

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 469 590**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12N 15/53 (2006.01)

C12N 15/63 (2006.01)

A01H 5/00 (2006.01)

A01H 5/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.06.2005 E 05756097 (1)**

97 Fecha y número de publicación de la concesión europea: **30.04.2014 EP 1756284**

54 Título: **Procedimiento de modificación de la morfología, la fisiología y la bioquímica de plantas que comprende la expresión de citoquinina oxidasa en las semillas**

30 Prioridad:

18.06.2004 US 871304

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

18.06.2014

73 Titular/es:

SCHMULLING, THOMAS (50.0%)

Preussenallee 30

14052 Berlin, DE y

WERNER, TOMAS (50.0%)

72 Inventor/es:

SCHMULLING, THOMAS y

WERNER, TOMAS

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 469 590 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

Procedimiento de modificación de la morfología, la fisiología y la bioquímica de plantas que comprende la expresión de citoquinina oxidasa en las semillas

5 La presente invención se refiere, en general, a procedimientos para modificar las propiedades o las características morfológicas de plantas, tales como procesos del desarrollo, y en particular el desarrollo de semillas y/o un mayor
 10 rendimiento de semillas de una planta. Los procedimientos comprenden expresar una proteína de control de la degradación de citoquinina, en particular citoquinina oxidasa 2, en la planta, operablemente bajo el control de una secuencia de promotor regulable, en el que el promotor regulable es una secuencia de promotor específica de la aleurona y/o del embrión. La presente invención se extiende a construcciones genéticas que son útiles para realizar
 15 el procedimiento de la invención y a plantas transformadas con estas, teniendo dichas plantas unas propiedades morfológicas alteradas, comparadas con sus homólogos por lo demás isogénicos.

Las semillas son la unidad de reproducción de las plantas superiores. Las semillas contienen compuestos de reserva para asegurar la nutrición del embrión después de la germinación. Estos compuestos de almacenamiento contribuyen significativamente a la nutrición humana, así como a la alimentación del ganado. Las semillas consisten
 20 en tres partes principales, a saber, el embrión, el endospermo y la cubierta de la semilla. Los compuestos de reserva se depositan en los tejidos de almacenamiento, que son el endospermo (que resulta de una fertilización doble, por ejemplo, en todos los cereales), el denominado perispermo (derivado del tejido de nucela) o los cotiledones (por ejemplo, variedades de legumbres). El endospermo está cubierto por una capa de células citoplásmicas densa, conocida como aleurona. Los compuestos de almacenamiento incluyen lípidos (aceite de colza), proteínas (por ejemplo, en la aleurona de cereales) o carbohidratos (almidón, oligosacáridos como la rafinosa).

El almidón es el compuesto de almacenamiento en las semillas de los cereales. Las especies de cereales más importante son el maíz (producción anual de aproximadamente 570 millones de toneladas), el arroz (540 millones de toneladas anuales) y el trigo (530 millones de toneladas anuales). Las semillas ricas en proteínas incluyen
 25 diferentes tipos de legumbres (*Phaseolus spec.*, *Vicia faba*, *Vigna spec.*; aproximadamente 20 millones de toneladas anuales), guisante (*Pisum sativum*; un millón de toneladas anuales) y soja (*Glycine max*; 136 millones de toneladas anuales). Las semillas de soja también son una importante fuente de lípidos, así como las semillas de diferentes especies de *Brassica* (aproximadamente 30 millones de toneladas anuales), algodón, sésamo oriental, lino, amapola, ricino, girasol, cacahuete, coco, palma y algunas otras plantas de menor importancia económica.

30 Después de la fertilización, la semilla en desarrollo se convierte en un órgano sumidero que atrae compuestos nutricionales desde órganos fuente de la planta y los utiliza para producir los compuestos de reserva en los órganos de almacenamiento, tales como semillas, bulbos o tubérculos. El concepto común predice que las citoquininas son un regulador positivo de la fuerza de sumidero.

Numerosos informes atribuyen una función estimuladora o inhibidora a las citoquininas en diferentes procesos del desarrollo, tales como el crecimiento y la ramificación de raíces, el control de la dominancia apical en el brote, el desarrollo de cloroplastos, y la senescencia de las hojas (Mok M.C. (1994), en Cytokines: Chemistry, Activity and Function, eds., Mok, D.W.S. & Mok, M.C. (CRC Boca Ratón, FL), pp.155-166). Se han extraído conclusiones acerca de las funciones biológicas de las citoquininas principalmente de estudios sobre las consecuencias de la aplicación de citoquininas exógenas o de niveles de citoquininas potenciados de modo endógeno (Klee, H.J. y Lanehon, M.B. (1995), en Plant Hormones: Physiology, Biochemistry and Molecular Biology, ed. Davies, P.J. (Kluwer, Dordrecht, Países Bajos), pp. 340-353; Schmülling, T., Rupp, H.M., Frank, M. y Schafer, S. (1999), en Advances in Regulation of Plant Growth and Development, eds. Surnad, M. Pac P. y Beck, E. (Peres, Praga), pp. 85-96).

La clonación de la citoquinina oxidasa ha permitido el estudio de la importancia de las citoquininas de tipo iP y Z durante el ciclo vital completo de las plantas superiores. La enzima catabólica citoquinina oxidasa (CKX) desempeña un papel principal en el control de los niveles de citoquinina en tejidos vegetales. Se ha descubierto actividad CKX en un gran número de plantas superiores y en diferentes tejidos vegetales. La enzima es una oxidoreductasa que contiene FAD que cataliza la degradación de citoquininas que portan cadenas laterales de isoprenoides insaturados y a veces se denomina citoquinina deshidrogenasa (Frébortova *et al.*, Biochem. J., 380, 121-130, 2004). Las bases libres iP y Z, y sus respectivos ribósidos, son los sustratos preferidos. Los productos de la reacción del catabolismo de iP son adenina y el aldehído insaturado 3-metil-2-butanal (Armstrong, D.J. (1994), en Cytokinin: Chemistry, Activity and Functions, eds. Mok, D.W.S. & Mok, M.C. (CRC Boca Ratón, FL), pp. 139-154).

Se han aislado los genes de citoquinina oxidasa de diversas especies vegetales: *Zea mays* (Morris, R.O., Bilyeu, K.D., Laskey, J.G. y Cherich, N.N. (1999), Biochem. Biophys. Res. Commun., 255, 328-333; Houba-Herlin, N., Pethe, C., d'Alayer, J. y Laloue, M. (1999), Plant J., 17, 615-626), *Arabidopsis thaliana* (documento WO 01/96580, documento WO 03/050287), y *Dendrobium* sp. (Yang *et al.*, J. Exp. Bot., 53, 1899-1907). La manipulación de la expresión del gen CKX se emplea como una herramienta poderosa para estudiar la importancia de las citoquininas

de tipo iP y Z durante el ciclo vital de plantas superiores; en particular, la manipulación de la expresión de CKX permite determinar la influencia de la citoquinina sobre el crecimiento de las raíces y el desarrollo y la morfología de brotes, sobre el tamaño de las semillas y sobre la senescencia de las hojas (documento WO 01/96580; documento WO 03/050287; Werner *et al.*, Plant Cell, 15, 2632-2550, 2003). Los transformantes que muestran una expresión de ARNm de *AtCKX* constitutiva y una mayor actividad citoquinina oxidasa manifiestan una mayor formación y crecimiento de las raíces. También se observaron efectos negativos sobre el crecimiento de brotes.

Se ha postulado que las citoquininas desempeñan un papel fundamental en el maíz para establecer el tamaño de las semillas, disminuir el aborto de la punta del grano, y aumentar la fecundación de las semillas durante condiciones ambientales desfavorables (documento WO 00/63401). Smigocki *et al.* (Proc. Natl. Acad. Sci. USA, 85, 5131-5135, 1988) han observado una mayor organogénesis de brotes tras la infección de tallos, trozos de hojas y plántulas con *Agrobacterium tumefaciens* que sobreexpresa la isopenteniltransferasa (ipt). Por tanto, también se ha postulado que unos mayores niveles de citoquininas en la semilla proporcionarán un mayor tamaño de las semillas y una mayor fertilización de las semillas (documento WO 00/63401). El mismo punto de partida se empleó en los documentos US 2004/237147 y WO 2005/097824, obteniéndose un mayor rendimiento de semillas mediante la sobreexpresión de una proteína de isopenteniltransferasa. A lo largo de las mismas líneas, en el documento WO 2004/044200, se ha predicho que una delección funcional de la citoquinina oxidasa mejoraría el rendimiento de semillas en plantas, y en particular en el arroz.

En la técnica se sabe que la expresión de CKX constitutiva en una planta conduce a aumentar el crecimiento radical, pero disminuye el crecimiento de brotes (documento WO 01/96580). También se ha demostrado que la expresión de CKX bajo el control de un promotor constitutivo fuerte en una planta da como resultado un mayor tamaño de las semillas, un mayor tamaño del embrión y un mayor tamaño de los cotiledones (documento WO 03/050287). Sin embargo, no se sabe qué partes de la semilla serían las más adecuadas para la expresión de CKX para lograr un mayor rendimiento de semillas. Los inventores ahora han demostrado, por primera vez, que puede obtenerse un mayor rendimiento de semillas con relación a plantas control cuando la citoquinina oxidasa (CKX), y en particular la citoquinina oxidasa 2 (CKX2) se sobreexpresa en ciertas partes de la semilla, concretamente la aleurona y/o el embrión. Los inventores también han demostrado que la sobreexpresión de CKX2 específica de la aleurona y/o del embrión conduce a un aumento mayor en el rendimiento de semillas, comparado con la sobreexpresión de CKX específica del endospermo.

Según la presente invención, se ha descubierto, de forma sorprendente, que plantas transgénicas que sobreexpresan un gen de citoquinina oxidasa, y en particular el gen de citoquinina oxidasa 2, en una parte confinada de la semilla, siendo dicha parte confinada la aleurona y/o el embrión, tienen un mayor rendimiento de semillas, comparado con las correspondientes plantas de tipo salvaje. Estos resultados son sorprendentes, puesto que se esperaría que un menor contenido en citoquininas estaría asociado con un crecimiento reducido de los órganos y, por tanto, con un menor rendimiento.

La presente memoria descriptiva proporciona un método para modificar las propiedades morfológicas de plantas, en particular para aumentar el rendimiento de semillas de una planta, que comprende expresar una citoquinina oxidasa bajo el control de un promotor regulable capaz de dirigir la expresión en la aleurona y/o el embrión de una semilla. La memoria descriptiva también proporciona composiciones para modificar las propiedades morfológicas de plantas, en particular para aumentar el rendimiento de semillas de una planta.

Los expertos en la técnica entenderán que la "morfología de una planta" o las "características morfológicas de una planta" o expresiones similares, cuando se utilizan en la presente, se refieren al aspecto externo de una planta. Más en concreto, la característica morfológica de la planta que se mejora utilizando los procedimientos según la presente memoria descriptiva es el rendimiento de semillas.

Tal como se emplea en la presente, la expresión "planta control" se refiere a plantas que, excepto por las características morfológicas modificadas o mejoradas, son muy similares o idénticas a la planta modificada. Estas plantas control también se denominan "líneas/plantas de tipo salvaje" o "líneas/plantas por lo demás isogénicas". Los inventores también han demostrado que el aumento en el rendimiento de semillas es mayor para plantas en las que la CKX se sobreexpresa en la aleurona y/o el embrión, comparado con plantas en las que la CKX se sobreexpresa en el endospermo. Por tanto, la expresión "plantas control" también incluye plantas en las que la CKX se sobreexpresa en el endospermo.

Existen varios parámetros conocidos que pueden utilizarse para medir el aumento en el rendimiento de semillas que incluyen, pero no se limitan al peso total de las semillas, el número total de semillas, el número total de semillas llenas, el índice de cosecha, y el peso de mil granos. Según se describe en los ejemplos de la presente, el peso total de las semillas puede medirse pesando todas las semillas llenas recolectadas de una planta. El número total de semillas puede medirse contando el número de semillas recolectadas de una planta. El número total de semillas llenas puede medirse contando el número de semillas llenas recolectadas de una planta. La expresión "índice de cosecha", tal como se emplea en la presente, se define como la proporción entre el peso total de las semillas y el

área sobre el suelo (mm^2), multiplicado por un factor de 10^6 . El peso de mil granos puede derivarse del número de semillas llenas contadas y su peso total.

De forma ventajosa, la actuación de los procedimientos según la presente memoria descriptiva resulta en plantas que tienen mayor rendimiento o biomasa, con relación a las correspondientes plantas control.

5 Por "rendimiento" se quiere significar la cantidad de material recolectado por área de producción. La expresión "mayor rendimiento" incluye un aumento en la biomasa en una o más partes de una planta con relación a la biomasa de las correspondientes plantas control. Dependiendo del cultivo, la parte recolectada de la planta puede ser diferente, por ejemplo, puede ser la semilla (por ejemplo, arroz, sorgo o maíz cuando se cultiva para obtener
10 semillas), la biomasa sobre el suelo total (por ejemplo, maíz cuando se emplea como ensilaje, caña de azúcar), la raíz (remolacha azucarera), el fruto (por ejemplo, tomate), las fibras de algodón, o cualquier otra parte de la planta que tenga un valor económico. Por ejemplo, los procedimientos de la presente memoria descriptiva se emplean para aumentar el rendimiento de semillas en el arroz y en el maíz. El aumento en el rendimiento incluye un aumento en el rendimiento de semillas, que incluye un aumento en la biomasa total de las semillas (peso total de las semillas), el número total de semillas y/o un aumento en el número de semillas (llenas). El aumento en el
15 rendimiento también se refleja como un aumento en el índice de cosecha, que se expresa como una proporción del rendimiento de las partes recolectables, tales como semillas, frente a la biomasa total, y también se refleja como un aumento en el peso de mil granos (derivado del número de semillas llenas contadas y su peso total).

Así, se proporciona un procedimiento para aumentar el rendimiento de semillas de una planta, aumentando el nivel y/o la actividad de una citoquinina oxidasa (CKX), y en particular la citoquinina oxidasa 2 (CKX2) en el embrión y/o
20 la aleurona de la semilla de una planta, en el que dicho embrión y/o aleurona tienen mayores niveles y/o actividad de citoquinina oxidasa, y en particular citoquinina oxidasa 2, con relación a otras partes de dicha semilla, y en el que el aumento en el rendimiento de semillas comprende al menos uno de un mayor peso total de las semillas, un mayor número total de semillas, un mayor número de semillas llenas, un mayor índice de cosecha o un mayor peso de mil granos, cada uno con relación a las correspondientes plantas control.

25 El rendimiento es, por naturaleza, un parámetro complejo en el que rendimiento total depende de una serie de componentes del rendimiento. Los parámetros para un mayor rendimiento de un cultivo son muy conocidos por los expertos en la técnica. Como ejemplo, los componentes clave del rendimiento para el trigo incluyen el número de plantas por hectárea o acre, el número de espigas por planta, el número de filas (de semillas) por espiga, el número de granos por fila, y el peso de mil granos. La mejora en el rendimiento, obtenida según los procedimientos de la
30 memoria descriptiva, puede obtenerse como el resultado de uno o más de estos componentes del rendimiento. Como ejemplo, los componentes clave del rendimiento para el arroz incluyen el número de plantas por hectárea o acre, el número de panículas por planta, el número de espiguillas por panícula, la tasa de llenado de semillas y el peso de mil granos. La mejora en el rendimiento, obtenida según los procedimientos de la memoria descriptiva, puede obtenerse como el resultado de uno o más de estos componentes del rendimiento, preferentemente la mejora en el rendimiento se obtiene principalmente basándose en un mayor número de flores por panícula y una mayor tasa de llenado de semillas.
35

La actuación de los procedimientos según la presente memoria descriptiva produce plantas que tienen un rendimiento de semillas modificado. El rendimiento de semillas modificados incluye al menos un aumento en uno cualquiera o más del peso total de las semillas, el número total de semillas, el número de semillas llenas, el peso de
40 mil granos y el índice de cosecha, cada uno con relación a plantas control. Por tanto, según la presente memoria descriptiva, se proporciona un procedimiento para aumentar uno o más del peso total de las semillas, el número total de semillas, el número de semillas llenas, el peso de mil granos y el índice de cosecha de plantas, comprendiendo dicho procedimiento modular la expresión de una molécula de ácido nucleico que codifica una proteína de CKX y/o modular la actividad de la propia CKX en una planta de una manera preferida en la aleurona
45 y/o el embrión, preferiblemente en el que la proteína de CKX es codificada por una secuencia de ácido nucleico representada por SEQ ID NO:26, SEQ ID NO:42, o una de sus porciones, o por secuencias capaces de hibridarse con estas, o en el que la CKX está representada por SEQ ID NO:4, SEQ ID NO:36, o un homólogo, derivado o fragmento activo de esta. Como alternativa, la CKX puede estar codificada por una secuencia de ácido nucleico representada por SEQ ID NO:37, o por una de sus porciones, o por secuencias capaces de hibridarse con esta, o
50 en el que la CKX está representada por SEQ ID NO:38, o un homólogo, derivado o fragmento activo de cualquiera de estas.

Los procedimientos de esta memoria descriptiva son aplicables a cualquier planta, en particular a plantas monocotiledóneas y plantas dicotiledóneas que incluyen una legumbre de forraje o pienso, una planta ornamental,
55 un cultivo alimentario, un árbol o un arbusto seleccionado de la lista que comprende *Acacia spp.*, *Acer spp.*, *Actinidia spp.*, *Aesculus spp.*, *Agathis australis*, *Albizia amara*, *Alsophila tricolor*, *Andropogon spp.*, *Arachis spp.*, *Areca catechu*, *Astelia fragrans*, *Astragalus cicer*, *Avena sativa*, *Baikiaea plurijuga*, *Betula spp.*, *Brassica spp.*, *Bruguiera gymnorhiza*, *Burkea africana*, *Butea frondosa*, *Cadaba farinosa*, *Calliandra spp.*, *Camellia sinensis*,

5 *Canna indica*, *Capsicum* spp., *Cassia* spp., *Centroema pubescens*, *Chaenomeles* spp., *Cinnamomum cassia*,
Coffea arabica, *Colophospermum mopane*, *Coronilla varia*, *Cotoneaster serotina*, *Crataegus* spp., *Cucumis* spp.,
Cupressus spp., *Cyathia dealbata*, *Cydonia oblonga*, *Cryptomeria japonica*, *Cymbopogon* spp., *Cynthea dealbata*,
Cydonia oblonga, *Dalbergia monetaria*, *Davallia divaricata*, *Desmodium* spp., *Dicksonia squarosa*, *Diheteropogon*
10 *amplectens*, *Dioclea* spp., *Dolichos* spp., *Dorycnium rectum*, *Echinochloa pyramidalis*, *Ehretia* spp., *Eleusine*
coracana, *Eragrostis* spp., *Erythrina* spp., *Eucalyptus* spp., *Euclea schimperii*, *Eulalia villosa*, *Fagopyrum* spp., *Feijoa*
sellowiana, *Fragaria* spp., *Flemingia* spp., *Freycinetia banksii*, *Geranium thunbergii*, *Ginkgo biloba*, *Glycine javanica*,
Gliricidia spp., *Gossypium hirsutum*, *Grevillea* spp., *Guibourtia coleosperma*, *Hedysarum* spp., *Hemarthia altissima*,
Heteropogon contortus, *Hordeum vulgare*, *Hyparrhenia rufa*, *Hypericum erectum*, *Hyperthelia dissoluta*, *Indigo*
15 *incarnata*, *Iris* spp., *Leptarrhena pyrolifolia*, *Lespedeza* spp., *Lettuca* spp., *Leucaena leucocephala*, *Loudetia simplex*,
Lotonous bainesii, *Lotus* spp., *Macrotyloma axillare*, *Malus* spp., *Manihot esculenta*, *Medicago sativa*, *Metasequoia*
glyptostroboides, *Musa sapientum*, *Nicotianum* spp., *Onobrychis* spp., *Ornithopus* spp., *Oryza* spp., *Peltophorum*
africanum, *Pennisetum* spp., *Persea gratissima*, *Petunia* spp., *Phaseolus* spp., *Phoenix canariensis*, *Phormium*
cookianum, *Photinia* spp., *Picea glauca*, *Pinus* spp., *Pisum sativum*, *Podocarpus totara*, *Pogonarthria fleckii*,
20 *Pogonarthria squarrosa*, *Populus* spp., *Prosopis cineraria*, *Pseudotsuga menziesii*, *Pterolobium stellatum*, *Pyrus*
communis, *Quercus* spp., *Rhaphiolepis umbellata*, *Rhopalostylis sapida*, *Rhus natalensis*, *Ribes grossularia*, *Ribes*
spp., *Robinia pseudoacacia*, *Rosa* spp., *Rubus* spp., *Salix* spp., *Schyzachyrium sanguineum*, *Sciadopitys verticillata*,
Sequoia sempervirens, *Sequoiadendron giganteum*, *Sorghum bicolor*, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus*
alopecuroides, *Stylosanthes humilis*, *Tadehagi* spp., *Taxodium distichum*, *Themeda triandra*, *Trifolium* spp., *Triticum*
25 *spp.*, *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vitis vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*,
Zea mays, amaranto, alcachofa, espárrago, brécol, coles de Bruselas, repollo, canola, zanahora, coliflor, apio,
verduras de hoja verde, lino, col rizada, lenteja, colza, oca, cebolla, patata, arroz, soja, remolacha azucarera,
azúcar de caña, girasol, tomate, calabaza, y té, entre otros, o las semillas de cualquier planta anteriormente
nombrada específicamente o un cultivo de tejido, de células o de órganos de cualquiera de las anteriores especies.

25 Según una realización preferida de la presente memoria descriptiva, la planta es una planta de cultivo, tal como
soja, girasol, canola, alfalfa, colza, algodón, tomate, patata o tabaco. También preferiblemente, la planta es una
planta monocotiledónea, tal como azúcar de caña. Más preferiblemente, la planta es un cereal, tal como arroz,
maíz, trigo, cebada, mijo, centeno, sorgo o avena.

30 El término “planta”, tal como se emplea en la presente, incluye plantas completas, los antecesores y la progenie de
las plantas y partes de las plantas, que incluyen semillas, brotes, tallos, hojas, raíces (que incluyen tubérculos),
flores, y tejidos y órganos, en los que cada uno de estos comprende el gen/ácido nucleico de interés. El término
“planta” también incluye cultivos en suspensión, tejido de callo, embriones, regiones meristemáticas, gametofitos,
esporofitos, polen y microesporas, de nuevo en los que cada uno de estos comprende el gen/ácido nucleico de
35 “plantas” y la expresión “partes de plantas”.

Debe resultar evidente que, aunque los procedimientos de esta memoria descriptiva están apoyados en la sección
de ejemplos por los genes y proteínas de AtCKX2, el concepto inventivo también se refiere al uso de citoquinina
oxidasa aisladas a partir de otras plantas y expresadas en la aleurona y/o el embrión de una semilla de estas otras
plantas para obtener efectos similares en plantas, según se describe en la sección de ejemplos.

40 La familia de genes de citoquinina oxidasa contiene al menos seis miembros de *Arabidopsis*. Se anticipa que
podrán aislarse homólogos funcionales de las citoquinina oxidasa de *Arabidopsis* descritas de otros organismos,
dada la prueba de la presencia de actividad citoquinina oxidasa en muchas plantas verdes (Hare y van Staden,
Physiol. Plant, 91:128-136, 1994; Jones y Schreiber, *Plant Growth Reg.*, 23:123-134, 1997), así como en otros
45 organismos (Armstrong, en *Cytokinins: Chemistry, Activity and Function*, eds. Mok y Mok, CRC Press, pp. 139-154,
1994). Por tanto, la secuencia de la citoquinina oxidasa, que es funcional en los procedimientos de esta memoria
descriptiva, no necesita ser idéntica a las descritas en la presente. Los procedimientos de esta memoria descriptiva
son particularmente útiles para cultivos de cereales y cultivos de monocotiledóneas en general, y también los genes
de citoquinina oxidasa que se pueden utilizar procedentes, por ejemplo, de trigo o maíz (Morris *et al.*, 1999; Rinaldi
50 y Comandini, 1999). Se contempla que otros genes con actividad citoquinina oxidasa o con cualquier otra actividad
metabolizante de citoquininas (véase Zažimalová *et al.*, *Biochemistry and Molecular Biology of Plant Hormones*,
Hooykaas, Hall y Libbenga (eds.), Elsevier Science, pp. 141-160, 1997) también puedan utilizarse para los fines de
esta descripción. De forma similar, los genes que codifican proteínas que aumentan la actividad metabolizante de
citoquininas endógenas también pueden utilizarse para los objetivos de esta descripción. En principio, también
55 pueden obtenerse fenotipos similares interfiriendo con genes que actúan cadena arriba de una citoquinina, tales
como receptores o proteínas implicados en la vías de transducción de señales de la citoquinina.

Cualquier secuencia de nucleótidos que codifica un polipéptido con actividad citoquinina oxidasa puede utilizarse en
los procedimientos de la memoria descriptiva. Por ejemplo, cualquiera de las diversas secuencias proporcionadas
en la presente que codifica un polipéptido con actividad citoquinina oxidasa puede utilizarse en los procedimientos

para aumentar el rendimiento de semillas.

5 Las expresiones “proteína o proteínas”, “péptido o péptidos”, “polipéptido o polipéptidos” u “oligopéptido u oligopéptidos”, cuando se emplean en la presente, se refieren a aminoácidos en una forma polimérica de cualquier longitud. Estas expresiones también incluyen modificaciones de aminoácidos conocidas, así como restos aminoácidos no naturales, restos L-aminoácidos y restos D-aminoácidos.

Las proteínas de CKX útiles en los procedimientos según la memoria descriptiva se definen en la presente como que tienen actividad citoquinina oxidasa/deshidrogenasa, y comprenden al menos 2 secuencias de 17 y 19 restos aminoácidos consecutivos, respectivamente, con una secuencia consenso como se muestra a continuación:

Secuencia consenso 1 (17 aminoácidos): hTDYLhhoIGGTLSSsG (SEQ ID NO:44)

10 Secuencia consenso 2 (19 aminoácidos): cLFxushGsLGOFGIstA (SEQ ID NO:45)

15 en las que las letras en mayúscula son los códigos IUPAC de una sola letra convencionales para los diversos aminoácidos, y las otras letras simbolizan la naturaleza de los aminoácidos, tal como se muestran en la tabla 1. La columna de la derecha lista, para cada clase, los aminoácidos concretos que se permiten en las secuencias consenso. Esta clasificación se basa en el agrupamiento de aminoácidos, según se define en la base de datos SMART (Schultz *et al.* (1998), Proc. Natl. Acad. Sci. USA, 95, 5857-5864; Letunic *et al.* (2002), Nucleic Acids Res., 30, 242-244).

Tabla 1

Clase	Clave	Aminoácidos permitidos
h	hidrófobos	G, H, L, R, T, W, Y
o	alcohol	S, T
l	alifático	V, I
s	pequeño	A, D, G, N, T, V
c	cargado	D, E, R
x	cualquiera	Y, R, N, F, D, H
u	muy pequeño	A, G, S
t	de tipo vuelta	R, N

20 La estructura cristalina de la CKX del maíz ha sido aclarada por Malito *et al.* (J. Mol. Biol., 341, 1237-1249, 2004), que ha permitido realizar un análisis de estructura-función y la identificación de aminoácidos que son importantes para la actividad catalítica de la CKX del maíz. Estos resultados pueden extrapolarse a otras proteínas de CKX mediante la comparación de las secuencias de proteínas.

25 Los procedimientos para medir la actividad citoquinina oxidasa/deshidrogenasa son muy conocidos en la técnica. Los procedimientos adecuados se basan en la conversión de [2-³H]iP en adenina (Motyka *et al.*, Plant Physiology, 112, 1035-1043, 1996), en ensayos colorimétricos (Libreros-Minotta y Tipton, Anal. Biochem., 231, 339-341, 1995), o en la medición de aceptores de electrones reducidos (Bilyeu *et al.*, Plant Physiol., 125, 378-386, 2001).

La presente memoria descriptiva se refiere a procedimientos para aumentar el rendimiento de semillas. En particular, los procedimientos comprenden la expresión en la aleurona y/o el embrión de una semilla de un ácido nucleico que codifica un citoquinina oxidasa seleccionada del grupo que consiste en:

30 a. ácidos nucleicos que comprenden una secuencia de ADN según una cualquiera de SEQ ID NO:42, 37, 3, 26, 31 o 33, o su complemento,

b. ácidos nucleicos que comprenden las secuencias de ARN que se corresponden con una cualquiera de SEQ ID NO:42, 37, 3, 26, 31 o 33, o su complemento,

35 c. ácidos nucleicos que se hibridan específicamente con cualquiera de SEQ ID NO:42, 37, 3, 26, 31 o 33, o su complemento,

d. ácidos nucleicos que codifican una proteína que comprende la secuencia de aminoácidos según una cualquiera de SEQ ID NO:4, 32, 36 o 38, o su complemento,

e. ácidos nucleicos según se define en uno cualquiera de (a) a (d) que se caracterizan porque dicho ácido nucleico es ADN, ADN genómico, ADNc, ADN sintético, o ARN en el que T es reemplazado por U,

5 f. un ácido nucleico que está degenerado comparado con un ácido nucleico según una cualquiera de SEQ ID NO:42, 37, 3, 26, 31 o 33, o que está degenerado comparado con un ácido nucleico según se define en uno cualquiera de (a) a (e) como resultado del código genético,

10 g. ácidos nucleicos que son divergentes de un ácido nucleico que codifica una proteína según cualquiera de SEQ ID NO:4, 36 o 38, que es divergente de un ácido nucleico según se define en uno cualquiera de (a) a (e), debido a diferencias en la utilización de codones entre los organismos,

h. ácidos nucleicos que codifican una proteína según SEQ ID NO:4, 36 o 38, o ácidos nucleicos según se define en (a) a (e) que son divergentes debido a las diferencias entre alelos,

i. ácidos nucleicos que codifican una proteína según cualquiera de SEQ ID NO:4, 36 o 38,

15 j. fragmentos funcionales de ácidos nucleicos según se define en (a) a (i) que tienen la actividad biológica de una citoquinina oxidasa, y

k. ácidos nucleicos que codifican una citoquinina oxidasa vegetal, que comprenden la secuencia consenso hTDYLhholGGTLSSsG y cLFxushGsLGQFGIIstA o que comprenden la expresión en semillas de un ácido nucleico que codifica una proteína que reduce el nivel de citoquininas activas en plantas o partes de plantas.

20 Más en concreto, la presente invención proporciona un procedimiento para aumentar el número de semillas llenas de una planta, consistiendo dicho procedimiento en (i) introducir en una célula vegetal una molécula de ácido nucleico aislada que codifica una citoquinina oxidasa unida operablemente a un promotor que dirige específicamente la expresión en el embrión y/o la aleurona de una semilla, y (ii) aumentar el nivel y/o la actividad de dicha citoquinina oxidasa en el embrión y/o la aleurona de la semilla de una planta, en el que dicha molécula de ácido nucleico aislada que codifica una citoquinina oxidasa es un ácido nucleico que codifica la proteína representada por SEQ ID NO:36, o es una molécula de ácido nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con dicho ácido nucleico que codifica la proteína de SEQ ID NO:36, y en el que dicho promotor comprende los nucleótidos 1-1256 de SEQ ID NO:41.

La expresión de un ácido nucleico que codifica un polipéptido de CKX o uno de sus homólogos también puede aumentar introduciendo una modificación genética (preferiblemente en el locus de un gen CKX).

30 Además de los genes de citoquinina oxidasa y las correspondientes proteínas descritas anteriormente, un gen de citoquinina oxidasa 2 (CKX2) es particularmente adecuado para su uso para aumentar el rendimiento de semillas en una planta. Además de CKX2 de *Arabidopsis thaliana* indicada en SEQ ID NO:36, pueden utilizarse otras proteínas de *Arabidopsis thaliana*, tales como las representadas en GenBank n.º de registro NP_181682, NP_200507, NP_849470, NP_194703, NP_850863 o AAG30909. También están disponibles proteínas de CKX procedentes de otras especies, tales como *Zea mays* (por ejemplo, GenBank n.º de registro CAE55202, CAE55200 o AAC27500), *Dendrobium* (GenBank CAC17752), *Hordeum vulgare* (GenBank AAN16383, AAO50082, AAM08400), o arroz (GenBank NP_913145, NP_916348, NP_922039) para su uso en los procedimientos y composiciones de la presente memoria descriptiva. Un homólogo procariota de SEQ ID NO:36 está representado por GenBank n.º de registro P46377.

40 Por tanto, la presente memoria descriptiva se refiere, de un modo más general, a un procedimiento para aumentar el rendimiento de semillas en una planta, comprendiendo dicho procedimiento aumentar el nivel y/o la actividad de una citoquinina oxidasa (CKX) en el embrión y/o la aleurona de la semilla de una planta, en el que dicho embrión y/o aleurona tiene un mayor nivel y/o actividad de citoquinina oxidasa con relación a otras partes de dicha semilla. Las citoquinina oxidadasas preferidas que se van a utilizar son codificadas por los ácidos nucleicos que codifican la citoquinina oxidasa, según se definió anteriormente. Más preferiblemente, las citoquinina oxidadasas que se van a utilizar proceden de *Arabidopsis thaliana*, lo más preferiblemente la citoquinina oxidasa que se va a utilizar es AtCKX2 codificada por una de SEQ ID NO:3, 26, 37 o 42.

45 Las expresiones "gen o genes", "ácido nucleico o ácidos nucleicos", "secuencia o secuencias de ácidos nucleicos", "secuencia o secuencias de nucleótidos", o "molécula o moléculas de ácidos nucleicos", cuando se emplean en la presente, se refieren a nucleótidos, bien ribonucleótidos o bien desoxirribonucleótidos, o una combinación de ambos, en una forma polimérica de cualquier longitud. Las expresiones también incluyen ARN y ADN bicatenario y monocatenario. Las expresiones también incluyen modificaciones de nucleótidos conocidas, tales como metilación, ciclación y formación de "casquetes", y la sustitución de uno o más de los nucleótidos naturales por un análogo, tal

como inosina. Las expresiones también incluyen ácido nucleicos peptídicos (PNA), un análogo de ADN en el que el esqueleto es un pseudopéptido que consiste en unidades de N-(2-aminoetil)glicina, en lugar de un azúcar.

Una “secuencia codificadora” o “marco de lectura abierto” u “ORF” se define como una secuencia de nucleótidos que puede transcribirse en ARNm y/o traducirse en un polipéptido cuando se coloca bajo el control de secuencias de control o secuencias reguladoras apropiadas, es decir, cuando la secuencia codificadora u ORF está presente en un formato expresable. Esta secuencia codificadora u ORF está limitada por un codón de inicio de la traducción 5' y un codón de fin de la traducción 3'. Una secuencia codificadora u ORF puede incluir, pero no se limita a ARN, ARNm, ADNc, secuencias de nucleótidos recombinantes, secuencias de nucleótidos fabricadas de modo sintético, o ADN genómico. Estas secuencias codificadoras u ORF pueden estar interrumpidas por secuencias de ácidos nucleicos intermedias.

Los genes y las secuencias codificadoras que codifican fundamentalmente la misma proteína pero que se aíslan de diferentes fuentes pueden consistir en secuencias de ácidos nucleicos sustancialmente divergentes. De forma recíproca, pueden diseñarse secuencias de ácidos nucleicos sustancialmente divergentes para realizar la expresión de fundamentalmente la misma proteína. Estas secuencias de ácidos nucleicos son el resultado, por ejemplo, de la existencia de diferentes alelos de un gen concreto, de la degeneración del código genético o de diferencias en la utilización de codones. Así, aminoácidos tales como la metionina y el triptófano son codificados por un único codón, mientras que otros aminoácidos, tales como arginina, leucina y serina, pueden ser traducidos a partir de hasta seis codones diferentes. Las diferencias en la utilización de codones preferidos son conocidas en la técnica. Por ejemplo, el codón GGC (para la glicina) es el codón que se utiliza con más frecuencia en *Agrobacterium tumefaciens* (36,2‰), es el segundo codón que se utiliza con más frecuencia en *Oryza sativa*, pero se utiliza con una frecuencia mucho más baja en *A. thaliana* y *Medicago sativa* (9‰ y 8,4‰, respectivamente). De los cuatro posibles codones que codifican la glicina, este codón GGC se emplea lo más preferiblemente en *A. tumefaciens* y *O. sativa*. Sin embargo, en *A. thaliana*, este sería el codón GGA (y GGU), mientras que en *M. sativa*, este sería el codón GGU (y GGA).

Los alelos existen en la naturaleza, y dentro de los procedimientos de la presente memoria descriptiva se incluye el uso de estos variantes alélicos naturales. Los variantes alélicos incluyen polimorfismos de un solo nucleótido (SNP), así como polimorfismos de inserción/delección pequeños (INDEL). El tamaño de los INDEL habitualmente es menor que 100 pb. Los SNP y los INDEL forman el mayor conjunto de variantes de secuencia en las cepas polimórficas naturales de la mayoría de los organismos.

Las secuencias de ADN, según se definen en la presente memoria descriptiva, pueden estar interrumpidas por secuencias intermedias. “Secuencias intermedias” significa cualquier secuencia de ácido nucleico que interrumpe una secuencia codificadora que comprende dicha secuencia de ADN de la invención, o que interrumpe el formato expresable de una secuencia de ADN que comprende la secuencia de ADN de la invención. La eliminación de la secuencia intermedia restablece la secuencia codificadora o el formato expresable. Los ejemplos de secuencias intermedias incluyen intrones y secuencias de ADN movilizables, tales como transposones. Una “secuencia de ADN movilizable” significa cualquier secuencia de ADN que puede moverse como resultado de un acontecimiento de recombinación.

Los procedimientos según la presente memoria descriptiva también pueden practicarse utilizando un variante de corte y empalme alternativo de una molécula de ácido nucleico que codifica una proteína de CKX. La expresión “variante de corte y empalme alternativo”, tal como se emplea en la presente, incluye variantes de una molécula de ácido nucleico en los que intrones y/o exones seleccionados han sido extirpados, reemplazados o añadidos. Estos variantes serán variantes en los que la actividad biológica de la proteína permanece sin afectar, que pueden obtenerse conservando selectivamente segmentos funcionales de la proteína. Estos variantes de corte y empalme pueden encontrarse en la naturaleza o pueden ser artificiales. Los procedimientos para preparar estos variantes de corte y empalme son muy conocidos en la técnica. Por tanto, según otro aspecto de la presente memoria descriptiva, se proporciona un procedimiento para modificar las características morfológicas de plantas, y en particular del rendimiento de semillas, que comprende la expresión en la aleurona y/o el embrión de una semilla de un variante de corte y empalme alternativo de una molécula de ácido nucleico que codifica una CKX. Preferiblemente, el variante de corte y empalme es un variante de corte y empalme de una secuencia representada por SEQ ID NO:1, 3, 5, 7, 9, 11 o 33. Un variante de corte y empalme preferido de *AtCKX2* está representado por SEQ ID NO:38.

También son útiles en los procedimientos de la presente memoria descriptiva los ácidos nucleicos capaces de hibridarse bajo condiciones de rigurosidad reducida, preferiblemente bajo condiciones rigurosas, con un ácido nucleico que codifica CKX. Se prefiere un ácido nucleico capaz de hibridarse con un ácido nucleico representado por SEQ ID NO:42 o 26, y es más preferida una molécula de ácido nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con un ácido nucleico que codifica la proteína de SEQ ID NO:36.

La "hibridación" es el procedimiento en el que secuencias de nucleótidos complementarias sustancialmente homólogas se asocian entre sí. El procedimiento de hibridación puede producirse totalmente en disolución, es decir, ambos ácidos nucleicos complementarios están en disolución. Las herramientas de la biología molecular que se basan en este procedimiento incluyen la PCR, la hibridación sustractiva y la determinación de la secuencia de ADN.

5 El procedimiento de hibridación también puede producirse con uno de los ácidos nucleicos complementarios inmovilizados sobre una matriz, tal como esferas magnéticas, esferas de Sepharose o cualquier otra resina. Las herramientas de la biología molecular que se basan en este procedimiento incluyen el aislamiento de ARNm poli(A+). El procedimiento de hibridación también puede producirse con uno de los ácidos nucleicos complementarios inmovilizados sobre un soporte sólido, tal como una membrana de nitrocelulosa o nailon, o
 10 inmovilizado, por ejemplo, mediante fotolitografía, por ejemplo, sobre un soporte de vidrio silíceo (esto último se conoce como matrices o micromatrices de ácidos nucleicos o chips de ácidos nucleicos). Las herramientas de la biología molecular que se basan en este procedimiento incluyen análisis de transferencia en gel de ARN y ADN, hibridación de colonias, hibridación en placas e hibridación de micromatrices. Para permitir que se produzca la hibridación, las moléculas de ácidos nucleicos en general se desnaturalizan de modo térmico o químico (por
 15 ejemplo, con NaOH) para disolver una doble cadena en dos cadenas sencillas y/o para eliminar horquillas u otras estructuras secundarias de ácidos nucleicos monocatenarios. La rigurosidad de la hibridación se ve influida por condiciones, tales como la temperatura, la concentración salina y la composición del tampón de hibridación. Unas condiciones de alta rigurosidad para la hibridación incluyen una alta temperatura y/o una baja concentración salina (las sales incluyen NaCl y Na₃-citrato) y/o la inclusión de formamida en el tampón de hibridación y/o la disminución de la concentración de compuestos, tales como SDS (detergente) en el tampón de hibridación y/o la exclusión de compuestos, tales como sulfato de dextrano o polietilenglicol (que estimulan el agrupamiento molecular) del tampón de hibridación. Las condiciones de hibridación convencionales se describen, por ejemplo, en Sambrook *et al.* (1989), pero los expertos en la técnica apreciarán que pueden diseñarse numerosas condiciones de hibridación
 20 diferentes en función de la homología y/o la longitud conocidas o esperadas de la secuencia del ácido nucleico. Unas condiciones de hibridación de rigurosidad suficientemente baja son particularmente preferidas para aislar ácidos nucleicos heterólogos con las secuencias de ADN de la memoria descriptiva definidas anteriormente. Los elementos que contribuyen a la heterología incluyen alelismo, degeneración del código genético y diferencias en la utilización de codones preferidos, según se analizó anteriormente.

30 La expresión "que se hibrida específicamente" se refiere a la unión, la formación de dúplex o la hibridación de una molécula con una secuencia de nucleótidos particular bajo condiciones de intermedias a rigurosas cuando esta secuencia se presenta en una mezcla compleja, por ejemplo, ARN o ADN celular total.

Unas "condiciones de hibridación rigurosas" y unas "condiciones de lavado de hibridación rigurosas", en el contexto de los experimentos de hibridación de ácidos nucleicos, tales como hibridaciones Southern y Northern, son dependientes de la secuencia y son diferentes bajo diferentes parámetros ambientales. Por ejemplo, las secuencias
 35 más largas se hibridan específicamente a temperaturas mayores. La T_m es la temperatura en la cual, con una fuerza iónica y un pH definidos, 50% de la secuencia diana se hibrida con una sonda perfectamente apareada. La especificidad generalmente es una función de los lavados tras la hibridación. Los factores críticos de estos lavados incluyen la fuerza iónica y la temperatura de la disolución de lavado final.

40 En general, las condiciones rigurosas se seleccionan para que sean aproximadamente 50 °C menores que el punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y un pH definidos. La T_m es la temperatura (bajo una fuerza iónica y un pH definidos) en la cual 50% de la secuencia diana se hibrida con una sonda perfectamente apareada. La T_m depende de las condiciones de la disolución y la composición de bases de la sonda, y puede calcularse utilizando la siguiente ecuación:

$$\begin{aligned}
 T_m = 79,8 \text{ }^\circ\text{C} & \quad + (18,5 \times \text{Log}[\text{Na}^+]) \\
 & \quad + (58,4 \text{ }^\circ\text{C} \times \%[\text{G}+\text{C}]) \\
 & \quad - (820/n.^{\circ} \text{ de pb en dúplex}) \\
 & \quad - (0,5 \times \% \text{ de formamida})
 \end{aligned}$$

45 Las condiciones de rigurosidad más preferidas son cuando la temperatura es 20 °C por debajo de T_m, y las condiciones rigurosas más preferidas son cuando la temperatura es 10 °C por debajo de T_m. La unión no específica también puede controlarse utilizando una cualquiera de una serie de técnicas conocidas tales como, por ejemplo, bloquear la membrana con disoluciones que contienen proteínas, la adición de ADN, ARN heterólogo y SDS al
 50 tampón de hibridación, y un tratamiento con ARNasa.

Las condiciones de lavado se realizan generalmente en la rigurosidad o por debajo. En general, unas condiciones rigurosas adecuadas para los ensayos de hibridación de ácidos nucleicos o los procedimientos de detección de la amplificación de genes son como se indican a continuación. También pueden seleccionarse condiciones más o menos rigurosas.

5 Para el objetivo de definir el nivel de rigurosidad, es posible referirse a Sambrook, J., E.F. Fritsch, *et al.*, 1989, Molecular Cloning: a Laboratory Manual, 2ª edición, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press. Un ejemplo de condiciones de baja rigurosidad es 4-6x SSC/SDS al 0,1-0,5% en p/v a 37^o-45 °C durante 2-3 horas. Dependiendo de la fuente y la concentración del ácido nucleico implicado en la hibridación, pueden emplearse otras condiciones de rigurosidad, tales como condiciones de rigurosidad intermedia. Los ejemplos de condiciones de rigurosidad intermedia incluyen 1-4x SSC/SDS al 0,25% en p/v SDS a \geq 45 °C durante 2-3 horas. Un ejemplo de condiciones de alta rigurosidad incluye 0,1-1x SSC/SDS al 0,1% en p/v SDS a 60 °C durante 1-3 horas. Los expertos en la técnica son conscientes de los diversos parámetros que pueden alterarse durante la hibridación y el lavado y que mantendrán o cambiarán las condiciones de rigurosidad. Por ejemplo, otra condición de hibridación rigurosa es una hibridación a 4x SSC a 65 °C, seguido de un lavado en 0,1x SSC a 65 °C durante aproximadamente una hora. Como alternativa, un ejemplo de una condición de hibridación rigurosas es en formamida al 50%, 4x SSC, a 42 °C. Otro ejemplo de condiciones rigurosas incluye una hibridación a 62 °C in 6x SSC, 0,05x BLOTTO, y un lavado a 2x SSC, SDS al 0,1% a 62 °C.

20 Claramente, la presente memoria descriptiva incluye el uso de secuencias de ADN que se hibridan con secuencias de ADN que codifican una citoquinina oxidasa, sus homólogos, derivados, o fragmentos inmunológicamente activos y/o funcionales, según se define a continuación, en procedimientos para aumentar el rendimiento de semillas, que comprende la expresión de estas secuencias de ADN en la aleurona y/o el embrión de una semilla. Preferiblemente, la citoquinina oxidasa es una citoquinina oxidasa vegetal, más en concreto (At)CKX de *Arabidopsis thaliana*.

Los "homólogos" de un polipéptido de CKX también pueden ser útiles en los procedimientos de esta memoria descriptiva.

25 Los "homólogos" de una CKX útiles en los procedimientos de la memoria descriptiva son los péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que contienen sustituciones, deleciones y/o adiciones de aminoácidos con relación a la proteína con respecto a la cual son homólogos, sin alterar una o más de sus propiedades funcionales, en particular sin reducir la actividad de la proteína resultante. Por ejemplo, un homólogo de una CKX consistirá en un variante de secuencia de aminoácidos bioactiva de esta CKX.

30 Dos formas especiales de homología, la homología ortóloga y paróloga, son conceptos evolutivos utilizados para describir relaciones ancestrales de genes. El término "parálogo" se refiere a genes homólogos que resultan de una o más duplicaciones de genes dentro del genoma de una especie. El término "ortólogo" se refiere a genes homólogos en diferentes organismos debidos a la relación ancestral de estos genes. El término "homólogos", tal como se emplea en la presente, también incluye los parálogos y los ortólogos de las proteínas útiles en los procedimientos según la memoria descriptiva. Los genes ortólogos pueden identificarse haciendo una búsqueda en una o más bases de datos de genes con un gen de interés utilizando, por ejemplo, el programa BLAST. Los genes en cuestión con mayor puntuación que resultan de la búsqueda de nuevo se someten a un análisis BLAST, y solo los genes en cuestión que vuelven a corresponderse con el gen utilizado en la búsqueda son conservados como genes ortólogos verdaderos. Por ejemplo, para encontrar un ortólogo del arroz de un gen de *Arabidopsis thaliana*, se puede realizar un análisis BLASTN o TBLASTX sobre una base de datos del arroz (tal como, pero sin limitarse a la base de datos Nipponbare de *Oriza sativa* disponible en NCBI (<http://www.ncbi.nlm.nih.gov>) o las secuencias genómicas del arroz (cultivares indica o japónica)). En la siguiente etapa, las secuencias del arroz obtenidas se emplean en un análisis BLAST inverso utilizando una base de datos de *Arabidopsis*. Los resultados pueden refinarse aún más si las secuencias resultantes se analizan con ClustalW y se visualizan en un árbol de reunión de vecinos. El procedimiento puede utilizarse para identificar ortólogos de muchas especies diferentes.

35 BLAST ("Basic Local Alignment Search Tool", herramienta de búsqueda de alineamientos locales básica) es una familia de programas (<http://www.ncbi.nlm.nih.gov/BLAST/>) cuyo objetivo es identificar regiones de alineamiento local óptimo, es decir, el alineamiento de algunas porciones de dos secuencias de ácidos nucleicos o proteínas, y detectar relaciones entre secuencias que comparten solo regiones aisladas de similitud (Altschul *et al.*, Nucleic Acids Res., 25: 3389-3402 (1997)).

55 Para producir estos homólogos, los aminoácidos presentes en la proteína pueden reemplazarse por otros aminoácidos que tengan propiedades similares, por ejemplo, hidrofobicidad, hidrofiliidad, momento hidrófobo, antigenicidad, propensión a formar o romper estructura de α -hélice o lámina β , etc. En la tabla 2 se ofrece una visión general de las propiedades físicas y químicas de aminoácidos.

55

Tabla 2. Propiedades de aminoácidos naturales

Propiedades de carga/hidrofobicidad	Grupo lateral	Aminoácido
Hidrófobo no polar	alifático	ala, ile, leu, val
	alifático, que contiene S	met
	aromático	phe, trp
	imino	pro
Polar no cargado	alifático	gly
	amida	asn, gln
	aromático	tyr
	hidroxilo	ser, thr
	sulfhidrilo	cys
Con carga positiva	básico	arg, his, lys
Con carga negativa	ácido	asp, glu

5 Los homólogos que son útiles en los procedimientos según la memoria descriptiva preferiblemente tienen actividad citoquinina oxidasa/deshidrogenasa y comprenden al menos 2 secuencias de 17 y 19 restos aminoácidos consecutivos, respectivamente, con una secuencia consenso como se muestra a continuación:

Secuencia consenso 1 (17 aminoácidos): hTDYLhhoLGGTLSSsG (SEQ ID NO:44)

Secuencia consenso 2 (19 aminoácidos): cLFxushGsLGQFGIlstA (SEQ ID NO:45)

10 en las que las letras en mayúscula son los códigos IUPAC de una sola letra convencionales para los diversos aminoácidos, y las otras letras simbolizan la naturaleza de los aminoácidos, tal como se muestran en la anterior tabla 2. La columna de la derecha lista, para cada clase, los aminoácidos concretos que se permiten en las secuencias consenso.

15 Los variantes de sustitución de una proteína según esta memoria descriptiva son los variantes en los que al menos un resto en la secuencia de la proteína o de aminoácidos se ha retirado y se ha insertado un resto diferente en su lugar. Las sustituciones de aminoácidos generalmente son de un solo resto, pero pueden agruparse dependiendo de las restricciones funcionales depositadas sobre el polipéptido; las inserciones habitualmente serán del orden de aproximadamente 1-10 restos aminoácidos, y las deleciones variarán de aproximadamente 1-20 restos. Preferiblemente, las sustituciones de aminoácidos comprenderán sustituciones de aminoácidos conservativas, tales como las descritas anteriormente.

20 Los variantes de inserción de una secuencia de aminoácidos de una proteína según esta memoria descriptiva son los variantes en los que uno o más restos aminoácidos se introducen en un sitio predeterminado en la proteína. Las inserciones pueden comprender fusiones amino-terminales y/o carboxi-terminales, así como inserciones intrasecuenciales de un único o de múltiples aminoácidos. En general, las inserciones dentro de la secuencia de aminoácidos serán más pequeñas que las fusiones amino- o carboxilo-terminales, del orden de aproximadamente 1 a 10 restos. Los ejemplos de proteínas o péptidos de fusión amino- o carboxi-terminal incluyen el dominio de unión o el dominio de activación de un activador transcripcional, según se emplean en un sistema de dos híbridos, 25 proteínas de la cubierta del fago, marcador de (histidina)₆, glutatión S-transferasa, proteína A, proteína de unión a maltosa, dihidrofolato reductasa, epitopo Tag•100 (EETARFQPGYRS), epitopo c-myc (EQKLISEEDL), epitopo FLAG® (DYKDDDK), lacZ, CMP (péptido de unión a calmodulina), epitopo HA (YPYDVPDYA), epitopo de proteína C (EDQVDPRLIDGK) y epitopo VSV (YTDIEMNRLGK).

30 Los variantes de deleción de una proteína según esta memoria descriptiva se caracterizan por la eliminación de uno o más aminoácidos de la secuencia de aminoácidos de la proteína.

Los variantes de aminoácidos de una proteína según esta memoria descriptiva pueden prepararse con facilidad utilizando técnicas sintéticas de péptidos muy conocidas en la técnica, tales como la síntesis de péptidos en fase

sólida y similares, o mediante manipulaciones de ADN recombinante. La manipulación de secuencias de ADN para producir proteínas variantes que se manifiestan como variantes de sustitución, inserción o delección es muy conocida en la técnica. Por ejemplo, las técnicas para realizar mutaciones de sustitución en sitios predeterminados en el ADN con una secuencia conocida son muy conocidas por los expertos en la técnica, tales como mediante mutagénesis de M13, kit de mutagénesis *in vitro* T7-Gen (USB, Cleveland, OH), kit de mutagénesis dirigida a sitio QuickChange (Stratagene, San Diego, CA), mutagénesis dirigida a sitio mediada por PCR, u otros protocolos de mutagénesis dirigida a sitio.

Los derivados de una proteína de CKX también pueden ser útiles en los procedimientos de la presente memoria descriptiva. Los “derivados” de una proteína, según esta memoria descriptiva, son los péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que comprenden al menos aproximadamente cinco restos aminoácidos contiguos en el polipéptido, pero que conservan la actividad biológica de esta proteína. Un “derivado” también puede comprender otros restos aminoácidos naturales, alterados por glicosilación o acilación, o no naturales, comparado con la secuencia de aminoácidos de una forma natural del polipéptido. Como alternativa, o además, un derivado puede comprender uno o más sustituyentes que no son aminoácidos, comparado con la secuencia de aminoácidos de una forma natural del polipéptido, por ejemplo, una molécula indicadora u otro ligando, unido de forma covalente o no covalente a la secuencia de aminoácidos tal como, por ejemplo, una molécula indicadora que está unida a esta para facilitar su detección.

“Inmunológicamente activo” significa que una molécula, o uno de sus fragmentos específicos, tales como epitopos específicos o haptenos, es reconocida, por ejemplo, mediante su unión a anticuerpos. Los epitopos específicos pueden determinarse utilizando, por ejemplo, técnicas de barrido de péptidos tales como las descritas en Geysen *et al.* (1996) (Geysen, H.M., Rodda, S.J. y Mason, T.J. (1986), A priori delineation of a peptide which mimics a discontinuous antigenic determinant, *Mol. Immunol.*, 23, 709-715).

La expresión “fragmento de una secuencia” o “parte de una secuencia” significa una secuencia truncada de la secuencia original indicada. La secuencia truncada (secuencia de ácido nucleico o proteína) puede variar mucho en longitud; el tamaño mínimo es una secuencia con un tamaño suficiente como para proporcionar una secuencia con al menos una función y/o actividad comparable con la secuencia original indicada (por ejemplo, “fragmento funcional”), mientras que el tamaño máximo no es crítico. En algunas aplicaciones, el tamaño máximo habitualmente no es sustancialmente mayor que el requerido para proporcionar la actividad y/o función o funciones deseadas de la secuencia original. Generalmente, la secuencia de aminoácidos truncada variará de aproximadamente 5 a aproximadamente 60 aminoácidos de longitud. Sin embargo, más generalmente, la secuencia tendrá un máximo de aproximadamente 50 aminoácidos de longitud, preferiblemente un máximo de aproximadamente 60 aminoácidos. Habitualmente, resulta deseable seleccionar secuencias de al menos aproximadamente 10, 12 o 15 aminoácidos, hasta un máximo de aproximadamente 20 o 25 aminoácidos. Los fragmentos funcionales comprenden al menos 50 aminoácidos, e incluyen un dominio de unión a FAD (según se define en Pfam (versión 14.0, junio, 2004), n.º de registro 1565, Bateman *et al.*, *Nucleic Acids Research Database Issue*, 32, D138-D141, 2004) y muestran actividad citoquinina oxidasa/deshidrogenasa. Lo más preferiblemente, los fragmentos funcionales muestran actividad citoquinina oxidasa/deshidrogenasa y comprenden una secuencia de aminoácidos que se corresponde con la secuencia que abarca de Leu 94 a Leu 461 de la proteína codificada por SEQ ID NO:26. Los fragmentos funcionales también pueden incluir los que comprenden un epitopo que es específico para las proteínas de CKX.

Así, debe entenderse que los fragmentos funcionales también pueden ser fragmentos inmunológicamente activos.

En los procedimientos de la memoria descriptiva, se contempla el uso de homólogos, derivados y/o fragmentos inmunológicamente activos y/o funcionales de las citoquinina oxidasas, según se definió anteriormente. Los homólogos, derivados y/o fragmentos inmunológicamente activos y/o funcionales de las proteínas de citoquinina oxidasa particularmente preferidos que se contemplan para su uso en los procedimientos de la presente memoria descriptiva se derivan de plantas, más específicamente de *Arabidopsis thaliana*, aún más específicamente las citoquinina oxidasas son (At)CKX de *Arabidopsis thaliana*, o son capaces de ser expresados en la aleurona y/o el embrión de una semilla de cereales, preferiblemente de arroz o maíz. La presente memoria descriptiva contempla claramente el uso de homólogos funcionales o derivados y/o fragmentos inmunológicamente activos de las proteínas de AtCKX y no se limita en su aplicación al uso de una secuencia de nucleótidos que codifica una de estas proteínas de AtCKX en los procedimientos de la presente memoria descriptiva.

Para realizar la expresión de una proteína de CKX en una célula, tejido u órgano, preferiblemente de origen vegetal, la proteína puede introducirse directamente en la célula, tal como mediante microinyección o medios balísticos o, como alternativa, una molécula de ácido nucleico aislada que codifica la proteína de CKX puede introducirse en la célula, tejido u órgano en un formato expresable.

En el contexto de la presente memoria descriptiva debe entenderse que los términos “expresión” y/o “sobreexpresión” se emplean de modo intercambiable y que ambos se relacionan con una “expresión ectópica y/o

potenciada" de una citoquinina oxidasa vegetal. Debe resultar evidente que esto significa una expresión potenciada de la citoquinina oxidasa vegetal, así como la expresión "de novo" de citoquinina oxidasas vegetales. Los métodos para aumentar la expresión de genes o proteínas están bien documentados en la técnica e incluyen, por ejemplo, la sobreexpresión dirigida por promotores apropiados, y el uso de potenciadores de la transcripción o potenciadores de la traducción. Pueden introducirse ácidos nucleicos aislados que actúen como elementos promotores o potenciadores en una posición apropiada (generalmente cadena arriba) de una forma no heteróloga de un polinucleótido para sobreregular la expresión de un ácido nucleico que codifica CKX o uno de sus variantes. Por ejemplo, pueden alterarse promotores endógenos *in vivo* mediante mutación, delección y/o sustitución (véase, Kmiec, patente de EEUU n.º 5.565.350; Zarlino *et al.*, documento PCT/US93/03868), o pueden introducirse promotores aislados en una célula vegetal en la orientación y la distancia adecuadas desde un gen de la presente memoria descriptiva para controlar la expresión del gen. Los procedimientos para reducir la expresión de genes o de productos génicos están bien documentados en la técnica.

La expresión de un ácido nucleico que codifica un polipéptido de CKX, o uno de sus homólogos, también puede aumentarse en la aleurona y/o el embrión de una semilla introduciendo una modificación genética (preferiblemente en el locus de un gen CKX). El locus de un gen, tal como se define en la presente, significa una región genómica, que incluye el gen de interés y 10 kb cadena arriba o debajo de la región codificadora.

La modificación genética puede introducirse, por ejemplo, mediante uno cualquiera (o más) de los siguientes procedimientos: activación de T-ADN, TILLING, mutagénesis específica dirigida a sitio, evolución dirigida y recombinación homóloga, o introduciendo y expresando en una planta un ácido nucleico que codifica un polipéptido de CKX, o uno de sus homólogos. Después de la introducción de la modificación genética, se sigue una etapa de seleccionar la expresión modificada de un ácido nucleico que codifica un polipéptido de CKX, o uno de sus homólogos, produciendo esta modificación en la expresión unas plantas con mayor rendimiento.

El marcaje de activación de T-ADN (Hayashi *et al.*, Science (1992), 1350-1353) implica la inserción de T-ADN, que habitualmente contiene un promotor (también puede ser un potenciador de la traducción o un intrón), en la región genómica del gen de interés o 10 kb cadena arriba o abajo de la región codificadora de un gen, en una configuración tal que el promotor dirige la expresión del gen implicado. Generalmente, la regulación de la expresión del gen implicado por su promotor natural está alterada y el gen cae bajo el control del promotor recién introducido. El promotor generalmente está inmerso en un T-ADN. Este T-ADN se inserta aleatoriamente en el genoma de la planta, por ejemplo, a través de una infección con *Agrobacterium* y conduce a la sobreexpresión de genes cerca del T-ADN insertado. Las plantas transgénicas resultantes muestran fenotipos dominantes debido a la sobreexpresión de genes cercanos al promotor introducido. Según la presente memoria descriptiva, el promotor que se va a introducir es un promotor capaz de conducir la expresión en la aleurona y/o el embrión de una semilla.

También puede introducirse una modificación genética en el locus de un gen CKX utilizando la técnica de TILLING ("Targeted Induced Local Lesions in Genomes", lesiones locales inducidas y dirigidas en genomas). Esta es una tecnología de mutagénesis útil para generar y/o identificar y, en último término, aislar variantes mutagenizadas de un ácido nucleico CKX capaz de mostrar actividad CKX. El TILLING también permite la selección de plantas que portan dichos variantes mutantes. Estos variantes mutantes pueden mostrar una actividad CKX incluso más alta que la que muestra en el gen en su forma natural. El TILLING combina una mutagénesis de alta densidad con procedimientos de búsqueda de alta capacidad de procesamiento. Las etapas que generalmente se siguen en el TILLING son: (a) mutagénesis EMS (Redei, G.P. y Koncz, C. (1992), en *Methods in Arabidopsis Research*, Koncz C., Chua N.H., Schell J., eds., Singapur, World Scientific Publishing Co, pp. 16-82; Feldmann *et al.* (1994), en Meyerowitz E.M., Somerville C.R., eds., *Arabidopsis*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 137-172; Lightner, J. y Caspar, T. (1998), en J. Martínez-Zapater, J. Salinas, eds., *Methods on Molecular Biology*, vol. 82. Humana Press, Totowa, NJ, pp. 91-104); (b) preparación del ADN y agrupamiento de los individuos; (c) amplificación con PCR de una región de interés; (d) desnaturalización y reasociado para permitir la formación de heterodúplex; (e) DHPLC, en el que la presencia de un heterodúplex en un agrupamiento se detecta como un pico adicional en el cromatograma; (f) identificación del mutante individual; y (g) secuenciación del producto de PCR mutante. Los procedimientos para el TILLING son muy conocidos en la técnica (McCallum *et al.* (2000), *Nat. Biotechnol.*, 18: 455-457; reseñado por Stemple (2004), *Nat. Rev. Genet.*, 5(2):145-150).

Puede utilizarse la mutagénesis específica dirigida a sitio para generar variantes de ácidos nucleicos de CKX. Están disponibles varios procedimientos para lograr la mutagénesis específica dirigida a sitio, siendo los más habituales los procedimientos basados en la PCR (Current Protocols in Molecular Biology, Wiley eds., <http://www.4ulr.com/products/currentprotocols/index.html>).

También puede utilizarse la evolución dirigida para generar variantes de ácidos nucleicos de CKX. Estos consisten en iteraciones de barajado de ADN, seguido de la búsqueda y/o selección apropiada para generar variantes de ácidos nucleicos de CKX o porciones de estos que codifican polipéptidos de CKX o sus homólogos o porciones que tienen una actividad biológica modificada (Castle *et al.* (2004), *Science*, 304(5674):1151-1154; patentes de EEUU

5.811.238 y 6.395.547).

La activación de T-ADN, el TILLING, la mutagénesis específica dirigida a sitio y la evolución dirigida son ejemplos de tecnologías que permiten la generación de nuevos alelos y variantes de CKX.

5 La recombinación homóloga permite la introducción en un genoma de un ácido nucleico seleccionado en una posición seleccionada definida. La recombinación homóloga es una tecnología convencional que se emplea de modo habitual en las ciencias biológicas para organismos inferiores, tales como levaduras o el musgo *Physcomitrella*. Los procedimientos para realizar la recombinación homóloga en plantas se han descrito no solo para plantas modelo (Offringa *et al.* (1990), EMBO J., 9(10):3077-3084), sino también para plantas de cultivo, por ejemplo, arroz (Terada *et al.* (2002), Nat. Biotech., 20(10):1030-1034; lida y Terada (2004), Curr. Opin. Biotech., 15(2):132-138). No es necesario que el ácido nucleico que se va a dirigir (que puede ser un ácido nucleico de CKX o uno de sus variantes, según se definió anteriormente en la presente) se dirija al locus de un gen CKX, sino que puede introducirse, por ejemplo, en regiones de alta expresión. El ácido nucleico que se va a dirigir puede ser un alelo mejorado utilizado para reemplazar al gen endógeno o puede introducirse además del gen endógeno.

15 La memoria descriptiva también se refiere a un procedimiento para la producción de plantas modificadas que tienen mayor rendimiento de semillas, que comprende la introducción de una modificación genética, dando como resultado dicha modificación genética una mayor expresión de citoquinina oxidasa en la aleurona y/o el embrión de una semilla.

20 La memoria descriptiva se refiere también a un procedimiento para aumentar el rendimiento de semillas en una planta que comprende la introducción de una molécula de ácido nucleico que codifica una CKX, según se definió anteriormente, unida operablemente a una o más secuencias de control, o un vector integrado de forma estable en el genoma de una célula vegetal capaz de dirigir la expresión en la aleurona y/o el embrión de una semilla.

25 Por tanto, la memoria descriptiva también se refiere a un vector que comprende un ácido nucleico que codifica una CKX, según se definió anteriormente, en el que el vector es un vector de expresión, y en el que el ácido nucleico que codifica una CKX está unido operablemente a una o más secuencias de control que permiten la expresión en la aleurona y/o el embrión de una semilla.

En la presente memoria descriptiva los inventores han demostrado, de forma sorprendente, que la expresión de citoquinina oxidasas en la aleurona y/o el embrión de una semilla da como resultado las características relacionadas con la semilla mencionadas anteriormente. Los ejemplos de promotores específicos de semillas incluyen, pero no se limitan a los listados en la tabla 3.

30 Tabla 3. Ejemplos de promotores expresables en plantas capaces de dirigir la expresión en la semilla

Fuente del gen	Patrón de expresión	Referencia bibliográfica
α -amilasa (<i>Amy32b</i>)	aleurona	Lanahan, M.B., <i>et al.</i> , Plant Cell, 4:203-211, 1992; Skriver, K., <i>et al.</i> , Proc. Natl. Acad. Sci. (USA), 88:7266-7270, 1991
gen similar a catepsina β	aleurona	Cejudo, F.J., <i>et al.</i> , Plant Molecular Biology, 20:849-856, 1992
genes específicos de semillas	semilla	Simon, <i>et al.</i> , Plant Mol. Biol., 5:191, 1985; Scofield, <i>et al.</i> , J. Biol. Chem., 262:12202, 1987; Baszczyński, <i>et al.</i> , Plant Mol. Biol., 14:633, 1990
albúmina de la nuez del Brasil	semilla	Pearson, <i>et al.</i> , Plant Mol. Biol., 18:235-245, 1992
legumina	semilla	Ellis, <i>et al.</i> , Plant Mol. Biol., 10:203-214, 1988
glutelina (arroz)	semilla	Takaiwa, <i>et al.</i> , Mol. Gen. Genet., 208:15-22, 1986; Takaiwa, <i>et al.</i> , FEBS Letts., 221:43-47, 1987
zeína	semilla	Matzke <i>et al.</i> , Plant Mol. Biol., 14(3):323-332, 1990
NapA	semilla	Stalberg, <i>et al.</i> , Planta, 199:515-519, 1996
glutenina 1 de bajo peso molecular	endospermo	Mol. Gen. Genet., 216:81-90, 1989; NAR, 17:461-

y de alto peso molecular de trigo		2, 1989
SPA de trigo	semilla	Albani <i>et al.</i> , Plant Cell, 9:171-184, 1997
α , ρ , γ -gliadinas de trigo	endospermo	EMBO, 3:1409-1415, 1984
promotor <i>ltr1</i> de cebada	endospermo	
B1, C, D, hordeína de cebada	endospermo	Theor. Appl. Gen., 98:1253-1262, 1999; Plant J., 4:343-355, 1993; Mol. Gen. Genet., 250:750-760, 1996
DOF de cebada	endospermo	Mena <i>et al.</i> , The Plant Journal, 116(1):53-62, 1998
<i>blz2</i>	endospermo	documento EP99106056.7
promotor sintético	endospermo	Vicente-Carbajosa <i>et al.</i> , Plant J., 13:629-640, 1998
prolamina NRP33 de arroz	endospermo	Wu <i>et al.</i> , Plant Cell Physiology, 39(8):885-889, 1998
α -globulina Glb-1 de arroz	endospermo	Wu <i>et al.</i> , Plant Cell Physiology, 39(8):885-889, 1998
OSH1 de arroz	embrión	Sato <i>et al.</i> , Proc. Natl. Acad. Sci. USA, 93:8117-8122, 1996
α -globulina REB/OHP-1 de arroz	endospermo	Nakase <i>et al.</i> , Plant Mol. Biol., 33:513-522, 1997
ADP-glucosa PP de arroz	endospermo	Trans. Res., 6:157-168, 1997
familia de genes ESR de maíz	endospermo	Plant J., 12:235-246, 1997
γ -karifina de sorgo	endospermo	PMB, 32:1029-1035, 1996
KNOX	embrión	Postma-Haarsma <i>et al.</i> , Plant Mol. Biol., 39:257-271, 1999
oleosina de arroz	embrión y aleurona	Wu <i>et al.</i> , J. Biochem., 123:386, 1998
oleosina de girasol	semilla (embrión y semilla seca)	Cummins, <i>et al.</i> , Plant Mol. Biol., 19:873-876, 1992

Otros ejemplos de promotores adecuados para la expresión específica de semillas en una planta pueden encontrarse en la sección de ejemplos, tabla 4.

5 Según la presente memoria descriptiva, se proporcionan procedimientos y composiciones para aumentar el rendimiento de semillas en una planta. El rendimiento de semillas puede aumentar aumentando la expresión de un gen de citoquinina oxidasa en el embrión y/o la aleurona de la semilla de una planta. Así, puede utilizarse un promotor preferido en el embrión y/o la aleurona para dirigir la expresión de una citoquinina oxidasa en estos componentes concretos de una semilla.

10 Según la presente memoria descriptiva, un gen de citoquinina oxidasa puede colocarse en una construcción genética, tal como un vector, bajo el control de un promotor preferido en el embrión y/o la aleurona. Un ejemplo de dicho promotor capaz de dirigir la expresión en el embrión y/o la aleurona es la secuencia representada en GenBank con el n.º de registro AF019212 (secuencia desde el nucleótido 1 hasta 1256, en lo sucesivo denominada PRO0218).

15 Por ejemplo, el gen AtCKX2 (SEQ ID NO:42) puede colocarse bajo el control de un promotor capaz de dirigir la expresión en la semilla, más en concreto un promotor capaz de dirigir la expresión en el embrión y/o la aleurona, por ejemplo, PRO0218.

Por tanto, se proporciona un vector que comprende la molécula de ácido nucleico aislada que tiene la secuencia

indicada en SEQ ID NO:41 o 43.

Estas construcciones después pueden utilizarse para transformar plantas, dicotiledóneas o monocotiledóneas. Según la presente memoria descriptiva, las plantas transformadas con un vector según la presente memoria descriptiva tienen un mayor rendimiento de semillas, cuando se comparan con plantas control nulicigóticas.

- 5 Preferiblemente, el vector según esta memoria descriptiva comprende una secuencia codificadora o marco de lectura abierto (ORF) que codifica una proteína de citoquinina oxidasa, o uno de sus homólogos o derivados, o uno de sus fragmentos inmunológicamente activos y/o funcionales, según se definió anteriormente. Preferiblemente, la citoquinina oxidasa es una citoquinina oxidasa vegetal, y más específicamente una (At)CKX de *Arabidopsis thaliana*. Lo más preferiblemente, la CKX es como la representada en una de SEQ ID NO:4, 36 o 38.
- 10 Un “vector” o “secuencia de vector” o “construcción genética” significa una secuencia de ADN que puede introducirse en un organismo mediante transformación y puede mantenerse de forma estable en este organismo. El mantenimiento del vector es posible, por ejemplo, en cultivos de *Escherichia coli*, *A. tumefaciens*, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*. Otros vectores, tales como fágmidos y vectores cósmidos, pueden mantenerse y multiplicarse en bacterias y/o virus. Las secuencias de vectores generalmente comprenden un conjunto de sitios exclusivos reconocidos por enzimas de restricción, el sitio de clonación múltiple (MCS), en el que una o más secuencias que no son del vector pueden insertarse. Como alternativa a los sitios de clonación múltiple, el vector también puede comprender sitios recombinantes. La clonación de genes a través de la recombinación es muy conocida en la técnica.
- 15 Por consiguiente, una “secuencia que no es del vector” significa una secuencia de ADN que está integrada en uno o más sitios del MCS comprendido dentro del vector.
- 20 Los “vectores de expresión” forman un subconjunto de vectores que, en virtud de comprender las secuencias de control o reguladoras apropiadas, permiten la creación de un formato expresable para la secuencia o secuencias que no son del vector insertadas, permitiendo así la expresión de la proteína codificada por esta secuencia o secuencias que no son del vector. En la técnica se conocen vectores de expresión que permiten la expresión de proteínas en organismos, que incluyen bacterias (por ejemplo, *E. coli*), hongos (por ejemplo, *S. cerevisiae*, *S. pombe*, *Pichia pastoris*), células de insecto (por ejemplo, vectores de expresión de baculovirus), células animales (por ejemplo, células COS o CHO), y células vegetales (por ejemplo, vectores de expresión basados en el virus X de la patata).
- 25 Un “formato expresable” significa que la molécula de ácido nucleico aislada está en una forma adecuada para ser transcrita en ARNm y/o traducida para producir una proteína, de modo constitutivo o después de la inducción por una señal intracelular o extracelular, tal como un estímulo ambiental o estrés (mitógenos, anoxia, hipoxia, temperatura, sales, luz, deshidratación, etc.) o un compuesto químico, tal como IPTG (isopropil-β-D-tiogalactopiranosido) o tal como un antibiótico (tetraciclina, ampicilina, rifampicina, kanamicina), hormona (por ejemplo, giberelina, auxina, citoquinina, glucocorticoide, brassinoesteroide, etileno, ácido abscísico, etc.), análogo de hormona (ácido indolacético (IAA), 2,4-D, etc.), metal (cinc, cobre, hierro, etc.), o dexametasona, entre otros. Tal como conocen los expertos en la técnica, la expresión de una proteína funcional también puede requerir una o más modificaciones postraduccionales, tales como glicosilación, fosforilación, desfosforilación, o una o más interacciones de proteína-proteína, entre otras. Todos estos procedimientos se incluyen dentro del alcance de la expresión “formato expresable”.
- 30 Debe entenderse que, para la expresión en monocotiledóneas de los genes de citoquinina oxidasa según los procedimientos de la memoria descriptiva, debe utilizarse una secuencia de ácido nucleico que se corresponde con la secuencia de ADNc, para evitar el corte y empalme incorrecto de intrones en monocotiledóneas. Las secuencias de ADNc preferidas para ser expresadas en monocotiledóneas tienen una secuencia de ácido nucleico representada por cualquiera de SEQ ID NO:25 a 30, SEQ ID NO:34, SEQ ID NO:37, o SEQ ID NO:42.
- 35 La expresión constitutiva del gen de citoquinina oxidasa en plantas da como resultado un mayor crecimiento de las raíces y un menor crecimiento de brotes, lo cual ilustra la importancia de la expresión confinada del gen de citoquinina oxidasa para las propiedades del crecimiento general de la planta. El confinamiento de la actividad citoquinina oxidasa puede lograrse utilizando promotores específicos de células, tejidos u órganos, puesto que la degradación de la citoquinina es un procedimiento limitado a los tejidos o células que expresan la proteína de CKX, en contraste con las estrategias que se basan en la síntesis de hormonas.
- 50 Preferiblemente, la expresión de una proteína en una célula, tejido u órgano específico, preferiblemente de origen vegetal, se realiza introduciendo y expresando una molécula de ácido nucleico aislada que codifica la proteína, tal como una molécula de ADNc, gen genómico, molécula oligonucleotídica sintética, molécula de ARNm o marco de lectura abierto, en esta célula, tejido u órgano, en el que la molécula de ácido nucleico está colocada operablemente en conexión con secuencias de control o reguladoras adecuadas, que incluyen un promotor,
- 55

preferiblemente un promotor expresable en plantas, y una secuencia terminadora. En particular, y según los procedimientos de la presente memoria descriptiva, una secuencia de ácido nucleico que codifica una CKX se une operablemente a un promotor capaz de dirigir la expresión en la aleurona y/o el embrión de una semilla.

5 La referencia en la presente a un "promotor" debe considerarse en su contexto más amplio e incluye las secuencias reguladoras de la transcripción derivadas de un gen genómico eucariota clásico, que incluye la caja TATA que se necesita para un inicio de la transcripción preciso, con o sin una secuencia de caja CCAAT y otros elementos de control o reguladores (es decir, secuencias activadoras cadena arriba, potenciadores y silenciadores) que alteran la expresión génica en respuesta a estímulos del desarrollo y/o externos, o de una manera específica de tejido.

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

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

15 Los promotores pueden contener copias adicionales de uno o más elementos reguladores específicos, para potenciar aún más la expresión y/o para alterar la expresión espacial y/o la expresión temporal de una molécula de ácido nucleico a la cual están unidos operablemente. Estos elementos reguladores pueden colocarse adyacentes a una secuencia de promotor heterólogo para dirigir la expresión de una molécula de ácido nucleico en respuesta, por ejemplo, a cobre, glucocorticoides, dexametasona, tetraciclina, giberelina, AMPc, ácido abscísico, auxina, heridas, etileno, jasmonato o ácido salicílico, o para conferir la expresión de una molécula de ácido nucleico a células, tejidos u órganos específicos, tales como meristemos, hojas, raíces, embrión, flores, semillas o frutos.

20 En el contexto de la presente memoria descriptiva, el promotor es una secuencia de promotor expresable en plantas, capaz de dirigir la expresión en la aleurona y/o el embrión de una semilla. "Expresable en plantas" significa que la secuencia de promotor, que incluye cualquier otro elemento regulador añadido a ella o contenido en ella, es al menos capaz de inducir, conferir, activar o potenciar la expresión en una célula, tejido u órgano vegetal, preferiblemente una célula, tejido u órgano de una planta monocotiledónea o dicotiledónea, *in casu* la aleurona y/o el embrión.

Las expresiones "operable en plantas" y "operable en una planta", cuando se emplean en la presente con respecto a una secuencia de promotor, deben considerarse equivalentes a una secuencia de promotor expresable en plantas.

30 Los promotores regulables, como parte de un sistema de expresión en plantas vírico binario, también son conocidos por los expertos en la técnica (Yadav, 1999, documento WO9922003; Yadav, 2000, documento WO0017365).

35 En el presente contexto, una "secuencia de promotor regulable" es un promotor que es capaz de conferir expresión a un gen estructural en una célula, tejido u órgano concreto, o a un grupo de células, tejidos u órganos de una planta, opcionalmente bajo condiciones específicas, aunque en general no confiere expresión a través de toda la planta bajo todas las condiciones. Por consiguiente, una secuencia de promotor regulable puede ser una secuencia de promotor que confiere expresión a un gen al que está unido operablemente en una localización concreta dentro de la planta o, como alternativa, a través de toda la planta bajo un conjunto específico de condiciones, tales como después de la inducción de la expresión del gen por un compuesto químico u otro inductor.

40 Preferiblemente, el promotor regulable utilizado en los procedimientos de la presente memoria descriptiva confiere expresión a una localización específica dentro de la planta, de forma constitutiva o después de una inducción, aunque no en la planta completa bajo cualquier circunstancia. En particular, el promotor regulable para su uso en la presente memoria descriptiva es un promotor capaz de dirigir la expresión en la aleurona y/o el embrión de una semilla. Otros tipos de promotores de este tipo son secuencias de promotor específicas de célula, secuencias de promotor específicas de tejido, secuencias de promotor específicas de órgano, secuencias de promotor de genes específicos del ciclo celular, secuencias de promotor inducible, y secuencias de promotor constitutivo que se han modificado para conferir expresión a una parte concreta de la planta en cualquier momento, tal como mediante la integración de dicho promotor constitutivo dentro de un elemento genético transponible (Ac, Ds, *Spm*, *En* u otro transposón).

45 La expresión "específico de célula" indica que la expresión se produce predominantemente en una célula o tipo de célula concreto, preferiblemente de origen vegetal, pero no necesariamente en esta célula o tipo de célula exclusivamente. La expresión "específico de tejido" indica que la expresión se produce predominantemente en un tejido o tipo de tejido concreto, preferiblemente de origen vegetal, pero no necesariamente en este tejido o tipo de tejido exclusivamente. De forma similar, la expresión "específico de órgano" indica que la expresión se produce predominantemente en un órgano concreto, preferiblemente de origen vegetal, pero no necesariamente en este órgano exclusivamente. De forma similar, la expresión "específico del ciclo celular" significa que la expresión es

predominantemente cíclica y se produce en una o más fases, no necesariamente consecutivas, del ciclo celular, aunque no necesariamente en las células, preferiblemente de origen vegetal, dentro de ciclo exclusivamente.

5 Los expertos en la técnica saben que un “promotor inducible” es un promotor cuya actividad transcripcional aumenta o es inducida en respuesta a estímulos del desarrollo, químicos, ambientales o físicos. De modo similar, los expertos en la técnica entenderán que un “promotor constitutivo” es un promotor que es transcripcionalmente activo en la mayoría, aunque no necesariamente todas, de las partes de un organismo, preferiblemente una planta, durante la mayoría, aunque no necesariamente todas, de las fases de su crecimiento y desarrollo.

10 Los expertos en la técnica pueden seleccionar con facilidad las secuencias de promotor apropiadas para su uso según la presente memoria descriptiva a partir de fuentes públicas o fácilmente disponibles, sin experimentación indebida.

15 La colocación de una molécula de ácido nucleico bajo el control regulador de una secuencia de promotor, o unirla operablemente a una secuencia de promotor, significa la colocación de la molécula de ácido nucleico de modo que la expresión es controlada por la secuencia de promotor. Un promotor se coloca de modo habitual, pero no necesariamente, cadena arriba, o en el extremo 5', y dentro de 2 kb del sitio de inicio de la transcripción de la molécula de ácido nucleico que regula. Para la construcción de combinaciones de promotor heterólogo/gen estructural, generalmente se prefiere colocar el promotor a una distancia desde el sitio de inicio de la transcripción del gen que sea aproximadamente la misma que la distancia entre ese promotor y el gen que controla en su emplazamiento natural (es decir, el gen del cual se deriva el promotor). Tal como se conoce en la técnica, pueden proveerse algunas variaciones en esta distancia sin pérdida de la función del promotor. De modo similar, el colocamiento preferido de un elemento de secuencia reguladora con respecto a un gen heterólogo que se va a colocar bajo su control se define por la colocación del elemento en su emplazamiento natural (es decir, el gen del cual se deriva). De nuevo, tal como se conoce en la técnica, también pueden producirse algunas variaciones en esta distancia.

20

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

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

35

Los expertos en la técnica conocerán otras secuencias de promotores y secuencias de terminadores que pueden resultar adecuadas para su uso para realizar los procedimientos de la presente memoria descriptiva. Estas secuencias pueden utilizarse con facilidad sin experimentación indebida.

40 En el contexto de la presente memoria descriptiva, las expresiones “expresión ectópica” o “sobrexpresión ectópica” de un gen o una proteína significan que se confieren patrones de expresión y/o niveles de expresión de este gen o proteína que normalmente no aparecen bajo condiciones naturales, y más concretamente significan una mayor expresión y/o un mayor nivel de expresión en uno o más de la aleurona y/o el embrión de una semilla.

45 Preferiblemente, la secuencia de promotor utilizada en el contexto de la presente memoria descriptiva está unida operablemente a una secuencia codificadora o un marco de lectura abierto (ORF) que codifica una proteína de citoquinina oxidasa, o uno de sus homólogos, derivados, o fragmentos inmunológicamente activos y/o funcionales, según se definió anteriormente.

50 El vector puede comprender opcionalmente un gen marcador seleccionable. Tal como se emplea en la presente, la expresión “gen marcador seleccionable” o “marcador seleccionable” o “marcador para la selección” o “marcador de búsqueda” incluye cualquier gen que confiere un fenotipo a una célula en el que se expresa para facilitar la identificación y/o la selección de células que están transfectadas o transformadas con una construcción génica según esta memoria descriptiva, o uno de sus derivados. Los genes marcadores seleccionables adecuados contemplados en la presente incluyen el gen de resistencia a ampicilina (Amp^r), el gen de resistencia a tetraciclina (Tc^r), el gen de resistencia a kanamicina (Kan^r), el gen de resistencia a fosfotricina, el gen de neomicina fosfotransferasa (*nptII*), el gen de resistencia a higromicina, el gen de p-glucuronidasa (GUS), el gen de cloranfenicol acetiltransferasa (CAT), el gen de la proteína fluorescente verde (*gfp*) (Haseloff *et al.*, 1977), y el gen

55

de luciferasa, entre otros.

La memoria descriptiva también se refiere a una célula hospedante que contiene cualquiera de las moléculas de ácidos nucleicos o vectores según la memoria descriptiva. Esta célula hospedante se elige del grupo que comprende células bacterianas, de insecto, fúngicas, animales o vegetales.

5 La memoria descriptiva también se refiere a un procedimiento para la producción de plantas, células vegetales o tejidos vegetales transgénicos, que comprende la introducción de una molécula de ácido nucleico según esta memoria descriptiva en un formato expresable o un vector según esta memoria descriptiva en esta planta, célula vegetal o tejido vegetal.

10 Por tanto, la invención proporciona un procedimiento para producir una planta que tiene un mayor número de semillas llenas y una mayor expresión de una citoquinina oxidasa en el embrión y/o la aleurona de una semilla, consistiendo dicho procedimiento en: (i) introducir en una célula vegetal una molécula de ácido nucleico aislada que codifica una citoquinina oxidasa unida operablemente a un promotor que dirige la expresión específicamente en el embrión y/o la aleurona de la semilla de una planta; y (ii) aumentar el nivel y/o la actividad de dicha citoquinina oxidasa en el embrión y/o la aleurona de la semilla de una planta, en el que dicha molécula de ácido nucleico
15 aislada que codifica una citoquinina oxidasa es un ácido nucleico que codifica la proteína representada por SEQ ID NO:36, o es una molécula de ácido nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con dicho ácido nucleico que codifica la proteína de SEQ ID NO:36, y en el que dicho promotor comprende los nucleótidos 1-1256 de SEQ ID NO:41.

20 La invención también proporciona un procedimiento para producir una planta que tiene un mayor número de semillas llenas, comprendiendo dicho procedimiento:

(a) introducir en una célula vegetal una construcción genética que comprende un ácido nucleico aislado que codifica una citoquinina oxidasa unida operablemente a un promotor que dirige la expresión específicamente en el embrión y/o la aleurona de una semilla, en el que dicha molécula de ácido nucleico aislada que codifica una citoquinina oxidasa es un ácido nucleico que codifica la proteína representada por SEQ ID NO:36, o es una molécula de ácido
25 nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con dicho ácido nucleico que codifica la proteína de SEQ ID NO:36, y en el que dicho promotor comprende los nucleótidos 1-1256 de SEQ ID NO:41;

(b) regenerar una planta a partir de esta;

(c) hacer crecer la planta regenerada hasta la fecundación de la semilla; y

30 (d) seleccionar una planta con un mayor número de semillas llenas, comparado con una correspondiente planta de tipo salvaje. En general, después de la transformación, las células vegetales o los agrupamientos de células se seleccionan para la presencia de uno o más marcadores que son codificados por genes expresables en plantas cotransferidos con el gen de interés, tras lo cual el material transformado se regenera para formar una planta completa.

35 Después de la transferencia del ADN y la regeneración, las plantas putativamente transformadas pueden evaluarse, por ejemplo, utilizando un análisis Southern, para la presencia del gen de interés, el número de copias y/o la organización genómica. Como alternativa, o además, los niveles de expresión del ADN recién introducido pueden controlarse utilizando un análisis Northern y/o Western, siendo ambas técnicas muy conocidas por los expertos en la técnica.

40 Las plantas transformadas generadas pueden propagarse mediante una diversidad de medios, tales como la propagación clónica o las técnicas de cruzamiento clásicas. Por ejemplo, una planta transformada de primera generación (T1) puede autofecundarse para producir transformantes de segunda generación (T2) homocigóticos, y las plantas T2 después pueden propagarse a través de técnicas de cruzamiento clásicas.

45 Los organismos transformados generados pueden tener una diversidad de formas. Por ejemplo, pueden ser quimeras de células transformadas y células no transformadas; transformantes clónicos (por ejemplo, todas las células se transforman para que contengan el módulo de expresión); injertos de tejidos transformados y no transformados (por ejemplo, en plantas, un rizoma transformado injertado en un vástago no transformado).

50 La memoria descriptiva también se refiere a una célula vegetal transgénica que comprende una secuencia de ácido nucleico según esta memoria descriptiva que está unida operablemente a elementos reguladores que permiten la transcripción y/o la expresión del ácido nucleico en la aleurona y/o el embrión de una semilla.

Según otra realización preferida, la memoria descriptiva se refiere a una célula vegetal transgénica, según se describió anteriormente en la presente, en la que el ácido nucleico según esta memoria descriptiva está integrado

de forma estable en el genoma de esta célula vegetal.

La memoria descriptiva también se refiere a una planta o un tejido vegetal transgénico que comprende células vegetales, según se describió anteriormente en la presente, y también una parte recolectable de dicha planta transgénica, preferiblemente seleccionada del grupo que consiste en semillas, hojas, frutos, cultivos de tallos, raíces, tubérculos, rizomas y bulbos. La presente memoria descriptiva también se refiere a productos directamente derivados de una parte recolectable de una planta transgénica según la memoria descriptiva, tales como gránulos o polvos secos, aceites, grasas y ácidos grasos, almidón o proteínas. La memoria descriptiva también se refiere a la progenie derivada de cualquiera de las plantas transgénicas según la memoria descriptiva.

Preferiblemente, se producen plantas transgénicas que expresan, en la aleurona y/o el embrión de una semilla, un ácido nucleico que codifica CKX, preferiblemente una CKX de *Arabidopsis*, lo más preferiblemente una CKX codificada por un ácido nucleico según se indica en una cualquiera de SEQ ID NO:3, 26, 37 o 42, o un ortólogo de dicho ácido nucleico. Preferiblemente, el ortólogo se deriva de una especie relacionada de la planta transgénica. Aún más preferiblemente, el ortólogo es específico (nativo o endógeno) de la especie de la planta transgénica.

Los medios para introducir ADN recombinante en células o tejidos vegetales incluyen, pero no se limitan a la transformación utilizando CaCl_2 y sus variaciones, en particular el procedimiento descrito por Hanahan (1983), la captación de ADN directa hacia protoplastos (Krens *et al.*, 1982; Paszkowski *et al.*, 1984), la captación mediada por PEG hacia protoplastos (Armstrong *et al.*, 1990), el bombardeo de micropartículas, la electroporación (Fromm *et al.*, 1985), la microinyección de ADN (Crossway *et al.*, 1986), el bombardeo de micropartículas de células o explantes de tejidos (Christou *et al.*, 1988; Sanford, 1988), la infiltración al vacío de tejido con ácidos nucleicos, o en el caso de plantas, la transferencia mediada por T-ADN desde *Agrobacterium* al tejido vegetal según se describe fundamentalmente en An *et al.* (1985), Dodds *et al.* (1985), Herrera-Estrella *et al.* (1983a, 1983b, 1985). Los procedimientos para la transformación de plantas monocotiledóneas son muy conocidos en la técnica e incluyen la transformación mediada por *Agrobacterium* (Cheng *et al.*, 1997, documento WO9748814; Hansen, 1998, documento WO9854961; Hiei *et al.*, 1994, documento WO9400977; Hiei *et al.*, 1998, documento WO9817813; Rikiishi *et al.*, 1999, documento WO9904618; Saito *et al.*, 1995, documento WO9506722), el bombardeo de microproyectiles (Adams *et al.*, 1999, documento US5969213; Bowen *et al.*, 1998, documento US5736369; Chang *et al.*, 1994, documento WO9413822; Lundquist *et al.*, 1999, documento US5874265/US5990390; Vasil y Vasil, 1995, documento US5405765; Walker *et al.*, 1999, documento US5955362), la captación de ADN (Eyal *et al.*, 1993, documento W09318168), la microinyección de células de *Agrobacterium* (von Holt, 1994, documento DE4309203) y la sonicación (Finer *et al.*, 1997, documento US5693512).

Una planta completa puede regenerarse a partir de la célula transformada o transfectada, según procedimientos muy conocidos en la técnica. Un tejido vegetal capaz de una posterior propagación clónica, por organogénesis o por embriogénesis, puede transformarse con una construcción génica según la presente memoria descriptiva y regenerarse una planta completa a partir de este. El tejido concreto elegido variará dependiendo de los sistemas de propagación clónica disponibles y que se adapten mejor a la especie concreta que se está transformando. Los ejemplos de tejidos diana incluyen discos foliares, polen, embriones, cotiledones, hipocotilos, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, capullos axilares, y meristemas radicales), y tejido meristemático introducido (por ejemplo, meristemo de cotiledones y meristemo de hipocotilos).

El término "organogénesis", tal como se emplea en la presente, significa un procedimiento mediante el cual se desarrollan brotes y raíces de una manera secuencial a partir de centros meristemáticos.

El término "embriogénesis", tal como se emplea en la presente, significa un procedimiento mediante el cual se desarrollan brotes y raíces juntos de una manera concertada (no secuencial) a partir de células somáticas o de gametos.

Preferiblemente, la planta transgénica producida según el procedimiento de la invención se transfecta o se transforma con una secuencia genética mediante cualquier medio reconocido en la técnica, tal como bombardeo de microproyectiles, microinyección, transformación mediada por *Agrobacterium* (que incluyen la transformación *in planta*), fusión de protoplastos, o electroporación, entre otros. Lo más preferiblemente, esta planta se produce mediante una transformación mediada por *Agrobacterium*.

La transformación mediada por *Agrobacterium* o transformación agrolística de plantas, levaduras, mohos u hongos filamentosos se basa en la transferencia de parte de las secuencias de vectores de transformación, denominada T-ADN, al núcleo, y la integración de este T-ADN en el genoma del eucariota.

Un "T-ADN" o ADN transferido, significa una parte del vector de transformación flanqueado por fronteras de T-ADN que, después de la activación de los genes *vir* de *Agrobacterium*, tiene mellas en las fronteras de T-ADN y se transfiere como un ADN monocatenario al núcleo de una célula eucariota.

Una “transformación agrolística” significa un procedimiento de transformación que combina características de la transformación mediada por *Agrobacterium* y del transporte de ADN biolístico. Así, un plásmido diana que contiene T-ADN se cotransporta con ADN/ARN que permite la producción en plantas de VirD1 y VirD2 con o sin VirE2 (Hansen y Chilton, 1996; Hansen *et al.*, 1997; Hansen y Chilton, 1997, documento W09712046).

5 Un “ADN extraño” significa cualquier secuencia de ADN que se introduce en el genoma del hospedante mediante técnicas recombinantes. El ADN extraño incluye, por ejemplo, una secuencia de T-ADN, o una de sus partes, tal como la secuencia de T-ADN que comprende el marcador seleccionable en un formato expresable. El ADN extraño incluye también secuencias de ADN intermedias, según se definió anteriormente.

10 La presente memoria descriptiva también incluye el uso de ácidos nucleicos que codifican CKX y el uso de polipéptidos de CKX unidos operablemente a un promotor capaz de dirigir la expresión en la aleurona y/o el embrión de una semilla. Uno de estos usos se relaciona con aumentar el rendimiento de la planta, en especial de las semillas. El rendimiento de semillas puede incluir uno o más de los siguientes: mayor número de flores por panícula, mayor peso total de las semillas, mayor número de semillas llenas, mayor peso de mil granos, y mayor índice de cosecha, cada uno con relación a las correspondientes plantas de tipo salvaje.

15 En particular, la presente memoria descriptiva proporciona el uso de ácidos nucleicos que codifican citoquinina oxidasas unidos operablemente a un promotor capaz de dirigir la sobreexpresión en la aleurona y/o el embrión con relación a otras partes de la semilla para aumentar el rendimiento de semillas de una planta. Uno de estos usos es en el que el ácido nucleico que codifica la citoquinina oxidasa está comprendido en una construcción genética que se introduce en una célula vegetal. En particular, se contempla el uso de ácidos nucleicos que codifican citoquinina oxidasas unidos operablemente a un promotor capaz de dirigir la sobreexpresión en la aleurona y/o el embrión con relación a otras partes de la semilla para aumentar el rendimiento de semillas de una planta, en el que el ácido nucleico que codifica la citoquinina oxidasa se selecciona del grupo que consiste en:

(a) ácidos nucleicos que comprenden una secuencia de ADN según una cualquiera de SEQ ID NO:42, 37, 27, 1, 3, 5, 7, 9, 11, 25, 26, 28 a 30, 33 o 34, o su complemento,

25 (b) ácidos nucleicos que comprenden las secuencias de ARN que se corresponden con una cualquiera de SEQ ID NO:42, 37, 27, 1, 3, 5, 7, 9, 11, 25, 26, 28 a 30, 33 o 34, o su complemento,

(c) ácidos nucleicos que se hibridan específicamente con una cualquiera de SEQ ID NO:42, 37, 27, 1, 3, 5, 7, 9, 11, 25, 26, 28 a 30, 33 o 34, o su complemento,

30 (d) ácidos nucleicos que codifican una proteína que comprende la secuencia de aminoácidos según una cualquiera de SEQ ID NO:2, 4, 6, 8, 10, 12, 35, 36 o 38, o su complemento,

(e) ácidos nucleicos según se define en uno cualquiera de (a) a (d) que se caracterizan porque dicho ácido nucleico es ADN, ADN genómico, ADNc, ADN sintético, o ARN en el que T es reemplazado por U,

35 (f) un ácido nucleico que está degenerado comparado con un ácido nucleico según una cualquiera de SEQ ID NO:42, 37, 27, 1, 3, 5, 7, 9, 11, 25, 26, 28 a 30, 33 o 34, o que está degenerado comparado con un ácido nucleico según se define en uno cualquiera de (a) a (e) como resultado del código genético,

(g) ácidos nucleicos que son divergentes de un ácido nucleico que codifica una proteína según cualquiera de SEQ ID NO:2, 4, 6, 8, 10, 12, 35, 36 o 38, que es divergente de un ácido nucleico según se define en uno cualquiera de (a) a (e), debido a diferencias en la utilización de codones entre los organismos,

40 (h) ácidos nucleicos que codifican una proteína según SEQ ID NO:2, 4, 6, 8, 10, 12, 35, 36 o 38, o ácidos nucleicos según se define en (a) a (e) que son divergentes debido a las diferencias entre alelos,

(i) ácidos nucleicos que codifican una proteína según cualquiera de SEQ ID NO:2, 4, 6, 8, 10, 12, 35, 36 o 38,

(j) fragmentos funcionales de ácidos nucleicos según se define en (a) a (i) que tienen la actividad biológica de una citoquinina oxidasa, y

45 (k) ácidos nucleicos que codifican una citoquinina oxidasa vegetal, que comprenden la secuencia consenso hTDYLhhoIGGTLSSsG y cLFXushGsLGQFGIIstA.

Además, se proporciona el uso de la SEQ ID NO:41 y el uso de un vector que comprende una citoquinina oxidasa según se define en los anteriores (a) a (k) para aumentar el rendimiento de semillas en una planta. En particular, se proporciona el uso de la SEQ ID NO:41 para aumentar el número de semillas llenas de una planta.

50 Los expertos en la técnica serán conscientes de que esta memoria descriptiva está sujeta a variaciones y modificaciones distintas de las que describen específicamente. Debe entenderse que esta memoria descriptiva

incluye todas estas variaciones y modificaciones. La memoria descriptiva también incluye todas las etapas, características, composiciones y compuestos mencionados o indicados en esta memoria descriptiva, de modo individual o colectivo, y cualquiera y todas las combinaciones de cualquiera o más de estas etapas o características.

5 A lo largo de esta memoria descriptiva, a menos que el contexto requiera lo contrario, la palabra “comprende” y sus variaciones, tales como “que comprende” y “comprendiendo”, implican la inclusión del número entero o etapa o grupo de números enteros o etapas mencionado, pero no la exclusión de cualquier otro número entero o etapa o grupo de números enteros o etapas.

10 Tal como se emplea en la presente, la expresión “se deriva de” indica que un número entero o un grupo de números enteros concretos se ha originado de la especie especificada, pero no se ha obtenido necesariamente de la fuente especificada directamente.

Los siguientes ejemplos se ofrecen como ilustración de la presente memoria descriptiva y no son limitantes. Los contenidos de todas las referencias incluidas en esta solicitud se incorporan por referencia en la presente como si se hubieran expuesto en su totalidad.

15 **Breve descripción de los dibujos**

Figura 1. Representación esquemática de genes de citoquinina oxidasa vegetales

Se muestran las estructuras de diferentes genes de citoquinina oxidasa aislados a partir del maíz (*ZmCKX1*, n.º de registro AF044603, Biochem. Biophys. Res. Com., 255:328-333, 1999) y *Arabidopsis* (*AtCKX1* a *AtCKX4*). Los exones se denominan “E” y están representados por recuadros grises. Los intrones están representados por recuadros blancos. También se indica el tamaño de los genes (en kb, en la parte superior de cada estructura), el número de registro de los genes (bajo los nombres) y una barra de tamaño que representa 0,5 kb.

Figura 2. Alineamiento de secuencias de aminoácidos de citoquinina oxidasas vegetales

Se alinean las secuencias de aminoácidos de citoquinina oxidasas de maíz (*ZmCKX1*) y *Arabidopsis* (*AtCKX1* a *AtCKX4*). Los restos aminoácidos idénticos se marcan con un recuadro negro, y los restos aminoácidos similares se presentan en un recuadro gris. Grupos de similitud de aminoácidos: (M,I,L,V), (F,W,Y), (G,A), (S,T), (R,K,H), (E,D), (N,Q).

Figura 3: Presentación esquemática del clon de entrada p41, que contiene CDS0427_2 dentro de los sitios AttL1 y AttL2 para la clonación Gateway® en el esqueleto pDONR201. CDS0427_2 es el código interno para la secuencia codificadora de CKX2 de *Arabidopsis thaliana* (SEQ ID NO:26). Este vector contiene también un módulo de resistencia a kanamicina bacteriano y un origen de la replicación bacteriano.

Figura 4: Vector binario p37 para la expresión en *Oryza sativa* del gen CKX2 de *Arabidopsis thaliana* bajo el control del promotor PRO0218. Este vector contiene un T-ADN derivado del plásmido Ti, limitado por una frontera a la izquierda (repetición LB, LB Ti C58) y una frontera a la derecha (repetición RB, RB Ti C58). Desde la frontera a la izquierda hasta la frontera a la derecha, este T-ADN contiene: un marcador seleccionable y un marcador de búsqueda para la selección de las plantas transformadas, cada uno bajo el control de un promotor constitutivo, el módulo terminador doble PRO0218-CDS0427_2-zeína y rbcS-deltaGA para la expresión del gen CKX2 de *Arabidopsis thaliana*. Este vector también contiene un origen de la replicación procedente de pBR322 para la replicación bacteriana y un marcador seleccionable (Spe/SmeR) para la selección bacteriana con espectinomicina y estreptomina.

Figura 5: Vector binario p35 para la expresión en *Oryza sativa* del gen CKX2 de *Arabidopsis thaliana* bajo el control del promotor PRO0090. Este vector contiene un T-ADN derivado del plásmido Ti, limitado por una frontera a la izquierda (repetición LB, LB Ti C58) y una frontera a la derecha (repetición RB, RB Ti C58). Desde la frontera a la izquierda hasta la frontera a la derecha, este T-ADN contiene: un marcador seleccionable y un marcador de búsqueda para la selección de las plantas transformadas, cada uno bajo el control de un promotor constitutivo, el módulo terminador doble PRO0090-CDS0427_2-zeína y rbcS-deltaGA para la expresión del gen CKX2 de *Arabidopsis thaliana*. Este vector también contiene un origen de la replicación procedente de pBR322 para la replicación bacteriana y un marcador seleccionable (Spe/SmeR) para la selección bacteriana con espectinomicina y estreptomina.

Ejemplos

50 **Ejemplo 1. Breve descripción de las secuencias descritas en la presente memoria descriptiva y ejemplos de promotores específicos de semillas**

SEQ ID NO:	Descripción
1	<i>AtCKX1</i> genómico
2	proteína de <i>AtCKX1</i>
3	<i>AtCKX2</i> genómico
4	proteína de <i>AtCKX2</i>
5	<i>AtCKX3</i> genómico
6	proteína de <i>AtCKX3</i>
7	<i>AtCKX4</i> genómico
8	proteína de <i>AtCKX4</i>
9	<i>AtCKX5</i> genómico (versión corta)
10	proteína de <i>AtCKX5</i> (versión corta)
11	<i>AtCKX6</i> genómico
12	proteína de <i>AtCKX6</i>
13	cebador 5' de <i>AtCKX1</i>
14	cebador 3' de <i>AtCKX1</i>
15	cebador 5' de <i>AtCKX2</i>
16	cebador 3' de <i>AtCKX2</i>
17	cebador 5' de <i>AtCKX3</i>
18	cebador 3' de <i>AtCKX3</i>
19	cebador 5' de <i>AtCKX4</i>
20	cebador 3' de <i>AtCKX4</i>
21	cebador 5' de <i>AtCKX5</i>
22	cebador 3' de <i>AtCKX5</i>
23	cebador 5' de <i>AtCKX6</i>
24	cebador 3' de <i>AtCKX6</i>
25	ADNc de <i>AtCKX1</i>
26	ADNc de <i>AtCKX2</i>
27	ADNc de <i>AtCKX3</i>
28	ADNc de <i>AtCKX4</i>
29	ADNc de <i>AtCKX5</i> (versión corta)
30	ADNc de <i>AtCKX6</i>
31	fragmento de ADNc de <i>AtCKX2</i>
32	fragmento del péptido de <i>AtCKX2</i>
33	<i>AtCKX5</i> genómico (versión larga)

ES 2 469 590 T3

34	ADNc de <i>AtCKX5</i> (versión larga)
35	proteína de <i>AtCKX5</i> (versión larga)
36	secuencia de proteína deducida CDS0427_2 de <i>AtCKX2</i>
37	variante de corte y empalme de <i>AtCKX2</i> , secuencia de ADN
38	variante de corte y empalme de <i>AtCKX2</i> , secuencia de proteína deducida
39	PRM3769 (sentido, codón de inicio en las posiciones 35 a 37)
40	PRM1526 (inverso, codón de fin complementario en las posiciones 30-32)
41	módulo de expresión con el terminador doble PRO0218-CDS0427_2-zeína y <i>rbcS-deltaGA</i>
42	ADNc de CDS0427_2 de <i>AtCKX2</i>
43	módulo de expresión con el terminador doble PRO0090-CDS0427_2-zeína y <i>rbcS-deltaGA</i>
44	firma HTDYLhholGGTLSssG
45	firma cLFxushGsLGOFGIstA

Tabla 4. Ejemplos de promotores adecuados para la expresión específica de semillas o preferida en plántulas

Nombre del gen	Expresión
metalotioneína Mte	embrión/escutelo + callo
beta-amilasa putativa	embrión/escutelo
desconocido	escutelo
inhibidor de proteinasa Rgpi9	semilla
proteína estructural	tejidos jóvenes, callo, embrión/escutelo
prolamina de 10 Kda	fuerte en el endospermo
alergeno RA2	semilla
prolamina RP7	endospermo
ML2 similar a metalotioneína	embrión/escutelo + callo
prolamina RM9	fuerte en el endospermo
prolamina RP5	fuerte en el endospermo
metionina aminopeptidasa putativa	embrión
proteína ribosómica 40S putativa	débil en el endospermo
alfa-globulina	fuerte en el endospermo
alanina aminotransferasa	aleurona, endospermo
ciclofilina 2	brote, embrión/endospermo
sacarosa sintasa SS1 (cebada)	constitutiva media, brote endospermo/aleurona
inhibidor de tripsina ITR1 (cebada)	débil en el endospermo
WSI18	embrión/aleurona

acuaporina	plántula
RAB21	embrión/aleurona
OSH1	plántula
arcelina 5A	semilla
cruciferina	semilla
albúmina 2S3	semilla
albúmina 2S2	semilla
FAE1	embrión
subunidad beta de faseolina	semilla
Lec1	embrión
gamma-zeína	semilla
proteína de transferencia de lípidos	semilla

Ejemplo 2: La expresión preferida en semillas o preferida en plántulas de un gen *CKX2* produce un aumento en el rendimiento de semillas

A) Manipulación del ADN y clonación de *AtCKX2*

5 A menos que se indique lo contrario, las técnicas de ADN recombinante se realizan según protocolos convencionales descritos en Sambrook y Russell (2001), *Molecular Cloning: a laboratory manual*, 3ª edición, Cold Spring Harbor Laboratory Press, CSH, Nueva York) o en los volúmenes 1 y 2 de Ausubel *et al.* (1994), *Current Protocols in Molecular Biology*, Current Protocols. Los materiales y los procedimientos convencionales para el trabajo molecular con plantas se describen en *Plant Molecular Biology Labfax* (1993), por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd. (Reino Unido) y Blackwell Scientific Publications (Reino Unido).

15 El gen *CKX2* de *Arabidopsis* (que se corresponde con SEQ ID NO:42) se amplificó mediante PCR utilizando como molde un banco de ADNc de plántulas de *Arabidopsis thaliana* (Invitrogen, Paisley, Reino Unido). Después de la transcripción inversa de ARN extraído de plántulas, los ADNc se clonaron en pCMV Sport 6.0. El tamaño medio de la inserción del banco fue de 1,5 kb, y el número original de clones fue de $1,59 \times 10^7$ cfu. Se determinó que la titulación original fue de $9,6 \times 10^5$ cfu/ml, y después de la primera amplificación fue de 6×10^{11} cfu/ml. Después de la extracción del plásmido se emplearon 200 ng de molde en una mezcla de PCR de 50 μ l. Se emplearon los cebadores prm3769 (SEQ ID NO:39) y prm1526 (SEQ ID NO:40), que incluyen los sitios AttB para la recombinación Gateway, para la amplificación por PCR.

20 La PCR se realizó utilizando ADN polimerasa Taq Hifi en condiciones convencionales. Se amplificó un fragmento de PCR de 1506 pb y se purificó utilizando también procedimientos convencionales. Después se realizó la primera etapa del procedimiento Gateway, la reacción de BP, durante la cual el fragmento de PCR se recombina *in vivo* con el plásmido pDONR para producir, según la terminología Gateway, un “clon de entrada”, p41 (figura 3). El pDONR se obtuvo en Invitrogen, como parte de la tecnología Gateway.

B) Construcción del vector

25 El clon de entrada p41 después se utilizó en una reacción LR con p831 o p830, ambos vectores de destino según la terminología Gateway TM, que se utiliza para la transformación del arroz.

30 El p831 contiene, como elementos funcionales dentro de las fronteras del T-ADN, un marcador seleccionable vegetal, un marcador de búsqueda y un módulo Gateway previsto para la recombinación *in vivo* de LR con la secuencia de interés que ya ha sido clonada en el vector donante. El promotor PRO0218 para la expresión preferida en el embrión y la aleurona está localizado cadena arriba de este módulo Gateway.

De forma similar, p830 contiene, como elementos funcionales dentro de las fronteras del T-ADN, un marcador seleccionable vegetal, un marcador de búsqueda y un módulo Gateway previsto para la recombinación *in vivo* de LR con la secuencia de interés que ya ha sido clonada en el vector donante. El promotor PRO0090 para la

expresión preferida en el endospermo está localizado cadena arriba de este módulo Gateway.

Después de la etapa de recombinación, los vectores de expresión resultantes p37 (que se origina de p831, figura 4) y p35 (que se origina de p830, figura 5) se transformaron en la cepa LBA4404 de *Agrobacterium* y después en plantas de *Oryza sativa*.

5 C) Transformación del arroz

Se descascarillaron semillas secas maduras del arroz Nipponbare cultivar japonesa. La esterilización se realizó incubando las semillas durante un minuto en etanol al 70%, seguido de 30 minutos en HgCl₂ al 0,2% y de 6 lavados de 15 minutos con agua destilada estéril. Las semillas estériles después se hicieron germinar sobre un medio que contenía 2,4-D (medio de inducción del callo). Después de una incubación durante 4 semanas en la oscuridad, los callos embriogénicos derivados del escutelo se extirparon y se propagaron en el mismo medio. Dos semanas después, los callos se multiplicaron o propagaron mediante subcultivo sobre el mismo medio durante 2 semanas más. Tres días antes del cocultivo, se subcultivaron trozos de callos embriogénicos sobre medio fresco para impulsar la actividad de división celular. La cepa LBA4404 de *Agrobacterium* que porta los vectores de T-ADN que comprenden un marcador de selección adecuado, se utilizó para el cocultivo. El *Agrobacterium* se cultivó durante 3 días a 28 °C sobre medio AB con los antibióticos apropiados. Las bacterias después se recogieron y se suspendieron en un medio de cocultivo líquido a una DO₆₀₀ de aproximadamente 1. La suspensión se trasladó a una placa Petri y los callos se sumergieron en la suspensión durante 15 minutos. Después, los tejidos del callo se secaron sobre un papel de filtro, se trasladaron a un medio de cocultivo solidificado y se incubaron durante 3 días en la oscuridad a 25 °C.

Después, el callo cocultivado se cultivó sobre medio que contiene 2,4-D durante 4 semanas en la oscuridad a 28 °C en presencia de un agente selectivo a una concentración adecuada. Durante este periodo, se desarrollaron islas de callo resistentes de crecimiento rápido. Tras la transferencia de este material a un medio de regeneración y la incubación en la luz se liberó el potencial embriogénico y se desarrollaron brotes en las siguientes cuatro a cinco semanas. Los brotes se extirparon del callo y se incubaron durante 2 a 3 semanas sobre un medio que contiene auxina, a partir del cual se trasladaron a tierra. Se cultivaron los brotes endurecidos con alta humedad y días cortos en un invernadero. Por último, las semillas se recolectaron de tres a cinco meses después del trasplante. El procedimiento produce transformantes de un solo locus a una tasa mayor que 50% (Aldemita y Hodges, 1996; Chan *et al.*, 1993; Hiei *et al.*, 1994).

D) Evaluación de transformantes: mediciones de crecimiento vegetativo

Se generaron aproximadamente 15 a 20 transformantes T0 independientes. Los transformantes primarios se trasladaron desde las cámaras de cultivo de tejido a un invernadero para su crecimiento y la recolección de la semilla T1. Fueron conservados cuatro acontecimientos (para los transformantes p37, promotor PRO0218) o cinco acontecimientos (para los transformantes p35, PRO0090) de los cuales se conservaron la progenie T1 que segregó 3:1 para la presencia/ausencia del transgén. Para cada uno de estos acontecimiento 10 plántulas T1 que contenían el transgén (hetero- u homocigotos) y 10 plántulas T1 que carecían del transgén (nulicigotos) se seleccionaron mediante el control de la expresión de marcadores visual. Las plantas T1 seleccionadas se trasladaron a un invernadero. Cada planta recibió una etiqueta de código de barras exclusiva para relacionar sin ambigüedades los datos de fenotipificación con la correspondiente planta. Las plantas T1 seleccionadas se cultivaron en tierra en macetas de 10 cm de diámetro bajo los siguientes ajustes ambientales: fotoperiodo, 11,5 h; intensidad de la luz diurna, 30.000 lux o más; temperatura diurna, 28 °C o mayor; temperatura nocturna, 22 °C; humedad relativa, 60-70%. Las plantas transgénicas y los correspondientes nulicigotos se cultivaron unos junto a otros en posiciones aleatorias. Después del etapa de la siembra hasta la etapa de la madurez, las plantas se hicieron pasar varias veces a través de una cabina de formación de imágenes digitales. En cada momento se tomaron imágenes digitales (2048 x 1536 píxeles, 16 millones de colores) de cada planta desde al menos 6 ángulos diferentes.

En la siguiente etapa, las panículas primarias maduras se recolectaron, se embolsaron, se marcaron con un código de barras y después se secaron durante tres días en una estufa a 37 °C. Las panículas después se trillaron y todas las semillas se recolectaron y se contaron. Las cáscaras llenas se separaron de las vacías utilizando un dispositivo de aire soplado. Las cáscaras vacías se desecharon y la fracción remanente se volvió a contar. Las cáscaras llenas se pesaron sobre una báscula analítica y se midió el área de la sección transversal utilizando formación de imágenes digitales. Este procedimiento permite obtener un conjunto de parámetros relacionados con la semilla.

Los parámetros descritos a continuación se obtuvieron de una forma automática a partir de las imágenes digitales utilizando programas informáticos de análisis de imágenes y se analizaron estadísticamente.

Se empleó un ANOVA (análisis de la varianza) de dos factores corregido para el diseño no equilibrado como modelo estadístico para la evaluación global de las características fenotípicas de la planta. Se realizó un ensayo de la F sobre todos los parámetros medidos en todas las plantas y de todos los acontecimientos transformados con

ese gen. El ensayo de la F se realizó para comprobar el efecto del gen sobre todos los acontecimientos de transformación y para verificar el efecto global del gen, también denominado en la presente “efecto global del gen”. Si el valor del ensayo de la F demuestra que los datos son significativos, entonces se concluye que existe un efecto “del gen”, lo cual significa que no solo la presencia o la posición del gen está provocando el efecto. El umbral para la significancia para un efecto global verdadero del gen se estableció al 5% de nivel de probabilidad para el ensayo de la F.

Para comprobar el efecto de los genes dentro de un acontecimiento, es decir, para un efecto específico de línea, se realizó un ensayo de la t dentro de cada acontecimiento utilizando conjuntos de datos procedentes de las plantas transgénicas y las correspondientes plantas nulas. Las “plantas nulas” o “segregantes nulos” o “nulicigotos” son las plantas tratadas de la misma manera que la planta transgénica, pero de las que el transgén se ha segregado. Las plantas nulas también pueden describirse como el homocigoto negativo de las plantas transformadas. El umbral para la significancia para el ensayo de la t se estableció al 10% de nivel de probabilidad. Los resultados para algunos acontecimiento pueden estar por encima o por debajo de este umbral. Esto se basa en la hipótesis de que un gen puede tener un efecto solo en ciertas posiciones en el genoma, y que la aparición de este efecto dependiente de la posición no es raro. Este tipo de efecto del gen también se denomina en la presente un “efecto de línea del gen”. El valor de p se obtuvo comparando el valor de t con la distribución de t o, como alternativa, comparando el valor de F con la distribución de F. El valor de p refleja la probabilidad de que la hipótesis nula (es decir, el transgén no tiene ningún efecto) sea correcta. El umbral para la significancia se estableció al 5% de valor de p para el ensayo de la F y al 10% de valor de p para el ensayo de la t.

El crecimiento vegetativo y el rendimiento de semillas se midieron según los procedimientos descritos anteriormente. Los inventores han descubierto, de modo sorprendente, que el rendimiento de semillas de las plantas transgénicas aumentó con las construcciones del promotor preferido en el endospermo y preferido en el embrión y/o la aleurona (expresado como peso total de las semillas, número de semillas (llenas) e índice de cosecha), cuando se compara con plantas nulas, y que las plantas transgénicas con las construcciones del promotor preferido en el embrión y/o la aleurona también habían aumentado el peso de mil granos, comparado con los transgénicos con la construcción del promotor preferido en el endospermo. Los inventores también observaron que el aumento del rendimiento fue mayor en las plantas transformadas con las construcciones del promotor preferido en el embrión y/o la aleurona que en las plantas con la construcción del promotor preferido en el endospermo. Los detalles se ofrecen en los párrafos E y F.

Los datos obtenidos en el experimento con plantas T1 después fueron confirmados en un experimento posterior con plantas T2. Lotes de semillas procedentes de las plantas positivas (hetero- y homocigotos) en T1 se seleccionaron controlando la expresión de marcadores. Para cada acontecimiento elegido, los lotes de semillas de heterocigotos se conservaron para la evaluación de T2. Dentro de cada lote de semillas se cultivó un número idéntico de plantas positivas y negativas en el invernadero para su evaluación. En particular, cuatro acontecimientos de los transformantes T2 p37 y tres acontecimientos de los transformantes T2 p35 fueron seleccionados para su posterior análisis. Para ambos transformantes T2 p37 y transformantes T2 p35 se ensayó un total de 120 plantas, distribuidas uniformemente a lo largo de cada acontecimiento.

E) Evaluación de los transformantes p37: medición de los parámetros relacionados con las semillas

Tras el análisis de las semillas descrito anteriormente, los inventores descubrieron que las plantas transformadas con el gen *AtCKX2* bajo el control del promotor preferido en el embrión y/o la aleurona presentaban un mayor peso total de las semillas, un mayor número de semillas llenas, un mayor índice de cosecha y un mayor peso de mil granos que las plantas que carecen del transgén *CKX2*. Estos descubrimientos fueron coherentes a través de 2 experimentos independientes con plantas T1, así como en un experimento con plantas T2, tal como se muestra en la tabla 5. Además de estos parámetros de rendimiento, 3 líneas en T1 también obtuvieron una puntuación positiva para el número total de semillas. Este aumento en el número total de semillas se confirmó en T2, en donde se demostró que el efecto es un efecto global del gen significativo (aumento medio +24%, valor de p del ensayo de la F de 0,0032).

Tabla 5. Análisis de los parámetros relacionados con las semillas para los transformantes p37

Parámetro	Generación T1, 1 ^{er} experimento	Generación T1, 2 ^o experimento	Generación T2	
	Diferencia con las plantas nulas	Diferencia con las plantas nulas	Diferencia con las plantas nulas	valor de p
Peso total de las semillas	+22%	+22%	+51%	0,0000

Número de semillas llenas	+22%	+17%	+46	0,0000
Índice de cosecha	+25%	+20%	+37%	0,0025
Peso de mil granos	+1%	+5%	+3%	0,0000

El peso total de las semillas se midió pesando todas las semillas llenas recolectadas de una planta de arroz transformada. El número de semillas llenas se determinó contando el número de semillas llenas recolectadas de una planta de arroz transformada. El número total de semillas se determinó contando el número de semillas recolectadas de una planta. El índice de cosecha se define como la proporción entre el peso total de las semillas y el área sobre el suelo (mm²), multiplicada por un factor de 10⁶. El peso de mil granos (TKW, "thousand kernel weight") se derivó del número de semillas llenas que se contaron y su peso total. Los números indican el aumento medio (en porcentaje) de cada parámetro calculado a partir de los transgenes frente a los correspondiente nulicigotos de 4 acontecimientos independientes en la generación T1, comprendiendo cada acontecimiento 10 plantas que portan el transgén y 10 nulicigotos, y de 4 acontecimientos independientes en la generación T2, comprendiendo cada acontecimiento 20 plantas que portan el transgén y 20 nulicigotos. Los valores de p del ensayo de la F listados para los datos de la generación T2 demuestran que los aumentos obtenidos para los diversos parámetros de rendimiento de semillas son todos significativos y que existe claramente un efecto global del gen.

15 F) Evaluación de los transformantes p35: medición de los parámetros relacionados con las semillas

Las plantas transformadas con el gen *AtCKX2* bajo el control del promotor preferido en el endospermo también presentaban una mayor rendimiento, comparado con las plantas nulicigóticas control, en particular para el peso total de las semillas, el número de semillas llenas y el índice de cosecha. El peso total de las semillas, el número de semillas llenas y el índice de cosecha se definieron anteriormente.

20 En un primer experimento se compararon plantas de la generación T1 de cinco acontecimientos independientes, en cada acontecimiento 10 plantas T1 que portan el transgén frente a 10 plantas T1 control correspondientes. Para el parámetro "peso total de las semillas", dos de los cinco acontecimiento mostraron un aumento significativo (58% y 67%, con un valor de p para el ensayo de la t de 0,0551 y 0,0211, respectivamente). Se obtuvieron resultados similares para el número de semillas llenas, para el cual estas dos líneas mostraron un aumento del 47% y 68%, con un valor de p de 0,0846 y 0,0166, respectivamente. Las dos líneas también obtuvieron una puntuación positiva para el índice de cosecha (aumentos del 41% (valor de p de 0,0223) y del 31%, respectivamente). Además de estas dos líneas, una tercera línea también obtuvo una puntuación significativamente mayor que las correspondientes plantas control nulicigóticas (+41%, valor de p de 0,035).

30 Los datos positivos para el rendimiento de las semillas observadas en la generación T1 se confirmaron en la generación T2. Los datos se ofrecen en la tabla 6.

Tabla 6. Análisis de los parámetros relacionados con las semillas para transformantes p37

Parámetro	Generación T2	
	Diferencia con las plantas nulas	valor de p
Peso total de las semillas	+24%	0,0484
Número de semillas llenas	+26%	0,0254
Índice de cosecha	+19%	0,0277

35 Los números indican el aumento medio (en porcentaje) de cada parámetro calculado a partir de transgenes frente a los correspondientes nulicigotos de 3 acontecimientos independientes en la generación T2, comprendiendo cada acontecimiento 20 plantas que portan el transgén y 20 nulicigotos. Los valores de p del ensayo de la F listados para los datos de la generación T2 demuestran que los aumentos obtenidos para los diversos parámetros de rendimiento de semillas son todos significativos y que existe un claro efecto global del gen.

Referencias bibliográficas

Aldemita, R.R. y Hodges, T.K. (1996), *Agrobacterium tumefaciens*-mediated transformation of japonica and indica

- rice varieties, *Planta*, 199, 612-617.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. y Lipman, D.J. (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucl. Acids Res.*, 25, 3389-3402.
- 5 Armstrong, D.J. (1994), en *Cytokinins: Chemistry, Activity and Functions*, eds. Mok, D.W.S y Mok, M.C. (CRC Boca Ratón, FL), pp. 139-154.
- An, G., Watson, B.D., Stachel, S., Gordon, M.P., y Nester, E.W. (1985), New cloning vehicles for transformation of higher plants, *EMBO J.*, 4, 277-284.
- Armstrong, C.L., Petersen, W.P., Buchholz, W.G., Bowen, B.A., y Sulc, S.L. (1990), Factors affecting PEG-mediated stable transformation of maize protoplasts, *Plant Cell Reports*, 9, 335-339.
- 10 Chan, M.T., Chang, H.H., Ho, S.L., Tong, W.F., y Yu, S.M. (1993), *Agrobacterium* mediated production of transgenic rice plants expressing a chimeric alpha-amylase promoter/beta-glucuronidase gene, *Plant Mol. Biol.*, 22, 491-506.
- Christou, P., McCabe, D.E., y Swain, W.F. (1988), Stable transformation of soybean callus by DNA-coated gold particles, *Plant Physiol.*, 87, 671-674.
- 15 Crossway, A., Oakes, J.V., Irvine, J.M., Ward, B., Knauf, V.C., y Shewmaker, C.K. (1986), Integration of foreign DNA following microinjection of tobacco mesophyll protoplasts, *Mol. Gen. Genet.*, 202, 179-185.
- Dodds, J.H. (1985), "Plant genetic engineering", Cambridge University Press.
- Ellis, J.G., Llewellyn, D.J., Dennis, E.S., y Peacock, W.J. (1987), Maize Adh-1 promoter sequences control anaerobic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco, *EMBO J.*, 6, 11-16.
- 20 Hanahan, D. (1983), Studies on transformation of *Escherichia coli* with plasmids, *J. Mol. Biol.*, 166, 557-580.
- Hansen, G. y Chilton, M.D. (1996), "Agrolistic" transformation of plant cells: integration of T-strands generated in planta, *Proc. Natl. Acad. Sci. U.S.A.*, 93, 14978-14983.
- Hansen, G., Shillito, R.D., y Chilton, M.D. (1997), T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes, *Proc. Natl. Acad. Sci. U.S.A.*, 94, 11726-11730.
- 25 Herrera-Estrella, L., De Block, M., Messens, E.H.J.P., Van Montagu, M., y Schell, J. (1983), Chimeric genes as dominant selectable markers in plant cells, *EMBO J.*, 2, 987-995.
- Hiei, Y., Ohta, S., Komari, T., y Kumashiro, T. (1994), Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA, *Plant J.*, 6, 271-282.
- Houba-Herlin, N., Pethe, C., d'Alayer, J. y Laloue, M. (1999), *Plant J.*, 17:615-626.
- 30 Klee, H.J. y Lanehon, M.B. (1995), en *Plant Hormones: Physiology, Biochemistry and Molecular Biology*, ed. Davies, P.J. (Kluwer, Dordrecht, Países Bajos), pp. 340-353.
- Krens, F.A., Molendijk, L., Wullems, G.J., y Schilperoort, R.A. (1982), *In vitro* transformation of plant protoplasts with Ti-plasmid DNA, *Nature*, 296, 72-74.
- 35 Mok M.C. (1994), en *Cytokines: Chemistry, Activity and Function*, eds., Mok, D.W.S. y Mok, M.C. (CRC Boca Ratón, FL), pp.155-166.
- Morris, R.O. *et al.* (1999), Isolation of a gene encoding a glycosylated cytokinin oxidase from maize, *Biochem. Biophys. Res. Commun.*, 255, 328-333.
- 40 Motyka, V., Faiss, M., Strnad, M., Kaminek, M. y Schmuelling, T. (1996), Changes in cytokinin content and cytokinin oxidase activity in response to derepression of ipt gene transcription in transgenic tobacco calli and plants, *Plant Physiol.*, 112, 1035-1043.
- Paszkowski, J., Shillito, R.D., Saul, M., Mandak, V., y Hohn, T.H.B.P.I. (1984), Direct gene transfer to plants, *EMBO J.*, 3, 2717-2722.
- Rinaldi, A.C. y Comandini, O. (1999), Cytokinin oxidase strikes again, *Trends in Plant Sc.*, 4, 300.
- Sambrook, J., Fritsch, E.F., y Maniatis, T. (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor

Laboratory Press.

Schmülling, T., Rupp, H.M. Frank, M. y Schafer, S. (1999), en *Advances in Regulation of Plant Growth and Development*, eds. Surnad, M. Pac P. y Beck, E. (Peres, Praga), pp. 85-96.

LISTA DE SECUENCIAS

- <110> Schmülling, Thomas Werner, Tomàs
- <120> Procedimiento para modificar la morfología, la bioquímica y la fisiología de plantas
- 5 <130> CD-029-PCT-3
- <150> documento US 10/871.304
- <151> 18-06-2004
- 10 <160> 45
- <170> PatentIn versión 3.3
- <210> 1
- 15 <211> 2236
- <212> ADN
- <213> *Arabidopsis thaliana*
- <400> 1
- 20 000
- <210> 2
- <211> 575
- 25 <212> PRT
- <213> *Arabidopsis thaliana*
- <400> 2
- 30 000
- <210> 3
- <211> 2991
- <212> ADN
- 35 <213> *Arabidopsis thaliana*
- <400> 3

atggctaatc	ttcgtttaat	gatcacttta	atcacggttt	taatgatcac	caaatcatca	60
aacggtatta	aaattgattt	acctaaatcc	cttaacctca	ccctctctac	cgatccttcc	120
atcatctccg	cagcctctca	tgacttcgga	aacataacca	ccgtgacccc	cggcggcgta	180
atctgcccct	cctccaccgc	tgatatctct	cgtctcctcc	aatacgccgc	aaacggaaaa	240
agtacattcc	aagtagcggc	tcgtggccaa	ggcactcct	taaacggcca	agcctcggtc	300
tccggcggag	taatcgtcaa	catgacgtgt	atcactgacg	tgggtggttc	aaaagacaag	360
aagtacgctg	acgtggcggc	cgggacgtta	tgggtggatg	tgcttaagaa	gacggcggag	420
aaaggggtgt	cgccggtttc	ttggacggat	tatttgcata	taaccgtcgg	aggaacgttg	480
tcgaatggtg	gaattggtgg	tcaagtgttt	cgaaacggtc	ctcttgttag	taacgtcctt	540
gaattggacg	ttattactgg	tacgcacatt	ctaaactttg	atgtacatac	aacaacaaaa	600
actgtttttg	ttttatagta	tttttcattt	tttgtacat	aggttttatg	ttttatagtt	660
gtgctaaact	tcttgcacca	cacgtaagtc	ttcgaaacac	aaaatgcgta	acgcacatct	720
atgttttttg	tacatatatt	atgttgttca	tgagaaataa	agtaattaca	tatacacaca	780
tttattgtcg	tacatatata	aataattaaa	gacaaatttt	cacaattggt	agcgtgttaa	840
tttgggattt	ttgtaattga	catgcatgac	gcatgcatat	ggagcttttc	ggttttctta	900
gatttgtgta	gtatttcaaa	tatatcattt	atcttctttc	gaataaagag	gtggtatatt	960
tttaaaatag	caacatttca	gaatttttct	ttgaatttac	actttttaaa	ttgttattgt	1020
taatattgat	tttgaataaa	taatttcagg	gaaagggtgaa	atgttgacat	gctcgcgaca	1080
gctaaaccce	gaattgttct	atggagtggt	aggagggttg	ggtcaatttg	gaattataac	1140
gagagccaga	attgttttgg	accatgcacc	taaacgggta	cgtatcatca	tattttacca	1200
tttgttttag	tcagcattca	tttttcatta	gtaattccgt	ttcaatttct	aaattttttt	1260
agtcaataga	aaatgattct	tatgtcagag	cttgattatt	tagtgatttt	tattgagata	1320
aaataaaaata	taacctaacg	gaaataatta	ttttactaat	cggataatgt	ctgattaaaa	1380
cattttatga	tattacacta	agagagttag	agacgtatgg	atcacaaaac	atgaagcttt	1440
cttagatggg	atcctaaaac	taaagttagg	tacaagtttg	gaatttaggt	caaattgctta	1500
agttgcatta	atgtgaacaa	aatctatgca	ttgaataaaa	aaaagatatg	gattatttta	1560
taaagtatag	tccttgtaat	cctaggactt	gttgtcta	cttgtcttat	gcgtgcaaat	1620

ctttttgatg	tcaatatata	atccttgttt	attagagtca	agctctttca	ttagtcaact	1680
actcaaata	actccaaagt	ttagaatata	gtcttctgac	taattagaat	cttacaaccg	1740
ataaacgtta	caatttggtt	atcattttta	aaaacagatt	tggtcataat	atacgatgac	1800
gttctgtttt	agtttcatct	attcacaat	tttatataat	tattttcaag	aaaatattga	1860
aatactatac	tgtaatatgg	tttctttata	tatgtgtgta	taaattaaat	gggattggtt	1920
tctctaaatg	aaattgtgta	ggccaaatgg	tttcggatgc	tctacagtga	tttcacaact	1980
tttacaagg	accaagaacg	tttgatatca	atggcaaacg	atattggagt	cgactattta	2040
gaagggtcaa	tatttctatc	aaacgggtgc	gttgacacct	ctttttccc	accttcagat	2100
caatctaaag	tcgctgatct	agtcaagcaa	cacggtatca	tctatgttct	tgaagtagcc	2160
aagtattatg	atgatcccaa	tctccccatc	atcagcaagg	tactacacat	ttacattttc	2220
atcatcgttt	ttatcatacc	ataagatatt	taaatgattc	atcattgcac	cacattaaga	2280
tattcatcat	catcatcggt	acattttttt	ttgcatctta	tgcttctcat	aatctactat	2340
tgtgtaggtt	attgacacat	taacgaaaac	attaagttac	ttgcccgggt	tcatatcaat	2400
gcacgacgtg	gcctacttcg	atttcttgaa	ccgtgtacat	gtcgaagaaa	ataaactcag	2460
atctttggga	ttatgggaac	ttcctcatcc	ttggcttaac	ctctacgttc	ctaaatctcg	2520
gattctcgat	tttcataacg	gtgttgctca	agacattctt	cttaagcaaa	aatcagcttc	2580
gggactcgct	cttctctatc	caacaaaccg	gaataagtac	atacttctct	tcattcatat	2640
ttatcttcaa	gaaccaaagt	aaataaattt	ctatgaactg	attatgctgt	tattgttaga	2700
tgggacaatc	gtatgtcggc	gatgatacca	gagatcgatg	aagatgttat	atatattatc	2760
ggactactac	aatccgctac	cccaaaggat	cttcagaag	tggagagcgt	taacgagaag	2820
ataattaggt	tttgcaagga	ttcaggtatt	aagattaagc	aatatcta	gcattatact	2880
agtaaagaag	attggattga	gcattttgga	tcaaatggg	atgatttttc	gaagaggaaa	2940
gatctatttg	atcccaagaa	actgttatct	ccagggcaag	acatcttttg	a	2991

<210> 4
 <211> 501
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 4

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

		275					280					285			
Leu	Val	Lys	Gln	His	Gly	Ile	Ile	Tyr	Val	Leu	Glu	Val	Ala	Lys	Tyr
	290					295					300				
Tyr	Asp	Asp	Pro	Asn	Leu	Pro	Ile	Ile	Ser	Lys	Val	Ile	Asp	Thr	Leu
305					310					315					320
Thr	Lys	Thr	Leu	Ser	Tyr	Leu	Pro	Gly	Phe	Ile	Ser	Met	His	Asp	Val
				325					330					335	
Ala	Tyr	Phe	Asp	Phe	Leu	Asn	Arg	Val	His	Val	Glu	Glu	Asn	Lys	Leu
			340					345					350		
Arg	Ser	Leu	Gly	Leu	Trp	Glu	Leu	Pro	His	Pro	Trp	Leu	Asn	Leu	Tyr
		355					360					365			
Val	Pro	Lys	Ser	Arg	Ile	Leu	Asp	Phe	His	Asn	Gly	Val	Val	Lys	Asp
	370					375					380				
Ile	Leu	Leu	Lys	Gln	Lys	Ser	Ala	Ser	Gly	Leu	Ala	Leu	Leu	Tyr	Pro
385					390					395					400
Thr	Asn	Arg	Asn	Lys	Trp	Asp	Asn	Arg	Met	Ser	Ala	Met	Ile	Pro	Glu
				405					410					415	
Ile	Asp	Glu	Asp	Val	Ile	Tyr	Ile	Ile	Gly	Leu	Leu	Gln	Ser	Ala	Thr
			420					425					430		
Pro	Lys	Asp	Leu	Pro	Glu	Val	Glu	Ser	Val	Asn	Glu	Lys	Ile	Ile	Arg
		435					440					445			
Phe	Cys	Lys	Asp	Ser	Gly	Ile	Lys	Ile	Lys	Gln	Tyr	Leu	Met	His	Tyr
	450				455						460				
Thr	Ser	Lys	Glu	Asp	Trp	Ile	Glu	His	Phe	Gly	Ser	Lys	Trp	Asp	Asp
465					470					475					480
Phe	Ser	Lys	Arg	Lys	Asp	Leu	Phe	Asp	Pro	Lys	Lys	Leu	Leu	Ser	Pro
				485					490					495	
Gly	Gln	Asp	Ile	Phe											
			500												

<210> 5

<211> 3302

<212> ADN

<213> *Arabidopsis thaliana*

5

<400> 5

000

10

<210> 6

<211> 523

<212> PRT

<213> *Arabidopsis thaliana*

15

<400> 6

000

20

<210> 7

<211> 2782

<212> ADN

<213> *Arabidopsis thaliana*

25

<400> 7

000

30

<210> 8

<211> 524

<212> PRT

<213> *Arabidopsis thaliana*

<400> 8

000

5 <210> 9
<211> 2805
<212> ADN
<213> *Arabidopsis thaliana*

10 <400> 9
000

15 <210> 10
<211> 536
<212> PRT
<213> *Arabidopsis thaliana*

20 <400> 10
000

25 <210> 11
<211> 1936
<212> ADN
<213> *Arabidopsis thaliana*

30 <400> 11
000

35 <210> 12
<211> 504
<212> PRT
<213> *Arabidopsis thaliana*

40 <400> 12
000

45 <210> 13
<211> 31
<212> ADN
<213> Secuencia artificial

50 <220>
<223> oligonucleótido: cebador o sonda

55 <400> 13
000

60 <210> 14
<211> 35
<212> ADN
<213> Secuencia artificial

<220>
<223> oligonucleótido: cebador o sonda

<400> 14
000

<210> 15
 <211> 33
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> oligonucleótido: cebador o sonda

 10 <400> 15

gcggtaccag agagagaaac ataaacaaat ggc **33**

 <210> 16
 <211> 31
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido: cebador o sonda

 20 <400> 16

gcggtaccca attttacttc caccaaaatg c **31**

 <210> 17
 <211> 34
 25 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido: cebador o sonda

 30 <400> 17

 000
 35
 <210> 18
 <211> 31
 <212> ADN
 <213> Secuencia artificial
 40
 <220>
 <223> oligonucleótido: cebador o sonda

 <400> 18
 45
 000

 <210> 19
 <211> 28
 50 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido: cebador o sonda
 55

 <400> 19

 000

<210> 20
<211> 32
<212> ADN
5 <213> Secuencia artificial

<220>
<223> oligonucleótido: cebador o sonda

10 <400> 20

000

<210> 21
15 <211> 28
<212> ADN
<213> Secuencia artificial

<220>
20 <223> oligonucleótido: cebador o sonda

<400> 21

000
25

<210> 22
<211> 31
<212> ADN
<213> Secuencia artificial
30

<220>
<223> oligonucleótido: cebador o sonda

<400> 22
35

000

<210> 23
<211> 32
40 <212> ADN
<213> Secuencia artificial

<220>
<223> oligonucleótido: cebador o sonda
45

<400> 23

000

50 <210> 24
<211> 32
<212> ADN
<213> Secuencia artificial

55 <220>
<223> oligonucleótido: cebador o sonda

<400> 24

60 000

<210> 25
 <211> 1728
 <212> ADN
 <213> *Arabidopsis thaliana*

5

<400> 25

000

<210> 26
 <211> 1506
 <212> ADN
 <213> *Arabidopsis thaliana*

15

<400> 26

atggctaato	ttcgtttaat	gatcacttta	atcacggttt	taatgatcac	caaatcatca	60
aacgggatta	aaattgattt	acctaaatcc	cttaacctca	ccctctctac	cgatccttcc	120
atcatctccg	cagcctctca	tgacttcgga	aacataacca	ccgtgacccc	cggcggcgta	180
atctgccctt	cctccaccgc	tgatatctct	cgtctcctcc	aatacgccgc	aaacggaaaa	240
agtacattcc	aagtagcggc	tcgtggccaa	ggccactcct	taaacggcca	agcctcggtc	300
tccggcggag	taatcgtcaa	catgacgtgt	atcactgacg	tgggtggttc	aaaagacaag	360
aagtacgctg	acgtggcggc	cgggacgtta	tgggtggatg	tgcttaagaa	gacggcggag	420
aaaggggtgt	cgccggttcc	ttggacggat	tatttgcata	taaccgtcgg	aggaacggtg	480
tcgaatggtg	gaattggtgg	tcaagtgttt	cgaaacggtc	ctcttgtag	taacgtcctt	540
gaattggacg	ttattactgg	gaaaggtgaa	atggtgacat	gctcgcgaca	gctaaaccca	600
gaattgttct	atggagtgtt	aggaggtttg	ggtcaatttg	gaattataac	gagagccaga	660
attgttttgg	accatgcacc	taaacggggc	aatgggttcc	ggatgctcta	cagtgatttc	720
acaactttta	caaaggacca	agaacgtttg	atatcaatgg	caaacgatat	tggagtcgac	780
tatttagaag	gtcaaatatt	tctatcaaac	ggtgtcgttg	acacctcttt	tttcccacct	840
tcagatcaat	ctaaagtcgc	tgatctagtc	aagcaacacg	gtatcatcta	tgttcttgaa	900
gtagccaagt	attatgatga	tcccaatctc	cccatcatca	gcaaggttat	tgacacatta	960
acgaaaacat	taagttactt	gcccgggttc	atatcaatgc	acgacgtggc	ctacttcgat	1020
ttcttgaacc	gtgtacatgt	cgaagaaaa	aaactcagat	ctttgggatt	atgggaactt	1080
cctcatcctt	ggcttaacct	ctacgttctt	aaactctcga	ttctcgattt	tcataacggt	1140
gttgtcaaa	acattcttct	taagcaaaaa	tcagcttcgg	gactcgctct	tctctatcca	1200
acaaaccgga	ataaatggga	caatcgtatg	tcggcgatga	taccagagat	cgatgaagat	1260
gttatatata	ttatcggact	actacaatcc	gctaccccaa	aggatcttcc	agaagtggag	1320
agcgttaacg	agaagataat	taggttttgc	aaggattcag	gtattaagat	taagcaatat	1380
ctaattgcatt	atactagtaa	agaagattgg	attgagcatt	ttggatcaaa	atgggatgat	1440
ttttcgaaga	ggaaagatct	atthgatccc	aagaaaactgt	tatctccagg	gcaagacatc	1500
ttttga						1506

<210> 27
 <211> 1572
 <212> ADN
 <213> *Arabidopsis thaliana*

20

<400> 27

000

<210> 28
 <211> 1575
 <212> ADN
 <213> *Arabidopsis thaliana*

30

<400> 28

000

<210> 29
 <211> 1611
 <212> ADN
 5 <213> *Arabidopsis thaliana*

 <400> 29

 000
 10
 <210> 30
 <211> 1515
 <212> ADN
 <213> *Arabidopsis thaliana*
 15
 <400> 30

 000
 20
 <210> 31
 <211> 84
 <212> ADN
 <213> *Arabidopsis thaliana*
 25
 <400> 31

tcagcttcgg gactcgctct totctatcca acaaaccgga ataaatggga caatcgatg 60
tcggcgatga taccagagat cgat 84

 <210> 32
 <211> 28
 30 <212> PRT
 <213> *Arabidopsis thaliana*

 <400> 32

Ser Ala Ser Gly Leu Ala Leu Leu Tyr Pro Thr Asn Arg Asn Lys Trp
1 5 10 15
Asp Asn Arg Met Ser Ala Met Ile Pro Glu Ile Asp
20 25
 35
 <210> 33
 <211> 2814
 <212> ADN
 <213> *Arabidopsis thaliana*
 40
 <400> 33

 000
 45
 <210> 34
 <211> 1620
 <212> ADN
 <213> *Arabidopsis thaliana*
 50
 <400> 34

 000
 55
 <210> 35
 <211> 539

<212> PRT
 <213> *Arabidopsis thaliana*

<400> 35

000

5 <210> 36
 <211> 501
 <212> PRT
 <213> *Arabidopsis thaliana*

10 <400> 36

```

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


				325						330					335	
Ala	Tyr	Phe	Asp	Phe	Leu	Asn	Arg	Val	His	Val	Glu	Glu	Asn	Lys	Leu	
			340					345					350			
Arg	Ser	Leu	Gly	Leu	Trp	Glu	Leu	Pro	His	Pro	Trp	Leu	Asn	Leu	Tyr	
		355					360					365				
Val	Pro	Lys	Ser	Arg	Ile	Leu	Asp	Phe	His	Asn	Gly	Val	Val	Lys	Asp	
	370					375					380					
Ile	Leu	Leu	Lys	Gln	Lys	Ser	Ala	Ser	Gly	Leu	Ala	Leu	Leu	Tyr	Pro	
385					390						395				400	
Thr	Asn	Arg	Asn	Lys	Trp	Asp	Asn	Arg	Met	Ser	Ala	Met	Ile	Pro	Glu	
			405						410					415		
Ile	Asp	Glu	Asp	Val	Ile	Tyr	Ile	Ile	Gly	Leu	Leu	Gln	Ser	Ala	Thr	
			420					425					430			
Pro	Lys	Asp	Leu	Pro	Glu	Val	Glu	Ser	Val	Asn	Glu	Lys	Ile	Ile	Arg	
		435					440					445				
Phe	Cys	Lys	Asp	Ser	Gly	Ile	Lys	Ile	Lys	Gln	Tyr	Leu	Met	His	Tyr	
	450					455					460					
Thr	Ser	Lys	Glu	Asp	Trp	Ile	Glu	His	Phe	Gly	Ser	Lys	Trp	Asp	Asp	
465					470					475					480	
Phe	Ser	Lys	Arg	Lys	Asp	Leu	Phe	Asp	Pro	Lys	Lys	Leu	Leu	Ser	Pro	
			485					490						495		
Gly	Gln	Asp	Ile	Phe												
			500													

<210> 37
 <211> 1455
 <212> ADN
 5 <213> *Arabidopsis thaliana*

<400> 37

atgggcta	atc	ttcgttta	gatcacttta	atcacggttt	taatgatcac	caaatcatca	60
aacggatta	aa	aatgattt	acctaaatcc	cttaacctca	ccctctctac	cgatccttcc	120
atcatctccg	cagcctctca	tgacttcgga	aacataacca	ccgtgacccc	cggcggcgta		180
atctgccct	cctccaccgc	tgatatctct	cgtctcctcc	aatacgccgc	aaacggaaaa		240
agtacattcc	aagtagcggc	tcgtggccaa	ggccactcct	taaacggcca	agcctcggtc		300
tccggcggag	taatcgtcaa	catgacgtgt	atcactgacg	tggtggtttc	aaaagacaag		360
aagtacgctg	acgtggcggc	cgggacggtt	tgggtggatg	tgcttaagaa	gacggcggag		420
aaaggggtgt	cgccggtttc	ttggacggat	tatttgcata	taaccgtccg	aggaacgttg		480
tcgaatggtg	gaattggtgg	tcaagtgttt	cgaaacggtc	ctcttgttag	taacgtcctt		540
gaattggacg	ttattactgg	gaaaggtgaa	atggtgacat	gctcgcgaca	gctaaacca		600
gaattgttct	atggagtgtt	aggaggtttg	ggtcaatttg	gaattataac	gagagccaga		660
attgttttgg	accatgcacc	taaacggcaa	gaacgtttga	tatcaatggc	aaacgatatt		720
ggagtcgact	attdagaagg	tcaaataatt	ctatcaaacg	gtgtcgttga	cacctctttt		780
ttcccacctt	cagatcaatc	taaagtcgct	gatctagtca	agcaacacgg	tatcatctat		840
gttcttgaag	tagccaagta	ttatgatgat	cccaatctcc	ccatcatcag	caaggttatt		900
gacacattaa	cgaaaacatt	aagttacttg	cccgggttca	tatcaatgca	cgacgtggcc		960
tacttcgatt	tcttgaaccg	tgtacatgtc	gaagaaaata	aactcagatc	tttgggatta		1020
tgggaacttc	ctcatccttg	gcttaacctc	tacgttccta	aatctcggat	tctcgatttt		1080
cataacgggtg	ttgtcaaaga	cattcttctt	aagcaaaaat	cagcttcggg	actcgtctct		1140
ctctatccaa	caaaccggaa	taaattgggac	aatcgtatgt	cggcgatgat	accagagatc		1200
gatgaagatg	ttatatatat	tatcggacta	ctacaatccg	ctaccccaaa	ggatcttcca		1260
gaagtggaga	gcgttaacga	gaagataaatt	aggttttgca	aggattcagg	tattaagatt		1320
aagcaatatc	taatgcatta	tactagtaaa	gaagattgga	ttgagcattt	tggatcaaaa		1380
tgggatgatt	tttcgaagag	gaaagatcta	tttgatccca	agaaactggt	atctccaggg		1440
caagacatct	tttga						1455

10 <210> 38
 <211> 484

<212> PRT

<213> *Arabidopsis thaliana*

<400> 38

5

```

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

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

```

<210> 39
 <211> 60
 <212> ADN
 <213> Artificial sequence
 5
 <220>
 <223> oligonucleótido: cebador o sonda
 <400> 39
ggggacaagt ttgtacaaa aagcaggctt cacaatggct aatcttcggt taatgatcac 60
 10
 <210> 40
 <211> 49
 <212> ADN
 <213> Artificial sequence
 15
 <220>
 <223> oligonucleótido: cebador o sonda
 <400> 40
 20
ggggaccact ttgtacaaga aagctggggt caaaagatgt cttgccctg 49
 <210> 41
 <211> 3328
 <212> ADN
 <213> Artificial sequence
 25
 <220>
 <223> módulo de expresión
 30
 <220>
 <221> característica_misc
 <222> (2495)..(2496)
 <223> n es a, c, g, o t
 35
 <400> 41

ggtcagccaa	tacattgatc	cgttgccaat	catgcaaagt	atthttggctg	tggccgagtg	60
ccggaattga	taattgtggt	ctgactaaat	taaatgacca	gaagtcgcta	tcttccaatg	120
tatccgaaac	ctggattaaa	caatcctggt	ctgttctcta	gcccctcctg	catggccgga	180
ttgttttttt	gacatgtttt	cttgactgag	gcctgtttgt	tctaaacttt	ttcttcaaac	240
ttttaacttt	ttcatcacat	cagaactttt	ctacacatat	aaacttttaa	cttttccgtc	300
acatcgttcc	aatttcaatc	aaactttcaa	ttttggcgtg	aactaaacac	accctgagtc	360
ttttattgct	cctccgtacg	ggttggctgg	ttgagaatag	gtatthttcag	agagaaaatc	420
tagatattgg	gaggaacttg	gcatgaatgg	ccactatatt	tagagcaatt	ctacggtcct	480
tgaggaggta	ccatgaggta	ccaaaatttt	agtgtaaatt	ttagtatctc	attataacta	540
ggtattatga	ggtaccaaat	ttacaataga	aaaaatagta	cttcatggta	ctttcttaag	600
taccgtaaaa	ttgctcctat	atthtaagggg	atgthttatat	ctatccatat	ccataaattg	660
atthttgataa	gaaaaaatgt	gagcacacca	agcatgtcca	tgaccttgca	ctcttggtc	720
actcgtcaac	tgtgaagaac	ctcaaaaatg	ctcaatatag	ctacaggtgc	ctgaaaaaat	780
aactthaaag	ttttgaacat	cgatttctact	aaacaacaat	tattatctcc	ctctgaaaga	840
tgatagthta	gaactctaga	atcattgtcg	gcgggagaaag	taaattattt	tccccaaatt	900
tccagctatg	aaaaaacctt	caccaaacac	catcaaacia	gagthtccca	aaccgccccat	960
gcggccatgc	tgtcacgcaa	cgcaccgcat	tgcctgatgg	ccgctcgatg	catgcatgct	1020
tccccgtgca	catatccgac	agacgcgccc	tgtcagcgag	ctcctcgacc	gacctgtgta	1080
gcccattgcaa	gcatccaccc	ccgccacgta	cacccccctc	tctctccctac	gtgtcaccgc	1140
tctctccacc	tatatatgcc	cacctggccc	ctctcctccc	atctccactt	caccogatcg	1200
cttcttcttc	ttcttcgthg	cattcatctt	gctagcattt	aaatcaacta	gggatatcac	1260
aagthttgtac	aaaaaagcag	gcttcacaat	ggctaattct	cgtthtaatga	tcactthtaat	1320
cacggtthta	atgatcacca	aatcatcaaa	cggtattaaa	attgatttac	ctaaatccct	1380
taacctcacc	ctctctaccg	atccttccat	catctccgca	gcctctcatg	acttcggaaa	1440
cataaccacc	gtgacccccg	gcgggcgtat	ctgccccctc	tccaccgctg	atatctctcg	1500
tctctcccaa	tacgcccga	acggaaaaag	tacattccaa	gtagcggctc	gtggccaagg	1560
ccactcctta	aacggccaag	cctcggctc	cggcggagta	atcgtcaaca	tgacgtgat	1620
cactgacgtg	gtgthttcaa	aagacaagaa	gtacgctgac	gtggcggccg	ggacgthtatg	1680
ggtggatgtg	cttaagaaga	cggcggagaa	aggggtgtcg	ccggtthctt	ggacggatta	1740
thttgcatata	accgtcggag	gaacgtgtc	gaatggthga	attggtggtc	aagthttctg	1800
aaacggtcct	cttgttagta	acgtccttga	attggacgth	attactggga	aagthgaaat	1860
gthtgacatgc	tgcgcacagc	taaaaccaga	attgthctat	ggagthtag	gagthttggg	1920
tcaattthga	attataacga	gagccagaat	tgtthttggac	catgcaccta	aacggcaaga	1980
acgthttgata	tcaatggcaa	acgatattgg	agtcgactat	ttagaaggth	aaatattthct	2040
atcaaaccgt	gtcgttgaca	cctctthttt	cccaccttca	gatcaatcta	aagthcgtga	2100
tctagthcaag	caacacggta	tcatctatgt	tctthgaagta	gccaagthatt	atgatgatcc	2160
caatctcccc	atcatcagca	aggtthattga	cacattaacg	aaaacattaa	gthtacttgcc	2220
cgggttcata	tcaatgcacg	acgtggccta	cttcgattth	thgaaccgth	tacatgtcga	2280
agaaaataaa	ctcagatctt	tgggattatg	ggaacttct	catccttggc	thaacctcta	2340
cgttctctaaa	tctcggattc	tcgattthtca	thaacggtgt	gtcaaagaca	ttcttcttaa	2400
gcaaaaatca	gcttccgggac	tgcctcttct	ctatccaaica	aaccggaata	agthacatact	2460
tctcttcatt	catatthtatc	thcaagaacc	aaagnnatgg	gacaatcgta	tgtcggcgat	2520
gataccagag	atcgatgaag	atgthtatata	tattatcgga	ctactacaat	ccgthacccc	2580
aaaggatctt	ccagaagthgg	agagcgtthaa	cgagaagata	attaggtthtt	gcaaggattc	2640
aggtatthaa	atthaaagcaat	atctaatgca	thataactagt	aaagaagatt	ggattgagca	2700
thtttgatca	aaatgggatg	atthtttccgaa	gaggaaagat	ctatthtgatc	ccaagaaact	2760
gthtatctcca	gggcaagaca	tctthttgaac	ccagctthtct	tgtacaaagt	ggtgatatca	2820
caagccccggg	cggthcttcta	gggataacag	ggttaattata	tccctctaga	tcacaagccc	2880
gggcggtctt	ctacgatgat	tgagthaaata	tgtgtcacgc	atcaccatgg	gtggcagthgt	2940
cagthgtgagc	atgacctgta	atgaacaatt	gaaatgaaaa	gaaaaaaagt	actccatctg	3000
thtccaaatta	aaactthttt	thaacctthta	ataggtthtat	acaataattg	atatagttht	3060
tctgtatatg	tctaatthgt	tatcatccgg	cggthcttct	agggataaca	gggthaatat	3120
atccctctag	acaacacaca	acaaataaga	gaaaaaacia	ataatattaa	thtgagaatg	3180
aacaaaagga	ccatatcatt	cattaactct	tctccatcca	thtccattth	acagthtccgat	3240
agcgaaaacc	gaataaaaaa	cacagthaaat	tacaagcaca	acaaatggta	caagaaaaac	3300
agthtttccca	atgccataat	actcgaac				3328

<210> 42
<211> 1506

<212> ADN
 <213> *Arabidopsis thaliana*

<400> 42

5

```

atggctaatc ttcgtttaat gatcacttta atcacggttt taatgatcac caaatcatca      60
aacggtatta aaattgattt acctaaatcc cttaacctca ccctctctac cgatccttcc      120
atcatctccg cagcctctca tgacttcgga aacataacca ccgtgacccc cggcggcgta      180
atctgcccct cctccaccgc tgatatctct cgtctcctcc aatacgccgc aaacggaaaa      240
agtacattcc aagtagcggc tcgtggccaa ggccactcct taaacggcca agcctcggtc      300
tccggcggag taatcgtcaa catgacgtgt atcactgacg tgggtggttc aaaagacaag      360
aagtagcgtg acgtggcggc cgggacgcta tgggtggatg tgcttaagaa gacggcggag      420
aaaggggtgt cgccggttcc ttggacggat tatttgcata taaccgtccg aggaacggtg      480
tcgaatgggt gaattgggtg tcaagtgtt cgaaacggtc ctcttgtag taacgtcctt      540
gaattggacg ttattactgg gaaaggtgaa atggtgacat gctcgcgaca gctaaacca      600
gaattgttct atggagtgtt aggaggtttg ggtcaatttg gaattataac gagagccaga      660
attgttttgg accatgcacc taaacgggcc aaatggtttc ggatgctcta cagtgatttc      720
acaacttta caaaggacca agaacgtttg atatcaatgg caaacgatat tggagtcgac      780
tatttagaag gtcaaatatt tctatcaaac ggtgtcgttg acacctctt tttcccacct      840
tcagatcaat ctaaagtcgc tgatctagtc aagcaacacg gtatcatcta tgttcttgaa      900
gtagccaagt attatgatga tcccaatctc cccatcatca gcaaggttat tgacacatta      960
acgaaaacat taagttactt gcccggttcc atatcaatgc acgacgtggc ctacttcgat     1020
ttcttgaaac gtgtacatgt cgaagaaaat aaactcagat ctttgggatt atgggaactt     1080
cctcatcctt ggcttaacct ctacgttccct aaatctcgga ttctcgattt tcataacggt     1140
gttgtcaaag acattcttct taagcaaaaa tcagcttcgg gactcgtctt tctctatcca     1200
acaaaccgga ataaatggga caatcgtatg tcggcgatga taccagagat cgatgaagat     1260
gttatatata ttatcggact actacaatcc gctaccccaa aggatcttcc agaagtggag     1320
agcgttaacg agaagataat taggttttgc aaggattcag gtattaagat taagcaatat     1380
ctaattgatt atactagtaa agaagattgg attgagcatt ttggatcaaa atgggatgat     1440
ttttcgaaga ggaaagatct atttgatccc aagaaactgt tatctccagg gcaagacatc     1500
ttttga                                           1506
    
```

<210> 43
 <211> 2746
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> módulo de expresión

15

<220>
 <221> característica_misc
 <222> (1913)..(1914)
 <223> n es a, c, g, o t

20

<400> 43

```

cttctacatc ggcttaggtg tagcaacacg actttattat tattattatt attattatta      60
ttattttaca aaaatataaa atagatcagt ccctcaccac aagtagagca agttggtgag      120
ttattgtaaa gttctacaaa gctaatttaa aagttattgc attaacttat ttcattattac      180
aaacaagagt gtcaatggaa caatgaaaac catatgacat actataaatt tgtttttatt      240
attgaaatta tataattcaa agagaataaa tccacatagc cgtaaagttc tacatgtggt      300
gcattaccaa aatatatata gcttacaaaa catgacaagc ttagtttgaa aaattgcaat      360
ccttatcaca ttgacacata aagtgagtga tgagtcataa tattattttc tttgctaccc      420
atcatgtata tatgatagcc acaaagttac tttgatgatg atatcaaaga acatttttag      480
gtgcacctaa cagaatatcc aaataatatg actcacttag atcataatag agcatcaagt      540
aaaactaaca ctctaagca accgatggga aagcatctat aaatagacaa gcacaatgaa      600
aatcctcatc atccttcacc acaattcaaa tattatagtt gaagcatagt agtaatttaa      660
atcaactagg gatatcacia gtttgtaaaa aaaagcaggc ttcacaatgg ctaatcttcg      720
ttaatgatc actttaatca cggttttaat gatcaccaaa tcatcaaacg gtattaaaat      780
tgatttacct aaatccctta acctcaccct ctctaccgat ccttccatca tctccgcagc      840
ctctcatgac ttcggaacaa taaccaccgt gacccccggc ggcgtaatct gccctcctc      900
caccgctgat atctctcgtc tctccaata cgccgcaaac ggaaaaagta cattccaagt      960
agcggctcgt ggccaaggcc actccttaa cggccaagcc tcggtctccg gcggagtaat      1020
cgtcaacatg acgtgtatca ctgacgtggt ggtttcaaaa gacaagaagt acgtgacgt      1080
ggcggccggg acgttatggg tggatgtgct taagaagacg gcggagaaag ggggtgcgcc      1140
ggtttcttgg acggattatt tgcataaac cgtcggagga acgttgcga atgggtggaat      1200
tgggtggtcaa gtgtttcgaa acggtcctct tgttagtaac gtccttgaat tggacgttat      1260
tactgggaaa ggtgaaatgt tgacatgctc ggcacagcta aaccagaat tgttctatgg      1320
agtgttagga ggtttgggtc aatttgggat tataacgaga gccagaattg ttttggacca      1380
tgcacctaaa cggcaagaac gtttgatata aatggcaaac gatattggag tgcactattt      1440
agaaggtcaa atatttctat caaacgggtg cgttgacacc tcttttttcc cacctcaga      1500
tcaatctaaa gtcgctgatc tagtcaagca acacgggatc atctatgttc ttgaagtagc      1560
caagtattat gatgatccca atctcccat catcagcaag gttattgaca cattaacgaa      1620
aacattaagt tacttgcccg ggttcatatc aatgcacgac gtggcctact tcgatttctt      1680
gaaccgtgta catgtcgaag aaaataaact cagatctttg ggattatggg aacttcctca      1740
tccttggctt aacctctacg ttctaaatc tcggattctc gattttcata acggtgttgt      1800
caaagacatt cttcttaagc aaaaatcagc ttcgggactc gctcttctct atccaacaaa      1860
ccggaataag tacatacttc tcttcattca tatttatctt caagaaccaa agnntggga      1920
caatcgtatg tcggcgatga taccagagat agatgaagat gttatatata ttatcgact      1980
actacaatcc gtaaccccaa aggatcttcc agaagtggag agcgttaacg agaagataat      2040
taggttttgc aaggattcag gtattaagat taagcaatat ctaatgcatt atactagtaa      2100
agaagattgg attgagcatt ttggatcaaa atgggatgat ttttcgaaga ggaaagatct      2160
atgtgatccc aagaaactgt tatctccagg gcaagacatc ttttgaacc agcttcttg      2220
tacaaagtgg tgatatcaca agcccgggcg gtottctagg gataacaggg taattatctc      2280
cctctagatc acaagcccgg gcggtcttct acgatgattg agtaataatg tgtcacgcat      2340
caccatgggt ggcagtgtca gtgtgagcaa tgacctgaat gaacaattga aatgaaaaga      2400
aaaaaagtac tccatctgtt ccaaattaaa attcatttta accttttaat aggtttatac      2460
aataattgat atatgttttc tgtatatgtc taatttgta tcatccgggc ggtcttctag      2520
ggataacagg gtaattatat ccctctagac aacacacaac aaataagaga aaaaacaaat      2580
aatattaatt tgagaatgaa caaaaggacc atatcattca ttaactcttc tccatccatt      2640
tccatttcac agttcgatag cgaaaaccga ataaaaaaca cagtaaatta caagcacaac      2700
aaatggtaca agaaaaacag ttttcccaat gccataatac tcgaac      2746

```

<210> 44

<211> 17

<212> PRT

5 <213> Secuencia artificial

<220>

<223> firma de proteína

10

<220>

<221> CARACTERÍSTICA_MISC

<222> (6)..(7)

<223> Xaa puede ser cualquier aminoácido hidrófobo de la lista: Gly, His, Leu, Arg, Thr, Trp o Tyr

<220>

<221> CARACTERÍSTICA_MISC

<222> (8)..(8)

<223> Xaa puede ser cualquier alcohol de la lista: Ser o Thr

5 <220>

<221> CARACTERÍSTICA_MISC

<222> (9)..(9)

<223> Xaa puede ser cualquier aminoácido de la lista: Val o Ile

10

<220>

<221> CARACTERÍSTICA_MISC

<222> (15)..(16)

<223> Xaa puede ser cualquier aminoácido pequeño de la lista: Ala, Asp, Gly, Asn, Thr o Val

15

<400> 44

His Thr Asp Tyr Leu Xaa Xaa Xaa Xaa Gly Gly Thr Leu Ser Xaa Xaa
1 5 10 15
Gly

<210> 45

<211> 19

20

<212> PRT

<213> Secuencia artificial

<220>

<223> firma de proteína

25

<220>

<221> CARACTERÍSTICA_MISC

<222> (1)..(1)

<223> Xaa puede ser cualquier aminoácido cargado de la lista: Asp, Glu o Arg

30

<220>

<221> CARACTERÍSTICA_MISC

<222> (4)..(4)

<223> Xaa puede ser cualquier aminoácido, preferiblemente de la lista: Tyr, Arg, Asn, Phe, Asp o His

35

<220>

<221> CARACTERÍSTICA_MISC

<222> (5)..(5)

<223> Xaa puede ser cualquier aminoácido muy pequeño de la lista: Ala, Gly o Ser

40

<220>

<221> CARACTERÍSTICA_MISC

<222> (6)..(6)

<223> Xaa puede ser cualquier aminoácido pequeño de la lista: Ala, Asp, Gly, Asn, Thr o Val

45

<220>

<221> CARACTERÍSTICA_MISC

<222> (7)..(7)

<223> Xaa puede ser cualquier aminoácido hidrófobo de la lista: Gly, His, Leu, Arg, Thr, Trp o Tyr

50

<220>

<221> CARACTERÍSTICA_MISC

<222> (9)..(9)

<223> Xaa puede ser cualquier aminoácido pequeño de la lista: Ala, Asp, Gly, Asn, Thr o Val

55

<220>

<221> CARACTERÍSTICA_MISC

<222> (15)..(16)

ES 2 469 590 T3

<223> Xaa puede ser cualquier aminoácido alifático de la lista: Val o Ile

<220>

<221> CARACTERÍSTICA_MISC

5

<222> (17)..(17)

<223> Xaa puede ser cualquier aminoácido pequeño de la lista: Ala, Asp, Gly, Asn, Thr o Val

<220>

<221> CARACTERÍSTICA_MISC

10

<222> (18)..(18)

<223> Xaa puede ser cualquier aminoácido de tipo vuelta de la lista: Arg o Asn

<400> 45

Xaa	Leu	Phe	Xaa	Xaa	Xaa	Xaa	Gly	Xaa	Leu	Gly	Gln	Phe	Gly	Xaa	Xaa
1				5					10					15	
Xaa	Xaa	Ala													

REIVINDICACIONES

5 1.- Un procedimiento para aumentar el número de semillas llenas de una planta, consistiendo dicho procedimiento en (i) introducir en una célula vegetal una molécula de ácido nucleico aislada que codifica una citoquinina oxidasa unida operablemente a un promotor que dirige específicamente la expresión en el embrión y/o la aleurona de una semilla, y (ii) aumentar el nivel y/o la actividad de dicha citoquinina oxidasa en el embrión y/o la aleurona de la semilla de una planta, en el que dicha molécula de ácido nucleico aislada que codifica una citoquinina oxidasa es un ácido nucleico que codifica la proteína representada por SEQ ID NO:36, o es una molécula de ácido nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con dicho ácido nucleico que codifica la proteína de SEQ ID NO:36, y en el que dicho promotor comprende los nucleótidos 1-1256 de SEQ ID NO:41.

15 2.- Un procedimiento de producción de una planta que tiene un mayor número de semillas llenas y una mayor expresión de una citoquinina oxidasa en el embrión y/o la aleurona de una semilla, consistiendo dicho procedimiento en (i) introducir en una célula vegetal una molécula de ácido nucleico aislada que codifica una citoquinina oxidasa unida operablemente a un promotor que dirige específicamente la expresión en el embrión y/o la aleurona de una semilla, y (ii) aumentar el nivel y/o la actividad de dicha citoquinina oxidasa en el embrión y/o la aleurona de la semilla de una planta, en el que dicha molécula de ácido nucleico aislada que codifica una citoquinina oxidasa es un ácido nucleico que codifica la proteína representada por SEQ ID NO:36, o es una molécula de ácido nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con dicho ácido nucleico que codifica la proteína de SEQ ID NO:36, y en el que dicho promotor comprende los nucleótidos 1-1256 de SEQ ID NO:41.

20 3.- Un procedimiento de producción de una planta que tiene un mayor número de semillas llenas, comprendiendo dicho procedimiento:

25 (a) introducir en una célula vegetal una construcción genética que comprende un ácido nucleico aislado que codifica una citoquinina oxidasa unida operablemente a un promotor que dirige la expresión específicamente en el embrión y/o la aleurona de una semilla, en el que dicha molécula de ácido nucleico aislada que codifica una citoquinina oxidasa es un ácido nucleico que codifica la proteína representada por SEQ ID NO:36, o es una molécula de ácido nucleico que codifica CKX2 que se hibrida específicamente en 0,1 a 1x SSC y SDS al 0,1% en p/v a 60 °C con dicho ácido nucleico que codifica la proteína de SEQ ID NO:36, y en el que dicho promotor comprende los nucleótidos 1-1256 de SEQ ID NO:41;

30 (b) regenerar una planta a partir de esta;

(c) hacer crecer la planta regenerada hasta la fecundación de la semilla; y

(d) seleccionar una planta con un mayor número de semillas llenas, comparado con una correspondiente planta de tipo salvaje.

35 4.- El uso de SEQ ID NO:41 para aumentar el número de semillas llenas de una planta.

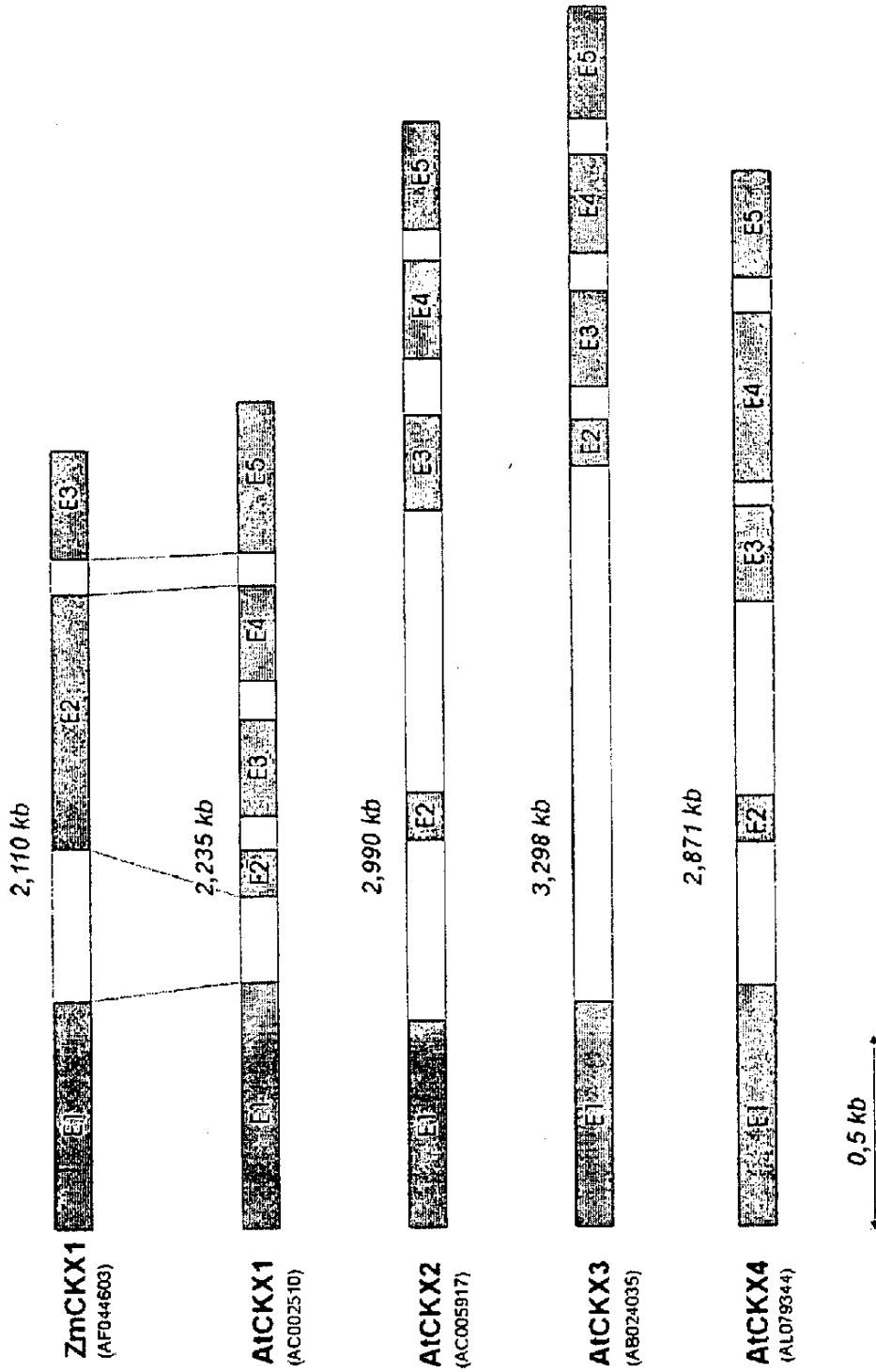


FIGURA 1

```

1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
ZmCKX1 1 -----MAVVYLLLAGLIACSHALAAGPALGDDRRGPPWASLAALADGKPRDSDNATAAASDFGNHTSALPAAVLYPSSTGDF
AtCKX1 1 MGLTSSLRHRQNNKFLGIFMILVLSICPGRTNLCNSHNSVSTPKELPSNPSPDIRSLSVSLDLEGIYSFDVHNVAKDDEGNRYOLFELAHFERSVFDH
AtCKX2 1 -----MANLRMITLITVIMITKSENGIKIDLPKSNLNTGSDPSISAASHDFGNHTVTPGCVIICPSSTADI
AtCKX3 1 MASNLRSOVRLIAITIVIIITLSPPTITNKSPOPWNILSHNEFAGKLSSSSSVESAATDRCHTKIFPSAVLHPSSEVEDI
AtCKX4 1 -----MTNLTCLSLITLITLIFSLTPTLIKDBEGIDVFLPISLNLVLTDPFSSAASHDFGNHTDENEGAVLCPSSITTEV

101 110 120 130 140 150 160 170 180 190 200
ZmCKX1 82 VALLSAANSTPG-----WPTIAERCGHSLMGQAFAPCGVVVNMASGDAAPRINVSADGRVYDAGGEQVHLDVIRASTARCVAPRVSNDYLY
AtCKX1 101 -----LTWAARGCHSLOGALAHQGVVIRKMESEERS--PDIRIYKGGQPIYDVSQGEIWMILREITLYKYLSPKSWFDYLDH
AtCKX2 70 -----STQVAARGCCHSINGOASVSGVIVNMTCITD-----VVSQKDKVYADAGTIVMVLVKKTAEKVSPVSMWTDYLDH
AtCKX3 83 -----PVAARGHSHRGOASAKDGVVNMRSVNV--RDRGIKVSRCLYVDVPAAMLHIEVINKITLITGLTPVSWTDYLY
AtCKX4 77 ARLERFANGGFSYNGKSTSPASTFKVAARGCCHSIRGOASAPGGVVVNMTCAMAAKPAAVVTSADGTYADVAAGTMMVLDVVKAAVDRGVSVPVWTDYLY

201 210 220 230 240 250 260 270 280 290 300
ZmCKX1 173 LTVGGTLSNAGISGQAFRRHGPOIENVLEMDVITGGHGVITCSKQINADLFDVVLGGGQFGMITRARIAPVPPAPARAKRWRVYHDFRAAFSADQERLTAP
AtCKX1 190 LTVGGTLSNAGISGQAFRRHGPOIENVYQLEIVTGKGEVITCSEKRNSELFSVLGGGQFGILITRARIISLEPAEHWKMPVLYSDFAFSRDOEVLISK
AtCKX2 154 LTVGGTLSNAGISGQAFRRHGQPLVSNVLEMDVITGKGEVITCSROINPELFFVGLGGQFGILITRARIIVLDHAPKRAKWRERMLYSDFTTETKDOERLISM
AtCKX3 171 LTVGGTLSNAGISGQAFRRHGQPLVSNVLEMDVITGKGEVITCSKQVNSDLFVVLGGGQFGILITRARIKLEVAAPKRAKWRFLYIDFSEFLRDOERVISK
AtCKX4 177 LSVGGTLSNAGISGQAFRRHGQPLVSNVLEMDVITGKGEVITCSKQVNSDLFVVLGGGQFGILITRARIIDHAPTRVKKSEILYSDFAFKRDOERLISM

301 310 320 330 340 350 360 370 380 390 400
ZmCKX1 273 RPOCGGASFGPSTYEGSVFVNOGLATDLANCFETDADVAIVLIVAGPQANATTVYSLEATLNYDNATAAAAADQELASLGLSLVTEGFAFORVAVYA
AtCKX1 290 EK-----TFDVAEGVYIINRTDLLNWMSSFSFSPINDSSTOASRPFKSDGKIVCLEVMKVNPE---EASSVDQETGKLSLNLNMTSTLFSSEVPIYI
AtCKX2 254 ANDIG-----VDMLEGOTFLSNG---VVDTSFPPSPQSKVADIVKQHGIVHLEVAKIVDDP---NLPITLISKVIDTITKTLVLPCHFSMDVAYF
AtCKX3 271 T--DG-----VDFJEGSIVVDHGPP-DNWRSTYVPPSDHLRIASGVKRRVRYCLEVMKIVDET---SOYTNNEEMEBESDLSLNRGFGMYEKDVTYM
AtCKX4 277 TNDIG-----VDFJEGQIMVSNG---FVDTSFPLSPQTRVASEVNDERIIVLEVAKIVDRT---TLPIEDQVIDITLSRITLCEAFGFMFVQDVPYF

401 410 420 430 440 450 460 470 480 490 500
ZmCKX1 373 AFIDRVGEEVALNKIGLNRPVPHPWLNIVPERSRIADPRGVFKGIIQOG-TDIVGPIIVVPLNKSMDDCMSAAATP--SEDFVYANSLIFSSVAP---N
AtCKX1 377 EFLDRVHIAERKLRKAKGLMVEVPHWLNILLPKSSIIYCFATEVFNNTITS--NNGGPIILYEVNOSKWKKHITSLTLP--NEDIFVLVAFEPESAVPNSSGKN
AtCKX2 340 DFLNRVHVEENKLRISIGLWELPHWLNIVPERSRIIDFINGVWKDILLKQKSASGLALHIFTRNKKQDNRMMSAWIPEIDEDVIYIIGLCSAATP-----K
AtCKX3 358 DFLNRVFTGELNLSKQGVVPHWLNIVPERSRIQISKDDGVFKGLLRNNTITSQVLIIVPMMSKNDRMMSAAIIP--EEDVYAVGCFERSAG-F-----D
AtCKX4 363 DFLNRVNEEDKLRISIGLWELVPHWLNIVPERSRIQDFHDGVINGILLNQTSTSGVITLYEYVTRNKKVNRMSIWTI--PDEDVYFVIGLICSAGGS-----Q

501 510 520 530 540 550 560 570 580 590 600
ZmCKX1 466 DLAPVQEQNSLIFCDLACIQYKVLVAHEDRSIMVRFHGAAMNREVEVEMKKNKVDPKRIISPGQDIF-----
AtCKX1 473 DLEVLLKQGRVMNFQAAANLVKQVLPHYETQKMKSHFG-KRMETBAQRQAVDPLARVAPGORIFRKTGKLSPIQLAKSKATCSPQRYHYASILPKPRTV
AtCKX2 435 DLPEVESVNEKLRIFCKDSGHLKIQKQIMLYTSKEDMTEHFG-SKWDDEFSKQKDLDFPKLLSPGQDIF-----
AtCKX3 451 NWEAFDQENMELIKFCEDANMCIQLIPYHSQEQEMRHFQ-PRNNIIVEYKRYKQDPKMLSPGNIEFKINSS-----
AtCKX4 457 NWQDENLNDKVIQFOENSGLIKHREVMYIRKEDMVKHFG-PKWDDELKRIKIMEIDPKRLLSPGQDIF-----

```

FIGURA 2

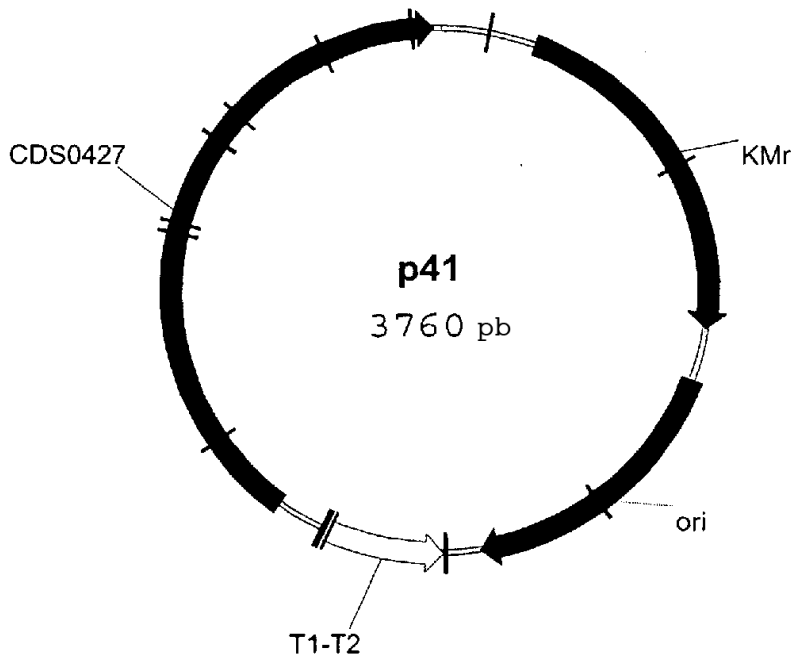


FIGURA 3

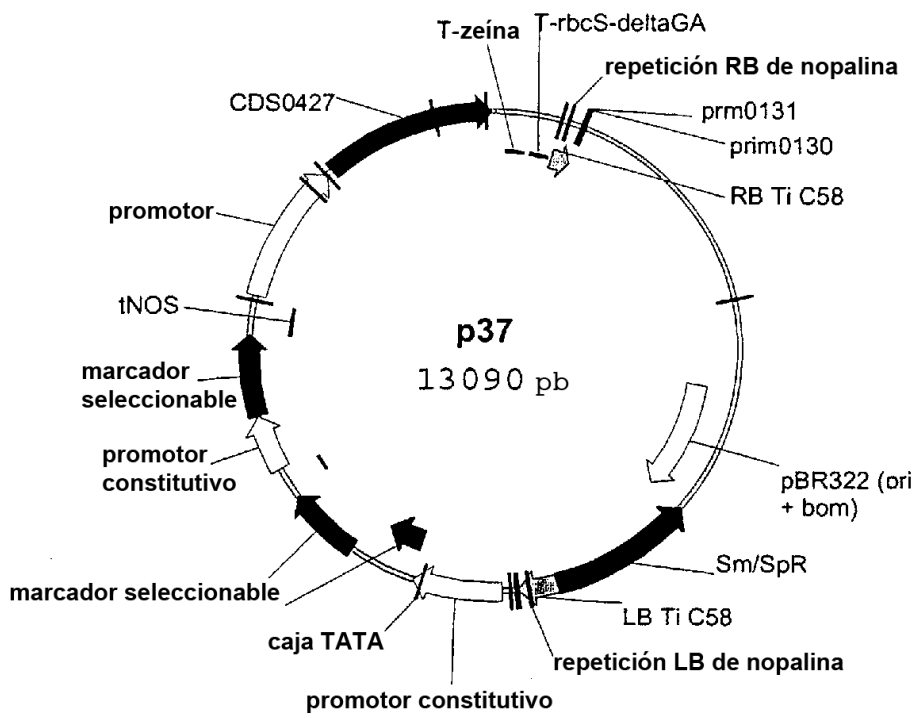


FIGURA 4

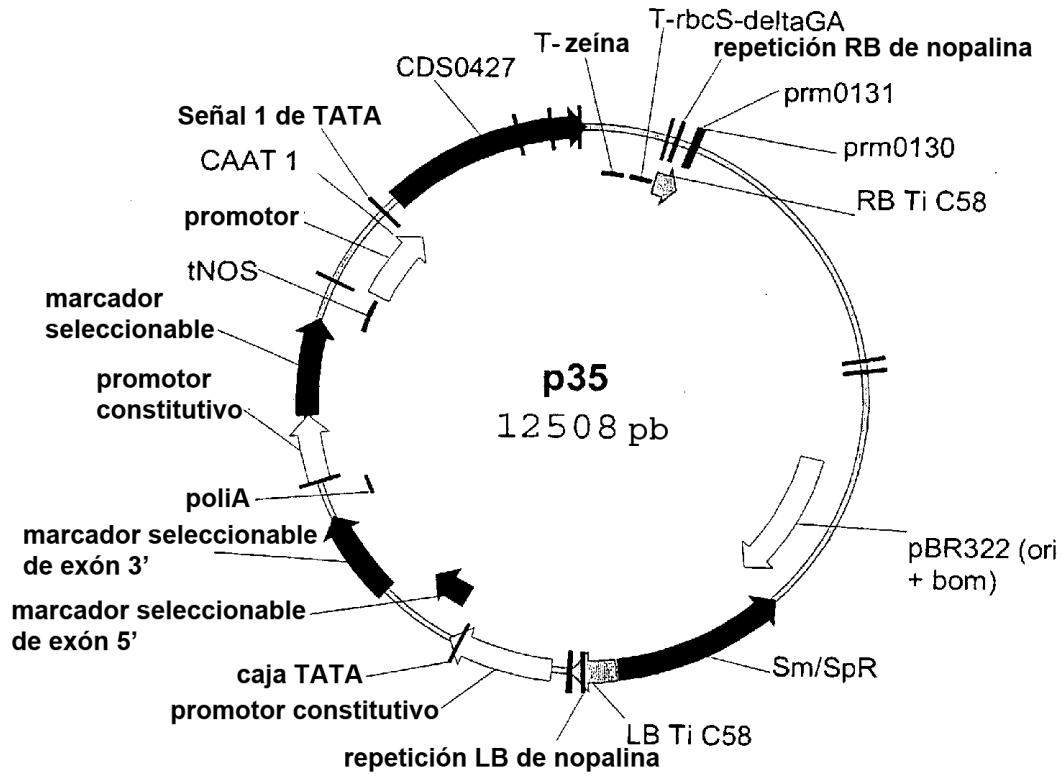


FIGURA 5