente de mamífero, levadura, o insecto. En una realización preferida, el vector está diseñado de tal manera que, por recombinación homóloga, el gen de TFSRP endógeno se rompe funcionalmente (es decir, ya no codifica una proteína funcional; a esto se hace referencia también como un vector puesto fuera de combate (“knock-out”). Alternativamente, el vector puede diseñarse de tal manera que, después de la recombinación homóloga, el gen TFSRP endógeno resulta mutado o alterado de cualquier otro modo pero codifica todavía una proteína funcional (v.g., la región reguladora de aguas arriba puede estar alterada alterando con ello la expresión de la TFSRP endógena). Para crear una mutación puntual por recombinación homóloga, pueden utilizarse híbridos DNA-RNA en una técnica conocida como quimeroplastia (Cole-Strauss *et al.* 1999 Nucleic Acids Research 27(5): 1323-1330 y Kmiec, 1999 Gene therapy American Scientist 87(3):240-247). Los procedimientos de recombinación homóloga en *Physcomitrella patens* son también bien conocidos en la técnica y se contemplan para uso en esta invención.

Por lo que respecta al vector de recombinación homóloga, la porción alterada del gen TFSRP está flanqueada en sus extremos 5' y 3' por una molécula de ácido nucleico adicional del gen TFSRP que permite que ocurra la recombinación homóloga entre el gen TFSRP exógeno transportado por el vector y un gen de TFSRP endógeno en un microorganismo o planta. La molécula de ácido nucleico de TFSRP adicional flanqueante tiene una longitud suficiente para la recombinación homóloga exitosa con el gen endógeno. Típicamente, se incluyen en el vector varios centenares de pares de bases hasta kilobases de DNA flanqueante (en ambos extremos 5' y 3') (véase, v.g., Thomas, K. R., y Capecchi, M.R., 1987 Cell 51:503 para una descripción de vectores de recombinación homóloga o Strepp *et al.* 1998 PNAS, 95(8):4368-4373 para recombinación basada en cDNA en *Physcomitrella patens*). El vector se introduce en un microorganismo o célula de planta (v.g., por DNA mediado por polietilén-glicol), y se seleccionan las células en las cuales el gen de TFSRP introducido se ha recombinado homológamente con el gen TFSRP endógeno utilizando métodos conocidos en la técnica.

En otra realización, pueden producirse microorganismos recombinantes que contienen sistemas seleccionados que permiten la expresión regulada del gen introducido. Por ejemplo, la inclusión de un gen de TFSRP en un vector poniéndolo bajo control del operón lac permite la expresión del gen TFSRP únicamente en presencia de IPTG. Tales sistemas reguladores son bien conocidos en la técnica.

Una célula hospedadora de la invención, tal como una célula hospedadora procariota o eucariota en cultivo, puede utilizarse para producir (es decir, expresar) una TFSRP. Un método alternativo puede aplicarse adicionalmente en las plantas por la transferencia directa de DNA en flores en desarrollo por la vía de electroporación o transferencia de genes en medio de *Agrobacterium*. De acuerdo con ello, la invención proporciona adicionalmente métodos para producir TFSRPs utilizando las células hospedadoras de la invención. En una realización, el método comprende cultivar la célula hospedadora de la invención (en la cual se ha introducido un vector de expresión recombinante que codifica una TFSRP, o en cuyo genoma se ha introducido un gen que codifica una TFSRP de tipo salvaje o alterada) en un medio adecuado hasta que se produce la TFSRP. En otra realización, el método comprende adicionalmente aislar TFSRPs del medio o la célula hospedadora.

Otro aspecto de la invención se refiere a TFSRPs aisladas, y porciones biológicamente activas de las mismas. Una proteína “aislada” o “purificada” o porción biológicamente activa de la misma está exenta de algo del material celular cuando se produce por técnicas de DNA recombinante, o de precursores químicos u otros productos químicos cuando se sintetiza químicamente. La expresión “sustancialmente exenta de material celular” incluye preparaciones de TFSRP en las cuales la proteína está separada de algunos de los componentes celulares de las células en las que se produce la misma natural o recombinantemente. En una realización, la expresión “sustancialmente exenta de material celular” incluye preparaciones de TFSRP que tienen menos de aproximadamente 30% (expresado en peso seco) de material distinto de TFSRP (a lo que se hace referencia también en esta memoria como “proteína contaminante”), de modo más preferible menos de aproximadamente 20% de material distinto de TFSRP, de modo todavía más preferible menos de aproximadamente 10% de material distinto de TFSRP, y de modo muy preferible menos de aproximadamente 5% de material distinto de TFSRP. Cuando la TFSRP o la porción biológicamente activa de la misma se produce recombinantemente, la misma está también de modo preferible sustancialmente exenta de medio de cultivo, es decir, el medio de cultivo representa menos de aproximadamente 20%, de modo más preferible menos de aproximadamente 10%, y de modo muy preferible menos de aproximadamente 5% del volumen de la preparación de proteína. La expresión “sustancialmente exenta de precursores químicos u otros productos químicos” incluye preparaciones de TFSRP en las cuales la proteína está separada de precursores químicos u otros productos químicos que están implicados en la síntesis de la proteína. En una realización, la expresión “sustancialmente exenta de precursores químicos u otros productos químicos” incluye preparaciones de una TFSRP que tienen menos de aproximadamente 30% (expresado en peso seco) de precursores químicos o productos químicos distintos de TFSRP, de modo más preferible menos de aproximadamente 20% de precursores químicos o productos químicos distintos de TFSRP, de modo todavía más preferible menos de aproximadamente 10% de precursores químicos o productos químicos distintos de TFSRP, y de modo muy preferible menos de aproximadamente 5% de precursores químicos o productos químicos distintos de TFSRP. En realizaciones preferidas, proteínas aisladas, o porciones biológicamente activas de las mismas carecen de proteínas contaminantes del mismo organismo del que se deriva la TFSRP. Típicamente, tales proteínas se producen por expresión recombinante de, por ejemplo, una TFSRP de *Physcomitrella patens* en plantas distintas de *Physcomitrella patens* o microorganismos tales como *C. glutamicum*, ciliados, algas u hongos.



Una TFSRP aislada o una porción de la misma de la invención puede participar en una respuesta de tolerancia al estrés en una planta, o más particularmente puede participar en la transcripción de una proteína implicada en una respuesta de tolerancia al estrés en una planta de *Physcomitrella patens*, o tiene una o más de las actividades indicadas en la Tabla 1. En realizaciones preferidas, la proteína o porción de la misma comprende una secuencia de aminoácidos que es suficientemente homóloga a una secuencia de aminoácidos codificada por un ácido nucleico de SEQ ID NO: 10, tal que la proteína o porción de la misma mantiene la capacidad para participar en el metabolismo de los compuestos necesarios para la construcción de membranas celulares en *Physcomitrella patens*, o en el transporte de moléculas a través de dichas membranas. La porción de la proteína es preferiblemente una porción biológicamente activa como se describe en esta memoria. En otra realización preferida, una TFSRP de la invención tiene una secuencia de aminoácidos de SEQ ID NO: 17.

En otra realización preferida adicional, la TFSRP tiene una secuencia de aminoácidos que está codificada por una secuencia de nucleótidos que se hibrida, v.g., se hibrida en condiciones severas, a una secuencia de nucleótidos de SEQ ID NO: 10.

En otra realización preferida adicional, la TFSRP tiene una secuencia de aminoácidos que es al menos aproximadamente 50-60%, con preferencia al menos aproximadamente 60-70%, de modo más preferible al menos aproximadamente 70-80%, 80-90%, 90-95%, y de modo todavía más preferible al menos aproximadamente 96%, 97%, 98%, 99% o más homóloga a la secuencia de aminoácidos de SEQ ID NO: 17.

Las TFSRPs preferidas de la presente invención poseen también preferiblemente al menos una de las actividades de TFSRP descritas en esta memoria. Por ejemplo, una TFSRP preferida de la presente invención incluye una secuencia de aminoácidos codificada por una secuencia de nucleótidos que se hibrida, v.g., se hibrida en condiciones severas, a una secuencia de nucleótidos de SEQ ID NO: 10 y que puede participar en una respuesta de tolerancia al estrés en una planta, o más particularmente puede participar en la transcripción de una proteína implicada en una respuesta de tolerancia al estrés en una planta de *Physcomitrella patens*, o que tiene una o más de las actividades indicadas en la Tabla 1.

En otras realizaciones, la TFSRP es sustancialmente homóloga a una secuencia de aminoácidos de SEQ ID NO: 17, y retiene la actividad funcional de la proteína de la secuencia de SEQ ID NO: 17, pero difiere en secuencia de aminoácidos debido a variación natural o mutagénesis, como se ha descrito en detalle anteriormente. De acuerdo con ello, en otra realización, la TFSRP es una proteína que comprende una secuencia de aminoácidos que es al menos aproximadamente 50-60%, de modo preferible al menos aproximadamente 60-70%, de modo más preferible al menos aproximadamente 70-80%, 80-90%, 90-95%, y de modo muy preferible al menos aproximadamente 96%, 97%, 98%, 99% o más homóloga a una secuencia entera de aminoácidos de SEQ ID NO: 17, y que tiene al menos una de las actividades TFSRP descritas en esta memoria. En otra realización, la invención se refiere a una proteína entera de *Physcomitrella patens* que es sustancialmente homóloga a la secuencia entera de aminoácidos codificada por un ácido nucleico de SEQ ID NO: 10.

Porciones biológicamente activas de una TFSRP incluyen péptidos que comprenden secuencias de aminoácidos derivadas de las secuencias de aminoácidos de una TFSRP, v.g., una secuencia de aminoácidos de SEQ ID NO: 17 o la secuencia de aminoácidos de una proteína homóloga a una TFSRP, que incluyen menos aminoácidos que una TFSRP de longitud total o la proteína de longitud total que es homóloga a una TFSRP, y exhiben al menos una actividad de una TFSRP. Por regla general, porciones biológicamente activas (péptidos, v.g., péptidos que tienen, por ejemplo, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 o más aminoácidos de longitud) comprenden un dominio o motivo que tiene al menos una actividad de una TFSRP. Además, otras porciones biológicamente activas, en las cuales se han delecionado otras regiones de la proteína, pueden prepararse por técnicas recombinantes y evaluarse respecto a una o más de las actividades descritas en esta memoria. Preferiblemente, las porciones biológicamente activas de una TFSRP incluyen uno o más dominios/motivos seleccionados o porciones de los mismos que tienen actividad biológica.

Las TFSRPs se producen preferiblemente por técnicas de DNA recombinante. Por ejemplo, una molécula de ácido nucleico que codifica la proteína se clona en un vector de expresión (como se ha descrito arriba), se introduce el vector de expresión en una célula hospedadora (como se ha descrito arriba) y se expresa la TFSRP en la célula hospedadora. La TFSRP puede aislarse luego de las células por un esquema de purificación apropiado utilizando técnicas estándar de purificación de proteínas. Como alternativa a la expresión recombinante, una TFSRP, un polipéptido o péptido puede sintetizarse químicamente utilizando técnicas estándar de síntesis de péptidos. Además, la TFSRP nativa puede aislarse de células (v.g., *Physcomitrella patens*), por ejemplo utilizando un anticuerpo anti-TFSRP, que puede producirse por técnicas estándar utilizando una TFSRP o fragmento de la misma de esta invención.

La invención proporciona también proteínas TFSRP quiméricas o de fusión. Como se utiliza en esta memoria, una "proteína quimérica" o "proteína de fusión" TFSRP comprende un polipéptido de TFSRP enlazado operativamente a un polipéptido que no es de TFSRP. Un "polipéptido de TFSRP" hace referencia a un polipéptido que tiene una secuencia de aminoácidos que corresponde a una TFSRP, en tanto que un "polipéptido que no es de TFSRP" hace referencia a un polipéptido que tiene una secuencia de aminoácidos correspondiente a una proteína que no es sustancialmente homóloga a la TFSRP, v.g., una proteína que es diferente de la TFSRP y se deriva del mismo organismo o de un organismo diferente. Dentro de la proteína de fusión, el término "enlazado operativamente" tiene por objeto indicar que el polipéptido de TFSRP y el polipéptido que no es de TFSRP están fusionados uno a otro de tal manera que ambas secuencias cumplen la función propuesta atribuida a la secuencia utilizada. El polipéptido que no es de TFSRP puede

estar fusionado al término N o al término C del polipéptido de TFSRP. Por ejemplo, en una realización, la proteína de fusión es una proteína de fusión GST-TFSRP en la cual las secuencias de TFSRP están fusionadas al término C de las secuencias de GST. Tales proteínas de fusión pueden facilitar la purificación de GSBRRPs recombinantes. En otra realización, la proteína de fusión es una TFSRP que contiene una secuencia de señal heteróloga en su término N. En ciertas células hospedadoras (v.g., células hospedadoras de mamífero), la expresión y/o secreción de una TFSRP pueden incrementarse por el uso de una secuencia señal heteróloga.

Preferiblemente, una proteína TFSRP quimérica o de fusión de la invención se produce por técnicas estándar de DNA recombinante. Por ejemplo, fragmentos de DNA que codifican las diferentes secuencias de polipéptidos se ligan unos a otros en marco de acuerdo con técnicas convencionales, por ejemplo empleando términos de extremos romos o extremos cohesivos para ligación, digestión con enzimas de restricción para proporcionar los términos apropiados, rellenado de extremos cohesivos en caso apropiado, tratamiento con fosfatasa alcalina para evitar la unión indeseable y ligación enzimática. En otra realización, el gen de fusión puede sintetizarse por técnicas convencionales que incluyen sintetizadores automáticos de DNA. Alternativamente, puede llevarse a cabo la amplificación de fragmentos de genes por PCR utilizando iniciadores de anclaje que dan lugar a salientes complementarios entre dos fragmentos de genes consecutivos que pueden reasociarse y re-amplificarse subsiguientemente para generar una secuencia de gen quimérica (véase, por ejemplo, *Current Protocols in Molecular Biology*, compiladores Ausubel *et al.* John Wiley & Sons: 1992). Además, están disponibles comercialmente muchos vectores de expresión que codifican ya un resto de fusión (v.g., un polipéptido GST). Un ácido nucleico codificante de TFSRP puede clonarse en un vector de expresión de este tipo de tal modo que el resto de fusión está enlazado en marco a la TFSRP.

Homólogos de la TFSRP pueden generarse por mutagénesis, v.g., mutación puntual discreta o truncación de la TFSRP. Tal como se utiliza en esta memoria, el término "homólogo" hace referencia a una forma variante de la TFSRP que actúa como agonista o antagonista de la actividad de la TFSRP. Un agonista de la TFSRP puede retener sustancialmente la misma, o un subconjunto, de las actividades biológicas de la TFSRP. Un antagonista de la TFSRP puede inhibir una o más de las actividades de la forma existente naturalmente de la TFSRP, por ejemplo por fijación competitiva a un miembro situado aguas abajo o aguas arriba de la cascada metabólica de componentes de la membrana celular que incluye la TFSRP, o por fijación a una TFSRP que media el transporte de compuestos a través de dichas membranas, impidiendo con ello que tenga lugar la translocación.

En una realización alternativa, los homólogos de la TFSRP pueden identificarse por escrutinio de genotecas combinatorias de mutantes, v.g., mutantes de truncación, de la TFSRP respecto a actividad agonista o antagonista de TFSRP. En una realización, se genera una genoteca diversificada de variantes de TFSRP por mutagénesis combinatoria al nivel del ácido nucleico y se codifica por una genoteca de genes diversificada. Una genoteca diversificada de variantes de TFSRP puede producirse mediante, por ejemplo, ligación enzimática de una mezcla de oligonucleótidos sintéticos en secuencias de genes de tal modo que pueda expresarse un conjunto degenerado de secuencias potenciales de TFSRP como polipéptidos individuales, o alternativamente, como una serie de proteínas de fusión mayores (v.g., para presentación de fago) que contienen la serie de secuencias de TFSRP en ellas. Existen una diversidad de métodos que pueden utilizarse para producir genotecas de homólogos potenciales de TFSRP a partir de una secuencia de oligonucleótidos degenerada. La síntesis química de una secuencia de genes degenerada puede realizarse en un sintetizador automático de DNA, y el gen sintético se liga luego a un vector de expresión apropiado. El uso de una serie degenerada de genes permite la provisión, en una mezcla, de todas las secuencias codificantes de la serie deseada de secuencias potenciales de TFSRP. Métodos para la síntesis de oligonucleótidos degenerados se conocen en la técnica (véase, v.g., Narang, S.A., 1983 *Tetrahedron* 39: 3; Itakura *et al.* 1984 *Annu. Rev. Biochem.* 53: 323; Itakura *et al.* 1984 *Science* 198: 1056; Ike *et al.* 1983 *Nucleic Acid Res.* 11: 477).

Adicionalmente, pueden utilizarse genotecas de fragmentos de la codificación de GSBRRP para generar una población diversificada de fragmentos de TFSRP para escrutinio y selección subsiguiente de homólogos de una TFSRP. En una realización, puede generarse una genoteca de fragmentos de secuencia codificantes por tratamiento de un fragmento PCR bicatenario de una secuencia codificante de TFSRP con una nucleasa en condiciones en las cuales se produce mellado sólo aproximadamente una vez por molécula, desnaturalización del DNA bicatenario, renaturalización del DNA para formar DNA bicatenario que puede incluir pares sentido/antisentido de productos mellados diferentes, eliminación de porciones monocatenarias de los dúplex reformados por tratamiento con nucleasa S1, y ligación de la genoteca de fragmentos resultante en un vector de expresión. Por este método, puede derivarse una genoteca de expresión que codifica fragmentos N-terminales, C-terminales e internos de diversos tamaños de la TFSRP.

Se conocen en la técnica varios métodos para escrutar productos génicos de genotecas combinatorias producidas por mutaciones puntuales o truncación, y para el escrutinio de genotecas de cDNA para productos génicos que tengan una propiedad seleccionada. Dichas técnicas son adaptables para escrutinio rápido de las genotecas de genes generadas por la mutagénesis combinatoria de homólogos de TFSRP. Las técnicas utilizadas más frecuentemente, que son susceptibles de análisis de alta capacidad, para escrutinio de grandes genotecas de genes incluyen típicamente clonación de la genoteca de genes en vectores de expresión replicables, transformación de células apropiadas con la genoteca de vectores resultante, y expresión de los genes combinatorios en condiciones en las cuales la detección de una actividad deseada facilita el aislamiento del vector que codifica el gen cuyo producto se detectó. La mutagénesis recursiva de conjunto (REM), una nueva técnica que mejora la frecuencia de mutantes funcionales en las genotecas, puede utilizarse en combinación con los ensayos de escrutinio para identificar homólogos de TFSRP (Arkin y Yourvan, 1992 *PNAS* 89: 7811-7815; Delgrave *et al.* 1993 *Protein Engineering* 6 (3): 327-331). En otra realización, pueden aprovecharse ensayos basados en células para analizar una genoteca de TFSRP diversificada, utilizando métodos bien conocidos en

la técnica. La presente invención proporciona adicionalmente un método de identificación de una nueva TFSRP, que comprende (a) generar una respuesta específica de anticuerpos para una TFSRP, o fragmento de la misma, como se describe en esta memoria; (b) escrutar material supuesto de TFSRP con el anticuerpo, en donde la fijación específica del anticuerpo al material indica la presencia de una TFSRP potencialmente nueva; y (c) analizar el material fijado en comparación con TFSRP conocidas, para determinar su novedad.

Las moléculas de ácido nucleico, proteínas, homólogos de proteínas, proteínas de fusión, iniciadores, vectores, y células hospedadoras descritas en esta memoria pueden utilizarse en uno o más de los métodos siguientes: identificación de *Physcomitrella patens* y organismos afines; mapeado de genomas de organismos afines a *Physcomitrella patens*; identificación y localización de secuencias de interés de *Physcomitrella patens*; estudios de evolución; determinación de regiones de TFSRP requeridas para función; modulación de una actividad de TFSRP; modulación del metabolismo de una o más funciones celulares; modulación del transporte transmembranal de uno o más compuestos; y modulación de la resistencia al estrés.

El musgo *Physcomitrella patens* representa un miembro de los musgos. El mismo es afín a otros musgos tales como *Ceratodon purpureus* que es capaz de crecer en ausencia de luz. Musgos tales como *Ceratodon* y *Physcomitrella* comparten un alto grado de homología en la secuencia de DNA y el nivel de polipéptidos permitiendo el uso de escrutinio heterólogo de moléculas de DNA con sondas procedentes de otros musgos u organismos, permitiendo así la derivación de una secuencia de consenso adecuada para escrutinio heterólogo o anotación funcional y predicción de funciones de genes en terceras especies. La posibilidad de identificar tales funciones puede tener por tanto importancia significativa, v.g. la predicción de especificidad de sustrato de enzimas. Adicionalmente, estas moléculas de ácido nucleico pueden servir como puntos de referencia para el mapeado de genomas de musgos, o de genomas de organismos afines.

Las moléculas de ácido nucleico de TFSRP de la invención tienen una diversidad de usos. Es muy importante que las secuencias de ácido nucleico y aminoácidos de la presente invención pueden utilizarse para transformar plantas, induciendo con ello tolerancia a estados de estrés tales como sequía. La presente invención proporciona por tanto una planta transgénica transformada por un ácido nucleico codificante de TFSRP, en donde la expresión de la secuencia de ácido nucleico en la planta da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta. La planta transgénica puede ser una monocotiledónea o una dicotiledónea. La invención estipula adicionalmente que la planta transgénica puede seleccionarse de maíz, trigo, centeno, avena, trigo híbrido, arroz, cebada, soja, cacahuete, algodón, colza, canola, mandioca, pimiento, girasol, tagetes, plantas solanáceas, patata, tabaco, berenjena, tomate, especies de Vicia, guisante, alfalfa, café, cacao, té, especies de *Sáliz*, palma de aceite, cocotero, céspedes perennes y cosechas forrajeras, por ejemplo. En particular, la presente invención describe la utilización de la expresión de CBF-1 (SEQ ID NO: 17) para producir por ingeniería genética plantas tolerantes a la sequía. Esta estrategia ha sido demostrada en esta invención para *Arabidopsis thaliana*, colza/canola, habas de soja, maíz y trigo, pero su aplicación no está restringida a estas plantas. De acuerdo con ello, la invención proporciona una planta transgénica que contiene una TFSRP seleccionada de CBF-1 como se define arriba, con inclusión de homólogos, en donde el estrés ambiental es sequía.

La presente invención proporciona también métodos de modificación de la tolerancia al estrés de una planta, que comprenden modificar la expresión de una TFSRP en la planta. La invención estipula que este método puede realizarse de tal manera que la tolerancia al estrés se aumenta o se reduce.

Adicionalmente, este método puede utilizarse en plantas transgénicas o no transgénicas. En los casos en que la planta es transgénica, la planta puede transformarse con un vector que contiene cualquiera de los ácidos nucleicos codificantes de TFSRP arriba descritos, o la planta puede transformarse con un promotor que dirige la expresión de TFSRP nativa en la planta, por ejemplo. La invención estipula que dicho promotor puede ser específico de tejido. Adicionalmente, un promotor de este tipo puede estar regulado por el desarrollo. En una alternativa, plantas no transgénicas pueden tener expresión de TFSRP nativa modificada por inducción de un promotor nativo. Adicionalmente, la invención estipula que la expresión de TFSRP puede modificarse por administración de una molécula antisentido que inhibe la expresión de TFSRP.

La expresión de CBF-1 (SEQ ID NO: 17), en plantas diana puede realizarse por uno de los ejemplos siguientes, pero sin carácter limitante: (a) promotor constitutivo, (b) promotor inducible por estrés, (c) promotor inducido por productos químicos, y (d) sobre-expresión de un promotor modificado por ingeniería genética con, por ejemplo, factores de transcripción derivados de dedos de cinc (Greisman y Pabo, 1997 Science 275:657). El último caso implica la identificación de los homólogos de CBF-1 (SEQ ID NO: 17), en la planta diana así como de su promotor. Se modifican por ingeniería genética factores de transcripción recombinantes que contienen dedos de cinc para interactuar específicamente con el homólogo de CBF-1 (SEQ ID NO: 17) y se activa la transcripción del gen correspondiente.

Como se muestra en esta memoria y se describe con mayor detalle más adelante, la expresión de la TFSRP CBF-1 (SEQ ID NO: 17) en *Arabidopsis thaliana* confiere a la planta un alto grado de tolerancia a la sequía.

En condiciones de estrés por sequía, las líneas que sobreexpresaban CBF-1 exhibían una tasa de supervivencia de 100%. Los controles no transformados exhibían una tasa de supervivencia de 10%. Es digno de mención que los análisis de estas líneas transgénicas se realizaron con plantas T1. Por tanto, los resultados serán mejores cuando se encuentre un expresante homocigótico fuerte. Una prueba adicional de la implicación de estos genes en la tolerancia

al estrés viene dada por el aumento en el nivel de material transcrito en respuesta al tratamiento a baja temperatura. La concentración de los transcritos para CBF-1 se incrementa al doble con respecto al ruido de fondo sin tratar después del tratamiento.

5 Además de introducir las secuencias de ácido nucleico de TFSRP en plantas transgénicas, estas secuencias pueden utilizarse también para identificar un organismo como perteneciente a la especie *Physcomitrella patens* o un pariente próximo de la misma. Asimismo, aquéllas pueden utilizarse para identificar la presencia de *Physcomitrella patens* o un pariente de la misma en una población mixta de microorganismos. La invención proporciona las secuencias de ácido nucleico de varios genes de *Physcomitrella patens*; por sonda del DNA genómico extraído de un cultivo de  
10 una población singular o mixta de microorganismos en condiciones severas con una sonda que abarca una región de un gen de *Physcomitrella patens* que es exclusiva de este organismo, puede comprobarse si está presente este organismo.

15 Adicionalmente, las moléculas de ácido nucleico y proteína de la invención pueden servir como marcadores para regiones específicas del genoma. Esto tiene utilidad no sólo en el mapeado del genoma, sino también en estudios funcionales de proteínas de *Physcomitrella patens*. Por ejemplo, para identificar la región del genoma a la que se fija una proteína particular de fijación de DNA de *Physcomitrella patens*, el genoma de *Physcomitrella patens* podía digerirse, e incubarse los fragmentos con la proteína de fijación de DNA. Aquellos fragmentos que fijan la proteína pueden son-  
20 darse adicionalmente con las moléculas de ácido nucleico de la invención, preferiblemente con marcadores fácilmente detectables; la fijación de una molécula de ácido nucleico de este tipo al fragmento del genoma permite la localización del fragmento al mapa genómico de *Physcomitrella patens*, y, cuando se realiza múltiples veces con enzimas diferentes, facilita una determinación rápida de la secuencia de ácido nucleico a la que se fija la proteína. Adicionalmente, las moléculas de ácido nucleico de la invención pueden ser suficientemente homólogas a las secuencias de especies afines de tal modo que estas moléculas de ácido nucleico pueden servir como marcadores para la construcción de un mapa  
25 genómico en musgos afines.

Las moléculas de ácido nucleico de TFSRP de la invención son también útiles para estudios evolutivos y estructurales de proteínas. Los procesos metabólicos y de transporte en los cuales participan las moléculas de la invención son utilizados por una diversidad de células procariotas y eucariotas; por comparación de las secuencias de las moléculas  
30 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, una comparación de este tipo 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 de la proteína que son esenciales para el funcionamiento de la enzima. Este tipo de determinación es valioso para estudios de ingeniería de proteínas y puede proporcionar una indicación de lo que puede tolerar la proteína en  
35 términos de mutagénesis sin pérdida de función.

La manipulación de las moléculas de ácido nucleico de TFSRP de la invención puede dar como resultado la producción de TFSRPs que tengan diferencias funcionales de las TFSRPs de tipo salvaje. Estas proteínas pueden mejorarse en eficiencia o actividad, pueden estar presentes en mayores números en la célula de lo que es usual, y pueden tener eficiencia o actividad reducidas. Existen varios mecanismos por los cuales la alteración de una TFSRP  
40 de la invención puede afectar directamente a la respuesta al estrés y/o la tolerancia al estrés. En el caso de plantas que expresan TFSRPs, el transporte incrementado puede conducir a un reparto mejorado de sal y/o solutos en el tejido y los órganos de la planta. Por aumento del número o la actividad de las moléculas transportadoras que exportan moléculas iónicas de la célula, puede ser posible afectar a la tolerancia de la célula a la sal.  
45

El efecto de la modificación genética en plantas, *C. glutamicum*, hongos, algas, o ciliados sobre la tolerancia al estrés puede evaluarse dejando crecer el microorganismo o planta modificado(a) en condiciones inferiores a las adecuadas y analizando posteriormente las características de crecimiento y/o metabolismo de la planta. Dichos métodos de análisis son bien conocidos por un experto en la técnica, e incluyen peso seco, peso húmedo, síntesis de proteínas, síntesis de carbohidratos, síntesis de lípidos, tasas de evapotranspiración, rendimiento de la planta en general 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).  
55

60 Por ejemplo, vectores de expresión de levadura que comprenden los ácidos nucleicos descritos en esta memoria, o fragmentos de los mismos, pueden construirse y transformarse en *Saccharomyces cerevisiae* utilizando protocolos estándar. Las células transgénicas resultantes pueden ensayarse luego en cuanto a fallo o alteración de su tolerancia a la sequía. Análogamente, vectores de expresión de plantas que comprenden los ácidos nucleicos descritos en esta memoria, o fragmentos de los mismos, pueden construirse y transformarse en una célula de planta apropiada tal como *Arabidopsis*, soja, colza, maíz, trigo, *Medicago truncatula*, etc., utilizando protocolos estándar. Las células transgénicas y/o plantas resultantes derivadas de ello pueden ensayarse luego en cuanto a fallo o alteración de su tolerancia a la sequía.  
65

La modificación por ingeniería genética de uno o más genes de TFSRP de la invención puede dar también como resultado TFSRPs que tienen actividades alteradas que afectan indirectamente a 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 (v.g., 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 la tirosina, desactivando 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 excretados típicamente, las células pueden alterarse genéticamente para transportar más productos de lo que es típico para una célula de tipo salvaje. Por optimización de la actividad de una o más TFSRPs de la invención que están involucradas 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 descritas en esta memoria, o fragmentos de las mismas, pueden utilizarse para generar mutaciones del tipo “knockout” (es decir 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 “knockout” resultantes pueden evaluarse luego respecto a su aptitud 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 desactivación de genes véase la Patente U.S. 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 TFSRPs que dan como resultado resistencia incrementada 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 descritos en esta memoria, las moléculas de ácido nucleico y proteínas de la invención pueden utilizarse para generar algas, ciliados, plantas, hongos u otros microorganismos tales como *C. glutamicum* que expresan moléculas de ácido nucleico y proteínas TFSRP mutadas tales que la tolerancia al estrés se mejora.

La presente invención proporciona también anticuerpos que se fijan específicamente a un polipéptido de TFSRP, o una porción del mismo, tal como es codificado(a) por un ácido nucleico dado a conocer en esta memoria o como se describe en esta memoria. Los anticuerpos pueden producirse por muchos métodos bien conocidos (véase, v.g. Harlow and Lane, “Antibodies; A Laboratory Manual” Cold Spring Harbor laboratory, Cold Spring Harbor, Nueva York, (1988)). Resumidamente, puede inyectarse antígeno purificado en un animal en una cantidad y en intervalos suficientes para provocar una respuesta inmunitaria. Los anticuerpos pueden purificarse directamente, o pueden obtenerse 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 escrutar genotecas de clones de ácido nucleico en cuanto a células secretoras del antígeno. Los 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 fija selectivamente” y “se fija específicamente” relacionadas con el polipéptido hacen referencia a una reacción de fijación que es determinante de la presencia de la proteína en una población heterogénea de proteínas y otros compuestos biológicos. Así, en condiciones de inmunoensayo diseñadas, los anticuerpos específicos fijados a una proteína particular no se fijan en una cantidad significativa a otras proteínas presentes en la muestra. La fijación selectiva de un anticuerpo en tales condiciones puede requerir un anticuerpo que se seleccione por su especificidad para una proteína particular. Pueden utilizarse una diversidad de formatos de inmunoensayo para seleccionar anticuerpos que se fijan selectivamente a una proteína particular. Por ejemplo, se utilizan rutinariamente inmunoensayos ELISA en fase sólida para seleccionar anticuerpos que son selectivamente inmunorreactivos con una proteína. Véase Harlow and 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 fijación selectiva.

En algunos casos, es deseable preparar anticuerpos monoclonales a partir de diversos hospedadores. Una descripción de técnicas para preparar tales anticuerpos monoclonales puede encontrarse en Stites *et al.* compiladores, “Basic and Clinical Immunology” (Lange Medical Publications, Los Altos, Calif., 4ª edición) y las referencias citadas en dicho lugar, y en Harlow and Lane, (“Antibodies; A Laboratory Manual” Cold Spring Harbor Publications, Nueva York, (1988)).

A lo largo de esta solicitud se han citado diversas publicaciones. Las descripciones de la totalidad de estas publicaciones y las referencias citadas en dichas publicaciones se incorporan en sus totalidades por la presente por 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 que antecede hace referencia a realizaciones 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 los ejemplos que siguen, que no deben interpretarse en modo alguno como imposición de limitaciones en cuanto al alcance de aquélla. Por el contrario, debe entenderse claramente que se puede recurrir a varias otras realizaciones, modificaciones, y equivalentes de las mismas que, después de la lectura de la presente descripción, pueden ser ideadas en sí mismas por los expertos en la técnica sin apartarse del espíritu de la presente invención y/o del

alcance de las reivindicaciones adjuntas. Adicionalmente, todas las referencias citadas en esta memoria se incorporan expresamente por la presente por referencia.

## 5 Ejemplos

### Ejemplo 1

#### Crecimiento de cultivos de *Physcomitrella patens*

10 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 recogida 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  
15 de regeneración de los gametofitos. El protonema se desarrolló a partir de la spora haploide como un cloronema rico en cloroplastos y caulonema pobre en cloroplastos, sobre el cual se formaron yemas al cabo de aproximadamente 12 días. Estas crecieron para dar gametóforos que llevaban 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.

20 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<sup>-1</sup>m<sup>2</sup> (luz blanca; tubo fluorescente Philips TL 65W/25) y un cambio luz/oscuridad de 16/8 horas. El musgo se modificó 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 de RNA y DNA se cultivaron en cultivos líquidos aireados. Los protonemas  
25 se trituraron cada nueve días y se transfirieron a medio de cultivo fresco.

### Ejemplo 2

#### Aislamiento de DNA total a partir de plantas

Los detalles para el aislamiento de DNA total se refieren al tratamiento de un gramo de peso fresco de material de planta. Los materiales utilizados incluyen los tampones siguientes: tampón CTAB: 2% (p/v) de bromuro de N-cetil-N,N,N-trimetilamonio (CTAB); Tris HCl 100 mM de pH 8,0; NaCl 1,4 M; EDTA 20 mM; tampón de N-laurilsarcosina: 10% (p/v) de N-laurilsarcosina; Tris HCl 100 mM de pH 8,0; EDTA 20 mM.  
35

El material de planta se trituró bajo nitrógeno líquido en un mortero para dar un polvo fino y se transfirió a matraces Eppendorf de 2 ml. El material de planta congelado se cubrió luego con una capa de 1 ml de tampón de descomposición (1 ml de tampón CTAB, 100  $\mu$ l de tampón de N-laurilsarcosina, 20  $\mu$ l de  $\beta$ -mercaptoetanol y 10  
40  $\mu$ l de solución de proteinasa K, 10 mg/ml) y se incubó a 60°C durante una hora con agitación continua mediante sacudidas. El homogeneizado obtenido se distribuyó en dos matraces Eppendorf (de 2 ml) y se extrajo dos veces mediante sacudidas 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 la temperatura ambiente durante 15 minutos en cada caso. El DNA se precipitó luego a -70°C durante 30 minutos utilizando isopropanol enfriado en hielo. El DNA precipitado se sedimentó  
45 a 4°C y 10.000 g durante 30 minutos y se resuspendió en 180  $\mu$ l de tampón TE (Sambrook *et al.* 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6). Para purificación ulterior, el DNA se trató 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 un paso de lavado con etanol al 70%, el DNA se secó y se recogió subsiguientemente en 50  $\mu$ l de H<sub>2</sub>O + RNAsa (concentración final 50 mg/ml). El DNA se disolvió durante una noche a 4°C y la digestión con RNAsa se  
50 llevó a cabo subsiguientemente a 37°C durante una hora. El almacenamiento del DNA tuvo lugar a 4°C.

### Ejemplo 3

#### Aislamiento de RNA total y RNA poli(A)<sup>+</sup> y construcción de una genoteca de cDNA de *Physcomitrella patens*

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

Para la construcción de la genoteca de cDNA, se realizó la síntesis de la primera cadena utilizando transcriptasa inversa del Virus de la Leucemia de los Murinos (Roche, Mannheim, Alemania) e iniciadores oligo-d(T), y la síntesis de la segunda cadena por incubación con DNA-polimerasa I, enzima Klenow y digestión con RNAsaH a 12°C (2 horas),  
65 16°C (1 hora) y 22°C (1 hora). La reacción se paró por incubación a 65°C (10 minutos) y se transfirió subsiguientemente a hielo. Las moléculas de DNA bicatenario se hicieron romas con DNA-polimerasa T4 (Roche, Mannheim) a 37°C (30 minutos). Los nucleótidos se separaron por extracción con fenol/cloroformo y columnas centrífugas Sepha-

## ES 2 279 777 T3

5 dex G50. Se ligaron adaptadores EcoRI (Pharmacia, Freiburg, Alemania) a los extremos del cDNA con DNA-ligasa T4 (Roche, 12°C, una noche) y se fosforilaron por incubación con polinucleótido-quinasa (Roche, 37°C, 30 minutos). Esta mezcla se sometió a separación en un gel de agarosa de punto de fusión bajo. Las moléculas de DNA mayores que 300 pares de bases se eluyeron del gel, se extrajeron con fenol, se concentraron en columnas Elutip-D (Schleicher y Schuell, Dassel, Alemania), se ligaron a ramas vectoras y se empaquetaron 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.

### 10 Ejemplo 4

#### *Secuenciación y anotación de la función de ESTs de Physcomitrella patens*

15 Se utilizaron genotecas de cDNA como se describe en el Ejemplo 2 para secuenciación de DNA de acuerdo con métodos estándar, y en particular, por el método de terminación de cadenas utilizando el Kit ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer, Weiterstadt, Alemania). Se realizó una secuenciación aleatoria subsiguientemente a la recuperación preparativa del plásmido a partir de genotecas de cDNA por escisión en masa *in vivo*, retransformación, y extensión subsiguiente de DH10B en placas de agar (detalles de material y protocolo de Stratagene, Ámsterdam, Países Bajos). Se preparó DNA plasmídico a partir de cultivos de *E. coli* desarrollados durante una noche que crecían en medio Caldo 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 DNA Qiagene (Qiagen, Hilden), de acuerdo con los protocolos del fabricante. Se utilizaron iniciadores de secuenciación con las secuencias de nucleótidos siguientes:

25                   5'-CAGGAAACAGCTATGACC-3'                   SEQ ID NO:24  
                  5'-CTAAAGGGAACAAAAGCTG-3'                   SEQ ID NO:25  
                  5'-TGTA AAAACGACGGCCAGT-3'                   SEQ ID NO:26

30 Las secuencias se procesaron y anotaron utilizando el paquete de soporte lógico EST-MAX proporcionado comercialmente por Bio-Max (Munich, Alemania). El programa incorpora prácticamente todos los métodos bioinformáticos importantes para la caracterización funcional y estructural de secuencias de proteínas. Para referencia, puede consultarse el sitio de la *web pedant.mips.biochem.mpg.de*. Los algoritmos más importantes incorporados en EST-MAX son: FASTA: Búsquedas muy sensibles de bases de datos de secuencias con estimaciones de significación estadística; Pearson W.R. (1990) Comparación de secuencias rápida y sensible con FASTP y FASTA. *Methods Enzymol.* 183:63-98; BLAST: Búsquedas muy sensibles de bases de datos de secuencias con estimaciones de significación estadística. Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. Herramienta de búsqueda de alineación local básica. *Journal of Molecular Biology* 215:403-10; PREDATOR: Predicción de alta exactitud de la estructura secundaria de secuencias simples y múltiples. Frishman, D. and Argos, P. (1997) 75% de exactitud en la predicción de la estructura secundaria de proteínas. *Proteins*, 27:329-335; CLUSTALW: Alineación de secuencias múltiples. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: Mejora de la sensibilidad de la alineación de secuencias progresivas múltiples por ponderación de la secuencia, penalidades por laguna específicas de las posiciones y elección de la matriz de peso. *Nucleic Acids Research*, 22: 4673-4680; TMAP: Predicción de la región transmembranal a partir de secuencias múltiplemente alineadas. Persson, B. and Argos, P. (1994) Predicción de segmentos transmembranales en proteínas utilizando alineaciones de secuencias múltiples. *J. Mol. Biol.* 237:182-192; ALOM2: Predicción de la región transmembranal de secuencias simples. 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. *Biochim. 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 semejanza contra una base de datos de bloques sin lagunas. J.C. Wallace and Henikoff S., (1992); PATMAT: Un programa de búsqueda y extracción para consultas y bases de datos de secuencia, patrón y bloque, *CABIOS* 8:249-254. Escrito por Bill Alford.

### 55 Ejemplo 5

#### *Identificación de ORFs de Physcomitrella patens correspondientes a CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1 y CABF-2*

60 Los cDNAs parciales de *Physcomitrella patens* (ESTs) que se muestran en la Tabla 1 siguiente se identificaron en el programa de secuenciación de *Physcomitrella patens* EST utilizando el programa EST-MAX por análisis BLAST. (Las Tablas 2-8 muestran algunos de estos resultados). Los Números de Identificación de Secuencia correspondientes a estos ESTs son como sigue: CABF-1 (SEQ ID NO: 1), DBF-1 (SEQ ID NO: 2), CBF-1 (SEQ ID NO: 3), HDZ-1 (SEQ ID NO: 4), ZF-1 (SEQ ID NO: 5), LZ-1 (SEQ ID NO: 6), y CABF-2 (SEQ ID NO: 7)

## ES 2 279 777 T3

Estos clones particulares se seleccionaron para análisis ulteriores, dado que codificaban factores de transcripción.

TABLA 1

Categoría Funcional	Función Supuesta	Código de Secuencia	Posición ORF	Nombre
Factor de Transcripción	proteína de fijación de DNA factor de transcripción, fijación de CCAAT, cadena factor de transcripción, proteína de dedo de cinc homeodominio de proteína con cremallera de leucina proteína de fijación de DNA VBP1 activador de transcripción CBF1	s_pp001031077f	1- 515	DBF-1 ejemplo comparativo
		c_pp004053131r	500-1	CABF-1 ejemplo comparativo
		s_pp004052093r	2-508	CABF-2 ejemplo comparativo
		c_pp001074039r	1154-447	ZF-1 ejemplo comparativo
		c_pp001058012r	364-750	HDZ-1 ejemplo comparativo
		s_pp013006061r	1-371	LZ-1 ejemplo comparativo
		c_pp004032055r	183-998	CBF-1 de acuerdo con la invención

TABLA 2

*Grado de identidad y semejanza de aminoácidos de PpHDZ-1 y otras proteínas homólogas (se utilizó el programa GCG Gap; penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)*

Swiss-Prot #	Q9LS31	Q9LS33	Q43529	Q9XH37	Q9SP47
Nombre de proteína	Proteína Homeobox PPHB7	Proteína Homeobox PPHB5	Homeobox	Homeodominio de proteína con cremallera de leucina	Homeodominio de proteína 57 con cremallera de leucina
Especie	Physcomitrella patens (musgo)	Physcomitrella patens (musgo)	Lycopersicon esculentum (tomate)	Oryza sativa (arroz)	Glycine max (soja)
Identidad %	71%	38%	30%	29%	30%
Semejanza %	72%	51%	40%	39%	36%



ES 2 279 777 T3

TABLA 3

Grado de identidad y semejanza de aminoácidos de PpZF-1 y otras proteínas homólogas (se utilizó el programa GCG Gap; penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)

Swiss-Prot #	Q9SK53	Q9ZTK7	Q9ZTK8	Q9XE47	O82431
Nombre de proteína	Proteína de dedo de cinc de caja B cuasi-constante	Proteína 2 cuasi-constante	Proteína 1 cuasi-constante	Proteína de dedo de cinc	Proteína 1 cuasi-constante
Especie	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Malus domestica Manzana (Malus silvestris)	Pinus radiata (Pino de Monterrey)	Raphanus sativus (Rábano)
Identidad %	40%	43%	42%	39%	41%
Semejanza %	50%	54%	54%	49%	53%

TABLA 4

Grado de identidad y semejanza de aminoácidos de PpCABF-1 y otras proteínas homólogas (se utilizó el programa GCG Gap; penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)

Swiss-Prot #	Q9ZQC3	O23310	P25209	Q9LFI3	O23633
Nombre de proteína	Factor de transcripción de fijación de CCAAT supuesto	Subunidad A del factor de transcripción de fijación de CCAAT	Subunidad A del factor de transcripción de fijación de CCAAT	Factor de transcripción NF-Y, proteína semejante a la de fijación de CCAAT	Factor de transcripción
Especie	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Zea mays (Maíz)	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)
Identidad %	47%	53%	49%	41%	46%
Semejanza %	58%	56%	57%	53%	52%

ES 2 279 777 T3

TABLA 5

Grado de identidad y semejanza de aminoácidos de PpDBF-1 y otras proteínas homólogas (se utilizó el programa GCG Gap; penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)

Swiss-Prot #	Q9ZUL5	O45609	Q9NPU9
Nombre de proteína	Proteína de fijación de DNA supuesta	Proteína M03C11.8	Proteína de 68,6 KDA hipotética
Especie	Arabidopsis thaliana (Berro de oreja de ratón)	Caenorhabditis elegans	Homo sapiens (Humana)
Identidad %	47%	24%	25%
Semejanza %	58%	35%	37%

TABLA 6

Grado de identidad y semejanza de aminoácidos de PpCABF-1 y otras proteínas homólogas (se utilizó el programa GCG Gap; penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)

Swiss-Prot #	O23636	Q9SNZ0	Q9SMP0	Q92869	O35088
Nombre de proteína	Factor de transcripción	Proteína activada por hemo	Factor de transcripción HAP5A	Subunidad del factor de transcripción NF-YC	Factor nuclear YC
Especie	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Homo sapiens (Humana)	Mus musculus (Ratón)
Identidad %	54%	40%	42%	26%	25%
Semejanza %	62%	49%	49%	31%	30%

ES 2 279 777 T3

TABLA 7

Grado de identidad y semejanza de aminoácidos de PpLZ-1 y otras proteínas homólogas (se utilizó el programa GCGGap: penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)

5

10

15

20

25

Swiss-Prot #	Q9SQK1	P43273	O24160	Q06979	Q41558
Nombre de proteína	Factor de transcripción BZIP	Factor de transcripción HBP-1B	Transcripción de cremallera de leucina	Factor 3,2 de fijación del elemento OCS	Factor de transcripción HBP-1B(C1)
Especie	Nicotiana tabacum (Tabaco común)	Arabidopsis thaliana (Tabaco común)	Nicotiana glauca 2.1 tabacum (Tabaco común)	Zea mays (Maiz)	Triticum aestivum (Trigo)
Identidad %	62%	73%	46%	46%	45%
Semejanza %	74%	61%	55%	53%	53%

30

TABLA 8

Grado de identidad y semejanza de aminoácidos de PpCBF-1 y otras proteínas homólogas (se utilizó el programa GCGGap: penalidad por laguna: 10; penalidad por extensión de laguna: 0,1; matriz de registro: blosum 62)

35

40

45

50

55

Swiss-Prot # o GenBank #	Q9M210	BAA33435	Q9LU18	Q9ZQP3	Q9SUK8
Nombre de proteína	Proteína semejante a factor de transcripción	DREB1B	Proteína semejante al factor de transcripción TINY	Proteína TINY supuesta	Dominio Apétala 2 de proteína semejante a TINY
Especie	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana (Berro de oreja de ratón)	Arabidopsis thaliana Berro de oreja de ratón
Identidad %	22%	21%	21%	20%	20%
Semejanza %	35%	32%	32%	30%	27%

Ejemplo 6

60 Clonación del cDNA de longitud total de *Physcomitrella patens* codificante de CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2

65 Para aislar CABF-1 (SEQ ID NO: 8), CABF-2 (SEQ ID NO: 14), CBF-1 (SEQ ID NO: 10) de longitud total se realizó una PCR como se describe más adelante bajo el título "Amplificación de Longitud Total" utilizando los ESTs originales descritos en el Ejemplo 4 como molde dado que los mismos eran de longitud total (véase la Tabla 9 para los iniciadores).

ES 2 279 777 T3

TABLA 9

*Esquema e iniciadores utilizados para clonación de clones de longitud total*

Gen	Sitios finales en producto	Método de aislamiento	Carrera de iniciadores	Iniciador RT-PCR
<b>DBF-1</b>	<b>XmaI/HpaI</b>	5' RACE y RT-PCR para clon de longitud total	RC056 (SEQ ID NO:27) 5'GCGATCCTCAGC CTGTCGATCCATT3'  RC116 (SEQ ID NO:28) 5'CCCTGAGGTATC GTTCTGTTCCCA 3'	RC279 (SEQ ID NO:29) 5'ATCCCGGGCGAT GGTGCGTTTCGAGAT CGTAAGG3'  RC280 (SEQ ID NO:30) 5'GCGTTAACGAGC TTTCTCGCAGTGCC AGATAA3'
<b>CABF-2</b>	<b>XmaI/SacI</b>	PCR del clon original EST		RC031 (SEQ ID NO:31) ATCCCGGGCTCTGC ACCCCAGATGTCGC ATCCT RC032: (SEQ ID NO:32) CTGAGCTCTAATGC ATTCACTGTTGCTG CTGCT

ES 2 279 777 T3

LZ-1	HpaI/EcoRV	5' RACE y RT-PCR para clon de longitud total	RC058 (SEQ ID NO:33) 5'CCTGTAGGGCCA CCCGGAGCTCACT3	RC108 (SEQ ID NO:34) 5'GAGTTAACGCAG TGGTCACAACGCA GAGTACGC3' RC109 (SEQ ID NO:35) 5'GCGATATCGCTTC CATACTGCGCCGA AGACTT3'
CBF-1	XmaI/HpaI	PCR del clon EST original		RC033 (SEQ ID NO:36) 5'GACCCGGGCCAT GTGATATGGCTTCA AAGTAT3' RC034 (SEQ ID NO:37) 5'GCGTTAACGACT CACTGAGAGTCAT AATGGTG3'
HDZ-1	XmaI/HpaI	5' RACE y RT-PCR para clon de longitud total	RC047 (SEQ ID NO:38) 5'CGTAGTCGOGCT CGAGCTGTTTGGT3'	RC321 (SEQ ID NO:39) 5'ATCCCGGGCACC AGGGCAAGAGGGG ATAGAGAC3' RC322 (SEQ ID NO:40) 5'GCGTTAACGCCG ATGGTGCAACTTTG GTTGAC3'

5

10

15

20

25

30

35

40

45

50

55

60

65

ZF-1	XmaI/SacI	5' RACE y RT-PCR para clon de longitud total	RC063 (SEQ ID NO:41) 5'COGTGTCCTCGG AGCATTCTGGCAT3	RC122 (SEQ ID NO:42) 5'ATCCCGGGAGGA GGGAGTTGGAATCT AGGAGAC3' RC124 (SEQ ID NO:43) 5'GCGAGCTCGACC TTGCTCGATGGAGA CTCCAAT3'
CABF-1	XmaI/SacI	PCR del clon EST original		RC019 (SEQ ID NO:44) 5'ATCCCGGGAATA GGACGGATGGCCG ACAGTTAC3' RC020 (SEQ ID NO:45) 5'ATGAGCTCAC TCTTACACTCCGCG GGGTTGGTT3'

Para aislar los clones codificantes de DBF-1 (SEQ ID NO: 9), HDZ-1 (SEQ ID NO: 11), ZF-1 (SEQ ID NO: 12) y LZ-1 (SEQ ID NO: 13) de *Physcomitrella patens*, se crearon genotecas de cDNA con el kit SMART RACE cDNA Amplification (Clontech Laboratories) siguiendo las instrucciones del fabricante. Se utilizó como molde RNA total aislado como se describe en el Ejemplo 2. Los cultivos se trataron antes del aislamiento del RNA como sigue: Estrés por Sal: 2, 6, 12, 24, 48 horas con medio complementado con NaCl 1-M; Estrés por Frío: 4°C para los mismos tiempos que en el caso de la sal; Estrés por Sequía: los cultivos se incubaron sobre papel de filtro seco durante los mismos tiempos anteriores. El RNA se retiró luego y se utilizó para aislamiento.

#### Protocolo 5' RACE

Se utilizaron las secuencias EST DBF-1 (SEQ ID NO: 2), HDZ-1 (SEQ ID NO: 4), ZF-1 (SEQ ID NO: 5) y LZ-1 (SEQ ID NO: 6) identificadas por la búsqueda de bases de datos como se describe en el Ejemplo 4 para diseñar oligos para RACE (véase la Tabla 9). Las secuencias extendidas para estos genes se obtuvieron por realización de la reacción en cadena de la polimerasa Rapid Amplification of cDNA Ends (RACE PCR) utilizando el kit Advantage 2 PCR (Clontech Laboratories) y el kit de amplificación de cDNA SMART RACE (Clontech Laboratories) empleando un termociclador Biometra T3 siguiendo las instrucciones del fabricante. Las secuencias obtenidas por las reacciones RACE correspondían a las regiones codificantes de longitud total para HDZ-1, ZF-1 y LZ-1 y se utilizaron para diseñar oligos para clonación de longitud total de los genes respectivos (véase a continuación bajo "Amplificación de Longitud total"). El producto RACE de DBF-1 no era de longitud total y fue necesaria una nueva reacción RACE (véase la Tabla 9 para los iniciadores).

#### Amplificación de Longitud total

Se obtuvieron clones de longitud total correspondientes a CABF-1 (SEQ ID NO: 8), CBF-1 (SEQ ID NO: 10) y CABF-2 (SEQ ID NO: 14) realizando la reacción en cadena de la polimerasa (PCR) con iniciadores específicos de genes (véase la Tabla 9) y el EST original como molde. Las condiciones para la reacción eran condiciones estándar con DNA-polimerasa PWO (Roche). La PCR se realizó 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.

## ES 2 279 777 T3

Los clones de longitud total para DBF-1 (SEQ ID NO: 9), HDZ-1 (SEQ ID NO: 11), ZF-1 (SEQ ID NO: 12) Y LZ-1 (SEQ ID NO: 13) se aislaron por repetición del método RACE pero utilizando los iniciadores específicos de genes que se dan en la Tabla 9.

5 Los fragmentos amplificados se extrajeron del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) y se ligaron al vector TOPO pCR 2.1 (Invitrogen) siguiendo las instrucciones del fabricante. Los vectores recombinantes se transformaron 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). Las células transformadas se seleccionaron 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 se había dejado crecer durante una noche a 37°C. Se seleccionaron las colonias blancas y se utilizaron 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 DNA plasmídico 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).

### Ejemplo 7

20 *Modificación por ingeniería genética de plantas de Arabidopsis tolerantes al estrés por sobre-expresión de los genes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1 y CABF-2*

#### *Construcción del vector binario: pGMSG*

25 El vector pLMNC53 (Mankin, 2000, Thesis PhD) se digirió con HindIII (Roche) y se rellenó en extremos romos con enzima Klenow y dNTPs 0,1 mM (Roche) de acuerdo con las instrucciones del fabricante. Este fragmento se extrajo del gel de agarosa con el Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. El fragmento purificado se digirió luego con EcoRI (Roche) de acuerdo con las instrucciones del fabricante. Este fragmento se extrajo del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. El fragmento de 1,4 kilobases resultante, la casete de gentamicina, incluía el promotor nos, el gen aacCI y el terminador g7.

35 El vector pBlueScript se digirió con EcoRI y SmaI (Roche) de acuerdo con las instrucciones del fabricante. El fragmento resultante se extrajo del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. El vector pBlue-Script digerido y los fragmentos de la casete de gentamicina se ligaron con DNA-ligasa T4 (Roche) de acuerdo con las instrucciones del fabricante, uniendo los dos sitios EcoRI respectivos y uniendo el sitio de extremos romos HindIII con el sitio SmaI.

40 El vector recombinante (pGMBS) se transformó en células Top10 (Invitrogen) utilizando condiciones estándar. Las células transformadas se seleccionaron 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), y se dejaron crecer durante una noche a 37°C. Se seleccionaron las colonias blancas y se utilizaron para inocular 3 ml de LB líquido que contenía 100 µg/ml de ampicilina y se dejaron crecer durante una noche a 37°C. El DNA plasmídico se extrajo utilizando el Kit QIAprep Spin Miniprep (Qiagen) siguiendo las instrucciones del fabricante. Los análisis de los clones subsiguientes y el mapeado de restricción se realizaron 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.).

50 Tanto el vector pGMBS como el vector p1bx SuperGUS se digirieron con XbaI y KpnI (Roche) de acuerdo con las instrucciones del fabricante, se escindió la casete de gentamicina de pGMBS y se produjo la cadena principal del vector p1bx SuperGUS. Los fragmentos resultantes se extrajeron del gel de agarosa con un Kit de Extracción de Gel QIAquick (Qiagen) de acuerdo con las instrucciones del fabricante. Estos dos fragmentos se ligaron con DNA-ligasa T4 (Roche) de acuerdo con las instrucciones del fabricante.

55 El vector recombinante resultante (pGMSG) se transformó en células Top10 (Invitrogen) utilizando condiciones estándar. Las células transformadas se seleccionaron 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 se había dejado crecer durante una noche a 37°C. Se seleccionaron las colonias blancas y se utilizaron 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 DNA plasmídico 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).

65 *Subclonación de CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1 y CABF-2 en el vector binario*

Los fragmentos que contenían los diferentes factores de transcripción de *Physcomitrella patens* se subclonaron a partir de los vectores recombinantes PCR2.1 TOPO por doble digestión con enzimas de rescisión (véase la Tabla 10)

de acuerdo con las instrucciones del fabricante. El fragmento de la subsecuencia se escindió del gel de agarosa con un Kit de Extracción de Gel QIAquick (QIAGEN) de acuerdo con las instrucciones del fabricante, y se ligó al vector binario pGMSG, se escindió con las enzimas apropiadas (véase Tabla 10) y se desfosforiló antes de la ligación. El pGMSG recombinante resultante contenía el factor de transcripción correspondiente en la orientación de sentido bajo el control del super-promotor constitutivo.

TABLA 10

Nombres de los diversos constructos de los factores de transcripción de *Physcomitrella patens* utilizados para transformación de plantas

Gen	Enzimas utilizadas para generar el fragmento de gen	Enzimas utilizadas para restringir pGMSG	Constructo Vector Binario
<b>CABF-1</b>	<b>XmaI/SacI</b>	<b>XmaI/SacI</b>	pBPSSH003 Ejemplo comparativo
<b>DBF-1</b>	<b>XmaI/HpaI</b>	<b>XmaI/Ecl136</b>	pBPSLVM009 Ejemplo comparativo
<b>CBF-1</b>	<b>XmaI/HpaI</b>	<b>XmaI/Ecl136</b>	pBPSSH002 de acuerdo con la invención
<b>HDZ-1</b>	<b>XmaI/HpaI</b>	<b>XmaI/Ecl136</b>	pBPSLVM007 Ejemplo comparativo
<b>ZF-1</b>	<b>XmaI/SacI</b>	<b>XmaI/SacI</b>	pBPSLVM008 Ejemplo comparativo
<b>LZ-1</b>	<b>HpaI/EcoRV</b>	<b>Ecl136</b>	pBPSLVM0012 Ejemplo comparativo
<b>CABF-2</b>	<b>XmaI/SacI</b>	<b>XmaI/SacI</b>	pBPSSI003 Ejemplo comparativo

### 35 Transformación en *Agrobacterium*

Los vectores recombinantes se transformaron en *Agrobacterium tumefaciens* C58C1 y PMP90 de acuerdo con condiciones estándar (Hoefgen y Willmitzer, 1990).

### 40 Transformación de plantas

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

### 45 Escrutinio de las Plantas Transformadas

Se utilizaron semillas T1 de acuerdo con protocolos estándar (Xiong *et al.* 1999, Plant Molecular Biology Reporter 17:159-170). Se extendieron las semillas sobre medio ½ MS con 0,6% de agar complementado con 1% de sacarosa, 150 µg/ml de gentamicina (Sigma-Aldrich) y 2 µg/ml de benomil (Sigma-Aldrich). Las semillas en las placas se vernalizaron durante 4 días a 4°C. Las semillas se dejaron germinar en una cámara climatizada a una temperatura del aire de 22°C y con intensidad lumínica de 40 micromoles<sup>-1m<sup>2</sup></sup> (luz blanca); tubo fluorescente Philips TL 65W/25) con 16 horas de luz y 8 horas de oscuridad como ciclo de longitud diurna. Las plantas de semillero transformadas se seleccionaron al cabo de 14 días y se transfirieron a medio ½ MS con placas de agar complementadas con 1% de sacarosa y se dejaron recuperar durante 5-7 días.

### 55 Escrutinio de Tolerancia a la Sequía

Se transfirieron plantas de semillero T1 a papel de filtro estéril seco en una cápsula Petri y se dejaron desecar durante 2 horas a 80% RH (humedad relativa) en una Cámara Sanyo Growth MLR-350H, micromoles<sup>-1m<sup>2</sup></sup> (luz blanca; tubo fluorescente Philips TL 65W/25). Se redujo luego la RH a 60% y se desecaron ulteriormente las plantas de semillero durante 8 horas. Las plantas de semillero se retiraron luego y se pusieron sobre placas de agar ½ MS 0,6% complementadas con 2 µg/ml de benomil y se evaluaron después de 5 días.

Los resultados del escrutinio de tolerancia a la sequía en plantas de *Arabidopsis thaliana* que sobre-expresaban el gen TFSRP se muestran en la Tabla 11. Es digno de mención que estos análisis se realizaron con plantas T1, por lo que los resultados deberían ser mejores cuando se encuentre un expresante homocigótico fuerte.



## ES 2 279 777 T3

TABLA 11

*Sumario de los ensayos de estrés por sequía*

Nombre del Gen	Ensayo de Estrés por Sequía		
	Número de supervivientes	Número total de plantas	Porcentaje de supervivientes
<b>HDZ-1</b>	<b>7</b>	<b>14</b>	<b>50%</b>
<b>ZF-1</b>	<b>25</b>	<b>45</b>	<b>53%</b>
<b>CABF-1</b>	<b>8</b>	<b>9</b>	<b>89%</b>
<b>DBF-1</b>	<b>4</b>	<b>5</b>	<b>80%</b>
<b>CABF-2</b>	<b>3</b>	<b>6</b>	<b>50%</b>
<b>LZ-1</b>	<b>11</b>	<b>14</b>	<b>79%</b>
<b>CBF-1</b>	<b>9</b>	<b>9</b>	<b>100%</b>
<b>Control</b>	<b>18</b>	<b>84</b>	<b>21%</b>

### *Escrutinio de Tolerancia a la Sal*

Se transfirieron plantas de semillero a papel de filtro impregnado en 1/2 MS y se pusieron sobre 1/2 MS con agar al 0,6% complementado con 2  $\mu\text{g/ml}$  de benomil la noche anterior al escrutinio de tolerancia a la sal. Para el escrutinio de tolerancia a la sal, el papel de filtro con las plantas de semillero se desplazó a pilas de papel de filtro estéril, impregnado en NaCl 50 mM, en una cápsula Petri. Después de 2 horas, el papel de filtro con las plantas de semillero se movió a pilas de papel de filtro estéril, impregnado con NaCl 200 mM, en una cápsula Petri. Después de 2 horas, el papel de filtro con las plantas de semillero se movió a pilas de papel de filtro estéril, impregnadas en NaCl 600 mM, en una cápsula Petri. Después de 10 horas, las plantas de semillero se movieron a cápsulas Petri que contenían 1/2 MS con agar al 0,6% complementado con 2  $\mu\text{g/ml}$  de benomil. Las plantas de semillero se evaluaron después de 5 días.

Los resultados del escrutinio y tolerancia a la sal en plantas de *Arabidopsis thaliana* que sobre-expresaba las TFSRPs se muestran en la Tabla 12. En particular, las plantas de *Arabidopsis thaliana* que sobre-expresaban ZF-1 exhibían una tasa de supervivencia de 52% (12 supervivientes de 23 plantas estresadas); LZ-1, 48% (10 supervivientes de 21 plantas estresadas); CABF-2, 56% (5 supervivientes de 9 plantas estresadas); mientras que el control sin transformar exhibía una tasa de supervivencia de 9% (2 supervivientes de 23 plantas estresadas). Es digno de mención que estos análisis se realizaron con plantas T1, y por consiguiente los resultados deberían ser mejores cuando se encuentre un expresante homocigótico fuerte.

TABLA 12

*Sumario de los ensayos de estrés por sal*

Nombre del Gen	Ensayo de Estrés por Sal		
	Número de supervivientes	Número total de plantas	Porcentaje de supervivientes
<b>ZF-1</b>	<b>12</b>	<b>23</b>	<b>52%</b>
<b>CABF-2</b>	<b>5</b>	<b>9</b>	<b>56%</b>
<b>LZ-1</b>	<b>10</b>	<b>21</b>	<b>48%</b>
<b>Control</b>	<b>2</b>	<b>23</b>	<b>9%</b>

## ES 2 279 777 T3

### Escrutinio de Tolerancia a la Helada

Se transportaron plantas de semillero a cápsulas Petri que contenían 1/2 MS con agar al 0,6% complementado con 2% de sacarosa y 2 µg/ml de benomil. Después de 4 días, las plantas de semillero se incubaron a 4°C durante una hora y se cubrieron luego con hielo machacado. Las plantas de semillero se pusieron luego en una Cámara Ambiental Environmental Specialist ES2000 y se incubaron durante 3,5 horas comenzando a -1,0°C y descendiendo -1°C/hora. Las plantas de semillero se incubaron luego a -5,0°C durante 24 horas y se dejaron descongelar después a 5°C durante 12 horas. El agua se retiró por vertido y las plantas de semillero se registraron al cabo de 5 días. Las plantas transgénicas se escrutan luego por su tolerancia al frío mejorada, demostrando que la expresión del transgén confiere tolerancia al frío.

### Ejemplo 8

#### 15 *Detección de los transgenes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2 en las líneas transgénicas de Arabidopsis*

Una hoja de una planta de tipo salvaje y una planta transgénica de *Arabidopsis* se homogeneizó en 250 µl de tampón de bromuro de hexadeciltrimetil-amonio (CTAB) (2% CTAB, NaCl 1,4 M, EDTA 8 mM y Tris 20 mM de pH 8,0) y 1 µl de β-mercaptoetanol. Las muestras se incubaron a 60-65°C durante 30 minutos y se añadieron luego 250 µl de cloroformo a cada muestra. Las muestras se agitaron enérgicamente durante 3 minutos y se centrifugaron durante 5 minutos durante 18.000 x g. Se retiró el sobrenadante de cada muestra y se añadieron 150 µl de isopropanol. Las muestras se incubaron a la temperatura ambiente durante 15 minutos, y se centrifugaron durante 10 minutos a 18.000 x g. Cada pellet se lavó con etanol a 70%, se secó, y se resuspendió en 20 µl de TE. Se utilizaron 4 µl de la suspensión anterior en una reacción PCR de 20 µl utilizando DNA-polimerasa Taq (Roche Molecular Biochemicals) de acuerdo con las instrucciones del fabricante. El plásmido del vector binario que contenía cada gen de TFSRP se utilizó como control positivo, y como control negativo se utilizó DNA genómico C24 de tipo salvaje en las reacciones PCR. La reacción PCR de 10 µl se analizó en gel de agarosa/bromuro de etidio al 0,8%.

Los iniciadores y los tiempos de reacción utilizados para amplificación de cada gen TFSRP se indican a continuación. Es notable que los transgenes se amplificaron con éxito a partir de las líneas transgénicas T1, pero no del tipo salvaje T24. Este resultado indica que las plantas transgénicas T1 contienen al menos una copia de los transgenes. No había indicación alguna de la existencia de genes idénticos o muy similares en el control de *Arabidopsis thaliana* sin transformar que pudo amplificarse por este método.

#### CABF-1

Los iniciadores utilizados en las reacciones fueron:

**5'GAATAGATACGCTGACACGC3'**

**SEQ ID NO:46**

**5'ATGAGCTCACTCTTACACTCCGCGGGGTTGGTT3'**

**SEQ ID NO:47**

El programa PCR fue: 1 ciclo de 1 minuto a 94°C, 1 minuto a 75°C y 3 minutos a 72°C, seguido por 14 ciclos del mismo ciclo excepto que la temperatura de reasociación disminuyó 1°C en cada ciclo hasta 62°C; y luego 16 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 3 minutos a 72°C. Se generó un fragmento de 600 pares de bases a partir del control positivo y las plantas transgénicas.

#### HDZ-1

Los iniciadores utilizados en las reacciones fueron:

**5'GAATAGATACGCTGACACGC3'**

**SEQ ID NO:46**

**5'GCGTTAACGCCGATGGTGCAACTTTGGTTGAC3'**

**SEQ ID NO:48**

El programa PCR fue como sigue: 30 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 4 minutos a 72°C, seguido por 10 minutos a 72°C. Se produjo un fragmento de 1,3 kb a partir del control positivo y las plantas transgénicas.

## ES 2 279 777 T3

### ZF-1

Los iniciadores utilizados en las reacciones fueron:

5                    **5'GAATAGATACGCTGACACGC3'**                    **SEQ ID NO:46**  
                      **5'GCGAGCTCGACCTTGCTCGATGGAGACTCCAAT3'**                    **SEQ ID NO:49**

10        El programa PCR fue como sigue: 1 ciclo de 1 minuto a 94°C, 1 minuto a 75°C y 3 minutos a 72°C, seguido por 14 ciclos del mismo ciclo excepto que la temperatura de reasociación disminuyó 1°C en cada ciclo hasta 62°C; y luego 16 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 3 minutos a 72°C. Se generó un fragmento de 1,3 kb a partir del control positivo y las plantas transgénicas T1.

### 15        CBF-1

Los iniciadores utilizados en las reacciones fueron:

20                    **5'GAATAGATACGCTGACACGC3'**                    **SEQ ID NO:46**  
                      **5'GCGTTAACGACTCACTGAGAGTCATAATGGTG3'**                    **SEQ ID NO:50**

25        El programa PCR fue como sigue: 1 ciclo de 1 minuto a 94°C, 1 minuto a 75°C y 3 minutos a 72°C, seguido por 14 ciclos del mismo ciclo excepto que la temperatura de reasociación disminuyó 1°C en cada ciclo hasta 62°C; y luego 16 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 3 minutos a 72°C. Se generó un fragmento de 1,1 kb a partir del control positivo y las plantas transgénicas T1.

### 30        DBF-1

Los iniciadores utilizados en las reacciones fueron:

35                    **5'CTAGTAACATAGATGACACC3'**                    **SEQ ID NO:51**  
                      **5'ATCCCGGGCGATGGTGCGTTCGAGATCGTAAGG3'**                    **SEQ ID NO:52**

40        El programa PCR fue como sigue: 30 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 4 minutos a 72°C, seguido por 10 minutos a 72°C. Se produjo un fragmento de 2,9 kb a partir del control positivo y las plantas transgénicas.

### 45        CABF-2

Los iniciadores utilizados en las reacciones fueron:

50                    **5'GAATAGATACGCTGACACGC3'**                    **SEQ ID NO:53**  
                      **5'CTGAGCTCTAATGCATTCCTGTTGCTGCTGCT3'**                    **SEQ ID NO:54**

55        El programa PCR fue como sigue: 30 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 4 minutos a 72°C, seguido por 10 minutos a 72°C. Se produjo un fragmento de 800 pb a partir del control positivo y las plantas transgénicas.

### LZ-1

60        Los iniciadores utilizados en las reacciones fueron:

**5'GAATAGATACGCTGACACGC3'**                    **SEQ ID NO:53**  
                      **5'GCGATATCGCTTCCATACCTGCGCCGAAGACTT3'**                    **SEQ ID NO:55**

65        El programa PCR fue como sigue: 30 ciclos de 1 minuto a 94°C, 1 minuto a 62°C y 4 minutos a 72°C, seguido por 10 minutos a 72°C. Se produjo un fragmento de 1,8 kb a partir del control positivo y las plantas transgénicas.

## ES 2 279 777 T3

### Ejemplo 9

#### *Detección del mRNA de los transgenes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2 en líneas transgénicas de Arabidopsis*

5 La expresión del transgén se detectó utilizando RT-PCR. Se aisló el RNA 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 se trituraron a un polvo fino en nitrógeno líquido. Se resuspendió el tejido triturado en 500  $\mu$ l de una mezcla 1:1 a 80°C de fenol a tampón de extracción (LiCl 100 mM, Tris 100 mM de pH 8, EDTA 10 mM, SDS al 1%), seguido por agitación enérgica breve para mezclar. Después de la adición de 250  $\mu$ l de cloroformo, se agitó brevemente cada muestra con intensidad. Se centrifugaron luego las muestras durante 5 minutos a 12.000 x g. La fase acuosa superior se llevó a un tubo Eppendorf nuevo. Se precipitó el RNA por adición de una décima parte de su volumen de acetato de sodio 3 M y dos volúmenes de etanol de 95%. Se mezclaron las muestras por inversión y se dejaron en hielo durante 30 minutos. El RNA se redujo a un sedimento por centrifugación a 12.000 x g durante 10 minutos. Se retiró el sobrenadante y los pelets se secaron brevemente al aire. Los pelets de las muestras de RNA se resuspendieron en 10  $\mu$ l de agua tratada con DEPC. Para eliminar el DNA contaminante de las muestras, se trató cada una con DNasa exenta de RNasa (Roche) de acuerdo con las recomendaciones del fabricante. Se sintetizó cDNA a partir del RNA total utilizando el kit de síntesis de cDNA de la Primera Cadena (Boehringer Mannheim) siguiendo las recomendaciones del fabricante. La amplificación por PCR de un fragmento específico de gen del cDNA sintetizado se realizó utilizando DNA polimerasa Taq (Roche) y los iniciadores específicos del gen (véase Tabla 13 para los iniciadores) en la reacción siguiente: tampón PCR 1X, MgCl<sub>2</sub> 1,5 mM, 0,2  $\mu$ M de cada iniciador, 0,2  $\mu$ M dNTPs, 1 unidad de polimerasa, y 5  $\mu$ l de cDNA de la reacción de síntesis. La amplificación se realizó en las condiciones siguientes: desnaturalización, 95°C, 1 minuto; reasociación, 62°C, 30 segundos; extensión, 72°C, 1 minuto, 35 ciclos; extensión, 72°C, 5 minutos; mantenimiento, 4°C, en todos los casos. Los productos PCR se corrieron en un gel de agarosa al 1%, se tiñeron con bromuro de etidio, y se visualizaron bajo luz UV utilizando el sistema de documentación de gel Quantity-One (Bio-Rad).

30 La expresión de los transgenes se detectó en la línea transgénica T1. Estos resultados indicaban que los transgenes se expresan en las líneas transgénicas y sugerían fuertemente que su producto génico mejoraba la tolerancia de la planta al estrés en las líneas transgénicas. De acuerdo con la exposición anterior, no pudo detectarse por este método expresión alguna de genes endógenos idénticos o muy similares. Estos resultados están de acuerdo con los datos del Ejemplo 7.

TABLA 7

*Iniciadores utilizados para la amplificación del mRNA del transgén respectivo en PCR utilizando RNA aislado de plantas transgénicas de Arabidopsis thaliana como molde*

Gen	Iniciador 5'	Iniciador 3'
<b>DBF-1</b>	<b>RC876 (SEQ ID NO:56)</b> 5'GGAGACGGTATCACACCAT CGAAGA3'	<b>RC877 (SEQ ID NO:57)</b> 5'TGCACAGACATCTGCCT GGCTCACA3'
<b>CABF-2</b>	<b>RC974 (SEQ ID NO:58)</b> 5'GATGATCGCAGCCGAAGCT	<b>RC975 (SEQ ID NO:60)</b> 5'GGCAGTCTGTGGAGGC

5		CCAGTG3' RC976: (SEQ ID NO:59) 5'GGGTGTGCCATGGACTGGT GTTCCAG3'	TGATACATCA3' RC977: (SEQ ID NO:61) 5'CCTGATCCTGTGACCCC TTTTGCCA3'
10	LZ-1	RC978 (SEQ ID NO:62) 5'GACATGGACGGTGATGCGA AGTTGG3' RC980: (SEQ ID NO:63) 5'CGGCAACAGCAGGGTCTAT ACCTTGG3'	RC979 (SEQ ID NO:64) 5'GCATACTCCAGGTCAA ATGCAGCAGC3' RC981: (SEQ ID NO:65) 5'GGGTCGGCAGCCTCCA ATCCATACA3'
15			
20	CBF-1	RC880 (SEQ ID NO:66) 5'GGCAGGGAATCTACGCATC GCTTTG3'	RC881 (SEQ ID NO:67) 5'CGACGAGATTCTCTGC AACATCTGAG3'
25			
30	HDZ-1	RC982 (SEQ ID NO:68) 5'GGAGCTTGGACTGCGACCTC GTCAAG3' RC984: (SEQ ID NO:70) 5'GTCATCGAGGAATCGCACA ACTCCT3'	RC983 (SEQ ID NO:69) 5'GGTGTGGCTCGTGCGA GGGCTATCAG3' RC985: (SEQ ID NO:71) 5'GGTTGACGTTGGATTGC ACATGGTGG3'
35			
40	ZF-1	RC874 (SEQ ID NO:72) 5'TGGATGTGCGAAGTGTGCG AGGTTG3'	RC875 (SEQ ID NO:73) 5'GCGCTGCCTCTGATAAT AGAGTTGG3'
45			
50	CABF-1	RC938 (SEQ ID NO:74) 5'GTGCAGGAGTGCATCCG AGTTCATC3'	RC939 (SEQ ID NO:75) 5'CGTACGGCTGTTGCATC ATCTGCATCG3'

## Ejemplo 10

55 *Elaboración por ingeniería genética de plantas de soja tolerantes al estrés por sobre-expresión de los genes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2*

Se utilizaron los constructos pBPSLVM111, pBPSLVM149, pBPSLVM157, pBPSLVM39, pBPSLVM12, pBPSLVM19, y pBPSLVM69 para transformar soja como se describe a continuación.

60 Se esterilizaron en superficie semillas de soja con etanol al 70% durante 4 minutos a la temperatura ambiente con agitación continua mediante sacudidas, seguido por 20% (v/v) de Clorox complementado con 0,05% (v/v) de Tween durante 20 minutos con agitación continua mediante sacudidas. Después de ello, se lavaron las semillas 4 veces con agua destilada y se pusieron sobre papel de filtro estéril humedecido en una cápsula Petri a la temperatura ambiente durante 6 a 39 horas. Se pelaron las cubiertas de las semillas, y se desprendieron los cotiledones del eje del embrión. El  
65 eje del embrión se examinó para asegurarse de que no se había deteriorado la región meristemática. Los ejes cortados del embrión se recogieron en una cápsula Petri estéril semi-abierta y se secaron al aire hasta un contenido de humedad menor que 20% (peso fresco) en una cápsula Petri sellada hasta su uso ulterior.

## ES 2 279 777 T3

Se preparó un cultivo de *Agrobacterium tumefaciens* a partir de una sola colonia en medio LB sólido más los antibióticos apropiados (v.g. 100 mg/l de estreptomina, 50 mg/l de kanamicina) seguido por crecimiento de la colonia simple en medio LB líquido hasta una densidad óptica a 600 nm de 0,8. A continuación, el cultivo de bacterias se redujo a un sedimento a 7000 rpm durante 7 minutos a la temperatura ambiente, y se resuspendió en medio MS (Murashige y Skoog, 1962) complementado con acetosiringona 100  $\mu\text{M}$ . Los cultivos bacterianos se incubaron en este medio de pre-inducción durante 2 horas a la temperatura ambiente antes de su utilización. Los ejes de los embriones de las semillas cigóticas de soja a aproximadamente 15% de contenido de humedad se embebieron durante 2 horas a la temperatura ambiente con el cultivo de suspensión de *Agrobacterium* pre-inducido. Los embriones se retiraron del cultivo de imbibición y se transfirieron a cápsulas Petri que contenían medio MS sólido complementado con 2% de sacarosa y se incubaron durante 2 días en la oscuridad, a la temperatura ambiente. Alternativamente, los embriones se pusieron sobre papel de filtro estéril húmedo (medio MS líquido) en una cápsula Petri y se incubaron en las mismas condiciones arriba descritas. Después de este periodo, los embriones se transfirieron a medio MS sólido o líquido complementado con 50 mg/l de carbenicilina o 300 mg/l de cefotaximo para destruir las agrobacterias. Se utilizó el medio líquido para humedecer el papel de filtro estéril. Los embriones se incubaron durante 4 semanas a 25°C, bajo 150  $\mu\text{mol m}^{-2}\text{s}^{-1}$  y fotoperiodo de 12 horas. Una vez que las plantas de semillero produjeron raíces, se transfirieron a suelo Metromix estéril. El medio de las plantas *in vitro* se eliminó por lavado antes de transferir las plantas al suelo. Las plantas se mantuvieron bajo una cubierta de plástico durante 1 semana para favorecer el proceso de aclimatación. Se transfirieron luego las plantas a una sala de crecimiento en la cual se incubaron a 25°C, bajo 150  $\mu\text{mol m}^{-2}\text{s}^{-1}$  de intensidad lumínica y fotoperiodo de 12 horas durante aproximadamente 80 días.

Las plantas transgénicas se escrutaron luego en cuanto a su tolerancia mejorada a la sequía, la sal y/o el frío, de acuerdo con el método de escrutinio descrito en el Ejemplo 7, demostrando que la expresión de los transgenes confiere tolerancia al estrés.

### Ejemplo 11

*Producción por ingeniería genética de plantas de Colza/Canola tolerantes al estrés por sobre-expresión de los genes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2*

Se utilizaron los constructos pBPSLVM111, pBPSLVM149, pBPSLVM157, pBPSLVM39, pBPSLVM12, pBPSLVM19, y pBPSLVM69 para transformar colza/canola como se describe a continuación.

El método de transformación de plantas descrito en esta memoria es también aplicable a *Brassica* y otras cosechas. Semillas de canola se esterilizan superficialmente con etanol al 70% durante 4 minutos a la temperatura ambiente con agitación continua mediante sacudidas, seguido por 20% (v/v) de Clorox complementado con 0,05% (v/v) de Tween durante 20 minutos, a la temperatura ambiente con agitación continua mediante sacudidas. A continuación, se lavan 4 veces las semillas con agua destilada y se dejan sobre papel de filtro estéril húmedo en una cápsula Petri a la temperatura ambiente durante 18 horas. Se apartan luego las cubiertas de las semillas y las semillas se secan al aire durante una noche en una cápsula Petri estéril semiabierta. Durante este periodo, las semillas pierden aproximadamente un 85% de su contenido de agua. Las semillas se guardan a la temperatura ambiente en una cápsula Petri herméticamente cerrada hasta su uso ulterior. Los constructos de DNA 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-DNA. Estos resultados se confirman por hibridación Southern en la cual el DNA se somete a electroforesis en un gel de agarosa al 1% y se transfiere a una membrana de nailon cargada positivamente (Roche Diagnostics). Se utiliza el Kit de Síntesis de Sonda PCR DIG (Roche Diagnostics) para preparar una sonda marcada con digoxigenina por PCR, y se utiliza de acuerdo con las recomendaciones del fabricante.

Las plantas transgénicas se escrutan luego respecto a su tolerancia mejorada al estrés de acuerdo con el método de escrutinio descrito en el Ejemplo 7, demostrando que la expresión del transgén confiere tolerancia a la sequía.

### Ejemplo 12

*Producción por ingeniería genética de plantas de maíz tolerantes al estrés por sobre-expresión de los genes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2*

Se utilizaron los constructos pBPSLVM111, pBPSLVM149, pBPSLVM157, pBPSLVM39, pBPSLVM12, pBPSLVM19, y pBPSLVM69 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. Los embriones inmaduros se co-cultivan con *Agrobacterium tumefaciens* que llevan vectores "super-binarios", y las plantas transgénicas se recuperan por organogénesis. Este procedimiento proporciona una eficiencia de transformación comprendida entre 2,5% y 20%. Las plantas transgénicas se escrutan luego respecto a su tolerancia mejorada a la sequía, la sal y/o el frío de acuerdo con el método de escrutinio descrito en el Ejemplo 7, demostrando que la expresión del transgén confiere tolerancia al estrés.

## ES 2 279 777 T3

### Ejemplo 13

*Producción por ingeniería genética de plantas de trigo tolerantes al estrés por sobre-expresión de los genes CABF-1, DBF-1, CBF-1, HDZ-1, ZF-1, LZ-1, y CABF-2*

5 Se utilizaron los constructos pBPSLVM111, pBPSLVM149, pBPSLVM157, pBPSLVM39, pBPSLVM12, pBPSLVM19, y pBPSLVM69 para transformar trigo como se describe a continuación.

10 La transformación del trigo se realiza con el método descrito por Ishida *et al.* 1996 Nature Biotech. 14745-50. Se co-cultivan embriones inmaduros con *Agrobacterium tumefaciens* que llevan vectores “super-binarios”, y se recuperan plantas transgénicas por organogénesis. Este procedimiento proporciona una eficiencia de transformación entre 2,5% y 20%. Las plantas transgénicas se escrutan luego respecto a su tolerancia mejorada al estrés de acuerdo con el método de escrutinio descrito en el Ejemplo 7, demostrando que la expresión de los transgenes confiere tolerancia a la sequía.

### 15 Ejemplo 14

*Monitorización de los cambios en la concentración de mRNA de CABF-1, CABF-2 y CBF-1 en cultivos de Physcomitrella patens tratados en frío*

#### 20 Preparación en portaobjetos de un Microsistema de DNA

Se realizó la amplificación por PCR en placas de 96 pocillos a partir de ESTs seleccionados de *Physcomitrella patens* clonados en el vector pBluescript. Se empleó el juego de tampones PCR (Boehringer Mannheim) para la reacción PCR. Cada mezcla de reacción PCR contiene 10  $\mu$ l de tampón PCR sin MgCl<sub>2</sub>, 10  $\mu$ l de MgSO<sub>4</sub>, 3  $\mu$ l del iniciador directo SK (MWG-Biotech, secuencia: 5'-CGCCAAGCGCGCAATTAACCCCTCACT-3', SEQ ID NO: 76), 3  $\mu$ l del iniciador inverso SK (MWG-Biotech, secuencia: 5-GCGTAATACGACTCACTATAGGG CGA-3', SEQ ID NO: 77), 2  $\mu$ l de dNTP, 1  $\mu$ l de DNA-polimerasa Taq (Roche), 72  $\mu$ l de agua y 1  $\mu$ l de molde de DNA. Después de desnaturalización a 95°C durante 3 minutos, las reacciones PCR se realizaron con 35 ciclos de 3 pasos consecutivos que incluían desnaturalización a 95°C durante 45 segundos, reasociación a 63°C durante 45 segundos, y elongación a 72°C durante 60 segundos. La última elongación se realizó a 72°C durante 10 minutos. Los productos PCR se purificaron luego con Kit de Purificación de PCR QIAquick (Qiagen, Inc.), se eluyeron con agua y se midió la concentración de DNA a 260 nm en un espectrofotómetro.

35 Se secaron 2 a 5  $\mu$ g de cada producto PCR y se disolvieron en 50  $\mu$ l de DMSO. Los productos PCR se formatearon luego desde placas de 96 pocillos a placas de 384 pocillos para la impresión. Se empleó el clasificador Micro-array GenIII (Molecular Dynamics) para imprimir los productos PCR en portaobjetos Microarray (Molecular Dynamics) con el formato recomendado por el fabricante. Los puntos impresos tenían un diámetro de aproximadamente 290  $\mu$ m y estaban espaciados aproximadamente 320  $\mu$ m de centro a centro. Después de la impresión, el portaobjetos se dejó en la cámara exenta de polvo durante 1 hora para secarlo. La reticulación UV se realizó con 600  $\mu$ J/mm. Los portaobjetos reticulados estaban listos para hibridación y se guardaron en cámaras oscuras y secas.

#### Síntesis de la sonda Microarray

45 Se extrajo el RNA total de cultivos de *Physcomitrella patens* tratados en frío (12 horas a 4°C en la oscuridad) siguiendo el método de extracción de RNA descrito en Ausubel *et al.* (1987 Curr. Prot. in Mol. Biol. J. Wiley and Sons, Nueva York). Se aplicó el kit Oligotex mRNA midi (Qiagen Inc.) para aislar el mRNA del RNA total con un método que combinaba a la vez un protocolo de lotes y el protocolo estándar recomendado por el fabricante. Después de fijar el RNA total con Oligotex, la muestra se centrifugó a 14000 x g para separar la mezcla Oligotex:mRNA con la fase líquida en lugar de pasarla a través de una columna. Después de 4 lavados con tampón OW2 como se describe en el protocolo de lotes, el Oligotex:mRNA se resuspendió en 400  $\mu$ l de OW2 y se recogió luego en la columna como en el protocolo estándar. El mRNA se eluyó siguiendo el protocolo estándar.

55 Se sintetizaron sondas de cDNA marcadas con Cy3 y Cy5 a partir de mRNA con el sistema Superscript Choice para síntesis de cDNA (Gibco BRL). Se mezclaron el iniciador oligo-(dT)<sub>25</sub> (Genosys Biotechnologies) y el iniciador Nonamer (Amersham Pharmacia Biotech) con mRNA para alcanzar un volumen total de 20  $\mu$ l. La mezcla se calentó primeramente a 70°C durante 10 minutos y se dejó luego a la temperatura ambiente durante 15 minutos antes de transferirla a hielo. Una vez que la muestra está introducida en el hielo, se añaden 8  $\mu$ l de Tampón de Síntesis de la Primera Cadena, 4  $\mu$ l de DTT 0,1M, 2  $\mu$ l de dNTP (Amersham Pharmacia Biotech), 2  $\mu$ l de Cy3- o Cy5-dCTP (Amersham Pharmacia Biotech), 2  $\mu$ l de Inhibidor de RNAsa (Gibco BRL) y 2  $\mu$ l de Transcriptasa Inversa SuperScript II. La síntesis de la primera cadena se realizó a 42°C durante 8 horas y la mezcla se calentó luego a 94°C durante 3 minutos después de la reacción.

65 Después de la síntesis de la primera cadena, se añadieron 4  $\mu$ l de hidróxido de sodio 2,5M a la mezcla de reacción y se incubó la mezcla a 37°C durante 10 minutos. Se añadieron luego a cada reacción 20  $\mu$ l de MOPS 2M (pH 5,0) y 500  $\mu$ l de tampón PB (Qiagen Inc.). Se purificó luego la sonda por medio del Kit de Purificación QIAquick PCR (Qiagen Inc.) con el protocolo proporcionado por el fabricante.

## ES 2 279 777 T3

### *Hibridación y lavados de cDNA Microarray*

Las sondas purificadas marcadas con Cy3 y Cy5 se mezclaron y se secaron a vacío para dar un volumen final de 9  $\mu$ l. Se añadieron luego a las sondas de cDNA 9  $\mu$ l de Solución de Hibridación Microarray (Amersham Pharmacia Biotech) y 18  $\mu$ l de formamida (Sigma) para dar un volumen final de 36  $\mu$ l. Se aplicó la muestra al portaobjetos Microarray impreso que se cubrió luego con un cubreobjetos limpio y exento de polvo sin cantidad alguna de aire atrapado. La hibridación se realizó 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, 0,2% de SDS a la temperatura ambiente durante 5 minutos y 15 minutos. Se realizaron 2 lavados en condiciones severas con 0,25 X SSC, 0,1 SDS a 55°C durante 10 y 30 minutos respectivamente. Después de los lavados, los portaobjetos se enjuagaron brevemente con agua Millipore y se secaron bajo nitrógeno comprimido.

### *Escaneo y análisis de los datos Microarray*

Los Microarrays de cDNA se escanearon utilizando el escáner Microarray GenIII ((Molecular Dynamics) equipado con dos canales láser. Las imágenes escaneadas se visionaron primeramente y se ajustaron en el soporte lógico Image-Quant (Molecular Dynamics) y se analizaron luego por medio del soporte lógico ArrayVision (Molecular Dynamics). La intensidad de señal para cada punto se extrajo por medio de soporte lógico ArrayVision (Molecular Dynamics) y se transfirió a Excel (Microsoft). Los datos obtenidos se normalizaron dividiendo la diferencia entre el valor de intensidad y el ruido de fondo y la diferencia entre el valor de control y el ruido de fondo. La relación se obtuvo luego por división de los datos normalizados.

### Ejemplo 15

#### 25 *Identificación de Genes Homólogos y Heterólogos*

Pueden utilizarse secuencias génicas para identificar genes homólogos o heterólogos a partir de cDNA o genotecas genómicas. Los genes homólogos (v.g. clones de cDNA de longitud total) pueden aislarse por hibridación de ácido nucleico utilizando por ejemplo genotecas de cDNA. Dependiendo de la abundancia del gen de interés, se extienden en placas 100.000 hasta 1.000.000 de bacteriófagos recombinantes y se transfieren a membranas de nailon. Después de desnaturalización con álcali, se inmoviliza el DNA en la membrana mediante, v.g., reticulación por UV. La hibridación se lleva a cabo en condiciones de severidad alta. En solución acuosa, la hibridación y el lavado se realizan con una fuerza iónica de NaCl 1M y una temperatura de 68°C. Se generan sondas de hibridación mediante v.g. marcación radiactiva (<sup>32</sup>P) por transcripción de la mella (High Prime, Roche, Mannheim, Alemania). Las señales se detectan por autorradiografía.

Los genes parcialmente homólogos o heterólogos que son afines pero no idénticos pueden identificarse de una manera análoga al procedimiento arriba descrito utilizando condiciones de hibridación y lavado de baja severidad. Para la hibridación acuosa, la fuerza iónica se mantiene normalmente a NaCl 1M mientras que la temperatura se reduce progresivamente desde 68 a 42°C.

El aislamiento de las secuencias génicas con homologías (o identidad/ semejanza de secuencia) solamente en un dominio diferenciado de (por ejemplo 10-20 aminoácidos) puede llevarse a cabo por utilización de sondas de oligonucleótidos sintéticos radio-marcadas. Los oligonucleótidos radio-marcados se preparan por fosforilación del extremo 5-prima de dos oligonucleótidos complementarios con polinucleótido-quinasa T4. Los oligonucleótidos complementarios se reasocian y se ligan para formar concatémoros. Los concatémoros bicatenarios se someten luego a radiomarcación mediante, por ejemplo, transcripción de la mella. La hibridación se lleva a cabo normalmente en condiciones de severidad baja utilizando concentraciones altas de oligonucleótidos.

50 Solución de hibridación de oligonucleótidos:

6 X SSC

fosfato de sodio 0,01 M

55 EDTA 1 mM (pH 8)

0,5% SDS

60 100  $\mu$ g/ml de DNA desnaturalizado de esperma de salmón

0,1% de leche desnatada seca

65 Durante la hibridación, la temperatura se reduce gradualmente hasta 5°C por debajo del valor T<sub>m</sub> estimado de los oligonucleótidos o hasta la temperatura ambiente seguido por pasos de lavado y autorradiografía. El lavado se realiza con severidad baja tal como tres pasos de lavado utilizando 4 X SSC. Detalles adicionales han sido 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

*Identificación de Genes Homólogos por Escrutinio de Genotecas de Expresión con Anticuerpos*

5 Pueden utilizarse clones de cDNA para producir proteína recombinante, por ejemplo en *E. coli* (v.g. sistema QIAexpress pQE de Qiagen). Las proteínas recombinantes se purifican luego normalmente por afinidad mediante cromatografía de afinidad Ni-NTA (Qiagen). Las proteínas recombinantes se utilizan luego 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.*  
 10 1994 *BioTechniques* 17:257-262. El anticuerpo puede utilizarse luego para escrutar genotecas de expresión de cDNA a fin de identificar genes homólogos o heterólogos por escrutinio inmunológico (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.

## 15 Ejemplo 17

*Mutagénesis in Vivo*

20 La mutagénesis *in vivo* de microorganismos puede realizarse mediante pasadas de DNA plasmídico (u otro vector) a través de *E. coli* u otros microorganismos (v.g. *Bacillus* spp o levaduras tales como *Saccharomyces cerevisiae*) que están deteriorados en sus capacidades para mantener la integridad de su información genética. Las cepas mutantes típicas tienen mutaciones en los genes para el sistema de reparación de DNA (v.g., mutHLS, mutD, mutT, etc.; para referencia, véase Rupp W.D. (1996) DNA repair mechanisms, en: *Escherichia coli* and *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 DNA mutadas en plantas se realiza 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

30

*Análisis in Vitro de la Función de los Genes de Physcomitrella en Organismos Transgénicos*

La determinación de las actividades y los parámetros cinéticos de las enzimas está bien establecida en la técnica. Los experimentos para determinar la actividad de cualquier enzima alterada dada tienen que adaptarse a la actividad específica de la enzima de tipo salvaje, lo cual está perfectamente dentro de la capacidad de un experto en la técnica. 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 enzimas pueden encontrarse, por ejemplo, en las referencias siguientes: Dixon, M., and Webb, E.C., (1979) *Enzymes*. Longmans: Londres; 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., GraBl, M., eds. (1983-1986) *Methods of Enzymatic Analysis*, 3rd ed., vol. I-XII, Verlag Chemie: Weinheim; and Ullmann's *Encyclopedia of Industrial Chemistry* (1987) vol. A9, *Enzymes*. VCH: Weinheim, p. 352-363.

45

La actividad de las proteínas que se fijan a DNA puede medirse por varios métodos bien establecidos, tales como ensayos de desplazamiento de bandas de DNA (denominados también ensayos de retardación en gel). El efecto de dichas proteínas sobre la expresión de otras moléculas puede medirse utilizando ensayos de genes informadores (tales como el descrito Kolmar, H. *et al.* (1995) *EMBO J.* 14: 3895-3904 y las referencias citadas en dicho lugar). Los sistemas de ensayo de genes informadores son bien conocidos y están bien establecidos para aplicaciones en células tanto pro- como eucariotas, utilizando enzimas tales como  $\beta$ -galactosidasa, proteína fluorescente verde, y varias otras.

50

La determinación de la actividad de las proteínas 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*, pp. 85-137, 199-234 y 270-322, Springer, Heidelberg (1989).

55

## Ejemplo 19

*Purificación del Producto Deseado a Partir de Organismos Transformados*

60

La recuperación del producto deseado a partir de material de plantas (a saber, *Physcomitrella patens* o *Arabidopsis thaliana*), hongos, algas, ciliados, células de *C. glutamicum*, u otras células bacterianas transformadas con las secuencias de ácido nucleico descritas en esta memoria, o el sobrenadante de los cultivos arriba descritos puede realizarse por diversos métodos bien conocidos en la técnica. Si el producto deseado no es secretado por las células, puede recogerse a partir del cultivo por centrifugación a baja velocidad, después de lo cual pueden lisarse las células por técnicas estándar, tales como fuerza mecánica o sonificación. Los órganos de las plantas pueden separarse mecánicamente de otros tejidos u órganos. Después de la homogeneización, los residuos celulares se separan por centrifugación, y la fracción sobrenadante que contiene las proteínas solubles se retiene para purificación ulterior del compuesto deseado.

65

## ES 2 279 777 T3

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 ulterior.

5 La fracción sobrenadante de cualquier método de purificación se somete a cromatografía con una resina adecuada, en la cual la molécula deseada es retenida 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 pasos de cromatografía pueden repetirse en caso necesario, utilizando la misma o diferentes resinas cromatográficas. Un experto en la técnica estaría plenamente versado en la selección de resinas cromatográficas apropiadas y en su aplicación más eficaz para una molécula particular a purificar. El producto purificado puede concentrarse por  
10 filtración o ultrafiltración, y guardarse a una temperatura a la que se maximiza la estabilidad del producto.

Existe una amplia serie de métodos de purificación conocidos en la técnica y el método de purificación que antecede no tiene por objeto ser limitante. Dichas técnicas de purificación se describen, por ejemplo, en Bailey J.E. & Ollis, D.F. *Biochemical Engineering Fundamentals*, McGraw-Hill: Nueva York (1986). Adicionalmente, la identidad y pureza de  
15 los compuestos aislados puede evaluarse por métodos estándar de la técnica. Éstos incluyen cromatografía líquida de alta resolución (HPLC), métodos espectroscópicos, métodos de tinción, 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,  
20 p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A. *et al.* (1987) *Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

- 5 1. Una planta transgénica transformada por un ácido nucleico codificante de una proteína relacionada con el estrés por factores de transcripción (TFSRP) en donde la TFSRP es un Factor de Fijación-1 CRT/DRE (CBF-1) como se define en SEQ ID NO: 17, y homólogos del mismo que tienen al menos 50%-60% de homología con la proteína que se define en SEQ ID NO: 17 y en donde la expresión de la secuencia de ácido nucleico en la planta da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta.
- 10 2. La planta transgénica de la reivindicación 1, en donde el ácido nucleico codificante de TFSRP es un Factor de Fijación-1 CRT-DRE (CBF-1) como se define en SEQ ID NO: 10 y homólogos del mismo que tienen al menos 50%-60% de homología con el ácido nucleico que se define en SEQ ID NO: 10.
- 15 3. La planta transgénica de cualquiera de las reivindicaciones 1 ó 2, en donde el ácido nucleico y la proteína son de una *Physcomitrella patens*.
4. La planta transgénica de cualquiera de las reivindicaciones 1-3, en donde la planta es una monocotiledónea.
- 20 5. La planta transgénica de cualquiera de las reivindicaciones 1-3, en donde la planta es una dicotiledónea.
6. La planta transgénica de cualquiera de las reivindicaciones 1-3, en donde la planta se selecciona de maíz, trigo, centeno, avena, trigo híbrido, arroz, cebada, soja, cacahuete, algodón, colza, canola, mandioca, pimiento, girasol, tagetes, plantas solanáceas, patata, tabaco, berenjena, tomate, especias de Vicia, guisante, alfalfa, café, cacao, té, especies de *Salix*, palma de aceite, cocotero, céspedes perennes y cosechas forrajeras.
- 25 7. Una semilla producida por una planta transgénica que expresa una proteína relacionada con el estrés por factores de transcripción (TFSRP) como se describe en las reivindicaciones 1 ó 3, en donde la semilla contiene dicha TFSRP y en donde la planta es progenie auténtica para tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta.
- 30 8. Una semilla producida por una planta transgénica transformada por un ácido nucleico codificante de una proteína relacionada con el estrés por factores de transcripción (TFSRP) como se describe en las reivindicaciones 1 a 3, en la cual la semilla contiene dicho ácido codificante de TFSRP de las reivindicaciones 2 ó 3 y en donde la planta es progenie auténtica para tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta.
- 35 9. Uso de una planta o semilla de cualquiera de las reivindicaciones 1-8 para la producción de un producto agrícola.
- 40 10. Una proteína relacionada con el estrés por factores de transcripción (TFSRP), en donde la TFSRP es como se describe en cualquiera de las reivindicaciones 1 ó 3.
11. Un ácido nucleico codificante de una proteína relacionada con el estrés por factores de transcripción (TFSRP), en donde el ácido nucleico codificante de TFSRP codifica una TFSRP como se describe en las reivindicaciones 1 ó 3.
- 45 12. Un ácido nucleico codificante de una proteína relacionada con el estrés por factores de transcripción (TFSRP) como se describe en las reivindicaciones 2 ó 3 que codifica una proteína relacionada con el estrés por factores de transcripción (TFSRP).
- 50 13. Un vector de expresión recombinante aislado que comprende un ácido nucleico de cualquiera de las reivindicaciones 11 y 12, en donde la expresión del vector en una célula hospedadora da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la célula hospedadora.
14. Una célula hospedadora que contiene el vector de la reivindicación 13.
- 55 15. La célula hospedadora de la reivindicación 14, en donde la célula pertenece a una planta.
16. Un método de producción de una planta transgénica con un ácido nucleico codificante de una proteína relacionada con el estrés por factores de transcripción (TFSRP), en donde la expresión del ácido nucleico en la planta da como resultado una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta que comprende:
- 60 (a) transformar una célula de planta con un vector de expresión que comprende el ácido nucleico de TFSRP, y
- 65 (b) generar a partir de la célula de la planta una planta transgénica con una tolerancia incrementada al estrés por sequía en comparación con una variedad de tipo salvaje de la planta, en donde la TFSRP es como se describe en las reivindicaciones 1 ó 3.

## ES 2 279 777 T3

17. El método de la reivindicación 16, en el cual el ácido nucleico codificante de TFSRP es como se describe en las reivindicaciones 2 ó 3.

5 18. Un método de aumentar la expresión de un gen de interés en la célula hospedadora en comparación con la variedad de tipo salvaje de la célula hospedadora, en el cual el gen de interés se transcribe en respuesta a una proteína relacionada con el estrés por factores de transcripción (TFSRP), que comprende

10 (a) transformar la célula hospedadora con un vector de expresión que comprende un ácido nucleico codificante de TFSRP, y

(b) expresar la TFSRP en la célula hospedadora, incrementando con ello la expresión del gen transcrito en respuesta a la TFSRP en comparación con la variedad de tipo salvaje de la célula hospedadora, en donde la TFSRP es como se describe en las reivindicaciones 1 ó 3.

15 19. El método de la reivindicación 18, en el cual el ácido nucleico codificante de TFSRP es como se describe en las reivindicaciones 2 ó 3.

20 20. Un método de identificación de una proteína nueva relacionada con el estrés por factores de transcripción (TFSRP) que comprende

(a) generar una respuesta específica de anticuerpos para una TFSRP como se describe en las reivindicaciones 1 ó 3;

25 (b) escrutar material supuesto de TFSRP con el anticuerpo, en donde la fijación específica del anticuerpo al material indica la presencia de una TFSRP potencialmente nueva y

(c) identificar el material fijado como una nueva TFSRP.

30 21. Un método de modificación de la tolerancia al estrés por sequía en una planta transgénica por modificación de la expresión de una secuencia de ácido nucleico que codifica una proteína relacionada con el estrés por factores de transcripción (TFSRP) como se describe en las reivindicaciones 1 ó 3.

22. El método de la reivindicación 21, en el cual la tolerancia a la sequía se incrementa.

35 23. El método de la reivindicación 21, en el cual la tolerancia a la sequía se reduce.

24. El método de la reivindicación 21, en el cual la planta se transforma con ácido nucleico codificante de TFSRP como se describe en las reivindicaciones 2 ó 3.

40 25. El método de la reivindicación 21, en el cual la planta se transforma con un promotor que dirige la expresión de TFSRP nativa.

26. El método de la reivindicación 25, en el cual el promotor es específico de tejido.

45 27. El método de la reivindicación 25, en el cual el promotor está regulado por el desarrollo.

28. El método de la reivindicación 21, en el cual la expresión de TFSRP se modifica por administración de una molécula anti-sentido que inhibe la expresión de TFSRP.

50

55

60

65

**FIGURA 1A**

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

GCACCAGCGAATCCGTCTCCGCCTCCGCCTTCTGCACGCGTGGTTGTGGTTCGACCTC  
TCGCCGGAGCAACAGGAACTAATCCCTTTTCCAGCACTAAACGATTGAAGCAATTT  
TTTTTTCTTGTGAACTGCTCACTCTCTCTCTGTTATGAGGGGATTCGAAGCTTGAAA  
GTTATGAGCTGAAGGTTGAGGACACGTAAGCACCAGAGGACGATCATACTACAATT  
AACCTTGCGGGGAAAAGCCCAGGCAAAATAGGACGGATGGCCGACAGTTACGGCC  
ACAACGCAGGTTACCCGAGAGCAGCCCGCATTCTGATAACGAGTCCGGCGGCCAT  
TACCGTGATCAGGACGCTTCTGTACGGGAGCAAGACCGGTTTTTGGCCATCGCAAAT  
GTGAGCCGAATCATGAAGAAAGCATTGCCATCTAATGCGAAGATATCGAAAGACGC  
CAAAGAGACTGTGCAGGAGTGCGTATCCGAGTTCATCAGTTTCATTACTGGTGAGGC  
GTCCGACAAGTGTCAGAGGGAAAAGAGGAAGACGATCAACGGGGATGACTTGCTGT  
GGGCCATGAGTACTCTTGGTTTTGAAGATTATGTGGAACCTCTGAAGGTGTACCTAC  
ACAAGTATCGTGAAGTGGAGGGGGAGAAGGCCTCTATGGCCAAGGGTGGTGATCAA  
GCAGGG

**FIGURA 1B**

Secuencia de nucleótidos del DBF-1 parcial de *Physcomitrella patens* (SEQ ID NO:2)

GGCACGAGGCTTGATGATGATCATGCACTAGCTTCTGCAAAGTGCCAGGCTTTAGCA  
CGTCTACTTCCCAAGTTACAGCAAGGTGGCCATCGCACATTGATATTCAGCCAGTGG  
ACAAGCATGCTGGATATTTTAGAATGGGCTCTTGACGTCATGGGTTTTTCTTACTC  
GCCTAGATGGAAGCACTCAAGTAAGTGAACGCCAAACCCTAGTGGACGAGTTCAAC  
AATGACCCTAGCATATTTGTGTTTCTCCTGTCTACTCGAGCTGGAGGTCAAGGTCTA  
AATTTAACAGGAGCAGACACAGTCATTTTACATGATTTGGACTTCAATCCCCAAATG  
GATCGACAGGCTGAGGATCGCTGTCATCGGATTGGCCAGTCTAAACCTGTTACGATA

TACCGACTTGTAACAAAAGATACGGTCGATGAAAGTATATACAAGATAGCCCAACA  
GAAGCTGGTCCTCGATGCGGCAGTTCTTGAANGAAAAGAGTCATCCTCTGATCTTAA  
TGATG

**FIGURA 1C**

Secuencia de nucleótidos del CBF-1 parcial de *Physcomitrella patens* (SEQ ID NO:3)

GGCACGAGCTGATACTAATTGCACGAGGTTTTCTCAAATGTGTTTTTGGGTACAGGA  
AGGTGGAGGGGAATTTGTAAGTGACAGAGACGTGGTGGGAGTGGGAGGAGTGTGA  
GGAATCGAGCTAGCACCTAAAGGATTTCTGGGGTGAAGGAAGGTGCGATTGAAGGCG  
TGCATGAAATTTTGACGCAGCGGGTTGAATCGGAAGGAGTTTTCAGCAGCTGGAAA  
GTACCTTCGAGGGTTGATTCATCCAAAGTTCCATGTGATATGGCTTCAAAGTATCC  
GCGAAAATGTAGAGAGCACGCATCACCTGGAGTTGGTGGCAGGGAATCTACGCATC  
GCTTTGATTCAAGGTCGCAGTCGTATTCGTTCTCGGAGAAACCTTACCACCGTCGTC  
GCCGGGATGCGTTTCGTGATATGATCAGCGATTTGGTGCATCAGCCTTCCGACACTG  
CCGTGCCTGGTTTTCAGGGGAGTGCCTATCGTCAGAACTGAACAAGTACGTGACA  
GAGATTCGGCCCACGAGGTGCTCGAAGAAAATTTGGCTTGGGACGTATGACACTGC  
AGAGGAAGCAGCGCGTGCCTTTGACATCGGAAATTTGTGTTGTAAGAAAACCTGC  
CGCTCAACTTTCCGGATTCGACTCAGATGTTGCAGAGAATCTCGTCGAAATTGACCC  
CCGAAGCGCAACGAAAAGCCATCGCGACGCTGGCGAAGGACGTAGTGCGAATGGA  
AAATGACAGGTCGAAGTTGGGTGGCGGTAACCTGACTACCACAGAGCCCCCGGTCC  
ATTCCGAGCCTATTACTCAACACCTTGCAGCAGCTGAGATTCGCGCGGTCACGTACA  
TTGAACAGCCCCTGGAAATTGTCTACGGAGTGGAAGAATCGGCGACGGCCATGTGCG  
GTAACGGAAGCAAATGCGCGCGATAACCACTCTTGGAGTTGGGACTTGGGCAAAGT  
GATCCTTGATGACGAGCTCTCTGAAATTCCTAACTTCGTCGGAG

FIGURA 1D

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

GGCACGAGGGCAAGAGGGGATAGAGACTTGAAAGGAAAGGGAGGGAAGGGTGTAA  
GGAGGCCACGGGCAGGGTCAAGGTGTCCAATGCACCTGCAAGATCAGGAAGCTTG  
AAGTAGATCAGGGAAAAACGATGGTAGTCCCTAGTTTACCCGCCTTCGGAGGACA  
GAACGCCATGCTCAGACGCAACATTGACAACAACACCGACACTCTGATTTCTCTGCT  
TCAAGGGTCTGCTCCCCTCGCGTGAGCATGCAACAAGGATATGTTGCAGTGCCGCG  
TTCATCGGAGAGTCTCGAAAACATGATGGGGGCTTGTGGGCAAAAACCTGCCTTACTT  
TTCGTCAATTTGATGGGCCGAGTGTAGAAGAGCAAGAGGATGTCGACGAAGGTATCG  
ACGAATTCGCACACCACGTGGAGAAAAAGGAGGAGATTGTCATTAGAACAAGTGCG  
ATCATTAGAACGGAATTTTGAAGTGGAAAACAAGCTTGAGCCCGAGAGGAAAATGC  
AACTAGCTAAGGAGCTTGGACTGCAACCTCGTCAAGTGGCGGTGTGGTTCCAGAAT  
AGACGGGCAAGGTGGAAAACCAAACAGCTCGAGCGCGACTACGAGACCCTGAAGA  
AAGCCTACGACAGGCTTAAAGCAGACTTCGAAGCCGTTACTCTAGACACAAGTGCT  
CTTAAAGCTGAGGTGAGTCGCCTCAAGGGAATCTCTAATGCACGACGTCAAGCCCG  
CCGAATTCGTTTCAGGGCAAGTGTGACACAACGAGTCACCCCTGCCTCCCCTGCGCAAT  
CGGAGAGGTCCGACATTGTGTTCATCGAGGAATCGCACAACTCCTACCATACATGTGG  
ATCCCGTGGCACCCGAGGAAGCCGGCGCTCACTTAACCATGAGCTCGGATAGCAAT  
TCCAGCGAGGTCATGGACGCTGATAGCCCTCGCACGAGCCACACCAGCGCTAGTAG  
GAGCACTTTGTCCACAAGTGTGGTGCAGCCTGACGAGGGCCTGGGAGTGGCCCAGT  
ACCCCCACTTTTCTCCCGAAAACCTTCGTGGGTCCCAATATGCCAGAGATTTGCGCTG  
ATCAGTCACTTGCATCTCAAGTGAAGCTGGAAGAGATCCACAGCTTCAATCCCGACC  
AAACCTTCCTGCTCTTGCCCAACTGGTGGGATTGGGCTTGATTCGTTTCTTCATCTGT  
ACCCATACACTTTTTCTTGAATCCAAGTTGAATTCACCTTAGGCAGTGTTTTTTCAC

GATGTACCACTTGTTATTCTTCCACCATGTGCAATCCAACGTCAANCAAAGTTGCAC  
ATCGGCGATCATTGGTGACGATGTCGAGCATCGATCGTCACATGC

**FIGURA 1E**

Secuencia de nucleótidos del ZF-1 parcial de *Physcomitrella patens* (SEQ ID NO:5)

GCACGAGCTCGGTTGTGGAAGCTGTCTCGTGGCTTCTTCCGCACCCTAAGATCTCGA  
CCAACTCTATTATCAGAGGCAGCGCTGCAGCCGACGAGATGGGTTTCGTTCGCTTTCC  
ACGACCGGCCCTTTAGTCCCAAGCCCAAGAAACAGAAGGTTGAATTGCCCGCGGAC  
ATATTCTCTGATGTGGACCCTTTCTAGACTTGGACGATGCAACCGTTACCGGAATT  
CAACCCGACAGCTTGGTACCAGTCCATATGCCAGAATGCTCCGAGGACACGGATTC  
GCTTGCTCACTCCATGGACCCTTCGTTTACTAAATTTCTCTCTCGGCCGAAGAGCGGT  
TACAGCTATGGCACATCTACCCTTACTCAGAGCATTCTTGTTCGTCTCTAGATGCCG  
CCGTTGTTCCAGACTCCAGTCTCAGCGACATTTCCACACCCTACCTAGACTCACAAA  
GCTCCCAAGATATGTCAGCTCGCCTGCCACACCAGACTGGAGGTCCCATTGACACCG  
TCGACCGTGAAGCTCGCGTGTGCGCTACAAGGAGAAGAGGCAGAAGCGCAAGTTT  
GAGAAAACAATTCGCTATGCATCAAGGAAGGCATATGCTGAGAGCCGGCCGAGGAT  
CAAAGGAAGGTTTCGCTAAGAGAACTGATTCCGACATGGAGCAGTTTGGCTCAGTGG  
ACTCAAGTTTCGGAGTGGTTCCAAGTTTTTGAGTTTTCTTGTGTATTGGAGTCTCCAT  
CGAGCAAGGTCATCTGAAATGGAAAGCTGCTGTGTAACATAGAGGAGCTGCTGTAA  
GAACTGTGTAGAGCCATCCAAGTGGTGAAGCACCTGAAAAAGTGGCAGCAATGTAA  
ATTGTTTCAGACTCTCAATGGTCACCAGTACCAAGTCATGCCATTCTATAATCCCTTTC  
AGAACACGATTAAATGCCTTGTGGACAGTACAGGATGTAGTCAGAGTTCTAGTAGT  
GGTTTTTTTCTATTTTTCTTTTTGTTGATTGAGAGCTTTCGGAACGGTGAGAACTTCGT  
GGCGCGAATCCTCTGTCCTGCGATCGTTATGATGCAGCGAATTCCTCCGATCTTGAT



GTATTTCAACACTTCCATAATGCTCTTGGATTTTTGGGTCATTTCTCAGAAGGTGTT  
GAGCTAACAAAAAAAAAA

**FIGURA 1F**

Secuencia de nucleótidos del nucleótido parcial LZ-1 de *Physcomitrella patens* (SEQ ID NO:6)

CTTCAAAGAACTCGGCAACAGCAGGGTCTATACCTTGGACCTGGTTCGTACAGTGAC  
CAAAATGGTCAGTCGGGTGGAGTTGGTGGAGCAAACGCATATAGTTCAGGAGCTGC  
TGCATTTGACCTGGAGTATGCAAGGTGGGTTGAAGATCATACCCGGCAGATGAGTG  
AGCTCCGGGTGGCCCTACAGGCTCATGTCGCTGACGCTGATTTACGATTACTAGTGG  
ATGGGAGTATGGCCCACTATGACGACCTCTTTCCGGCTCAAGGACGCTGCTGCAAAG  
CCGACGTGTTTCATCTCGTGTCCGGCATGTGGAAACTCCTGCAGAGCGATGCTTTG  
TATGGATTGGAGGCTGCCGACCCTCTGAGTTACT

**FIGURA 1G**

Secuencia de nucleótidos del CABF-2 parcial de *Physcomitrella patens* (SEQ ID NO:7)

CATCCTGGCGCCGTCATGCCTTTACAGATGCACTACCCGCAAGCCCAGCAACAGATG  
ATGCCGCAGCTTGGTGATCAGCAGATGCAGCCGCAGCTTCATTATCAGCAAATTCAG  
AAACAGCAGCTGTCCCAGTTCTGGCAGCAGCAAATGCAGGAAATGGAGCAAGTCAA  
TGATTTTAAGACCCATCAGCTACCACTGGCACGCATCAAAAAAATCATGAAGTCGG  
ATGAAGATGTTAAGATGATCGCAGCCGAAGCTCCAGTGCTGTTTTCAAAGCTTGTG  
AGATGTTTATTTTAGAATTGACACTGCGCTCTTGGATTCATACGGAGGAAAATAAGC  
GAAGGACACTACAAAGAAATGATATTGCAGGGGCTATCACTAGGGGAGACATCTTC  
GACTTTCTTGTGACATCGTTCCACGTGACGAGTTGAAGGAAGAAGATTTGGGTGTG  
CCATGGAC

**FIGURA 2A**

Secuencia de nucleótidos del CABF-1 de longitud total de *Physcomitrella patens* (SEQ ID NO:8)

GGCACCAGCGAATCCGTCTCCGCTCCGCCCTTCTGCACGCGTGGTTGTGGTTCGACCT  
CTCGCCGGAGCAACAGGAACTAATCCCTTTTCCAGCACTAAACGATTGAAGCAATT  
TTTTTTTTCTTGTAAGTCTCACTCTCTCTCTGTTATGAGGGGATTCTGAAGCTTGAA  
AGTTATGAGCTGAAGGTTGAGGACACGTAAGAGCGAAGGACGATCATACTACAATT  
AACCTTGCGGGGAAAAGCCCAGGCAAAATAGGACGGATGGCCGACAGTTACGGCC  
ACAACGCAGGTTACCCGAGAGCAGCCCGCATTCTGATAACGAGTCCGGCGGCCAT  
TACCGTGATCAGGACGCTTCTGTACGGGAGCAAGACCGGTTTTTGGCCATCGCAAAT  
GTGAGCCGAATCATGAAGAAAGCATTGCCATCTAATGCGAAGATATCGAAAGACGC  
CAAAGAGACTGTGCAGGAGTGCGTATCCGAGTTCATCAGTTTCATTACTGGTGAGGC  
GTCCGACAAGTGTGAGAGGAAAAGAGGAAGACGATCAACGGGGATGACTTGCTGT  
GGCCATGAGTACTCTTGGTTTTGAAGATTATGTGGAACCTCTGAAGGTGTACCTAC  
ACAAGTATCGTGAAGTGGAGGGGAGAAGGCCTCTATGGCCAAGGGTGGTGATCAG  
CAGGGAGGAAAAGAGAGCAACCAAGGAGGTATGGGGTCGATGGGCATGGCAGGCG  
GAATCAACGGCATGAACGGAACGATGAACGGGAACATGCATGGGCATGGAATCCC  
GTATCGATGCAGATGATGCAACAGCCGTACGCGCAGCAGGCACCTCCGGGGATGAT  
ATATTCCTCATCAAATGATGCCGCAATACCAGATGCCGATGCAGTCTGGTGGAAA  
CCAACCCCGCGGAGTGTAAGAGTTTTCACTGGCAGGAGGCTTTGGAAGTGGGGATA  
TTGTCGACAGCGTGATGGGGTGTTTTGGAGCATGGGCAGGGCATTATGGTGCTGTTG  
AAACAGTGATGGGTGGGTTCATGTGAAGTGTGGCGACTGTTGAATGATGAAAACAT  
AGAAGTGATGTCGTTGAAGCTCGGGGAGTTTCAAGTGAAAGGAGGAGCACTTTTTG  
TTTGGAAAGGAGCGTACCGGGTCTGGCAGTGTACATTCTGAATGATAGTTATCTGTG  
CTGATTTTTCTTGGCCTTGGCAATACGAGGGGGTTGAATATTTTGCTTTGAATTCGTT  
GACATTTCAACCTTTTCTATGTGAAAAGGCTCTGTAGGATGCAAGATAAGGAAAGAC  
ATGCAGATTGATAA

**FIGURA 2B**

Secuencia de nucleótidos del DBF-1 de longitud total de *Physcomitrella patens* (SEQ ID NO:9)

GCCCTTATCCCGGGCGATGGTGCGTTCGAGATCGTAAGGTTGCCGACGAAGGCGTA  
ACTTGAAGTCTCTGTGTCCCGGCGATGTCCCAATGTTGGCCCGATTTCTGTTTTT

ES 2 279 777 T3

AGCGAGCTGTGGGCTAGTTTGTGGGTATGATCCGGGGAATGAGACGAGATGTCTGT  
CTGAGTGAGACCACTCTAGGGGCTGTTGGAGGATGAGGAGGGAAGCGCAGAAGTTG  
GCCATTCTTTTCAGTGACTGGACTCTGTGCGAGTGGTCAGCTTTCGGGAGCTGCTGTT  
GCATTGACCGGTGATTCTTTCGAGATCGTAGAGACAGCAGCTGGCAAGGGTTTTGGG  
AGGCTTTTCAAATGAAGGGCATTCAAGAGCTTTCAGATGATGAAGATTATATCCGC  
CTGTCAATGCATCGCGATATTTCAACAGGGGCAAAGCGCTCTCAAAGACATCATCCA  
ATCATGCCAATGGAAATGGAAATCCAAACGGAACGAGTTTTGGAGTTTCAACTTCTT  
CAGCAAGTGACTCTGACAAAAGATAAGAAATCCGAAGTTTCAGGTTCTTTACTAAGCG  
ATTCTGGCAAGAATCAAAAGTCCGTTACTGAATTGGATTTCGTTTCGCATTTAACCGCA  
AGTCCAGAATTGCCAAGCGACCTATCGAGCTACTCGAAGACGAGGAGGACGTGGAC  
GTTGGAGCTGCAAAGGTTGTAGACATTGAGCCGACTAACGGAAACAAGAGGCGGAG  
ACGGTATCACACCATCGAAGACAGTGACGATGAAGAGTTGGATGAGAAGAAATCGT  
TTGGTGATAATCTGACCCCAGGAACGGAAATCGATCAATGTGCAGCCGATGAATCCT  
TAGCAAAAAGGTTGCAGGATTTAGAGCACCGGGCAGTTTCTGGCCGTAATCGCCTG  
GTTCAAATTTTGTGAGATTCCGATGAAGAAGAAGAGGAAGAAGTAAATCCCATAAC  
CATCACCTACAAAGGTGTGACCAGATTGCAGCATCATTGCGAGAAGAGCTGCAGG  
CCAGCAGITCAAGTGATAACTCGGTTAATGAAGATCGTTATGCAGAGGTTGATGTAG  
CAGCAGCAAAAATTGTGAGCCAGGCAGATGTCTGTGCAGCTTGTGGCATTGCCGAG  
AATGATACACAACGAATGCTCAAGCCATATCAGCTTGTAGGCGTCAATTCATGCTG  
CTACTTCACCGCAAACATGTTGGGGGTGCAGTTGCGTATCTTGCCCTTCTGAAACAT  
CTTGATGGAGATGCTGGTCCTCATCTTTTAGTTGCACCTGCTTCTCTTTTAGAAAAC  
GGCAAAGAGAACTCAAGAAGTGGTGTCTGCATTTAAGGTGGAGCTCTATCATGGC  
TCAGGAAGGGCAGCTTTAAACAGGAGGCTTCAGTATGCTGCAAATCTAAAGGGCC  
TGCACCCTTTAAACGTGATGCTGACGTGCTACTCCCTTTTTGAGAGGCAGAGTGCTCA  
GACAAAAGATGACCGCAAATTCCTTAAGAAATGGAATTGGCGCTGTGTGGTTATGG

ES 2 279 777 T3

ACGAGGCTCATCTTTTGAAGGACAGAAGCAGCTTTCGCAGCAAAAAGTTGCGAGAT  
ATAGCTCACAAAGCAATCAAAGACTGATGCTGACTGGTACTCCACTCCAGAACGAT  
TTGCAGGAGCTATGGTCACTTCTGGAGTTCATGATGCCTGATGTGTTCAACACAAAT  
GGCGTTGATTTAGATCAATATCTGGGAACCAGGAACGATACCTCAGGGATTGTGTG  
CAGGATACGAACTTGATGACTCGGATCAAAGGAATACTAGGACCTTTTGTATTACGG  
AGAATGAAAAGCTGATGTTATGCGCCAGCTTGTATCAAAGATTCAGGAGGTGGAGTG  
TGTGGAGATGCTAGACGAGCAATCAATGGCATATAAAAAAGCTGTGAATGAGTATA  
GAGCCCTTGCTGAGTCCGCACGTGCCGCTAAAGCTGCAAAGAAATCCTCAGTTAGC  
GTAGTAGATGTCCTTCTCGTGCACAAGTGACCAATATCTTTACTCAATTGAGAAAG  
GTCAAGAAATTGGCTAAGAAATTCATCCATTAGGAGTTTTTGGATATGAATGCGAT  
TTGCAGCGTGTGGAGGAAGAATTGACTAGTTACAGCGATTTTGACCTCCACAAGTTG  
TGTATTCAATATGGAGGCGCTGCGGGAGGGCAAGGAAAGCTTGATGATGATCATGC  
ACTAGCTTCTGCAAAGTGCCAGGCTTTAGCACGTCTACTTCCCAAGTTACAGCAAGG  
TGGCCATCGCACATTGATATTCAGCCAGTGGACAAGCATGCTGGATATTTTAGAATG  
GGCTCTTGACGTCATGGGTTTTTCTTACACTCGCCTAGATGGAAGCACTCAAGTAAG  
TGAACGCCAAACCCTAGTGGACGAGTTCAACAATGACCCTAGCATATTTGTGTTTCT  
CCTGTCTACTCGAGCTGGAGGTCAAGGTCTAAATTTAACAGGAGCAGACACAGTCAT  
TTTACATGATTTGGACTTCAATCCCCAAATGGATCGACAGGCTGAGGATCGCTGTCA  
TCGGATTGGCCAGTCTAAACCTGTTACGATATAACCGACTTGTAACAAAAGATACGGT  
CGATGAAAGTATATACAAGATAGCCCAACAGAAGCTGGTCCTCGATGCGGCAGTTC  
TTGAAGGAAAAGAGTCATCCTCTGATCTTAATGATGGTGATGCTCGCACGATGGGTG  
AAATCTTTCTGCATTATTGGATGTTCCACCGACATGATCCTGGAGTCCAGAACACA  
TTTTTAATTTATTTTCATTATCTTTATCTGGCACTGCGAGAAAGCTCGTTAACGCAAG  
GGC

FIGURA 2C

Secuencia de nucleótidos de la variante de DBF-1 de longitud total de *Physcomitrella patens*  
(SEQ ID NO:22)

GCCCTTATCCCGGGCGATGGTGC GTTCGAGATCGTAAGGTTGCCGACGAAGGCGTA  
ACTTGGAAGTCCTCTGTGTCCCGGCGATGTCCAATGTTGGCCCGATTTTCTGTTTTT  
AGCGAGCTGTGGGCTAGTTTGTGGGTATGATCCGGGGAATGAGACGAGATGTCTGT  
CTGAGTGAGACCACTCTAGGGGCTGTTGGAGGATGAGGAGGGAAGCGCAGAAGTTG  
GCCATTCTTTTCAGTGACTGGACTCTGTGCGAGTGGTCAGCTTTCGGGAGCTGCTGTT  
GCATTGACCGGTGATTCTTTCGAGATCGTAGAGACAGCAGCTGGCAAGGGTTTTGGG  
AGGCTTTTCAAATGAAGGGCATTCAAGAGCTTTCAGATGATGAAGATTATATCCGC  
CTGTCAATGCATCGCGATATTTCAACAGGGGCAAAGCGCTCTCAAAGACATCATCCA  
ATCATGCCAATGGAAATGGAAATCCAAACGGAACGAGTTTTGGAGTTTCAACTTCTT  
CAGCAAGTGACTCTGACAAAGATAAGAAATCCGAAGTTTCAGGTTCTTTACTAAGCG  
ATTCTGGCAAGAATCAAAGTCCGTTACTGAATTGGATTTCGTTTCGCATTTAACCGCA  
AGTCCAGAATTGCCAAGCGACCTATCGAGCTACTCGAAGACGAGGAGGACGTGGAC  
GTTGGAGCTGCAAAGGTTGTAGACATTGAGCCGACTAACGGAAACAAGAGGCGGAG  
ACGGTATCACACCATCGAAGACAGTGACGATGAAGAGTTGGATGAGAAGAAATCGT  
TTGGTGATAATCTGACCCCAGGAACGGAAATCGATCAATGTGCAGCCGATGAATCCT  
TAGCAAAAAGGTTGCAGGATTTAGAGCACCGGGCAGTTTCTGGCCGTAATCGCCTG  
GTTCAAATTTTGTGAGATTCCGATGAAGAAGAAGAGGAAGAAGTAAATCCCATAAC  
CATCACCTACAAAGGTGTGACCAGATTGCAGCATCATTGCGAGAAGAGCTGCAGG  
CCAGCAGTTCAAGTGATAACTCGGTTAATGAAGATCGTTATGCAGAGGTTGATGTAG  
CAGCAGCAAAAATTGTGAGCCAGGCAGATGTCTGTGCAGCTTGTGGCATTGCCGAG  
AATGATACACAACGAATGCTCAAGCCATATCAGCTTGTAGGCGTCAATTTTCATGCTG  
CTACTTCACCGCAAACATGTTGGGGGTGGCAGTIGCGTATCTTGCCCTTCTGAAACA

ES 2 279 777 T3

TCTTGATGGAGATGCTGGTCCTCATCTTTTAGTTGCACCTGCTTCTCTTTTAGAAAAC  
TGGCAAAGAGAACTCAAGAAGTGGTGTCTGCATTTAAGGTGGAGCTCTATCATGG  
CTCAGGAAGGGCAGCTTTAAACAGGAGGCTTCAGTATGCTGCAAAAATCTAAAGGGC  
CTGCACCCTTTAACGTGATGCTGACGTGCTACTCCCTTTTTGAGAGGCAGAGTGCTC  
AGACAAAAGATGACCGCAAATTCCTTAAGAAATGGAATTGGCGCTGTGTGGTTATG  
GACGAGGCTCATCTTTGAAGGACAGAAGCAGCTTTCGCAGCAAAAAGTTGCGAGA  
TATAGCTCACAAAGCAATTCAAAGACTGATGCTGACTGGTACTCCACTCCAGAACGA  
TTGCAGGAGCTATGGTCACTTCTGGAGTTCATGATGCCTGATGTGTTCAACACAAA  
TGGCGTTGATTTAGATCAATATCTGGGAACCAGGAACGATACCTCAGGGATTGTGT  
GCAGGATACGAACTTGATGACTCGGATCAAAGGAATACTAGGACCTTTTGTATTACG  
GAGAATGAAAAGTATGTTATGCGCCAGCTTGTATCAAAGATTCAGGAGGTGGAGT  
GTGTGGAGATGCTAGACGAGCAATCAATGGCATATAAAAAAGCTGTGAATGAGTAT  
AGAGCCCTTGCTGAGTCCGCACGTGCCGCTAAAGCTGCAAAGAAATCCTCAGTTAG  
CGTAGTAGATGTCCTTCCTCGTCGACAAGTGACCAATATCTTTACTCAATTGAGAAA  
GTCAAGAAATTGGCTAAGAAAATTCATCCATTAGGAGTTTTTGGATATGAATGCGAT  
TTGCAGCGTGTGGAGGAAGAATTGACTAGTTACAGCGATTTTGACCTCCACAAGTTG  
TGTATTCAATATGGAGGCGCTGCGGGAGGGCAAGGAAAGCTTGATGATGATCATGC  
ACTAGCTTCTGCAAAGTGCCAGGCTTTAGCACGTCTIACTTCCCAAGTTACAGCAAGG  
TGGCCATCGCACATTGATATTCAGCCAGTGGACAAGCATGCTGGATATTTTAGAATG  
GGCTCTTGACGTGATGGGTTTTTCTTACACTCGCCTAGATGGAAGCACTCAAGTAAG  
TGAACGCCAAACCCTAGTGGACGAGTTCAACAATGACCCTAGCATAATTTGTGTTTCT  
CCTGTCTACTCGAGCTGGAGGTCAAGGTCTAAATTTAACAGGAGCAGACACAGTCAT  
TTTACATGATTTGGACTTCAATCCCCAAATGGATCGACAGGCTGAGGATCGCTGTCA  
TCGGATTGGCCAGTCTAAACCTGTTACGATATACCGACTTGTAACAAAAGATACGGT  
CGATGAAAGTATATACAAGATAGCCCAACAGAAGCTGGTCCTCGATGCGGCAGTTC

TTGAAGGAAAAGAGTCATCCTCTGATCTTAATGATGGTGATGCTCGCACGATGGGTG  
AAATTCCTTCTGCATTATTGGATGTTCCACCGACATGATCCTGGAGTCCAGAACACA  
TTTTTAATTTATTTTCATTATCTTTATCTGGCACTGCGAGAAAGCTCGTTAACGCAAG  
GGC

**FIGURA 2D**

Secuencia de nucleótidos del CBF-1 de longitud total de *Physcomitrella patens* (SEQ ID NO:10)

GGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCT  
TTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTC  
ACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCTCACTAA  
AGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATC  
CCCCGGGCTGCAGGAATTCGGCACCAGAGGATTTCCGGGTGAAGGAAGGTGCGATT  
GAAGGCGTGCATGAAATTTGACGCAGCGGGTTGAATCGGAAGGAGTTTTCAGCAG  
CTGGAAAGTACCTTCGAGGGTTGATTCATCCAAGTTTCCATGTGATATGGCTTCAA  
AGTATCCGCGAAAATGTAGAGAGCACGCATCACCTGGAGTTGGTGGCAGGGAATCT  
ACGCATCGCTTTGATTCAAGGTCGCAGTCGTATTCGTTCTCGGAGAAACCTTACCAC  
CGTCGTCCCGGGATGCGTTTCGTGATATGATCAGCGATTTGGTGCATCAGCCTTCC  
GACACTGCCGTGCCTGGTTTCAGGGGAGTGCCTATCGTCAGAACTGAACAAGTA  
CGTGACAGAGATTCGGCCCACGAGGTGCTCGAAGAAAATTTGGCTTGGGACGTATG  
ACACTGCAGAGGAAGCAGCGCGTGCCTTTGACATCGGAAAATTTGTGTTGTAAGAAA  
AACCTGCCGCTCAACTTCCGGATTCGACTCAGATGTTGCAGAGAATCTCGTCGAAA  
TTGACCCCCGAAGCGCAACGAAAAGCCATCGCGACGCTGGCGAAGGACGTAGTGCG  
AATGGAAAATGACAGGTGCAAGTTGGGTGGCGGTAACCTGACTACCACAGAGCCCC  
CGGTCCATTCCGAGCCTATTACTCAACACCTTGCAGCAGCTGAGATTCGCGCGGTCA  
CGTACATTGAACAGCCCCTGGAAAATTGTCTACGGAGTGGAAGAATCGGCGACGGCC

ATGTCGGTAACGGAAGCAAATGCGCGGATAACCACTCTTGGAGTTGGGACTTGGG  
CAAAGTGATCCTTGATGACGAGCTCTCTGAAATTCCTAACTTCGTCGGAGAACTAGA  
TCACGAGGCTATGGATTTTCAGTAGTCATGGAGAGGTTTACTACCACCATTATGACTC  
TCAGTGAGTCCTACAAGCATATTTTCAACTAGTCAACATCCTCAGTAGATTTAGTCC  
ATTACTTTCTGTGTCAGAGCCACGCCTGCGGCTTAGACCGGGAAAGCTTGTATAAAC  
TGTA AATTGAGCTCTCGTAGACATGATGTAACACCCAATCACCTGTAAACCCCCCAG  
CTTGAGATCACAAGGAGTAGAAAACCTGATAGCTTCAAGAGTTTCAACCAAAAAAA  
AAAAA

**FIGURA 2E**

Secuencia de nucleótidos del HDZ-1 de longitud total de *Physcomitrella patens* (SEQ ID NO:11)

GGGAAGGGTGTAAGGAGGCCACGGGCAGGGTCAAGGTGTCCAATGCACCTGCAAG  
ATCAGGAAGCTTGAAGTAGATCAGGGAAAAACGATGGTAGTCCCTAGTTTACCCG  
CCTTCGGAGGACAGAACGCCATGCTCAGACGCAACATTGACAACAACACCGACT  
CTGATTTCTCTGCTTCAAGGGTCTGCTCCCTCGCGTGAGCATGCAACAAGTGCCG  
CGTTCATCGGAGAGTCTCGAAAACATGATGGGGGCTTGTGGGCAAAAACCTGCCTTA  
CTTTTCGTCAATTTGATGGGCCGAGTGTAAGAAGCAAGAGGATGTCGACGAAGGTA  
TCGACGAATTCGCACACCACGTGGAGAAAAAGAGGAGATTGTCAATTAGAACAAGTG  
CGATCATTAGAACGGAATTTTGAAGTGAAAAACAAGCTTGAGCCCGAGAGGAAAAAT  
GCAACTAGCTAAGGAGCTTGGACTGCGACCTCGTCAAGTGGCGGTGTGGTTCCAGA  
ATAGACGGGCAAGGTGGAAAACCAAACAGCTCGAGCACGACTACGAGACCCTGAA  
GAAAGCCTIACGACAGGCTTAAAGCAGACTTCGAAGCCGTTACTCTAGACACAAATG  
CTCTTAAAGCTGAGGTGAGTCGCCTCAAGGGAATCTCTAATGACGACGTCAAGCCCG  
CCGAATTCGTTCAAGGCAAGTGTGACACAACGAGTCACCCTGCCTCCCTGCGCAAT



ES 2 279 777 T3

CGGAGAGGTCCGACATTGTGTCATCGAGGAATCGCACAACTCCTACCATACATGTGG  
ATCCCGTGGCACCCGAGGAAGCCGGCGCTCACTTAACCATGAGCTCGGATAGCAAT  
TCCAGCGAGGTCATGGACGCTGATAGCCCTCGCACGAGCCACACCAGCGCTAGTAG  
GAGCACTTTGTCCACAAGTGTGGTGCAGCCTGACGAGGGCCTGGGAGTGGCCCAGT  
ACCCCCACTTTTCTCCCGAAAACCTTCGTGGGTCCCAATATGCCAGAGATTTGCGCTG  
ATCAGTCACTTGCATCTCAAGTGAAGCTGGAAGAGATCCACAGCTTCAATCCCGACC  
AAACCTTCCTGCTCTTGCCAACTGGTGGGATTGGGCTTGATTTCGTTTCTTCATCTGT  
ACCCATACACTTTTTCTTGAATCCAAGTTGAATCACTTTAGGCAGTGTTTTTTAC  
GATGTACCACTTGTTATTCTTCCACCATGTGCAATCCAACGTCAACCAAAGTTGCAC  
CATCGGCGTTAACGCAAGGGC

**FIGURA 2F**

Secuencia de nucleótidos del ZF-1 de longitud total de *Physcomitrella patens* (SEQ ID NO:12)

GAGGAGGGAGTTGGAATCTAGGAGACGTGCATGTGCTGTGGGAGGAATTCTCTGGG  
GATTCGAGGCCTTGTGTATGTTGTTCAAGTAAAGGGAGTAGCTTTTTTCACTTGAAG  
GGGCTGGTGCTGCTGTTGTTGCAAGTCTTTTGACATTGAAAGAGGCGGGGTTGCACG  
CCCCGGTGTGAGGAAGAGTCTTGTAGTAGGTGGGTTGTGTTGTGCCGTGGTATAGTA  
TGCCGAAGCCTTGTGATGCATGCCATGTTTCCAGCGCGGGTGTCTGCCGAGCGG  
ACGCTGCCTACCTGTGCGTAGGCTGCGATGGGAAGGTCCACGGGGCCAACAACTA  
GCGTCTCGACACGAGCGCGTGTGGATGTGCGAAGTGTGCGAGGTTGCTCCAGCCGT  
GGTGACCTGCAAGGCGGATGCGGCTTCTCTCTGTGTGGCCTGTGACACAGACATCCA  
CTCCGCCAACCCGCTAGCGCAGCGTCACGAGAGAGTGCCGGTGACACCTCTGTTCCG  
AGAGTGCGAGTCCTTTGCGTGGGCCAGATTTCTGCGTGTGGTGTGAGAGAATGGGT  
GCCATGATCTGCTGAAGGGCTGTGAGGACGCCTCGGTTGTGGAAGCTGTCTCGTGGC  
TTCTTCCGCACCCTAAGATCTCGACCAACTCTATTATCAGAGGCAGCGCTGCAGCCG

ACGAGATGGGTTTCGTCGCCTTCCACGACCGGCCCTTTAGTCCCAAGCCCAAGAAAC  
AGAAGGTTGAATTGCCCGCGGACATATTCTCTGATGTGGACCCTTTCCTAGACTTGG  
ACGATGCAACCGTTACCGGAATTCAACCCGACAGCTTGGTACCAGTCCATATGCCAG  
AATGCTCCGAGGACACGGATTCGCTTGCTCACTCCATGGACCCTTCGTTTACTAAAT  
TTCCTCTCTCGGCGAAGAGCGGTTACAGCTATGGCACATCTACCCTTACTCAGAGCA  
TTTCTTGTTTCGTCTCTAGATGCCGCCGTTGTTCCAGACTCCAGTCTCAGCGACATTC  
CACACCCTACCTAGACTCACAAAGCTCCCAAGATATGTCAGCTCGCCTGCCACACCA  
GACTGGAGGTCCCATTGACACCGTCGACCGTGAAGCTCGCGTGTGCGCTACAAGG  
AGAAGAGGCAGAAGCGCAAGTTTGAGAAAACAATTCGCTATGCATCAAGGAAGGC  
ATATGCTGAGAGCCGGCCGAGGATCAAAGGAAGGTTTCGCTAAGAGAAGTATTCCG  
ACATGGAGCAGTTTGGCTCAGTGGACTCAAGTTTCGGAGTGGTTCCAAGTTTTTGAG  
TTTTCTTGTTATTGGAGTCTCCATCGAGCAAGGTC

**FIGURA 2G**

Secuencia de nucleótidos del LZ de longitud total de *Physcomitrella patens* (SEQ ID NO:13)

GCCCTTATCCCGGGTGCTCTGGCAGTGGGACGGATTTGGAAGCAACAGGAGGTGGG  
CTTGTTGAGCTGCGGAGTATGGAAAAAAGCGGGAAGGTGACGTGAGAGCTGGAATG  
ATGGCCGAGTGAGCGTGTGTTTGTGTTTGGAGGGGTAATTAGATGGGAAGATAGAGGTC  
GGATGAGTCTGGGCGGTTTGCGTAGAGACGTCGAGGAAAAGGAAAGTGGCGAGGTG  
TAGGATCTTGGTGGATTTTTCTCCCCTGAAGCTAGAGACTTCCGGTGCAGAATGTGG  
TTAAATGGAACCTAACAGGTGGAATTCATGACATGGAAACCTACTGGGTCTTGTTG  
GAATAAATCTCACGCTGTCGGCTTCTTTACGTCATTTTCTTAGGTTTCAGAGATAT  
AGTAGAAAGGTTTGTGGAATTATCAAGATGGGTGACAACAGTGCAAGTGCAAGGAC  
GGATTCATCTTCTGACATGGACGGTGATGCGAAGTTGGATGATGGGCAGCACTTAGC  
TAGTGGCGGTGGAAACTCAAACGATTCCAGTCTCGAAACTGGAACGAAGAATGGCG

ATTCTAAGGTACTAAGGCGGTAGCACAAAATCGTGAGGCAGCCCGAAAAAGTAGG  
CTCAGAAAAAAGGCATATGTGCAGCAGTTGGAGTCCAGCCGCATAAAGCTGAACCA  
ACTCGAGCAAGAGCTTCAAAGAACTCGGCAACAGCAGGGTCTATACCTTGGACCTG  
GTTTCGTACAGTGACCAAAATGGTCAGTCGGGTGGAGTTGGTGGAGCAAACGCATAT  
AGTTCAGGAGCTGCTGCATTTGACCTGGAGTATGCAAGGTGGGTTGAAGATCATAACC  
CGGCAGATGAGTGAGCTCCGGGTGGCCCTACAGGCTCATGTCGCTGACGCTGATTTA  
CGATTACTAGTGGATGGGAGTATGGCCCACTATGACGACCTCTTTCGGCTCAAGGAC  
GCTGCTGCAAAAGCCGACGTGTTTCATCTCGTGTCCGGCATGTGGAAAACTCCTGCA  
GAGCGATGCTTTGTATGGATTGGAGGCTGCCGACCCTCTGAGTTACTAAAGATATTA  
GTACCTCAAATAGAACCTTTGACAGAGCAGCAGTTGTTAAACATCTGCAATCTGCAG  
CAGTCCTCTCAACAGGGTGAAGAGGCCCTCTCTCAAGGGATGGAACAACCTCAGCA  
GTCGCTTGCCGAAACACTGTCTGCCGGTCTCTTGGCTCAGCAGCAAATGTTGCCAA  
CTACATGGGACAGATGGCTGTGGCCATGGGACAACCTTGGGAACCTCGAAGGTTTCG  
TGCGTCAGGCTGATCATTGCGACAACAGACGTTACAACAAATGCACCGGGTATTAA  
CCATTCCCAAGTAGCCCGAGGACTTCTTGGCATGGGTGATTACTTTGCTCGTCTTCG  
AGCTCTTAGTTCTCTATGGTCCGCCAGGCCTCGTGAATGAGAAACATTGTCGTTTCA  
GGCGATGGTGAAGTCTTCGGCGCAGGTATGGAAGCGATATCGCAAGGGC

**FIGURA 2H**

Secuencia de nucleótidos del CABF-1 de longitud total de *Physcomitrella patens* (SEQ ID NO:14)

CAGCATCCTCACATCCCGCCTTCTCTGCACCCAGATGTCGCATCCTGGCGCCGTC  
ATGCCTTTACAGATGCACTACCCGCAAGCCAGCAACAGATGATGCCGCAGCTTGGT  
GATCAGCAGATGCAGCCGAGCTTCATTATCAGCAAATTCAGAAACAGCAGCTGTC  
CCAGTTCTGGCAGCAGCAAATGCAGGAAATGGAGCAAGTCAATGATTTTAAGACCC  
ATCAGCTACCACTGGCACGCATCAAAAAATCATGAAGTCGGATGAAGATGTTAAG

ES 2 279 777 T3

ATGATCGCAGCCGAAGCTCCAGTGCTGTTTTCAAAAGCTTGTGAGATGTTTATTTA  
GAATTGACACTGCGCTCTTGGATTCATACGGAGGAAAATAAGCGAAGGACACTACA  
AAGAAATGATATTGCAGGGGCTATCACTAGGGGAGACATCTTCGACTTTCTTGTGTA  
CATCGTTCCACGTGACGAGTTGAAGGAAGAAGATTTGGGTGTGCCATGGACTGGTGT  
TCCAGGGGATGGTTCTGTACCTTACGGAGGAATATTCTATCCACCCATGGCTGGACA  
GCAAATGCATCATTCTATGGGGGCTCCTGAGATGATGGTTGGGCAGCCACCAAACCC  
ACAAATGATGTATCAGCCTCCACAGACTGCCTTTGTCCCCGAGCAGCAGCAACAGTG  
AATGCATTACCACCTAGAGAACGCTGAGCATCGAAGACGGGACAACCTCAAGGAAAG  
GGCTATCGCATCGAGATTCTTTCGTCACGTGGGAATGGTATTTATCATACTGTTGCTA  
CCATCTGTCAATTCTTATGGCAAAGGGGTACAGGATCAGGATTTTACCTTTCACTA  
CAGCGCTTTTGTGTTGGCTTTCAACTATAATTTAAGGAAATCGTAGCTGTAGGCGGT  
GATGCGACAGTTCTGAGCACTGCTAATTCTAGCAGAGTTTATGTTTGGTTTAGCAAG  
TCATGAAGGGCACAAAGGGACCCGACCCCTCCATGGATCTGGTAGAAATTTGTGAA  
TAGTGATACTAGTGCAGGCATAATTATTAGCATGTGCAGGAGTTGCTCTTAATGTTA  
GGTTCGAGGATCGGGTATCCATTTCTTAGTACCATTGTTTCTTTTATGTCTCCCTGG  
TTTTATCTTTCAGACTGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURA 3A

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

SKDAKETVQECVSEFISFITGEASDKCQREKRKTINGDDLWAMSTLGFEDYVEPLKVY  
LHKYRELEGEKASMAKGGDQGGKESNQGGMGSMGMAGGINGMNGTMNGNMHGH  
GIPVSMQMMQPYAQQAPPGMIYSPHQMMPQYQMPMQSGGNQPRGV\*

**FIGURA 3B**

Secuencia deducida de aminoácidos del DBF-1 de *Physcomitrella patens* (SEQ ID NO:16)

MKGIQELSDDEDYTPPVNASRYFNRGKALSKTSSNHANGNGNPNGTSFGVSTSSASDSD  
KDKKSEVSGSLLSDSGKNQKSVTELDSEAFNRKSRIAKRPIELLEDEEDVDVGA AKVVDI  
EPTNGNKRRRRYHTIEDSDDEELDEKKSFGDNLTPGTEIDQCAADESLAKRLQDLEHRA  
VSGRNRLVQILSDSDEEEEEVNPTITLQRCQIAASLREELQASSSSDNSVNEDRYAEV  
DVAAAKIVSQADVCAACGLAENDTQRMLKPYQLVGVNFMLLLHRKHVGGAVAYLALL  
KHLGDAGPHLLVAPASLLENWQRELKKWCPAFKVELYHGSGRAALNRRLQYAAKSK  
GPAPFNVMLTCYSLFERQSAQTKDDRKFLKKWNWRCVVMDEAHLKDRSSFRSKKLR  
DIAHKAIQRLMLTGTPQLQNDLQELWSLLEFMMPDVFNNTNGVDLDQYLGTRNDTSGIVV  
QDTNLMTRIKGILGPFVLRMKTDVMRQLVSKIQEVECVEMLDEQSMAYKKA VNEYR  
ALAESARA AKA AKKSSVSVVDVLP RRQVTNIFTQLRKVKKLAKKFHPLGVFGYECDLQ  
RVEEELTSYSDFDLHKLCIQYGGAAGGQGLDDDHALASAKQALARLLPKLQGGHR  
TLIFSQWTSMLDILEWALDVMGFSYTRLDGSTQVSERQTLVDEFNNDPSIFVLLSTRAG  
GQGLNLTGADTVILHDLDFNPQMDRQAEDRCHRIGQSKPVITYRLVTKDTVDESIYKIA  
QQKLVLDAAVLEGKESSDLNDGDARTMGEILSALLDVPPT\*

**FIGURA 3C**

Secuencia deducida de aminoácidos de la variante de DBF-1 de *Physcomitrella patens*  
(SEQ ID NO:23)

KDKKSEVSGSLLSDSGKNQKSVTELDSEAFNRKSRIAKRPIELLEDEEDVDVGA AKVVDI  
EPTNGNKRRRRYHTIEDSDDEELDEKKSFGDNLTPGTEIDQCAADESLAKRLQDLEHRA  
VSGRNRLVQILSDSDEEEEEVNPTITLQRCQIAASLREELQASSSSDNSVNEDRYAEV  
DVAAAKIVSQADVCAACGLAENDTQRMLKPYQLVGVNFMLLLHRKHVGG\*

**FIGURA 3D**

Secuencia deducida de aminoácidos del CBF-1 de *Physcomitrella patens* (SEQ ID NO:17)

MASKYPRKCREHASPGVGGRESTHREFDSRSQSYSEKPYHRRRRDAFRDMISDLVHQP  
SDTAVPGFRGVRYRQKLNKYVTEIRPTRCSKKIWLGTYDTAEAAARAFDIGNLCCKKNL  
PLNFPDSTQMLQRISKLTPEAQRKAIAITLAKDVVRMENDRSKLGGGNLTTEPPVHSEP  
ITQHLAAAEIRAVTYIEQPLEIVYGVEESATAMSVTEANARDNHSWSWDLGKVILDDEL  
SEIPNFVGELDHEAMDFSSHGEVYYHHYDSQ\*

**FIGURA 3E**

Secuencia deducida de aminoácidos del HDZ-1 de *Physcomitrella patens* (SEQ ID NO:18)

MVVPSLPAFGGQNAMLRRNIDNNTDTLISLLQGSCSPRVSMQQVPRSSSELENMMGAC  
GQKLPYFSSFDGPSVEEQEDVDEGIDEFAHHVEKKRRLSLEQVRSLEARNFEVENKLEPER  
KMQLAKELGLRPRQVAVWFQNRARRWTKQLEHDYETLKKAYDRLKADFEAVTLDT  
NALKAEVSRKLGISNDDVKPAEFVQGKCDTTSHPASPAQSERSDIVSSRNRTTPTIHVDP  
VAPEEAGAHLMSSDSNSSEVMDADSPRTSHTSASRSTLSTSVVQPDEGLGVAQYPHFS  
PENFVGPNMPEICADQSLASQVKLEEIHSFNPDQTFLLLPNWWDDWA\*

**FIGURA 3F**

Secuencia deducida de aminoácidos del ZF-1 de *Physcomitrella patens* (SEQ ID NO:19)

MPKPCDACHVSSAAVFCRADAA YLCVGC DGKVHGANKLASRHERVWMCEVCEVAPA  
VVTCKADAASLCVACDTDIHSANPLAQRHERVPVTPLFESASPLRGPDFCVLVSENGCH  
DLLKGCEDASVVEAVSWLLPHPKISTNSIIRGSAAADEMGSSPFHDPFSPKPKKQKVEL  
PADIFSDVDPFLDLDDATVTGIQPDSLVPVHMPECEDTDSLASHMDPSFTKFPKSAKSG  
YSYGTSTLTQSISCSLLDAAVVPDSSLSDISTPYLDSQSSQDMSARLPHQTGGPIDTVDRE

ARVLRVYKEKRQKRKFEKTIRYASRKAYAESRPRIKGRFAKRTDSDMEQFGSVDSSFGVV  
PSF\*

**FIGURA 3G**

Secuencia deducida de aminoácidos del LZ-1 de *Physcomitrella patens* (SEQ ID NO:20)

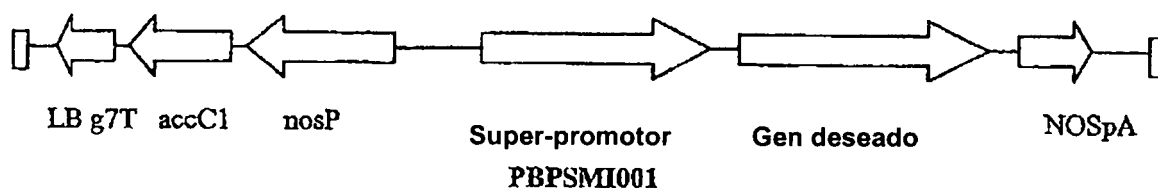
MGDNSASARTDSSSDMDGDAKLDDGQHLASGGGNSNDSSLETGTKNGDSKVLRRLAQ  
NREAARKSRLRKKAYVQQLESSRIKLNQLEQELQRTTRQQGLYLGPGSYSDQNGQSGG  
VGGANAYSSGAAAFDLEYARWVEDHTRQMSELRVALQAHVADADLRLLDVGSMAHY  
DDLFRDKDAAAKADVHFLVSGMWKTPAERCFVWIGGCRPSELLKILVPQIEPLTEQQLL  
NICNLQSSQQGEEALSQGMEQLQQLAETLSAGSLGSAANVANYMGQMAVAMGQLG  
NLEGFVRQADHLRQQTTLQQMHRVLTIRQVARGLLAMGDYFARLRALSSLWSARPRE\*

**FIGURA 3H**

Secuencia deducida de aminoácidos de CABF-2 de *Physcomitrella patens* (SEQ ID NO:21)

MSHPGAVMPLQMHPYQAQQQMMPQLGDQQMQPQLHYQQIQKQQLSQFWQQQM  
EMEQVNDFKTHQLPLARIKKIMKSDEDVKMIAAEAPVLFKACEMFILELTRSWIHTEE  
NKRRTLQRNDIAGAITRGDIFDFLVDIVPRDELKEEDLGVPWTGVPGDGSVPYGGIFYP  
PMAGQQMHSMGAPEMMVGQPPNPQMMYQPPQTAFFVPEQQQQ\*

**FIGURA 4**



**Figura 5**





**Figura 6**

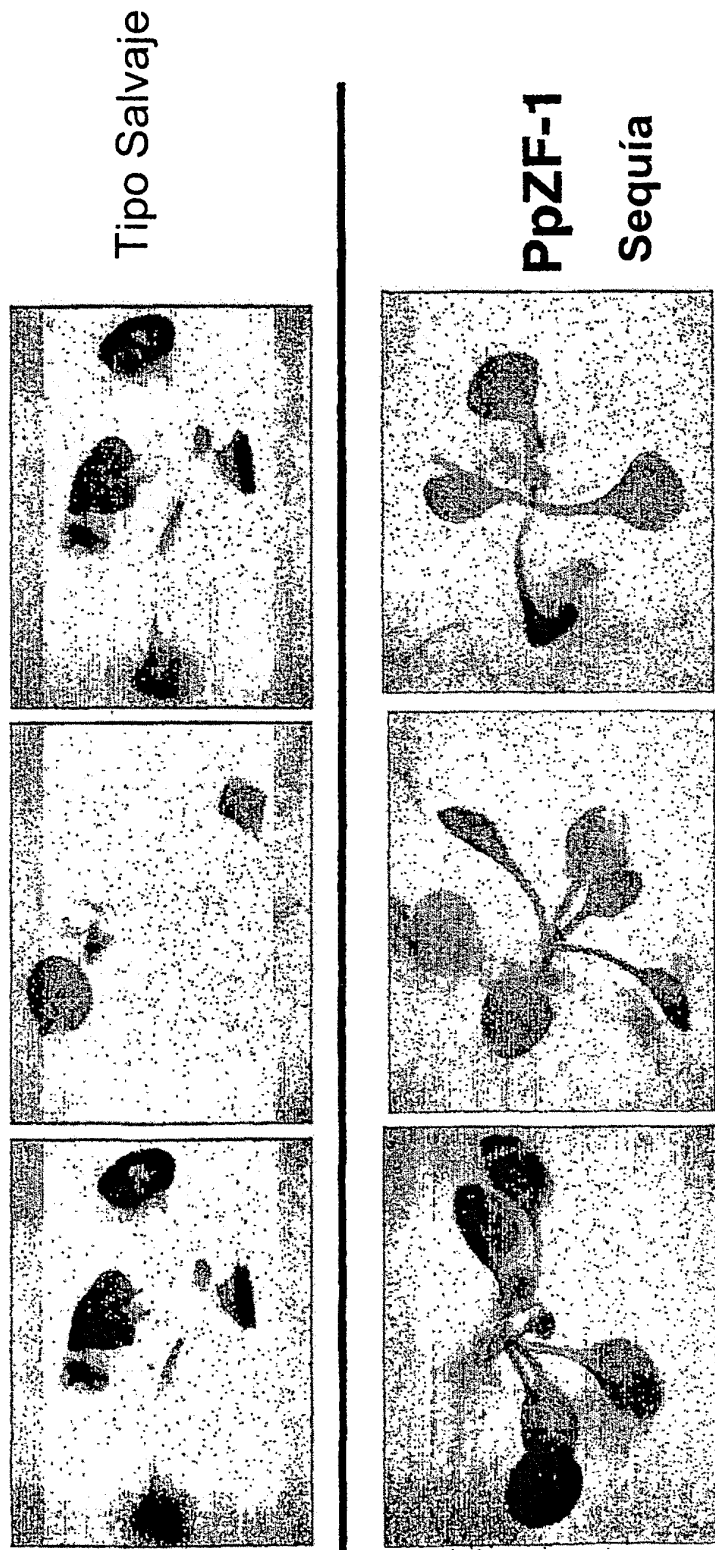


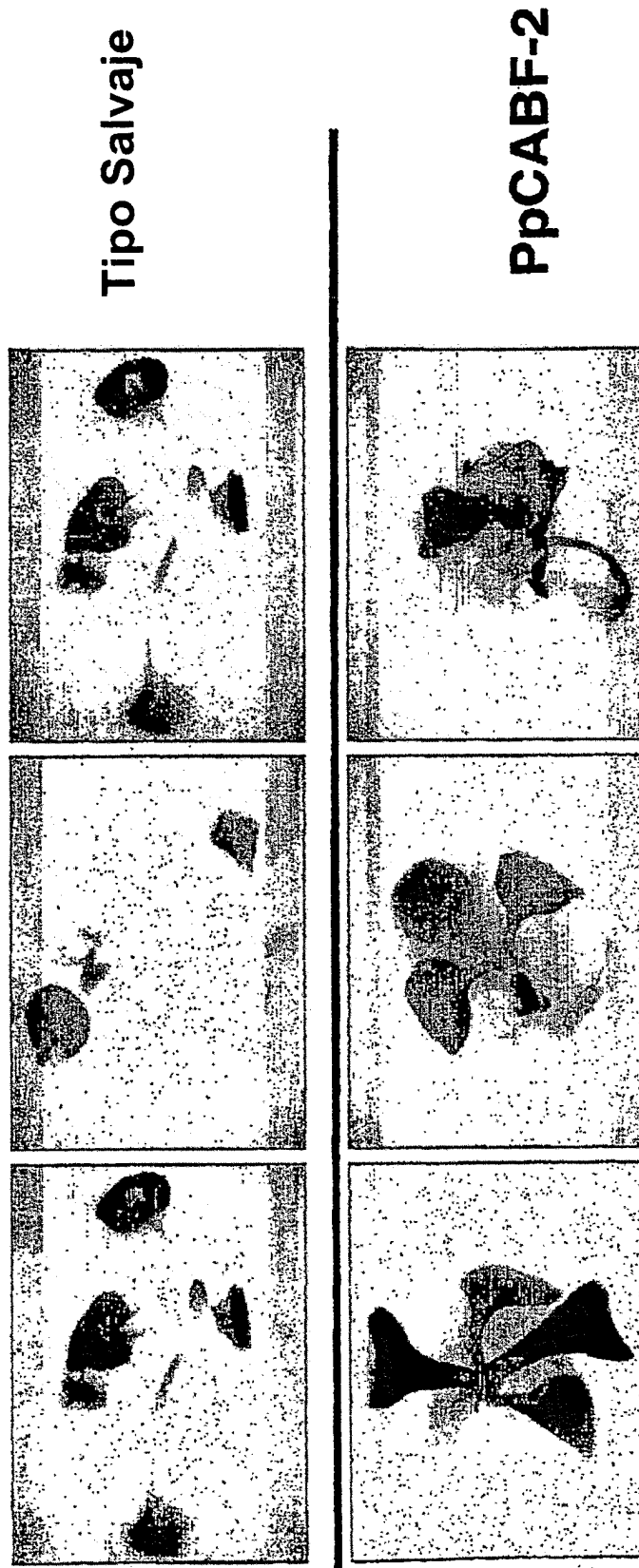
Figura 7



**Figura 8**



**Figura 9**



**Figura 10**

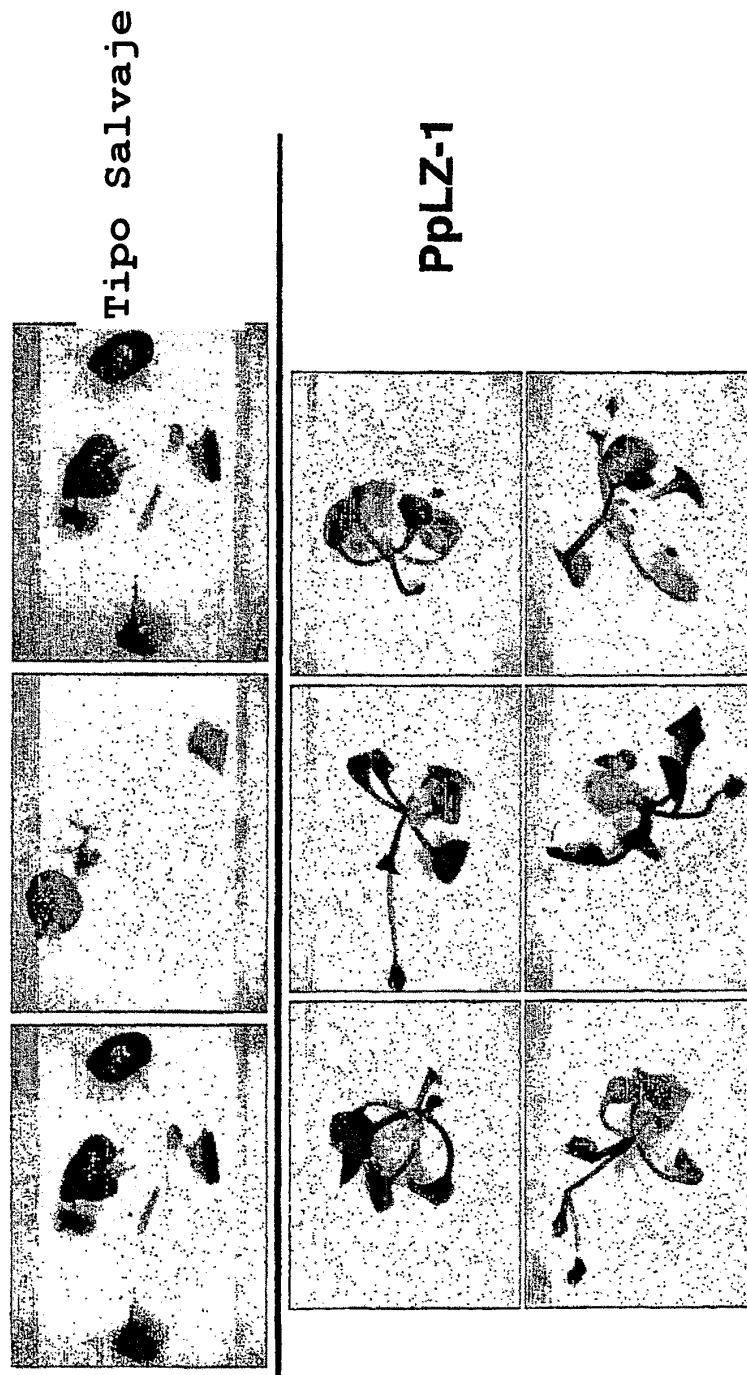
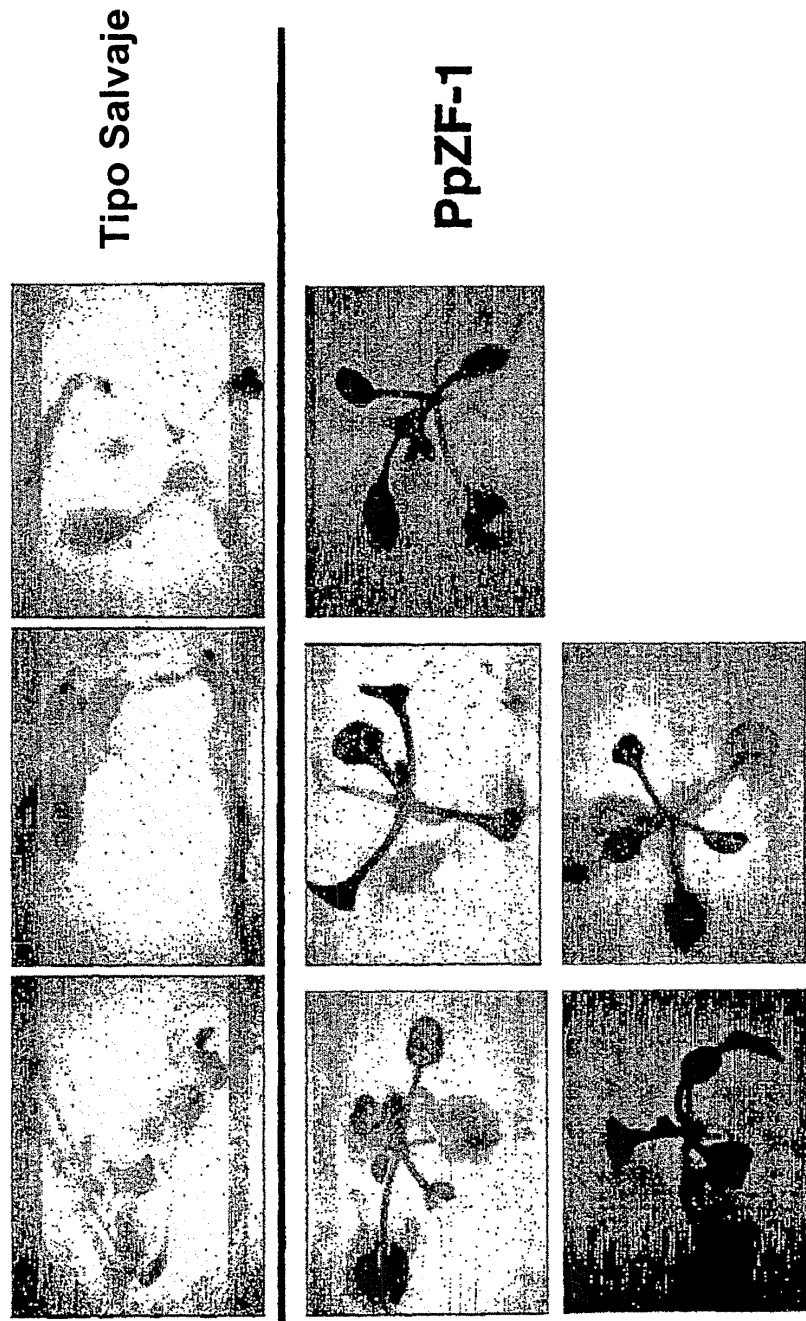


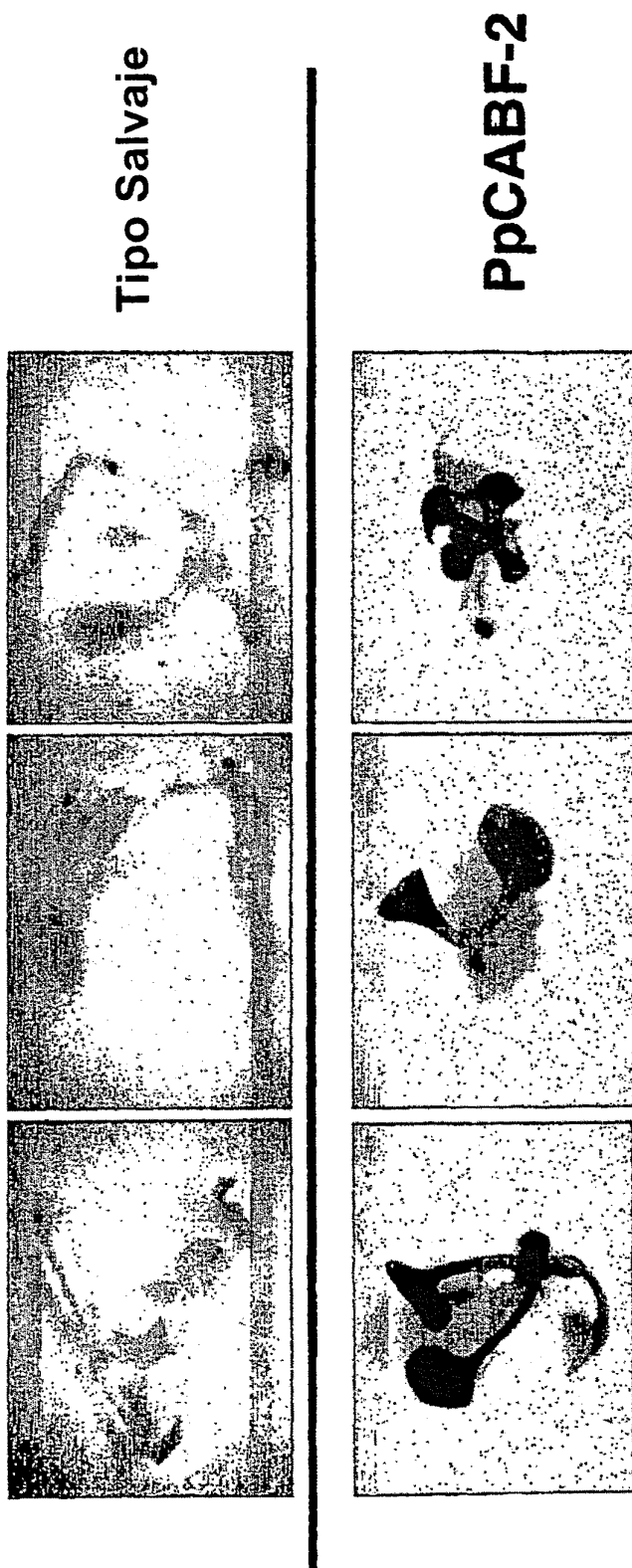
Figura 11



Figura 12



**Figura 13**





**Figura 14**



# ES 2 279 777 T3

## LISTA DE SECUENCIAS

<110> BASF PLANT SCIENCE GMBH

5 <120> PROTEÍNAS RELACIONADAS CON EL ESTRÉS POR FACTORES DE TRANSCRIPCIÓN Y MÉTODOS DE UTILIZACIÓN EN LAS PLANTAS

10 <130> 16313-0005

<140> PCT/US00/34972

<141> 2000-12-22

15 <150> 60/171,745

<151> 1999-12-22

20 <160> 77

<170> PatentIn Ver. 2.1

<210> 1

25 <211> 685

<212> DNA

<213> *Physcomitrella patens*

30 <400> 1

35 **gcaccagcga atccgtctcc gcctcegcct tctgcacgcg tggttgtggg cgacctctcg 60**  
**cgggagcaac aggaaactaa tcccttttcc agcactaaac gattgaagca atttttttt 120**  
**tcttgtgaac tgctcaactc ctctctgtta tgaggggatt cgaagcttga aagttatgag 180**  
**ctgaaggttg aggacacgta agcaccagag gacgatcata ctacaattaa ccettgcggg 240**  
**gaaaagccca ggcaaaatag gacggatggc cgacagttac ggccacaacg caggttcacc 300**  
**cgagagcagc ccgcattctg ataacgagtc cggcggccat taccgtgatc aggacgett 360**  
40 **tgtacgggag caagaccggg ttttgcccat cgcaaagtg agccgaatca tgaagaaagc 420**  
**attgccatct aatgcgaaga tatcgaaaga cgccaaagag actgtgcagg agtgcgtatc 480**  
**cgagttcatc agtttcatta ctggtgaggc gtccgacaag tgtcagaggg aaaagaggaa 540**  
**gacgatcaac ggggatgact tgctgtgggc catgagtact cttggttttg aagattatgt 600**  
**ggaacctctg aaggtgtacc tacacaagta tctgtaactg gagggggaga aggcctctat 660**  
45 **ggccaagggg ggtgatcaag cagggg 685**

<210> 2

50 <211> 517

<212> DNA

<213> *Physcomitrella patens*

55 <220>

<221> base modificada

<222> (1) .. (517)

<223> "n" representa a, t, c, g, otro o desconocido

60

65

## ES 2 279 777 T3

<400>

```

5      ggcacgaggc ttgatgatga tcatgcacta gcttctgcaa agtgccaggc tttagcacgt 60
      ctacttccca agttacagca aggtggccat cgcacattga tattcagcca gtggacaagc 120
      atgctggata ttttagaatg ggctcttgac gtcacgggtt tttcttacac tcgcctagat 180
      ggaagcactc aagtaagtga acgccaacc ctagtggacg agttcaaca tgaccctagc 240
      atatttgtgt ttctcctgtc tactcgagct ggagggtcaag gtctaaattt aacaggagca 300
      gacacagtca ttttacatga tttggacttc aatccccaaa tggatcgaca ggctgaggat 360
10     cgctgtcatc ggattggcca gtctaaacct gttacgatat accgacttgt aacaaaagat 420
      acggtcgatg aaagtatata caagatagcc caacagaagc tggtcctcga tgcggcagtt 480
      cttgaangaa aagatcattc ctctgatctt aatgatg                                     517

```

15 <210> 3

<211> 1000

<212> DNA

<213> *Physcomitrella patens*

20

<400> 3

```

25     ggcacgagct gatactaatt gcacgaggtt ttctcaaatg tgtttttggg tacaggaagg 60
      tggaggggaa tttgtaagtg acagagacgt ggtgggagtg ggaggagtgt gaggaatcga 120
      gctagcacct aaaggatttc ggggtgaagg aagggtgcgat tgaaggcgtg catgaaattt 180
      tgacgcagcg ggttgaatcg gaaggagttt tcagcagctg gaaagtacct tcgagggttg 240
      attcatccaa agtttccatg tgatatggct tcaaagtatc cgcgaaaatg tagagagcac 300
30     gcatcacctg gagttgggtg caggaatct acgcacgct ttgattcaag gtcgcagtcg 360
      tattcgttct cggagaaacc ttaccacctg cgtcgccggg atgctgttcg tgatatgatc 420
      agcgatttgg tgcacagcc ttccgacct gccgtgcctg gtttcagggg agtgcgctat 480
      cgtcagaaac tgaacaagta cgtgacagag attcggccca cgagggtctc gaagaaaatt 540
      tggcttggga cgtatgacac tgcagaggaa gcagcgcctg cctttgacat cggaaatttg 600
35     tgttgtaaga aaaacctgcc getcaacttt ccggattcga ctcagatgtt gcagagaatc 660
      tcgtcgaat tgacccccga agcgaacga aaagccatcg cgacgctggc gaaggacgta 720
      gtgcgaatgg aaaatgacag gtcgaagtgg ggtggcggtg acctgactac cacagagccc 780
      ccggtccatt ccgagcctat tactcaacac cttgcagcag ctgagattcg cgcggtcacg 840
      tacattgaac agcccctgga aattgtctac ggagtggaag aatcggcgac ggccatgtcg 900
40     gtaacggaag caaatgcgcg cgataaccac tcttgagttt gggacttggg caaagtgatc 960
      cttgatgacg agctctctga aattcetaac ttcgtcggag                                     1000

```

45 <210> 4

<211> 1343

<212> DNA

<213> *Physcomitrella patens*

50

<220>

<221> base modificada

<222> (1) .. (1343)

55 <223> "n" representa a, t, c, g, otro o desconocido

60

65

ES 2 279 777 T3

<400> 4

```

5      ggcacgaggg caagagggga tagagacttg aaaggaagg gaggggaaggg tgtaaggagg 60
      cccacgggca ggggtcaaggt gtccaatgca cctgcaagat caggaagcct gaagtagatc 120
      agggaaaaaa cgatggtagt ccctagttta cccgccttcg gaggacagaa cgccatgctc 180
      agacgcaaca ttgacaacaa caccgacact ctgattttctc tgcttcaagg gtccctgctcc 240
      cctcgcgtga gcatgcaaca aggatatgtt gcagtgccgc gttcatcgga gagtctcgaa 300
      aacatgatgg gggcttgtgg gcaaaaactg ccttactttt cgtcatttga tgggcccagt 360
10     gtagaagagc aagaggatgt cgacgaaggt atcgacgaat tcgcacacca cgtgggaaaa 420
      aaggaggaga ttgtcattag aacaagtgcg atcattagaa cgggaattttg aagtggaaaa 480
      caagcttgag cccgagagga aaatgcaact agctaaggag cttggactgc aacctcgtca 540
      agtggcgtg tggttccaga atagacgggc aaggtggaaa accaaacagc tcgagcgcga 600
      ctacgagacc ctgaagaaag cctacgacag gcttaaagca gacttegaag ccgttactct 660
15     agacacaagt gctcttaaag ctgaggtgag tcgcctcaag ggaatctcta atgcacgacg 720
      tcaagcccgc cgaattcgtt cagggcaagt gtgacacaac gagtccoct gcctcccctg 780
      cgcaatcgga gaggtccgac atttgtgcat cgaggaatcg cacaactcct accatacatg 840
      tggatcccgt ggcacccgag gaagccggcg ctcaactaac catgagctcg gatagcaatt 900
      ccagtcaggt cctggacgct gatagccctc gcacgagcca caccagecgt agtaggagca 960
20     ctttgctcac aagtgtggtg cagcctgacg agggcctggg agtggcccag taeccccact 1020
      tttctcccga aaacttcgtg ggtcccata tggcagagat ttgcgctgat cagtcatttg 1080
      catctcaagt gaagctggaa gagatccaca gcttcaatcc cgaccaaacc ttctgctct 1140
      tgcccaactg gtgggattgg gcttgattcg tttcttcac tgtaccata cactttttcc 1200
      ttgaatccaa gttgaattca ctttaggcag tgtttttca cgatgtacca cttgttattc 1260
25     ttccaccatg tgcaatccaa cgtaancaa agttgcacat cggcgatcat tggtgacgat 1320
      gtcgagcatc gatcgtcaca tgc                                     1343

```

<210> 5

30 <211> 1157

<212> DNA

<213> *Physcomitrella patens*

35 <400> 5

```

40     gcacgagctc ggttgtggaa gctgtctcgt ggcttcttcc gcaccctaag atctcgacca 60
      actctattat cagagggcagc gctgcagccg acgagatggg ttcgctgcct ttccacgacc 120
      ggcccttag tcccaagccc aagaaacaga aggttgaatt gcccgcggac atattctctg 180
      atgtggacc tttcctagac ttggacgatg caaccgttac cggaaatcaa cccgacagct 240
      tggtagcagt ccataatgcca gaatgctccg aggacacgga ttcgcttgc cactccatgg 300
      accctcgtt tactaaattt cctctctcgg cgaagagcgg ttacagctat ggcacatcta 360
45     cccttactca gagcatttct tgttcgtctc tagatgccgc cgttgttcca gactccagt 420
      tcagcgacat ttccacacc tacctagact cacaagctc ccaagatag tcagctcgc 480
      tgccacacca gactggaggt ccattgaca ccgtcgaccg tgaagctcgc gtgttcgct 540
      acaaggagaa gaggcagaag cgcaagtgtg agaaaacaat tcgctatgca tcaaggaagg 600
      catatgctga gagccggccg aggatcaaag gaaggttcgc taagagaact gattccgaca 660
50     tggagcagtt tggctcagtg gactcaagtt tcggagtggg tccaagttt tgagtttct 720
      tgtgtattgg agtctccatc gagcaaggtc atctgaaatg gaaagctgct gtgtaacata 780
      gaggagctgc tgtaagaact gtgtagagcc atccaagtgg tgaagcacct gaaaaagtgg 840
      cagcaatgta aattgttcag actctcaatg gtcaccagta ccaagtcag ccattctata 900
      atccctttca gaacacgatt aaatgccttg tggacagtac aggatgtagt cagagtcta 960
55     gtagtggttt ttttctattt ttcttttgg tgaattgagag ctttcggaac ggtgagaact 1020
      tcgtggcgcg aatcctctgt cctgcgatcg ttatgatgca gcgaattctt ccgatcttga 1080
      tgaatttcaa cacttcata atgctcttgg atttttgggt catttctca gaagtggtg 1140
      agctaacaaa aaaaaaa                                     1157

```

60 <210> 6

<211> 374

<212> DNA

65 <213> *Physcomitrella patens*

ES 2 279 777 T3

<400> 6

```

5      cttcaaagaa ctcggcaaca gcagggctca taccttggac ctggttcgta cagtgaccaa 60
      aatggtcagt cgggtggagt tgggtggagca aacgcatata gttcaggagc tgctgcattt 120
      gacctggagt atgcaaggtg ggttgaagat catacccggc agatgagtga gctccgggtg 180
      gccctacagg ctcatgtcgc tgacgctgat ttacgattac tagtggatgg gagtatggcc 240
      cactatgaag acctctttcg gctcaaggac gctgctgcaa aagccgacgt gtttcatctc 300
      gtgtccggca tgtggaaaac tcctgcagag cgatgctttg tatggattgg aggctgccga 360
10     ccctctgagt tact                                     374

```

<210> 7

<211> 461

<212> DNA

15 <213> *Physcomitrella patens*

<400> 7

```

20     catcctggcg ccgctcatgcc tttacagatg cactaccgac aagcccagca acagatgatg 60
      ccgcagcttg gtgatcagca gatgcagccg cagcttcatt atcagcaaat tcagaaacag 120
      cagctgtccc agttctggca gcagcaaatg caggaaatgg agcaagtcaa tgattttaag 180
      acccatcagc taccactggc acgcatcaaa aaaatcatga agtcggatga agatgttaag 240
25     atgatcgcag ccgaagetcc agtgctgttt tcaaaagctt gtgagatgtt tattttagaa 300
      ttgacactgc gctcttggat tcatacggag gaaaataagc gaaggacact acaaagaaat 360
      gatattgcag gggctatcac taggggagac atcttcgact ttcttgttga catcgttcca 420
      cgtgacgagt tgaaggaaga agatttgggt gtgccatgga c                                     461

```

30 <210> 8

<211> 1347

<212> DNA

<213> *Physcomitrella patens*

35

<400> 8

```

40     ggcaccagcg aatccgtctc cgcctcogcc ttctgcacgc gtggttgtgg tcgacctctc 60
      gccggagcaa caggaaacta atcccttttc cagcactaaa cgattgaagc aatttttttt 120
      ttcttgtgaa ctgctcactc tctctctgtt atgaggggat togaagcttg aaagtatatga 180
      gctgaagggt gaggacacgt aagagcgaag gacgatcata ctacaattaa cccttgccgg 240
      gaaaagccca ggcaaaatag gacggatggc cgacagttac ggccacaacy caggttcacc 300
      cgagagcagc ccgcattctg ataacgagtc cggcggccat taccgtgatc aggacgcttc 360
45     tgtacgggag caagaccggt ttttgcccat cgcaaagtgt agccgaatca tgaagaaagc 420
      attgccatct aatgcgaaga tatcgaaga cgccaaagag actgtgcagg agtgcgtatc 480
      cgagttcatc agtttcatc ctggtgagyc gtccgacaag tgtcagaggg aaaagaggaa 540
      gacgatcaac ggggatgact tgctgtgggc catgagtaact ctgtgttttg aagattatgt 600
      ggaacctctg aaggtgtacc tacacaagta togtgaactg gagggggaga aggcctctat 660
50     ggccaagggg ggtgatcagc agggaggaaa agagagcaac caaggaggta tggggtcgat 720
      gggcatggca ggcggaatca acggcatgaa cggaaacgatg aacgggaaca tgcattggca 780
      tgggaattccc gtatcgatgc agatgatgca acagccgtac gcgcagcagg cacctccggg 840
      gatgatatat tctcctcatc aaatgatgcc gcaataccag atgccgatgc agtctgggtg 900
      aaaccaaccc cgcggagtgt aagagttttc actggcagga ggctttggaa gtggggatat 960
55     tgtcgacagc gtgatggggt gttttggagc atgggcaggg cattatgggtg ctgttgaaac 1020
      agtgatgggt gggatcatgtg aagtgttggc gactgttgaa tgatgaaaac atagaagtga 1080
      tgtcgttgaa gctcggggag tttcaagtga aaggaggagc actttttgtt tggaaaggag 1140
      cgtaccgggt ctggcagtgt acattctgaa tgatagttat ctgtgctgat ttttcttggc 1200
      cttggcaata cgaggggggt gaatattttg ctttgaattc gttgacattt caaccttttc 1260
60     tatgtgaaaa ggctctgtag gatgcaagat aaggaaagac atgcagattg ataaaaaaaa 1320
      aaaaaaaaaa aaaaaaaaaa aaaaaaa                                     1347

```

<210> 9

<211> 2838

65 <212> DNA

<213> *Physcomitrella patens*

ES 2 279 777 T3

<400> 9

5 ggccttatcc cgggcgatgg tgcgttcgag atcgtaaagg tgccgacgaa ggcgtaactt 60  
 ggaagtcctc tgtgtcccgg cgatgtccca atgttggccc gattttctgt ttttagcgag 120  
 ctgtgggcta gtttgtgggt atgatccggg gaatgagacg agatgtctgt ctgagtgaga 180  
 ccactctagg ggctgttggg ggatgaggag ggaagcgcag aagtggcca ttcttttcag 240  
 tgactggact ctgtgcgagt ggtcagcttt cgggagctgc tgttgcaattg accggtgatt 300  
 ctttcgagat cgtagagaca gcagctggca agggttttgg gaggctttc aaatgaaggg 360  
 10 cattcaagag ctttcagatg atgaagatta tattccgct gtcaatgcat cgcgatattt 420  
 caacaggggc aaagcgcctc caaagacatc atccaatcat gccaatggaa atggaaatcc 480  
 aaacggaacg agttttggag tttcaacttc ttcagcaagt gactctgaca aagataagaa 540  
 atccgaagt tccaggttctt tactaagcga ttctggcaag aatcaaaagt ccgttactga 600  
 attggattcg ttcgcattta accgcaagtc cagaattgcc aagecgaccta tcgagctact 660  
 cgaagacgag gaggacgtgg acgttggagc tgcaaagggt gtagacattg agccgactaa 720  
 cggaaacaag agcgggagac ggatcacac catcgaagc agtgacgatg aagagttgga 780  
 tgagaagaaa tgcgttgggtg ataactctgac cccaggaacg gaaatcgcac aatgtgcagc 840

20 cgatgaatcc ttagcaaaaa gggtgcagga ttagagcac cgggcagttt ctggccgtaa 900  
 tcgectgggt caaattttgt cagattccga tgaagaagaa gaggaagaag taaatcccat 960  
 aaccatcacc ctacaaagggt gtgaccagat tgcaatca ttgagagaag agctgcaggc 1020  
 cagcagttca agtgataact cggttaatga agatcgttat gcagaggtt atgtagcagc 1080  
 agcaaaaatt gtgagccagg cagatgtctg tgcagcttgt ggcattgccg agaatgatac 1140  
 acaacgaatg ctcaagccat atcagcttgt aggcgtcaat ttcatgctgc tacttcaccg 1200  
 25 caaacatggt ggggggtgcag ttgcgtatct tgccctctg aaacatctt atggagatgc 1260  
 tggctcctcat cttttagttg cacctgcttc tcttttagaa aactggcaaa gagaactcaa 1320  
 gaagtgggtg cctgcattta aggtggagct ctatcatggc tcaggaaggg cagctttaa 1380  
 caggaggctt cagtatgctg caaaatctaa aggcctgca cccttaacg tgatgtgac 1440  
 gtgctactcc ctttttgaga ggcagagtgc tcagacaaaa gatgaccgca aattccttaa 1500  
 30 gaaatggaat tggcgtctg tggttatgga cgaggtcat cttttgagg acagaagcag 1560  
 ctttcgcagc aaaaagttgc gagatatagc tcacaaagca attcaaagc tgatgctgac 1620  
 tggactcca ctccagaacg atttgcagga gctatggtea cttctggagt tcatgatgcc 1680  
 tgatgtgttc aacacaaatg gcgttgattt agatcaatat ctgggaacca ggaacgatac 1740  
 35 ctccagggatt gttgtgcagg atacgaactt gatgactcgg atcaaaaggaa tactaggacc 1800  
 ttttgtatta cggagaatga aaactgatgt tatgcgccag cttgtatcaa agattcagga 1860  
 ggtggagtgt gtggagatgc tagacgagca atcaatggca tataaaaaag ctgtgaatga 1920  
 gtatagagcc cttgctgagt ccgcacgtgc cgctaaagct gcaaagaaat cctcagttag 1980  
 cgtagtagat gtccttctc gtgcacaagt gaccaatatac tttactcaat tgagaaagg 2040  
 40 caagaaatg gctaagaaat ttcattccatt aggagttttt ggatatgaat gcgatttga 2100  
 gcgtgtggag gaagaattga ctagttagc cgattttgac ctccacaagt tgtgtattca 2160  
 atatggaggc gctgcgggag ggcaaggaaa gcttgatgat gatcatgac tagcttctgc 2220  
 aaagtgccag gctttagcac gtctacttcc caagttacag caagtggee atcgcacatt 2280  
 gatattcagc cagtggacaa gcatgctgga tatttttagaa tgggctctt acgtcatggg 2340  
 45 ttttcttac actcgcctag atggaagcac tcaagtaagt gaacgcaaa ccctagtgga 2400  
 cgagttcaac aatgacccta gcatatttgt gtttctcctg tctactcgag ctggaggcca 2460  
 aggtctaaat ttaacaggag cagacacagt cattttacat gatttgact tcaatccca 2520  
 aatggatcga caggctgagg atcgtgtca tcggattggc cagtctaac ctgttacgat 2580  
 ataccgactt gtaacaaaag atacggtcga tgaaagtata tacaagatag cccaacagaa 2640  
 50 gctggctctc gatgcggcag ttcttgaagg aaaagagtca tctctgatc ttaatgatgg 2700  
 tgatgctcgc acgatgggtg aaattctttc tgcattattg gatgtccac cgcacatgatc 2760  
 ctggagtcca gaacacattt ttaatttatt ttcattatct ttatctggca ctgcgagaaa 2820  
 gctcgttaac gcaagggc 2838

55 <210> 10  
 <211> 1424

60 <212> DNA  
 <213> *Physcomitrella patens*

65

ES 2 279 777 T3

<400> 10

```

5   ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta 60
    cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcgataaca atttcacaca 120
    ggaaacagct atgaccatga ttacgccaaag ctcgaaatta accctcacta aagggaaaca 180
    aagctggagc tccaccgocg tggcggccgc tctagaacta gtggatcccc cgggctgcag 240
    gaattcggca ccagaggatt tcgggggtgaa ggaaggtgoc attgaaggcg tgcataaat 300
    tttgacgcag cgggttgaat cggaaaggagt tttcagcagc tggaaaytac cttcgagggt 360
    tgattcatcc aaagtttcca tgtgatatgg cttcaaagta tccgcgaaaa tgtagagagc 420
10  acgcatcacc tggagttggt ggcagggaaat ctacgcacgc ctttgattca aggtcgcagt 480
    cgtattcgtt ctccggagaaa ccttaccacc gtcgtcgcgc ggatgcgctt cgtgatatga 540
    tcagcgattt ggtgcatcag ccttccgaca ctgcctgccc tggtttcagg ggagtgocgt 600
    atcgtcagaa actgaacaag tacgtgacag agattcggcc cacgaggtgc tccaagaaaa 660
15  tttggcttgg gacgtatgac actgcagagg aagcagcgcg tgcctttgac atccgaaatt 720
    tgtgtttaa gaaaaacctg ccgctcaact ttccggatcc gactcagatg ttgcagagaa 780
    tctcgtcga aattgaccccc gaagcgaac gaaaagccat cgcgacgctg gcgaaggagc 840
    tagtgcgaat ggaaaatgac aggtcgaagt tgggtggcgg taacctgact accacagagc 900
    ccccggtcca ttccgagcct attactcaac accttgcagc agctgagatt cgcgcggtca 960
20  cgtacattga acagcccctg gaaattgtct acggagtggg agaatcggcg acggccatgt 1020
    cggtaacgga agcaaatgcg cgcgataacc actctggag ttgggacttg ggcaaatgta 1080
    tccttgatga cgagctctct gaaattccta actctcgcg agaactagat cacgaggcta 1140
    tggatttcag tagtcatgga gaggtttact accaccatta tgactctcag tgagtcctac 1200
    aagcatattt tcaactagtc aacatcctca gtagatttag tccattactt tctgtgtcag 1260
25  agccacgcct gcggttaga ccgggaaagc ttgtataaac tgtaaatga gctctcgtag 1320
    acatgatgta acaccaatc acctgtaaac cccccagctt gagatcaca ggagtagaaa 1380
    acctgatagc ttcaagagtt tcaacccaaa aaaaaaaaaa aaaa 1424

```

<210> 11

30 <211> 1318

<212> DNA

<213> *Physcomitrella patens*

35 <400> 11

```

    gcccttatcc cgggcacgag ggcaagaggg gatagagact tgaagggaaa ggggagggaa 60
    ggggtgaagg aggccacgag gcagggtcaa ggtgtccaat gcacctgcaa gatcaggaag 120
    cttgaagtag atcagggaaa aaacgatggt agtccctagt ttaccgcctc tcggaggaca 180
40  gaagcccatg ctcagacgca acattgaca caacaccgac actctgattt ctctgcttca 240
    aggttccctg tcccctcgcg tgagcatgca acaagtgccg cgttcatcgg agagtctcga 300
    aaacatgatg ggggcttgtg ggcaaaaact gccttacttt tcgtcatttg atgggcccag 360
    tgtagaagag caagaggatg tcgacgaagg tctcagcaga ttccgcacacc acgtggagaa 420
    aaagaggaga ttgtcattag aacaagtgcg atcattagaa cggaaatctg aagtggaaaa 480
45  caagcttgag cccgagagga aaatgcaact agctaaggag cttggactgc gacctcgtca 540
    agtggcggtg tggttccaga atagacgggc aaggtggaaa accaaacagc tcgagcacga 600
    ctacgagacc ctgaagaaag cctacgacag gcttaaagca gacttcgaag ccgttactct 660
    agacacaaat gctcttaaag ctgaggtgag tgcctcaag ggaatctcta atgacgagct 720
    caagcccgcc gaattcgttc agggcaagtg tgacacaacg agtcaccctg cctccccctg 780
50  gcaatcggag aggtccgaca ttgtgtcctc gaggaatcgc acaactccta ccatacatgt 840
    ggatcccgtg gcacccgagg aagccggcgc tcaactaacc atgagctcgg atagcaattc 900
    cagcagaggtc atggacgctg atagccctcg cacgagccac accagcgtca gtaggagcac 960
    tttgtccaca agtgtggtgc agcctgacga gggcctggga gtggccagat acccccactt 1020
55  ttctcccgaa aacttcgtgg gtcccaatat gccagagatt tgcgctgatc agtcaactgc 1080
    atctcaagtg aagctggaag agatccacag cttcaatccc gaccacacct cctctctctt 1140
    gcccaactgg tgggattggg cttgattcgt ttctcatct gtaccatac acttttctct 1200
    tgaatccaag ttgaattcac tttaggcagt gtttttcaac gatgtaccac ttgttattct 1260
    tccaccatgt gcaatccaac gtcaacccaaa gttgcaccat cggcggttaac gcaagggc 1318

```

60 <210> 12

<211> 1341

<212> DNA

65 <213> *Physcomitrella patens*

ES 2 279 777 T3

<400> 12

5 gaggagggag ttggaatcta ggagacgtgc atgtgctgtg ggaggaattc tctggggatt 60  
 tcgaggcctt gttgtatggt gttcagtaaa gggagtagct ttttccactt gaaggggctg 120  
 gtgctgctgt tgttgcaagt cttttgacat taaaagaggc ggggttgca cccccggtgt 180  
 gaggaagagt cttgtagtag gtgggttgtg ttgtgccgtg gtatagtatg ccgaagcctt 240  
 gtgatgcatg ccatgtttcc agcgcggcgg tgttctgccg agcggacgct gcctacctgt 300  
 gcgtaggctg cgatgggaag gtccacgggg ccaacaaact agcgtctcga cacgagcgcg 360  
 10 tgtggatgtg cgaagtgtgc gaggttgctc cagccgttgg gacctgcaag gcggatgcgg 420  
 cttctctctg tgtggcctgt gacacagaca tccactccgc caaccgccta gcgcagcgtc 480  
 acgagagagt gccggtgaca cctctgttcg agagtgcgag tcctttgcgt gggccagatt 540  
 tctgcgtggt ggtgtcagag aatgggtgcc atgatctgct gaagggctgt gaggacgct 600  
 cggttgtgga agctgtctcg tggcttcttc cgcacctaa gatctcgacc aactctatta 660  
 15 tcaagggcag cgtgcagcc gacgagatgg gttcgtgcc tttccacgac cggcccttta 720

gtcccaagcc caagaaacag aaggttgaat tgcccgcgga catattctct gatgtggacc 780  
 ctttccctaga cttggacgat gcaaccgta ccggaattca acccgacagc ttggtaccag 840  
 20 tccatagtc agaatgctcc gaggacacgg attcgcttgc tcaactccatg gacccttctg 900  
 ttactaaatt tectctctcg gcgaagagcg gttacagcta tggcacatct acccttactc 960  
 agagcatttc ttgttcgtct ctagatgccg ccggtgttcc agactccagt ctacgagaca 1020  
 ttccacacac ctaccatgac tcacaaagct cccaagatat gtcagctcgc ctgccacacc 1080  
 agactggagg tcccattgac accgtcgacc gtgaagctcg cgtgttgcgc tacaaggaga 1140  
 25 agagccggcc gaggatcaaa ggaaggttcg ctaagagaac tgattccgac atggagcagt 1200  
 agagccggcc gaggatcaaa ggaaggttcg ctaagagaac tgattccgac atggagcagt 1260  
 ttggctcagt ggactcaagt ttcggagtgg ttccaagttt ttgagttttc ttgtgtattg 1320  
 30 gagtctccat cgagcaaggt c 1341

<210> 13

<211> 1521

<212> DNA

35 <213> *Physcomitrella patens*

<400> 13

40 gcccttatcc cgggtgctct ggcagtggga cggatttggga agcaacagga ggtgggcttg 60  
 ttgagctgcy gagtatggaa aaaagcggga aggtgacgtg agagctggaa tgatggccga 120  
 gtgagcgtgt ttgttttgag ggggtaatta gatgggaaga tagaggtcgg atgagtctgg 180  
 gcggtttgcy tagagcgtc gaggaaaagg aaagtggcga ggtgtaggat cttggtggat 240  
 45 ttttctcccc tgaagctaga gacttccggt gcagaatgtg gttaaatgga actcaacagg 300  
 tggaaattcat gacatggaaa cctactgggt cttgttggga atacaatctc acgctgtcgg 360  
 cttctcttta cgtcattttc ttaggttcag agatatagta gaaaggtttg tggaaattatc 420  
 aagatgggtg acaacagtgc aagtgcagg acggattcat cttctgacat ggacggtgat 480  
 gcgaagttgg atgatgggca gcacttagct agtggcggtg gaaactcaaa cgattccagt 540  
 50 ctcgaaactg gaacgaagaa tggcagattct aaggtactaa ggcggttagc acaaaatcgt 600  
 gaggcagccc gaaaaagtag gctcagaaaa aaggcatatg tgcagcagtt ggagtccagc 660  
 cgcataaagc tgaaccaact cgagcaagag cttcaagaa ctccggcaaca gcagggtcta 720  
 taccttggac ctggttcgta cagtgaccaa aatggtcagt cgggtggagt tgggtggagca 780  
 aacgcatata gttcaggagc tgctgcattt gacctggagt atgcaagggtg ggttgaagat 840  
 55 catacccggc agatgagtga gctccgggtg gccctacagg ctcatgtcgc tgacgctgat 900  
 ttacgattac tagtggatgg gagtatggcc cactatgacg acctcttctg gctcaaggac 960  
 gctgctgcaa aagccgacgt gtttcatctc gtgtccggca tgtggaaaac tctcgagag 1020  
 cgatgctttg tatggattgg aggetgccga ccctctgagt tactaaagat attagtaacct 1080  
 caaatagaac ctttgacaga gcagcagttg ttaaaccatct gcaatctgca gcagtcctct 1140  
 60 caacagggty aagaggccct ctctcaaggg atggaacaac ttcagcagtc gcttgcgaa 1200  
 aactgtctg ccggttctct tggctcagca gcaaatgtg ccaactacat gggacagatg 1260  
 gctgtggcca tgggacaact tgggaacctc gaaggtttcg tgcgtcaggc tgatcatttg 1320  
 cgacaacaga cgttacaaca aatgcaccgg gtattaacca ttcgccaagt agcccagga 1380  
 65 cttcttgcca tgggtgatta ctttgcctgt cttcgagctc ttagttctct atggtccgcc 1440  
 aggcctcgtg aatgagaaac attgtcgttt caggcgatgg tgaagtcttc ggcgcaggtg 1500  
 tggaaagcag atcgcaaggg c 1521



ES 2 279 777 T3

<210> 14

<211> 1176

<212> DNA

5 <213> *Physcomitrella patens*

<400> 14

```

10      cagcatcctc acatcccgcc ttcctctgca cccagatgt cgcacctcgg cgccgtcatg 60
      cctttacaga tgcactaccc gcaagcccag caacagatga tgccgcagct tggatgatcag 120
      cagatgcagc cgcagcttca ttatcagcaa attcagaaac agcagctgtc ccagttctgg 180
      cagcagcaaa tgcaggaaat ggagcaagtc aatgatttta agacccatca gctaccactg 240
15      gcacgcatca aaaaaatcat gaagtcggat gaagatgtta agatgatcgc agccgaagct 300

      ccagtgcctg tttcaaaagc ttgtgagatg tttatcttag aattgacact gcgctcctgg 360
      attcatacgg aggaaaataa gcgaaggaca ctacaaagaa atgatattgc aggggctatc 420
      actaggggag acatcttcga ctttcttgtt gacatogttc cacgtgacga gttgaaggaa 480
20      gaagatttgg gtgtgccatg gactgggtgt ccaggggatg gttctgtacc ttacggagga 540
      atattctatc caccatggc tggacagcaa atgcatcatt ctatgggggc tcctgagatg 600
      atgggtgggc agccaccaaa ccacaaaatg atgtatcagc ctccacagac tgcctttgtc 660
      cccgagcagc agcaacagtg aatgcattac cacctagaga acgctgagca tcgaagacgg 720
      gacaactcaa ggaaagggtc atcgcacgca gattctttcg tcacgtggga atggtatctt 780
25      tcatactggt gctaccatct gtcattctta tggcaaaagg ggtcacagga tcaggatttt 840
      acctttcact acagcgtctt tgtgttggct ttcaactata ttttaaggaa atcgtagctg 900
      taggcgggtg tgcgacagt ctgagcactg ctaattctag cagagtttat gtttgggtta 960
      gcaagtcatg aagggcacia agggaccgca cccctccatg gatctggtag aaatttgtga 1020
      atagtgatac tagtgagcgc ataattatta gcatgtgcag gagttgctct taatgttagg 1080
30      ttcgaggatc gggtatccat tttcttagta ccattgttcc ttttatgtct ccctgggttt 1140
      atctttcaga ctgaaaaaaaa aaaaaaaaaa aaaaaa 1176

```

<210> 15

35 <211> 218

<212> PRT

<213> *Physcomitrella patens*

40

45

50

55

60

65

ES 2 279 777 T3

<400> 15

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

<210> 16

<211> 801

50 <212> PRT

<213> *Physcomitrella patens*

55

60

65

ES 2 279 777 T3

<400> 16

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

50

55

60

65

ES 2 279 777 T3

Val Ala Ala Ala Lys Ile Val Ser Gln Ala Asp Val Cys Ala Ala Cys  
 245 250 255  
 5 Gly Ile Ala Glu Asn Asp Thr Gln Arg Met Leu Lys Pro Tyr Gln Leu  
 260 265 270  
 Val Gly Val Asn Phe Met Leu Leu Leu His Arg Lys His Val Gly Gly  
 275 280 285  
 10 Ala Val Ala Tyr Leu Ala Leu Leu Lys His Leu Asp Gly Asp Ala Gly  
 290 295 300  
 Pro His Leu Leu Val Ala Pro Ala Ser Leu Leu Glu Asn Trp Gln Arg  
 15 305 310 315 320  
 Glu Leu Lys Lys Trp Cys Pro Ala Phe Lys Val Glu Leu Tyr His Gly  
 325 330 335  
 20 Ser Gly Arg Ala Ala Leu Asn Arg Arg Leu Gln Tyr Ala Ala Lys Ser  
 340 345 350  
 Lys Gly Pro Ala Pro Phe Asn Val Met Leu Thr Cys Tyr Ser Leu Phe  
 25 355 360 365  
 Glu Arg Gln Ser Ala Gln Thr Lys Asp Asp Arg Lys Phe Leu Lys Lys  
 370 375 380  
 30 Trp Asn Trp Arg Cys Val Val Met Asp Glu Ala His Leu Leu Lys Asp  
 385 390 395 400  
 Arg Ser Ser Phe Arg Ser Lys Lys Leu Arg Asp Ile Ala His Lys Ala  
 405 410 415  
 35 Ile Gln Arg Leu Met Leu Thr Gly Thr Pro Leu Gln Asn Asp Leu Gln  
 420 425 430  
 Glu Leu Trp Ser Leu Leu Glu Phe Met Met Pro Asp Val Phe Asn Thr  
 40 435 440 445  
 Asn Gly Val Asp Leu Asp Gln Tyr Leu Gly Thr Arg Asn Asp Thr Ser  
 450 455 460  
 45 Gly Ile Val Val Gln Asp Thr Asn Leu Met Thr Arg Ile Lys Gly Ile  
 465 470 475 480  
 Leu Gly Pro Phe Val Leu Arg Arg Met Lys Thr Asp Val Met Arg Gln  
 485 490 495  
 50 Leu Val Ser Lys Ile Gln Glu Val Glu Cys Val Glu Met Leu Asp Glu  
 500 505 510  
 Gln Ser Met Ala Tyr Lys Lys Ala Val Asn Glu Tyr Arg Ala Leu Ala  
 55 515 520 525  
 Glu Ser Ala Arg Ala Ala Lys Ala Ala Lys Lys Ser Ser Val Ser Val  
 530 535 540  
 60  
 65

ES 2 279 777 T3

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

Thr

55 <210> 17  
 <211> 268  
 <212> PRT  
 <213> *Physcomitrella patens*

60

65

ES 2 279 777 T3

<400> 17

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

<210> 18

<211> 339

<212> PRT

<213> *Physcomitrella patens*

65

ES 2 279 777 T3

<400> 18

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

<210> 19

ES 2 279 777 T3

<211> 358

<212> PRT

<213> *Physcomitrella patens*

5

<400> 19

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

50

55

60

65



ES 2 279 777 T3

Phe Thr Lys Phe Pro Leu Ser Ala Lys Ser Gly Tyr Ser Tyr Gly Thr  
 225 230 235 240  
 5 Ser Thr Leu Thr Gln Ser Ile Ser Cys Ser Ser Leu Asp Ala Ala Val  
 245 250 255  
 Val Pro Asp Ser Ser Leu Ser Asp Ile Ser Thr Pro Tyr Leu Asp Ser  
 260 265 270  
 10 Gln Ser Ser Gln Asp Met Ser Ala Arg Leu Pro His Gln Thr Gly Gly  
 275 280 285  
 15 Pro Ile Asp Thr Val Asp Arg Glu Ala Arg Val Leu Arg Tyr Lys Glu  
 290 295 300  
 Lys Arg Gln Lys Arg Lys Phe Glu Lys Thr Ile Arg Tyr Ala Ser Arg  
 305 310 315 320  
 20 Lys Ala Tyr Ala Glu Ser Arg Pro Arg Ile Lys Gly Arg Phe Ala Lys  
 325 330 335  
 25 Arg Thr Asp Ser Asp Met Glu Gln Phe Gly Ser Val Asp Ser Ser Phe  
 340 345 350  
 Gly Val Val Pro Ser Phe  
 355  
 30 <210> 20  
 <211> 343  
 <212> PRT  
 35 <213> *Physcomitrella patens*  
 <400> 20  
 40 Met Gly Asp Asn Ser Ala Ser Ala Arg Thr Asp Ser Ser Ser Asp Met  
 1 5 10 15  
 Asp Gly Asp Ala Lys Leu Asp Asp Gly Gln His Leu Ala Ser Gly Gly  
 20 25 30  
 45 Gly Asn Ser Asn Asp Ser Ser Leu Glu Thr Gly Thr Lys Asn Gly Asp  
 35 40 45  
 50 Ser Lys Val Leu Arg Arg Leu Ala Gln Asn Arg Glu Ala Ala Arg Lys  
 50 55 60  
 Ser Arg Leu Arg Lys Lys Ala Tyr Val Gln Gln Leu Glu Ser Ser Arg  
 65 70 75 80  
 55 Ile Lys Leu Asn Gln Leu Glu Gln Glu Leu Gln Arg Thr Arg Gln Gln  
 85 90 95  
 60 Gln Gly Leu Tyr Leu Gly Pro Gly Ser Tyr Ser Asp Gln Asn Gly Gln  
 100 105 110  
 Ser Gly Gly Val Gly Gly Ala Asn Ala Tyr Ser Ser Gly Ala Ala Ala  
 115 120 125  
 65

ES 2 279 777 T3

Phe Asp Leu Glu Tyr Ala Arg Trp Val Glu Asp His Thr Arg Gln Met  
 130 135 140  
 5 Ser Glu Leu Arg Val Ala Leu Gln Ala His Val Ala Asp Ala Asp Leu  
 145 150 155 160  
 Arg Leu Leu Val Asp Gly Ser Met Ala His Tyr Asp Asp Leu Phe Arg  
 165 170 175  
 10 Leu Lys Asp Ala Ala Ala Lys Ala Asp Val Phe His Leu Val Ser Gly  
 180 185 190  
 Met Trp Lys Thr Pro Ala Glu Arg Cys Phe Val Trp Ile Gly Gly Cys  
 195 200 205  
 Arg Pro Ser Glu Leu Leu Lys Ile Leu Val Pro Gln Ile Glu Pro Leu  
 210 215 220  
 20 Thr Glu Gln Gln Leu Leu Asn Ile Cys Asn Leu Gln Gln Ser Ser Gln  
 225 230 235 240  
 Gln Gly Glu Glu Ala Leu Ser Gln Gly Met Glu Gln Leu Gln Gln Ser  
 245 250 255  
 25 Leu Ala Glu Thr Leu Ser Ala Gly Ser Leu Gly Ser Ala Ala Asn Val  
 260 265 270  
 Ala Asn Tyr Met Gly Gln Met Ala Val Ala Met Gly Gln Leu Gly Asn  
 275 280 285  
 30 Leu Glu Gly Phe Val Arg Gln Ala Asp His Leu Arg Gln Gln Thr Leu  
 290 295 300  
 35 Gln Gln Met His Arg Val Leu Thr Ile Arg Gln Val Ala Arg Gly Leu  
 305 310 315 320  
 40 Leu Ala Met Gly Asp Tyr Phe Ala Arg Leu Arg Ala Leu Ser Ser Leu  
 325 330 335  
 Trp Ser Ala Arg Pro Arg Glu  
 340  
 45 <210> 21  
 <211> 214  
 <212> PRT  
 50 <213> *Physcomitrella patens*  
 <400> 21  
 55 Met Ser His Pro Gly Ala Val Met Pro Leu Gln Met His Tyr Pro Gln  
 1 5 10 15  
 Ala Gln Gln Gln Met Met Pro Gln Leu Gly Asp Gln Gln Met Gln Pro  
 20 25 30  
 60 Gln Leu His Tyr Gln Gln Ile Gln Lys Gln Gln Leu Ser Gln Phe Trp  
 35 40 45  
 65

ES 2 279 777 T3

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

35 <210> 22  
 <211> 2838  
 <212> DNA  
 40 <213> *Physcomitrella patens*  
 <400> 22

45 gcccttatcc cgggcgatgg tgcgttcgag atcgtaaggt tgccgacgaa ggcgtaactt 60  
 ggaagtccctc tgtgtcccgg cgatgtccca atggtggccc gattttctgt ttttagcgag 120  
 ctgtgggcta gtttgtgggt atgatccggg gaatgagacg agatgtctgt ctgagtgaga 180  
 ccactctagg ggctgttggg ggtgaggag ggaagcgcag aagttggcca tcttttcag 240  
 tgactcgact ctgtgagagt ggtcagcttt cgggagctgc tgttgattg accggtgatt 300  
 ctttcgagat cgtagagaca gcagctggca agggttttgg gaggcttttc aatgaaggg 360  
 50 cattcaagag ctttcagatg atgaagatta tattccgct gtcaatgcat cgcgatattt 420  
 caacaggggc aaagecgtct caaagacatc atccaatcat gccaatggaa atggaaatcc 480  
 aaacggaacg agttttggag tttcaacttc ttcagcaagt gactctgaca aagataagaa 540  
 atccgaagtt tcaggttctt tactaagcga ttctggcaag aatcaaaagt ccgttactga 600  
 attggattcg ttgcattta accgcaagtc cagaattgcc aagcgacctc tcgagctact 660  
 55 cgaagacgag gaggacgtgg acgttggagc tgcaaagggt gtagacattg agccgactaa 720  
 cggaaacaag aggcggagac ggtatcacac catcgaagac agtgacgatg aagagttgga 780  
 tgagaagaaa tcgtttgggtg ataactctgac ccaggaacg gaaatcgatc atgtgacagc 840  
 cgatgaatcc ttagcaaaaa ggttgacgga tttagagcac cgggcagttt ctggccgtaa 900  
 60 tcgcttggtt caaattttgt cagattccga tgaagaagaa gaggaagaag taaatcccat 960  
 aacctcacc ctacaaaggt gtgaccagat tgcagcatca ttgagagaag agctgcaggc 1020  
 cagcagttca agtgataact cggttaatga agatcgttat gcagaggttg atgtagcagc 1080

65

ES 2 279 777 T3

agcaaaaatt gtgagccagg cagatgtctg tgcagcttgt ggcattgcog agaatgatac 1140  
 acaacgaatg ctcaagccat atcagcttgt aggcgtcaat ttcatgctgc tacttcaccg 1200  
 caaacatggt ggggggtggca gttgcgtatc ttgcccttct gaaacatctt gatggagatg 1260  
 5 ctggtcctca ttttttagtt gcacctgctt ctcttttaga aaactggcaa agagaactca 1320  
 agaagtgggtg tcctgcattt aagggtggagc tctatcatgg ctccaggaagg gcagctttaa 1380  
 acaggaggct tcagtatgct gcaaaatcta aagggcctgc accctttaac gtgatgctga 1440  
 cgtgctactc ccttttttag aggcagagtg ctccagacaaa agatgaccgc aaattcctta 1500  
 agaaaatggaa ttggcgctgt gtgggttatgg acgaggctca tcttttgaag gacagaagca 1560  
 10 gctttcgcag caaaaagttg cgagatatag ctccaaaagc aattcaaaga ctgatgctga 1620  
 ctggtactcc actccagaac gattttgcagg agctatggtc acttctggag ttcgatgctg 1680  
 ctgatgtggt caacacaaat ggcggttgatt tagatcaata tctgggaacc aggaacgata 1740  
 cctcagggat tgttgtgcag gatacgaact tgatgactcg gatcaaagga atactaggac 1800  
 cttttgtatt acgggaatg aaaactgatg ttatgcgcca gcttgtatca aagattcagg 1860  
 15 aggtggagtg tgtggagatg cttagcagagc aatcaatggc atataaaaaa gctgtgaatg 1920  
 agtatagagc ccttgctgag tccgcacgtg ccgctaaagc tgcaaagaaa tcctcagtta 1980  
 gcgtagtaga tgtccttctt cgtcgacaag tgaccaatat ctttactcaa ttgagaaagt 2040  
 caagaaattg gctaagaaat ttcattccatt aggagttttt ggatatgaat gcgatttgca 2100  
 gogtgtggag gaagaattga ctagttacag cgattttgac ctccacaagt tgtgtattca 2160  
 20 atatggaggc getgcggggag ggcaaggaaa gcttggatgat gatcatgcac tagcttctgc 2220  
 aaagtgccag gcttttagcac gtctacttcc caagttacag caaggtggcc atcgcacatt 2280  
 gatattcagc cagtggacaa gcattgctgga ttttttagaa tgggctcttg acgtcatggg 2340  
 tttttcttac actgccttag atggaagcac tcaagtaagt gaacgccaaa ccctagtgga 2400  
 cgagttcaac aatgacccta gcattttgt gttctcctg tctactcgag ctggaggctca 2460  
 25 aggtctaaat ttaacaggag gcagacacagt cattttacat gatttggact tcaatcccca 2520  
 aatggatcga caggctgagg atcgtctgca teggattggc cagtctaaac ctgttacgat 2580  
 ataccgactt gtaacaaaag ataccggtcga tgaaagtata tacaagatag cccaacagaa 2640  
 gctggtcctc gatgcggcag ttcttgaagg aaaagagtca tctctgata ttaatgatgg 2700  
 30 tgatgctcgc acgatgggtg aaattctttc tgcattattg gatgttccac cgacatgata 2760  
 ctggagtcga gaacacattt ttaattttatt ttcattatct ttatctggca ctgcgagaaa 2820  
 gctcgttaac gcaagggc 2838

<210> 23

35 <211> 288

<212> PRT

<213> *Physcomitrella patens*

40 <400> 23

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

65

ES 2 279 777 T3

Ala Ala Lys Val Val Asp Ile Glu Pro Thr Asn Gly Asn Lys Arg Arg  
 115 120 125

5 Arg Arg Tyr His Thr Ile Glu Asp Ser Asp Asp Glu Glu Leu Asp Glu  
 130 135 140

Lys Lys Ser Phe Gly Asp Asn Leu Thr Pro Gly Thr Glu Ile Asp Gln  
 145 150 155 160

10 Cys Ala Ala Asp Glu Ser Leu Ala Lys Arg Leu Gln Asp Leu Glu His  
 165 170 175

Arg Ala Val Ser Gly Arg Asn Arg Leu Val Gln Ile Leu Ser Asp Ser  
 180 185 190

Asp Glu Glu Glu Glu Glu Val Asn Pro Ile Thr Ile Thr Leu Gln  
 195 200 205

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

Ser Ser Ser Asp Asn Ser Val Asn Glu Asp Arg Tyr Ala Glu Val Asp  
 225 230 235 240

25 Val Ala Ala Ala Lys Ile Val Ser Gln Ala Asp Val Cys Ala Ala Cys  
 245 250 255

Gly Ile Ala Glu Asn Asp Thr Gln Arg Met Leu Lys Pro Tyr Gln Leu  
 260 265 270

30 Val Gly Val Asn Phe Met Leu Leu Leu His Arg Lys His Val Gly Gly  
 275 280 285

35 <210> 24

<211> 18

<212> DNA

40 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Iniciador

45 <400> 24

caggaaacag ctatgacc

18

50 <210> 25

<211> 19

<212> DNA

55 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Iniciador

60 <400> 25

ctaaaggga caaaagctg

19

65 <210> 26

<211> 18

## ES 2 279 777 T3

	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 26	
10	tgtaaacga cggccagt	18
	<210> 27	
15	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 27	
25	gcgatcctca gcctgcgat ccatt	25
	<210> 28	
30	<211> 26	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 28	
40	ccctgaggta tegtctctgg ttcca	26
	<210> 29	
45	<211> 33	
	<212> DNA	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 29	
55	atcccgggcg atggtgcgtt cgagatcgta agg	33
	<210> 30	
60	<211> 32	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	

## ES 2 279 777 T3

<400> 30  
gcgtaacga gctttctcgc agtgccagat aa 32

5  
<210> 31  
<211> 33  
<212> DNA  
10 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador

15  
<400> 31  
atcccgggct ctgcaccca gatgtcgcat cct 33

20  
<210> 32  
<211> 33  
<212> DNA  
25 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador

30  
<400> 32  
ctgagctcta atgcattcac tgtgtctgct gct 33

35  
<210> 33  
<211> 25  
<212> DNA  
40 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador

45  
<400> 33  
cctgtagggc caccggagc tcaact 25

50  
<210> 34  
<211> 33  
<212> DNA  
55 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador

60  
<400> 34  
gagtaacgc agtggtcaca acgcagagta cgc 33

65  
<210> 35  
<211> 33

## ES 2 279 777 T3

	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 35	
10	gcgatatcgc tccatacct gcgccgaaga ctt	33
	<210> 36	
15	<211> 32	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 36	
25	gacccgggcc atgtgatatg gcttcaaagt at	32
	<210> 37	
30	<211> 32	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 37	
40	gcgttaacga ctactgaga gtcataatgg tg	32
	<210> 38	
45	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 38	
55	cgtagtcgcg ctcgagctgt ttggt	25
	<210> 39	
60	<211> 33	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	



## ES 2 279 777 T3

<400> 39  
atcccgggca cgagggcaag aggggataga gac 33

5  
<210> 40  
<211> 32  
<212> DNA  
10 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
15  
<400> 40  
gcgttaacgc cgatggtgca acttgggtg ac 32

20  
<210> 41  
<211> 25  
<212> DNA  
25 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
30  
<400> 41  
ccgtgtcctc ggagcattct ggcac 25

35  
<210> 42  
<211> 33  
<212> DNA  
40 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
45  
<400> 42  
atcccgggag gagggagtg gaatctagga gac 33

50  
<210> 43  
<211> 33  
<212> DNA  
55 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
60  
<400> 43  
gcgagctcga ccttgctcga tggagactcc aat 33

65  
<210> 44  
<211> 33

## ES 2 279 777 T3

	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 44	
10	atcccgggaa taggacggat ggccgacagt tac	33
	<210> 45	
15	<211> 33	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 45	
25	atgagctcac tcttacctc cgcggggttg gtt	33
	<210> 46	
30	<211> 20	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 46	
40	gaatagatac gctgacacgc	20
	<210> 47	
45	<211> 33	
	<212> DNA	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 47	
55	atgagctcac tcttacctc cgcggggttg gtt	33
	<210> 48	
60	<211> 32	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	

## ES 2 279 777 T3

	<400> 48	
	gcttaaacgc cgatggtgca actttggtg ac	32
5	<210> 49	
	<211> 33	
	<212> DNA	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
15	<400> 49	
	gcgagctcga cttgctcga tggagactcc aat	33
20	<210> 50	
	<211> 32	
	<212> DNA	
25	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
30	<400> 50	
	gcgtaaaca ctcactgaga gtcataatgg tg	32
35	<210> 51	
	<211> 20	
	<212> DNA	
40	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
45	<400> 51	
	ctagtaacat agatgacacc	20
50	<210> 52	
	<211> 33	
	<212> DNA	
55	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
60	<400> 52	
	atcccgggcg atggtgcgtt cgagatcgta agg	33
65	<210> 53	
	<211> 20	

## ES 2 279 777 T3

	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 53	
10	gaatagatac gctgacacgc	20
	<210> 54	
15	<211> 33	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<221> Descripción de la Secuencia Artificial: Iniciador	
	<400> 54	
25	ctgagctcta atgcattcac tgttgetget gct	33
	<210> 55	
30	<211> 33	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 55	
40	gcgatatgc ttccatacct ggcggaaga ctt	33
	<210> 56	
45	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 56	
55	ggagacggta tcacaccatc gaaga	25
	<210> 57	
60	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	

## ES 2 279 777 T3

<400> 57  
tgcacagaca tctgcctggc tcaca 25

5  
<210> 58  
<211> 25  
<212> DNA  
10 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
15  
<400> 58  
gatgatcgca gccgaagctc cagtg 25

20  
<210> 59  
<211> 26  
<212> DNA  
25 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
30  
<400> 59  
gggtgtgcca tggactggtg ttccag 26

35  
<210> 60  
<211> 26  
<212> DNA  
40 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
45  
<400> 60  
ggcagtctgt ggaggctgat acatca 26

50  
<210> 61  
<211> 25  
<212> DNA  
55 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Iniciador  
60  
<400> 61  
cctgatcctg tgacccttt tgcca 25

65  
<210> 62  
<211> 25

## ES 2 279 777 T3

	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 62	
10	gacatggacg gtgatgcgaa gttgg	25
	<210> 63	
15	<211> 26	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 63	
25	cggcaacagc aggtctata ccttgg	26
	<210> 64	
30	<211> 26	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 64	
40	gcatactcca gtcaaatgc agcagc	26
	<210> 65	
45	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 65	
55	gggtcggcag cctccaatcc ataca	25
	<210> 66	
60	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	

## ES 2 279 777 T3

	<400> 66	
	ggcaggggaat ctacgcatcg ctttg	25
5	<210> 67	
	<211> 26	
	<212> DNA	
10	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
15	<400> 67	
	cgacgagatt ctctgcaaca tctgag	26
20	<210> 69	
	<211> 26	
	<212> DNA	
25	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
30	<400> 68	
	ggagcttgga ctgcgacctc gtcaag	26
35	<210> 69	
	<211> 26	
	<212> DNA	
40	<213> Secuencia Artificial	
	<220>	
	<223> Secuencia Artificial	
45	<400> 69	
	ggtgtggctc gtgcgagggc taccag	26
50	<210> 70	
	<211> 25	
	<212> DNA	
55	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
60	<400> 70	
	gtcatcgagg aatcgacaaa ctctt	25
65	<210> 71	
	<211> 26	

## ES 2 279 777 T3

	<212> DNA	
	<213> Secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 71	
10	ggttgacggt ggattgcaca tggtag	26
	<210> 72	
15	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 72	
25	tggatgtgcg aagtgtgcca ggttg	25
	<210> 73	
30	<211> 25	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 73	
40	gcgctgcctc tgataataga gttgg	25
	<210> 74	
45	<211> 27	
	<212> DNA	
	<213> Secuencia Artificial	
50	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	
	<400> 74	
55	gtgcaggagt gcgtatcca gttcatc	27
	<210> 75	
60	<211> 27	
	<212> DNA	
	<213> Secuencia Artificial	
65	<220>	
	<223> Descripción de la Secuencia Artificial: Iniciador	



## ES 2 279 777 T3

<400> 75

cgtagggctg ttgcatcatc tgcacg

27

5

<210> 76

<211> 26

<212> DNA

10 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Iniciador

15

<400> 76

cgccaagcgc gcaattaacc ctact

26

20

<210> 77

<211> 26

<212> DNA

25 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Iniciador

30

<400> 77

gcgtaatacg actcactata gggcga

26

35

40

45

50

55

60

65