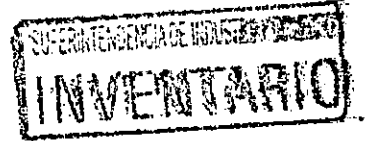


SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



PCT



DELEGATURA PROPIEDAD INDUSTRIAL

División de Nuevas Creaciones

11-9269

SOLICITANTE
PATENTE DE INVENCION

ENTRADA E

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269-00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

21. EXPEDIENTE No.

54. TITULO

PRODUCCION Y ALMACENAMIENTO DE PROTEINAS MEJORADOS

EN PLANTAS

51. CLASIFICACION INTERNACIONAL

DONALD DANFORTH PLANT SCIENCE CENTER y THE UNITED STATES OF AMERICA, AS represented by THE SECRETARY OF AGRICULTURE

71. SOLICITANTE

SAINT LOUIS, MO, US y WASHINGTON, D.C., US.

DOMICILIO

JUAN PABLO CADENA SARMIENTO

74. APODERADO

22. BOGOTÁ, D.C.

28-01-2011



SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269- -00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

ón

**FORMULARIO ÚNICO DE SOLICITUD DE PATENTE
PCT
ENTRADA EN FASE NACIONAL**

1

SOLICITUD DE:



Patente de Invención.



Patente de Modelo de Utilidad.



Capítulo I.



Capítulo II.

2

SOLICITANTE

Nombre: DONALD DANFORTH PLANT
SCIENCE CENTER y THE UNITED STATES
OF AMERICA, AS represented by THE
SECRETARY OF AGRICULTUREDirección: 975 North Warson Road Saint Louis,
MO 63132, US; Washington, D.C. 20250, US.Nacionalidad o Domicilio: Estados Unidos de
AméricaLugar de Constitución: Estados Unidos de
América

Teléfono:

Fax:

E-mail

IDENTIFICACIÓN

C.C.



NIT



C.E.



Otro



Cual

Número

3

REPRESENTANTE
O APODERADO

Nombre: JUAN PABLO CADENA SARMIENTO

Dirección: Calle 70A #4 - 41, Bogotá, Colombia.

Teléfono: 744-2200

Fax: 742-2200

E-mail:

IDENTIFICACIÓN

C.C.



NIT



C.E.



Otro



Cual

Número

80.410.337 TP 57492

En caso de haber presentado poder general para asuntos que se adelanten ante la Delegatura
de Propiedad Industrial, sírvase indicar el numero de radicación (del poder). _____

4

INVENTOR (ES)
(72)

Nombre: Eliot M.HERMAN, Monica A. SCHMIDT.

Dirección: 315 Meramec #33, Clayton, MO 63105, US; 8140 Whitburn Drive
Clayton, MO 63105, US (respectivamente).

Nacionalidad o Domicilio: Estados Unidos de América (respectivamente).

5 PCT (86)

Solicitud Internacional No.

PCT/US2009/049097

Fecha

29.06.2009

Publicación Internacional No

WO 2009/158716 A1

Fecha

30.12.2009

6 Título (54): PRODUCCION Y ALMACENAMIENTO DE PROTEINAS MEJORADOS EN PLANTAS

7 Clasificación Internacional (51): _____

8 Prioridad

(33) País de Origen

(32) Fecha

(31) Número de Solicitud

US

28.06.2008

61/076,616

SI ☒NO ☐

9 Comprobante de pago No.

10-102.847; 10-100.368; 10-100.364;
10-100.337; 10-84.599; 11-2.721

Fecha

26-01-2011

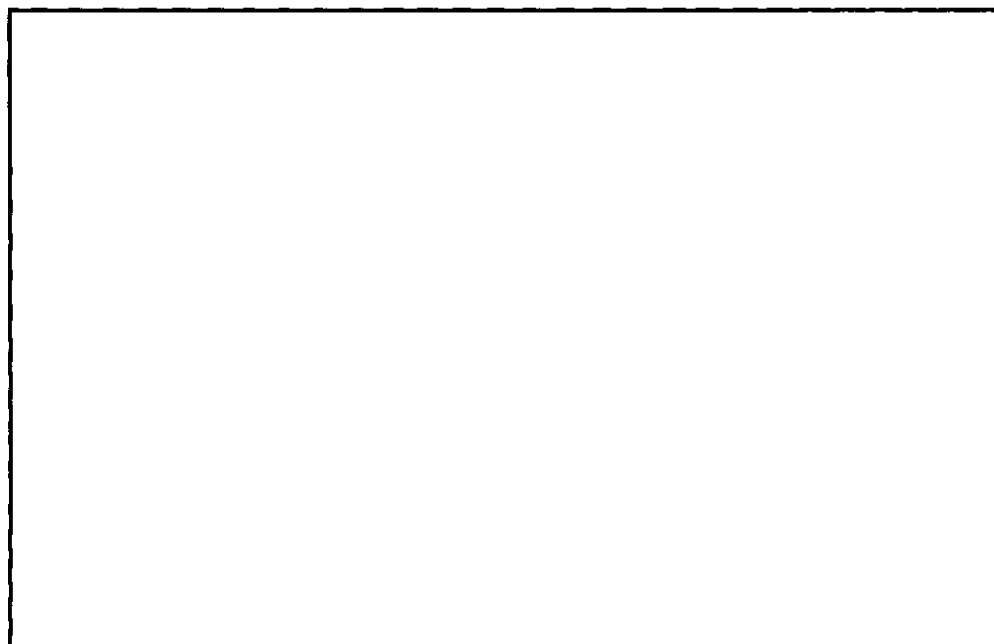
Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página.

10

ANEXOS

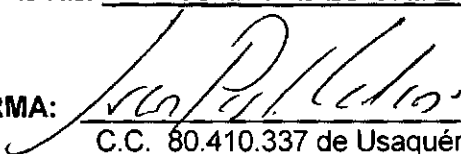
- ☒ Comprobante de pago de la tasa de presentación de la solicitud. \$ 571.000
- ☐ Documento que acredita la existencia y representación legal de la persona jurídica peticionaria, en cuanto no estén inscritas en alguna Cámara de Comercio Colombiana.
- ☐ Poderes, si fuere el caso. **Fotocopia autenticada.**
- ☐ Documento de cesión del inventor al solicitante o a su causante. **Fotocopia autenticada.**
- ☒ Traducción de la solicitud internacional al castellano (Resumen, descripción, reivindicaciones, dibujos).
- ☒ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen preliminar (IPEA).
- ☐ Traducción de examen preliminar internacional efectuado por la IPEA.
- ☐ De ser el caso, copia del contrato de acceso.
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- ☐ De ser el caso, certificado de depósito del material biológico.
- ☐ Arte final 12x12.
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☒ De ser el caso, modificaciones efectuadas a la solicitud internacional bajo el Art. 34.
- ☒ Otros: - Comprobante(s) de pago de la tasa por concepto de Reivindicaciones adicionales \$ 459.000 (17*\$27.000).

11 FIGURA CARACTERÍSTICA:



12 Solicito la concesión de la patente,

NOMBRE: JUAN PABLO CADENA SARMIENTO

FIRMA: 
C.C. 80.410.337 de Usaquén TP 57492 del C.S.J.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800176089-2

~~114047~~

115611

RECIBO OFICIAL DE CAJA : 10 - 102,847
FECHA : DICIEMBRE 2 DE 2010

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269-00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
BRIGARD & CASTRO LTDA	CONSIGNACION	BANCO DE BOGOTA	062754387	392749382	02/12/2010	27,674,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE IN	554,000.00
		TOTAL :	\$ 554,000.00

SON: QUINIENTOS CINCUENTA Y CUATRO MIL PESOS

RESPONSABLE :

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

115611 2

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800176089-2

CERTIFICACION DE EXCEDENTE : 10 - 100,368
FECHA : NOVIEMBRE 25 DE 2010

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269- -00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eje: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

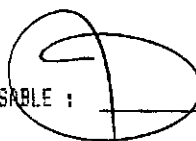
***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
BRIGARD & CASTRO LTDA	CONSIGNACION	BANCO DE BOGOTA	062754387	329790456	25/11/2010	53,309,000.00

***** CONCEPTO *****

CANT. RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1 50005-01-09 OTROS SERVICIOS DE PROPIEDAD I	99 OTROS SERVICIOS ADMINISTRATIVOS	260,000.00
	TOTAL :	\$ 260,000.00

SON: DOSCIENTOS SESENTA MIL PESOS

RESPONSABLE : 

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

ESTA CERTIFICACION DE EXCEDENTE DE PAGO SOLO PODRA SER UTILIZADA
COMO PARTE DE LA CANCELACION DE UNA TARIFA VIGENTE Y NO PODRA SER
SUSTITUIDA POR NINGUN RECIBO.

115611

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

MIT : B00176089-2

CERTIFICACION DE EXCEDENTE : 10 - 100,364
 FECHA : NOVIEMBRE 25 DE 2010

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269-00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
 Tra. 11 PATENTEIYII Eve: 379 FASENACIONALI
 Act. 411 PRESENTACION Folios: 167

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
BRIGARD & CASTRO LTDA	CONSIGNACION	BANCO DE BOGOTA	062754387	329790486	25/11/2010	58,309,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-09	OTROS SERVICIOS DE PROPIEDAD I	
		99 OTROS SERVICIOS ADMINISTRATIVOS	130,000.00
		TOTAL :	\$ 130,000.00

SON: CIENTO TREINTA MIL PESOS

RESPONSABLE :

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

ESTA CERTIFICACION DE EXCEDENTE DE PAGO SOLO PODRA SER UTILIZADA
 COMO PARTE DE LA CANCELACION DE UNA TARIFA VIGENTE Y NO PODRA SER
 SUSTITUIDA POR NINGUN RECIBO.

11560.16

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

MIT : 900176089-2

CERTIFICACION DE EXCEDENTE : 10 - 100,337
FECHA : NOVIEMBRE 25 DE 2010

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269- -00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	Nº. PAGO	FECHA PGO	Vr. PAGO
BRIGARD & CASTRO LTDA	CONSIGNACION	BANCO DE BOGOTA	062754387	329790486	25/11/2010	53,309,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-09	OTROS SERVICIOS DE PROPIEDAD I	
99		OTROS SERVICIOS ADMINISTRATIVOS	52,000.00
		TOTAL :	\$ 52,000.00

SON: CINCUENTA Y DOS MIL PESOS

RESPONSABLE :

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

ESTA CERTIFICACION DE EXCEDENTE DE PAGO SOLO PODRA SER UTILIZADA
COMO PARTE DE LA CANCELACION DE UNA TARIFA VIGENTE Y NO PODRA SER
SUSTITUIDA POR NINGUN RECIBO.

115611 7

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800176089-2

CERTIFICACION DE EXCEDENTE : 10 - 84,599
FECHA : OCTUBRE 6 DE 2010

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269- -00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

**** CONSIGNACION ****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
ERIBARD & CASTRO LTDA	CONSIGNACION	BANCO DE BOGOTA	062754387	382752226	07/10/2010	34,127,400.00

**** CONCEPTO ****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD	99 OTROS SERVICIOS ADMINISTRATIVOS	24,000.00
TOTAL :			\$ 24,000.00

SON: VEINTICUATRO MIL PESOS

RESPONSABLE :

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

ESTA CERTIFICACION DE EXCEDENTE DE PAGO SOLO PODRA SER UTILIZADA
COMO PARTE DE LA CANCELACION DE UNA TARIFA VIGENTE Y NO PODRA SER
SUSTITUIDA POR NINGUN RECIBO.

115611 A

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

-/-

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 002721

Bogotá D.C., 13 de 2011 - 07:45:52

RECIBIDO DE : BRIGARD & CASTRO S.A.

NI 860.047.372

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	391765861	12/01/2011	61.870.100.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNIDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	10.000.00	10.000.00
				\$10.000.00

SON: **DIEZ MIL PESOS MONEDA CORRIENTE**

Responsable: 

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269- -00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 167

PATENT COOPERATION TREATY

PCT/US2009/049097

From the INTERNATIONAL BUREAU

PCTNOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

RANDALL, David, A.
Dickstein Shapiro LLP
2049 Century Park East Suite 700
Los Angeles, CA 90067-3109
ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year)

06 January 2011 (06.01.2011)

Applicant's or agent's file reference

AGRIUS228001

International application No.

PCT/US2009/049097

IMPORTANT NOTIFICATION

International filing date (day/month/year)

29 June 2009 (29.06.2009)

1. The following indications appeared on record concerning:

☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative

Name and Address

UNITED STATES DEPARTMENT OF AGRICULTURE
5601 Sunnyside Avenue
4-1159
Beltsville, MD 20705
United States of America

State of Nationality

US

State of Residence

US

Telephone No.

Facsimile No.

E-mail address

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☒ the name ☒ the address ☐ the nationality ☐ the residence

Name and Address

THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE
DEPARTMENT OF AGRICULTURE
Washington, D.C. 20250
United States of America

State of Nationality

US

State of Residence

US

Telephone No.

Facsimile No.

E-mail address

☐ Notifications by e-mail authorized

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☐ the Authority(ies) specified for supplementary search☐ the International Preliminary Examining Authority
☐ the designated Offices concerned
☒ the elected Offices concerned
☐ other:The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Authorized officer

Sontag Frederic

e-mail pt12.pct@wipo.int

Telephone No. +41 22 338 74 12

Facsimile No. +41 22 338 70 90

PATENT COOPERATION TREATY

PCT/US2009/049097

From the INTERNATIONAL BUREAU

PCTNOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

RANDALL, David, A.
Dickstein Shapiro LLP
2049 Century Park East Suite 700
Los Angeles, CA 90067-3109
ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year) 14 January 2011 (14.01.2011)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference AGRIUS228001	
International application No. PCT/US2009/049097	International filing date (day/month/year) 29 June 2009 (29.06.2009)

1. The following indications appeared on record concerning:

☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative

Name and Address THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE DEPARTMENT OF AGRICULTURE Washington, D.C. 20250 United States of America	State of Nationality US	State of Residence US
	Telephone No.	
	Facsimile No.	
	E-mail address	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☒ the name ☐ the address ☐ the nationality ☐ the residence

Name and Address THE UNITED STATES OF AMERICA, AS represented by THE SECRETARY OF AGRICULTURE Washington, D.C. 20250 United States of America	State of Nationality US	State of Residence US
	Telephone No.	
	Facsimile No.	
	E-mail address ☐ Notifications by e-mail authorized	

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☐ the Authority(ies) specified for supplementary search☐ the International Preliminary Examining Authority
☐ the designated Offices concerned
☒ the elected Offices concerned
☐ other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No. +41 22 338 70 90	Authorized officer Clemente Rodolfo e-mail pt12.pct@wipo.int Telephone No. +41 22 338 74 12
---	--

PATENT COOPERATION TREATY

PCT/US2009/049097

From the INTERNATIONAL BUREAU

PCTNOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

RANDALL, David, A.
Dickstein Shapiro LLP
2049 Century Park East Suite 700
Los Angeles, CA 90067-3109
ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year) 06 January 2011 (06.01.2011)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference AGRIUS228001	
International application No. PCT/US2009/049097	International filing date (day/month/year) 29 June 2009 (29.06.2009)

1. The following indications appeared on record concerning:

☒ the applicant ☒ the inventor ☐ the agent ☐ the common representative

Name and Address SCHMIDT, Monica, S. 975 North Warson Road Saint Louis, MO 63132 United States of America	State of Nationality US	State of Residence US
	Telephone No.	
	Facsimile No.	
	E-mail address	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person ☒ the name ☒ the address ☐ the nationality ☐ the residence

Name and Address SCHMIDT, Monica, A. 8140 Whitburn Drive Clayton, MO 63105 United States of America	State of Nationality US	State of Residence US
	Telephone No.	
	Facsimile No.	
	E-mail address ☐ Notifications by e-mail authorized	

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☐ the Authority(ies) specified for supplementary search☐ the International Preliminary Examining Authority
☐ the designated Offices concerned
☒ the elected Offices concerned
☐ other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No. +41 22 338 70 90	Authorized officer Sontag Frederic e-mail pt12.pct@wipo.int Telephone No. +41 22 338 74 12
---	---

PATENT COOPERATION TREATY

PCT/US2009/049097

From the INTERNATIONAL BUREAU

PCTNOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

To:

RANDALL, David, A.
Dickstein Shapiro LLP
2049 Century Park East Suite 700
Los Angeles, CA 90067-3109
ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year) 06 January 2011 (06.01.2011)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference AGRIUS228001	
International application No. PCT/US2009/049097	International filing date (day/month/year) 29 June 2009 (29.06.2009)

1. The following indications appeared on record concerning:		
☒ the applicant	☒ the inventor	☐ the agent
☐ the common representative		
Name and Address HERMAN, Eliot, M. 975 North Warson Road Saint Louis, MO 63132 United States of America	State of Nationality US	State of Residence US
	Telephone No.	
	Facsimile No.	
	E-mail address	
2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:		
☐ the person	☐ the name	☒ the address
☐ the nationality		
☐ the residence		
Name and Address HERMAN, Eliot, M. 315 Meramec #33 Clayton, MO 63105 United States of America	State of Nationality US	State of Residence US
	Telephone No.	
	Facsimile No.	
	E-mail address ☐ Notifications by e-mail authorized	
3. Further observations, if necessary:		
4. A copy of this notification has been sent to:		
☒ the receiving Office	☐ the International Preliminary Examining Authority	
☐ the International Searching Authority	☐ the designated Offices concerned	
☐ the Authority(ies) specified for supplementary search	☒ the elected Offices concerned	
☐ other:		

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Sontag Frederic e-mail pt12.pct@wipo.int Telephone No. +41 22 338 74 12
---	---

- (51) International Patent Classification:
C12N 15/11 (2006.01) C12N 5/14 (2006.01)

(21) International Application Number:
PCT/US2009/049097

(22) International Filing Date:
29 June 2009 (29.06.2009)

(25) Filing Language:
English

(26) Publication Language:
English

(30) Priority Data:
61/076,616 28 June 2008 (28.06.2008) US

(71) Applicants (for all designated States except US): THE DONALD DANFORTH PLANT SCIENCE CENTER [US/US]; 975 North Warson Road, Saint Louis, MO 63132 (US). UNITED STATES DEPARTMENT OF AGRICULTURE [US/US]; 5601 Sunnyside Avenue, 4-1159, Beltsville, MD 20705 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): HERMAN, Elliot, M. [US/US]; 975 North Warson Road, Saint Louis, MO 63132 (US). SCHMIDT, Monica, S. [US/US]; 975 North Warson Road, Saint Louis, MO 63132 (US).

(74) Agent: RANDALL, David, A.; Jones Day, 555 South Flower Street, Fifteenth Floor, Los Angeles, CA 90071 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: IMPROVED PROTEIN PRODUCTION AND STORAGE IN PLANTS

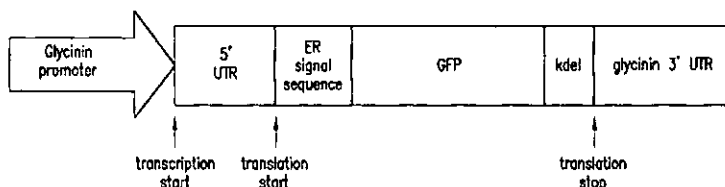


FIG. 7

(57) Abstract: A transgenic dicotyledonous plant having a deficiency of one or more plant seed storage proteins, further having a transgenic polynucleotide construct comprising an open reading frame operably linked to a storage protein promoter and an ER signal sequence. The polynucleotide construct encodes a protein product that can accumulate at high levels in the seed. Also provided are methods of producing a heterologous protein in a plant seed.

(19) Organización Mundial De La Propiedad Intelectual
Oficina Internacional

(10) Número de Publicación Internacional:

(43) Fecha de publicación internacional:
30 diciembre 2009 (30.12.2009)

PCT WO 2009/158716 A1

(51) Clasificación⁷ de patente internacional:
C12N 15/11 (2006.01) C12N 5/14 (2006.01)

(81) Estados designados: (a menos que se indique lo contrario, para cada tipo de protección nacional disponible): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(21) Número de solicitud internacional: PCT/US2009/049097

(22) Fecha de presentación internacional: 29 junio 2009 (29.06.2009)

(25) Idioma de Presentación: Inglés

(26) Idioma de Publicación: Inglés

(30) Fecha de Prioridad:
61/076,616 28 Junio 2008 (28.06.2008) US

(84) Estados designados: (a menos que se indique lo contrario, para cada tipo de protección regional disponible): Patente ARIPO (BW, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Patente Euroasiática (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Patente Europea (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), Patente OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Solicitante e (para todos los Estados designados excepto US): THE DONALD DANFORTH PLANT SCIENCE CENTER [US/US]; 975 North Warson Road, Saint Louis, MO 63132 (US). UNITED STATES DEPARTMENT OF AGRICULTURE [US/US]; 5601 Sunnyside Avenue, 4-1159, Beltsville, MD 20705 (US).

Publicada:
-- con el reporte de búsqueda internacional (Art. 21(3))
-- con parte del listado de secuencias de la descripción (Regla 5.2(a))

(72) Inventores; e
(75) Inventores/Solicitantes (para US solamente): HERMAN, Eliot, M. [US/US]; 975 North Warson Road, Saint Louis, MO 63132 (US). SCHMIDT, Monica, S. [US/US]; 975 North Warson Road, Saint Louis, MO 63132 (US).

(74) Agentes: RANDALL, David, A.; Jones Day, 555 South Flower Street, Fiftieth Floor, Los Angeles, CA 90071 (US).

(54) Título: PRODUCCION Y ALMACENAMIENTO MEJORADOS DE PROTEINAS EN PLANTAS

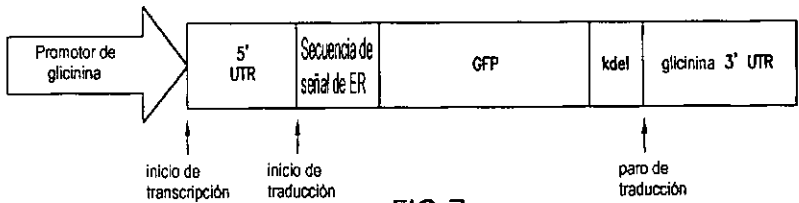


FIG. 7

(57) Resumen: Se describe una planta dicotiledónea transgénica que tiene una deficiencia de una o más proteínas de almacenamiento de semilla vegetal, que tiene además una construcción de polinucleótido transgénica que comprende un marco de lectura abierto enlazado operablemente a un promotor de proteína de almacenamiento y una secuencia de señal ER. La construcción de polinucleótido codifica para un producto de proteína que puede acumularse a altos niveles en la semilla. Se proporcionan también métodos para producir una proteína heteróloga en una semilla vegetal.

PRODUCCION Y ALMACENAMIENTO MEJORADOS DE PROTEINAS EN PLANTAS

Referencias Cruzadas de las Solicitudes Relacionadas

Esta solicitud reclama prioridad a la Solicitud Provisional de los Estados Unidos con no. de serie 61/076,616, presentada el 28 de junio de 2008, la cual se incorpora en la presente como referencia.

Declaración Relacionada con la Investigación o el Desarrollo Patrocinado

Federalmente

Por una parte, esta investigación se realizó con la ayuda del gobierno de los Estados Unidos de acuerdo con la cesión de investigación especial #2006-06103 otorgada por el USDA y administrada por la Universidad de Missouri, Columbia y por la ayuda de investigación intramuros del Servicio de Investigación Agrícola de USDA, proyecto número "3622-210000-025-00. El gobierno de los Estados Unidos puede tener ciertos derechos en la invención.

Nombres y Partes que se Unirán al Acuerdo de Investigación

Esta invención que resulta de un acuerdo de unión de investigación entre el Centro de Ciencias de las Plantas de Donald Danforth y el gobierno de los Estados Unidos, como es representado por el Departamento de Agricultura de los Estados Unidos, el Servicio de Investigación Agrícola ("USDA" por sus siglas en inglés). Así, el gobierno de los Estados Unidos puede tener ciertos derechos en la invención.

Campo de la Invención

La presente invención se refiere al campo de genética vegetal. Más específicamente, la presente invención se refiere a construcciones y métodos genéticos para usar las construcciones para modificar semillas vegetales y producir una cantidad incrementada de una proteína de interés.

Antecedentes de la Invención

Las semillas proporcionan una fuente importante de proteína dietaria para humanos y ganado. Ciertos tipos de semillas, tales como soya, son capaces de acumular un nivel relativamente alto de proteína endógena, haciendo la soya una buena opción para modificación genética para producir proteínas introducidas. A pesar de la

disponibilidad de muchas herramientas moleculares, la modificación genética de semillas comúnmente es restringida por una acumulación insuficiente de la proteína manipulada. Muchos procesos intracelulares pueden impactar la acumulación de proteínas total, incluyendo transcripción, traducción, ensamble y doblez de proteínas, metilación, fosforilación, transporte y proteólisis. La intervención en uno o más de estos procesos puede incrementar la cantidad de proteína producida en semillas genéticamente manipuladas.

La introducción de un gen puede causar efectos dañinos en el crecimiento y desarrollo de las plantas. Bajo tales circunstancias, la expresión del gen puede tener que ser limitada al tejido objetivo deseado. Por ejemplo, podría ser necesario expresar un gen de desregulación de aminoácidos de una manera específica de semilla para evitar un fenotipo no deseado que pudiera afectar el rendimiento u otros rasgos agronómicos.

Las semillas de soya contienen de 35% a 43% de proteína sobre una base de peso seco; la mayoría de esta proteína es proteína de almacenamiento. Hay dos proteínas de almacenamiento de semilla de soya principales: glicinina (también conocida como las globulinas 11S) y beta-conglicinina (conocidas también como las globulinas 7S). Juntas, comprenden 70 a 80% de la proteína total de la semilla, o 25 a 35% del peso en seco de la semilla.

La glicinina es una proteína grande con un peso molecular de aproximadamente 360 kDa. Es un hexámero compuesto de las diferentes combinaciones de cinco subunidades mayores identificadas como G1, G2, G3, G4 y G5.

La beta-conglicinina es una glicoproteína heterogénea con un peso molecular que varía de 150 y 240 kDa. Está compuesta de combinaciones variables de tres subunidades altamente cargadas negativamente identificadas como alfa, alfa' y beta.

Kinney y Herman ("Cosuppression of the α Subunits of beta-conglycinin in Transgenic Soybean Seeds Induces the Formation of Endoplasmic Reticulum-Derived Protein Bodies" Plant Cell 13:1165-1178 (2001)) reportan que la transformación con una construcción que contiene una región que puede transcribirse a una secuencia líder no traducida 5' de beta-conglicinina se traduce en una reducción de las subunidades alfa y alfa' de la proteína beta-conglicinina. Kinney y Herman hacen notar que "el incremento

en la proteína beta-conglicinina fue aparentemente compensado por una acumulación incrementada de glicinina y otras proteínas vacuolares en el ER" llevándoles a especular que "tal vez al co-expresar otras proteínas, posiblemente como una proteína de fusión de glicinina con un separador escindible. Puede ser posible configurar soyas para expresar y

5 acumular a altos niveles proteínas extrañas que requieran de eventos de doblez y procesamiento mediados por ER". Kinney y Herman no enseñan reducir la expresión de beta-conglicinina en una semilla de soya mientras se expresa bajo el control del promotor de glicinina, una proteína extraña fusionada a un péptido de señal ER.

Oulmassov *et al.* (solicitud de patente de E.U.A. 2005/0079494) describe la

10 expresión de glicinina mutada bajo el control de un promotor, tal como un promotor de glicinina. Oulmassov *et al.* describen además la supresión mediada antisentido de secuencias que contienen un bajo contenido de aminoácidos esenciales, pero que son expresadas a niveles relativamente altos en tejidos particulares, tales como beta-conglicinina y glicinina. Oulmassov *et al.* Wu no enseñan ninguna consecuencia posible

15 de reducir la expresión con beta-conglicinina con respecto a la expresión y acumulación de proteínas que son expresadas bajo el control de un promotor de glicinina ni tampoco enseñan alguna consecuencia posible de suprimir la expresión de beta-conglicinina en una semilla de soya mientras se expresa bajo el control del promotor de glicinina, una proteína extraña fusionada a un péptido de señal ER.

20 Wu (solicitud de patente de E.U.A. 2007/0067871) describe proporcionar una soya con un contenido de beta-conglicinina en semilla incrementado, que comprende mutaciones no transgénicas que proporcionan un fenotipo nulo de al menos dos de las subunidades de glicinina. Uno enseña reducir la expresión de beta-conglicinina o glicinina, o la expresión de una proteína exógena.

25 Lo que se requiere en la técnica es un método para usar soya para producir altos niveles de una proteína de interés, tal como para alimento, combustible, pienso, enzimas industriales, enzimas de bioprocesamiento, vacunas, proteínas terapéuticas, anticuerpos o similares.

Breve Descripción de la Invención

Se proporcionan en la presente métodos para producir cantidades incrementadas de una proteína heteróloga de interés en una semilla. En una modalidad, suprimir genéticamente la producción de una proteína de semilla causa que la semilla reequilibre su composición proteínica al incrementar la producción de otras proteínas para mantener un contenido de proteína de semilla normal. Este efecto puede ser combinado con el uso de un "imitador de alelo" de los genes que sean subregulados para reequilibrar el contenido proteínico y de esta manera conducir la expresión de la proteína heteróloga.

En una modalidad, se proporciona en la presente una dicotiledónea transgénica que tiene una deficiencia de una o más proteínas de almacenamiento vegetal y un polinucleótido heterólogo que tiene un marco de lectura abierto enlazado operablemente a un promotor de proteína de almacenamiento y una secuencia de señal ER. Opcionalmente, el polinucleótido heterólogo comprende además un dominio potenciador de traducción 5' (TED) y/o un 3' TED. Opcionalmente, el polinucleótido heterólogo comprende además una secuencia de retención de ER para inducir la acreción del polinucleótido heterólogo en el lumen del ER o una vesícula derivada del ER.

En otra modalidad, se proporciona una planta dicotiledónea transgénica que tiene una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se traduce en una reducción de al menos 50% en proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre, y un polinucleótido heterólogo que tiene un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que tiene una secuencia de señal ER, una secuencia de codificación de proteína deseada y una señal de retención de ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla, y en donde la semilla de la planta transgénica es capaz de producir una proteína heteróloga a un nivel que sea más alto que el 5% del peso en seco total de la semilla. El polinucleótido heterólogo también puede tener un dominio potenciador de traducción 5' y/o un dominio potenciador de traducción 3'. La secuencia de retención ER puede inducir la acreción de la proteína heteróloga en el lumen de ER o una vesícula derivada del ER. La dicotiledónea puede ser, por ejemplo, un miembro de la

familia de las Fabáceas, y opcionalmente del orden Fabales, y opcionalmente del género soya. La dicotiledónea puede ser un miembro del género glicina, tal como una soya. El promotor puede derivarse, por ejemplo, del inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglucina. El dominio potenciador de traducción puede ser derivado de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglucina. La deficiencia en proteína de almacenamiento puede ser, por ejemplo, uno o más de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglucina. En una modalidad, la deficiencia puede deberse a, por ejemplo, la presencia de un ARNi, un fragmento de antisentido o sentido de un ácido nucleico que codifique para una proteína de almacenamiento de semilla.

La semilla dicotiledónea provista en la presente puede tener, por ejemplo, una deficiencia de más del 75% de las proteínas de almacenamiento endógenas de la semilla. La proteína heteróloga puede acumularse en una semilla de la dicotiledónea hasta un nivel que sea mayor que aproximadamente 2% o mayor de aproximadamente 4% o mayor de aproximadamente 5% del peso en seco total de la semilla. En otra modalidad, se proporciona una semilla transgénica, o una proteína obtenida de la semilla. La proteína heteróloga puede ser purificada.

En una modalidad, la secuencia de codificación de proteína codifica para una enzima o fragmento de la misma. La enzima puede ser una enzima celulolítica, tal como una β -glucosidasa, una exoglucanasa I, una exoglucanasa II, una endoglucanasa, una xilanasa, una hemicelulasa, una ligninasa, una lignina peroxidasa o una manganeso peroxidasa. En una modalidad, se proporciona una composición enzimática comercialmente útil.

En una modalidad, se proporciona en la presente una planta dicotiledónea transgénica, que tiene una deficiencia de una o más proteínas de almacenamiento vegetal endógenas, en donde la deficiencia se traduce en una reducción de al menos 50% en el

nivel de la proteína de almacenamiento vegetal endógena en comparación con una planta tipo silvestre, y un polinucleótido heterólogo que tiene una región reguladora génica de una proteína compensadora enlazada operablemente a un marco de lectura abierto que codifica para una secuencia que tiene una secuencia de señal ER, una secuencia de
5 codificación de proteína deseada y una secuencia de retención ER, en donde la semilla de la planta dicotiledónea transgénica es capaz de producir la proteína heteróloga a un nivel que sea mayor que el 5% del peso en seco total de la semilla.

En otra modalidad, se proporciona en la presente un método para almacenar establemente una enzima antes de usar, al almacenar la enzima en una
10 semilla proveniente de una planta dicotiledónea transgénica que tiene una deficiencia de una o más proteínas de almacenamiento vegetal, en donde la deficiencia se traduce en una reducción de al menos 50% en la proteína de semilla endógena, y un polinucleótido heterólogo que tiene un promotor de proteína de almacenamiento de semilla, un marco de lectura de abierto que tiene un ácido nucleico que codifica para una secuencia de señal
15 ER, una enzima de interés y una señal de retención ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla, y en donde la semilla de la planta transgénica es capaz de producir la enzima a un nivel que es mayor que 5% del peso en seco total de la semilla, y almacenar la enzima en la semilla de la dicotiledónea transgénica.

20 En otra modalidad más, se proporciona en la presente un método para producir una cantidad incrementada de una proteína heteróloga en una planta dicotiledónea, que incluye transformar establemente una célula vegetal con un polinucleótido que tenga un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que tenga una secuencia de señal ER, una secuencia de
25 codificación de proteína deseada y una secuencia de retención ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla, obtener una línea vegetal homocigótica de la célula vegetal transformada establemente, introducir la línea vegetal transformada establemente en una planta que tenga una deficiencia en una proteína de almacenamiento de semilla endógena, en donde
30 la deficiencia se traduce en una reducción de al menos 50% en la proteína de

almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre, cultivar las semillas de la planta transgénica introducida, y obtener la proteína heteróloga de las semillas de la planta transgénica introducida, en donde la semilla de la planta transgénica introducida es capaz de producir una proteína heteróloga a un nivel que es más alto que el 5% del peso en seco total de la semilla. La deficiencia en una proteína de almacenamiento de semilla endógena puede ser debida a, por ejemplo, la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que codifique para una proteína de almacenamiento de semilla.

En otra modalidad más, se proporciona un método para producir una cantidad incrementada de una proteína heteróloga en una planta dicotiledónea, al transformar establemente una célula vegetal con un polinucleótido que tenga un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprenda una secuencia de señal ER, una secuencia de codificación de proteína deseada y una señal de retención ER; en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla; en donde el polinucleótido contiene también una secuencia de ARNi que es capaz de la subregulación de una proteína de almacenamiento de semilla endógena, obtener una línea vegetal homocigótica y cultivar las semillas de la planta, para obtener una proteína heteróloga.

Breve Descripción de las Figuras

La figura 1 es un modelo de las diferentes vías de ubicación subcelular del retículo endoplásmico (ER) a una vacuola de almacenamiento de proteínas o un cuerpo proteínico (PB).

La figura 2 es un diagrama esquemático que muestra una construcción de ARNi diseñado para suprimir glicinina. Los segmentos para el gen de glicinina como para el gen fad2 (ácido graso desaturasa) fueron incluidos en la construcción de ARNi. El segmento fad2 fue añadido como una característica opcional de la construcción, proporcionando un marcador para tamizado adicional.

Las figuras 3A y 3B son micrografías electrónicas que muestran que células de plantas "SP-" de línea deficiente en proteína de semilla forman vacuolas de almacenamiento de proteínas (PSVs) (Figura 3A) que son notoriamente similares en

tamaño y apariencia a las PSVs formadas en células de semillas WT (Figura 3B). PSV: vacuola de almacenamiento de proteína; OB: cuerpo oleo; AV: vacuola autofágica; ER: retículo endoplásmico; Nucl: núcleo; G: aparato de golgi. La barra es igual a 1 μ m.

Las figuras 4A, 4B y 4C son fotografías de una comparación por enfoque isoelectrico bidimensional/electroforesis en dodecilsulfato de sodio-gel de poliacrilamida (IEF/SDS-PAGE) entre el proteoma de semillas tipo silvestre (WT) variedad "Jack" y SP-.

La figura 5 es una gráfica de dispersión de un ensayo de transcrito a gran escala de la variedad SP- en comparación con la variedad "Jack" WT usando un ensayo de plataforma de disposición de ADN Affymetrix.

La figura 6 esa gráfica de tarta que demuestra los cambios en la composición de proteínas de semilla en semillas de líneas de soya WT ("Jack") vs SP-. La composición porcentual de las diferentes proteínas de semilla se muestra.

La figura 7 es un diagrama esquemático que muestra los detalles de la construcción GFP-kdel. El promotor de glicinina, glicinina 5' UTR ("TED"), secuencia de señal ER, secuencia de codificación GFP, la secuencia de señal de retención de ER kdel y la glicinina 3' UTR ("TED") son indicados. El sitio de inicio de transcripción, el sitio de inicio de traducción y el sitio de detención de traducción son indicados.

Las figuras 8A, 8B, 8C y 8D son un panel de fotografías que muestran imágenes de luz blanca (Figura 8A) y azul (Figura 8B) de semillas de soya enteras de las dos líneas progenitoras homocigóticas y la progenie homocigótica de la cruce. Las semillas mostradas han sido fragmentadas para exponer el tejido cotiledóneo (Figuras 8A y 8B). Las Figuras 8C y 8D son imágenes GFP pseudocoloreadas de células de parénquima de almacenamiento de GFP-kdel (proteína GFP con un marcador de retención ER KDEL C-terminal añadido) en un fondo WT (Figura 8C) y semillas GFP-kdel x β CS. La barra es igual a 10 μ m.

Las figuras 9A, 9B, 9C y 9D son un panel de fotografías de la separación 2D IEF-PAGE del lisado de proteínas de semilla β CS (suprimida en β -conglucina; Figura 9A), semilla GFP-kdel (Figura 9B) y semilla GFP-dkel x β CS (Figura 9C), e inmunoblot de un gel de lisado replicado (9 Figura 9D) sondeado para GFP. Los puntos de proteína de control GFP-kdel (Figuras 9A y 9B) o GFP son encerrados en los cuadros como se marca.



La introducción del rasgo GFP-kdel en la línea β CS se tradujo en la acumulación incrementada de la GFP-kdel. La identidad de los puntos de GFP se determinó mediante inmunorreactividad en blots usando una sonda de anticuerpos monoclonales comercial (Figura 9D).

5 Las figuras 10A, 10B y 10C muestran microscopias de fluorescencia (Figura 10A), ID PAGE (Figura 10B), e inmunoblot de β -conglucina (Figura 10C) ya sea para β BC, GFP-kdel en WT, o GFP-kdel x β CS.

La figura 11 es una gráfica de barras de un análisis fluorométrico de abundancia de GFP-kdel en lisados de semilla preparados de semillas ya sea β CS, GFP-kdel en WT, o GFP-kdel x β CS, y ensayados usando GFP comercial como un parámetro de control.

La figura 12 es una fotografía de un PAGE unidimensional de proteínas de semilla fraccionadas que muestra el procesamiento de glicina en WT ("Jack"), β CS y GFP-kdel x β CS. El gel teñido resultante fue escaneado y la distribución relativa de la fracción de proglicina de la proglicina sumada, glicina A4, subunidad ácida de glicina y subunidad básica de glicina. Los resultados muestran una reducción de más de tres veces de la fracción de proglicina de la población de proteínas de glicina. Marcadores de peso molecular e isoformas de proteína de almacenamiento son indicados.

Las figuras 13A, 13B y 13C son fotografías de un análisis en gel bidimensional de la producción de proteínas en una planta SP- (Figura 13A), GFP-kdel transformada en un fondo WT (Figura 13B) y una planta SP- introducida con GFP-kdel (Figura 13C).

La figura 14 es una gráfica de barras de un análisis fluorométrico de la abundancia de GFP-kdel en lisados de semilla preparados a partir de semillas de una planta SP-, una planta WT transformada con GFP-kdel y una cruce GFP-kdel x SP-. GFP comercial se usó como una norma de control.

Las figuras 15A y 15B es una micrografía electrónica que demuestra la abundancia de cuerpos de proteínas en el citoplasma de células de semilla de maduración tardía de plantas BCS x GFP-kdel. Figura 15A: Los cuerpos de proteína contienen una matriz dispersa y están unidos por una membrana ER. Figura 15B: imagen

que demuestra el origen ER de los cuerpos de proteína. PSV; vacuola de almacenamiento de proteína; OB = cuerpo óleo. La barra es igual a 1 μ m.

Descripción Detallada de la Invención

Se proporciona en la presente una planta dicotiledónea modificada genéticamente que tiene una semilla que produce una alta cantidad de una proteína heteróloga de interés. En ciertas modalidades, la semilla es deficiente en al menos una proteína de almacenamiento de semilla endógena, permitiendo que se produzca en la misma una cantidad incrementada de una proteína extraña. La secuencia de ácido nucleico que codifica para la proteína heteróloga puede ser enlazada operablemente a una región reguladora de una proteína de semilla. En algunas modalidades, esta región reguladora se deriva de una proteína de semilla que es sobrerregulada naturalmente en respuesta a la deficiencia de la proteína de semilla endógena.

En ciertas modalidades, la programación genética en dicotiledóneas puede utilizarse exitosamente para producir una proteína de interés, por ejemplo, una proteína cualitativa y cuantitativamente superior (por ejemplo, proteína recombinante).

En ciertas modalidades, el antecedente genético de la planta es modificado de tal manera que haya una deficiencia en la cantidad de una o más proteínas de almacenamiento (por ejemplo, en peso). Al usar uno o más promotores de proteína de almacenamiento para conducir la transcripción de la proteína objetivo, los mecanismos de reequilibrio de la planta pueden traducirse en niveles especialmente altos de una producción de proteína heteróloga.

En esta modalidad, y sin desear ser limitados por teoría, en respuesta a una deficiencia genética que causa la pérdida de una proteína de almacenamiento de semilla mayor, la semilla se "reequilibra" al incrementar la producción de otras proteínas de almacenamiento de semilla. Al enlazar un gen heterólogo de interés a elementos reguladores génicos de una proteína de semilla endógena que es sobre-regulada para reequilibrar la cantidad total de la proteína en la semilla, se puede producir un alto nivel de proteína heteróloga en la semilla. Gracias a este alto nivel de acumulación de la proteína extraña, este método "mímico de alelo" es útil para producir proteínas, particularmente proteínas comercialmente valiosas, en soya u otras semillas dicotiledóneas.

25

Además, la dirección opcional de la proteína heteróloga al ER le permite acumularse establemente a niveles todavía más altos en la semilla. Pueden diseñarse señales en la proteína heteróloga que puedan traducirse en el secuestro de la proteína objetivo en vesículas derivadas del ER, no obstante de si la planta produce naturalmente

5 cuerpos proteicos fisiológicamente. Estas vesículas derivadas del ER están sorprendentemente libres de otras proteínas, tales como proteasas, glucosidasas, etc., produciendo acumulación de niveles más altos de proteína en una forma menos degradada o degradable. Según se usa en la presente, aplican las siguientes abreviaturas y definiciones.

10 El término "glicinina", se refiere a una proteína de almacenamiento de semilla mayor, conocida también como IIS globulina, que está presente en semillas de soya.

El término " β -conglucina" se refiere a una proteína de almacenamiento de semilla mayor, también conocida como 7S globulina que está presente en semillas de

15 soya.

El término "GBP" se refiere a proteína de unión a glucosa.

El término "KTI" se refiere a inhibidor de tripsina Kunitz.

El término "LE" se refiere a lectina de soya.

El término "P34" se refiere al alérgeno de soya inmunodominante P34 o Gly

20 m Bd 3.

La abreviatura "fad2" se refiere a una secuencia que codifica para una porción de un gen de "ácido graso desaturasa".

Según se usa en la presente, el término "vegetal" incluye células vegetales, protoplastos vegetales, callos vegetales, grupos vegetales y células vegetales que están

25 intactas en plantas o partes de plantas, tales como polen, flores, embriones, semillas, vainas, hojas, tallos y similares.

El término "PB" se refiere a un cuerpo proteínico. Estas vesículas de una sola membrana son capaces de almacenar proteínas, y se derivan directamente del retículo endoplásmico (en contraste con las PSVs descritas abajo).

30 El término "PSVs" se refiere a vacuolas de almacenamiento de proteínas.

En semillas, estas vesículas se forman típicamente a partir de una división de la vacuola durante el proceso de maduración y secado de la semilla. Así, las enzimas degradadoras de proteínas normalmente presentes en la vacuola también pueden estar presentes en la PSV. En semillas de soya normales, la mayoría de las proteínas acumuladas se ubican en estos organelos.

El término "SMP" se refiere a una proteína de maduración de semilla.

El término "proteína de almacenamiento" se refiere a una proteína que se acumula específicamente en una planta, por ejemplo, en semillas.

La abreviatura "BBI" se refiere a "inhibidor bowman birk", que es un inhibidor de serina proteasa.

El término " β CS" se refiere a una planta que es deficiente en la proteína de almacenamiento β -conglucina únicamente.

El término "SP-" se refiere a supresión de proteína de almacenamiento, es decir, una planta que es deficiente tanto en glicina como en β -conglucina.

El término "semilla" incluye generalmente el núcleo de la semilla, la cápsida de la semilla y/o la cáscara de la semilla, o cualquier porción de la misma.

"Maduración de semilla" se refiere al periodo que inicia con la fertilización en la cual las reservas metabolizables, por ejemplo, almidón, azúcares, oligosacáridos, fenólicos, aminoácidos y proteínas, se depositan en varios tejidos en la semilla, llevando a agrandamiento de la semilla, relleno de la semilla y concluyendo con la desecación de la semilla.

El término "WT" se refiere a tipo silvestre y se refiere a un antecedente de origen natural de una planta, o, como es aparente del contexto de uso, WT puede referirse a una planta que tenga un antecedente genético de origen natural excepto para la manipulación genética de la presente invención.

El término "ORF" se refiere a un marco de lectura abierto; es decir, una secuencia que codifica para un péptido (por ejemplo, la "proteína objetivo"). En general, esta secuencia es ininterrumpida por intrones entre los codones de inicio y terminación que codifica para una secuencia de aminoácidos.

El término "polinucleótido heterólogo" se refiere generalmente a un

polinucleótido que no existe en forma idéntica en la planta anfitriona excepto como resultado de un evento de transformación. Los términos "ADN heterólogo", "gen heterólogo" o "ADN extraño" se refieren a ADN, y típicamente a una secuencia de codificación de ADN ("secuencia de codificación heteróloga"), que ha sido introducida en

5 células vegetales de otra fuente, es decir, una fuente no vegetal o de otra especie de planta, o una secuencia de codificación de la misma especie que es puesta bajo el control de un promotor vegetal que controla normalmente otra secuencia de codificación.

El término gen "endógeno" se refiere a un gen nativo encontrado normalmente en su ubicación natural en el genoma y no está aislado. Un gen "extraño"

10 se refiere a un gen que no se encuentra normalmente en el organismo anfitrión pero que es introducido mediante transferencia génica.

El término "secuencia de codificación" o "región de codificación" se refiere a una secuencia de ADN que codifica para una proteína específica.

Un "gen quimérico" o "casete de expresión" en el contexto de la presente

15 invención, se refiere a una secuencia promotora enlazada operablemente a una secuencia de ADN que codifica para un producto génico deseado, y de preferencia una secuencia terminadora de transcripción. En una modalidad preferida, el gen quimérico contiene también una región de codificación de péptidos de señal enlazado operablemente entre el promotor y la secuencia de codificación de producto génico en

20 cuadro de producción con la secuencia de codificación de producto génico. Esta secuencia de señal ayuda ubicar la proteína en el ER. La secuencia puede contener además elementos reguladores de transcripción, tales como las señales de terminación de transcripción mencionadas arriba, así como señales reguladoras de traducción, tales como codones de terminación.

25 "Enlazado operablemente" se refiere a componentes de un casete de expresión, que están enlazados de tal manera que funcionen como una unidad para expresar una proteína heteróloga. Por ejemplo, un promotor enlazado operablemente a un ADN heterólogo, el cual codifica para una proteína, promueve la producción de ARNm funcional que corresponde al ADN heterólogo.

30 "Transcrito de ARN" se refiere al producto que resulta de una transcripción

catalizada por ARN polimerasa de una secuencia de ADN.

El término "ARN mensajero (ARNm)" se refiere generalmente al ARN que puede ser traducido en proteína por la célula.

El término ARN de "sentido" se refiere generalmente a un transcrito de ARN que incluye el ARNm. El término "ARN de antisentido" se refiere a un transcrito de ARN que es complementario a todo o parte de un transcrito primario objetivo o ARN y que puede inhibir la expresión de un gen objetivo. La complementariedad de un ARN de antisentido puede ser con cualquier parte del transcrito génico específico, es decir, en la secuencia de no codificación 5', secuencia de no codificación 3', intrones o la secuencia de codificación. El término "inhibición de antisentido" se refiere a la producción de transcritos de ARN de antisentido capaces de impedir la expresión de la proteína objetivo.

El término "ARNi" o "interferencia de ARN" se refiere generalmente a métodos de inhibición de expresión de una proteína de introducir un fragmento de ARN en la célula. El ARN puede ser codificado por un fragmento de ADN que sea integrado en el genoma. El ARNi también puede prepararse mediante cualquier otro medio conocido en la técnica. El fragmento de ARNi puede ser de cualquier longitud adecuada, y puede ser de una sola o de doble hebra.

En general, "secuencias reguladoras" son secuencias de nucleótidos ya sea en genes endógenos o heterólogos (quiméricos) que se ubican hacia el extremo 5' (5'), dentro, o hacia el extremo 3' hacia la región de codificación de la proteína. Estas secuencias reguladoras o "regiones reguladoras" pueden regular la transcripción y/o traducción.

Una "región reguladora de transcripción" o "promotor" se refiere a secuencias de ácido nucleico que influyen y/o promueven el inicio de transcripción. Los promotores se consideran típicamente que incluyen regiones reguladoras, tales como elementos potenciadores o inductores.

El término "regiones reguladoras hacia el extremo 5'" se refieren generalmente a una región 5' del codón de inicio de traducción de proteína. Así, las "regiones reguladoras hacia el extremo 5'" pueden abarcar, por ejemplo, un promotor, o pueden abarcar un promotor y una 5' UTR. La región reguladora hacia el extremo 5'

también se puede referir a regiones más allá del extremo 5' de la secuencia promotora típica, tales como "elementos potenciadores".

El término "TED" se refiere a un dominio potenciador de traducción.

5 El 5' TED (o región no traducida 5' (UTR)) se refiere generalmente a la región que codifica para un ARNm que es 5' (hacia el extremo 5') hacia el sitio de inicio de traducción. De esta manera, la región está entre el sitio de inicio de transcripción y el sitio de inicio de traducción. El 5' TED puede ser una parte de la "región reguladora hacia el extremo 5'".

10 La 3' TED (o región no traducida 3' (UTR)) se refiere generalmente a la región en el ARNm que está hacia el extremo 3' del codón de paro de la secuencia de codificación de proteína. Esta región puede contener, por ejemplo, una señal de poliadenilación y/o cualquier otra señal reguladora capaz de afectar procesamiento de ARNm o expresión génica.

15 "Codón de inicio" y "codón de terminación" se refieren a una unidad de tres nucleótidos adyacentes en una secuencia de codificación que especifica el inicio y terminación de cadena, respectivamente, de una secuencia de proteínas.

El alcance de la presente invención se ilustra abajo con varios ejemplos, características técnicas opcionales y enseñanzas genéricas.

Proteínas de almacenamiento

20 Las semillas de muchas especies de plantas contienen proteínas de almacenamiento. Estas proteínas han sido clasificadas sobre la base de su tamaño y solubilidad (Higgins, T. J. (1984) *Ann. Rev. Plant Physiol.* 35:191-221). Aunque no todas las clases se encuentran en cada especie, las semillas de la mayoría de las especies de plantas contienen proteínas de más de una clase. Las proteínas dentro de una clase de solubilidad o tamaño particular generalmente están más estructuralmente emparentadas a
25 miembros de la misma clase en otras especies que con miembros de una clase diferente dentro de la misma especie. En muchas especies, las proteínas de semilla de una clase dada son codificadas comúnmente por familias de varios genes, algunas veces de tal complejidad que las familias pueden dividirse en subclases con base en la homología de
30 secuencia.

Las semillas de soya poseen un contenido de proteína relativamente alto, que consiste ampliamente en dos proteínas de almacenamiento, β -conglucina y glicina. En semillas tipo silvestre, la β -conglucina comprende alrededor de 15-20% de la proteína de soya total. Ambas de estas proteínas están constituidas de varias isoformas derivadas de familias génicas.

Las proteínas de almacenamiento pivotal han sido identificadas en la presente como estando implicadas en el desarrollo programado de una planta. Sin embargo, la persona normalmente versada en la materia puede identificar ahora fácilmente otras proteínas de almacenamiento en otras plantas objetivo mediante homologías funcional, estructural o de secuencia. Habiendo identificado estas proteínas de almacenamiento, experimentos de supresión como los descritos aquí pueden identificar otras proteínas de almacenamiento que estén implicadas en el reequilibrio de la proteína. Los elementos reguladores (por ejemplo, promotor, TEDs, etc.) pueden usarse para producir altos niveles de una proteína deseada de acuerdo con la presente invención. De esta manera, los diferentes ejemplos demostrados en la presente pueden ser aplicados ahora a otras plantas de importancia potencial para el comercio o humanidad.

Deficiencias genéticas

Una planta de la presente invención comprende una deficiencia, tal como por ejemplo, una deficiencia genética, dando como resultado una reducción de una porción sustancial del contenido de proteína de almacenamiento de semilla endógeno de la planta. Por ejemplo, en varias modalidades específicas, las semillas de la planta pueden comprender menos de aproximadamente 75%, 70% o menos del 60% o menos de alrededor de 50% o menos de aproximadamente 40% o menos de alrededor de 25% o menos de 15% de la cantidad de proteína soluble total en la semilla.

En otras modalidades específicas, la deficiencia genética se traduce en una cantidad de una proteína de almacenamiento de semilla específica que es menor de 1%, aproximadamente 2%, alrededor de 5%, aproximadamente 10%, alrededor de 25%, aproximadamente 50%, alrededor de 75%, o aproximadamente 85% de la cantidad de la proteína de almacenamiento de semilla endógena que normalmente está presente en una semilla de soya WT.

A manera de ejemplo, una planta de la presente invención puede ser deficiente en una o dos o tres o cuatro o más de glicinina, conglucina, KTI, LE, P34, GBP y SMP, u otras proteínas de almacenamiento de semilla.

En una modalidad, la manipulación genética para crear una deficiencia de una proteína de almacenamiento de semilla se puede obtener mediante métodos tales como co-supresión, antisentido, ARNi u otros métodos. La patente de E.U.A. No. 5,190,931 describe métodos ejemplares del uso de una construcción de antisentido para subregular un gen. La patente de E.U.A. No. 5,231,020 describe métodos ejemplares del uso de una construcción de ácido nucleico de sentido para subregular un gen. La inhibición genética de la expresión de un producto génico mediante el uso de ARNm de doble hebra se describe en la patente de E.U.A. No. 6,506,559.

En otras modalidades, una deficiencia de una proteína de almacenamiento de semilla particular, o proteínas de almacenamiento de semilla en el agregado, también se puede lograr mediante métodos de reproducción convencionales seguidos por el análisis para un bajo nivel de una o más proteínas de semilla. Una deficiencia de una proteína de almacenamiento de semilla también puede lograrse mediante mutaciones naturales o mutaciones inducidas, seguida por métodos de tamizado para identificar las plantas que tengan un bajo nivel de una o más proteínas de almacenamiento de semilla. En otra modalidad, una planta que tenga una deficiencia de una proteína de almacenamiento de semilla puede obtenerse, por ejemplo, a partir de un banco o depósito de semillas disponible públicamente.

Una deficiencia genética de una proteína en la semilla lleva a compensación (reequilibrio) por otras proteínas de semilla

Como se describe en la presente, la supresión de una proteína de semilla puede llevar a una compensación por un incremento en la producción de otras proteínas de semilla, llamada "compensación" o "reequilibrio". Este reequilibrio es demostrado en la presente con líneas vegetales que tienen una deficiencia tanto en glicinina como en β -conglucina ("SP-"; ejemplo 1) como en plantas que tienen una deficiencia nada más en β -conglucina (ejemplo 11).

La supresión de la proteína de semilla β -conglucina mediante

silenciamiento génico mediado por secuencia fue compensada por una abundancia incrementada de glicinina (ejemplo 11). La supresión de β -conglucina α/α' también se logró usando tecnología de ARNi, como también se describe en el ejemplo 11. Este método dio como resultado el silenciamiento completo de α/α' β -conglucina. Una fracción de la producción incrementada de glicinina fue conservada en forma de su precursor, proglicinina, y fue secuestrada en PBs. La acumulación de proteínas en un cuerpo de proteína, en lugar de la PSV, demuestra dos puntos importantes: 1) que PBs derivados del ER pueden ser inducidas y acumular proteínas en semillas de soya y 2) que la supresión de una proteína de almacenamiento endógena se traduce en la acumulación incrementada de otra proteína de almacenamiento para compensar la pérdida de masa. Este fenómeno mantiene el contenido proteico total de la semilla de soya en menos de 40% y es llamado "reequilibrio".

Las líneas vegetales que tienen una deficiencia tanto en β -conglucina como en glicinina ("SP-") fueron preparadas, como se describe en el ejemplo 1. En estas semillas que fueron genéticamente deficientes tanto en β -conglucina como en glicinina, la pérdida de proteína fue compensada por la producción de otras proteínas. Los cambios de la producción de proteína pueden verse en las figuras 4A a 4C, la cual es un IEF-PAGE de extractos de proteína de semilla de la línea WT ("Jack") en comparación con la SP-. La figura 5 muestra las diferencias en el transcriptoma entre las dos líneas. La figura 6 es un diagrama de tarta que muestra el porcentaje de varias proteínas de almacenamiento mayores presentes en semilla de soya de WT ("Jack") contra la línea SP-. La remoción de las proteínas de semilla β -conglucina y glicinina en la línea SP- claramente se traduce en compensación por otras proteínas.

El fenómeno de reequilibrio puede usarse para producir grandes cantidades de proteínas extrañas en semillas

En la presente se proporcionan semillas que poseen una biología intrínseca que pueden ser explotadas como la base de una plataforma de producción de proteínas al hacer que una proteína extraña tenga una participación en el proceso de reequilibrio y al acumular la proteína extraña en una población estable de PBs. Esto junto sienta las bases para desarrollar semillas dicotiledóneas como una plataforma de



producción de proteínas.

Así, en algunas modalidades, una (o más) proteínas de almacenamiento de semilla se reducen como se describió arriba, y una proteína heteróloga deseada se produce en la semilla. En una modalidad, cualquier promotor adecuado es enlazado operablemente a la secuencia que codifica para la proteína heteróloga. En otra modalidad, la secuencia que codifica para la proteína heteróloga es un promotor específico de semilla. El promotor puede ser, por ejemplo, seleccionado de los promotores de KTI, LE, P34, GBP o SMP.

En una modalidad preferida, la expresión incrementada de la proteína heteróloga puede ocurrir cuando su secuencia génica sea enlazada operablemente a un promotor del gen que codifique para una proteína que sea sobrerregulada en respuesta a la deficiencia genética descrita arriba en una semilla de soya. Por cada proteína específica que sea removida en una semilla, se puede producir otra proteína (o proteínas) en su lugar. Mediante el uso de la región reguladora génica (tal como la región promotora, terminadora y opcionalmente otras regiones) de esta proteína "compensadora" para conducir la expresión del gen heterólogo de interés, puede obtenerse un nivel todavía más alto de producción de proteína en la semilla.

En consecuencia, en una modalidad, la proteína de semilla que es suprimida es β -conglucina, mientras que la expresión de la proteína heteróloga es controlada por al menos una porción de la región reguladora de glicina. En una modalidad, esta región reguladora está hacia el extremo 5' de la secuencia heteróloga. En otra modalidad, la región reguladora está hacia el extremo 3' de la secuencia heteróloga. En una modalidad la región reguladora comprende el promotor de glicina. En una modalidad la región reguladora es la región reguladora hacia el extremo 5' de glicina, la cual puede incluir, por ejemplo, el promotor y/o la 5' UTR. En otra modalidad, la región reguladora de glicina también incluye la región reguladora de glicina 3'.

En otra modalidad, la proteína de semilla que es suprimida es tanto β -conglucina como glicina, mientras que la proteína heteróloga es controlada por al menos una porción de las regiones reguladoras (5', 3' o ambas) de una de KTI, LE, P34, GBP o SMP.

El marco conceptual del reequilibrio de proteína y secuestro de proteína en PBs fue probado al construir un transgen con la proteína reportera GFP, flanqueada por la secuencia de señal de tránsito ER y secuencia de señal de retención (KDEL), bajo el control regulador de elementos de glicinina (ejemplo 9). Al poner la construcción GFP-kdel bajo elementos genéticos de glicinina, la expresión del transcrito GFP-kdel imitará aquella de la expresión en regulación génica y de esta manera probablemente participará en la asignación de nutrientes que incluye la sobreexpresión de genes de glicinina. Las soyas que expresan este transgen acumularon 1.6% de GFP-kdel en la semilla. Sin embargo, cuando las plantas GFP-kdel fueron genéticamente cruzadas con las plantas β CS, el nivel de expresión de GFP-kdel fue incrementado casi 4 veces hasta aproximadamente 7% en las semillas (ejemplo 12). Así, el incremento en la acumulación de GFP-kdel en las semillas β CS demuestra que imitar el alelo del gen que participa en reequilibrio de proteínas puede traducirse en un grave incremento en la acumulación de la proteína heteróloga de interés.

Así, en una modalidad, una gran cantidad de una proteína de interés puede producirse en la semilla. Por ejemplo, la proteína heteróloga puede ser expresada de aproximadamente 1%, 2%, 5%, 7%, 10%, 12%, 15%, 18%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 70% o más de la proteína soluble total en la semilla.

Además, en una modalidad, el peso en seco de la proteína heteróloga puede expresarse de alrededor de 0.5%, 1%, 2%, 4%, 5%, 7%, 10%, 12%, 15%, 18%, 20%, 25%, 30%, 35%, 40%, 45%, 50% o más del peso en seco total de la semilla. La proteína heteróloga puede producirse, por ejemplo, a una cantidad de alrededor de 50, 100, 150, 200, 250 o más mg de proteína por semilla.

En una modalidad, la cantidad de proteína heteróloga producida puede medirse sobre una base por planta. La proteína heteróloga puede producirse, por ejemplo, en una cantidad de alrededor de 3 g, 6 g, 8 g, 10 g, 15 g o 20 g o más de proteína heteróloga por planta.

La proteína heteróloga puede producirse, por ejemplo, a una cantidad de alrededor de 25, 50, 100, 200, 300, 400, 500, 600, 700, 850, 1,000, 1,500, 2,000 o más libras por acre por temporada. El rendimiento real puede depender de muchos



parámetros tales como plantas por acre, variedad vegetal, calidad en suelo, prácticas de cultivo, estrés vegetal y también el nivel de pureza de la proteína heteróloga que será producida.

Promotores

5 En una modalidad, el polinucleótido heterólogo provisto en la presente comprende un promotor obtenido de, o derivado de, un gen de proteína de almacenamiento vegetal. En varias modalidades, el promotor se deriva de la planta del mismo orden, familia, género o especie de la planta transformada por una construcción de la presente invención.

10 Puede usarse cualquier promotor adecuado. En una modalidad, el promotor es un promotor específico de semilla. En otra modalidad, el promotor es un promotor específico de semilla temprano. En otra modalidad más, el promotor es un promotor específico de semilla tardío. En otra modalidad, el promotor es de un gen que compensa la deficiencia genética de la proteína de semilla.

15 Las secuencias promotoras pueden concluir en o cerca del codón de inicio e incluir nucleótidos contiguos hacia el extremo 5' (5'). Las secuencias promotoras pueden ser de al menos aproximadamente 500 nucleótidos o por lo menos alrededor de 1,000 nucleótidos o al menos aproximadamente 1,500 nucleótidos. Aunque la longitud exacta no es crítica para la invención, un experto en la técnica puede determinar y
20 optimizar fácilmente la longitud del promotor (por ejemplo, al medir y comparar los niveles de transcripción).

En una modalidad, la secuencia reguladora hacia el extremo 5' o la secuencia promotora puede tener una identidad de ácido nucleico de al menos 80%, 85%, 90%, 95%, 97%, 99.5% o más con por lo menos una porción de una de SEQ ID NO: 1,
25 SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8.

En una modalidad específica, la región reguladora hacia el extremo 5', que comprende el promotor y 5' UTR como se describe abajo, se selecciona de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID
30 NO: 7. SEQ ID NO: 1 y 2 se derivan de glicinina, en tanto que SEQ ID NO: 3-8

representan regiones reguladoras hacia el extremo 5' de conglucina, KTI, LE, P34, GBP y SMP, respectivamente.

Dominios potenciadores de traducción UTR

Un polinucleótido heterólogo de la presente invención comprende 5 opcionalmente un dominio potenciador de traducción (TED) de una proteína de almacenamiento vegetal. En algunas modalidades, esta es la región no traducida del transcrito de ARNm. Estas regiones no traducidas pueden estar en la región 5' o 3' del gen. La secuencia del TED o UTR también se puede derivar de otro organismo, o puede ser una secuencia completamente sintética.

10 5' UTR (o "5' TED"): Una construcción de la presente invención puede comprender una 5' TED de la región 5' de un ARNm que codifique para una proteína de almacenamiento vegetal tal como glicina, conglucina, KTI, LE, P34, GBP o SMP. Un 5' TED puede identificarse generalmente del promotor y el codón de inicio de una proteína de almacenamiento vegetal. Un 5' TED puede tener al menos aproximadamente 5, 10, 15 25, 30, 35, 40 o más nucleótidos o al menos alrededor de 50 nucleótidos o por lo menos aproximadamente 100 nucleótidos, o al menos alrededor de 150 nucleótidos. Aunque la longitud exacta no es crítica para la invención, un experto en la técnica puede determinar y optimizar fácilmente la longitud del terminador (por ejemplo, al medir y comparar los niveles de traducción). En modalidades específicas, una porción del extremo 3' de las 20 secuencias reguladoras hacia el extremo 5' descritas en SEQ ID NO: 1 (glicina), SEQ ID NO: 2 (una secuencia de glicina alternativa), SEQ ID NO: 3 (conglucina), SEQ ID NO: 4 (KTI), SEQ ID NO: 5 (LE), SEQ ID NO: 6(P34), SEQ ID NO: 7 (GBP) y SEQ ID NO: 8 (SMP) comprenden, respectivamente, secuencias 5' UTR.

3' UTR (o "3' TED" o "terminador"): Una TED puede derivarse, por ejemplo, 25 de la región 3' de un ARNm que codifique para una proteína de almacenamiento vegetal tal como glicina, conglucina, KTI, LE, P34, GBP o SMP. Este 3' TED puede iniciar en o cerca del codón de paro e incluir nucleótidos contiguos hacia el extremo 3'. El 3' TED puede tener aproximadamente alrededor de 10, 25, 40, 50, 75, 100, 150 nucleótidos o al menos aproximadamente 250 nucleótidos o al menos aproximadamente 500 nucleótidos. 30 Aunque la longitud exacta no es crítica para la invención, un experto en la técnica puede

determinar y optimizar fácilmente la longitud del terminador (por ejemplo, al medir y comparar niveles de producción).

En una modalidad, la secuencia terminadora se selecciona de SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20 y SEQ ID NO: 21. Estas secuencias representan las secuencias 3' TED para glicinina, una secuencia terminadora de glicinina alternativa, β -conglucina, una secuencia terminadora de β -conglucina alternativa, KTI, LE, P34, GBP y SMP, respectivamente.

En una modalidad, la secuencia terminadora puede tener una identidad de ácido nucleico de al menos 80%, 85%, 90%, 95%, 97%, 99.5% o más con una de SEQ ID NOS: 13-21.

En modalidades adicionales, las construcciones quiméricas también pueden comprender secuencias reguladoras génicas que estén más allá del extremo 5' o más allá del extremo 3' del gen de interés, tales como secuencias potenciadoras. Por ejemplo, las regiones "potenciadoras" de transcripción pueden estar presentes más allá del extremo 5' o extremo 3' de un gen de interés. Por ejemplo, una región potenciadora puede estar 1,000 -2,000 nucleótidos o incluidos 1 o más kb hacia el extremo 5' de una secuencia, o también puede estar presente en una ubicación de 1,000-2,000 nucleótidos o incluso 1 o más kb hacia el extremo 3' de la región transcrita de un gen. Así, en algunas modalidades, estas secuencias potenciadoras más distantes de una proteína de almacenamiento vegetal, tal como KTI, LE, P34, GBP y SMP también están presentes en la secuencia quimérica. En ciertas modalidades, la presencia de estas secuencias potenciadoras incrementa la cantidad de la proteína heteróloga que se produce en la semilla.

25 El retículo endoplásmico

El retículo endoplásmico (ER) de plantas es una parte del sistema de endomembrana, un sistema altamente conservado en eucariontes. Dirigir proteínas al ER es de considerable interés toda vez que las proteínas producidas en el ER son traficadas a otros organelos y también permanecen asociadas con el propio ER. Los compartimientos derivados de ER tienen diversas funciones, tales como almacenamiento

de proteínas, aceites y enzimas hidrolíticas usadas en respuesta a ataques de patógenos. Incrementar el conocimiento de los mecanismos de tráfico de proteínas de almacenamiento al ER ha llevado a mejoras en el uso de plantas como biofábricas de proteínas. Las plantas pueden almacenar proteínas exógenas dentro del ER además de
5 otros compartimientos de la endomembrana.

Vesículas derivadas de ER – cuerpos de proteínas contra vacuola de almacenamiento de proteínas

La figura 1 muestra un diagrama esquemático de las diferentes vías de tráfico de proteínas subcelulares que llevan a la ubicación de una proteína en un cuerpo
10 de proteína o una vacuola de almacenamiento de proteína en una semilla de soya. En muchas especies vegetales, pero no en soya, las PBs derivadas del ER permiten la acumulación estable de proteína no glicosilada, ya que las proteínas no siguen una vía de endomembrana típica del ER al aparato de Golgi y de ahí en adelante a la prevacuola.

En plantas de soya WT, la mayoría de las proteínas de semilla se acumulan
15 en la vacuola de almacenamiento de proteínas (PSV), en su lugar. Las proteínas que son dirigidas a la vacuola de almacenamiento de proteínas (PSV) o a la prevacuola (la cual se dirige después al PSV) de semilla de soya es probable que sean degradadas rápidamente, sin embargo, ya que la vacuola típicamente contiene muchas enzimas
líticas.

En contraste, como se describe en la presente con respecto a semillas de
20 soya transgénicas, el tiempo de residencia en el ER de proteínas que usan una secuencia de dirección al ER C-terminal (KDEL) induce el tráfico de grandes cantidades de proteína extraña a las PBs producidas *de novo*. Así, las proteínas pueden ser secuestradas en PBs derivadas del ER en semillas de soya. Los cuerpos de proteína resultante son una
25 población estable de organelos que persisten a través de la maduración de la semilla y permanecen en la semilla madura seca. Esto se demuestra en la presente con la proteína fluorescente verde de proteína reportera (GFP-kdel) de 27 kDa, como la mostrada en el
ejemplo 9.

En una modalidad opcional de la presente invención, la secuencia
30 heteróloga comprende además una secuencia de retención ER para inducir la acreción

39

del polipéptido heterólogo en el lumen del ER o una vesícula derivada del ER. Este ER o vesículas derivadas del ER incluyen las ER, PBs, PSVs, vesículas de transporte, etc. Estas vesículas pueden identificarse estructuralmente como comprendiendo una membrana (por ejemplo, membrana de bicapa de lípidos), un lumen, en donde la proteína
5 objetivo es un componente soluble o insoluble que reside al menos parcialmente en el lumen. Sin desear ser limitados por teoría, se cree que el ER o la ubicación de vesícula derivada de ER de la proteína de interés es, en parte, responsable de los altos niveles de proteína heteróloga producidos en plantas de acuerdo con la presente invención. En una modalidad, el polipéptido heterólogo se acumula en el aparato de Golgi o vesículas de
10 Golgi.

Secuencia de señal ER

En algunas modalidades, el polinucleótido heterólogo comprende una secuencia de señal ER. Una secuencia de señal ER es cualquier secuencia de polinucleótidos que codifica para una secuencia de aminoácidos que permite el
15 reconocimiento de la proteína por la partícula de reconocimiento de señal en el retículo endoplásmico dando como resultado la translocación de la proteína dentro del lumen del ER. Esta secuencia está presente típicamente en la región N-terminal de la proteína.

En algunas modalidades, la secuencia de señal ER puede añadirse a proteínas que no tengan naturalmente una secuencia de dirección al ER. La
20 proteína heteróloga puede ya tener una secuencia de señal ER, sin embargo. Si se desea, esta secuencia de señal puede reemplazarse con otra secuencia de señal ER, tal como aquellas mostradas abajo en la tabla 1. Como alternativa, se puede usar la secuencia de señal original de la proteína. También puede usarse una secuencia de señal completamente sintética. Ejemplos de secuencias de señal que
25 dirigen a proteínas recién sintetizadas al retículo endoplásmico en células vegetales incluyen secuencias de lectina de cebada (Dombrowski *et al.*, 1993, *Plant Cell* 5:587-596), aleuraina de cebada (Holwerda *et al.*, 1992, *Plant Cell* 4:307-318), esporamina de boniato (Matsuoka *et al.*, 1991, *proc. Natl. Acad. Sci. E.U.A.* 88:834-838), patatina (Sonnewald *et al.*, 1991, *Plant J.* 1:95-106), proteínas de almacenamiento vegetal de
30 soya (Mason *et al.*, 1988, *Plant Mol. Biol.* 11:845-856), y beta-fructosidasa (Faye *et*

al. 1989, *Plant Physiol.* 89:845-851).

Aunque la persona versada en la materia puede determinar fácilmente secuencias de señal ER útiles, otros ejemplos se muestran en la tabla 1.

Tabla 1

SEQ ID NO: 9	MKIMMMIKLCFFSMSLICIAPADA
SEQ ID NO: 10	MAASHGNAIFVLLLCTFLPSLAC
SEQ ID NO: 11	MAARIGIFSVFVAVLLSISAFSSA
SEQ ID NO: 12	MKTNLFLFLIFSLLLSSLSSAE (secuencia de señal de quitinasa básica de <i>A. thaliana</i>)

Otros ejemplos de secuencias de señal ER se describen por Emanuelsson et al. (*J. Mol. Biol.* 300,1005-1016 (2000)).

Secuencia de retención ER

Opcionalmente, el polinucleótido heterólogo comprende además una secuencia de retención ER. Se ha descubierto que cuando esta secuencia se añade a un polinucleótido heterólogo que también contiene una secuencia de señal ER, el producto de proteína será retenido en las vesículas derivadas del ER cuando el producto sea secuestrado a partir de cierta acción de procesamiento tal como degradación proteolítica. Sorprendentemente, las presentes construcciones dirigen el producto de proteína de polinucleótido heterólogo a vesículas derivadas del ER llamadas “cuerpos de proteína” sin importar de si la planta anfitriona produce naturalmente cuerpos de proteínas. Así, es ahora posible estabilizar el producto de péptido heterólogo y acumularlo a más altos niveles.

Una secuencia de retención ER es cualquier secuencia de polinucleótidos que codifiquen para una secuencia de aminoácidos que se sepa se traduzca en la retención de una proteína dada en o asociada con el retículo endoplásmico tales como las secuencias que codifican para los aminoácidos (representados por el código de aminoácido de una sola letra) KDEL (SEQ ID NO: 23), KHDEL (SEQ ID NO: 25), HDEL (SEQ ID NO: 26), KEEL (SEQ ID NO: 27) SEKDEL (SEQ ID NO: 28) y SEHDEL (SEQ ID NO: 29). Los ácidos nucleicos ejemplares que codifican para el KDEL o KHDEL, respectivamente, se muestran en SEQ ID NO: 22 y 24. Típicamente, estas secuencias

son C-terminales en proteínas vesiculares y están generalmente 3' en el ORF.

Como alternativa, una secuencia de retención ER opcional se deriva de la región C-terminal de una proteína vacuolar (en donde estas secuencias juegan un papel en suministrar proteínas vacuolares a vacuola vegetal). Ejemplos no limitantes de estas secuencias vacuolares se muestran en la patente de E.U.A. No. 6,054,637 incorporada en la presente a manera de referencia. Otras secuencias pueden ser fácilmente identificadas por la persona experta mediante el uso de un ensayo funcional.

Marco de lectura abierto del polinucleótido heterólogo que será expresado

Un polinucleótido heterólogo, de acuerdo con la presente invención comprende un marco de lectura abierto (ORF). El ORF, que codifica para una proteína de interés que será expresada en la semilla, puede ser cualquier ORF. Típicamente, un ORF de la presente invención puede codificar para una porción o una proteína de almacenamiento de semilla completa, enzima de la vía de ácido graso, enzima biosintética de tocoferol, enzimas degradadoras celulósicas, una vacuna, un péptido terapéutico, una proteína o péptido usado en cosméticos, enzima biosintética de aminoácidos o una enzima de ramificación de almidón. Típicamente, el ORF incluye, por ejemplo, el ácido nucleico que codifica para una proteína de interés objetivo, junto con una secuencia de señal ER de flanqueo en la región N-terminal, y una secuencia de señal de retención ER en la secuencia carboxi-terminal.

Opcionalmente, el ORF es optimizado en codones vegetales para un patrón de uso de codones preferido. La modificación de un ORF para uso de codones óptimo en plantas se describe en la patente de E.U.A. No. 5,689,052.

Elección de la proteína que será expresada

La proteína de interés que será codificada para la construcción quimérica puede ser cualquier proteína deseada. La proteína puede ser una proteína de longitud completa, o puede ser un fragmento de una proteína de longitud completa. La secuencia puede derivarse, por ejemplo, de una fuente vegetal, una fuente animal, una fuente fúngica, una fuente viral, una fuente bacteriana o puede ser una secuencia completa o parcialmente sintética.

Cualquier proteína deseada puede manipularse usando el sistema descrito

en la presente, independientemente de su especie de origen o su ubicación celular normal. Los tipos ejemplares de proteínas que pueden producirse en el sistema descrito en la presente incluyen pero no están limitados a una cinasa, una proteína estructural, una proteasa, una enzima, una amilasa, una enzima celulolítica, un inhibidor, una proteína
5 de valor nutricional incrementado, una proteína farmacéutica, una proteína o fragmento de proteína usado en cosméticos, una proteína útil para bioprocesamiento, una proteína comercialmente útil, un anticuerpo o fragmento del mismo, una proteína de membrana, una proteína nuclear, una proteína de transporte, una proteína de señalización, proteína de almacenamiento, una proteína receptora, un precursor de hormona, una hormona, un
10 péptido y una secuencia de proteínas, polipéptidos o péptidos completamente sintética.

Los genes sintéticos que codifican para proteínas de interés que incluyen pero no están limitadas a enzimas industriales, enzimas y proteínas terapéuticas, vacunas y anticuerpos pueden insertarse en el casete de expresión génica específico de semilla de soya descrito en la presente que contiene los elementos reguladores 5' y 3' de glicinina, KTI, P34, SBP, SMP o LE.
15

En una modalidad, para sacar ventaja del mecanismo de compensación de proteína para lograr expresión de proteína incrementada, los elementos de regulación de glicinina pueden usarse para conducir la expresión de proteínas en un antecedente β CS. En una modalidad, los elementos reguladores de KTI, P34, SBP, SMP o LE se pueden
20 usar para conducir la expresión de proteínas en un antecedente SP-. Las construcciones descritas en la presente pueden inducir a las proteínas a participar en el proceso del equilibrio de proteína que resulta de la supresión de conglucina y/o glicinina incrementando la síntesis y acumulación de proteínas.

En algunas modalidades, el ORF tiene una secuencia de nucleótidos que
25 codifica para la secuencia de señal de dirección al ER del gen básico de quitinasa de *Arabidopsis* fusionado 5' y una secuencia de nucleótidos que codifica para una secuencia de retención KDEL ER carboxi-terminal fusionada 3' al gen. Algunas de las proteínas que serán expresadas poseen sus propias secuencias de señal ER intrínsecas; si así es, estas secuencias pueden ser reemplazadas, si se desea, con las secuencias de señal ER
30 descritas en la presente.

Los plásmidos también pueden contener un marcador de resistencia a higromicina bajo el promotor constitutivo fuerte derivado del gen de ubiquitina 3 de papa para selección de transformantes. La transformación y producción de líneas homocigóticas que contienen los genes de interés puede producirse como se describe en la presente. Las plantas transgénicas pueden ser introducidas ya sea en la SP-, β CS u otra línea deficiente en proteínas de almacenamiento en semilla.

En una modalidad, la proteína que será expresada es una enzima celulolítica que es útil en la industria de biocombustibles. La industria de los biocombustibles está madurando rápidamente. Sin embargo, los costos de obtener muchas de las enzimas necesarias para varios procesos de producción de biocombustible pueden ser prohibitivos. Obtener las enzimas de una soya u otro cultivo de dicotiledónea, en su lugar, puede traducirse en costos más bajos asociados con la producción de biocombustibles, y también puede ser más ambientalmente amigable que los métodos tradicionales para obtener estas enzimas.

Como un ejemplo, las plantas de soya transgénicas descritas en la presente pueden usarse para crear una "biofábrica" para producir numerosas proteínas implicadas en la producción de etanol celulósico. Así, en una modalidad, la proteína que será expresada es una enzima celulósica. Ejemplos de estas enzimas incluyen pero no están limitadas a β -glucosidasa, exoglucanasa 1, exoglucanasa II, endoglucanasa, xilanasa, hemicelulasa, y ligninasa (tales como lignina peroxidasa o manganeso peroxidasa), y similares. La proteína que será expresada también puede ser cualquier otra enzima útil en la industria de biocombustibles.

En una modalidad, la proteína que será expresada es una β -glucosidasa. Una secuencia de ácido nucleico o una secuencia de aminoácidos de una β -glucosidasa de cualquier especie puede ser usada. Las β -glucosidasas ejemplares pertenecen a la familia de proteínas EC=3.21.21. La secuencia para esta enzima puede derivarse de cualquier especie adecuada, tal como de una especie de *Aspergillus*, por ejemplo, *Aspergillus niger*. Las secuencias de ácido nucleico y aminoácidos de β -glucosidasa ejemplares se muestran en SEQ ID NOS: 35-42.

En una modalidad, la proteína que será expresada es una β -glucosidasa de

Aspergillus kawachii, tal como la mostrada en SEQ ID NO: 36, o una forma modificada de β -glucosidasa de *Aspergillus kawachii*, tal como la mostrada en SEQ ID NOS: 37, 38 y 39. En otra modalidad, la proteína que será expresada es una β -glucosidasa de *Aspergillus niger* (SEQ ID NO: 40). En otra modalidad más, la proteína que será expresada es una β -glucosidasa de *Aspergillus terreus* (por ejemplo, XM_001212222; SEQ ID NO: 42).

En una modalidad, la proteína que será expresada es exoglucanasa 1, tal como en la familia de proteínas EC=3.2.1.91, también conocida como exocelobiohidrolasa I, CBH1, o 1,4- β -celobiohidrolasa. La secuencia para esta enzima puede derivarse de cualquier fuente adecuada, tal como, por ejemplo, *Trichoderma reesei* (*Hypocrea jecorina*). Una secuencia de aminoácidos de exoglucanasa 1 ejemplar se muestra en SEQ ID NOS: 43 y 44.

En una modalidad, la proteína que será expresada es exoglucanasa II, tal como en la familia de proteínas (EC=3.2.1.91), también conocida como exocelobiohidrolasa II, CBHII, CBH2 y 1,4- β -celobiohidrolasa. La secuencia para esta enzima puede derivarse de cualquier fuente adecuada, tal como, por ejemplo, *Trichoderma reesei* (*Hypocrea jecorina*). Una secuencia de aminoácidos de exoglucanasa II ejemplar se muestra en SEQ ID NOS: 45 y 46.

En una modalidad, la proteína que será expresada es una endoglucanasa. Las endoglucanasas pertenecen generalmente a la familia de proteínas EC=3.2.1.4, y también se conocen como endo-1,4- β -glucanasa E1, celulasa E1 y endocelulasa E1. La secuencia para esta enzima puede derivarse de cualquier fuente adecuada, tal como, por ejemplo, *Acidothermus cellulolyticus*. Una secuencia de aminoácidos de endoglucanasa ejemplar se muestra en SEQ ID NO: 47.

En una modalidad, la proteína que será expresada es una xilanasas. Ciertas xilanasas, tales como aquellas que pertenecen al grupo enzimático EC 3.2.1.8 (1,4-beta-D-xilan xilanohidrolasa), pueden catalizar la endohidrólisis de enlaces (1,4)-beta-D-xilosídicos en xilanos. Algunas xilanasas, tales como 1,3-beta xilanasas (EC 3.2.1.32) catalizan la degradación de enlaces 1,3-beta-D-glicosídicos. Una secuencia de aminoácidos de xilanasas ejemplar de *Aspergillus niger* se muestra en SEQ ID NO: 48.

Una secuencia de proteínas de xilanasas ejemplar de *aspergillus niger* se muestra en SEQ

ID NO: 49.

En una modalidad, la proteína que será expresada es una hemicelulasa. La secuencia de esta enzima se puede derivar de cualquier fuente adecuada.

En una modalidad, la proteína que será expresada es una ligninasa. Estas enzimas catalizan la degradación de lignina de las paredes de células vegetales. La secuencia de esta enzima puede derivarse de cualquier fuente adecuada, tal como, por ejemplo, un basidiomicetes degradador de lignina, tal como *Phanerochaete chrysosporium*. Una secuencia de aminoácidos de ligninasa ejemplar de *Phanerochaete chrysosporium* se muestra en SEQ ID NO: 50.

En una modalidad, la proteína que será expresada es una enzima lignasa tal como una enzima manganeso peroxidasa (EC 1.11.1.13). La secuencia de proteínas para esta enzima puede derivarse de cualquier fuente adecuada, tal como, por ejemplo, *Trametes versicolor*. Una secuencia de aminoácidos de manganeso peroxidasa ejemplar se muestra en SEQ ID NO: 51.

Otro tipo de enzima degradadora de lignina es lignina peroxidasa (por ejemplo, grupo enzimático EC 1.11.1.14), una hemoproteína que puede catalizar el corte oxidante de enlaces C-C y enlaces de éter (C-O-C) en un número de compuestos de lignina. Esta enzima puede catalizar la siguiente reacción: 1,2-bis(3,4-dimetoxifenil)propan-1,3-diol + H₂O. Una secuencia de aminoácidos de lignina peroxidasa ejemplar del microorganismo *Phanerochaete chrysosporium* (No. de registro P49012) se muestra en SEQ ID NO: 52.

En una modalidad, la proteína que será expresada puede tener al menos 80%, 85%, 90%, 95%, 97%, 99.5% de identidad con una de las secuencias anteriores. Como alternativa, la proteína que será expresada puede ser cualquier otra proteína de interés adecuada.

Métodos para la transformación estable de plantas

Los ácidos nucleicos pueden ser incorporados en construcciones de ácido nucleico recombinantes, típicamente construcciones de ADN, capaces de ser introducidas establemente en una célula vegetal.

Para la práctica de la presente invención, las composiciones y métodos

convencionales para preparar y usar vectores y células hospederas son empleadas, como las descritas, por ejemplo, en Sambrook *et al.* (eds.) (1989), *Molecular Cloning: A Laboratory Manual*, 2ª edición, volumen 1-3, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., y Ausubel *et al.*, ediciones, (1992) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, Nueva York.

Un número de vectores adecuados para la transformación estable de células vegetales o para el establecimiento de plantas transgénicas han sido descritos en, por ejemplo, Pouwels *et al.* (1985, supp. 1987) *Cloning Vectors: A Laboratory Manual*; Weissbach *et al.*, (1989) *Methods for Plant Molecular Biology*, Academic Press: Nueva York; y Gelvin *et al.* (1990) *Plant Molecular Biology Manual*, Kluwer Academic Publishers. Típicamente, los vectores de expresión de plantas incluyen, por ejemplo, uno o más genes vegetales clonados bajo el control transcripcional de secuencias reguladoras 5' y 3' y un marcador seleccionable dominante. Estos vectores de expresión vegetal también contienen una región reguladora de promotor (por ejemplo, una región reguladora que controla la expresión específica de células o tejidos inducible o constitutiva, regulada ambientalmente o en desarrollo), un sitio de inicio de transcripción, un sitio de unión a ribosomas, una señal de procesamiento de ARN, un sitio de terminación de transcripción y/o una señal de poliadenilación.

Existen varios métodos de transformación estable de la construcción de interés en el genoma vegetal. Los métodos ejemplares incluyen, pero no están limitados a, bombardeo de microproyectiles, electroporación, transformación mediada por *Agrobacterium* y absorción de ADN directa por protoplastos.

Los métodos de transformación a base de electroporación pueden utilizar un cultivo en suspensión de células, tallo embrionario o transformación directa de un tejido tal como un embrión inmaduro u otro tejido vegetal. Los protoplastos también pueden emplearse para la transformación por electroporación de plantas (Lazzeri *et al.*, 1985, "A procedure for plant regeneration from immature cotyledon tissue of soybean", *Plant Mol. Biol. Rep.*, 3:160-167).

Transformación por bombardeo de partículas

Un método particularmente eficiente para suministrar segmentos de ADN

de transformación a células vegetales es llamado bombardeo de microproyectiles. Este método ha sido empleado exitosamente para la transformación de un número de especies vegetales. En este método, partículas son recubiertas con ácidos nucleicos y suministradas al interior de células por una fuerza propulsora. Las partículas ejemplares incluyen aquellas que comprenden tungsteno, platino u oro. Para el bombardeo, células en suspensión son concentradas sobre filtros o medio de cultivo sólido. Como alternativa, embriones inmaduros u otras células objetivo pueden ser dispuestos en medio de cultivo sólido.

Muchos tipos de sistemas de bombardeo de partículas pueden usarse para el proceso de transformación. En un escenario de bombardeo de partículas típico, partículas de oro o tungsteno son recubiertas con la construcción de ADN de interés, y se ponen sobre una plataforma en donde se usa una fuerza potente (típicamente un gas) para acelerar las partículas en células en espera que hayan sido puestas de tal forma que acepten las partículas recubiertas con ADN. Un sistema ejemplar es el Biolistics Particle Delivery System (Bio-Rad Laboratories, Hercules, CA).

La transformación exitosa mediante bombardeo de partículas requiere generalmente que las células objetivo tengan división activa, sean accesibles a microproyectiles, puedan cultivarse *in vitro* y sean totipotentes, es decir, capaces de regeneración para producir plantas fértiles maduras. Los métodos de bombardeo de partículas adecuados se describen, por ejemplo, en patente de E.U.A. No. 5,100,792, patente de E.U.A. No. 5,179,022 y patente de E.U.A. No. 5,204,253. Además, la patente de E.U.A. No. 5,015,580 describe un método de transformación mediado por partículas de plantas de soya al bombardear el eje embrionario de una semilla de soya.

Los tejidos objetivo para el bombardeo de microproyectiles pueden incluir, pero no están limitados a células individuales, agregaciones de células, embriones inmaduros, callo embrionario joven de embriones inmaduros, microesporas, embriones derivados de microesporas y tejido de meristemo apical.

En una modalidad, el procedimiento de transformación incluye el bombardeo con partículas combinado con embriogénesis somática, como se describe, por ejemplo, en Schmidt *et al.*, (2008), *In Vitro Cellular & Developmental Biology Plant*, 44:162-168. La embriogénesis somática se describe en Bailey, 1993, *In Vitro Cellular and*

Developmental Biology Plant 29(3): 102-108. Métodos adicionales se describen en Parrott, *et al.*, 2004, Transgenic soybean, In: J.E. Specht and H.R. Boerma (eds). Soybeans: Improvement, Production and Uses, 3ª edición, -Agronomy Monograph No. 16. ASA-CSA-SSSA, Madison, WI. Pág. 265-302.

5 La transformación estable mediada por *Agrobacterium* también se puede usar para generar plantas transformadas establemente que contengan el transgén de interés. El uso de vectores de integración de plantas mediado por *Agrobacterium* para introducir ADN en células vegetales se conoce bien en la técnica (Fraley *et al.*, 1985; patente de E.U.A. No. 5,563,055). Los métodos de transformación de soya usando
10 sistemas a base de *Agrobacterium* han sido descritos en la patente de E.U.A. No. 5,569,834, la descripción de la cual se incorpora específicamente en la presente a manera de referencia en su totalidad. La patente de E.U.A. No. 5,932,782 describe métodos de transformación usando construcciones de *Agrobacterium* recubiertas sobre micropartículas que se usarán para el bombardeo de partículas. La solicitud de patente
15 de E.U.A. No. US2006260012 describe métodos para transformar células o tejidos de soya usando métodos a base de *Agrobacterium*.

La transformación de protoplastos vegetales también se puede lograr usando varios métodos, tales como precipitación de fosfato de calcio, tratamiento con polietilenglicol, electroporación y combinaciones de estos tratamientos (Potrykus *et al.*,
20 1985, Direct gene transfer to protoplasts: an efficient and generally applicable method for stable alterations of plant genomes, *UCLA Sym Mol Cell Biol*, 35:181-199). Además, la transformación de plantas mediante transferencia génica a base de electroporación o polen se describe en la patente de E.U.A. No. 5,629,183.

Marcadores seleccionables

25 En modalidades de la invención, un gen marcador seleccionable se usa para detectar las células que han completado exitosamente la transformación de la construcción génica ajena. Típicamente, se tamizan células para transformación exitosa dentro de pocos días hasta pocas semanas después del procedimiento de transformación. El tamizado para transformantes exitosos puede llevarse a cabo más
30 rápidamente al co-transformar uno o más casetes de expresión transgénicos con un

casete de expresión de marcador seleccionable y convenientemente al tamizar células de callo tomadas a través del proceso de transformación para un marcador seleccionable en cultivo o en placas de medios.

El gen marcador seleccionable en el casete de expresión de marcador seleccionable es enlazado operablemente a elementos reguladores de marcador seleccionable incluyendo un promotor y terminador. La expresión en la célula vegetal transgénica del gen marcador seleccionable generalmente codifica para una proteína que confiere resistencia a un antibiótico o herbicida. Los genes marcadores seleccionables comunes incluyen, por ejemplo, el gen de resistencia a *nptII*/kanamicina, para la selección en los medios que contienen kanamicina, el gen de fosfinotricina acetiltransferasa, para selección en medios que contengan fosfinotricina (PPT), o el gen de *hph* higromicina fosfotransferasa, para selección en medios que contengan higromicina B. Otros marcadores seleccionables incluyen los genes marcadores de resistencia de bleomicina y los genes marcadores de resistencia a glufosinato.

En una modalidad, el marcador seleccionable es higromicina fosfotransferasa. Otra secuencia de higromicina fosfotransferasa ejemplar se muestra en SEQ ID NO: 33. En otra modalidad, la secuencia marcadora seleccionable es flanqueada por elementos reguladores hacia el extremo 3' de ubiquitina de papa 3 (SEQ ID NO: 30). En una modalidad adicional, la secuencia de marcador seleccionable es flanqueada por una secuencia terminadora de ubiquitina de papa 3, tal como aquellas mostradas en SEQ ID NO: 31 y 32).

Regeneración de plantas transformadas

Cualquier método adecuado para regenerar el material vegetal transformado puede ser usado. Los métodos generales de embriogénesis somática se describen en Parrott *et al.*, 1994, "Somatic embryogenesis in legumes", pág. 199-227. En Y.P.S. Bajaj (ed) *Biotechnology in Agriculture and Forestry. Somatic embryogenesis*, vol. 31. Springer Verlag, Berlin & Heidelberg.

Cualquier medio de regeneración bien conocido puede usarse para regenerar plantas a partir del material genéticamente transformado. Según se usa en la presente, "medio de cultivo de plantas" se refiere a cualquier medio usado en la técnica

para soportar viabilidad y crecimiento de una célula o tejido vegetal, o para el crecimiento de especímenes vegetales enteros. Estos medios incluyen comúnmente componentes definidos incluyendo, pero no limitados a: compuestos macronutrientes que proporcionan fuentes nutricionales de nitrógeno, fósforo, potasio, azufre, calcio, magnesio y hierro; 5 micronutrientes, tales como boro, molibdeno, manganeso, cobalto, zinc, cobre, cloro y yodo; carbohidratos; vitaminas; fitohormonas; agentes de selección (para células o tejidos transformados, por ejemplo, antibióticos o herbicidas) y agentes de gelificación (por ejemplo, agar, Bactoagar, agarosa, fitagel, Gelrite, etc.). El medio puede ser ya sea sólido o líquido. Los medios y métodos de regeneración adecuados se describen, por ejemplo, 10 en Schmidt, *et al.*, 2005, "Towards normalization of soybean somatic embryo naturation," *Plant Cell Rep.* 24:383-391.

Una vez que las plantas transformadas putativas son identificadas, el tejido puede ser probado para confirmar que el transgen haya sido transformado establemente en el genoma vegetal.

15 La producción de una línea homocigótica que tiene el gen heterólogo típicamente requerirá de varias etapas de cruce adicionales. En algunas modalidades, el tejido transformado es luego cultivado en plantas maduras o "transformantes primarios" (T0), los cuales son después auto-cruzados. Las semillas de esta auto-cruza se cultivan (generación T1) y los transformantes positivos se identifican y se auto-cruzan. Estas 20 semillas se cultivan después para crear plantas maduras (generación T2) y se tamizan para identificar la presencia homocigótica de la secuencia transformada, para producir una línea vegetal homocigótica.

Dicotiledóneas

Una planta dicotiledónea (o "dicot"), según se proporciona en la presente, 25 puede ser cualquier dicotiledónea. Opcionalmente, la dicotiledónea es un miembro del orden Fabales. Opcionalmente, la dicotiledónea es un miembro de la familia de las Fabáceas (conocidas comúnmente como legumbres). Opcionalmente, la dicotiledónea es un grano de soya, tal como *Glycine max*.

Ejemplos de dicotiledóneas útiles en las composiciones y métodos 30 proporcionados en la presente son: *Abrus Adans.* (por ejemplo, semilla de abrus), *Acacia*

- P. Mill.*, (por ejemplo, acacia), *Adenanthera* L. (por ejemplo, árbol de baya china), *Aeschynomene* L. (por ejemplo, jointvetch), *Azalia* Smith (por ejemplo, caoba), *Albizia* Durazz. (por ejemplo, albizia), *Alhagi Gagnebin* (por ejemplo, alhagi), *Alysicarpus* Neck, ex Desv. (por ejemplo, moneda), *Amorpha* L. (por ejemplo, false índigo, arbusto índigo),
- 5 *Amphicarpa*), *Amphicarpea* Ell. ex Nutt. (por ejemplo, *amphicarpea*, cacahuete), *Anadenanthera* Speg. (por ejemplo, *anadenanthera*), *Andira* Juss. (por ejemplo, andira), *Anthyllis* L. (por ejemplo, kidneyvetch), *Apios* Fabr. (por ejemplo, cacahuete), *Arachis* L. (por ejemplo, cacahuete), *Aspalathus* L. (por ejemplo, aspalato), *Aspalthium* Medik.), *Astragalus* L. (por ejemplo, astragales, *astragalus* spp., locoweed, locoweed species,
- 10 *milkvetch*), *Baphia* Lodd. (por ejemplo, bafia), *Baptisia* Vent. (por ejemplo, baptisia, False índigo, índigo silvestre), *Barbieria* DC. (por ejemplo, barbieria), *Bauhinia* L. (por ejemplo, bauhinia), *Bituminaria* Heister ex Fabr. (por ejemplo, bituminaria), *Bonaveria* Scop.), *Brongniartia* Kunth (por ejemplo, greentwig), *Brya* P. Br. (por ejemplo, coccuswood), *Butea* Roxb. ex Willd. (por ejemplo, butea), *Caesalpinia* L. (por ejemplo, *caesalpinia*, nicker,
- 15 *poinciana*), *Caiaandra* Benth.), *Cajanus* Adans. (por ejemplo, cájano), *Calliandra* Benth. (por ejemplo, callandra, mesquita falsa, *stickpea*), *Calopogonium* Desv. (por ejemplo, calopogonio), *Camelina* sp. (por ejemplo, "lino falso"), *Canavaia* Adans. Mut. Dc, *Canavalia* Adans. (por ejemplo, jackbean), *Caragana* Fabr. (por ejemplo, *peashrub*), *Cassia* L. (por ejemplo, casia, especies de casia), *Centrosema* (DC.) Benth. (por ejemplo,
- 20 *guisante mariposa*, *centrosema*), *Ceratonia* L. (por ejemplo, *ceratonia*), *Cercidium* L. R. Tulasne, *Cercis* L. (por ejemplo, *redbud*), *Chamaecrista* (L.) Moench (por ejemplo, *guisante sensible*), *Chamaecystis* Link (por ejemplo, *chamaecystis*), *Chamaesenna* (Dc.) Raf. Ex Pittier), *Chapmannia* Torr. & Gray (por ejemplo, *chapmannia*), *Christia* Moench (por ejemplo, *guisante isleño*), *Cicer* L. (por ejemplo, *cicer*), *Cladrastis* Raf. (por ejemplo,
- 25 *podo sudafricano*), *Clitoria* L. (por ejemplo, *clitoria*, alas de pichón), *Codariocalyx* Hassk. (por ejemplo, *loto*), *Cojoba* Britt. & Rose (por ejemplo, *cojoba*), *Cologania* Kunth (por ejemplo, *cologania*), *Colutea* L. (por ejemplo, *colutea*), *Copaifera* L. (por ejemplo, *copaifera*), *Coronilla* L. (por ejemplo, *crownvetch*), *Corynella* DC. (por ejemplo, *corynella*), *Coursetia* DC. (por ejemplo, *babybonnets*, *coursetia*), *Cracca* Benth.), *Crotalaria* L. (por
- 30 ejemplo, *rattlebox*), *Crudia* Schreb.), *Cullen* Medik. (por ejemplo, *scurfpea*), *Cyamopsis*

- DC. (por ejemplo, *cyamopsis*), *Cynometra* L. (por ejemplo, *cynometra*), *Cytiscus Linnaeus*), *Cytisus* Desf. (por ejemplo, rusco), *Dalbergia* L. f. (por ejemplo, palo rosa de la India), *Dalea* L. (por ejemplo, dalea, *dalea* spp., trébol del prado, trébol del prado, tréboles del prado), *Daniellia Bennett* (por ejemplo, danielia), *Delonix Raf.* (por ejemplo, *delonix*),
- 5 *Derris Lour.* (por ejemplo, derris), *Desmanthus Willd.* (por ejemplo, *bundleflower*), *Desmodium* Desv. (por ejemplo, legumbres perennes, loto, trébol, trébol amarillo), *Dialium* L. (por ejemplo, dialio), *Dichrostachys (DC.) Wight & Arn.* (por ejemplo, *dichrostachys*), *Dioclea Kunth* (por ejemplo, dioclea), *Diphysa Jacq.* (por ejemplo, *diphysa*), *Dipogon Lieb.* (por ejemplo, dipogo), *Dipteryx Schreber* (por ejemplo, dipterix), *Ebenopsis Britt. & Rose*
- 10 (por ejemplo, ébano de Texas), *Entada Adans.* (por ejemplo, vid de callingcard), *Enterolobium Mart.* (por ejemplo, enterolobio), *Eriosema (DC.) D. Don* (por ejemplo, guisante arenoso), *Errazurizia Phil.* (por ejemplo, rusco), *Erythrina L.* (por ejemplo, eritrina), *Erythrophleum Afzel. ex R. Br.* (por ejemplo, sasswood), *Eysenhardtia Kunth* (por ejemplo, kidneywood), *Faidherbia A. Chev.* (por ejemplo, acacia), *Falcataria (Nielsen)*
- 15 *Barneby & Grimes* (por ejemplo, pluma de pavo real), *Flemingia Roxb. ex Ait. f.* (por ejemplo, flemingia), *Galactia P. Br.* (por ejemplo, guisante lechoso), *Galega L.* (por ejemplo, mala hierba), *Genista L.* (por ejemplo, rusco), *Genistidium LM. Johnston* (por ejemplo, guisante de brocha, *genistidium*), *Gleditsia L.* (por ejemplo, algarrobo melífero, algarrobo), *Gliricidia Kunth* (por ejemplo, quickstick), *Glottidium Desv.* (por ejemplo, glotidio), *Glycine max* (por ejemplo, soya), *Glycine Willd.* (por ejemplo, soya), *Glycyrrhiza*
- 20 L. (por ejemplo, regaliz), *Gymnocadus Lam.*), *Gymnocladus Lam.* (por ejemplo, cafetal), *Haematoxylum L.* (por ejemplo, *haematoxylum*), *Halimodendron Fischer ex DC.* (por ejemplo, *halimodendron*), *Havardia Small* (por ejemplo, havardia), *Hedysarum L.* (por ejemplo, veza dulce, *sweetvetch*), *Hippocrepis L.* (por ejemplo, hipocrepis),
- 25 *Hoffmannseggia Cav.* (por ejemplo, *hoffmanseggia*, guisante, especies de guisante, *Hoffmanseggia Cavanilles.*), *Hoita Rydb.* (por ejemplo, *leather-roof*), *Hymenaea L.* (por ejemplo, *hymenaea*), *Indigofera L.* (por ejemplo, índigo), *Inga P. Mill.* (por ejemplo, inga), *Inocarpus J. R. & G. Forst.*), *Kanaloa D.H. Lorence & K. R. Wood* (por ejemplo, kanaloa), *Kummerowia Schindl.* (por ejemplo, *kummerowia*), *Lablab Adans.* (por ejemplo, *lablab*),
- 30 *Laburnum Medik.* (por ejemplo, árbol dorado), *Lathyrus L.* (por ejemplo, chícharo, vid,



- peavine* spp.), *Lens* P. Mill, (por ejemplo, lenteja), *Lespedeza Michx.* (por ejemplo, lespedeza, lespedeza perenne), *Leucaena Benth.* (por ejemplo, leadtree), *Lonchocarpus Kunth* (por ejemplo, lancepod), *Lotononis (DC.) Ecklon & Zeyh.* (por ejemplo, lotononis), *Lotus L.* (por ejemplo, veza, deervetch spp., trébol), *Lupinus L.* (por ejemplo, lupina, lupinas), *Lysiloma Benth.* (por ejemplo, tamarindo falso, lysiloma), *Maackia Rupr.* (por ejemplo, maackia), *Machaerium Pers.* (por ejemplo, machaerium), *Macroptilium (Benth.) Urban* (por ejemplo, bushbean, macroptilium), *Macrotyloma (Wight & Amort) Verde,* (por ejemplo, macrotyloma), *Marina Liebm.* (por ejemplo, trébol del prado falso, marina), *Medicago L.* (por ejemplo, alfalfa), *Meibomia Heist. Ex Fabr.), Melilotus P. Mill,* (por ejemplo, trébol dulce, sweetclover), *Mimosa L.* (por ejemplo, mimosa, planta sensible), *Mucuna Adans.* (por ejemplo, mucuna), *Myrospermum Jacq.* (por ejemplo, mirospermo), *Myroxylon L. f.* (por ejemplo, myroxylon), *Neonotonia Lackey* (por ejemplo, neonotonia), *Neorudolphia Britt.* (por ejemplo, neorudolphia), *Neptunia Lour,* (por ejemplo, neptunia, puff), *Nissolia Jacq.* (por ejemplo, nissolia, yellowhood), *Olneya Gray* (por ejemplo, olneya), *Onobrychis P. Mill,* (por ejemplo, esparceta), *Ononis L.* (por ejemplo, parabueyes), *Orbexilum Raf.* (por ejemplo, leather-root, orbexilum), *Ormosia G. Jackson* (por ejemplo, ormosia), *Ornithopus L.* (por ejemplo, pie de pájaro), *Oxyrhynchus Brandeg.* (por ejemplo, oxyrhynchus), *Oxytropis DC.* (por ejemplo, maleza loca), *Pachyrhizus L. C. Rich, ex DC.* (por ejemplo, pachyrhizus), *Paraserianthes I. Nielsen* (por ejemplo, paraserianthes), *Parkia R. Br.* (por ejemplo, parquia), *Parkinsonia L.* (por ejemplo, paloverde, parkinsonia), *Parryella Torr. & Gray ex Gray* (por ejemplo, parryella), *Pediomelum Rydb.* (por ejemplo, raíz, raíz de la India, pediomelo, scurfpea), *Peltophorum (T. Vogel) Benth.* (por ejemplo, peltophorum), *Pentaclethra Benth.* (por ejemplo, pentaclethra), *Pericopsis Thwaites), Peteria Gray* (por ejemplo, peteria), *Phaseolus Linnaeus), Phaseolus L.* (por ejemplo, judía, haba), *Physostigma Balf.* (por ejemplo, physostigma), *Pickeringia Nutt. ex Torr. & Gray* (por ejemplo, guisante chaparral), *Pictetia DC.* (por ejemplo, pictetia), *Piscidia L.* (por ejemplo, piscidia), *Pisum L.* (por ejemplo, chícharo), *Pitcheria Nutt.), Pithecellobium Mart,* (por ejemplo, blackbead, pithecellobium), *Poitea Vent.* (por ejemplo, wattapama), *Pongamia Ventenat), Prosopis L.* (por ejemplo, mesquite), *Psophocarpus Necker ex DC.* (por ejemplo, psophocarpus), *Psoralea*

Linnaeus), *Psoralidium* Rydb. (por ejemplo, *breadroot*, *scurfpea*), *Psorothamnus* Rydb. (por ejemplo, *dalea*, *smokebush*), *Pterocarpus* Jacq. (por ejemplo, *pterocarpus*), *Pueraria* DC. (por ejemplo, kudzu), *Retama* Raf., *nom. cons.*), *Rhynchosia* Lour, (por ejemplo, *snoutbean*), *Robinia* L. (por ejemplo, *algarrobo*), *Rupertia* J. Grimes (por ejemplo, *rupertia*), *Sabinea* DC. (por ejemplo, *sabinea*), *Samanea* Merr. (por ejemplo, *raintree*), *Schizolobium* Vogel (por ejemplo, *Brazilian firetree*), *Schrankia* Willd. (por ejemplo, *schrankia*), *Scorpiurus* L. (por ejemplo, *cola de escorpión*), *Secula* Small), *Senna* P. Mill, (por ejemplo, *sena*), *Sesbania* Scop, (por ejemplo, *cáñamo ribereño*, *sesbania*), *Sophora* L. (por ejemplo, *necklacepod*, *sophora*), *Spartium* L. (por ejemplo, *rusco*), *Sphaerophysa* DC. (e.g., *sphaerophysa*), *Sphenostylis* E. Meyer (por ejemplo, *sphenostylis*), *Sphinctospermum* Rose (por ejemplo, *sphinctospermum*), *Sphinctospermum* Rose), *Stahlia* Bello (por ejemplo, *stahlia*), *Strongylodon* Vogel (por ejemplo., *strongylodon*), *Strophostyles* Ell. (por ejemplo, *haba peluda*, *haba silvestre fuzzy bean*, *fuzzybean*, *wildbean*), *Stryphnodendron* C. Martius (por ejemplo, *stryphnodendron*), *Stylosanthes* Sw. (por ejemplo, *pencilflower*), *Sutherlandia* R. Br. (por ejemplo, *sutherlandia*), *Swainsonia* Salisb.), *Tamarindus* L. (por ejemplo, *tamarindo*), *Taralea* Aublet (por ejemplo, *taralea*), *Tephrosia* Pers. (por ejemplo, *hoarypea*, *tephrosia*), *Teramnus* P. Br. (por ejemplo, *teramnus*), *Tetragonolobus* Scop, (por ejemplo, *tetragonolobus*), *Thermopsis* R. Br. ex Ait. f. (por ejemplo, *goldenbanner*, *goldenpea* spp. (*golden banner*), *thermopsis*), *Ticanto* Adans. (por ejemplo, *gray nicker*), *Trifolium* L. (por ejemplo., *trébol*, *clover* spp., *trefles*), *Trigonella* L. (por ejemplo, *fenogreco*), *Ulex* L. (por ejemplo., *tojo*), *Vexillifera*), *Vicia* L. (por ejemplo, *veza*, *vetch* spp.), *Vigna* Savi (por ejemplo., *frijol de vaca*, *vigna*), *Wisteria* Nutt. (por ejemplo, *wisteria*), *Zapoteca* H. Hernandez (por ejemplo, *guisante blanco*), *Zornia* J. F. Gmel. (por ejemplo, *zornia*), y similares.

25 Procesamiento de proteínas

Las proteínas producidas mediante los métodos descritos en la presente pueden extraerse de las semillas y usarse directamente en un proceso comercial. Como alternativa, las proteínas pueden ser purificadas parcial o completamente, luego ya sea almacenadas o usadas inmediatamente. Además, una "harina" o molienda de soya puede prepararse a partir de las plantas transgénicas, y el material se puede almacenar hasta

SS

que se requiera.

Aislamiento, purificación y análisis de proteínas

Los niveles de proteínas pueden probarse usando procedimientos proteómicos estándares. El contenido total de proteínas puede determinarse, por ejemplo, usando un ensayo a base del método de Bradford (Bradford, 1976, *Analytical Biochem*, 72:248-254).

El análisis de proteínas puede proceder de acuerdo con métodos ampliamente conocidos. La identidad de la proteína puede confirmarse, por ejemplo, mediante el uso de ensayos a base de gel, inmunoblots y espectrometría de masas. El tamaño de la proteína puede determinarse, por ejemplo, usando cromatografía de líquidos de alto rendimiento.

La actividad enzimática de una masa específica de material crudo o de enzima purificada puede llevarse a cabo usando protocolos estándares para ensayar la enzima de interés. Las mediciones de la estabilidad de la enzima, perfiles de temperatura, perfiles de pH y vida media útil de la enzima pueden determinarse también usando métodos estándares.

La proteína transgénica de interés puede purificarse a cualquier grado que se requiera para uso adicional. El grado de purificación requerido puede depender de muchos parámetros, tales como costo, estabilidad de la proteína purificada contra la proteína que permanezca en la semilla, requerimientos hacia abajo, remoción de contaminantes, requerimiento de procesamiento adicional de la proteína (es decir, doblez de proteína adecuado o modificaciones post-traduccionales), etc. Un ejemplo de un método para aislar y purificar una proteína producida en semillas de soya se muestra en el ejemplo 16.

En modalidades de la invención, las semillas de soya pueden ser procesadas, por ejemplo, por molienda o pulverización. Un extracto crudo del material molido puede prepararse al añadir un líquido y agitar durante un tiempo, seguido por filtración opcional para remover las partículas grandes. Como alternativa, en algunas modalidades, podría no ser necesario purificar la proteína de interés en absoluto – la molienda de semilla que contenga la proteína de interés, junto con el resto de la semilla

de soya, simplemente puede ser usada.

Ensayo inmunoabsorbente ligado a enzimas (ELISA)

En algunas modalidades, un método de ELISA que se conoce generalmente en la técnica, puede usarse para el análisis de la proteína expresada. Usando la producción de beta-glucosidasa en soya como un ejemplo, se puede usar el siguiente escenario. Placas de varios pocillos son recubiertas con anticuerpo anti-beta-glucosidasa de conejo, después el extracto de semilla de soya y muestras de control se añaden a pocillos individuales de la placa y se incuban durante 1 hora a 35°C. Conjugado peroxidasa de rábano anti-conejo es luego añadido a cada pocillo e incubado durante 1 hora a 35°C, seguido por la adición y el substrato de tetrametilbencidina (Sigma, E.U.A.) e incubación durante 3 minutos a temperatura ambiente. La reacción se detiene al añadir H₂SO₄ 1N a cada pocillo. Las placas son leídas a 450 nm en un lector de microplaca (Bio-Rad, modelo 3550) y los datos se procesan, por ejemplo, usando MICROPLATE MANAGER™ III (Bio-Rad). Los resultados de un análisis de varias líneas homocigóticas se miden para determinar la cantidad de proteína expresada por semilla.

Almacenamiento de plantas transgénicas en la semilla de soya

Las semillas de soya transgénicas descritas en la presente también se pueden usar como recipientes de almacenamiento de proteínas naturales. Semillas de soya secas y maduras que contienen las proteínas transgénicas pueden almacenarse a temperatura ambiente o menos, hasta que se requieran. Así, el procesamiento adicional de la proteína puede retrasarse hasta el momento de uso. Este método puede ser un medio eficiente y económico de almacenar enzimas transgénicas que se usarán para propósitos comerciales.

Características técnicas inesperadas de la presente invención

Sorprendentemente, la presente invención se traduce en una planta con características superiores. Por ejemplo, la invención permite la producción ya sea de una proteína de interés que se presente naturalmente en la naturaleza, se modifique sobre una proteína natural, sea sintética (nuevo diseño) o una proteína que tenga combinaciones de las mismas. Puede producirse como una proteína que tenga propiedades fisicoquímicas deseables (tales como solubilidad o estabilidad bajo diferentes

58

condiciones; puede producirse como una proteína de fusión enlazada a residuos que ayuden a la purificación o incrementen su utilidad).

La presente invención es especialmente útil para producir proteínas que de otra manera sean sensibles a degradación y, a través del secuestro mediante compartimentalización enseñada en la presente, puedan demostrar estabilidad notable.

La presente invención es especialmente útil para producir proteínas que requieran baja humedad para conservar su función. Estas proteínas, por ejemplo, pueden ser dirigidas a vesículas derivadas del ER en una semilla y almacenarse durante meses o años y conservar su valor nutricional, actividad enzimática o una propiedad deseada.

La presente invención es especialmente útil para producir proteínas que se pueden usar comercialmente en forma no purificada o forma parcialmente purificada gracias al alto nivel de abundancia en una semilla u otra parte de la planta. En algunas modalidades, se pueden añadir porciones que proporcionen o eviten la agregación.

Las construcciones de la presente invención pueden combinarse con otras características útiles, tales como elementos reguladores adicionales que permitan al gen ser encendido o apagado mediante señales temporales o externas. Ejemplos de modalidades útiles se muestran en la tabla 2.

Tabla 2
Secuencias quiméricas ejemplares

PLANTA	GENÉTICA	REGIÓN PROMOTORA/REG U-LADORA 5' DE	SECUENCIA DE SEÑAL ER	ORF	SECUENCIA DE RETENCION ER	3'TED DE
Soya	Deficiente en glicinina y congloncinina	KTI	SEQ ID NO: 12	Enzima industrial	KDEL	KTI
Soya	Deficiente en glicinina, congloncinina y KTI	LE	SEQ ID NO: 12	Enzima industrial	KDEL	LE
Soya	Deficiente en glicinina y congloncinina	P34	SEQ ID NO: 9	β-glucosidasa	KDEL	P34
Soya	Deficiente en B-congloncinina	Glicinina	SEQ ID NO: 11	Fragmento de anticuerpo	HDEL	Glicinina
Soya	Deficiente en glicinina	SMP	SEQ ID NO: 12	Exoglucanasa I	SEKDEL	Conglicini na
Soya	Deficiente en B-congloncinina	Glicinina	SEQ ID NO: 9	Exoglucanasa II	KDEL	Glicinina
Haba común (Phaseolus sp.)	Deficiente en glicinina	KTI	SEQ ID NO: 10	Endoglucanasa	HDEL	GBP
Acacia sp.	Deficiente en glicinina	KTI	SEQ ID NO: 12	β-glucosidasa	KDEL	KTI
Camelina sp.	Deficiente en congloncinina	Glicinina	SEQ ID NO: 10	Ligninasa	KDEL	glicinina

Ejemplos

Los siguientes ejemplos se llevan a cabo usando técnicas estándares, las cuales se conocen bien y son de rutina para aquellos expertos en la técnica, excepto cuando se describa lo contrario en detalle. Los ejemplos son ilustrativos, más no limitan la invención.

Ejemplo 1

Generación de la línea de supresión de proteína de almacenamiento "SP-"

Una construcción de ARNi diseñada de acuerdo con la figura 2 para suprimir el contenido de proteína de almacenamiento en semillas fue transferida a soya usando protocolos de transformación biolística (Parrott, *et al.*, 2004, "Transgenic soybean", *In*: J.E. Specht and H. R. Boerma (eds). Soybeans: Improvement, Production, and Uses, 3ª edición. – Agronomy Monograph No. 16. ASA-CSA-SSSA, Madison, WI. Pág. 265-302). Un casete de ARNi específico para la supresión simultánea de las proteínas de almacenamiento de soya endógenas y FAD2-1 omega-6 ácido graso desaturasa se produjo al invertir secuencias específicas para estos marcos de lectura abiertos que flanquean un intrón bajo el promotor de glicinina y terminador 3'. Una región de 331pb del gen A1bB2 de glicinina (SEQ ID NO: 55) se puso adyacente a una región de 128 pb del gen FAD2-1a (SEQ ID NO: 57). Este ADN heterólogo de 459 pb fue luego puesto en repeticiones invertidas aproximadamente un intrón. El intrón derivado sintéticamente se obtuvo de una porción de vector de silenciamiento p3UTR12850S. Este casete (SEQ ID NO: 56) se puso después bajo los elementos reguladores de glicinina (figura 2). El ARNi de FAD2 se añadió como una característica opcional de la construcción para proporcionar un marcador para potencial de tamizado adicional para fenotipo alto oleico y para mantener consistencia con la supresión de conglucina anterior que también incluyó la supresión de FAD2 (Kinney and Herman ("Cosuppression of the α Subunits of beta-conglycinin in Transgenic Soybean Seeds Induces the Formation of Endoplasmic Reticulum-Derived Protein Bodies", *Plant Cell* 13:1165-1178 (2001)).

Los embriones somáticos regenerados y semillas T0 fueron tamizados mediante ID SDS/PAGE para distribución de proteínas total y con inmunoblots que

ensayan para reactividad cruzada con anticuerpos anti-glicinina y anti-conglicinina.

Las líneas transgénicas recuperadas no sólo exhibieron el fenotipo de contenido de glicinina suprimido sino que también exhibieron una supresión esencialmente completa de las subunidades α/α' y β de conglicinina. Las líneas generadas en la presenta con una supresión tanto de glicinina como de conglicinina serán conocidas como SP- (supresión de proteína de almacenamiento).

Ejemplo 2

Contenido de proteínas y aceite en SP-

Las líneas SP- fueron regeneradas en plantas de soya. Las plantas resultantes crecieron y dieron semillas equivalentes en comparación con los controles. Las semillas T0 fueron fragmentadas para ensayar el fenotipo y las semillas SP- se volvieron a cultivar y se reseleccionaron dos veces más para producir una población homocigótica. El fenotipo SP- fue estable a lo largo de cada generación subsecuente con subunidades de α/α' y β -conglicinina no siendo detectadas y los niveles de glicinina ampliamente reducidos. El nivel de ácido oleico en las semillas SP- fue > 94% indicando que la supresión del marcador de tamizado de FAD2 también estaba presente.

El tamaño y peso en seco para las semillas durmientes SP- cultivadas en invernadero que promedian 146 mg es similar al promedio de semilla durmiente "Jack" de variedad cultivado en invernadero tipo silvestre (WT) de 163 mg. El contenido de proteínas y aceite total de la SP- (40.2%, 19.1%) y la variedad "Jack" WT (37.5%, 20.5%) es similar. Así, los ensayos demuestran que la supresión de proteínas que corresponden a una mayoría de la proteína total de soya se traduce en el reequilibrio de la composición de proteína de soya a un contenido de proteína/aceite casi idéntico y tamaño de semilla.

Ejemplo 3

Microscopia electrónica e inmunocitoquímica con inmuno-oro

Muestras de tejido fueron criofijadas con un dispositivo de alta presión Balzer (Bal-Tech, Principality of Liechtenstein), sustituidas por congelación con acetona/OsO4 e incrustadas en plástico epon. Secciones ultradelgadas fueron teñidas tanto con acetato de uranilo acuoso saturado como con citrato de plomo (33 mg/ml) antes de la observación. Se llevó a cabo después el análisis inmunocitoquímico. Muestras

paralelas fueron criofijadas y luego procesadas mediante sustitución por congelación sin ningún fijador. Las muestras sustituidas fueron transferidas a resina Lowicryl HM-20 que se polimerizó mediante iluminación con luz UV. Secciones delgadas fueron marcadas con anti-GFP MAb (Clontech) o anti-glicinina policlonal de conejo producida previamente por este laboratorio. Las secciones fueron marcadas indirectamente con anti-IgG (conejo o ratón) de oro coloidal 10 nm (Sigma), luego se contrastaron con 5% de acetato de uranilo antes de la observación EM. Toda la TEM se llevó a cabo con un microscopio LEO 912AB con imágenes capturadas usando una cámara CCD 2k x 2k operada en modo de montaje.

10

Ejemplo 4

Morfología bruta normal de SP-

Para examinar la estructura celular de la SP- en comparación con la WT, cotiledóneas maduras de ambas fueron preparadas mediante criofijación de alta presión y las muestras resultantes se sustituyeron por congelación con acetona/OsO₄ y luego se incrustaron en plástico Epon, y se procesaron como se detalló arriba. Los granos de soya SP- forman PSVs (figura 3A) que son notoriamente similares en tamaño y apariencia a las PSVs formadas en semillas WT (figura 3B). Las PSVs en la SP- poseen una matriz amorfa rellena de proteínas típica de soya.

15

Ejemplo 5

Análisis de proteínas bidimensional

20

Las proteínas totales fue aislada de semillas de soya maduras como se describe en Joseph *et al.*, (2006), "Evaluation of Glycine germplasm for nulls of the immunodominant allergen P34/Gly m Bd 30k", *Crop Science* 46:1755-1763. Brevemente, un total de 150 µg de proteína se cargaron en una tira de gel (IPG) con gradiente de pH inmovilizado de 11 cm (pH 3-10 nL) (BioRad, Hercules CA) y luego se hidrataron durante la noche. Se llevó a cabo enfoque isoeléctrico (IEF) para un total de 40kVh usando Protean IEF Cell (BioRad) y luego se corrieron en el gel SDS-PAGE de segunda dimensión (gradiente lineal 8-16%). Las células fueron teñidas durante la noche en azul de coomassie al 0.1% en 40% (v/v) de metanol y 10% (v/v) de ácido acético mientras se absorbía y la inmunodetección subsecuente usando anticuerpo monoclonal GFP

25

30

(Clontech Inc, Mountain View, CA) como se describe en Joseph et al., (citado arriba). Cada muestra se corrió en geles triplicados y se escaneó y analizó en Phoretix 2D Evolution (versión 2005; NOnlinear Dynamics Ltd.). Los puntos de GFP identificados en el inmunoblot dejaron que los puntos correspondientes fueran ubicados en los geles replicados y el volumen de estos puntos fue normalizado contra el volumen de puntos de proteoma completo para determinar el volumen porcentual de la proteína GFP en el proteoma de semilla de soya soluble completo.

Ejemplo 6

Identificación de proteínas expresadas a un nivel incrementado debido a SP-

10 El análisis proteómico de las soyas SP- muestra que otras proteínas de almacenamiento compensan la ausencia de polipéptidos de proteína de almacenamiento.

El fraccionado por IEF/SDS-PAGE bidimensional (2D) de la SP- en comparación con la WT muestra un dramático cambio en la distribución de puntos de las proteínas que resulta a partir de la supresión de las proteínas de almacenamiento.

15 Las figuras 4A a 4C muestran la comparación entre la supresión de proteína WT y de almacenamiento usando geles 2D de segunda dimensión de amplia gama, pH 3-10. La tinción de proteína total muestra que hay un cambio a gran escala en la distribución de proteínas en SP- con las proteínas de almacenamiento ausentes reemplazadas por otros puntos de proteínas abundantes.

20 La supresión de las proteínas de almacenamiento se confirmó más al sondear un inmunoblot replicado con anticuerpos específicos para fracción de proteína de almacenamiento de conglucina y glicina que mostraron una ausencia de las subunidades e isoformas de conglucina y una reducción significativa de las subunidades e isoformas de glicina. Los geles 2D de SP- por triplicado fueron evaluados mediante el
25 volumen de punto de las proteínas totales en comparación con la tipo silvestre. Las proteínas alteradas significativamente fueron graduadas por examen visual con la ayuda de volumen de punto/escaneo de gel. Los puntos de proteína seleccionados fueron extirpados, sometidos a fragmentación triptica y analizados mediante espectrometría de masas MS/MS tándem. El mapa de los puntos de proteínas numerados seleccionados
30 para análisis de espectroscopia de masas se muestra en la figura 4C.

62

Los datos de proteómica compilados a partir de los puntos de proteínas en la figura 4C se muestran en la tabla 3. Gran parte del reequilibrio de contenido de proteínas se debe a un contenido incrementado del inhibidor de tripsina Kunitz (KTI), lectina de soya (LE), también conocida como “aglutinina”, y el alérgeno de soya inmunodominante P34 o Gly m Bd 30k. Otras proteínas con contenido incrementado incluyen proteína de unión a glucosa (GBP) y proteína de maduración de semilla (SMP). La identificación de proteómica y espectrometría de masas muestra que el reequilibrio del acortamiento de las proteínas de almacenamiento ocurre por acumulación incrementada sólo de unas pocas proteínas.

Tabla 3

Punto No.	KDal	pT	Registro No.	Nombre de proteína	% promedio	Puntuación Mascot	No. de péptidos
1	60.9	6.42	170064	Proteína de unión a glucosa	37	604	22
2	60.9	6.42	170064	Proteína de unión a glucosa	15	416	8
3							
4	50.6	6.33	70010	Proteína madura	14	246	9
5	23.7	6.07	70024	Proteína asociada a maduración	67	308	9
6	60.9	6.42	170064	Proteína de unión a glucosa	12	373	7
7							
8	40.7	6.19	22597178	Alcohol deshidrogenasa	20	278	10
9	58.6	5.52	4249566	Glicinina	27	171	9
10	40.7	6.19	22597178	Alcohol deshidrogenasa	38	395	17
11	35.3	5.96	4102190	Proteína de maduración de semilla de 35 kDa	41	353	19
12	32.0	6.60	9622153	Proteína de maduración de semilla PM34	31	292	12
13							
14	32	6.6	9622153	Proteína de maduración de semilla PM34	35	274	13
15	64.4	5.21	18641	Glicinina	20	291	7
16	43.1	5.65	1199563	Proteasa vacuolar de semilla madura de 34 kDa	10	124	5
17	27.5	5.15	3114258	Aglutinina	45	436	14
18							
19	27.6	5.15	3114258	Aglutinina	41	345	11
20	64.3	5.21	18641	Glicinina G4	8	141	5
21							
22	24.3	4.99	18770	Inhibidor de tripsina A	17	141	5
23	47.1	8.68	434061	Globulina 7S	28	192	9
24	27.6	5.15	3114258	Aglutinina	28	225	9
25	64.4	5.21	18641	Glicinina	10	204	6
26	64.4	5.21	18641	Glicinina	19	353	12
27	20.3	4.61	354134	Inhibidor de tripsina Kunitz	73	629	23

La figura 6 muestra una gráfica circular que demuestra las cantidades de

varias proteínas en semillas de soya WT ("Jack") contra las semillas SP-. La gráfica demuestra claramente el reequilibrio o proceso de compensación, en donde un incremento de ciertas proteínas ocurre cuando otras proteínas se disminuyen en la semilla.

5

Ejemplo 7

Semillas SP- forman PSVs en una morfología y patrón correctos por desarrollo

Las vacuolas de almacenamiento de proteínas (PSVs) de semillas dicotiledóneas tales como soya se forman por la subdivisión de la vacuola central en forma coordinada con la síntesis y deposición de las proteínas de almacenamiento. Esto se traduce en PSVs rellenas de proteínas que llenan gran parte del citoplasma de semillas que están acumulando una proteína de almacenamiento.

Las células de parénquima de almacenamiento de plantas que tienen una supresión tanto de proteínas de almacenamiento de glicinina como de β -conglucina, tales como la línea SP- descrita en los ejemplos anteriores, parecen contener sólo PSV. Esto indica que las proteínas PSV compensadoras son y siguen siendo vacuolares sin ninguna redirección de las proteínas en los cuerpos proteicos derivados de ER. La estructura y distribución de todos los demás organelos subcelulares y estructuras parecieron ser idénticas en el SP- y control de wt.

En contraste, la supresión individual de conglucina, ya sea mediante manipulación genética dirigida como se muestra abajo, por ejemplo, en el ejemplo 11, o la mutación de origen natural se traduce en una gran fracción de esa glicinina que permanece en la forma de proglicina precursora que es acumulada en cuerpos de proteínas derivados de ER.

Ejemplo 8

25

Semillas de soya SP- exhiben pocos cambios en transcripción

Las soyas SP- de maduración tardía se compararon con la WT (tipo silvestre) usando el chip génico Affymetrix DNA con replicas tanto biológicas como técnicas. Los datos de transcrito resultantes se analizaron y mostraron pocos transcritos sobre o subregulados usando un límite superior/inferior de dos veces relativamente severo con una relación de correlación positiva.

El chip génico Affymetrix se basa en las ESTs de soya por Shoemaker *et al.*, 2002, A compilation of soybean ESTs: generation and analysis, *Genome*, 45(2):329-38, y una gran fracción de estas ESTs no son anotadas más allá de ser genes expresados. Notablemente, entre los transcritos que no mostraron ninguna variación significativa, estaban aquellos de las proteínas que sí demostraron abundancia de proteínas incrementada en el proteoma de la semilla SP-, en particular KTI, P34 y LE.

El transcrito que codifica para la proteína principal asociada con cuerpos oleos, oleosina, estuvo entre aquellos que no han diferido en la semilla SP-. Una mayoría del proteoma de la semilla de soya es remodelado con poca consecuencia paralela en el transcrito. La estrecha similitud de los datos de disposición de los transcritos de SP- y WT se ilustra en la gráfica de dispersión mostrada en la figura 5.

Ejemplo 9

GFP-KDEL conducida por promotor de glicinina

Para demostrar cómo se puede usar el método para producir altos niveles de proteína en una semilla, se hizo una construcción que tenía el ORF del polinucleótido heterólogo que codifica para GFP, junto con un péptido de señal ER y señal de retención ER ejemplar "KDEL" conducida por el promotor de glicinina y una 3'TED derivada de glicinina (véase figura 7). La transformación de embriones somáticos de soya (*Glycine max* WT variedad "Jack") por biolística se llevó a cabo como se describe en Trick *et al.* 1997, "Recent advances in soyebean transformation", *Plant Tissue Culture and Biology*, 3(1):9-26), y la regeneración como se describe en Schmidt *et al.*, 2004, "Towards normalization of soybean somatic embryo maturation", *Plant Cell Reports* 24:383-391. The hygromicine resistance gene under the control of the potato ubiquitin 3 regulatory elements (Garbarino y Belknap, 1994, "Isolation of a ubiquitin-ribosomal protein gene (ubi3) de papa y la expresión de su promotor en plantas transgénicas, "*Plant Mol. Biol.* 24(1):119-127), se usó como un marcador seleccionable en cultivo de tejidos. Un marco de lectura abierto (SEQ ID NO: 34) de GFP disponible comercialmente (Clontech Inc., Mountain View, CA), menos los codones de inicio y codones de paro, se pusieron en un casete que contenía los elementos reguladores de glicinina específicos de semilla (Nielsen *et al.*, 1989, "Characterization of the glycinin gene family in soybean", *Plant Cell*,

65

1(3):313-328), una secuencia de señal ER de 21 aminoácidos del gen de quitinasa Arabidopsis (SEQ ID NO: 12), y un marcador de retención KDEL (SEQ ID NO: 23) como el descrito en Moravec *et al.*, 2007, "Production of Escherichia coli heat labile toxin (LT) B subunit in soybean seed and analysis of its immunogenicity as an oral vaccine", *Vaccine*, 25(9):1647-57. La construcción se muestra en la figura 7.

Semillas secas maduras fueron cosechadas y observadas visualmente para la expresión de GFP-kdel bajo un microscopio de disección por fluorescencia usando luz azul (450nm) para excitación. Las plantas positivas a GFP-kdel se cultivaron hasta la generación T1 para obtener semillas homocigóticas. Las semillas de GFP-kdel se examinaron a nivel celular usando un microscopio de excitación de dos fotones Zeiss LSM 510 con excitación a 488 nm usando un filtro de emisión de 512 nm.

La construcción se introdujo en soya mediante transformación biolística seguida por selección y regeneración de las plantas. Las semillas positivas a GFP-kdel fueron recultivadas para las generaciones T1 y T2 produciendo una línea homocigótica de semillas que expresaban GFP-KDEL específicas de semilla.

La expresión de GFP-KDEL en las semillas homocigóticas progenitoras se traduce en la acumulación de 1.6% de GFP-kdel en la soya ensayada mediante la comparación del volumen de puntos a partir de 2D IEF/SDS-PAGE y mediante el análisis de los lisados de semilla usando un ensayo de fluorómetro con un control de curva estándar usando GFP obtenida comercialmente. Las imágenes de microscopia de luz fluorescente mostraron que PBs de GFP-kdel se distribuyen a lo largo del citoplasma. Los ensayos de TEM-inmuno-oro de las PBs de GFP-kdel usando un MAb comercial específico para estructuras tipo PB unidas a ER de 0.2-0.3 um marcadas con GFP como se observó previamente.

Ejemplo 10

GFP-kdel conducida por el promotor de glicinina en la línea SP-

El impacto de la expresión de una proteína extraña se probó más en el contexto del proceso de reequilibrio de proteína que ocurre en la SP-. La construcción que contiene el GFP-kdel conducido por el promotor de glicinina se introdujo en soya mediante transformación biolística seguida por selección y regeneración de las plantas, como se

describe en el ejemplo 9. Las semillas positivas para GFP-kdel fueron recultivadas para las generaciones T1 y T2 produciendo una línea homocigótica de semillas que expresaban GFP-kdel específicas de semilla.

La expresión de GFP-kdel en las semillas homocigóticas progenitoras se traduce en la acumulación de 1.6-2% de GFP-kdel en la soya ensayada mediante comparación de volumen de punto a partir de 2D IEF/SDS-PAGE y por ensayos de lisado de semillas usando un ensayo de fluorómetro con un control de curva estándar usando GFP obtenida comercialmente. Estos datos se muestran en comparación con datos producidos en las plantas SP- introducidas se muestran en las figuras 13A a 13C y figura 14.

La microscopia de luz de fluorescencia muestra que las PBs de GFP-kdel son distribuidas a lo largo de citoplasma. Ensayos de TEM-inmuno-oro de las PBs de GFP-kdel usando un MAb comercial específico para estructuras tipo PB unidas a ER de 0.2-0.3 μm de diámetro marcadas con GFP como se describió previamente.

Ejemplo 11

Producción de líneas de soya βCS

Para demostrar más el método para reducir el nivel de una proteína de almacenamiento de semilla, la línea de soya transgénica (" βCS "), deficiente en subunidad α/α de β -conglucina, se preparó mediante dos tipos diferentes de supresión genética. En un método, la planta fue transformada con una construcción de "sentido" que tenía un promotor de β -conglucina que conducía la expresión del gen FAD2. Esto dio como resultado la supresión de la β -conglucina.

En otro ejemplo, una línea βCS de soya se hizo al transformar una planta de soya con una construcción de ARNi diseñada para suprimir β -conglucina. Para preparar la construcción de ARNi, un fragmento de secuencia del gen de β -conglucina (genbankAB030495) (SEQ ID NO: 53) se puso adyacente a un fragmento de 128 pb del gen FAD2 (SEQ ID NO: 57). La región de 256 pb completa se puso después en repeticiones invertidas alrededor de un intrón que se clonó usando un vector pKannibal siguiendo el método descrito en Wesley *et al.*, (2001) "Construct design for efficient, effective and high-throughput gene silencing in plantas", *Plant J.* 27, 581-590. La secuencia de casete completo se muestra en SEQ ID NO: 54.

La transformación usando estas secuencias se encontró que suprimía los niveles de β -conglucina en semillas de soya, dando como resultado el silenciamiento completo de la α/α' β -conglucina. Una fracción de la producción incrementada de glicina se retuvo en forma de su precursor, proglicina, y fue secuestrada en PBs.

5

Ejemplo 12

GFP-kdel conducida por el promotor de glicina en una línea de soya β CS

La expresión de una proteína extraña como un producto génico extrínseco debe ser relativamente independiente del proceso intrínseco de reequilibrio de contenido de proteínas. La proteína extraña seleccionada, una GFP modificada para incluir una secuencia de retención KDEL ER, se diseña para acumularse en los cuerpos de proteína derivados de ER formadores de ER que son organelos inertes, creados de nuevo y normalmente no encontrados en soya. Debido a que los cuerpos GFP-kdel son acumulados establemente a través de la maduración de la semilla (Schmidt y Herman 2008), la GFP-kdel puede ser cuantificada para medir su acumulación a través de la maduración de semilla.

15

La línea GFP-kdel fue introducida en una línea de soya transgénica (" β CS") deficiente en la subunidad α/α' de β -conglucina como resultado de supresión genética. Las cruas resultantes fueron tamizadas visualmente como semillas secas maduras analizadas para la expresión de GFP-kdel.

20

Las figuras 8A a 8D muestran las imágenes de luz blanca (Figura 8A) y luz azul (Figura 8B) de la expresión de GFP-kdel en soya. Los granos de soya fueron fragmentos e hidratados y usados para formar imágenes de la distribución subcelular de GFP-kdel usando un microscopio Zeiss LSM 510 de excitación de dos fotones con excitación a 488 nm usando un filtro de emisión de 512 nm. Las figuras 8A a 8D (panel inferior) muestra GFP-kdel en el fondo genético WT (Figura 8C) y en el fondo de β CS (Figura 8D).

25

La porción restante del chip hidratado se usó para producir geles IEF/SDS-PAGE 2D de amplio alcance. En referencia a las figuras 9A a 9D, lisado de proteínas de la semilla β CS se muestra en la Figura 9A, lisado de proteínas de semilla de GFP-kdel expresada en un fondo WT se muestra en el panel B, lisado de proteínas de β CS x GFP-kdel se muestra en la Figura 9C y un inmunoblot de gel de lisado replicado (Figura 9D) se

30

usó para identificar qué puntos eran GFP-kdel (en cuadro).

GFP-kdel equivale a aproximadamente 2% del proteoma de semilla en el antecedente WT y cuando se cruza en el antecedente β CS, se incrementa a >7% del proteoma de la semilla.

5 Como se muestra en las figuras 10A a 10C, (Figura 10A), porciones de los fragmentos hidratados se visualizaron mediante microscopia de fluorescencia y las imágenes resultantes mostraron que la GFP-kdel es expresado a un nivel más alto en un fondo de β CS en relación a un fondo de WT.

10 Los lisados de los fragmentos de semilla se fraccionaron después mediante 1D PAGE. La figura 10B muestra geles 1D de proteínas de plantas homocigóticas transformadas expresando GFP-kdel en un fondo de WT en comparación con un fondo de β CS. La figura 10C es un inmunoblot que usa un anticuerpo contra β -conglucina, confirmando la falta de proteína β -conglucina en los lisados de semillas de β CS y β CS x GFP-kdel.

15 Para obtener una evaluación adicional de la abundancia de GFP-kdel, lisados de semilla se prepararon y ensayaron usando un fluorómetro con GFP comercial como un parámetro de control. Los resultados mostrados en la figura 11 confirman tanto la impresión visual de la fluorescencia de GFP-kdel (figura 10A) como la abundancia de volumen de puntos (figuras 9A a 9D) que la construcción de GFP-kdel introducida en una
20 planta β CS se traduce en aproximadamente una mejora de 3.5 a 4.0 veces de la acumulación de GFP-kdel en comparación con una planta WT transformada con GFP-kdel.

25 La co-producción y acumulación de proglicina y GFP-kdel que se acumulan para formar cuerpos de proteína derivados de ER se traduce en la formación de dos poblaciones distintas de cuerpos de proteína unidos a ER. Cuerpos de proteínas derivados de ER formados de nuevo y el secuestro de ER de productos transgénicos incrementan la estabilidad de proteínas de otra manera inestables después de la traducción. La subunidad α/α de la supresión de β -conglucina de granos de soya produce cuerpos de proteína de proglicina y la introducción de la línea GFP-kdel
30 produce también cuerpos de proteínas que se traducen en la formación de cuerpos de

proteína GFP-kdel que son tanto más abundantes en número como en tamaño en comparación con los cuerpos de proteínas en el progenitor GFP-kdel. Proteínas estrechamente emparentadas tales como las α y γ -zeínas acumuladas, formando una sola población de cuerpos de proteínas. Esto en contraste con los resultados de proglicinina y GFP-kdel, proteínas que cuando son expresadas formaron dos cuerpos de proteína distintos con la GFP-kdel siendo favorecida para agregación en cuerpos de proteína. Esto indica que la agregación de proteínas en los cuerpos de proteínas que forman ER es un proceso dependiente de especie de proteína que forma cuerpos de proteína sólo con un tipo de proteína incluso si más de una proteína secuestrada de cuerpo de proteína es sintetizada. Esto es potencialmente adecuado para usar la formación de cuerpos de proteínas como una plataforma para biotecnología toda vez que aseguran que la proteína secuestrada en las agregaciones sea relativamente pura. Los cuerpos de proteína pueden ser después aislados directamente de lisados, simplificando ampliamente la trayectoria de purificación hacia abajo.

Ejemplo 13

Procesamiento de proteínas heterólogas en la línea de soya β CS

El procesamiento del polinucleótido heterólogo que codifica para glicinina se examinó en plantas de soya WT y β CS.

Como se muestra en la figura 12 (línea 1), en soya WT (no transgénica), ninguna cantidad de pro-glicinina pudo ser detectada en semillas. La línea 2 muestra que en la planta β CS, niveles elevados de glicinina son almacenados en cuerpos de proteínas como pro-glicinina. El porcentaje de pro-glicinina/glicinina es de aproximadamente ~24%.

Ejemplo 14

Producción de β -glucosidasa en semillas de soya

Las construcciones descritas en la presente pueden usarse para producir enzimas tales como β -glucosidasa en semillas de soya. Por ejemplo, la secuencia de β -glucosidasa *Aspergillus kawachii* (SEQ ID NO: 35) puede insertarse en una construcción genética que tenga las secuencias reguladoras de glicinina de soya, junto con una señal ER y secuencias de retención. La construcción puede ser después transformada en tejido de soya usando bombardeo biolístico, plantas son regeneradas y cruzadas y se obtienen

plantas homocigóticas estables. Estas plantas son después introducidas en las líneas de soya β CS. La producción de proteína de interés a partir de esas plantas es luego determinada. Las plantas con un alto nivel de expresión de β -glucosidasa se seleccionan para la producción de escala ascendente de la proteína de interés.

5

Ejemplo 15

Producción de exocelobiohidrolasa I en semillas de soya

Las construcciones descritas en la presente pueden usarse para producir enzimas tales como exocelobiohidrolasa I en semillas de soya. Por ejemplo, la exocelobiohidrolasa I de secuencia de *Trichoderma reesei* modificada (No. de registro P26294), puede modificarse para incluir la secuencia de señal ER de quitinasa de *Arabisopsis* y una secuencia de señal de retención de KDEL, como se muestra en SEQ ID NO: 22. La región reguladora de genes de KTI (o la región reguladora de genes de otra proteína que se encuentra que compensa la proteína reducida en la línea de soya que tiene la deficiencia genética) puede usarse para conducir la expresión de la proteína. La construcción puede ser después transformada en tejido de soya usando bombardeo biolístico, las plantas son después regeneradas a la población homocigótica para la expresión de la enzima. Las plantas homocigóticas que expresan enzima pueden ser después introducidas con plantas SP- homocigóticas para formar plantas que sean homocigóticas para la enzima exocelobiohidrolasa I y para el rasgo SP-. La producción de exocelobiohidrolasa I a partir de estas plantas es luego determinada. Plantas con un alto nivel de expresión de exocelobiohidrolasa I se seleccionan de preferencia para la producción de escala ascendente de la proteína.

20

Ejemplo 16

Procesamiento de la proteína de interés a partir de semillas

25

30

La proteína puede utilizarse directamente, no purificada de la harina de semilla de soya, o la proteína de interés se puede purificar parcial o sustancialmente. El método exacto dependerá del tipo de proteína que se produzca, así como de la pureza que se requerirá. Como un ejemplo, una proteína de interés transgénica puede extraerse de granos de soya secos y maduros al mezclar la molienda de semilla de soya con NaCl 0.35 M en PBS a 75 g/L a temperatura ambiente durante 2.5 horas. El extracto puede ser

después pasado a través de varias capas de estopilla, y centrifugarse a 12,000 g durante 1 hora a 4°C. El sobrenadante es luego recuperado y la concentración de NaCl se ajusta a 0.4 M (pH 8.0). Después de una segunda centrifugación a 12,000 g durante 10 minutos a 4°C, el sobrenadante puede ser recogido y filtrado a través de una membrana de nitrocelulosa de 0.45 µm. El filtrado puede ser cargado en una columna SP SEPHAROSE™ (Bio-Rad, Hercules, Calif.) la cual sea previamente equilibrada con NaCl 0.4 M en 50 mM de fosfato de sodio, pH 8.0 (regulador de pH de unión) a una velocidad de flujo de 5 ml/min. La columna puede ser lavada con el regulador de pH de unión hasta que ya no eluya más contaminantes. La proteína de interés es luego eluida mediante un gradiente lineal, seguido por la diálisis contra PBS. La proteína purificada puede ser después analizada, por ejemplo, mediante SDS-PAGE y/o inmunoblot, y almacenada a -80°C, o, alternativamente, liofilizada y almacenada a temperatura ambiente o menos, hasta que se requiera.

Ejemplo 17

Producción de otras enzimas en semillas de soya

Un gen sintético que codifica para una enzima industrial de interés seleccionada puede insertarse en el casete de expresión génica específico de semilla de soya descrito en la presente que contiene los elementos reguladores 5' y 3' ya sea de glicinina, KTI, P34, SBP, SMP o LE. Los elementos reguladores de KTI, P34, SBP, SMP o LE se usarán para conducir la expresión de las enzimas industriales, tales como β-glucosidasa, en un fondo SP-. La construcción induce a la enzima industrial a precipitarse en el proceso de reequilibrio de proteína que resulta de la supresión de conglucina y/o glicinina incrementando la síntesis y acumulación de la enzima industrial.

El ORF tiene de preferencia una secuencia de nucleótidos que codifica para la señal de dirección a ER del gen básico de quitinasa de Arabidopsis fusionado 5' y una secuencia de nucleótidos que codifica para una secuencia de retención KDEL ER carboxi-terminal fusionada 3' al gen. El plásmido puede también contener un marcador de resistencia a higromicina para la selección de transformantes. La transformación y producción de líneas homocigóticas que contienen el gen de interés se producirá como se describe en la presente. Las plantas transgénicas son después introducidas en su SP-,

βCS u otra línea deficiente en proteínas de almacenamiento de semilla. La cantidad de la proteína transgénica y la semilla será después determinada. Las plantas que tengan un alto nivel de expresión de la proteína industrial de interés se usarán para la producción a escala ascendente de la proteína. Mediante el uso de este método, se puede obtener un

5 rendimiento incrementado de la proteína de interés industrial.

Ejemplo 18

Producción de un fragmento de anticuerpo humano en semillas de soya

Los métodos descritos en la presente pueden usarse para producir fragmentos de anticuerpo en semillas de soya. Por ejemplo, un fragmento de anticuerpo que se producirá se selecciona. El ácido nucleico que codifica para el fragmento de anticuerpo es luego insertado en una construcción de vector que tiene elementos reguladores hacia el extremo 5' y hacia el extremo 3' de glicinina, como se describió en la presente. La construcción es luego transformada en semillas de soya usando electroporación. Las

10 plantas transformadas son luego regeneradas y cruzadas, y las plantas homocigóticas estables son obtenidas. Estas plantas son luego introducidas en una línea βCS. La producción del fragmento de anticuerpo se confirma. Las plantas se cultivan para producción a escala ascendente de la proteína de interés. Los fragmentos de anticuerpo son luego aislados y purificados de las semillas de soya maduras.

20 A partir de lo anterior, se apreciará que, aunque modalidades específicas de la invención han sido descritas en la presente por motivos de ilustración, pueden hacerse varias modificaciones sin desviarse del espíritu y alcance de la invención. Todas las referencias citadas en la presente se incorporan a manera de referencia en sus totalidades.

Tabla 4
Secuencias de ácido nucleico y proteínas

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
Glicinina	Glycine max	región promotora/ reguladora 5'	caaaacaaattaataaaacacttacaacaccggattttt ttaaftaaaatgtgccatttaggataaatagtttaatttt aataattatttaaaaagccgtatctactaaaatgattttat ttggttgaaaatattataatgtttaaatcaacacaatctat caaaattaaactaaaaaaaataaagtgtacgtggttaa cattagtacagtaataataaggagaaaagagaaattaa gaaattgaaagcgagtcataattttaattatgaacctgc atataaaaaggaaagaaagaatccaggagaaaag aaatgaaacctgcattggtccctcgatcacgagttt ctgccatttgaatagaacacactgaacacacctttcttt gtcacttaattgagatgccgaagccacctcacacctg aacttcattgaggttagcacccaaggcttccatagcca tgcatactgaagaatgtctcaagctcagcacctacttc tgtgacgtgtccctcattcaccttctctctccataaa taaccacgcctcaggttctcgcttcacaaactcaaacat tctctccattggctcctaaacactcatcagtcacccgc	SEQ ID NO: 1
Glicinina (gen Gyl de soya para sub-unidad de glicina G1; número de registro X15121.1)	Glycine max	región promotora/ reguladora 5'	caaaacaaattaataaaacacttacaacaccggattttt ttaaftaaaatgtgccatttaggataaatagtttaattttt aataattatttaaaaagccgtatctactaaaatgattttat ttggttgaaaatattataatgtttaaatcaacacaatctat caaaattaaactaaaaaaaataaagtgtacgtggttaa cattagtacagtaataataaggagaaaatgagaaattaa gaaattgaaagcgagtcataattttaattatgaacctgc atataaaaaggaaagaaagaatccaggagaaaag aaatgaaacctgcattggtccctcgatcacgagttt ctgccatttgaatagaacacactgaacacacctttcttt gtcacttaattgagatgccgaagccacctcacacctg aacttcattgaggttagcacccaaggcttccatagcca tgcatactgaagaatgtctcaagctcagcacctacttc tgtgacgtgtccctcattcaccttctctctccataaa ataaccacgcctcaggttctcgcttcacaaactcaaac atttctctccattggctcctaaacactcatcagtcacac c	SEQ ID NO: 2
Conglicinina	Glycine max	región promotora/ reguladora 5'	gttttcaatttgaattttaatgtgtgtgtaagtalaattt aaaataaaaataaaaacaattattatatacaaatggcaa aaacatttaaacgtattattttaaaaaataatgataaa tatatttataatttaataatctattctatgtatttttaaaaatct attatataatgatcaactaaaatttttatactacattatt ttgcattttatcaattttcttgcgtttttggcataatttaata tgactattctttaataatcaatcattattcttacatggtacat attgttgaaccalatgaaggttcattgcattgactatg tggatagtgtttgatccatgcccttcatttgcgctatta attaatttggtaacagatttcttcaatcagttacttaatcc ttcctcatcataaataatctgtagttcgaatgccataata ttgattagttttggaccataagaaaaagccaagggaac aaaagaagacaaaacacaatgagagtatcctttgcata gcaatgtctaagttcataaaattcaaacaaaaagcgaat cacacacagtggaatcattatccactagctgatcag gatcgccgcgtcaagaaaaaaactggaccccaaa agccatgcacaacaacacgtactcacaagggcgicaa tcgagcagcccaaaacattcaccaactcaacccatcat gagccacacatttgtgttttaacccaacctcaaacl cgtaattctctccgccaccttatttgttttcaaacacc cgtaaacctgcattccaccccggtggccaaatgttcattg catgttaacaagacctatgactataaatatcgaatctc ggcccaagtttcatcatcaagaaccagttcaatatcct agtacgcgtaattaaagaatttaagatatatt	SEQ ID NO: 3

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
KTI	Glycine max	región promo- tora/reguladora 5'	attgatactgataaaaaaatatcatgtcttctggactg atgatgcagtatacttttgacattgcccttattttttca gaaaagcttcttagttctgggtcttcattattgttccc atctccattgtgaattgaatcatttgcttcgtgcacaaat acatttagntaggtagcatgcattggcagattcacgggtt attatgcatgacttaagttcatggtagtacctgcc acgcatgcattatattggtagattgataggcaatttg gtgtcaacaataataataataatgttttatattatga aataacagtgatcaaaacaaacagttttatctttattaac aagatttgttttggatgacgtttttaagttacgctt cccccttcttgaatttagaacctttatcatcataaaatc aaatactaaaaaattacattttcataaataaataacaca aatattttaaaaaatctgaataataatgaacaattatc atattatcacgaaaaatcatttaataaataattatataat aaaatgtaatagtattatattaggttttttactgcac gcataatatacaaaaaagattaaatgaactattataaa taataacactaaattaatggatcatatcaaaataatg aaaaagtaataaaattgttaattacttctatatgtatta cacacacaaataataataatagtaaaaaaattatgat aaataattaccatctcataaagatttaaaataatgataa aaatatagattttttatgcaactagctagccaaaaag agaacacgggtatatataaaagagtagctttaaatct actgtacttcttatttctgacgtttttatatcaaggtag atagctgaagattttaattatcagcttaaatattcattagc acttaatacttttctgttttcttctatcctataagtagtccc gatttcccaacattgcttattcacacaactaactaagaa agtttccatagcccccaaaa	SEQ ID NO: 4
LE	Glycine max	región promo- tora/reguladora 5'	aatgccatcgtatcgtgtcacaaatggaatcacgcaatg aacaatgctatcccttgagaaaagtgaaatgcagca gcagcagcagactagagtgctacaaatgcttactctct gagaaaaagtgaatgcagcggcagcagacctgagtg ctatatacaattagacacagggctatttaattgaaattgt cttattattaaattttctgttttatatttttttaaaatttaatt aaatttatattattatttaagacagatatattttttgtg attataaatgtgtcactttttcttttagtccatgtattcttcta tttttcaatttaactttttatttttttttaagtcactctgac aagaaaaattgttgacataaaactattaacataaaatta tgtaacatgtgataacatcatattttactaataaacgtc gcattttaacgttttttaacaaatcgcactgtaagagta aaaatgaaatgttgaagggttaattgcactactaactat tttttctataagtaacttttttgggataaattgtataica ttagatagcatatttaataatgggtaccttttcacaaaac ctaccctgttagtcaaacacacataagagaggatgg atttaaacagtcagccacgttaagtatatagtgaaagaa ggctgataacacactctattatgttagtacgtactgtatt ccrttttgttagttttgaatttaatttaaaatataatg ctaacaacattaaattttaatttacgtctaattatattgt gatgtataaaaaattgtcaaccttttaaaattataaaaga aataattatttgataacaacttttgaaaagtacccaata atgctagtataaataggggcatgactccccatgcatca cagtcgaatttagctgaagcaagc	SEQ ID NO: 5
P34	Glycine max	región promo- tora/reguladora 5'	atggataaaaaatctagcatttciccttttctactagcat attaaattaacgatccagaaatatttataaattttttta gttaatgactcaatacacggccagtcagggtcaacct tggctggataaacaacccicataacatgacatacaatta taattgaaaattctcatatagcacaattatgaaggcaaa aacatggcacacaaaagggtacttgttttaactataaacg attatgaattttcagggaaaatagcatgttcgtctgati ctctgaagaacttttttaggtacttttactaagttggacg tgattttgtctatgttcaggagaaaaataaggataaaatg	SEQ ID NO: 6

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			tgattttaaaaaaatatgacagatgcaatcttatttcac ttttatactttcactactgcttcataatgacctaaagtcaga gaaatattttaaaaagalaataacgataaagaatacagat gagaaagaaacctcacacaatgaatagaccaaatag acctatttatttcccttagaataaagaaaaataattatttt atttttcacattacatttatattttctatcacttctcattta ggtaattgattggcataatgagtgatgaatttttttttaa aaaaagcgtaaatattaattatattcatgcattattgttt ctgtcttcattttctttaatcttacgttaataaattatt attaaattttatagttgatgaataatataagagatataa ataaaaaaataaattattataataaaaaataaaaaata attattttgagalaaaaaatttaagagaacaattataaac ggagagattattatatttagtttatgtaccgtgtacgtgtc actaacatgggtgtctctccatcatttctaggaaaaaac attataggagatgaaaaagcaaaagtgtgtctgttta tggtttgtatataccagctcacttggcagcaattacc cgtcttgctgtacttacgagacacgtacattaacatt gtcctagctagtgcatgcaattgcccccattcatcac tcctccctttctctctttatattatataatataatata atatatatatatatatatatatatataaacaagcac aatgcatcatctcaagaaattaagagagttttttgttc ctcactgaccaagcc	
SMP	Glycine max	región promo- tora/reguladora 5'	gtttcacatgatcttcattctgtgtgcttaggagactc aacttcagagtcctgtgatgatcaatgacicttaagtgtg acttatggcttcattttaataaattactcalagacaatg atgcccttcattatccatcccaatttaactaatgttaa gattttctttacaaaaactagaaaataaatttttaagag aattaatatttagttgggggalaatttaattatagaatat tctgttttctctcctaatttcaataaataattatagtaaattt taaaacaacattatgtatatatatgtgaaattagaata aataatttgaataattcaatatcaaccaatttaatttttt aaagggttaaaattatagtcatttggacatagacagta atatggagctagagttatgttaaaacacaggtcagaaaa ataaaactcatgaaaattgtlaaccaacgaacattaca ataagttcagcagtgatctttgtctcattttttacatttct tagtcctttatttttggtaaacgttgccttttattattat gttttagtaatttttttactttattacattaaattttaaatt ctattttgttttttattgttttgaattttattttgttaatttt ttcataagaacactaatacttagaaaaagataaaaaaa aaaaaggaaagttcgalggaattcaagctcatgaaatt tatgttaaaacatgagtaattcattatattttaaaattta aaataaaaaaaaaggaaagttcgtggaattcaag ctcatgaaattttatgttaaaacatgagtaattcattaat attttaaaatttaaaaataaattgttttacttatattaata gttgggtgaaaatttaaaaaagatgcacgtttaaacaaag gttggaaatcgttttattttaatttcaccacitcagtgga tgcacatttagacaaggatggaatcgttgacttggatt ggttactccttccccaacacgtgccacctctagg aaggtaaggaccgagtgaccgataacaacaccgat ttccaaatatatatctcattcctaagctcacacacatttt acgttacatttcattatagatgttcaatcgttaagaca catgtcaccacaaaagagcatctcataacaacgtgtc acacctcccaagcacacgtgtcactcacaacaacacc ctcacctatatataaattatcaaaccctccttccattctc cacatctcaatcicaatatctacacaaaaggtgttccatt gagtgaaaagtagtgtgttaagaactaaacaattttca	SEQ ID NO: 8
		secuencia de señal ER	mkimmmiklcffsmliciapada	SEQ ID NO: 9
		secuencia de señal ER	maashgnaifvllctflpslac	SEQ ID NO: 10

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
		Secuencia de señal ER	maarigifsvfvavllsisafssa	SEQ ID NO: 11
	Arabidopsis thaliana	Secuencia de señal ER de quitinasa básica de A. thaliana	mktnlflifslslssae	SEQ ID NO: 12
Glicinina	Glycine max	Terminador (3'TED)	agcccttttgatgtgctacccacattgtcttttggca atagtgctagcaaccaataaataataataataatga ataagaaaacaaaggcttagcttgccctttgtcactgt aaaataataalgaagtactctataatgagtcacgaa acttttgcgggaataaaaggagaaattccaatgagttt ctgtcaaatctctttgtctctctctctctctctttttt ctttctctgagcttctgcaaaacaaaggcaacaat aacgattggccaatgatagttagctgacgataatc tttaggaagtgttggcaggacaggacatgatgagaa gactaaaattgaagtattgcagaccaatagttgaag attaaactttaagaatgaagacgtcttatcaggttctcatg acttggga	SEQ ID NO: 13
Glicinina (Gen Gy1 de soya para subunidad G1 de glicinina; número de registro X15121.1)	Glycine max	Terminador (3'TED)	agcccttttgatgtgctacccacattgtcttttggca atagtgctagcaaccaataaataataataataatga ataagaaaacaaaggcttagcttgccctttgtcactgt aaaataataalgaagtactctataatgagtcacgaa acttttgcgggaataaaaggagaaattccaatgagttt ctgtcaaatctctttgtctctctctctctctctttttt ttctctgagcttctgcaaaacaaaggcaacaataa cgattggccaatgatagttagctgacgataatcttt agggaagtgttggcaggacaggacatgatgagaa ctaaaattgaagtattgcagaccaatagttgaagatt aacctttaagaatgaagacgtcttatcaggttctcatgac ttggga	SEQ ID NO: 14
Beta-conglicinina	Glycine max	Terminador (3'TED)	aataagtatgtagtactaaaatgatgctgtaatagctca tagtgagcggagaaagtatcgggclatttaactatgact tgagctccatctatgaataataaatcagcataatgatgct ttgttttggtagac	SEQ ID NO: 15
proteína de almacenamiento de beta-congli- cinina (gen alfa'- bosp); número de registro M13759.1	Glycine max	Terminador (3'TED)	ataagtatgtagtactaaaatgatgctgtaatagctcat agtgagcggagaaagtatcgggclatttaactatgact gagctccatctatgaataataaatcagcataatgatgct ttgttttggtagac	SEQ ID NO: 16
KTI	Glycine max	Terminador (3'TED)	gacacaagtgtgagagtactaaaatgctttggtgt acgaaatcattacactaaaataaatcaagcttatat atgccttclaaaggccgaatgcaaagaaattggtcttct cgttatctttgccactttactagtacgtattaattactactt aatcaacttgttacggctcattataacc	SEQ ID NO: 17
LE	Glycine max	Terminador (3'TED)	atgtgacagatcgaaggagaaagtgttaataagcga ctctcactactcgatcgtagtgattgtcattgttatatai aataatgttactttcacacttatcgtatgcatgtgaa actataacacatttaactctacttgcataatgataacactc tccccatttaaaactctgtcaatttaagataataagattc tttaaatgattaaaaaaatattataaattcaatcactc ctactaataaattattaattattattgattaaaaaaat acttatactaatttagtctgaatagaataatgattctag	SEQ ID NO: 18
P34	Glycine max	Terminador (3'TED)	gccgtaagggttcaataacacgagtgcttgttttctnag ggacaagcattgtacttatgtatgattctgtgaaccatg agtctccacgtgtactaatgtgaagggcaaaaataaa acacagaacaagtcgtttttctcaataatgtgaagggt	SEQ ID NO: 19

5

10

15

20

25

30

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			agaaaaatggaaccatgccctcctctcttgcattgattta aaataattaggcagatggt	
5 GBP	Glycine max	Terminador (3' TED)	gaggtttagaacaatcaagaaaagggtgcatgtggct gaagatcacggggaatgtattaagcttcagagactctt aaattaaatttctgtatttggttataigttagtcttta aattagccagatggagttatgtgtactaaatgcaggg atgctaattggaataaaatggccacttgattgttagctat ctcttaggtagcagaataagacgtaaacgtgttcttgc tccaa	SEQ ID NO: 20
10 SMP	Glycine max	Terminador (3' TED)	ttaaacgtgatctatgatacaacaataattagtatatata gacgcatgcagtttatagtatatattgtcatgttgtag ttttacatttggtttgcgtttlacattcttcaaaaaaaa aaaaatgtgtagtagctgtgaaggtttgaagattgttct aggctccgtgggaaccatttcaacaataaacatttgcg cgttctgtacacgtatgatgagaagagatgccttatg ggcagtatcatctaaaacttatttcatccatcatagaatt tggatctattggactggactgaactgaactgaatgatcc tttttcttttaatttcttactaacaataacataaaaaa ccagataattaaacttagccagtatgaatttaactatttgc taatgctatgacttatcactgtctgtatcatcttaattcttt ttcatattttatataataa	SEQ ID NO: 21
15 marcador de retención ER (codifica KDEL, seguido por 2 codones de paro	Sintético	Secuencia de señal de retención ER-CDS	aaggatgaactttaatga	SEQ ID NO: 22
Marcador de retención ER KDEL	Sintético	Retención ER-secuencia de aminoácidos (codificada por secuencia de ácido nucleico anterior)	kdel	SEQ ID NO: 23
20 marcador de retención ER (codifica KDEL, seguido por 2 codones de paro	Sintético	retención ER	aagcatgatgaactttaatga	SEQ ID NO: 24
Marcador de retención ER KDEL	Sintético	Retención ER-secuencia de aminoácidos (codificada por secuencia de ácido nucleico anterior)	khdel	SEQ ID NO: 25
Marcador de retención ER	Sintético	retención ER	hdel	SEQ ID NO: 26
Marcador de retención ER	Sintético	retención ER	kcei	SEQ ID NO: 27
Marcador de retención ER	Sintético	retención ER	sekdel	SEQ ID NO: 28
Marcador de retención ER	Sintético	retención ER	sehdel	SEQ ID NO: 29
25 Ubiquitina 3	Solanum tuberosum	región pro-motora/ reguladora 5'	ccaaagcacatacttatcgatttaatttcatcgaagag attaatatcgaataatcatatacatactttaaacataaac aaattttaaacataatcttggtatataaatttttaaa gtcatgaagtatgtaicaaatcacatatgaaaaaatt aactattcataatttaaaaaatagaaaagatacatctagt gaaatttaggtgcatgtatcaaatacattaggaaaaggg	SEQ ID NO: 30

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
Fluorescente Verde (GFP)	secuencia - (derivada originalmente de Aequorea Victoria		ccaatctgtgtaattagatggtgatgtaatgggcaca aatcttctgtcagtgaggagggtgaagtgatgcaaca tacggaaaacttacccttaatttattgactactggaa aactacctgttccatggccaacactgtcactacttctct tatgggtgtcaatgcttttcaagataccagatcatatga agcggcacgacttctcaagagcgccatgctgaggg atacgtgcaggagaggaccatcttccaaggacgac gggaactacaagacacgtgctgaagtcagtttgagg gagacacccctgtcaacaggatcagccttaagggaat cgatttcaaggaggacggaacatcctcggccacaa gttggatacaactacaactcccacaacgtatataca cggcagacaaaagaatggaatcaagcctaact caaaattagacacaacattgaagatggaagcgttaac tagcagaccattatcaacaaaactccaattggcga ggccctgtcctttaccagacaaccattaccgtccaca caatcggcctttcgaagatcccaacgaaaagagag accacatggtcctcttgagttgtaacagctgctggga ttacacatggcatggatgaactatactga	
Beta glucosidasa	Aspergillus kawachii	(secuencia CDS original - AB003470	atgagggtcacttgattgaggcgggtgctctcactgct gtctcgtcggccagcgtgatgaattggcttactcccc accgtattaccatccccttggccaatggccaggge gactgggcgcaggcataccagcgcgtgtgtatattgt ctcgcagatgacattgctgagaaggtaacttgacca caggaaactggatgggaattggagctatgttggtcag actggcggggtccccgattgggagttccgggaatgt gtttacaggatagccctctgggcgttcgcgactccgac tacaactctgttcttccctccggtatgaacgtggcgtca acctgggacaagaatctggcatacctccgcggcaag gctatgggtcaggaaattagtacaagggtgccgat ccaattgggtccagctccggccctctcgttagaagt cccgcagggtgtcgaactgggagggtcttcccccg acccggccctaaagtgtgtgtctttgcagagaccalc aagggtatccaagatgtgtgtgtgtgcgcagcgcta agcactacattgcctacgagcaagagcatttccgtcag gcgcctgaagcccaagggttatggatttaacattccga gagtggaagcgcgaacctcgcagataagactatgca cgagctgtaccttggcccttcgcggatgccatccgtg cgggtgtcgtgcgtgtgtgtgtctctacaaccagatc aacaacagctatggctgccagaacagctacacttga acaagctgtcaaggccgagctgggttccagggtctt gtcatgagtattggcggtcaccatgctgtgtgtgag tgggtcttggcaggattggatatgtctatgccaggaga cgtcgcactacgacagtgtgtactgtgggtacaa acctgaccgttagcgtgtctaacggaacgggtgcccc atggcgtgtgtatgacatggctgtccgcatcaggccg cctactacaaggctcggccgtgaccgtctgtgactcct cccaacttcagctcatggaccagagatgaatacggct acaagtactactatgtgtcggaggaccgtacgagaa ggtcaaccactacgtgaacgtgcaacgcaaccacag cgaactgatccgccgattggagcggacagcacggt gtctctcaagaacgacggcgctctgcttactgtgta aggagcgcctgtgtcgccttatcggagaagatgcgg gtcctaacccttatgggtccaacggctgcagtaccgt ggatgcgacaatggaacattggcgtatggctgggga agtgggtactgccaacttcccatcctgtgacccccga gcaggccatctcaaacgaggtgtcicaagaacaagaat gggttattcaccgccaccgataactgggtatcgtatca gattgaggcgcttctaagaccgaggtgtctcttctgt ctttgtcaacgccgactctggcggagggttacatcaatgt cgacgggaacctgggtgaccgcaagaacctgaccct	SEQ ID NO: 35

5

10

15

20

25

30

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCION	SECUENCIA	SEQ ID NO:
			gtggaggaaacggcgataatgtgatcaaggctgctgt agcaactgaacaacaccattgttatcattcactctgtc ggcccagtccttggttaacgaatggtacgacaaccca atgttaccgctattctctgggggtggtctgcccggtcagg agtcgtggaactctcttgcgacgtcctctatggccgtg tcaaccccgggtgccaagtcgccctttaccctggggcaa gactcgtgagcctaccaagattacttggtcaccgagc ccaacaacggcaatggagccccccagggaagacttcg tcgaggggcgtcttcattgactaccgaggattcgacaag cgcaacgagacccccgatctacgagttcggctatggtct gagctacaccactttcaactactcgaacctgaggtgc agggtctgagcggcccgctacgagcctgcttcggg tgagactgaggcagcgccaactttggagaggttggga aatgctcgaattaccttaccggacggactgcaga aaatcaccagttcatctacccctggctcaacagtacc gatctcaggcgtctctggggatgctagctacggaca ggactcctcgactatctcccgaggagccaccgat ggctctgcgaaccgatcctgctgctgggctgctc ctggcggaacccctgcctgtacgacgagctcatccg cgtgtcgggtaccatcaagaacaccggcaaggtgct ggtgatgaagttcccaactgtatgttcccttggcggc cccaacgagcccaagatcgtgctgctcaattcgagc gcatcacgctgcagccgtcagaggagacgaagtgga gcacgactctgacgcccgtgaccttgaactggaa tgttgagaagcaggactgggagattacgtctatccca agatggtgtttgtcgggaagctcctcgcggaagccgc gctccggcgctctgctactgttactaa	
Beta glucosidasa	Aspergillus kawachii (No. de registro BAA19913)	(proteína de secuencia original de longitud completa)	mrflieavaltavslasadelaysppyyppswang qgdwaqayqravdivsqmtlaekvnltgtgwelel cvgtgqgvprlgvpgmclqdsplgvrdsdynsafp sgmnvaatwdknlaylrgkamgqefsdkgadiql gpaagplgrspdggrnwefspdpalsgvlfaetik giqdagvvatakhyaieqehfrqapeaagygnfis esgsanlddktmhelylwpfadairagagavmcsy nqinnsygcqnsytlknllkaelgfgfvmsdwaa hhagvsgalagdmsmpgdvdydsgtsywgtnlt vsvlntgtpqwrddmavrimaaykvgdrdlwt ppnfsswtrdeygykyyyvsegpyekvnhyvvn qrmhselirrigadstvlkndgalpltkgerlvaliged agsnpygangcsdrgcdngtlamgwsgtanfpyl vtpeqaisnevlknkgvftatdnwaidqicalakta svslvfvnadsgegyinvdgnlgrknltlwrngdn vikaaasnccntiviihsvgpvlvnewydnpnvtai lwgglpqgesgnsladvlygrvnpgakspftwgktr eayqdylvtepnngngapqedfvegvfdyrgfdk metpiyefgyglsyttfnysnlevqvlsapayepasg eteaaptfgevgnasnyltpdglqkitkfiypwlnt dleassgdasygqdsdyldpegatdgsaqpilpagg ggpgnprlydelirsvtikntgkvagdevpqlyvsl ggpnepkivlrqferitlpseetkwtlttrrdlanw nvekqdwetsypkmvfvgsrrkpplralsptvh	SEQ ID NO: 36
Beta glucosidasa	Aspergillus kawachii (secuencia parcial de No. de registro BAA19913)	Secuencia de arriba con 19aa (Proteína removida de secuencia de señal)	delaysppyyppswangqgdwaqayqravdivs qmtlaekvnltgtgwelelcvgtgqgvprlgvpgm clqdsplgvrdsdynsafpsgmvaatwdknlayl rgkamgqefsdkgadiqlgpaagplgrspdggrm wegfspdpalsgvlfaetikgiqdagvvatakhya ieqehfrqapeaagygnfisesgsanlddktmhely lwpfadairagagavmcsynqinnsygcqnsytl killkaelgfgfvmsdwaaahhagvsgalagdms mpgdvdydsgtsywgtnltvsvlntgtpqwrdd	SEQ ID NO: 37

5

10

15

20

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			catatattcttgatctagataaataacgattttgattatgtat aatctcaaatgaagggttlatctacttcagaaataaca atatacttttatcagaacatccaacaaagcaacaaccaa ctagagtgaaaaatacacattgttcttagacatacaaaa attgagaaaagaatctcaaaatttagagaacaaatct gaatttctagaagaaaaataattatgcactttgctatt gctcgaataaataatgaagaaattagactttttaaaa gatgttagactagataactcaaaagctattaaaggagt aatattcttctacatlaagatttttagttacagtcctgtaat taaagacacatttttagattgtatctaaacttaaatgtatct agaatacatatatttgaatgcatcatatacatgtatccga cacaccaattctcataaaaaacgtaatatctaaactaa tttatccctcaagtcacttaagcccaatatacattttcatc tctaaaggcccaagtgccacaaaatgcaggcccaat tacgaagaaaagggtgttaaaaccctaataaagggtg cactggcagagcttactctcattccatcaacaaga aacccataaaggccgcagccactgatttctctctcc aggcgaag	
Ubiquitina 3	Solanum tuberosum	terminador	ttgattttaatttagcaaaatgtcttatcagtttctctttt gtcgaacggtaatttagagttttttgctatatggatttct gttttgatgtatgtgacaacccctcgggattgtgattatt tcaaaactaagagttttgtctattgttctcgtctatttgg atatcaa	SEQ ID NO: 31
Monómero de ubiquitina/pro- teína ribosómica (Genbank No. de registro Z11669.1)	Solanum tuberosum	terminador	ttttaattgtagcaaaatgtcttatcagtttctcttttgcg aacggtaatttagagttttttgctatatggatttctgttt tgatgtatgtgacaacccctcgggattgtgatttattcaa aacaaagagttttgtctattgttctcgtctatttggatac aa	SEQ ID NO: 32
Gen de resistencia a higromicina (hph)	Escherichia coli	ORF	atgaaaagcctgaactcaccgcgacgtctgtcgaga agtttctgacgaaaagttcgacagcgtctccgacctg atgcagctctcggaggcggaagaatctcgtgttctcag cttcgatgtaggaggcggtggatatgtcctcgggttaa atagctgcgccgaiggtttctacaaagatcgttatgttta tcggcactttgcatcggccgcgtcccgattccggaa gtgcttgacattggggcattcagcgagagcctgacctta ttgcatctcccggtgcacagggtgtcacgttgcaag acctgcctgaaaccgaactgcgcgtgttctgcagcc ggtcgcggaggccatggatcgatcgctcggccga tcttagccagacgagcgggttcggccattcggaccg caaggaaatcgggtcaatacactacatggcgtgatttcata tgccgcatgtcgtatcccatgtgtatcactggcaact gtgatggacgacacgtcagtgctgtcgtcgcgcag gctctcgtatgagctgaltgtttggccgaggactgccc cgaagtcggcaccctcgtgcacgggatttcggctcc aacaaatgtcctgacggacaatggccgcataacagcg gtcattgactggagcgaggcgatgttcggggattccc aatacagggtcgccaactctctctcggaggccgtgg ttggcttgtatggagcagcagacgcgtacttcgagcg gaggcatccgagcttgaggatcgcgcggctccg ggcgtatatgtcctcgatgtgttgaccaactctatca gagcttggtgacggcaatttcgatgatgcagcttggg cgagggtcgtatcgacgcaatcgtccgatccggag ccgggactgtcggcgtaacaaaicgcccgcagaa gcgcggccgtcggaccgatggctgtgtagaagtact cgccgatagtggaaaccgacgcccagcactcgtcc gagggcaagggaatag	SEQ ID NO: 33
Proteína	Sintético		atgttcagtaaggagaaagaacttttactggagttgtc	SEQ ID NO: 34

82

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			mavrimaaykvgdrilwtpnfsswtrdeygyk yyyvsegpyekvnhyvnmhseilirrigadstvl kndgalpltgkerlvaligedagsnpygangcsdrg cdngtlamgwsgtanfpylvtpqaisnevlknk ngvftatdnwaidqiealaktasvslfvnadsgegy invdgnlgdrknltlwmgdnvikaasncntivii hsvgpvlvnewydnpnvtailwgglpqgesgnsla dvlygrvnpgakspftwgktreayqdylvtepnn ngapqedfvegvidyrgfdkmetpiyefgyglsyt tfmysnlevqvlsapayepasgeteaaptfgevgnas nylypdglqkitkfiypwlinstdleassgdasygqds sdylpegatdgsaqpilpagggpggnprlydelirvs vtikntgkvagdevpqlyvslggpnepkivlrqferi tlqpseetkwsttltrrdlanwnvekqdwetsypk mvfvgsssrkpplrslptvh	
Beta glucosidasa - "GSF V1"	Aspergillus kawachii/ con secuencia sintética	(Secuencia de señal sustituida y secuencias de KDEL en letras MAYÚSCULAS)	MKTNLFLFLIFSLLSLSSAEdelaysp pyypspwangqgdwaqayqravdivsqmtlaek vnlttgtwelcvgqtggvprlgvpgmclqdspl gvrdsdynsafpsgrmnvaatwdknlaylrgkamg qefsdkgadiqlgpaagplgrspdggmwegfspd palsgvlfaetikgiqdagvvatakhyaieqehfrq apeaaggyfnisesgsanlddktmhelylwpfada ragagavmcsynqinnsygcqnsytlnklkaelgf qgfvmstdwaahhagvsgalagldmsmpgdvdy dsqtsywgtnltvsvlntgvpqwrddmavrimaa yykvgrdrilwtpnfsswtrdeygykyyyvsegg yekvnhyvnmhseilirrigadstvlkndgalplt gkerlvaligedagsnpygangcsdrgcdngtlam gwsgtanfpylvtpqaisnevlknkngvftatdn waidqiealaktasvslfvnadsgegyinvdgnlg drknltlwmgdnvikaasncntiviihsvgpvlv newydnpnvtailwgglpqgesgnsladvlygrvn pgakspftwgktreayqdylvtepnnngapqedf vegvidyrgfdkmetpiyefgyglsytfmysnlev qvlsapayepasgeteaaptfgevgnasnylypdgl qkitkfiypwlinstdleassgdasygqdsdylpega tdgsaqpilpagggpggnprlydelirvsvtikntgk vagdevpqlyvslggpnepkivlrqferitlqpseet kwsttltrrdlanwnvekqdwetsypkmvfvgss srkpplrslptvhKDEL	SEQ ID NO: 38
Beta glucosidasa - "GSF V2"	Aspergillus kawachii con secuencia sintética	Secuencias de señal sustituida y de KDEL en MAYÚSCULAS y las sustituciones de 13 aminoácidos en MAYÚSCULAS y subrayadas)	MKTNLFLFLIFSLLSLSSAEd elayspyyppspwangqgdwaqayqravdi vsqmtlDekvnlttgtwelcvgqtggv prlgvpgmclqdsplgvrdsdynsafpAgm nvaatwdknlaylrgkamgqefsdkgadiq lgpaagplgrspdggmwegfspdalsgv lfaetikgiqdagvvatakhyaieqehfr qapeaagFgf nisesgsanlddktmhelyl wpfadairagagavmcsynqinnsygcqns ytlnklkaelgfgqgfvmstdwaahhagvsg alagldmsmpgdvdydsqtsywgtnltvsv lntgvpqwrddmavrimaa yykvgrdrilw tpnfsswtrdeygykyyyvsegpyekvnQ yvnvqmhseilirrigadstvlkndgalp ltgkerlvaligedagsnpygangcsdrgc dngtlamgwsgtanfpylvtpqaisnev lkHknkngvftatdnwaidqiealaktasvsl fvfnadsgegyinvdgnlgdrknltlwmg dnvikaasncntivYihsvgpvlvnewy dnpnvtailwgglpqgesgnsladvlygrv	SEQ ID NO: 39

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			cccagtggcgtgtggatgacatggccgtccggatcat ggctgcctactacaaggttggccgcgaccgicctgg actcctcccaatttcagctcgtggactcgcgatgaatat ggcttcgcgcacttcttccctccgaaggcgcttatgaa cgtgtcaatgaattcgtcaacgtgcagcgtgaccaatgc ccaggatgatccgtcggattggcgcggatagtgtcgtg ctcttgaaagacgacggtgcccttcccttgacgggcca ggagaagactgttggcattctgggcgaagacccgg gtcgaalccgaaggagcaaatggttcagtaccgt ggctgtgacaagggtactctggccatggcttgggta gtggtactgccaaacttcccttactgtgactccgaac aggccattcagaacgaggtctgaaggccgtgga atgtcttggccgtgacggacaactatgatacacagcag atgccgcggttgccttcaatccacggttcatggttt cgtgaacgcagacgcgggtgaaggtttccctaatgtg acggaaacatgggtgatcgcaagaacctcacctctg gcagaacggagaggaaigtatcaagactgtcacgga gcactgcaacaacaccgttgtgtgatccattcgggtg gacctgttctcatcgatgagtggtatgcgaccccaat gtcacccgcatctgtgggtggtctccgggccaagg agtctggcaacgccattgcggacgtgctgacggccg cgtcaacctggcggcaagaccccttacctggggt aagacgcgcgcgtcctacggcgactacctctcacgg agcccaacaacggcaacggtgtctctcaagacaact caacgaggggcgtgtttactaccgtcgtctgaca agtacaatgagacgccatctacgagttcggtcatggt ctgagctacacgacgttgagctgtggtccctcaggt ccagcttatcaacggatccagctatgttccactacgg gtcagacgagcgcgccaggcatttggtaaagtcca ggacgcgtctagctacctgtaccctgagggactgaag aggatticcaagttcatctatccctggctgaacttacc gatcttaagcgtctaccgcgacccctgaatacggaga gccaacttcgagtatattcctgaagggtctaccgatg gctctcctcagccccgtctgctgccagcgggggtcc tggcggcaaccccggtctctatgaggatctcttcagg tttctgtgaccatcacaacacccggcaagggtgctgt gatgagggtgcctcagctgtatgttctgctgggtgccc caacgagccgaagcgggtgctgcgcaagttcgagcg cctgcacatgcccttggcagcaaaaggctcggacg actacctgaaccgcgtagcctagccaactgggatg tcgtggcccaggactggaagatcactccctatgctaag accatcttgttggcacctcttcgcgcaagctgcctc gctggtcgttggccacgggtgcagtaa	
XM_001212225 de beta-glucosidasa	Aspergillus terreus NIH2624	Traducción de CDS anterior	mklsileaaaltaasvsaqddlaysppypspwad ghgewsnaykravdivsqmtllekvnlttwtgwele rcvgqtsvprlgipsicldspgirmsdynsafpa ginvaatwdkklayqrkamgeefsdkgidvqlgp aagplgrspdggrmwegfspdpaltgvlfaetikgiq dagviatakhyilneqehfrqvgeaqgygfnitets snvddkthmelylwpfadavragvgavmcsynqi nnsygcqnsllnkllkaelgfqfvmstdwsahhsg vgaalagldmsmpgdifdsdgsfygntltvgvlngt ipqwrddmavrimaaykvgrdrwtppnfssw trdeygfahffpsegayervnefvnvqrdhaqviri gadsvllkndgalpltgqektvgilgedagsnpg angcsdrgcdkgilamawsgtanfpylvtpesai qnevlkgrgnvfavtdnydtqjaavasqstvsivfv nadagegflnvdmngdrknltlwqngeeviktvt ehcnntvvihsvgpvlidewyahpnvtgilwagl pgqesgnaiadvlygrvnpgektpftwgktrasyg	SEQ ID NO: 42

5

10

15

20

25

30

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			dylttepnnngapqdnfnegvfidyrrfdkynetpi yefghglstyttfelsinglvqlingssyvpttgqtsaaqa fgkvedassylpeglkriskfiypwlnstdlkastg dpeygepnfeyipegaidsppqrlpasggpggnp glyedlfqvsvtiintgkvagdevpqlyvslggpnep krvlrkferlhiapgqqkvwttnlrrdlnwdvvaq dwkitpyaktifvgtsrrklplagrlprvq	
Exoglucanasa 1 (No. de registro P62694)	Trichoderma reesei	Proteína de longitud completa (aa 1-17 = señal)	myrklavisaflataraqsactlqsethppltwqkcss ggctctqqtgsvvidanwrwthatnsstncydntws stlcpdnetcaknccldgaayastygvttsgnslsigf vtqsaqknvgarlylmasdttqefllgnecfsfdvd vsqllpcglnqalyfvsmdadggvskypnttagaky gtgycdsqcprdlkfingqanvegwepssnnantgi ggghgscscsemiweansisealtphpcttvqgeice gdgcggtydnryggtdcpdcdwnpyrlgntsfy gpgssfldttklrtvtqfetsgainryyvqngvtfq qpnaelgsysgneinddyctaeaeefgssfsdkgg ltqfkkatsggmvlvmslwdyyanmlwldstyp tnetsstpgavrgscstssgvpqavcsqspnakvtfs nikfgpigstgnpsggpppggnrgtttrpatttgssp gptqshygqcgigysgptvcasgttcqvlnpysq cl	SEQ ID NO: 43
Proteína exocelobiohidro- lasa 1 modificada	Trichoderma reesei	Proteína - (sec anterior con señal ER reemplazada y KDEL añadida en MAYÜSCU- LAS)	MKTNLFLFLIFSLLSLSSAEqsactlqs cthppltwqkcssggctctqqtgsvvidanwrwthat nsstncydntwsstlcpdnetcaknccldgaayast ygvttsgnslsigfvtqsaqknvgarlylmasdttq efllgnecfsfdvdvsqllpcglnqalyfvsmdadgg vskypnttagakygtgycdsqcprdlkfingqanve gwepssnnantgiggghgscscsemiweansiseal tphpcttvqgeicegdgcggtydnryggtdcpdcd dwnpyrlgntsfyggssfldttklrtvtqfetsgai nryyvqngvtfqpnaelgsysgneinddyctaeae efgssfsdkggltqfkkatsggmvlvmslwdyy anmlwldstyptnetsstpgavrgscstssgvpqav esqspnakvtfsnikfgpigstgnpsggpppggnrg tttrpatttgsspgptqshygqcgigysgptvcas gttcqvlnpysqclKDEL	SEQ ID NO: 44
Exoglucanasa 2 (No. de Registro P07987)	Trichoderma reesei	Proteína de longitud completa (1-24 = secuencia de señal)	mivgilatlalaaavpleerqacssvwgqcggn wsgptccasgstcvysndyysqclpgaassssstra sttsrvspitrsstsatpppgsttrvpvgshtatysgn pfvgvtpwanayyasevsslaipsltgamataaaav akvpsfmwldtldktplmeqladirtankngny gqfvvydlpdrdcaalasngeysiadggvakykny idtirqivveysdirtllviepdsianlvtnlgtpkcna qsaylecinyavtqlnlpnvamyldaghagwlgw panqdpaaqlfanvyknasspralrglatnvanyng wnitsppsytqgnavyneklyihaigpllanhgwsn affitdqrsgkqptgqqqwgdcwcnvigtgfgirps antgdsllidsfvvwpgecdgtssdssaprfdshcal pdalqpapqagawfqayfvqiltanpsfl	SEQ ID NO: 45
Exoglucanasa 2 Modificada	Trichoderma reesei	Proteína - (sec anterior con señal ER reemplazada y KDEL añadida en MAYÜSCU- LAS)	MKTNLFLFLIFSLLSLSSAEqacssv wgqcggnwsgptccasgstcvysndyysqclpg aassssstraasttsrvspitrsstsatpppgsttrvpv gshtatysgnpfvgvtpwanayyasevsslaipsltg amataaaavakvpsfmwldtldktplmeqladirt ankngnyagqfvvydlpdrdcaalasngeysiad ggvakyknyidtirqivveysdirtllviepdsianlv tnlgtpkcnaqsaylecinyavtqlnlpnvamylda ghagwlgwpanqdpaaqlfanvyknasspralrgl atnvanyngwnitsppsytqgnavyneklyihaig	SEQ ID NO: 46

85

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			npgakspftwgktreayqdy lvtepnnng apqedfvegv fidyrgfdkr netpiyefgy glsyttfnysnlevqvlsap ayepasgete aaptfgevgn asDylypSgl qRitkfiypw lnGtdleassgdasygqds dylpegatdg saqpilpagg gpggnprlyd elirvsvtik ntgkvagdevpqlyvslggp nepkivlrqf eritlpsee tkwsttlrr dlanwnvekq dweitsypkmvfvgsrrkL piraslpvh KDEL	
Secuencia de beta glucosidasa #2 de Patente de E.U.A. No. 7223902 (No. Reg. ABT13410.1)	Aspergillus niger	proteína	mrflieavaltavslasadelaysppypspwang qgdwaqayqravdivsqmtldekvnltgtgwelel cvgqtggvprlgvpgmclqdsplgvrdsdynsafp agmnvaatwdknlaylrkamgqcfsgkadiql gpaagplgrspdggrnwegfspdpalsgvlfaetik giqdagvvatakhyaieqehfrqapeaaggfgfnis esgsanlddktmhelylwpfadairagagavmcsy nqinnsygcqnsytlnklkaelgfgfvmwdwaa hhagvsgalagldmsmpgdvdydsywgntli svlntgvpqwrddmavrimaaykykgrdrwtp pnfsswtrdeygykyyyvsegpyekvnqyvnvqr nhselirrigadstvlkndgalpltgkerlvaligeda gsnpygangcsdrgcdngtlamgwsgtanfpyl vtpeqaisnevlkhkngvftatdnwaidqiealakta svslvfnadsgcgyinvdgnlgrdmilwrmgdn vikaaasncntivvihsvgpvlvnewydnpnvtai lwgglpqgesgnsldvlygrvnpgakspftwgktr eayqdylvtepnnngapqedfvegvfdyrgfdk metpiyefgyglsyttfnysnlevqvlsapayepasg cteaaptfgevgnasdylypsglritkfiypwlngtd leassgdasygqdsdylpegatdgsaqpilpaggg pggnprlydelirvsvtikntgkvagdevpqlyvslg gpnepkivlrqferitlpsee tkwsttlrrdlanwn vekqdwetsypkmvfvgsrrklpiraslpvh	SEQ ID NO: 40
Precursor XM_00121225 de beta- glucosidasa I	Aspergillus terreus NIH2624	CDS	atgaagcttccatttggaggcagcagcttgacagct gcctccgtagtcagcgacagcagcgcgcatact cccgcgcgtactacccttccctgggcgcatggcca cggtagtggtcgaacgcgtacaagcgcgcttagat atcgtctctcagatgacattgacggagaaggtcaatct caccaccggtactggatggaggtggagaggtgtgtc ggtcagacgggcagtgccctagactgggaatccaa gcctctgtctgcaggatagccctctgggtattcgcatgt cggactataaactcgccctccctcgccgtattaacgttg cggccacctgggacaagaagcttgccaccaacgcg gcaaggcaatgggcgaggaattcagtgacaagggtta ttgatgttcagttggccctgtctccggtcctctggca gggtcccccgatggaggccgaaactgggagggtctct ctcctgatccccccctgactggtgtgtgttcgccgaga cgatcaagggtatccaggacgcccggagtattgttac cgcgaaagcactacattctcaacgaacaagagcattcc gccagggtcggcgaagcccagggtatggcttcaacat caccgaaccgtgagctcaaatgtggtgacaagacc atgcacgagctgtatctctggccctcgccgatcggt gcgcgcggcggtggcgctgtgatgtctcctacaac cagatcaacaacagctacggatgccaaaacagttga ccctgaacaagctctgaaagccgaactcggatttcag ggatttgtcatgagtactggagtgtcaccacagcgg tgttggcgccgcttggctgttggacatgtccatgcc gggagatatacttgcagacgcgccacttctctatgg cacgaacctgactgttggcgtcctcaacggcaccattc	SEQ ID NO: 41

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
			pllanhgwsnaffitdqgrsgkqptgqqqwgdwcnvigtgfgirpsantgdsldsfvwwkpggcedgtsdssaprfdshcalpdalqpapqagawfqayfvqlitnanspsfIKDEL	
5 Endoglucanasa E1 (No. de registro P54583)	Acidothermus cellulolyticus	Proteína de longitud completa (1-41 = secuencia de señal	mpralrrvpgsrvmlrvgvvavilvaalanlavp rparaagggwyhtsgreildannvpvriaginwfgf etcnyvvhglwsrdyrsmldqikslgyntirlpysd dilkpgtมป์nsinfyqmndlqgltslqvmdkiva yagqiglrilrdhrpdcsgqsalywtssvseatwisd lqalaqrykgnptvvgfdlhncphdpacwgcgdp idwrlaaeragnavlsvnpnllifvegvsyngdsy wwggnlqgagqypvvlvnpnlrvysahdyatsvy pqtwfsdptfpnnmpgiwnknwgylfhqnipv wlgefgtltqsttdqtwlklvqylrptaagyadsfqw tfvswnpdsgdtggilkddwqvtvdkdgylapik ssifdpvgasaspssqpspsvpspspsasrtptpt ptptasptplltptatptasptpsptaasgarclasyq vnsdwngngftvravtnsgsvatktwtvswtfggn qtiitnswnaavtqngqsvtarnmsynnviqpgqntt fgfqasytgsnaaptvacaaas	SEQ ID NO: 47
10 Xilanasa - U39784	Aspergillus niger	CDS	atgaaggtcactgcggctttgcaagtctcttgcctacg gccttcgcggccctgctccggagccgttctgtgtc gcgaagtgcggatcaactacgtgcagaactacaac ggcaaccttggtgacttcacctacgagagtagccg ggacattttccatgtactgggaggatggagtcagttcc gacttcgtcgttggttggctggaccactggctcctcl aaatctatcacctactctgccaatacagcgttctagc tccagctcctacctggctgtctacggctgggtcaactct cctcagcgccgaatactacatcgtcaggattacgggta ttacaaccttgtagctcggccacgagccttggtaccg tgtactctgatggaagcacctaccagtctgcaccgac actcgacgaacgcggccatctatcacaggaacaagc acgttcacgcagtagtctcctcgtcgtgaaagtacacgc acatccggaacagttagctatcgccaaccatttcaattc tggcgcgagcatgggttcggcaatagcaactcaatta tcagggtcatggcggtggaggcatggaacggtgtcgg cagtgccaggtgtcagcatctcctcttaa	SEQ ID NO: 48
Xilanasa - AAA99065.1	Aspergillus niger	Traducción de proteína de secuencia anterior	mkvtaaasllltafaapapepvlvsrsaginyvqny ngnlgdfdydestgtfsmywedgvssdfvvglgwt gssksitysaqysassssylavgywnspqacyi vedygdynpcssatslgtvysdgstyqvctdtrtps itgtstftqyfsvrestrstgtvtianhfnfwaqhgfens nfnyqvmaveawngvgsasvtiss	SEQ ID NO: 49
15 Ligninasa 1508163A	Phanerochaete chrysosporium	Enzima	mafkkqlfaaislalslaanaaaviekratcsngktvg dasccawfdvlddiqqnlfhggqcgacahesirlvf hdsiaispameaqqkfggggadsimifddietafh pnigldeivklqkpfvqkhgvtpgdfiafagavalsn cpgapqmnfftrapatqpapdglvpcpfhtvdqii nrndagefdelelvwmlsahsvaavndvdpvtvq glpfdstpgifdsqffvetqlrgtafpgsggnqgeves plpgeiriqsdhtiaryrtacewqsfvnnqsklvdd fqfiflaltqlgdpnamtdcsdvipqskpipgnlpfs ffpagktikdveqacaetpfptltlpggpetsvqri	SEQ ID NO: 50
Ligninasa: manganeso peroxidasa (EC 1.11.1.13)	Trametes versicolor	Enzima	vaxpdgvntatnaaxxqlfdggccgeevhesiarhx aigvsnpcgapqigvsnxpgapqlardsrtaxewq sllixselvpxpppalsnadveqaxactpf	SEQ ID NO: 51
25 Lignina peroxidasa	Phanerochaete chrysosporium	Enzima	mafkkqlfaaitvalsitaanaavvkekratcangktv gdasccawfdvlddiqanmfhggqcgacahesirl	SEQ ID NO: 52

87

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCIÓN	SECUENCIA	SEQ ID NO:
(Accession No. P49012)	(un basidiomicetes)		vfhdslaispameakgkfggggadsimifdtietaf hpniglddevvamqkpfvqkhgvtpgdfiafagava lsncpgapqmnfftrkpatqpapdglvpepfhtvd qiiarvndagefdelevwmlsahsvaavndvdp vqglpfdstpgifdsqffvetqfrgtlfpsggnqgev esgmageiriqtdhltlardsrtacewqsfvgnqsklv ddfqfiflaltqlgqdpnamtdcsdviplskipgng pfsffppgkshsdiqacaetpfpvltpgpatsvar ipphka	
Secuencia de ARNi para BCS (genbank AB237643.1)	Sintético	Fragmento de ARNi-suprime conglucina	ccacagacttcaatctggtgatgccctgagagtcacct cagggaaccacatactatgtggtcaacctgacaacaa cgaaaatctcagattaataacactcgccataccgttaa caagcctgtagatttgag	SEQ ID NO: 53
Secuencia de ARNi química	Sintético	El casete de ARNi completo para producir BCS	gcgccgcctcgcagcctacagacttcaatctggtgat gccctgagagtcacctcaggaaccacatactatgtggt caaccttgacaacaacgaaaatctcagattaataacac tcgccatacccgtaacaagcctgtagatttgaggat atcgagctcgggctgtttcttctcgtcacacacacaat aggggtggcctatgtattagccttcaatgtctctgtaga ccctatgatagtttgcaagccactaccaccttatgctc ccatatacttaaccgtgaaagcttactagtctagtacc caattgtaaggaaataattttctttttcttttagtat aaaatagttaagtgtattagtagtattataataata tagttgtataaattgtgaaaaataattataaatattgtt tacataaacaacatagtaattgtaaaaaatagacaagt gatgtgtaagacgaagaataaaagttagagtaag tatatttttaattgaattgacgaacatgtaagatgata tactagcattaattttgttttaatacataatgtaattcag ctggttgatgaattaaatacaatgataaaatactatagt aaaaataagaataaataaataaataaataatttttatgat taatagtttattataataataataatctataccactactaat attttagtttaaaagttaataaataattttgttagaattcca atctgctgttaatttaataaataaataaataaataaaca agctaaagtaacaataaatacaactaataagaacag taactaattgtaacaaaacataatctaattgtaataaac aaagcgtaaagctttcacggttagaataataggagca taagggtgtagtggttgcaaaactatcatagggtct accagagacattgaaggctaaatacataggccacct attgtgagtgtgacgagaagagaacagcccgagctc gatacctcaaatctaccaggctgttaacgggtatggc gagtgttaatactgagatttctgtgtgtcagggttga ccacatagtagtggttctgaggggactctcagggca tcaccagattgaagtcgtaggctcaggtctagagcgg ccgc	SEQ ID NO: 54
Fragmento de ARNi de glicina (de glicina A1bB2) (genbank AB030495)	Sintético	Secuencia de ARNi de proteína de almacenamiento (para producir línea SP-)	cattgacgagaccatttgacaatgggacttcgccaca acataggccagacttcatcacctgacatctcaacctc aagctggtagcatcacaaccgctaccagcctcgacttc ccagccctctcgtggctcaaacctagtgccagtttg atcactccgcaagaatgctatgtcgtgccactaca acctgaacgcaaacagcataatatacgattgaatgga cgggcattggtacaagtgggtaattgcaatgggtgaga gagtgtttgatggagagctgcaagaggagacaggtgt aactgtgccacaaaactttgcggtggcig	SEQ ID NO: 55
Inserto de ARNi de glicina completo	Sintético	Inserto de casete completo usado para producir línea SP-	gcgccgcctcgcagcattgacgagaccattgcacaa tgggacttcgccacaacataggccagacttcatcacct gacatctcaacctcaagctgtagcatcacaaccgc taccagcctcgacttccagccctctcgtggctcaaac tcagtcccagtttgatcactccgaagaatgctatgt tcgtgccactacaacctgaacgcaaacagcataat atacgcattgaatggacggcattggtacaagtgggta	SEQ ID NO: 56

ORIGEN (GEN)	ORIGEN (ESPECIE)	FUNCION	SECUENCIA	SEQ ID NO:
			attgcaatgggtgagagagtgttgatggagagctgcaa gaggacaggtgttaactgtgccacaaaacttgcggt ggctggatatcgagctcgggctgttctctctgcac actcacaatagggtggcctatgtatttagcctcaatgc tcggtagaccctatgatagtttgaagccactaccac cctatgctcccatatattctaaccgtgaaagcttactag tctagtacccaattggtgaagaaataattttctttt ccttttagtataaaatagttaagtgatgtaattagtagat tataataatagttgttataattgtgaaaaataattata aatatattgtttacataaacaacatagtaataaaaaa tatgacaagtgtgtgaagacgaagaagataaaagt gagagtaagtataatttttaataatgaaatgacgaacatg taagatgatactagcattaatattgttttaatacataata gtaattcagctgtgttgatgaattaaatcaatgataaa atactatagtaaaaaataagaataaaataaataaata ttttttatgattaatagttttatataaataaataatctatacc attactaaatatttttagtttaaaagtttaataaataattttgtta gaaattccaatctgctgtgaattatcaataaacaataat taataacaagctaaagtaacaataatacaaaactaat agaacaagtaataatgaacaaaacataatctaagc taataataacaagcgtaaagcttcacgggttagaatat gggagcataagggtgtagtggttcgcaaaactatcat agggtctaccagagacattgaaggctaaatacatagg ccaccctattgtgagtgacgagaagagaacagcc cgagctcgataccagccaccgcaaagtgttggtgcac agttaacacctgtccctctgcagctcctcatcaaacac tcctcaccattgcaatlcaccactgtaccaatgccgt ccattcaatgcgtatattatgctgtttgcgttcagggtgta gtgtggcacgaacatagcattcttcggagtgatccaa actgggcactgagtttagccacgagagggtggga agtcgaggctggtagcgggttgatgctaccagctga gggtgaagatgtcagggtgaagagctgcctatgt gtggcgaagtccattgtgcaaatggctcgtcaatgct cgagctagagcggccgc	
ARNi de FAD2 (genbank ab188250)	sintético	ARNi - supprime FAD2	gggctgttctctctctgcacactcacaatagggtggc ctatgtatttagccttcaatgtctctgtagaccctatgat agttttgcaagccactaccaccctatgctcccatatatt ctaaccgtga	SEQ ID NO: 57

REIVINDICACIONES

1. Una planta dicotiledónea transgénica, caracterizada porque comprende:

a. una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se resulta en una reducción de al menos 50% en proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre; y

b. un polinucleótido heterólogo que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla; y

en donde la semilla de la planta transgénica es capaz de producir una proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

2. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque el polinucleótido heterólogo comprende además un dominio potenciador de traducción 5' y/o un dominio potenciador de traducción 3'.

3. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la secuencia de señal de retención en ER induce la agregación de la proteína heteróloga en el lumen del ER o una vesícula derivada del ER.

4. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque es un miembro de la familia de las Fabáceas, y opcionalmente del orden Fabales, y opcionalmente del género soya.

5. La planta dicotiledónea de conformidad con la reivindicación 4, caracterizada porque es un miembro del género Glycine.

6. La planta dicotiledónea de conformidad con la reivindicación 5, caracterizada porque es una soya.

7. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque el promotor se deriva de inhibidor de tripsina Kunitz, lectinas de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglucina.

8. La planta dicotiledónea de conformidad con la reivindicación 2, caracterizada porque el dominio potenciador de traducción se deriva de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglicinina.

5 9. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la deficiencia de proteína de almacenamiento es de uno o más de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglicinina.

10 10. La planta dicotiledónea de conformidad con la reivindicación 9, caracterizada porque la semilla tiene una deficiencia de más del 75% de las proteínas de almacenamiento endógenas de la semilla.

15 11. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque comprende además un dominio potenciador de traducción 5' y un dominio potenciador de traducción 3', en donde el promotor y los dominios potenciadores de traducción 3' y 5' se derivan de la misma proteína de almacenamiento.

20 12. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la proteína heteróloga se acumula en una semilla de la planta dicotiledónea hasta un nivel que es mayor que aproximadamente 2% o mayor que aproximadamente 4% o mayor que aproximadamente 5% del peso en seco total de la semilla.

20 13. Una semilla caracterizada porque es de la planta dicotiledónea de conformidad con la reivindicación 1.

14. Una proteína transgénica caracterizada porque se obtiene de la semilla de conformidad con la reivindicación 13.

25 15. La proteína transgénica de conformidad con la reivindicación 14, caracterizada porque la proteína heteróloga ha sido purificada.

16. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la secuencia de codificación de proteína deseada codifica para una enzima o fragmento de la misma.

30 17. La planta dicotiledónea de conformidad con la reivindicación 16, caracterizada porque la enzima es una enzima celulolítica.

18. La planta dicotiledónea de conformidad con la reivindicación 17, caracterizada porque la enzima celulolítica se deriva de una fuente fúngica, una fuente bacteriana, una fuente animal o una fuente vegetal.

5 19. La planta dicotiledónea de conformidad con la reivindicación 17, caracterizada porque la enzima celulolítica es una β -glucosidasa, una exoglucanasa I, una exoglucanasa II, una endoglucanasa, una xilanas, una hemicelulosa, una ligninasa, una ligina peroxidasa o una manganeso peroxidasa.

20. Un producto caracterizado porque comprende la proteína de conformidad con la reivindicación 14.

10 21. Una composición de enzima comercialmente útil, caracterizada porque comprende la proteína de conformidad con la reivindicación 14.

22. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la deficiencia de una o más proteínas de almacenamiento de semilla se debe a la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que
15 codifica para una proteína de almacenamiento de semilla.

23. Una planta dicotiledónea caracterizada porque comprende:

a. una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se resulta en una reducción de al menos 50% en proteína de almacenamiento de planta endógena en comparación con aquella de una planta tipo silvestre; y
20

b. un polinucleótido heterólogo que comprende una región reguladora génica de una proteína compensadora enlazada operablemente a un marco de lectura abierto que codifica para una secuencia que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en
25 ER;

en donde la semilla de la planta dicotiledónea transgénica es capaz de producir la proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

24. Un método para almacenar establemente una enzima antes de usar, almacenar dicha enzima en una semilla de una planta dicotiledónea transgénica, en
30 donde:

a. tenga una deficiencia en una o más proteína de almacenamiento de semilla, en donde la deficiencia se resulta en una reducción de al menos 50% en una proteína de semilla endógena; y

5 b. un polinucleótido heteróloga que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, un enzima deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla; y

10 donde la semilla de la planta transgénica es capaz de producir dicha enzima a un nivel que es al menos aproximadamente 5% del peso en seco total de la semilla; y

almacenar dicha enzima en dicha la semilla de la planta transgénica.

25. Un método para producir una cantidad incrementada de una proteína heteróloga en semillas de una planta dicotiledónea, caracterizado porque comprende:

15 a. transformar establemente una célula vegetal con un polinucleótido que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento
20 de semilla;

b. obtener una línea vegetal homocigótica de la célula vegetal establemente transformada;

c. introducir la línea vegetal establemente transformada en una planta que tenga una deficiencia en una o más proteínas de almacenamiento de semilla endógenas,
25 en donde la deficiencia se resulta en una reducción de al menos 50% en la proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre;

d. cultivar semillas de la planta transgénica introgrida; y

e. obtener la proteína heteróloga de las semillas de la planta transgénica
30 introgrida,

donde dicha semilla de la planta transgénica introgrida es capaz de producir una proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

26. El método de conformidad con la reivindicación 25, caracterizado porque la
5 deficiencia en una proteína de almacenamiento de semilla endógena se debe a la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que codifica para una proteína de almacenamiento de semilla.

27. Un método para producir una cantidad incrementada de una proteína heteróloga en semillas de una planta dicotiledónea, caracterizado porque comprende:

10 a. transformar establemente una célula vegetal con un polinucleótido que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento
15 de semilla, en donde el polinucleótido comprende además una secuencia de ARNi que es capaz de la subregulación de una proteína de almacenamiento de semilla endógena;

b. obtener una línea vegetal homocigótica de la célula vegetal establemente transformada;

c. cultivar las semillas de la línea vegetal homocigótica; y

20 d. obtener la proteína heteróloga de las semillas de la planta homocigótica.

REIVINDICACIONES

1. Una planta dicotiledónea transgénica, caracterizada porque comprende:

a. una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se traduce en una reducción de al menos 50% en proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre; y

b. un polinucleótido heterólogo que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla; y

en donde la semilla de la planta transgénica es capaz de producir una proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

2. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque el polinucleótido heterólogo comprende además un dominio potenciador de traducción 5' y/o un dominio potenciador de traducción 3'.

3. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la secuencia de señal de retención en ER induce la agregación de la proteína heteróloga en el lumen del ER o una vesícula derivada del ER.

4. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque es un miembro de la familia de las Fabáceas, y opcionalmente del orden Fabales, y opcionalmente del género soya.

5. La planta dicotiledónea de conformidad con la reivindicación 4, caracterizada porque es un miembro del género Glycine.

6. La planta dicotiledónea de conformidad con la reivindicación 5, caracterizada porque es una soya.

7. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque el promotor se deriva de inhibidor de tripsina Kunitz, lectinas de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de

maduración de semilla, glicinina o conglucina.

8. La planta dicotiledónea de conformidad con la reivindicación 2, caracterizada porque el dominio potenciador de traducción se deriva de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglucina.

9. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la deficiencia de proteína de almacenamiento es de uno o más de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglucina.

10. La planta dicotiledónea de conformidad con la reivindicación 9, caracterizada porque la semilla tiene una deficiencia de más del 75% de las proteínas de almacenamiento endógenas de la semilla.

11. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque comprende además un dominio potenciador de traducción 5' y un dominio potenciador de traducción 3', en donde el promotor y los dominios potenciadores de traducción 3' y 5' se derivan de la misma proteína de almacenamiento.

12. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la proteína heteróloga se acumula en una semilla de la planta dicotiledónea hasta un nivel que es mayor que aproximadamente 5% del peso en seco total de la semilla.

13. Una semilla caracterizada porque es de la planta dicotiledónea de conformidad con la reivindicación 1.

14. Una proteína transgénica caracterizada porque se obtiene de la semilla de conformidad con la reivindicación 13.

15. La proteína transgénica de conformidad con la reivindicación 14, caracterizada porque la proteína heteróloga ha sido purificada.

16. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la secuencia de codificación de proteína deseada codifica para una enzima o fragmento de la misma.

17. La planta dicotiledónea de conformidad con la reivindicación 16, caracterizada

porque la enzima es una enzima celulolítica.

18. La planta dicotiledónea de conformidad con la reivindicación 17, caracterizada porque la enzima celulolítica se deriva de una fuente fúngica, una fuente bacteriana, una fuente animal o una fuente vegetal.

5 19. La planta dicotiledónea de conformidad con la reivindicación 17, caracterizada porque la enzima celulolítica es una β -glucosidasa, una exoglucanasa I, una exoglucanasa II, una endoglucanasa, una xilanas, una hemicelulosa, una ligninasa, una ligina peroxidasa o una manganeso peroxidasa.

10 20. Un producto caracterizado porque comprende la proteína de conformidad con la reivindicación 14.

21. Una composición de enzima comercialmente útil, caracterizada porque comprende la proteína de conformidad con la reivindicación 14.

15 22. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la deficiencia de una o más proteínas de almacenamiento de semilla se debe a la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que codifica para una proteína de almacenamiento de semilla.

23. Una planta dicotiledónea caracterizada porque comprende:

20 a. una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se traduce en una reducción de al menos 50% en proteína de almacenamiento de planta endógena en comparación con aquella de una planta tipo silvestre; y

25 b. un polinucleótido heterólogo que comprende una región reguladora génica de una proteína compensadora enlazada operablemente a un marco de lectura abierto que codifica para una secuencia que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER;

en donde la semilla de la planta dicotiledónea transgénica es capaz de producir la proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

24. Un método para almacenar establemente una enzima antes de usar,

caracterizado porque comprende:

almacenar una enzima endógena en semillas de una planta dicotiledónea transgénica, en donde la enzima endógena comprende al menos aproximadamente 5% del peso en seco total de las semillas y la enzima endógena se acumula predominantemente en cuerpos de proteínas ER dentro de las semillas.

25. Un método para producir una cantidad incrementada de una proteína heteróloga en semillas de una planta dicotiledónea, caracterizado porque comprende:

a. transformar establemente una célula vegetal con un polinucleótido que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla;

b. obtener una línea vegetal homocigótica de la célula vegetal establemente transformada;

c. introducir la línea vegetal establemente transformada en una planta que tenga una deficiencia en una o más proteínas de almacenamiento de semilla endógenas, en donde la deficiencia se traduce en una reducción de al menos 50% en la proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre;

d. cultivar semillas de la planta transgénica introgrida, las semillas de la planta transgénica introgrida produciendo una proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla; y

e. obtener la proteína heteróloga de las semillas de la planta transgénica introgrida.

26. El método de conformidad con la reivindicación 25, caracterizado porque la deficiencia en una proteína de almacenamiento de semilla se debe a la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que codifica para una proteína de almacenamiento de semilla.

27. Un método para producir una cantidad incrementada de una proteína heteróloga en semillas de una planta dicotiledónea, caracterizado porque comprende:

- 5 a. transformar establemente una célula vegetal con un polinucleótido que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla, en donde el polinucleótido comprende además una secuencia de ARNi que es capaz de la subregulación de una proteína de almacenamiento de semilla endógena;
- 10 b. obtener una línea vegetal homocigótica de la célula vegetal establemente transformada;
- c. cultivar las semillas de la línea vegetal homocigótica; y
- d. obtener la proteína heteróloga de las semillas de la planta homocigótica.

RESUMEN DE LA INVENCION

Se describe una planta dicotiledónea transgénica que tiene una deficiencia de una o más proteínas de almacenamiento de semilla vegetal, que tiene además una construcción de polinucleótido transgénica que comprende un marco de lectura abierto
5 enlazado operablemente a un promotor de proteína de almacenamiento y una secuencia de señal ER. La construcción de polinucleótido codifica para un producto de proteína que puede acumularse a altos niveles en la semilla. Se proporcionan también métodos para producir una proteína heteróloga en una semilla vegetal.

1/15

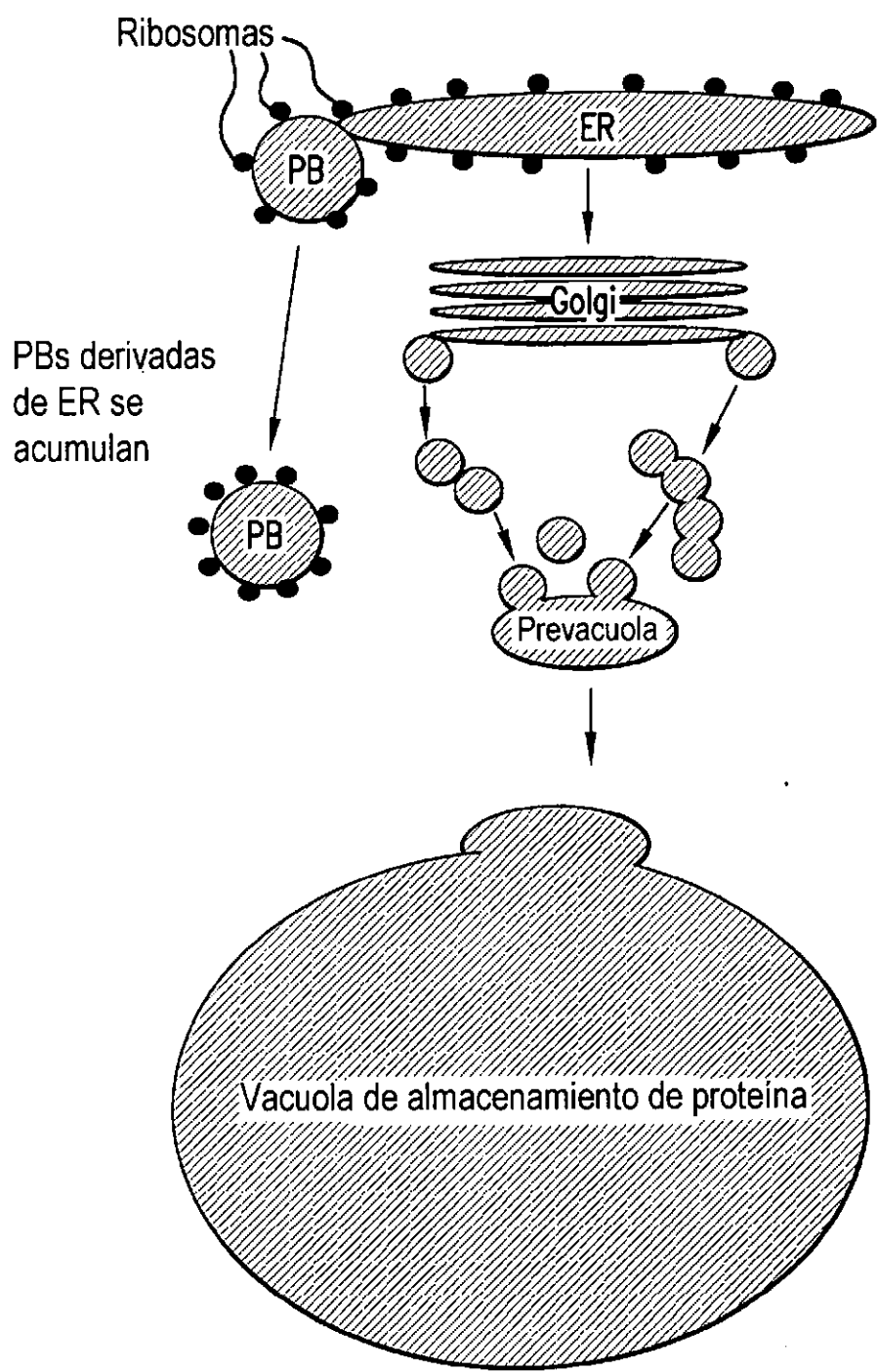


FIG.1

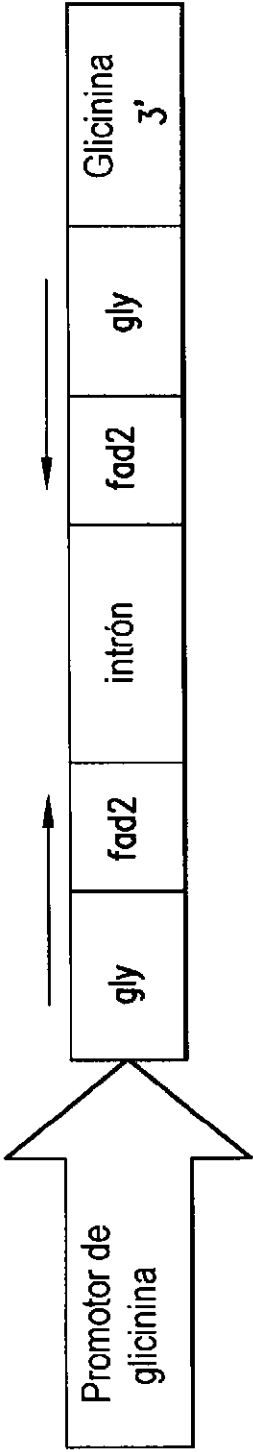


FIG.2

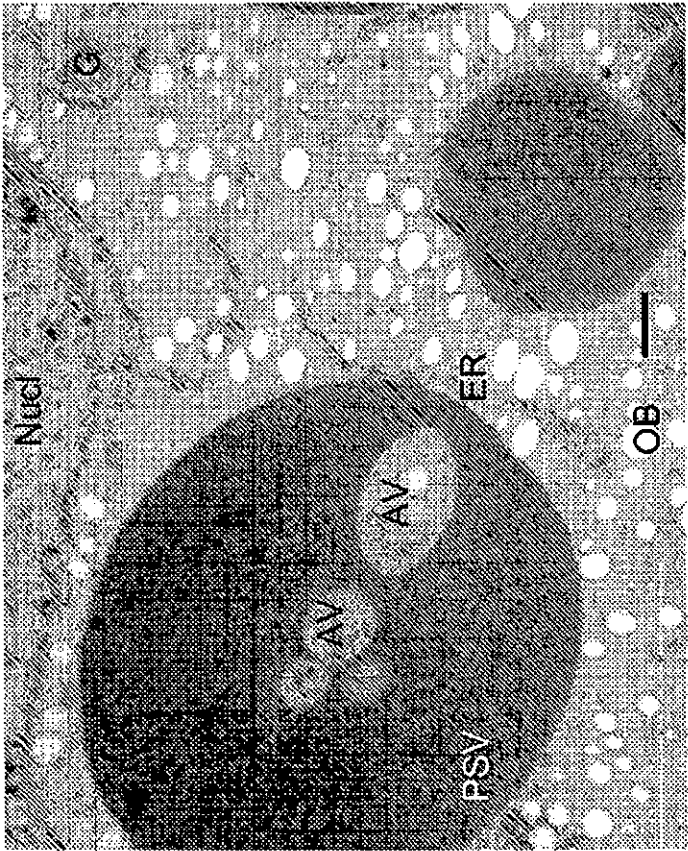


FIG.3B

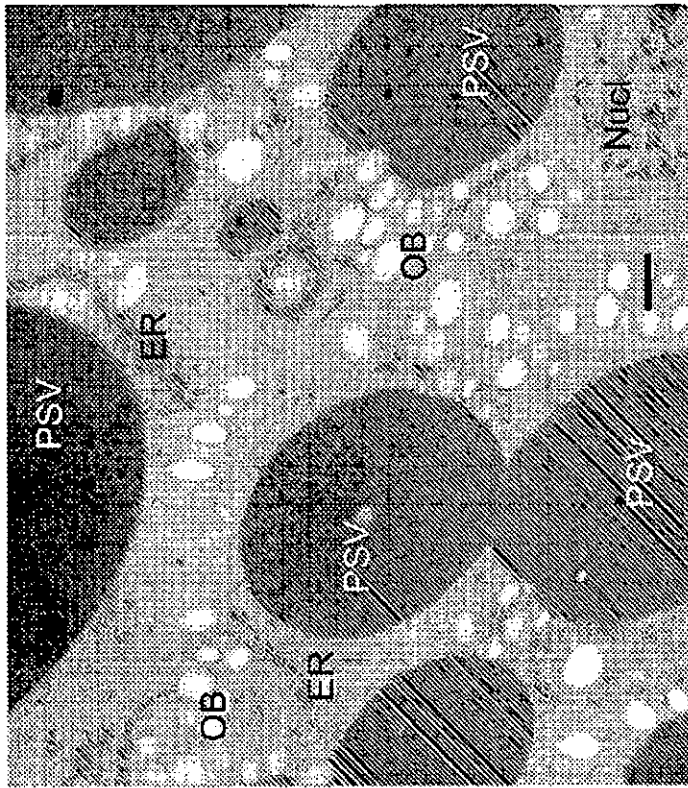


FIG.3A

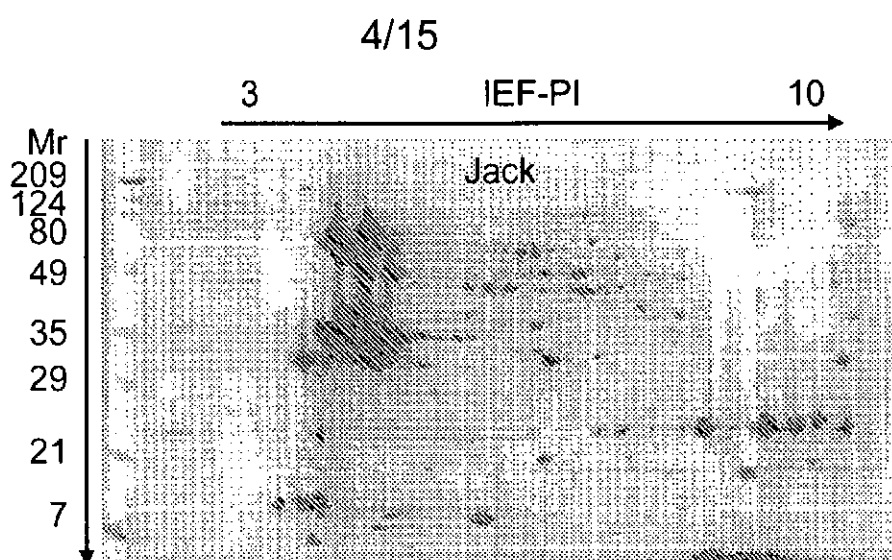


FIG.4A

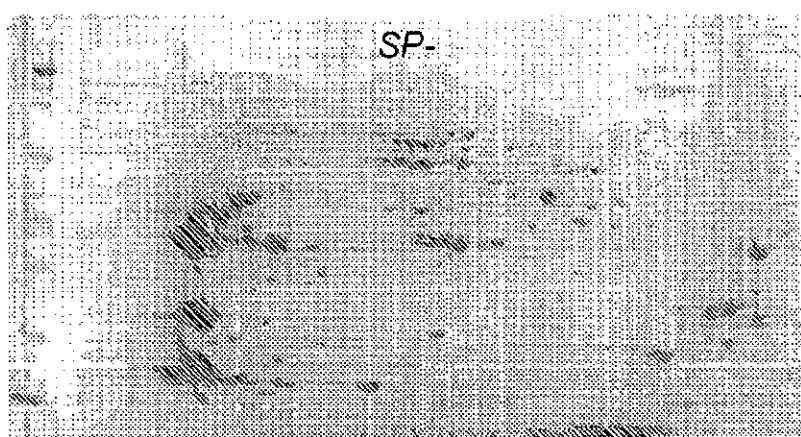


FIG.4B

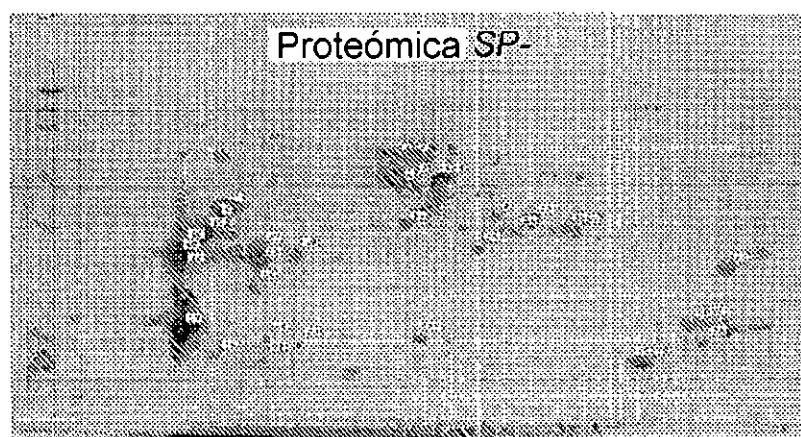
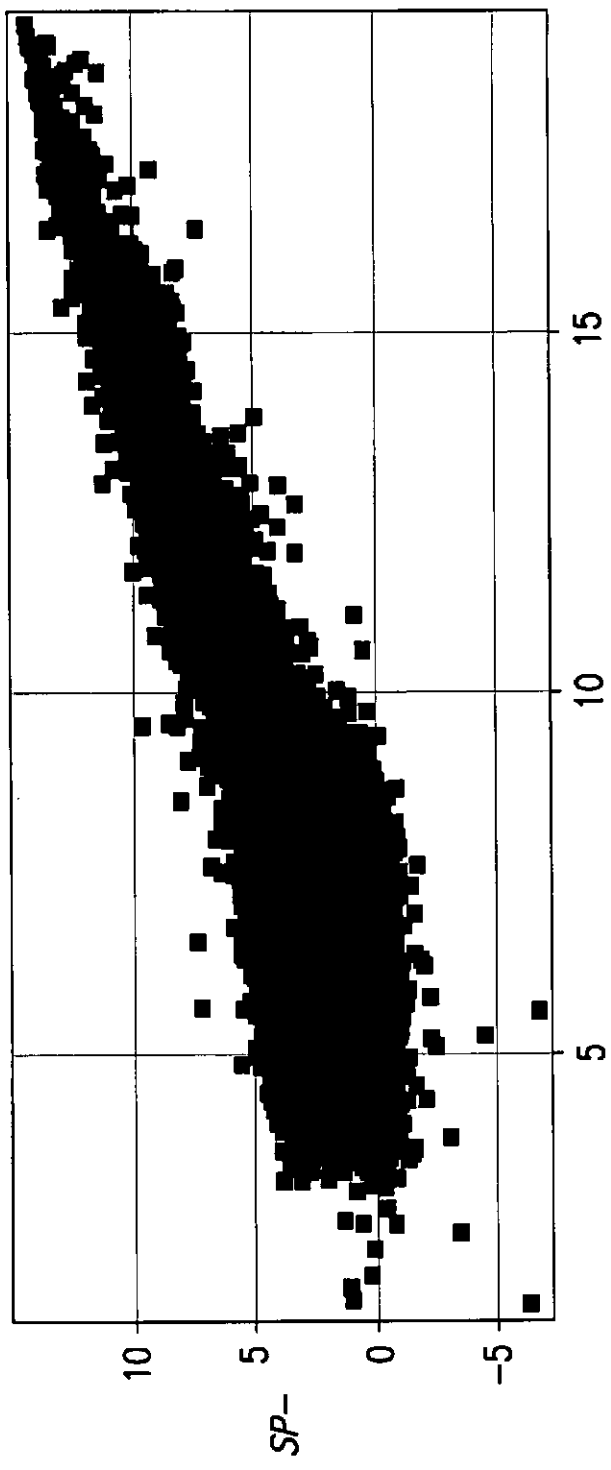


FIG.4C

5/15



Jack Tipo Silvestre

FIG.5

105

6/15

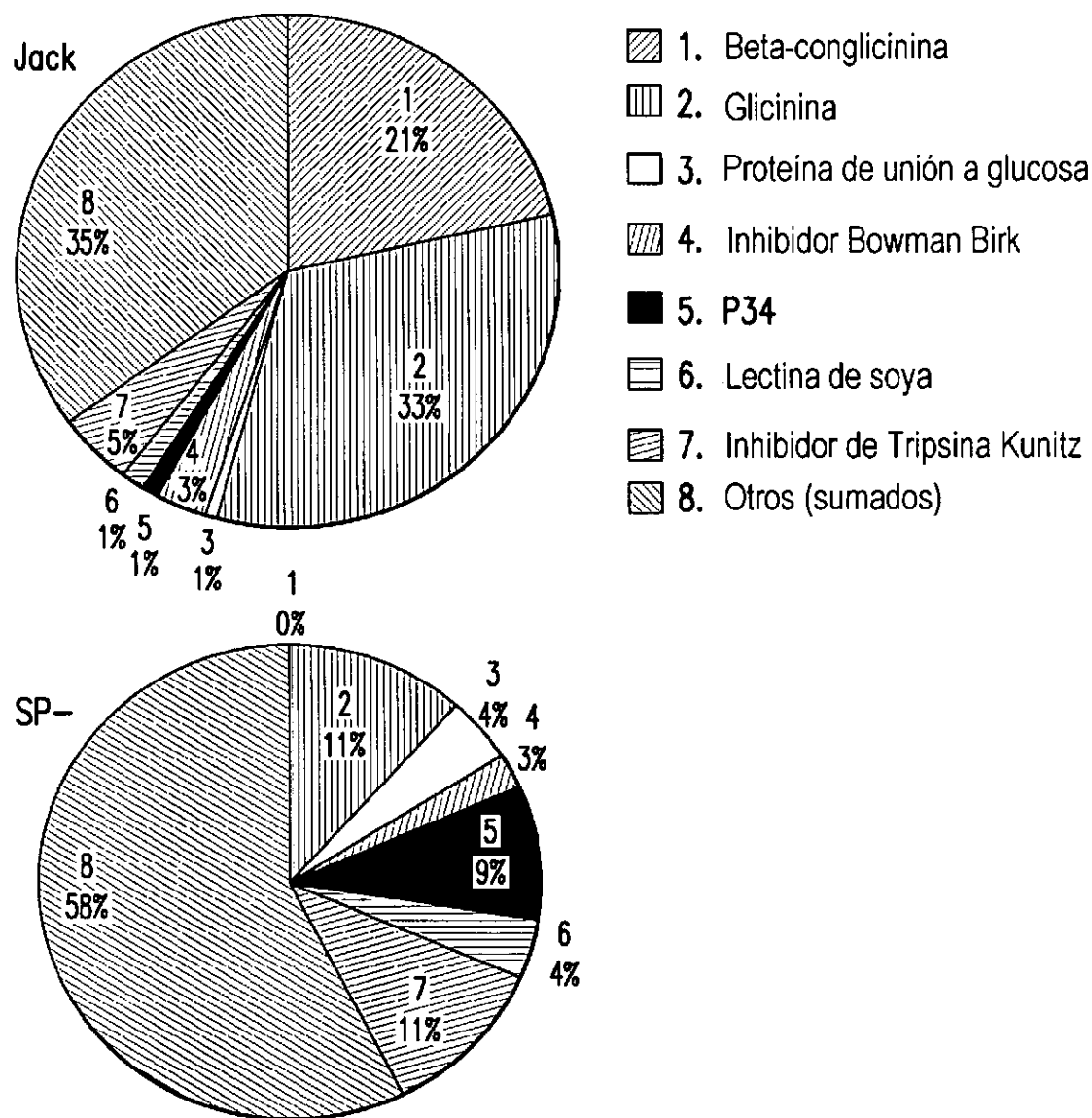


FIG.6

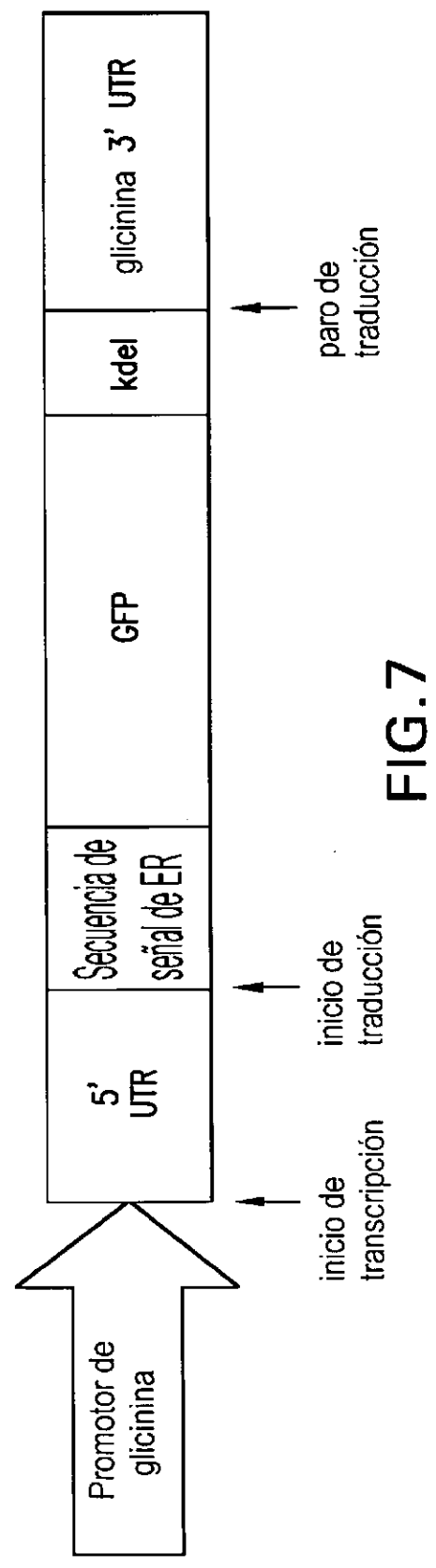


FIG. 7

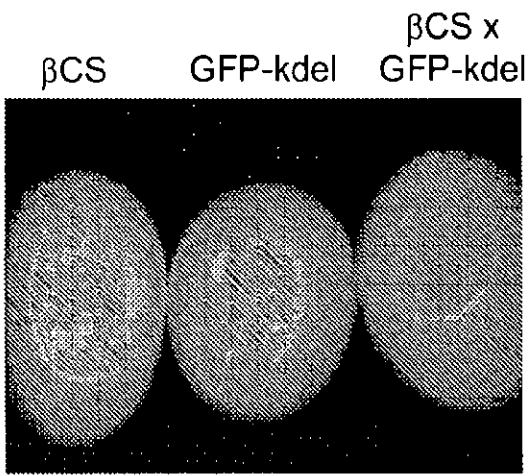


FIG.8A

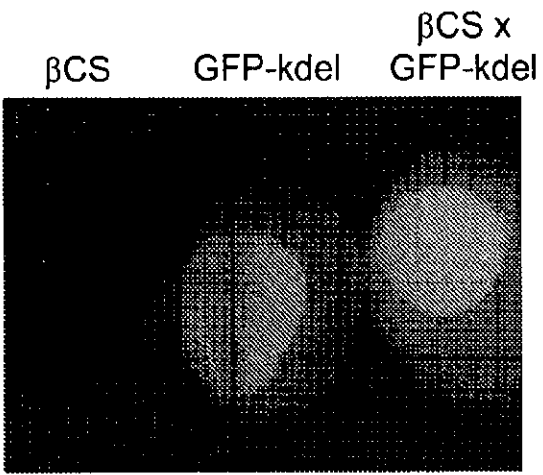


FIG.8B

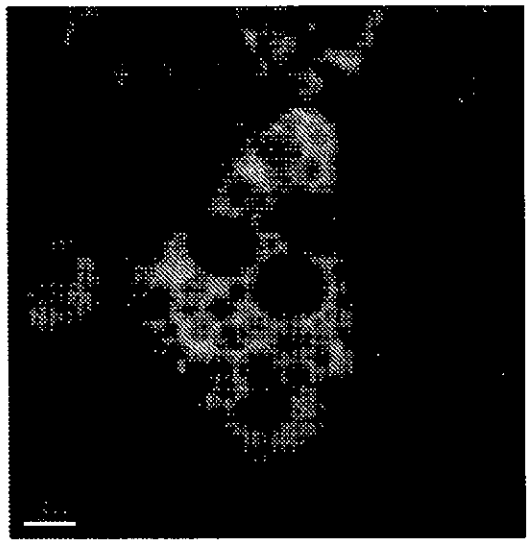


FIG.8C

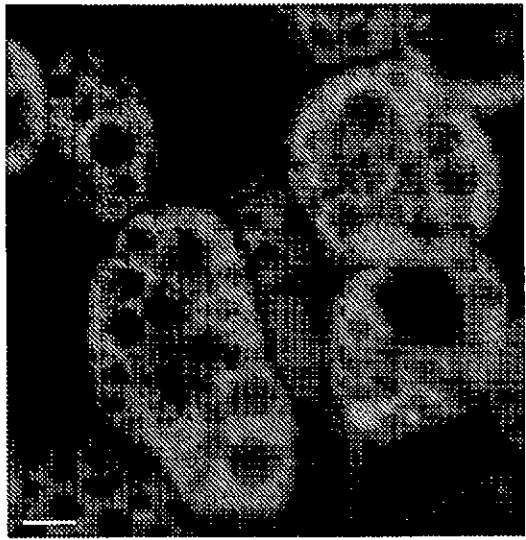


FIG.8D

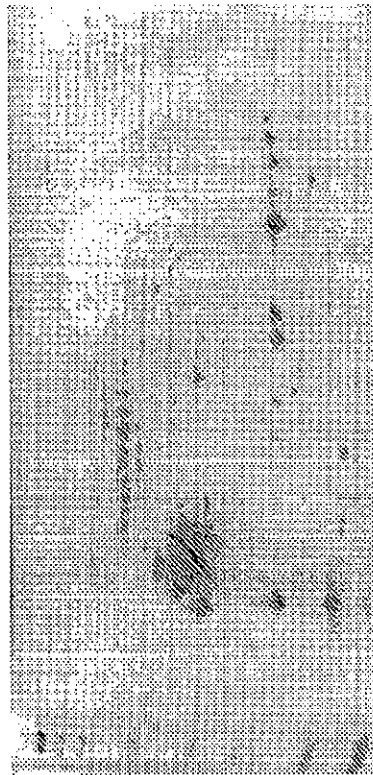


FIG. 9A

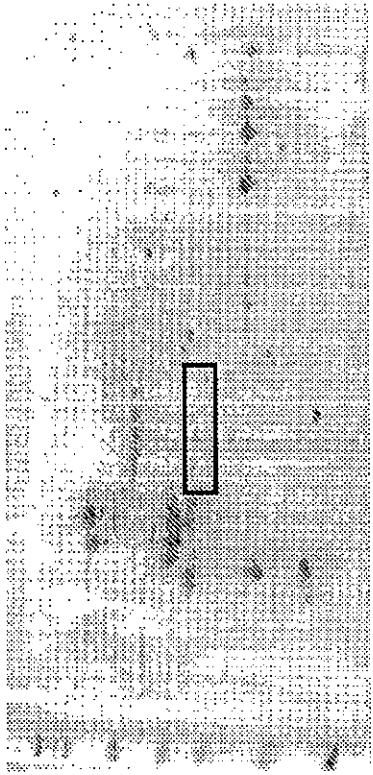


FIG. 9B

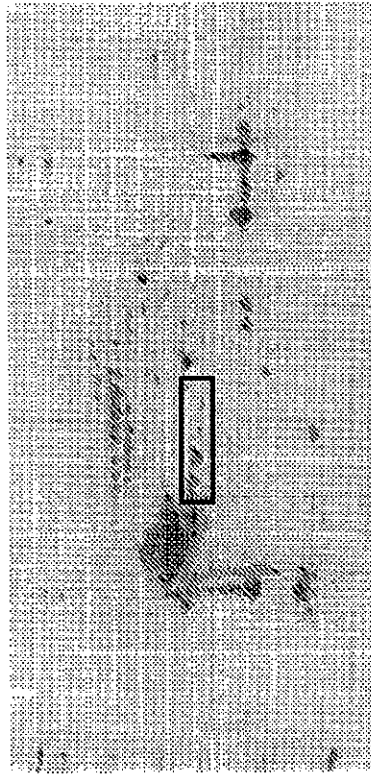


FIG. 9C

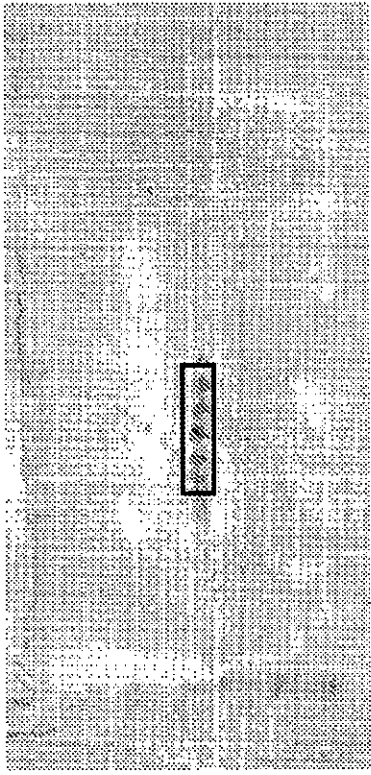


FIG. 9D

10/15

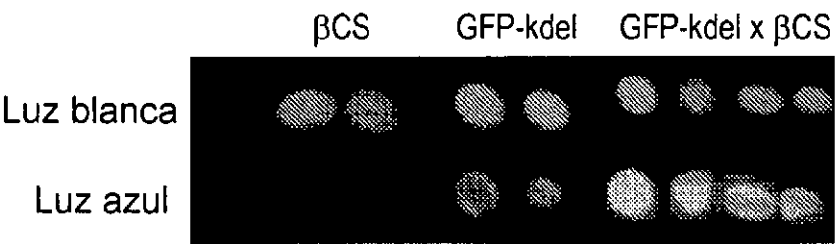


FIG.10A

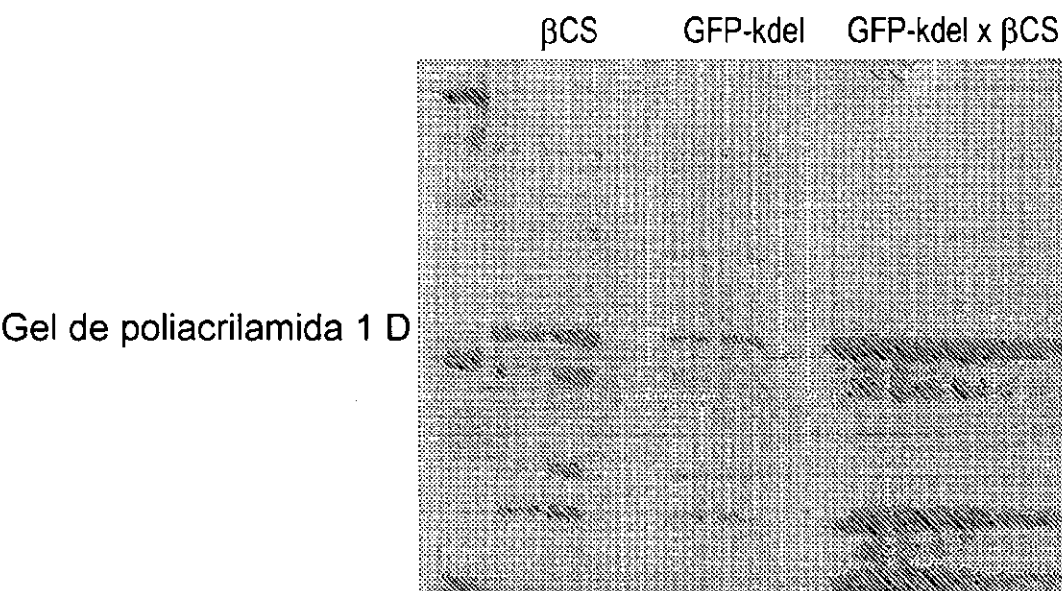


FIG.10B

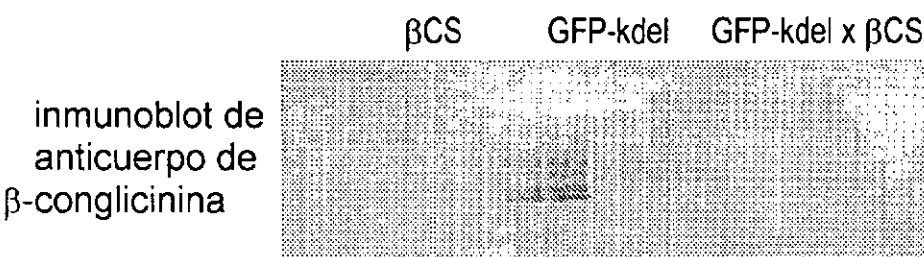


FIG.10C

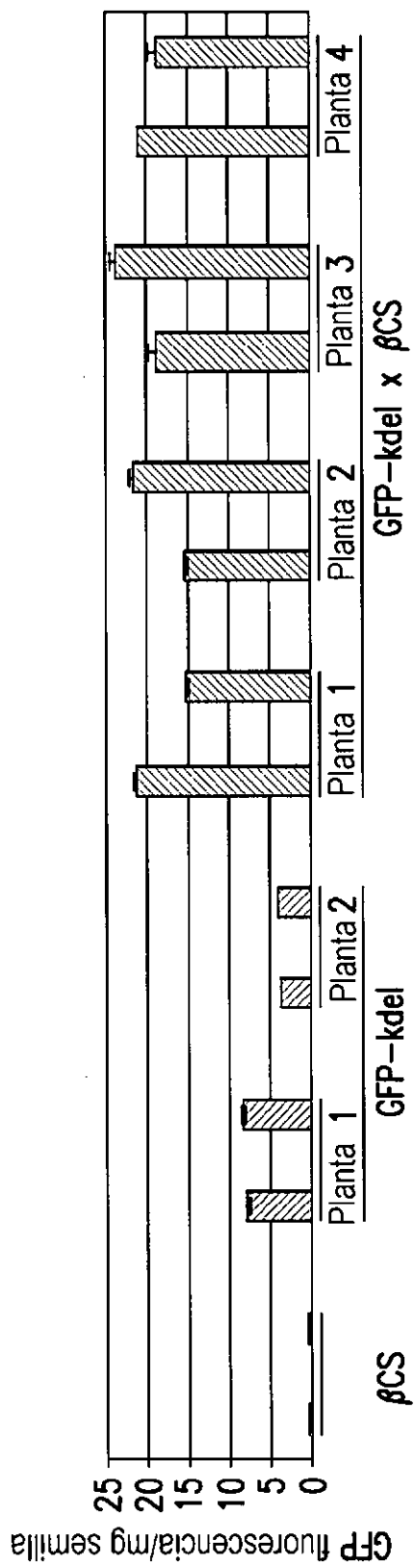


FIG.11

12/15

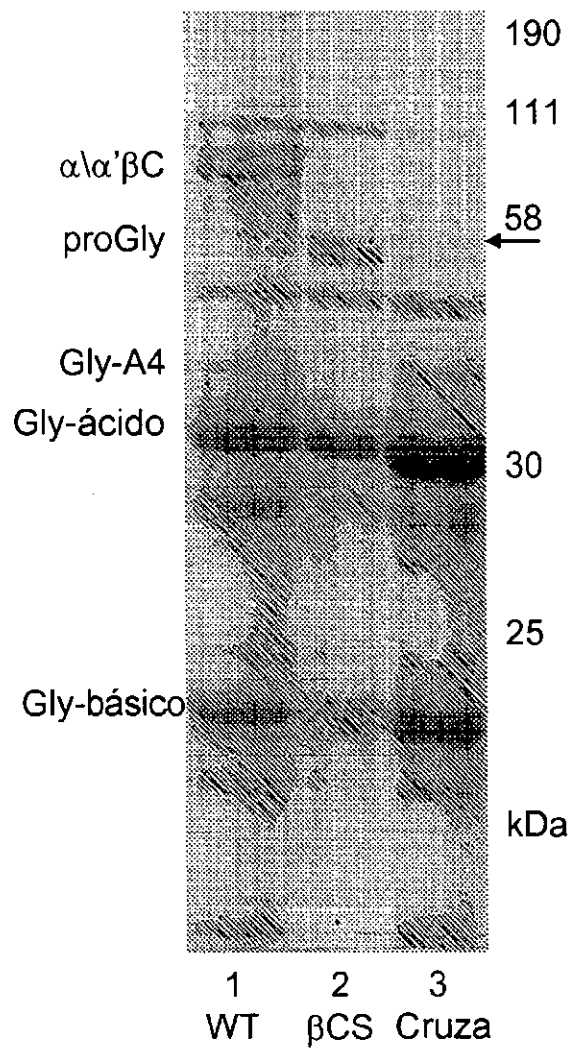


FIG.12

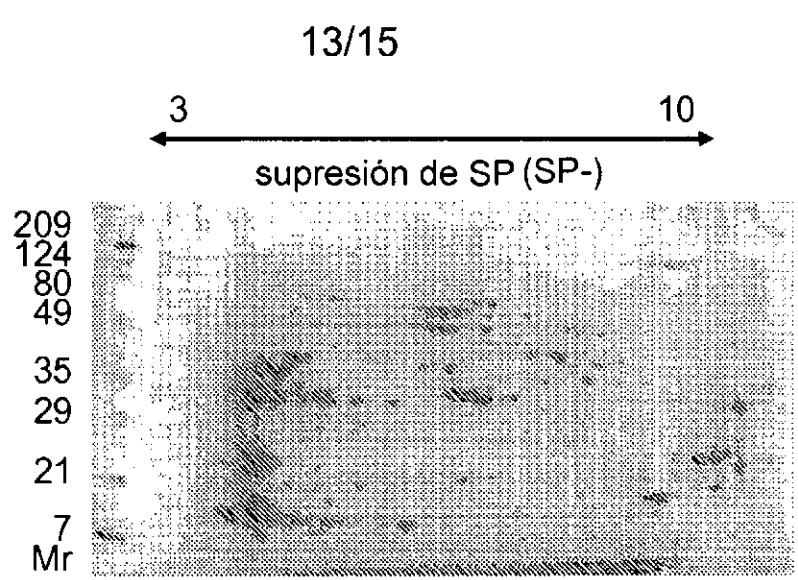


FIG.13A

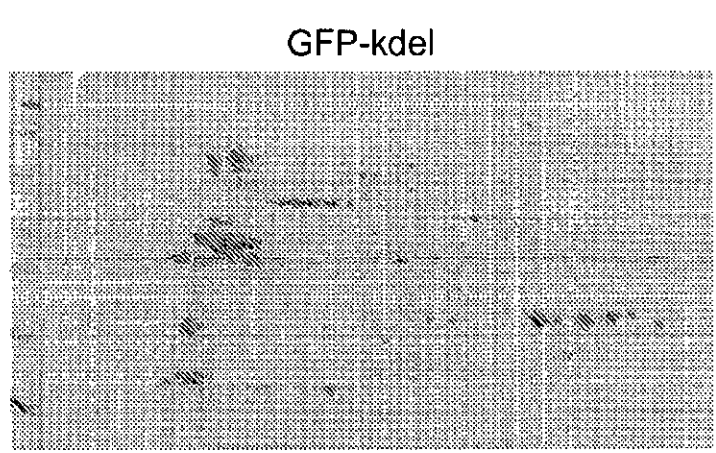


FIG.13B

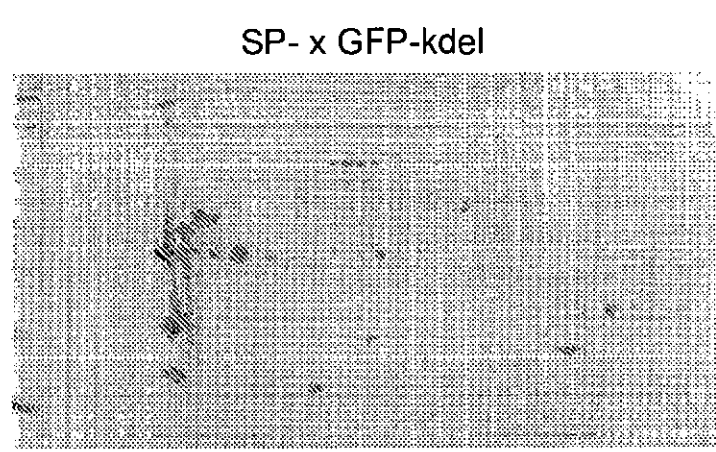


FIG.13C

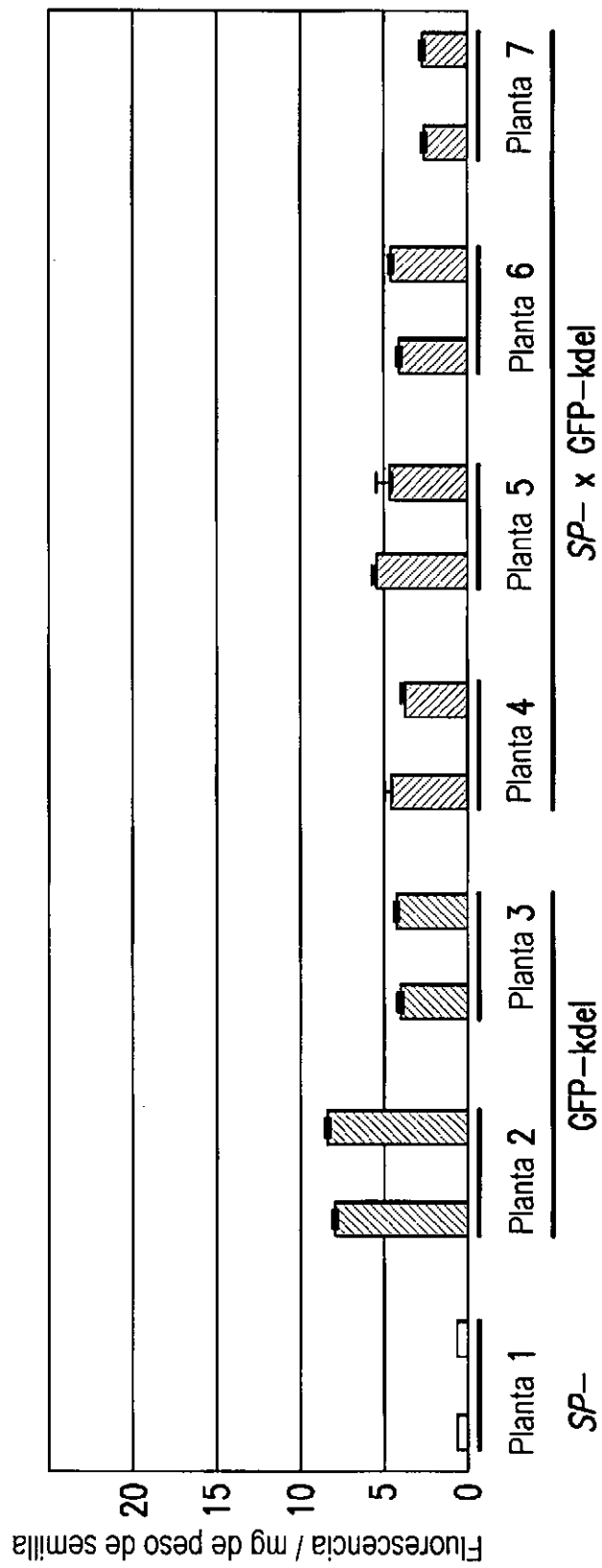


FIG.14

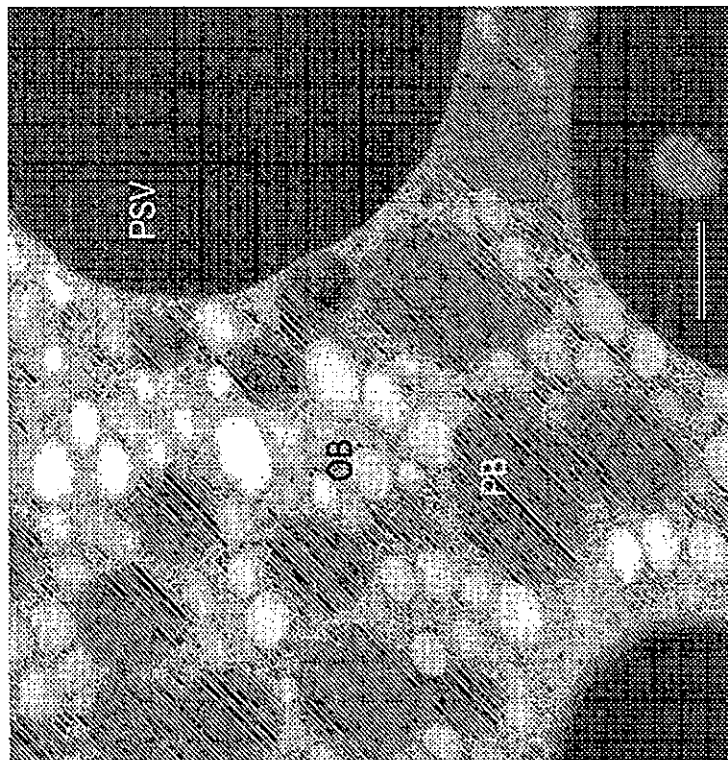


FIG.15A

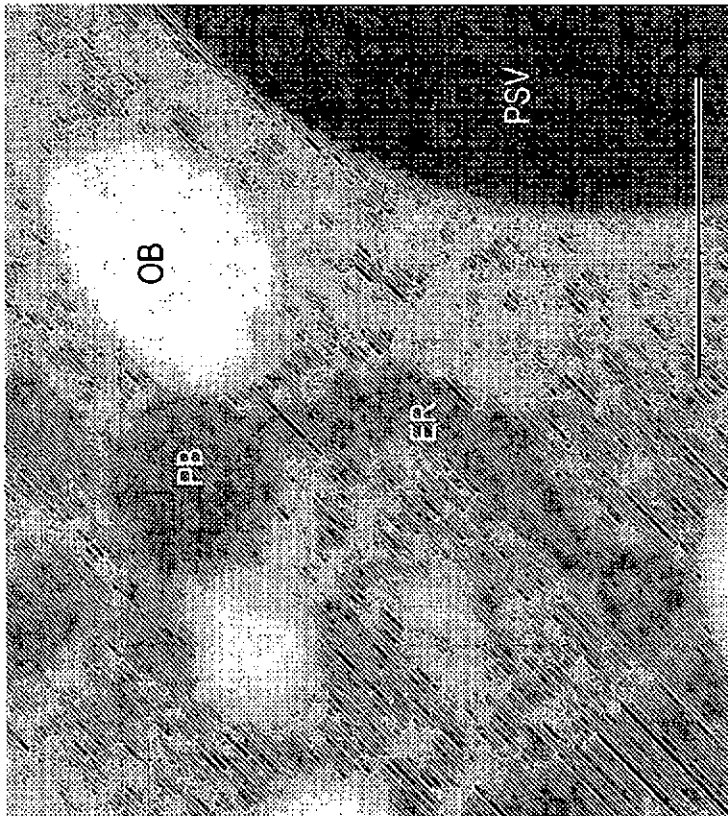


FIG.15B

<110> The Donald Danforth Plant Science Center
Departamento de Agricultura de los Estados Unidos

<120> PRODUCCIÓN Y ALMACENAMIENTO DE PROTEÍNAS MEJORADOS EN PLANTAS

<130> AGRIUS-228001-PCT

<140>

<141>

5 <150> 61/076,616
<151> 2008-06-28

<160> 57

<170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 632

<212> ADN

<213> Glycine max (soybean)

10

<220>

<223> promotor/región reguladora 5' de Glicinina

<400> 1

caaaacaaat taataaaaca cttacaacac cggatTTTTT ttaattaaaa tgtgccattt 60
aggataaata gttaatatTT ttaataatta tttaaaaagc cgtatctact aaaatgattt 120
ttatttggtt gaaaatatta atatgtTTaa atcaacacaa tctatcaaaa ttaactaaa 180
aaaaaaataa gtgtacgtgg ttaacattag tacagtaata taagaggaaa atgagaaatt 240
aagaaattga aagcgagtct aattttttaa ttatgaacct gcatatataa aaggaaagaa 300
agaatccagg aagaaaagaa atgaaacat gcatgggtccc ctggtcatca cgagtttctg 360
ccatttgcaa tagaaacact gaaacacctt tctctttgtc acttaattga gatgccgaag 420
ccacctcaca ccatgaactt catgaggtgt agcaccacag gcttccatag ccatgcatac 480
tgaagaatgt ctcaagctca gcacctact tctgtgacgt gtccctcatt caccttctc 540
tcttccctat aaataaccac gcctcagggt ctccgcttca caactcaaac attctctcca 600
ttgggtcctta aacactcacc agtcatcacc gc 632

15

<210> 2

<211> 632

<212> ADN

<213> Glycine max (soybean)

<220>

<223> Gen Gyl de soya para subunidad G1 de glicinina (No. reg. X15121.1)

20

<400> 2

caaaacaaat taataaaaca cttacaacac cggatTTTTT ttaattaaaa tgtgccattt 60
aggataaata gttaatatTT ttaataatta tttaaaaagc cgtatctact aaaatgattt 120
ttatttggtt gaaaatatta atatgtTTaa atcaacacaa tctatcaaaa ttaactaaa 180
aaaaaaataa gtgtacgtgg ttaacattag tacagtaata taagaggaaa atgagaaatt 240
aagaaattga aagcgagtct aattttttaa ttatgaacct gcatatataa aaggaaagaa 300
agaatccagg aagaaaagaa atgaaacat gcatgggtccc ctggtcatca cgagtttctg 360
ccatttgcaa tagaaacact gaaacacctt tctctttgtc acttaattga gatgccgaag 420
ccacctcaca ccatgaactt catgaggtgt agcaccacag gcttccatag ccatgcatac 480
tgaagaatgt ctcaagctca gcacctact tctgtgacgt gtccctcatt caccttctc 540
ctcttcccta taaataacca cgctcagggt tctccgcttc acaactcaaa cattctctc 600
cattgggtcct taaacactca tcagtcaccc cc 632

25

<210> 3
 <211> 962
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> promotor/región reguladora 5' de Conglicinina

<400> 3
 5 gttttcaaatt ttgaatttta atgtgtgttg taagtataaa tttaaaataa aaataaaaaac 60
 aattattata tcaaaatggc aaaaacattt aatagctatt atttattaaa aaaatatgta 120
 ataatatatt tatattttta tatctattct tatgtatttt ttaaaaatct attatatatt 180
 gatcaactaa aatattttta tatctacact tattttgcat ttttatcaat tttcttgctg 240
 tttttggcat atttaataat gactattctt taataatcaa tcattattct tacatggtag 300
 atattgttgg aaccatatga agtgttcatt gcatttgact atgtggatag tgttttgatc 360
 catgcccttc atttgccgct attaattaat ttggtaacag attcgttcta atcagttact 420
 taatccttcc tcatcataat taatctggtg gtgcgaatgc cataatattg attagttttt 480
 tggaccataa gaaaaagcca aggaacaaaa gaagacaaaa cacaatgaga gtatcctttg 540
 catagcaatg tctaagttca taaaattcaa acaaaaacgc aatcacacac agtggacatc 600
 acttatccac tagctgatca ggatcgccgc gtcaagaaaa aaaaactgga ccccaaaaagc 660
 catgcacaac aacacgtact cacaagggcg tcaatcgagc agcccaaaac attcaccaac 720
 tcaacccatc atgagccac acatttgttg tttctaacc aacctcaaac tcgtattctc 780
 10 ttcgccacc tcatttttgt ttatttcaac acccgtaaaa ctgcatcca ccccgtaggc 840
 aaatgttcat gcatgttaac aagacctatg actataaata tctgcaatct cggcccaagt 900
 tttcatcatc aagaaccagt tcaatatcct agtacgccgt attaagaat ttaagatata 960
 ct 962

<210> 4
 <211> 1088
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> promotor/región reguladora 5' de KTI

15 <220>
 <221> característica_misc
 <222> 175
 <223> n = A,T,C o G

<400> 4
 20 attgatactg ataaaaaaat atcatgtgct ttctggactg atgatgcagt atacttttga 60
 cattgcccctt attttatttt tcagaaaagc tttcttagtt ctgggttctt cattatttgt 120
 ttcccatctc cattgtgaat tgaatcattt gcttcgtgtc acaatacat ttagnatagg 180
 acatgcattg gtcagattca cggtttatta tgtcatgact taagttcatg gtagtacatt 240
 acctgccacg catgcattat attggttaga ttgataggc aaatttgggt gtcaacaata 300
 taaatataaa taatgttttt atattatgaa ataacagtga tcaaaacaaa cagttttatc 360
 tttattaaca agattttgtt ttgtttgat gacgtttttt aatgtttacg ctttccccct 420
 tcttttgaat ttagaacact ttatcatcat aaaatcaaat actaaaaaaa ttacatattt 480
 cataaataat aacacaaaata tttttaaaaa atctgaaata ataatgaaca atattacata 540
 ttatcacgaa aattcattaa taaaaatatt atataaataa aatgtaatag tagttatatg 600
 taggtttttt gtactgcacg cataatatat acaaaaagat taaaatgaac tattataaat 660
 aataacacta aattaatggt gaatcatatc aaaataatga aaaagtaaat aaaatttgta 720
 attaacttct atatgtatta cacacacaaa taataaataa tagtaaaaaa aattatgata 780
 aatatttacc atctcataaa gatattttaa ataatgataa aaatatagat tattttttat 840
 gcaactagct agccaaaaag agaacacggg tatatataaa aagagtacct ttaaattcta 900
 ctgtacttcc ttatttcttg acgtttttat atcaagtgga catacgtgaa gatttttaatt 960
 atcagtctaa atatttcatt agcacttaat acttttctgt tttattccta tcctataagt 1020
 25 agtcccgatt ctcccaacat tgcttattca cacaactaac taagaaagtc ttccatagcc 1080
 ccccaaaa 1088

117

<210> 5
 <211> 966
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> promotor/región reguladora 5' de LE

<400> 5
 5 aatgccatcg tatcggtgtca caatggaata cagcaatgaa caaatgctat cctcttgaga 60
 aaagtgaat gcagcagcag cagcagacta gagtgtctaca aatgcttata ctcttgagaa 120
 aagtgaatg cagcggcagc agacctgagt gctatataca attagacaca gggctctatta 180
 attgaaattg tcttattatt aaatatctcg ttttatatta attttttaaa ttttaattaa 240
 atttatatat atttatatta agacagatat atttatttgt gattataaat gtgtcacttt 300
 ttcttttagt ccatgtattc ttctattttt tcaatttaac tttttatttt tatttttaag 360
 tcactctgat caagaaaaca ttgttgacat aaaactatta acataaaatt atgttaacat 420
 gtgataacat catattttac taatataacg tcgcatttta acgttttttt aacaaatata 480
 gactgtaaga gtaaaaatga aatgtttgaa aagggttaatt gcatactaac tatttttttt 540
 cctataagta atcttttttg ggatcaattg tatatcattg agatacgata ttaaatatgg 600
 gtaccttttc acaaaaccta ccttggttag tcaaacacata cataagagag gatggattta 660
 aaccagtcag caccgtaagt atatagttaa gaaggctgat aacacactct attattgtta 720
 10 gtacgtacgt atttcctttt ttgttttagt tttgaattta attaatataa atatatatgc 780
 taacaacatt aaatttttaa ttacgtcta attatatatt gtgatgtata ataaattgtc 840
 aaccttttaa aattataaaa gaaatattaa ttttgataaa caacttttga aaagtaccca 900
 ataagtctag tataaatagg ggcattgact cccattgcat acagtgaat ttactggaag 960
 caaagc 966

<210> 6
 <211> 1369
 <212> ADN
 <213> Glycine max (soya)

<220>
 15 <223> promotor/región reguladora 5' de P34

<400> 6
 atggataaaa aatctagcat tctctctttt ctactagca tattaaatta acgatccaga 60
 aatatttata aatatttttt taatgcttaa tgactcaata cagggccagt caaggtaaac 120
 cttggtcgga taaacaaccc tcataacatg acatacaatt ataattgaaa attctcatat 180
 agcacaatta tgaaggcaaa aacatggcac acaaaaggta ctgtgtttta ctataaacga 240
 ttatgaattt ttcagggaat atagcatgtt cgtcttgatt ctcttgaaga acttttttagg 300
 taccctttac taagttggac gtgattttgt ctatgttcag gagaaaataa aggataaaat 360
 gtcttttgtg gacatgattt tgattgtgat ttcatgttaa gggagtaagg atgatgtagt 420
 tttatgtaga gttttaggtt cttggatctt ttcoattaat gaagtcattt gcttcttgaa 480
 gatcaatgac aacaaaatga agaagtaaga aagggtgatt gagacctcat ttttaagaaa 540
 20 aaatgagtca agaataagct taccaccata ggaagtcatt aataagagct tgaagtaag 600
 agaagatgag tggagggaga gggagagaga tgacacaaaa ttatgcctc aaataaggctc 660
 tgaacattga agtctaattt cttaaatgat caaaattgaa aaaatacaca cacaagacct 720
 ctatagttta agtgtcattc aaaattggag aaaaatttag atttctattc aaatttcact 780
 tgaatttgaa tttatagagt caaattttga gtcaaaattt catlaattat aatcagtga 840
 tttcaactat ggtttagtct gctaattcaa tatcaagtc aaagttcttc actaagtgtg 900
 cttaggtgtc atgaggcatg taaaatataa aggacatgta caaagtatga ccatatgatg 960
 tgacaatgag gtgtaacaag caaatgctca ccttcccttt aggtgtgtcc aaaattta 1020
 tggattgagc ttctcccaat tcaattaaat ttctttttta acacacacat caaatagtgc 1080
 actgaatgca cgtgaaatta caaaactatc tcaaatataa aaactagtct aggtgtccta 1140
 aaatacaaaag actgaaaaat cctatatatt cagagtaccc tctttacact atggagtcc 1200
 25 aaatacaaga ctcaaaaata atgaaatcct aatataatat atgtacaaaa ataagtggat 1260
 tcatacttgg tctattgatc gaaaatctac cttaaggctc atgagaatcc taagggtctt 1320
 tcctgcatct ctgactcaat cttttaagtc tccaacctg acttttgta 1369

<210> 7
 <211> 2262
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> promotor/región reguladora 5' de GBP

<400> 7

5	tggaatattta agtcttctat aatatttcat ttagagccag aagccagggt caaaggaata 60
	ggtaattcac atgaattcat tctcttggtt ctatacagct attatttttc catcttagtg 120
	ttgcaggaaa ctacctcagt tgtgttagat gtgcaaaact tgtatggata tatatactgt 180
	tcagtgttgg gaaacccatg ctttcttaac tcacagagat acattttaaac tttttttaga 240
	aacttgctta gtatcttata ctgtttgtta ttcatttttg gcagttgggc cttaaagatac 300
	tcctatgaat cttgtgctag agaagactta cgtatgctaaa acaggacggg gcatgcctga 360
	actttaagga gacgttgccc tgttccactt ccaattaggt aactgotatc gtgatgaaca 420
	aaaaattggg gagtttatca ccttgccctt tgccatgatt caattaaaag cgtgtttgga 480
	ctttggaacc tcattctaac accaccctat gatgggttag acgcaaaatc tagactgggt 540
	agtgtttaac gtgtatctgt gtgaacacag ttacaaacgc attccttggt taatgtctacc 600
	atgcctagga gttgaatcat ttgtaacttt accaatttag tcattactac tagcattctt 660
	ttccctattc aagttagtgt tagctccagt tagtgatggg catttctctc tataaacttt 720
	aattgttaga tgagtggag aggaacctgt ttgattgtta tggttctagt tctagtgtatt 780
10	tttattaatt gggttcgacc atattagtgt ttgatttgag ctatagatag ttttttcccc 840
	aaaagatcag tcttctctca tgcagattc atgggttggg actcttttta tccagttcca 900
	acaaaacttg tgttcgaact acgaagtcag tcttacttat tgggtaacat gtgggttttg 960
	gtgtttaatg gatctagaat actgtttgta gctaaacctt tcttatcata aagggcctaa 1020
	aaagtaaaat tggttattac atttggaataa aaagaaataa tctaggccca ctggcacact 1080
	gagaaacgtt ttcaatgaat aatttaatag ttttttttta taaaaaataa ttaataaaaa 1140
	ataatggagt ttttaaaaaa attacaacaa tctgtttctc taagggtttt taatagtcca 1200
	gataattcat agcttagagc aatacgacat ggttaggaag cataaaaaaa atatacgaca 1260
	tggttaggaa ttttttttta gtatgtctga cataattttt taaatgtttt ggcttcatat 1320
	gaatttaaca gtgcgtcata tgaacttaca cactcattat attttttaac cttttaaatg 1380
	atttttaaaa aaatatgaca gatgcaatct tattttcact ttttatactt tcaactactg 1440
	ttcatatgac ctaaagtcag agaaatattt taaaaagata aatacgataa agaatacgat 1500
15	gagaaagaaa cctcacacaa tgaatagacc aaattagacc tattttattt ccttagaaat 1560
	aaagaaaata attatttttt attttttcac attacattta tatttttcta tcaactttctc 1620
	tatttaggta ttgattggca tatgagtgtt catgaatttt tttttttaaa aaaagcgtaa 1680
	atattaatta tattcatgca ttatttgttt tctgtctttc attttctatt taatcttacg 1740
	ttatcaataa ttattatta aattttatag ttgatgatga atataaaga gatataaata 1800
	aaaaaaataa ttaattttat aataaaaatt aaaaaataa ttttttgaga taaaaaattt 1860
	taagagaaca attataaacg gagagtatta tatttagttt tatgtaccgt gtacgtgtct 1920
	actaacatgg tgtctctcca tcattttcgt aggaaaaaac attataggag tatgaaaaaa 1980
	gcaaaagttt tgtctgttta tggttttgta tataccagc tctacttggc agcaattacc 2040
	cgtcttgctt gctacttacg agacacgtac attaacactt gtccatgcta gtgcacgaa 2100
	ttgccacccc attcatcact cctccctttt ccttctcttt atatttatat atatatatat 2160
	atatatatat atatatatat atatatatat atatatatac aagcacaatg catcatctca 2220
20	aagaaattaa gagagttttt tttgttcttc actgaccaag cc 2262

<210> 8
 <211> 1387
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> promotor/región reguladora 5' de SMP

<400> 8

25	gtttcacatg atccttcatt ctgtgtggct taggagactc aacttcagag tccgtgatga 60
	tcaatgactc ttaagttgtg acttatggct tcatgtttta taaattactt catagacaat 120
	gatgcccttt cattattcca tcccaaatta actaatgttt aagattttct ttacacaaaa 180
	ctagaaaaata aatatttttaa gagaattaat atttagttgg gggataattt taattattag 240

aatattcttg ttttctctct taatttcaat aaatatatta gtaaattttt taaaacaaca 300
 ttatgtatat atatatgtga aaattagaat aaatattttg aatatttcaa tatcaacca 360
 tttaaatatt tttttaagg ttaaaattat agtcattttg gaacatagac agtaatatgg 420
 agctagagtt atgttaaata caggtcagaa aaataaaaact catgaaaatt tgtaaaccaa 480
 cgaaacttac aataagttca gcagtgatct tttgtctcat ttttttacat ttcttttagtc 540
 tctttatttt ttgggttaaac gttgcttttt agtttattat gttttagtaa ttttttttta 600
 ctttattaca ttaaatttta aaatttctat tttgtttttt ttatgttttt gaaaatttta 660
 ttttggttaat ttttttcata agaacactaa tacttagaaa agaataaaaa aaaaaaagga 720
 aagttcgatg gaattcaagc tcatgaaatt ttatgttaaa aacatgagta attcattaat 780
 atttttaa ataaataaaa aaaaaaagg aaagttcgat ggaattcaag ctcatgaaat 840
 5 tttatgttaa aaacatgagt aattcattaa ttttttaaaa tttaaaataa ttaattgttt 900
 tacttatatt aaatagttgg tgaaaattta aaaaaagatg cacgttttaa caaggttgga 960
 atcgttttga ttttaatttc accactcgag tgggatgcac atttagacaa ggatggaatc 1020
 gtttgacttg gatttggtta ctcttcccc caacacgctg ccaccttcta gggaaggtaa 1080
 gggaccgagt gaccgataac aacaccgatt tccaaatata tatctcattc ctaagctcac 1140
 acacatcttt tacgttacat ttcatatta gatgctttca atcgttaaga cacatgtcac 1200
 caccaaaaga gcatttcata acaacgtgtc acacctccca agcacacgtg tcaactcaca 1260
 cacaaccctc acctatatat aaattatcaa accctccttc cattcctcca catctcaatc 1320
 tcaatatcta cacaaaagtg ttccacttga gtgaaaagta gtgtgttaag aactaaacaa 1380
 tttttca 1387

<210> 9
 <211> 24
 10 <212> PRT
 <213> Arabidopsis thaliana

<220>
 <223> Secuencia de señal ER de un gen anhidrasa carbónica alfa (NP_850685)

<400> 9
 Met Lys Ile Met Met Met Ile Lys Leu Cys Phe Phe Ser Met Ser Leu
 1 5 10 15
 Ile Cys Ile Ala Pro Ala Asp Ala
 20

15 <210> 10
 <211> 24
 <212> PRT
 <213> Oryza sativa Japonica

<220>
 <223> Secuencia de señal ER de una proteína sin nombre (BAG87020)

<400> 10
 Met Ala Ala Ser His Gly Asn Ala Ile Phe Val Leu Leu Leu Cys Thr
 1 5 10 15
 20 Leu Phe Leu Pro Ser Leu Ala Cys
 20

<210> 11
 <211> 24
 <212> PRT
 <213> Arabidopsis thaliana

<220>
 <223> Secuencia de señal ER de riboforina I (AT2G01720)

25 <400> 11
 Met Ala Ala Arg Ile Gly Ile Phe Ser Val Phe Val Ala Val Leu Leu

1 5 10 15
 Ser Ile Ser Ala Phe Ser Ser Ala
 20

<210> 12
 <211> 21
 <212> PRT
 <213> Arabidopsis thaliana

5 <220>
 <223> Secuencia de señal ER de quitinasa básica de A. thaliana

<400> 12
 Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
 1 5 10 15
 Leu Ser Ser Ala Glu
 20

<210> 13
 <211> 447
 <212> ADN
 <213> Glycine max (soya)

10 <220>
 <223> Terminador (TED 3-prima) de Glicinina

<400> 13
 agcccttttt gtatgtgcta cccacttttt gtcttttttg caatagtgtc agcaaccaat 60
 aaataataat aataataatg aataagaaaa caaaggcttt agcttgccct ttgttcactg 120
 taaaataata atgtaagtac tctctataat gagtcacgaa acttttgccg gaataaaaagg 180
 agaaattcca atgagttttc tgtcaaatct tcttttgtct ctctctctct ctcttttttt 240
 tttttctttc ttctgagctt cttgcaaaac aaaaggcaaa caataacgat tggccaatg 300
 atagttagct tgatcgatga tatcttttag aagtgttgag aggacaggac atgatgtaga 360
 agactaaaat tgaaagtatt gcagacccaa tagttgaaga ttaactttaa gaatgaagac 420
 gtcttatcag gttcttcatt acttggga 447

15 <210> 14
 <211> 445
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> Terminador (TED 3-prima) de gen Gyl de soya para subunidad
 G1 de glicinina (No. Registro X15121.1)

<400> 14
 agcccttttt gtatgtgcta cccacttttt gtcttttttg caatagtgtc agcaaccaat 60
 aaataataat aataataatg aataagaaaa caaaggcttt agcttgccct ttgttcactg 120
 taaaataata atgtaagtac tctctataat gagtcacgaa acttttgccg gaataaaaagg 180
 agaaattcca atgagttttc tgtcaaatct tcttttgtct ctctctctct ctcttttttt 240
 tttctttctt ctgagcttct tgcaaaacaa aaggcaaaaca ataacgattg gtccaatgat 300
 agtttagctg atcgatgata tcttttaggaa gtgttgagcag gacaggacat gatgtagaag 360
 actaaaattg aaagtattgc agacccaata gttgaagatt aactttaaga atgaagacgt 420
 cttatcaggt tcttcattgac ttgga 445

<210> 15
 <211> 133
 <212> ADN
 <213> Glycine max (soya)

25

<220>
 <223> Terminador (TED 3-prima) de Beta conglucininina

 <400> 15
 aataagtatg tagtactaaa atgtatgctg taatagctca tagtgagcga ggaaagtatc 60
 gggctatttta actatgactt gagctccatc tatgaataaa taaatcagca tatgatgctt 120
 ttgttttgtg tac 133

 <210> 16
 <211> 132
 <212> ADN
 <213> Glycine max (soya)

 <220>
 <223> Terminador (TED 3-prima) de gen (alfa-prima-bcsp) de proteina
 de almacenamiento de Beta-conglucininina (No. de Registro M13759.1)

 <400> 16
 ataagtatgt agtactaaaa tgtatgctgt aatagctcat agtgagcgag gaaagtatcg 60
 ggctatttta ctatgacttg agctccatct atgaataaat aaatcagcat atgatgctt 120
 tgttttgtgt ac 132

 <210> 17
 <211> 188
 <212> ADN
 <213> Glycine max (soya)

 <220>
 <223> Terminador (TED 3-prima) de KTI

 <400> 17
 gacacaagtg tgagagtact aaataaatgc tttggttgta cgaaatcatt aactaaata 60
 aaataatcaa agcttatata tgccttctaa ggccgaatgc aaagaaattg gttcttctcg 120
 ttatctttgc cactttacta gtacgtatta attactactt aatcatcttg ttacggctca 180
 ttatatcc 188.

 <210> 18
 <211> 325
 <212> ADN
 <213> Glycine max (soya)

 <220>
 <223> Terminador (TED 3-prima) de LE

 <400> 18
 atgtgacaga tcgaaggaag aaagtgtaat aagacgactc tcactactcg atcgctagtg 60
 attgtcattg ttatatataa taatgttatc tttcacaact tatcgtaatg cattgtgaaa 120
 ctataacaca tttaatccta cttgtcatat gataacactc tccccattta aaactcttgt 180
 caatttaaag atataagatt ctttaaataa ttataaaaaa tatattataa attcaatcac 240
 tcctactaat aaattattaa ttaatatatta ttgattaaaa aaatacttat actaatttag 300
 tctgaataga ataattagat tctag 325

 <210> 19
 <211> 213
 <212> ADN
 <213> Glycine max (soya)

 <220>
 <223> Terminador (TED 3-prima) de P34

 <400> 19

172

gccgtaaagg ttcaatacaa cgagtgttg tttcttagg gacaagcatt gtacttatgt 60
 atgattctgt gtaaccatga gtcttcacg ttgtactaat gtgaaggga aaaataaaac 120
 acagaacaag ttctgttttc tcaaataatg tgaaggtaga aaatggaacc atgcctcttc 180
 tcttgcattg gatttaaaat attagcagat ggt 213

<210> 20
 <211> 246
 <212> ADN
 <213> Glycine max (soya)

5 <220>
 <223> Terminador (TED 3-prima) de GBP

<400> 20
 gaggtttaga acaatcaaga aaagggtgtg atgtggctga agatcacggg gaatgtatta 60
 agcttcagag actctttaa tttaattttc tgtatattgt gttatatgtt actagttctt 120
 taaattagcc agatggagtt tatgtgtatc taaatgcagg gatgctaag gaataaaatg 180
 gccacttgta ttgttagcta tctcttatgg tagcagaata agacgtaaac tggttctttg 240
 ctccaa 246

<210> 21
 <211> 475
 <212> ADN
 <213> Glycine max (soya)

10 <220>
 <223> Terminador (TED 3-prima) de SMP

<400> 21
 ttaaaacgtg atctatgata caacaatatt agtatatata gacgcatgca gtttatatag 60
 tatatatgt catgttgat gtttttacat ttgggttggt ttgtttacat tctcttcaaa 120
 aaaaaaaaaa tgttagtac gtgtaagggt ttgaagattg gttctaggct ccgtgggaac 180
 catttcaaca ataaacattt tgcgcgttct tgtacacgta gtgatgagaa gagatgcctt 240
 atgggcagta tcatctaaaa cttaattttc tccatcatag aatttggatc tattggactg 300
 gactgaactg aactgaatga tcttttttc ttttttaatt tcattcacta acaaatacat 360
 aaaacaccag atattaactt agccagtatg aattttaact attttgtcta atgctatgac 420
 ttatcactgt ctgtatcatc tttaattctt ttttcatatt atttatatta aataa 475

15 <210> 22
 <211> 18
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia de señal de retención en ER (marcador de retención codifica KDEL seguido por 2 codones de paro)

20 <400> 22
 aaggatgaac tttaatga 18

<210> 23
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Marcador de retención en ER - KDEL

25 <400> 23
 Lys Asp Glu Leu
 1

<210> 24
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> Secuencia de señal de retención en ER (marcador de
 retención codifica KHDEL seguido por 2 codones de paro)

 5 <400> 24
 aagcatgatg aactttaatg a 21

 <210> 25
 <211> 5
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Marcador de retención en ER - KHDEL

 10 <400> 25
 Lys His Asp Glu Leu
 1 5

 <210> 26
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Secuencia de señal de retención en ER (marcador de
 retención codifica HDEL)

 15 <400> 26
 His Asp Glu Leu
 1

 <210> 27
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Marcador de retención en ER - KEEL

 20 <400> 27
 Lys Glu Glu Leu
 1

 <210> 28
 <211> 6
 <212> PRT
 <213> Secuencia Artificial

 <220>
 25 <223> Marcador de retención en ER - SEKDEL

124

<400> 28
Ser Glu Lys Asp Glu Leu
1 5

<210> 29
<211> 6
<212> PRT
<213> Secuencia Artificial

5 <220>
<223> Marcador de retención en ER - SEHDEL

<400> 29
Ser Glu His Asp Glu Leu
1 5

<210> 30
<211> 920
<212> ADN
<213> Solanum tuberosum

10 <220>
<223> promotor/región reguladora 5' de Ubiquitina 3

<400> 30
ccaaagcaca tacttatoga tttaaatttc atcgaagaga ttaatatoga ataatcatat 60
acataacttta aatacataac aaatttttaa tacatatatc tggatatataa ttaatttttt 120
aaagtcatga agtatgtatc aaatacacat atggaaaaaa ttaactattc ataattttaa 180
aaatagaaaa gatacatcta gtgaaattag gtgcatgtat caaatacatt aggaaaagg 240
catatatctt gatctagata attaacgatt ttgatattatg tataatttcc aaatgaagg 300
ttatatctac ttcagaaata acaatatact tttatcagaa cattcaacaa agcaacaacc 360
aactagagtg aaaaatacac attgttctct agacatacaa aattgagaaa agaattctca 420
aatttagaga aacaaatctg aatttctaga agaaaaaaat aattatgcac tttgctattg 480
ctcgaaaaat aaatgaaaga aatttagactt ttttaaaaga tgtagacta gatatactca 540
aaagctatta aaggagtaat attcttctta cattaagtat tttagttaca gtctgtaat 600
taaagacaca ttttagattg tatctaaact taaatgtatc tagaatacat atatttgaat 660
gcatcatata catgtatccg acacaccaat tctcataaaa aacgtaatat cctaaactaa 720
tttatcttc aagtcaactt aagccaata tacattttca tctctaaagg cccaagtggc 780
acaaaatgtc aggcccaatt acgaagaaaa gggcttgtaa aaccctaata aagtggcact 840
ggcagagctt acactctcat tccatcaaca aagaaaccct aaaagccgca gcgccactga 900
tttctctct ccaggcgaag 920

15 <210> 31
<211> 178
<212> ADN
20 <213> Solanum tuberosum

<220>
<223> terminador de Ubiquitina 3

<400> 31
ttgatattta tgtagcaaa atgtcttata agttttctct ttttgcgaa cggtaattta 60
gagttttttt tgctatatgg attttcgttt ttgatgtatg tgacaaccct cgggattggt 120
gatttatctt aaaactaaga gtttttgcct tattgttctc gtctattttg gatataca 178

25 <210> 32
<211> 174
<212> ADN
<213> Solanum tuberosum

125

<220>

<223> terminador de monómero de Ubiquitina/proteína ribosómica
(Genbank No. de Registro Z11669.1)

<400> 32

```

ttttaatggt tagcaaatgt cttatcagtt ttctcttttt gtcgaacggt aatttagagt 60
ttttttttgc tatatggatt ttctgtttttg atgtatgtga caaccctcgg gattgttgat 120
ttattttcaaa actaagagtt ttgtcttatt gttctcgtct attttgata tcaa 174

```

5

<210> 33

<211> 1026

<212> ADN

<213> Escherichia coli

<220>

<223> ORF de gen de resistencia a Higromicina (hph)

<400> 33

10

```

atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga aaagttcgac 60
agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt cagcttcgat 120
gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatgggtt ctacaaagat 180
cgttatgttt atcggcactt tgcctcggcc gcgctcccga ttccggaagt gcttgacatt 240
ggggcattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg tgtaacgttg 300
caagacctgc ctgaaaccga actgcccgct gttctgcagc cggctcgcga ggccatggat 360
gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg accgcaagga 420
atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc ccatgtgtat 480
cactggcaaa ctgtgatgga cgacaccgtc agtgctcctg tcgcgcaggg tctcgatgag 540
ctgatgcttt ggcccgagga ctgcccga aaacacagcg tcattgactg gaggcaggcg 600
tccaacaatg tcctgacgga caatggcgcg aacatcttct tctggaggcc gtggttggt 660
atgttcgggg attccaata cgaggtcgcc aacatcttct tctggaggcc gtggttggt 720
tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc aggatcgccg 780
cggctccggg cgtatatgct ccgattggt cttgaccaac tctatcagag cttggttgac 840
ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgag acgcaatcgt ccgatccgga 900
gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg gaccgatggc 960
tgtgtagaag tactcgccga tagtggaaac cgacgccccg gcactcgtcc gagggcaag 1020
gaatag 1026

```

15

<210> 34

<211> 717

<212> ADN

<213> Secuencia Artificial

<220>

<223> Secuencia sintética de Proteína Fluorescente Verde (GFP)
derivada originalmente de Aequorea Victoria

20

<400> 34

```

atgttcagta aaggagaaga acttttcact ggagttgtcc caattcttgt tgaattagat 60
ggtgatgtta atgggcacaa attttctgtc agtggagagg gtgaagggtga tgcaacatac 120
ggaaaactta ccttaaat tatttgact actggaaaac tacctgttcc atggccaaca 180
cttgtcacta ctttctctta tgggtttcaa tgcttttcaa gataccaga tcatatgaag 240
cggcacgact tctcaagag cgccatgcct gagggatacg tgcaggagag gaccatctct 300
ttcaaggacg acgggaacta caagacacgt gctgaagtca agtttgagg agacaccctc 360
gtcaacagga tcgagcttaa gggaatcgat ttcaaggagg acggaaacat cctcgccac 420
aagttggaat acaactaca ctcccacaac gtatacatca cggcagacaa acaaaagaat 480
ggaatcaaag ctaacttcaa aattagacac aacattgaag atggaagcgt tcaactagca 540
gaccattatc acaaaaatac tccaattggc gatggccctg tccttttacc agacaacat 600
tacctgtcca cacaatctgc cttttcgaaa gatcccaacg aaaagagaga ccatatggtc 660
cttcttgagt ttgtaacagc tgctgggatt acacatggca tggatgaact atactga 717

```

25

nb

<210> 35
 <211> 2583
 <212> ADN
 <213> Aspergillus kawachii

<220>
 <223> secuencia CDS original de Beta glucosidasa (No. Reg. AB003470)

<400> 35
 5 atgagggttca ctttgattga ggcgggtggct ctcactgctg tctcgtctggc cagcgtgat 60
 gaattggctt actccccacc gtattaccca tccccctggg ccaatggcca gggcgactgg 120
 ggcgaggcat accagcgcgc tgttgatatt gtctcgcaga tgacattggc tgagaaggtc 180
 aatctgacca caggaactgg atgggaattg gagctatgtg ttggtoagac tggcgggggt 240
 ccccgattgg gagttccggg aatgtgttta caggatagcc ctctgggcgt tcgcgactcc 300
 gactacaact ctgctttccc ttccggtatg aacgtggctg caacctggga caagaatctg 360
 gcataacctc gcggcaaggc tatgggtcag gaatttagtg acaagggtgc cgatatccaa 420
 ttgggtccag ctgcgcggcc tctcggtaga agtcccgacg gtggtcgtaa ctgggagggc 480
 ttctcccccg acccggccct aagtgtgtg ctctttgcag agaccatcaa gggatatcaa 540
 gatgctggtg tggtcgcgac ggctaagcac tacattgcct acgagcaaga gcatttccgt 600
 caggcgctcg aagcccaagg ttatggattt aacatttccg agagtggag cgcgaaacctc 660
 gacgataaga ctatgcacga gctgtacctc tggcccttcg cggatgccat ccgtgcgggt 720
 gctggcgctg tgatgtgctc ctacaaccag atcaacaaca gctatggctg ccagaacagc 780
 10 tacactctga acaagctgct caagccgag ctgggtttcc agggctttgt catgagtgat 840
 tgggcggctc accatgctgg tgtgagtgtt gctttggcag gattggatat gtctatgcca 900
 ggagacgtcg actacgacag tggtagctct tactgggta caaacctgac cgtagcgtg 960
 ctcaacggaa cgggtgcccc atggcgtgtt gatgacatgg ctgtccgcat catggccgcc 1020
 tactacaagg tcggccgtga ccgtctgtgg actctccca acttcagctc atggaccaga 1080
 gatgaatacg gctacaagta ctactatgtg tcggaggagc cgtacgagaa ggtcaaccac 1140
 tacgtgaacg tgcaacgcaa ccacagcga ctgatccgc gcattggagc ggacagcacg 1200
 gtgctcctca agaacgacgg cgctctgcct ttgactggta aggagcgccg ggtcgcgctt 1260
 atcggagaag atgcgggctc caacccttat ggtgccaacg gctgcagtga ccgtggatgc 1320
 gacaatggaa cattggcgat gggctgggga agtggtactg ccaacttccc atacctggtg 1380
 acccccgagc aggccatctc aaacgaggtg ctcaagaaca agaattggtg attcaccgcc 1440
 accgataact gggctatcga tcagattgag gcgcttgcta agaccgcca ggtctctctt 1500
 15 gtctttgtca acgcccactc tggcgagggt tacatcaatg tcgacggaaa cctgggtgac 1560
 cgcaagaacc tgaccctgtg gaggaaacggc gataatgtga tcaaggctgc tgcagcaac 1620
 tgcaacaaca ccattgttat cattcactct gtcggccag tcttgggttaa cgaatggtac 1680
 gacaacccca atgttaccgc tattctctgg ggtggtctgc ccggtcagga gtctggcaac 1740
 tctcttgccg acgtcctcta tggccgtgtc aaccccggtg ccaagtgcgc ctttacctgg 1800
 ggcaagactc gtgagccta ccaagattac ttggtcaccg agcccaacaa cggcaatgga 1860
 gccccccagg aagacttcgt cgaggcgctc ttcatlgact accgcgatt cgacaagcgc 1920
 aacgagaccc cgatctacga gttcgctat ggtctgagct acaccacttt caactactcg 1980
 aaccttgagg tgcaggttct gagcgcctcc gcgtacgagc ctgcttcggg tgagactgag 2040
 gcagcgccaa cttttggaga ggttggaat gcgtcgaatt acctctacc cgacggactg 2100
 cagaaaaatca ccaagttcat ctacccctgg ctcaacagta ccgatctcga ggcattctct 2160
 ggggatgcta gctacggaca ggactcctcg gactatcttc ccgagggagc caccgatggc 2220
 20 tctgcgcaac cgatcctgcc tgcgtgtggc ggtcctggcg gcaaccctcg cctgtacgac 2280
 gagctcatcc gcgtgtcggg gaccatcaag aacaccggca aggttgctgg tgatgaagtt 2340
 ccccaactgt atgtttccct tggcgccccc aacgagccca agatcgtgct gcgtcaattc 2400
 gagcgcatca cgtgcagcc gtcagaggag acgaagtga gcacgactct gacgcgccgt 2460
 gaccttgcaa actggaatgt tgagaagcag gactgggaga ttacgtcgt tcccaagatg 2520
 gtgtttgtcg gaagctcctc gcggaagccg ccgctccggg cgtctctgcc tactgttcac 2580
 taa 2583

<210> 36
 <211> 860
 <212> PRT
 <213> Aspergillus kawachii

25 <220>
 <223> secuencia de proteínas de longitud completa de Beta glucosidasa

128

(No. de Registro BAA19913)

<400> 36
 Met Arg Phe Thr Leu Ile Glu Ala Val Ala Leu Thr Ala Val Ser Leu
 1 5 10 15
 Ala Ser Ala Asp Glu Leu Ala Tyr Ser Pro Pro Tyr Tyr Pro Ser Pro
 20 25 30
 Trp Ala Asn Gly Gln Gly Asp Trp Ala Gln Ala Tyr Gln Arg Ala Val
 35 40 45
 Asp Ile Val Ser Gln Met Thr Leu Ala Glu Lys Val Asn Leu Thr Thr
 50 55 60
 Gly Thr Gly Trp Glu Leu Glu Leu Cys Val Gly Gln Thr Gly Gly Val
 65 70 75 80
 Pro Arg Leu Gly Val Pro Gly Met Cys Leu Gln Asp Ser Pro Leu Gly
 85 90 95
 Val Arg Asp Ser Asp Tyr Asn Ser Ala Phe Pro Ser Gly Met Asn Val
 100 105 110
 Ala Ala Thr Trp Asp Lys Asn Leu Ala Tyr Leu Arg Gly Lys Ala Met
 115 120 125
 Gly Gln Glu Phe Ser Asp Lys Gly Ala Asp Ile Gln Leu Gly Pro Ala
 130 135 140
 Ala Gly Pro Leu Gly Arg Ser Pro Asp Gly Gly Arg Asn Trp Glu Gly
 145 150 155 160
 Phe Ser Pro Asp Pro Ala Leu Ser Gly Val Leu Phe Ala Glu Thr Ile
 165 170 175
 Lys Gly Ile Gln Asp Ala Gly Val Val Ala Thr Ala Lys His Tyr Ile
 180 185 190
 Ala Tyr Glu Gln Glu His Phe Arg Gln Ala Pro Glu Ala Gln Gly Tyr
 195 200 205
 Gly Phe Asn Ile Ser Glu Ser Gly Ser Ala Asn Leu Asp Asp Lys Thr
 210 215 220
 Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Ile Arg Ala Gly
 225 230 235 240
 Ala Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly
 245 250 255
 Cys Gln Asn Ser Tyr Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu Gly
 260 265 270
 Phe Gln Gly Phe Val Met Ser Asp Trp Ala Ala His His Ala Gly Val
 275 280 285
 Ser Gly Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val Asp
 290 295 300
 Tyr Asp Ser Gly Thr Ser Tyr Trp Gly Thr Asn Leu Thr Val Ser Val
 305 310 315 320
 Leu Asn Gly Thr Val Pro Gln Trp Arg Val Asp Asp Met Ala Val Arg
 325 330 335
 Ile Met Ala Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Trp Thr Pro
 340 345 350
 Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Tyr Lys Tyr Tyr
 355 360 365
 Tyr Val Ser Glu Gly Pro Tyr Glu Lys Val Asn His Tyr Val Asn Val
 370 375 380
 Gln Arg Asn His Ser Glu Leu Ile Arg Arg Ile Gly Ala Asp Ser Thr
 385 390 395 400
 Val Leu Leu Lys Asn Asp Gly Ala Leu Pro Leu Thr Gly Lys Glu Arg
 405 410 415
 Leu Val Ala Leu Ile Gly Glu Asp Ala Gly Ser Asn Pro Tyr Gly Ala
 420 425 430
 Asn Gly Cys Ser Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Gly
 435 440 445
 Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu Gln
 450 455 460

Ala Ile Ser Asn Glu Val Leu Lys Asn Lys Asn Gly Val Phe Thr Ala
 465 470 475 480
 Thr Asp Asn Trp Ala Ile Asp Gln Ile Glu Ala Leu Ala Lys Thr Ala
 485 490 495
 Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile
 500 505 510
 Asn Val Asp Gly Asn Leu Gly Asp Arg Lys Asn Leu Thr Leu Trp Arg
 515 520 525
 Asn Gly Asp Asn Val Ile Lys Ala Ala Ala Ser Asn Cys Asn Asn Thr
 530 535 540
 5 Ile Val Ile Ile His Ser Val Gly Pro Val Leu Val Asn Glu Trp Tyr
 545 550 555 560
 Asp Asn Pro Asn Val Thr Ala Ile Leu Trp Gly Gly Leu Pro Gly Gln
 565 570 575
 Glu Ser Gly Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val Asn Pro
 580 585 590
 Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala Tyr Gln
 595 600 605
 Asp Tyr Leu Val Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln Glu
 610 615 620
 Asp Phe Val Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp Lys Arg
 625 630 635 640
 10 Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr Thr
 645 650 655
 Phe Asn Tyr Ser Asn Leu Glu Val Gln Val Leu Ser Ala Pro Ala Tyr
 660 665 670
 Glu Pro Ala Ser Gly Glu Thr Glu Ala Ala Pro Thr Phe Gly Glu Val
 675 680 685
 Gly Asn Ala Ser Asn Tyr Leu Tyr Pro Asp Gly Leu Gln Lys Ile Thr
 690 695 700
 Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Glu Ala Ser Ser
 705 710 715 720
 Gly Asp Ala Ser Tyr Gly Gln Asp Ser Ser Asp Tyr Leu Pro Glu Gly
 725 730 735
 15 Ala Thr Asp Gly Ser Ala Gln Pro Ile Leu Pro Ala Gly Gly Gly Pro
 740 745 750
 Gly Gly Asn Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser Val Thr
 755 760 765
 Ile Lys Asn Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln Leu Tyr
 770 775 780
 Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Ile Val Leu Arg Gln Phe
 785 790 795 800
 Glu Arg Ile Thr Leu Gln Pro Ser Glu Glu Thr Lys Trp Ser Thr Thr
 805 810 815
 Leu Thr Arg Arg Asp Leu Ala Asn Trp Asn Val Glu Lys Gln Asp Trp
 820 825 830
 20 Glu Ile Thr Ser Tyr Pro Lys Met Val Phe Val Gly Ser Ser Ser Arg
 835 840 845
 Lys Pro Pro Leu Arg Ala Ser Leu Pro Thr Val His
 850 855 860

<210> 37
 <211> 841
 <212> PRT
 <213> Secuencia Artificial

<220>
 25 <223> secuencia parcial de Beta glucosidasa (No. de Registro
 BAA19913) con secuencia de señal (19 aa) suprimida

<400> 37
 Asp Glu Leu Ala Tyr Ser Pro Pro Tyr Tyr Pro Ser Pro Trp Ala Asn
 1 5 10 15
 Gly Gln Gly Asp Trp Ala Gln Ala Tyr Gln Arg Ala Val Asp Ile Val
 20 25 30
 Ser Gln Met Thr Leu Ala Glu Lys Val Asn Leu Thr Thr Gly Thr Gly
 35 40 45
 Trp Glu Leu Glu Leu Cys Val Gly Gln Thr Gly Gly Val Pro Arg Leu
 50 55 60
 Gly Val Pro Gly Met Cys Leu Gln Asp Ser Pro Leu Gly Val Arg Asp
 65 70 75 80
 Ser Asp Tyr Asn Ser Ala Phe Pro Ser Gly Met Asn Val Ala Ala Thr
 85 90 95
 Trp Asp Lys Asn Leu Ala Tyr Leu Arg Gly Lys Ala Met Gly Gln Glu
 100 105 110
 Phe Ser Asp Lys Gly Ala Asp Ile Gln Leu Gly Pro Ala Ala Gly Pro
 115 120 125
 Leu Gly Arg Ser Pro Asp Gly Gly Arg Asn Trp Glu Gly Phe Ser Pro
 130 135 140
 Asp Pro Ala Leu Ser Gly Val Leu Phe Ala Glu Thr Ile Lys Gly Ile
 145 150 155 160
 Gln Asp Ala Gly Val Val Ala Thr Ala Lys His Tyr Ile Ala Tyr Glu
 165 170 175
 Gln Glu His Phe Arg Gln Ala Pro Glu Ala Gln Gly Tyr Gly Phe Asn
 180 185 190
 Ile Ser Glu Ser Gly Ser Ala Asn Leu Asp Asp Lys Thr Met His Glu
 195 200 205
 Leu Tyr Leu Trp Pro Phe Ala Asp Ala Ile Arg Ala Gly Ala Gly Ala
 210 215 220
 Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr Gly Cys Gln Asn
 225 230 235 240
 Ser Tyr Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu Gly Phe Gln Gly
 245 250 255
 Phe Val Met Ser Asp Trp Ala Ala His His Ala Gly Val Ser Gly Ala
 260 265 270
 Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val Asp Tyr Asp Ser
 275 280 285
 Gly Thr Ser Tyr Trp Gly Thr Asn Leu Thr Val Ser Val Leu Asn Gly
 290 295 300
 Thr Val Pro Gln Trp Arg Val Asp Asp Met Ala Val Arg Ile Met Ala
 305 310 315 320
 Ala Tyr Tyr Lys Val Gly Arg Asp Arg Leu Trp Thr Pro Pro Asn Phe
 325 330 335
 Ser Ser Trp Thr Arg Asp Glu Tyr Gly Tyr Lys Tyr Tyr Tyr Val Ser
 340 345 350
 Glu Gly Pro Tyr Glu Lys Val Asn His Tyr Val Asn Val Gln Arg Asn
 355 360 365
 His Ser Glu Leu Ile Arg Arg Ile Gly Ala Asp Ser Thr Val Leu Leu
 370 375 380
 Lys Asn Asp Gly Ala Leu Pro Leu Thr Gly Lys Glu Arg Leu Val Ala
 385 390 395 400
 Leu Ile Gly Glu Asp Ala Gly Ser Asn Pro Tyr Gly Ala Asn Gly Cys
 405 410 415
 Ser Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met Gly Trp Gly Ser
 420 425 430
 Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu Gln Ala Ile Ser
 435 440 445
 Asn Glu Val Leu Lys Asn Lys Asn Gly Val Phe Thr Ala Thr Asp Asn
 450 455 460
 Trp Ala Ile Asp Gln Ile Glu Ala Leu Ala Lys Thr Ala Ser Val Ser
 465 470 475 480

130

Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile Asn Val Asp
 485 490 495
 Gly Asn Leu Gly Asp Arg Lys Asn Leu Thr Leu Trp Arg Asn Gly Asp
 500 505 510
 Asn Val Ile Lys Ala Ala Ala Ser Asn Cys Asn Asn Thr Ile Val Ile
 515 520 525
 Ile His Ser Val Gly Pro Val Leu Val Asn Glu Trp Tyr Asp Asn Pro
 530 535 540
 Asn Val Thr Ala Ile Leu Trp Gly Gly Leu Pro Gly Gln Glu Ser Gly
 545 550 555 560
 5 Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val Asn Pro Gly Ala Lys
 565 570 575
 Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala Tyr Gln Asp Tyr Leu
 580 585 590
 Val Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln Glu Asp Phe Val
 595 600 605
 Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp Lys Arg Asn Glu Thr
 610 615 620
 Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr Thr Phe Asn Tyr
 625 630 635 640
 Ser Asn Leu Glu Val Gln Val Leu Ser Ala Pro Ala Tyr Glu Pro Ala
 645 650 655
 10 Ser Gly Glu Thr Glu Ala Ala Pro Thr Phe Gly Glu Val Gly Asn Ala
 660 665 670
 Ser Asn Tyr Leu Tyr Pro Asp Gly Leu Gln Lys Ile Thr Lys Phe Ile
 675 680 685
 Tyr Pro Trp Leu Asn Ser Thr Asp Leu Glu Ala Ser Ser Gly Asp Ala
 690 695 700
 Ser Tyr Gly Gln Asp Ser Ser Asp Tyr Leu Pro Glu Gly Ala Thr Asp
 705 710 715 720
 Gly Ser Ala Gln Pro Ile Leu Pro Ala Gly Gly Gly Pro Gly Gly Asn
 725 730 735
 Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser Val Thr Ile Lys Asn
 740 745 750
 15 Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln Leu Tyr Val Ser Leu
 755 760 765
 Gly Gly Pro Asn Glu Pro Lys Ile Val Leu Arg Gln Phe Glu Arg Ile
 770 775 780
 Thr Leu Gln Pro Ser Glu Glu Thr Lys Trp Ser Thr Thr Leu Thr Arg
 785 790 795 800
 Arg Asp Leu Ala Asn Trp Asn Val Glu Lys Gln Asp Trp Glu Ile Thr
 805 810 815
 Ser Tyr Pro Lys Met Val Phe Val Gly Ser Ser Ser Arg Lys Pro Pro
 820 825 830
 Leu Arg Ala Ser Leu Pro Thr Val His
 835 840
 20 <210> 38
 <211> 866
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Beta glucosidasa GSF V1 - Beta glucosidasa (No. de
 Registro BAA19913) con secuencia de señal sustituida
 y KDEL añadida
 <400> 38
 Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
 1 5 10 15
 25 Leu Ser Ser Ala Glu Asp Glu Leu Ala Tyr Ser Pro Pro Tyr Tyr Pro

131

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

132

Tyr Ile Asn Val Asp Gly Asn Leu Gly Asp Arg Lys Asn Leu Thr Leu
 515 520 525
 Trp Arg Asn Gly Asp Asn Val Ile Lys Ala Ala Ala Ser Asn Cys Asn
 530 535 540
 Asn Thr Ile Val Ile Ile His Ser Val Gly Pro Val Leu Val Asn Glu
 545 550 555 560
 Trp Tyr Asp Asn Pro Asn Val Thr Ala Ile Leu Trp Gly Gly Leu Pro
 565 570 575
 Gly Gln Glu Ser Gly Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val
 580 585 590
 5 Asn Pro Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala
 595 600 605
 Tyr Gln Asp Tyr Leu Val Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro
 610 615 620
 Gln Glu Asp Phe Val Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp
 625 630 635 640
 Lys Arg Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr
 645 650 655
 Thr Thr Phe Asn Tyr Ser Asn Leu Glu Val Gln Val Leu Ser Ala Pro
 660 665 670
 Ala Tyr Glu Pro Ala Ser Gly Glu Thr Glu Ala Ala Pro Thr Phe Gly
 675 680 685
 10 Glu Val Gly Asn Ala Ser Asn Tyr Leu Tyr Pro Asp Gly Leu Gln Lys
 690 695 700
 Ile Thr Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Glu Ala
 705 710 715 720
 Ser Ser Gly Asp Ala Ser Tyr Gly Gln Asp Ser Ser Asp Tyr Leu Pro
 725 730 735
 Glu Gly Ala Thr Asp Gly Ser Ala Gln Pro Ile Leu Pro Ala Gly Gly
 740 745 750
 Gly Pro Gly Gly Asn Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser
 755 760 765
 Val Thr Ile Lys Asn Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln
 770 775 780
 15 Leu Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Ile Val Leu Arg
 785 790 795 800
 Gln Phe Glu Arg Ile Thr Leu Gln Pro Ser Glu Glu Thr Lys Trp Ser
 805 810 815
 Thr Thr Leu Thr Arg Arg Asp Leu Ala Asn Trp Asn Val Glu Lys Gln
 820 825 830
 Asp Trp Glu Ile Thr Ser Tyr Pro Lys Met Val Phe Val Gly Ser Ser
 835 840 845
 Ser Arg Lys Pro Pro Leu Arg Ala Ser Leu Pro Thr Val His Lys Asp
 850 855 860
 Glu Leu
 865
 20
 <210> 39
 <211> 866
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Beta glucosidasa GSF V2 - Beta glucosidasa (No. Reg.
 BAA19913) con secuencia de señal sustituida y KDEL añadida,
 así como 13 sustituciones de aminoácido en posiciones 59, 110,
 210, 320, 382, 475, 524, 549, 695, 700, 704, 715 y 852
 25 <400> 39
 Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser

133

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

134

5 Thr Ala Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly
500 505 510
Tyr Ile Asn Val Asp Gly Asn Leu Gly Asp Arg Arg Asn Leu Thr Leu
515 520 525
Trp Arg Asn Gly Asp Asn Val Ile Lys Ala Ala Ala Ser Asn Cys Asn
530 535 540
Asn Thr Ile Val Val Ile His Ser Val Gly Pro Val Leu Val Asn Glu
545 550 555 560
Trp Tyr Asp Asn Pro Asn Val Thr Ala Ile Leu Trp Gly Gly Leu Pro
565 570 575
10 Gly Gln Glu Ser Gly Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val
580 585 590
Asn Pro Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala
595 600 605
Tyr Gln Asp Tyr Leu Val Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro
610 615 620
Gln Glu Asp Phe Val Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp
625 630 635 640
Lys Arg Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr
645 650 655
Thr Thr Phe Asn Tyr Ser Asn Leu Glu Val Gln Val Leu Ser Ala Pro
660 665 670
15 Ala Tyr Glu Pro Ala Ser Gly Glu Thr Glu Ala Ala Pro Thr Phe Gly
675 680 685
Glu Val Gly Asn Ala Ser Asp Tyr Leu Tyr Pro Ser Gly Leu Gln Arg
690 695 700
Ile Thr Lys Phe Ile Tyr Pro Trp Leu Asn Gly Thr Asp Leu Glu Ala
705 710 715 720
Ser Ser Gly Asp Ala Ser Tyr Gly Gln Asp Ser Ser Asp Tyr Leu Pro
725 730 735
Glu Gly Ala Thr Asp Gly Ser Ala Gln Pro Ile Leu Pro Ala Gly Gly
740 745 750
Gly Pro Gly Gly Asn Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser
755 760 765
20 Val Thr Ile Lys Asn Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln
770 775 780
Leu Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Ile Val Leu Arg
785 790 795 800
Gln Phe Glu Arg Ile Thr Leu Gln Pro Ser Glu Glu Thr Lys Trp Ser
805 810 815
Thr Thr Leu Thr Arg Arg Asp Leu Ala Asn Trp Asn Val Glu Lys Gln
820 825 830
Asp Trp Glu Ile Thr Ser Tyr Pro Lys Met Val Phe Val Gly Ser Ser
835 840 845
Ser Arg Lys Leu Pro Leu Arg Ala Ser Leu Pro Thr Val His Lys Asp
850 855 860
Glu Leu
865

25 <210> 40
<211> 860
<212> PRT
<213> Aspergillus niger
<220>
<223> Beta glucosidasa (SEQ ID NO.2 de US Pat. No. 7223902,
también No. de Reg. ABT13410)
<400> 40
Met Arg Phe Thr Leu Ile Glu Ala Val Ala Leu Thr Ala Val Ser Leu

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

135

Ser Val Ser Leu Val Phe Val Asn Ala Asp Ser Gly Glu Gly Tyr Ile
 500 505 510
 Asn Val Asp Gly Asn Leu Gly Asp Arg Arg Asn Leu Thr Leu Trp Arg
 515 520 525
 Asn Gly Asp Asn Val Ile Lys Ala Ala Ala Ser Asn Cys Asn Asn Thr
 530 535 540
 Ile Val Val Ile His Ser Val Gly Pro Val Leu Val Asn Glu Trp Tyr
 545 550 555 560
 Asp Asn Pro Asn Val Thr Ala Ile Leu Trp Gly Gly Leu Pro Gly Gln
 565 570 575
 5 Glu Ser Gly Asn Ser Leu Ala Asp Val Leu Tyr Gly Arg Val Asn Pro
 580 585 590
 Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ala Tyr Gln
 595 600 605
 Asp Tyr Leu Val Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln Glu
 610 615 620
 Asp Phe Val Glu Gly Val Phe Ile Asp Tyr Arg Gly Phe Asp Lys Arg
 625 630 635 640
 Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr Thr
 645 650 655
 Phe Asn Tyr Ser Asn Leu Glu Val Gln Val Leu Ser Ala Pro Ala Tyr
 660 665 670
 10 Glu Pro Ala Ser Gly Glu Thr Glu Ala Ala Pro Thr Phe Gly Glu Val
 675 680 685
 Gly Asn Ala Ser Asp Tyr Leu Tyr Pro Ser Gly Leu Leu Arg Ile Thr
 690 695 700
 Lys Phe Ile Tyr Pro Trp Leu Asn Gly Thr Asp Leu Glu Ala Ser Ser
 705 710 715 720
 Gly Asp Ala Ser Tyr Gly Gln Asp Ser Ser Asp Tyr Leu Pro Glu Gly
 725 730 735
 Ala Thr Asp Gly Ser Ala Gln Pro Ile Leu Pro Ala Gly Gly Gly Pro
 740 745 750
 Gly Gly Asn Pro Arg Leu Tyr Asp Glu Leu Ile Arg Val Ser Val Thr
 755 760 765
 15 Ile Lys Asn Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln Leu Tyr
 770 775 780
 Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Ile Val Leu Arg Gln Phe
 785 790 795 800
 Glu Arg Ile Thr Leu Gln Pro Ser Glu Glu Thr Lys Trp Ser Thr Thr
 805 810 815
 Leu Thr Arg Arg Asp Leu Ala Asn Trp Asn Val Glu Lys Gln Asp Trp
 820 825 830
 Glu Ile Thr Ser Tyr Pro Lys Met Val Phe Val Gly Ser Ser Ser Arg
 835 840 845
 Lys Leu Pro Leu Arg Ala Ser Leu Pro Thr Val His
 850 855 860
 20 <210> 41
 <211> 2586
 <212> ADN
 <213> Aspergillus terreus NIH2624
 <220>
 <223> CDS de precursor de Beta-glucosidasa I (No. Reg. XM_001212225)
 <400> 41
 atgaagcttt ccatttttggg ggcagcagct ttgacagctg cctccgtagt cagcgcacag 60
 gacgatctcg catactccc gcggtactac ccttctccct gggccgatgg ccacggtgag 120
 25 tggtcgaacg cgtacaagcg cgctgtagat atcgctctctc agatgacatt gacggagaag 180
 gtcaatctca ccaccggtac tggatgggag ttggagaggt gtgtcggtca gacgggcagt 240
 gtccctagac tgggaatccc aagcctctgt ctgcaggata gccctctggg tattcgcatg 300

136

tcggactata actcggcctt cctcgccgggt attaacggtt cgccacactg ggacaagaag 360
 cttgcctacc aacgcggcaa ggcaatgggc gaggaattca gtgacaaggg tattgatgtt 420
 cagttggggc ctgctgccgg tcctcttggc aggtccccc atggaggccg aaactgggag 480
 ggcttctctc ctgatcccg cctgactggt gtgttgttcg ccgagacgat caagggtatc 540
 caggacgccg gaggttattgc taccgcgaag cactacattc tcaacgaaca agagcatttc 600
 cgcaggtcgc gcgaagccca gggctatggc ttcaacatca ccgaaaccgt gagctcaaatt 660
 gtgatgaca agaccatgca cgagctgtat ctctggccct tcgccgatgc ggtgcgcgcg 720
 ggcgtgggcg ctgtgatgtg ctctacaac cagatcaaca acagctacgg atgccaaaac 780
 agtttgaccc tgaacaagct cttgaaagcc gaactcggat ttcagggatt tgtcatgagt 840
 gactggagtg ctcaccacag cgggtgttggc gccgccttgg ctggttttga catgtccatg 900
 ccgggagata tcagtttcga cagcggcact tccttctatg gcacgaacct gactgttggc 960
 gtcctcaacg gcaccattcc ccagtggcgt gtggatgaca tggccgtccg gatcatggct 1020
 gcctactaca aggttggccg cgaccgtctc tggactcctc ccaatttcag ctctgggact 1080
 cgcgatgaat atggcttcgc gcacttcttc ccttccgaag gcgcttatga acgtgtcaat 1140
 gaattcgtca acg!gcagcg tgaccatgcc caggtgatcc gtcggattgg cgcggatagt 1200
 gtcgtgctct tgaagaacga cggtgccctt cccttgacgg gccaggagaa gactgttggc 1260
 attctgggcg aagacgccgg gtcgaatccg aagggagcaa atggttgacg tgaccgtggc 1320
 tgtgacaagg gtactctggc catggcttgg ggtagtggta ctgccaactt cccttacctt 1380
 gtgactcccg aacaggccat tcagaacgag gttctgaagg gccgtggaaa tgtctttgcc 1440
 gtgacggaca actatgatac acagcagatt gccgcggtt cctctcaatc caccgtttca 1500
 ttggttttcg tgaacgcaga cgcgggtgaa gggttcttca atgtggacgg aaacatgggt 1560
 gatcgcaaga acctcaccct ctggcagaac ggagaggaa tgatcaagac tgtcacggag 1620
 cactgcaaca acaccgttgt tgtgatccat tcgggtgggac ctgttctcat cgatgagtgg 1680
 tatcgcgacc ccaatgtcac cggcattctg tgggctggtc tcccgggcca ggagtctggc 1740
 aacgccattg cggacgtgct gtacggcgcg ttcgaacctg gggcaagac cccctttacc 1800
 tggggtgaaga cgcgcgcgct ctacggcgac tacctctca ccgagcccaa caacggcaac 1860
 ggtgctctc aagacaactt caacgagggc gtgtttattg actaccgtcg cttcgacaag 1920
 tacaatgaga cgcctatcta cgagttcggc catggtctga gctacacgac gtttgagctg 1980
 tctggcctcc aggtccagct tatcaacgga tccagctatg ttccactac gggtcagacg 2040
 agcgcgcgcc aggcatttgg taaagtcgag gacgcgtcta gctacctgta cctgagggga 2100
 ctgaagagga ttccaagtt catctatccc tggctgaact ctaccgatct taaagcgtct 2160
 accggcgatc ctgaatacgg agagcccaac ttcagataa ttccgaagg tgctaccgat 2220
 ggctctctc agcccgctct gcctgccagc gggggtcctg gcggcaacc cggctctat 2280
 gaggatctct tccaggttct tgtgaccatc accaaccacg gcaaggttgc tggatgatgag 2340
 gtgcctcagc tgtatgtttc gctgggtggc cccaacgagc cgaagcgggt gctgcgcaag 2400
 ttcgagcgcc tgcacatcgc ccctggtcag caaaaggtct ggacgactac cctgaaccgc 2460
 cgtgacctag ccaactggga tgtcgtggcc caggactgga agatcactcc ctatgctaag 2520
 accatcttgg ttggcacctc ttcgcgcaag ctgcctctcg ctggtcgctt gccacgggtg 2580
 cagtaa 2586

<210> 42
 <211> 861
 <212> PRT
 <213> *Aspergillus terreus* NIH2624

<220>
 <223> precursor de Beta-glucosidasa I (traducción de No. de Registro XM_001212225)

<400> 42
 Met Lys Leu Ser Ile Leu Glu Ala Ala Ala Leu Thr Ala Ala Ser Val
 1 5 10 15
 Val Ser Ala Gln Asp Asp Leu Ala Tyr Ser Pro Pro Tyr Tyr Pro Ser
 20 25 30
 Pro Trp Ala Asp Gly His Gly Glu Trp Ser Asn Ala Tyr Lys Arg Ala
 35 40 45
 Val Asp Ile Val Ser Gln Met Thr Leu Thr Glu Lys Val Asn Leu Thr
 50 55 60
 Thr Gly Thr Gly Trp Glu Leu Glu Arg Cys Val Gly Gln Thr Gly Ser
 65 70 75 80
 Val Pro Arg Leu Gly Ile Pro Ser Leu Cys Leu Gln Asp Ser Pro Leu

25

5 Gln Glu Ser Gly Asn Ala Ile Ala Asp Val Leu Tyr Gly Arg Val Asn
 580 585 590
 Pro Gly Gly Lys Thr Pro Phe Thr Trp Gly Lys Thr Arg Ala Ser Tyr
 595 600 605
 Gly Asp Tyr Leu Leu Thr Glu Pro Asn Asn Gly Asn Gly Ala Pro Gln
 610 615 620
 Asp Asn Phe Asn Glu Gly Val Phe Ile Asp Tyr Arg Arg Phe Asp Lys
 625 630 635 640
 Tyr Asn Glu Thr Pro Ile Tyr Glu Phe Gly His Gly Leu Ser Tyr Thr
 645 650 655
 10 Thr Phe Glu Leu Ser Gly Leu Gln Val Gln Leu Ile Asn Gly Ser Ser
 660 665 670
 Tyr Val Pro Thr Thr Gly Gln Thr Ser Ala Ala Gln Ala Phe Gly Lys
 675 680 685
 Val Glu Asp Ala Ser Ser Tyr Leu Tyr Pro Glu Gly Leu Lys Arg Ile
 690 695 700
 Ser Lys Phe Ile Tyr Pro Trp Leu Asn Ser Thr Asp Leu Lys Ala Ser
 705 710 715 720
 Thr Gly Asp Pro Glu Tyr Gly Glu Pro Asn Phe Glu Tyr Ile Pro Glu
 725 730 735
 Gly Ala Thr Asp Gly Ser Pro Gln Pro Arg Leu Pro Ala Ser Gly Gly
 740 745 750
 15 Pro Gly Gly Asn Pro Gly Leu Tyr Glu Asp Leu Phe Gln Val Ser Val
 755 760 765
 Thr Ile Thr Asn Thr Gly Lys Val Ala Gly Asp Glu Val Pro Gln Leu
 770 775 780
 Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Arg Val Leu Arg Lys
 785 790 795 800
 Phe Glu Arg Leu His Ile Ala Pro Gly Gln Gln Lys Val Trp Thr Thr
 805 810 815
 Thr Leu Asn Arg Asp Leu Ala Asn Trp Asp Val Val Ala Gln Asp
 820 825 830
 Trp Lys Ile Thr Pro Tyr Ala Lys Thr Ile Phe Val Gly Thr Ser Ser
 835 840 845
 20 Arg Lys Leu Pro Leu Ala Gly Arg Leu Pro Arg Val Gln
 850 855 860

<210> 43
 <211> 513
 <212> PRT
 <213> Trichoderma reesei

<220>
 <223> Proteína de longitud completa de Exoglucanasa (No. de
 Registro P62694)
 <400> 43

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

139

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

<210> 44

<211> 521

<212> PRT

<213> Secuencia Artificial

<220>

25

<223> Proteína de exocelobiohidrolasa I modificada (señal ER

140

reemplazada y secuencia KDEL añadida a No. Reg. P62694)

<400> 44

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

Thr Thr Thr Thr Arg Arg Pro Ala Thr Thr Thr Gly Ser Ser Pro Gly
465 470 475 480
Pro Thr Gln Ser His Tyr Gly Gln Cys Gly Gly Ile Gly Tyr Ser Gly
485 490 495
Pro Thr Val Cys Ala Ser Gly Thr Thr Cys Gln Val Leu Asn Pro Tyr
500 505 510
Tyr Ser Gln Cys Leu Lys Asp Glu Leu
515 520

5 <210> 45
<211> 471
<212> PRT
<213> Trichoderma reesei

<220>
<223> Proteína de longitud completa de Exoglucanasa 2 (No. Reg.
P07987)

<400> 45
Met Ile Val Gly Ile Leu Thr Thr Leu Ala Thr Leu Ala Thr Leu Ala
1 5 10 15
Ala Ser Val Pro Leu Glu Glu Arg Gln Ala Cys Ser Ser Val Trp Gly
20 25 30
Gln Cys Gly Gly Gln Asn Trp Ser Gly Pro Thr Cys Cys Ala Ser Gly
35 40 45
Ser Thr Cys Val Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu Pro Gly
50 55 60
Ala Ala Ser Ser Ser Ser Thr Arg Ala Ala Ser Thr Thr Ser Arg
65 70 75 80
Val Ser Pro Thr Thr Ser Arg Ser Ser Ser Ala Thr Pro Pro Pro Gly
85 90 95
Ser Thr Thr Thr Arg Val Pro Pro Val Gly Ser Gly Thr Ala Thr Tyr
100 105 110
Ser Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr
115 120 125
15 Ala Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met
130 135 140
Ala Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu
145 150 155 160
Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
165 170 175
Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
180 185 190
Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
195 200 205
Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
210 215 220
20 Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu
225 230 235 240
Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr
245 250 255
Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
260 265 270
Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
275 280 285
Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala
290 295 300
Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
305 310 315 320
25 Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr
325 330 335

142

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

<210> 46
 <211> 472
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Exoglucanasa 2 modificada (señal ER reemplazada y
 secuencia KDEL añadida a No. de Registro P07987)

<400> 46
 Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
 1 5 10 15
 Leu Ser Ser Ala Glu Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly
 20 25 30
 Gly Gln Asn Trp Ser Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys
 35 40 45
 Val Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser
 50 55 60
 Ser Ser Ser Ser Thr Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro
 65 70 75 80
 Thr Thr Ser Arg Ser Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr
 85 90 95
 Thr Arg Val Pro Pro Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn
 100 105 110
 Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu
 115 120 125
 Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala
 130 135 140
 Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu
 145 150 155 160
 Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala
 165 170 175
 Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu
 180 185 190
 Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile
 195 200 205
 Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg
 210 215 220
 Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu
 225 230 235 240
 25 Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys

5
10
15
20
25

245 250 255
Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr
260 265 270
Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala
275 280 285
Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe
290 295 300
Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu
305 310 315 320
Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro
325 330 335
Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His
340 345 350
Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe
355 360 365
Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln
370 375 380
Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro
385 390 395 400
Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys
405 410 415
Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe
420 425 430
Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala
435 440 445
Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn
450 455 460
Pro Ser Phe Leu Lys Asp Glu Leu
465 470

<210> 47
<211> 562
<212> PRT
<213> Acidothermus cellulolyticus

<220>
<223> secuencia de longitud completa de Endoglucanasa E1
(No. de Registro P54583)
<400> 47
Met Pro Arg Ala Leu Arg Arg Val Pro Gly Ser Arg Val Met Leu Arg
1 5 10 15
Val Gly Val Val Val Ala Val Leu Ala Leu Val Ala Ala Leu Ala Asn
20 25 30
Leu Ala Val Pro Arg Pro Ala Arg Ala Ala Gly Gly Gly Tyr Trp His
35 40 45
Thr Ser Gly Arg Glu Ile Leu Asp Ala Asn Asn Val Pro Val Arg Ile
50 55 60
Ala Gly Ile Asn Trp Phe Gly Phe Glu Thr Cys Asn Tyr Val Val His
65 70 75 80
Gly Leu Trp Ser Arg Asp Tyr Arg Ser Met Leu Asp Gln Ile Lys Ser
85 90 95
Leu Gly Tyr Asn Thr Ile Arg Leu Pro Tyr Ser Asp Asp Ile Leu Lys
100 105 110
Pro Gly Thr Met Pro Asn Ser Ile Asn Phe Tyr Gln Met Asn Gln Asp
115 120 125
Leu Gln Gly Leu Thr Ser Leu Gln Val Met Asp Lys Ile Val Ala Tyr
130 135 140
Ala Gly Gln Ile Gly Leu Arg Ile Ile Leu Asp Arg His Arg Pro Asp
145 150 155 160
Cys Ser Gly Gln Ser Ala Leu Trp Tyr Thr Ser Ser Val Ser Glu Ala

5
10
15
20

Thr	Trp	Ile	Ser	Asp	Leu	Gln	Ala	Leu	Ala	Gln	Arg	Tyr	Lys	Gly	Asn
			180					185					190		
Pro	Thr	Val	Val	Gly	Phe	Asp	Leu	His	Asn	Glu	Pro	His	Asp	Pro	Ala
			195				200					205			
Cys	Trp	Gly	Cys	Gly	Asp	Pro	Ser	Ile	Asp	Trp	Arg	Leu	Ala	Ala	Glu
	210					215				220					
Arg	Ala	Gly	Asn	Ala	Val	Leu	Ser	Val	Asn	Pro	Asn	Leu	Leu	Ile	Phe
	225				230					235					240
Val	Glu	Gly	Val	Gln	Ser	Tyr	Asn	Gly	Asp	Ser	Tyr	Trp	Trp	Gly	Gly
				245					250					255	
Asn	Leu	Gln	Gly	Ala	Gly	Gln	Tyr	Pro	Val	Val	Leu	Asn	Val	Pro	Asn
			260					265					270		
Arg	Leu	Val	Tyr	Ser	Ala	His	Asp	Tyr	Ala	Thr	Ser	Val	Tyr	Pro	Gln
		275					280					285			
Thr	Trp	Phe	Ser	Asp	Pro	Thr	Phe	Pro	Asn	Asn	Met	Pro	Gly	Ile	Trp
	290					295					300				
Asn	Lys	Asn	Trp	Gly	Tyr	Leu	Phe	Asn	Gln	Asn	Ile	Ala	Pro	Val	Trp
	305				310					315					320
Leu	Gly	Glu	Phe	Gly	Thr	Thr	Leu	Gln	Ser	Thr	Thr	Asp	Gln	Thr	Trp
				325					330					335	
Leu	Lys	Thr	Leu	Val	Gln	Tyr	Leu	Arg	Pro	Thr	Ala	Gln	Tyr	Gly	Ala
				340				345					350		
Asp	Ser	Phe	Gln	Trp	Thr	Phe	Trp	Ser	Trp	Asn	Pro	Asp	Ser	Gly	Asp
		355					360					365			
Thr	Gly	Gly	Ile	Leu	Lys	Asp	Trp	Gln	Thr	Val	Asp	Thr	Val	Lys	
	370					375				380					
Asp	Gly	Tyr	Leu	Ala	Pro	Ile	Lys	Ser	Ser	Ile	Phe	Asp	Pro	Val	Gly
	385				390					395					400
Ala	Ser	Ala	Ser	Pro	Ser	Ser	Gln	Pro	Ser	Pro	Ser	Val	Ser	Pro	Ser
				405					410					415	
Pro	Ser	Pro	Ser	Pro	Ser	Ala	Ser	Arg	Thr	Pro	Thr	Pro	Thr	Pro	Thr
			420					425					430		
Pro	Thr	Ala	Ser	Pro	Thr	Pro	Thr	Leu	Thr	Pro	Thr	Ala	Thr	Pro	Thr
		435					440					445			
Pro	Thr	Ala	Ser	Pro	Thr	Pro	Ser	Pro	Thr	Ala	Ala	Ser	Gly	Ala	Arg
	450					455				460					
Cys	Thr	Ala	Ser	Tyr	Gln	Val	Asn	Ser	Asp	Trp	Gly	Asn	Gly	Phe	Thr
	465				470					475					480
Val	Thr	Val	Ala	Val	Thr	Asn	Ser	Gly	Ser	Val	Ala	Thr	Lys	Thr	Trp
				485					490					495	
Thr	Val	Ser	Trp	Thr	Phe	Gly	Gly	Asn	Gln	Thr	Ile	Thr	Asn	Ser	Trp
			500					505					510		
Asn	Ala	Ala	Val	Thr	Gln	Asn	Gly	Gln	Ser	Val	Thr	Ala	Arg	Asn	Met
		515					520					525			
Ser	Tyr	Asn	Asn	Val	Ile	Gln	Pro	Gly	Gln	Asn	Thr	Thr	Phe	Gly	Phe
	530					535					540				
Gln	Ala	Ser	Tyr	Thr	Gly	Ser	Asn	Ala	Ala	Pro	Thr	Val	Ala	Cys	Ala
	545				550					555					560
Ala	Ser														

<210> 48
<211> 636
<212> ADN
<213> Aspergillus niger

25
<220>
<223> CDS de Xilanasa (No. de Registro U39784)

145

<400> 48
 atgaagggtca ctgcgggttt tgcaagtctc ttgcttacgg ccttcgcggc ccttgctccg 60
 gagcctgttc tgggtgtcgcg aagtgcgggt atcaactacg tgcagaacta caacggcaac 120
 cttggtgact tcacctacga cgagagtacc gggacatttt ccatgtactg ggaggatgga 180
 gtcagttccg acttcgtcgt tggtttgggc tggaccactg gtcctcttaa atctatcacc 240
 tactctgccc aatacagcgc ttctagctcc agctcctacc tggctgtcta cggctgggtc 300
 aactctcttc aggcggaata ctacatcgtc gaggattacg gtgattacaa cccttgacgc 360
 tcggccacga gccttgggtac cgtgtactct gatggaagca cctaccaagt ctgcaccgac 420
 actcgacgaa cgcggccatc tatcacagga acaagcacgt tcacgcagta cttctccggt 480
 cgtgaaagta caccacatc cggaacagtg actatcgcca accatttcaa tttctyggcg 540
 5 cagcatgggt tcggcaatag caacttcaat tatcagggtc tggcgggtgga ggcatggaac 600
 ggtgtcggca gtgccagtgt caccatctcc tcttaa 636

<210> 49
 <211> 211
 <212> PRT
 <213> Aspergillus niger

<220>
 <223> secuencia de proteínas de Xilanasa (No. de Registro AAA99065.1)

<400> 49
 10 Met Lys Val Thr Ala Ala Phe Ala Ser Leu Leu Thr Ala Phe Ala
 1 5 10 15
 Ala Pro Ala Pro Glu Pro Val Leu Val Ser Arg Ser Ala Gly Ile Asn
 20 25 30
 Tyr Val Gln Asn Tyr Asn Gly Asn Leu Gly Asp Phe Thr Tyr Asp Glu
 35 40 45
 Ser Thr Gly Thr Phe Ser Met Tyr Trp Glu Asp Gly Val Ser Ser Asp
 50 55 60
 Phe Val Val Gly Leu Gly Trp Thr Thr Gly Ser Ser Lys Ser Ile Thr
 65 70 75 80
 Tyr Ser Ala Gln Tyr Ser Ala Ser Ser Ser Ser Tyr Leu Ala Val
 85 90 95
 15 Tyr Gly Trp Val Asn Ser Pro Gln Ala Glu Tyr Tyr Ile Val Glu Asp
 100 105 110
 Tyr Gly Asp Tyr Asn Pro Cys Ser Ser Ala Thr Ser Leu Gly Thr Val
 115 120 125
 Tyr Ser Asp Gly Ser Thr Tyr Gln Val Cys Thr Asp Thr Arg Arg Thr
 130 135 140
 Arg Pro Ser Ile Thr Gly Thr Ser Thr Phe Thr Gln Tyr Phe Ser Val
 145 150 155 160
 Arg Glu Ser Thr Arg Thr Ser Gly Thr Val Thr Ile Ala Asn His Phe
 165 170 175
 Asn Phe Trp Ala Gln His Gly Phe Gly Asn Ser Asn Phe Asn Tyr Gln
 180 185 190
 20 Val Met Ala Val Glu Ala Trp Asn Gly Val Gly Ser Ala Ser Val Thr
 195 200 205
 Ile Ser Ser
 210

<210> 50
 <211> 366
 <212> PRT
 <213> Phanerochaete chrysosporium

<220>
 <223> Ligninasa 1508163A

25 <400> 50

Met Ala Phe Lys Gln Leu Phe Ala Ala Ile Ser Leu Ala Leu Ser Leu
 1 5 10 15
 Ser Ala Ala Asn Ala Ala Val Ile Glu Lys Arg Ala Thr Cys Ser
 20 25 30
 Asn Gly Lys Thr Val Gly Asp Ala Ser Cys Cys Ala Trp Phe Asp Val
 35 40 45
 Leu Asp Asp Ile Gln Gln Asn Leu Phe His Gly Gly Gln Cys Gly Ala
 50 55 60
 Glu Ala His Glu Ser Ile Arg Leu Val Phe His Asp Ser Ile Ala Ile
 65 70 75 80
 Ser Pro Ala Met Glu Ala Gln Gly Lys Phe Gly Gly Gly Gly Ala Asp
 85 90 95
 Gly Ser Ile Met Ile Phe Asp Asp Ile Glu Thr Ala Phe His Pro Asn
 100 105 110
 Ile Gly Leu Asp Glu Ile Val Lys Leu Gln Lys Pro Phe Val Gln Lys
 115 120 125
 His Gly Val Thr Pro Gly Asp Phe Ile Ala Phe Ala Gly Ala Val Ala
 130 135 140
 Leu Ser Asn Cys Pro Gly Ala Pro Gln Met Asn Phe Phe Thr Gly Arg
 145 150 155 160
 Ala Pro Ala Thr Gln Pro Ala Pro Asp Gly Leu Val Pro Glu Pro Phe
 165 170 175
 His Thr Val Asp Gln Ile Ile Asn Arg Val Asn Asp Ala Gly Glu Phe
 180 185 190
 Asp Glu Leu Glu Leu Val Trp Met Leu Ser Ala His Ser Val Ala Ala
 195 200 205
 Val Asn Asp Val Asp Pro Thr Val Gln Gly Leu Pro Phe Asp Ser Thr
 210 215 220
 Pro Gly Ile Phe Asp Ser Gln Phe Phe Val Glu Thr Gln Leu Arg Gly
 225 230 235 240
 Thr Ala Phe Pro Gly Ser Gly Gly Asn Gln Gly Glu Val Glu Ser Pro
 245 250 255
 Leu Pro Gly Glu Ile Arg Ile Gln Ser Asp His Thr Ile Ala Arg Asp
 260 265 270
 Tyr Arg Thr Ala Cys Glu Trp Gln Ser Phe Val Asn Asn Gln Ser Lys
 275 280 285
 Leu Val Asp Asp Phe Gln Phe Ile Phe Leu Ala Leu Thr Gln Leu Gly
 290 295 300
 Gln Asp Pro Asn Ala Met Thr Asp Cys Ser Asp Val Ile Pro Gln Ser
 305 310 315 320
 Lys Pro Ile Pro Gly Asn Leu Pro Phe Ser Phe Phe Pro Ala Gly Lys
 325 330 335
 Thr Ile Lys Asp Val Glu Gln Ala Cys Ala Glu Thr Pro Phe Pro Thr
 340 345 350
 Leu Thr Thr Leu Pro Gly Pro Glu Thr Ser Val Gln Arg Ile
 355 360 365

20

<210> 51
 <211> 102
 <212> PRT
 <213> Trametes versicolor

<220>
 <223> Ligninasa: manganeso peroxidasa (EC 1.11.1.13)

<220>
 <221> VARIANTE
 <222> 3, 15, 16, 36, 54, 68, 77, 83, 97
 <223> Xaa = Cualquier Aminoácido

25

<400> 51
Val Ala Xaa Pro Asp Gly Val Asn Thr Ala Thr Asn Ala Ala Xaa Xaa
1 5 10 15
Gln Leu Phe Asp Gly Gly Glu Cys Gly Glu Glu Val His Glu Ser Ile
20 25 30
Ala Arg His Xaa Ala Ile Gly Val Ser Asn Cys Pro Gly Ala Pro Gln
35 40 45
Ile Gly Val Ser Asn Xaa Pro Gly Ala Pro Gln Leu Ala Arg Asp Ser
50 55 60
Arg Thr Ala Xaa Glu Trp Gln Ser Leu Leu Ile Glu Xaa Ser Glu Leu
65 70 75 80
Val Pro Xaa Pro Pro Ala Leu Ser Asn Ala Asp Val Glu Gln Ala
85 90 95
Xaa Ala Glu Thr Pro Phe
100

<210> 52
<211> 371
<212> PRT
<213> Phanerochaete chrysosporium (un basidiomicetes)

10 <220>
<223> Lignina peroxidasa (No. de Registro P49012)

<400> 52
Met Ala Phe Lys Gln Leu Phe Ala Ala Ile Thr Val Ala Leu Ser Leu
1 5 10 15
Thr Ala Ala Asn Ala Ala Val Val Lys Glu Lys Arg Ala Thr Cys Ala
20 25 30
Asn Gly Lys Thr Val Gly Asp Ala Ser Cys Cys Ala Trp Phe Asp Val
35 40 45
Leu Asp Asp Ile Gln Ala Asn Met Phe His Gly Gly Gln Cys Gly Ala
50 55 60
Glu Ala His Glu Ser Ile Arg Leu Val Phe His Asp Ser Ile Ala Ile
65 70 75 80
Ser Pro Ala Met Glu Ala Lys Gly Lys Phe Gly Gly Gly Ala Asp
85 90 95
Gly Ser Ile Met Ile Phe Asp Thr Ile Glu Thr Ala Phe His Pro Asn
100 105 110
Ile Gly Leu Asp Glu Val Val Ala Met Gln Lys Pro Phe Val Gln Lys
115 120 125
His Gly Val Thr Pro Gly Asp Phe Ile Ala Phe Ala Gly Ala Val Ala
130 135 140
Leu Ser Asn Cys Pro Gly Ala Pro Gln Met Asn Phe Phe Thr Gly Arg
145 150 155 160
Lys Pro Ala Thr Gln Pro Ala Pro Asp Gly Leu Val Pro Glu Pro Phe
165 170 175
His Thr Val Asp Gln Ile Ile Ala Arg Val Asn Asp Ala Gly Glu Phe
180 185 190
Asp Glu Leu Glu Leu Val Trp Met Leu Ser Ala His Ser Val Ala Ala
195 200 205
Val Asn Asp Val Asp Pro Thr Val Gln Gly Leu Pro Phe Asp Ser Thr
210 215 220
Pro Gly Ile Phe Asp Ser Gln Phe Phe Val Glu Thr Gln Phe Arg Gly
225 230 235 240
Thr Leu Phe Pro Gly Ser Gly Gly Asn Gln Gly Glu Val Glu Ser Gly
245 250 255
Met Ala Gly Glu Ile Arg Ile Gln Thr Asp His Thr Leu Ala Arg Asp
260 265 270
Ser Arg Thr Ala Cys Glu Trp Gln Ser Phe Val Gly Asn Gln Ser Lys

148

275 280 285
 Leu Val Asp Asp Phe Gln Phe Ile Phe Leu Ala Leu Thr Gln Leu Gly
 290 295 300
 Gln Asp Pro Asn Ala Met Thr Asp Cys Ser Asp Val Ile Pro Leu Ser
 305 310 315 320
 Lys Pro Ile Pro Gly Asn Gly Pro Phe Ser Phe Phe Pro Pro Gly Lys
 325 330 335
 Ser His Ser Asp Ile Glu Gln Ala Cys Ala Glu Thr Pro Phe Pro Ser
 340 345 350
 Leu Val Thr Leu Pro Gly Pro Ala Thr Ser Val Ala Arg Ile Pro Pro
 355 360 365
 5 His Lys Ala
 370

<210> 53
 <211> 134
 <212> ADN
 <213> Glycine max (soya)

<220>
 <223> secuencia para subunidad alfa de beta-conglicinina (BCS)
 (genbank No. de Registro AB237643.1)

10 <400> 53
 cctacagact tcaatctggt gatgccctga gagtccccctc aggaaccaca tactatgtgg 60
 tcaaccctga caacaacgaa aatctcagat taataacact cgccataccc gttaacaagc 120
 ctggtagatt tgag 134

<210> 54
 <211> 1195
 <212> ADN
 <213> Secuencia Artificial

15 <220>
 <223> casete de ARNi completo para producir BCS

20 <400> 54
 gcggccgcoct cgagcctaca gacttcaatc tggatgatgcc ctgagagtcc cctcaggaac 60
 cacatactat gtggtcaacc ctgacaacaa cgaaaaatctc agattaataa cactcgccat 120
 acccggttaac aagcctggta gatttgagga tatcgagctc gggtctgttc tcttctcgtc 180
 acactcacia taggggtggcc tatgtattta gcccttcaatg tctctggtag accctatgat 240
 agttttgcaa gccactacca cccttatgct ccctatatatt ctaaccgtga aagcttacta 300
 gtctagtacc ccaattggta aggaaataat tattttcttt ttctctttta gtataaaata 360
 gttaaagtgt gtttaattagt atgattataa taatatagtt gttataattg tgaaaaaata 420
 atttataaat atattgttta cataaacaac atagtaattg aaaaaaatat gacaagtgat 480
 gtgtaagacg aagaagataa aagttgagag taagtatat atttttaattg aatttgatcg 540
 aacatgtaag atgatatact agcattaata tttgttttaa tcataatagt aattctagct 600
 ggtttgatga attaaatatc aatgataaaa tactatagta aaaataagaa taaataaatt 660
 aaaataatat ttttttatga ttaatatgtt attatataat taaatatcta taccattact 720
 aaatatatta gtttaaaagt taataaatat tttgttagaa attccaatct gcttgaatt 780
 tatcaataaaa caaaatatta aataacaagc taaagtaaca aataatatca aactaataga 840
 aacagtaatc taatgtaaca aaacataatc taatgctaataa ataacaagc gtaaagcttt 900
 cacggttaga atatatggga gcataagggt ggtagtggtc tgcaaaacta tcatagggtc 960
 taccagagac attgaaggct aaatacatag gccaccctat tgtgagtgtg acgagaagag 1020
 aaacagcccg agctcgatat cctcaaactc accaggcttg ttaacgggta tggcagagtgt 1080
 tattaatctg agattttcgt tgggtgcagg gttgaccaca tagtatgtgg ttcttgaggg 1140
 gactctcagg gcatcaccag attgaagtct gtaggctcga gtctagagcg gccgc 1195

25 <210> 55
 <211> 332

<212> ADN

<213> Secuencia Artificial

<220>

<223> Fragmento de ARNi de glicinina de glicinina AlbB2 (Genbank No. de Registro AB030495)

<400> 55

5 cattgacgag accatttgca caatgggact tgcacacaac ataggccaga cttcatcacc 60
tgacatcttc aaccttcaag ctggtagcat cacaaccgct accagcctcg acttcccagc 120
cctctcgttg ctcaaactca gtgcccagtt tggatcactc cgcaagaatg ctatgttcgt 180
gccacactac aacctgaacg caaacagcat aatatacgca ttgaatggac gggcattggt 240
acaagtgggtg aattgcaatg gtgagagagt gtttgatgga gagctgcaag agggacaggt 300
gttaactgtg ccacaaaact ttgcgggtggc tg 332

<210> 56

<211> 1591

<212> ADN

<213> Secuencia Artificial

<220>

10 <223> Casete de ARNi completo para producir SP-

<400> 56

15 gcggcgcgct cgagcattga cgagaccatt tgcacaatgg gacttcgcca caacataggc 60
cagacttcat cactgacat cttcaaccct caagctggta gcatcacaac cgctaccagc 120
ctcgacttcc cagccctctc gtggctcaaa ctcaagtgcc agtttgatc actccgcaag 180
aatgctatgt tctgtccaca ctacaacctg aacgcaaaaca gcataatata cgcattgaat 240
ggacgggcat tggtagaagt ggtgaattgc aatgggtgaga gagggtttga tggagagctg 300
caagagggac aggtgttaac tgtgccacaa aactttgcgg tggctggata tgcagctcgg 360
gctgtttctc ttctcgtcac actcacaata ggggtggccta tgtatttagc cttcaatgtc 420
tctggtagac cctatgatag ttttgcaagc cactaccacc cttatgctcc catataattc 480
aaccgtgaaa gcttactagt ctagtacctc aattggtaag gaaataatta ttttcttttt 540
tcccttttagt ataaaatagt taagttagt taattagtat gattataata atatagttgt 600
tataattgtg aaaaaataat ttataaatat attgtttaca taaacaacat agtaattgaa 660
aaaaatatga caagtgatgt gtaagacgaa gaagataaaa gttgagagta agtatattat 720
ttttaatgaa tttgatcgaa catgtaagat gatatactag cattaatatt tgttttaatc 780
ataatagtaa ttctagctgg tttgatgaat taaatatcaa tgataaaaata ctatagtaaa 840
aataagaata aataaattaa aataatattt ttttatgatt aatagtttat tatataatta 900
aataatctata ccattactaa atatttttagt ttaaaagtta ataaatattt tgttagaaat 960
tccaatctgc ttgtaattta tcaataaaca aaatattaaa taacaagcta aagtaacaaa 1020
taatatacaa ctaatagaaa cagtaattcta atgtaacaaa acataatcta atgctaatat 1080
aacaagcgt aaagctttca cggttagaat atatgggagc ataagggtgg tagtggcttg 1140
caaaactatc atagggtcta ccagagacat tgaaggctaa atacatagc caccctattg 1200
tgagtgtgac gagaagagaa acagcccagc ctcgatatcc agccaccgca aagttttgtg 1260
20 gcacagttaa cactgttccc tcttgagct ctccatcaa cactctctca ccattgcaat 1320
tcaccacttg taccaatgcc cgtccattca atgcgtatat tatgctgttt gcgttcaggt 1380
tgtagtgtgg caggaacata gcattcttgc ggagtgatcc aaactgggca ctgagtttga 1440
gccacgagag ggctgggaag tgcaggctgg tagcggttgt gatgctacca gcttgaggtg 1500
tgaagatgtc aggtgatgaa gtctggccta tgttggtggc aagtcccatt gtgcaaatgg 1560
tctcgtcaat gctcaggtct agagcggccg c 1591

<210> 57

<211> 130

<212> ADN

<213> Secuencia Artificial

<220>

25 <223> Gen FAD2-1 (Genbank No. de Registro AB188250)

150

<400> 57

```
gggctgtttc ttttctcgtc acactcacia taggggtggcc tatgtattta gccttcaatg 60
tctctggtag accctatgat agttttgcaa gccactacca cccttatgct cccatatatt 120
ctaaccgtga                                     130
```

151

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:
DAVID A. RANDALL
DICKSTEIN SHAPIRO LLP
2049 CENTURY PARK EAST SUITE 700
LOS ANGELES, CA 90067-3109

PCT

COMMUNICATION IN CASES FOR WHICH
NO OTHER FORM IS APPLICABLE

Date of mailing
(day/month/year) 01 DEC 2010

Applicant's or agent's file reference

A0098.1001-PC

REPLY DUE

See paragraph 1 below

International application No.

PCT/US09/49097

International filing date

(day/month/year) 29 June 2009 (29.06.2009)

Applicant

THE DONALD DANFORTH PLANT SCIENCE CENTER

1. ☐ REPLY DUE within ____ months/days from the above date of mailing

☒ NO REPLY DUE

2. COMMUNICATION:

The International Preliminary Report on Patentability (Chapter II) mailed by the IPEA/US on 22 November 2010 included incorrect indications regarding the claims in Box No. I (Basis of the Report). Attached hereto is the corrected IPRP (Chapter II) accompanied by the amended sheets as the annexes.

Name and mailing address of the IPEA/ US

Mail Stop PCT, Attn: IPEA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer

CATHY K. WORLEY

Telephone No. 571-272-0500

PCT/US2009/049097 22.11.2010

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:
DAVID A. RANDALL
DICKSTEIN SHAPIRO LLP
2049 CENTURY PARK EAST SUITE 700
LOS ANGELES, CA 90067-3109

PCT

NOTIFICATION OF TRANSMITTAL OF
INTERNATIONAL PRELIMINARY
REPORT ON PATENTABILITY
(Chapter II of the Patent Cooperation Treaty)

(PCT Rule 71.1)

Date of mailing
(day/month/year)

22 NOV 2010

Applicant's or agent's file reference

A0098.1001-PC

IMPORTANT NOTIFICATION

International application No.

International filing date (day/month/year)

Priority date (day/month/year)

PCT/US09/49097

29 June 2009 (29.06.2009)

28 June 2008 (28.06.2008)

Applicant

THE DONALD DANFORTH PLANT SCIENCE CENTER

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary report on patentability and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.
4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary report on patentability. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the *PCT Applicant's Guide*.

The applicant's attention is drawn to Article 33(5), which provides that the criteria of novelty, inventive step and industrial applicability described in Article 33(2) to (4) merely serve the purposes of international preliminary examination and that "any Contracting State may apply additional or different criteria for the purposes of deciding whether, in that State, the claimed invention is patentable or not" (see also Article 27(5)). Such additional criteria may relate, for example, to exemptions from patentability, requirements for enabling disclosure, clarity and support for the claims.

Name and mailing address of the IPEA/ US

Mail Stop PCT, Attn: IPEA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Facsimile No.

Authorized officer

CATHY K. WORLEY

Telephone No. (571) 272-0500

Form PCT/IPEA/416 (January 2004)

CORRECTED

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

(Chapter II of the Patent Cooperation Treaty)

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference A0098.1001-PC	FOR FURTHER ACTION		See Form PCT/IPEA/416																								
International application No. PCT/US09/49097	International filing date (day/month/year) 29 June 2009 (29.06.2009)	Priority date (day/month/year) 28 June 2008 (28.06.2008)																									
International Patent Classification (IPC) or national classification and IPC IPC: C12N 15/11(2006.01), 15/14(2006.01), 15/82(2006.01); A01H 5/00(2006.01) USPC: 800/287, 286, 288, 278, 298, 312; 435/410, 419																											
Applicant THE DONALD DANFORTH PLANT SCIENCE CENTER																											
1. This report is the international preliminary examination report, established by this International Preliminary Examining Authority under Article 35 and transmitted to the applicant according to Article 36. 2. This REPORT consists of a total of <u>6</u> sheets, including this cover sheet. 3. This report is also accompanied by ANNEXES, comprising: a. ☒ (sent to the applicant and to the International Bureau) a total of <u>4</u> sheets, as follows: ☒ sheets of the description, claims and/or drawings which have been amended and are the basis of this report and/or sheets containing rectifications authorized by this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions). ☐ sheets which supersede earlier sheets, but which this Authority considers contain an amendment that goes beyond the disclosure in the international application as filed, as indicated in item 4 of Box No. I and the Supplemental Box. b. ☐ (sent to the International Bureau only) a total of (indicate type and number of electronic carrier(s)) _____, containing a sequence listing and/or tables related thereto, in electronic form only, as indicated in the Supplemental Box Relating to Sequence Listing (see Section 802 of the Administrative Instructions).																											
4. This report contains indications relating to the following items: <table style="width: 100%; border: none;"> <tr> <td style="width: 10%;">☒</td> <td style="width: 20%;">Box No. I</td> <td>Basis of the report</td> </tr> <tr> <td>☐</td> <td>Box No. II</td> <td>Priority</td> </tr> <tr> <td>☐</td> <td>Box No. III</td> <td>Non-establishment of opinion with regard to novelty, inventive step and industrial applicability</td> </tr> <tr> <td>☐</td> <td>Box No. IV</td> <td>Lack of unity of invention</td> </tr> <tr> <td>☒</td> <td>Box No. V</td> <td>Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement</td> </tr> <tr> <td>☐</td> <td>Box No. VI</td> <td>Certain documents cited</td> </tr> <tr> <td>☐</td> <td>Box No. VII</td> <td>Certain defects in the international application</td> </tr> <tr> <td>☐</td> <td>Box No. VIII</td> <td>Certain observations on the international application</td> </tr> </table>				☒	Box No. I	Basis of the report	☐	Box No. II	Priority	☐	Box No. III	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability	☐	Box No. IV	Lack of unity of invention	☒	Box No. V	Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement	☐	Box No. VI	Certain documents cited	☐	Box No. VII	Certain defects in the international application	☐	Box No. VIII	Certain observations on the international application
☒	Box No. I	Basis of the report																									
☐	Box No. II	Priority																									
☐	Box No. III	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability																									
☐	Box No. IV	Lack of unity of invention																									
☒	Box No. V	Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement																									
☐	Box No. VI	Certain documents cited																									
☐	Box No. VII	Certain defects in the international application																									
☐	Box No. VIII	Certain observations on the international application																									
Date of submission of the demand 28 April 2010 (28.04.2010)		Date of completion of this report 19 November 2010 (19.11.2010)																									
Name and mailing address of the IPEA/ US Mail Stop PCT, Attn: IPEA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No.		Authorized officer CATHY K. WORLEY Telephone No. (571) 272-0500																									

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.

PCT/US09/49097

Box No. I Basis of the report

1. With regard to the language, this report is based on:

- ☒ the international application in the language in which it was filed.
- ☐ a translation of the international application into English, which is the language of a translation furnished for the purposes of:
- ☐ international search (under Rules 12.3(a) and 23.1(b))
- ☐ publication of the international application (under Rule 12.4(a))
- ☐ international preliminary examination (under Rules 55.2(a) and/or 55.3(a))

2. With regard to the elements of the international application, this report is based on *(replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report)*:

- ☐ the international application as originally filed/furnished
- ☒ the description:
- pages 1-63 as originally filed/furnished
- pages* NONE received by this Authority on _____
- pages* NONE received by this Authority on _____
- ☒ the claims:
- pages NONE as originally filed/furnished
- pages* NONE as amended (together with any statement) under Article 19
- pages* 64-67 received by this Authority on 28 October 2010 (28.10.2010)
- pages* NONE received by this Authority on _____
- ☒ the drawings:
- pages 1/15 - 15/15 as originally filed/furnished
- pages* NONE received by this Authority on _____
- pages* NONE received by this Authority on _____
- ☒ a sequence listing and/or any related table(s) - see Supplemental Box Relating to Sequence Listing.

3. ☐ The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheets/figs _____
- ☐ the sequence listing (*specify*): _____
- ☐ any table(s) related to the sequence listing (*specify*): _____

4. ☐ This report has been established as if (some of) the amendments annexed to this report and listed below had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheets/figs _____
- ☐ the sequence listing (*specify*): _____
- ☐ any table(s) related to the sequence listing (*specify*): _____

5. ☐ This report has been established taking into account the rectification of an obvious mistake authorized by or notified to this Authority under Rule 91 (Rule 70.2(e)).

* If item 4 applies, some or all of those sheets may be marked "superseded."

155

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.
PCT/US09/49097

Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims <u>1-27</u>	YES
	Claims <u>NONE</u>	NO
Inventive Step (IS)	Claims <u>NONE</u>	YES
	Claims <u>1-27</u>	NO
Industrial Applicability (IA)	Claims <u>1-27</u>	YES
	Claims <u>NONE</u>	NO

2. Citations and Explanations (Rule 70.7)
Please See Continuation Sheet

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.

PCT/US09/49097

Supplemental Box Relating to Sequence Listing

Continuation of Box No. 1, item 2:

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, this report was established on the basis of:

a. type of material

- ☒ a sequence listing
☐ table(s) related to the sequence listing

b. format of material

- ☒ on paper
☒ in electronic form

c. time of filing/furnishing

- ☒ contained in the international application as filed
☒ filed together with the international application in electronic form
☐ furnished subsequently to this Authority for the purposes of search and/or examination
☐ received by this Authority as an amendment* on _____

2. ☒ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

* If item 4 in Box No. 1 applies, the listing and/or table(s) related thereto, which form part of the basis of the report, may be marked "superseded."

157

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.
PCT/US09/49097

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:

V. 2. Citations and Explanations:

Claims 1-27 lack an inventive step under PCT Article 33(3) as being obvious over Tada et al (Plant Biotechnology Journal (2003) Vol. 1, pages 411-422) in view of Takaiwa et al (US 2004/0031075; pages 1-11), further in view of Kinney et al (US 2004/0139504; pages 1-23), further in view of Baulcombe et al (US 2007/0130653; pages 1-43, further in view of Maruta (US Patent No. 5,516,668), further in view of Huang (WO 02/064750), and further in view of Budworth et al (US 2006/0120750).

The claims are directed to transgenic dicots in which expression of one or more seed storage proteins has been inhibited or suppressed and in which a "desired" protein is expressed under the control of a seed storage protein promoter.

The general inventive concept is following the "Source" / "Sink" theory wherein the inhibition of expression of endogenous seed storage proteins will make more energy and free amino acids available as a "source", and the simultaneous expression of a protein of interest under the control of a seed-preferred promoter will allow for high expression levels due to providing a "sink" for the extra amino acids to go into. This general idea is taught in Tawaiwa (see left column on page 2) and Tada (see entire article).

Plants that are deficient in one or more seed storage proteins were known in the art (see Tada, Takaiwa, Kinney, Baulcombe, and Maruta), with Kinney specifically teaching soybean plants and antisense inhibition of expression of glycinin and conglycinin (see paragraph 0012 on page 2) although they are silent with regard to the percent reduction in accumulation of the proteins. Baulcombe specifically claim RNAi inhibition of expression of glycinin (see claim 49). Tada teaches rice plants in which storage proteins are significantly reduced, including over a 50% reduction in the amount of glutelin in LGC-1 compared to wild type (see Figure 4 on page 414). It would have been obvious to screen for plants in which the endogenous seed storage proteins were inhibited the most which would have resulted in plants with at least a 75% reduction in the endogenous seed storage proteins.

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.
PCT/US09/49097

Supplemental Box

Plants that express a heterologous protein of interest at high levels were well-known in the art, and are specifically taught by Huang et al (see entire document); including the use of seed-preferred promoters and signal peptides for trafficking into the secretory pathway via the endoplasmic reticulum (see pages 35 - 36 of Huang et al). Huang et al teach over 40% total soluble protein for some of the transgenic lines they produced (see Figure 6). Given that soybean seeds have 35% to 43% protein on a dry weight basis (see paragraph 0007 on page 2 of the instant specification), then if one were to practice this method in soybeans, one would get a yield 14% to 17% of the total dry weight of the seed. Huang teaches purified protein for which specific activities were calculated (see data for lysozyme produced in seeds). Budworth teaches the production of recombinant proteins in plants, and specifically mentions a carboxy-terminal extension that causes a protein to be retained in the ER (see paragraph 0060). Such ER-retention signals (ie. KDEL or HDEL or SEKDEL) were well known in the art. Budworth teaches the production of cellulolytic enzymes including beta-glucosidase (see Table 5 on page 60).

Kinney specifically teaches seed-preferred promoters for use in soybeans, such as the kunitz trypsin inhibitor promoter (KTI) (see Figure 1), the beta-conglycinin promoter (see Figure 2), and glycinin promoter (see paragraph 0055 on page 5). The instant specification discloses glycinin as a protein that compensates for a decrease in beta-conglycinin (paragraph 0010 on page 2). Kinney teaches enhancers (see paragraph 0048 on page 4), and teaches that they can be an innate element of a promoter or it may be heterologous to the promoter.

Takaiwa et al teach the 5'UTR of the GluB-1 gene as having enhancer activity (see right column on page 1), and the GluB-1 gene encodes a seed maturation protein as required in the instant claim 8.

The property of stably storing an enzyme prior to use in plant seeds was well known in the art, because seeds have naturally dried to the point where protease activity is very low, therefore, proteins that are stored in dry seeds are inherently stable.

Claims 1-27 meet the criteria set out in PCT Article 33(4), and thus have industrial applicability because the subject matter claimed can be made or used in industry.

159

REPLACEMENT SHEET

WHAT IS CLAIMED:

1. A transgenic dicot plant comprising:
 - a. a deficiency of one or more seed storage proteins, wherein the deficiency results in an at least 50% reduction in endogenous seed storage protein compared to that of a wild type plant; and
 - b. a heterologous polynucleotide comprising a seed storage protein promoter, an open reading frame comprising an ER signal sequence, a desired protein coding sequence, and an ER retention signal sequence, wherein the open reading frame is operably linked to the seed storage protein promoter; andwherein the seed of the transgenic plant is capable of producing a heterologous protein at a level that is greater than 5% of the total dry weight of the seed.
2. The dicot plant of claim 1, wherein the heterologous polynucleotide further comprises a 5' translational enhancer domain and/or a 3' translational enhancer domain.
3. The dicot plant of claim 1, wherein the ER retention signal sequence induces accretion of the heterologous protein in the lumen of the ER or an ER-derived vesicle.
4. The dicot plant of claim 1, wherein said dicot plant is a member of the Fabaceae family, and optionally Fabales order, and optionally of soya genus.
5. The dicot plant of claim 4, wherein said dicot plant is a member of the Glycine genus.
6. The dicot plant of claim 5, wherein said dicot plant is a soybean.
7. The dicot plant of claim 1, wherein the promoter is derived from Kunitz trypsin inhibitor, soybean lectin, immunodominant soybean allergen P34 or Gly m Bd 30k, glucose binding protein, seed maturation protein, glycinin, or conglycinin.

REPLACEMENT SHEET

8. The dicot plant of claim 2, wherein the translational enhancer domain is derived from Kunitz trypsin inhibitor, Soybean lectin, immunodominant soybean allergen P34 or Gly m Bd 30k, glucose binding protein, seed maturation protein, glycinin, or conglycinin.
9. The dicot plant of claim 1, wherein the storage protein deficiency is of one or more of Kunitz trypsin inhibitor, soybean lectin, immunodominant soybean allergen P34 or Gly m Bd 30k, glucose binding protein, seed maturation protein, glycinin, or conglycinin.
10. The dicot plant of claim 9, wherein the seed has more than a 75% deficiency of the seed's endogenous storage proteins.
11. The dicot plant of claim 1, further comprising a 5' translational enhancer domain and a 3' translational enhancer domain, wherein the promoter and the 3' and the 5' translational enhancer domains are derived from the same storage protein.
12. The dicot plant of claim 1, wherein the heterologous protein accumulates in a seed of the dicot plant to a level that is greater than about 5% of the seed's total dry weight.
13. A seed of the dicot plant of claim 1.
14. A transgenic protein obtained from the seed of claim 13.
15. The transgenic protein of claim 14, wherein the heterologous protein has been purified.
16. The dicot plant of claim 1, wherein the desired protein coding sequence encodes an enzyme or fragment thereof.
17. The dicot plant of claim 16, wherein the enzyme is a cellulolytic enzyme.
18. The dicot plant of claim 17, wherein the cellulolytic enzyme is derived from a fungal source, a bacterial source, an animal source, or a plant source.

REPLACEMENT SHEET

19. The dicot plant of claim 17, wherein the cellulolytic enzyme is a β glucosidase, an Exoglucanase I, an Exoglucanase II, an endoglucanase, a xylanase, a hemicellulase, a ligninase, a ligin peroxidase, or a manganese peroxidase.
20. A product comprising the protein of claim 14.
21. A commercially useful enzyme composition comprising the protein of claim 14.
22. The dicot plant of claim 1, wherein said deficiency of one or more seed storage proteins is due to the presence of an RNAi, an antisense, or a sense fragment of a nucleic acid encoding a seed storage protein.
23. A transgenic dicot plant comprising:
 - a. a deficiency of one or more endogenous plant storage proteins, wherein the deficiency results in an at least 50% reduction in the level of said endogenous plant storage protein compared to a wild type plant; and
 - b. a heterologous polynucleotide comprising a gene regulatory region of a compensating protein operably linked to an open reading frame encoding a sequence comprising an ER signal sequence, a desired protein coding sequence, and an ER retention signal sequence; wherein the seed of the transgenic dicot plant is capable of producing the heterologous protein at a level that is greater than 5% of the total dry weight of the seed.
24. A method of stably storing an enzyme prior to use, the method comprising: storing an endogenous enzyme in seeds of a transgenic dicot plant, wherein the endogenous enzyme comprises at least about 5% of the total dry weight of the seeds and the endogenous enzyme is predominately accumulated in ER protein bodies within the seeds.
25. A method of producing an enhanced amount of a heterologous protein in seeds of a dicot plant, comprising:
 - a. stably transforming a plant cell with a polynucleotide comprising a seed storage protein promoter, an open reading frame comprising an ER signal sequence, a desired

REPLACEMENT SHEET

protein coding sequence, and an ER retention signal sequence, wherein the open reading frame is operably linked to the seed storage protein promoter;

- b. obtaining a homozygous plant line from said stably transformed plant cell;
- c. introgressing said stably transformed plant line to a plant having a deficiency in an endogenous seed storage protein, wherein the deficiency results in an at least 50% reduction in said endogenous seed storage protein compared to that of a wild type plant;
- d. growing seeds of said introgressed transgenic plant, said seeds of the introgressed transgenic plant producing a heterologous protein at a level that is greater than 5% of the total dry weight of the seeds; and
- e. obtaining the heterologous protein from the seeds of the introgressed transgenic plant.

26. The method of claim 25, wherein said deficiency in an endogenous seed storage protein is due to the presence of an RNAi, an antisense, or a sense fragment of a nucleic acid encoding a seed storage protein.

27. A method of producing an enhanced amount of a heterologous protein in seeds of a dicot plant, comprising:

- a. stably transforming a plant cell with a polynucleotide comprising a seed storage protein promoter, an open reading frame comprising an ER signal sequence, a desired protein coding sequence, and an ER retention signal sequence, wherein the open reading frame is operably linked to the seed storage protein promoter, wherein said polynucleotide further comprises an RNAi sequence that is capable of downregulation of an endogenous seed storage protein and the seed storage protein promoter comprises a promoter of a compensating protein;
- b. obtaining a homozygous plant line from said stably transformed plant cell;
- c. growing the seeds of said homozygous plant line; and
- d. obtaining the heterologous protein from the seeds of the homozygous dicot plant.

REIVINDICACIONES

1. Una planta dicotiledónea transgénica, caracterizada porque comprende:

a. una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se traduce en una reducción de al menos 50% en proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre; y

b. un polinucleótido heterólogo que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla; y

en donde la semilla de la planta transgénica es capaz de producir una proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

2. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque el polinucleótido heterólogo comprende además un dominio potenciador de traducción 5' y/o un dominio potenciador de traducción 3'.

3. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la secuencia de señal de retención en ER induce la agregación de la proteína heteróloga en el lumen del ER o una vesícula derivada del ER.

4. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque es un miembro de la familia de las Fabáceas, y opcionalmente del orden Fabales, y opcionalmente del género soya.

5. La planta dicotiledónea de conformidad con la reivindicación 4, caracterizada porque es un miembro del género Glycine.

6. La planta dicotiledónea de conformidad con la reivindicación 5, caracterizada porque es una soya.

7. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque el promotor se deriva de inhibidor de tripsina Kunitz, lectinas de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de

maduración de semilla, glicinina o conglícinina.

8. La planta dicotiledónea de conformidad con la reivindicación 2, caracterizada porque el dominio potenciador de traducción se deriva de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglícinina.

9. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la deficiencia de proteína de almacenamiento es de uno o más de inhibidor de tripsina Kunitz, lectina de soya, alérgeno de soya inmunodominante P34 o Gly m Bd 30k, proteína de unión a glucosa, proteína de maduración de semilla, glicinina o conglícinina.

10. La planta dicotiledónea de conformidad con la reivindicación 9, caracterizada porque la semilla tiene una deficiencia de más del 75% de las proteínas de almacenamiento endógenas de la semilla.

11. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque comprende además un dominio potenciador de traducción 5' y un dominio potenciador de traducción 3', en donde el promotor y los dominios potenciadores de traducción 3' y 5' se derivan de la misma proteína de almacenamiento.

12. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la proteína heteróloga se acumula en una semilla de la planta dicotiledónea hasta un nivel que es mayor que aproximadamente 5% del peso en seco total de la semilla.

13. Una semilla caracterizada porque es de la planta dicotiledónea de conformidad con la reivindicación 1.

14. Una proteína transgénica caracterizada porque se obtiene de la semilla de conformidad con la reivindicación 13.

15. La proteína transgénica de conformidad con la reivindicación 14, caracterizada porque la proteína heteróloga ha sido purificada.

16. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la secuencia de codificación de proteína deseada codifica para una enzima o fragmento de la misma.

17. La planta dicotiledónea de conformidad con la reivindicación 16, caracterizada

porque la enzima es una enzima celulolítica.

18. La planta dicotiledónea de conformidad con la reivindicación 17, caracterizada porque la enzima celulolítica se deriva de una fuente fúngica, una fuente bacteriana, una fuente animal o una fuente vegetal.

5 19. La planta dicotiledónea de conformidad con la reivindicación 17, caracterizada porque la enzima celulolítica es una β -glucosidasa, una exoglucanasa I, una exoglucanasa II, una endoglucanasa, una xilanas, una hemicelulosa, una ligninasa, una ligina peroxidasa o una manganeso peroxidasa.

10 20. Un producto caracterizado porque comprende la proteína de conformidad con la reivindicación 14.

21. Una composición de enzima comercialmente útil, caracterizada porque comprende la proteína de conformidad con la reivindicación 14.

15 22. La planta dicotiledónea de conformidad con la reivindicación 1, caracterizada porque la deficiencia de una o más proteínas de almacenamiento de semilla se debe a la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que codifica para una proteína de almacenamiento de semilla.

23. Una planta dicotiledónea caracterizada porque comprende:

20 a. una deficiencia de una o más proteínas de almacenamiento de semilla, en donde la deficiencia se traduce en una reducción de al menos 50% en proteína de almacenamiento de planta endógena en comparación con aquella de una planta tipo silvestre; y

25 b. un polinucleótido heterólogo que comprende una región reguladora génica de una proteína compensadora enlazada operablemente a un marco de lectura abierto que codifica para una secuencia que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER;

en donde la semilla de la planta dicotiledónea transgénica es capaz de producir la proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla.

24. Un método para almacenar establemente una enzima antes de usar,

caracterizado porque comprende:

almacenar una enzima endógena en semillas de una planta dicotiledónea transgénica, en donde la enzima endógena comprende al menos aproximadamente 5% del peso en seco total de las semillas y la enzima endógena se acumula predominantemente en cuerpos de proteínas ER dentro de las semillas.

25. Un método para producir una cantidad incrementada de una proteína heteróloga en semillas de una planta dicotiledónea, caracterizado porque comprende:

a. transformar establemente una célula vegetal con un polinucleótido que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla;

b. obtener una línea vegetal homocigótica de la célula vegetal establemente transformada;

c. introducir la línea vegetal establemente transformada en una planta que tenga una deficiencia en una o más proteínas de almacenamiento de semilla endógenas, en donde la deficiencia se traduce en una reducción de al menos 50% en la proteína de almacenamiento de semilla endógena en comparación con aquella de una planta tipo silvestre;

d. cultivar semillas de la planta transgénica introgrida, las semillas de la planta transgénica introgrida produciendo una proteína heteróloga a un nivel que es mayor que 5% del peso en seco total de la semilla; y

e. obtener la proteína heteróloga de las semillas de la planta transgénica introgrida.

26. El método de conformidad con la reivindicación 25, caracterizado porque la deficiencia en una proteína de almacenamiento de semilla se debe a la presencia de un ARNi, un fragmento de antisentido o de sentido de un ácido nucleico que codifica para una proteína de almacenamiento de semilla.

27. Un método para producir una cantidad incrementada de una proteína heteróloga en semillas de una planta dicotiledónea, caracterizado porque comprende:

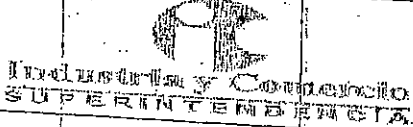
- 5 a. transformar establemente una célula vegetal con un polinucleótido que comprende un promotor de proteína de almacenamiento de semilla, un marco de lectura abierto que comprende una secuencia de señal ER, una secuencia de codificación de proteína deseada y una secuencia de señal de retención en ER, en donde el marco de lectura abierto está enlazado operablemente al promotor de proteína de almacenamiento de semilla, en donde el polinucleótido comprende además una secuencia de ARNi que es capaz de la subregulación de una proteína de almacenamiento de semilla endógena;
- 10 b. obtener una línea vegetal homocigótica de la célula vegetal establemente transformada;
- c. cultivar las semillas de la línea vegetal homocigótica; y
- d. obtener la proteína heteróloga de las semillas de la planta homocigótica.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-009269-00000-0000

Fecha: 2011-01-27 17:02:15 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALI
Act. 411 PRESENTACION Folios: 167



LISTA DE CHEQUEO
ADMISIÓN A TRÁMITE - NUEVAS CREACIONES

PATENTE DE INVENCION ☐ MODELO DE UTILIDAD ☐ Art 33 Decisión 486/00

☐ Indicación que se solicita una patente.
☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
☐ Descripción de la invención
☐ Dibujos de ser estos pertinentes
☐ Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)
Completa ☐ Incompleta ☐

PATENTE DE INVENCION PCT ☒ MODELO DE UTILIDAD PCT ☐ Art.33 Decisión 486/00, Circular Única

☒ Indicación que se solicita una PCT
☒ Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)
☒ Dibujos de ser estos pertinentes
☒ Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)
Completa ☒ Incompleta ☐

DISEÑO INDUSTRIAL ☐ (Art. 119 Decisión 486/00)

☐ Indicación que se solicita Diseño industrial
☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
☐ Representación gráfica y fotográfica del Diseño Industrial o muestra del material que incorpora el diseño
☐ Comprobante de pago de las tasas establecidas
Completa ☐ Incompleta ☐

ESQUEMA DE TRAZADO ☐ (Art. 92 Decisión 486/00)

☐ Indicación que se solicita un esquema de trazado
☐ Datos de identificación del solicitante o de la persona que presenta la solicitud
☐ Representación gráfica de un esquema de trazado
☐ Comprobante de pago de las tasas establecidas
Completa ☐ Incompleta ☐



Bogotá D.C.,
2020

Doctor(a)

CADENA SARMIENTO JUAN PABLO

Apoderado(a) y/o Representante de

THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY OF AGRICULTURE

REFERENCIA	Radicación No.	11 009269
	Solicitud internacional	PCT/US2009/049097
	Fecha inicio fase nacional	28/01/2011
	Trámite	11
	Evento	379
	Actuación	430
	Oficio No.	

03963

Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del Reglamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

1. La solicitud no contiene el poder debidamente otorgado, con el lleno de los requisitos exigidos por los artículos 63 y subsiguientes del Código de Procedimiento Civil. Adicionalmente, deberá presentar el documento que acredita la existencia y representación legal de la sociedad solicitante, cuando ésta fuere persona jurídica.

2. La solicitud no contiene la copia del documento en que consta la cesión del derecho a la patente del inventor al solicitante o a su causante, conforme a lo establecido por el artículo 26 literal k) de la Decisión 486 de la Comisión de la Comunidad Andina. Cuando se trate de documentos otorgados en el extranjero, estos deberán legalizarse de conformidad con lo dispuesto por los artículos 259 y 260 del Código de Procedimiento Civil. Cuando se trate de las declaraciones a que hace referencia la Regla 4.17 ii) del Tratado de Cooperación en Materia de Patentes PCT., el alcance de las tales declaraciones se determinará según indiquen claramente una cesión de derechos del inventor al solicitante o su causante, de conformidad con lo dispuesto en el numeral 5.2.15 del Capítulo Quinto, Título X de la Circular Única de Superintendencia de Industria y Comercio.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Continúa en la página siguiente...



Continuación radicado No: 11 009269

172

Industria y Comercio
SUPERINTENDENCIA

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Unica No.10 de 2001.

JOSE LUIS SALAZAR LOPEZ
Director de Nuevas Creaciones (c)

El anterior requerimiento fue notificado por fijación en lista de fecha 04 MAR. 2011.
adpa11