



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 365 902**

51 Int. Cl.:
C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **02757104 .1**

96 Fecha de presentación : **09.08.2002**

97 Número de publicación de la solicitud: **1485490**

97 Fecha de publicación de la solicitud: **15.12.2004**

54 Título: **Polinucleótidos y polipéptidos en plantas con el fin de mejorar sus características.**

30 Prioridad: **09.08.2001 US 310847 P**
19.11.2001 US 336049 P
11.12.2001 US 338692 P
14.06.2002 US 171468

45 Fecha de publicación de la mención BOPI:
13.10.2011

45 Fecha de la publicación del folleto de la patente:
13.10.2011

73 Titular/es: **MENDEL BIOTECHNOLOGY, Inc.**
3935 Point Eden Way
Hayward, California 94545, US

72 Inventor/es: **Ratcliffe, Oliver;**
Riechmann, Jose, Luis;
Adam, Luc, J.;
Dubell, Arnold, N.;
Heard, Jacqueline, E.;
Pilgrim, Marsha, L.;
Jiang, Cai-Zhong;
Reuber, T., Lynne;
Creelman, Robert, A.;
Pineda, Omaira;
Yu, Guo-Liang y
Broun, Pierre, E.

74 Agente: **Ponti Sales, Adelaida**

ES 2 365 902 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

Polinucleótidos y polipéptidos en plantas con el fin de mejorar sus características

5 CAMPO DE LA INVENCION

[0001] Esta invención se refiere al campo de la biología de plantas, más particularmente, la presente invención abarca composiciones y procedimientos para modificar fenotípicamente una planta.

10 INTRODUCCION

[0002] Los rasgos de una planta, tales como sus características bioquímicas, de desarrollo, o fenotípicas, se pueden controlar a través de una serie de procesos celulares. Una forma importante de manipular ese control es mediante los factores de transcripción - proteínas que influyen en la expresión de un gen o grupos de genes particulares. Las plantas transformadas y transgénicas que comprenden células con niveles alterados de al menos un factor de transcripción seleccionado, por ejemplo, poseen rasgos ventajosos o deseables. Por tanto, las estrategias para la manipulación de rasgos alterando el contenido en factores de transcripción de las células de una planta pueden dar como resultado plantas y cultivos con propiedades comercialmente valiosas. Los solicitantes han identificado polinucleótidos que codifican factores de transcripción, han desarrollado numerosas plantas transgénicas usando estos polinucleótidos, y han analizado las plantas para una variedad de rasgos importantes. De este modo, los solicitantes han identificado secuencias de polinucleótidos y polipéptidos importantes para la producción de plantas y cultivos comercialmente valiosos así como los procedimientos para su preparación y su uso. Otros aspectos y formas de realización de la invención se describen a continuación y se pueden inferir de las enseñanzas de esta descripción en su totalidad.

25

ANTECEDENTES DE LA INVENCION

[0003] Los factores de transcripción (FTs) pueden modular la expresión de genes, ya sea incrementando o disminuyendo (induciendo o reprimiendo) la tasa de transcripción. Esta modulación da como resultado niveles diferenciales de expresión génica en diversas fases del desarrollo, en tejidos y tipos celulares diferentes, y en respuesta a diferentes estímulos exógenos (por ejemplo, ambientales) y endógenos a lo largo del ciclo de vida del organismo.

30

[0004] Debido a que los factores de transcripción son elementos de control claves de las rutas biológicas, la alteración de los niveles de expresión de uno o más factores de transcripción puede variar todas las rutas biológicas de un organismo. Por ejemplo, la manipulación de los niveles de factores de transcripción seleccionados puede dar como resultado una expresión incrementada de proteínas o compuestos químicos metabólicos en plantas económicamente útiles o mejorar otras características de relevancia agrícola. Por el contrario, la expresión bloqueada o reducida de un factor de transcripción puede reducir la biosíntesis de compuestos no deseados o la eliminación de un rasgo no deseable. Por tanto, la manipulación de los niveles de los factores de transcripción en una planta ofrece un potencial tremendo en biotecnología agrícola para la modificación de los rasgos de una planta.

40

[0005] Los documentos WO98/07842, WO01/35727 y EP-A-1033405 se refieren a genes de plantas que incluyen factores de transcripción de la familia AP2. La presente descripción proporcionan nuevos factores de transcripción útiles para la modificación del fenotipo de una planta de formas deseables.

45

RESUMEN DE LA INVENCION

[0006] La presente invención proporciona un procedimiento para la producción de una planta transgénica que tenga un rasgo alterado seleccionado del grupo constituido por una mayor resistencia al estrés osmótico, una arquitectura alterada, una estructura de los tejidos vasculares alterada, una floración retardada, una morfología del tallo alterada comparada con una planta de tipo silvestre de la misma especie, el procedimiento que comprende:

50

(a) la proporción de un polinucleótido que codifica un polipéptido con una identidad de al menos el 95% con

**MDYRESTGESQSKYKGI RRRKWKWVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFFCL
 HQPDSLESLNFP HLLNPSLVSRTSPRSIQQAASNAGMAIDAGIVHSTSVNSGCGDTTTY
 ENGADQVEPLNISVYDYLGGHDHV***

5 y que comprende un dominio conservado con una identidad de secuencia de al menos el 95% a los aminoácidos con las coordenadas 11 a 80 de

**MDYRESTGESQSKYKGI RRRKWKWVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFFCL
 HQPDSLESLNFP HLLNPSLVSRTSPRSIQQAASNAGMAIDAGIVHSTSVNSGCGDTTTY
 ENGADQVEPLNISVYDYLGGHDHV***

10 (b) la introducción del polinucleótido en una planta para generar una planta transgénica; y

(c) la selección de la planta transgénica que comprende la primera secuencia de nucleótidos y que tiene dicho rasgo seleccionado alterado en comparación con la planta de tipo silvestre.

15 **[0007]** En el presente documento también se describen plantas transgénicas que comprenden un vector o casete de expresión que comprende los polinucleótidos definidos anteriormente donde la planta transgénica presenta un rasgo alterado comparado con una planta de tipo silvestre de la misma especie, dicho rasgo alterado que se selecciona del grupo constituido por una mayor resistencia al estrés osmótico, una arquitectura alterada, una estructura de los tejidos vasculares alterada, una floración retardada, y una morfología del tallo alterada.

20

[0008] La planta puede ser cualquier planta, incluyendo, pero no limitado a, *Arabidopsis*, mostaza, soja, trigo, maíz, patata, algodón, arroz, colza, girasol, alfalfa, caña de azúcar, hierba, plátano, mora, arándano, fresa, frambuesa, cantalupo, zanahoria, coliflor, café, pepino, berenjena, uva, melón Galia, lechuga, mango, melón, cebolla, papaya, guisantes, pimientos, piña, calabaza, espinaca, calabacín, maíz dulce, tabaco, tomate, sandía, frutas rosáceas, coles vegetales y menta u otras lamiáceas. En otro aspecto más, la invención es un material vegetal aislado de una planta, incluyendo pero no limitado a, tejidos vegetales, frutos, semillas, células vegetales, embriones, protoplastos, polen, y similares. En otro aspecto más, la invención es un cultivo de tejidos vegetales transgénicos de células regenerables, incluyendo pero no limitado a, embriones, células meristemáticas, microsporas, protoplastos, polen, y similares.

30

[0009] En el presente documento se describe una planta transgénica que comprende uno o más de los polinucleótidos descritos anteriormente donde el polipéptido codificado se expresa y regula la transcripción de un gen.

35 **[0010]** En un aspecto adicional la presente invención proporciona un procedimiento para el uso de la composición de polinucleótidos para el cultivo de una progenie de plantas a partir de una planta transgénica que incluye el cruzamiento de plantas, la producción de semillas a partir de plantas transgénicas, y procedimientos de cultivo usando plantas transgénicas, el procedimiento que comprende la transformación de una planta con la composición de polinucleótidos para crear una planta transgénica, el cruzamiento de la planta transgénica con otra
 40 planta, la selección de la semilla, y el crecimiento de la progenie de la planta a partir de la semilla.

[0011] En un aspecto adicional, la presente descripción proporciona una progenie de una planta derivada de una planta parental donde dicha progenie de la planta presenta niveles de ARN mensajero al menos tres veces superiores a los de dicha planta parental, donde el ARN mensajero codifica una proteína de unión al ADN que es capaz de unirse a una secuencia reguladora del ADN y que induce la expresión de un gen que codifica un rasgo de
 45 la planta, donde la progenie de la planta se caracteriza por una variación en el rasgo de la planta comparada con dicha planta parental. En otro aspecto adicional, la progenie de la planta presenta niveles de ARN mensajero al menos 50 veces superiores comparados a los de dicha planta parental.

[0012] En un aspecto adicional, la descripción se refiere a un vector de clonación o expresión que comprende el polinucleótido aislado o recombinante descrito anteriormente o células que comprenden el vector de clonación o expresión.

5 **[0013]** En otro aspecto adicional, la presente descripción se refiere a una composición producida incubando un polinucleótido de la invención con una nucleasa, una enzima de restricción, una polimerasa; una polimerasa y un cebador; un vector de clonación; o con una célula.

[0014] Además, la descripción se refiere a un procedimiento para la producción de una planta que tiene un rasgo modificado, donde dicho rasgo se define en las reivindicaciones adjuntas. El procedimiento comprende la alteración de la expresión de un polinucleótido aislado o recombinante como se describe en el presente documento o la alteración de la expresión o actividad de un polipéptido como se describe en el presente documento en una planta para producir una planta modificada, y la selección de la planta modificada por un rasgo modificado. En un aspecto, la planta es una planta monocotiledónea. En otro aspecto, la planta es una planta dicotiledónea. En otro aspecto, el polinucleótido recombinante procede de una planta dicotiledónea y la planta es una planta monocotiledónea. En otro aspecto más el polinucleótido recombinante procede de una planta monocotiledónea y la planta es una planta dicotiledónea. En otro aspecto más el polinucleótido recombinante procede de una planta monocotiledónea y la planta es una planta monocotiledónea. En otro aspecto adicional, el polinucleótido recombinante procede de una planta dicotiledónea y la planta es una planta dicotiledónea.

20 **[0015]** Adicionalmente se describe una planta transgénica que comprende un polinucleótido aislado o recombinante que codifica un polipéptido donde el polipéptido es la SEQ ID N°: 32.

[0016] Adicionalmente se describe una planta con niveles de expresión alterados de un polipéptido descrito anteriormente o una planta con niveles de expresión o actividad alterados de un polipéptido descrito anteriormente.

[0017] La planta puede ser cualquier planta, incluyendo, pero no limitado a, *Arabidopsis*, mostaza, soja, trigo, maíz, patata, algodón, arroz, colza, girasol, alfalfa, caña de azúcar, hierba, plátano, mora, arándano, fresa, frambuesa, cantalupo, zanahoria, coliflor, café, pepino, berenjena, uva, melón Galia, lechuga, mango, melón, cebolla, papaya, guisantes, pimientos, piña, calabaza, espinaca, calabacín, maíz dulce, tabaco, tomate, sandía, frutas rosáceas, coles vegetales y menta u otras lamiáceas. En otro aspecto más, la invención es un material vegetal aislado de una planta, incluyendo pero no limitado a, tejidos vegetales, frutos, semillas, células vegetales, embriones, protoplastos, polen, y similares.

35 **[0018]** Adicionalmente se describe un cultivo de tejidos vegetales transgénicos de células regenerables, incluyendo pero no limitado a, embriones, células meristemáticas, microsporas, protoplastos, polen, y similares.

BREVE DESCRIPCIÓN DEL LISTADO DE SECUENCIAS, TABLAS, Y FIGURAS

40 **[0019]** La última página de la descripción proporciona secuencias de ejemplo de los polinucleótidos y polipéptidos de la invención. Los rasgos asociados al uso de las secuencias están incluidos en los Ejemplos.

[0020] La Tabla 4 muestra los polinucleótidos y polipéptidos identificados por la SEQ ID N°; Mendel Gene ID N°; dominio conservado del polipéptido; y si el polinucleótido se probó en un ensayo transgénico. La primera columna muestra la SEQ ID N° del polinucleótido; la segunda columna muestra el Mendel Gene ID N°, GID; la tercera columna muestra el rasgo(s) que resulta de la inactivación o sobreexpresión del polinucleótido en la planta transgénica; la cuarta columna muestra la categoría del rasgo; la quinta columna muestra la familia del factor de transcripción a la cual pertenece el polinucleótido; la sexta columna ("Comentario"), incluye efectos y utilidades específicos conferidos por el polinucleótido de la primera columna; la séptima columna muestra la SEQ ID N° del polipéptido codificado por el polinucleótido; y la octava columna muestra las posiciones de los restos aminoácidos del dominio conservado en las coordenadas de aminoácidos (AA).

[0021] La Figura 1 muestra un árbol filogenético de familias de plantas relacionadas adaptadas de Daly y col. (2001 Plant Physiology 127:1328-1333).

Descripción detallada de formas de realización ejemplares

- [0022]** En un aspecto importante, la presente descripción se refiere a polinucleótidos y polipéptidos, por ejemplo, para la modificación de fenotipos de plantas. A lo largo de esta descripción, se mencionan y/o se incorporan específicamente diversas fuentes de información. Las fuentes de información incluyen artículos de revistas científicas, documentos de patentes, libros de texto, y direcciones de páginas web inactivas, por ejemplo. La referencia a estas fuentes de información indica claramente que pueden ser usadas por alguien con conocimientos en la materia.
- [0023]** Se debe indicar que tal y como se usan en el presente documento y en las reivindicaciones anexas, las formas en singular "un" y "el" incluyen la referencia en plural, a menos que el contexto dicte claramente otra cosa. Así, por ejemplo, una referencia a "una planta" incluye una pluralidad de dichas plantas, y una referencia a "un estrés" es una referencia a uno o más estreses o sus equivalentes conocidos por aquellos expertos en la materia, etcétera.
- [0024]** Las secuencias de polinucleótidos de la descripción codifican polipéptidos que son miembros de familias de factores de transcripción muy conocidas, incluyendo familias de factores de transcripción de plantas, como se describe en la Tabla 4. En general, los factores de transcripción codificados por las presentes secuencias están involucrados en la diferenciación y proliferación celular y en la regulación del crecimiento. Por consiguiente, alguien con conocimientos en la materia reconocerá que con la expresión de las presentes secuencias en una planta, se puede variar la expresión de genes autólogos o inducir la expresión de genes introducidos. Afectando la expresión de secuencias autólogas similares en una planta que tengan la actividad biológica de las presentes secuencias, o introduciendo las presentes secuencias en una planta, se puede alterar el fenotipo de una planta a uno con rasgos mejorados. Las secuencias como las descritas en el presente documento también se pueden usar para transformar una planta e introducir rasgos deseables no encontrados en cultivos o variedades silvestres. A continuación las plantas se pueden seleccionar para aquellas que producen el grado de sobre- o infra-expresión más deseable de genes diana de interés y mejora del rasgo coincidente.
- [0025]** Las secuencias descritas en el presente documento pueden proceder de cualquier especie, particularmente especies de plantas, en una forma de origen natural o de cualquier fuente, ya sea natural, sintética, semisintética o recombinante. Las secuencias también pueden incluir fragmentos de las presentes secuencias de aminoácidos. En este contexto, un "fragmento" se refiere a un fragmento de una secuencia de polipéptidos con una longitud de al menos 5 a 15 aminoácidos aproximadamente, más preferentemente de al menos 14 aminoácidos, y que retiene parte de la actividad biológica de un factor de transcripción. Cuando se menciona "secuencia de aminoácidos" para referirse a una secuencia de aminoácidos de una molécula de proteínas de origen natural, no se pretende que la "secuencia de aminoácidos" y términos similares se limiten a la secuencia de aminoácidos nativa completa asociada a la molécula de proteína mencionada.
- [0026]** Como reconocerán aquellos con conocimientos ordinarios en la materia, los factores de transcripción se pueden identificar por la presencia de una región o dominio de similitud o identidad estructural con una secuencia consenso específica o la presencia de un sitio de unión al ADN consenso específico o un motivo para un sitio de unión al ADN (véase, por ejemplo, Riechmann y col., (2000) Science 290: 2105-2110). Los factores de transcripción de plantas de la invención pertenecen a la familia del factor de transcripción con el dominio AP2 (APETALA2) (Riechmann y Meyerowitz (1998) Biol. Chem. 379:633-646).
- [0027]** Además de los procedimientos para la modificación del fenotipo de una planta empleando uno o más polinucleótidos y polipéptidos como los descritos en el presente documento, los polinucleótidos y polipéptidos descritos en el presente documento presentan una variedad de usos adicionales. Estos usos incluyen su utilización en la producción recombinante (es decir, expresión) de proteínas; como reguladores de la expresión génica de la planta, como sondas diagnósticas para la presencia de ácidos nucleicos complementarios o parcialmente complementarios (incluyendo la detección de ácidos nucleicos codificantes naturales); como sustratos para otras reacciones, por ejemplo, reacciones de mutación, reacciones de PCR, o similares; como sustratos para clonación, por ejemplo, incluyendo reacciones de digestión o ligación; y para la identificación de moduladores exógenos o endógenos de los factores de transcripción. Un "polinucleótido" es una secuencia de ácidos nucleicos que comprende una pluralidad de nucleótidos polimerizados, por ejemplo, al menos 15 nucleótidos polimerizados consecutivos aproximadamente, opcionalmente al menos 30 nucleótidos consecutivos aproximadamente, y al menos 50 nucleótidos consecutivos aproximadamente. En muchos casos, un polinucleótido comprende una secuencia de nucleótidos que codifica un polipéptido (o proteína) o uno de sus dominios o fragmentos. Adicionalmente, el polinucleótido puede comprender un promotor, un intrón, una región potenciadora, un sitio de poliadenilación, un sitio de iniciación de la traducción, regiones sin traducir en 5' o 3', un gen informador, un marcador seleccionable, o similar. El polinucleótido puede ser ADN o ARN de cadena sencilla o de doble cadena. El polinucleótido opcionalmente comprende bases modificadas o un esqueleto modificado. El polinucleótido puede ser, por ejemplo,

ADN o ARN genómico, un transcrito (tal como ARNm), un ADNc, un producto de la PCR, un ADN clonado, un ADN o ARN sintético, o similar. El polinucleótido puede comprender una secuencia en cualquiera de las dos orientaciones en dirección directa o dirección inversa.

5 **[0028]** Un "polinucleótido recombinante" es un polinucleótido que no se encuentra en su estado nativo, por ejemplo, el polinucleótido comprende una secuencia de nucleótidos no encontrada en la naturaleza, o el polinucleótido está en un contexto distinto del que se encuentra en su forma natural, por ejemplo, separado de secuencias de nucleótidos a las cuales normalmente está próximo en la naturaleza, o adyacente (o contiguo) a secuencias de nucleótidos a las cuales normalmente no está próximo. Por ejemplo, la secuencia en cuestión puede estar clonada dentro de un vector, o combinada de otra forma con uno o más ácidos nucleicos adicionales.

10 **[0029]** Un "polinucleótido aislado" es un polinucleótido de origen natural o recombinante, que está presente fuera de la célula en la que se encuentra normalmente en la naturaleza, ya sea purificado o no. Opcionalmente, un polinucleótido aislado se somete a uno o más procedimientos de enriquecimiento o purificación, por ejemplo, lisis celular, extracción, centrifugación, precipitación, o similar.

15 **[0030]** Un "polipéptido" es una secuencia de aminoácidos que comprende una pluralidad de restos aminoácidos polimerizados consecutivos, por ejemplo, al menos 15 restos aminoácidos polimerizados consecutivos aproximadamente, opcionalmente al menos 30 restos aminoácidos polimerizados consecutivos aproximadamente, y al menos 50 restos aminoácidos polimerizados consecutivos aproximadamente. En muchos casos, un polipéptido comprende una secuencia de restos aminoácidos polimerizados que es un factor o un dominio de transcripción o una de sus porciones o fragmentos. Adicionalmente, el polipéptido puede comprender un dominio de localización, 2) un dominio de activación, 3) un dominio de represión, 4) un dominio de oligomerización o 5) un dominio de unión al ADN, o similar. El polipéptido opcionalmente comprende restos aminoácidos modificados, restos aminoácidos de origen natural no codificados por un codón, y restos aminoácidos que no son de origen natural.

20 **[0031]** Un "polipéptido recombinante" es un polipéptido producido por la traducción de un polinucleótido recombinante. Un "polipéptido sintético" es un polipéptido creado por la polimerización consecutiva de restos aminoácidos aislados usando procedimientos muy conocidos en la materia. Un "polipéptido aislado", ya sea un polipéptido de origen natural o recombinante, está más enriquecido en (o fuera de) una célula que el polipéptido en su estado natural en una célula silvestre, por ejemplo, enriquecido en más del 5% aproximadamente, enriquecido en más del 10% aproximadamente, o enriquecido en más del 20% aproximadamente, o más del 50% aproximadamente, o superior, es decir, indicado alternativamente: enriquecido en un 105%, 110%, 120%, 150% o superior, en relación al tipo silvestre estandarizado al 100%. Dicho enriquecimiento no es el resultado de una respuesta natural de una planta silvestre. Alternativa, o adicionalmente, el polipéptido aislado se separa de otros componentes celulares a los cuales está asociado normalmente, por ejemplo, mediante cualquiera de los diversos procedimientos de purificación de proteínas del presente documento.

30 **[0032]** "Identidad" o "similitud" se refiere a la similitud de secuencia entre dos secuencias de polinucleótidos o entre dos secuencias de polipéptidos, siendo la identidad una comparación más estricta. Las frases "porcentaje de identidad" y "% de identidad" se refieren al porcentaje de similitud de secuencia encontrado en una comparación de dos o más secuencias de polinucleótidos o dos o más secuencias de polipéptidos. La identidad o similitud se puede determinar comparando una posición en cada secuencia que se pueden alinear con fines comparativos. Cuando una posición en la secuencia comparada está ocupada por el mismo nucleótido, base o aminoácido, entonces las moléculas son idénticas en esa posición. El grado de similitud o identidad entre secuencias de polinucleótidos es una función del número de nucleótidos idénticos o coincidentes en posiciones compartidas por las secuencias de polinucleótidos. El grado de identidad de las secuencias de polipéptidos es una función del número de aminoácidos idénticos en posiciones compartidas por las secuencias de polipéptidos. El grado de homología o similitud de las secuencias de polipéptidos es una función del número de aminoácidos, es decir, estructuralmente relacionado, en posiciones compartidas por las secuencias de polipéptidos.

40 **[0033]** Las secuencias de ácidos nucleicos "alteradas" que codifican polipéptidos incluyen aquellas secuencias con eliminaciones, inserciones, o sustituciones de diferentes nucleótidos, que dan como resultado un polinucleótido que codifica un polipéptido con al menos una característica funcional del polipéptido. Incluidos dentro de esta definición están los polimorfismos que pueden o no ser fácilmente detectables usando una sonda de oligonucleótidos particular del polinucleótido que codifica el polipéptido, e hibridaciones impropias o inesperadas a variantes alélicas, con un locus distinto del locus cromosómico normal para la secuencia de polinucleótidos que codifica el polipéptido. La proteína polipeptídica codificada también puede estar "alterada", y puede contener eliminaciones, inserciones o sustituciones de restos aminoácidos que producen un cambio silencioso y que dan como resultado un polipéptido funcionalmente equivalente. Se pueden introducir sustituciones de aminoácidos deliberadas en base a la similitud de las propiedades químicas de la cadena lateral del resto, incluyendo, pero no limitado a, polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad, y/o la naturaleza anfipática de los restos, siempre que se mantenga la actividad biológica del polipéptido. Por ejemplo, los aminoácidos cargados

negativamente pueden incluir ácido aspártico y ácido glutámico, los aminoácidos cargados positivamente pueden incluir lisina y arginina, y los aminoácidos con grupos de cabeza polares sin carga que tienen valores de hidrofiliidad similares pueden incluir leucina, isoleucina, y valina; glicina y alanina; asparagina y glutamina; serina y treonina; y fenilalanina y tirosina. Los alineamientos entre diferentes secuencias de polipéptidos se pueden usar para calcular "el porcentaje de similitud de secuencia".

[0034] El término "planta" incluye plantas completas, órganos/estructuras vegetativas geminales (por ejemplo, hojas, tallos y tubérculos), raíces, flores y órganos/estructuras florales (por ejemplo, brácteas, sépalos, pétalos, estambres, carpelos, anteras y óvulos), semillas (incluyendo embrión, endospermo, y tegumento) y fruto (el ovario maduro), tejido de la planta (por ejemplo, tejido vascular, tejido conjuntivo, y similares) y células (por ejemplo, células de guarda, óvulos, y similares), y la progenie de las mismas. La clase de plantas que se pueden utilizar en el procedimiento de la invención es por lo general tan amplia como la clase de plantas superiores e inferiores susceptibles a las técnicas de transformación, incluyendo las angiospermas (plantas monocotiledóneas y dicotiledóneas), gimnospermas, helechos, colas de caballo, psilofitas, licofitas, briofitas y algas pluricelulares. (Véase, por ejemplo, la Figura 1, adaptada de Daly y col. 2001 *Plant Physiology* 127:1328-1333; y véase también Tudge, C., *The Variety of Life*, Oxford University Press, Nueva York, 2000, pp. 547-606).

[0035] Una "planta transgénica" se refiere a una planta que contiene material genético no encontrado en una planta de tipo silvestre de la misma especie, variedad o cultivo. El material genético puede incluir un transgen, un acontecimiento de mutagénesis de inserción (tal como mutagénesis de inserción mediante transposón o T-ADN), una secuencia de marcaje de la activación, una secuencia mutada, un acontecimiento de recombinación homóloga o una secuencia modificada por quimeroplastia. Normalmente, el material genético extraño ha sido introducido en la planta por manipulación humana, pero como reconoce alguien con conocimientos en la materia se puede usar cualquier procedimiento.

[0036] Una planta transgénica puede contener un vector o casete de expresión. El casete de expresión normalmente comprende una secuencia que codifica un polipéptido unida de manera operable a (es decir, bajo el control regulador de) secuencias reguladoras inducibles o constitutivas apropiadas que permiten la expresión del polipéptido. El casete de expresión se puede introducir en una planta por transformación o por cultivo después de la transformación de una planta parental. Una planta se refiere a una planta completa así como a una parte de una planta, tal como la semilla, el fruto, la hoja, o la raíz, el tejido de una planta, células de la planta o cualquier otro material vegetal, por ejemplo, un explante de la planta, así como su progenie, y a sistemas *in vitro* que mimetizan los componentes o procesos bioquímicos o celulares de una célula.

[0037] La "expresión ectópica o expresión alterada" en referencia a un polinucleótido indica que el patrón de expresión en, por ejemplo, una planta o un tejido de planta transgénica, es diferente del patrón de expresión en una planta silvestre o en una planta de referencia de la misma especie. El patrón de expresión también se puede comparar con un patrón de expresión de referencia en una planta silvestre de la misma especie. Por ejemplo, el polinucleótido o polipéptido se expresa en un tipo celular o de tejido distinto del tipo celular o de tejido en el que se expresa la secuencia en la planta silvestre, o como respuesta a agentes inducibles diferentes, tales como hormonas o señales ambientales, o niveles de expresión diferentes (ya sea superiores o inferiores) comparados con aquellos encontrados en una planta silvestre. El término también se refiere a patrones de expresión alterados que se producen reduciendo los niveles de expresión por debajo del nivel de detección o aboliendo completamente la expresión. El patrón de expresión resultante puede ser transitorio o estable, constitutivo o inducible. En referencia a un polipéptido, el término "expresión ectópica o expresión alterada" además puede estar relacionado con los niveles de actividad alterados que resultan de las interacciones de los polipéptidos con moduladores exógenos o endógenos o de interacciones con factores o como resultado de la modificación química de los polipéptidos.

[0038] Un "fragmento" o "dominio", con respecto a un polipéptido, se refiere a una secuencia del polipéptido. En algunos casos, el fragmento o dominio es una secuencia del polipéptido que desempeña al menos una de las funciones biológicas del polipéptido intacto esencialmente de la misma forma, o en un grado similar, que la que desempeña el polipéptido intacto. Por ejemplo, un fragmento polipeptídico puede comprender un motivo estructural o un dominio funcional reconocible tal como un sitio de unión al ADN o un dominio que se une a una región promotora del ADN, un dominio de activación, o un dominio para interacciones proteína-proteína. Los fragmentos pueden variar de tamaño desde tan sólo seis aminoácidos hasta la longitud completa del polipéptido intacto, pero preferentemente tienen una longitud de al menos 30 aminoácidos aproximadamente y más preferentemente una longitud de al menos 60 aminoácidos aproximadamente. En referencia a una secuencia de polinucleótidos, "un fragmento" se refiere a cualquier secuencia de un polinucleótido, normalmente, de al menos 15 nucleótidos consecutivos aproximadamente, preferentemente de al menos 30 nucleótidos aproximadamente, más preferentemente de al menos 50 nucleótidos aproximadamente, de cualquiera de las secuencias proporcionadas en el presente documento.

[0039] También se describe la producción de secuencias de ADN que codifican factores de transcripción y derivados de factores de transcripción, o sus fragmentos, íntegramente mediante química sintética. Después de su

producción, las secuencias sintéticas se pueden insertar en cualquiera de los numerosos vectores de expresión y sistemas celulares disponibles usando reactivos muy conocidos en la materia. Además, la química sintética se puede usar para introducir mutaciones en una secuencia que codifica factores de transcripción o cualquiera de sus fragmentos.

5

[0040] Un "dominio conservado", con respecto a un polipéptido, se refiere a un dominio dentro de una familia de factores de transcripción que presenta un mayor grado de homología de secuencia, tal como una identidad de secuencia de al menos el 65% que incluyen sustituciones conservativas, y preferentemente una identidad de secuencia de al menos el 80%, y más preferentemente una identidad de secuencia de restos aminoácidos de al menos el 85%, o al menos del 86% aproximadamente, o al menos del 87% aproximadamente, o al menos del 88% aproximadamente, o al menos del 90% aproximadamente, o al menos del 95% aproximadamente, o al menos del 98% aproximadamente de un polipéptido de restos aminoácidos consecutivos. Se puede hacer referencia a que un fragmento o dominio está fuera de una secuencia consenso o fuera de un sitio de unión al ADN consenso que se sabe que existe, o que existe para una clase, familia, o subfamilia de factores de transcripción particulares. En este caso, el fragmento o dominio no incluirá los aminoácidos exactos de una secuencia consenso o un sitio de unión al ADN consenso de una clase, familia o subfamilia de factores de transcripción, o los aminoácidos exactos de una secuencia consenso o un sitio de unión al ADN consenso de un factor de transcripción particular. Además, un fragmento, región, o dominio particular de un polipéptido, o un polinucleótido que codifica un polipéptido, puede estar "fuera de un dominio conservado" si todos los aminoácidos del fragmento, región, o dominio están fuera de un dominio(s) conservado definido para un polipéptido o proteína. Los dominios conservados para cada una de las secuencias de polipéptidos del clon G 47 y del clon G 2133 se listan en la Tabla 4 como se describe en el Ejemplo VII. Estos polipéptidos de la Tabla 4 presentan dominios conservados indicados específicamente por sitios de iniciación y de detención. Una comparación de las regiones de los polipéptidos de la Tabla 4 permite a alguien experto en la materia identificar un dominio(s) conservado(s) para cualquiera de los polipéptidos listados o mencionados en esta descripción, incluyendo los de la Tabla 4.

[0041] Un "rasgo" se refiere a una característica fisiológica, morfológica, bioquímica, o física de una planta o un material vegetal o una célula particular. En algunos casos, esta característica es visible a simple vista, tal como el tamaño de la semilla o de la planta, o se puede medir con técnicas bioquímicas, como la detección del contenido en proteína, fécula o aceite de semillas u hojas, o por la observación de un proceso metabólico o fisiológico, por ejemplo, midiendo la captación de dióxido de carbono, o por la observación del nivel de expresión de un gen o genes, por ejemplo, empleando análisis de transferencia de Northern, RT-PCR, ensayos de expresión de genes con biochips, o sistemas de expresión de un gen informador, o mediante observaciones agrícolas como la tolerancia al estrés, el rendimiento, o la tolerancia a patógenos. No obstante, se puede usar cualquier técnica para medir la cantidad, el nivel comparativo, o la diferencia de cualquier compuesto químico o macromolécula seleccionados en las plantas transgénicas.

[0042] La "modificación del rasgo" se refiere a una diferencia detectable en una característica de una planta que expresa ectópicamente un polinucleótido o polipéptido de la presente invención con respecto a una planta que no lo expresa, tal como una planta silvestre. En algunos casos, la modificación del rasgo se puede evaluar cuantitativamente. Por ejemplo, la modificación del rasgo puede suponer un incremento o un descenso de al menos el 2% aproximadamente en un rasgo observado (diferencia), una diferencia de al menos el 5%, una diferencia de al menos el 10% aproximadamente, una diferencia de al menos el 20% aproximadamente, de al menos el 30% aproximadamente, de al menos el 50% aproximadamente, de al menos el 70% aproximadamente, o de al menos el 100% aproximadamente, o una diferencia incluso mayor comparada con una planta silvestre. Es sabido que puede existir una variación natural en el rasgo modificado. Por tanto, la modificación del rasgo observada supone una variación de la distribución normal del rasgo en las plantas comparada con la distribución observada en la planta silvestre.

50 I. Rasgos que se pueden modificar

[0043] Las modificaciones de rasgos de interés particular incluyen las de semillas (como el embrión o endospermo), frutos, raíces, flores, hojas, tallos, brotes, plántulas o similares, incluyendo: una mayor tolerancia a condiciones ambientales como son la congelación, el frío, el calor, la sequía, la saturación de agua, la radiación y el ozono, una mayor tolerancia a las enfermedades microbianas, fúngicas o víricas, mayor tolerancia a las plagas, incluyendo nematodos, micoplasmas, parásitos de plantas superiores o similares; disminución de la sensibilidad a herbicidas, mayor tolerancia a metales pesados o mayor capacidad para captar metales pesados, mejora del crecimiento en malas condiciones de luz (por ejemplo, poca luz y/o día de corta duración), o cambios en los niveles de expresión de genes de interés. Otros fenotipos que se pueden modificar están relacionados con la producción de metabolitos vegetales, tales como las variaciones en la producción de taxol, tocoferol, tocotrienol, esteroides, fitoesteroides, vitaminas, monómeros de ceras, antioxidantes, aminoácidos, ligninas, celulosa, taninos, lípidos de prenilo (como las clorofilas y carotenoides), glucosinolatos, y terpenoides, producción mejorada de proteínas o aceites, o con la composición alterada (sobre todo en semillas), o composición modificada de azúcares (insolubles o

solubles) y/o de almidón. Las características físicas de las plantas que se pueden modificar incluyen el desarrollo celular (por ejemplo, el número de tricomas), el tamaño y el número de frutos y semillas, los rendimientos de partes de la planta tales como tallos, hojas, inflorescencias y raíces, la estabilidad de las semillas durante el almacenamiento, las características de la vaina de la semilla (por ejemplo, la susceptibilidad a la rotura), la longitud y la cantidad de los pelos de la raíz, las distancias entre los nodos, o la calidad de la cáscara de la semilla. Las características del crecimiento de las plantas que se pueden modificar incluyen la velocidad de crecimiento, la velocidad de germinación de las semillas, el vigor de las plantas y plántulas, la senescencia de hojas y flores, la esterilidad masculina, la apomixis, el tiempo de floración, la abscisión de las flores, la tasa de absorción de nitrógeno, la sensibilidad osmótica a las concentraciones de azúcares solubles, la biomasa o las características de transpiración, así como características de la arquitectura de la planta como la dominancia apical, los patrones de ramificación, el número de órganos, la identidad de los órganos, y la forma o el tamaño del órgano.

II. Los factores de transcripción modifican la expresión de genes endógenos

15 **[0044]** La expresión de genes que codifican los factores de transcripción que modifican la expresión de genes endógenos, polinucleótidos, y proteínas es muy conocida en la materia. Además, las plantas transgénicas que comprenden polinucleótidos aislados que codifican factores de transcripción también pueden modificar la expresión de genes endógenos, polinucleótidos, y proteínas. Los ejemplos incluyen Peng y col. (1997, *Genes and Development* 11:3194-3205) y Peng y col. (1999, *Nature*, 400:256-261). Además, muchos otros han demostrado que 20 un factor de transcripción de *Arabidopsis* expresado en una especie de planta exógena desencadena la misma respuesta fenotípica o una respuesta fenotípica muy similar. Véase, por ejemplo, Fu y col. (2001, *Plant Cell* 13:1791-1802); Nandi y col. (2000, *Curr. Biol.* 10:215-218); Coupland (1995, *Nature* 377:482-483); y Weigel y Nilsson (1995, *Nature* 377:482-500).

25 **[0045]** En otro ejemplo, Mandel y col. (1992, *Cell* 71:133-143) y Suzuki y col. (2001, *Plant J.* 28:409-418) enseñan que un factor de transcripción expresado en otra especie de planta desencadena la misma respuesta fenotípica o una respuesta fenotípica muy similar de la secuencia endógena, como se había predicho a menudo en estudios previos sobre factores de transcripción de *Arabidopsis* en *Arabidopsis* (véase Mandel y col., 1992, más arriba; Suzuki y col., 2001, más arriba).

30 **[0046]** Otros ejemplos incluyen Müller y col. (2001, *Plant J.* 28:169-179); Kim y col. (2001, *Plant J.* 25:247-259); Kyojuka y Shimamoto (2002, *Plant Cell Physiol.* 43:130-135); Boss y Thomas (2002, *Nature*, 416:847-850); He y col. (2000, *Transgenic Res.*, 9:223-227); y Robson y col. (2001, *Plant J.* 28:619-631).

35 **[0047]** En otro ejemplo más, Gilmour y col. (1998, *Plant J.* 16:433-442) enseña un factor de transcripción AP2 de *Arabidopsis*, el CBF1, que cuando se sobreexpresa en plantas transgénicas, incrementa la tolerancia de la planta a la congelación. Jaglo y col. (2001, *Plant Physiol.* 127:910-917) ha identificado secuencias adicionales en *Brassica napus* que codifican genes de tipo CBF y que los transcritos de estos genes se acumulan rápidamente en respuesta a bajas temperaturas. También se ha encontrado que los transcritos que codifican proteínas de tipo CBF se 40 acumulan rápidamente en respuesta a bajas temperaturas en el trigo, así como en el tomate. Un alineamiento de las proteínas CBF procedentes de *Arabidopsis*, *B. napus*, trigo, centeno, y el tomate reveló la presencia de secuencias de aminoácidos conservadas, PKK/RPAGRxKFxETRHP y DSAWR, que relaciona los dominios de unión al ADN AP2/EREBP de las proteínas y los distingue de otros miembros de la familia de proteínas AP2/EREBP. (Véase Jaglo y col., más arriba).

45 III. Polipéptidos y polinucleótidos

[0048] La presente descripción proporciona, entre otras cosas, factores de transcripción (FTs), y polipéptidos homólogos a factores de transcripción, y polinucleótidos aislados o recombinantes que codifican los polipéptidos, o 50 nuevas variantes de polipéptidos o polinucleótidos que codifican nuevas variantes de factores de transcripción derivados de las secuencias específicas proporcionadas en el presente documento. Estos polipéptidos y polinucleótidos se pueden emplear para modificar las características de una planta.

[0049] Ejemplos de polinucleótidos que codifican los polipéptidos de la descripción fueron identificados en la 55 *Arabidopsis thaliana* en la base de datos del GenBank usando programas y parámetros de análisis de secuencias disponibles para el público. Las secuencias identificadas inicialmente a continuación se caracterizaron en profundidad para identificar secuencias que comprenden cadenas de secuencias específicas correspondientes a motivos de secuencias presentes en familias de factores de transcripción conocidas. Además, se identificaron ejemplos de polinucleótidos adicionales que codifican los polipéptidos de la invención en la base de datos de plantas del GenBank usando programas y parámetros de análisis de secuencias disponibles para el público. Las secuencias 60 identificadas inicialmente a continuación se caracterizaron en profundidad para identificar secuencias que comprenden cadenas de secuencias específicas correspondientes a motivos de secuencias presentes en familias de factores de transcripción conocidos. Las secuencias de polinucleótidos que cumplen dichos criterios fueron

confirmadas como factores de transcripción.

[0050] Se identificaron polinucleótidos adicionales de la descripción por selección de *Arabidopsis thaliana* y/u otras librerías de ADNc de plantas con sondas correspondientes a factores de transcripción conocidos en condiciones de hibridación de baja rigurosidad. Posteriormente se recuperaron secuencias adicionales, incluyendo secuencias codificantes de longitud completa, mediante el procedimiento de amplificación rápida de los extremos del ADNc (RACE), usando un kit disponible comercialmente según las instrucciones del fabricante. Cuando sea necesario, se llevan a cabo múltiples rondas de RACE para aislar los extremos 5' y 3'. A continuación se recuperó el ADNc de longitud completa mediante una reacción en cadena de la polimerasa (PCR) de extremo a extremo rutinaria usando cebadores específicos para los extremos 5' y 3' aislados. En el Listado de secuencias se proporcionan secuencias de ejemplo.

[0051] Los polinucleótidos de la descripción pueden ser o fueron expresados ectópicamente en plantas que sobreexpresan o están inactivados y se observaron los cambios en la(s) característica(s) o rasgo(s) de las plantas.

[0052] Los polinucleótidos de la descripción pueden ser o fueron expresados ectópicamente en células de plantas que sobreexpresan y se observaron los cambios en los niveles de expresión de una serie de genes, polinucleótidos, y/o proteínas de las células de la planta. Por tanto, los polinucleótidos y polipéptidos se pueden emplear para variar los niveles de expresión de genes, polinucleótidos, y/o proteínas de las plantas.

IV. Producción de polipéptidos

[0053] Los polinucleótidos de la descripción incluyen secuencias que codifican factores de transcripción y polipéptidos homólogos a factores de transcripción y secuencias complementarias a ellos, así como fragmentos únicos de secuencias codificantes, o secuencias complementarias a ellos. Dichos polinucleótidos pueden ser, por ejemplo, ADN o ARN, por ejemplo, ARNm, ARNc, ARN sintético, ADN genómico, ADNc, ADN sintético, oligonucleótidos, etc. Los polinucleótidos son de doble cadena o de cadena sencilla, e incluyen cualquiera o ambas secuencias en dirección directa (es decir, codificante) y secuencias en dirección inversa (es decir, no codificante, complementaria). Los polinucleótidos incluyen la secuencia codificante de un factor de transcripción, o un polipéptido homólogo a un factor de transcripción, aislada, en combinación con secuencias codificantes adicionales (por ejemplo, un marcador para su purificación, una señal de localización, en forma de proteína de fusión, en forma de preproteína, o similar), en combinación con secuencias no codificantes (por ejemplo, intrones o inteínas, elementos reguladores como promotores, potenciadores, terminadores, y similares), y/o en un vector o en el entorno de un hospedador en el que el polinucleótido que codifica un factor de transcripción o un polipéptido homólogo al factor de transcripción es un gen endógeno o exógeno.

[0054] Existen una variedad de procedimientos para la producción de los polinucleótidos de la descripción. Los procedimientos para la identificación y el aislamiento de clones de ADN son muy conocidos por los expertos en la materia, y se describen en, por ejemplo, Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA ("Berger"); Sambrook y col., Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, 1989 ("Sambrook") y Current Protocols in Molecular Biology, F. M. Ausubel y col., eds., Current Protocols, un proyecto conjunto entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (suplementos hasta el año 2000) ("Ausubel").

[0055] Alternativamente, polinucleótidos como los descritos en el presente documento se pueden producir mediante una variedad de procedimientos de amplificación *in vitro* adaptados a la presente invención mediante la selección apropiada de cebadores específicos o degenerados. Ejemplos de protocolos suficientes para dirigir a personas expertas en los procedimientos de amplificación *in vitro*, incluyendo la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR), la amplificación con la Q-beta-replicasa y otras técnicas mediadas por la ARN polimerasa (por ejemplo, NASBA), por ejemplo, para la producción de los ácidos nucleicos homólogos de la invención se pueden encontrar en Berger (más arriba), Sambrook (más arriba), y Ausubel (más arriba), así como en Mullis y col., (1987) PCR Protocols A Guide to Methods and Applications (Innis y col. eds) Academic Press Inc. San Diego, CA (1990) (Innis). Procedimientos mejorados para la clonación *in vitro* de ácidos nucleicos amplificados se describen en Wallace y col., patente de EE.UU. N° 5.426.039. Procedimientos mejorados para la amplificación por PCR de ácidos nucleicos grandes se resumen en Cheng y col. (1994) Nature 369: 684-685 y las referencias allí citadas, en los que se generan amplicones de PCR de hasta 40 kb. La persona experta apreciará que esencialmente cualquier ARN se puede convertir en un ADN de doble cadena adecuado para su digestión por restricción, su expansión por PCR y su secuenciación usando la transcriptasa inversa y una polimerasa. Véase, por ejemplo, Ausubel, Sambrook y Berger, todos ellos más arriba.

[0056] Alternativamente, se pueden ensamblar polinucleótidos y oligonucleótidos como los descritos en el presente documento a partir de fragmentos producidos mediante procedimientos de síntesis en fase sólida. Normalmente, se sintetizan individualmente fragmentos de hasta 100 bases aproximadamente y a continuación se ligan enzimática o químicamente para producir una secuencia deseada, por ejemplo, un polinucleótido que codifica todo o parte de un factor de transcripción. Por ejemplo, se describe la síntesis química usando el método de la fosforoamidita, por ejemplo, de Beaucage y col. (1981) *Tetrahedron Letters* 22:1859-1869; y Matthes y col. (1984) *EMBO J.* 3:801-805. Según dichos procedimientos, los oligonucleótidos se sintetizan, se purifican, se hibridan a sus cadenas complementarias, se ligan y a continuación opcionalmente se clonan en vectores adecuados. Y si así se desea, polinucleótidos y polipéptidos como los descritos en el presente documento se pueden encargar en cualquiera de una serie de distribuidores comerciales.

V. Secuencias homólogas

[0057] Las secuencias homólogas, es decir, que comparten una identidad o similitud de secuencia significativa, con aquellas proporcionadas en el Listado de secuencias, derivadas de *Arabidopsis thaliana* o de otras plantas seleccionadas también son un aspecto de la invención. Las secuencias homólogas pueden proceder de cualquier planta incluyendo monocotiledóneas y dicotiledóneas y en particular especies de plantas de importancia agrícola, incluyendo pero no limitado a, cultivos como soja, trigo, maíz, patata, algodón, arroz, colza (incluyendo canola), girasol, alfalfa, caña de azúcar y hierba; o frutas y vegetales, tales como plátano, mora, arándano, fresa, y frambuesa, cantalupo, zanahoria, coliflor, café, pepino, berenjena, uva, melón Galia, lechuga, mango, melón, cebolla, papaya, guisantes, pimientos, piña, calabaza, espinaca, calabacín, maíz dulce, tabaco, tomate, sandía, frutas rosáceas (como manzanas, melocotones, peras, cerezas y ciruelas), y coles vegetales (como brócoli, repollo, coliflor, coles de Bruselas, y colinabo). Otros cultivos, frutas y vegetales cuyo fenotipo se puede cambiar incluyen la cebada, el centeno, el mijo, el sorgo, la grosella, el aguacate, los cítricos como naranjas, limones, pomelos y mandarinas, alcachofas, cerezas, frutos secos como nueces y cacahuets, escarola, puerro, raíces, tales como arrurruz, la remolacha, la yuca, el nabo, el rábano, el ñame y la batata, y las alubias. Las secuencias homólogas también pueden proceder de especies leñosas, como el pino, el álamo y el eucalipto, o la menta u otros labiados.

Ortólogos y parálogos

[0058] Los expertos en la materia conocen diferentes procedimientos para la identificación y definición de estas secuencias funcionalmente homólogas. Hay descritos tres procedimientos generales para la definición de parálogos y ortólogos; se puede identificar un parólogo u ortólogo u homólogo mediante uno o más de los procedimientos descritos a continuación.

[0059] Los ortólogos y parálogos son genes evolutivamente relacionados que tienen secuencias similares y funciones similares. Los ortólogos son genes estructuralmente relacionados en especies diferentes que proceden de un evento de especiación. Los parálogos son genes estructuralmente relacionados dentro de una misma especie que proceden de un evento de duplicación.

[0060] Dentro de una misma especie de plantas, la duplicación génica puede producir dos copias de un gen particular, dando lugar a dos o más genes con secuencias similares y funciones similares conocidos como parálogos. Por tanto, un parólogo es un gen similar con una función similar dentro de la misma especie. Los parálogos normalmente están agrupados juntos o en el mismo clado (un grupo de genes similares) cuando se analiza la filogenia de una familia de genes usando programas como CLUSTAL (Thompson y col. (1994) *Nucleic Acids Res.* 22:4673-4680; Higgins y col. (1996) *Methods Enzymol.* 266 383-402). También se pueden identificar grupos de genes similares con un análisis BLAST por pares (Feng y Doolittle (1987) *J. Mol. Evol.* 25:351-360). Por ejemplo, los factores de transcripción de un clado con un dominio muy similar a MADS procedente de *Arabidopsis* comparten una función común en el tiempo de floración (Ratcliffe y col. (2001) *Plant Physiol.* 126:122-132), y un grupo de factores de transcripción con un dominio muy similar a AP2 procedente de *Arabidopsis* están involucrados en la tolerancia de las plantas a la congelación (Gilmour y col. (1998) *Plant J.* 16:433-442). El análisis de grupos de genes similares con funciones similares que están dentro de un clado puede dar subsecuencias que son particulares al clado. Estas subsecuencias, conocidas como secuencias consenso, se pueden usar no sólo para definir las secuencias dentro de cada clado, sino que definen las funciones de estos genes; los genes dentro de un clado pueden contener secuencias parálogas u ortólogas que comparten la misma función. (Véase también, por ejemplo, Mount, D.W. (2001) *Bioinformatics: Sequence and Genome Analysis* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, página 543).

[0061] La especiación, la producción de nuevas especies a partir de especies parentales, también puede dar lugar a dos o más genes con secuencias similares y funciones similares. Estos genes, denominados ortólogos, a menudo presentan una función idéntica en sus plantas hospedadoras y a menudo se pueden intercambiar entre especies sin perder su función. Debido a que las plantas tienen ancestros comunes, muchos genes de cualquier

especie de plantas tendrá un gen ortólogo correspondiente en otra especie de planta. Una vez construido un árbol filogenético para una familia de genes de una especie usando un programa como CLUSTAL (Thompson y col. (1994) *Nucleic Acids Res.* 22:4673-4680; Higgins y col. (1996) *Methods Enzymol.* 266:383-402), las secuencias ortólogas potenciales se pueden colocar en el árbol filogenético y se puede determinar su relación con genes de la especie de interés. Una vez identificado el par ortólogo, se puede determinar la función del ortólogo de prueba determinando la función del ortólogo de referencia.

[0062] La degeneración del código genético permite variaciones importantes en la secuencia de nucleótidos de un polinucleótido mientras se mantiene la secuencia de aminoácidos de la proteína codificada. Los dominios conservados dentro de una familia de factores de transcripción presenta una identidad de secuencia de al menos el 95%, o de al menos el 98% aproximadamente. Los factores de transcripción que son homólogos a las secuencias listadas comparten una identidad de secuencia de aminoácidos de al menos el 95% con respecto a la longitud total del polipéptido o del homólogo.

[0063] El porcentaje de identidad se puede determinar electrónicamente, por ejemplo, usando el programa MEGALIGN (DNASTAR, Inc. Madison, Wis.). El programa MEGALIGN puede crear alineamientos entre dos o más secuencias según los diferentes procedimientos, por ejemplo, el método de clustal. (Véase, por ejemplo, Higgins, D. G. y P. M. Sharp (1988) *Gene* 73:237-244). El algoritmo clustal agrupa secuencias en agrupamientos examinando las distancias entre todos los pares. Los agrupamientos se alinean por pares y a continuación en grupos. Se pueden usar otros algoritmos o programas de alineamiento, incluyendo FASTA, BLAST, o ENTREZ, FASTA y BLAST. Éstos están disponibles como parte del paquete para el análisis de secuencias GCG (Universidad de Wisconsin, Madison, Wis.), y se pueden usar con o sin los ajustes por defecto. ENTREZ está disponible en el National Center for Biotechnology Information. En una forma de realización, el porcentaje de identidad de dos secuencias se puede determinar mediante el programa GCG con una ponderación de la separación de 1, por ejemplo, cada separación entre aminoácidos se pondera como si fuera un solo desapareamiento del aminoácido o nucleótido entre las dos secuencias (véase el documento USPN 6.262.333).

[0064] En *Methods in Enzymology*, vol. 266: *Computer Methods for Macromolecular Sequence Analysis* (1996), ed. Doolittle, Academic Press, Inc., San Diego, Calif., USA, se describen otras técnicas para el alineamiento. Preferentemente, para alinear las secuencias se utiliza un programa de alineamiento que permita separaciones en la secuencia. El algoritmo Smith-Waterman es un tipo de algoritmo que permite separaciones en los alineamientos de secuencias. Véase, *Methods Mol. Biol.* 70: 173-187 (1997). Además, para alinear secuencias se puede usar el programa GAP que utiliza el método de alineamiento de Needleman y Wunsch. Una estrategia de búsqueda alternativa utiliza el *software* MPSRCH, que corre sobre un ordenador MASPAR. El MPSRCH utiliza un algoritmo Smith-Waterman para puntuar secuencias en un ordenador de computación masiva en paralelo. Este enfoque mejora la capacidad de captar apareamientos remotamente relacionados, y es especialmente tolerante con separaciones pequeñas y errores en la secuencia de nucleótidos. Se pueden usar secuencias de aminoácidos codificadas por ácidos nucleicos para la búsqueda tanto en bases de datos de proteínas como de ADN.

[0065] El porcentaje de similitud entre dos secuencias de polipéptidos, por ejemplo, la secuencia A y la secuencia B, se calcula dividiendo la longitud de la secuencia A, menos el número de restos separados en la secuencia A, menos el número de restos separados en la secuencia B, por la suma de los restos apareados entre la secuencia A y la secuencia B, multiplicado por 100. La separaciones con poca o ninguna similitud entre las dos secuencias de aminoácidos no están incluidas en la determinación del porcentaje de similitud. El porcentaje de identidad entre secuencias de polinucleótidos también se puede contar o calcular mediante otros procedimientos conocidos en la materia, por ejemplo, el método de Jotun Hein. (Véase, por ejemplo, Hein, J. (1990) *Methods Enzymol.* 183:626-645). La identidad entre secuencias también se puede determinar mediante otros procedimientos conocidos en la materia, por ejemplo, variando las condiciones de hibridación (véase Solicitud de patente de EE.UU. N° 20010010913).

[0066] Así, en el presente documento se describen procedimientos para la identificación de una secuencia similar o paróloga u ortóloga u homóloga a uno o más polinucleótidos como los indicados en el presente documento, o uno o más polipéptidos diana codificados por los polinucleótidos, o indicados en otra parte del presente documento y puede incluir la unión o asociación del fenotipo o la función génica de una planta dada con una secuencia. En los procedimientos, se proporciona una base de datos de secuencias (local o remota en una intranet o en internet) y se realiza una búsqueda contra la base de datos de secuencias usando las secuencias relevantes del presente documento y se asocian los fenotipos o funciones génicas de la planta.

[0067] Además, se pueden usar una o más secuencias de polinucleótidos o uno o más polipéptidos codificados por las secuencias de polinucleótidos para buscar contra BLOCKS (Bairoch y col. (1997) *Nucleic Acids Res.* 25:217-221), PFAM, y otras bases de datos que contienen motivos, secuencias y funciones génicas previamente identificados y anotados. Se pueden usar procedimientos que buscan patrones en la secuencia primaria con penalizaciones de separación en la estructura secundaria (Smith y col. (1992) *Protein Engineering* 5:35-51)

además de algoritmos como Basic Local Alignment Search Tool (BLAST; Altschul, S. F. (1993) *J. Mol. Evol.* 36:290-300; Altschul y col. (1990) más arriba), BLOCKS (Henikoff, S. y Henikoff, G. J. (1991) *Nucleic Acids Research* 19:6565-6572), Hidden Markov Models (HMM; Eddy, S. R. (1996) *Cur. Opin. Str. Biol.* 6:361-365; Sonnhammer y col. (1997) *Proteins* 28:405-420), y similares, para manipular y analizar secuencias de polinucleótidos y polipéptidos codificados por los polinucleótidos. Estas bases de datos, algoritmos y otros procedimientos son muy conocidos en la materia y están descritos en Ausubel y col. (1997; *Short Protocols in Molecular Biology*, John Wiley & Sons, Nueva York N.Y., unidad 7,7) y en Meyers, R. A. (1995; *Molecular Biology and Biotechnology*, Wiley VCH, Nueva York N.Y., p 856-853).

10 **[0068]** Además, se pueden utilizar procedimientos que usan el alineamiento manual de secuencias similares u homólogas a una o más secuencias de polinucleótidos o uno o más polipéptidos codificados por las secuencias de polinucleótidos para identificar regiones de dominios similares y conservados. Dichos procedimientos manuales son muy conocidos por los expertos en la materia y pueden incluir, por ejemplo, la comparación de la estructura terciaria entre una secuencia de polipéptidos codificada por un polinucleótido que tiene una función conocida con una
15 secuencia de polipéptidos codificada por una secuencia de polinucleótidos que tiene una función aún por determinar. Dichos ejemplos de estructuras terciarias pueden comprender estructuras predichas de hélices alfa, láminas beta, hélices anfipáticas, motivo de cremallera de leucinas, motivo de dedos de cinc, regiones ricas en prolina, motivos con repetición de cisteínas, y similares.

20 VI. Identificación de polinucleótidos o ácidos nucleicos por hibridación

[0069] Los polinucleótidos homólogos a las secuencias ilustradas en el Listado de secuencias y tablas se pueden identificar, por ejemplo, mediante hibridación entre sí en condiciones rigurosas o muy rigurosas. Los polinucleótidos de cadena sencilla se hibridan cuando se asocian en función de una variedad de fuerzas físico-químicas bien caracterizadas, tales como los puentes de hidrógeno, la exclusión del disolvente, el apilamiento de bases y similares. La rigurosidad de una hibridación refleja el grado de identidad de secuencia de los ácidos nucleicos involucrados, de manera que cuanto más rigurosas son las condiciones, más similares son las dos cadenas de polinucleótidos. La rigurosidad está influida por una variedad de factores, incluyendo la temperatura, la concentración y composición salina, la presencia de aditivos orgánicos y no orgánicos; los disolventes, etc.
25 presentes tanto en las disoluciones de hibridación como de lavado y en las incubaciones (y su número), como se describe con más detalle en las referencias citadas anteriormente. La invención engloba las secuencias de polinucleótidos que son capaces de hibridarse a las secuencias de polinucleótidos reivindicadas, y en particular, a las mostradas en las SEQ ID N°: 860; 802; 240; 274; 558; 24; 1120; 44; 460; 286; 120; 130; 134; 698; 832; 580; 612; 48, y sus fragmentos en diferentes condiciones de rigurosidad. (Véase, por ejemplo, Wahl, G. M. y S. L. Berger
30 (1987) *Methods Enzymol.* 152:399-407; Kimmel, A. R (1987) *Methods Enzymol.* 152: 507-511). Las estimaciones de la homología se obtienen con la hibridación de ADN-ADN o ADN-ARN en condiciones de rigurosidad como es bien sabido por los expertos en la materia (Hames y Higgins, Eds. (1985) *Nucleic Acid Hybridisation*, IRL Press, Oxford, R.U.). Las condiciones de rigurosidad se pueden ajustar para seleccionar fragmentos moderadamente similares, tales como secuencias homólogas procedentes de organismos remotamente relacionados, a fragmentos muy
40 similares, tales como genes que duplican enzimas funcionales procedentes de organismos íntimamente relacionados. Los lavados después de la hibridación determinan las condiciones de rigurosidad.

[0070] Además de las secuencias de nucleótidos listadas en la Tabla 4, se pueden identificar y aislar ADNc de longitud completa, ortólogos, parálogos y homólogos de las presentes secuencias de nucleótidos usando
45 procedimientos muy conocidos. Los ortólogos, parálogos y homólogos de librerías de ADNc de las presentes secuencias de nucleótidos se pueden seleccionar usando procedimientos de hibridación para determinar su utilidad como dianas de hibridación o sondas de amplificación.

[0071] Un ejemplo de condiciones de hibridación rigurosas para la hibridación de ácidos nucleicos complementarios que tienen más de 100 restos complementarios en un filtro de transferencia de Southern o Northern es de 5°C a 20°C inferior aproximadamente al punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y pH definidos. La T_m es la temperatura (en condiciones de fuerza iónica y pH definidos) a la que el 50% de la secuencia diana se hibrida con una sonda perfectamente apareada. Las moléculas de ácidos nucleicos que se hibridan en condiciones rigurosas normalmente se hibridarán a una sonda basada en el ADNc completo o en
50 porciones seleccionadas, por ejemplo, a una subsecuencia única del ADNc en condiciones de lavado de 0,2x SSC a 2,0 x SSC, SDS al 0,1% a 50-65°C. Por ejemplo, en condiciones de rigurosidad elevada son 0,2 x SSC, SDS al 0,1% a 65°C aproximadamente. Las condiciones de rigurosidad ultraelevada serán las mismas condiciones excepto que la temperatura de lavado se incrementará de 3°C aproximadamente a 5°C aproximadamente, y las condiciones de rigurosidad ultra-ultraelevada serán las mismas condiciones excepto que la temperatura de lavado se incrementará
55 de 6°C aproximadamente a 9°C aproximadamente. Para la identificación de homólogos menos íntimamente relacionados los lavados se pueden llevar a cabo a una temperatura inferior, por ejemplo, a 50°C. En general, las condiciones de rigurosidad se incrementan aumentando la temperatura de lavado y/o reduciendo la concentración de SSC, como es sabido en la materia.

[0072] En otro ejemplo, una concentración salina rigurosa normalmente será inferior a NaCl 750 mM y citrato trisódico 75 mM aproximadamente, preferentemente inferior a NaCl 500 mM y citrato trisódico 50 mM aproximadamente, y lo más preferentemente inferior a NaCl 250 mM y citrato trisódico 25 mM aproximadamente. La hibridación en condiciones de baja rigurosidad se puede obtener en ausencia del disolvente orgánico, por ejemplo, formamida, mientras que la hibridación en condiciones de alta rigurosidad se puede obtener en presencia de al menos el 35% de formamida aproximadamente, y lo más preferentemente de al menos el 50% de formamida aproximadamente. Las condiciones de temperatura rigurosas normalmente incluirán temperaturas de al menos 30°C aproximadamente, más preferentemente de al menos 37°C aproximadamente, y lo más preferentemente de al menos 42°C aproximadamente. Parámetros variables adicionales, tales como el tiempo de hibridación, la concentración de detergente, por ejemplo, dodecilsulfato sódico (SDS), y la inclusión o exclusión de un ADN portador, son muy conocidos por los expertos en la materia. Combinando estas condiciones diferentes se consiguen diversos niveles de rigurosidad según se necesite. En una forma de realización preferida, la hibridación se producirá a 30°C en NaCl 750 mM, citrato trisódico 75 mM, y SDS al 1%. En una forma de realización más preferida, la hibridación se producirá a 37°C en NaCl 500 mM, citrato trisódico 50 mM, SDS al 1%, formamida al 35%, y 100 µg/ml de ADN de esperma de salmón desnaturalizado (ADNss). En la forma de realización más preferida, la hibridación se producirá a 42°C en NaCl 250 mM, citrato trisódico 25 mM, SDS al 1%, formamida al 50%, y 200 µg/ml de ADN de esperma de salmón desnaturalizado (ADNss). Variaciones útiles de estas condiciones serán obvias para los expertos en la materia.

20

[0073] En las etapas de lavado que siguen a la hibridación también se pueden variar las condiciones de rigurosidad. Las condiciones de rigurosidad del lavado se pueden definir por la concentración salina y por la temperatura. Como anteriormente, la rigurosidad del lavado se puede incrementar reduciendo la concentración salina o incrementando la temperatura. Por ejemplo, una concentración salina rigurosa para las etapas de lavado preferentemente será inferior a NaCl 30 mM y citrato trisódico 3 mM aproximadamente, y lo más preferentemente inferior a NaCl 15 mM y citrato trisódico 1,5 mM aproximadamente. Las condiciones de temperatura rigurosas para las etapas de lavado normalmente incluirán temperaturas de al menos 25°C aproximadamente, y más preferentemente de al menos 42°C aproximadamente. Otro conjunto preferido de condiciones muy rigurosas usa dos lavados finales de 0,1X SSC, SDS al 0,1% a 65°C. Los lavados con las condiciones de alta rigurosidad más preferidos son al menos a 68°C aproximadamente. Por ejemplo, en una forma de realización preferida, las etapas de lavado se producirán a 25°C en NaCl 30 mM, citrato trisódico 3 mM, y SDS al 0,1%. En una forma de realización más preferida, las etapas de lavado se producirán a 42°C en NaCl 15 mM, citrato trisódico 1,5 mM, y SDS al 0,1%. En la forma de realización más preferida, las etapas de lavado se producirán a 68°C en NaCl 15 mM, citrato trisódico 1,5 mM, y SDS al 0,1%. Variaciones útiles de estas condiciones serán obvias para los expertos en la materia (véase Solicitud de patente de EE.UU. N° 20010010913).

35

[0074] Como otro ejemplo más, se pueden seleccionar condiciones rigurosas de manera que un oligonucleótido que es perfectamente complementario con el oligonucleótido codificante se hibride al oligonucleótido codificante con una relación de señal a ruido al menos 5-10x superior aproximadamente a la relación para la hibridación del oligonucleótido perfectamente complementario a un ácido nucleico que codifica un factor de transcripción conocido a partir de la fecha de presentación de la solicitud. Las condiciones se pueden seleccionar de manera que se observa una relación de señal a ruido superior en el ensayo particular usado, por ejemplo, 15x, 25x, 35x, 50x aproximadamente o superior. Por consiguiente, el ácido nucleico objeto se hibrida al único oligonucleótido codificante con una relación de señal a ruido al menos 2x superior comparada con la hibridación del oligonucleótido codificante con un ácido nucleico que codifica un polipéptido conocido. De nuevo, se pueden seleccionar relaciones de señal a ruido superiores, por ejemplo, de 5x, 10x, 25x, 35x, 50x aproximadamente o superior. La señal particular dependerá del marcador usado en el ensayo pertinente, por ejemplo, un marcador fluorescente, un marcador colorimétrico, un marcador radiactivo, o similar.

50

[0075] Alternativamente, se pueden obtener polipéptidos homólogos al factor de transcripción seleccionando una librería de expresión usando anticuerpos específicos para uno o más factores de transcripción. Con la presente disposición del factor de transcripción descrito, y secuencias de ácidos nucleicos homólogas al factor de transcripción, el polipéptido(s) codificado se puede expresar y purificar en un sistema de expresión heterólogo (por ejemplo, *E. coli*) y se puede usar para generar anticuerpos (monoclonales y policlonales) específicos para el polipéptido(s) en cuestión. También se puede generar anticuerpos contra péptidos sintéticos derivados de secuencias de un factor de transcripción, o de secuencias homólogas a un factor de transcripción o de secuencias de aminoácidos. Los procedimientos para generar anticuerpos son muy conocidos en la materia y están descritos en Harlow y Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Nueva York. A continuación se pueden usar dichos anticuerpos para seleccionar una librería de expresión producida a partir de la planta de la que se desea clonar homólogos del factor de transcripción adicionales, usando los procedimientos descritos anteriormente. Los ADNc seleccionados se pueden confirmar por secuenciación y actividad enzimática.

60

VII. Variaciones de secuencia

[0076] Los expertos en la materia se darán cuenta fácilmente que cualquiera de una variedad de secuencias de polinucleótidos son capaces de codificar los factores de transcripción y polipéptidos homólogos a los factores de transcripción. Debido a la degeneración del código genético, muchos polinucleótidos diferentes pueden codificar polipéptidos idénticos y/o sustancialmente similares, además de las secuencias ilustradas en el Listado de secuencias. Los ácidos nucleicos que tienen una secuencia que difiere de las secuencias mostradas en el Listado de secuencias, o secuencias complementarias, que codifican péptidos funcionalmente equivalentes (es decir, péptidos que tienen algún grado de actividad biológica equivalente o similar), pero que difieren en secuencia de la secuencia mostrada en el Listado de secuencias debido a la degeneración del código genético, también están dentro del alcance de la invención.

[0077] Las secuencias de polinucleótidos alteradas que codifican polipéptidos incluyen aquellas secuencias con eliminaciones, inserciones, o sustituciones de diferentes nucleótidos, que dan como resultado un polinucleótido que codifica un polipéptido con al menos una característica funcional de los presentes polipéptidos. Dentro de esta definición están incluidos polimorfismos que pueden o pueden no ser fácilmente detectables usando una sonda de oligonucleótidos particular del polinucleótido que codifica los presentes polipéptidos, y la hibridación impropia o inesperada a variantes alélicas, con un locus distinto del locus cromosómico normal para la secuencia de polinucleótidos que codifica los presentes polipéptidos.

[0078] Variante alélica se refiere a cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica aparece de forma natural por mutación, y puede dar como resultado un polimorfismo fenotípico en poblaciones. Las mutaciones de los genes pueden ser silenciosas (es decir, sin ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos que tienen la secuencia de aminoácidos alterada. El término variante alélica también se usa en el presente documento para denotar una proteína codificada por una variante alélica de un gen. Variante de procesamiento alternativo se refiere a formas alternativas del ARN transcrito a partir de un gen. La variación por procesamiento alternativo aparece de forma natural con el uso de sitios de procesamiento alternativo dentro de la molécula de ARN transcrita, o menos habitualmente entre moléculas de ARN transcritas por separado, y puede dar como resultado varios ARNm transcritos a partir del mismo gen. Las variantes de procesamiento alternativo pueden codificar polipéptidos que tienen una secuencia de aminoácidos alterada. El término variante de procesamiento alternativo también se usa en el presente documento para denotar una proteína codificada por una variante de procesamiento alternativo de un ARNm transcrito a partir de un gen.

[0079] Los ADNc generados a partir de ARNm que han sufrido procesamiento alternativo, que retienen las propiedades del factor de transcripción, están incluidos dentro del alcance de la presente invención, como lo están los polipéptidos codificados por dichos ADNc y ARNm. Las variantes alélicas y las variantes de procesamiento alternativo de estas secuencias se pueden clonar con sondas de ADNc o librerías genómicas procedentes de diferentes organismos o tejidos individuales según los procedimientos habituales conocidos en la materia (véase USPN 6.388.064).

[0080] Por ejemplo, la Tabla 1 ilustra que los codones AGC, AGT, TCA, TCC, TCG, y TCT codifican todos el mismo aminoácido: serina. Por consiguiente, en cada posición de la secuencia donde haya un codón que codifique serina, se puede usar cualquiera de las secuencias de trinucleótidos anteriores sin alterar el polipéptido codificado.

Tabla 1

Aminoácido		Posibles codones					
Alanina	Ala A	GCA	GCC	GCG	GCU		
Cisteína	Cys C	TGC	TGT				
Ácido aspártico	Asp D	GAC	GAT				
Ácido glutámico	Glu E	GAA	GAG				
Fenilalanina	Phe F	TTC	TTT				
Glicina	Gly G	GGA	GGC	GGG	GGT		
Histidina	His H	CAC	CAT				
Isoleucina	Ile I	ATA	ATC	ATT			
Lisina	Lys K	AAA	AAG				
Leucina	Leu L	TTA	TTG	CTA	CTC	CTG	CTT
Metionina	Met M	ATG					
Asparagina	Asn N	AAC	AAT				
Prolina	Pro P	CCA	CCC	CCG	CCT		
Glutamina	Gln Q	CAA	CAG				
Arginina	Arg R	AGA	AGG	CGA	CGC	CGG	CGT

Serina	Ser	S	AGC	AGT	TCA	TCC	TCG	TCT
Treonina	Thr	T	ACA	ACC	ACG	ACT		
Valina	Val	V	GTA	GTC	GTG	GTT		
Triptófano	Trp	W	TGG					
Tirosina	Tyr	Y	TAC	TAT				

[0081] Las alteraciones de secuencia que no varían la secuencia de aminoácidos codificada por el polinucleótido se denominan variaciones "silenciosas". Con la excepción de los codones ATG y TGG, que codifican metionina y triptófano, respectivamente, se puede sustituir cualquiera de los posibles codones para un mismo aminoácido mediante una variedad de técnicas, por ejemplo, mutagénesis dirigida de sitio, disponibles en la materia. Por consiguiente, todas y cada una de dichas variaciones de una secuencia seleccionadas en la Tabla anterior son una característica de la descripción.

[0082] Además de las variaciones silenciosas, se pueden introducir otras variaciones conservativas que alteren uno, o unos pocos aminoácidos en el polipéptido codificado, sin alterar la función del polipéptido. Estas variantes conservativas son, asimismo, una característica de la invención.

[0083] Por ejemplo, también están contempladas por la invención las sustituciones, eliminaciones e inserciones introducidas en las secuencias proporcionadas en el Listado de secuencias. Dichas modificaciones de la secuencia se pueden introducir mediante ingeniería genética en una secuencia por mutagénesis dirigida de sitio (Wu (ed.) Meth. Enzymol. (1993) vol. 217, Academic Press) o los otros procedimientos indicados a continuación. Las sustituciones de aminoácidos normalmente son de restos únicos; las inserciones habitualmente serán del orden de entre 1 y 10 restos aminoácidos aproximadamente; y las eliminaciones comprenderán entre 1 y 30 restos aproximadamente. En formas de realización preferidas, las eliminaciones o inserciones se introducen en pares adyacentes, por ejemplo, una eliminación de dos restos o una inserción de dos restos. Se pueden combinar sustituciones, eliminaciones, inserciones o cualquiera de sus combinaciones para llegar a una secuencia. Las mutaciones que se introducen en el polinucleótido que codifica el factor de transcripción no deben situar la secuencia fuera del marco de lectura y no deben crear regiones complementarias que pudieran producir una estructura de ARNm secundaria. Preferentemente, el polipéptido codificado por el ADN desempeña la función deseada.

[0084] Las sustituciones conservativas son aquellas en las que se ha eliminado al menos un resto de la secuencia de aminoácidos y se ha insertado en su lugar un resto diferente. Generalmente dichas sustituciones se introducen de acuerdo con la Tabla 2 cuando se desea mantener la actividad de la proteína. La Tabla 2 muestra aminoácidos que se pueden sustituir por otro aminoácido en una proteína y que normalmente se consideran sustituciones conservativas.

[0085]

Tabla 2

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

[0086] Sustituciones similares son aquellas en las que se ha eliminado al menos un resto en la secuencia de aminoácidos y se ha insertado en su lugar un resto diferente. Generalmente dichas sustituciones se introducen de acuerdo con la Tabla 3 cuando se desea mantener la actividad de la proteína. La Tabla 3 muestra aminoácidos que se pueden sustituir por otro aminoácido en una proteína y que normalmente se consideran sustituciones estructurales y funcionales. Por ejemplo, un resto de la columna 1 en la Tabla 3 se puede sustituir por un resto de la columna 2; además, un resto de la columna 2 en la Tabla 3 se puede sustituir por el resto de la columna 1.

Tabla 3

Resto	Substituciones similares
Ala	Ser; Thr; Gly; Val; Leu; Ile
Arg	Lys; His; Gly
Asn	Gln; His; Gly; Ser; Thr
Asp	Glu; Ser; Thr
Gln	Asn; Ala
Cys	Ser; Gly
Glu	Asp
Gly	Pro; Arg
His	Asn; Gln; Tyr; Phe; Lys; Arg
Ile	Ala; Leu; Val; Gly; Met
Leu	Ala; Ile; Val; Gly; Met
Lys	Arg; His; Gln; Gly; Pro
Met	Leu; Ile; Phe
Phe	Met; Leu; Tyr; Trp; His; Val; Ala
Ser	Thr; Gly; Asp; Ala; Val; Ile; His
Thr	Ser; Val; Ala; Gly
Trp	Tyr; Phe; His
Tyr	Trp; Phe; His
Val	Ala; Ile; Leu; Gly; Thr; Ser; Glu

10

[0087] Se pueden seleccionar sustituciones que sean menos conservativas que las de la Tabla 2 escogiendo restos que difieran más significativamente en su efecto de mantener (a) la estructura del esqueleto polipeptídico en el área de sustitución, por ejemplo, como conformación en lámina o en hélice, (b) la carga o hidrofobicidad de la molécula en el sitio diana, o (c) el volumen de la cadena lateral. Las sustituciones que de manera general se espera que produzcan los mayores cambios en las propiedades de la proteína serán aquellas en las que (a) un resto hidrófilo, por ejemplo, serilo o treonilo, es sustituido por (o mediante) un resto hidrófobo, por ejemplo, leucilo, isoleucilo, fenilalanilo, valilo o alanilo; (b) una cisteína o prolina es sustituida por (o mediante) cualquier otro resto; (c) un resto que tiene una cadena lateral electropositiva, por ejemplo, lisilo, arginilo, o histidilo, es sustituido por (o mediante) un resto electronegativo, por ejemplo, glutamilo o aspartilo; o (d) un resto que tiene una cadena lateral voluminosa, por ejemplo, fenilalanina, es sustituido por (o mediante) uno que no tiene cadena lateral, por ejemplo, glicina.

VIII. Modificación adicional de secuencias - mutación/evolución forzada

[0088] Además de generar sustituciones silenciosas o conservativas como las indicadas anteriormente, la presente invención incluye opcionalmente procedimientos para la modificación de las secuencias del Listado de secuencias. En los procedimientos, se usan procedimientos de modificación de ácidos nucleicos o proteínas para alterar la secuencia dada para producir nuevas secuencias y/o para modificar química o enzimáticamente secuencias dadas y así variar las propiedades de los ácidos nucleicos o proteínas.

30

[0089] De esta forma, se modifican secuencias dadas de ácidos nucleicos, por ejemplo, según los procedimientos de mutagénesis o evolución artificial habituales para producir secuencias modificadas. Las secuencias modificadas se pueden generar usando polinucleótidos naturales aislados y purificados a partir de cualquier organismo o se pueden sintetizar a partir de composiciones purificadas y compuestos químicos usando medios químicos muy conocidos por los expertos en la materia. Por ejemplo, Ausubel, más arriba, suministra detalles adicionales sobre los procedimientos de mutagénesis. Los procedimientos de evolución artificial forzada están descritos, por ejemplo, en Stemmer (1994) Nature 370:389-391, Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751, y las patentes de EE.UU. 5.811.238, 5.837.500, y 6.242.568. los procedimientos para la manipulación genética de factores de transcripción sintéticos y otros polipéptidos están descritos, por ejemplo, en Zhang y col. (2000) J. Biol. Chem. 275:33850-33860, Liu y col. (2001) J. Biol. Chem. 276:11323-11334, e Isalan y col. (2001) Nature Biotechnol. 19:656-660. También están disponibles muchos otros procedimientos de mutación y

40

evolución y se espera que estén dentro de las competencias de la persona experta.

[0090] De manera similar, mediante procedimientos habituales se puede llevar a cabo la alteración química o enzimática de ácidos nucleicos y polipéptidos expresados. Por ejemplo, la secuencia se puede modificar con la adición de lípidos, azúcares, péptidos, compuestos orgánicos o inorgánicos, con la inclusión de nucleótidos o aminoácidos modificados, o similares. Por ejemplo, las técnicas para la modificación de proteínas están ilustradas en Ausubel, más arriba. Detalles adicionales sobre las modificaciones químicas y enzimáticas se pueden encontrar en el presente documento. Estos procedimientos de modificación se pueden usar para modificar cualquier secuencia dada, o para modificar cualquier secuencia producida mediante los diversos procedimientos de modificación por mutación o evolución artificial indicados en el presente documento.

[0091] Por consiguiente, la descripción prevé la modificación de cualquier ácido nucleico dado por mutación, evolución, modificación química o enzimática, o cualquier otro procedimiento disponible, así como de los productos producidos poniendo en práctica dichos procedimientos, por ejemplo, usando las secuencias del presente documento como sustrato de partida para las diversas aproximaciones de modificación.

[0092] Por ejemplo, se puede usar una secuencia de codificación optimizada que contiene codones preferidos para un hospedador eucariota o procariota particular, por ejemplo, para incrementar la velocidad de traducción o para producir transcritos de ARN recombinante que tengan propiedades deseables, como una semivida más prolongada, comparada con transcritos producidos usando una secuencia no optimizada. Los codones de detención de la traducción también se pueden modificar para reflejar las preferencias del hospedador. Por ejemplo, codones de detención preferidos para *Saccharomyces cerevisiae* y mamíferos son TAA y TGA, respectivamente. El codón de detención preferido para plantas monocotiledóneas es TGA, mientras que los insectos y *E. coli* prefieren usar TAA como codón de detención.

[0093] Las secuencias de polinucleótidos descritas en el presente documento también se pueden manipular genéticamente con el fin de alterar una secuencia de codificación por multitud de razones, incluyendo pero no limitado a, alteraciones que modifiquen la secuencia para facilitar la clonación, procesamiento y/o expresión del producto génico. Por ejemplo, opcionalmente se introducen alteraciones usando técnicas que son muy conocidas en la materia, por ejemplo, mutagénesis dirigida de sitio, para insertar nuevos sitios de restricción, para alterar los patrones de glicosilación, para cambiar la preferencia de los codones, para introducir sitios de procesamiento alternativo, etc.

[0094] Además, un fragmento o dominio derivado de cualquiera de los polipéptidos descritos en el presente documento se puede combinar con dominios derivados de otros factores de transcripción o dominios sintéticos para modificar la actividad biológica de un factor de transcripción. Por ejemplo, un dominio de unión al ADN derivado de un factor de transcripción descrito en el presente documento se puede combinar con el dominio de activación de otro factor de transcripción o con un dominio de activación sintético. Un dominio de activación de la transcripción ayuda en la iniciación de la transcripción a partir de un sitio de unión al ADN. Los ejemplos incluyen la región de activación de la transcripción de VP16 o GAL4 (Moore y col. (1998) Proc. Natl. Acad. Sci. USA 95: 376-381; y Aoyama y col. (1995) Plant Cell 7:1773-1785), péptidos derivados de secuencias bacterianas (Ma y Ptashne (1987) Cell 51; 113-119) y péptidos sintéticos (Giniger y Ptashne, (1987) Nature 330:670-672).

IX. Expresión y modificación de polipéptidos

[0095] Normalmente, secuencias de polinucleótidos como las descritas en el presente documento se incorporan a moléculas de ADN (o ARN) recombinantes que dirigen la expresión de polipéptidos como los descritos en el presente documento en células hospedadoras apropiadas, plantas transgénicas, sistemas de traducción *in vitro*, o similares. Debido a la degeneración inherente del código genético, las secuencias de ácidos nucleicos que codifican secuencias de aminoácidos sustancialmente idénticas o funcionalmente equivalentes se pueden sustituir por cualquiera de las secuencias listadas para obtener la clonación y expresión del homólogo pertinente.

X. Vectores, promotores, y sistemas de expresión

[0096] También se describen construcciones recombinantes que comprenden una o más de las secuencias de ácidos nucleicos del presente documento. Las construcciones normalmente comprenden un vector, como un plásmido, un cósmido, un fago, un virus (por ejemplo, un virus de planta), un cromosoma bacteriano artificial (BAC), un cromosoma de levadura artificial (YAC), o similar, en el que se ha insertado una secuencia de ácidos nucleicos de la invención, en orientación directa o inversa. En un aspecto preferido de esta forma de realización, la construcción comprende adicionalmente secuencias reguladoras, incluyendo por ejemplo, un promotor unido de manera operable a la secuencia. Los expertos en la materia conocen un gran número de vectores y promotores adecuados, y éstos están disponibles comercialmente.

[0097] Los textos generales que describen técnicas de biología molecular son útiles en el presente documento, incluyendo el uso y la producción de vectores, promotores y muchos otros aspectos relevantes, incluyendo Berger, Sambrook y Ausubel, más arriba. Cualquiera de las secuencias identificadas se puede incorporar en un casete o vector, por ejemplo, para su expresión en plantas. Se han descrito una serie de vectores de expresión adecuados para la transformación estable de células vegetales o para el establecimiento de plantas transgénicas, incluyendo aquellos descritos en Weissbach y Weissbach, (1989) *Methods for Plant Molecular Biology*, Academic Press, y Gelvin y col., (1990) *Plant Molecular Biology Manual*, Kluwer Academic Publishers. Ejemplos específicos incluyen los vectores derivados de un plásmido Ti de *Agrobacterium tumefaciens*, así como los descritos por Herrera-Estrella y col. (1983) *Nature* 303: 209, Bevan (1984) *Nucl Acid Res.* 12: 8711-8721, Klee (1985) *Bio/Technology* 3: 637-642, para plantas dicotiledóneas.

[0098] Alternativamente, se pueden usar vectores no Ti para transferir el ADN a plantas y células monocotiledóneas usando técnicas de introducción de ADN libre. Dichos procedimientos pueden suponer el uso de, por ejemplo, liposomas, electroporación, bombardeo con microproyectiles, filamentos de carburo de silicio, y virus. Con el uso de estos procedimientos se pueden producir plantas transgénicas como trigo, arroz (Christou (1991) *Bio/Technology* 9: 957-962) y maíz (Gordon-Kamm (1990) *Plant Cell* 2: 603-618). Un embrión inmaduro también puede ser un buen tejido diana de monocotiledóneas para técnicas de introducción directa de ADN usando una pistola de partículas (Weeks y col. (1993) *Plant Physiol* 102: 1077-1084; Vasil (1993) *Bio/Technology* 10: 667-674; Wan y Lemeaux (1994) *Plant Physiol* 104: 37-48, y para la transferencia de ADN mediada por *Agrobacterium* (Ishida y col. (1996) *Nature Biotech* 14: 745-750).

[0099] Normalmente, los vectores para la transformación de plantas incluyen una o más secuencias (genómicas o de ADNc) que codifican plantas clonadas bajo el control transcripcional de secuencias reguladoras en 5' y 3' y un marcador seleccionable dominante. Dichos vectores para la transformación de plantas normalmente también contienen un promotor (por ejemplo, una región reguladora que controla la expresión inducible o constitutiva, regulada por el ambiente o durante el desarrollo, o específica de una célula o tejido), un sitio de inicio de la transcripción, una señal para el procesamiento de ARN (tal como sitios para el procesamiento alternativo de intrones), un sitio de terminación de la transcripción, y/o una señal de poliadenilación.

[00100] Ejemplos de promotores de plantas constitutivos que pueden ser útiles para la expresión de la secuencia de FT incluyen: el promotor 35S del virus del mosaico de la coliflor (CaMV), que confiere un elevado nivel de expresión constitutivo en la mayoría de tejidos vegetales (véase, por ejemplo, Odell y col. (1985) *Nature* 313:810-812); el promotor de la nopalina sintasa (An y col. (1988) *Plant Physiol* 88:547-552); y el promotor de la octopina sintasa (Fromm y col. (1989) *Plant Cell* 1: 977-984).

[0100] Se pueden usar una variedad de promotores génicos de plantas que regulan la expresión génica en respuesta a señales ambientales, hormonales, químicas, de desarrollo, y de forma activa en tejido para la expresión de una secuencia de FT en plantas. La selección de un promotor se basa principalmente en el fenotipo de interés y está determinada por factores tales como el tejido (por ejemplo, semilla, fruto, raíz, polen, tejido vascular, flor, carpelo, etc.), la capacidad de inducción (por ejemplo, en respuesta a lesiones, calor, frío, sequía, luz, patógenos, etc.), la sincronización, la fase de desarrollo, y similares. Se han caracterizado numerosos promotores conocidos y se pueden emplear de forma favorable para promover la expresión de un polinucleótido de la invención en una planta o célula transgénica de interés. Por ejemplo, los promotores específicos de tejido incluyen: promotores específicos de semillas (tales como el promotor de la napina, faseolina o el DC3 descritos en la patente de EE.UU. N° 5.773.697), promotores específicos de frutos que son activos durante la maduración de la fruta (como el promotor dru 1 (patente de EE.UU. N° 5.783.393), o el promotor A11 (patente de EE.UU. N° 4.943.674) y el promotor de la poligalacturonasa del tomate (Bird y col. (1988) *Plant Mol Biol* 11:651), promotores específicos de las raíces, como los descritos en las patentes de EE.UU. N° 5.618.988, 5.837.848 y 5.905.186, promotores activos en el polen como el PTA29, PTA26 y PTA13 (patente de EE.UU. N° 5.792.929), promotores activos en el tejido vascular (Ringli y Keller (1998) *Plant Mol Biol* 37:977-988), promotores específicos de las flores (Kaiser y col., (1995) *Plant Mol Biol* 28:231-243), del polen (Baerson y col. (1994) *Plant Mol Biol* 26:1947-1959), de los carpelos (Ohl y col. (1990) *Plant Cell* 2:837-848), del polen y los óvulos (Baerson y col. (1993) *Plant Mol Biol* 22:255-267), promotores inducibles por auxina (como los descritos en van der Kop y col. (1999) *Plant Mol Biol* 39:979-990 o Baumann y col. (1999) *Plant Cell* 11:323-334), promotores inducibles por citoquinas (Guevara-García (1998) *Plant Mol Biol* 38:743-753), promotores sensibles a giberelina (Shi y col. (1998) *Plant Mol Biol* 38:1053-1060, Willmott y col. (1998) 38:817-825) y similares. Promotores adicionales son aquellos que desencadenan la expresión en respuesta al calor (Ainley y col. (1993) *Plant Mol Biol* 22: 13-23), a la luz (por ejemplo, el promotor rbcS-3A del guisante, Kuhlemeier y col. (1989) *Plant Cell* 1:471, y el promotor rbcS del maíz, Schaffner y Sheen (1991) *Plant Cell* 3: 997); a lesiones (por ejemplo, wunl, Siebertz y col. (1989) *Plant Cell* 1: 961); a patógenos (como el promotor PR-1 descrito en Buchel y col. (1999) *Plant Mol. Biol.* 40:387-396, y el promotor PDF1.2 descrito en Manners y col. (1998) *Plant Mol. Biol.* 38:1071-80), y compuestos químicos como el metil jasmonato o el ácido salicílico (Gatz y col. (1997) *Plant Mol Biol* 48: 89-108). Además, la sincronización de la expresión se puede controlar usando promotores como los que actúan en la senescencia (An y Amazon (1995) *Science* 270: 1986-1988); o en el desarrollo tardío de la semilla (Odell y col.

(1994) Plant Physiol 106:447-458).

[0101] Los vectores para la expresión en plantas también pueden incluir señales de procesamiento del ARN que pueden estar situadas dentro, en dirección 5' o en dirección 3' de la secuencia codificante. Además, los vectores de expresión pueden incluir secuencias reguladoras adicionales a partir de la región 3' sin traducir de genes vegetales, por ejemplo, una región terminadora en 3' para incrementar la estabilidad del ARNm, tal como la región terminadora PI-II de la patata o las regiones terminadoras en 3' de la octopina o nopalina sintasa.

Elementos de expresión adicionales

10 **[0102]** Las señales de iniciación específicas pueden ayudar a la traducción eficiente de secuencias codificantes. Estas señales pueden incluir, por ejemplo, el codón de iniciación ATG y secuencias adyacentes. En los casos en los que una secuencia codificante, su codón de iniciación y secuencias en dirección 5' están insertadas en el vector de expresión adecuado, pueden no ser necesarias señales adicionales para el control de la traducción. No obstante, en los casos en los que sólo está insertada la secuencia codificante (por ejemplo, la secuencia codificante de una proteína madura), o uno de sus fragmentos, se pueden suministrar señales exógenas por separado para el control de la transcripción que incluyen el codón de iniciación ATG. El codón de iniciación se suministra en el marco de lectura correcto para facilitar su transcripción. Los elementos transcripcionales exógenos y los codones de iniciación pueden tener diversos orígenes, tanto naturales como sintéticos. La eficacia de la expresión se puede mejorar con la inclusión de potenciadores adecuados en el sistema celular utilizado.

20

Hospedadores para la expresión

[0103] La presente descripción también se refiere a células hospedadoras que son transducidas con vectores como los descritos en el presente documento, y la producción de polipéptidos de la invención (incluyendo sus fragmentos) mediante técnicas recombinantes. Las células hospedadoras están manipuladas genéticamente (es decir, se han introducido ácidos nucleicos, por ejemplo, transducido, transformado o transfectado) con los vectores de esta descripción, que pueden ser, por ejemplo, un vector de clonación o un vector de expresión que comprende los ácidos nucleicos pertinentes de la presente invención. Opcionalmente el vector es un plásmido, una partícula vírica, un fago; un ácido nucleico desnudo, etc. Las células hospedadoras manipuladas genéticamente se pueden cultivar en medios nutrientes convencionales modificados según sea adecuado para la activación de los promotores, la selección de los transformantes, o la amplificación del gen pertinente. Las condiciones de cultivo, tales como la temperatura, el pH y similares, son las usadas previamente con la célula hospedadora seleccionada para la expresión, y serán evidentes para los expertos en la materia y en las referencias citadas en el presente documento, incluyendo, Sambruk y Ausubel.

35

[0104] La célula hospedadora puede ser una célula eucariota, tal como una célula de levadura, o una célula vegetal, o la célula hospedadora puede ser una célula procariota, como una célula bacteriana. Los protoplastos vegetales también son adecuados para algunas aplicaciones. Por ejemplo, los fragmentos de ADN se introducen en tejidos vegetales, en células vegetales cultivadas o en protoplastos vegetales mediante procedimientos habituales incluyendo electroporación (Fromm y col., (1985) Proc. Natl. Acad. Sci. USA 82, 5824, infección con vectores víricos como el virus del mosaico de la coliflor (CaMV) (Hohn y col., (1982) Molecular Biology of Plant Tumors, (Academic Press, Nueva York) pp. 549-560; documento de EE.UU. 4.407.956), penetración balística a alta velocidad por partículas pequeñas con el ácido nucleico dentro de la matriz de cuentas o partículas pequeñas, o sobre la superficie (Klein y col., (1987) Nature 327, 70-73), el uso de polen como vector (documento WO 85/01856), o el uso de *Agrobacterium tumefaciens* o *A. rhizogenes* que portan un plásmido de T-ADN en el cual se clonan fragmentos de ADN. El plásmido de T-ADN se transmite a las células vegetales después de la infección por *Agrobacterium tumefaciens*, y una fracción se integra de manera estable en el genoma de la planta (Horsch y col. (1984) Science 233:496-498; Fraley y col. (1983) Proc. Natl. Acad. Sci. USA 80, 4803).

50 **[0105]** La célula puede incluir un ácido nucleico como los descritos que codifica un polipéptido, en donde la célula expresa un polipéptido de la invención. La célula también puede incluir secuencias de vectores, o similares. Además, son una característica adicional de esta descripción células y plantas transgénicas que incluyen cualquier polipéptido o ácido nucleico como el anterior o a lo largo de esta memoria descriptiva, por ejemplo, producido mediante la transducción de un vector como el descrito.

55

[0106] Para la producción prolongada y con un rendimiento elevado de proteínas recombinantes, se puede usar la expresión estable. Las células hospedadoras transformadas con una secuencia de nucleótidos que codifica un polipéptido como el descrito opcionalmente se cultivan en condiciones adecuadas para la expresión y recuperación de la proteína codificada a partir del cultivo celular. La proteína o uno de sus fragmentos producidos por una célula recombinante se pueden secretar, unirse a membrana, o estar contenidos dentro de la célula, dependiendo de la secuencia y/o del vector usado. Como comprenderán los expertos en la materia, los vectores de expresión que contienen polinucleótidos que codifican proteínas maduras como se ha descrito en el presente

documento se pueden diseñar con secuencias señal que dirigen la secreción de los polipéptidos maduros a través de la membrana celular procariota o eucariota.

XI. Restos aminoácidos modificados

5

[0107] Los polipéptidos descritos en el presente documento pueden contener uno o más restos aminoácidos modificados. La presencia de aminoácidos modificados puede ser ventajosa, por ejemplo, para incrementar la semivida del polipéptido, reducir la antigenicidad o toxicidad del polipéptido, incrementar la estabilidad de almacenamiento del polipéptido, o similares. Los resto(s) aminoácidos se modifican, por ejemplo, co-

10

traduccionalmente o post-traduccionalmente durante la producción recombinante o se modifican por medios químicos o sintéticos.

[0108] Ejemplos no limitantes de un resto aminoácido modificado incluyen la incorporación u otro uso de aminoácidos acetilados, aminoácidos glicosilados, aminoácidos sulfatados, aminoácidos prenilados (por ejemplo, farnesilados, geranilgeranilados), aminoácidos modificados con PEG (por ejemplo, "PEGilados"), aminoácidos biotilados, aminoácidos carboxilados, aminoácidos fosforilados, etc. La bibliografía está repleta de referencias adecuadas para guiar a alguien experto en la materia en la modificación de restos aminoácidos.

15

[0109] Los restos aminoácidos modificados pueden evitar o incrementar la afinidad del polipéptido por otra molécula, incluyendo, pero no limitado a, polinucleótidos, proteínas, carbohidratos, lípidos y derivados lipídicos, y otros compuestos orgánicos o sintéticos.

20

XII. Identificación de factores adicionales

25

[0110] También se puede usar un factor de transcripción como se ha descrito en el presente documento para identificar moléculas endógenas o exógenas adicionales que pueden afectar al fenotipo o a un rasgo de interés. Por otra parte, dichas moléculas incluyen compuestos orgánicos (moléculas pequeñas o grandes) y/o inorgánicos que afectan a la expresión (es decir, regulan) de un factor de transcripción particular. Alternativamente dichas moléculas incluyen moléculas

30

[0111] endógenas que actúan a nivel transcripcional mediante un factor de transcripción como se ha descrito en el presente documento para modificar un fenotipo según se desee. Por ejemplo, los factores de transcripción se pueden emplear para identificar uno o más genes en dirección 3' con el cual se somete a un efecto regulador del factor de transcripción. En una aproximación, un factor de transcripción o un homólogo del factor de transcripción como se ha descrito en el presente documento se expresa en una célula hospedadora, por ejemplo, una célula vegetal transgénica, tejido o explante, y se controlan los productos de expresión, ya sean ARN o proteínas, de dianas probables o aleatorias, por ejemplo, mediante hibridación a un biochip de sondas de ácidos nucleicos correspondientes a genes expresados en tejidos o tipos celulares de interés, mediante electroforesis en gel bidimensional de productos proteicos, o mediante cualquier otro procedimiento conocido en la materia para valorar la expresión de productos génicos a nivel de ARN o proteínas. Alternativamente, se puede usar un factor de transcripción como se ha descrito en el presente documento para identificar secuencias promotoras (es decir, sitios de unión) involucrados en la regulación de una diana en dirección 3'. Después de identificar una secuencia promotora, se pueden modificar las interacciones entre el factor de transcripción y la secuencia promotora cambiando nucleótidos específicos en la secuencia promotora o aminoácidos específicos en el factor de transcripción que interaccionan con la secuencia promotora para alterar un rasgo de la planta. Normalmente, los sitios de unión al ADN del factor de transcripción se identifican mediante ensayos de desplazamiento en gel. Después de identificar las regiones promotoras, las secuencias de las regiones promotoras se pueden emplear en ensayos de ADN de doble cadena para identificar moléculas que afectan a las interacciones de los factores de transcripción con sus promotores (Bulyk y col. (1999) Nature Biotechnology 17:573-577).

50

[0112] Los factores de transcripción identificados también son útiles para identificar proteínas que modifican la actividad del factor de transcripción. Dicha modificación se puede producir por modificación covalente, tal como por fosforilación, o mediante interacciones proteína-proteína (homo o heteropoliméricas). Se puede emplear cualquier procedimiento adecuado para la detección de interacciones proteína-proteína. Entre los procedimientos que se pueden emplear están la co-inmunoprecipitación, entrecruzamiento y co-purificación a través de gradientes o columnas cromatográficas, y el sistema de levaduras de dos híbridos.

55

[0113] El sistema de dos híbridos detecta interacciones de proteínas *in vivo* y ha sido descrito por Chien y col. ((1991), Proc. Natl. Acad. Sci. USA 88:9578-9582) y está disponible comercialmente en Clontech (Palo Alto, Calif.). En dicho sistema, se construyen plásmidos que codifican dos proteínas híbridas: una consta de un dominio de unión al ADN de una proteína activadora de la transcripción fusionada al polipéptido del FT y la otra consta del dominio de activación de la proteína activadora de la transcripción fusionada a una proteína desconocida que está codificada por un ADNc que ha sido recombinado en el plásmido como parte de una librería de ADNc. El plásmido

60

de fusión con el dominio de unión al ADN y la librería de ADNc se transforman en una cepa de la levadura *Saccharomyces cerevisiae* que contiene un gen informador (por ejemplo, lacZ) cuya región reguladora contiene el sitio de unión del activador de la transcripción. Ninguna de las dos proteínas híbridas pueden activar la transcripción del gen informador por sí solas. La interacción de las dos proteínas híbridas reconstituye la proteína activadora funcional y da como resultado la expresión del gen informador, que se detecta mediante un ensayo para el producto del gen informador. A continuación, los plásmidos de la librería sensibles a la expresión del gen informador se aíslan y se secuencian para identificar las proteínas codificadas por los plásmidos de la librería. Después de identificar las proteínas que interaccionan con los factores de transcripción, se pueden llevar a cabo ensayos para compuestos que interfieran con las interacciones proteína-proteína del FT.

10

XIII. Identificación de moduladores

[0114] Además de las moléculas intracelulares descritas anteriormente, se pueden identificar moléculas extracelulares que alteran la actividad o expresión de un factor de transcripción, ya sea directa o indirectamente. Por ejemplo, los procedimientos pueden suponer la puesta en contacto en primer lugar de una molécula candidato con una planta o una célula vegetal. La molécula se puede introducir mediante administración tópica, tal como pulverización o inmersión de una planta, y a continuación se controla el efecto de la molécula sobre la expresión o actividad del polipéptido del FT o la expresión del polinucleótido. Los cambios en la expresión del polipéptido del FT se pueden controlar mediante el uso de anticuerpos policlonales o monoclonales, electroforesis en gel o similares. Los cambios en la expresión de la secuencia de polinucleótidos correspondiente se pueden detectar mediante el uso de biochips, transferencias de Northern, PCR cuantitativa, o cualquier otra técnica para controlar cambios en la expresión del ARNm. Estas técnicas están ejemplificadas en Ausubel y col. (eds) Current Protocols in Molecular Biology, John Wiley & Sons (1998, y suplementos hasta 2001). Dichos cambios en los niveles de expresión se pueden correlacionar con rasgos modificados de las plantas y así las moléculas identificadas pueden ser útiles para sumergirlas o pulverizarlas sobre la fruta, vegetales y cultivos de grano para modificar rasgos en las plantas.

[0115] Esencialmente se puede probar cualquier composición disponible para la actividad moduladora de la expresión o la actividad de cualquier ácido nucleico o polinucleótido del presente documento. Así, se pueden probar librerías disponibles de compuestos como productos químicos, polipéptidos, ácidos nucleicos y similares para su actividad moduladora. A menudo, los potenciales compuestos moduladores se pueden disolver en disoluciones acuosas u orgánicas (por ejemplo, a base de DMSO) para comodidad de su administración a la célula o planta de interés en la que se debe probar la actividad del modulador. Opcionalmente, los ensayos están diseñados para seleccionar grandes librerías de composiciones moduladoras automatizando las etapas de ensayo y proporcionando compuestos de cualquier fuente adecuada a los ensayos, que normalmente se corren en paralelo (por ejemplo, en formatos de microtitulación sobre placas de microtitulación en ensayos robotizados).

[0116] En un aspecto, los procedimientos de selección de alto rendimiento suponen la proporción de una librería combinatoria que contiene un gran número de compuestos potenciales (compuestos moduladores potenciales). Dichas "librerías químicas combinatorias" a continuación se someten a selección en uno o más ensayos, como se ha descrito en el presente documento, para identificar aquellos miembros de la librería (especies químicas o subclases particulares) que presentan una actividad característica deseada. Los compuestos identificados de esta manera pueden servir como compuestos diana.

[0117] Una librería química combinatoria puede ser, por ejemplo, una colección de diversos compuestos químicos generados por síntesis química o síntesis biológica. Por ejemplo, una librería química combinatoria tal como una librería de polipéptidos se forma combinando un conjunto de bloques de construcción químicos (por ejemplo, en un ejemplo, aminoácidos) de todas las formas posibles para una longitud de compuesto dada (es decir, el número de aminoácidos en un compuesto polipeptídico de una longitud establecida). Ejemplos de librerías incluyen librerías de péptidos, librerías de ácidos nucleicos, librerías de anticuerpos (véase, por ejemplo, Vaughn y col. (1996) Nature Biotechnology, 14(3):309-314 y el documento PCT/US96/10287), librerías de carbohidratos (véase, por ejemplo, Liang y col. Science (1996) 274:1520-1522 y patente de EE.UU. 5.593.853), librerías de ácidos nucleicos peptídicos (véase, por ejemplo, patente de EE.UU. 5.539.083), y librerías de moléculas orgánicas pequeñas (véase, por ejemplo, benzodiazepinas, Baum C&EN Jan 18, pág. 33 (1993); isoprenoides, patente de EE.UU. 5.569.588; tiazolidinodionas y metatiazanonas, patente de EE.UU. 5.549.974; pirrolidinas, patentes de EE.UU. 5.525.735 y 5.519.134; compuestos morfolino, patente de EE.UU. 5.506.337) y similares.

[0118] La preparación y selección de librerías combinatorias y otros tipos de librerías es muy conocido por los expertos en la materia. Dichas librerías químicas combinatorias incluyen, pero no están limitadas a, librerías de péptidos (véase, por ejemplo, patente de EE.UU. 5.010.175; Furka, (1991) Int. J. Pept. Prot. Res. 37:487-493; y Houghton y col. (1991) Nature 354:84-88). También se pueden utilizar otros procesos químicos para generar librerías con diversidad química.

[0119] Además, como se ha indicado, está disponible para el público en general el equipamiento para la

selección de alto rendimiento de un compuesto, por ejemplo, usando cualquiera de una serie de sistemas robóticos muy conocidos que también han sido desarrollados para procesos químicos en fase de disolución útiles en sistemas de ensayo. Estos sistemas incluyen estaciones de trabajo automatizadas que tienen un aparato de síntesis automatizado y sistemas robóticos que utilizan brazos robóticos. Cualquiera de los dispositivos anteriores es adecuado para su uso, por ejemplo, en la selección de alto rendimiento de moduladores potenciales. La naturaleza e implementación de las modificaciones de estos dispositivos (si los hubiera) para que puedan funcionar como se describe en el presente documento será evidente para las personas expertas en la materia relevante.

[0120] De hecho, hay disponibles comercialmente sistemas completos de selección de alto rendimiento. Estos sistemas normalmente automatizan todos los procedimientos incluyendo la toma de todas las muestras y reactivos, la dispensación de líquidos, las incubaciones programadas, y las lecturas finales de la microplaca en un detector(es) adecuado para el ensayo. Estos sistemas configurados proporcionan un alto rendimiento y un inicio rápido, así como un elevado grado de flexibilidad y personalización. De manera similar, también están disponibles comercialmente implementaciones de selección microfluidizadas.

[0121] Los fabricantes de dichos sistemas proporcionan protocolos detallados de los diversos sistemas de selección de alto rendimiento. Así, por ejemplo, Zymark Corp proporciona boletines técnicos que describen sistemas de selección para detectar la modulación de la transcripción génica, la unión de ligandos, y similares. Los sistemas integrados del presente documento, además de proporcionar el alineamiento de secuencias y, opcionalmente, la síntesis de ácidos nucleicos relevantes, pueden incluir dichos aparatos de selección para identificar moduladores que tengan un efecto sobre uno o más de los polinucleótidos o polipéptidos descritos.

[0122] En algunos de los ensayos es deseable tener controles positivos para asegurar que los componentes de los ensayos funcionan correctamente. Esto es, se pueden incubar activadores o inhibidores transcripcionales conocidos con células/plantas/etc. en una muestra del ensayo, y el incremento/reducción resultante en la transcripción se puede detectar midiendo el incremento resultante en la expresión de ARN/proteína, etc., según los procedimientos del presente documento. Se apreciará que los moduladores también se pueden combinar con activadores o inhibidores transcripcionales para encontrar moduladores que inhiban la activación transcripcional o la represión transcripcional. Se puede controlar la expresión de los ácidos nucleicos y proteínas del presente documento o de cualquier ácido nucleico o proteína adicional activada por los ácidos nucleicos o proteínas del presente documento, o de ambos.

[0123] Además, se proporciona un procedimiento para identificar composiciones que modulan la actividad o expresión de un polinucleótido o polipéptido como se ha descrito en el presente documento. Por ejemplo, un compuesto de prueba, ya sea una molécula grande o pequeña, se pone en contacto con una célula, planta (o tejido vegetal o explante), o con una composición que comprende el polinucleótido o polipéptido de interés y se evalúa el efecto resultante sobre la célula, planta (o tejido o explante), controlando, directa o indirectamente, uno o más de: el nivel de expresión del polinucleótido o polipéptido, la actividad (o modulación de la actividad) del polinucleótido o polipéptido. En algunos casos, se puede detectar una alteración en el fenotipo de una planta después del contacto de una planta (o célula vegetal, o tejido o explante) con el modulador putativo, por ejemplo, por la modulación de la expresión o la actividad de un polinucleótido o polipéptido de la invención. La modulación de la expresión o la actividad de un polinucleótido o polipéptido de la invención también puede estar causada por elementos moleculares en una vía de transducción de señales con un segundo mensajero y dicha modulación puede afectar a elementos similares en la misma u otra vía de transducción de señales con un segundo mensajero.

XIV. Subsecuencias

[0124] También está contemplado el uso de polinucleótidos, también denominados oligonucleótidos en el presente documento, que normalmente tienen al menos 12 bases, preferentemente al menos 15 bases, más preferentemente al menos 20, 30 ó 50 bases, que se hibridan al menos en condiciones de rigurosidad elevada (o condiciones de rigurosidad ultra-elevadas o rigurosidad ultra-ultraelevadas) a una secuencia de polinucleótidos descrita anteriormente. Los polinucleótidos se pueden usar como sondas, cebadores, agentes en dirección directa y en dirección inversa, y similares, según los procedimientos indicados anteriormente.

[0125] Las subsecuencias de los polinucleótidos descritos, incluyendo fragmentos de polinucleótidos y oligonucleótidos, son útiles como sondas y cebadores de ácidos nucleicos. Un oligonucleótido adecuado para su uso como sonda o cebador tiene una longitud de al menos 15 nucleótidos aproximadamente, más habitualmente de al menos 18 nucleótidos aproximadamente, habitualmente de al menos 21 nucleótidos aproximadamente, con frecuencia una longitud de al menos 30 nucleótidos aproximadamente, o 40 nucleótidos aproximadamente, o superior. Una sonda de ácidos nucleicos es útil en protocolos de hibridación, por ejemplo, para identificar homólogos adicionales del polipéptido, incluyendo protocolos para experimentos con biochips. Los cebadores se pueden hibridar a una hebra de ADN diana complementaria mediante hibridación de ácidos nucleicos para formar un híbrido entre el cebador y la hebra de ADN diana, y a continuación extenderse a lo largo de la hebra de ADN diana

mediante una enzima ADN-polimerasa. Se pueden usar pares de cebadores para la amplificación de una secuencia de ácidos nucleicos, por ejemplo, mediante la reacción en cadena de la polimerasa (PCR) u otros procedimientos de amplificación de ácidos nucleicos. Véase Sambrook y Ausubel, más arriba.

5 **[0126]** Además se describe un polipéptido aislado o recombinante que incluye una subsecuencia de al menos 15 aminoácidos contiguos aproximadamente codificados por los polinucleótidos recombinantes o aislados descritos. Por ejemplo, dichos polipéptidos, o parte de sus dominios o fragmentos, se pueden usar como inmunógenos, por ejemplo, para producir anticuerpos específicos para la secuencia de polipéptidos, o como sondas para la detección de una secuencia de interés. Una subsecuencia puede tener un tamaño entre 15 aminoácidos aproximadamente y la
10 longitud completa del polipéptido.

[0127] Un polipéptido expresado que comprende dicha subsecuencia de polipéptidos desempeña al menos una función biológica del polipéptido intacto sustancialmente de la misma forma, o en un grado similar, a la que desempeña el polipéptido intacto. Por ejemplo, un fragmento del polipéptido puede comprender un motivo estructural
15 o un dominio funcional reconocible tal como un dominio de unión al ADN que se une a una región promotora del ADN específica, un dominio de activación o un dominio para las interacciones proteína-proteína.

XV. Producción de plantas transgénicas

20 Modificación de rasgos

[0128] Los polinucleótidos descritos se emplean de manera favorable para producir plantas transgénicas con diversos rasgos, o características, que se han modificado de una manera deseable, por ejemplo, para mejorar las características de la semilla de una planta. Por ejemplo, para modificar los rasgos de una planta se puede usar la
25 alteración de los niveles de expresión o patrones de expresión (por ejemplo, patrones de expresión espacial o temporal) de uno o más de los factores de transcripción descritos (u homólogos de los factores de transcripción), comparados con los niveles de la misma proteína encontrados en una planta silvestre. Un ejemplo ilustrativo de la modificación de un rasgo, de mejora de las características, alterando los niveles de expresión de un factor de transcripción particular se describe a continuación en los Ejemplos y en el Listado de secuencias.

30

Arabidopsis como sistema modelo

[0129] *Arabidopsis thaliana* es objeto de interés creciente como modelo genético y del metabolismo en plantas. La *Arabidopsis* tiene un genoma pequeño, y hay disponibles estudios bien documentados. Es fácil de crecer
35 a gran escala y hay disponibles mutantes que definen mecanismos importantes genéticamente controlados, o se pueden obtener fácilmente. Existen diversos procedimientos para introducir y expresar genes homólogos aislados (véase Koncz, y col., eds. *Methods in Arabidopsis Research*. y col. (1992), World Scientific, Nueva Jersey, Nueva Jersey, en "Preface"). Debido a su pequeño tamaño, un ciclo de vida corto, una autogamia obligada y una elevada fertilidad, la *Arabidopsis* también es un organismo seleccionado para el aislamiento de mutantes y para estudios en
40 rutas morfogenéticas y de desarrollo, y el control de estas rutas mediante factores de transcripción (Koncz, más arriba, p. 72). Una serie de estudios que han introducido factores de transcripción en *A. thaliana* han demostrado la utilidad de esta planta para la comprensión de los mecanismos de la regulación génica y de la alteración de rasgos en plantas. Véase, por ejemplo, Koncz, más arriba, y patente de EE.UU. N° 6.417.428).

45 Genes de Arabidopsis en plantas transgénicas

[0130] La expresión de genes que codifican factores de transcripción que modifican la expresión de genes, polinucleótidos, y proteínas endógenas es muy conocida en la materia. Además, plantas transgénicas que comprenden polinucleótidos aislados que codifican factores de transcripción también pueden modificar la expresión
50 de genes, polinucleótidos, y proteínas endógenas. Los ejemplos incluyen Peng y col. (1997, *Genes and Development* 11:3194-3205) y Peng y col. (1999, *Nature*, 400:256-261). Además, muchos otros autores han demostrado que un factor de transcripción de *Arabidopsis* expresado en una especie vegetal exógena desencadena una respuesta fenotípica idéntica o muy similar. Véase, por ejemplo, Fu y col. (2001, *Plant Cell* 13:1791-1802); Nandi y col. (2000, *Curr. Biol.* 10:215-218); Coupland (1995, *Nature* 377:482-483); y Weigel y Nilsson (1995, *Nature*
55 377:482-500).

Genes homólogos introducidos en plantas transgénicas

[0131] Genes homólogos que pueden proceder de cualquier planta, o de cualquier fuente natural, sintética,
60 semisintética o recombinante, y que comparten una identidad o similitud de secuencia significativa con los proporcionados en el presente documento se pueden introducir en plantas, por ejemplo, plantas de cultivo, para conferir rasgos deseables o mejorados. En consecuencia, se pueden producir plantas transgénicas que comprendan un vector o casete de expresión recombinante con un promotor unido de manera operable a una o más secuencias

homólogas a las secuencias descritas en el presente documento. El promotor puede ser, por ejemplo, un promotor vegetal o vírico.

[0132] Así, la invención proporciona procedimientos para la preparación de plantas transgénicas, y para la modificación de los rasgos de una planta. Estos procedimientos incluyen la introducción en una planta de un vector o casete de expresión recombinante que comprende un promotor funcional unido de manera operable a una o más secuencias homólogas a las secuencias descritas en el presente documento.

[0133] Las descripciones completas de los rasgos asociados a cada polinucleótido de la invención se describen en profundidad en la Tabla 4, y Tabla 5.

Tabla 4

SEQ ID N° del polinucleótido	GID N°	Rasgo	Categoría	Familia	Comentario	SEQ N° del polinucleótido	Dominios conservados
31	G47	Arquitectura; tallo; tiempo de floración; contenido en aceite de la semilla alterado	Desarrollo y morfología; tiempo de floración; bioquímica de la semilla	AP2	Arquitectura alterada y desarrollo de inflorescencias, estructura de los tejidos vasculares; floración tardía; contenido en aceite de la semilla alterado	32	(11-80)
439	G2133	Morfología; otro; tiempo de floración; contenido en proteína de la semilla	Desarrollo y morfología; tiempo de floración	AP2	Múltiples alteraciones del desarrollo; floración tardía; contenido en proteínas de la semilla alterado	440	(11-83)

Rasgos de interés

[0134] Ejemplos de algunos de los rasgos que pueden ser deseables en plantas, y que se pueden proporcionar transformando las plantas con las secuencias descritas en el presente documento, se listan en la Tabla 5.

Tabla 5. Genes, rasgos y utilidades que afectan a las características de la planta

<u>Categoría del rasgo</u>	<u>Rasgos</u>	<u>Genes de los factores de transcripción que tienen un impacto en los rasgos</u>	<u>Efectos del gen de utilidad sobre:</u>
Resistencia y tolerancia	Resistencia al estrés osmótico	G47;	Velocidad de germinación, supervivencia, rendimiento
Desarrollo, morfología	Arquitectura general de la planta	G47; G2133;	Tejidos vasculares, contenido en lignina; contenido en la pared celular; apariencia
	Estructura de la flor, inflorescencia	G47;	Horticultura ornamental; producción de azafrán u otras flores comestibles
	Morfología del tallo	G47;	Ornamental; digestibilidad
Crecimiento, reproducción	Floración retardada	G47; G2133	Retraso en la producción de polen de plantas OGM; sincronía de la floración; mayor rendimiento
<u>Significación de los rasgos de la planta modificada</u>			

[0135] Actualmente, la existencia de una serie de grupos de maduración para diferentes latitudes representa una barrera importante para la introducción de nuevos rasgos valiosos. Cualquier rasgo (por ejemplo, resistencia a enfermedades) debe ser introducido por separado en cada uno de los diferentes grupos de maduración, un ejercicio laborioso y caro. La disponibilidad de una única cepa, que pudiera crecer en cualquier latitud, por tanto incrementaría enormemente el potencial para la introducción de nuevos rasgos para cultivar especies como la soja y el algodón.

[0136] Para muchos de los rasgos, listados en la Tabla 5 y a continuación, que se pueden conferir a plantas, se puede usar un único gen de factores de transcripción para incrementar o reducir, adelantar o retrasar, o mejorar o perjudicar un rasgo dado. Por ejemplo, la sobreexpresión de un gen de factores de transcripción que aparezca de manera natural en una planta puede provocar la floración temprana con respecto a plantas no transformadas o de tipo silvestre. Inactivando el gen, o suprimiendo el gen (con, por ejemplo, supresión en dirección inversa) la planta puede experimentar una floración retardada. De manera similar, la sobreexpresión o supresión de uno o más genes puede impartir diferencias significativas en la producción de productos vegetales, tales como diferentes relaciones de ácidos grasos. Así, la supresión de un gen que provoca que una planta sea más sensible al frío puede mejorar la tolerancia de la planta al frío.

[0137] Resistencia al estrés osmótico. Los genes de los factores de transcripción descritos en el presente documento que confieren resistencia al estrés osmótico pueden incrementar la velocidad de germinación en condiciones adversas, que podría tener un impacto en la supervivencia y en la producción de semillas y plantas.

[0138] Tolerancia a la sequía y al bajo contenido en agua. Las estrategias que permitan a las plantas sobrevivir en condiciones con poca agua pueden incluir, por ejemplo, un área superficial reducida o la producción de aceites o ceras superficiales. Una serie de genes de factores de transcripción descritos en el presente documento incrementan la tolerancia de una planta a condiciones de bajo contenido en agua y proporcionan los beneficios de una supervivencia mejorada, un rendimiento incrementado y la extensión geográfica y temporal de la extensión de plantación.

[0139] Estrés oxidativo. En plantas, así como en todos los seres vivos, las tensiones bióticas y abióticas inducen la formación de radicales oxígeno, incluyendo radicales superóxido y peróxido. Esto tiene el efecto de acelerar la senescencia, particularmente en las hojas, con la pérdida de rendimiento resultante y el efecto adverso sobre la apariencia. En general, las plantas que tienen el nivel más alto de mecanismos de defensa, tales como, por ejemplo, restos poliinsaturados de lípidos de membrana, tienen más probabilidades de prosperar en condiciones que

inducen estrés oxidativo (por ejemplo, mucha luz, ozono, déficit de agua, particularmente en combinación). La introducción de los genes de los factores de transcripción descritos en el presente documento que incrementan el nivel de los mecanismos de defensa al estrés oxidativo proporcionarían efectos beneficiosos sobre el rendimiento y la apariencia de las plantas. Se ha demostrado que un agente oxidante específico, el ozono, provoca lesiones foliares significativas, que tiene un impacto sobre el rendimiento y la apariencia de plantas de cultivo y ornamentales. Además de una reducción en las lesiones foliares que se encontraría en plantas resistentes al ozono creadas transformando plantas con algunos de los genes de los factores de transcripción descritos en el presente documento, también se ha demostrado que esto último incrementa la fluorescencia de la clorofila (Yu-Sen Chang y col. Bot. Bull. Acad. Sin. (2001) 42: 265-272).

10

[0140] Arquitectura general de la planta. Varios de los genes de los factores de transcripción descritos en el presente documento se han introducido en plantas para alterar numerosos aspectos de la morfología de la planta. Por ejemplo, se ha demostrado que se pueden usar una serie de factores de transcripción para manipular la ramificación, tales como medios para modificar la ramificación lateral, una posible aplicación en la industria forestal.

15

También se han producido plantas transgénicas que presentan un contenido de la pared celular, una producción de lignina, un número de órganos de floración, o una forma general de las plantas alterados. Los genes de los factores de descripción descritos en el presente documento transformados en plantas se pueden usar para afectar a la morfología de la planta incrementando o reduciendo la distancia entre los nodos, ambas cosas que pueden ser ventajosas en diferentes circunstancias. Por ejemplo, para el crecimiento rápido de plantas leñosas para proporcionar más biomasa, o un menor número de nudos, generalmente es deseable una mayor distancia entre nodos. Para una mejor protección contra el viento de arbustos o árboles, o características de recolección de, por ejemplo, miembros de la familia de las gramíneas, puede ser ventajoso una menor distancia entre nodos. Estas modificaciones también demuestran ser útiles en la industria de la horticultura ornamental para la creación de características fenotípicas únicas de plantas ornamentales.

20

[0141] Mayor estatura. Para algunas plantas ornamentales, puede ser muy deseable la capacidad de proporcionar variedades más grandes. Para muchas plantas, incluyendo árboles frutales o árboles y arbustos que sirven como protección contra el viento o como pantalla visual, una mayor estatura proporciona unos beneficios obvios. Las especies de cultivo también pueden producir rendimientos más elevados en cultivos más grandes.

25

[0142] Menor estatura o enanismo. Se pueden usar genes de los factores de transcripción descritos en el presente documento que reducen la estatura de la planta para producir plantas que sean más resistentes al daño por el viento y la lluvia, más resistentes al calor o a una baja humedad o la falta de agua. Las plantas enanas también tienen un interés importante en la industria de la horticultura ornamental, y particularmente en aplicaciones para jardines domésticos en los que la disponibilidad de espacio puede estar limitada.

30

[0143] Desarrollo y modificaciones de las raíces. Modificando la estructura o el desarrollo de las raíces con la transformación en una planta de uno o más de los genes de los factores de transcripción descritos en el presente documento, se pueden producir plantas que tengan la capacidad de prosperar en suelos de otro modo improductivos. Por ejemplo, las raíces de las vides que se extienden más profundamente en suelos rocosos, o que siguen siendo viables en suelos anegados, incrementarían la extensión de plantación eficaz del cultivo. Puede ser ventajoso manipular una planta para producir raíces cortas, como cuando un suelo en el que crecerá la planta está ocasionalmente inundado, o cuando son frecuentes hongos patógenos o nematodos que causan enfermedades.

35

[0144] Forma, color y modificaciones de la hoja. En experimentos de laboratorio se ha demostrado que la sobreexpresión de algunos de los factores de transcripción descritos en el presente documento produce efectos notables sobre el desarrollo de la hoja. En las fases tempranas del crecimiento, estas plántulas transgénicas desarrollaron hojas estrechas que apuntan hacia arriba con peciolo largos, lo que indica posiblemente una interrupción en los procesos controlados por el ciclo circadiano o movimientos nictinásticos. Se pueden usar otros genes de factores de transcripción para incrementar la biomasa vegetal; sería útil un mayor tamaño en cultivos en los que la porción vegetativa de la planta es la porción comercializable.

45

[0145] Velocidad de crecimiento y desarrollo de la planta. Se ha demostrado que una serie de los genes de los factores de transcripción descritos en el presente documento tienen efectos significativos sobre la velocidad de crecimiento y el desarrollo de la planta. Estas observaciones han incluido, por ejemplo, un crecimiento y un desarrollo más rápido o retardado de los órganos reproductores. Esto sería útil para regiones con temporadas de crecimiento corto o largo, respectivamente. La aceleración del crecimiento de la planta también mejoraría el rendimiento temprano o incrementaría la biomasa en una fase más temprana, cuando esto es deseable (por ejemplo, en la producción de productos forestales).

50

- [0146]** Planta, vigor de la plántula. Se ha demostrado que las plántulas transformadas con los factores de transcripción descritos en el presente documento poseen cotiledones más grandes y parecen algo más avanzadas que las plantas control. Esto indica que las plántulas se desarrollan más rápidamente que las plantas control. El desarrollo rápido de las plántulas es probable que reduzca la pérdida debida a enfermedades, particularmente frecuente en la fase de plántula (por ejemplo, caída de las plántulas) y así es importante para la supervivencia de plantas que germinan en el campo o en entornos controlados.
- [0147]** Floración temprana y retardada. Los genes de los factores de transcripción descritos en el presente documento que aceleran la formación podrían tener aplicaciones valiosas en dichos programas puesto que permiten unos tiempos de generación más rápidos. En un número de especies, por ejemplo, brócoli, o coliflor, donde las partes reproductoras de la planta constituyen el cultivo y los tejidos vegetativos se descartan, sería ventajoso acelerar el tiempo de floración. La aceleración de la floración podría acortar los programas de cultivo y de cría de árboles. Adicionalmente, en algunos casos, un tiempo de generación más rápido podría permitir que se realizasen cosechas adicionales de un cultivo en una temporada de siembra dada. Ya se ha demostrado que una serie de genes de *Arabidopsis* aceleran la floración cuando se expresan constitutivamente. Estos incluyen LEAFY, APETALA1 y CONSTANS (Mandel, M. y col., 1995, Nature 377, 522-524; Weigel, D. y Nilsson, O., 1995, Nature 377, 495-500; Simon y col., 1996, Nature 384, 59-62).
- [0148]** Regulando la expresión de la floración potencial usando promotores inducibles, se puede desencadenar la floración con la aplicación de un inductor químico. Dichos sistemas inducibles también se podrían usar para ajustar la floración de variedades de cultivos a diferentes latitudes. Actualmente, especies como la soja y el algodón están disponibles en forma de grupos de maduración que son adecuados para diferentes latitudes en base a sus tiempos de floración (que están regidos por la duración del día). Un sistema en el que la floración se pudiera controlar químicamente permitiría un único grupo de maduración septentrional de alto rendimiento que se puede crecer en cualquier latitud. En regiones meridionales dichas plantas podrían ser cultivadas durante más tiempo, incrementando así los rendimientos, antes de que se indujese la floración. En áreas más septentrionales, la inducción se usaría para asegurar que el cultivo florece antes de las primeras heladas invernales.
- [0149]** En un considerable número de especies, por ejemplo, cultivos radiculares, donde las partes vegetativas de las plantas constituyen el cultivo y los tejidos reproductivos se descartan, sería ventajoso retrasar o evitar la floración. La extensión del desarrollo vegetativo con los genes de los factores de transcripción descritos en el presente documento podría provocar de esta forma un gran aumento de los rendimientos. La prevención de la floración podría ayudar a maximizar los rendimientos vegetativos y evitar la fuga del polen del organismo genéticamente modificado (OGM).
- [0150]** En la Tabla 4 se proporciona un listado de efectos específicos y utilidades que los genes de los factores de transcripción descritos en el presente documento tienen sobre plantas, determinados por observación directa y análisis de ensayos.
- [0151]** También se puede modificar el rasgo de una planta usando el sistema Cre-lox (por ejemplo, como se describe en la patente de EE.UU. Nº 5.658.772). Se puede modificar el genoma de una planta para incluir un primer y segundo sitio lox que a continuación se pone en contacto con una recombinasa Cre. Si los sitios lox están en la misma orientación, la secuencia de ADN interpuesta entre los dos sitios se escinde. Si los sitios lox están en orientaciones opuestas, la secuencia interpuesta se invierte.
- [0152]** Los polinucleótidos y polipéptidos descritos también se pueden expresar en una planta en ausencia de casete de expresión manipulando la actividad o el nivel de expresión del gen endógeno por otros medios. Por ejemplo, expresando ectópicamente un gen mediante el marcaje de activación de T-ADN (Ichikawa y col. (1997) Nature 390 698-701; Kakimoto y col. (1996) Science 274: 982-985). Este procedimiento supone la transformación de una planta con un marcador génico que contiene múltiples potenciadores transcripcionales y una vez se haya insertado el marcador en el genoma, la expresión de una secuencia que codifica el gen adyacente queda desregulada. En otro ejemplo, la maquinaria transcripcional de una planta se puede modificar para así incrementar los niveles de transcripción de un polinucleótido (véase, por ejemplo, Publicaciones PCT WO 96/06166 y WO 98/53057 que describen la modificación de la especificidad de unión al ADN de las proteínas con el motivo de dedos de cinc cambiando aminoácidos particulares en el motivo de unión al ADN).
- [0153]** La planta transgénica también puede incluir la maquinaria necesaria para expresar o alterar la actividad de un polipéptido codificado por un gen endógeno, por ejemplo, alterando el estado de fosforilación del

polipéptido para mantenerlo en un estado activado.

[0154] Se pueden producir plantas transgénicas (o células vegetales, o explantes vegetales, o tejidos vegetales) que incorporen los polinucleótidos descritos y/o expresen los polipéptidos descritos mediante una variedad de técnicas bien establecidas como se ha descrito anteriormente. Después de la construcción de un vector, lo más habitualmente un casete de expresión, que incluye un polinucleótido, por ejemplo, que codifica un factor de transcripción o un homólogo del factor de transcripción, se pueden usar técnicas habituales para introducir el polinucleótido en una planta, una célula vegetal, un explante vegetal o un tejido vegetal de interés. Opcionalmente, la célula, explante o tejido vegetal se puede regenerar para producir una planta transgénica.

10

[0155] La planta puede ser cualquier planta superior, incluyendo plantas gimnospermas, monocotiledóneas y dicotiledóneas. Hay protocolos adecuados disponibles para leguminosas (alfalfa, soja, trébol, etc.), umbelíferas (zanahoria, apio, chirivía), crucíferas (col, rábano, colza, brócoli, etc.), cucurbitáceas (melón y pepino), gramíneas (trigo, maíz, arroz, cebada, mijo, etc.), solanáceas (patata, tomate, tabaco, pimientos, etc.), y otros cultivos. Véanse los protocolos descritos en Ammirato y col. (1984) Handbook of Plant Cell Culture - Crop Species, Macmillan Publ. Co. Shimamoto y col. (1989) Nature 338:274-276; Fromm y col. (1990) Bio/Technology 8:833-839; y Vasil y col. (1990) Bio/Technology 8:429-434.

15

[0156] La transformación y regeneración de células vegetales tanto monocotiledóneas como dicotiledóneas es ahora un proceso rutinario, y el experto seleccionará la técnica de transformación más apropiada. La elección del procedimiento variará con el tipo de planta a transformar; los expertos en la materia reconocerán la conveniencia de procedimientos particulares para tipos vegetales dados. Los procedimientos adecuados pueden incluir, pero no están limitados a: electroporación de protoplastos vegetales; transformación mediada por liposomas; transformación mediada por polietilenglicol (PEG); transformación usando virus; microinyección de células vegetales; bombardeo de células vegetales con microproyectiles; infiltración sobre vacío; y transformación mediada por *Agrobacterium tumefaciens*. Transformación significa la introducción de una secuencia de nucleótidos en una planta de forma que provoca una expresión estable o transitoria de la secuencia.

25

[0157] Ejemplos de éxitos en la modificación de las características vegetales mediante transformación con secuencias clonadas que sirven para ilustrar los conocimientos actuales en este campo tecnológico, incluyen: patentes de EE.UU. N° 5.571.706; 5.677.175; 5.510.471; 5.750.386; 5.597.945; 5.589.615; 5.750.871; 5.268.526; 5.780.708; 5.538.880; 5.773.269; 5.736.369 y 5.610.042.

30

[0158] Después de la transformación, las plantas se seleccionan preferentemente usando un marcador seleccionable dominante incorporado al vector de transformación. Normalmente, dicho marcador conferirá resistencia a antibióticos o a herbicidas a las plantas transformadas, y la selección de los transformantes se puede conseguir exponiendo las plantas a concentraciones apropiadas del antibiótico o herbicida.

35

[0159] Después de que las plantas transformadas se seleccionen y se crezcan hasta su madurez, se identifican las plantas que presenten un rasgo modificado. El rasgo modificado puede ser cualquiera de los rasgos descritos anteriormente. Adicionalmente, para confirmar que el rasgo modificado es debido a cambios en los niveles de expresión o de actividad del polipéptido o polinucleótido descrito, se puede determinar analizando la expresión de ARNm utilizando transferencias de Northern, RT-PCR o biochips, o la expresión de proteínas usando inmunotransferencias o transferencias de Western o ensayos de desplazamiento en gel.

45

Ejemplos

[0160] Se pretende que los siguientes ejemplos ilustren pero no limiten la presente invención. Las descripciones completas de los rasgos asociados a cada polinucleótido descrito se describen en profundidad en la Tabla 4 y Tabla 5.

50

Ejemplo I: Identificación y clonación del gen de longitud completa

[0161] Se identificaron secuencias del factor de transcripción putativo (genómicas o ESTs) relacionadas con factores de transcripción conocidos en la base de datos del GenBank de *Arabidopsis thaliana* usando el programa de análisis de secuencias tblastn con los parámetros por defecto y un umbral límite para el valor de P de -4 o -5 o inferior, dependiendo de la longitud de la secuencia cotejada. A continuación se buscaron los aciertos en la secuencia del factor de transcripción putativo para identificar aquéllas que contienen cadenas de la secuencia particular. Si los aciertos de la secuencia contenían dichas cadenas de la secuencia, las secuencias se confirmaron

55

como factores de transcripción.

[0162] Alternativamente, se seleccionaron librerías de ADNc de *Arabidopsis thaliana* procedentes de diferentes tejidos o tratamientos, o librerías genómicas para identificar nuevos miembros de una familia de transcripción usando una aproximación de hibridación de baja rigurosidad. Las sondas se sintetizaron usando cebadores específicos para el gen en una reacción de PCR estándar (temperatura de hibridación de 60°C) y marcadas con dCTP ³²P usando el kit High Prime DNA Labeling (Boehringer Mannheim). Las sondas purificadas radiomarcadas se añadieron a filtros sumergidos en medio de hibridación de Church (NaPO₄ 0,5 M a pH 7,0, SDS al 7%, seroalbúmina bovina al 1% en p/v) y se hibridaron durante toda la noche a 60°C con agitación. Los filtros se lavaron dos veces durante 45 a 60 minutos con 1xSCC, SDS al 1% a 60°C.

[0163] Para identificar una secuencia adicional 5' o 3' de una secuencia parcial de ADNc en una librería de ADNc, se llevó a cabo la amplificación rápida en 5' y 3' de los extremos de ADNc (RACE) usando el kit de amplificación de ADNc U.C. Marathon (Clontech, Palo Alto, CA). En general, el procedimiento supone el aislamiento en primer lugar de ARNm poli(A), la síntesis de la primera y segunda hebra de ADNc para generar ADNc de doble cadena, la escisión de los extremos no apareados del ADNc, seguido por la ligación del U.C. Darathon Adaptor al ADNc para formar una librería de ADNc de doble cadena ligado al adaptador.

[0164] Se diseñaron cebadores específicos para el gen para su uso junto con cebadores específicos para el adaptador para ambas reacciones RACE en 5' y 3'. Se usaron cebadores anidados, en lugar de cebadores únicos, para incrementar la especificidad de la PCR. Usando las reacciones RACE en 5' y 3', se obtuvieron, se secuenciaron y se clonaron fragmentos de RACE en 5' y 3'. El proceso se puede repetir hasta que se identificaron los extremos 5' y 3' del gen de longitud completa. A continuación, se generó el ADNc de longitud completa mediante PCR usando cebadores específicos para los extremos 5' y 3' del gen mediante PCR de extremo a extremo.

Ejemplo II: Construcción de vectores de expresión

[0165] La secuencia se amplificó a partir de una librería genómica o de ADNc usando cebadores específicos para las secuencias en dirección 5' y en dirección 3' de la región codificante. El vector de expresión fue pMEN20 o pMEN65, ambos que proceden de pMON316 (Sanders y col., (1987) Nucleic Acids Research 15:1543-1558) y contienen el promotor 35S de CaMV para expresar transgenes. Para clonar la secuencia en el vector, tanto pMEN20 como el fragmento de ADN amplificado se sometieron a digestión por separado con las enzimas de restricción Sall y NotI a 37°C durante 2 horas. Los productos de digestión se sometieron a electroforesis en un gel de agarosa al 0,8% y se visualizaron por tinción con bromuro de etidio. Los fragmentos de ADN que contienen la secuencia y el plásmido linealizado se escindieron y se purificaron usando un kit de extracción en gel de Qiaquick (Qiagen, Valencia, CA). Los fragmentos de interés se ligaron a una relación de 3:1 (vector a inserto). Las reacciones de ligación se llevaron a cabo 16°C durante 16 horas usando la ADN ligasa de T4 (New England Biolabs, Beverly MA). Los ADNs ligados se transformaron en células competentes de la cepa DH5-alfa de *E. coli* usando el procedimiento de choque térmico. Las transformaciones se pusieron en placas LB que contienen 50 mg/l de kanamicina (Sigma, St. Louis, MO). Se crecieron colonias individuales durante toda la noche en 5 ml de caldo LB que contiene 50 mg/l de kanamicina a 37°C. El ADN plasmídico se purificó usando los kits Qiaquick Mini Prep (Qiagen).

Ejemplo III: Transformación de *Agrobacterium* con el vector de expresión

[0166] Después de construir el vector plasmídico que contiene el gen, el vector se usó para transformar células de *Agrobacterium tumefaciens* que expresan los productos génicos. El lote de células de *Agrobacterium tumefaciens* para su transformación se preparó como se describe por Nagel y col. (1990) FEMS Microbiol Letts. 67: 325-328. Se creció la cepa ABI de *Agrobacterium* en 250 ml de medio LB (Sigma) durante toda la noche a 28°C con agitación hasta que se alcanzó una absorbancia (A₆₀₀) de 0,5-1,0. Las células se recogieron por centrifugación a 4000g durante 15 minutos a 4°C. A continuación las células se resuspendieron en 250 µl de tampón frío (HEPES 1 mM, y pH ajustado a 7,0 con KOH). Las células se centrifugaron de nuevo como se ha descrito anteriormente y se resuspendieron en 125 µl de tampón frío. A continuación las células se centrifugaron y se resuspendieron dos veces más en el mismo tampón HEPES como se ha descrito anteriormente a un volumen de 100 µl y 750 µl, respectivamente. Las células resuspendidas se distribuyeron entonces en alícuotas de 40 µl, congeladas rápidamente en nitrógeno líquido y almacenadas a -80°C.

[0167] Las células de *Agrobacterium* se transformaron con plásmidos preparados como se ha descrito anteriormente siguiendo el protocolo descrito por Nagel y col. Por cada construcción de ADN a transformar, se mezclaron 50-100 ng de ADN (generalmente resuspendidos en Tris-HCl 10 mM, EDTA 1 mM, pH 8,0) con 40 µl de

células de *Agrobacterium*. A continuación la mezcla de ADN/células se transfirió a una cubeta enfriada con una separación de los electrodos de 2 mm y se sometió a un voltaje de 2,5 kV disipado a 25 μ F y 200 μ F usando un aparato Gene Pulser II (Bio-Rad, Hercules, CA). Después de la electroporación, las células se resuspendieron inmediatamente en 1,0 ml de LB y se dejó que se recuperaran sin selección con antibióticos durante 2-4 horas a 28°C en una incubadora con agitación. Después de la recuperación, las células se sembraron en placa sobre medio selectivo de caldo LB que contiene 100 μ g/ml de estreptomina (Sigma) y se incubó durante 24-48 horas a 28°C. A continuación se tomaron colonias únicas y se inocularon en medio fresco. La presencia de la construcción plasmídica se verificó mediante amplificación por PCR y análisis de secuencias.

10 Ejemplo IV: Transformación de plantas *Arabidopsis* con *Agrobacterium tumefaciens* con el vector de expresión

[0168] Después de la transformación de *Agrobacterium tumefaciens* con vectores plasmídicos que contienen el gen, se identificaron colonias de *Agrobacterium* únicas, se propagaron, y se usaron para transformar plantas *Arabidopsis*. Resumiendo, a cultivos de 500 ml de medio LB que contiene 50 mg/l de kanamicina se les inoculó las colonias y se crecieron a 28°C con agitación durante 2 días hasta que se alcanzó una absorbancia óptica a una longitud de onda de 600 nm por encima de 1 cm (A_{600}) de > 2,0. A continuación las células se recogieron por centrifugación a 4000g durante 10 minutos, y se resuspendieron en medio de infiltración (1/2 X de sales de Murashige y Skoog (Sigma), 1 X vitaminas B-5 de Gamborg (Sigma), sacarosa al 5,0% (p/v) (Sigma), 20 bencilaminopurina 0,044 μ M (Sigma), 200 μ l/l de Silwet L-77 (Lehle Seeds) hasta que se alcanzó una A_{600} de 0,8.

[0169] Antes de la transformación, se sembraron semillas de *Arabidopsis thaliana* (ecotipo Columbia) a una densidad de ~10 plantas por cada maceta de 4" en medio para macetas Pro-Mix BX (Hummert International) cubiertas con una red de fibra de vidrio (18 mm X 16 mm). Las plantas se crecieron con iluminación continua (50-75 μ E/m²/sec) a 22-23°C con una humedad relativa del 65-70%. Después de 4 semanas aproximadamente, los tallos de inflorescencia primarios (yemas) se cortan para estimular el crecimiento de múltiples yemas secundarias. Después de la floración de las yemas secundarias maduras, las plantas se prepararon para la transformación quitando todas las silicuas y flores abiertas.

[0170] A continuación las macetas se sumergieron boca abajo en la mezcla de medio de infiltración de *Agrobacterium* como se ha descrito anteriormente durante 30 segundos, y se pusieron sobre su costado para permitir el drenaje en una superficie plana de 1' x 2' cubierta con una envuelta de plástico. Después de 24 horas, se retiró la envuelta de plástico y las macetas se pusieron de pie. El procedimiento de inmersión se repitió una semana más tarde, para un total de dos inmersiones por maceta. A continuación se recogieron las semillas de cada maceta de transformación y se analizaron siguiendo el protocolo descrito a continuación.

Ejemplo V: Identificación de transformantes primarios de *Arabidopsis*

[0171] Las semillas recogidas de las macetas de transformación se esterilizaron esencialmente de la manera siguiente. Las semillas se dispersaron en una disolución que contiene Triton X-100 al 0,1% (v/v) (Sigma) y H₂O estéril y se lavaron agitando la suspensión durante 20 minutos. A continuación la disolución de lavado se drenó y se sustituyó por una disolución de lavado fresca para lavar las semillas durante 20 minutos con agitación. Después de la retirada de la segunda disolución de lavado, se añadió una disolución que contiene Triton X-100 al 0,1% (v/v) y etanol al 70% (Equistar) a las semillas y la suspensión se agitó durante 5 minutos. Después de la retirada de la disolución de etanol/detergente, se añadió una disolución que contiene Triton X-100 al 0,1 % (v/v) y el 30% (v/v) de lejía (Clorox) a las semillas, y la suspensión se agitó durante 10 minutos. Después de la retirada de la disolución de lejía/detergente, las semillas se lavaron cinco veces en H₂O destilada estéril. Las semillas se almacenaron en el último agua de lavado a 4°C durante 2 días en oscuridad antes de ponerse en placa con medio de selección con antibióticos (1 X de sales de Murashige y Skoog (el pH se ajustó a 5,7 con KOH 1 M), 1 X vitaminas B5 de Gamborg, 50 fitagar al 0,9% (Life Technologies), y 50 mg/l de kanamicina). Las semillas se germinaron en condiciones de iluminación continua (50-75 μ E/m²/sec) a 22-23°C. Después de 7-10 días de crecimiento en estas condiciones, fueron visibles y se obtuvieron los transformantes primarios resistentes a kanamicina (generación T1). Estas plántulas se transfirieron primero a placas con medio de selección fresco donde las plántulas siguieron creciendo durante 3-5 días más, y a continuación a tierra (medio para macetas Pro-Mix BX).

[0172] Los transformantes primarios se cruzaron y las semillas de la progenie (T2) se recogieron; se seleccionaron y analizaron las plántulas resistentes a kanamicina. Los niveles de expresión de los polinucleótidos recombinantes en los transformantes varía desde un incremento en el nivel de expresión del 5% aproximadamente hasta un incremento en el nivel de expresión de al menos el 100%. Se hicieron observaciones similares con

respecto al nivel de expresión del polipéptido.

Ejemplo VI: Identificación de plantas *Arabidopsis* con genes inactivados de los factores de transcripción

- 5 **[0173]** La selección de colecciones de *Arabidopsis* mutadas para su inserción por mutantes defectuosos en un gen diana conocido se realizó esencialmente como se describe en Krysan y col. (1999) Plant Cell 11:2283-2290. Resumiendo, se diseñaron cebadores específicos del gen, anidados en 5-250 pares de bases entre sí, a partir de las regiones 5' y 3' de un gen diana conocido. De manera similar, también se crearon grupos de cebadores anidados específicos para cada uno de los extremos de T-ADN o del transposón (los bordes "derecho" e "izquierdo"). Se usaron todas las posibles combinaciones de cebadores específicos para el gen y para el T-ADN/transposón para detectar mediante PCR un caso de inserción dentro o cerca del gen diana. Los fragmentos de ADN amplificados se secuenciaron a continuación, lo que permite la determinación precisa del punto de inserción del T-ADN/transposón en relación al gen diana. Los casos de inserción dentro de la secuencia codificante o interpuesta de los genes fueron desentrañados a partir de un conjunto que comprende una pluralidad de casos de inserción en una única planta mutante para su caracterización funcional. El procedimiento se describe con más detalle en Yu y Adam, solicitud de EE.UU. N° de serie 09/177.733 presentada el 23 de octubre de 1998.

Ejemplo VII: Identificación de fenotipos modificados en plantas con sobreexpresión o con genes inactivados

- 20 **[0174]** Se llevaron a cabo experimentos para identificar aquellos transformantes o con genes inactivados que presentaban una mayor tolerancia al estrés ambiental. Para dichos estudios, los transformantes se expusieron a una variedad de condiciones de estrés ambientales. Las plantas se expusieron a estrés por frío (6 horas de exposición a 4-8°C), estrés por calor (6 horas de exposición a 32-37°C), estrés salino elevado (6 horas de exposición a NaCl 200 mM), estrés por sequía (168 horas después de retirar el agua de las bandejas), estrés osmótico (6 horas de exposición a manitol 3 M), o limitación de nutrientes (nitrógeno, fosfato, y potasio) (Nitrógeno: todos los componentes del medio MS permanecieron constantes excepto el N que se redujo a 20 mg/l de NH₄NO₃, o Fosfato: todos los componentes del medio MS excepto el KH₂PO₄, que se sustituyó por K₂SO₄, Potasio: todos los componentes del medio MS excepto la eliminación de KNO₃ y KH₂PO₄, que se sustituyeron por NaH₄PO₄).
- 30 **[0175]** Se llevaron a cabo experimentos para identificar aquellos transformantes o con genes inactivados que presentaban unas características estructurales y de desarrollo modificadas. Para dichos estudios, los transformantes se observaron ocularmente para identificar nuevas características estructurales o de desarrollo asociadas a la expresión ectópica de los polinucleótidos o polipéptidos.
- 35 **[0176]** Se midió el tiempo de floración por el número de hojas en el botón presentes cuando es evidente una inflorescencia visible de aproximadamente 3 cm. El botón y el número total de hojas en el tallo de la progenie están íntimamente correlacionados con el momento de la floración (Koornneef y col. (1991) Mol. Gen. Genet 229:57-66). Se midió la respuesta de vernalización. Para los tratamientos de vernalización, las semillas se plantaron en placas de agar MS, se precintaron con cinta de microporos, y se pusieron en una sala fría a 4°C con bajos niveles de luz durante 6-8 semanas. A continuación las placas se transfirieron a las salas de crecimiento junto a las placas que contienen los controles no vernalizados recién plantados. Se contaron las hojas del botón cuando fue evidente una inflorescencia visible de aproximadamente 3 cm.
- 45 **[0177]** En la Tabla 4 se proporcionan los fenotipos modificados observados para plantas que sobreexpresan o con genes inactivados. Para una planta que sobreexpresa particular que presenta una característica menos beneficiosa, puede ser más útil seleccionar una planta con una expresión reducida del factor de transcripción particular. Para una planta con genes inactivados particular que presenta una característica menos beneficiosa, puede ser más útil seleccionar una planta con una expresión incrementada del factor de transcripción particular.
- 50 **[0178]** Las secuencias descritas en el presente documento se pueden usar para preparar plantas transgénicas y plantas con rasgos alterados. Las plantas transgénicas específicas listadas a continuación se produjeron a partir de las secuencias del Listado de secuencias, como se ha indicado. La Tabla 4 proporciona ejemplos de secuencias de polinucleótidos y polipéptidos. La Tabla 4 incluye, de izquierda a derecha para cada secuencia: la primera columna muestra la SEQ ID N° del polinucleótido; la segunda columna muestra el Mendel Gene ID N°, GID; la tercera columna muestra el rasgo(s) resultante de la inactivación del gen o sobreexpresión del polinucleótido en la planta transgénica; la cuarta columna muestra la categoría del rasgo; la quinta columna muestra la familia de factores de transcripción a la que pertenece el polinucleótido; la sexta columna ("Comentario"), incluye efectos y utilidades específicos conferidos por el polinucleótido de la primera columna; la séptima columna muestra la SEQ ID N° del polipéptido codificado por el polinucleótido; y la octava columna muestra las posiciones de los

restos aminoácidos del dominio conservado en coordenadas de aminoácidos (AA).

Ejemplo VIII: Identificación de secuencias homólogas

5 **[0179]** Se identificaron secuencias homólogas de *Arabidopsis* y especies de plantas distintas de *Arabidopsis* usando herramientas para búsqueda de secuencias en bases de datos, tales como Basic Local Alignment Search Tool (BLAST) (Altschul y col. (1990) J. Mol. Biol. 215:403-410; y Altschul y col. (1997) Nucl. Acid Res. 25: 3389-3402). Los programas para el análisis de secuencias tblastx se emplearon usando la matriz de puntuación BLOSUM-62 (Henikoff, S. y Henikoff, J. G. (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919).

10

Ejemplo IX: Introducción de polinucleótidos en plantas dicotiledóneas

[0180] Las secuencias SEQ ID N°: 31 y 439, parálogas, ortólogas, y homólogas recombinadas en los vectores de expresión pMEN20 o pMEN65 se transformaron en una planta con el fin de modificar los rasgos de la planta. El vector de clonación se puede introducir en una variedad de plantas de cereal por medios muy conocidos en la técnica tales como, por ejemplo, transferencia directa de ADN o transformación mediada por *Agrobacterium tumefaciens*. Ahora es un proceso rutinario producir plantas transgénicas usando la mayoría de plantas dicotiledóneas (véase Weissbach y Weissbach, (1989) más arriba; Gelvin y col., (1990) más arriba; Herrera-Estrella y col. (1983) más arriba; Bevan (1984) más arriba; y Klee (1985) más arriba). Los procedimientos para el análisis de rasgos son rutinarios en la materia y más arriba se han descrito ejemplos de los mismos.

20

Ejemplo X: Transformación de plantas de cereal con un vector de expresión

[0181] Plantas de cereal como el maíz, trigo, arroz, sorgo o cebada, también se pueden transformar con las presentes secuencias de polinucleótidos en vectores de expresión pMEN20 o pMEN65 con el fin de modificar los rasgos de la planta. Por ejemplo, se puede modificar pMEN020 para sustituir la región codificante NptII con el gen BAR de *Streptomyces hygroscopicus* que confiere resistencia a fosfotricina. Los sitios KpnI y BgIII del gen Bar se eliminan por mutagénesis dirigida de sitio con cambios de codones silenciosos.

25

[0182] El vector de clonación se puede introducir en una variedad de plantas de cereal por medios muy conocidos en la materia tales como, por ejemplo, transferencia directa de ADN o transformación mediada por *Agrobacterium tumefaciens*. Ahora es un proceso rutinario producir plantas transgénicas de la mayoría de cultivos de cereales (Vasil, I., Plant Molec. Biol. 25: 925-937 (1994)) tales como maíz, trigo, arroz, sorgo (Cassas, A. y col., Proc. Natl. Acad. Sci USA 90: 11212-11216 (1993) y cebada (Wan, Y. y Lemeaux, P. Plant Physiol. 104:37-48 (1994)). Se pueden usar procedimientos de transferencia de ADN, tales como microproyectiles, para el maíz (Fromm, y col. Bio/Technology 8: 833-839 (1990); Gordon-Kamm y col. Plant Cell 2: 603-618 (1990); Ishida, Y., Nature Biotechnology 14:745-750 (1990)), el trigo (Vasil, y col. Bio/Technology 10:667-674 (1992); Vasil y col., Bio/Technology 11:1553-1558 (1993); Weeks y col., Plant Physiol. 102:1077-1084 (1993)), y el arroz (Christou Bio/Technology 9:957-962 (1991); Hiei y col. Plant J. 6:271-282 (1994); Aldemita y Hodges, Planta 199:612-617; Hiei y col., Plant Mol Biol. 35:205-18 (1997)). Para la mayoría de plantas de cereal, las células embriogénicas derivadas de tejidos del escutelo inmaduro son las dianas celulares preferidas para la transformación (Hiei y col., Plant Mol Biol. 35:205-18 (1997); Vasil, Plant Molec. Biol. 25: 925-937 (1994)).

30

35

40

[0183] Los vectores según la presente descripción se pueden transformar en células embriogénicas de maíz derivadas del tejido escutelar inmaduro usando bombardeo con microproyectiles, siendo el genotipo A188XB73 el genotipo preferido (Fromm, y col., Bio/Technology 8: 833-839 (1990); Gordon-Kamm y col., Plant Cell 2: 603-618 (1990)). Después del bombardeo con microproyectiles los tejidos se seleccionan en fosfotricina para identificar las células embriogénicas transgénicas (Gordon-Kamm y col., Plant Cell 2: 603-618 (1990)). Las plantas transgénicas se regeneran mediante técnicas estándar de regeneración del maíz (Fromm, y col., Bio/Technology 8: 833-839 (1990); Gordon-Kamm y col., Plant-Cell 2: 603-618 (1990)).

45

50

[0184] Los plásmidos preparados como se ha descrito anteriormente también se pueden usar para producir plantas de trigo y arroz transgénicas (Christou, Bio/Technology 9:957-962 (1991); Hiei y col., Plant J. 6:271-282 (1994); Aldemita y Hodges, Planta 199:612-617 (1996); Hiei y col., Plant Mol Biol. 35:205-18 (1997)) que expresan de manera coordinada genes de interés siguiendo los protocolos de transformación estándar conocidos por los expertos en la materia para el arroz y el maíz (Vasil, y col. Bio/Technology 10:667-674 (1992); Vasil y col., Bio/Technology 11:1553-1558 (1993); Weeks y col., Plant Physiol. 102:1077-1084 (1993)), donde se usa el gen bar como marcador seleccionable.

55

<110> Mendel Biotechnology, Inc

<120> POLINUCLEÓTIDOS Y POLIPÉPTIDOS EN PLANTAS CON EL FÍN DE MEJORAR SUS CARACTERÍSTICAS

<150> 60/310,847

5 <151> 2001-08-09

<150> 60/336,049

<150> 2001-11-19

10 <150> 60/338,692

<150> 2011-12-11

<150> 10/171,468

<150> 2022-06-14

15 >G47 (38 .. 472)

CTTCTTCTTCACATCGATCATCATACAACAACAAAAATGGATTACAGAGAATCCACCGG
TGAAAGTCAGTCAAAGTACAAAGGAATCCGTCGTCGGAAATGGGGCAAATGGGTATCAGA
GATTAGAGTTCCGGGAACTCGTGACCGTCTCTGGTTAGGTTTCATTCTCAACAGCAGAAGG
TGCCGCCGTAGCACACGACGTTGCTTTCTTCTGTTTACACCAACCTGATTCTTTAGAATC
TCTCAATTTCCCTCATTGCTTAATCCTTCACTCGTTTCCAGAACTTCTCCGAGATCTAT
CCAGCAAGCTGCTTCTAACGCCGGCATGGCCATGACGCCGGAATCGTCCACAGTACCAG
CGTGAACCTCTGGATGCGGAGATACGACGACGTATTACGAGAATGGAGCTGATCAAGTGA
GCCGTTGAATATTTGAGTGTATGATTATCTGGGCGGCCACGATCACGTTTGATTTATCTC
GACGGTCATGATCACGTTTGATCTTCTTTTGAGTAAGATTTTGTACCATAATCAAAACAG
GTGTGGTGCTAAAACTTACTCAAAACAAGATTAGGTACCACAGAGAAACAATCAAATGG
TTGTGAATATACATTATAAGGTTTGATTAATGTTTGTTCCTGATTTAGTGAAGTTG
GTCCATTGTATACAAATCTATTCAAGAAACCTAGCGCGAGATCATGTTTCGTGATTGAAG
ATTGAGATTTTTAAGTATTCGTAATATTTTTGTAAAATACAAATAAAAAAAAAAAAAA
AAAAA

>Secuencia de aminoácidos G47 (Dominio en las coordenadas AA: 11-80)
MDYRESTGESQSKYKGIRRRKWKWVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFVCL
HQPDSLES LNFP HLLNPSLVSR TSPRSIQQAASNAGMAIDAGIVHSTSVNSGCGDPTTY
ENGADQVEPLNISVYDYLGGHDHV*

>G2133 (26 .. 457)

20

ATCTCATCTTCATCCACCCAAAAACATGGATTCAAGAGACACCGGAGAAACTGACCAGAG
CAAGTACAAAGGTATCCGTCGTCGGAAATGGGGAAAATGGGTATCAGAGATTCGTGTCCC
GGGAACTCGTCAACGCTCTCTGGTTAGGCTCTTTCTCCACCGCAGAAGGCGCTGCCGTAGC
CCACGACGTCGCTTTTTACTGCTTGCACCGACCATCTTCCCTCGACGACGAATCTTTTAA
CTTCCCTCACTTACTTACAACCTCCCTCGCTCCAATATATCTCCTAAGTCCATCCAAAA
AGCTGCTTCCGACGCCGGCATGGCCGTGGACGCCGATTCCATGGTGCTGTGCTGTTGGGAG
TGGTGGTTGTGAAGAGAGATCTTCCATGGCGAATATGGAGGAGGAGGACAACTTAGTAT
CTCCGTGTATGATTATCTTGAAGACGATCTCGTTGATCTATACGAGTACGTTTTTAGCA
GTTAA

>Secuencia de aminoácidos G2133 (Dominio en las coordenadas AA: 11-83)
MDSRDTGETDQSKYKGIRRRKWKWVSEIRVPGTRQRLWLGSFSTAEGAAVAHDVAFYCL
HRPSSLDDSEFNFP HLLTSLASNISPKSIQKAASDAGMAVDAGFHGAVSGSGGCEERS
MANMEEEDKLSISVYDYLEDDLIV*

REIVINDICACIONES

1. Un procedimiento para la producción de una planta transgénica que tiene un rasgo alterado seleccionado del grupo constituido por una mayor resistencia al estrés osmótico, una arquitectura alterada, una estructura de los tejidos vasculares alterada, una floración retardada, y una morfología del tallo alterada comparada con una planta de tipo silvestre de la misma especie, el procedimiento que comprende:

(a) la proporción de un polinucleótido que codifica un polipéptido con una identidad de al menos el 95% con

**MDYRESTGESQSKYKGIRRRKWGKVVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFFCL
 HQPDSLESLNFPHLLNPSLVSRTSPRSIQQAASNAGMAIDAGIVHSTSVNSGCGDTTTY
 ENGADQVEPLNISVYDYLGGHDHV***

10

y que comprende un dominio conservado con una identidad de secuencia de al menos el 95% a los aminoácidos con las coordenadas 11 a 80 de

**MDYRESTGESQSKYKGIRRRKWGKVVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFFCL
 HQPDSLESLNFPHLLNPSLVSRTSPRSIQQAASNAGMAIDAGIVHSTSVNSGCGDTTTY
 ENGADQVEPLNISVYDYLGGHDHV***

15

- (b) la introducción del polinucleótido en una planta para producir una planta transgénica; y
 (c) la selección de la planta transgénica que comprende la primera secuencia de nucleótidos y que tiene dicho rasgo seleccionado alterado en comparación con la planta de tipo silvestre.

20 2.

El procedimiento de la reivindicación 1 donde dicho polipéptido comprende

**MDYRESTGESQSKYKGIRRRKWGKVVSEIRVPGTRDRLWLGSFSTAEGAAVAHDVAFFCL
 HQPDSLESLNFPHLLNPSLVSRTSPRSIQQAASNAGMAIDAGIVHSTSVNSGCGDTTTY
 ENGADQVEPLNISVYDYLGGHDHV***

3. El procedimiento de la reivindicación 1 ó 2, donde el polinucleótido comprende un promotor constitutivo, inducible, o específico de tejido unido de manera operable a la primera secuencia de nucleótidos.

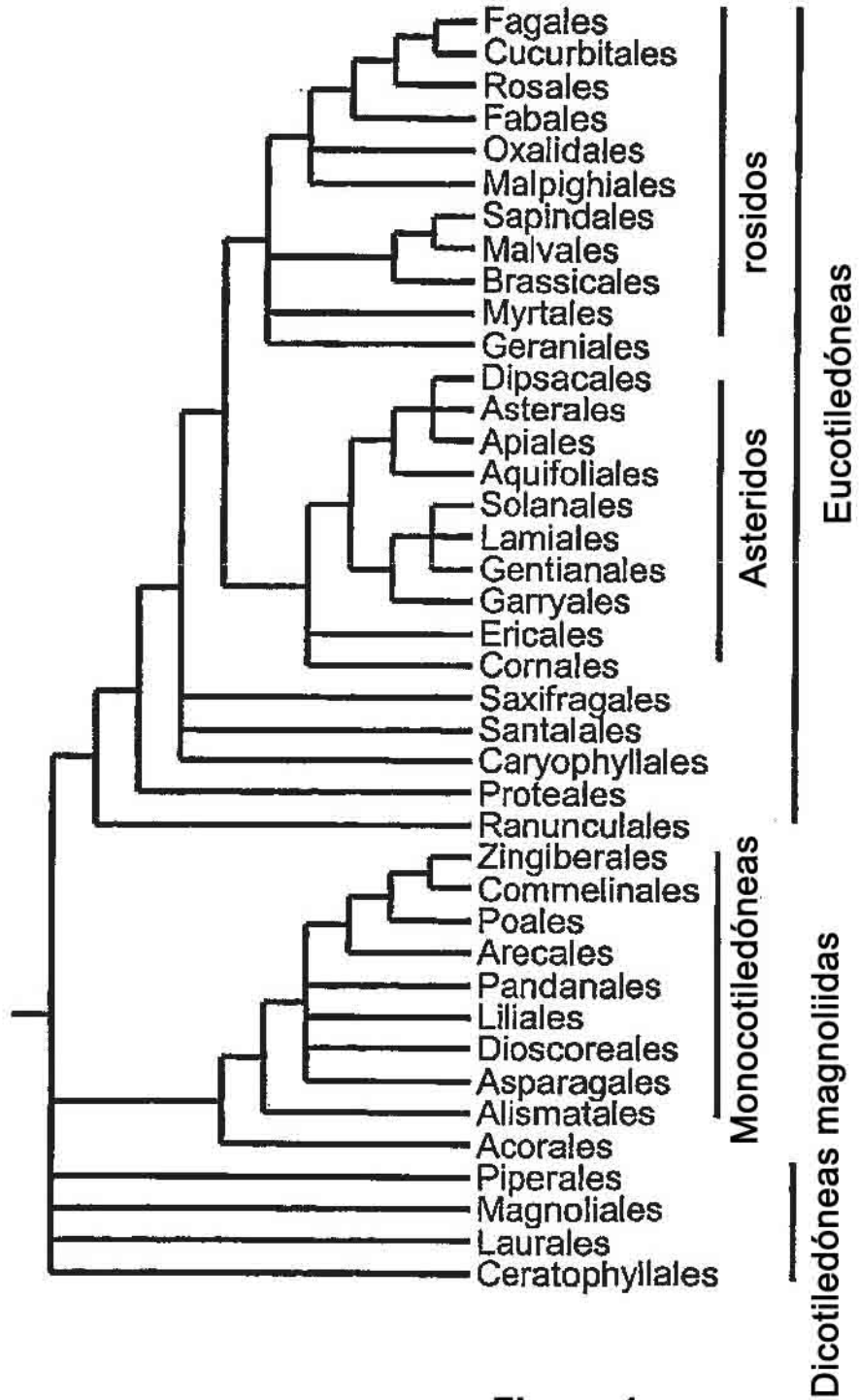


Figura 1