

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 480 665**

51 Int. Cl.:

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.09.2009 E 09792298 (3)**

97 Fecha y número de publicación de la concesión europea: **30.04.2014 EP 2334795**

54 Título: **Composiciones y métodos para la expresión de una secuencia de nucleótidos heteróloga en plantas**

30 Prioridad:

08.09.2008 US 95134 P

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

28.07.2014

73 Titular/es:

**ATHENIX CORPORATION (100.0%)
2500 Paramount Parkway
Morrisville, NC 27560 , US**

72 Inventor/es:

**HAMMER, PHILIP, E.;
BEILINSON, VADIM y
HINSON, TODD, K.**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 480 665 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

Composiciones y métodos para la expresión de una secuencia de nucleótidos heteróloga en plantas5 **Campo de la invención**

[0001] La presente invención se refiere al campo de la biología molecular de plantas, más concretamente a la identificación y uso de elementos reguladores en plantas.

10 **Antecedentes de la invención**

[0002] La biogénesis del cloroplasto en las plantas depende de las actividades coordinadas de dos sistemas genéticos independientes localizados en el cloroplasto y el núcleo (véase Cline y Henry (1996), Annu. Rev. Cell Dev. Biol. 12, 1-26). Las grandes proteínas constituyentes del cloroplasto son codificadas por los genes nucleares y se sintetizan citoplásmicamente como formas precursoras que contienen prolongaciones N-terminales conocidas como péptidos de tránsito. El péptido de tránsito es decisivo para el reconocimiento específico de la superficie del cloroplasto y en la mediación de la translocación post-traducciona de las pre-proteínas a través de la envoltura del cloroplasto y de allí a los diversos subcompartimentos diferentes dentro del cloroplasto (por ejemplo, el estroma, el tilacoide y la membrana del tilacoide).

[0003] Los genes que se ha informado que tienen secuencias de péptidos de tránsito codificadas naturalmente en su extremo N-terminal incluyen la subunidad pequeña del cloroplasto de la ribulosa-1,5-bisfosfatocarboxilasa (RuBisCo), de Castro Silva Filho et al. (1996) Plant Mol. Biol. 30: 769-780; Schnell, D. J. et al. (1991) J. Biol. Chem. 266 (5): 3335-3342; la 5-(enolpiruvil)shikimato-3-fosfato sintasa (EPSPS), Archer et al. (1990) J. Bioenerg. y Biomemb. 22 (6):789-810; la triptófano sintetasa. Zhao, J. et al. (1995) J. Biol. Chem. 270 (11):6081-6087; la plastocianina, Lawrence et al. (1997) J. Biol. Chem. 272 (33):20357-20363; la corismato sintasa, Schmidt et al. (1993) J. Biol. Chem. 268 (36):27477-27457; y la proteína de unión a clorofila a/b que recoge la luz (LHBP), Lamppa y col. (1988) J. Biol. Chem. 263: 14996-14999, aunque no todas estas secuencias han sido útiles para la expresión heteróloga de proteínas dirigidas a cloroplastos en plantas superiores.

30 **Compendio de la invención**

[0004] Se proporcionan composiciones y métodos para dirigir al cloroplasto polipéptidos en una planta. Las composiciones comprenden casetes de expresión que comprenden una secuencia de nucleótidos que codifica una secuencia peptídica dirigida a cloroplastos (o péptido de tránsito al cloroplasto, "CTP") derivada de un organismo de algas conectada operablemente a la secuencia de nucleótidos de interés.

[0005] Por consiguiente, la presente invención proporciona una célula vegetal monocotiledónea que tiene incorporado de manera estable en su genoma un casete de expresión que comprende una secuencia de nucleótidos que codifica un péptido dirigido al cloroplasto (CTP), en donde dicha secuencia de nucleótidos que codifica dicho CTP está conectada operablemente a una secuencia de nucleótidos de interés, y en donde dicha secuencia de nucleótidos que codifica dicho CTP se selecciona del grupo que consiste en: a) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6; b) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto; c) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y, d) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto.

[0006] Los constructos de expresión son útiles para la expresión y el correcto direccionamiento de la secuencia de nucleótidos de interés en una planta monocotiledónea. También se mencionan los vectores que comprenden los casetes de expresión, y las plantas y células de plantas que tienen incorporado establemente o expresado transitoriamente en sus genomas un casete de expresión ha descrito anteriormente. Adicionalmente, las composiciones incluyen semillas transgénicas de tales plantas.

[0007] Por lo tanto la invención proporciona adicionalmente:

- una planta que comprende la célula vegetal monocotiledónea de la invención; y
- una semilla derivada de tal planta en donde la semilla comprende la secuencia de nucleótidos que codifica dicho péptido de tránsito al cloroplasto.

[0008] También se proporcionan métodos para la expresión de una secuencia de nucleótidos en una planta o célula vegetal, así como métodos para identificar secuencias de CTP de algas para uso en una planta. En particular, la invención proporciona un método para expresar una secuencia de nucleótidos de interés en una planta monocotiledónea, comprendiendo dicho método:

- a) introducir en una célula vegetal un casete de expresión que comprende una secuencia de nucleótidos que codifica un péptido dirigido al cloroplasto (CTP), en donde dicha secuencia de nucleótidos que codifica dicho CTP está conectada operablemente a dicha secuencia de nucleótidos de interés, y en donde dicha secuencia de nucleótidos que codifica dicho CTP se selecciona del grupo que consiste en i) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6; ii) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto; iii) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y, iv) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto; y,
- b) regenerar una planta transformada a partir de dicha célula vegetal; en donde dicha planta ha incorporado de forma estable en su genoma dicho casete de expresión.

Descripción de las figuras

La Figura 1 demuestra la expresión de TagGFP en protoplastos de tabaco.

La Figura 2 demuestra la expresión y el procesamiento del péptido de tránsito al cloroplasto de EPSPS de *Chlamydomonas* en células de maíz. Calle 1: Línea Hi-II de maíz no transgénico; Calles 2-6: Eventos To individuales transformados con el constructo de CTP de EPSPS de *Chlamydomonas*/GRG23(ace3)(R173K); Calle 7: Marcador de peso molecular de proteína; Calle 8: Proteína GRG23(ace3)(R173K) purificada (4 ng).

La Figura 3 demuestra el cálculo del peso molecular de la proteína EPSPS de *Chlamydomonas* - GRG23(ace3)(R173K) procesada expresada en el maíz.

[0010] Se utilizó una regresión lineal de la representación gráfica del Log del Peso molecular vs Distancia de Migración de los patrones de peso molecular de las proteínas de la Figura 2 para calcular el peso molecular aparente del patrón de proteína GRG23(ace3)(R173K) y EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K) procesada detectada en extractos de plantas.

Descripción detallada

[0011] En la producción de plantas transgénicas es a menudo útil dirigir proteínas foráneas a localizaciones subcelulares específicas, por ejemplo, el cloroplasto, la vacuola, la mitocondria, o el RE. Trabajadores previos han fusionado secuencias de ADN que codifican péptidos de tránsito de varios genes de plantas a genes de interés. Cuando se traduce el gen la proteína resultante tiene el péptido de tránsito de planta fusionado al extremo amino terminal de la proteína de interés, y por lo tanto la proteína se dirige, con diferente eficacia, al compartimento subcelular deseado.

[0012] Por lo tanto, la presente invención se refiere a composiciones y métodos para la dirigir al cloroplasto polipéptidos en plantas monocotiledóneas superiores o células de plantas monocotiledóneas. Las composiciones comprenden casetes de expresión que comprenden una secuencia de nucleótidos que codifica un péptido de tránsito al cloroplasto (CTP) derivado de un organismo de algas conectada operablemente a una secuencia de nucleótidos de interés, donde el CTP se selecciona del grupo que consiste en:

- a) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6;
- b) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto;
- c) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y,
- d) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto.

[0013] En una realización, el CTP deriva de *Chlamydomonas* sp. En otra realización, el CTP comprende la secuencia de aminoácidos mostrada en la SEC ID NO: 7 o una secuencia de aminoácidos codificada por el SEQ ID NO: 6. Se proporcionan plantas, células vegetales, y semillas transformadas.

[0014] Las secuencias que codifican el CTP, cuando se ensambla dentro de un constructo de ADN de tal manera que la secuencia que codifica CTP está conectada operablemente a una secuencia de nucleótidos de interés, facilitan el transporte co-traducciona l o post-traduccion del péptido de interés al cloroplasto de una célula vegetal transformada de manera estable con este constructo de ADN. Los métodos para expresar una secuencia de nucleótidos en una planta comprenden introducir en células vegetales un casete de expresión que comprende una secuencia de nucleótidos que codifica el CTP descrito en la presente memoria conectada operablemente a una secuencia de nucleótidos de interés, y regenerar una planta transformada a partir de la célula vegetal.

[0015] Los artículos "un" y "una" se utilizan en la presente memoria para referirse a uno o más de uno (es decir, a al

menos uno) del objeto gramatical del artículo. A modo de ejemplo, "un elemento" significa uno o más elementos.

[0016] Tal como se utiliza en la presente memoria, se pretende que el término "molécula de ácido nucleico" incluya moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de ARN (por ejemplo, ARNm) y análogos del ADN o ARN generados utilizando análogos de nucleótidos. La molécula de ácido nucleico puede ser de hebra sencilla o de doble hebra, pero preferiblemente es ADN de doble hebra.

Péptidos de tránsito al cloroplasto

[0017] Los cloroplastos son orgánulos que se encuentran en células vegetales y algas eucarióticas que llevan a cabo la fotosíntesis. El cloroplasto es un orgánulo celular complejo compuesto de tres membranas: la membrana de la envoltura interna, la membrana de la envoltura externa, y la membrana del tilacoide. Las membranas encierran conjuntamente tres compartimentos acuosos denominados espacio intermedio, estroma, y el lumen de los tilacoides. Si bien los cloroplastos contienen su propio genoma circular, muchas proteínas constituyentes de cloroplastos están codificadas por los genes nucleares y se sintetizan citoplásmicamente como formas precursoras que contienen prolongaciones N-terminales conocidas como péptidos de tránsito al cloroplasto (CTP). El CTP es decisivo para el reconocimiento específico de la superficie del cloroplasto y en la mediación de la translocación post-traducciona de las pre-proteínas a través de la envoltura del cloroplasto y a los diversos subcompartimentos diferentes dentro del cloroplasto (por ejemplo, el estroma, el tilacoide y la membrana del tilacoide).

[0018] Se han identificado al menos dos dominios funcionales distintos en péptidos de tránsito al cloroplasto: el dominio de direccionamiento al estroma (DDE) y el dominio de direccionamiento al lumen (DDL). Los DDE regulan el acceso a la ruta de importación general y son necesarios y suficientes para la importación de la proteína pasajera al estroma. Los precursores de proteínas del estroma poseen péptidos de tránsito que contienen sólo un DDE, mientras que los precursores de proteínas luminales del tilacoide tienen tanto un DDE como un DDL.

[0019] Los DDE varían en longitud de aproximadamente 30 a 120 residuos y son ricos en residuos hidroxilados y deficientes en residuos ácidos. Tienden a compartir varios motivos compositivos: un amino terminal de 10-15 residuos carece de residuos de Gly, Pro y cargados; una región media variable rica en Ser, Thr, Lys y Arg; y una región carboxi-proximal con una secuencia débilmente conservada (Ile/Val-X-Ala/Cys-Ala; SEC ID NO: 17) para el procesamiento proteolítico. Sin embargo, no hay bloques extensos de conservación de la secuencia, ni ningún motivos estructural secundario conservado. Los análisis teóricos sugieren que los DDE adoptan conformaciones helicoidales predominantemente al azar.

[0020] Los genes que se ha informado que tienen secuencias de péptidos de tránsito codificadas naturalmente en su extremo N-terminal incluyen la subunidad pequeña del cloroplasto de la ribulosa-1,5-bisfosfato carboxilasa (RuBisCo), de Castro Silva Filho et al. (1996) *Plant Mol. Biol.* 30: 769-780; Schnell, D. J. et al. (1991) *J. Biol. Chem.* 266 (5): 3335-3342; la 5-(enolpiruvil)shikimato-3-fosfato sintasa (EPSPS), Archer et al. (1990) *J. Bioenerg. and Biomemb.* 22 (6): 789-810; la triptófano sintetasa. Zhao, J. et al. (1995) *J. Biol. Chem.* 270 (11):6081-6087; la plastocianina, Lawrence et al. (1997) *J. Biol. Chem.* 272 (33):20357-20363; la corismato sintasa, Schmidt et al. (1993) *J. Biol. Chem.* 268 (36):27477-27457; y la proteína de unión a clorofila a/b que recoge la luz (LHBP), Lamppa y col. (1988) *J. Biol. Chem.* 263:14996-14999. Aunque se han descrito varios CTP, sólo unos pocos se han utilizado con éxito en los intentos de dirigir moléculas quiméricas a los cloroplastos en plantas superiores.

[0021] La presente invención describe el uso de los CPT derivados de especies de algas, particularmente *Chlamydomonas sp.*, en plantas superiores donde los CTP se seleccionan del grupo que consiste en: a) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6; b) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto; c) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y, d) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto.

[0022] Para los propósitos de la presente invención, las "plantas superiores" son consideradas miembros del subreino Embryophytae. En una realización, el CTP útil en los métodos y composiciones descritos en la presente memoria deriva de *Chlamydomonas*. En otra realización, el CTP se muestra en el SEQ ID NO: 7, o es codificado por la SEC ID NO: 6.

[0023] Los CTP descritos en la presente memoria son útiles para dirigir un polipéptido al cloroplasto de una célula vegetal. En una realización, los CTP descritos en esta memoria proporcionan una mejor translocación en comparación con los CTP derivados de, por ejemplo, organismos de plantas superiores. Los CTP descritos en la presente memoria pueden dar como resultado al menos aproximadamente 20%, al menos aproximadamente 30%, al menos aproximadamente 40%, al menos aproximadamente 50%, al menos aproximadamente 60%, al menos aproximadamente 70%, al menos aproximadamente 80%, al menos aproximadamente 90%, al menos

aproximadamente 100%, o más, o al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, o más de mejora en la translocación del polipéptido al cloroplasto cuando se compara con un CTP de referencia. Se puede medir una mejora en términos de la cantidad de polipéptido que se transloca al cloroplasto, la cantidad de polipéptido activo que se transloca al cloroplasto, o ambos. También se puede medir una mejora en términos de una mejora en el fenotipo de un organismo transformado con la proteína dirigida al cloroplasto de interés. Por ejemplo, cuando el CTP se utiliza para dirigir una proteína de resistencia a herbicidas al cloroplasto de la planta, se puede medir una mejora en la actividad en términos de una mejora en la resistencia a los herbicidas.

10 *Casetes de expresión*

15 **[0024]** Las secuencias que codifican CTP se pueden proporcionar en un casete de expresión que permite dirigir la expresión y localización de un polipéptido codificado por la secuencia de nucleótidos de interés al cloroplasto de las células vegetales. Por "casete de expresión" se quiere significar un constructo de ADN que es capaz de dar como resultado la expresión de una proteína de un marco de lectura abierto en una célula. El casete incluirá, en la dirección de transcripción 5'-3', una región de inicio de la transcripción que comprende preferiblemente un promotor adecuado para la expresión en una célula vegetal de interés, conectado operablemente a una secuencia que codifica CTP como se describe en la presente memoria, que está conectada operablemente a una secuencia de nucleótidos de interés, y una región de terminación de la traducción y de la transcripción (es decir, región de terminación) funcional en plantas. La secuencia de nucleótidos que codifica CTP y la secuencia de nucleótidos de interés pueden estar separadas entre sí por secuencias de nucleótidos que codifican uno o más aminoácidos "conectores" como se comenta en otra parte en la presente memoria.

25 **[0025]** El casete puede contener adicionalmente al menos un gen adicional que se va a co-transformar en el organismo, tal como un gen marcador seleccionable. Alternativamente, el gen o los genes adicionales pueden proporcionarse en múltiples casetes de expresión. Tal casete de expresión se proporciona con una pluralidad de sitios de restricción para la inserción de la secuencia de nucleótidos de interés para estar bajo la regulación transcripcional de las regiones reguladoras.

30 **[0026]** El casete de expresión puede comprender adicionalmente una o varias regiones 3' y/o 5' no traducidas. Por "región 3' no traducida" se quiere significar una secuencia de nucleótidos localizada aguas abajo de una secuencia codificante. Las secuencias señal de poliadenilación y otras secuencias que codifican señales reguladoras capaces de afectar a la adición de extensiones de ácido poliadenílico al extremo 3' del precursor de ARNm son las regiones 3' no traducidas. Por "región 5' no traducida" se quiere significar una secuencia de nucleótidos localizada aguas arriba de una secuencia codificante. Otros elementos no traducidos aguas arriba o aguas abajo incluyen los potenciadores. Los potenciadores son secuencias de nucleótidos que actúan incrementando la expresión de una región promotora. Los potenciadores son bien conocidos en la técnica e incluyen, pero no se limitan a, la región potenciadora de SV40 y el elemento potenciador 35S.

40 **[0027]** La región de terminación puede ser nativa con la secuencia de nucleótidos que codifica CTP, puede ser nativa con la secuencia de nucleótidos de interés, o puede derivar de otra fuente. Las regiones de terminación convenientes están disponibles en el plásmido Ti de *A. tumefaciens*, tales como las regiones de terminación de la octopina sintasa y de la nopalina sintasa. Ver también Guerineau et al. (1991) Mol. Gen. Genet. 262:141-144; Proudfoot (1991) Cell 64:671-674; Sanfacon et al. (1991) Genes Dev. 5:141-149; Mogen et al. (1990) Plant Cell 2:1261-1272; Munroe et al. (1990) Gene 91:151-158; Ballas et al. (1989) Nucleic Acids Res. 17:7891-7903; y Joshi et al. (1987) Nucleic Acid Res. 15:9627-9639.

50 **[0028]** Los casetes de expresión descritos en la presente memoria pueden comprender adicionalmente uno o más elementos reguladores distintos de CTP, así como los CTP adicionales conocidos en la técnica. Por "elemento regulador" o "región reguladora" se quiere significar una porción de ácido nucleico encontrada aguas arriba o aguas abajo de un gen, que puede estar compuesta por ADN o ARN, o por ADN y ARN y que está implicada en la expresión génica. Los elementos reguladores pueden ser capaces de mediar la especificidad de órgano, o controlar la activación de genes de desarrollo o temporales e incluir elementos promotores, elementos promotores esenciales, elementos que son inducibles en respuesta a un estímulo externo, elementos que se activan constitutivamente, terminadores de la transcripción, señales de poliadenilación, y elementos que disminuyen o aumentan la actividad del promotor tales como elementos reguladores negativos o potenciadores de la transcripción, respectivamente. Por "que actúa en cis" se quiere significar una secuencia que es físicamente contigua a la secuencia transcrita. Las secuencias que actúan en cis interactúan típicamente con proteínas u otras moléculas para llevar a cabo (activar/desactivar, regular, modular, etc.) la transcripción. Por "promotor de la transcripción" se quiere significar una secuencia de ácido nucleico que, cuando se coloca próxima a un promotor y está presente en un medio de transcripción capaz de soportar la transcripción, confiere una mayor actividad de transcripción en comparación con la que resulta a partir del promotor en ausencia del potenciador. Los potenciadores pueden actuar aguas arriba, dentro, o aguas abajo de un gen, incluso tan lejos como 50 kilobases del sitio de inicio de la transcripción. Los potenciadores también pueden funcionar independientemente de su orientación. Por "terminador de la transcripción"

se quiere significar una secuencia de ADN que incluye una secuencia de pares de bases de nucleótidos necesaria para reducir o eliminar la transcripción. Por "señal de poliadenilación" se quiere significar una secuencia que controla la terminación de la transcripción y la traducción.

5 **[0029]** En un aspecto, las secuencias de ADN sintéticas están diseñadas para un polipéptido dado, tales como los polipéptidos dirigidos al cloroplasto útiles en los métodos descritos en la presente memoria. La expresión del marco de lectura abierto de la secuencia de ADN sintético en una célula da como resultado la producción del polipéptido. Las secuencias de ADN sintéticas pueden ser útiles para eliminar simplemente los sitios de reconocimiento de endonucleasas de restricción no deseados, para facilitar estrategias de clonación de ADN, para alterar o eliminar cualquier potencial sesgo codónico, para alterar o mejorar el contenido de GC, para suprimir o alterar los marcos de lectura alternativos, y/o para alterar o eliminar sitios de reconocimiento de corte y empalme de intrones/exones, sitios de poliadenilación, secuencias de Shine-Delgarno, elementos promotores no deseados y similares que pueden estar presentes en una secuencia de ADN nativa. También es posible que las secuencias de ADN sintéticas se puedan utilizar para introducir otras mejoras en una secuencia de ADN, tales como la introducción de una secuencia de intrón, la creación de una secuencia de ADN que es expresada como una proteína de fusión a secuencias de direccionamiento a orgánulos, tales como péptidos de tránsito al cloroplasto, péptidos del direccionamiento al apoplasto/vacuolas, o secuencias de péptidos que dan como resultado la retención del péptido resultante en el retículo endoplásmico. Los genes sintéticos también se pueden sintetizar utilizando codones de células anfitrionas preferidas para la expresión mejorada, o se pueden sintetizar utilizando codones con una frecuencia de uso de codones anfitriones preferidos. Véase, por ejemplo, Campbell y Gowri (1990) Plant Physiol. 92:1-11; Patentes de los Estados Unidos Núms. 6.320.100; 6.075.185; 5.380.831; y 5.436.391, Solicitudes de Patente de los Estados Unidos Publicadas Núms. 20040005600 y 20010003849, y Murray et al. (1989) Nucleic Acids Res. 17:477-498.

25 **[0030]** Los ácidos nucleicos de interés que se van a dirigir al cloroplasto también pueden ser optimizados para la expresión en el cloroplasto para dar cuenta de las diferencias en el uso de codones entre el núcleo de la planta y este orgánulo. De esta manera, los ácidos nucleicos de interés pueden ser sintetizados utilizando codones preferidos de cloroplastos. Véase, por ejemplo, La Patente de los Estados Unidos Núm. 5.380.831.

Variantes y fragmentos

30 **[0031]** También se hace referencia a las moléculas de ácido nucleico que son fragmentos de las secuencias de CTP descritas. Los empleados en la invención tienen por lo menos 90% de identidad de secuencia con el SEQ ID NO: 6. Mediante "fragmento" se quiere significar una porción de la secuencia de CTP. Un fragmento de una secuencia de nucleótidos puede ser biológicamente activo y por lo tanto puede ser capaz de facilitar la translocación de un polipéptido de interés al cloroplasto de una planta, o puede ser un fragmento que puede ser utilizado como una sonda de hibridación o cebador de PCR usando los métodos descritos a continuación. Los análisis para determinar si tales fragmentos tienen actividad CTP son bien conocidos en la técnica.

40 **[0032]** Las moléculas de ácido nucleico que son fragmentos de una secuencia de nucleótidos que codifica CTP descritas en la presente memoria pueden comprender al menos aproximadamente 90, 100, 125, 150, 175, 200, 225, 250, 275, 300, nucleótidos contiguos, o hasta el número de nucleótidos presente en una secuencia de CTP completa descrita en la presente memoria (por ejemplo, 306 nucleótidos para el SEQ ID NO: 1) dependiendo del uso previsto. Mediante nucleótidos "contiguos" se quieren significar residuos de ácido nucleico que son inmediatamente adyacentes entre sí. Los fragmentos biológicamente activos de las secuencias que codifican los CTP codificará un CTP que conserva la actividad. Mediante "conserva la actividad de CTP" se quiere significar que el fragmento dirigirá la translocación al cloroplasto de al menos aproximadamente 30%, al menos aproximadamente 50%, al menos aproximadamente 70%, o al menos aproximadamente 80% del polipéptido codificado por la secuencia de nucleótidos de interés. En un caso, un fragmento de una secuencia de nucleótidos que codifica los CTP descrita en la presente memoria puede comprender una o más delecciones del SEQ ID NO: 6, incluyendo hasta 50 aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 12, aproximadamente 15, aproximadamente 18 o aproximadamente 21 delecciones. En otra realización, un fragmento de una secuencia de nucleótidos que codifica los CTP descrita en la presente memoria puede codificar un aminoácido que comprende una o más delecciones del SEQ ID NO: 7, incluyendo hasta aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6 o aproximadamente 7 delecciones de aminoácidos.

60 **[0033]** Se puede preparar una porción biológicamente activa de un CTP aislando una porción de una de las secuencias de CTP y evaluando la actividad de la porción de la CTP. Los métodos para medir la actividad de los CTP son bien conocidos en la técnica. Véase la sección titulada "Evaluación de la actividad de CTP" para los ejemplos de los métodos adecuados.

[0034] También están incluidas las variantes de las secuencias de nucleótidos que codifican CTP o las secuencias de aminoácidos de los CTP descritas en la presente memoria. Mediante "variante" se quiere significar una secuencia suficientemente idéntica, o una secuencia que difiere en al menos un aminoácido de un péptido de tránsito al

cloroplasto nativo. Las secuencias que codifican CTP incluidas en la presente invención son suficientemente idénticas a la secuencia de nucleótidos de SEQ ID NO: 6. Las secuencias de CTP incluidas en la presente memoria son suficientemente idénticas a la secuencia de aminoácidos de SEQ ID NO: 7. Mediante "suficientemente idéntica" se quiere significar una secuencia de nucleótidos que tiene al menos 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% de identidad de secuencia en comparación con una secuencia de referencia que utiliza uno de los programas de alineamiento como se describe en la presente memoria.

[0035] En un ejemplo, las variantes descritas en la presente memoria incluyen nucleótidos o sustituciones de aminoácidos, deleciones, truncamientos e inserciones de uno o más nucleótidos del SEQ ID NO: 6, o uno o más aminoácidos del SEQ ID NO: 7, incluyendo hasta aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7 sustituciones, deleciones o inserciones de aminoácidos.

[0036] Las variantes de origen natural se pueden identificar mediante el uso de técnicas de biología molecular bien conocidas, tales como reacción en cadena de la polimerasa (PCR) y técnicas de hibridación como se describe a continuación. Las secuencias de nucleótidos variantes también incluyen secuencias de nucleótidos derivadas sintéticamente que se han generado, por ejemplo, mediante el uso de mutagénesis dirigida al sitio pero que todavía tiene la actividad de CTP definida en la presente memoria.

[0037] Las variantes empleadas por la presente invención son biológicamente activas, es decir continúan poseyendo la actividad biológica deseada de la secuencia nativa, esto es, conservando de la actividad de CTP (es decir, facilitando la translocación del polipéptido expresado al cloroplasto). Mediante "conserva la actividad de CTP" se quiere significar que la variante dirigirá la translocación al cloroplasto de al menos aproximadamente 30%, al menos aproximadamente 50%, al menos aproximadamente 70%, o al menos aproximadamente 80% del polipéptido codificado por la secuencia nucleótidos de interés. Los métodos para medir la actividad de CTP son bien conocidos en la técnica. Véase la sección titulada "Evaluación de la actividad de CTP" para los ejemplos de los métodos adecuados.

[0038] El experto en la técnica apreciarán adicionalmente que se pueden introducir cambios en el CTP por medio de mutación en la secuencia de nucleótidos que codifica los CTP sin alterar la capacidad del CTP para conducir la translocación de un polipéptido en el cloroplasto de una célula vegetal. Por lo tanto, se pueden crear moléculas aisladas de ácido nucleico variantes mediante la introducción de una o más sustituciones, adiciones o deleciones de nucleótidos en la secuencia de nucleótidos correspondiente descritas en la presente memoria. Las mutaciones pueden introducirse por medio de técnicas convencionales, tales como mutagénesis dirigida al sitio y mutagénesis mediada por PCR.

[0039] Alternativamente, se pueden elaborar secuencias de nucleótidos variantes mediante la introducción de mutaciones al azar a lo largo de la totalidad o parte de la secuencia de CTP, tal como mediante mutagénesis de saturación, y los mutantes resultantes se pueden escrutar para determinar la capacidad para conducir la translocación de una secuencia de polipéptido conectada operablemente al cloroplasto de una célula vegetal.

[0040] Mediante "conectada operablemente" se quiere significar una conexión funcional entre un elemento regulador (p. ej., un CTP) y una segunda secuencia, donde la secuencia de CTP dirige la translocación del polipéptido de interés al cloroplasto de una célula vegetal. Generalmente, pero no siempre, conectada operablemente significa que las secuencias de ácidos nucleicos que se unen son contiguas y, cuando es necesario unir dos regiones codificantes de proteínas, contiguas y en el mismo marco de lectura.

[0041] Para determinar el porcentaje de identidad de dos ácidos nucleicos, las secuencias se alinean para propósitos de comparación óptima. El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias (es decir, porcentaje de identidad = número de posiciones idénticas/número total de posiciones (p. ej., posiciones solapantes) x 100). En un caso, las dos secuencias tienen la misma longitud. En otro ejemplo, la comparación es a través de la totalidad de la secuencia de referencia (p. ej., SEQ ID NO: 1, 2, 4, o 6). El porcentaje de identidad entre dos secuencias se puede determinar utilizando técnicas similares a las descritas a continuación, con o sin permitir huecos. Al calcular el porcentaje de identidad, se cuentan los emparejamientos típicamente exactos.

[0042] La determinación del porcentaje de identidad entre dos secuencias se puede lograr utilizando un algoritmo matemático. Un ejemplo no limitante de un algoritmo matemático utilizado para la comparación de dos secuencias es el algoritmo de Karlin y Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264, modificado como en Karlin y Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5877. Tal algoritmo se incorpora en el programa BLASTN de Altschul et al. (1990) J. Mol. Biol. 215:403. Las búsquedas de nucleótidos BLAST se pueden realizar con el programa BLASTN, puntuación = 100, longitud de palabra = 12, para obtener secuencias de nucleótidos homólogas a las secuencias de la invención. Para obtener alineaciones con huecos con fines de comparación, se puede utilizar Gapped BLAST como describen Altschul et al. (1997) Nucleic Acids Res. 25:3389. Alternativamente, se puede utilizar PSI-BLAST

para realizar una búsqueda iterada que detecta relaciones distantes entre moléculas. Véase Altschul *et al.* más arriba. Cuando se utilizan los programas BLAST, Gapped BLAST y PSI-Blast, se pueden utilizar los parámetros por defecto de los respectivos programas (p. ej., BLASTN). Véase, www.ncbi.nlm.nih.gov. Otro ejemplo no limitante de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo ClustalW (Higgins et al. (1994) Nucleic Acids Res. 22:4673-4680). ClustalW compara secuencias y alinea la totalidad de la secuencia de ADN, y por lo tanto puede proporcionar datos sobre la conservación de la secuencia de la secuencia de nucleótidos completa. El algoritmo ClustalW se utiliza en varios paquetes de soporte lógico de análisis de ADN disponibles en el mercado, tales como el módulo ALIGNX del programa Vector NTI Suite (Informax, Inc). Un ejemplo no limitante de un programa de soporte lógico útil para el análisis de alineamientos ClustalW es GeneDoc™. Genedoc™ (Karl Nicholas) permite la evaluación de la similitud del ADN y la identidad entre múltiples genes. Otro ejemplo no limitante preferido de un algoritmo matemático utilizado para la comparación de secuencias es el algoritmo de Myers y Miller (1988) CABIOS 4:11-17. Tal algoritmo se incorpora en el programa ALIGN (versión 2.0), que es parte del paquete de soporte lógico de alineamiento de secuencia GCG (disponible en Accelrys, Inc., 9865 Scranton Rd., San Diego, California, USA).

[0043] A menos que se indique lo contrario, se utilizará GAP Versión 10, que utiliza el algoritmo de Needleman y Wunsch (1970) J. Mol. Biol. 48(3):443-453, para determinar la identidad o similitud de secuencia con los siguientes parámetros: % de identidad y % de similitud para una secuencia de nucleótidos utilizando un Peso del HUECO de 50 y un Peso de Longitud de 3, y la matriz de puntuación *nwsgapdna.cmp*; % de identidad o % de similitud para una secuencia de aminoácidos utilizando un peso del HUECO de 8 y un peso de longitud de 2, y el programa de puntuación BLOSUM62. También se pueden utilizar programas equivalentes. Por "programa equivalente" se quiere significar cualquier programa de comparación de secuencia que, para dos secuencias cualesquiera en cuestión, genera un alineamiento que tiene emparejamientos de residuos de nucleótidos idénticos y un porcentaje de identidad de secuencia idéntico cuando se compara con el alineamiento correspondiente generado por GAP Version 10.

[0044] Utilizan do métodos tales como PCR, hibridación, y similares, se pueden identificar secuencias correspondientes de otros organismos, en particular de otros organismos de algas, tales secuencias que tienen una identidad sustancial con las secuencias de la invención. Véanse, por ejemplo, Sambrook J., y Russell, DW (2001) Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York) e Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, Nueva York).

[0045] En otro ejemplo, los CTP se pueden identificar basándose en la identificación de secuencias conocidas por comprender los CTP. Por ejemplo, las secuencias dirigidas al cloroplasto se pueden identificar basándose en la similitud de las proteínas dirigidas al cloroplasto descritas en la presente memoria o conocidas en la técnica (p. ej., acetolactato sintasa (AHAS), subunidad pequeña (SSU), y EPSPS). La secuencia de CTP de estas proteínas dirigidas se puede identificarse utilizando métodos conocidos en la técnica. Véase, por ejemplo, Emanuelsson y von Heijne (2001) Biochimica et Biophysica Acta 1541:114-119; Nielson et al. (1997) Protein Eng..10:1-6; y, Nielson y Krogh (1998) Intell. Syst. Mol. Biol. 6:122-130. También están disponibles para la identificación una variedad de programas de ordenador. Véanse, por ejemplo, ChloroP (que se puede encontrar en la dirección de internet www.cbs.dtu.dk/services/ChloroP/); Predotar (que se puede encontrar en la dirección de internet www.inra.fr/Internet/Produits/Predotar/); y, SignalP (que se puede encontrar en la dirección de internet www.cbs.dtu.dk/services/SignalP/).

[0046] Los cebadores oligonucleotídicos se pueden diseñar para uso en reacciones de PCR para amplificar secuencias de ADN correspondientes a partir de ADNc o ADN genómico de un organismo de interés. Los métodos para diseñar cebadores de PCR y clonación por PCR son generalmente conocidos en la técnica y se describen en Sambrook y col. (1989) Molecular Cloning: A Laboratory Manual (2ª ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York). Véase también Innis et al., Eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, Nueva York); Innis y Gelfand, eds. (1995) PCR Strategies (Academic Press, Nueva York); e Innis y Gelfand, eds. (1999) PCR Methods Manual (Academic Press, Nueva York). Los métodos conocidos de PCR incluyen, pero no se limitan a, los métodos que utilizan cebadores emparejados, cebadores anidados, cebadores específicos individuales, cebadores degenerados, cebadores específicos de genes, cebadores específicos del vector y cebadores con emparejamientos erróneos parciales.

[0047] En un método de hibridación, la totalidad o parte de una secuencia de nucleótidos conocida se puede utilizar para escrutar bibliotecas de ADNc o genómicas. Los métodos para la constructo de tales genotecas de ADNc y genómicas son generalmente conocidos en la técnica y se describen en Sambrook y Russell, 2001, más arriba. Las sondas de hibridación pueden ser fragmentos genómicos de ADN, fragmentos de ADNc, fragmentos de ARN, u otros oligonucleótidos, y pueden marcarse con un grupo detectable tal como ³²P, o cualquier otro marcador detectable, tal como otros radioisótopos, un compuesto fluorescente, una enzima, o un cofactor enzimático. Las sondas para la hibridación se pueden elaborar marcando oligonucleótidos sintéticos basándose en la secuencia que codifica CTP conocida descrita en la presente memoria o cebadores para la proteína dirigida al cloroplasto conocida. Se pueden utilizar adicionalmente cebadores degenerados diseñados sobre la base de nucleótidos conservados en la

secuencia de nucleótidos. La sonda comprende típicamente una región de secuencia de nucleótidos que hibrida en condiciones rigurosas con al menos aproximadamente 12, al menos aproximadamente 20, al menos aproximadamente 25, 30, 35, 40, 45, 50, 55, 60, 75, 100, 125, 150, 175, 200, 250, 300, 350, o 400 nucleótidos consecutivos de la secuencia que codifica el CTP de la invención, una secuencia de nucleótidos que codifica una proteína dirigida al cloroplasto, o un fragmento o variante de la misma. La preparación de sondas de hibridación es generalmente conocida en la técnica y se describe en Sambrook y Russell, 2001, más arriba.

[0048] Por ejemplo, la secuencia que codifica CTP completa descrita en la presente memoria (o una secuencia codificadora para la proteína dirigida al cloroplasto), o una o más porciones de la misma, se puede utilizar como una sonda capaz de hibridar específicamente a las secuencias de tipo CTP correspondientes. Para conseguir una hibridación específica bajo una variedad de condiciones, tales sondas incluyen secuencias que son únicas y tienen al menos aproximadamente 10 nucleótidos de longitud, o al menos aproximadamente 20 nucleótidos de longitud. Tales sondas se pueden utilizar para amplificar secuencias que codifican los CTP correspondientes a partir de un organismo seleccionado mediante PCR. Esta técnica se puede utilizar para aislar secuencias codificantes adicionales de un organismo deseado o como un ensayo de diagnóstico para determinar la presencia de secuencias codificantes en un organismo. Las técnicas de hibridación incluyen escrutinio por hibridación de bibliotecas de ADN cultivados en placa (ya sea placas o colonias; véase, por ejemplo, Sambrook y col. (1989) *Molecular Cloning: A Laboratory Manual* (2^a ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York).

[0049] La hibridación de tales secuencias se puede llevar a cabo en condiciones rigurosas. Mediante "condiciones rigurosas" o "condiciones de hibridación rigurosas" se quieren significar condiciones bajo las cuales una sonda hibridará con su secuencia diana en un grado detectablemente mayor que con otras secuencias (por ejemplo, al menos 2 veces más sobre el fondo). Las condiciones rigurosas dependen de la secuencia y serán diferentes en diferentes circunstancias. Al controlar la rigurosidad de la hibridación y/o las condiciones de lavado, se pueden identificar las secuencias diana que son 100% complementarias a la sonda (sondeo homólogo). Alternativamente, las condiciones de rigurosidad se pueden ajustar para permitir algunos emparejamientos erróneos en las secuencias de manera que se detectan menores grados de similitud (sondeo heterólogo). Generalmente, una sonda tiene menos de aproximadamente 1000 nucleótidos de longitud, o menos de 500 nucleótidos de longitud.

[0050] Típicamente, las condiciones rigurosas serán aquellas en las que la concentración de sal es menor de aproximadamente 1,5 M de ión de Na, típicamente de aproximadamente 0,01 a 1,0 M de concentración de ión de Na (u otras sales) a pH de 7,0 a 8,3 y la temperatura es de al menos aproximadamente 30°C para sondas cortas (p. ej., de 10 a 50 nucleótidos) y al menos aproximadamente 60°C para sondas largas (p. ej., mayor que 50 nucleótidos). Las condiciones rigurosas también se pueden lograr con la adición de agentes desestabilizantes tales como formamida. Las condiciones de baja rigurosidad ilustrativas incluyen hibridación con una solución tampón de formamida de 30 a 35%, NaCl 1 M, SDS (dodecil sulfato de sodio) al 1% a 37°C, y un lavado en 1X a 2X SSC (20X SSC = NaCl 3,0 M/citrato trisódico 0,3 M) de 50 a 55°C. Los ejemplos de condiciones de rigurosidad moderada incluyen hibridación en formamida de 40 a 45%, NaCl 1,0 M, SDS al 1% a 37°C, y un lavado en 0,5 X a 1X SSC de 55 a 60°C. Las condiciones de alta rigurosidad ilustrativas incluyen hibridación en formamida al 50%, NaCl 1 M, SDS al 1% a 37°C, y un lavado en 0,1 X SSC de 60 a 65°C. Opcionalmente, los tampones de lavado pueden comprender SDS de aproximadamente 0,1% a aproximadamente 1%. La duración de la hibridación es generalmente menor de aproximadamente 24 horas, normalmente de aproximadamente 4 a aproximadamente 12 horas. Opcionalmente, los tampones de lavado pueden comprender SDS de aproximadamente 0,1% a aproximadamente 1%.

[0051] La especificidad es típicamente la función de los lavados de post-hibridación, siendo los factores críticos la fuerza iónica y la temperatura de la solución de lavado final. Para los híbridos ADN-ADN, la T_m se puede aproximar de la ecuación de Meinkoth y Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81,5^\circ\text{C} + 16,6 (\log M) + 0,41 (\%GC) - 0,61 (\% \text{form}) - 500/L$; donde M es la molaridad de cationes monovalentes, % GC es el porcentaje de nucleótidos de guanosina y citosina en el ADN, % form es el porcentaje de formamida en la solución de hibridación, y L es la longitud del híbrido en pares de bases. La T_m es la temperatura (bajo una fuerza iónica y pH definidos) a la que 50% de una secuencia diana complementaria hibrida con una sonda perfectamente emparejada. La T_m se reduce en aproximadamente 1°C por cada 1% de emparejamiento erróneo; por lo tanto, la T_m , las condiciones de hibridación, y/o lavado se pueden ajustar para hibridar con secuencias de la identidad deseada. Por ejemplo, si se buscan secuencias con $\geq 90\%$ de identidad, la T_m se puede disminuir 10°C. Generalmente, se seleccionan condiciones rigurosas para que sean aproximadamente 5°C más bajas que el punto de fusión térmico (T_m) para la secuencia específica y su complemento a una fuerza iónica y pH definidos. Sin embargo, las condiciones severamente rigurosas pueden utilizar una hibridación y/o lavado 1, 2, 3, o 4°C inferiores que el punto de fusión térmico (T_m); las condiciones moderadamente rigurosas pueden utilizar una hibridación y/o lavado 6, 7, 8, 9, o 10°C inferiores que el punto de fusión térmico (T_m); las condiciones de baja rigurosidad pueden utilizar una hibridación y/o lavado 11, 12, 13, 14, 15, o 20°C inferiores que el punto de fusión térmico (T_m). Utilizando la ecuación, la hibridación y las composiciones lavado, y T_m deseada, los expertos normales entenderán que las variaciones en la rigurosidad de la hibridación y/o las soluciones de lavado se describen inherentemente. Si el grado deseado de emparejamientos erróneos da como resultado una T_m de menos de 45°C (solución acuosa) o 32°C (solución de formamida), la concentración de SSC se puede aumentar de manera que se puede utilizar una temperatura más alta. Una guía

extensa para la hibridación de ácidos nucleicos se encuentra en Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, Parte I, Capítulo 2 (Elsevier, Nueva York); y Ausubel y col., Eds. (1995) *Current Protocols in Molecular Biology*, Capítulo 2 (Greene Publishing y Wiley-Interscience, Nueva York). Véase Sambrook y col. (1989) *Molecular Cloning: A Laboratory Manual* (2^a ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York).

[0052] También se conocen las secuencias aisladas que tienen actividad de CTP y que hibridan en condiciones rigurosas con las secuencias de CTP descritas en esta memoria, o a fragmentos de las mismas.

10 *Métodos de uso*

[0053] Los métodos de la presente invención se refieren a la expresión, translocación, y procesamiento adecuados de secuencias dirigidas al cloroplasto en plantas superiores y células vegetales bajo el control de las secuencias de CTP. Por lo tanto, la presente invención proporciona un método para expresar una secuencia de nucleótidos de interés en una planta monocotiledónea, comprendiendo dicho método:

- a) introducir en una célula vegetal un casete de expresión que comprende una secuencia de nucleótidos que codifica un péptido dirigido al cloroplasto (CTP), en donde dicha secuencia de nucleótidos que codifica dicho CTP está conectada operablemente a dicha secuencia de nucleótidos de interés, y en donde dicha secuencia de nucleótidos que codifica dicho CTP se selecciona del grupo que consiste en
 - i) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6;
 - ii) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto;
 - iii) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y,
 - iv) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto; y,
- b) regenerar una planta transformada a partir de dicha célula vegetal;

en donde dicha planta ha incorporado de forma estable en su genoma dicho casete de expresión.

[0054] Para los propósitos de la presente invención, una proteína dirigida al cloroplasto "procesada" es una en la que se ha eliminado el CTP. En el momento de la translocación de una proteína dirigida al cloroplasto al cloroplasto de una célula vegetal, el CTP se elimina de la proteína específica mediante escisión en un "sitio de escisión" concreto entre el CTP y la proteína madura. El sitio de escisión se puede determinar experimentalmente, o se puede pronosticar basándose en la estructura de la secuencia (p. ej., mediante el alineamiento de la proteína no procesada con proteínas dirigidas al cloroplastos en las que se conoce el sitio de escisión, mediante el análisis de la secuencia para determinar la presencia de dominios CTP característicos, y similares) o mediante el uso de uno o más algoritmos para la predicción de sitio de escisión comentado en otra parte en la presente memoria (por ejemplo, SignalP).

[0055] Las plantas transgénicas pueden tener un cambio en el fenotipo, incluyendo, pero no limitado a, un mecanismo de defensa de patógenos o insectos alterado, un aumento de la resistencia a uno o más herbicidas, una mayor capacidad para resistir las condiciones ambientales de estrés, una capacidad modificada para producir almidón, un nivel modificado de la producción de almidón, un contenido y/o composición de aceite modificado, una capacidad modificada para utilizar, repartir y/o almacenar nitrógeno, y similares. Estos resultados se pueden lograr a través de la expresión y el direccionamiento de un polipéptido de interés a los cloroplastos en las plantas, en donde el polipéptido de interés funciona en el cloroplasto. Las secuencias de CTP son útiles para dirigir secuencias nativas, así como secuencias heterólogas (no nativas) en plantas superiores. Para los propósitos de la presente invención, "plantas superiores" se consideran miembros del subreino Embryophytae. La planta es una monocotiledónea.

[0056] En general, la secuencia de nucleótidos que codifica el CTP se proporciona en un casete de expresión con una secuencia de nucleótidos de interés para la expresión en la planta de interés. En una realización, las secuencias que codifican CTP son útiles para la mejora de la translocación de secuencias nativas en una planta. En otras realizaciones, las secuencias que codifican CTP son útiles para la expresión y translocación de polipéptidos codificados por secuencias de nucleótidos heterólogas. Mediante "secuencia de nucleótidos heteróloga" se quiere significar una secuencia que no está conectada operablemente de forma natural a la secuencia que codifica CTP de la invención, incluyendo múltiples copias de origen no natural de una secuencia de ADN de origen natural. Si bien esta secuencia de nucleótidos es heteróloga a la secuencia que codifica CTP, puede ser homóloga, o "nativa", o heteróloga, o "foránea", a la planta huésped. En algunos casos, la planta transformada puede tener un cambio en el fenotipo. "Heteróloga" generalmente se refiere a las secuencias de ácido nucleico que no son endógenas a la célula o parte del genoma nativo en el que están presentes, y se han añadido a la célula por medio de infección, transfección, microinyección, electroporación, microproyección, o similar .

[0057] Se puede utilizar cualquier secuencia de nucleótidos de interés con las secuencias que codifican CTP, siempre y cuando el polipéptido codificado por la secuencia de nucleótidos de interés (es decir, el "polipéptido de interés") sea funcional en un cloroplasto. Dichas secuencias de nucleótidos incluyen, pero no se limitan a, secuencias que codifican tolerancia a herbicidas, secuencias que codifican insecticidas, secuencias que codifican nematocidas, secuencias que codifican antimicrobianos, secuencias que codifican antifúngicos, secuencias que codifican antivirales, secuencias que codifican tolerancia al estrés abiótico y biótico, o secuencias que modifican características de las plantas tales como el rendimiento, la calidad del grano, el contenido de nutrientes, la calidad y la cantidad de almidón, la fijación y/o la utilización de nitrógeno, y el contenido y/o la composición de aceite. Más genes específicos de interés para la presente invención incluyen, pero no se limitan a, genes que mejoran el rendimiento del cultivo, genes que mejoran la conveniencia de los cultivos, genes que codifican proteínas que confieren resistencia a estrés abiótico, tal como sequía, temperatura, salinidad, metales tóxicos u oligoelementos, o aquellos que confieren resistencia a toxinas tales como plaguicidas y herbicidas, o a estrés biótico, tal como ataques por hongos, virus, bacterias, insectos y nematodos, y al desarrollo de enfermedades asociadas con estos organismos. Se reconoce que cualquier gen de interés puede conectarse operablemente a las secuencias que codifican CTP y se expresa en una planta, siempre y cuando el polipéptido codificado por el gen sea funcional en cloroplastos.

Estas secuencias de nucleótidos de interés pueden codificar proteínas implicadas en la comunicación de resistencia a enfermedades o a plagas. Mediante "resistencia a enfermedades" o "resistencia a las plagas" se quiere significar que las plantas evitan los síntomas nocivos que son el resultado de las interacciones planta-patógeno. Los genes de resistencia a enfermedades y de resistencia de insectos tales como las lisozimas o las cecropinas para la protección antibacteriana, o proteínas tales como las defensinas, las glucanasas o las quitinasas para la protección antifúngica, o las endotoxinas de *Bacillus thuringiensis*, los inhibidores de proteasas, las colagenasas, las lectinas, o las glicosidasas para el control de nematodos o insectos son todos los ejemplos de productos génicos útiles. Los ejemplos de los genes de interés se pueden encontrar, por ejemplo, en www.nbiap.vt.edu/cfdocs/fieldtests2.cfm.

[0058] "Plaga" incluye, pero se limita a, insectos, hongos, bacterias, virus, nematodos, ácaros, garrapatas, y similares. Las plagas de insectos incluyen insectos seleccionados de los órdenes *Coleoptera*, *Diptera*, *Hymenoptera*, *Lepidoptera*, *Mallophaga*, *Homoptera*, *Hemiptera*, *Orthoptera*, *Thysanoptera*, *Dermaptera*, *Isoptera*, *Anoplura*, *Siphonaptera*, *Trichoptera*, etc., particularmente *Coleoptera*, *Lepidoptera*, y *Diptera*. Los virus incluyen, pero no se limitan a virus del mosaico del tabaco o del pepino, el virus de la mancha anular, el virus de la necrosis, el virus del mosaico enano del maíz, etc. Los nematodos incluyen, pero no se limitan a nematodos parásitos tales como nematodos del nudo de la raíz, del quiste, y de las lesiones, incluyendo *Heterodera* spp., *Meloidogyne* spp., y *Globodera* spp.; particularmente miembros de los nematodos del quiste, incluyendo, pero no limitados a, *Heterodera glycines* (Nematodo del quiste de soja); *Heterodera Schachtii* (Nematodo del quiste de la remolacha); *Heterodera avenae* (nematodo del quiste de los cereales); y *Globodera rostochiensis* y *Globodera pallida* (nematodos del quiste de la patata). Las lesiones por Nematodos incluyen, pero no se limitan a *Pratylenchus* spp. Las plagas fúngicas incluyen las que causan las royas de las hojas, amarilla, estriada y del tallo.

[0059] Una "proteína de resistencia a herbicidas" o una proteína resultante de la expresión de una "molécula de ácido nucleico que codifica resistencia a herbicidas" incluye proteínas que confieren a una célula la capacidad para tolerar una mayor concentración de un herbicida que las células que no expresan la proteína, o para tolerar una cierta concentración de un herbicida durante un período de tiempo más largo que las células que no expresan la proteína. Los rasgos de resistencia a herbicidas se pueden introducir en las plantas mediante los genes que codifican la resistencia a herbicidas que actúan inhibiendo la acción de la acetolactato sintasa (ALS), en particular los herbicidas de tipo sulfonilurea, los genes que codifican para resistencia a herbicidas que actúan inhibiendo la acción de la glutamina sintasa, tales como fosfinotricina o Basta (por ejemplo, el gen *bar*), glifosato (por ejemplo, el gen de la EPSP sintasa y el gen GAT) u otros tales genes conocidos en la técnica.

[0060] Los genes que mejoran el rendimiento del cultivo incluyen genes de enanismo, como Rht1 y Rht2 (Peng et al. (1999) Nature 400:256-261), y los que incrementan en crecimiento vegetal, tales como la glutamato deshidrogenasa inducida por amonio. Los genes que mejoran la conveniencia de los cultivos incluyen, por ejemplo, los que permiten que las plantas tengan una reducción del contenido de grasas saturadas, los que refuerzan el valor nutricional de las plantas, y aquellos que aumentan la proteína de los cereales. Los genes que mejoran la tolerancia a la sal son aquellos que aumentan o permiten el crecimiento de la planta en un entorno de mayor salinidad que el entorno nativo de la planta en el que se ha introducido o se han introducido el gen o los genes tolerantes a las sales.

Vectores de transformación de plantas

[0061] Típicamente el casete de expresión de plantas se insertará en un "vector de transformación de plantas". Mediante "vector de transformación" se quiere significar una molécula de ADN que es necesaria para la transformación eficaz de una células. Tal molécula puede consistir en uno o más casetes de expresión, y se puede organizar en más de una molécula de ADN "vectora". Por ejemplo, los vectores binarios son vectores de transformación de plantas que utilizan dos vectores de ADN no contiguos para codificar todas las funciones que actúan en cis y en trans requisito para la transformación de células vegetales (Hellens and Mullineaux (2000) Trends

in Plant Science 5:446-451). "Vector" se refiere a un constructo de ácido nucleico diseñado para la transferencia entre diferentes células anfitrionas. "Vector de expresión" se refiere a un vector que tiene la capacidad de incorporar, integrar y expresar secuencias o fragmentos de ADN heterólogo en una célula foránea. Mediante "introduciendo" se quiere significar presentar al organismo que está siendo transformado el constructo nucleotídico de manera que el constructo obtiene acceso al interior de al menos una célula del organismo.

[0062] Este vector de transformación de plantas puede estar compuesto de uno o más vectores con ADN necesarios para lograr la transformación de la planta. Por ejemplo, es una práctica habitual en la técnica utilizar vectores de transformación de plantas que están compuestos de más de un segmento de ADN contiguo. Estos vectores son referidos a menudo en la técnica como 'vectores binarios'. Los vectores binarios así como los vectores con plásmidos coadyuvantes se utilizan muy a menudo para la transformación mediada por *Agrobacterium*, donde el tamaño y la complejidad de los segmentos de ADN necesarios para lograr una transformación eficaz es bastante grande, y es ventajoso separar las funciones sobre moléculas de ADN separadas. Los vectores binarios contienen típicamente un vector plasmídico que contiene las secuencias que actúan en cis requeridas para la transferencia de ADN-T (tal como el borde derecho y el borde izquierdo), un marcador seleccionable que se diseña para que sea susceptible de expresión en una célula vegetal, y un "gen de interés" (un gen diseñado para que sea susceptible de expresión en una célula vegetal para la que se desea la generación de plantas transgénicas). También están presentes en este vector plasmídico las secuencias requeridas para la replicación bacteriana.

[0063] Las secuencias que actúan en cis se disponen de manera que permiten la transferencia eficaz a células vegetales y su expresión en las mismas. Por ejemplo, el gen marcador seleccionable y el gen de interés están localizados entre los bordes izquierdo y derecho. Con frecuencia un segundo vector plasmídico contiene los factores que actúan en trans que median la transferencia de ADN-T de *Agrobacterium* a células vegetales. Este plásmido contiene con frecuencia las funciones de virulencia (genes Vir) que permiten la infección de las células vegetales por *Agrobacterium*, y la transferencia de ADN mediante la escisión en las secuencias de los bordes y la transferencia de ADN mediada por vir, como se entiende en la técnica (Hellens and Mullineaux (2000) Trends in Plant Science, 5:446-451). Se pueden utilizar varios tipos de cepas de *Agrobacterium* (p. ej. LBA4404, GV3101, EHA101, EHA105, etc.) para la transformación de plantas. El segundo vector plasmídico no es necesariamente para la transformación de las plantas mediante otros métodos tales como microproyección, microinyección, electroporación, polietilenglicol, etc.

Transformación de plantas

[0064] Los métodos de la invención implican introducir un constructo nucleotídico en una planta. Mediante "introducir" se quiere significar la presentación a la planta del constructo nucleotídico de tal manera que el constructo obtenga acceso al interior de una célula de la planta. Los métodos de la invención no requieren el uso de un método concreto de introducción de un constructo nucleotídico en una planta, únicamente que el constructo nucleotídico obtenga acceso al interior de al menos una célula de la planta. Los métodos para introducir constructos nucleotídicos en plantas son conocidos en la técnica incluyendo, pero no limitados a, métodos de transformación estable, métodos de transformación transitoria, y métodos mediados por virus.

[0065] La transformación de células vegetales se puede completar por medio de uno de varios mecanismos conocidos en la técnica. Mediante "planta" que quieren significar plantas completas, órganos de plantas (p. ej., hojas, tallos, raíces, etc.), semillas, células vegetales, propágulos, embriones y su progenie. Las células vegetales pueden ser diferenciadas o no diferenciadas (p. ej. callos, células de cultivo en suspensión, protoplastos, células de la hoja, células de la raíz, células del floema, polen). "Plantas transgénicas" o "plantas transformadas" o plantas o células o tejidos "transformados establemente" se refieren a plantas que han incorporado o integrado secuencias de ácido nucleico exógeno o fragmentos de ADN a la célula vegetal. Por "transformación estable" se quiere significar que el constructo nucleotídico introducido en una planta se integra en el genoma de la planta y es susceptible de ser heredado por su progenie.

[0066] In general, los métodos de transformación de plantas implican la transferencia de ADN heterólogo a células vegetales diana (p. ej. embriones inmaduros o maduros, cultivos en suspensión, callos no diferenciados, protoplastos, etc.), aplicando a continuación un nivel umbral máximo de la selección apropiada (dependiendo del gen marcado seleccionable) para recuperar las células vegetales transformadas de un grupo de masas de células no transformadas. Los explantes se transfieren típicamente a un suministro de nueva aportación del mismo medio y se cultivan rutinariamente. Con posterioridad, las células transformadas se diferencian a brotes después de colocarlas en un medio de regeneración con un suplemento de nivel umbral máximo de agente de selección. Los brotes se transfieren a continuación a medio de enraizamiento selectivo para la recuperación del brote de raíz o plántula. La plántula transgénica crece a continuación a una planta madura y produce semillas fértiles (p. ej. Hiei et al. (1994) The Plant Journal 6:271-282; Ishida et al. (1996) Nature Biotechnology 14:745-750). Una descripción general de las técnicas y los métodos para generar plantas transgénicas se encuentra en Ayres y Park (1994) Critical Reviews in Plant Science 13:219-239 and Bommineni and Jauhar (1997) Maydica 42:107-120. Puesto que el material transformado contiene muchas células; las células transformadas y no transformadas están presentes en cualquier parte del callo o tejido o grupo de células diana sujeto. La capacidad para destruir las células no transformadas y

5 permitir que las células transformadas proliferen da como resultado cultivos de plantas transformados. A menudo, la capacidad para eliminar células no transformadas es un limitación para la recuperación rápida de las células vegetales transformadas y la generación satisfactoria de plantas transgénicas. Se pueden utilizar métodos moleculares y bioquímicos para confirmar la presencia del gen heterólogo de interés integrado en el genoma de la planta transgénica.

10 **[0067]** La generación de plantas transgénicas se puede llevar a cabo por medio de uno de varios métodos, incluyendo, pero no limitados a, introducción de ADN heterólogo mediante *Agrobacterium* en células vegetales (transformación mediada por *Agrobacterium*), bombardeo de células vegetales con ADN foráneo heterólogo adherido a partículas, y otros diversos métodos no mediados por partículas dirigidas (p. ej. Hiei et al. (1994) *The Plant Journal* 6:271-282; Ishida et al. (1996) *Nature Biotechnology* 14:745-750; Ayres and Park (1994) *Critical Reviews in Plant Science* 13:219-239; Bommineni and Jauhar (1997) *Maydica* 42:107-120) para transferir ADN.

15 **[0068]** Los métodos para la transformación de cloroplastos son conocidos en la técnica. Véanse, por ejemplo, Svab et al. (1990) *Pro. Natl. Acad. Sci. USA* 87:8526-8530; Svab y Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab y Maliga (1993) *EMBO J.* 12:601-606. El método se basa en la liberación con pistola de partículas de ADN que contiene un marcador seleccionable y el direccionamiento del ADN al genoma del plástido a través de recombinación homóloga. Adicionalmente, la transformación del plástido se puede completar por medio de transactivation de un transgen silencioso que porta un plástido mediante expresión preferida del tejido de una ARN polimerasa codificada en el núcleo y dirigida al plástido. Tal sistema ha sido referendo por McBride et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:7301-7305.

20 **[0069]** Las células que han sido transformadas se pueden desarrollar en las plantas de acuerdo con métodos convencionales. Véase, por ejemplo, McCormick et al. (1986) *Plant Cell Reports* 5:81-84. Estas plantas se pueden cultivar a continuación, y polinizar con la misma cepa transformada o con cepas diferentes, e identificar el híbrido resultante que tiene la expresión constitutiva de la característica fenotípica deseada. Se pueden desarrollar dos o más generaciones para asegurar que se mantiene y se hereda establemente la expresión de la característica fenotípica deseada y a continuación se cosechan las semillas para asegurar que se la logrado la expresión de la característica fenotípica deseada. De esta manera, la presente invención proporciona una semilla transformada (también referida como "semilla transgénica") que tiene un constructo nucleotídico de la invención, por ejemplo, un casete de expresión de la invención, incorporado establemente a su genoma.

Plantas

35 **[0070]** La presente invención se puede utilizar para la transformación de cualquier especie de planta superior, incluyendo, pero no limitada a, monocotiledoneas y dicotiledoneas. En una realización, el CTP incluido en la presente memoria es activo tanto en monocotiledoneas como en dicotiledoneas. En otra realización, el CTP es activo únicamente en monocotiledoneas o únicamente en dicotiledoneas. Los ejemplos de las plantas de interés incluyen, pero no están limitadas a, maíz (maíz), sorgo, trigo, girasol, tomate, crucíferas, pimientos, patata, algodón, arroz, soja, remolacha azucarera, caña de azúcar, tabaco, cebada, y colza, *Brassica* sp., alfalfa, centeno, mijo, cártamo, cacahuetes, batata, mandioca, café, coco, piña, cítricos, cacao, te, banana, aguacate, higo, guayaba, mango, olivo, papaya, anacardo, macadamia, almendra, avenas, hortalizas, plantas ornamentales, y coníferas.

40 **[0071]** Las hortalizas incluyen, pero no están limitadas a, tomates, lechuga, judías verdes, judías de lima, guisantes, y miembros del género *Curcumis* tales como el pepino, el melón galia y el melón de almizcle. Las plantas ornamentales incluyen, pero no están limitadas a, azalea, hortensia, hibisco, rosas, tulipanes, narcisos, petunias, claveles, flor de pasqua, y el crisantemo. Preferiblemente, las plantas de la presente invención son plantas de cultivo (por ejemplo, maíz, sorgo, trigo, arroz, caña de azúcar, cebada, etc.).

45 **[0072]** Esta invención es adecuada para cualquier miembro de la familia de las plantas monocotiledóneas, incluyendo, pero no limitadas a, maíz, arroz, cebada, avena, trigo, sorgo, centeno, caña de azúcar, piña, ñame, cebolla, plátano, coco y dátiles.

Evaluación de la Transformación de la planta

50 **[0073]** Después de la introducción de ADN foráneo heterólogo en células vegetales monocotiledóneas, la transformación o integración de ADN heterólogo en el genoma de la planta se confirma por medio de diversos métodos tales como el análisis de ácidos nucleicos o proteínas y metabolitos asociados con el ADN integrado.

55 **[0074]** El análisis mediante PCR es un método rápido para detectar células, tejido o brotes transformados para determinar la presencia de ADN incorporado en la etapa anterior antes de su trasplante al suelo (Sambrook y Russell, 2001. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York). La PCR se lleva a cabo utilizando cebadores oligonucleotídicos específicos para el gen de interés o un fondo de vector de *Agrobacterium*, etc.

[0075] La transformación de plantas se puede confirmar por análisis de transferencia Southern de ADN genómico (Sambrook y Russell, 2001, más arriba). En general, el ADN total se extrae del transformante, se digiere con enzimas de restricción apropiadas, se fracciona en un gel de agarosa y se transfiere a una membrana de nitrocelulosa o de nailon. La membrana o "transferencia" se sondea a continuación con, por ejemplo, un fragmento de ADN diana radiomarcado con ³²P para confirmar la integración del ADN introducido en el genoma de la planta de acuerdo con técnicas estándar (Sambrook y Russell, 2001, más arriba).

[0076] En el análisis de transferencia de Northern, se aísla ARN a partir de tejidos específicos de transformante, se fracciona en un gel de agarosa-formaldehído, se transfiere sobre un filtro de nailon de acuerdo con procedimientos convencionales que se utilizan rutinariamente en la técnica (Sambrook y Russell, 2001, más arriba). La expresión del ARN codificado por un gen heterólogo conectado operablemente a la secuencia que codifica CTP se somete a ensayo a continuación hibridando el filtro con una sonda radioactiva derivada del gen heterólogo, por medio de métodos conocidos en la técnica (Sambrook y Russell, 2001, más arriba).

15 *Evaluación de la Actividad CTP*

[0077] Se conocen análisis para determinar la eficacia con la cual las secuencias de CTP dirigen una proteína de interés a un cloroplasto. Véase, por ejemplo, Mishkind et al. (1985) J of Cell Biol. 100:226-234. Un gen informador tal como glucuronidasa (GUS), cloranfenicol acetil transferasa (CAT), o proteína fluorescente verde (GFP) está conectada operablemente a la secuencia de CTP. Esta fusión se coloca detrás del control de un promotor adecuado, se liga a un vector de transformación, y se transforma en una planta o célula vegetal. Después de un período de tiempo adecuado para la expresión y localización en el cloroplasto, la fracción del cloroplasto se extrae y se somete a ensayo la actividad informadora. La capacidad de las secuencias aisladas para dirigir y liberar la proteína informadora al cloroplasto se puede comparar con otras secuencias conocidas de CTP. Véase de Castro Silva Filho et al. (1996) Plant Mol. Biol. 30: 769-780. La importación de proteínas también se puede *verificar in vitro* a través de la adición de proteasas a la fracción de cloroplastos aislada. Las proteínas que se importaron con éxito al cloroplasto son resistentes a las proteasas añadidas externamente mientras que las proteínas que permanecen en el citosol son susceptibles de digestión. La importación de proteínas también se puede verificar por la presencia de proteína funcional en el cloroplasto utilizando técnicas moleculares convencionales para la detección, o evaluando el fenotipo resultante de la expresión de una proteína dirigida al cloroplasto.

[0078] Los siguientes ejemplos se ofrecen a modo de ilustración y no a modo de limitación.

35 **SECCIÓN EXPERIMENTAL**

Ejemplo 1. CTP de AHAS de *Chlamydomonas*.

[0079] El CTP de AHAS de *Chlamydomonas* (SEC ID NO: 3) se identificó a partir de la comparación del péptido AHAS de *Chlamydomonas* ("ChlamyAHAS") (Acceso GENBANK AF022816; SEC ID NO: 8) con varios péptidos AHAS de plantas, bacterias, levaduras y hongos. Este alineamiento muestra que el CTP de ChlamyAHAS muestra muy poca conservación de la secuencia con los CTP de la planta, mientras que las proteínas maduras muestran elementos conservados. Por lo tanto, mediante la comparación de las múltiples proteínas AHAS con la proteína ChlamyAHAS, se dedujo que los aminoácidos 1-99 de la proteína de AHAS de *Chlamydomonas* comprendían el CTP. El sitio de procesamiento previsto para este CTP después de la importación al cloroplasto de *Chlamydomonas* es aproximadamente en el aminoácido 92 del SEQ ID NO: 8.

Ejemplo 2. CTP de la subunidad pequeña de RuBisCo de *Chlamydomonas* (SSU).

[0080] El CTP de SSU de *Chlamydomonas* (SEC ID NO: 5) se identificó a partir de la comparación del péptido SSU completo de *Chlamydomonas* ("ChlamySSU") (Acceso GenBank CAA28160; SEQ ID NO: 9) con péptidos de la subunidad pequeña de Rubisco de varias plantas. Este alineamiento muestra que el CTP de ChlamySSU muestra muy poca conservación de la secuencia con los CTP de la planta de, mientras que las proteínas maduras muestran elementos conservados. Por lo tanto, mediante la comparación de los sitios de procesamiento conocidos de varias proteínas RuBisCo de plantas con la proteína ChlamySSU, se dedujo que los aminoácidos 1-45 de la Proteína SSU de *Chlamydomonas* (SEC ID NO: 9) comprendía el CTP.

[0081] El sitio de escisión para el Precursor de la subunidad pequeña de RuBisCo de *Chlamydomonas* ("ChlamySSU") se ha determinado empíricamente que está entre los residuos 44 y 45 de SEQ ID NO: 9 (Schmidt et al 1979, Journal of Cell Biology 83:615-662).

Ejemplo 3. CTP de EPSPS de *Chlamydomonas*.

[0082] El CTP de EPSPS de *Chlamydomonas* (SEQ ID NO: 7) se identificó a partir de la comparación del péptido EPSPS completo de *Chlamydomonas* (GENBANK XP_001702942; SEQ ID NO: 10) con varias plantas y EPSPS

bacterianos. Este alineamiento muestra que el CTP de ChlamyEPSPS muestra muy poca conservación de secuencia con los CTP de la planta de, mientras que las proteínas maduras muestran elementos conservados. Por lo tanto, mediante la comparación de varias proteínas EPSPS con la proteína ChlamyEPSPS, se dedujo que los aminoácidos 1-75 de la proteína EPSPS de *Chlamydomonas* (SEQ ID NO: 10) comprendía el CTP. El sitio de procesamiento previsto para este CTP después de la importación al cloroplasto de *Chlamydomonas* es aproximadamente en el aminoácido 61 del SEQ ID NO: 10.

Tabla 1. CTP de *Chlamydomonas*

Proteína	SEC ID NO:	Residuos de la Proteína Precursora Utilizados para CTP	Sitio de Escisión Pronosticado
ChlamyAHAS	8	1-99	92
ChlamySSU	9	1-45	Aminoácido 44 o Aminoácido 45
ChlamyEPSPS	10	1-75	61

10 Ejemplo 4. Constructos de ADN para el uso de los CTP de *Chlamydomonas* para el direccionamiento de proteínas

15 **[0083]** Un elemento de ADN que utiliza CTP de *Chlamydomonas*, incluyendo los CTP de AHAS de *Chlamydomonas*, SSU de *Chlamydomonas*, o EPSPS de *Chlamydomonas* para generar múltiples constructos para el direccionamiento de proteínas a los cloroplastos puede implicar la inclusión de sitios de reconocimiento de endonucleasa de restricción convenientes (por ejemplo, un sitio de restricción *Bam*HI), así como pequeños conectores peptídicos entre el CTP del cloroplasto y la proteína (por ejemplo, un CTP tripeptídico Gly-Ser-Gly; SEQ ID NO: 18). Además, tales elementos de ADN se pueden diseñar y fabricar de forma sintética, de manera que la secuencia de ADN varíe del ADN original, pero codifique el péptido idéntico.

20 **[0084]** Alternativamente, se pueden diseñar constructos de ADN de tal manera que no se necesiten sitios de enzimas de restricción, y la fusión de CTP/proteína se pueda lograr mediante síntesis total de la región codificante combinada, o mediante estrategias basadas en PCR, incluyendo "costura de PCR" y similares.

25 **[0085]** También se puede diseñar la fusión de CTP/proteína de una manera en la que alguna de cualquiera de las proteínas esté truncada. Por ejemplo, se pueden eliminar uno o más aminoácidos del extremo N-terminal de una proteína expresada en bacterias, y aún así lograr una fusión funcional a un CTP de *Chlamydomonas*.

30 **[0086]** Se diseñó un casete que contenía una secuencia de ADN diseñada sintéticamente que codifica el CTP de AHAS de *Chlamydomonas* que incorpora un sitio de restricción *Bam*HI y un conector Gly-Ser-Gly. Estos constructos de ADN contienen (de 5' a 3') (1) un sitio de clonación de *Pst* I, (2) las bases ACC para proporcionar el contexto "Kozak" para una traducción eficaz, (3) la porción del gen que codifica la metionina amino terminal a través del sitio de escisión del péptido de tránsito conocido de ChlamyAHAS, y que incluye una región de ADN pequeña que codifica los aminoácidos C-terminales con respecto al sitio de escisión, (4) bases de ADN que codifica los residuos Gly - Ser - Gly con un sitio de clonación *Bam*HI incorporado, y, (5) la región codificante del gen de interés.

35 **[0087]** Se diseñó un casete que contenía una secuencia de ADN diseñada sintéticamente que codifica el CTP de SSU de *Chlamydomonas* que incorpora un sitio de restricción *Bam*HI y un conector Gly-Ser-Gly. Estos constructos de ADN contienen (de 5' a 3') (1) Un sitio de clonación de *PST* I, (2) las bases ACC para proporcionar el contexto "Kozak" para una traducción eficaz, (3) la porción del gen que codifica la metionina amino terminal a través del sitio de escisión de péptido de tránsito conocido de ChlamySSU, y que incluye una región de ADN pequeña que codifica los aminoácidos C-terminales con respecto al sitio de escisión, (4) bases de ADN que codifica los residuos Gly - Ser - Gly con un sitio de clonación *Bam*HI incorporado, y, (5) la región codificante del gen de interés.

40 **[0088]** Se diseñó un casete que contenía una secuencia de ADN diseñada sintéticamente que codifica el CTP de EPSPS de *Chlamydomonas* que incorpora un sitio de restricción *Bam*HI y un conector Gly-Ser-Gly. Estos constructos de ADN contienen (de 5' a 3') (1) un sitio de clonación *Pst* I, (2) las bases ACC para proporcionar el contexto "Kozak" para una traducción eficaz, (3) la porción del gen que codifica la metionina amino terminal a través del sitio de escisión de péptido de tránsito conocido de ChlamyEPSPS, y que incluye una región de ADN pequeña que codifica los aminoácidos C-terminales con respecto al sitio de escisión, (4) bases de ADN que codifica los residuos Gly - Ser - Gly con un sitio de clonación *Bam*HI incorporado, y, (5) la región codificante del gen de interés.

Ejemplo 5. Fusión de un péptido de tránsito de una especie no vegetal a una proteína heteróloga, y correcta localización y la escisión en monocotiledóneas: CTP de AHAS de *Chlamydomonas* funciona en monocotiledóneas.

55 **[0089]** Se diseñaron constructos de ADN de tal manera que la proteína resultante codificara el péptido de tránsito de AHAS de *Chlamydomonas* ("ChlamyAHAS") en el extremo N-terminal, seguido de una proteína de fusión a un gen que confiere resistencia a herbicida sobre las células (GRG-1; Patente de los Estados Unidos Núm. 7.405.347).

Para el precursor de ChlamyAHAS, los sitios de escisión del péptido de tránsito se dedujeron de los alineamientos de las secuencias de proteínas con las proteínas ALS de bacterias, hongos y levaduras.

5 [0090] Estos constructos de ADN contienen (de 5' a 3') un sitio de clonación *Pst* I, (2) las bases ACC para proporcionar el contexto "Kozak" para una traducción eficaz, (3) la porción del gen que codifica la metionina amino terminal a través del sitio de escisión de péptido de tránsito conocido del CTP de ChlamyAHAS y que incluye una región de ADN pequeña que codifica los aminoácidos C-terminales con respecto al sitio de escisión, (4) bases de ADN que codifica los residuos Gly - Ser - Gly con un sitio de clonación *Bam*H I incorporado, y, (5) la región codificante del gen de interés (en este caso GRG-1).

10 [0091] Estas moléculas de ADN se elaboraron sintéticamente (DNA 2.0 de Menlo Park, CA). La secuencia de ADN de la región que contiene este constructo se proporciona como SEQ ID NO: 11, y la secuencia de aminoácidos resultante se proporciona como SEQ ID NO: 12.

15 [0092] Se elaboró un constructo de control (pAG250), que contenía GRG-1 expresado a partir del promotor *TrpPro5*, en donde la proteína GRG-1 no tiene un CTP de cloroplasto.

20 [0093] Este constructo no CTP/GRG-1 fue diseñado en un vector para su uso en la transformación mediada por *Agrobacterium* de embriones de maíz, y se generaron plantas de maíz transgénicas que contenían este constructo. Las plantas transformadas *To* fueron analizadas por PCR para confirmar la presencia del constructo en las líneas de maíz, y estas plantas *To* se sometieron a cruzamiento natural con una línea no transgénica para generar una progenie *T*₁ hemizigota. Las plantas *T*₁ transgénicas resultantes producen grandes cantidades de proteína GRG-1. Sin embargo, las plantas transformadas con pAG250 y que expresan GRG-1 no localizada no son resistentes al glifosato.

25 [0094] El constructo de CTP de ChlamyAHAS/GRG-1 de algas se diseñó en un vector para su uso en la transformación mediada por *Agrobacterium* de embriones de maíz, y se generaron las plantas de maíz transgénicas que contenían este constructo.

30 [0095] Las plantas *To* transformadas se analizaron mediante PCR para confirmar la presencia del constructo en las líneas de maíz, y estas plantas *To* se sometieron a cruzamiento natural con una línea no transgénica para generar una progenie *T*₁ hemizigota. Las plantas *T*₁ transgénicas resultantes son resistentes a aplicaciones mediante pulverización de glifosato (en comparación con los controles no transgénicos).

35 [0096] Las transferencias Western de tejidos de hoja de plantas de maíz transgénicas muestran que estas plantas expresan la proteína GRG-1. Además, el tamaño de la proteína identificada mediante transferencia Western es congruente con la importación de la proteína a los cloroplastos, y el procesamiento de la proteína ChlamyAHAS/GRG-1 en o cerca del sitio de escisión.

40 [0097] De este modo, el CTP de ChlamyAHAS es suficiente para dirigir GRG-1 al cloroplasto de maíz, y dar como resultado un fenotipo (resistencia a los herbicidas) que no es conferido por GRG-1 en ausencia de direccionamiento al cloroplasto.

45 Ejemplo 6. Fusión de un péptido de tránsito de una especie no vegetal a una proteína heteróloga, y correcto direccionamiento y escisión en monocotiledóneas: CTP de SSU de *Chlamydomonas* funciona en monocotiledóneas

[0098] Para someter a ensayo si un CTP de cloroplastos de algas puede funcionar en monocotiledóneas, se generaron plantas monocotiledóneas transgénicas y se evaluaron la expresión y la escisión de un CTP de algas por medio de análisis de transferencia Western.

50 [0099] Se diseñaron constructos de ADN de tal manera que la proteína resultante codificara el péptido de tránsito de algas en el extremo N-terminal, fusionado a una proteína que confiere resistencia a herbicida sobre las células (en este caso la proteína GRG-8; Publicación de Patente de los Estados Unidos Núm. 20060150270)

55 [0100] Estos constructos de ADN contienen (de 5' a 3') (1) un sitio de clonación *Pst* I, (2) las bases ACC para proporcionar el contexto "Kozak" para una traducción eficaz, (3) la porción del gen que codifica la metionina amino terminal a través del sitio de escisión del péptido de tránsito conocido de ChlamySSU, y que incluye una región de ADN pequeña que codifica los aminoácidos C-terminales con respecto al sitio de escisión, (4) bases de ADN que codifica los residuos Gly - Ser - Gly con un sitio de clonación *Bam*H I incorporado, y, (5) la región codificante del gen de interés (en este caso GRG-8).

60 [0101] La secuencia de ADN de la región que contiene este constructo se proporciona como SEQ ID NO: 13, y la secuencia de aminoácidos resultante se proporciona como SEQ ID NO: 14.

[0102] Este constructo de CTP/GRG-8 (pAG1675) se diseñó en un vector para su uso en la transformación mediada por *Agrobacterium* de embriones de maíz, y se generaron e identificaron las plantas de maíz transgénicas.

5 **[0103]** Las plantas To transformadas con pAG1675 se analizaron mediante PCR para confirmar la presencia del constructo en las líneas de maíz, y estas plantas To se sometieron a cruzamiento natural a continuación con una línea no transgénica para generar una progenie T₁ hemizigota. La plantas T₁ transgénicas exhibieron resistencia a aplicaciones mediante pulverización de glifosato en comparación con los controles no transgénicos.

10 Ejemplo 7. Expresión y procesamiento de CTP de cloroplasto de SSU de *Chlamydomonas* fusionado a la proteína GRG-8 en células de maíz

15 **[0104]** Se encontró que las transferencias Western de tejido de hoja de plantas de maíz transgénico expresaban la proteína CTP/GRG-8. La proteína total de la hoja se extrajo de las hojas de maíz (tampón de extracto de proteínas P-PER de Pierce) y se separó en un gel Bis-Tris al 4-12%. La proteína GRG-8 se visualizó utilizando anticuerpos policlonales de cabra anti-GRG8. Un extracto de maíz no transgénico se comparó junto (calle 3). Para evaluar el procesamiento de CTP, se purificó un patrón de proteína GRG-8 etiquetada con HIS de una cepa de *E. coli*. El tamaño de la proteína identificada mediante transferencia Western es congruente con la importación de la proteína a los cloroplastos, y el procesamiento de la proteína CTP/GRG-8 dentro del CTP de *ChlamySSU*, en o cerca del sitio de escisión pronosticado.

20 Ejemplo 8. Evaluación de la tolerancia al glifosato de la planta de maíz que expresa la proteína SSU-GRG8 de *Chlamydomonas*

25 **[0105]** La tolerancia a la pulverización de glifosato de un evento de maíz transgénico que expresa la proteína SSU-GRG8 de *Chlamydomonas* se comparó con varias plantas To no transgénicas de control. Las plantas individuales se transfirieron al invernadero y se cultivaron en planos durante 10 días. Al cabo de 10 días, se aplicó una concentración de glifosato que se aproximó a una velocidad de pulverización 1x campo (7 mM con un suplemento de Tween 20 al 0,1% como agente tensioactivo) a los planos. El glifosato se aplicó mediante una tabla de pulverización para permitir una aplicación consistente del herbicida a las plantas individuales. Las plantas se clasificaron al cabo de 3 semanas para determinar si las plantas toleraban la pulverización de glifosato (material de la hoja sobre todo verde: <50% de daño) o no toleraban la pulverización de glifosato (>75% daño, o muerte de la planta). La planta transgénica mostró tolerancia a glifosato, mientras que cada una de las plantas de control no logró mostrar la tolerancia.

35 Ejemplo 9. Fusión de un péptido de tránsito de una especie no vegetal a una proteína heteróloga, y correcta localización y escisión en monocotiledóneas: CTP de EPSPS de *Chlamydomonas* funciona en monocotiledóneas

40 **[0106]** El ADN y las secuencias de aminoácidos para el precursor de EPSPS de *Chlamydomonas* se obtuvieron de bases de datos públicas. El sitio de escisión del péptido de tránsito se pronosticó basándose en los alineamientos de las secuencias de proteínas con proteínas EPSPS de bacterias, hongos y levaduras. Se construyó un gen sintético que codificaba el CTP de la metionina amino terminal a través del sitio de escisión pronosticado. Este ADN se ligó para crear una fusión en marco con el codón de inicio de un gen *GRG-23(ace3)(R173K)* sintético (Publicación de la Solicitud de Patente de los Estados Unidos Núm. 20080127372). A continuación este casete CTP-GRG-23(ace3)(R173K) se ligó a un vector de transformación de plantas.

45 **[0107]** Se diseñaron constructos de ADN de tal manera que la proteína resultante codificó el péptido de tránsito de EPSPS de *Chlamydomonas* ("ChlamyEPSPS") en el extremo N-terminal, seguido de la fusión a una proteína que confiere resistencia a herbicidas sobre las células (GRG-23(ace3)(R173K)).

50 **[0108]** Estas moléculas de ADN se elaboraron sintéticamente (DNA 2.0 de Menlo Park, CA). La secuencia de ADN de la región que contiene este constructo se proporciona como SEQ ID NO: 15, y la secuencia de aminoácidos resultante se proporciona como SEQ ID NO: 16.

55 **[0109]** Estos constructos de ADN contienen (de 5' a 3'), (1) un sitio de clonación *Pst* I, (2) las bases ACC para proporcionar el contexto "Kozak" para una traducción eficaz, (3) la porción del gen que codifica la metionina amino terminal a través del sitio de escisión del péptido de tránsito conocido de *ChlamyEPSPS* e incluyendo una pequeña región de ADN que codifica los aminoácidos C-terminales con respecto al sitio de escisión, (4) la región codificante del gen de interés (en este caso GRG-23(ace3)(R173K)).

60 **[0110]** Este constructo de CTP de EPSPS de *Chlamydomonas/GRG-23(ace3)(R173K)* se diseñó en un vector para su uso en la transformación mediada por *Agrobacterium* de embriones de maíz, y se identificaron los eventos transgénicos.

[0111] Las plantas transformadas To se analizaron mediante PCR para confirmar la presencia del constructo en las

líneas de maíz, y estas plantas To se sometieron a cruzamiento natural con una línea no transgénica para generar una progenie T₁ hemizigota. El plantas T₁ transgénicas resultante son resistentes a aplicaciones mediante pulverización de glifosato (en comparación con los controles no transgénicos). Se encontró que las transferencias Western de tejido de hoja de plantas de maíz transgénico expresaba la proteína CTP de EPSPS de *Chlamydomonas*/GRG-23(ace3)(R173K). La proteína detectada en tejidos de la planta es menor que la proteína EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K) completa, y tiene un tamaño similar a la proteína GRG-23(ace3)(R173K) nativa (Figura 2). La capacidad de la proteína CTP de EPSPS de *Chlamydomonas*/GRG-23(ace3)(R173K) para conferir resistencia a herbicidas y el tamaño de la proteína GRG-23(ace3)(R173K) madura resultante en las plantas resistentes a herbicidas son a la vez congruentes con la importación de la proteína a los cloroplastos, y procesamiento de la proteína CTP de EPSPS de *Chlamydomonas*/GRG-23(ace3)(R173K) en o cerca del sitio de escisión.

Ejemplo 10. Análisis del peso molecular de la proteína EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K) expresada en el maíz.

[0112] A partir de una transferencia Western de líneas de maíz transgénicas que expresan la proteína EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K), se representó gráficamente la distancia de migración de los patrones de de peso molecular de proteínas para generar un gráfico lineal de Log de Peso Molecular vs Distancia de Migración (Figura 3). Se utilizó la ecuación de regresión lineal de este gráfico para calcular el peso molecular aparente del patrón de proteína de GRG23(ace3)(R173K) (Figura 2, calle 8) y de la proteína EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K) procesada detectada en extractos de plantas. Por medio de este método, se determinó que el peso molecular aparente de la proteína GRG23(ace3)(R173K) purificada era de 45,347 gramos/mol, y se determinó que el peso molecular aparente de EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K) procesada (calle 6, distancia de migración = 48,5 mm) era de 46,227 gramos/mol. Los pesos moleculares estimados de las proteínas procesadas son congruentes con el procesamiento de varios aminoácidos del CTP de *Chlamydomonas* aguas arriba de su unión con GRG-23(ace3)(R173K) (se estima que el peso molecular de GRG23(ace3)(R173K) es de 45,570 gramos/mol). No se detecta proteína de un tamaño coincidente con la proteína EPSPS de *Chlamydomonas* - GRG-23(ace3)(R173K) sin procesar (PM = 52,012 gramos/mol) mediante transferencia Western. Por lo tanto, el CTP de la EPSPS de *Chlamydomonas* se procesa en el maíz en un sitio de reconocimiento discreto dentro del CTP de *Chlamydomonas*. La purificación y el análisis de aminoácidos N-terminal de esta proteína por medio de métodos conocidos en la técnica permitirían la determinación inequívoca del sitio exacto de escisión dentro del CTP de *Chlamydomonas*.

Ejemplo 11. Evaluación de secuencias de CTP de algas en dicotiledóneas

[0113] Para evaluar la capacidad de los CTP de cloroplastos de algas para funcionar en células dicotiledóneas, el CTP de cloroplasto de AHAS de *Chlamydomonas* (SEC ID NO: 3) se coloca en marco 5' del gen TagGFP (Evrogen, Moscú, Rusia). Un vector de control contenía el gen TagGFP sin un péptido de tránsito al cloroplasto. El constructo de CTP de ChlamyAHAS (pAX3517) y el constructo de control (pAX3521) se organizaron para iniciar la transcripción a partir del promotor UBQ3 de *Arabidopsis* (Norris et. al, 1993, Plant Molecular Biology 21:895-906). Los constructos utilizaron terminadores de la transcripción 35S o PinII.

[0114] Se utilizaron aproximadamente 12 µg de cada plásmido purificado en experimentos de transformación de protoplastos de tabaco mediada por polietilenglicol. Después de la transformación, los protoplastos se incubaron en una cámara de crecimiento a 25°C durante 23 horas. Se verificó la expresión y la localización de la proteína TagGFP después de la expresión transitoria con un microscopio fluorescente invertido.

[0115] El constructo que expresa la TagGFP sin un CTP de cloroplasto se detecto únicamente en el citoplasma de los protoplastos. Sin embargo, el constructo de CTP de AHAS de *Chlamydomonas* liberó correctamente TagGFP en el cloroplasto dando como resultado la acumulación de fluorescencia en el cloroplasto de estos protoplastos (Figura 1).

Ejemplo 12. Evaluación de una secuencia de CTP de algas en células de soja

[0116] Para evaluar la capacidad de los CTP de cloroplastos de algas para funcionar en células de soja, el constructo de CTP ChlamyAHAS (pAX3517), el constructo ChlamyEPSPS (pAX4562), y un constructo de control (pAX3521) que contenía el gen TagGFP sin un péptido de tránsito al cloroplasto se utilizaron en la transformación mediada por polietilenglicol de protoplastos de soja. Expresión y localización de la proteína TagGFP después de la expresión transitoria se verificó con un microscopio fluorescente invertido.

[0117] Para la constructo de control que carecía de un CTP de cloroplasto, se observó la fluorescencia de TagGFP únicamente en el citoplasma. Del mismo modo, no se observó expresión en los cloroplastos a partir de dos constructos independientes que expresaban TagGFP con el CTP de cloroplasto de ChlamyEPSPS. Sin embargo, se detectó TagGFP en el cloroplasto de protoplastos que habían sido transformados con el constructo de CTP de

cloroplasto ChlamyAHAS. Por lo tanto, este CTP funciona en células de soja y así como de tabaco.

Tabla 2. Localización de proteínas conectadas a CTP de cloroplastos de algas en células dicotiledóneas.

Construeto	CTP	Gen	Localización de TagGFP en protoplastos de tabaco	La localización de TagGFP en protoplastos de soja
pAX3517	ChlamyAHAS	TagGFP	Cloroplasto	Cloroplasto
pAX4562	ChlamyEPSPS	TagGFP	<i>no sometido a ensayo</i>	No cloroplasto
pAX3521	Ninguno	TagGFP	No cloroplasto	No cloroplasto

5

Tabla 3. Resumen de la función de los CTP de Chlamydomonas en células vegetales

CTP	Función en células monocotiledóneas	Función en las células Dicotiledóneas
ChlamySSU	++	<i>no sometido a ensayo</i>
ChlamyEPSPS	+++	-
ChlamyAHAS	+++	+++

[0118] Todas las publicaciones y solicitudes de patente mencionadas en la memoria descriptiva son indicativas del nivel de conocimiento práctico de los expertos en la técnica a la que pertenece esta invención.

10

[0119] Aunque la invención anterior se ha descrito con cierto detalle a modo de ilustración y ejemplo para propósitos de claridad de comprensión, resultará obvio que se pueden poner en práctica ciertos cambios y modificaciones dentro del alcance de las reivindicaciones adjuntas.

15

LISTA DE SECUENCIAS

<110> Philip E. Hammer Vadim Beilinson Todd K. Hinson

<120> COMPOSICIONES Y METODOS PARA LA EXPRESIÓN

DE UNA SECUENCIA DE NUCLEÓTIDOS HETERÓLOGA EN PLANTAS

20

<130> 045 600/377 480

<150> 61/095.134

<151> 2008-09-08

<160> 18

<170> FastSEQ para Windows versión 4.0

25

<210> 1

<211> 297

<212> ADN

<213> Chlamydomonas reinhardtii

<400> 1

```

atgaaggccc tgcaagtgg aaccgctgtg gcgcggggcc aagcgggctg tgtttctccc 60
gtcccgccgct ctgtgcctat gtcgtctcag gcatgattc cgagcaccag ctcccagca 120
gtcgtgacac ccgcccggtc cggtcgcccgc gccctcgtg tgtcggccaa gctggctgat 180
gggtctcgtc gcatgcagtc cgaggagggtg cgccgcgcca aggagggtggc ccaggctgcg 240
ctggccaagg acagccctgc cgactgggtg gaccgctacg gctcggagcc gcgcaag 297

```

30

<210> 2

<211> 297

<212> DNA

<213> Chlamydomonas reinhardtii

35

<400> 2

ES 2 480 665 T3

atgaaggccc tgcgaagtgg aaccgctgtg gcgcggggcc aagcgggctg tgtttctccc 60
 gctccgcgcc ctgtgcctat gtcgtctcag gcgatgattc cgagcaccag ctcccagca 120
 gctcgtgcac ccgcccgtc cggtcgccc gcctcgcctg tgtcggccaa gctggctgat 180
 gggtcacgtc gcatgcagtc cgaggagggtg cgccgcgcca aggagggtgc ccaggctgcg 240
 ctggctaagg acagccctgc cgactgggta gaccgctacg gctcggagcc gcgcaag 297

<210> 3

<211> 99

<212> PRT

5 <213> Chlamydomonas reinhardtii

<400> 3

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

Pro Arg Lys

<210> 4

10 <211> 135

<212> ADN

<213> Chlamydomonas reinhardtii

<400> 4

atggccgccc tcattgccaa gtcctccgtc tccgcggccc tggcccggcc ggcccgctcc 60
 agcgtgcgcc ccatggccgc gctgaagccc gccgtcaagg ccgccccctg ggctgccccg 120
 gctcaggcta accag 135

15 <210> 5

<211> 45

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 5

Met Ala Ala Val Ile Ala Lys Ser Ser Val Ser Ala Ala Val Ala Arg
 1 5 10 15
 Pro Ala Arg Ser Val Arg Pro Met Ala Ala Leu Lys Pro Ala Val
 20 25 30
 Lys Ala Ala Pro Val Ala Ala Pro Ala Gln Ala Asn Gln
 35 40 45

20 <210> 6

<211> 225

<212> ADN

<213> Chlamydomonas reinhardtii

25 <400> 6

atgcagctgc tcaaccagcg gcaggcgtg cggtgggaa gaagctccgc cagcaagaac 60
 cagcaggtgg cgccgctggc atcaaggccc gcaagcagcc tctccgtctc cgctcctcc 120
 gtggcgccgg cgccggcctg ctcggcgccc gccggcgccc gccgcccgcg cgtgggtggg 180
 cgcgctccg ccaccaagga gaaggtggag gagctcacca tccag 225

<210> 7

<211> 75

<212> PRT

ES 2 480 665 T3

<213> Chlamydomonas reinhardtii

<400> 7

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

<210> 8

5

<211> 683

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 8

ES 2 480 665 T3

Met Lys Ala Leu Arg Ser Gly Thr Ala Val Ala Arg Gly Gln Ala Gly
 1 5 10 15
 Cys Val Ser Pro Ala Pro Arg Pro Val Pro Met Ser Ser Gln Ala Met
 20 25 30
 Ile Pro Ser Thr Ser Ser Pro Ala Ala Arg Ala Pro Ala Arg Ser Gly
 35 40 45
 Arg Arg Ala Leu Ala Val Ser Ala Lys Leu Ala Asp Gly Ser Arg Arg
 50 55 60
 Met Gln Ser Glu Glu Val Arg Arg Ala Lys Glu Val Ala Gln Ala Ala
 65 70 75 80
 Leu Ala Lys Asp Ser Pro Ala Asp Trp Val Asp Arg Tyr Gly Ser Glu
 85 90 95
 Pro Arg Lys Gly Ala Asp Ile Leu Val Gln Ala Leu Glu Arg Glu Gly
 100 105 110
 Val Asp Ser Val Phe Ala Tyr Pro Gly Ala Ser Met Glu Ile His
 115 120 125
 Gln Ala Leu Thr Arg Ser Asp Arg Ile Thr Asn Val Leu Cys Arg His
 130 135 140
 Glu Gln Gly Glu Ile Phe Ala Ala Glu Gly Tyr Ala Lys Ala Ala Gly
 145 150 155 160
 Arg Val Gly Val Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu
 165 170 175
 Val Thr Gly Leu Ala Asp Ala Met Met Asp Ser Ile Pro Leu Val Ala
 180 185 190
 Ile Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala Phe Gln
 195 200 205
 Glu Thr Pro Ile Val Glu Val Thr Arg Ala Ile Thr Lys His Asn Tyr
 210 215 220
 Leu Val Leu Asp Ile Lys Asp Leu Pro Arg Val Ile Lys Glu Ala Phe
 225 230 235 240
 Tyr Leu Ala Arg Thr Gly Arg Pro Gly Pro Val Leu Val Asp Val Pro
 245 250 255
 Lys Asp Ile Gln Gln Gln Leu Ala Val Pro Asp Trp Glu Ala Pro Met
 260 265 270
 Ser Ile Thr Gly Tyr Ile Ser Arg Leu Pro Pro Pro Val Glu Glu Ser
 275 280 285
 Gln Val Leu Pro Val Leu Arg Ala Leu Gln Gly Ala Ala Lys Pro Val
 290 295 300
 Ile Tyr Tyr Gly Gly Gly Cys Leu Asp Ala Gln Ala Glu Leu Arg Glu
 305 310 315 320
 Phe Ala Ala Arg Thr Gly Ile Pro Leu Ala Ser Thr Phe Met Gly Leu
 325 330 335
 Gly Val Val Pro Ser Thr Asp Pro Asn His Leu Gln Met Leu Gly Met
 340 345 350
 His Gly Thr Val Phe Ala Asn Tyr Ala Val Asp Gln Ala Asp Leu Leu
 355 360 365
 Val Ala Leu Gly Val Arg Phe Asp Asp Arg Val Thr Gly Lys Leu Asp
 370 375 380
 Ala Phe Ala Ala Arg Ala Arg Ile Val His Ile Asp Ile Asp Ala Ala
 385 390 395 400
 Glu Ile Ser Lys Asn Lys Thr Ala His Val Pro Val Cys Gly Asp Val
 405 410 415
 Lys Gln Ala Leu Ser His Leu Asn Arg Leu Leu Ala Ala Glu Pro Leu
 420 425 430
 Pro Ala Asp Lys Trp Ala Gly Trp Arg Ala Glu Leu Ala Ala Lys Arg
 435 440 445
 Ala Glu Phe Pro Met Arg Tyr Pro Gln Arg Asp Asp Ala Ile Val Pro
 450 455 460
 Gln His Ala Ile Gln Val Leu Gly Glu Glu Thr Gln Gly Glu Ala Ile
 465 470 475 480
 Ile Thr Thr Gly Val Gly Gln His Gln Met Trp Ala Ala Gln Trp Tyr
 485 490 495
 Pro Tyr Lys Glu Thr Arg Arg Trp Ile Ser Ser Gly Gly Leu Gly Ser

ES 2 480 665 T3

Met Gly Phe Gly Leu Pro Ala Ala Leu Gly Ala Ala Val Ala Phe Asp
 Gly Lys Asn Gly Arg Pro Lys Lys Thr Val Val Asp Ile Asp Gly Asp
 Gly Ser Phe Leu Met Asn Val Gln Glu Leu Ala Thr Ile Phe Ile Glu
 Lys Leu Asp Val Lys Val Met Leu Leu Asn Asn Gln His Leu Gly Met
 Val Val Gln Trp Glu Asp Arg Phe Tyr Lys Ala Asn Arg Ala His Thr
 Tyr Leu Gly Lys Arg Glu Ser Glu Trp His Ala Thr Gln Asp Glu Glu
 Asp Ile Tyr Pro Asn Phe Val Asn Met Ala Gln Ala Phe Gly Val Pro
 Ser Arg Arg Val Ile Val Lys Glu Gln Leu Arg Gly Ala Ile Arg Thr
 Met Leu Asp Thr Pro Gly Pro Tyr Leu Leu Glu Val Met Val Pro His
 Ile Glu His Val Leu Pro Met Ile Pro Gly Gly Ala Ser Phe Lys Asp
 Ile Ile Thr Glu Gly Asp Gly Thr Val Lys Tyr

<210> 9

<211> 185

<212> PRT

5

<213> Chlamydomonas reinhardtii

<400> 9

Met Ala Ala Val Ile Ala Lys Ser Ser Val Ser Ala Ala Val Ala Arg
 Pro Ala Arg Ser Ser Val Arg Pro Met Ala Ala Leu Lys Pro Ala Val
 Lys Ala Ala Pro Val Ala Ala Pro Ala Gln Ala Asn Gln Met Met Val
 Trp Thr Pro Val Asn Asn Lys Met Phe Glu Thr Phe Ser Tyr Leu Pro
 Pro Leu Ser Asp Glu Gln Ile Ala Ala Gln Val Asp Tyr Ile Val Ala
 Asn Gly Trp Ile Pro Cys Leu Glu Phe Ala Glu Ser Asp Lys Ala Tyr
 Val Ser Asn Glu Ser Ala Ile Arg Phe Gly Ser Val Ser Cys Leu Tyr
 Tyr Asp Asn Arg Tyr Trp Thr Met Trp Lys Leu Pro Met Phe Gly Cys
 Arg Asp Pro Met Gln Val Leu Arg Glu Ile Val Ala Cys Thr Lys Ala
 Phe Pro Asp Ala Tyr Val Arg Leu Val Ala Phe Asp Asn Gln Lys Gln
 Val Gln Ile Met Gly Phe Leu Val Gln Arg Pro Lys Ser Ala Arg Asp
 Trp Gln Pro Ala Asn Lys Arg Ser Val

<210> 10

<211> 512

<212> PRT

10

<213> Chlamydomonas reinhardtii

<400> 10

Met Gln Leu Leu Asn Gln Arg Gln Ala Leu Arg Leu Gly Arg Ser Ser
1 5 10 15
Ala Ser Lys Asn Gln Gln Val Ala Pro Leu Ala Ser Arg Pro Ala Ser
20 25 30
Ser Leu Ser Val Ser Ala Ser Ser Val Ala Pro Ala Pro Ala Cys Ser
35 40 45
Ala Pro Ala Gly Ala Gly Arg Ala Val Val Val Arg Ala Ser Ala
50 55 60
Thr Lys Glu Lys Val Glu Glu Leu Thr Ile Gln Pro Val Lys Lys Ile
65 70 75 80
Ala Gly Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile
85 90 95
Leu Leu Leu Ala Ala Leu Ser Glu Gly Thr Thr Leu Val Lys Asn Leu
100 105 110
Leu Asp Ser Asp Ile Arg Tyr Met Val Gly Ala Leu Lys Ala Leu
115 120 125
Asn Val Lys Leu Glu Glu Asn Trp Glu Ala Gly Glu Met Val Val His
130 135 140
Gly Cys Gly Gly Arg Phe Asp Ser Ala Gly Ala Glu Leu Phe Leu Gly
145 150 155 160
Asn Ala Gly Thr Ala Met Arg Pro Leu Thr Ala Ala Val Val Ala Ala
165 170 175
Gly Arg Gly Lys Phe Val Leu Asp Gly Val Ala Arg Met Arg Glu Arg
180 185 190
Pro Ile Glu Asp Leu Val Asp Gly Leu Val Gln Leu Gly Val Asp Ala
195 200 205
Lys Cys Thr Met Gly Thr Gly Cys Pro Pro Val Glu Val Asn Ser Lys
210 215 220
Gly Leu Pro Thr Gly Lys Val Tyr Leu Ser Gly Lys Val Ser Ser Gln
225 230 235 240
Tyr Leu Thr Ala Leu Leu Met Ala Ala Pro Leu Ala Val Pro Gly Gly
245 250 255 260
Ala Gly Gly Asp Ala Ile Glu Ile Ile Ile Lys Asp Glu Leu Val Ser
265 270 275
Gln Pro Tyr Val Asp Met Thr Val Lys Leu Met Glu Arg Phe Gly Val
280 285
Val Val Glu Arg Leu Asn Gly Leu Gln His Leu Arg Ile Pro Ala Gly
290 295 300
Gln Thr Tyr Lys Thr Pro Gly Glu Ala Tyr Val Glu Gly Asp Ala Ser
305 310 315 320
Ser Ala Ser Tyr Phe Leu Ala Gly Ala Thr Ile Thr Gly Gly Thr Val
325 330 335
Thr Val Glu Gly Cys Gly Ser Asp Ser Leu Gln Gly Asp Val Arg Phe
340 345 350
Ala Glu Val Met Gly Leu Leu Gly Ala Lys Val Glu Trp Ser Pro Tyr
355 360 365
Ser Ile Thr Ile Thr Gly Pro Ser Ala Phe Gly Lys Pro Ile Thr Gly
370 375 380
Ile Asp His Asp Cys Asn Asp Ile Pro Asp Ala Ala Met Thr Leu Ala
385 390 395 400
Val Ala Ala Leu Phe Ala Asp Arg Pro Thr Ala Ile Arg Asn Val Tyr
405 410 415
Asn Trp Arg Val Lys Glu Thr Glu Arg Met Val Ala Ile Val Thr Glu
420 425 430
Leu Arg Lys Leu Gly Ala Glu Val Glu Glu Gly Arg Asp Tyr Cys Ile
435 440 445
Val Thr Pro Pro Pro Gly Gly Val Lys Gly Val Lys Ala Asn Val Gly
450 455 460
Ile Asp Thr Tyr Asp Asp His Arg Met Ala Met Ala Phe Ser Leu Val
465 470 475 480
Ala Ala Ala Gly Val Pro Val Val Ile Arg Asp Pro Gly Cys Thr Arg
485 490 495
Lys Thr Phe Pro Thr Tyr Phe Lys Val Phe Glu Ser Val Ala Gln His

500 505 510

<210> 11

5 <211> 1602

<212> ADN

<213> Secuencia Artificial

<220>

<223> constructo sintético ChlamyAHAS/GRG1

<400> 11

```

atgaaggccc tgcgaagtgg aaccgctgtg gcgcggggcc aagcgggctg tgtttctccc 60
gctccgcgcc ctgtgcctat gtcgtctcag gcgatgattc cgagcaccag ctccccagca 120
gctcgtgcac ccgcccggtc cggtcgccgc gcctcgtctg tgtcggccaa gctggctgat 180
gggtctcgtc gcatgcagtc cgaggagggtg cgccgcgccca aggagggtggc ccaggctgcg 240
ctggccaagg acagccctgc cgactgggtg gaccgctacg gctcggagcc gcgcaagga 300
tccggcatga aggtgacaat ccagcctggc gatctcacag gcatcattca gagcccagcg 360
tcaaagtctt caatgcagag agcgtgcgcy gcgccctgg tggcgaaggg gatctcagaa 420
atcatcaacc ttgggcatag caacgatgat aaggccgcga gagatctcgt gagccgtctt 480
ggggccagac ttgaagatca gccagatggc agcctccaga tcacttcaga agcggttaag 540
ccagtggcgc ctttcatcga ttgcggggaa tcagggctgt ctatccgat gttcacacca 600
atcgtggcgc tctcaaagga agaagtgaca atcaaggggt cagggtcact cgttactcgc 660
cctatggatt tcttcgatga aatcctgcc caactgggcy tgaagtgaa gtcaaatcag 720
gggaagctcc ctctggttat ccaggggccca cttaagccag cggatgttac agttgatggg 780
tctctctcat ctcagttcct gacagccctc ctgcttgcc acgcccgcc ggatgccagc 840
gatgttgcca tcaaggtgac taacctgaag tcacgtcctt acatcgatct tactcttgat 900
gttatgaagc gtttcggcct caagactcct gaaaaccgca actacgaaga gttctactt 960
aaggccggga acgtgtacga cgaacaaga atgcagcgtt acactgttga aggggattgg 1020
tcagggggcg cgttcctgct cgttcggggg gccatcgccg ggccaatcac tgttcgtggc 1080
cttgatatcg cgtcaactca ggcgataaag gcgatcgctt aggcgctcat gagcgccaac 1140
gccgggatcg cgatcgatgc caaggaaatc aagctgcatc ctgccgatct gaacgccttc 1200
gagttcgatg ccaactgatg ccctgatctc tccccaccac tcgtggccct cgctcatac 1260
tgcaaggggg aaacaaagat caagggcgtg agccgccttg cgcataagga atctgataga 1320
gggctgactc ttcaggatga gttcgggaag atgggcyttg aaatccatct tgaaggggat 1380
ctcatgcgtg tgatcggcgg gaagggggtg aagggcgccg aagttagctc acgtcatgat 1440
catcgcaccg ccatggcgtg cgccgtggcg gcgctcaagg ccgttgggga aacaacaatc 1500
gaacatgccg aagcgggtaa caagtcttac cctgatttct actcagattt gaagcagctc 1560
ggggcgctgg tgtctctgaa ccatcagttc aacttctctt ag 1602
    
```

<210> 12

<211> 533

5

<212> PRT

<213> Secuencia Artificial

<220>

<223> constructo sintético ChlamyAHAS/GRG1

<400> 12

```

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

10

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

<210> 13

<211> 1413

<212> ADN

5

<213> Secuencia Artificial

<220>

<223> constructo sintético ChlamyAHAS/GRG8

<400> 13

```

ctgcagacca tggccgccgt cattgccaag tcctccgtct ccgcggccgt ggccccccg 60
gcccgcctca gCGTgCGccc catggccgCG ctgaagcccG ccgtcaaggc cgcccccgTg 120
gtgccccggG ctCaggctaa ccagggatcc ggcatgatga tgggtagagc caaactcagc 180
attatcccGc cgggcaagcC tttgaccgga cgcgccatgc cgccgggatc gaagtcgatc 240
accaaccgCG cattgctgct cgccggcctc gccaaaggca cgagccggct aaccggTgCG 300
ctgaagagCG acgataccCG ctatattggcc gaagcgctgc gtgCGatggg tGtaacgatc 360
gacgagcccc acgacaccac gttcatcgtc aaaggcagCG gcaagctgca gccgcccggca 420
gccccgcttt tcctcgGcaa tgccggcagc gcaacgcgct tcctgaCGgc ggccgCGgca 480
ctggTggagc gcaaggTcat cgtcGacggc gatgcccata tgcGcaagcG gccgatcGga 540
ccgctagtcG acgcgTtGcG ctCGctcggc atcGatgcct cggctGaaac cggctGccccG 600
ccagTcacga tcaacggcac cggccgcttc gaggcaagcc gcgtGcagat cGatggcggc 660
ctgtccagcc agtatgtctc ggcgctcctg atgatggccG ccggcGcgca tcgcgctgtc 720
gatgtcGagc ttctcggcga acatatcggc gctctcggct atatcGacct gaccgTtgcc 780
gccatgCGcG ctttcggcGc gaaggttgag cgtgtgagcc cggTcGcctg gcgCGtcgag 840
cccaccggct atcatgCGgc cgacttcgTg atcGagccgg atgcctctgc tgcGacctat 900
ctctgggccc ccgaagtTct gagcggcggc aagatcGatc tcggcacgcc ggCGgaacag 960
ttctcGcaac cggatGcgaa agcctatgat ctGatTtcga aattccCGca tctGcctgct 1020
gtcatcGagc gctcGcagat gcaggacgCC atcccGacgc tcgcccTtct cGccgctttc 1080
aacGaaatgc ctgtgCGctt cgtcGgtatc gaaaacctgc gcgtcaagga atgcGatcgt 1140
atcCGcGcGc tctcGagcGg cctatcccGc atcgtTccga acctcGgcac ggaagagggc 1200
gacgatctca tcatcGcctc cGatccgagc ctTgcccggca aaatcctgac cGcagagatc 1260
gatagctttg ccgatcaccG catcGccatg agctttGcGc tggccggcct gaagatcggc 1320
ggcattacca ttctcGaccC cgactgCGtc gccaaGacat tcccgtccta ctggaatgTg 1380
ctgtcttcGc tgggggTcGc ctacgaagac tga 1413

```

<210> 14

<211> 467

<212> PRT

5

<213> Secuencia Artificial

<220>

<223> constructo sintético ChlamyAHAS/GRG8

<400> 14

```

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

```

10

ES 2 480 665 T3

210 215 220
 Ala Leu Leu Met Met Ala Ala Gly Gly Asp Arg Ala Val Asp Val Glu
 225 230 235
 Leu Leu Gly Glu His Ile Gly Ala Leu Gly Tyr Ile Asp Leu Thr Val
 245 250 255
 Ala Ala Met Arg Ala Phe Gly Ala Lys Val Glu Arg Val Ser Pro Val
 260 265 270
 Ala Trp Arg Val Glu Pro Thr Gly Tyr His Ala Ala Asp Phe Val Ile
 275 280 285
 Glu Pro Asp Ala Ser Ala Ala Thr Tyr Leu Trp Ala Ala Glu Val Leu
 290 295 300
 Ser Gly Gly Lys Ile Asp Leu Gly Thr Pro Ala Glu Gln Phe Ser Gln
 305 310 315
 Pro Asp Ala Lys Ala Tyr Asp Leu Ile Ser Lys Phe Pro His Leu Pro
 320 325 330 335
 Ala Val Ile Asp Gly Ser Gln Met Gln Asp Ala Ile Pro Thr Leu Ala
 340 345 350
 Val Leu Ala Phe Asn Glu Met Pro Val Arg Phe Val Gly Ile Glu
 355 360 365
 Asn Leu Arg Val Lys Glu Cys Asp Arg Ile Arg Ala Leu Ser Ser Gly
 370 375 380
 Leu Ser Arg Ile Val Pro Asn Leu Gly Thr Glu Glu Gly Asp Asp Leu
 385 390 395 400
 Ile Ile Ala Ser Asp Pro Ser Leu Ala Gly Lys Ile Leu Thr Ala Glu
 405 410 415
 Ile Asp Ser Phe Ala Asp His Arg Ile Ala Met Ser Phe Ala Leu Ala
 420 425 430
 Gly Leu Lys Ile Gly Gly Ile Thr Ile Leu Asp Pro Asp Cys Val Ala
 435 440 445
 Lys Thr Phe Pro Ser Tyr Trp Asn Val Leu Ser Ser Leu Gly Val Ala
 450 455 460
 Tyr Glu Asp
 465

<210> 15

<211> 1548

<212> ADN

5

<213> Secuencia Artificial

<220>

<223> constructo sintético ChlamyEPS/SPS/GRG23(ace3)(R173K)

<400> 15

atgaaggccc tgcgaagtgg aaccgctgtg gcgcggggcc aagcgggctg tgtttctccc 60
 gtcctcgccc ctgtgcctat gtcgtctcag gcgatgattc cgagcaccag ctccccagca 120
 gctcgtgcac ccgcccggtc cggctcggcg gccctcgtctg tgcctggccaa gctggctgat 180
 ggtctctgtc gcatgcagtc cgaggagggtg cgccgcgcca aggagggtggc ccaggctgcg 240
 ctggccaagg acagccctgc cgactgggtg gaccgctacg gctcggagcc gcgcaaggga 300
 tccggcatgg aaactgatcg ccttgtgatc ccaggatcga aaagcatcac caaccgggct 360
 ttgcttttgg ctgccgcagc gaagggcagc tcggtcctgg tgagaccatt ggtcagcgc 420
 gatacctcag cattcaaac tgcaatccag gccctcgggtg ccaacgtctc agccgacggg 480
 gacgattggg tcggtgaagg cctgggtcag gcaccaacc tcgacgccga catctggtgc 540
 gaggacgcag gtactgtggc ccggttcctc cctccattcg tagccgcagg tcaggggaag 600
 ttcaccgtcg acggatcaga gcagctgcgg cggcgcccgc ttcggcccgt ggtcgacggc 660
 atccgccacc tgggcgcccg cgtctcctcc gagcagctgc cccttacaat tgaagcgagc 720
 gggctggcag gcggggagta cgaattgaa gccatcaga gcagccagtt cgcctccggc 780
 ctgatcatgg ccgccccgta cgcgagacaa ggctcgcgtg tgcggatacc aaatcccgtg 840
 tcacagccct acctcacgat gacactgcgg atgatgaggg acttcggcat tgagaccagc 900
 accgacggag ccaccgtcag cgtccctcca gggcgtaca cagcccggcg gtatgaaata 960
 gaaccggatg cgtcaactgc gtcgtacttc gccgcccgtt ccgcccgtctc tggcagggcg 1020
 ttcgaatttc aaggccttgg cacagacagc atccaaggcg acacgtcatt cttcaatgta 1080
 cttgggcggc tcggtgcgga ggtccactgg gcatccaact cggtcacat acggggaccg 1140

gaaaggctga ccggcgacat tgaagtggat atgggcgaga tttcggācac cttcatgaca 1200
 ctgcggcgca ttgccccttt ggccgatgga cccatcacga taaccaacat tggctatgca 1260
 cggttgaagg aatccgaccg catctcagcg atggaaagca acctgcgcac gctcgggtgta 1320
 caaacggagc tcggacacga ctggatgaga atctaccctt ctaccgccca cggcggtaga 1380
 gtgaattgcc accgggacCa caggatcgct atggcgtttt caatcctggg actgagagtg 1440
 gacgggatta ccctcgacga ccctcaatgc gtcgggaaga ctttccctgg cttcttcgac 1500
 taccttggac gccttttccc cgaaaaggcg cttacgctcc ccggctag 1548

<210> 16

<211> 515

10

<212> PRT

<213> Secuencia Artificial

<220>

<223> constructo sintético ChlamyEPSPS/GRG23(ace3)(R173K)

5

<400> 16

Met Lys Ala Leu Arg Ser Gly Thr Ala Val Ala Arg Gly Gln Ala Gly
 1 5 10 15
 Cys Val Ser Pro Ala Pro Arg Pro Val Pro Met Ser Ser Gln Ala Met
 20 25 30
 Ile Pro Ser Thr Ser Ser Pro Ala Arg Ala Pro Ala Arg Ser Gly
 35 40 45
 Arg Arg Ala Leu Ala Val Ser Ala Lys Leu Ala Asp Gly Ser Arg Arg
 50 55 60
 Met Gln Ser Glu Glu Val Arg Arg Ala Lys Glu Val Ala Gln Ala Ala
 65 70 75 80
 Leu Ala Lys Asp Ser Pro Ala Asp Trp Val Asp Arg Tyr Gly Ser Glu
 85 90 95
 Pro Arg Lys Gly Ser Gly Met Glu Thr Asp Arg Leu Val Ile Pro Gly
 100 105 110
 Ser Lys Ser Ile Thr Asn Arg Ala Leu Leu Leu Ala Ala Ala Lys
 115 120 125
 Gly Thr Ser Val Leu Val Arg Pro Leu Val Ser Ala Asp Thr Ser Ala
 130 135 140
 Phe Lys Thr Ala Ile Gln Ala Leu Gly Ala Asn Val Ser Ala Asp Gly
 145 150 155 160
 Asp Asp Trp Val Val Glu Gly Leu Gly Gln Ala Pro Asn Leu Asp Ala
 165 170 175
 Asp Ile Trp Cys Glu Asp Ala Gly Thr Val Ala Arg Phe Leu Pro Pro
 180 185 190
 Phe Val Ala Ala Gly Gln Gly Lys Phe Thr Val Asp Gly Ser Glu Gln
 195 200 205
 Leu Arg Arg Arg Pro Leu Arg Pro Val Val Asp Gly Ile Arg His Leu
 210 215 220
 Gly Ala Arg Val Ser Ser Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser
 225 230 235 240
 Gly Leu Ala Gly Gly Glu Tyr Glu Ile Glu Ala His Gln Ser Ser Gln
 245 250 255
 Phe Ala Ser Gly Leu Ile Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu
 260 265 270
 Arg Val Arg Ile Pro Asn Pro Val Ser Gln Pro Tyr Leu Thr Met Thr
 275 280 285
 Leu Arg Met Met Arg Asp Phe Gly Ile Glu Thr Ser Thr Asp Gly Ala
 290 295 300
 Thr Val Ser Val Pro Pro Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile
 305 310 315 320
 Glu Pro Asp Ala Ser Thr Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val
 325 330 335
 Ser Gly Arg Arg Phe Glu Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln
 340 345 350
 Gly Asp Thr Ser Phe Phe Asn Val Leu Gly Arg Leu Gly Ala Glu Val
 355 360 365
 His Trp Ala Ser Asn Ser Val Thr Ile Arg Gly Pro Gln Arg Leu Thr
 370 375 380
 Gly Asp Ile Glu Val Asp Met Gly Glu Ile Ser Asp Thr Phe Met Thr
 385 390 395 400
 Leu Ala Ala Ile Ala Pro Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn
 405 410 415
 Ile Gly His Ala Arg Leu Lys Glu Ser Asp Arg Ile Ser Ala Met Glu
 420 425 430
 Ser Asn Leu Arg Thr Leu Gly Val Gln Thr Asp Val Gly His Asp Trp
 435 440 445
 Met Arg Ile Tyr Pro Ser Thr Pro His Gly Gly Arg Val Asn Cys His
 450 455 460
 Arg Asp His Arg Ile Ala Met Ala Phe Ser Ile Leu Gly Leu Arg Val
 465 470 475 480
 Asp Gly Ile Thr Leu Asp Asp Pro Gln Cys Val Gly Lys Thr Phe Pro
 485 490 495
 Gly Phe Phe Asp Tyr Leu Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr
 500 505 510
 Leu Pro Gly
 515

<210> 17

<211> 4
 <212> PRT
 <213> Secuencia Artificial
 <220>
 5 <223> dominio de direccionamiento al estroma
 <220>
 <221> VARIANTE
 <222> 1
 <223> Xaa = Ile o Val
 10 <220>
 <221> VARIANTE
 <222> 2
 <223> Xaa = cualquier aminoácido
 <220>
 15 <221> VARIANTE
 <222> 3
 <223> Xaa = Ala o Cys
 <400> 17
 Xaa Xaa Xaa Ala
 1
 20 <210> 18
 <211> 3
 <212> PRT
 <213> Secuencia Artificial
 <220>
 25 <223> secuencia conectora
 <400> 18
 Gly Ser Gly
 1

REIVINDICACIONES

1. Una célula vegetal monocotiledónea que tiene incorporado establemente en su genoma un casete de expresión que comprende una secuencia de nucleótidos que codifica un péptido dirigido al cloroplasto (CTP), en donde dicha secuencia de nucleótidos que codifica dicho CTP está conectada operablemente a una secuencia de nucleótidos de interés, y en donde dicha secuencia de nucleótidos que codifica dicho CTP se selecciona del grupo que consiste en:
 - a) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6;
 - b) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto;
 - c) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y,
 - d) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto.
2. Una planta que comprende la célula vegetal monocotiledónea de la reivindicación 1.
3. Una semilla derivada de la planta de la reivindicación 2, en donde dicha semilla comprende la secuencia de nucleótidos que codifica dicho péptido de tránsito al cloroplasto.
4. La célula vegetal monocotiledónea de la reivindicación 1, en donde la secuencia de nucleótidos de interés codifica un producto génico que confiere resistencia a herbicidas, agentes patógenos, o insectos.
5. Un método para expresar una secuencia de nucleótidos de interés en una planta monocotiledónea, comprendiendo dicho método:
 - a) introducir en una célula vegetal un casete de expresión que comprende una secuencia de nucleótidos que codifica un péptido dirigido al cloroplasto (CTP), en donde dicha secuencia de nucleótidos que codifica dicho CTP está conectada operablemente a dicha secuencia de nucleótidos de interés, y en donde dicha secuencia de nucleótidos que codifica dicho CTP se selecciona del grupo que consiste en
 - i) la secuencia de nucleótidos mostrada en el SEQ ID NO: 6;
 - ii) una secuencia de nucleótidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 6, en donde dicha secuencia de nucleótidos codifica un péptido de tránsito al cloroplasto;
 - iii) una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada en el SEQ ID NO: 7; y,
 - iv) una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con el SEQ ID NO: 7, en donde dicha secuencia de aminoácidos es un péptido de tránsito al cloroplasto; y,
 - b) regenerar una planta transformada a partir de dicha célula vegetal; en donde dicha planta ha incorporado de forma estable en su genoma dicho casete de expresión.
6. El método de la reivindicación 5, en donde dicha secuencia de nucleótidos de interés codifica un producto génico que confiere resistencia a herbicidas o plagas.

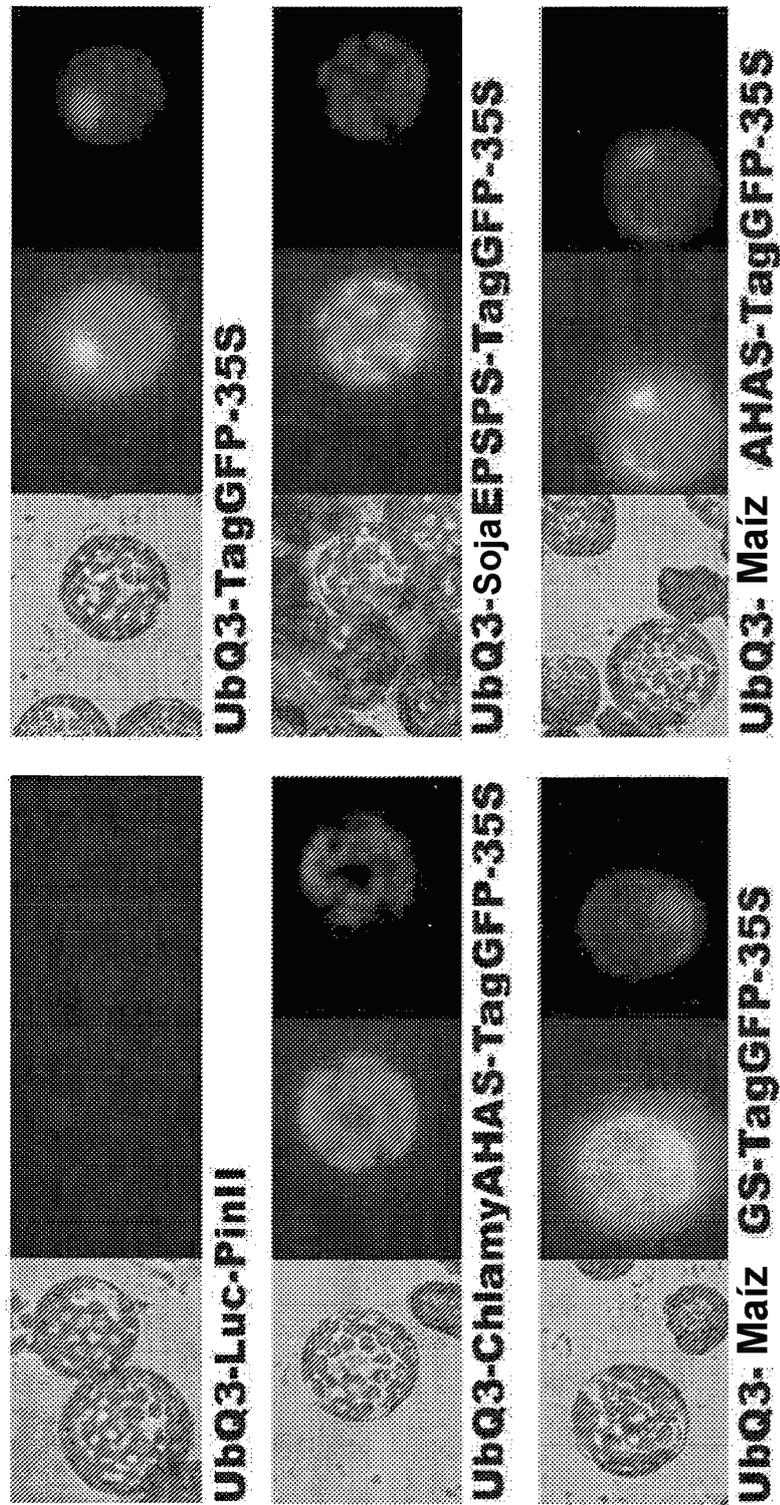


FIG. 1

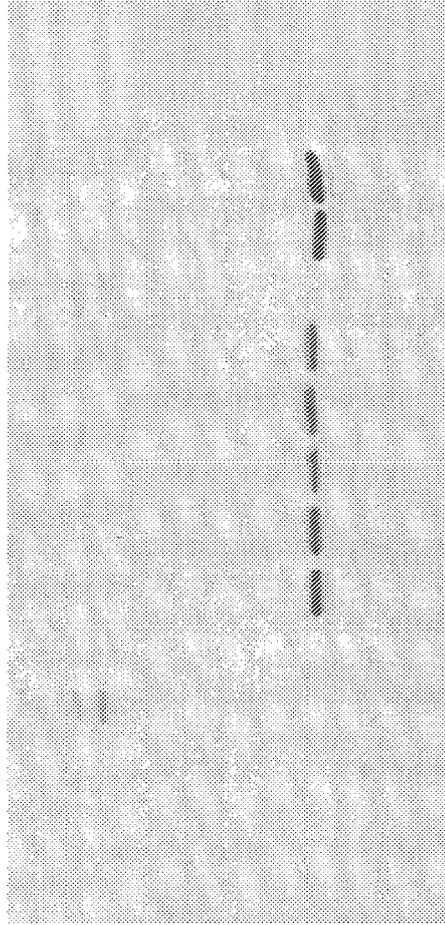


FIG. 2

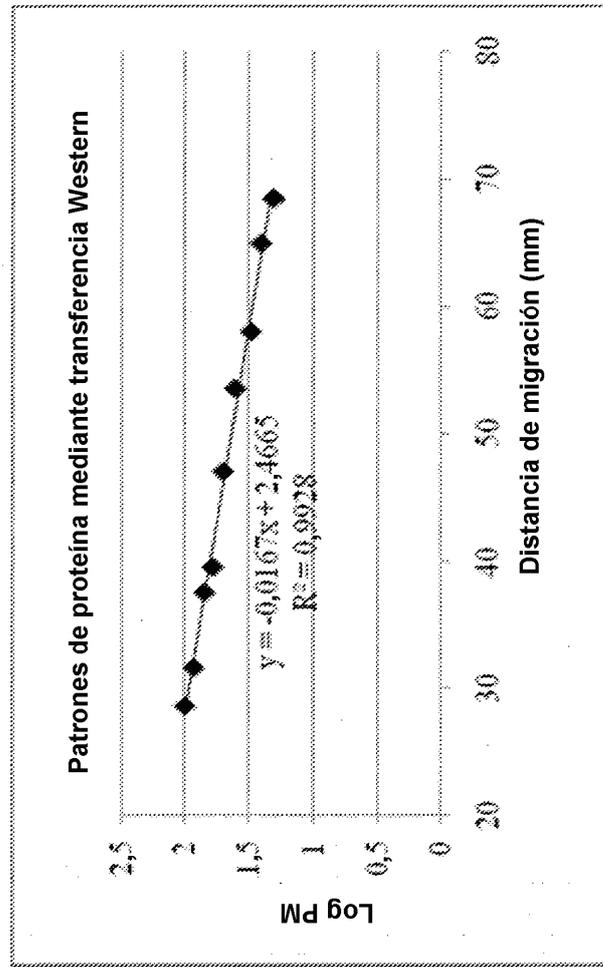


FIG. 3