

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
25. September 2003 (25.09.2003)

PCT

(10) Internationale Veröffentlichungsnummer
WO 03/077643 A2

(51) Internationale Patentklassifikation⁷: **A01H 5/00**,
C12N 15/82, 15/11, 15/29, C07K 14/415

(21) Internationales Aktenzeichen: PCT/EP03/02733

(22) Internationales Anmeldedatum:
17. März 2003 (17.03.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 12 893.6 20. März 2002 (20.03.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): **BASF PLANT SCIENCE GMBH** [DE/DE]; .,
67056 Ludwigshafen (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **BAUER, Jörg**
[DE/DE]; Friedrich-Profit-Str. 56, 67063 Ludwigshafen
(DE).

(74) Anwalt: **DÖRPER, Thomas**; c/o BASF Aktiengesellschaft, ., 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: METHODS FOR INCREASING THE OIL CONTENT IN PLANTS

(54) Bezeichnung: VERFAHREN ZUM ERHÖHEN DES ÖLGEHALTES IN PFLANZEN

(57) Abstract: The invention relates to methods for increasing the oil content in plants by reducing the content of storage proteins by one or more. The invention also relates to the use of plants having a reduced content of storage proteins for producing foodstuffs, feedstuffs, seeds, pharmaceuticals or fine chemicals, particularly for producing oils.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zum Erhöhen des Ölgehaltes in Pflanzen durch Verminderung eines oder mehrerer Speicherproteine. Die Erfindung betrifft ferner die Verwendung von Pflanzen mit einer verminderten Speicherproteingehalt zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien, insbesondere zur Herstellung von Ölen.



WO 03/077643 A2

Verfahren zum Erhöhen des Ölgehaltes in Pflanzen

Beschreibung

Die Erfindung betrifft Verfahren zum Erhöhen des Ölgehaltes in Pflanzen durch Verminderung eines oder mehrerer Speicherproteine. Die Erfindung betrifft ferner die Verwendung von Pflanzen mit einer verminderten Speicherproteingehalt zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien, insbesondere zur Herstellung von Ölen.

Die Erhöhung des Ölgehaltes in Pflanzen und insbesondere in Pflanzensamen ist für die klassische wie für die modernen Pflanzenzüchtung und insbesondere die pflanzliche Biotechnologie von großem Interessen. Bedingt durch den steigenden Verbrauch von Pflanzenölen für Ernährung bzw. industrielle Anwendungen sind Möglichkeiten zur Steigerung bzw. Modifikation von Pflanzenölen zunehmend Gegenstand aktueller Forschung (z.B. Töpfer et al. (1995) Science 268:681-686). Ziel ist dabei insbesondere die Erhöhung des Gehaltes an Fettsäuren in Samenölen.

Auch die aus den pflanzlichen Ölen erhältlichen Fettsäuren sind von besonderem Interesse. Sie kommen beispielsweise als Grundstoffe für Weichmacher, Schmierstoffe, Tenside, Kosmetika usw. zum Einsatz oder werden in der Lebens- und Futtermittelindustrie als wertvoll Grundstoffe eingesetzt. So ist beispielsweise die Bereitstellung von Rapsölen mit Fettsäuren mittlerer Kettenlänge von besonderem Interesse, da diese besonderes in der Tensidherstellung begehrt sind.

Durch die gezielte Modulation pflanzlicher Stoffwechselwege mittels gentechnische Verfahren kann der pflanzlichen Metabolismus in einer Weise vorteilhaft verändert werden, die durch klassische Züchtungsmethoden nur über langwierige Schritte bzw. überhaupt nicht zu erreichen wären. So werden ungewöhnliche Fettsäuren, beispielsweise bestimmte polyungesättigte Fettsäuren, nur in bestimmten Pflanzen bzw. überhaupt nicht in Pflanzen synthetisiert und können deshalb nur durch Expression des entsprechenden Enzyms in transgenen Pflanzen hergestellt werden (z.B. Millar et al. (2000) Trends Plant Sci 5:95-101).

Lipide werden aus Fettsäuren synthetisiert und ihre Synthese kann in zwei Teilmechanismen unterteilt werden, einen quasi "prokaryotischen" und einen quasi "eukaryotischen" (Browse et al. (1986) Biochemical J 235:25-31; Ohlrogge & Browse (1995) Plant Cell 7:957-970). Der prokaryotische Mechanismus ist in den

2

Plastiden lokalisiert und umfasst die Biosynthese der freien Fettsäuren, die in das Cytosol exportiert werden, wo sie als Fettsäureacyl-CoA-Ester in den eukaryotischen Mechanismus eingehen und mit Glycerol-3-phosphat zu Phosphatidsäure (PA) verestert werden. PA ist der Ausgangspunkt für die Synthese von neutralen und polaren Lipiden. Die neutralen Lipide werden dabei über den Kennedy-Weg synthetisiert (Voelker (1996) Genetic Engineering ed.: Setlow 18:111-113; Shankline & Cahoon (1998) Annu Rev Plant Physiol Plant Mol Biol 49:611-649; Frentzen (1998) Lipids 100:161-166).

Samen sind insbesondere darauf angewiesen Energie und Grundbausteine zu speichern, um z.B. eine spätere Keimung zu gewährleisten. Die Speicherung erfolgt in Form von Speicherlipiden, Speicherproteinen oder Stärke (Speicherkohlenhydrat). Je nach Pflanze variieren dabei die Verhältnisse der drei Speichermoleküle zueinander. So enthalten Rapssorten durchschnittlich etwa 48 % Speicherlipide, 19 % Stärke und 21 % Speicherproteine, während Sojabohne 22 % Lipide, 12 % Stärke und 37 % Proteine (jeweils bezogen auf die Trockenmasse) enthält (Biochemistry & Molecular Biology of the Plant ed. Buchanan, Gruissem, Jones 2000, American Society of Plant Physiologists). Die Speichermoleküle werden während der Embryoentwicklung des Samens akkumuliert.

Speicherproteine (infolge auch SP) im Embryo dienen zur Speicherung von Kohlenstoff, Stickstoff und Schwefel, die für das schnelle heterotrophe Wachstum bei Keimung von Samen oder Pollen benötigt werden. Sie haben meist keine enzymatische Aktivität. Diese Proteine werden dabei nur im Embryo während der Samenentwicklung synthetisiert. SP akkumulieren dabei zum einen in Proteinspeichervakuolen (PSV) von unterschiedlich differenzierten Zellen im Embryo bzw. Endosperm. Als weitere Form können sie auch als Proteinkörper assoziiert mit den endoplasmatischen Retikulum (ER) vorliegen (Herman & Larkins (1999) Plant Cell 11:601-613). Dabei werden alle Speicherproteine ursprünglich am rauen ER synthetisiert (Bollini & Chrispeels (1979) Planta 146:487-501). Sie können im ER verbleiben oder können über den Golgiapparat zu anderen Kompartimenten in der Zelle transportiert werden. Im ER erfolgt die Prozessierung und korrekte Faltung der Speicherproteine, sowie deren Oligomerisierung (Vitale & Denecke (1999) Plant Cell 11:615-628). Speicherproteine in Samen wurden historisch gesehen auf Grund ihrer verschiedenen Löslichkeiten eingeordnet. Proteine wurden dabei sequentiell mit verschiedenen Lösungsmitteln aus den Samen extrahiert. Wasserlösliche Proteine wurden dabei als Albumine bezeichnet. Proteine, die in einer Salzlösung löslich sind, wurden als Globuline, und Proteine, die in einem Ethanol-Wasser-

Gemisch löslich sind, wurden als Prolamine bezeichnet. Gluteline, die letzte Gruppe, können nur durch Behandlung mit verdünnten Säuren extrahiert werden (Biochemistry & Molecular Biology of the Plant ed. Buchanan, Gruissem, Jones 2000, American Society of Plant Physiologists). Prolamine kommen dabei spezifisch nur im Endosperm von Gräsern (Poaceae) vor, wo sie die Hauptspeicherproteine darstellen (Ausnahmen sind Reis und Hafer, wo Glutelin-ähnliche und Globuline überwiegen). In Dikotyledonen dominieren dagegen Globuline.

Insgesamt können vier große Genfamilien für Speicherproteine aufgrund ihrer Sequenzen zugeordnet werden: 2S-Albumine (Napin-ähnlich), 7S-Globuline (Phaseolin-ähnlich), 11S/12S-Globuline (Legumin-/Cruciferin-ähnlich) und die Zein-Prolamine.

2S Albumine sind weit verbreitet in Samen von Dikotyledonen, einschließlich wichtiger kommerzieller Pflanzenfamilien wie Fabaceae (z.B. Sojabohne), Brassicaceae (z.B. Raps), Euphorbiaceae (z.B. Rizinus) oder Asteraceae (z.B. Sonnenblume). 2S Albumine sind kompakte globuläre Proteine mit konservierten Cysteinresten, die oft Heterodimere bilden.

7S-Globuline liegen in trimerer Form vor und enthalten keine Cysteinreste. Nach ihrer Synthese werden sie wie die 2S-Albumine in kleinere Fragmente gespalten und glykosyliert. Trotz Unterschiede in der Polypeptidgröße sind die verschiedenen 7S-Globuline hoch konserviert und gehen vermutlich wie die 2S-Albumine auf ein gemeinsames Vorläuferprotein zurück. Die 7S-Globuline sind nur in geringen Mengen in Monokotyledonen vorhanden. In Dikotyledonen ist ihr Anteil immer kleiner verglichen mit den 11S/12S-Globulinen.

11S/12S-Globuline stellen neben den 2S-Albuminen die Hauptfraktion der Speicherproteine in Dikotyledonen. Die hohe Ähnlichkeit der verschiedenen 11S-Globuline aus verschiedenen Pflanzengattungen lassen wiederum auf ein gemeinsames Vorläuferprotein in der Evolution schließen.

Alle Speichermoleküle in Samen werden direkt oder indirekt aus Kohlenhydrat-Vorstufen gebildet. Dabei ist Saccharose die Primärquelle von Kohlenstoff und Energie welche von den Blättern in die sich entwickelnden Samen transportiert wird. So wird z.B. Saccharose zu Glucose-6-phosphat und Pyruvat umgesetzt, die in die Plastiden transportiert und dort zur Synthese von Acetyl-CoA dienen, welches das Ausgangsprodukt für die Synthese der Fettsäuren ist.

Wie in den einzelnen Pflanzen das Gleichgewicht zwischen den einzelnen Speichermolekülen reguliert wird, ist nicht im einzelnen bekannt. So führt die Reduktion der Lipidsynthese zur Erhöhung der Speicherproteine (O'Hara et al. (2000) Trans Biochem Soc 613- 615).

EP-A 0 591 530 beschreibt die Verminderung der Expression eines Speicherproteins in Samen, insbesondere des Speicherproteins Glutelin in Reis, mit dem Ziel die Verwertbarkeit von Reis in Fermentationsprozessen zur Herstellung alkoholischer Getränke zu optimieren. In diesen Prozessen sind Proteine als solche hinderlich. Eine Auswirkung der Verminderung auf den Gehalt anderen pflanzlicher Inhaltstoffe ist nicht beschrieben.

WO 87/47731 beschreibt die Verminderung von einem oder mehreren Speicherproteinen im Samen von Sojabohnen durch Antisense-Technologie. Weiterhin beschrieben ist ein Verfahren, wobei neben dem Speicherprotein noch die Expression eines Gen der Fettsäurebiosynthese (mikrosomale Δ -12 Desaturase [Fad 2-1] Gene) vermindert wird. In Beispiel 2 (S.25/ Z.4-9) geschieht dies durch Cosuppression. Es resultieren transgene Sojabohnenpflanzen mit verminderten Gehalt eines Speicherproteins und einem Gehalt an Ölsäure an der Gesamtmenge von Fettsäuren, das im Verhältnis zu anderen Fettsäuren relativ höher ist als in nicht transgenen Sojabohnenpflanzen. Die Veränderung im Fettsäureprofil ist auf die Suppression des Fad2-1 Gens und nicht des Speicherproteins zurückzuführen. Auch wird keine Veränderung im Gesamtgehalt der Fettsäuren beschrieben.

WO 97/35023 beschreibt Speicherproteine und Verfahren zur Erhöhung des Gehaltes bestimmter Aminosäuren in Pflanzen durch Expression besagter Speicherproteine. Eine Beschreibung der Auswirkung auf andere pflanzliche Metaboliten ist nicht offenbart. WO 97/41239 beschreibt ähnliche Verfahren bezogen auf schwefelreiche Speicherproteine.

WO 98/26064 beschreibt Verfahren zur Verminderung eines oder mehrerer Speicherproteine, bevorzugt in Mais. Beschrieben sind ferner Pflanzen mit einem erhöhten oder veränderten Gehalt an Aminosäuren oder Stärke. Eine Auswirkung auf den Ölgehalt ist nicht beschrieben.

WO 99/15004 beschreibt die Modifikation des Gehaltes von Metaboliten in den Speicherorganen von Pflanzen durch Expression von schwefelreichen Proteinen mit mehr als 10 % schwefelhaltigen Aminosäuren, insbesondere dem Samenalbumin der Sonnenblume (SSA; sunflower seed albumin). Dabei bewirkt die Expression in Lupine

eine Erhöhung des Ölgehaltes (Beispiel 1; S.28/Z.4-5, 17). Umgekehrt bewirkt die Expression des gleichen Gen in Erbse (*Pisum sativum*) eine Erniedrigung (Beispiels 2; S.32/Z.21).

- 5 EP-A 0 620 281 beschreibt eine Veränderung in der Lipidzusammensetzung (Fettsäuremuster) in Raps durch Verminderung der Expression eines Speicherproteins (Napin) mittels Antisense-Technologie. Beispiel 6 (S.11/Z.12-15) beschreibt, dass sich lediglich das Verhältnis der einzelnen Fettsäuren dahingehend
- 10 veränderte, dass der Gehalt an Ölsäure sank und der Gehalt an Linol- und Linolensäure stieg. Es wird explizit daraufhin gewiesen, dass der Gesamtgehalt an Fettsäure unverändert blieb. Entsprechende Daten sind auch in der korrespondierenden Veröffentlichung der Erfinder offenbart (Kohno-Murase J et al.
- 15 (1994) Plant Mol Biol 26(4): 1115-1124). WO 01/81604 beschreibt transgene Pflanzen die eine zytosolische Acetyl-CoA-carboxylase (ACCase) exprimieren.

Es stellte sich daher die Aufgabe alternative Verfahren zur

20 Erhöhung des Ölgehaltes in Pflanzen bereitzustellen.

Diese Aufgabe wird durch die vorliegende Erfindung gelöst. Ein erster Gegenstand der Erfindung umfasst ein Verfahren zum Erhöhen des Gesamtölgehalt in pflanzlichen Organismen, dadurch gekennzeichnet,

25 zeichnet, dass nachfolgende Arbeitsschritte umfasst sind

- a) Verminderung der Proteinmenge eines oder mehrerer Speicherproteine in einem pflanzlichen Organismus oder einem Gewebe, Organ, Teil oder Zelle des besagten pflanzlichen Organismus
- 30 und
- b) Auswahl von pflanzlichen Organismen, bei denen - im Unterschied oder Vergleich zur Ausgangsorganismus - der Gesamtölgehalt in dem besagten pflanzlichen Organismus oder einem
- 35 Gewebe, Organ, Teil oder Zelle des besagten pflanzlichen Organismus erhöht ist.

Das erfindungsgemäße Verfahren kann im Prinzip auf alle Pflanzenarten angewendet werden. Bevorzugt auf solche in denen natür-

40 licherweise ein Speicherprotein exprimiert wird. Die zu den im Rahmen dieser Erfindung offenbarten Speicherproteinsequenzen homologen Sequenzen aus anderen Pflanzen können z.B. durch Datenbanksuche oder Durchmustern von Gen-Banken leicht aufgefunden werden.

"Pflanzlicher Organismus oder von diesem abgeleitete Zellen" meint allgemein jede Zelle, Gewebe, Teile oder Vermehrungsgut (wie Samen oder Früchte) eines Organismus, der zur Photosynthese befähigt ist. Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niederer Pflanzen des Pflanzenreiches. Einjährige, mehrjährige, monocotyledone und dicotyledone Pflanzen sind bevorzugt. Eingeschlossen sind reife Pflanze, Saatgut, Sprosse und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut (zum Beispiel Knollen, Samen oder Früchte) und Kulturen, zum Beispiel Zell- oder Kalluskulturen. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.

"Pflanze" im Rahmen der Erfindung meint alle Gattungen und Arten höherer und niederer Pflanzen des Pflanzenreiches. Eingeschlossen unter dem Begriff sind die reifen Pflanzen, Saatgut, Sprossen und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut, Pflanzenorgane, Gewebe, Protoplasten, Kallus und andere Kulturen, zum Beispiel Zellkulturen, sowie alle anderen Arten von Gruppierungen von Pflanzenzellen zu funktionellen oder strukturellen Einheiten. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium. "Pflanze" umfasst alle einjährigen und mehrjährige, monokotyledonen und dikotyledonen Pflanzen und schließt beispielhaft jedoch nicht einschränkend solche der Gattungen Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solarium, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Zea, Avena, Hordeum, Secale, Triticum, Sorghum, Picea und Populus ein.

Bevorzugt sind Pflanzen nachfolgender Pflanzenfamilien: Amarantaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Sterculiaceae, Tetragoniaceae, Theaceae, Umbelliferae.

Bevorzugte monokotyle Pflanzen sind insbesondere ausgewählt aus den monokotylen Kulturpflanzen, wie zum Beispiel der Familie der Gramineae wie Reis, Mais, Weizen oder andere Getreidearten wie Gerste, Hirse, Roggen, Triticale oder Hafer sowie dem Zuckerrohr sowie alle Arten von Gräsern.

Die Erfindung wird ganz besonders bevorzugt aus dikotyledone pflanzliche Organismen angewendet. Bevorzugte dikotyle Pflanzen sind insbesondere ausgewählt aus den dikotylen Kulturpflanzen, wie zum Beispiel

- Asteraceae wie Sonnenblume, Tagetes oder Calendula und andere mehr,
- Compositae, besonders die Gattung Lactuca, ganz besonders die Art sativa (Salat) und andere mehr,
- Cruciferae, besonders die Gattung Brassica, ganz besonders die Arten napus (Raps), campestris (Rübe), oleracea cv Tastie (Kohl), oleracea cv Snowball Y (Blumenkohl) und oleracea cv Emperor (Broccoli) und weitere Kohlarten; und der Gattung Arabidopsis, ganz besonders die Art thaliana sowie Kresse oder Canola und andere mehr,
- Cucurbitaceae wie Melone, Kürbis oder Zucchini und andere mehr,
- Leguminosae besonders die Gattung Glycine, ganz besonders die Art max (Sojabohne) Soja sowie Alfalfa, Erbse, Bohnengewächsen oder Erdnuss und andere mehr
- Rubiaceae, bevorzugt der Unterklasse Lamiidae wie beispielsweise Coffea arabica oder Coffea liberica (Kaffeestrauch) und andere mehr,
- Solanaceae besonders die Gattung Lycopersicon, ganz besonders die Art esculentum (Tomate) und die Gattung Solanum, ganz besonders die Art tuberosum (Kartoffel) und melongena (Aubergine) sowie Tabak oder Paprika und andere mehr,
- Sterculiaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Theobroma cacao (Kakaostrauch) und andere mehr,
- Theaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Camellia sinensis oder Thea sinensis (Teestrauch) und andere mehr,

- Umbelliferae, besonders die Gattung *Daucus* (ganz besonders die Art *carota* (Karrotte)) und *Apium* (ganz besonders die Art *graveolens dulce* (Selderie)) und andere mehr; und die Gattung *Capsicum*, ganz besonders die Art *annuum* (Pfeffer) und andere mehr,

sowie Lein, Soja, Baumwolle, Hanf, Flachs, Gurke, Spinat, Möhre, Zuckerrübe und den verschiedenen Baum-, Nuss- und Weinarten, insbesondere Banane und Kiwi.

Umfasst sind ferner Schmuckpflanzen, Nutz- oder Zierbäume, Blumen, Schnittblumen, Sträucher oder Rasen. Beispielfhaft aber nicht einschränkend seien zu nennen Angiospermen, Bryophyten wie zum Beispiel *Hepaticae* (Leberblümchen) und *Musci* (Moose); Pteridophyten wie Farne, Schachtelhalm und Lycopoden; Gymnospermen wie Koniferen, Cycaden, Ginkgo und Gnetales, die Familien der *Rosaceae* wie Rose, *Ericaceae* wie Rhododendrons und Azaleen, *Euphorbiaceae* wie Weihnachtssterne und Krokus, *Caryophyllaceae* wie Nelken, *Solanaceae* wie Petunien, *Gesneriaceae* wie das Usambaraveilchen, *Balsaminaceae* wie das Springkraut, *Orchidaceae* wie Orchideen, *Iridaceae* wie Gladiolen, *Iris*, *Freesie* und Krokus, *Compositae* wie Ringelblume, *Geraniaceae* wie Geranien, *Liliaceae* wie der Drachenbaum, *Moraceae* wie Ficus, *Araceae* wie Philodendron und andere mehr.

Pflanzliche Organismen im Sinne der Erfindung sind weiterhin weitere photosynthetisch aktive befähigte Organismen, wie zum Beispiel Algen, Cyanobakterien sowie Moose. Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* oder *Dunaliella*. Insbesondere bevorzugt ist *Synechocystis*.

Am meisten bevorzugt sind Pflanzen, die zur Ölproduktion geeignet sind, wie beispielsweise *Arabidopsis*, Raps, Sonnenblume, Sesam, Färberdistel, Ölbaum, Soja, Mais, Weizen oder verschiedene Nussarten wie beispielsweise Walnuss Mandel. Unter diesen wieder besonders bevorzugt sind die dikotyledonen Pflanzen, insbesondere Raps, Soja und Sonnenblume.

"Öle" umfasst neutrale und/oder polare Lipiden und Mischungen derselben. Beispielfhaft jedoch nicht einschränkend seien die in Tabelle 1 aufgeführten zu nennen.

Tab. 1: Pflanzliche Lipidklassen

Neutrale Lipide	Triacylglycerol (TAG)
	Diacylglycerol (DAG)
	Monoacylglycerol (MAG)
Polare Lipide	Monogalactosyldiacylglycerol (MGDG)
	Digalactosyldiacylglycerol (DGDG)
	Phosphatidylglycerol (PG)
	Phosphatidylcholine (PC)
	Phosphatidylethanolamine (PE)
	Phosphatidylinositol (PI)
	Phosphatidylserin (PS)
	Sulfoquinovosyldiacylglycerol

Neutrale Lipide meint bevorzugt Triacylglyceride. Sowohl die neutralen als auch die polaren Lipide können ein breites Spektrum an verschiedenen Fettsäuren enthalten. Beispielhaft jedoch nicht einschränkend seien die in Tabelle 2 aufgeführten Fettsäuren zu nennen.

Tab. 2: Übersicht über verschiedene Fettsäuren (Auswahl)

¹ Kettenlänge:Anzahl der Doppelbindungen

* nicht natürlicherweise in Pflanzen vorkommend

Nomenklatur ¹	Name
16:0	Palmitinsäure
16:1	Palmitoleinsäure
16:3	Roughaninsäure
18:0	Stearinsäure
18:1	Ölsäure
18:2	Linolsäure
18:3	Linolensäure
γ-18:3	Gamma-Linolensäure*
20:0	Arachidinsäure
22:6	Docosahexanonsäure (DHA) *
20:2	Eicosadienonsäure
20:4	Arachidonsäure (AA) *
20:5	Eicosapentaenosäure (EPA) *
22:1	Erucasäure

Öle meint bevorzugt Samenöle.

„Ölgehalt“ meint die Summe aller Öle nach obiger Definition, bevorzugt die Summe die Triacylglyceride.

"Erhöhung" des Ölgehaltes meint die Steigerung des Gehaltes an Ölen in einer Pflanze oder einem Teil, Gewebe oder Organ derselben, bevorzugt in den Samenorganen der Pflanze. Dabei ist der Ölgehalt im Vergleich zu einer nicht dem erfindungsgemäßen Verfahren unterworfenen, aber ansonsten unveränderten Ausgangspflanze unter ansonsten gleichen Rahmenbedingungen um mindestens 5 %, bevorzugt mindestens 10 %, besonders bevorzugt mindestens 15 %, ganz besonders bevorzugt mindestens 20 %, am meisten bevorzugt mindestens 25 % erhöht. Rahmenbedingungen meint dabei alle für die Keimung, Kultivierung oder Wachstum der Pflanze relevanten Bedingungen wie Boden-, Klima- oder Lichtverhältnisse, Düngung, Bewässerung, Pflanzenschutzmaßnahmen usw.

"Speicherprotein" meint allgemein ein Protein, das mindestens eine der nachfolgenden wesentlichen Eigenschaften aufweist:

- a) Speicherproteine werden im wesentlichen nur im Embryo während der Samenentwicklung exprimiert. "Im wesentlichen" bedeutet dabei, dass in dem besagten Stadium mindestens 50 %, bevorzugt mindestens 70 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % der Gesamtexpression über die Lebensdauer einer Pflanze hinweg stattfindet.
- b) Speicherproteine werden während der Keimung des Samens wieder abgebaut. Dabei beträgt der Abbau während der Keimung mindestens 20 %, bevorzugt mindestens 50 %, ganz besonders bevorzugt mindestens 80 %.
- c) Speicherproteine machen einen wesentlichen Anteil am Gesamtproteingehalt des nicht-keimenden Samens aus. Bevorzugt macht das Speicherprotein in dem nicht-keimenden Samen der Wildtyp-Pflanze mehr als 5 Gew.-% des Gesamtproteins aus, besonders bevorzugt mindestens 10 Gew.-%, ganz besonders bevorzugt mindestens 20 Gew.-%, am meisten bevorzugt mindestens 30 Gew.-%.

Bevorzugt weisen Speicherproteine 2 oder alle der oben genannten wesentlichen Eigenschaften a), b) oder c) auf.

Speicherproteine können in Untergruppen entsprechend weiterer charakteristischer Eigenschaften, wie beispielsweise ihrem Sedimentationskoeffizienten oder der Löslichkeit in verschiedenen Lösungen (Wasser, Salzlösung, Alkohol) aufgeteilt werden. Die Bestimmung des Sedimentationskoeffizienten kann in der dem Fachmann vertrauten Weise mittels Ultrazentrifugation durchgeführt werden (z.B. beschrieben bei Correia JJ (2000) *Methods in Enzymology* 321:81-100).

11

Bevorzugt ist das Speicherprotein ausgewählt aus den Klassen der 2S-Albumine (Napin-ähnlich), 7S-Globuline (Phaseolin-ähnlich), 11S/12S-Globuline (Legumin-/Cruciferin-ähnlich) oder Zein-Prolamine.

Besonders bevorzugte 2S-Albumine umfassen

- a) 2S-Albumine aus Arabidopsis, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 2, 4, 6 oder 8, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 1, 3, 5 oder 7 kodierten Proteine,
- b) 2S-Albumine aus Arten der Gattung Brassica, wie beispielsweise Brassica napus, Brassica nigra, Brassica juncea, Brassica oleracea oder Sinapis alba, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 32, 34, 36, 38, 40, 46 oder 48, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 31, 33, 35, 37, 39, 45 oder 47 kodierten Proteine,
- c) 2S-Albumine aus Soja, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 42 oder 44, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 41 oder 43 kodierten Proteine,
- d) 2S-Albumine aus Sonnenblume (Helianthus annuus), ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 50 oder 52, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 49 oder 51 kodierten Proteine,

sowie die entsprechenden Homologen und funktionellen Äquivalente zu a) oder b) oder c) oder d) aus identischen oder anderen Pflanzenarten, insbesondere Raps, Sonnenblume, Lein, Sesam, Färberdistel, Ölbaum, Soja oder verschiedene Nussarten. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 2S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Wasser aus.

Funktionelle Äquivalente der 2S-Albumine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60 %, bevorzugt mindestens 80 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % zu einer der Proteinsequenzen mit der SEQ ID NO: 2, 4, 6, 8, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 oder 52 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am

meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 2S-Speicherproteins auf.

Weiterhin bevorzugte funktionelle Äquivalente enthalten in den für sie kodierenden Nukleinsäuresequenzen mindestens eins, bevorzugt zwei, besonders bevorzugt 3, am meisten bevorzugt alle der Sequenzmotive ausgewählt aus jeweils einer bestimmten der nachfolgenden Gruppen I , II, III oder IV:

Gruppe I:

1. 5'-TTCCTCGTCTGCGCA-3' (SEQ ID NO: 53)
2. 5'-TCCTCACCAACGCTTCC-3' (SEQ ID NO: 54)
3. 5'-CACCTAAGAGCTTGCCAG-3' (SEQ ID NO: 55)
4. 5'-GTTTGCCCCACCTTGA-3' (SEQ ID NO: 56)

Gruppe II:

1. 5'-ATGGCGAACAAGCTCTTCCTCG-3' (SEQ ID NO: 57)
2. 5'-GCCTTCTTCTTCCTTCTCACCAA-3' (SEQ ID NO: 58)
3. 5'-CCAAAATGTAGGAAGGAGTTTCA-3' (SEQ ID NO: 59)
4. 5'-CTTTGCGTTTGCCCAACCTTGAAAG-3' (SEQ ID NO: 60)
5. 5'-GCAAGTTAGCGTTTGTCCCTTCCAG-3' (SEQ ID NO: 61)
6. 5'-GTCCCTTCCAGAAGACCATGCCTGG-3' (SEQ ID NO: 62)

Gruppe III:

1. 5'-CACATCATGGAGAAGATCCAAG-3' (SEQ ID NO: 63)
2. 5'-AGAAGAAGGACACATGCAGAAAGTGCTGCA-3' (SEQ ID NO: 64)
3. 5'-ATGCCAGTGCAAAGCGCT-3' (SEQ ID NO: 65)
4. 5'-AGATCATCACTGGTGAATGGTGATCGTGTA-3' (SEQ ID NO: 66)
5. 5'-GGCAATGGAAGCACTTAGAGTGTGCTTTGTG-3' (SEQ ID NO: 67)

Gruppe IV:

1. 5'-GACGAGAACCCGATCTCCG-3' (SEQ ID NO: 68)
2. 5'-TGAACCAGTGTCGCATGTTCTCCAGCA-3' (SEQ ID NO: 69)
3. 5'-CTCCAGCAGTGTTGTCAAGAGCT-3' (SEQ ID NO: 70)
4. 5'-CAGTGCCAATGTGAGGCGGTGAAGCAGGTG-3' (SEQ ID NO: 71)
5. 5'-GGCTCAGATTCTCCCTAACGTA-3' (SEQ ID NO: 72)
6. 5'-TGCAACCTTCAATCAAGACGATG-3' (SEQ ID NO: 73)

13

Besonders bevorzugt weisen die funktionelle Äquivalente in ihren Nukleinsäuresequenzen mindestens eines der nachfolgenden Sequenzmotive ausgewählt aus einer bestimmten der nachfolgenden Gruppen V, VI, VII oder VIII auf:

Gruppe V:

1. 5'-ATGGCDAACAAGYTSTTCCTCGTCTGCGCRACCTYTCGCTCTSTGYTTTCMTCTCACCACCGCTTCCTRTCTAYCGCACYGTYGTYGARTTCGAMGAAGATGACGCCASYAACCCCRTRGGYCCAADA-3' (SEQ ID NO: 74)
2. 5'-AAGGAGTTTTCAGMAAKMMCARCACCTAAGAGCTTGCCAG-3' (SEQ ID NO: 75)
3. 5'-TCCAGMAGTGCTGCAACGAGCTTCGCCARGAAGAGCCAGWTTGYGTTTGCCCCACCTTGARACAAGCTGCCARGGCMGTAGRYTCCAGGGACARCA-3' (SEQ ID NO: 76)
4. 5'-CAGGAAAATTTACMAGDCAGCYAAGYACTTGCCYAACTTTGCRAVATCCMGCAAGTTG-3' (SEQ ID NO: 77)

Gruppe VI:

1. 5'-AGCCGTATCTACCAGACYGCTACGCACTTACCTARAGTTTGCAACATYMSGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCATGCCTGGRCCCTCC-3' (SEQ ID NO: 78)
2. 5'-ACCAGGAAGAGCCMCTTTGCGTTTGCCCAACCTTGAAAGGAGCATCCAAAGCSGTTAAACA-3' (SEQ ID NO: 79)
3. 5'-CAGCCGGCCCCATTTTRGGATYCCAAAATGTAGGAAGGAGTTTCARCAAGCACAACACCT-3' (SEQ ID NO: 80)

Gruppe VII:

1. 5'-ATCGCCACACACYTGCWGCCTCCAAATGGCA-3' (SEQ ID NO: 81)
2. 5'-GAAATGAGCGAGCTGARAAGCCCCAWATGCCAGTGCAAAGCGCT-3' (SEQ ID NO: 82)

Gruppe VIII:

1. 5'-ATCACCACCACCATCGAMGACGAGAACCCGATCTCCG-3' (SEQ ID NO: 83)
2. 5'-TACAGGGACAAAGGYTGAACAGTGTCGCATGTTCTCCAGCA-3' (SEQ ID NO: 84)
3. 5'-CAAAACATCGAAGGRCAGTGCCAATGTGAGGCGGTGAAGCAGGTG-3' (SEQ ID NO: 85)
4. 5'-CAACAGTTGAAGCAGARGGCTCAGATTCTCCCTAACGTATGCAACCTTCAAT-3' (SEQ ID NO: 86)

Besonders bevorzugte 7S-Globuline umfassen solche aus Arabidopsis oder Soja, ganz besonders bevorzugt die Proteine mit der SEQ ID NO: 155 oder 157, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 154 oder 156 kodierten Proteine. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische

Eigenschaften wie einen 7S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Salzlösung aus. Als weitere charakteristische Eigenschaft können 7S-Globuline keine Cysteinreste enthalten.

Funktionelle Äquivalente der 7S-Globuline haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60 %, bevorzugt mindestens 80 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % zu einer der Proteinsequenzen mit der SEQ ID NO: 155 oder 157 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 7S-Speicherproteins auf.

Besonders bevorzugte 11S/12S-Globuline umfassen bevorzugt 11S-Globuline aus Raps, Soja und Arabidopsis insbesondere

- a) 11S-Globuline aus Raps mit der SEQ ID NO: 10, 12, 14, 16 oder 18, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 9, 11, 13, 15 oder 17 kodierten Proteine,
- b) die 11S-Globuline aus Soja mit der SEQ ID NO: 20, 22, 24, 26 oder 28, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 19, 21, 23, 25 oder 27 kodierten Proteine,
- c) die 11S-Globuline aus Arabidopsis thaliana mit der SEQ ID NO: 112, 114, 116, 118, 120 oder 122 am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 111, 113, 115, 117, 119 oder 121 kodierten Proteine,

sowie die entsprechenden Homologen und funktionellen Äquivalente aus anderen Pflanzenarten, insbesondere Raps, Sonnenblume, Lein, Sesam, Färberdistel, Ölbaum, Soja oder verschiedene Nussarten, wie beispielsweise das Sonnenblume 11S Speicherprotein (SEQ ID NO: 30), insbesondere das durch die Nukleinsäuresequenz gemäß SEQ ID NO: 29 kodierte Protein. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 11S- oder 12S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Salzlösung (PBS; phosphatgepufferte Salzlösung) und/oder eine schlechte Löslichkeit in Wasser aus.

Funktionelle Äquivalente der 11S- oder 12S Albumine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60 %, bevorzugt mindestens 80 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % zu einer der Proteinsequenzen mit der SEQ ID NO: 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 112, 114, 116, 118, 120 oder 122 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt - die charakteristischen Eigenschaften eines 11S- oder 12S-Speicherproteins auf.

Insbesondere bevorzugt ist das Arabidopsis thaliana 12S Cruciferin Speicherprotein (ATCRU3) gemäß SEQ ID NO: 112 sowie seine Homologen aus anderen Pflanzenarten wie beispielsweise Raps, Soja oder Sonnenblume, wobei diese bevorzugt eine Homologie von mindestens 65 %, bevorzugt mindestens 80 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % zu einer der Proteinsequenzen mit der SEQ ID NO: 112 aufweisen.

Weiterhin bevorzugte funktionelle Äquivalente der Raps 11S/12S-Speicherproteine enthalten in den für sie kodierenden Nukleinsäuresequenzen mindestens eins, bevorzugt zwei, besonders bevorzugt 3 der Sequenzmotive ausgewählt aus der Gruppe IX oder ausgewählt aus der Gruppe X:

Gruppe IX:

1. 5'-GAGGCTGGTCGCATCGA-3' (SEQ ID NO: 87)
2. 5'-TTCCGTGATATGCACCAGAA-3' (SEQ ID NO: 88)
3. 5'-TGGTTAACGACAACGGTGA-3' (SEQ ID NO: 89)

Gruppe X:

1. 5'-GAGGCTGGTCGGGTCTGA-3' (SEQ ID NO: 90)
2. 5'-TTCCGTGACATGCACCAGAA-3' (SEQ ID NO: 91)
3. 5'-TGGTGAACGACAACGGAcA-3' (SEQ ID NO: 92)

Besonders bevorzugt weisen die funktionelle Äquivalente der Raps 11S-Speicherproteine in ihren Nukleinsäuresequenzen ein Sequenzmotive mit der SEQ ID NO: 93 auf:

5'-CARGGBTTCGGTGAYATGCACCAGAAARGTVGARCA-3' (SEQ ID NO: 93)

Weiterhin bevorzugte funktionelle Äquivalente der Soja 11S/12S-Speicherproteine enthalten in den für sie kodierenden Nukleinsäuresequenzen mindestens eins, bevorzugt zwei, besonders bevorzugt 3, am meisten bevorzugt 4, 5 oder 6 der Sequenzmotive ausgewählt aus der Gruppe XI oder ausgewählt aus der Gruppe XII:

Gruppe XI:

1. 5'-ATTGAGACATGGAACCCTAA-3' (SEQ ID NO: 94)
2. 5'-TGGATGTACAACAATGAAGA-3' (SEQ ID NO: 95)
3. 5'-CACTCCTGTTGTTGCCGTTTCT-3' (SEQ ID NO: 96)
4. 5'-ATTCTATCTTGCTGGGAACCAAGA-3' (SEQ ID NO: 97)
5. 5'-GCCATTGTGACAGTGAAAGGAGGTCT-3' (SEQ ID NO: 98)
6. 5'-AGGGAATGCAGTGTTTCGACGGTGAGC-3' (SEQ ID NO: 99)

Gruppe XII:

1. 5'-AAACATGGAACCTCTCAACAC-3' (SEQ ID NO: 100)
2. 5'-TGGACCTATAACACTGGCGA-3' (SEQ ID NO: 101)
3. 5'-TGAACCAGTTGTTGCCATCAGT-3' (SEQ ID NO: 102)
4. 5'-ATTTTACCTTGCTGGGAACCCAGA-3' (SEQ ID NO: 103)
5. 5'-GGTGAGAGAGTGTTTGATGGAGAGCT-3' (SEQ ID NO: 104)

Weiterhin bevorzugte funktionelle Äquivalente der Arabidopsis thaliana 11S/12S-Speicherproteine enthalten in den für sie kodierenden Nukleinsäuresequenzen mindestens eins, bevorzugt zwei, besonders bevorzugt 3, am meisten bevorzugt 4, 5 oder 6 der Sequenzmotive ausgewählt aus der Gruppe XIII:

Gruppe XIII:

1. 5'-AACGAGTGCCAGCTCGA-3' (SEQ ID NO: 123)
2. 5'-TTCCGTGACATGCACAG-3' (SEQ ID NO: 124)
3. 5'-CAGAACCAGCTTGACCGCA-3' (SEQ ID NO: 125)
4. 5'-TAGCCGGAACAACCCACAAGG-3' (SEQ ID NO: 126)
5. 5'-AACCTCGATGACCCGTC-3' (SEQ ID NO: 127)
6. 5'-TGCTGACGTGTACAAGCCA-3' (SEQ ID NO: 128)

Besonders bevorzugt weisen die funktionelle Äquivalente der Arabidopsis 11S/12S-Speicherproteine in ihren Nukleinsäuresequenzen ein Sequenzmotive mit der SEQ ID NO: 129 auf:

- 5'-GACATGCACCAGAARKTRGA-3' (SEQ ID NO: 129)

Besonders bevorzugte Zein-Prolamine umfassen bevorzugt solche aus monokotyledonen Pflanzen, insbesondere Mais, Reis, Hafer, Gerste oder Weizen. Ganz besonders bevorzugt sind die Mais

Zein-Prolamine beschrieben durch SEQ ID NO: 159, 161, 163 oder 165 - insbesondere die durch SEQ ID NO 158, 160, 162 oder 164 kodierten Protein -, das Reis Prolamin gemäß SEQ ID NO: 167 - insbesondere das durch SEQ ID NO 166 kodierte Protein -, das Hafer Prolamin gemäß SEQ ID NO: 169 - insbesondere das durch SEQ ID NO 168 kodierte Proteine-, das Gerste Prolamin gemäß SEQ ID NO: 171 und/oder 172 - insbesondere das durch SEQ ID NO: 170 kodierte Protein - und das das Weizen Prolamin gemäß SEQ ID NO: 174 - insbesondere das durch SEQ ID NO 173 kodierte Protein. Funktionelle Äquivalente zeichnen sich bevorzugt durch eine Löslichkeit in 70%iger ethanolischer Lösung und eine schlechte Löslichkeit in Wasser oder Salzlösung aus.

Funktionelle Äquivalente der Zein-Prolamine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60 %, bevorzugt mindestens 80 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % zu einer der Proteinsequenzen mit der SEQ ID NO: 159, 161, 163, 165, 167, 169, 171, 172 oder 174 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines Zein-Prolamine auf.

Dabei haben die einzelnen Buchstaben in den genannten Sequenzen nachfolgende dem Fachmann vertraute Bedeutung:

Kodierung	Bedeutung		Kodierung	Bedeutung
G	Guanin			
A	Adenin		S (Strong)	C oder G
T	Thymin		W (Weak)	A oder T
C	Cytosin		B (not A)	C, G oder T
R (puRine)	A oder G		D (not C)	G, A oder T
Y (pYrimidine)	C oder T		H (not G)	A, C oder T
M (aMino)	A oder C		V (not T)	A, C oder G
K (Keto)	G oder T		N (aNy)	A, C, G oder T

Funktionelle Äquivalente meint insbesondere natürliche oder künstliche Mutationen der obengenannten Speicherproteine sowie homologe Polypeptide aus anderen Pflanzen, die die gleichen wesentlichen und - bevorzugt - charakteristischen Eigenschaften aufweisen. Bevorzugt sind homologe Polypeptide aus oben beschriebenen bevorzugten Pflanzen. Die zu den im Rahmen

dieser Erfindung offenbarten Speicherproteinen homologen Sequenzen aus anderen Pflanzen - beispielsweise solchen deren genomische Sequenz ganz oder teilweise bekannt ist, wie beispielsweise aus *Arabidopsis thaliana*, *Brassica napus*, *Nicotiana tabacum* oder *Solanum tuberosum* - durch Homologievergleiche aus Datenbanken auffinden. können z.B. durch Datenbanksuche oder Durchmustern von Gen-Banken - unter Verwendung der beispielhaft aufgeführten Speicherprotein-Sequenzen als Suchsequenz bzw. Sonde - leicht aufgefunden werden.

Mutationen umfassen Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Aminosäurereste.

Unter Homologie zwischen zwei Nukleinsäuresequenzen wird die Identität der Nukleinsäuresequenz über die jeweils gesamte Sequenzlänge verstanden, die durch Vergleich mit Hilfe des Programmalgorithmus GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA; Altschul et al. (1997) *Nucleic Acids Res.* 25:3389ff) unter Einstellung folgender Parameter berechnet wird:

Gap Weight: 50

Length Weight: 3

Average Match: 10

Average Mismatch: 0

Beispielhaft wird unter einer Sequenz, die eine Homologie von mindestens 80 % auf Nukleinsäurebasis mit der Sequenz SEQ ID NO: 1 aufweist, eine Sequenz verstanden, die bei einem Vergleich mit der Sequenz SEQ ID NO: 1 nach obigem Programmalgorithmus mit obigem Parametersatz eine Homologie von mindestens 80 % aufweist.

Unter Homologie zwischen zwei Polypeptiden wird die Identität der Aminosäuresequenz über die jeweils gesamte Sequenzlänge verstanden, die durch Vergleich mit Hilfe des Programmalgorithmus GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA) unter Einstellung folgender Parameter berechnet wird:

Gap Weight: 8

Length Weight: 2

Average Match: 2,912

Average Mismatch: -2,003

Beispielhaft wird unter einer Sequenz, die eine Homologie von mindestens 80 % auf Proteinbasis mit der Sequenz SEQ ID NO: 2 aufweist, eine Sequenz verstanden, die bei einem Vergleich mit

der Sequenz SEQ ID NO: 2 nach obigem Programmalgorithmus mit obigem Parametersatz eine Homologie von mindestens 80 % aufweist.

Funktionelle Äquivalente umfasst auch solche Proteine, die durch Nukleinsäuresequenzen kodiert werden, die unter Standardbedingungen mit einer der durch SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119, 121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170 oder 173 beschriebenen, für ein Speicherproteine kodierenden Nukleinsäuresequenz, der zu dieser komplementären Nukleinsäuresequenz oder Teilen der vorgenannten hybridisieren und die wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt - weitere charakteristische Eigenschaften aufweisen.

"Standardhybridisierungsbedingungen" ist breit zu verstehen und meint stringente als auch weniger stringente Hybridisierungsbedingungen. Solche Hybridisierungsbedingungen sind unter anderem bei Sambrook J, Fritsch EF, Maniatis T et al., in Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57) oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit ungefähr 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit ungefähr 0,2X SSC bei 50°C bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M NaCl, pH 7,0). Darüberhinaus kann die Temperatur während des Waschschrilles von niedrig stringenten Bedingungen bei Raumtemperatur, ungefähr 22°C, bis zu stärker stringenten Bedingungen bei ungefähr 65°C angehoben werden. Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50 % Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt. Einige beispielhafte Bedingungen für Hybridisierung und Waschschrill sind infolge gegeben:

(1) Hybridisierungsbedingungen zum Beispiel aus nachfolgenden Bedingungen ausgewählt sein:

- a) 4X SSC bei 65°C,
- b) 6X SSC bei 45°C,

- c) 6X SSC, 100 µg/ml denaturierter, fragmentierte Fischsperma-DNA bei 68°C,
- f) 50 % Formamid, 4X SSC bei 42°C,
- h) 2X oder 4X SSC bei 50°C (schwach stringente Bedingung),
- i) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (schwach stringente Bedingung).

(2) Waschschrirte können zum Beispiel aus nachfolgenden Bedingungen ausgewählt sein:

- a) 0,015 M NaCl/0,0015 M Natriumcitrat/0,1 % SDS bei 50°C.
- b) 0,1X SSC bei 65°C.
- c) 0,1X SSC, 0,5 % SDS bei 68°C.
- d) 0,1X SSC, 0,5 % SDS, 50 % Formamid bei 42°C.
- e) 0,2X SSC, 0,1 % SDS bei 42°C.
- f) 2X SSC bei 65°C (schwach stringente Bedingung).

Die Verminderung der Expression eines Speicherproteins kann auf vielfältige Art und Weise realisiert werden.

"Proteinmenge" meinte die Menge eines Speicherprotein-Polypeptides in einem Organismus, einem Gewebe, einer Zelle oder einem Zellkompartiment. Bevorzugt meint Proteinmenge die Menge eines bestimmten Speicherproteins im Samen einer Pflanze.

"Verminderung" der Proteinmenge meint die mengenmäßige Verminderung der Menge eines Speicherproteins in einem Organismus, einem Gewebe, einer Zelle oder einem Zellkompartiment - beispielsweise durch eines der unten beschriebenen Verfahren - im Vergleich zu dem Wildtyp derselben Gattung und Art auf den dieses Verfahren nicht angewendet wurde, unter ansonst gleichen Rahmenbedingungen (wie beispielsweise Kulturbedingungen, Alter der Pflanzen etc.). Bevorzugt meint Verminderung die Verminderung der Proteinmenge im Samen einer Pflanze.

Der Verminderung beträgt dabei mindestens 10 %, bevorzugt mindestens 10 % oder mindestens 20 %, besonders bevorzugt um mindestens 40 % oder 60 %, ganz besonders bevorzugt um mindestens 70 % oder 80 %, am meisten bevorzugt um mindestens 90 % oder 95 %. Verfahren zur Bestimmung der Proteinmenge sind dem Fachmann bekannt. Beispielhaft seien zu nennen: Das Mikro-Biuret Verfahren (Goa, J (1953) Scand J Clin Lab Invest 5:218-222), die Folin-Ciocalteu-Methode (Lowry OH et al. (1951) J Biol Chem 193:265-275) oder die Messung der Adsorption von CBB G-250 (Bradford MM (1976) Analyt Biochem 72:248-254). Bevorzugt erfolgt

die Verminderung des oder der Speicherproteine und/oder die Erhöhung der Ölproduktion im Samen einer Pflanze.

In einer weiteren bevorzugten Ausführungsform wird die Proteinmenge von mehr als einem Speicherprotein vermindert. Dabei können die verminderten Speicherproteine der gleichen oder unterschiedlichen Klassen, wie 2S-Albuminen, 7S-Globulinen, 11S/12S-Globulinen oder Zein-Prolaminen, angehören. Bevorzugt werden Speicherproteine aus mehr als einer dieser Klassen gleichzeitig in ihrer Proteinmenge vermindert. Die zu vermindernenden Speicherproteine können stark homolog oder weniger stark homolog zu einander sein. Bevorzugt haben mindestens zwei der in ihrer Proteinmenge verminderten Speicherproteine eine Homologie geringer als 90 %, bevorzugt geringer als 70 %, besonders bevorzugt geringer als 60 %, ganz besonders bevorzugt geringer als 50 %.

"Verminderung" oder "vermindern" ist im Zusammenhang mit einem Speicherprotein (bzw. der Menge eines Speicherprotein oder der für dieses kodierende mRNA Menge) weit auszulegen und umfasst die teilweise oder im wesentlichen vollständige, auf unterschiedliche semasiologische Mechanismen beruhende Unterbindung oder Blockierung der Expression eines Speicherproteins in einer Pflanze oder einem davon abgeleiteten Teil, Gewebe, Organ, Zellen oder Samen.

Eine Verminderung im Sinne der Erfindung umfasst die mengenmäßige Verringerung eines Speicherproteins bis hin zu einem im wesentlichen vollständigen Fehlen des Speicherproteins (d.h. fehlende immunologische Nachweisbarkeit des Speicherproteins). Dabei wird die Expression eines bestimmten Speicherproteins in einer Zelle oder einem Organismus bevorzugt um mehr als 50 %, besonders bevorzugt um mehr als 80 %, ganz besonders bevorzugt um mehr als 90 % vermindert.

Erfindungsgemäß sind verschiedene Strategien zur Verminderung der Expression eines Speicherproteins umfasst. Der Fachmann erkennt, dass eine Reihe verschiedener Methoden zur Verfügung stehen, um die Expression eines Speicherproteins in gewünschter Weise zu beeinflussen. Eine Verminderung der Speicherproteinmenge kann beispielsweise jedoch nicht einschränkend unter Verwendung nachfolgender Verfahren realisiert werden:

- a) Einbringung einer doppelsträngigen RNA-Nukleinsäuresequenz (infolge "SP-dsRNA"), wobei die doppelsträngige RNA-Sequenz nachfolgende Elemente umfasst

- i) mindestens eine "sense"-Ribonukleotidsequenz, die im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz und
- ii) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter i) im wesentlichen komplementären sind

oder einer die SP-dsRNA Expression gewährleistenden Expressionskassette oder Expressionskassetten;

- b) Einbringung einer Speicherprotein antisense-RNA-Nukleinsäuresequenzen oder einer deren Expression gewährleistenden Expressionskassette, wobei die Speicherprotein antisense-RNA-Nukleinsäuresequenz im wesentlichen komplementär ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz. Umfasst sind solche Verfahren bei denen die antisense-RNA-Nukleinsäuresequenz gegen ein Speicherprotein-Gen (also genomische DNA-Sequenzen) oder ein Speicherprotein-Gentranskript (also RNA-Sequenzen) gerichtet ist. Umfasst sind auch α -anomere Nukleinsäuresequenzen.
- c) Einbringung einer Speicherprotein antisense-Nukleinsäuresequenzen kombiniert mit einem Ribozym oder einer deren Expression gewährleistenden Expressionskassette
- d) Einbringung von Speicherprotein sense-Nukleinsäuresequenzen zur Induktion einer Kosuppression oder einer deren Expression gewährleistenden Expressionskassette und Verminderung der Expression über eine Cosuppression, wobei die Speicherprotein sense-RNA-Nukleinsäuresequenz im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz
- f) Einbringung DNA-bindender Faktoren gegen Speicherprotein-Gene oder -RNAs oder einer deren Expression gewährleistenden Expressionskassette
- g) Einbringung von den Speicherprotein RNA-Abbau bewirkende virale Nukleinsäuresequenzen und Expressionskonstrukten oder einer deren Expression gewährleistenden Expressionskassette

- h) Einbringung von Konstrukten zur Induktion einer homologen Rekombination an endogenen Speicherprotein-Genen beispielsweise zur Erzeugung von Knockout-Mutanten.
- i) Einführung von Mutationen in endogenen Speicherprotein-Gene zur Erzeugung eines Funktionsverlustes (z.B. Generierung von Stopp-Kodons, Verschiebungen im Leseraster etc.)

Dabei kann jedes einzelne dieser Verfahren eine Verminderung der Speicherprotein-Expression im Sinne der Erfindung bewirken. Auch eine kombinierte Anwendung ist denkbar. Weitere Methoden sind dem Fachmann bekannt und können die Behinderung oder Unterbindung der Prozessierung des Speicherproteins, des Transports des Speicherproteins oder dessen mRNA, Hemmung der Ribosomenanlagerung, Hemmung des RNA-Spleißens, Induktion eines Speicherprotein-RNA abbauenden Enzyms und/oder Hemmung der Translationselongation oder -termination umfassen.

Die einzelnen bevorzugten Verfahren seien infolge kurz beschrieben:

- a) Einbringung einer doppelsträngigen Speicherprotein RNA-Nukleinsäuresequenz (SP-dsRNA)

Das Verfahren der Genregulation mittels doppelsträngiger RNA ("double-stranded RNA interference"; dsRNAi) ist vielfach in tierischen und pflanzlichen Organismen beschrieben (z.B. Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). Auf die in den angegebenen Zitaten beschriebenen Verfahren und Methoden wird ausdrücklich Bezug genommen. Eine effiziente Gensuppression kann auch bei transienter Expression oder nach transienter Transformation beispielsweise infolge einer biolistischen Transformation gezeigt werden (Schweizer P et al. (2000) Plant J 2000 24: 895-903). dsRNAi-Verfahren beruhen auf dem Phänomen, dass durch gleichzeitiges Einbringen von komplementären Strang- und Gegenstrang eines Gentranskriptes eine hocheffiziente Unterdrückung der Expression des entsprechenden Gens bewirkt wird. Der bewirkte Phänotyp kommt dem einer entsprechenden knock-out Mutanten sehr ähnlich (Waterhouse PM et al. (1998) Proc Natl Acad Sci USA 95:13959-64).

24

Das dsRNAi-Verfahren hat sich bei der Verminderung der Speicherprotein-Expression als besonders effizient und vorteilhaft erwiesen. Wie u.a. in WO 99/32619 beschrieben sind dsRNAi-Ansätze klassischen antisense-Ansätzen deutlich überlegen.

Ein weiterer Gegenstand der Erfindung bezieht sich daher auf doppelsträngige RNA-Moleküle (dsRNA-Moleküle), die bei Einführung in eine Pflanze (oder eine davon abgeleitete Zelle, Gewebe, Organ oder Samen) die Verminderung eines Speicherprotein bewirken.

Das doppelsträngiges RNA-Molekül zur Verminderung der Expression eines Speicherprotein Proteins ist dadurch gekennzeichnet, dass es enthält

- a) mindestens eine "sense"-Ribonukleotidsequenz, die im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119, 121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170 oder 173 und
- b) "antisense"-Ribonukleotidsequenzen, die zu den "sense"-Ribonukleotidsequenz unter a) im wesentlichen - bevorzugt vollständig - komplementären sind.

"Im wesentlichen identisch" meint, dass die dsRNA Sequenz auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu der Speicherprotein Zielsequenz aufweisen kann und dennoch eine effizient Verminderung der Expression bewirken. Bevorzugt beträgt die Homologie nach obiger Definition mindestens 65 %, bevorzugt mindestens 75 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt 95 % zwischen der "sense"-Ribonukleotidsequenz einer dsRNA und mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz.

"Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz" meint Fragmente einer RNA oder mRNA transkribiert von einer für ein Speicherprotein kodierenden Nukleinsäuresequenz, bevorzugt von einem Speicherprotein-Gen. Dabei haben die Fragmente bevorzugt eine Sequenzlänge von mindestens 20 Basen, bevorzugt mindestens 50 Basen, besonders bevorzugt mindestens 100 Basen, ganz besonders bevorzugt mindestens 200 Basen, am meisten bevorzugt mindestens

500 Basen. Umfasst ist auch die vollständige transkribierte RNA oder mRNA.

"Im wesentlichen komplementär" meint, dass die "antisense"-Ribonukleotidsequenz auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu dem Komplement der "sense"-Ribonukleotidsequenz aufweisen kann. Bevorzugt beträgt die Homologie mindestens 80 %, bevorzugt mindestens 90 %, ganz besonders bevorzugt mindestens 95 %, am meisten bevorzugt 100 % zwischen der "antisense"-Ribonukleotidsequenz und dem Komplement der "sense"-Ribonukleotidsequenz.

Alternativ, kann eine "im wesentlichen identische" dsRNA auch als Nukleinsäuresequenz definiert werden, die befähigt ist, mit einem Teil eines Speicherprotein Gentranskriptes zu hybridisieren (z.B. in 400 mM NaCl, 40 mM PIPES pH 6,4, 1 mM EDTA bei 50°C oder 70°C für 12 bis 16 h).

Eine 100%ige Sequenzidentität zwischen dsRNA und dem Speicherprotein Gentranskript ist nicht zwingend erforderlich, um eine effiziente Verminderung der Speicherprotein Expression zu bewirken. Das Verfahren ist demnach tolerant gegenüber Sequenzabweichungen, wie sie infolge genetischer Mutationen, Polymorphismen oder evolutionärer Divergenzen vorliegen können. So ist es beispielsweise möglich mit einer einzigen dsRNA, die ausgehend von einer bestimmten Speicherprotein Sequenz eines Organismus generiert wurde, die Expression weiterer homologer Speicherproteine des gleichen Organismus oder aber auch die Expression von Speicherproteinen in anderen verwandten Arten zu unterdrücken. Zu diesem Zweck umfasst die dsRNA bevorzugt Sequenzbereich von Speicherprotein-Gentranskripten, die konservierten Bereichen der einzelnen Speicherproteinfamilien entsprechen. Besagte konservierte Bereiche können aus Sequenzvergleichen abgeleitet werden (vgl. Fig. 1a-7d). Bevorzugt enthält die dsRNA in ihrer "sense"-Ribonukleotidsequenz mindestens eines der Sequenzmotive ausgewählt aus einer der oben definierten Gruppen von Sequenzmotiven I, II, III, IV, V, VI, VII, VIII, IX, X, XI oder XII oder den Sequenzmotiven mit der SEQ ID NO: 93 oder 129.

Am meisten bevorzugt sind doppelsträngige RNA Moleküle beschrieben durch die Ribonukleinsäuresequenz gemäß SEQ ID NO: 106, 108 oder 110.

26

In einer weiteren bevorzugten Ausführungsform enthält die dsRNA mehrere Sequenzabschnitte, die eine gleichzeitige Suppression mehrerer Speicherproteine, bevorzugt von Speicherproteinen aus verschiedenen Klassen - wie beispielsweise einem 2S-Albumin, 7S-Globuline, 11S/12S-Globulin oder die Zein-Prolamine - bewirken. Dazu umfasst die dsRNA bevorzugt

- a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jede dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119, 121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170 oder 173 und wobei zumindest zwei der Speicherprotein-Nukleinsäuresequenzen, zu deren "sense"-RNA-Transkript die besagte "sense"-Ribonukleotidsequenzen im wesentlichen identisch sind, untereinander eine Homologie von unter 90 %, bevorzugt unter 80 %, ganz besonders bevorzugt unter 60 % am meisten bevorzugt unter 50 % über die gesamte Länge ihrer kodierenden Nukleotidsequenz haben, und
- b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen - bevorzugt vollständig - komplementären sind.

Wie oben beschrieben, können die "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen als separate Moleküle oder - bevorzugt - als ein einzelnes, selbstkomplementäres RNA Molekül vorliegen, wobei im letzteren Fall die beiden Stränge bevorzugt über eine verbindende Sequenz ("Linker"), die ganz besonders bevorzugt ein Intron darstellt, miteinander verbunden sind.

Am meisten bevorzugt sind doppelsträngige RNA Moleküle beschrieben durch die Ribonukleinsäuresequenz gemäß SEQ ID NO: 145 oder 147.

Natürlich können, um den gleichen Zweck zu erreichen, auch mehrere individuelle dsRNA Moleküle, die jeweils einen der oben definierten Ribonukleotidsequenzabschnitte umfassen, in die Zelle oder den Organismus eingebracht werden.

Die dsRNA kann aus einem oder mehr Strängen polymerisierter Ribonukleotide bestehen. Es können ferner Modifikationen sowohl des Zucker-Phosphat-Gerüsts als auch der Nukleoside vorliegen. Beispielsweise können die Phosphodiesterbindungen der natürlichen RNA dahingehend modifiziert sein, dass sie zumindest ein Stickstoff oder Schwefel-Heteroatom umfassen. Basen können dahingehend modifiziert werden, dass die Aktivität beispielsweise von Adenosindeaminase eingeschränkt wird. Solche und weitere Modifikationen sind weiter unten bei den Verfahren zur Stabilisierung von antisense-RNA beschrieben.

Die dsRNA kann enzymatisch oder ganz oder teilweise chemisch-synthetisch hergestellt werden.

Die doppelsträngige dsRNA Struktur kann ausgehend von zwei komplementären, separaten RNA-Strängen oder - bevorzugt - ausgehend von einem einzelnen, selbstkomplementären RNA-Strang gebildet werden.

Sollen die zwei separaten RNA-Stränge der dsRNA in einer Zelle oder Pflanze zusammengebracht werden, so kann dies auf verschiedene Art geschehen:

- a) Transformation der Zelle oder Pflanze mit einem Vektor, der beide Expressionskassetten umfasst,
- b) Kotransformation der Zelle oder Pflanze mit zwei Vektoren, wobei der eine die Expressionskassetten mit der "sense"-Ribonukleotidsequenz, der andere die Expressionskassetten mit der "antisense"-Ribonukleotidsequenzen umfasst,
- c) Kreuzung von zwei Pflanzen, die mit jeweils einem Vektor transformiert wurden, wobei der eine die Expressionskassetten mit der "sense"-Ribonukleotidsequenz, der andere die Expressionskassetten mit der "antisense"-Ribonukleotidsequenz umfasst.

In der bevorzugten Ausführungsform wird die dsRNA-Struktur durch einem einzelnen, selbstkomplementären RNA-Strang gebildet. Hier können "sense"- und "antisense"-Ribonukleotidsequenzen durch eine verbindende Sequenz ("Linker") verknüpft sein und beispielsweise eine Haarnadelstruktur ausbilden. Bevorzugt ist die verbindende Sequenz ein Intron.

Die Nukleinsäuresequenz kodierend für eine dsRNA kann weitere Elemente beinhalten, wie beispielsweise Transkriptionsterminationssignale oder Polyadenylierungssignale.

Die Bildung der RNA Duplex kann entweder außerhalb der Zelle oder innerhalb derselben initiiert werden. Wie in WO 99/53050 kann die dsRNA auch eine Haarnadelstruktur umfassen, indem "sense"- und "antisense"-Strang durch einen "Linker" (beispielsweise ein Intron) verbunden werden. Die selbstkomplementären dsRNA-Strukturen sind bevorzugt, da sie lediglich die Expression eines Konstruktes erfordern und die komplementären Stränge stets in einem äquimolaren Verhältnis umfassen.

Die Expressionskassetten kodierend für den "antisense"- oder "sense"-Strang einer dsRNA oder für den selbstkomplementären-Strang der dsRNA, werden bevorzugt in einen Vektor insertiert und mit den unten beschriebenen Verfahren stabil (beispielsweise unter Verwendung von Selektionsmarkern) in das Genom einer Pflanze insertiert, um eine dauerhafte Expression der dsRNA zu gewährleisten.

Die dsRNA kann unter Verwendung einer Menge eingeführt werden, die zumindest ein Kopie pro Zelle ermöglicht. Höhere Mengen (z.B. mindestens 5, 10, 100, 500 oder 1000 Kopien pro Zelle) können ggf. eine effizienter Verminderung bewirken.

Die dsRNA kann entweder in vivo oder in vitro synthetisiert werden. Dazu kann eine DNA-Sequenz kodierend für eine dsRNA in eine Expressionskassette unter Kontrolle mindestens eines genetischen Kontrollelementes (wie beispielsweise Promotor, Enhancer, Silencer, Splice-Donor oder -Akzeptor, Polyadenylierungssignal) gebracht werden. Entsprechend vorteilhafte Konstruktionen sind weiter unten beschrieben. Eine Polyadenylierung ist nicht erforderlich, ebenso müssen keine Elemente zur Initiierung einer Translation vorhanden sein.

Eine dsRNA kann chemisch oder enzymatisch synthetisiert werden. Dazu können zelluläre RNA Polymerasen oder Bakteriophagen RNA Polymerasen (wie z.B. T3-, T7- oder SP6 RNA-Polymerase) verwendet werden. Entsprechende Verfahren zu in vitro Expression von RNA sind beschrieben (WO 97/32016; US 5,593,874; US 5,698,425, US 5,712,135, US 5,789,214, US 5,804,693). Eine chemisch oder enzymatisch in vitro synthetisierte dsRNA kann vor der Einführung in eine Zelle, Gewebe oder Organismus aus dem Reaktionsgemisch beispielsweise durch Extraktion, Präzipitation, Elektrophorese,

Chromatographie oder Kombinationen dieser Verfahren ganz oder teilweise aufgereinigt werden. Die dsRNA kann unmittelbar in die Zelle eingeführt werden oder aber auch extrazellulär (z.B. in den interstitiellen Raum) appliziert werden.

Bevorzugt wird die Pflanze jedoch stabil mit einem Expressionskonstrukt, das die Expression der dsRNA realisiert, transformiert. Entsprechende Verfahren sind weiter unten beschrieben.

- b) Einbringung einer Speicherprotein antisense-Nukleinsäuresequenz

Verfahren zur Suppression eines bestimmten Proteins durch Verhinderung der Akkumulation seiner mRNA durch die "antisense"-Technologie sind vielfach auch in Pflanzen - beschrieben (Sheehy et al. (1988) Proc Natl Acad Sci USA 85: 8805-8809; US 4,801,340; Mol JN et al. (1990) FEBS Lett 268(2):427-430). Das antisense Nukleinsäuremolekül hybridisiert bzw. bindet mit der zellulären mRNA und/oder genomischen DNA kodierend für das zu supprimierende Speicherprotein-Zielprotein. Dadurch wird die Transkription und/oder Translation des Zielproteins unterdrückt. Die Hybridisierung kann auf konventionelle Art über die Bildung einer stabilen Duplex oder - im Fall von genomischer DNA - durch Bindung des antisense Nukleinsäuremoleküls mit der Duplex der genomischen DNA durch spezifische Wechselwirkung in der großen Furche der DNA-Helix entstehen.

Eine antisense Nukleinsäuresequenz geeignet zur Verminderung eines Speicherproteins enthält einen "antisense"-RNA-Strang umfassend mindestens eine Ribonukleotidsequenz, die im wesentlichen komplementär ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119, 121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170 oder 173.

"Im wesentlichen komplementär" meint, dass die antisense-RNA Sequenz auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu dem Komplement der Speicherprotein Zielsequenz aufweisen kann. Bevorzugt beträgt die Homologie nach obiger Definition mindestens 75 %, bevorzugt mindestens 85 %, ganz besonders bevorzugt mindestens 95 %, am meisten bevorzugt 98 % zwischen dem "antisense"-RNA-Molekül und dem Komplement mindestens eines Teil des "sense"-RNA-Transkriptes

einer Speicherprotein-Nukleinsäuresequenz. Das Komplement kann entsprechend den Basenpaarregeln von Watson und Crick in der dem Fachmann geläufigen Weise aus den entsprechenden Sequenzen abgeleitet werden.

"Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz" meint Fragmente einer RNA oder mRNA transkribiert von einer für ein Speicherprotein kodierenden Nukleinsäuresequenz, bevorzugt von einem Speicherprotein-Gen. Dabei haben die Fragmente bevorzugt eine Sequenzlänge von mindestens 20 Basen, bevorzugt mindestens 50 Basen, besonders bevorzugt mindestens 100 Basen, ganz besonders bevorzugt mindestens 200 Basen, am meisten bevorzugt mindestens 500 Basen. Umfasst ist auch die vollständige transkribierte RNA oder mRNA.

Die antisense Nukleinsäuresequenz kann zu der gesamten transkribierten mRNA des besagten Proteins komplementär sein, sich auf die kodierende Region beschränken oder nur aus einem Oligonukleotid bestehen, das zu einem Teil der kodierenden oder nicht-kodierenden Sequenz der mRNA komplementär ist. So kann das Oligonukleotid beispielsweise komplementär zu der Region sein, die den Translationsstart für das besagte Protein umfasst. Antisense-Nukleinsäuresequenzen können eine Länge von zum Beispiel 5, 10, 15, 20, 25, 30, 35, 40, 45 oder 50 Nukleotide haben, können aber auch länger sein und mindestens 100, 200, 500, 1000, 2000 oder 5000 Nukleotide umfassen. Antisense-Nukleinsäuresequenzen können rekombinant exprimiert oder chemisch bzw. enzymatisch unter Verwendung von dem Fachmann bekannten Verfahren synthetisiert werden. Bei der chemischen Synthese können natürlich oder modifizierte Nukleotide verwendet werden.

In einer weiteren bevorzugten Ausführungsform kann die Expression eines Speicherproteins durch Nukleotidsequenzen inhibiert werden, die komplementär zu der regulatorischen Region eines Speicherprotein-Gens (z.B. einem Speicherprotein Promoter und/oder Enhancer) sind und triple-helikale Strukturen mit der dortigen DNA-Doppelhelix ausbilden, so dass die Transkription des Speicherprotein-Gens vermindert wird. Entsprechende Verfahren sind beschrieben (Helene C (1991) Anticancer Drug Res 6(6):569-84; Helene C et al. (1992) Ann NY Acad Sci 660:27-36; Maher LJ (1992) Bioassays 14(12):807- 815).

In einer weiteren Ausführungsform kann das antisense Nukleinsäuremolekül eine α -anomere Nukleinsäure sein. Derartige α -anomere Nukleinsäuremoleküle bilden spezifische doppelsträngige Hybride mit komplementärer RNA in denen - im Unterschied zu den konventionellen β -Nukleinsäuren - die beiden Stränge parallel zueinander verlaufen (Gautier C et al. (1987) Nucleic Acids Res 15:6625-6641). Das antisense Nukleinsäuremolekül kann ferner auch 2'-O-Methylribonukleotide (Inoue et al. (1987) Nucleic Acids Res 15:6131-6148) oder chimäre RNA-DNA Analoge beinhalten (Inoue et al. (1987) FEBS Lett 215:327-330).

- c) Einbringung einer Speicherprotein antisense-Nukleinsäuresequenz kombiniert mit einem Ribozym

Vorteilhaft kann die oben beschriebene antisense-Strategie mit einem Ribozym-Verfahren gekoppelt werden. Katalytische RNA-Moleküle oder Ribozyme können an jede beliebige Ziel-RNA angepasst werden und spalten das Phosphodiester-Gerüst an spezifischen Positionen, wodurch die Ziel-RNA funktionell deaktiviert wird (Tanner NK (1999) FEMS Microbiol Rev 23(3):257-275). Das Ribozym wird dadurch nicht selber modifiziert, sondern ist in der Lage, weitere Ziel-RNA-Moleküle analog zu spalten, wodurch es die Eigenschaften eines Enzyms erhält. Der Einbau von Ribozymsequenzen in "antisense"-RNAs verleiht eben diesen "antisense"-RNAs diese enzymähnliche, RNA-spaltende Eigenschaft und steigert so deren Effizienz bei der Inaktivierung der Ziel-RNA. Die Herstellung und Verwendung entsprechender Ribozym-"antisense"-RNA-Moleküle ist beispielsweise beschrieben bei Haseloff et al. (1988) Nature 334: 585-591.

Auf diese Art können Ribozyme (z.B. "Hammerhead"-Ribozyme; Haselhoff und Gerlach (1988) Nature 334:585-591) verwendet werden, um die mRNA eines zu supprimierenden Enzyms - z.B. Speicherprotein - katalytisch zu spalten und die Translation zu verhindern. Die Ribozym-Technologie kann die Effizienz einer antisense-Strategie erhöhen. Verfahren zur Expression von Ribozymen zur Verminderung bestimmter Proteine sind beschrieben in (EP 0 291 533, EP 0 321 201, EP 0 360 257). In pflanzlichen Zellen ist eine Ribozym-Expression ebenfalls beschrieben (Steinecke P et al. (1992) EMBO J 11(4):1525-1530; de Feyter R et al. (1996) Mol Gen Genet. 250(3):329-338). Geeignete Zielsequenzen und Ribozyme können zum Beispiel wie bei "Steinecke P, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds, Academic Press, Inc. (1995), S.449-460" beschrieben, durch Sekundärstrukturberechnungen von Ribozym-

und Ziel-RNA sowie durch deren Interaktion bestimmt werden (Bayley CC et al. (1992) Plant Mol Biol 18(2):353-361; Lloyd AM and Davis RW et al. (1994) Mol Gen Genet 242(6):653-657). Beispielsweise können Derivate der Tetrahymena L-19 IVS RNA konstruiert werden, die komplementäre Bereiche zu der mRNA des zu supprimierenden Speicherprotein Proteins aufweisen (siehe auch US 4,987,071 und US 5,116,742). Alternativ können solche Ribozyme auch über einen Selektionsprozess aus einer Bibliothek diverser Ribozyme identifiziert werden (Bartel D und Szostak JW (1993) Science 261:1411-1418).

- d) Einbringung einer Speicherprotein sense-Nukleinsäuresequenz zur Induktion eines Kosuppression

Die Expression einer Speicherprotein Nukleinsäuresequenz - oder eines teils derselben - in sense-Orientierung kann zu einer Kosuppression des entsprechenden homologen, endogenen Gens führen. Die Expression von sense-RNA mit Homologie zu einem endogenen Gen kann die Expression desselben vermindern oder ausschalten, ähnlich wie es für antisense Ansätze beschrieben wurde (Jorgensen et al. (1996) Plant Mol Biol 31(5):957-973; Goring et al. (1991) Proc Natl Acad Sci USA 88:1770-1774; Smith et al. (1990) Mol Gen Genet 224:447-481; Napoli et al. (1990) Plant Cell 2:279-289; Van der Krol et al. (1990) Plant Cell 2:291-99). Dabei kann das eingeführte Konstrukt das zu vermindern, homologe Gen ganz oder nur teilweise repräsentieren. Die Möglichkeit zur Translation ist nicht erforderlich. Die Anwendung dieser Technologie auf Pflanzen ist beispielsweise beschrieben bei Napoli et al. (1990) The Plant Cell 2: 279-289 und in US 5,034,323.

Eine "sense"-Nukleinsäuresequenz geeignet zur Verminderung eines Speicherproteins enthält einen "sense"-RNA-Strang umfassend mindestens eine Ribonukleotidsequenz, die im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkript einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119, 121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170 oder 173.

"Im wesentlichen identisch" meint, dass die sense-RNA Sequenz auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu dem Komplement der Speicherprotein Zielsequenz aufweisen kann. Bevorzugt beträgt die Homologie nach obiger Definition mindestens 65 %, bevorzugt mindestens

75 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt 95 % zwischen dem "sense"-RNA-Molekül und dem "sense"-RNA-Transkript mindestens eines Teils einer Speicherprotein-Nukleinsäuresequenz.

"Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz" meint Fragmente einer RNA oder mRNA transkribiert von einer für ein Speicherprotein kodierenden Nukleinsäuresequenz, bevorzugt von einem Speicherprotein-Gen. Dabei haben die Fragmente bevorzugt eine Sequenzlänge von mindestens 20 Basen, bevorzugt mindestens 50 Basen, besonders bevorzugt mindestens 100 Basen, ganz besonders bevorzugt mindestens 200 Basen, am meisten bevorzugt mindestens 500 Basen. Umfasst ist auch die vollständige transkribierte RNA oder mRNA.

- e) Einbringung DNA-bindende Faktoren gegen Speicherprotein Gene oder RNAs

Eine Verminderung einer Speicherprotein Genexpression ist auch mit spezifischen DNA-bindenden Faktoren z.B. mit Faktoren vom Typus der Zinkfingertranskriptionsfaktoren möglich. Diese Faktoren lagern sich an die genomische Sequenz des endogenen Zielgens, bevorzugt in den regulatorischen Bereichen, an und bewirken eine Repression des endogenen Gens. Die Verwendung eines solchen Verfahrens ermöglicht die Verminderung der Expression eines endogenen Speicherprotein Gens, ohne dass dessen Sequenz gentechnisch manipuliert werden muss. Entsprechende Verfahren zur Herstellung entsprechender Faktoren sind beschrieben (Dreier B et al. (2001) J Biol Chem 276(31):29466-78; Dreier B et al. (2000) J Mol Biol 303(4):489-502; Beerli RR et al. (2000) Proc Natl Acad Sci USA 97 (4):1495-1500; Beerli RR et al. (2000) J Biol Chem 275(42):32617-32627; Segal DJ and Barbas CF 3rd. (2000) Curr Opin Chem Biol 4(1):34-39; Kang JS and Kim JS (2000) J Biol Chem 275(12):8742-8748; Beerli RR et al. (1998) Proc Natl Acad Sci USA 95(25):14628-14633; Kim JS et al. (1997) Proc Natl Acad Sci USA 94(8):3616-3620; Klug A (1999) J Mol Biol 293(2):215-218; Tsai SY et al. (1998) Adv Drug Deliv Rev 30(1-3):23-31; Mapp AK et al. (2000) Proc Natl Acad Sci USA 97(8):3930-3935; Sharrocks AD et al. (1997) Int J Biochem Cell Biol 29(12):1371-1387; Zhang L et al. (2000) J Biol Chem 275(43):33850-33860).

Die Selektion dieser Faktoren kann unter Verwendung eines beliebigen Stückes eines Speicherprotein-Gens erfolgen. Bevorzugt liegt dieser Abschnitt in einem innerhalb der

Speicherproteine konservierten Sequenzbereichen oder im Bereich der Promotorregion. Für eine Genunterdrückung kann er aber auch im Bereich der kodierenden Exons oder Introns liegen.

Die DNA-bindenden Faktoren können beispielsweise gegen Sequenzen gerichtet sein, die in verschiedenen Speicherprotein-Genen konserviert vorliegen. Bevorzugt enthält ist der DNA-bindende Faktor gegen ein Sequenzmotive ausgewählt aus einer der oben definierten Gruppen von Sequenzmotiven I, II, III, IV, V, VI, VII, VIII, IX, X, XI oder XII oder den Sequenzmotiven mit der SEQ ID NO: 93 oder 129 gerichtet. Ferner können aber auch Sequenzen im Promotorbereich genutzt werden, die bei vielen Speicherproteinen auftreten. Beispielsweise jedoch nicht einschränkend seinen nachfolgende Sequenzen zu nennen:

i) 5'-CATGCATG-3' (SEQ ID NO: 130)

ii) 5'-GCCACYTC-3' (SEQ ID NO: 131)

Weitere geeignete Abschnitte sind für den Fachmann mittels Datenbankabfrage aus der Genbank oder - ausgehend von einer Speicherprotein cDNA, deren Gen nicht in der Genbank vorhanden ist, durch Durchmusterung einer genomischen Bibliothek nach korrespondierenden genomischen Klonen erhältlich.

Die Genexpression kann auch durch maßgeschneiderte, niedermolekulare synthetische Verbindungen unterdrückt werden, beispielsweise vom Polyamid-Typ (Dervan PB und Bürli RW (1999) Current Opinion in Chemical Biology 3:688-693; Gottesfeld JM et al. (2000) Gene Expr 9(1-2):77-91). Diese Oligomere bestehen aus den Bausteinen 3-(Dimethylamino)propylamin, N-Methyl-3-hydroxypyrrol, N-Methylimidazol und N-Methylpyrrole und können an jedes Stück doppelsträngiger DNA so angepasst werden, dass sie sequenzspezifisch in die große Furche binden und die Expression der dortigen Gensequenzen blockieren. Entsprechende verfahren sind beschrieben (siehe unter anderem Bremer RE et al. (2001) Bioorg Med Chem. 9(8):2093-103; Ansari AZ et al. (2001) Chem Biol. 8(6):583-92; Gottesfeld JM et al. (2001) J Mol Biol 309(3):615-29; Wurtz NR et al. (2001) Org Lett 3(8):1201-3; Wang CC et al. (2001) Bioorg Med Chem 9(3):653-7; Urbach AR und Dervan PB (2001) Proc Natl Acad Sci USA 98(8):4343-8; Chiang SY et al. (2000) J Biol Chem 275(32):24246-54).

- g) Einbringung von den Speicherprotein RNA-Abbau bewirkende virale Nukleinsäuresequenzen und Expressionskonstrukten

Die Speicherprotein Expression kann effektiv auch durch Induktion des spezifischen Speicherprotein RNA-Abbaus durch die Pflanze mit Hilfe eines viralen Expressionssystems (Amplikon) (Angell SM et al. (1999) Plant J 20(3):357-362) realisiert werden. Diese Systeme - auch als "VIGS" (viral induced gene silencing) bezeichnet - bringen Nukleinsäuresequenzen mit Homologie zu den zu supprimierenden Transkripten mittels viraler Vektoren in die Pflanze ein. Die Transkription wird sodann - vermutlich vermittelt durch pflanzliche Abwehrmechanismen gegen Viren - abgeschaltet. Vermutlich sind die zugrunde liegenden Mechanismen denen der durch doppelsträngige RNA bewirkten Effekten ähnlich. Entsprechende Techniken und Verfahren sind beschrieben (Ratcliff F et al. (2001) Plant J 25(2):237-45; Fagard M und Vaucheret H (2000) Plant Mol Biol 43(2-3):285-93; Anandalakshmi R et al. (1998) Proc Natl Acad Sci USA 95(22):13079-84; Ruiz MT (1998) Plant Cell 10(6): 937-46). Für die Auswahl der entsprechenden Sequenzen gelten die gleichen Regeln wie für die Bestimmung des "sense"-Stranges bei doppelsträngiger RNA oder der Cosuppression (s.o.). Entsprechend geeignete Systeme zur Suppression der Genexpression unter Verwendung viraler Expressionssysteme sind beispielsweise beschrieben in WO 99/15682 und WO 98/36083.

- h) Einbringung von Konstrukten zur Induktion einer homologen Rekombination an endogenen Speicherprotein-Genen beispielsweise zur Erzeugung von Knockout-Mutanten.

Zur Herstellung eines homolog rekombinanten Organismus mit verminderter Speicherprotein-Aktivität verwendet man beispielsweise ein Nukleinsäurekonstrukt, das zumindest einen Teil eines endogenen Speicherprotein Gens enthält, das durch eine Deletion, Addition oder Substitution mindestens eines Nukleotids so verändert wird, so dass die Funktionalität vermindert oder gänzlich aufgehoben wird. Die Veränderung kann auch die regulativen Elemente (z.B. den Promotor) des Gens betreffen, so dass die kodierende Sequenz unverändert bleibt, eine Expression (Transkription und/oder Translation) jedoch unterbleibt und vermindert wird.

Bei der konventionellen homologen Rekombination ist die veränderte Region an ihrem 5'- und 3'-Ende von weiteren Nukleinsäuresequenzen flankiert, die eine ausreichende Länge für die Ermöglichung der Rekombination aufweisen müssen. Die Länge

liegt in der Regel in einem Bereich von mehreren einhundert Basen bis zu mehreren Kilobasen (Thomas KR und Capecchi MR (1987) Cell 51:503; Strepp et al. (1998) Proc Natl Acad Sci USA 95(8):4368-4373). Für die homologe Rekombination wird der Wirtsorganismus - zum Beispiel eine Pflanze - mit dem Rekombinationskonstrukt unter Verwendung der unten beschriebenen Verfahren transformiert und erfolgreich rekombinierte Klone unter Verwendung zum Beispiel einer Antibiotika- oder Herbizidresistenz selektioniert.

Homologe Rekombination ist ein relativ seltenes Ereignis in höheren Eukaryoten, vor allem in Pflanzen. Zufällige Integrationen in das Wirtsgenom überwiegen. Eine Möglichkeit die zufällig integrierten Sequenzen zu entfernen und so Zellklone mit einer korrekten homologen Rekombination anzureichern, besteht in der Verwendung eines sequenz-spezifischen Rekombinationssystems wie in US 6,110,736 beschrieben, durch welche unspezifisch integrierte Sequenzen wieder deletiert werden können, was die Selektion erfolgreich über homologe Rekombination integrierter Ereignisse erleichtert. Eine Vielzahl von sequenzspezifischen Rekombinationssystemen kann verwendet werden, beispielhaft sind das Cre/lox-System des Bacteriophagen P1, das FLP/FRT System der Hefe, die Gin Rekombinase des Mu Phagen, die Pin Rekombinase aus E. coli und das R/RS System des pSR1 Plasmids genannt. Bevorzugt sind das Bacteriophagen P1 Cre/lox und das Hefe FLP/FRT System. Das FLP/FRT und cre/lox Rekombinasesystem wurde bereits in pflanzlichen Systemen angewendet (Odell et al. (1990) Mol Gen Genet 223: 369-378)

- i) Einführung von Mutationen in endogenen Speicherprotein Gene zur Erzeugung eines Funktionsverlustes (z.B. Generierung von Stopp-Kodons, Verschiebungen im Leseraster etc.)

Weitere geeignete Methoden zur Verminderung der Speicherprotein-Aktivität sind die Einführung von Nonsense-Mutationen in endogene Speicherprotein Gene zum Beispiel mittels Einführung von RNA/DNA-Oligonukleotiden in die Pflanze (Zhu et al. (2000) Nat Biotechnol 18(5):555-558) sowie die Generierung von Knockout-Mutanten mit Hilfe von z.B. T-DNA-Mutagenese (Koncz et al. (1992) Plant Mol Biol 20(5):963-976), ENU-(N-Ethyl-N-nitrosoharnstoff) - Mutagenese oder homologer Rekombination (Hohn B und Puchta (1999) H Proc Natl Acad Sci USA 96:8321-8323.). Punktmutationen können auch mittels DNA-RNA Hybriden erzeugt werden, die auch als "chimeraplasty" bekannt sind (Cole-Strauss et al. (1999)

Nucl Acids Res 27(5):1323-1330; Kmiec (1999) Gene therapy American Scientist 87(3):240-247).

Die Methoden der dsRNAi, der Kosuppression mittels sense-RNA und der "VIGS" ("virus induced gene silencing") werden auch als "post-transcriptional gene silencing" (PTGS) bezeichnet. PTGS-Verfahren sind besonders vorteilhaft, weil die Anforderungen an die Homologie zwischen dem zu supprimierenden endogenem Gen und der transgen exprimierten sense- oder dsRNA-Nukleinsäuresequenz (bzw. zwischen dem endogenen Gen und seiner dominant-negativen Variante) geringer sind als beispielsweise bei einem klassischen antisense-Ansatz. Entsprechende Homologie-Kriterien sind bei der Beschreibung des dsRNAI-Verfahrens genannt und allgemein für PTGS-Verfahren oder dominant-negative Ansätze übertragbar. Aufgrund der hohen Homologie zwischen den pflanzlichen Speicherproteinen (s. Fig 1a bis 7d) kann man voraussichtlich unter Verwendung einer bestimmten Speicherprotein-Nukleinsäuresequenzen auch die Expression von homologen Speicherproteinen in der gleichen oder anderen Arten effektiv supprimieren, ohne dass die Isolierung, Strukturaufklärung und Konstruktion entsprechender Suppressionskonstrukte für dort vorkommenden Speicherprotein-Homologen zwingend erforderlich wäre. Dies erleichtert erheblich den Arbeitsaufwand.

Alle Substanzen und Verbindungen die direkt oder indirekt eine Verminderung der Proteinmenge, RNA-Menge oder Genaktivität zumindest eines Speicherproteins bewirken, und so direkt oder indirekt eine Verminderung der Proteinmenge zumindest eines Speicherproteins bewirken, seien infolge unter der Bezeichnung "anti-SP"-Verbindungen zusammengefasst. Der Begriff "anti-SP"-Verbindung schließt explizit die in den oben beschriebenen Verfahren zum Einsatz kommenden Nukleinsäuresequenzen, Peptide, Proteine oder andere Faktoren ein.

"Einbringung" umfasst im Rahmen der Erfindung alle Verfahren, die dazu geeignet eine "anti-SP"-Verbindung, direkt oder indirekt, in eine Pflanze oder eine Zelle, Kompartiment, Gewebe, Organ oder Samen derselben einzuführen oder dort zu generieren. Direkte und indirekte Verfahren sind umfasst. Die Einbringung kann zu einer vorübergehenden (transienten) Präsenz einer "anti-SP"-Verbindung (beispielsweise einer dsRNA) führen oder aber auch zu einer dauerhaften (stabilen).

Gemäß der unterschiedlichen Natur der oben beschriebenen Ansätze kann die "anti-SP"-Verbindung ihre Funktion direkt ausüben (zum Beispiel durch Insertion in ein endogenes Speicherprotein Gen). Die Funktion kann aber auch indirekt nach Transkription in eine

RNA (zum Beispiel bei antisense Ansätzen) oder nach Transkription und Translation in ein Protein (zum Beispiel bei Bindungsfaktoren) ausgeübt werden. Sowohl direkte als auch indirekt wirkende "anti-SP"-Verbindungen sind erfindungsgemäß umfasst.

Einführen umfasst beispielsweise Verfahren wie Transfektion, Transduktion oder Transformation.

"anti-SP" Verbindungen umfasst somit beispielsweise auch rekombinante Expressionskonstrukte, die eine Expression (d.h. Transkription und ggf. Translation) beispielsweise einer Speicherprotein-dsRNA oder einer Speicherprotein "antisense"-RNA - bevorzugt in einer Pflanze oder einem Teil, Gewebe, Organ oder Samen derselben - bedingen.

In besagten Expressionskonstrukten steht ein Nukleinsäuremolekül, dessen Expression (Transkription und ggf. Translation) eine "anti-SP"-Verbindung generiert, bevorzugt in funktioneller Verknüpfung mit mindestens einem genetischen Kontrollelement (beispielsweise einem Promotor), das eine Expression in einem Organismus, bevorzugt in Pflanzen, gewährleistet. Soll das Expressionskonstrukt direkt in die Pflanze eingeführt und die "anti-SP"-Verbindung (beispielsweise die Speicherprotein dsRNA) dort in planta erzeugt werden, so sind pflanzenspezifische genetische Kontrollelemente (beispielsweise Promotoren) bevorzugt. Die "anti-SP"-Verbindung kann jedoch auch in anderen Organismen oder in vitro erzeugt und dann in die Pflanze eingebracht werden. In diesem sind all prokaryotischen oder eukaryotischen genetischen Kontrollelemente (beispielsweise Promotoren) bevorzugt, die die Expression in den jeweils für die Herstellung gewählten Organismus erlauben.

Unter einer funktionellen Verknüpfung versteht man zum Beispiel die sequentielle Anordnung eines Promotors mit der zu exprimierenden Nukleinsäuresequenz (zum Beispiel einer "anti-SP"-Verbindung) und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator derart, dass jedes der regulativen Elemente seine Funktion bei der transgenen Expression der Nukleinsäuresequenz, je nach Anordnung der Nukleinsäuresequenzen zu sense oder anti-sense RNA, erfüllen kann. Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die transgen zu exprimierende Nukleinsäuresequenz hinter der als Promoter fungierenden Sequenz positioniert wird, so dass beide

Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der transgen zu exprimierende Nukleinsäuresequenz geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare.

Die Herstellung einer funktionellen Verknüpfung als auch die Herstellung einer Expressionskassette kann mittels gängiger Rekombinations- und Klonierungstechniken realisiert werden, wie sie beispielsweise in Maniatis T, Fritsch EF und Sambrook J (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Silhavy TJ, Berman ML und Enquist LW (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Ausubel FM et al. (1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience und bei Gelvin et al. (1990) In: Plant Molecular Biology Manual beschrieben sind. Zwischen beide Sequenzen können aber auch weitere Sequenzen positioniert werden, die zum Beispiel die Funktion eines Linkers mit bestimmten Restriktionsenzymchnittstellen oder eines Signalpeptides haben. Auch kann die Insertion von Sequenzen zur Expression von Fusionsproteinen führen. Bevorzugt kann die Expressionskassette, bestehend aus einer Verknüpfung von Promoter und zu exprimierender Nukleinsäuresequenz, integriert in einem Vektor vorliegen und durch zum Beispiel Transformation in ein pflanzliches Genom insertiert werden.

Unter einer Expressionskassette sind aber auch solche Konstruktionen zu verstehen, bei denen ein Promoter - zum Beispiel durch eine homologe Rekombination - hinter ein endogenes Speicherprotein-Gen platziert wird, und durch Expression einer antisense Speicherprotein-RNA die erfindungsgemäße Verminderung eines Speicherproteins bewirkt wird. Analog kann auch eine "anti-SP" Verbindung (zum Beispiel eine Nukleinsäuresequenz kodierend für eines Speicherprotein dsRNA oder eine Speicherprotein antisense RNA) derart hinter einen endogenen Promotor platziert werden, dass der gleiche Effekt auftritt. Beide Ansätze führen zu Expressionskassetten im Sinne der Erfindung.

Pflanzenspezifische Promotoren meint grundsätzlich jeden Promotor, der die Expression von Genen, insbesondere Fremdgenen, in Pflanzen oder Pflanzenteilen, -zellen, -geweben, -kulturen steuern kann. Dabei kann die Expression beispielsweise konstitutiv, induzierbar oder entwicklungsabhängig sein.

Bevorzugt sind:

a) Konstitutive Promotoren

"Konstitutive" Promotoren meint solche Promotoren, die eine Expression in zahlreichen, bevorzugt allen, Geweben über einen größeren Zeitraum der Pflanzenentwicklung, bevorzugt zu allen Zeitpunkten der Pflanzenentwicklung, gewährleisten (Benfey et al. (1989) EMBO J 8:2195-2202). Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Insbesondere bevorzugt ist der Promotor des 35S-Transkriptes des CaMV Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228) oder der 19S CaMV Promotor (US 5,352,605; WO 84/02913; Benfey et al. (1989) EMBO J 8:2195-2202). Ein weiterer geeigneter konstitutiver Promotor ist der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028), der LeguminB-Promotor (GenBank Acc.-Nr. X03677), der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), den Ubiquitin 1 Promotor (Christensen et al. (1992) Plant Mol Biol 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), den Sma Promotor, den Cinnamylalkoholdehydrogenase-Promotor (US 5,683,439), die Promotoren der vakuolärer ATPase Untereinheiten oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991), sowie weitere Promotoren von Genen, deren konstitutive Expression in Pflanzen dem Fachmann bekannt ist.

b) Gewebespezifische Promotoren

Bevorzugt sind ferner Promotoren mit Spezifitäten für Samen, wie zum Beispiel der Promotor des Phaseolins (US 5,504,200; Bustos MM et al. (1989) Plant Cell 1(9):839-53), des 2S Albumingens (Joseffson LG et al. (1987) J Biol Chem 262:12196-12201), des Legumins (Shirsat A et al. (1989) Mol Gen Genet 215(2): 326-331), des USP (unknown seed protein; Bäumlein H et al. (1991) Mol Gen Genet 225(3):459-67), des Napin Gens (US 5,608,152; Stalberg K et al. (1996) L Planta 199:515-519), des Saccharosebindepoteins (WO 00/26388) oder der Legumin B4-Promotor (LeB4; Bäumlein H et al. (1991) Mol Gen Genet 225: 121-128; Baeumlein et al. (1992) Plant Journal 2(2):233-9; Fiedler U et al. (1995) Biotechnology (NY) 13(10):1090f), der Oleosin-Promoter aus Arabidopsis (WO 98/45461), der Bce4-Promoter aus Brassica (WO 91/13980). Weitere geeignete samenspezifische Promotoren sind die der

41

Gene kodierend für das "High Molecular Weight Glutenin" (HMWG), Gliadin, Verzweigungsenzym, ADP Glucose Pyrophosphatase (AGPase) oder die Stärkesynthase. Bevorzugt sind ferner Promotoren, die eine samenspezifische Expression in Monokotyledonen wie Mais, Gerste, Weizen, Roggen, Reis etc. erlauben. Vorteilhaft eingesetzt werden können der Promoter des lpt2 oder lpt1-Gen (WO 95/15389, WO 95/23230) oder die Promotoren beschrieben in WO 99/16890 (Promotoren des Hordein-Gens, des Glutelin-Gens, des Oryzin-Gens, des Prolamin-Gens, des Gliadin-Gens, des Glutelin-Gens, des Zein-Gens, des Kasirin-Gens oder des Secalin-Gens).

c) Chemisch induzierbare Promotoren

Die Expressionskassetten können auch einen chemisch induzierbaren Promotor enthalten (Übersichtsartikel: Gatz et al. (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108), durch den die Expression des exogenen Gens in der Pflanze zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige Promotoren, wie z.B. der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366), durch Salicylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzolsulfonamid-induzierbarer Promotor (EP 0 388 186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure induzierbarer Promotor (EP 0 335 528) bzw. ein durch Ethanol- oder Cyclohexanon-induzierbarer Promotor (WO 93/21334) können ebenfalls verwendet werden.

Besonders bevorzugt sind konstitutive sowie samenspezifische Promotoren.

Es können ferner weitere Promotoren funktionell mit der zu exprimierenden Nukleinsäuresequenz verknüpft sein, die eine Expression in weiteren Pflanzengewebe oder in anderen Organismen, wie zum Beispiel *E.coli* Bakterien ermöglichen. Als Pflanzen Promotoren kommen im Prinzip alle oben beschriebenen Promotoren in Frage.

Die in den erfindungsgemäßen Expressionskassetten oder Vektoren enthaltenen Nukleinsäuresequenzen können mit weiteren genetischen Kontrollsequenzen neben einem Promoter funktionell verknüpft sein. Der Begriff der genetischen Kontrollsequenzen ist breit zu verstehen und meint all solche Sequenzen, die einen Einfluss auf das Zustandekommen oder die Funktion der erfindungsgemäßen Expressionskassette haben. Genetische Kontrollsequenzen modifizieren zum Beispiel die Transkription und Translation in

prokaryotischen oder eukaryotischen Organismen. Vorzugsweise umfassen die erfindungsgemäßen Expressionskassetten 5'-stromaufwärts von der jeweiligen transgen zu exprimierenden Nukleinsäuresequenz einen pflanzenspezifischen Promoter und 3'-stromabwärts eine Terminatorsequenz als zusätzliche genetische Kontrollsequenz, sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils funktionell verknüpft mit der transgen zu exprimierenden Nukleinsäuresequenz.

Genetische Kontrollsequenzen umfassen auch weitere Promotoren, Promotorelemente oder Minimalpromotoren, die die expressionsteuernden Eigenschaften modifizieren können. So kann durch genetische Kontrollsequenzen zum Beispiel die gewebespezifische Expression zusätzlich abhängig von bestimmten Stressfaktoren erfolgen. Entsprechende Elemente sind zum Beispiel für Wasserstress, Abscisinsäure (Lam E und Chua NH, J Biol Chem 1991; 266(26): 17131 -17135) und Hitzestress (Schoffl F et al. (1989) Mol Gen Genetics 217(2-3):246-53) beschrieben.

Weitere vorteilhafte Kontrollsequenzen sind beispielsweise in den gram-positiven Promotoren amy und SPO2, in den Hefe- oder Pilzpromotoren ADC1, MFa, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH.

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für das erfindungsgemäße Verfahren verwendet werden. Darüberhinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Regionen, Introns oder nichtkodierende 3'-Region von Genen wie beispielsweise das Actin-1 Intron, oder die Adh1-S Introns 1, 2 und 6 (allgemein: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994)). Es ist gezeigt worden, dass diese eine signifikante Funktionen bei der Regulation der Genexpression spielen können. So wurde gezeigt, dass 5'-untranslatierte Sequenzen die transiente Expression heterologer Gene verstärken können. Beispielhaft für Translationsverstärker sei die 5'-Leadersequenz aus dem Tabak-Mosaik-Virus zu nennen (Gallie et al. (1987) Nucl Acids Res 15:8693-8711) und dergleichen. Sie können ferner die Gewebespezifität fördern (Rouster J et al. (1998) Plant J 15:435-440).

Die Expressionskassette kann vorteilhafterweise eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promoter enthalten, die eine erhöhte transgene Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der transgen zu exprimierenden Nukleinsäuresequenzen können zusätzliche vorteil-

hafte Sequenzen inseriert werden, wie weitere regulatorische Elemente oder Terminatoren. Die transgen zu exprimierenden Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

Als Kontrollsequenzen geeignete Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA Polyadenylierungssignale aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACHS entsprechen (Gielen et al. (1984) EMBO J 3:835 ff) oder funktionelle Äquivalente davon. Beispiele für besonders geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopalin-Synthase)-Terminator.

Als Kontrollsequenzen sind weiterhin solche zu verstehen, die eine homologe Rekombination bzw. Insertion in das Genom eines Wirtsorganismus ermöglichen oder die Entfernung aus dem Genom erlauben. Bei der homologen Rekombination kann zum Beispiel die kodierende Sequenz eines bestimmten endogenen Gens gegen die für eine dsRNA kodierende Sequenz gezielt ausgetauscht werden. Methoden wie die cre/lox-Technologie erlauben eine gewebespezifische, unter Umständen induzierbare Entfernung der Expressionskassette aus dem Genom des Wirtsorganismus (Sauer B (1998) Methods. 14(4):381-92). Hier werden bestimmte flankierende Sequenzen dem Zielgen angefügt (lox-Sequenzen), die später eine Entfernung mittels der cre-Rekombinase ermöglichen.

Eine Expressionskassetten und die von ihr abgeleiteten Vektoren können weitere Funktionselemente enthalten. Der Begriff Funktionselement ist breit zu verstehen und meint all solche Elemente, die einen Einfluss auf Herstellung, Vermehrung oder Funktion der erfindungsgemäßen Expressionskassetten, Vektoren oder transgenen Organismen haben. Beispielhaft aber nicht einschränkend seien zu nennen:

- a) Selektionsmarker, die eine Resistenz gegen einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456), Antibiotika oder Biozide, bevorzugt Herbizide, wie zum Beispiel Kanamycin, G 418, Bleomycin, Hygromycin, oder Phosphinotricin etc. verleihen. Besonders bevorzugte Selektionsmarker sind solche die eine Resistenz gegen Herbizide verleihen. Beispielhaft seien genannt: DNA Sequenzen, die für Phosphinothricinacetyltransferasen (PAT) kodieren und Glutaminsynthaseinhibitoren inaktivieren (bar und pat Gen), 5-Enolpyruvylshikimat-3-phosphatsynthasegene (EPSP Synthasegene), die eine Resistenz gegen Glyphosat® (N-(phosphonomethyl)glycin) verleihen, das für das Glyphosat®

degradierende Enzyme kodierende gox Gen (Glyphosatoxido-reduktase), das deh Gen (kodierend für eine Dehalogenase, die Dalapon inaktiviert), Sulfonylharnstoff- und Imidazolinon inaktivierende Acetolactatsynthasen sowie bxn Gene, die für Bromoxynil degradierende Nitrilaseenzyme kodieren, das aasa-Gen, das eine Resistenz gegen das Antibiotikum Apectinomycin verleiht, das Streptomycinphosphotransferase (SPT) Gen, das eine Resistenz gegen Streptomycin gewährt, das Neomycinphosphotransferase (NPTII) Gen, das eine Resistenz gegen Kanamycin oder Geneticidin verleiht, das Hygromycinphosphotransferase (HPT) Gen, das eine Resistenz gegen Hygromycin vermittelt, das Acetolactatsynthase Gen (ALS), das eine Resistenz gegen Sulfonylharnstoff-Herbizide verleiht (z.B. mutierte ALS-Varianten mit z.B. der S4 und/oder Hra Mutation).

- b) Reportergene, die für leicht quantifizierbare Proteine kodieren und über Eigenfarbe oder Enzymaktivität eine Bewertung der Transformationseffizienz oder des Expressionsortes oder -zeitpunktes gewährleisten. Ganz besonders bevorzugt sind dabei Reporter-Proteine (Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) wie das "green fluorescence protein" (GFP) (Sheen et al. (1995) Plant Journal 8(5):777-784), die Chloramphenicoltransferase, eine Luziferase (Ow et al. (1986) Science 234:856-859), das Aequorin-Gen (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), die β -Galactosidase, ganz besonders bevorzugt ist die β -Glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).
- c) Replikationsursprünge, die eine Vermehrung der erfindungsgemäßen Expressionskassetten oder Vektoren in zum Beispiel E.coli gewährleisten. Beispielhaft seien genannt ORI (origin of DNA replication), der pBR322 ori oder der P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
- d) Elemente, die für eine Agrobakterium vermittelte Pflanzen-transformation erforderlich sind, wie zum Beispiel die rechte oder linke Begrenzung der T-DNA oder die vir-Region.

Zur Selektion erfolgreich homolog rekombinierter oder auch transformierter Zellen ist es in der Regel erforderlich, einen selektionierbaren Marker zusätzlich einzuführen, der den erfolgreich rekombinierten Zellen eine Resistenz gegen ein Biozid (zum Beispiel ein Herbizid), einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456) oder ein Antibiotikum verleiht.

Der Selektionsmarker erlaubt die Selektion der transformierten Zellen von untransformierten (McCormick et al. (1986) Plant Cell Reports 5:81-84).

Zusätzlich kann besagte transgene Expressionskassette oder Expressionsvektoren Nukleinsäuresequenzen enthalten, die nicht zu einer Verminderung mindestens eines Speicherproteins führen, und deren transgene Expression zu einer zusätzlichen Steigerung der Fettsäure-Biosynthese führt (infolge proOIL). Diese zusätzlich transgen exprimierte proOIL Nukleinsäuresequenz kann beispielhaft aber nicht einschränkend ausgewählt sein aus Nukleinsäuren kodierend für Acetyl-CoA-Carboxylase (ACCase), Glycerol-3-Phosphat Acyltransferase (GPAT), Lysophosphatidat-Acyltransferase (LPAT), Diacylglycerol-Acyltransferase (DAGAT) und Phospholipid:diacylglycerol-acyltransferase (PDAT). Entsprechende Sequenzen sind dem Fachmann bekannt und aus Datenbanken oder entsprechenden cDNA-Banken der jeweiligen Pflanzen leicht zugänglich.

Die Einführung einer erfindungsgemäßen Expressionskassette in einen Organismus oder Zellen, Geweben, Organe, Teile bzw. Samen desselben (bevorzugt in Pflanzen bzw. pflanzliche Zellen, Gewebe, Organe, Teile oder Samen), kann vorteilhaft unter Verwendung von Vektoren realisiert werden, in denen die Expressionskassetten enthalten sind. Vektoren können beispielhaft Plasmide, Cosmide, Phagen, Viren oder auch Agrobakterien sein. Die Expressionskassette kann in den Vektor (bevorzugt ein Plasmidvektor) über eine geeignete Restriktionsschnittstelle eingeführt werden. Der entstandene Vektor wird zunächst in E.coli eingeführt. Korrekt transformierte E.coli werden selektioniert, gezüchtet und der rekombinante Vektor mit dem Fachmann geläufigen Methoden gewonnen. Restriktionsanalyse und Sequenzierung können dazu dienen, den Klonierungsschritt zu prüfen. Bevorzugt sind solche Vektoren, die eine stabile Integration der Expressionskassette in das Wirtsgenom ermöglichen.

Die Herstellung eines transformierten Organismus (bzw. einer transformierten Zelle oder Gewebes) erfordert, dass die entsprechende DNA (z.B. der Expressionsvektor), RNA oder Protein in die entsprechende Wirtszelle eingebracht wird. Für diesen Vorgang, der als Transformation (oder Transduktion bzw. Transfektion) bezeichnet wird, steht eine Vielzahl von Methoden zur Verfügung (Keown et al. (1990) Methods in Enzymology 185:527-537). So kann die DNA oder RNA beispielhaft direkt durch Mikroinjektion oder durch Bombardierung mit DNA-beschichteten Mikropartikeln eingeführt werden. Auch kann die Zelle chemisch, zum Beispiel mit Polyethylenglycol, permeabilisiert werden, so dass die DNA durch

Diffusion in die Zelle gelangen kann. Die DNA kann auch durch Protoplastenfusion mit anderen DNA-enthaltenden Einheiten wie Mitocysten, Zellen, Lysosomen oder Liposomen erfolgen. Elektroporation ist eine weitere geeignete Methode zur Einführung von DNA, bei der die Zellen reversibel durch einen elektrischen Impuls permeabilisiert werden. Entsprechende Verfahren sind beschrieben (beispielsweise bei Bilang et al. (1991) *Gene* 100:247-250; Scheid et al. (1991) *Mol Gen Genet* 228:104-112; Guerche et al. (1987) *Plant Science* 52:111-116; Neuhaus et al. (1987) *Theor Appl Genet* 75:30-36; Klein et al. (1987) *Nature* 327:70-73; Howell et al. (1980) *Science* 208:1265; Horsch et al. (1985) *Science* 227:1229-1231; DeBlock et al. (1989) *Plant Physiology* 91:694-701; *Methods for Plant Molecular Biology* (Weissbach and Weissbach, eds.) Academic Press Inc. (1988); and *Methods in Plant Molecular Biology* (Schuler and Zielinski, eds.) Academic Press Inc. (1989)).

Bei Pflanzen werden dabei die beschriebenen Methoden zur Transformation und Regeneration von Pflanzen aus Pflanzengewebe oder Pflanzenzellen zur transienten oder stabilen Transformation genutzt. Geeignete Methoden sind vor allem die Protoplastentransformation durch Polyethylenglykol-induzierte DNA-Aufnahme, das biolistische Verfahren mit der Genkanone, die sogenannte "particle bombardment" Methode, die Elektroporation, die Inkubation trockener Embryonen in DNA-haltiger Lösung und die Mikroinjektion.

Neben diesen "direkten" Transformationstechniken kann eine Transformation auch durch bakterielle Infektion mittels *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* durchgeführt werden. Die *Agrobacterium*-vermittelte Transformation ist am besten für dicotyledone Pflanzenzellen geeignet. Die Verfahren sind beispielsweise beschrieben bei Horsch et al. (1985) *Science* 225: 1229f).

Werden *Agrobacterien* verwendet, so ist die Expressionskassette in spezielle Plasmide zu integrieren, entweder in einen Zwischenvektor (englisch: shuttle or intermediate vector) oder einen binären Vektor. Wird ein Ti oder Ri Plasmid zur Transformation verwendet werden soll, ist zumindest die rechte Begrenzung, meistens jedoch die rechte und die linke Begrenzung der Ti oder Ri Plasmid T-DNA als flankierende Region mit der einzuführenden Expressionskassette verbunden.

Bevorzugt werden binäre Vektoren verwendet. Binäre Vektoren können sowohl in *E.coli* als auch in *Agrobacterium* replizieren. Sie enthalten in der Regel ein Selektionsmarkergen und einen

Linker oder Polylinker flankiert von der rechten und linken T-DNA Begrenzungssequenz. Sie können direkt in Agrobacterium transformiert werden (Holsters et al. (1978) Mol Gen Genet 163:181-187). Das Selektionsmarkergen erlaubt eine Selektion transformierter Agrobakteria und ist zum Beispiel das nptII Gen, das eine Resistenz gegen Kanamycin verleiht. Das in diesem Fall als Wirtsorganismus fungierende Agrobacterium sollte bereits ein Plasmid mit der vir-Region enthalten. Diese ist für die Übertragung der T-DNA auf die pflanzliche Zelle erforderlich. Ein so transformiertes Agrobacterium kann zur Transformation pflanzlicher Zellen verwendet werden. Die Verwendung von T-DNA zur Transformation pflanzlicher Zellen ist intensiv untersucht und beschrieben (EP 120 516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasterdam, Chapter V; An et al. (1985) EMBO J 4:277-287). Verschiedene binäre Vektoren sind bekannt und teilweise kommerziell erhältlich wie zum Beispiel pBI101.2 oder pBIN19 (Clontech Laboratories, Inc. USA).

Weitere zur Expression in Pflanzen geeignet Promotoren sind beschrieben (Rogers et al. (1987) Meth in Enzymol 153:253-277; Schardl et al. (1987) Gene 61:1-11; Berger et al. (1989) Proc Natl Acad Sci USA 86:8402-8406).

Direkte Transformationstechniken eignen sich für jeden Organismus und Zelltyp. Im Falle von Injektion oder Elektroporation von DNA bzw. RNA in pflanzliche Zellen sind keine besonderen Anforderungen an das verwendete Plasmid gestellt. Einfache Plasmide wie die der pUC-Reihe können verwendet werden. Sollen vollständige Pflanzen aus den transformierten Zellen regeneriert werden, so ist es erforderlich, dass sich auf dem Plasmid ein zusätzliches selektionierbares Markergen befindet.

Stabil transformierte Zellen d.h. solche, die die eingeführte DNA integriert in die DNA der Wirtszelle enthalten, können von untransformierten selektioniert werden, wenn ein selektionierbarer Marker Bestandteil der eingeführten DNA ist. Als Marker kann beispielhaft jedes Gen fungieren, dass eine Resistenz gegen Antibiotika oder Herbizide (wie Kanamycin, G 418, Bleomycin, Hygromycin oder Phosphinotricin etc.) zu verleihen vermag (s.o.). Transformierte Zellen, die ein solches Markergen exprimieren, sind in der Lage, in der Gegenwart von Konzentrationen eines entsprechenden Antibiotikums oder Herbizides zu überleben, die einen untransformierten Wildtyp abtöten. Beispiel sind oben genannt und umfassen bevorzugt das bar Gen, dass Resistenz gegen das Herbizid Phosphinotricin verleiht (Rathore KS et al. (1993) Plant Mol Biol 21(5):871-884), das nptII Gen, dass Resistenz gegen Kanamycin verleiht, das hpt Gen, das Resistenz gegen Hygromycin verleiht,

oder das EPSP-Gen, das Resistenz gegen das Herbizid Glyphosat verleiht. Der Selektionsmarker erlaubt die Selektion von transformierten Zellen von untransformierten (McCormick et al. (1986) Plant Cell Reports 5:81-84). Die erhaltenen Pflanzen können in üblicher Weise gezüchtet und gekreuzt werden. Zwei oder mehr Generationen sollten kultiviert werden, um sicherzustellen, dass die genomische Integration stabil und vererblich ist.

Die oben genannten Verfahren sind beispielsweise beschrieben in Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von SD Kung und R Wu, Academic Press, S.128-143 sowie in Potrykus (1991) Annu. Rev Plant Physiol Plant Molec Biol 42:205-225). Vorzugsweise wird das zu exprimierende Konstrukt in einen Vektor kloniert, der geeignet ist, Agrobacterium tumefaciens zu transformieren, beispielsweise pBin19 (Bevan et al. (1984) Nucl Acids Res 12:8711f).

Sobald eine transformierte Pflanzenzelle hergestellt wurde, kann eine vollständige Pflanze unter Verwendung von dem Fachmann bekannten Verfahren erhalten werden. Hierbei geht man beispielhaft von Kalluskulturen aus. Aus diesen noch undifferenzierten Zellmassen kann die Bildung von Spross und Wurzel in bekannter Weise induziert werden. Die erhaltenen Sprösslinge können ausgepflanzt und gezüchtet werden.

Dem Fachmann sind auch Verfahren bekannt, um aus Pflanzenzellen, Pflanzenteile und ganze Pflanzen zu regenerieren. Beispielsweise werden hierzu Verfahren beschrieben von Fennell et al. (1992) Plant Cell Rep. 11: 567-570; Stoeger et al (1995) Plant Cell Rep. 14:273-278; Jahne et al. (1994) Theor Appl Genet 89:525-533 verwendet.

"Transgen" meint bezüglich zum Beispiel einer Nukleinsäuresequenz, einer Expressionskassette oder einem Vektor enthaltend besagte Nukleinsäuresequenz oder einem Organismus transformiert mit besagter Nukleinsäuresequenz, Expressionskassette oder Vektor alle solche durch gentechnische Methoden zustandegekommene Konstruktionen, in denen sich entweder

- a) die Speicherprotein Nukleinsäuresequenz, oder
- b) eine mit der Speicherprotein Nukleinsäuresequenz funktionell verknüpfte genetische Kontrollsequenz, zum Beispiel ein Promotor, oder

c) (a) und (b)

sich nicht in ihrer natürlichen, genetischen Umgebung befinden oder durch gentechnische Methoden modifiziert wurden, wobei die Modifikation beispielhaft eine Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Nukleotidreste sein kann. Natürliche genetische Umgebung meint den natürlichen chromosomalen Locus in dem Herkunftsorganismus oder das Vorliegen in einer genomischen Bibliothek. Im Fall einer genomischen Bibliothek ist die natürliche, genetische Umgebung der Nukleinsäuresequenz bevorzugt zumindest noch teilweise erhalten. Die Umgebung flankiert die Nukleinsäuresequenz zumindest an einer Seite und hat eine Sequenzlänge von mindestens 50 bp, bevorzugt mindestens 500 bp, besonders bevorzugt mindestens 1000 bp, ganz besonders bevorzugt mindestens 5000 bp. Eine natürlich vorkommende Expressionskassette - beispielsweise die natürlich vorkommende Kombination des Speicherprotein-Promotors mit dem entsprechenden Speicherprotein-Gen - wird zu einer transgenen Expressionskassette, wenn diese durch nicht-natürliche, synthetische ("künstliche") Verfahren wie beispielsweise einer Mutagenisierung geändert wird. Entsprechende Verfahren sind beschrieben (US 5,565,350; WO 00/15815; siehe auch oben).

Als transgene Organismen bevorzugte Wirts- oder Ausgangsorganismen sind vor allem Pflanzen gemäß der oben genannten Definition. Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches, insbesondere Pflanzen, die für die Gewinnung von Ölen verwendet werden wie beispielsweise Raps, Sonnenblume, Sesam, Färberdistel, Ölbaum, Soja, Mais, Weizen und Nussarten. Eingeschlossen sind ferner die reifen Pflanzen, Saatgut, Sprossen und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut und Kulturen, zum Beispiel Zellkulturen. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.

Die Herstellung der transgenen Organismen kann mit den oben beschriebenen Verfahren zur Transformation oder Transfektion von Organismen realisiert werden.

Ein weiterer Gegenstand der Erfindung betrifft die Verwendung der erfindungsgemäßen, transgenen Organismen und der von ihnen abgeleitete Zellen, Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln, Blätter etc.-, und transgenes Vermehrungsgut wie Saaten oder Früchte, zur Herstellung von

Nahrungs- oder Futtermitteln, Pharmazeutika oder Feinchemikalien, insbesondere von Ölen, Fetten, Fettsäuren oder Derivaten der vor-
genannten.

Sequenzen

1. SEQ ID NO: 1
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S subunit 1 (GenBank Acc.-No.: M22032)
2. SEQ ID NO: 2
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 1
3. SEQ ID NO: 3
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S subunit 3 (GenBank Acc.-No.: M22035)
4. SEQ ID NO: 4
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 3
5. SEQ ID NO: 5
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S subunit 2 (GenBank Acc.-No.: M22034)
6. SEQ ID NO: 6
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 2
7. SEQ ID NO: 7
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S subunit 4 (GenBank Acc.-No.: M22033)
8. SEQ ID NO: 8
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 4
9. SEQ ID NO: 9
Nukleinsäuresequenz kodierend für B.napus Cruciferin Speicherprotein (GenBank Acc.-No.: X59294)
10. SEQ ID NO: 10
Proteinsequenz kodierend für B.napus Cruciferin Speicherprotein
11. SEQ ID NO: 11
Nukleinsäuresequenz kodierend für Brassica napus Cruciferin (GenBank Acc.-No.: X14555)
12. SEQ ID NO: 12
Proteinsequenz kodierend für Brassica napus Cruciferin

13. SEQ ID NO: 13
Nukleinsäuresequenz kodierend für B.napus BnC2 Cruciferin
Speicherprotein (GenBank Acc.-No.: X59295)
14. SEQ ID NO: 14
Proteinsequenz kodierend für B.napus BnC2 Cruciferin
Speicherprotein
15. SEQ ID NO: 15
partielle Nukleinsäuresequenz kodierend für B.napus
Cruciferin cru4 subunit (GenBank Acc.-No.: X57848)
16. SEQ ID NO: 16
partielle Proteinsequenz kodierend für B.napus Cruciferin
cru4 subunit
17. SEQ ID NO: 17
Nukleinsäuresequenz kodierend für B.napus cru1 Cruciferin
subunit (GenBank Acc.-No.: X62120)
18. SEQ ID NO: 18
Proteinsequenz kodierend für B.napus cru1 Cruciferin subunit
19. SEQ ID NO: 19
Nukleinsäuresequenz kodierend für Glycinin A-1a-B-x subunit
aus des Sojabohne (GenBank Acc.-No.: M36686)
20. SEQ ID NO: 20
Proteinsequenz kodierend für Glycinin A-1a-B-x subunit aus
des Sojabohne
21. SEQ ID NO: 21
Nukleinsäuresequenz kodierend für Sojabohne Glycinin subunit
G2 (GenBank Acc.-No.: X15122)
22. SEQ ID NO: 22
Proteinsequenz kodierend für Sojabohne Glycinin subunit G2
23. SEQ ID NO: 23
Nukleinsäuresequenz kodierend für Sojabohne A5A4B3 Glycinin
subunits (GenBank Acc.-No.: X02626)
24. SEQ ID NO: 24
Proteinsequenz kodierend für Sojabohne A5A4B3 Glycinin
subunits

25. SEQ ID NO: 25
Nukleinsäuresequenz kodierend für Sojabohne (G.max) Glycinin
Speicherprotein subunit A3-B4 (GenBank Acc.-No.: M10962)
26. SEQ ID NO: 26
Proteinsequenz kodierend für Sojabohne (G.max) Glycinin
Speicherprotein subunit A3-B4
27. SEQ ID NO: 27
Nukleinsäuresequenz kodierend für Sojabohne Glycinin subunit
G3 (GenBank Acc.-No.: X15123)
28. SEQ ID NO: 28
Proteinsequenz kodierend für Sojabohne Glycinin subunit G3
29. SEQ ID NO: 29
Nukleinsäuresequenz kodierend für Sonnenblume 11S Speicher-
protein (G3-D1) (GenBank Acc.-No.: M28832)
30. SEQ ID NO: 30
Proteinsequenz kodierend für Sonnenblume 11S Speicherprotein
(G3-D1)
31. SEQ ID NO: 31
Nukleinsäuresequenz kodierend für Raps (B.napus) Napin
(GenBank Acc.-No.: J02586)
32. SEQ ID NO: 32
Proteinsequenz kodierend für Raps (B.napus) Napin
33. SEQ ID NO: 33
Nukleinsäuresequenz kodierend für Brassica juncea 2S
Speicherprotein (GenBank Acc.-No.: X65040)
34. SEQ ID NO: 34
Proteinsequenz kodierend für Brassica juncea 2S Speicher-
protein
35. SEQ ID NO: 35
Nukleinsäuresequenz kodierend für Brassica oleracea 2S
Speicherprotein (GenBank Acc.-No.: X65038)
36. SEQ ID NO: 36
Proteinsequenz kodierend für Brassica oleracea 2S Speicher-
protein

37. SEQ ID NO: 37
Nukleinsäuresequenz kodierend für *Brassica napus* cv. Topas Napin (GenBank Acc.-No.: U04945)
38. SEQ ID NO: 38
Proteinsequenz kodierend für *Brassica napus* cv. Topas Napin
39. SEQ ID NO: 39
partielle Nukleinsäuresequenz kodierend für *Sinapis alba* sin1 Speicherprotein (GenBank Acc.-No.: X91799)
40. SEQ ID NO: 40
partielle Proteinsequenz kodierend für *Sinapis alba* sin1 Speicherprotein
41. SEQ ID NO: 41
Nukleinsäuresequenz kodierend für Sojabohne (*Glycine max*) napin-type 2S Albumin 1 (GenBank Acc.-No.: U71194)
42. SEQ ID NO: 42
Proteinsequenz kodierend für Sojabohne (*Glycine max*) napin-type 2S Albumin 1
43. SEQ ID NO: 43
Nukleinsäuresequenz kodierend für Sojabohne (*Glycine max*) 2S Albumin (GenBank Acc.-No.: AF005030)
44. SEQ ID NO: 44
Proteinsequenz kodierend für Sojabohne (*Glycine max*) 2S Albumin
45. SEQ ID NO: 45
Nukleinsäuresequenz kodierend für *Brassica nigra* 2S Speicherprotein (GenBank Acc.-No.: X65039)
46. SEQ ID NO: 46
Proteinsequenz kodierend für *Brassica nigra* 2S Speicherprotein
47. SEQ ID NO: 47
Nukleinsäuresequenz kodierend für *Sinapis alba* sin5 Speicherprotein (GenBank Acc.-No.: X91798)
48. SEQ ID NO: 48
Proteinsequenz kodierend für *Sinapis alba* sin5 Speicherprotein

49. SEQ ID NO: 49
Nukleinsäuresequenz kodierend für Sonnenblume HaG5 2 S
Albumin (GenBank Acc.-No.: X06410)
50. SEQ ID NO: 50
proteinsequenz kodierend für Sonnenblume HaG5 2 S Albumin
51. SEQ ID NO: 51
partielle Nukleinsäuresequenz kodierend für Sonnenblume
(*Helianthus annuus*) 2S Albumin (GenBank Acc.-No.: X76101)
52. SEQ ID NO: 52
partielle Proteinsequenz kodierend für Sonnenblume
(*Helianthus annuus*) 2S Albumin
- 53.-104:
SEQ ID NO: 51 - 104 Sequenzmotive aus verschiedenen Speicher-
proteinklassen
105. SEQ ID NO: 105
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von
Arabidopsis thaliana 12S Speicherprotein AtCru3 (Insert von
Vektor pCR2.1-AtCRU3-RNAi)
106. SEQ ID NO: 106
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression
von *Arabidopsis thaliana* 12S Speicherprotein AtCru3
107. SEQ ID NO: 107
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von
Arabidopsis thaliana 12S Speicherprotein AtCra1
108. SEQ ID NO: 108
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression
von *Arabidopsis thaliana* 12S Speicherprotein AtCra1
109. SEQ ID NO: 109
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von
Arabidopsis thaliana 2S Speicherprotein At2S2
110. SEQ ID NO: 110
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression
von *Arabidopsis thaliana* 2S Speicherprotein At2S2

111. SEQ ID NO: 111
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* 12S
Cruciferin Speicherprotein (ATCRU3; GenBank Acc.-No.: U66916)
112. SEQ ID NO: 112
Proteinsequenz kodierend für *Arabidopsis thaliana* 12S
Cruciferin Speicherprotein (ATCRU3)
113. SEQ ID NO: 113
Nukleinsäuresequenz kodierend für *A.thaliana* 12S Speicher-
protein (CRA1; GenBank Acc.-No.: M37247)
114. SEQ ID NO: 114
Proteinsequenz kodierend für *A.thaliana* 12S Speicherprotein
(CRA1)
115. SEQ ID NO: 115
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* 12S
Speicherprotein AT5g44120/MLN1_4 (GenBank Acc.-No.: AY070730)
116. SEQ ID NO: 116
Proteinsequenz kodierend für *Arabidopsis thaliana* 12S
Speicherprotein AT5g44120/MLN1_4
117. SEQ ID NO: 117
Nukleinsäuresequenz kodierend für *Arabidopsis* 12S Speicher-
protein (CRB; GenBank Acc.-No.: X14313; M37248)
118. SEQ ID NO: 118
Proteinsequenz kodierend für *Arabidopsis* 12S Speicherprotein
(CRB)
119. SEQ ID NO: 119
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana*
putatives 12S Speicherprotein (aus GenBank Acc.-No.:
AC003027)
120. SEQ ID NO: 120
Proteinsequenz kodierend für *Arabidopsis thaliana* putatives
Speicherprotein (Protein_id="AAD10679.1)
121. SEQ ID NO: 121
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana*
Cruciferin 12S Speicherprotein (At1g03890) (GenBank Acc.-No.:
AY065432)

122. SEQ ID NO: 122
Proteinsequenz kodierend für *Arabidopsis thaliana* Cruciferin
12S Speicherprotein (At1g03890)
- 123.-131:
SEQ ID NO: 123-131 Sequenzmotive aus Nukleinsäuresequenzen
verschiedener Speicherproteinklassen
132. SEQ ID NO: 132
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana*
Prohibitin 1 (Atphb1) (GenBank Acc.-No.: U66594)
133. SEQ ID NO: 133
Proteinsequenz kodierend für *Arabidopsis thaliana*
Prohibitin 1 (Atphb1)
134. SEQ ID NO: 134 Oligonukleotidprimer OPN1
135. SEQ ID NO: 135 Oligonukleotidprimer OPN2
136. SEQ ID NO: 136 Oligonukleotidprimer OPN3
137. SEQ ID NO: 137 Oligonukleotidprimer OPN4
138. SEQ ID NO: 138 Oligonukleotidprimer OPN5
139. SEQ ID NO: 139 Oligonukleotidprimer OPN6
140. SEQ ID NO: 140 Oligonukleotidprimer OPN7
141. SEQ ID NO: 141 Oligonukleotidprimer OPN8
142. SEQ ID NO: 142 Oligonukleotidprimer OPN9
143. SEQ ID NO: 143 Oligonukleotidprimer OPN10
144. SEQ ID NO: 144
Nukleinsäuresequenz kodierend für sRNAi4-dsRNA zur
Suppression mehrerer Speicherproteine
145. SEQ ID NO: 145
Ribonukleinsäuresequenz kodierend für sRNAi4-dsRNA zur
Suppression mehrerer Speicherproteine

146. SEQ ID NO: 146
Nukleinsäuresequenz kodierend für sRNAi8-dsRNA zur
Suppression mehrerer Speicherproteine
147. SEQ ID NO: 147
Ribonukleinsäuresequenz kodierend für sRNAi8-dsRNA zur
Suppression mehrerer Speicherproteine
148. SEQ ID NO: 148 Oligonukleotidprimer OPN11
149. SEQ ID NO: 149 Oligonukleotidprimer OP12
150. SEQ ID NO: 150 Oligonukleotidprimer OPN13
151. SEQ ID NO: 151 Oligonukleotidprimer OPN15
152. SEQ ID NO: 152 Oligonukleotidprimer OPN16
153. SEQ ID NO: 153 Oligonukleotidprimer OPN17
154. SEQ ID NO: 154
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana*
"globulin-like protein" (GenBank Acc.-No.: NM_119834)
155. SEQ ID NO: 155
Proteinsequenz kodierend für *Arabidopsis thaliana* "globulin-
like protein" (Protein_id="NP_195388.1)
156. SEQ ID NO: 156
Nukleinsäuresequenz kodierend für Glycine max 7S Samen-
globulin (GenBank Acc.-No.: U59425)
157. SEQ ID NO: 157
Proteinsequenz kodierend für für Glycine max 7S Samenglobulin
158. SEQ ID NO: 158
Nukleinsäuresequenz kodierend für Zea mays 19kD Zein (GenBank
Acc.-No.: E01144)
159. SEQ ID NO: 159
Proteinsequenz kodierend für Zea mays 19kD Zein
160. SEQ ID NO: 160
Nukleinsäuresequenz kodierend für Zea mays 19kD alpha Zein B1
(GenBank Acc.-No.: AF371269)

161. SEQ ID NO: 161
Proteinsequenz kodierend für Zea mays 19kD alpha Zein B1
162. SEQ ID NO: 162
Nukleinsäuresequenz kodierend für Zea mays 19kD alpha Zein B2
(GenBank Acc.-No.: AF371270)
163. SEQ ID NO: 163
Proteinsequenz kodierend für Zea mays 19kD alpha Zein B2
164. SEQ ID NO: 164
Nukleinsäuresequenz kodierend für Zea mays 22kD alpha-zein
(GenBank Acc.-No.: X61085)
165. SEQ ID NO: 165
Proteinsequenz kodierend für Zea mays 22kD alpha-zein
166. SEQ ID NO: 166
Nukleinsäuresequenz kodierend für Oryza sativa Prolamin
(GenBank Acc.-No.: AB016503)
167. SEQ ID NO: 167
Proteinsequenz kodierend für Oryza sativa Prolamin
168. SEQ ID NO: 168
Nukleinsäuresequenz kodierend für A.sativa Avenin (GenBank
Acc.-No.: M38446)
169. SEQ ID NO: 169
Proteinsequenz kodierend für A.sativa Avenin
170. SEQ ID NO: 170
Nukleinsäuresequenz kodierend für Hordeum vulgare C-Hordein
(GenBank Acc.-No.: M36941)
171. SEQ ID NO: 171
Proteinsequenz Teil 1 kodierend für Hordeum vulgare C-Hordein
172. SEQ ID NO: 172
Proteinsequenz Teil 2 kodierend für Hordeum vulgare C-Hordein
173. SEQ ID NO: 173
Nukleinsäuresequenz kodierend für Triticum aestivum LMW
Glutenin-1D1 (GenBank Acc.-No.: X13306)

174. SEQ ID NO: 174

Proteinsequenz kodierend für *Triticum aestivum* LMW
Glutenin-1D1

175. SEQ ID NO: 175

Expressionskassette für *Arabidopsis thaliana* ACCase (GenBank
Acc.-No.: D34630) als Fusionsprotein mit der plastidären
Signalsequenz der Transketolase aus Tabak unter Kontrolle
des Napin-Promotors

176. SEQ ID NO: 176

Proteinsequenz kodierend ein Fusionsprotein aus der
Arabidopsis thaliana ACCase (GenBank Acc.-No.: D34630) und
der plastidären Signalsequenz der Transketolase aus Tabak

177. SEQ ID NO: 177

Nukleotidsequenz kodierend für *Arabidopsis thaliana* ACCase
(GenBank Acc.-No.: D34630) als Fusionsprotein mit der
plastidären Signalsequenz der Transketolase aus Tabak
unter Kontrolle des Napin-Promotors

178. SEQ ID NO: 178

Binärer Expressionsvektor für Agrobakterium vermittelte
Pflanzentransformation pSUN2-USP1,2,3.

179. SEQ ID NO: 179

Binärer Expressionsvektor für Agrobakterium vermittelte
Pflanzentransformation pSUN2-USP.

Abbildungen

1. Fig. 1a-b: Alignment von Brassica 2S Albuminen

Angegeben sind die für die einzelnen Speicherproteinen
kodierenden Nukleinsäuresequenzen mit ihren jeweiligen
GenBank Acc.-No.:

J02586 *Brassica napus* 2S Albumin (SEQ ID NO: 31)

X65040 *Brassica juncea* 2S Albumin (SEQ ID NO: 33)

X65038 *Brassica oleracea* 2S Albumin (SEQ ID NO: 35)

U04945 *Brassica napus* cv. Topas 2S Albumin (SEQ ID NO: 37)

X91799 *Sinapis alba* 2S Albumin (SEQ ID NO: 39)

61

2. Fig. 2: Alignment von Soja (Glycine max) 2S Albuminen

Angegeben sind die für die einzelnen Speicherproteinen
kodierenden Nukleinsäuresequenzen mit ihren jeweiligen
GenBank Acc.-No.:

U71194 Glycine max 2S Albumin 1	(SEQ ID NO: 41)
AF005030 Glycine max 2S Albumin	(SEQ ID NO: 43)

3. Fig. 3a-b: Alignment von Brassica nigra, Sinapis alba und
Helianthus annuus 2S Albuminen.

Angegeben sind die für die einzelnen Speicherproteinen
kodierenden Nukleinsäuresequenzen mit ihren jeweiligen
GenBank Acc.-No.:

X65039 Brassica napus 2S Albumin	(SEQ ID NO: 45)
X91798 Sinapis alba 2S Albumin	(SEQ ID NO: 47)
X06410 Helianthus annuus 2S Albumin HaG5	(SEQ ID NO: 49)
X76101 Helianthus annuus 2S Albumin	(SEQ ID NO: 51)

4. Fig. 4a-b: Alignment von Helianthus annuus 2S Albuminen.

Angegeben sind die für die einzelnen Speicherproteinen
kodierenden Nukleinsäuresequenzen mit ihren jeweiligen
GenBank Acc.-No.:

X06410 Helianthus annuus 2S Albumin HaG5	(SEQ ID NO: 49)
X76101 Helianthus annuus 2S Albumin	(SEQ ID NO: 51)

5. Fig. 5: Alignment von Arabidopsis thaliana 2S Albuminen.

Angegeben sind die für die einzelnen Speicherproteinen
kodierenden Nukleinsäuresequenzen mit ihren jeweiligen
GenBank Acc.-No.:

gi 166609:951-1445 (M22032) At2S1	(SEQ ID NO: 1)
gi 166611:212-706 (M22035) At2S3	(SEQ ID NO: 3)
gi 166610:375-887 (M22034) At2S2	(SEQ ID NO: 5)
gi 166612:500-1000 (M22033) At2S4	(SEQ ID NO: 7)

6. Fig. 6a-c: DNA-Alignment von Brassica napus 11S Speicherproteinen

Angegeben sind die für die einzelnen Speicherproteinen kodierenden Nukleinsäuresequenzen mit ihren jeweiligen

GenBank Acc.-No.:

X59294 CRU1_Bn	(SEQ ID NO: 9)
X14555 CRUA_Bn	(SEQ ID NO: 11)
X59295 CRU2_Bn	(SEQ ID NO: 13)
X57848 CRU4_Bn	(SEQ ID NO: 15)
X62120 CRU3_BN	(SEQ ID NO: 17)

(sequence ID from soybean 11S storage proteins after alignement)

7. Fig.7a-d:

DNA-Alignment von Soja 11S Speicherproteinen

Angegeben sind die für die einzelnen Speicherproteinen kodierenden Nukleinsäuresequenzen mit ihren jeweiligen

GenBank Acc.-No.:

GLC1: M36686; D00566 GMGLYBSU_2	(SEQ ID NO: 19)
GLC2: X15122 GMGY2_8	(SEQ ID NO: 21)
GLC3: X02626 GMGLYR1_3	(SEQ ID NO: 23)
GLC4: M10962 GMGLYAB_4	(SEQ ID NO: 25)
GLC5: X15123; S44896 GMGY3_7	(SEQ ID NO: 27)

8. Fig. 8:

Ergebnisse der Messung des Lipidgehalts (TAG) in T3 Samen transgener Arabidopsis Pflanzen, die mit dem Konstrukt Napin-TK-AtACCcase transformiert wurden. Samen von jeweils 10 individuellen T2 Pflanzen pro unabhängiger transgener Linie wurden wie in Beispiel 9 beschrieben gemessen. Die Menge an TAG in Prozent entspricht dem Anteil an der Gesamtmasse des Trockengewichts der Samen. Wt bezeichnet die nicht-transformierte Kontrolle, 5L, 3L, 12L und 1L sind die unabhängigen transgenen Linien. Die Federbalken beschreiben den Standardfehler aus allen gemessenen Werten.

9. Fig.9:

Schematische Darstellung der Speicherprotein-Suppressionskonstrukte.

- A: Insert (1155 bp) aus Vektor pCR2.1-AtCRU3-RNAi (5025 bp) kodierend für eine die AtCru3-Expression supprimierende dsRNA.
- B: Insert aus Vektor pCR2.1-sRNAi4 (1) und pCR2.1-sRNAi8 (2) kodierend für eine die AtCru3-, AtCRB und At2S3-Expression supprimierende dsRNA. In den beiden Konstrukten sind die "sense"-RNA-Stränge und "anti-sense"-RNA-Stränge für die einzelnen zu supprimierenden Ziele unterschiedlich angeordnet. Schraffierte Bereiche stellen Intronsequenzen (Linker) dar.

10. Fig.10: Ergebnis der quantitativen Ölbestimmung.

Samen von unabhängigen transgenen Pflanzen der T1 Generation wurden analysiert. Angegeben ist der Fettsäuregehalt in Gewichtsprozent des Samens (%FA [w/w]). Mehrere unabhängige Linien (2-1; 3-1; 5-1; 12-1; 14-1; 16-1; 17-1) zeigten einen signifikant erhöhten Ölgehalt (bis zu 38 % erhöhten Ölgehalten für Linie 16-1) verglichen zu Kontrollpflanzen (C).

Beispiele

Allgemeine Methoden:

Alle Chemikalien, wenn nicht anders erwähnt, stammen von den Firmen Fluka (Buchs), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deisenhofen). Restriktionsenzyme, DNA-modifizierende Enzyme und Molekularbiologie-Kits wurden von den Firmen Amersham-Pharmacia (Freiburg), Biometra (Göttingen), Roche (Mannheim), New England Biolabs (Schwalbach), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Qiagen (Hilden), Stratagen (Amsterdam, Niederlande), Invitrogen (Karlsruhe) und Ambion (Cambridgeshire, United Kingdom). Die verwendeten Reagenzien wurden entsprechend der Angaben des Herstellers eingesetzt.

Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die im Rahmen der vorliegenden Erfindung durchgeführten Klonierungsschritte wie z.B. Restriktionsspaltungen, Agarosegelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrozellulose und Nylonmembranen, Verknüpfen von

DNA-Fragmenten, Transformation von *E. coli* Zellen, Anzucht von Bakterien, Vermehrung von Phagen und Sequenzanalyse rekombinanter DNA werden wie bei Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6 beschrieben durchgeführt. Die Sequenzierung rekombinanter DNA-Moleküle erfolgt mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma ABI nach der Methode von Sanger (Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-5467).

Beispiel 1: Allgemeine Verfahren

Die Pflanze *Arabidopsis thaliana* repräsentiert ein Mitglied der höheren Pflanzen (Samenpflanzen). Diese Pflanze ist eng verwandt mit anderen Pflanzenarten aus der Familie der Cruciferen wie z.B. *Brassica napus*, aber auch mit anderen Pflanzenfamilien der Dikotyledonen. Aufgrund des hohen Grades an Homologie ihrer DNA-Sequenzen bzw. Polypeptidsequenzen kann *Arabidopsis thaliana* als Modellpflanze für andere Pflanzenarten eingesetzt werden.

a) Anzucht von *Arabidopsis* Pflanzen

Die Pflanzen werden entweder auf Murashige-Skoog Medium mit ,5 % Saccharose (Ogas et al. (1997) Science 277:91-94) oder auf Erde gezogen (Focks & Benning (1998) Plant Physiol 118:91-101). Um einheitliche Keimungs- und Blühzeiten zu erreichen, werden die Samen nach Ausplattieren bzw. Ausstreuen auf Erde zwei Tage bei 4°C stratifiziert. Nach der Blüte werden die Schoten markiert. Entsprechend der Markierungen werden dann Schoten mit einem Alter von 6 bis 20 Tagen nach der Blüte geerntet.

b) Isolierung von total RNA und poly-A⁺ RNA aus Pflanzen

Für die Herstellung von Suppressionskonstrukten wird RNA bzw. polyA⁺ RNA isoliert. RNA wurde aus Schoten von *Arabidopsis* Pflanzen nach folgender Vorschrift isoliert: Schotenmaterial im Alter von 6 bis 20 Tage nach Blüte wurde geerntet und in flüssigem Stickstoff schockgefroren. Das Material wurde vor der weiteren Verwendung bei -80°C gelagert. 75 mg des Materials wurde im gekühlten Mörser zu einem feinem Pulver gemahlen und mit 200 µL des Lysis-Puffers aus dem Ambion RNAqueos-Kit versetzt. Die Isolierung der totalen RNA wurde dann nach Herstellerangaben durchgeführt. Die RNA wurde mit 50 µL Elutionspuffer (Ambion) eluiert und die Konzentration durch Absorption einer 1 zu 100 verdünnten Lösung am Photometer (Eppendorf) bei 260 nm bestimmt. 40 µg/ml RNA entspricht dabei einer Absorption von 1. Die RNA-Lösungen wurden mit RNase

freiem Wasser auf eine Konzentration von 1 µg/µL eingestellt. Die Konzentrationen wurden durch Agarosegelelektrophorese überprüft. Zur Isolierung von polyA⁺ RNA wurde oligo(dT)-Zellulose von Amersham Pharmacia nach Herstellerangaben verwendet. RNA bzw. polyA⁺ RNA wurde bei -70°C gelagert.

c) Konstruktion der cDNA-Bank

Zur Konstruktion der cDNA-Bank aus *Arabidopsis* Schoten-RNA wurde die Erststrangsynthese unter Verwendung von Reverser Transkriptase aus Maus-Leukämie-Virus (Clontech) und Oligo-d(T)-Primern, die Zweitstrangsynthese durch Inkubation mit DNA-Polymerase I, Klenow-Enzym und RNase H-Spaltung bei 12°C (2 Std.), 16°C (1 Std.) und 22°C (1 Std.) erzielt. Die Reaktion wurde durch Inkubation bei 65°C (10 min) gestoppt und anschließend auf Eis überführt. Doppelsträngige DNA-Moleküle wurde mit T4-DNA-Polymerase (Roche, Mannheim) bei 37°C (30 min) mit glatten Enden versehen. Die Nukleotide wurden durch Phenol/Chloroform-Extraktion und Sephadex-G50-Zentrifugiersäulen entfernt. EcoRI/XhoI-Adapter (Pharmacia, Freiburg, Deutschland) wurden mittels T4-DNA-Ligase (Roche, 12°C, über Nacht) an die cDNA-Enden ligiert, mit XhoI nachgeschnitten und durch Inkubation mit Polynukleotidkinase (Roche, 37°C, 30 min) phosphoryliert. Dieses Gemisch wurde der Trennung auf einem Low-Melting-Agarose-Gel unterworfen. DNA-Moleküle über 300 Basenpaaren wurden aus dem Gel eluiert, Phenol-extrahiert, auf Elutip-D-Säulen (Schleicher und Schüll, Dassel, Deutschland) konzentriert und an Vektorarme ligiert und in lambda-ZAPII-Phagen oder lambda-ZAP-Express-Phagen unter Verwendung des Gigapack Gold-Kits (Stratagene, Amsterdam, Niederlande) verpackt, wobei Material des Herstellers verwendet und seine Anweisungen befolgt wurden.

d) Isolierung von genomischer DNA aus Pflanzen wie *Arabidopsis thaliana* oder *Brassica napus* (CTAB-Methode)

Zur Isolierung genomischer DNA aus Pflanzen wie *Arabidopsis thaliana* oder *Brassica napus* werden ca. 0,25 g Blattmaterial junger Pflanzen im vegetativen Stadium in flüssigem Stickstoff zu feinem Pulver gemörsert. Das pulverisierte Pflanzenmaterial wird zusammen mit 1 ml 65°C-warmem CTAB I-Puffer (CTAB: Hexadecyltrimethylammoniumbromid, auch genannt Cetyltrimethylammoniumbromid; Sigma Kat.-Nr.: H6269) und 20 µl β-Mercaptoethanol in einen vorgewärmten zweiten Mörser gegeben und nach vollständiger Homogenisierung wird der Extrakt in ein 2 ml Eppendorf-Gefäß überführt und für 1 h bei 65°C unter regelmäßiger, vorsichtiger Durchmischung

inkubiert. Nach Abkühlung auf Raumtemperatur wird der Ansatz mit 1 ml Chloroform/Octanol (24:1, mit 1M Tris/HCl, pH 8,0 ausgeschüttelt) durch langsames Invertieren extrahiert und zur Phasentrennung für 5 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Anschließend wird die wässrige Phase erneut mit 1 ml Chloroform/Octanol extrahiert, zentrifugiert und durch Invertieren mit 1/10 Volumen auf 65°C vorgewärmtem CTAB II-Puffer sorgfältig gemischt. Anschließend wird der Ansatz durch vorsichtiges Schwenken mit 1 ml Chloroform/Octanol-Gemisch (siehe oben) versetzt und zur erneuten Phasentrennung für 5 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Die wässrige untere Phase wird in ein frisches Eppendorf-Gefäß überführt und die obere organische Phase wird in einem frischen Eppendorf-Gefäß erneut für 15 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Die hieraus resultierende wässrige Phase wird mit der wässrigen Phase des vorherigen Zentrifugationsschrittes vereinigt und der gesamte Ansatz mit exakt demselben Volumen vorgewärmtem CTAB III-Puffer versetzt. Es folgt eine Inkubation bei 65°C, bis die DNA in Flocken ausfällt. Dies kann bis zu 1 h dauern oder durch Inkubation bei 37°C über Nacht erfolgen. Das aus dem anschließenden Zentrifugationsschritt (5 min, 2000 rpm (500 x g), 4°C) resultierende Sediment wird mit 250 µl auf 65°C vorgewärmtem CTAB IV-Puffer versetzt und für mindestens 30 min bzw. bis zur vollständigen Auflösung des Sediments bei 65°C inkubiert. Anschließend wird die Lösung zur Fällung der DNA mit 2,5 Volumina eiskaltem Ethanol vermischt und für 1h bei -20°C inkubiert. Alternativ kann der Ansatz mit 0.6 Volumina Isoopropanol vermischt und ohne weitere Inkubation sofort für 15 min bei 8,500 rpm (7,500 x g) und 4°C zentrifugiert werden. Die sedimentierte DNA wird durch Invertieren des Eppendorf-Gefäßes zweimal mit je 1 ml 80%igem eiskaltem Ethanol gewaschen, nach jedem Waschschrift erneut zentrifugiert (5 min, 8,500 rpm (7,500 x g), 4°C) und anschließend für ca. 15 min luftgetrocknet. Abschließend wird die DNA in 100 µl TE mit 100 µg/ml RNase resuspendiert und für 30 min bei Raumtemperatur inkubiert. Die DNA Lösung ist nach einer weiteren Inkubationsphase über Nacht bei 4°C homogen und kann für weiterführende Experimente verwendet werden.

Lösungen für CTAB:

Lösung I (für 200 ml):

- 100 mM Tris/HCl pH 8,0 (2,42 g)
- 1,4 M NaCl (16,36 g)
- 20 mM EDTA (8,0 ml von 0,5 M Stammlösung)

67

2 % (w/v) CTAB (4,0 g)

Jeweils vor der Verwendung werden frisch zugesetzt:
2 % β -Mercaptoethanol (20 μ l für 1 ml Lösung I).

Lösung II (für 200 ml):

0,7 M NaCl (8,18 g)

10 % (w/v) CTAB (20 g)

Lösung III (für 200 ml):

50 mM Tris/HCl pH 8,0 (1,21 g)

10 mM EDTA (4 ml 0,5 M von 0,5 M Stammlösung)

1 % (w/v) CTAB (2,0 g)

Lösung IV (High-salt TE) (für 200 ml):

10 mM Tris/ HCl pH 8,0 (0,242 g)

0,1 mM EDTA (40 μ l 0,5 M Stammlösung)

1 M NaCl (11, 69 g)

Chloroform/Octanol (24:1) (für 200 ml):

192 ml Chloroform

8 ml Octanol

Die Mischung wird 2x mit 1 M TrisHCl pH 8,0 ausgeschüttelt
und vor Licht geschützt gelagert.

Beispiel 2: Herstellung von Suppressionskonstrukten

Ausgehend von der genomischer Arabidopsis thaliana DNA oder cDNA wurden über PCR mittels der aufgeführten Oligonukleotide folgende Fragmente von Speicherproteinsequenzen amplifiziert. Dabei kam nachfolgendes PCR Protokoll zum Einsatz:

Zusammensetzung des PCR-Ansatzes (50 μ L):

5,00 μ L Template cDNA oder genomische DNA (ca. 1 μ g)

5,00 μ L 10x Puffer (Advantage-Polymerase) + 25 mM $MgCl_2$

5,00 μ L 2mM dNTP

1,25 μ L je Primer (10 pmol/ μ L)

0,50 μ L Advantage-Polymerase (Clontech)

PCR-Programm: Anfangsdenaturierung für 2 min bei 95°C, dann 35 Zyklen mit 45 sec 95°C, 45 sec 55°C und 2 min 72°C. Abschließende Extension von 5 min bei 72°C.

- a) Suppressionskonstrukt für *Arabidopsis thaliana* 12S Speicherprotein AtCRU3 (SEQ ID NO: 111 bzw. 112; GenBank Acc.-No: U66916):

Für die den sense-Strang der dsRNA und den Linker kodierende Sequenz wurde aus genomischer *Arabidopsis thaliana* DNA mit nachfolgendem Oligonukleotid-Primerpaar ein Exonbereich mit dem vollständigen anschließenden Intron einschließlich der an das Intron anschließenden Spleiß-Akzeptorsequenz (Basenpaar 1947 bis 2603 der Sequenz mit der GenBank Acc.-No: U66916) amplifiziert:

ONP1 (SEQ ID NO: 134):

5'-ATAAGAATGCGGCCGCGTGTTCATTTGGCCGGAACAAC-3'

ONP2 (SEQ ID NO: 135):

5'-CCCGGATCCTTCTGTAACATTTGACAAAACATG-3'

Das PCR-Produkt wurden in den pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in dem pCR2.1-1 Vektor und die Sequenz überprüft.

Für die den antisense-Strang der dsRNA kodierende Sequenz wurde aus *Arabidopsis thaliana* cDNA lediglich das gleiche Exon wie oben (Basenpaar 1947 bis 2384) mit dem nachfolgenden Primerpaar amplifiziert:

ONP3 (SEQ ID NO: 136):

5'ATAAGAATGCGGCCGCGTGTTCATTTGGCCGGAACAAC-3'

ONP4 (SEQ ID NO: 137):

5'ATAAGAATGCGGCCGCGGATCCACCCTGGAGAACGCCACGAGTG-3'

Das PCR-Produkt wurden in den pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in dem pCR2.1-2 Vektor und die Sequenz überprüft.

0,5 µg von Vektor pCR2.1-1 wurden mit dem Restriktionsenzym BamHI (New England Biolabs) für 2 Stunden nach Herstellerangaben inkubiert und dann für 15 min mit alkalischer Phosphatase (New England Biolabs) dephosphoryliert. Der so präparierte Vektor (1 µL) wurde dann mit dem aus Vektor pCR2.1-2 gewonnenen Fragment ligiert. Dazu wurden 0,5 µg von Vektor pCR2.1-2 2 Stunden mit BamHI (New England Biolabs) verdaut und die DNA-Fragmente per Gelelektrophorese aufgetrennt. Das neben dem Vektor (3,9 kb) entstandene 489 bp große Stück wurde aus dem Gel ausgeschnitten und mit dem

"Gelpurification"-Kit (Qiagen) nach Herstellerangaben aufgereinigt und mit 50 µL Elutionspuffer eluiert. 10 µL des Eluats wurden mit Vektor pCR2.1-1 (s.o.) über Nacht bei 16°C ligiert (T4 Ligase, New England Biolabs). Die Ligationsprodukte wurden dann in TOP10 Zellen (Stratagene) nach Herstellerangaben transformiert und entsprechend selektioniert. Positive Klone wurden mit dem Primerpaar ONP1 und ONP2 durch PCR verifiziert. Der erhaltene Vektor pCR2.1-AtCRU3-RNAi wurde dann für 2 Stunden mit NotI (New England Biolabs) inkubiert, mit Klenow-Fragment "geblunted" und die DNA-Fragmente über Gelelektrophorese analysiert. Das 1155 bp große Fragment wurde dann in den mit StuI geschnittenen, dephosphorylierten binären Vektor pSUN2-USP1,2,3 (SEQ ID NO: 178; siehe Beispiel 5) ligiert. Bei dem Vektor pSUN2-USP1,2,3 handelt es sich um ein Derivat des Vektors pPZP111 ((Hajdukiewicz, P et al. 1994 Plant Mol Biol 25:989-994), in dem drei Expressionskassetten, jeweils unter Kontrolle des USP-Promotors, vorliegen. Vektoren mit der gewünschten Orientierung des Inserts wurden mittels Restriktionsverdau und Sequenzierung ermittelt. Das entstehende Konstrukt trägt die Bezeichnung pSUN2-USP-RNAi-a. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 105 beschrieben.

- b) Suppressionskonstrukt für *Arabidopsis thaliana* 12S Speicherprotein AtCRB (SEQ ID NO: 117 bzw. 118; GenBank Acc.-No: X14313; M37248):

Für die den sense-Strang und antisense-Strang der dsRNA wird mit nachfolgendem Oligonukleotid-Primerpaar ein Exonbereich (Basenpaar 601 bis 1874 der Sequenz mit der GenBank Acc.-No: M37248) aus *Arabidopsis thaliana* cDNA amplifiziert:

ONP5 (SEQ ID NO: 138):

5' ATAAGAATGCGGCCGCGGATCCCTCAGGGTCTTTTCTTGCCCACT-3'

ONP6 (SEQ ID NO: 139):

5'-CCGCTCGAGTTTACGGATGGAGCCACGAAG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in dem pCR2.1-3 Vektor und die Sequenz überprüft.

Für den als Linker fungierenden Bereich wird aus *Arabidopsis thaliana* genomischer DNA ein Intron mit den entsprechenden Spliceakzeptor und -donorsequenzen der flankierenden Exons

(Basenpaar 1874 bis 2117 der Sequenz mit der GenBank Acc.-No: M37248) mit dem nachfolgenden Primerpaar amplifiziert:

ONP7 (SEQ ID NO: 140):

5'-CCGCTCGAGGTAAGCTCAACAAATCTTTAG-3'

ONP8 (SEQ ID NO: 141):

5'-ACGCGTCGACGCGTTCTGCGTGCAAGATATT-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in dem pCR2.1-4 Vektor und die Sequenz überprüft.

Das Konstrukt für AtCRB wurde in einer ähnlichen Strategie wie für AtCRU3 erläutert, erstellt. Vektor pCR2.1-3 wird mit XhoI (New England Biolabs) für 2 Stunden inkubiert und dephosphoryliert (alkalische Phosphatase, New England Biolabs). Vektor pCR2.1-4 wird ebenfalls mit XhoI in derselben Weise inkubiert und die Gelfragmente per Gelelektrophorese aufgetrennt. Die entsprechenden Fragmente werden in der unter AtCRU3 beschriebenen Art und Weise aufgereinigt und ligiert, resultierend nach Bakterientransformation in dem Vektor pCR2.1-AtCRB Exon/Intron. Dieser Vektor wird für 2 Stunden mit XbaI (NEB), anschließend für 15 min mit Klenow-Fragment (NEB), dann für 2 Stunden mit SalI inkubiert und zuletzt 15 min mit alkalischer Phosphatase (NEB) behandelt. Parallel wird der Vektor pCR2.1-3 mit BamHI (NEB), dann 15 min mit Klenow-Fragment und anschließend 2 Stunden mit XhoI (NEB) inkubiert. Das Exon-Fragment von AtCRB wird nach Gelelektrophorese isoliert, gereinigt und zur Ligation eingesetzt. Beide Fragmente werden dann ligiert und der Vektor pCR2.1-AtCRB-RNAi resultiert.

Der erhaltene Vektor pCR2.1-AtCRB-RNAi wird dann für 2 Stunden mit HindIII und PvuI verdaut und 15 min mit Klenow-Fragment inkubiert (blunten). Das ausgeschnittene Fragment wird über Gelelektrophorese isoliert und dann für die Ligation eingesetzt. Das entsprechende Fragment wird dann in den mit EcoRV geschnittenen, dephosphorylierten Vektor pSUN2-USP-RNAi-a (siehe oben) ligiert. Vektoren mit der gewünschten Orientierung des Inserts werden mittels Restriktionsverdau und Sequenzierung ermittelt. Das entstehende Konstrukt trägt die Bezeichnung pSUN2-USP-RNAi-b. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 107 beschrieben.

- c) Suppressionskonstrukt für *Arabidopsis thaliana* 2S Speicherprotein At2S3 (SEQ ID NO: 3 bzw. 4; GenBank Acc.-No: M22035):

Für die den sense-Strang und antisense-Strang der dsRNA wird aus *Arabidopsis thaliana* cDNA mit nachfolgendem Oligonukleotid-Primerpaar ein Exonbereich (Basenpaar 212 bis 706 der Sequenz mit der GenBank Acc.-No: M22035) amplifiziert:

ONP9 (SEQ ID NO: 142):

5'-ATAAGAATGCGGCCGCGGATCCATGGCTAACAAGCTCTTCCTCGTC-3'

ONP10 (SEQ ID NO: 143):

5'-ATAAGAATGCGGCCGCGGATCCCTAGTAGTAAGGAGGGAAGAAAG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in dem pCR2.1-5 Vektor und die Sequenz überprüft.

Für den als Linker fungierenden Bereich wird das gleiche Intron wie unter b) mit den Primern OPN 7 und OPN 8 amplifiziert eingesetzt.

Das Konstrukt für At2S3 wird in einer ähnlichen Strategie wie für AtCRU3 erläutert, erstellt. Vektor pCR2.1-5 wird mit mit XhoI (New England Biolabs) für 2 Stunden inkubiert und dephosphoryliert (alkalische Phosphatase, New England Biolabs). Vektor pCR2.1-3 wird ebenfalls mit XhoI in derselben Weise inkubiert und die Gelfragmente per Gelelektrophorese aufgetrennt. Die entsprechenden Fragmente werden in der unter AtCRU3 beschriebenen Art und Weise aufgereinigt und ligiert, resultierend nach Bakterientransformation in dem Vektor pCR2.1-At2S3 Exon/Intron. Dieser Vektor wird für 2 Stunden mit SalI (NEB), anschließend für 15 min mit Klenow-Fragment (NEB) inkubiert und zuletzt 15 min mit alkalischer Phosphatase (NEB) behandelt. Parallel wird der Vektor pCR2.1-5 mit BamHI (NEB) und dann 15 min mit Klenow-Fragment inkubiert. Das Exon-Fragment von At2S3 wird nach Gelelektrophorese isoliert, gereinigt und zur Ligation eingesetzt. Beide Fragmente werden dann ligiert und der Vektor pCR2.1-At2S3-RNAi resultierte.

Der erhaltene Vektor pCR2.1-At2S3-RNAi wird dann für 2 Stunden mit HindIII und XbaI (New England Biolabs) verdaut und 15 min mit Klenow-Fragment inkubiert (blunten). Das ausgeschnittene Fragment wird über Gelelektrophorese isoliert und dann für die Ligation eingesetzt. Das entsprechende Fragment wird dann in den mit SmaI verdauten und dephosphorylierten Vektor pSUN2-USP-RNAi-b (siehe oben) ligiert. Vekto-

ren mit der gewünschten Orientierung des Inserts werden mittels Restriktionsverdau und Sequenzierung ermittelt. Das entstehende Konstrukt trägt die Bezeichnung pSUN2-USP-RNAi1. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 109 beschrieben.

Durch die oben erwähnten Klonierungsschritte befinden sich alle drei Konstrukte AtCRU3, AtCRB und At2S3 im binären Vektor pSUN2-USP 1,2,3 jeweils unter der Kontrolle des USP-Promotors (Bäumlein et al. 1991, Mol Gen Genet 225(3):459-67).

Beispiel 3: Plasmide für die Pflanzentransformation

Zur Pflanzentransformation können binäre Vektoren, wie pBinAR verwendet werden (Höfgén und Willmitzer (1990) Plant Science 66: 221-230). Die Konstruktion der binären Vektoren kann durch Ligation der cDNA in Sense- oder Antisense-Orientierung in T-DNA erfolgen. 5' der cDNA aktiviert ein Pflanzenpromotor die Transkription der cDNA. Eine Polyadenylierungssequenz befindet sich 3' von der cDNA.

Die gewebespezifische Expression lässt sich unter Verwendung eines gewebespezifischen Promotors erzielen. Beispielsweise kann die samenspezifische Expression erreicht werden, indem der Napin- oder der LeB4- oder der USP-Promotor 5' der cDNA einkloniert wird. Auch jedes andere samenspezifische Promotorelement kann verwendet werden. Zur konstitutiven Expression in der ganzen Pflanze lässt sich der CaMV-35S-Promotor verwenden.

Ein weiteres Beispiel für einen binären Vektor ist der Vektor pSUN2-USP1,2,3, in welchen die Fragmente aus Beispiel 2 kloniert wurden, sowie pSUN2-USP. Der Vektor pSUN2-USP enthält den USP-Promotor sowie den OCS Terminator. pSUN2-USP1,2,3, enthält dreimal den USP-Promotor. Die Fragmente aus Beispiel 2 wurden in die multiple Klonierungsstelle der Vektors pSUN2-USP1,2,3 kloniert, um die samenspezifische Expression der Suppressionskonstrukte zu ermöglichen.

Beispiel 4: Transformation von Agrobacterium

Die Agrobacterium-vermittelte Pflanzentransformation kann zum Beispiel unter Verwendung der Agrobacterium tumefaciens-Stämme GV3101 (pMP90) (Koncz und Schell (1986) Mol Gen Genet 204: 383- 396) oder LBA4404 (Clontech) durchgeführt werden. Die Transformation kann durch Standard-Transformationstechniken durchgeführt werden (Deblaere et al. (1984) Nucl Acids Res 13:4777-4788).

Beispiel 5: Pflanzentransformation

Die Agrobacterium-vermittelte Pflanzentransformation kann unter Verwendung von Standard-Transformations- und Regenerations-techniken durchgeführt werden (Gelvin, Stanton B., Schilperoort, Robert A., Plant Molecular Biology Manual, 2. Aufl., Dordrecht: Kluwer Academic Publ., 1995, in Sect., Ringbuc Zentrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

Die Transformation mittels Agrobacterium von Arabidopsis thaliana wurde durch die Methode nach Bechthold et al., 1993 (C.R. Acad. Sci. Ser. III Sci. Vie., 316, 1194-1199) durchgeführt. Beispielsweise kann Raps mittels Kotyledonen- oder Hypokotyltransformation transformiert werden (Moloney et al., Plant Cell Report 8 (1989) 238-242; De Block et al., Plant Physiol. 91 (1989) 694-701). Die Verwendung von Antibiotika für die Agrobacterium- und Pflanzenselektion hängt von dem für die Transformation verwendeten binären Vektor und Agrobacterium-Stamm ab. Die Rapsselektion wird gewöhnlich unter Verwendung von Kanamycin als selektierbarem Pflanzenmarker durchgeführt.

Der Agrobacterium-vermittelte Gentransfer in Lein (Linum usitatissimum) lässt sich unter Verwendung von beispielsweise einer von Mlynarova et al. (1994) Plant Cell Report 13:282-285 beschriebenen Technik durchführen.

Die Transformation von Soja kann unter Verwendung von beispielsweise einer in EP-A-0 0424 047 (Pioneer Hi-Bred International) oder in EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo) beschriebenen Technik durchgeführt werden.

Die Pflanzentransformation unter Verwendung von Teilchenbeschuss, Polyethylenglycol-vermittelter DNA-Aufnahme oder über die Siliziumcarbonatfaser-Technik ist beispielsweise beschrieben von Freeling und Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

Beispiel 6: Untersuchung der Expression eines rekombinanten Genproduktes in einem transformierten Organismus

Die Aktivität eines rekombinanten Genproduktes im transformierten Wirtsorganismus wurde auf der Transkriptions- und/oder der Translationsebene gemessen.

Ein geeignetes Verfahren zur Bestimmung der Menge an Transkription des Gens (ein Hinweis auf die Menge an RNA, die für die Translation des Genproduktes zur Verfügung steht) ist die Durchführung eines Northern-Blots wie unten ausgeführt (als Bezugsstelle siehe Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York, oder den oben erwähnten Beispielteil), wobei ein Primer, der so gestaltet ist, dass er an das Gen von Interesse bindet, mit einer nachweisbaren Markierung (gewöhnlich radioaktiv oder chemilumineszent) markiert wird, so dass, wenn die Gesamt-RNA einer Kultur des Organismus extrahiert, auf einem Gel aufgetrennt, auf eine stabile Matrix transferiert und mit dieser Sonde inkubiert wird, die Bindung und das Ausmaß der Bindung der Sonde das Vorliegen und auch die Menge der mRNA für dieses Gen anzeigt. Diese Information zeigt den Grad der Transkription des transformierten Gens an. Zelluläre Gesamt-RNA kann aus Zellen, Geweben oder Organen mit mehreren Verfahren, die alle im Fachgebiet bekannt sind, wie zum Beispiel das von Bormann, E.R., et al. (1992) Mol. Microbiol. 6:317-326 beschriebene, präpariert werden.

Northern-Hybridisierung:

Für die RNA-Hybridisierung wurden 20 µg Gesamt-RNA oder 1 µg poly(A)⁺-RNA mittels Gelelektrophorese in Agarosegelen mit einer Stärke von 1,25 % unter Verwendung von Formaldehyd, wie beschrieben in Amasino (1986, Anal. Biochem. 152, 304) aufgetrennt, mittels Kapillaranziehung unter Verwendung von 10 x SSC auf positiv geladene Nylonmembranen (Hybond N+, Amersham, Braunschweig) übertragen, mittels UV-Licht immobilisiert und 3 Stunden bei 68°C unter Verwendung von Hybridisierungspuffer (10 % Dextransulfat Gew./Vol., 1 M NaCl, 1 % SDS, 100 mg Heringsperma-DNA) vorhybridisiert. Die Markierung der DNA-Sonde mit dem Highprime DNA labeling-Kit (Roche, Mannheim, Deutschland) erfolgte während der Vorhybridisierung unter Verwendung von alpha-³²P-dCTP (Amersham Pharmacia, Braunschweig, Deutschland). Die Hybridisierung wurde nach Zugabe der markierten DNA-Sonde im gleichen Puffer bei 68°C über Nacht durchgeführt. Die Waschschritte wurden zweimal für 15 min unter Verwendung von 2 X SSC und zweimal für 30 min unter Verwendung von 1 X SSC, 1 % SDS, bei 68°C durchgeführt. Die Exposition der verschlossenen Filter wurde bei -70°C für einen Zeitraum von 1 bis 14 T durchgeführt.

Zur Untersuchung des Vorliegens oder der relativen Menge an von dieser mRNA translatiertem Protein können Standardtechniken, wie ein Western-Blot, eingesetzt werden (siehe beispielsweise Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York). Bei diesem Verfahren werden die zellulären Gesamt-

Proteine extrahiert, mittels Gelelektrophorese aufgetrennt, auf eine Matrix, wie Nitrozellulose, übertragen und mit einer Sonde, wie einem Antikörper, der spezifisch an das gewünschte Protein bindet, inkubiert. Diese Sonde ist gewöhnlich mit einer chemilumineszenten oder kolorimetrischen Markierung versehen, die sich leicht nachweisen lässt. Das Vorliegen und die Menge der beobachteten Markierung zeigt das Vorliegen und die Menge des gewünschten, in der Zelle vorliegenden mutierten Proteins an.

Beispiel 7: Analyse der Auswirkung der rekombinanten Proteine auf die Produktion des gewünschten Produktes

Die Auswirkung der genetischen Modifikation in Pflanzen, Pilzen, Algen, Ciliaten oder auf die Produktion einer gewünschten Verbindung (wie einer Fettsäure) kann bestimmt werden, indem die modifizierten Mikroorganismen oder die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen) gezüchtet werden und das Medium und/oder die zellulären Komponenten auf die erhöhte Produktion des gewünschten Produktes (d.h. von Lipiden oder einer Fettsäure) untersucht wird. Diese Analysetechniken sind dem Fachmann bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie (siehe beispielsweise Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et al. (1993) Biotechnology, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988) Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

Neben den oben erwähnten Verfahren werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145, beschrieben extrahiert. Die qualitative und quantitative Lipid- oder Fettsäureanalyse ist beschrieben bei Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide -

Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

Zusätzlich zur Messung des Endproduktes der Fermentation ist es auch möglich, andere Komponenten der Stoffwechselwege zu analysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamteffizienz der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (z.B. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere Ionen), Messungen der Biomassezusammensetzung und des Wachstums, Analyse der Produktion üblicher Metabolite von Biosynthesewegen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F. Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192 (ISBN: 0199635773) und darin angegebenen Literaturstellen beschrieben.

Ein Beispiel ist die Analyse von Fettsäuren (Abkürzungen: FAME, Fettsäuremethylester; GC-MS, Gas-Flüssigkeitschromatographie-Massenspektrometrie; TAG, Triacylglycerin; TLC, Dünnschichtchromatographie).

Der unzweideutige Nachweis für das Vorliegen von Fettsäureprodukten kann mittels Analyse rekombinanter Organismen nach Standard-Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: Advances on Lipid Methodology, Vierte Aufl.: Christie, Oily Press, Dundee, 119-169; 1998, Gaschromatographie-Massenspektrometrie-Verfahren, Lipide 33:343-353).

Das zu analysierende Material kann durch Ultraschallbehandlung, Mahlen in der Glasmühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren aufgebrochen werden. Das Material muss nach dem Aufbrechen zentrifugiert werden. Das Sediment wird in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit 2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl- und Lipidverbindungen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester werden in Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapillarsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 mikrom, 0,32 mm) bei einem Temperaturgradienten zwischen 170°C und 240°C für 20 min

und 5 min bei 240°C unterworfen. Die Identität der erhaltenen Fettsäuremethylester muss unter Verwendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma), definiert werden.

Für die Öl-Analyse der mit den Suppressionskonstrukten transformierten Arabidopsis Pflanzen wurde folgendes Protokoll angewendet:

Die Extraktion der Lipide aus Samen wird nach der Methode von Bligh & Dyer (1959) Can J Biochem Physiol 37:911 durchgeführt. Dazu werden 5 mg Arabidopsis Samen in 1,2 ml Qiagen-Microtubes (Qiagen, Hilden) auf einer Sartorius (Göttingen) Mikrowaage abgewogen. Das Samenmaterial wird mit 500 µL Chloroform/Methanol (2:1; enthält Mono-C17-glycerin von Sigma als internen Standard) in der Rätschmühle MM300 der Firma Retsch (Haan) homogenisiert und 20 min bei RT inkubiert. Nach Zugabe von 500 µL 50 mM Kaliumphosphatpuffer pH 7,5 erfolgt die Phasentrennung. Von der organischen Phase werden 50 µL abgenommen, mit 1500 µL Chloroform verdünnt und 5 µL auf die Kapillaren Chromarods SIII der Firma Iatroscan (SKS, Bechenheim) aufgetragen. Nach Auftrag der Proben werden diese für 15 min in einer Dünnschichtkammer, die gesättigt ist mit 6:2:2 Chloroform: Methanol: Toluol in einem ersten Schritt aufgetrennt. Nach Ablauf der Zeit werden die Kapillaren 4 min bei Raumtemperatur getrocknet und dann für 22 min in eine Dünnschichtkammer, die gesättigt ist mit 7:3 n-Hexan:Diethylether gestellt. Nach einem weiteren Trocknungsschritt für 4 min bei Raumtemperatur werden die Proben in einem Iatroscan MK-5 (SKS, Bechenheim) entsprechend Fraser & Taggart, 1988 J. Chromatogr. 439:404 analysiert. Folgende Parameter wurden für die Messungen eingestellt: Slice width 50 msec, Treshold 20 mV, Noise 30, Skim ratio 0. Die Quantifizierung der Daten erfolgte anhand des internen Standards Mono-C17-glycerin (Sigma) sowie einer erstellten Eichkurve mit Tri-C17-glycerin (Sigma) mittels des Programms ChromStar (SKS, Bechenheim).

Für die quantitative Bestimmung der Ölgehalte wurden Samen von unabhängigen transgenen Pflanzen der T1 Generation analysiert. Die geernteten Samen wurde entsprechend der beschriebenen Analysemethoden aufgearbeitet und die Ölgehalte quantitativ bestimmt (Fig. 10). Dabei konnten mehrere unabhängige Linien identifiziert werden, die einen signifikant erhöhten Ölgehalt verglichen zu Kontrollpflanzen zeigen, die mit dem Vektor ohne das Suppressionskonstrukt transformiert wurden. So ergeben sich für die transgenen Linien Samen mit einem um bis zu 38 % erhöhten Ölgehalten (Linie 16-1 zeigt die höchste Ölmenge). Der Vergleich mit den parallel gewachsenen Kontrollpflanzen verdeutlicht die Spezifität

und damit die Funktionalität des eingesetzten Konstruktes. Damit konnte gezeigt werden, dass eine spezifische Verringerung einer Speicherproteinklasse zu einer Erhöhung des Ölgehaltes in Samen führt.

Beispiel 8: Klonierung der Super-RNAi Konstrukte für die Suppression von Speicherproteinen aus verschiedenen Klassen

a) Herstellung von Super-Suppressionskonstrukt 1

Die Vektoren pCR2.1-AtCRU3-RNAi und pCR2.1-4 (siehe Beispiel 2) werden mit den Restriktionsenzymen XhoI und SalI für 2 Stunden bei 37°C inkubiert, die DNA-Fragmente durch Agarose-Gelelektrophorese aufgetrennt und sowohl der Vektor als auch das PCR-Insert aus pCR2.1-4 ausgeschnitten und mit dem "Gelpurification"-Kit von Qiagen nach Herstellerangaben aufgereinigt und mit 50 µL Elutionspuffer eluiert. Vom Vektor wird 1 µL, vom PCR-Insert aus pCR2.1-4 8 µL der Eluate für die Ligation eingesetzt, resultierend in dem Konstrukt pCR2.1-sRNAi1. Dieser Vektor wird für 2 Stunden mit dem Restriktionsenzym XhoI und dann für 15 min mit Klenow-Fragment inkubiert. Der Vektor pCR2.1-AtCRB-RNAi (siehe Beispiel 2) wird mit dem Enzym EcoRI für 2 Stunden inkubiert und ebenfalls 15 min mit Klenow-Fragment behandelt. Beide Inkubationsansätze werden durch Gelelektrophorese aufgetrennt und jeweils der Vektor (pCR2.1-sRNAi1) bzw. das Insert (aus pCR2.1-AtCRB-RNAi) aus dem Agarosegel ausgeschnitten und die DNA-Fragmente wie oben beschrieben aufgereinigt. Für die Ligation werden 1 µL des Eluates vom Vektor und 8 µL des Eluates vom Insert eingesetzt und bei 4°C über Nacht inkubiert. Das resultierende Konstrukt wird mit pCR2.1-sRNAi2 bezeichnet. Der resultierende Vektor wird mit dem Enzym XbaI und anschließend mit Klenow-Fragment inkubiert. Der Vektor pCR2.1-4 wird mit den Enzymen EcoRV und XbaI und anschließend mit Klenow-Fragment inkubiert. Nach Gelelektrophorese und -reinigung wird das Fragment aus pCR2.1-4 mit dem Vektor pCR2.1-sRNAi2 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi3. Der resultierende Vektor wird dann mit dem Enzym ApaI für 2 Stunden und dann mit Klenow-Fragment für 15 min inkubiert. Als Insert wird der Vektor pCR2.1-At2S3-RNAi mit dem Enzym EcoRI für 2 Stunden und dann mit Klenow-Fragment für 15 min inkubiert. Nach Gelelektrophorese und -reinigung werden die Eluate ligiert, resultierend in dem Vektor pCR2.1-sRNAi4. Aus diesem Vektor wird dann das sRNAi4-Fragment (SEQ ID NO: 144), kodierend für die super-supprimierende dsRNA, durch Inkubation mit HindIII und

PvuI ausgeschnitten und in den binären Vektor pSUN-USP (SEQ ID NO: 179) ligiert. Das Konstrukt dient der gleichzeitigen Suppression von *Arabidopsis thaliana* Speicherproteinen CRB (SEQ ID NO: 4), CRU3 (SEQ ID NO: 112) und At2S3 (SEQ ID NO: 118).

b) Herstellung von Super-Suppressionskonstrukt 2

Ausgehend von *Arabidopsis thaliana* cDNA wird ein Fragment aus dem Speicherprotein AtCRU3 (SEQ ID NO: 111, 112) mit dem nachfolgenden Oligonukleotid-Primerpaar unter den in Beispiel 2 angegebenen PCR-Bedingungen amplifiziert:

OPN 11: 5'-AAAAGGCCTGTGTTCCATTTGGCCGAAACAAC-3'
(SEQ ID NO: 148)

OPN 12: 5'-AAAGATATCACCTGGAGAAGCCACGAGTG-3'
(SEQ ID NO: 149).

Das erhaltene Fragment wird in den Vektor pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in den pCR2.1-6 und die Sequenzen überprüft.

Ausgehend von *Arabidopsis thaliana* cDNA wird ein Fragment aus dem Speicherprotein At2S3 (SEQ ID NO: 3, 4) mit dem nachfolgenden Oligonukleotid-Primerpaar unter den in Beispiel 2 angegebenen PCR-Bedingungen amplifiziert:

OPN 13: 5'-AAAAGGCCTATGGCTAACAAGCTCTTCCTCGTC-3'
(SEQ ID NO: 150)

OPN 14: 5'-AAAGATATCCTAGTAGTAAGGAGGGAAGAAAG-3'
(SEQ ID NO: 151).

Das erhaltene Fragment wird in den Vektor pCR2.1-TOPO Vektor (Invitrogen) gemäß Herstellerangaben kloniert, resultierend in den pCR2.1-7 und die Sequenzen überprüft.

Aus den pCR2.1-3, pCR2.1-4 (siehe Beispiel 2) und pCR2.1-6 und pCR2.1-7 werden dann die Konstrukte folgendermaßen zusammen ligiert: Der Vektor pCR2.1-3 wird 2 Stunden mit EcoRV inkubiert und anschließend 15 min mit alkalischer Phosphatase dephosphoryliert. Der Vektor pCR2.1-6 wird mit den Enzymen StuI und EcoRV für 2 Stunden inkubiert und das PCR-Insert über Gelelektrophorese und -reinigung isoliert. Vektor pCR2.1-3 und Insert aus pCR2.1-6 werden dann über Nacht bei 4°C ligiert, resultierend in dem Konstrukt

pCR2.1-sRNAi5. Dieser Vektor wird dann mit EcoRV inkubiert und dephosphoryliert und mit dem StuI/ EcoRV inkubierten und gelaufgereinigten Fragment aus pCR2.1-7 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi6. Dieser Vektor wird dann mit XhoI inkubiert und dephosphoryliert. Der Vektor pCR2.1-4 wird mit SalI und XhoI inkubiert und das Insert aus pCR2.1-4 mit dem vorbereiteten Vektor pCR2.1-sRNAi6 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi7. Ausgehend von pCR2.1-sRNAi7 wird eine PCR mit den nachfolgenden Primerpaar unter den in Beispiel 2 gegebenen Bedingungen durchgeführt:

OPN 15: 5' CCGCTCGAGCTCAGGGTCTTTTCTTGCCCACT
(SEQ ID NO: 152)

OPN 16: 5'-CCGGTCGACCTAGTAGTAAGGAGGGAAGAAAG
(SEQ ID NO: 153).

Das resultierende PCR-Produkt wird mit den Enzymen XhoI und SalI inkubiert. Das Fragment wird dann in den Vektor pCR2.1-sRNAi7 (inkubiert mit XhoI) ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi8. Aus diesem Vektor wird dann das sRNAi8-Fragment (SEQ ID NO: 146), kodierend für die super-supprimierende dsRNA, durch Inkubation mit HindIII und XbaI ausgeschnitten und in den binären Vektor pSUN-USP (SEQ ID NO: 179) ligiert. Das Konstrukt dient der gleichzeitigen Suppression von *Arabidopsis thaliana* Speicherproteinen CRB (SEQ ID NO: 4), CRU3 (SEQ ID NO: 112) und At2S3 (SEQ ID NO: 118).

Vergleichsversuch 1:

Die Auswirkung der transgenen Expression der cytosolischer, nicht-lichtregulierten *Arabidopsis thaliana* ACCase (GenBank Acc.-No.: D34630) als Fusionsprotein mit mit Signalsequenz für Transport in Chloroplasten aus der Transketolase aus Tabak (Henkes S et al. (2001) Plant Cell 13(3):535-51) auf den Ölgehalt der Pflanzen wurde untersucht. Dazu wurde das Fusionsprotein (vgl. SEQ ID NO: 176, 177) unter Kontrolle des Napin-Promotors in *Arabidopsis thaliana* Pflanzen exprimiert. Die für dieses Expressionskonstrukt kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 175 beschrieben. Die Herstellung und Untersuchung der Pflanzen erfolgte mit den oben genannten Verfahren. Es konnte keine signifikante Änderung des Ölgehaltes in Samen verglichen mit Wildtypkontrollpflanzen festgestellt werden (siehe Fig. 8).

Patentansprüche

1. Verfahren zum Erhöhen des Gesamtölgehalt in einem pflanzlichen Organismus, dadurch gekennzeichnet, dass nachfolgende Arbeitsschritte umfasst sind
 - a) Verminderung der Proteinemenge eines oder mehrerer Speicherproteine in einem pflanzlichen Organismus oder einem Gewebe, Organ, Teil oder Zelle des besagten pflanzlichen Organismus und
 - b) Auswahl von pflanzlichen Organismen, bei denen - im Unterschied oder Vergleich zur Ausgangsorganismus - der Gesamtölgehalt in dem besagten pflanzlichen Organismus oder einem Gewebe, Organ, Teil oder Zelle des besagten pflanzlichen Organismus erhöht ist.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass der Gesamtölgehalt im Samen einer Pflanze erhöht wird.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass das Speicherprotein ausgewählt ist aus den Speicherprotein-Klassen der 2S-Albumine, 7S-Globuline, 11S/12S-Globuline oder Zein-Prolamine.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass das Speicherprotein ausgewählt ist aus
 - a) der Gruppe der Polypeptid gemäß SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 112, 114, 116, 118, 120, 122, 133, 155, 157, 159, 161, 163, 165, 167, 169, 171, 172 oder 174 oder
 - b) einem funktionellen Äquivalent eines der Polypeptide unter a) mit einer Homologie zu diesem von mindestens 65 %.
5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass das Speicherprotein beschrieben ist durch ein Polypeptid gemäß SEQ ID NO: 112 oder einem funktionellen Äquivalent desselben mit einer Homologie zu diesem von mindestens 65 %.

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass die Proteinmenge von mindestens zwei Speicherproteinen vermindert wird, wobei die Homologie von mindestens zwei der in ihrer Proteinmenge verminderten Speicherproteine untereinander geringer als 90 % ist.
7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass mindestens zwei der zu vermindernenden Speicherproteine zu unterschiedlichen Klassen ausgewählt aus der Gruppe der Klassen der 2S-Albumine, 7S-Globuline, 11S/12S-Globuline oder Zein-Prolamine zählen.
8. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass die Verminderung der Proteinmenge mindestens eines Speicherproteins gewährleistet wird durch Anwendung eines Verfahrens ausgewählt aus der Gruppe bestehend aus
 - a) Einbringung einer doppelsträngigen RNA-Nukleinsäuresequenz oder einer deren Expression gewährleistenden Expressionskassette oder Expressionskassetten, wobei die doppelsträngige RNA-Sequenz nachfolgende Elemente umfasst
 - i) mindestens eine "sense"-Ribonukleotidsequenz, die im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz und
 - ii) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenz unter i) im wesentlichen komplementären sind
 - b) Einbringung einer Speicherprotein antisense-RNA-Nukleinsäuresequenz oder einer deren Expression gewährleistenden Expressionskassette, wobei die Speicherprotein antisense-RNA-Nukleinsäuresequenz im wesentlichen komplementär ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz
 - c) Einbringung einer Speicherprotein antisense-RNA-Nukleinsäuresequenz kombiniert mit einem Ribozym oder einer deren Expression gewährleistenden Expressionskassette,
 - d) Einbringung von Speicherprotein sense-RNA-Nukleinsäuresequenzen zur Induktion einer Kosuppression oder einer deren Expression gewährleistenden Expressionskassette, wobei die Speicherprotein sense-RNA-Nukleinsäuresequenz im wesentlichen identisch ist zu mindestens einem Teil

des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz

- e) Einbringung DNA-bindende Faktoren gegen Speicherprotein-Gene oder -RNAs oder einer deren Expression gewährleistenden Expressionskassette,
- f) Einbringung von den Speicherprotein RNA-Abbau bewirkende virale Nukleinsäuresequenzen und Expressionskonstrukten oder einer deren Expression gewährleistenden Expressionskassette,
- g) Einbringung von Konstrukten zur Induktion einer homologen Rekombination an endogenen Speicherprotein-Genen und
- h) Einführung von Mutationen in ein endogenes Speicherprotein Gen.

9. Verfahren nach einem der Ansprüche 1 bis 8, umfassend

- I) die stabile Transformation einer pflanzlichen Zelle mit einer rekombinanten Expressionskassette enthaltend in funktioneller Verknüpfung mit einem in Pflanzen aktiven Promotor eine Nukleinsäuresequenz kodierend für
 - a) eine doppelsträngigen Speicherprotein RNA-Nukleinsäuresequenz, wobei die doppelsträngige RNA-Sequenz nachfolgende Elemente umfasst
 - i) mindestens eine "sense"-Ribonukleotidsequenz, die im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz und
 - ii) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter i) im wesentlichen komplementären sind,
 - oder
 - b) eine Speicherprotein antisense-RNA-Nukleinsäuresequenz, wobei die Speicherprotein antisense-RNA-Nukleinsäuresequenz im wesentlichen komplementär ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz, oder

- c) eine Speicherprotein antisense-Nukleinsäuresequenz kombiniert mit einem Ribozym oder
- d) eine Speicherprotein sense-Nukleinsäuresequenzen zur Induktion einer Kosuppression, wobei die Speicherprotein sense-RNA-Nukleinsäuresequenz im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz, oder
- f) DNA-bindende Faktoren gegen Speicherprotein-Gene oder -RNAs
- g) den Speicherprotein RNA-Abbau bewirkende virale Nukleinsäuresequenzen

II) Regeneration der Pflanze aus der pflanzlichen Zelle, und

III) Expression besagter Nukleinsäuresequenz in einer Menge und für eine Zeit hinreichend um den Gesamtölgehalt in besagter Pflanze zu erhöhen.

10. Doppelsträngiges RNA-Molekül, umfassend

- a) mindestens eine "sense"-Ribonukleotidsequenz, die im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119, 121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170 oder 173 und
- b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.

11. Doppelsträngiges RNA-Molekül nach Anspruch 10, umfassend

- a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei
 - i) jede dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 111, 113, 115, 117, 119,

121, 132, 154, 156, 158, 160, 162, 164, 166, 168, 170
oder 173 und wobei

- ii) zumindest zwei der Speicherprotein-Nukleinsäuresequenzen, zu deren "sense"-RNA-Transkript die besagten "sense"-Ribonukleotidsequenzen im wesentlichen identisch sind, untereinander eine Homologie von unter 90 % haben, und
 - b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.
12. Doppelsträngiges RNA-Molekül nach Anspruch 10 oder 11, dadurch gekennzeichnet, dass zumindest zwei der Speicherprotein-Nukleinsäuresequenzen, zu deren "sense"-RNA-Transkript die besagten "sense"-Ribonukleotidsequenzen im wesentlichen identisch sind, zu zwei unterschiedlichen Klassen von Speicherproteinen zählen ausgewählt aus der Gruppe von Speicherproteinklassen bestehend aus 2S-Albuminen, 7S-Globulinen, 11S/12S-Globulinen und Zein-Prolaminen.
13. Doppelsträngiges RNA-Molekül nach einem der Ansprüche 10 bis 12, dadurch gekennzeichnet, dass die "sense"-Ribonukleotidsequenzen und die "antisense"-Ribonukleotidsequenzen kovalent miteinander verbunden sind.
14. Transgene Expressionskassette enthalten in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für doppelsträngiges RNA-Molekül gemäß einem der Ansprüche 10 bis 13.
15. Transgenes Expressionssystem enthaltend
- a) in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für "sense"-Ribonukleotidsequenzen eines doppelsträngigen RNA-Molekül gemäß einem der Ansprüche 10 bis 12 und
 - b) in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für "antisense"-Ribonukleotidsequenzen eines doppelsträngigen RNA-Molekül gemäß einem der Ansprüche 10 bis 12,

wobei die Promotoren so gewählt sind, das in einem bestimmten Organismus die gleichzeitige Expression von "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen gewährleistet ist.

16. Transgener Vektor enthaltend eine transgene Expressionskassette gemäß Anspruch 14 oder ein transgenes Expressionssystem gemäß Anspruch 15.
17. Transgener Organismus enthaltend eine transgene Expressionskassette gemäß Anspruch 14 oder ein transgenes Expressionssystem gemäß Anspruch 15 oder einen transgenen Vektor gemäß Anspruch 16.
18. Transgener Organismus nach Anspruch 17 ausgewählt aus der Gruppe bestehend aus Bakterien, Hefen, nicht-menschlichen Tieren und Pflanzen.
19. Transgener Organismus nach einem der Ansprüche 17 oder 18, dadurch gekennzeichnet, dass die Pflanze ausgewählt ist aus der Gruppe bestehend aus Arabidopsis, Raps, Sonnenblume, Sesam, Färberdistel, Ölbaum, Soja, Mais, Weizen und Nussarten.
20. Verwendung eines transgenen Organismus nach einem der Ansprüche 17 bis 19 zur Herstellung von Ölen, Fetten, freien Fettsäuren oder Derivaten der vorgenannten.

```

      *           20           *           40           *           60
J02586 : ACCAAAACTCATCAATACAAACAAGATTAAAAACATACACGAATGGCGAACAAGCTCTTC : 60
X65040 : -----ATGGCGAACAAGCTCTTC : 18
X65038 : -----ATGGCGAACAAGCTCTTC : 18
U04945 : -----AAAAAACATACACACAAATAGCAAATGGCGAACAAGCTCTTC : 42
X91799 : ----- : -

      *           80           *           100          *           120
J02586 : CTCGTCTCGGCAACTCTCGCCTTCTTCTTCCTTCTCACCATGCCTCCATCTACCGGACG : 120
X65040 : CTCGTCTCGGCAACTCTCGCCTTCTTCTTCCTTCTCACCATGCCTCCATCTACCGGACG : 78
X65038 : CTCGTCTCGGCAACTCTCGCCTTCTTCTTCCTTCTCACCATGCCTCCATCTACCGGACG : 78
U04945 : CTCGTCTCGGCAACTCTTGCCTTCTTCTTCCTTCTCACCACGCCTCCATCTACCGCACC : 102
X91799 : ----- : -

      *           140          *           160          *           180
J02586 : GTCGTCTCGAGTTCGACGAAGATGATGCCACAGACTCAGCCGGGCCATTTAGGATTCCAAAA : 180
X65040 : GTCGTCTCGAGTTCGACGAAGATGATGCCACAGACTCAGCCGGGCCATTTAGGATTCCAAAA : 138
X65038 : GTGGTCTCGAGTTCGACGAAGATGATGCCACAAACCCAGCCGGGCCATTTAGGATCCAAAA : 138
U04945 : ATCGTGGAAGTCGACGAAGATGATGCCACAAACCCAGCCGGGCCATTTAGGATTCCAAAA : 162
X91799 : -----CCAGCCGGGCCATTTGGGATTCCAAAA : 27

      *           200          *           220          *           240
J02586 : TGTAGGAAGGAGTTTCAGCAAGCACAACACCTAAGAGCTTGCCAGCAGTGGCTCCACAAG : 240
X65040 : TGTAGGAAGGAGTTTCAGCAAGCACAACACCTAAGAGCTTGCCAGCAGTGGCTCCACAAG : 198
X65038 : TGTAGGAAGGAGTTTCAGCAAGCACAACACCTAAGAGCTTGCCAGCAGTGGCTCCACAAG : 198
U04945 : TGTAGGAAGGAGTTTCAGCAAGCACAACACCTGAAAGCTTGCCAACAATGGCTCCACAAG : 222
X91799 : TGTAGGAAGGAGTTTCAACAAGCACAACACCTAAGAGCTTGCCAGCAATGGCTCCACAAG : 87

      *           260          *           280          *           300
J02586 : CAAGCAATGCAGTCTGGCGGTGGTCTCAGCTGGACCCTCGACGGTGAGTTTGACTTTGAA : 300
X65040 : CAAGCAATGCAGTCTGGCGGTGGTCTCAGCTGGACCCTCGACGGTGAGTTTGACTTTGAA : 258
X65038 : CAAGCAATGCAGTCTGGCGGTGGTCTCAGCTGGACCCTCGACAGTGAGTTTGACTTTGAA : 258
U04945 : CAGGCAATGCAGTCCGGTAGTGGCCCAAGCTGGACCCTCGACGGTGAGTTTGATTTTGAA : 282
X91799 : CAGGCAATGCAGTCTGGTAGTGGTCCAAGCTGGACCCTCGACGATGAGTTTGATTTTGAA : 147

      *           320          *           340          *           360
J02586 : GACGACATGGAGAACCCGCGAGGGTCCACAGCAGAGACCGCCTCTACTCCAGCAGTGCTGT : 360
X65040 : GACGACATGGAGAACTCGCAGGGTCCACAGCAGAGACCGCCTCTACTCCAGCAGTGCTGT : 318
X65038 : GACGACATGGAGAACCCGCGAGGGTCCACAGCAGAGACCGCCTCTACTCCTGCAATGCTGT : 318
U04945 : GATGACATGGAGAACCCCCAGGGCCCAACAGAGGCCGCCACTACTCCAGCAGTGCTGC : 342
X91799 : GATGACATGGAGAACCCACAGGGACCACAGCAGAGGCCACCACTACTCCAGCAGTGCTGC : 207

      *           380          *           400          *           420
J02586 : AACGAGCTCCACCAGGAAGAGCCCCCTTTGCGTTTGCCCAACCTTGAAAGGAGCATCCAAA : 420
X65040 : AACGAGCTCCACCAGGAAGAGCCCCCTTTGCGTTTGCCCAACCTTGAAAGGAGCATCCAAA : 378
X65038 : AACGAGCTGGACCAGGAAGAGCCCCCTTTGCGTTTGCCCAACCTTGAAAGGAGCATCCAAA : 378
U04945 : AACGAGCTCCACCAGGAAGAGCCACTTTGCGTTTGCCCAACCTTGAAAGGAGCATCCAAA : 402
X91799 : AACGAGCTCCACCAGGAAGAGCCACTTTGCGTTTGCCCAACCTTGAAAGGAGCATCCAAA : 267

```

Fig. 1a

```

          *           440           *           460           *           480
J02586 : GCGGTTAAACAACAAATTCAACAACAGGGACAACAGCAAGGAAAGCAG----- : 468
X65040 : GCGGTTAAACAACAAATTCAACAACAGGGACAACAGCAAGGAAAGCAG----- : 426
X65038 : GCGGTTAAACAACAAATTCAACAACAGGGACAACAGCAAGGAAAGCAG----- : 426
U04945 : GCCGTTAAACAACAGGTTTCGACAACAGCAAGGACAGCAGGGACAGCAG-----CTG : 453
X91799 : GCCGTTAAACAGCAGGTTAGACAACAGCTGGAGCAGCAGGGACAGCAGGGACCGCACCTG : 327

```

```

          *           500           *           520           *           540
J02586 : ---CAAATGGTGAGCCGTATCTACCAGACCGCTACGCACTTACCTAAAGTTTGCAACATC : 525
X65040 : ---CAAATGGTGAGCCGTATCTACCAGACCGCTACGCACTTACCTAAAGTTTGCAACATC : 483
X65038 : ---CAAATGGTGAGCCGTATCTACCAGACCGCTACGCACTTACCTAAAGTTTGCAACATC : 483
U04945 : CAGCAAGTAATTAGCCGTATCTACCAGACTGCTACGCACTTACCTAAAGTTTGCAACATC : 513
X91799 : CAGCATGTAATTAGCCGTATCTACCAGACCGCTACGCACTTACCTAGAGTTTGCAACATT : 387

```

```

          *           560           *           580           *           600
J02586 : CCGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCATGCCTGGGCCCTCCTACTAGATTCCA : 585
X65040 : CCGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCATGCCTGGGCCCTCCTACTAG----- : 537
X65038 : CCGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCATGCCTGGGCCCTCCTACTAG----- : 537
U04945 : CCGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCATGCCTGGACCCTCCTACTAGATTCCA : 573
X91799 : AGGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCATGCCTGGACCCTCC----- : 435

```

Fig. 1b

```

          *           20           *           40           *           60
U71194   : ----GGCACGAG--AAATGACCAAGCTTACAATTCTCCTCATCGCTCTTCTCTTC---AT : 51
AF005030 : ACCCATCAATAGCAAAATGACCAAGTTCACAATCCTCCTCATCTCTCTTCTCTTCTGCAT : 60

          *           80           *           100          *           120
U71194   : CGCCACACCTGCTGCGCCTCCAAATGGCAACAGCACCAGCAAGAGAGCTGCCGCGAGCA : 111
AF005030 : CGCCACACTTGCAGCGCCTCCAAATGGCA---GCACCAGCAAGATAGCTGCCGCAAGCA : 117

          *           140          *           160          *           180
U71194   : GCTCAAGGGGATCAACCTCAACCCCTGTGAG---CACATCATGGAGAAGATCCAAGCTGG : 168
AF005030 : GCTCCAGGGGTGAACCTCACGCCCTGCGAGAAGCACATCATGGAGAAGATCCAAGGCCG : 177

          *           200          *           220          *           240
U71194   : CCGCCGCGGCGAGGACGGCAGCGACGAAGATCACATTCTCATCAGGACCATGCCGGGAAG : 228
AF005030 : CGGCGATGACGATGATGATGATGACGACGACAATCACATTCTCAGGACCATGCCGGGAAG : 237

          *           260          *           280          *           300
U71194   : AATCAACTACATCAGGAAGAAGGAAGGAAAAGAAGA-----AGAAGAAGGACACAT : 282
AF005030 : AATCAACTACATAAGGAGGAACGAAGGAAAAGACGAAGACGAAGAAGAAGGACACAT : 297

          *           320          *           340          *           360
U71194   : GCAGAAGTGCTGCAGCGAAATGAGCGAGCTGAAAAGCCCCATATGCCAGTGCAAAGCGCT : 342
AF005030 : GCAGAAGTGCTGCACAGAAATGAGCGAGCTGAGAAGCCCCAAATGCCAGTGCAAAGCGCT : 357

          *           380          *           400          *           420
U71194   : ACAGAAGATAATGGATAACCAGAGCGAGCAACTGGAGGGGAAGGAGAAGAAGCAGATGGA : 402
AF005030 : GCAGAAGATAATGGAGAACCAGAGCGAGGAAGCTGGAGGAGAAGCAGAAGAAGAAAATGGA : 417

          *           440          *           460          *           480
U71194   : GAGAGAGCTCATGAACCTTGGCTATTAGGTGCAGGTTGGGACCCATGATAGGGTGCGACTT : 462
AF005030 : GAAGGAGCTCATTAACCTTGGCTACTATGTGCAGGTTTGGACCCATGATCCAGTGCGACTT : 477

          *           500          *           520          *           540
U71194   : GTCCTCCGATGACTGAAAA---AAAAGTACT-----ACTAACACATATAT : 504
AF005030 : GTCCTCCGATGACTAAGAAGTTAAAAGCAATGTTGTCACTTGTCTGCTACTAACACATG-AT : 536

          *           560          *           580          *           600
U71194   : GTGTTAGTTTATGCTAGCTAGAAGACGTATAAGCTATCTCCG--TATGTTGTATATTAA : 562
AF005030 : GTGATAGTTTATGCTAGCTAGCTATA-ACATAAGCTGTCTCTGAGTGTGTTGTATATTAA : 595

          *           620          *           640          *           660
U71194   : TAAAAAGATCATCACTGGTGAATGGTGATCGTGTATGTA----ACGTAGTGGGCAATGGA : 618
AF005030 : TAA--AGATCATCACTGGTGAATGGTGATCGTGTACGTACCCTACTTAGTAGGCAATGGA : 653

          *           680          *           700          *           720
U71194   : AGCACTTAGAGTGTGCTTTGTG----GCCTTGCCCTCTGTTTGTATACT--GAGACTTT : 672
AF005030 : AGCACTTAGAGTGTGCTTTGTGCATGGCCTTGCC-TCTGTTTGTGAGACTTTTGTAAATGTT : 712

          *           740          *           760          *
U71194   : TGCGAATACCGTTCGTTTTTCCCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA----- : 724
AF005030 : TTCGAGTTTAAATC-TTTGCCTTTGCAAAAAAAAAAAAAAAAAAAAAAAAAA : 770

```

Fig. 2

	*	800	*	820	*	840	
X65039	:	-----ATGGCGAACAAGC				:	13
X91798	:	-----				:	-
X76101	:	-----				:	-
X06410	:	ATAACGCCACCGATTAAACTCACCTCTCCACGTATGAACCTCCACCCACCATATATACGC				:	840

	*	860	*	880	*	900	
X65039	:	TCTTCCTCGTCTCGGCAACTCTCGCCTTCTTCTTCCTGCTCACCAATGCCTCCATCTACC				:	73
X91798	:	-----				:	-
X76101	:	-----GGCAA--AGA				:	8
X06410	:	ACCACCACCACACCATAATTACACAACCACAACACCA-TCTCCCACAATGGCAA--AGC				:	897

	*	920	*	940	*	960	
X65039	:	GGACGGTCGTCGAGTTCGACGAAGATGATGACACAAACCAAGCCGGACCATTTAGGATTTC				:	133
X91798	:	-----CCAGCCGGCCCATTTGGGATTTC				:	22
X76101	:	TAACACTTCTCTTGCTCGCCTTAGCTGCTCTTGTAGCCTTGGCTACAGCCCACACAACCA				:	68
X06410	:	AAATAGTTCTCGCACTCGCTTTCGCCGCCCTTGTAGCCTTTGCTACCGCCCACACAACCA				:	957

	*	980	*	1000	*	1020	
X65039	:	CAAGATGTCGAAAGGAGTTTCGGCAAGCACAAACATCTAAGAGCTT-GCCAGCAATGGCTC				:	192
X91798	:	CAAAATGTAGGAAGGAGTTTCAACAAGCACAAACACCTAAGAGCTT-GCCAGCAATGGCTC				:	81
X76101	:	TCAT-CACCACCACCATCGACGACGAGAACCCGATCTCCGAACAAAGGCAATGTTGGCAA				:	127
X06410	:	TAAT-CACCACCACCATCGAAGACGAGAACCCGATCTCCGGACAAAGGCAAGTGAGCCAA				:	1016

	*	1040	*	1060	*	1080	
X65039	:	CACAGGCAGGCAATGCAGTCCGGTAGTGGTCCAAGCTGGACCCTGGACGGTGAGTTTGAC				:	252
X91798	:	CACAAGCAGGCAATGCAGTCTGGTAGTGGTCCAAGCTGGACCCTCGACGATGAGTTTGAT				:	141
X76101	:	CAGGTACAGGGACAAAGGTTGAACCAAGTGTGCGATGTTCTCCAGCAAGGTCAGAGGGGG				:	187
X06410	:	CGGATACAGGGACAAAGGCTGAACCAAGTGTGCGATGTTCTCCAGCAGGGTCAGAACATT				:	1076

	*	1100	*	1120	*	1140	
X65039	:	TTTGAAGACGACATGGAGAACCAACAGGGCCCACAGCAGAGGCCACCTCTACTCCAGCAA				:	312
X91798	:	TTTGAAGACGACATGGAGAACCCCCAGGGACCACAGCAGAAGCCGCCACTACTCCAGCAA				:	201
X76101	:	CA-----GCAACACCAACAGCAACAGCATCAG-CAGCAGGAGCAGCAGCTCCTCCAGCAG				:	241
X06410	:	CCT---CGCGAATTCGATAACCCCTCAGATGGGGCGGCAGCAGGAGCAGCAGCTCCAGCAG				:	1133

	*	1160	*	1180	*	1200	
X65039	:	TGCTGCAACGAGCTCCACCAGGAAGAGGCACCTTTGTGTTTGCCCAACCTTGAAAGGAGCA				:	372
X91798	:	TGCTGCAACGAGCTTCACCAGGAGGAGCCACTTTGCGTTTGCCCAACCTTGAAAGGAGCT				:	261
X76101	:	TGTTGTCAAGAGCTTCAAAACATCGAAGGACAGTGCCAATGTGAGGCGGTGAAGCAGGTG				:	301
X06410	:	TGTTGTCAAGAGCTCCAAAACATCGAAGGGCAGTGCCAATGTGAGGCGGTGAAGCAGGTG				:	1193

	*	1220	*	1240	*	1260	
X65039	:	TCCAAAGCGGTTAGACAACAGGTTGACAACAG-----G-----				:	406
X91798	:	TCCAAAGCCGTTAAACAACAGGTTGACAACAGTTGGGG-----C-----AGCAG				:	306
X76101	:	GTCCGAGATGCTCAGCGACACGAGCAACAGCGACCGCGA-----GTGCCCTTCCAG				:	352
X06410	:	TTCCGAGAAGCCCAGCAGCAAGTACAACAGCAACAGGGACGGCAGCTTGTACCTTCCGC				:	1253

Fig. 3a

	*	1280	*	1300	*	1320	
X65039	:	G----	ACACCAG---	CAGCAGATGCAGCATGTAATTAGCCGTATCTACCAGACCGCTACG	:	459	
X91798	:	GG---	ACAGCAGGGACCGCAGGTGCAGCATGTAATTAGCCGTATCTACCAGACCGCTACG	:	363		
X76101	:	GGTTCTCAGCAGTCTCAACAGTTGAAGCAGAGGGCT--	CAGATTCTCCCTAACGTATGCA	:	410		
X06410	:	GGTTCGCAGCAGACCCAACAGTTGAAGCAGAAGGCT--	CAGATTCTCCCTAACGTATGCA	:	1311		

	*	1340	*	1360	*	1380	
X65039	:	CACTT--	ACCTAGAGTTTGCAACATCCCGCAAGTTAGCGTTTGTCCCTTCCAGAAGACCA	:	517		
X91798	:	CACTT--	ACCTAAAGTTTGCAACATCCCCAAGTAAGCGTTTGTCCCTTCAAGAAGACCA	:	421		
X76101	:	ACCTTCAATCAAGACGATGCGAAATCGAAAGCGTCAGGAGT----	GTTGCTGAGAGCAA	:	465		
X06410	:	ACCTTCAATCAAGACGATGTGAAATCGGAACCATCACCACCA--	CCGTCACCGAGAGCAA	:	1369		

	*	1400	*	1420	*	1440	
X65039	:	TGCCTGGGCCCCTCCTACTAG-----		:	537		
X91798	:	TGCCTGGACCCCTCC-----		:	435		
X76101	:	TTTTGAAATCCCATT--	TGATATGCCGTTTGATATCCCTTGGC--	CCTTTCGCCCAAGC	:	520	
X06410	:	TATCGATATCCCCTTCCGTGACAGGCCCTTTGGCACTGGATCACAACAGTGCAGAGAAAC		:	1429		

Fig. 3b

```

X06410 : ACCACCACCACACCATAATTACACACAACCACAACACCATCTCCCACAATGGCAAAGCAAA : 900
X76101 : -----GGCAAAGATAA : 11

          *          920          *          940          *          960
X06410 : TAGTTCTCGCACTCGCTTTTCGCCGCCCTTGTAGCCTTTGCTACCGCCCACACAACCATAA : 960
X76101 : CACTTCTCTTGCTCGCCTTAGCTGCTCTTGTAGCCTTGGCTACAGCCCACACAACCATCA : 71

          *          980          *          1000          *          1020
X06410 : TCACCACCACCATCGAAGACGAGAACCCGATCTCCGGACAAAGGCAAGTGAGCCAACGGA : 1020
X76101 : TCACCACCACCATCGACGACGAGAACCCGATCTCCGAACAAAGGCAATGTTGGCAACAGG : 131

          *          1040          *          1060          *          1080
X06410 : TACAGGGACAAAGGCTGAACCAGTGTTCGCATGTTCCCTCCAGCAGGGTCAGAACATTCCCTC : 1080
X76101 : TACAGGGACAAAGGTTGAACCAGTGTTCGCATGTTCCCTCCAGCAAGGTCAGAGGGGGC--A : 189

          *          1100          *          1120          *          1140
X06410 : GCGAATTCGATAACCCCTCAGATGGGGCGGCAGCAGGAGCAGCAGCTCCAGCAGTGTGTC : 1140
X76101 : GCAACACCAACAGCAACAGCATCAG-CAGCAGGAGCAGCAGCTCCTCCAGCAGTGTGTC : 248

          *          1160          *          1180          *          1200
X06410 : AAGAGCTCCAAAACATCGAAGGGCAGTGCCAATGTGAGGCGGTGAAGCAGGTGTTCCGAG : 1200
X76101 : AAGAGCTTCAAAAACATCGAAGGACAGTGCCAATGTGAGGCGGTGAAGCAGGTGGTCCGAG : 308

          *          1220          *          1240          *          1260
X06410 : AAGCCCAGCAGCAAGTACAACAGCAACAGGGACGGCAGCTTGTACCCTTCCGCGGTTCGC : 1260
X76101 : ATGCTCAGCGACACGAGCAACAGCGACCGCGA-----GTGCCCTTCCAGGGTTCCTC : 359

          *          1280          *          1300          *          1320
X06410 : AGCAGACCCAACAGTTGAAGCAGAAGGCTCAGATTCTCCCTAACGTATGCAACCTTCAAT : 1320
X76101 : AGCAGTCTCAACAGTTGAAGCAGAGGGCTCAGATTCTCCCTAACGTATGCAACCTTCAAT : 419

          *          1340          *          1360          *          1380
X06410 : CAAGACGATGTGAAATCGGAACCATCACCACCACCGTCCCGAGAGCAATATCGATATCC : 1380
X76101 : CAAGACGATGCGAAATCGAAAGCGTCAGGAGT---GTTGCTGAGAGCAATTTTGAAATCC : 476

          *          1400          *          1420          *          1440
X06410 : CCTTCCGTGACAGGCCCTTTGGCACTGGATCACAACAGTGCAGAGAACTGAAATCCAAC : 1440
X76101 : CATT---TGATATGCCGTTTGATATCCCTTGGC--CCTTTCGCCCAAGCTCAGAGTCA-C : 530

```

Fig. 4a

```

*      1460      *      1480      *      1500
X06410 : GACCCGTTGGTGAATGCCAAAGGTATAATATACTGTAGGAAATTCATAAACATTTCTTTA : 1500
X76101 : AGC-----AATGC-----AGACAGAGT-----GAAATCCAAAGGCCTGTGAGTC : 569

*      1520      *      1540      *      1560
X06410 : ATATATTTATATAATGAAAACCTACGTATTAAATATTAATGAACATATATGTTCTTTTA : 1560
X76101 : A-GTGCCAAAGGTATGTGGAGCAGCAAATTCAGTCCTCCAGGCC----- : 612

*      1580      *      1600      *      1620
X06410 : TATATAAAATAAAAACCTACTATGTCCTTAGAATAATGAAAACCTATGTATTATTAATC : 1620
X76101 : -ATACCAACAGAGCCCGTACGA---C-CGGAGGCAACAGAGCCCATACG----ACCGGA- : 662

*      1640      *      1660      *      1680
X06410 : AACACATATCTTCTAATGACGATTTCCAACAGGTTTCGTGGAGCAGCAAATGCAGCAGTC : 1680
X76101 : GGCAACAGAGCCCATATA-----TG--AACAGAGGCAAGGACCA-TACGAACAGAGGCC : 711

*      1700      *      1720      *      1740
X06410 : TCCGAGGTCCACTAGACCATACCAACAGCGGCCAGGACAACAGCAGCAGCAGCAGAGAGG : 1740
X76101 : AT--ACGAACAG-AGGCCATACCAACAGCGAGGAGGACGACAGCAGGAGCAGCA---AGG : 765

*      1760      *      1780      *      1800
X06410 : GCTCCAACAACAATGCTGCAACGAGCTACAAAACGTGAAGAGGGAGTGTCAATTGCGAGGC : 1800
X76101 : GCTCCAGCAA---TGCTGCAACGAGCTCCAAAACGTGAGGAGGGAGTGTCACTGCGAGGC : 822

*      1820      *      1840      *      1860
X06410 : AATTCAAGAACTGGCTAGGAGAGTGATGAGGCAGCCACAGCAGCAGCAGCAGCAACGTCG : 1860
X76101 : GATTAAAGAACTGGGCCAAAGAATGA-----GG-----CAGCAGCAACAACAACAG : 870

*      1880      *      1900      *      1920
X06410 : TGGGCAGTTTCGGTGGGCAGGAGATGGAACCGCGAGGAGGGTGATTTCAGAAATCTGCCCAA : 1920
X76101 : TAGGCAGTATGGTGGGCAGCAGACACAACTGTGGAGAGAATTCTTGAGAATCTGCCTAA : 930

*      1940      *      1960      *      1980
X06410 : CCACTGCGACTTGGAGTCCAGCAATGCA-CAACCTGTACGGGATGATGTGAAGTATGAA : 1979
X76101 : CCAATGCGACCTAGATGTCCAGCAATGCAACATCCCCTAC-----TGAA : 974

*      2000      *      2020      *      2040
X06410 : GGAGCTATAAATAAAGGAGTGGCATCTCTAGTCGTTGGA---GGTCTTTGACTTTGAGAT : 2036
X76101 : G-AA-----AAATAAATGAGTTGCATCTCTAGTTGATCGATGAGGTTTATGACTT-GAGAT : 1028

*      2060      *      2080      *      2100
X06410 : GCAGGGAGGGTAGTTTCTCCTTTCTAATAATAATAAGAATAAATGACTTGATGAGCCA : 2096
X76101 : GCATGGAGGGTAGTTTCTACTTTCTGATGATAATAATAAGAATAATGA-----G : 1077

*      2120      *      2140      *      2160
X06410 : CAAATGTGTACTTATAGCTCATGTGTA----ATCGTTGCTTTTCATGTTATTTTTGATATG : 2152
X76101 : CACATGTGTATCTATAGCTCATGTGTACGTAATCGTGGCTTTTCATGTTATCCTTTT-TTCC : 1136

*      2180      *      2200      *      2220
X06410 : TTGAATGGCACAAATGAGTGATTTAATGCCCATGAAAAATCTCAAACCTAGGTTCTTAAA : 2212
X76101 : TTGCATC---CATGCCCCGAAATCTAGTATCTAGA----- : 1167

```

Fig. 4b

```

                *           20           *           40           *           60           *
At2S2 : ATGGCAAACAAGCTCTTCCTCGTCTGCGCAACTTTGCGCCTCTGCTTCCTCCTCACCAACGCTTCCATCT : 70
At2S3 : ATGGCTAACAAGCTCTTCCTCGTCTGCGCAACTCTCGCCCTCTGCTTCCTCCTCACCAACGCTTCCATCT : 70
At2S4 : ATGGCGAACAAGCTCTTCCTCGTCTGCGCAGCTCTCGCCCTGTGTTTCATCCTCACCAACGCTTCCGTCT : 70
At2S1 : ATGGCAAACAAGTTGTTCTCGTCTGCGCAGCTCTCGCTCTCTGCTTCCTCCTCACCAACGCTTCCATCT : 70

                80           *           100           *           120           *           140
At2S2 : ACCGCACTGTTGTCGAGTTCGACGAAGATGACGCCAGCAACCCCATGGGCCCAAGAC---AGAAATGTCA : 137
At2S3 : ACCGCACCGTTGTCGAATTCGAAGAAGATGACGCCAGCAACCCCGTAGGTCCAAGAC---AGAGATGCCA : 137
At2S4 : ATCGCACCGTTGTCGAGTTCGACGAAGATGACGCCAGTAACCCCATAGGCCCAATAC---AGAAATGTCA : 137
At2S1 : ACCGCACCGTCGTTGAGTTCGAAGAAGATGACGCCACTAACCCCATAGGCCCAAAAATGAGGAAATGCCG : 140

                *           160           *           180           *           200           *
At2S2 : GAAGGAGTTTTCAGCAATCACAGCACCTAAGAGCTTGCCAGAAATTGATGCGCATGCAAATGAGGCAAGGC : 207
At2S3 : GAAGGAGTTTTCAGCAATCACAACACCTAAGAGCTTGCCAGAGATGGATGAGCAAGCAAATGAGGCAAGGA : 207
At2S4 : GAAGGAGTTTTCAGCAAGACCAGCACCTAAGAGCTTGCCAGAGATGGATGCGCAAGCAAATGTGGCAAGGA : 207
At2S1 : CAAGGAGTTTTCAGAAAGAACAACACCTAAGAGCTTGCCAGCAATTGATGCTCCAGCAAGCAAGGCAAGGC : 210

                220           *           240           *           260           *           280
At2S2 : CGTGGTGGTGGTCCCTCCCTCGACGATGAGTTTCGATTTGGAAGACGACATCGAGAACCCACAAGGCCCCC : 277
At2S3 : CGTGGTGGTGGTTCCTTCCCTCGACGATGAGTTTCGATTTTCGAGG-----GCCCCC : 256
At2S4 : CGTGGTGGTGGTTCCTTCCCTCGACGATGAGTTTCGATATGGAAGACGACATCG-----AGAACC : 265
At2S1 : CGTAGCGATGAGTTTGATTTTGAAGACGACATGGAGAACCCACAGG-----GACAAC : 262

                *           300           *           320           *           340           *
At2S2 : AGCAGGGACACCAGATCCTCCAGCAGTGCTGCAGCGAGCTTCGCCAGGAAGAGCCAGTTTGTGTTTGCCC : 347
At2S3 : AGCAGGGATACCAGCTACTCCAGCAGTGCTGCAACGAGCTTCGCCAGGAAGAGCCAGTTTGCCTTTGCCC : 326
At2S4 : CGCAGAGACGACAGCTACTCCAGAAGTGCTGCAGCGAGCTTCGCCAAGAAGAGCCAGTTTGCCTTTGCCC : 335
At2S1 : AGCAGGAACAACAGCTATTCCAGCAGTGCTGCAACGAGCTTCGCCAGGAAGAGCCAGATTGTGTTTGCCC : 332

                360           *           380           *           400           *           420
At2S2 : CACCTTGAGACAAGCTGCCAGGGCCGTTAGCCTCCAGGGACAACA---CGGACCATTTCCAATCCAGGAAA : 414
At2S3 : CACCTTGAAACAAGCTGCCAGGGCAGTTAGCCTCCAGGGACAGCA---CGGACCATTTCCAATCCAGGAAA : 393
At2S4 : CACCTTGAGACAAGCTGCCAAGGCCGTTAGATTCCAGGGACAGCAACACCAACCAGAGCAAGTCAGGAAA : 405
At2S1 : CACCTTGAAACAAGCTGCCAAGGCCGTTAGACTCCAGGGACAGCA---CCAACCAATGCAAGTCAGGAAA : 395

                *           440           *           460           *           480           *
At2S2 : ATTTACAAGACAGCTAAGTACTTGCCTAACATTTGCAAGATCCAGCAAGTTGGTGAATGCCCCTTCCAGA : 484
At2S3 : ATTTACCAGTCAGCTAAGTACTTGCCTAACATTTGCAAGATCCAGCAAGTTGGTGAATGTCCCTTCCAGA : 463
At2S4 : ATTTACCAGGCAGCTAAGTACTTGCCTAACATTTGCAAAATCCAGCAAGTTGGTGTGTTGCCCCTTCCAGA : 475
At2S1 : ATTTACCAGACAGCCAAGCACTTGCCCAACGTTTGCGACATCCCGCAAGTTGATGTTTGTCCCTTCAACA : 465

                500           *           520
At2S2 : CCACCATCCCTTTCTTCCCTCCTTACTAA--- : 513
At2S3 : CCACCATCCCTTTCTTCCCTCCTTACTACTAG : 495
At2S4 : -----TCCCTTCAATCCCTTCTTACTACTAA : 501
At2S1 : -----TCCCTTCATTCCTTCTTTCTACTAA : 495

```

Fig. 5

	*	20	*	40	*	60	
CRU1_Bn	:	-----	ATGGGCTCGGCTCTCATCTCTTCTCTCTTTTTCCTTAGCACTTT	:		43	
CRUA_Bn	:	-----	ATGGGCTCGGCTCTCATCTCTTCTCTCTTTTTCCTTAGCACTTT	:		43	
CRU2_Bn	:	-----	ATGGGCTCGACTCTCGTCTCTTCTCTATTTTTCGATAACAGTTT	:		43	
CRU4_Bn	:	-----		:		-	
CRU3_Bn	:	ATGGTTAAAGTTTCCTCATCTCCTCGTCGCAACGTTTCGGGGTTCTCCTCGTCTCAACGGCTGTCT	:			65	
	*	80	*	100	*	120	*
CRU1_Bn	:	TGATCTTTTCTCCATGGCTCTACAGCTCAACAGTTTCCAAACGAGTGT	CAGCTAGACCAGCTCAAT	:		108	
CRUA_Bn	:	TGACTTTTCTCCATGGCTCTACAGCTCAACAGTTTCCAAACGAGTGT	CAGCTAGACCAGCTCAAT	:		108	
CRU2_Bn	:	TGATCTTTTCTCCATGGCTCTACAGCTCAACAGTTTCCAAACGAGTGCCA	ACTAGACCAGCTCAAT	:		108	
CRU4_Bn	:	-----		:		-	
CRU3_Bn	:	CGCAAGGCAGTCGCTAGGG-GTTCCTCCTCAGCTAGGGAACGCGTGT	AACCTCGATAACTTAGAC	:		129	
		140	*	160	*	180	*
CRU1_Bn	:	GCACTGGAGCCGTCACACGTACTTAAGGCTGAGGCTGGTTCGCATCGAGGTGT	GGGACCACCACGC	:		173	
CRUA_Bn	:	GCACTGGAGCCGTCACACGTACTTAAGGCTGAGGCTGGTTCGCATCGAGGTGT	GGGACCACCACGC	:		173	
CRU2_Bn	:	GCGCTGGAGCCGTCACACGTACTTAAGGCCGAGGCTGGTTCGCATCGAAGTGT	GGGACCACCACGC	:		173	
CRU4_Bn	:	-----		:		-	
CRU3_Bn	:	GTTCTCCAGCCTACCGAACTATCAAGAGCGAGGCTGGTTCGGGTCGAGTACT	GGGATCACAACAA	:		194	
		200	*	220	*	240	*
CRU1_Bn	:	TCCTCAGCTACGTTGCTCTGGTGTCTCCTTTGTACGTTACATCATCGAGTCT	AAGGGTCTCTACT	:		238	
CRUA_Bn	:	TCCTCAGCTACGTTGCTCTGGTGTCTCCTTTGTACGTTACATCATCGAGTCT	AAGGGTCTCTACT	:		238	
CRU2_Bn	:	TCCTCAGCTACGCTGCTCTGGTGTCTCCTTCGTACGTTACATAATCGAGTCT	CAGGGTCTATACT	:		238	
CRU4_Bn	:	-----		:		-	
CRU3_Bn	:	TCCTCAGATCCGATGTGCTGGTGTCTCTGTCTCTCGTGTATAATCGAACAAGGCGGTCTCTACC	:			259	
	*	280	*	300	*	320	
CRU1_Bn	:	TGCCCTCTTTCTTTAGCACCGCGAAGCTCTCCTTCGTGGCTAAAGGAGAAGGTCT	TATGGGGAGA	:		303	
CRUA_Bn	:	TGCCCTCTTTCTTTAGCACCGCGAGGCTCTCCTTCGTGGCTAAAGGAGAAGGTCT	TATGGGGAGA	:		303	
CRU2_Bn	:	TGCCCTCTTTCTTTAAATACCGCGAAGCTCTCTTTTCGTGGCTAAAGGACAAGGCCT	TATGGGGAGA	:		303	
CRU4_Bn	:	-----		:		-	
CRU3_Bn	:	TTCCTACCTTCTTCAGCTCCCCAAAATTCATACGTTGTTCAAGGAATGGGTATTAGCGGAAGA	:			324	
	*	340	*	360	*	380	*
CRU1_Bn	:	GTGGTCCCTGGATGCGCCGAGACATTCCAGGACTCATCAGTGTTC	CAACCAAGCGGTGGTAGCCC	:		368	
CRUA_Bn	:	GTGGTCCCT---GTGCGCCGAGACATTCCAGGACTCATCAGTGTTC	CAACCAAGCGGTGGTAGCCC	:		365	
CRU2_Bn	:	GTGGTCCCTGGATGCGCTGAGACTTTCCAGGACTCATCAGTATTC	CAACCAAGGCAGTGGCAGCCC	:		368	
CRU4_Bn	:	-----		:		-	
CRU3_Bn	:	GTGGTCCCTGGATGCGCGGAAACCTTCATGGACTCGCAGCCTATG	CAAGGACAACAACAAGGTCA	:		389	
		400	*	420	*	440	*
CRU1_Bn	:	CTCGGGAGAAGGTCAGGGCCAAGGACAACAAGGTCAAGGGCCAAGGCCA---	CCAAGGTCAAGGCC	:		430	
CRUA_Bn	:	CTTCGGAGAAGGTCAGGGCCAAGGACAACAAGGTCAAGGGCCAAGGCCA---	CCAAGGTCAAGGCC	:		427	
CRU2_Bn	:	CTTCGGAGAAGGTCAAGGGCCAAGGTCAAG					

Fig. 6a

	460	*	480	*	500	*	520	
CRU1_Bn	: AAGGACAACAGGG-----CCAACAAGGTCAG---CAAGGACAACAGAGTCAA-GGC--							: 477
CRUA_Bn	: AAGGACAACAGGG-----CCAACAAGGTCAG---CAAGGACAACAGAGTCAA-GGC--							: 474
CRU2_Bn	: AGGGCCAACAGGGTCAAGGCAAGGGCCAACAGGGTCAATCCCAGGGCCAACAGGGTCAA-GGT--							: 495
CRU4_Bn	: AGG----ACAGGGCCA-----GGAACAAGGTCAGGGACAGGGTCA---GGGTCAA-GGC--							: 105
CRU3_Bn	: AGGGACAACAAGGTCAAC---AGGGTCAGCAGGGTCAA---CAGGGACAGCAGGGTCAGCAGCAG							: 510
		*	540	*	560	*	580	
CRU1_Bn	: CAGGGCTTCCGTGATATGCACCAGAAAGTGGAGCACATAAGGACTGGGGACACCATCGCTACACA							: 542
CRUA_Bn	: CAGGGTTTCCGTGATATGCACCAGAAAGTGGAGCACATAAGGACTGGGGACACCATCGCTACACA							: 539
CRU2_Bn	: CAAGGTTTCCGTGATATGCACCAGAAAGTAGAGCACATAAGGAGCGGCGACACCATTGCTACACA							: 560
CRU4_Bn	: CAGGGTTTCCGTGACATGCACCAGAAAGTAGAGCACCTACGGTCGGGTGACACCATAGCGACACC							: 170
CRU3_Bn	: CAAGGTTTCCGTGACATGCACCAGAAGTCTGAACATGTTTCGACATGGAGACATCATTGCCATTAC							: 575
		*	600	*	620	*	640	*
CRU1_Bn	: TCCCGGTGTAGCCCAATGGTTCTACAACGACGGAAACCAACCACTTGTTCATCGTTTCCGTCTCTCG							: 607
CRUA_Bn	: TCCCGGTGTAGCCCAATGGTTCTACAACGACGGAAACCAACCACTTGTTCATCGTTTCCGTCTCTCG							: 604
CRU2_Bn	: TCCCGGTGTAGCTCAATGGTTCTACAACAATGGAAACCAACCTCTTGTTCATCGTTGCCGTCTATGG							: 625
CRU4_Bn	: ACCTGGTGTAGCTCAATGGTTCTACAACAATGGAAATGAGCCTCTCATTTCTTGTGTCAGCCGCTG							: 235
CRU3_Bn	: TGCAGGCTCTTCCATTGGATCTACAACACCGGTGACCAGCCACTTGTTCATTATCTGCCTTCTCTCG							: 640
		660	*	680	*	700	*	
CRU1_Bn	: ATTTAGCCAGCCACCAGAATCAGCTCGACCGCAACCCAAGGCCATTTTACTTAGCCGGAACAAC							: 672
CRUA_Bn	: ATTTAGCCAGCCACCAGAATCAGCTCGACCGCAACCCAAGGCCATTTTACTTAGCCGGAACAAC							: 669
CRU2_Bn	: ATTTAGCTAGCCACCAGAACCAGCTTGACCGCAACCCAAGCCAATTTTACTTAGCAGGAAAAAAC							: 690
CRU4_Bn	: ACATCGCAAATAACCTGAACCAGCTTGACCGCAACCTCAGACCGTTTGTGTTAGCCGGAACAAC							: 300
CRU3_Bn	: ACATTGCCAACTACCAAAACCAACTCGACCGCAACCCAAGAACGTTCCGTCTGGCCGGAACAAC							: 705
		720	*	740	*	760	*	780
CRU1_Bn	: CCACAAGGCCAAGTATGGATAGAAGGACGCGAGCAACAGCCACAAAAAGAACATCCTTAATGGCTT							: 737
CRUA_Bn	: CCACAAGGCCAAGTATGGATAGAAGGACGCGAGCAACAGCCACAAAAAGAACATCCTTAATGGCTT							: 734
CRU2_Bn	: CCACAAGGCCAATCATGGCTACACGGACGAGGGCAACAGCCACAAAAACAACATCCTTAATGGCTT							: 755
CRU4_Bn	: CCACAAGGACAGCAATGGCTACAAGGCCGACAGCAACAGAAAAACAACAACATCTTCAATGGCTT							: 365
CRU3_Bn	: CCACAGGGC-----GGTTCCTCAG--CAGCAGCAGCAACAACAACAGAACATGTTGAGCGGGTT							: 761
		*	800	*	820	*	840	
CRU1_Bn	: CACACCAGAGGTTCTTGCTAAAGCTTTCAAGATCGATGTTAGGACAGCGCAACAACCTTCAGAACC							: 802
CRUA_Bn	: CACACCAGAGGTTCTTGCTAAAGCTTTCAAGATCGATGTTAGGACAGCGCAACAACCTTCAGAACC							: 799
CRU2_Bn	: CTCTCCAGAGGTTCTTGCTCAAGCGTTCAAGATCGATGTTAGGACAGCGCAACAACCTTCAGAACC							: 820
CRU4_Bn	: CGCACCTCAGATCTTGGCTCAAGCCTTTAAATCAGCGTCGAGACAGCTCAAAAACCTTCAAAACC							: 430
CRU3_Bn	: CGACCCTCAGGTCCTAGCCCAGGCATTGAAAAATCGACGTTAGGTTGGCTCAGGAGCTTCAGAACC							: 826
		*	860	*	880	*	900	*
CRU1_Bn	: AGCAAGACAACCGTGGAAACATTATCCGAGTCCAAGGCCCATTCAGTGTCTATTAGGCCGCCTTTG							: 867
CRUA_Bn	: AGCAAGACAACCGTGGAAACATTATCCGAGTCCAAGGCCCATTCAGTGTCTATTAGGCCGCCTTTG							: 864
CRU2_Bn	: AGCAAGATAACCGGGGAAACATTGTCCGTGTCCAAGGCCCTTCCGGTGTATTAGGCCGCCATTG							: 885
CRU4_Bn	: AGCAAGTCAACCGTGGCAACATCGTCAAGGTCCAAGGTCAATTCCGGCGTCATTAGACCACCTTG							: 495
CRU3_Bn	: AACAAGACAGCAGAGGAAACATCGTTCTGTGTTAAGGGACCTTTCCAGGTTGTGAGGCCGCCTCTT							: 891
		920	*	940	*	960	*	
CRU1_Bn	: AGG-----AGTCAGAG-----ACCGCAGGAGACAGAAGTTAA							: 899
CRUA_Bn	: AGG-----AGTCAGAG-----ACCGCAGGAG---GAAGTTAA							: 893
CRU2_Bn	: AAA-----AGCCAGAG-----ACCACAGGAGACAGAAGCTAA							: 917
CRU4_Bn	: AGACAAGGCCAAGGCGGTCAGCA-----ACCACAGGAA---GAAGTTAA							: 536
CRU3_Bn	: AGACAGCCATACGAGAGTGAGCAGTGGAGACACCCCCGTGGCCACCACAAAGCCCACAAGACAA							: 956
		980	*	1000	*	1020	*	1040
CRU1_Bn	: CGGTTTGAAGAGACCATATGCAGCGGAGGTGCACCGATAACCTCGATGACCCATCTAATGCTG							: 964
CRUA_Bn	: CGGTTTGAAGAGACCATATGCAGCGGAGGTGCACCGATAACCTCGATGACCCATCTAATGCTG							: 958
CRU2_Bn	: CGGTCTGAAGAGACCATATGCAGCGCAAGGTGCACCGATAACCTCGATGACCCATCTAAGCGCG							: 982
CRU4_Bn	: TGGTTTGGAGGAGACTTTGTGCACAATGCGATGCACTGAAAACCTTGATGACCCGTCAAGTGCTG							: 601
CRU3_Bn	: CGGCTTGGAGGAGACTATCTGCAGCATGAGGACCCACGAGAACATTGATGACCCAGCCCGTGCTG							: 1021

Fig. 6b

	*	1060	*	1080	*	1100	
CRU1_Bn	:	ACGTATACAAGCCACAGCTCGGTTACATCAGCACTCTGAACAGCTATGATCTCCCCATCCTTCGC	:	1029			
CRUA_Bn	:	ACGTATACAAGCCACAGCTCGGTTACATCAGCACTCTGAACAGCTATGATCTCCCCATCCTTCGC	:	1023			
CRU2_Bn	:	ATGTGTATAAGCCACAGCTTGGTTACATCAGCACTCTTAACAGTTATGATCTACCCATCCTTCGC	:	1047			
CRU4_Bn	:	ATGTCTACAAGCCATCGCTCGGATACATTAGCACACTCAACAGCTACAATCTCCCTATCCTCAGA	:	666			
CRU3_Bn	:	ACGTGTATAAGCCCAACCTCGGCCGTGTGACTAGCGTCAACAGCTACACTTTACCCATCTTGCAG	:	1086			
	*	1120	*	1140	*	1160	*
CRU1_Bn	:	TTCCTTCGTCTCTCAGCCCTCCGTGGATCTATCCGTCAAAACGCGATGGTGCTTCCACAGTGGAA	:	1094			
CRUA_Bn	:	TTCCTTCGTCTCTCAGCCCTCCGTGGATCTATCCGTCAAAACGCGATGGTGCTTCCACAGTGGAA	:	1088			
CRU2_Bn	:	GTACTTCGCCCTCTCAGCCCTCCGTGGATCAATCCGTCAAAATGCAATGGTTCTTCCACAGTGGAA	:	1112			
CRU4_Bn	:	TTCCTCCGCCCTTAGCGCTCTTCGTGGCTCCATCCATAACAACGCTATGGTGCTACCACAATGGAA	:	731			
CRU3_Bn	:	TATATCAGACTCAGCGCCACCCGTGGCATTTCTCCAGGGTAATGCGATGGTGCTTCCGAAATACAA	:	1151			
		1180	*	1200	*	1220	*
CRU1_Bn	:	CGCAAACGCAAACGCGGTTCTCTACGTGACAGACGGGGAAGCCCATGTGCAGGTGGTTAACGACA	:	1159			
CRUA_Bn	:	CGCAAACGCAAACGCGGTTCTCTACGTGACAGACGGGGAAGCCCATGTGCAGGTGGTTAACGACA	:	1153			
CRU2_Bn	:	GTCAAAGTCAAACGCGGTTCTCTACGTGACAGACGGGGAAGCCCAAATACAGGTGGTTAACGACA	:	1177			
CRU4_Bn	:	CGTGAACGCAAACGCGGCGCTCTACGTGACAAAGGGGAAGGCTCATATACAGATGGTGAACGACA	:	796			
CRU3_Bn	:	CATGAACGCGAACGAGATCTTGTACTGCACTCAAGGACAAGCAAGGATTCAAGTGGTGAACGACA	:	1216			
		1240	*	1260	*	1280	*
CRU1_Bn	:	ACGGTGACAGAGTGTTTCGACGGACAAGTCTCTCAAGGACAGCTACTTTCCATACCACAAGGTTTC	:	1224			
CRUA_Bn	:	ACGGTGACAGAGTGTTTCGACGGACAAGTCTCTCAAGGACAGCTACTTTCCATACCACAAGGTTTC	:	1218			
CRU2_Bn	:	ACGGTGACAGAGTGTTTCGATGGACAAGTCTCTCAAGGGCAGCTACTTTCCATTCACAAGGATTC	:	1242			
CRU4_Bn	:	ACGGACAAAGAGTGTTTCGACCAAGAGATCTCCAAGGGACAGTTACTTGTCTGCGCCACAAGGCTTC	:	861			
CRU3_Bn	:	ACGGACAGAACGTGCTGGACCAGCAGGTGCAGAAGGGACAGCTCGTGGTCATCCCACAAGGATTC	:	1281			
	*	1320	*	1340	*	1360	
CRU1_Bn	:	TCCGTGGTGAAACGCGCAACAAGCGAACAGTTCCGG-TGGATCGAGTTCAAGACAAACGCAAACG	:	1288			
CRUA_Bn	:	TCCGTGGTGAAACGCGCAACAAGCGAACAGTTCCGG-TGGATCGAGTTCAAGACAAACGCAAACG	:	1282			
CRU2_Bn	:	TCCGTGGTGAAACGCGCAACAAGCGATCAGTTTCAGG-TGGATAGAATTCAAGACAAACGCAAACG	:	1306			
CRU4_Bn	:	GCGGTCGTGAAACGTGCCACAAGCCAACAGTTTCAG-TGGATCGAGTTCAAGAGCAACGACAACG	:	925			
CRU3_Bn	:	GCCTATGTTGTCCA-GTCCACCAAAAACAACCTTCGAATGGATTTCTTTCAAGACAAACGCTAACG	:	1345			
	*	1380	*	1400	*	1420	*
CRU1_Bn	:	CACAGATCAACACACTTGTCTGGACGAACCTCGGTCTTGAGAGGTTTACCATTAGAGGTCATATCC	:	1353			
CRUA_Bn	:	CACAGATCAACACACTTGTCTGGACGAACCTCGGTCTTGAGAGGTTTACCATTAGAGGTCATATCC	:	1347			
CRU2_Bn	:	CCCAGATCAACACTCTTGTCTGGACGTACCTCAGTCATGAGAGGTTTACCATTAGAGGTCATAGCC	:	1371			
CRU4_Bn	:	CACAGATCAACACACTCGCGGGACGCACCTCAGTCATGAGAGGTTTACCATTGAAGTTATATCC	:	990			
CRU3_Bn	:	CGATGGTCAGCACTTTGGCCGGTAGAACCTCGGCCTTGAGGGCATTGCCACTAGAGGTCATAACC	:	1410			
		1440	*	1460	*	1480	*
CRU1_Bn	:	AATGGGTACCAAATCTCACTCGAAGAAGCAAGAAGGGTTAAGTTCAACACGATCGAGACCACTTT	:	1418			
CRUA_Bn	:	AATGGGTACCAAATCTCACTCGAAGAAGCAAGAAGGGTTAAGTTCAACACGATCGAGACCACTTT	:	1412			
CRU2_Bn	:	AATGGGTACCAAATCTCACTTGAAGAAGCAAGAAGGGTTAAGTTCAACACAATAGAGACCACTTT	:	1436			
CRU4_Bn	:	AACGGGTATCAGATCTCACCCCAAGAAGCTAGAAGTGTTAAGTTCAAGCACTCTTGAGACCACATT	:	1055			
CRU3_Bn	:	AACGCTTTCCAAATTTCTCTCGAGGAAGCTAGAAGGATCAAGTTCAACACGCTTGAGACCACTTT	:	1475			
		1500	*	1520	*	1540	*
CRU1_Bn	:	GACGCACAGCAGTGGCCAGCTAGCTACGGAGGGCCAAGGAAG--GCTGATGCTTAA	:	1473			
CRUA_Bn	:	GACGCACAGCAGTGGCCAGCTAGCTACGGAGGGCCAAGGAAG--GCTGATGCTTAA	:	1467			
CRU2_Bn	:	GACCCACAGTAGTGGCCAGCGAGCTACGGAAGGCCAAGGAAG--GCTGATGCTTGA	:	1491			
CRU4_Bn	:	GACTCAAAGCAGTGGTCTTATGGGCTACGGTATGCC TAGG-----GTCGAGGCTTGA	:	1107			
CRU3_Bn	:	GACTCGTGCGCGCGGT--GGACAACCCAGTTGATCGAGGAGATAGTCGAGGCTTAA	:	1530			

Fig. 6c

	*	20	*	40	*	60	
GLC2	:	ATGGCCAAGCTTGTTC	TTTCCCTTTGT	-TTCCTTCTTTT	CAGTGGCTGCT	-----TCGCT	: 54
GLC3	:	ATGGCTAAGCTTGTTC	TTTCCCTTTGT	-TTTCTGCTTTT	CAGTGGCTGCTGCT	-TCGCTTTT	: 63
GLC1	:	ATGGCCAAGCTAGT	TTTTCCTTTGT	-TTTCTGCTTTT	CAGTGGCTGCTGCT	-TCGCTTTT	: 63
GLC4	:	ATGGGGAAGCCCTTC	---ACTCTCTCTCT	TTTCTTCCCTTT	GCTTGCTACTCTT	TGTCGAGTGCATG	: 62
GLC5	:	ATGGGGAAGCCCTTC	TTCACTCTCTCTCT	TTTCTTCCCTTT	GCTTGCTACTCTT	TGTCGAGTGCATG	: 65

	*	80	*	100	*	120	*	
GLC2	:	CTGAGAGAGCAGGC	CACAGCAA----	AATGAGTGCCAGAT	CCAAAAGCTGAAT	GCCCTCAAACCGG	:	115
GLC3	:	TTCAGAGAGCAGCC	CACAGCAA----	AACGAGTGCCAGAT	CCAACGCCTCAAT	GCCCTAAAACCGG	:	124
GLC1	:	TCCAGAGAGCAGCC	TACAGCAA----	AACGAGTGCCAGAT	CCAAAACTCAAT	GCCCTCAAACCGG	:	124
GLC4	:	CTTTGCTATTAGCT	CCAGCAAGCTCA	ACGAGTGCCAACT	CAACAACCTCA	ACGC GTTGAAACCGG	:	127
GLC5	:	CTTTGCTATTACCT	CCAGCAAGTTCA	ACGAGTGCCAACT	CAACAACCTCA	ACGC GTTGAAACCGG	:	130

	140	*	160	*	180	*	
GLC2	:	ATAACCGTATAGAGT	CGGAAGGTGGGTT	TCATTGAGACAT	TGGAACCCTAACA	ACAAGCCATTCCAG	: 180
GLC3	:	ATAACCGTATAGAGT	CAGAAGGTGGCTT	CATTGAGACAT	TGGAACCCTAACA	ACAAGCCATTCCAG	: 189
GLC1	:	ATAACCGTATAGAGT	CAGAAGGAGGGCT	CATTGAGACAT	TGGAACCCTAACA	ACAAGCCATTCCAG	: 189
GLC4	:	ACCACCGCGTTGAGT	CCGAAGGTGGTTT	TGATTCAAACAT	TGGAACCTCTCA	ACACCCGTGAGCTGAAA	: 192
GLC5	:	ACCACCGCGTTGAGT	CCGAAGGTGGTCT	TATTGAAACAT	TGGAACCTCTCA	ACACCCGTGAGCTGCAA	: 195

	200	*	220	*	240	*	260	
GLC2	:	TGTGCCGGTGTTGCC	CTCTCTCGCTGC	ACCCTTAACCGCA	ATGCCCTTCGTAG	ACCTTCCTACAC	: 245	
GLC3	:	TGTGCCGGTGTTGCC	CTCTCTCGCTGC	ACCCTCAACCGCA	ACGCCCTTCGCAG	ACCTTCCTACAC	: 254	
GLC1	:	TGTGCCGGTGTTGCC	CTCTCTCGCTGC	ACCCTCAACCGCA	ACGCCCTTCGTAG	ACCTTCCTACAC	: 254	
GLC4	:	TGCGCCGGTGTCAC	TGTTCCAAACTC	ACCCTCAACCGCA	ATGGCCTCCACT	CGCCATCTTACTC	: 257	
GLC5	:	TGCGCCGGTGTCAC	TGTTCCAAACGC	ACCCTCAACCGCA	ACGGCTCCCACT	TGCCATCTTACTT	: 260	

	*	280	*	300	*	320	
GLC2	:	CAACGGTCCCCAGGA	AATCTACATACA	ACAAGGTAATGGT	ATTTTTTGGCAT	GATATTCCCGGGTT	: 310
GLC3	:	CAACGTCCCCAGGAG	ATCTACATCCA	ACAAGGTAGTGGT	ATTTTTTGGCAT	GATATTCCCGGGTT	: 319
GLC1	:	CAACGGTCCCCAGGA	AATCTACATCCA	ACAAGGTAAGGGT	ATTTTTTGGCAT	GATATACCCGGGTT	: 319
GLC4	:	ACCTTATCCCCGGAT	GATCATCATCG	CCCAAGGGAAAGG	AGCACTTGGAGTT	TGCAATTCAGGAT	: 322
GLC5	:	ACCTTATCCCCAAAT	GATCATTTGTC	GTTCAGGGAAGGG	AGCAATTGGATT	TGCATTTCCGGGAT	: 325

	*	340	*	360	*	380	*	
GLC2	:	GTCCTAGCACTTAT	CAAGAGCCGCA	AGAATC-TCAGCA	ACGAGGA---CGA	AGCCAGAG-----	: 365	
GLC3	:	GTCCTAGCACATTT	GAAAGAGCC-----	TCAACAAAAGGA	---CAAAGCAGCAG	-----	: 365	
GLC1	:	GTCCTAGCACATTT	GAAAGAGCCTCA	ACAACC-TCAACA	AAAGAGGA---CAA	AGCAGCAG-----	: 374	
GLC4	:	GTCCTGAGACGTTT	GAGGAGCCACA	AGAACAATCAA	ACAGAAGAGGCT	CAAGGTCGCAGAAGCAG	: 387	
GLC5	:	GTCCTGAGACGTTT	GAGAAGCCACA	ACAACAATCA	AGCAGAAGAGGCT	CAAGGTCAC---AGCAG	: 387	

	400	*	420	*	440	*	
GLC2	:	--GCCCAAGACCGT	CACCAAAAGGTAC	ATCGCTTCAGAG	AGGGTGATTTGAT	CGCAGTGCCTAC	: 428
GLC3	:	--GCCCAAGACCGT	CACCAAGATCTAT	CACCTTCAGAG	AGGGTGATTTGAT	TGCAGTGCCTAC	: 428
GLC1	:	--ACCACAAGACCGT	CACCAAGATCTAT	AACTTCAGAG	AGGGTGATTTGAT	CGCAGTGCCTAC	: 437
GLC4	:	CAGCTACAGGACAGT	CACCAAGATTTC	GTCACTTCAAT	GAAGGAGACGTAC	TCTGTGATTCTCT	: 452
GLC5	:	CAACTACAAGACAGT	CACCAAGATTTC	GTCACTTCAAT	GAAGGAGACGTAC	TAGTGATTCTCT	: 452

Fig. 7a

	460	*	480	*	500	*	520	
GLC2 :	TGGTGTTCATGGTGGATGTACAACAATGAAGACACTCCTGTTGTTGCCGTTTCTATTATTGACA	:	493					
GLC3 :	CGGTTTTGCATACTGGATGTACAACAATGAAGACACTCCTGTTGTTGCCGTTTCTCTTATTGACA	:	493					
GLC1 :	TGGTGTTCATGGTGGATGTACAACAATGAAGACACTCCTGTTGTTGCCGTTTCTATTATTGACA	:	502					
GLC4 :	TAGTGTTCCTTACTGGACCTATAACACTGGCGATGAACCAGTTGTTGCCATCAGTCTTCTTGACA	:	517					
GLC5 :	TGGTGTTCCTTACTGGACCTATAACACTGGCGATGAACCAGTTGTTGCCATCAGTCTCTTGACA	:	517					
	*	540	*	560	*	580		
GLC2 :	CCAACAGCTTGGAGAACCAGCTCGACCAGATGCCTAGGAGATTCTATCTTGCTGGGAACCAAGAG	:	558					
GLC3 :	CCAACAGCTTCCAGAACCAGCTCGACCAGATGCCTAGGAGATTCTATCTTGCTGGGAACCAAGAG	:	558					
GLC1 :	CCAACAGCTTGGAGAACCAGCTCGACCAGATGCCTAGGAGATTCTATCTTGCTGGGAACCAAGAG	:	567					
GLC4 :	CCTCTAACTTCAATAACCAGCTTGATCAAACCCCTAGGGTATTTTACCTTGCTGGGAACCCAGAT	:	582					
GLC5 :	CCTCCAACCTTCAACAATCAGCTTGATCAAACCCAGAGTATTTTACCTTGCTGGGAACCCAGAT	:	582					
	*	600	*	620	*	640	*	
GLC2 :	CAAGAGTTTCTAAAA-----TATCAGC---AGCAGCAGCAAGGAGGTTCCTCAAAGCCAGAAAGG	:	614					
GLC3 :	CAAGAGTTTCTACAG-----TATCAGCCACAGAAGCAGCAAGGAGGTACTCAAAGCCAGAAAGG	:	617					
GLC1 :	CAAGAGTTTCTAAAA-----TATCAGC-----AAGAGCAAGGAGGTTCATCAAAGCCAGAAAGG	:	620					
GLC4 :	ATAGAGTACCCAGAGACCATGCAACAACAACAACAGCAGAAAAGTCATGGTGGACGCAAGCAGGG	:	647					
GLC5 :	ATAGAGCATCCCAGAGACCATGCAACAACAGCAGCAGCAGAGAAGAGTCATGGTGGACGCAAGCAGGG	:	647					
	660	*	680	*	700	*		
GLC2 :	AAAGC---AACAAGAAGAAGAAAACGAAGGAAGCAACATATTGAGTGGCTTCGCCCCCTGAATTCT	:	676					
GLC3 :	AAAGCGTCAGCAAGAAGAAGAAAACGAAGGAGGCAGCATATTGAGTGGCTTCGCCCCCGAATTCT	:	682					
GLC1 :	AAAGCATCAGCAAGAAGAAGAAAACGAAGGAGGCAGCATATTGAGTGGCTTCACCCCTGGAATTCT	:	685					
GLC4 :	GCAACACCAGCAGGAGGAAGAGGAAGAAGGTGGCAGCGTGCTCAGTGGCTTCAGCAAACACTTCT	:	712					
GLC5 :	GCAACACCGACAGCAGGA---GGAAGAAGGTGGCAGTGTGCTCAGTGGCTTCAGCAAACATTTCT	:	709					
	720	*	740	*	760	*	780	
GLC2 :	TGAAAGAAGCGTTTCGGCGTGAACATGCAGATAGTGAGAAACCTACAAGGTGAGAACGAAGAGGAG	:	741					
GLC3 :	TGGAACATGCGTTTCGTCGTGGACAGGCAGATAGTGAGAAAGCTACAAGGTGAGAACGAAGAGGAA	:	747					
GLC1 :	TGGAACATGCATTTCAGCGTGGACAAGCAGATAGCGAAAAACCTACAAGGAGAGAACGAAGGGGAA	:	750					
GLC4 :	TGGCACAATCCTTCAACACCAACGAGGACATAGCTGAGAACT-----TGAGTCTCCAGACGAC	:	771					
GLC5 :	TAGCACAATCCTTCAACACCAACGAGGACACAGCTGAGAACT-----TCGGTCTCCAGATGAC	:	768					
	*	800	*	820	*	840		
GLC2 :	GATAGTGGAGCCATTGTGACAGTGAAAGGAGGTCTAAGAGTCACAGCTCCAGCCATGAGGAAGCC	:	806					
GLC3 :	GAGAAGGGTGCCATTGTGACAGTGAAAGGAGGTCTCAGCGTGATAAGCCCACCCACGGAAGAGCA	:	812					
GLC1 :	GACAAGGGAGCCATTGTGACAGTGAAAGGAGGTCTGAGCGTGATAAAACCCACCCACGGACGAGCA	:	815					
GLC4 :	GAAAGGAAGCAGATCGTGACAGTGGAAGGAGGTCTCAGCGTTATCAGCCCCAAGTGGCAAGAACA	:	836					
GLC5 :	GAAAGGAAGCAGATCGTGACAGTGGAGGGAGGCCTCAGCGTTATCAGCCCCAAGTGGCAAGAACA	:	833					
	*	860	*	880	*	900	*	
GLC2 :	ACAGCAAGAAGAA-----GATGATGATGATGAGGAAGAG-----	:	840					
GLC3 :	GCAACAAAGACC-----CGAGG-----AAGAGGAG-----	:	837					
GLC1 :	GCAACAAAGACCCACGGAAGAGGAAGAAGAAGAGGATGAG-----	:	858					
GLC4 :	ACAAGATGAAGATGAAGATGAAGACGAAGATGATGAAGATGAACAAATTCCTCTCACCTCCTC	:	901					
GLC5 :	A-----GAAGACGAAGACGAAGACGAAGAATATGGACGGACGCCCTCTTATCCTCCAC	:	892					
	920	*	940	*	960	*		
GLC2 :	-----CAGCCA-----CAGTGCCTGGAGACAGACAAAGGTTGC	:	873					
GLC3 :	-----AAGCCA-----GATTGTGACGAGAAAGACAAACATTGC	:	870					
GLC1 :	-----AAGCCA-----CAGTGCAAGGGTAAAGACAAACACTGC	:	891					
GLC4 :	GCCGACCAAGCCATGGAAAGCGTGAACAAGACGAGGACGAGGACGAAGATGAAGATAAACCTCGT	:	966					
GLC5 :	GACGACCAAGCCATGGAAAGCATGAAGATGACGAGGACGAGGACGAAGAAGAAGATCAACCTCGT	:	957					
	980	*	1000	*	1020	*	1040	
GLC2 :	CAACGCC-----	:	880					
GLC3 :	CAAAGCC-----	:	877					
GLC1 :	CAACGCC-----	:	899					
GLC4 :	CCTAGTCGACCAAGCCAAGGAAAGCGGAACAAGACAGGACAGGACGAGGACGAAGATGAAGATGA	:	1031					
GLC5 :	CCTGATCA-----CCC-----	:	968					

Fig. 7b

```

      *           1060           *           1080           *           1100
GLC2 : -----AAAGCAAA : 888
GLC3 : -----AAAGCA-- : 883
GLC1 : -----C---CGAGGAAGCCAAAGCAAA : 918
GLC4 : AGATCAACCTCGCAAGAGCCGCGAATGGAGATCGAAAAAGACACAACCCAGAAGACCTAGACAAG : 1096
GLC5 : -----TCCACA-----G-----CG-----ACCAAGCAGGCCCGAACAAC : 997

      *           1120           *           1140           *           1160           *
GLC2 : AGGAGC-----AGAAATGGCATTGATGAGACCATTTGCACAATGAGA : 930
GLC3 : -G-----AAATGGCATTGACGAGACCATTTGCACAATGAGA : 918
GLC1 : AGCAGA-----AGAAATGGCATTGACGAGACCATATGCACCATGAGA : 960
GLC4 : AAGAACCACGTGAAAGAGGATGCGAGACAAGAAACGGGGTTGAGGAAAATATCTGCACCTTGAAG : 1161
GLC5 : AAGAACCACGTGGAAGAGGATGTCAGACTAGAAATGGGGTTGAGGAAAATATTTGCACCATGAAG : 1062

      1180           *           1200           *           1220           *
GLC2 : CTTCCGCCAAAACATTGGTCAGAATTCATCACCTGACATCTACAACCCTCAAGCTGGTAGCATCAC : 995
GLC3 : CTTCCGCCACAACATTGGCCAGACTTCATCACCTGACATCTTCAACCCTCAAGCTGGTAGCATCAC : 983
GLC1 : CTTCCGCCACAACATTGGCCAGACTTCATCACCTGACATCTACAACCCTCAAGCCGGTAGCGTCAC : 1025
GLC4 : CTTACGAGAACATTGCTCGCCCTTCACGCGCTGACTTCTACAACCCTAAAGCTGGTCGCATTAG : 1226
GLC5 : CTTACGAGAACATTGCTCGCCCTTCACGTGCTGACTTCTACAACCCAAAAGCTGGTCGCATTAG : 1127

      1240           *           1260           *           1280           *           1300
GLC2 : AACCGCCACCAGCCTTGACTTCCCAGCCCTCTGGCTTCTCAAACCTCAGTGCCCAAGTATGGATCAC : 1060
GLC3 : AACCCTACCAGCCTCGACTTCCCAGCCCTCTCGTGGCTCAAACCTCAGTGCCCAAGTATGGATCAC : 1048
GLC1 : AACCGCCACCAGCCTTGACTTCCCAGCCCTCTCGTGGCTCAGACTCAGTGCTGAGTTTGGATCTC : 1090
GLC4 : TACCCTCAACAGCCTCACCCCTCCCAGCCCTCCGCCAATTCCAACCTCAGTGCCCAATATGTTGTCC : 1291
GLC5 : CACCCTCAACAGTCTCACCCCTCCCAGCCCTCCGCCAATTCCGACTCAGTGCCCAATATGTTGTCC : 1192

      *           1320           *           1340           *           1360
GLC2 : TCCGCAAGAATGCTATGTTTCGTGCCACACTACACCCTGAACGCGAACAGCATAATATACGCATTG : 1125
GLC3 : TCCGCAAGAATGCTATGTTTCGTGCCACACTACAACCTGAACGCAAACAGCATAATATACGCATTG : 1113
GLC1 : TCCGCAAGAATGCAATGTTTCGTGCCACACTACAACCTGAACGCGAACAGCATAATATACGCATTG : 1155
GLC4 : TCTACAAGAATGGAATTTACTCTCCACATTGGAATCTGAATGCAAACAGTGTGATCTATGTGACT : 1356
GLC5 : TCTACAGGAATGGAATTTACTCTCCAGATTGGAACCTGAACGCGAACAGTGTGACG---ATGACT : 1254

      *           1380           *           1400           *           1420           *
GLC2 : AATGGGCGGGCATTGGTACAAGTGGTGAATTGCAATGGTGAGAGAGTGTTTGTATGGAGAGCTGCA : 1190
GLC3 : AATGGACGGGCATTGGTACAAGTGGTGAATTGCAATGGTGAGAGAGTGTTTGTATGGAGAGCTGCA : 1178
GLC1 : AATGGACGGGCATTGATACAAGTGGTGAATTGCAACGGTGAGAGAGTGTTTGTATGGAGAGCTGCA : 1220
GLC4 : CGAGGACAAGGAAAGGTTAGAGTTGTGAACCTGCCAAGGGAATGCAGTGTTCGACGGTGAGCTTAG : 1421
GLC5 : CGAGGGAAAGGAAGAGTTAGAGTGGTGAACCTGCCAAGGGAATGCAGTGTTCGACGGTGAGCTAAG : 1319

      1440           *           1460           *           1480           *
GLC2 : AGAGGGAGGGGTGCTGATCGTTCCACAAAACTTTGCGGTGGCTGCA-AAATCCCAGAGCGATAAC : 1254
GLC3 : AGAGGGACAGGTGTTAATTGTGCCACAAAACTTTGCGGTGGCTGCA-AGATCACAGAGCGACAAC : 1242
GLC1 : AGAGGGACGGGTGCTGATCGTGCCACAAAACTTTGTGGTGGCTGCA-AGATCACAGAGTGACAAC : 1284
GLC4 : GAGGGGACAATTGCTGGTGGTACCACAGAACTTCGTGGTGGCGGAGCAAGCCGGAGAACAAGGA- : 1485
GLC5 : GAGGGGACAATTGCTAGTGGTGGCGCAGAACCCCGCGGTGGCTGAGCAAGGGGGAGAACAAGGA- : 1383

      1500           *           1520           *           1540           *           1560
GLC2 : TTTGAGTATGTGTCATTCAAGACCAATGATAGACCCTCGATCGGAAACCTTGCAGGGGCAAACCTC : 1319
GLC3 : TTCGAGTATGTTTCATTCAAGACCAATGATAGACCCTCGATCGGCAACCTTGCAGGTGCAAACCTC : 1307
GLC1 : TTCGAGTATGTGTCATTCAAGACCAATGATACCCATGATCGGCACTCTTGCAGGGGCAAACCTC : 1349
GLC4 : TTCGAATACATAGTATTCAGACACACCACAACGCAGTCACTAGCTACTTGAAGGATG----- : 1543
GLC5 : TTGGAATATGTAGTGTTCAAGACACACCACAACGCCGTGAGCAGCTACATTAAGGATG----- : 1441

```

Fig. 7c

```

      *      1580      *      1600      *      1620
GLC2 : ATTGTTGAACGCATTGCCAGAGGAAGTGATTTCAGCACACTTTTAACTAAAGAGCCAGCAGGCCA : 1384
GLC3 : ATTGTTGAACGCATTGCCGGAGGAAGTGATTTCAGCAAACCTTTTAACTAAGGAGGCAGCAGGCCA : 1372
GLC1 : ATTGTTGAACGCATTACCAGAGGAAGTGATTTCAGCACACTTTCAACCTAAAAAGCCAGCAGGCCA : 1414
GLC4 : --TGTTTAGGGCAATTCCTTCAGAGGTTCTTGCCCATCTTACAACCTTCGACAGAGTCAAGTGT : 1606
GLC5 : --TGTTTAGGGTAATCCCTTCGGAGGTTCTTTCCAATTCTTACAACCTTGGCCAGAGTCAAGTGC : 1504

      *      1640      *      1660      *      1680      *
GLC2 : GGCAGGTGAAGAACAACAACCCTTTCAGCTTCCTTGTTCCACCTCAGGAGTCTCAGAGGAGAGCT : 1449
GLC3 : GGCAGGTCAAGAACAACAACCCTTTCAGCTTCCTGGTTCCACCTAAGGAGTCTCAGAGGAGAGTT : 1437
GLC1 : GGCAGATAAAGAACAACAACCCTTTCAGTTCCTGGTTCCACCTCAGGAGTCTCAGAAGAGAGCT : 1479
GLC4 : CTGAGCTTAAGTATGA-AGGAAATTGGGGTCCTTTGGTCAACCCTGAGTCTCAACAAGGCTCACC : 1670
GLC5 : GTCAGCTCAAGTATCA-AGGAAACTCCGGCCCTTTGGTCAACCCATAA----- : 1551

      1700
GLC2 : GTGGCTTAG----- : 1458
GLC3 : GTGGCTTAG----- : 1446
GLC1 : GTGGCTTAG----- : 1488
GLC4 : CCGTGTTAAAGTCGCATAA : 1689
GLC5 : ----- : -

```

Fig. 7d

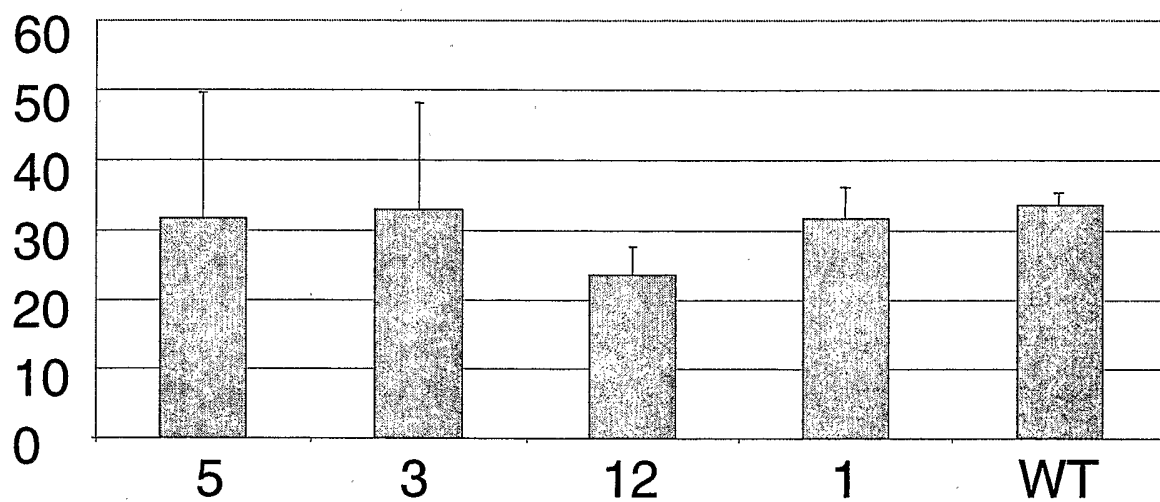


Fig. 8

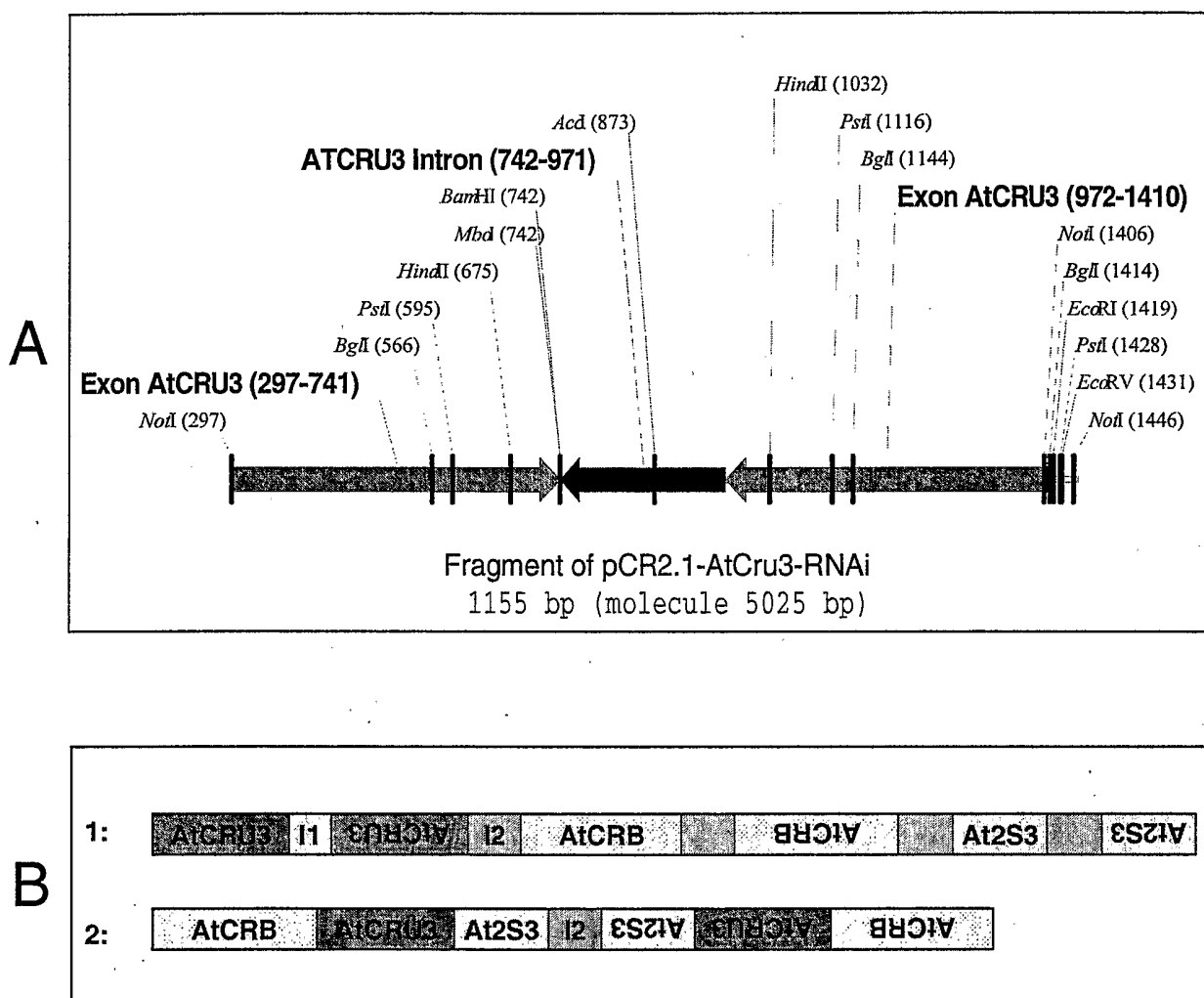


Fig. 9

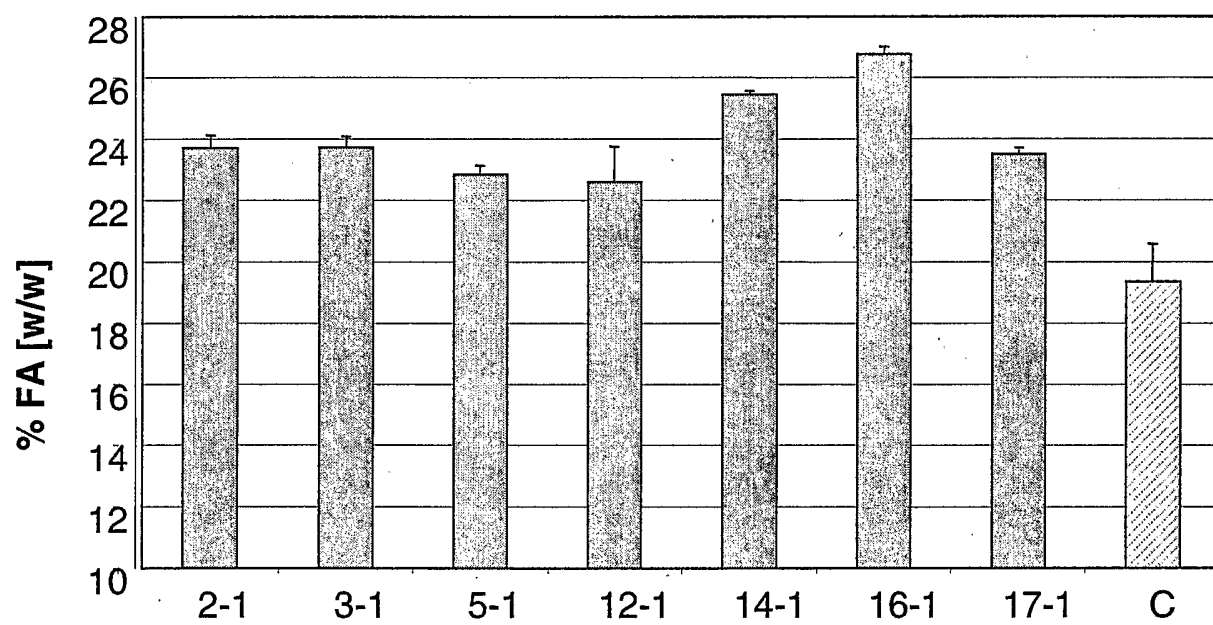


Fig. 10

SEQUENZPROTOKOLL

<110> BASF Plant Science GmbH

<120> Verfahren zum Erhöhen des Ölgehaltes in Pflanzen

<130> NAE_877/2001

<140>

<141>

<160> 179

<170> PatentIn Ver. 2.1

<210> 1

<211> 495

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(492)

<223> albumine 2S subunit 1

<400> 1

atg gca aac aag ttg ttc ctc gtc tgc gca gct ctc gct ctc tgc ttc	48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe	
1 5 10 15	
ctc ctc acc aac gct tcc atc tac cgc acc gtc gtt gag ttc gaa gaa	96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu	
20 25 30	
gat gac gcc act aac ccc ata ggc cca aaa atg agg aaa tgc cgc aag	144
Asp Asp Ala Thr Asn Pro Ile Gly Pro Lys Met Arg Lys Cys Arg Lys	
35 40 45	
gag ttt cag aaa gaa caa cac cta aga gct tgc cag caa ttg atg ctc	192
Glu Phe Gln Lys Glu Gln His Leu Arg Ala Cys Gln Gln Leu Met Leu	
50 55 60	
cag caa gca agg caa ggc cgt agc gat gag ttt gat ttc gaa gac gac	240
Gln Gln Ala Arg Gln Gly Arg Ser Asp Glu Phe Asp Phe Glu Asp Asp	
65 70 75 80	
atg gag aac cca cag gga caa cag cag gaa caa cag cta ttc cag cag	288
Met Glu Asn Pro Gln Gly Gln Gln Gln Glu Gln Gln Leu Phe Gln Gln	
85 90 95	
tgc tgc aac gag ctt cgc cag gaa gag cca gat tgt gtt tgc ccc acc	336
Cys Cys Asn Glu Leu Arg Gln Glu Glu Pro Asp Cys Val Cys Pro Thr	
100 105 110	
ttg aaa caa gct gcc aag gcc gtt aga ctc cag gga cag cac caa cca	384
Leu Lys Gln Ala Ala Lys Ala Val Arg Leu Gln Gly Gln His Gln Pro	
115 120 125	
atg caa gtc agg aaa att tac cag aca gcc aag cac ttg ccc aac gtt	432
Met Gln Val Arg Lys Ile Tyr Gln Thr Ala Lys His Leu Pro Asn Val	
130 135 140	
tgc gac atc ccg caa gtt gat gtt tgt ccc ttc aac atc cct tca ttc	480
Cys Asp Ile Pro Gln Val Asp Val Cys Pro Phe Asn Ile Pro Ser Phe	
145 150 155 160	
cct tct ttc tac taa	495
Pro Ser Phe Tyr	

<210> 2
 <211> 164
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 2
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe
 1 5 10 15
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu
 20 25 30
 Asp Asp Ala Thr Asn Pro Ile Gly Pro Lys Met Arg Lys Cys Arg Lys
 35 40 45
 Glu Phe Gln Lys Glu Gln His Leu Arg Ala Cys Gln Gln Leu Met Leu
 50 55 60
 Gln Gln Ala Arg Gln Gly Arg Ser Asp Glu Phe Asp Phe Glu Asp Asp
 65 70 75 80
 Met Glu Asn Pro Gln Gly Gln Gln Gln Glu Gln Gln Leu Phe Gln Gln
 85 90 95
 Cys Cys Asn Glu Leu Arg Gln Glu Glu Pro Asp Cys Val Cys Pro Thr
 100 105 110
 Leu Lys Gln Ala Ala Lys Ala Val Arg Leu Gln Gly Gln His Gln Pro
 115 120 125
 Met Gln Val Arg Lys Ile Tyr Gln Thr Ala Lys His Leu Pro Asn Val
 130 135 140
 Cys Asp Ile Pro Gln Val Asp Val Cys Pro Phe Asn Ile Pro Ser Phe
 145 150 155 160
 Pro Ser Phe Tyr

<210> 3
 <211> 495
 <212> DNA
 <213> Arabidopsis thaliana
 <220>
 <221> CDS
 <222> (1)..(492)
 <223> albumine 2S subunit 3
 <400> 3
 atg gct aac aag ctc ttc ctc gtc tgc gca act ctc gcc ctc tgc ttc 48
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Leu Ala Leu Cys Phe
 1 5 10 15
 ctc ctc acc aac gct tcc atc tac cgc acc gtt gtc gaa ttc gaa gaa 96
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu
 20 25 30
 gat gac gcc agc aac ccc gta ggt cca aga cag aga tgc cag aag gag 144
 Asp Asp Ala Ser Asn Pro Val Gly Pro Arg Gln Arg Cys Gln Lys Glu
 35 40 45
 ttt cag caa tca caa cac cta aga gct tgc cag aga tgg atg agc aag 192
 Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Arg Trp Met Ser Lys
 50 55 60

3

```

caa atg agg caa gga cgt ggt ggt ggt cct tcc ctc gac gat gag ttc 240
Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
65 70 75 80

gat ttc gag ggc ccc cag cag gga tac cag cta ctc cag cag tgc tgc 288
Asp Phe Glu Gly Pro Gln Gln Gly Tyr Gln Leu Leu Gln Gln Cys Cys
85 90 95

aac gag ctt cgc cag gaa gag cca gtt tgc gtt tgc ccc acc ttg aaa 336
Asn Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro Thr Leu Lys
100 105 110

caa gct gcc agg gca gtt agc ctc cag gga cag cac gga cca ttc caa 384
Gln Ala Ala Arg Ala Val Ser Leu Gln Gly Gln His Gly Pro Phe Gln
115 120 125

tcc agg aaa att tac cag tca gct aag tac ttg cct aac att tgc aag 432
Ser Arg Lys Ile Tyr Gln Ser Ala Lys Tyr Leu Pro Asn Ile Cys Lys
130 135 140

atc cag caa gtt ggt gaa tgt ccc ttc cag acc acc atc cct ttc ttc 480
Ile Gln Gln Val Gly Glu Cys Pro Phe Gln Thr Thr Ile Pro Phe Phe
145 150 155 160

cct cct tac tac tag 495
Pro Pro Tyr Tyr

<210> 4
<211> 164
<212> PRT
<213> Arabidopsis thaliana

<400> 4
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Leu Ala Leu Cys Phe
1 5 10 15
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu
20 25 30
Asp Asp Ala Ser Asn Pro Val Gly Pro Arg Gln Arg Cys Gln Lys Glu
35 40 45
Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Arg Trp Met Ser Lys
50 55 60
Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
65 70 75 80
Asp Phe Glu Gly Pro Gln Gln Gly Tyr Gln Leu Leu Gln Gln Cys Cys
85 90 95
Asn Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro Thr Leu Lys
100 105 110
Gln Ala Ala Arg Ala Val Ser Leu Gln Gly Gln His Gly Pro Phe Gln
115 120 125
Ser Arg Lys Ile Tyr Gln Ser Ala Lys Tyr Leu Pro Asn Ile Cys Lys
130 135 140
Ile Gln Gln Val Gly Glu Cys Pro Phe Gln Thr Thr Ile Pro Phe Phe
145 150 155 160
Pro Pro Tyr Tyr

```

<210> 5

<211> 513

<212> DNA
 <213> Arabidopsis thaliana
 <220>
 <221> CDS
 <222> (1)..(510)
 <223> albumine 2S subunit 2
 <400> 5
 atg gca aac aag ctc ttc ctc gtc tgc gca act ttc gcc ctc tgc ttc 48
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Phe Ala Leu Cys Phe
 1 5 10 15
 ctc ctc acc aac gct tcc atc tac cgc act gtt gtc gag ttc gac gaa 96
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
 20 25 30
 gat gac gcc agc aac ccc atg ggc cca aga cag aaa tgt cag aag gag 144
 Asp Asp Ala Ser Asn Pro Met Gly Pro Arg Gln Lys Cys Gln Lys Glu
 35 40 45
 ttt cag caa tca cag cac cta aga gct tgc cag aaa ttg atg cgc atg 192
 Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Lys Leu Met Arg Met
 50 55 60
 caa atg agg caa ggc cgt ggt ggt ggt ccc tcc ctc gac gat gag ttc 240
 Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
 65 70 75 80
 gat ttg gaa gac gac atc gag aac cca caa ggc ccc cag cag gga cac 288
 Asp Leu Glu Asp Asp Ile Glu Asn Pro Gln Gly Pro Gln Gln Gly His
 85 90 95
 cag atc ctc cag cag tgc tgc agc gag ctt cgc cag gaa gag cca gtt 336
 Gln Ile Leu Gln Gln Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val
 100 105 110
 tgt gtt tgc ccc acc ttg aga caa gct gcc agg gcc gtt agc ctc cag 384
 Cys Val Cys Pro Thr Leu Arg Gln Ala Ala Arg Ala Val Ser Leu Gln
 115 120 125
 gga caa cac gga cca ttc caa tcc agg aaa att tac aag aca gct aag 432
 Gly Gln His Gly Pro Phe Gln Ser Arg Lys Ile Tyr Lys Thr Ala Lys
 130 135 140
 tac ttg cct aac att tgc aag atc cag caa gtt ggt gaa tgc ccc ttc 480
 Tyr Leu Pro Asn Ile Cys Lys Ile Gln Gln Val Gly Glu Cys Pro Phe
 145 150 155 160
 cag acc acc atc cct ttc ttc cct cct tac taa 513
 Gln Thr Thr Ile Pro Phe Phe Pro Pro Tyr
 165 170
 <210> 6
 <211> 170
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 6
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Phe Ala Leu Cys Phe
 1 5 10 15
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
 20 25 30
 Asp Asp Ala Ser Asn Pro Met Gly Pro Arg Gln Lys Cys Gln Lys Glu
 35 40 45

Phe	Gln	Gln	Ser	Gln	His	Leu	Arg	Ala	Cys	Gln	Lys	Leu	Met	Arg	Met
50			55			60									
Gln	Met	Arg	Gln	Gly	Arg	Gly	Gly	Gly	Pro	Ser	Leu	Asp	Asp	Glu	Phe
65			70			75			80						
Asp	Leu	Glu	Asp	Asp	Ile	Glu	Asn	Pro	Gln	Gly	Pro	Gln	Gln	Gly	His
			85			90			95						
Gln	Ile	Leu	Gln	Gln	Cys	Cys	Ser	Glu	Leu	Arg	Gln	Glu	Glu	Pro	Val
			100			105			110						
Cys	Val	Cys	Pro	Thr	Leu	Arg	Gln	Ala	Ala	Arg	Ala	Val	Ser	Leu	Gln
115						120			125						
Gly	Gln	His	Gly	Pro	Phe	Gln	Ser	Arg	Lys	Ile	Tyr	Lys	Thr	Ala	Lys
130						135			140						
Tyr	Leu	Pro	Asn	Ile	Cys	Lys	Ile	Gln	Gln	Val	Gly	Glu	Cys	Pro	Phe
145			150			155			160						
Gln	Thr	Thr	Ile	Pro	Phe	Phe	Pro	Pro	Tyr						
			165			170									

<210> 7

<211> 501

<212> DNA

<213> Arabidopsis thaliana

 $\langle 220 \rangle$

<221> CDS

<222> (1) .. (498)

<223> albumine 2S subunit 4

 $\langle 400 \rangle \cdot 7$

atg gcg aac aag ctc ttc ctc gtc tgc gca gct ctc gcc ctg tgt ttc 48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe
1 5 10 15

atc ctc acc aac gct tcc gtc tat cgc acc gtt gtc gag ttc gac gaa 96
Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu
20 25 30

gat gac gcc agt aac ccc ata ggc cca ata cag aaa tgt cag aag gag 144
Asp Asp Ala Ser Asn Pro Ile Gly Pro Ile Gln Lys Cys Gln Lys Glu
35 40 45

ttt cag caa gac cag cac cta aga gct tgc cag aga tgg atg cgc aag 192
Phe Gln Gln Asp Gln His Leu Arg Ala Cys Gln Arg Trp Met Arg Lys
50 55 60

caa atg tgg caa gga cgt ggt ggt ggt cct tcc ctc gac gat gag ttc 240
Gln Met Trp Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
65 70 75 80

gat atg gaa gac gac atc gag aac ccg cag aga cga cag cta ctc cag 288
Asp Met Glu Asp Asp Ile Glu Asn Pro Gln Arg Arg Gln Leu Leu Gln
85 90 95

aag tgc tgc agc gag ctt cgc caa gaa gag cca gtt tgc gtt tgc ccc 336
Lys Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro
100 105 110

acc ttg aga caa gct gcc aag gcc gtt aga ttc cag gga cag caa cac 384
Thr Leu Arg Gln Ala Ala Lys Ala Val Arg Phe Gln Gly Gln Gln His
115 120 125

caa cca gag caa gtc agg aaa att tac cag gca gct aag tac ttg cct 432
 Gln Pro Glu Gln Val Arg Lys Ile Tyr Gln Ala Ala Lys Tyr Leu Pro
 130 135 140

aac att tgc aaa atc cag caa gtt ggt gtt tgc ccc ttc cag atc cct 480
 Asn Ile Cys Lys Ile Gln Gln Val Gly Val Cys Pro Phe Gln Ile Pro
 145 150 155 160

tca atc cct tct tac tac taa 501
 Ser Ile Pro Ser Tyr Tyr
 165

<210> 8

<211> 166

<212> PRT

<213> Arabidopsis thaliana

<400> 8

Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe
 1 5 10 15

Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu
 20 25 30

Asp Asp Ala Ser Asn Pro Ile Gly Pro Ile Gln Lys Cys Gln Lys Glu
 35 40 45

Phe Gln Gln Asp Gln His Leu Arg Ala Cys Gln Arg Trp Met Arg Lys
 50 55 60

Gln Met Trp Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
 65 70 75 80

Asp Met Glu Asp Asp Ile Glu Asn Pro Gln Arg Arg Gln Leu Leu Gln
 85 90 95

Lys Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro
 100 105 110

Thr Leu Arg Gln Ala Ala Lys Ala Val Arg Phe Gln Gly Gln Gln His
 115 120 125

Gln Pro Glu Gln Val Arg Lys Ile Tyr Gln Ala Ala Lys Tyr Leu Pro
 130 135 140

Asn Ile Cys Lys Ile Gln Gln Val Gly Val Cys Pro Phe Gln Ile Pro
 145 150 155 160

Ser Ile Pro Ser Tyr Tyr
 165

<210> 9

<211> 1473

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (1)..(1470)

<223> cruciferin

<400> 9

atg gct cgg ctc tca tct ctt ctc tct ttt tcc tta gca ctt ttg atc 48
 Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Ile
 1 5 10 15

ttt ctc cat ggc tct aca gct caa cag ttt cca aac gag tgt cag cta	96
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu	
20 25 30	
gac cag ctc aat gca ctg gag ccg tca cac gta ctt aag gct gag gct	144
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala	
35 40 45	
ggg cgc atc gag gtg tgg gac cac cac gct cct cag cta cgt tgc tct	192
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	
ggg gtc tcc ttt gta cgt tac atc atc gag tct aag ggt ctc tac ttg	240
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu	
65 70 75 80	
ccc tct ttc ttt agc acc gcg aag ctc tcc ttc gtg gct aaa gga gaa	288
Pro Ser Phe Phe Ser Thr Ala Lys Leu Ser Phe Val Ala Lys Gly Glu	
85 90 95	
ggg ctt atg ggg aga gtg gtc cct gga tgc gcc gag aca ttc cag gac	336
Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp	
100 105 110	
tca tca gtg ttt caa cca agc ggt ggt agc ccc tcg gga gaa ggt cag	384
Ser Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Ser Gly Glu Gly Gln	
115 120 125	
ggc caa gga caa caa ggt cag ggc caa ggc cac caa ggt caa ggc caa	432
Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln	
130 135 140	
gga caa cag ggc caa caa ggt cag caa gga caa cag agt caa ggc cag	480
Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln	
145 150 155 160	
ggc ttc cgt gat atg cac cag aaa gtg gag cac ata agg act ggg gac	528
Gly Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp	
165 170 175	
acc atc gct aca cat ccc ggt gta gcc caa tgg ttc tac aac gac gga	576
Thr Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly	
180 185 190	
aac caa cca ctt gtc atc gtt tcc gtc ctc gat tta gcc agc cac cag	624
Asn Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln	
195 200 205	
aat cag ctc gac cgc aac cca agg cca ttt tac tta gcc gga aac aac	672
Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn	
210 215 220	
cca caa ggc caa gta tgg ata gaa gga cgc gag caa cag cca caa aag	720
Pro Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys	
225 230 235 240	
aac atc ctt aat ggc ttc aca cca gag gtt ctt gct aaa gct ttc aag	768
Asn Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys	
245 250 255	
atc gat gtt agg aca gcg caa caa ctt cag aac cag caa gac aac cgt	816
Ile Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg	
260 265 270	
gga aac att atc cga gtc caa ggc cca ttc agt gtc att agg ccg cct	864
Gly Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro	
275 280 285	

ttg agg agt cag aga ccg cag gag aca gaa gtt aac ggt tta gaa gag 912
 Leu Arg Ser Gln Arg Pro Gln Glu Thr Glu Val Asn Gly Leu Glu Glu
 290 295 300
 acc ata tgc agc gcg agg tgc acc gat aac ctc gat gac cca tct aat 960
 Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn
 305 310 315 320
 gct gac gta tac aag cca cag ctc ggt tac atc agc act ctg aac agc 1008
 Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser
 325 330 335
 tat gat ctc ccc atc ctt cgc ttc ctt cgt ctc tca gcc ctc cgt gga 1056
 Tyr Asp Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly
 340 345 350
 tct atc cgt caa aac gcg atg gtg ctt cca cag tgg aac gca aac gca 1104
 Ser Ile Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala
 355 360 365
 aac gcg gtt ctc tac gtg aca gac ggg gaa gcc cat gtg cag gtg gtt 1152
 Asn Ala Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val
 370 375 380
 aac gac aac ggt gac aga gtg ttc gac gga caa gtc tct caa gga cag 1200
 Asn Asp Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln
 385 390 395 400
 cta ctt tcc ata cca caa ggt ttc tcc gtg gtg aaa cgc gca aca agc 1248
 Leu Leu Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser
 405 410 415
 gaa cag ttc cgg tgg atc gag ttc aag aca aac gca aac gca cag atc 1296
 Glu Gln Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile
 420 425 430
 aac aca ctt gct gga cga acc tcg gtc ttg aga ggt tta cca tta gag 1344
 Asn Thr Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu
 435 440 445
 gtc ata tcc aat ggg tac caa atc tca ctc gaa gaa gca aga agg gtt 1392
 Val Ile Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val
 450 455 460
 aag ttc aac acg atc gag acc act ttg acg cac agc agt ggc cca gct 1440
 Lys Phe Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala
 465 470 475 480
 agc tac gga ggg cca agg aag gct gat gct taa 1473
 Ser Tyr Gly Gly Pro Arg Lys Ala Asp Ala
 485 490

<210> 10

<211> 490

<212> PRT

<213> Brassica napus

<400> 10

Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Ile
 1 5 10 15

Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu
 20 25 30

Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala
 35 40 45

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

10

Asn Thr Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu
 435 440 445
 Val Ile Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val
 450 455 460
 Lys Phe Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala
 465 470 475 480
 Ser Tyr Gly Gly Pro Arg Lys Ala Asp Ala
 485 490

<210> 11

<211> 1467

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (1)..(1464)

<223> cruciferin

<400> 11

atg gct cgg ctc tca tct ctt ctc tct ttt tcc tta gca ctt ttg act	48
Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Thr	
1 5 10 15	
ttt ctc cat ggc tct aca gct caa cag ttt cca aac gag tgt cag cta	96
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu	
20 25 30	
gac cag ctc aat gca ctg gag ccg tca cac gta ctt aag gct gag gct	144
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala	
35 40 45	
ggg cgc atc gag gtg tgg gac cac cac gct cct cag cta cgt tgc tct	192
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	
ggg gtc tcc ttt gta cgt tac atc atc gag tct aag ggt ctc tac ttg	240
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu	
65 70 75 80	
ccc tct ttc ttt agc acc gcg agg ctc tcc ttc gtg gct aaa gga gaa	288
Pro Ser Phe Phe Ser Thr Ala Arg Leu Ser Phe Val Ala Lys Gly Glu	
85 90 95	
ggg ctt atg ggg aga gtg gtc ctg tgc gcc gag aca ttc cag gac tca	336
Gly Leu Met Gly Arg Val Val Leu Cys Ala Glu Thr Phe Gln Asp Ser	
100 105 110	
tca gtg ttt caa cca agc ggt ggt agc ccc ttc gga gaa ggt cag ggc	384
Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Phe Gly Glu Gly Gln Gly	
115 120 125	
caa gga caa caa ggt cag ggc caa ggc cac caa ggt caa ggc caa gga	432
Gln Gly Gln Gln Gly Gln Gln Gly His Gln Gly Gln Gly Gln Gly	
130 135 140	
caa cag ggc caa caa ggt cag caa gga caa cag agt caa ggc cag ggt	480
Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln Gly	
145 150 155 160	
ttc cgt gat atg cac cag aaa gtg gag cac ata agg act ggg gac acc	528
Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp Thr	
165 170 175	

11

atc gct aca cat ccc ggt gta gcc caa tgg ttc tac aac gac gga aac	576
Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly Asn	
180 185 190	
caa cca ctt gtc atc gtt tcc gtc ctc gat tta gcc agc cac cag aat	624
Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln Asn	
195 200 205	
cag ctc gac cgc aac cca agg cca ttt tac tta gcc gga aac aac cca	672
Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn Pro	
210 215 220	
caa ggc caa gta tgg ata gaa gga cgc gag caa cag cca caa aag aac	720
Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys Asn	
225 230 235 240	
atc ctt aat ggc ttc aca cca gag gtt ctt gct aaa gct ttc aag atc	768
Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys Ile	
245 250 255	
gat gtt agg aca gcg caa caa ctt cag aac cag caa gac aac cgt gga	816
Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg Gly	
260 265 270	
aac att atc cga gtc caa ggc cca ttc agt gtc att agg ccg cct ttg	864
Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro Leu	
275 280 285	
agg agt cag aga ccg cag gag gaa gtt aac ggt tta gaa gag acc ata	912
Arg Ser Gln Arg Pro Gln Glu Glu Val Asn Gly Leu Glu Glu Thr Ile	
290 295 300	
tgc agc gcg agg tgc acc gat aac ctc gat gac cca tct aat gct gac	960
Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn Ala Asp	
305 310 315 320	
gta tac aag cca cag ctc ggt tac atc agc act ctg aac agc tat gat	1008
Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp	
325 330 335	
ctc ccc atc ctt cgc ttc ctt cgt ctc tca gcc ctc cgt gga tct atc	1056
Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly Ser Ile	
340 345 350	
cgt caa aac gcg atg gtg ctt cca cag tgg aac gca aac gca aac gcg	1104
Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala	
355 360 365	
gtt ctc tac gtg aca gac ggg gaa gcc cat gtg cag gtg gtt aac gac	1152
Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val Asn Asp	
370 375 380	
aac ggt gac aga gtg ttc gac gga caa gtc tct caa gga cag cta ctt	1200
Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Leu	
385 390 395 400	
tcc ata cca caa ggt ttc tcc gtg gtg aaa cgc gca aca agc gaa cag	1248
Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Glu Gln	
405 410 415	
ttc cgg tgg atc gag ttc aag aca aac gca aac gca cag atc aac aca	1296
Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr	
420 425 430	
ctt gct gga cga acc tcg gtc ttg aga ggt tta cca tta gag gtc ata	1344
Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile	
435 440 445	

12

tcc aat ggg tac caa atc tca ctc gaa gaa gca aga agg gtt aag ttc 1392
 Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val Lys Phe
 450 455 460

aac acg atc gag acc act ttg acg cac agc agt ggc cca gct agc tac 1440
 Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr
 465 470 475 480

gga ggg cca agg aag gct gat gct taa 1467
 Gly Gly Pro Arg Lys Ala Asp Ala
 485

<210> 12

<211> 488

<212> PRT

<213> Brassica napus

<400> 12

Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Thr
 1 5 10 15

Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu
 20 25 30

Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala
 35 40 45

Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser
 50 55 60

Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu
 65 70 75 80

Pro Ser Phe Phe Ser Thr Ala Arg Leu Ser Phe Val Ala Lys Gly Glu
 85 90 95

Gly Leu Met Gly Arg Val Val Leu Cys Ala Glu Thr Phe Gln Asp Ser
 100 105 110

Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Phe Gly Glu Gly Gln Gly
 115 120 125

Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln Gly
 130 135 140

Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln Gly
 145 150 155 160

Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp Thr
 165 170 175

Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly Asn
 180 185 190

Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln Asn
 195 200 205

Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn Pro
 210 215 220

Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys Asn
 225 230 235 240

Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys Ile
 245 250 255

Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg Gly
 260 265 270

13

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

<210> 13

<211> 1491

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (1)..(1488)

<223> cruciferin BnC2

<400> 13

atg gct cga ctc tcg tct ctt ctc tat ttt tcg ata aca gtt ttg atc 48
 Met Ala Arg Leu Ser Ser Leu Leu Tyr Phe Ser Ile Thr Val Leu Ile
 1 5 10 15
 ttt ctc cat ggc tct aca gct caa cag ttt cca aac gag tgc caa cta 96
 Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu
 20 25 30
 gac cag ctc aat gcg ctg gag ccg tca cac gta ctt aag gcc gag gct 144
 Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala
 35 40 45
 ggt cgc atc gaa gtg tgg gac cac cac gct cct cag cta cgc tgc tct 192
 Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser
 50 55 60

14

ggt gtc tcc ttc gta cgt tac ata atc gag tct cag ggt cta tac ttg	240
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Gln Gly Leu Tyr Leu	
65 70 75 80	
ccc tct ttc tta aat acc gcg aac gtc tct ttc gtt gct aaa gga caa	288
Pro Ser Phe Leu Asn Thr Ala Asn Val Ser Phe Val Ala Lys Gly Gln	
85 90 95	
ggc ctt atg ggg aga gtg gtc cct gga tgc gct gag act ttc cag gac	336
Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp	
100 105 110	
tca tca gta ttc caa cca ggc agt ggc agc ccc ttc gga gaa ggt caa	384
Ser Ser Val Phe Gln Pro Gly Ser Gly Ser Pro Phe Gly Glu Gly Gln	
115 120 125	
ggc caa ggt cag cag ggt cag ggc caa ggt cag ggt cag ggt caa ggc	432
Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly Gln Gly Gln Gly Gln Gly	
130 135 140	
aag ggc caa cag ggt caa ggc aag ggc caa cag ggt caa tcc cag ggc	480
Lys Gly Gln Gln Gly Gln Gly Lys Gly Gln Gln Gly Gln Ser Gln Gly	
145 150 155 160	
caa cag ggt caa ggt caa ggt ttc cgt gat atg cac cag aaa gta gag	528
Gln Gln Gly Gln Gly Gln Gly Phe Arg Asp Met His Gln Lys Val Glu	
165 170 175	
cac ata agg agc ggc gac acc att gct aca cat ccc ggt gta gct caa	576
His Ile Arg Ser Gly Asp Thr Ile Ala Thr His Pro Gly Val Ala Gln	
180 185 190	
tgg ttc tac aac aat gga aac caa cct ctt gtc atc gtt gcc gtc atg	624
Trp Phe Tyr Asn Asn Gly Asn Gln Pro Leu Val Ile Val Ala Val Met	
195 200 205	
gat tta gct agc cac cag aac cag ctt gac cgc aac cca agc caa ttt	672
Asp Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Ser Gln Phe	
210 215 220	
tac tta gca gga aaa aac cca caa ggc caa tca tgg cta cac gga cga	720
Tyr Leu Ala Gly Lys Asn Pro Gln Gly Gln Ser Trp Leu His Gly Arg	
225 230 235 240	
ggg caa cag cca caa aac aac atc ctt aat ggc ttc tct cca gag gtt	768
Gly Gln Gln Pro Gln Asn Asn Ile Leu Asn Gly Phe Ser Pro Glu Val	
245 250 255	
ctt gct caa gcg ttc aag atc gat gtt agg aca gcg caa caa ctt cag	816
Leu Ala Gln Ala Phe Lys Ile Asp Val Arg Thr Ala Gln Gln Leu Gln	
260 265 270	
aac cag caa gat aac cgg gga aac att gtc cgt gtc caa ggc ccc ttc	864
Asn Gln Gln Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe	
275 280 285	
ggt gtt att agg ccg cca ttg aaa agc cag aga cca cag gag aca gaa	912
Gly Val Ile Arg Pro Pro Leu Lys Ser Gln Arg Pro Gln Glu Thr Glu	
290 295 300	
gct aac ggt cta gaa gag acc ata tgc agc gca agg tgc acg gat aac	960
Ala Asn Gly Leu Glu Glu Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn	
305 310 315 320	
ctc gat gac cca tct aac gcg gat gtg tat aag cca cag ctt ggt tac	1008
Leu Asp Asp Pro Ser Asn Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr	
325 330 335	

15

atc agc att ctt aac agt tat gat cta ccc atc ctt cgc gta ctt cgc 1056
 Ile Ser Ile Leu Asn Ser Tyr Asp Leu Pro Ile Leu Arg Val Leu Arg
 340 345 350

ctc tca gcc ctc cgt gga tca atc cgt caa aat gca atg gtt ctt cca 1104
 Leu Ser Ala Leu Arg Gly Ser Ile Arg Gln Asn Ala Met Val Leu Pro
 355 360 365

cag tgg aag tca aag tca aac gcg gtt ctc tac gtg aca gac ggg gaa 1152
 Gln Trp Lys Ser Lys Ser Asn Ala Val Leu Tyr Val Thr Asp Gly Glu
 370 375 380

gcc caa ata cag gtg gtt aac gac aac ggt gac aga gtg ttc gat gga 1200
 Ala Gln Ile Gln Val Asn Asp Asn Gly Asp Arg Val Phe Asp Gly
 385 390 395 400

caa gtc tct caa ggg cag cta ctt tcc att cca caa gga ttc tcc gtt 1248
 Gln Val Ser Gln Gly Gln Leu Leu Ser Ile Pro Gln Gly Phe Ser Val
 405 410 415

gtg aaa cgc gca aca agc gat cag ttc agg tgg ata gaa ttc aag aca 1296
 Val Lys Arg Ala Thr Ser Asp Gln Phe Arg Trp Ile Glu Phe Lys Thr
 420 425 430

aac gca aac gcc cag atc aac act ctt gct gga cgt acc tca gtc atg 1344
 Asn Ala Asn Ala Gln Ile Asn Thr Leu Ala Gly Arg Thr Ser Val Met
 435 440 445

aga ggt tta cca tta gag gtc ata gcc aat ggg tac caa atc tca ctt 1392
 Arg Gly Leu Pro Leu Glu Val Ile Ala Asn Gly Tyr Gln Ile Ser Leu
 450 455 460

gaa gaa gca aga agg gtt aag ttc aac aca ata gag acc act ttg acc 1440
 Glu Glu Ala Arg Arg Val Lys Phe Asn Thr Ile Glu Thr Thr Leu Thr
 465 470 475 480

cac agt agt ggc cca gcg agc tac gga agg cca agg aag gct gat gct 1488
 His Ser Ser Gly Pro Ala Ser Tyr Gly Arg Pro Arg Lys Ala Asp Ala
 485 490 495

tga 1491

<210> 14
 <211> 496
 <212> PRT
 <213> Brassica napus

<400> 14
 Met Ala Arg Leu Ser Ser Leu Leu Tyr Phe Ser Ile Thr Val Leu Ile
 1 5 10 15
 Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu
 20 25 30
 Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala
 35 40 45
 Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser
 50 55 60
 Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Gln Gly Leu Tyr Leu
 65 70 75 80
 Pro Ser Phe Leu Asn Thr Ala Asn Val Ser Phe Val Ala Lys Gly Gln
 85 90 95
 Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp
 100 105 110

16

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

<210> 15
 <211> 555
 <212> DNA
 <213> Brassica napus
 <220>
 <221> CDS
 <222> (1)..(552)
 <223> cruciferin cru4

<400> 15
 ttg tgc aca atg aga tgc acc gaa aac ctt gat gac ccg tca agt gct 48
 Leu Cys Thr Met Arg Cys Thr Glu Asn Leu Asp Asp Pro Ser Ser Ala
 1 5 10 15
 gat gtc tac aag cca tcg ctc gga tac att agc aca ctc aac agc tac 96
 Asp Val Tyr Lys Pro Ser Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr
 20 25 30
 aac ctc cct atc ctc aga ttc ctc cgc ctt agc gct ctt cgt ggc tcc 144
 Asn Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly Ser
 35 40 45
 atc cat aac aac gct atg gtg ctg ccg caa tgg aac gtg aac gca aac 192
 Ile His Asn Asn Ala Met Val Leu Pro Gln Trp Asn Val Asn Ala Asn
 50 55 60
 gcg gca ctc tac gtg aca aag ggg aag gct cat ata cag atg gtg aac 240
 Ala Ala Leu Tyr Val Thr Lys Gly Lys Ala His Ile Gln Met Val Asn
 65 70 75 80
 gac aac gga caa aga gtg ttt gac caa gag atc tcc cag gga cag tta 288
 Asp Asn Gly Gln Arg Val Phe Asp Gln Glu Ile Ser Gln Gly Gln Leu
 85 90 95
 ctt gtc gtg cca caa ggc ttc gcg gtc gtg aaa cgt gcc aca agc caa 336
 Leu Val Val Pro Gln Gly Phe Ala Val Val Lys Arg Ala Thr Ser Gln
 100 105 110
 cag ttc cag tgg atc gag ttc aag agc aac gac aac gca cag atc aac 384
 Gln Phe Gln Trp Ile Glu Phe Lys Ser Asn Asp Asn Ala Gln Ile Asn
 115 120 125
 aca ctc gcg gga cgc acc tca gtc atg aga ggt tta cca ctt gag gtt 432
 Thr Leu Ala Gly Arg Thr Ser Val Met Arg Gly Leu Pro Leu Glu Val
 130 135 140
 ata tcc aac ggg tat cag atc tca ccc caa gaa gct aga agt gtt aag 480
 Ile Ser Asn Gly Tyr Gln Ile Ser Pro Gln Glu Ala Arg Ser Val Lys
 145 150 155 160
 ttc agc act ctt gag acc aca ttg act caa agc agt ggt cct atg ggc 528
 Phe Ser Thr Leu Glu Thr Thr Leu Thr Gln Ser Ser Gly Pro Met Gly
 165 170 175
 tac ggt atg cct aga gtc gag gct tga 555
 Tyr Gly Met Pro Arg Val Glu Ala
 180

<210> 16
 <211> 184
 <212> PRT
 <213> Brassica napus

<400> 16
 Leu Cys Thr Met Arg Cys Thr Glu Asn Leu Asp Asp Pro Ser Ser Ala
 1 5 10 15

18

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

<210> 17

<211> 1530

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (1)..(1527)

<223> cruciferin cru4

<400> 17

atg gtt aaa gtt cct cat ctc ctc gtc gca acg ttc ggg gtt ctc ctc 48
 Met Val Lys Val Pro His Leu Leu Val Ala Thr Phe Gly Val Leu Leu
 1 5 10 15
 gtc ctc aac ggc tgt ctc gca agg cag tcg cta ggg gtt cct cct cag 96
 Val Leu Asn Gly Cys Leu Ala Arg Gln Ser Leu Gly Val Pro Pro Gln
 20 25 30
 cta ggg aac gcg tgt aac ctc gat aac tta gac gtt ctc cag cct acc 144
 Leu Gly Asn Ala Cys Asn Leu Asp Asn Leu Asp Val Leu Gln Pro Thr
 35 40 45
 gaa act atc aag agc gag gct ggt cgg gtc gag tac tgg gat cac aac 192
 Glu Thr Ile Lys Ser Glu Ala Gly Arg Val Glu Tyr Trp Asp His Asn
 50 55 60
 aat cct cag atc cga tgt gct ggt gtc tct gtc tct cgt gtt ata atc 240
 Asn Pro Gln Ile Arg Cys Ala Gly Val Ser Val Ser Arg Val Ile Ile
 65 70 75 80
 gaa caa ggc ggt ctc tac ctt cct acc ttc ttc agc tcc ccc aaa att 288
 Glu Gln Gly Gly Leu Tyr Leu Pro Thr Phe Phe Ser Ser Pro Lys Ile
 85 90 95

19

tca tac gtt gtt caa gga atg ggt att agc gga aga gtg gtc cct gga	336
Ser Tyr Val Val Gln Gly Met Gly Ile Ser Gly Arg Val Val Pro Gly	
100 105 110	
tgc gcg gaa acc ttc atg gac tcg cag cct atg caa gga caa caa caa	384
Cys Ala Glu Thr Phe Met Asp Ser Gln Pro Met Gln Gly Gln Gln Gln	
115 120 125	
ggt caa cca tgg cag gga caa caa gga caa cag ggt cag cag gga caa	432
Gly Gln Pro Trp Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln	
130 135 140	
caa ggt caa cag ggt cag cag gga caa caa ggt caa cag ggt cag cag	480
Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln	
145 150 155 160	
ggt caa cag gga cag cag ggt cag cag cag caa ggg ttc cgt gac atg	528
Gly Gln Gln Gly Gln Gln Gly Gln Gln Gln Gln Gly Phe Arg Asp Met	
165 170 175	
cac cag aag gtc gaa cat gtt cga cat gga gac atc att gcc att act	576
His Gln Lys Val Glu His Val Arg His Gly Asp Ile Ile Ala Ile Thr	
180 185 190	
gca ggc tct tcc cat tgg atc tac aac acc ggt gac cag cca ctt gtc	624
Ala Gly Ser Ser His Trp Ile Tyr Asn Thr Gly Asp Gln Pro Leu Val	
195 200 205	
att atc tgc ctt ctc gac att gcc aac tac caa aac caa ctc gac cgc	672
Ile Ile Cys Leu Leu Asp Ile Ala Asn Tyr Gln Asn Gln Leu Asp Arg	
210 215 220	
aac cca aga acg ttc cgt ctg gcc gga aac aac cca cag ggc ggt tcc	720
Asn Pro Arg Thr Phe Arg Leu Ala Gly Asn Asn Pro Gln Gly Gly Ser	
225 230 235 240	
cag cag cag cag caa caa caa cag aac atg ttg agc ggg ttc gac cct	768
Gln Gln Gln Gln Gln Gln Gln Gln Asn Met Leu Ser Gly Phe Asp Pro	
245 250 255	
cag gtc cta gcc cag gca ttg aaa atc gac gtt agg ttg gct cag gag	816
Gln Val Leu Ala Gln Ala Leu Lys Ile Asp Val Arg Leu Ala Gln Glu	
260 265 270	
ctt cag aac caa caa gac agc aga gga aac atc gtt cgt gtt aag gga	864
Leu Gln Asn Gln Gln Asp Ser Arg Gly Asn Ile Val Arg Val Lys Gly	
275 280 285	
cct ttc cag gtt gtg agg ccg cct ctt aga cag cca tac gag agt gag	912
Pro Phe Gln Val Val Arg Pro Pro Leu Arg Gln Pro Tyr Glu Ser Glu	
290 295 300	
cag tgg aga cac ccc cgt ggc cca cca caa agc cca caa gac aac ggc	960
Gln Trp Arg His Pro Arg Gly Pro Pro Gln Ser Pro Gln Asp Asn Gly	
305 310 315 320	
ttg gag gag act atc tgc agc atg agg acc cac gag aac att gat gac	1008
Leu Glu Glu Thr Ile Cys Ser Met Arg Thr His Glu Asn Ile Asp Asp	
325 330 335	
cca gcc cgt gct gac gtg tat aag ccc aac ctc ggc cgt gtg act agc	1056
Pro Ala Arg Ala Asp Val Tyr Lys Pro Asn Leu Gly Arg Val Thr Ser	
340 345 350	
gtc aac agc tac act tta ccc atc ttg cag tat atc aga ctc agc gcc	1104
Val Asn Ser Tyr Thr Leu Pro Ile Leu Gln Tyr Ile Arg Leu Ser Ala	
355 360 365	

20

acc cgt ggc att ctc cag ggt aat gcg atg gtg ctt ccg aaa tac aac 1152
 Thr Arg Gly Ile Leu Gln Gly Asn Ala Met Val Leu Pro Lys Tyr Asn
 370 375 380

atg aac gcg aac gag atc ttg tac tgc act caa gga caa gca agg att 1200
 Met Asn Ala Asn Glu Ile Leu Tyr Cys Thr Gln Gly Gln Ala Arg Ile
 385 390 395 400

caa gtg gtg aac gac aac gga cag aac gtg ctg gac cag cag gtg cag 1248
 Gln Val Val Asn Asp Asn Gly Gln Asn Val Leu Asp Gln Gln Val Gln
 405 410 415

aag gga cag ctc gtg gtc atc cca caa gga ttc gcc tat gtt gtc cag 1296
 Lys Gly Gln Leu Val Val Ile Pro Gln Gly Phe Ala Tyr Val Val Gln
 420 425 430

tcc cac caa aac aac ttc gaa tgg att tct ttc aag aca aac gct aac 1344
 Ser His Gln Asn Asn Phe Glu Trp Ile Ser Phe Lys Thr Asn Ala Asn
 435 440 445

gcg atg gtc agc act ttg gcc ggt aga acc tcg gcc ttg agg gca ttg 1392
 Ala Met Val Ser Thr Leu Ala Gly Arg Thr Ser Ala Leu Arg Ala Leu
 450 455 460

cca cta gag gtc ata acc aac gct ttc caa att tct ctc gag gaa gct 1440
 Pro Leu Glu Val Ile Thr Asn Ala Phe Gln Ile Ser Leu Glu Glu Ala
 465 470 475 480

aga agg atc aag ttc aac acg ctt gag acc act ttg act cgt gcg cgc 1488
 Arg Arg Ile Lys Phe Asn Thr Leu Glu Thr Thr Leu Thr Arg Ala Arg
 485 490 495

ggt gga caa ccc cag ttg atc gag gag ata gtc gag gct taa 1530
 Gly Gly Gln Pro Gln Leu Ile Glu Glu Ile Val Glu Ala
 500 505

<210> 18
 <211> 509
 <212> PRT
 <213> Brassica napus

<400> 18
 Met Val Lys Val Pro His Leu Leu Val Ala Thr Phe Gly Val Leu Leu
 1 5 10 15
 Val Leu Asn Gly Cys Leu Ala Arg Gln Ser Leu Gly Val Pro Pro Gln
 20 25 30
 Leu Gly Asn Ala Cys Asn Leu Asp Asn Leu Asp Val Leu Gln Pro Thr
 35 40 45
 Glu Thr Ile Lys Ser Glu Ala Gly Arg Val Glu Tyr Trp Asp His Asn
 50 55 60
 Asn Pro Gln Ile Arg Cys Ala Gly Val Ser Val Ser Arg Val Ile Ile
 65 70 75 80
 Glu Gln Gly Gly Leu Tyr Leu Pro Thr Phe Phe Ser Ser Pro Lys Ile
 85 90 95
 Ser Tyr Val Val Gln Gly Met Gly Ile Ser Gly Arg Val Val Pro Gly
 100 105 110
 Cys Ala Glu Thr Phe Met Asp Ser Gln Pro Met Gln Gly Gln Gln Gln
 115 120 125
 Gly Gln Pro Trp Gln Gly Gln Gln Gly Gln Gln Gly Gln Gly Gln
 130 135 140

21

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

<210> 19
 <211> 1488
 <212> DNA
 <213> Glycine max
 <220>
 <221> CDS
 <222> (1)..(1485)
 <223> Gycinin A-1a-B-x subunit
 <400> 19

atg gcc aag cta gtt ttt tcc ctt tgt ttt ctg ctt ttc agt ggc tgc	48
Met Ala Lys Leu Val Phe Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys	
1 5 10 15	
tgc ttc gct ttc agt tcc aga gag cag cct cag caa aac gag tgc cag	96
Cys Phe Ala Phe Ser Ser Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln	
20 25 30	
atc caa aaa ctc aat gcc ctc aaa ccg gat aac cgt ata gag tca gaa	144
Ile Gln Lys Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu	
35 40 45	
gga ggg ctc att gag aca tgg aac cct aac aac aag cca ttc cag tgt	192
Gly Gly Leu Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys	
50 55 60	
gcc ggt gtt gcc ctc tct cgc tgc acc ctc aac cgc aac gcc ctt cgt	240
Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg	
65 70 75 80	
aga cct tcc tac acc aac ggt ccc cag gaa atc tac atc caa caa ggt	288
Arg Pro Ser Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly	
85 90 95	
aag ggt att ttt ggc atg ata tac ccg ggt tgt cct agc aca ttt gaa	336
Lys Gly Ile Phe Gly Met Ile Tyr Pro Gly Cys Pro Ser Thr Phe Glu	
100 105 110	
gag cct caa caa cct caa caa aga gga caa agc agc aga cca caa gac	384
Glu Pro Gln Gln Pro Gln Gln Arg Gly Gln Ser Ser Arg Pro Gln Asp	
115 120 125	
cgt cac cag aag atc tat aac ttc aga gag ggt gat ttg atc gca gtg	432
Arg His Gln Lys Ile Tyr Asn Phe Arg Glu Gly Asp Leu Ile Ala Val	
130 135 140	
cct act ggt gtt gca tgg tgg atg tac aac aat gaa gac act cct gtt	480
Pro Thr Gly Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val	
145 150 155 160	
gtt gcc gtt tct att att gac acc aac agc ttg gag aac cag ctc gac	528
Val Ala Val Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp	
165 170 175	
cag atg cct agg aga ttc tat ctt gct ggg aac caa gag caa gag ttt	576
Gln Met Pro Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe	
180 185 190	
cta aaa tat cag caa gag caa gga ggt cat caa agc cag aaa gga aag	624
Leu Lys Tyr Gln Gln Glu Gln Gly Gly His Gln Ser Gln Lys Gly Lys	
195 200 205	
cat cag caa gaa gaa gaa aac gaa gga ggc agc ata ttg agt ggc ttc	672
His Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe	
210 215 220	

23

acc ctg gaa ttc ttg gaa cat gca ttc agc gtg gac aag cag ata gcg	720
Thr Leu Glu Phe Leu Glu His Ala Phe Ser Val Asp Lys Gln Ile Ala	
225 230 235 240	
aaa aac cta caa gga gag aac gaa ggg gaa gac aag gga gcc att gtg	768
Lys Asn Leu Gln Gly Glu Asn Glu Gly Glu Asp Lys Gly Ala Ile Val	
245 250 255	
aca gtg aaa gga ggt ctg agc gtg ata aaa cca ccc acg gac gag cag	816
Thr Val Lys Gly Gly Leu Ser Val Ile Lys Pro Pro Thr Asp Glu Gln	
260 265 270	
caa caa aga ccc cag gaa gag gaa gaa gaa gag gat gag aag cca	864
Gln Gln Arg Pro Gln Glu Glu Glu Glu Glu Glu Asp Glu Lys Pro	
275 280 285	
cag tgc aag ggt aaa gac aaa cac tgc caa cgc ccc cga gga agc caa	912
Gln Cys Lys Gly Lys Asp Lys His Cys Gln Arg Pro Arg Gly Ser Gln	
290 295 300	
agc aaa agc aga aga aat ggc att gac gag acc ata tgc acc atg aga	960
Ser Lys Ser Arg Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr Met Arg	
305 310 315 320	
ctt cgc cac aac att ggc cag act tca tca cct gac atc tac aac cct	1008
Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Tyr Asn Pro	
325 330 335	
caa gcc ggt agc gtc aca acc gcc acc agc ctt gac ttc cca gcc ctc	1056
Gln Ala Gly Ser Val Thr Thr Ala Thr Ser Leu Asp Phe Pro Ala Leu	
340 345 350	
tcg tgg ctc aga ctc agt gct gag ttt gga tct ctc cgc aag aat gca	1104
Ser Trp Leu Arg Leu Ser Ala Glu Phe Gly Ser Leu Arg Lys Asn Ala	
355 360 365	
atg ttc gtg cca cac tac aac ctg aac gcg aac agc ata ata tac gca	1152
Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile Tyr Ala	
370 375 380	
ttg aat gga cgg gca ttg ata caa gtg gtg aat tgc aac ggt gag aga	1200
Leu Asn Gly Arg Ala Leu Ile Gln Val Val Asn Cys Asn Gly Glu Arg	
385 390 395 400	
gtg ttt gat gga gag ctg caa gag gga cgg gtg ctg atc gtg cca caa	1248
Val Phe Asp Gly Glu Leu Gln Glu Gly Arg Val Leu Ile Val Pro Gln	
405 410 415	
aac ttt gtg gtg gct gca aga tca cag agt gac aac ttc gag tat gtg	1296
Asn Phe Val Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu Tyr Val	
420 425 430	
tca ttc aag acc aat gat aca ccc atg atc ggc act ctt gca ggg gca	1344
Ser Phe Lys Thr Asn Asp Thr Pro Met Ile Gly Thr Leu Ala Gly Ala	
435 440 445	
aac tca ttg ttg aac gca tta cca gag gaa gtg att cag cac act ttc	1392
Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln His Thr Phe	
450 455 460	
aac cta aaa agc cag cag gcc agg cag ata aag aac aac aac cct ttc	1440
Asn Leu Lys Ser Gln Gln Ala Arg Gln Ile Lys Asn Asn Asn Pro Phe	
465 470 475 480	
aag ttc ctg gtt cca cct cag gag tct cag aag aga gct gtg gct tag	1488
Lys Phe Leu Val Pro Pro Gln Glu Ser Gln Lys Arg Ala Val Ala	
485 490 495	

<210> 20
 <211> 495
 <212> PRT
 <213> Glycine max
 <400> 20
 Met Ala Lys Leu Val Phe Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys
 1 5 10 15
 Cys Phe Ala Phe Ser Ser Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln
 20 25 30
 Ile Gln Lys Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu
 35 40 45
 Gly Gly Leu Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys
 50 55 60
 Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg
 65 70 75 80
 Arg Pro Ser Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly
 85 90 95
 Lys Gly Ile Phe Gly Met Ile Tyr Pro Gly Cys Pro Ser Thr Phe Glu
 100 105 110
 Glu Pro Gln Gln Pro Gln Gln Arg Gly Gln Ser Ser Arg Pro Gln Asp
 115 120 125
 Arg His Gln Lys Ile Tyr Asn Phe Arg Glu Gly Asp Leu Ile Ala Val
 130 135 140
 Pro Thr Gly Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val
 145 150 155 160
 Val Ala Val Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp
 165 170 175
 Gln Met Pro Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe
 180 185 190
 Leu Lys Tyr Gln Gln Glu Gln Gly Gly His Gln Ser Gln Lys Gly Lys
 195 200 205
 His Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe
 210 215 220
 Thr Leu Glu Phe Leu Glu His Ala Phe Ser Val Asp Lys Gln Ile Ala
 225 230 235 240
 Lys Asn Leu Gln Gly Glu Asn Glu Gly Glu Asp Lys Gly Ala Ile Val
 245 250 255
 Thr Val Lys Gly Gly Leu Ser Val Ile Lys Pro Pro Thr Asp Glu Gln
 260 265 270
 Gln Gln Arg Pro Gln Glu Glu Glu Glu Glu Glu Asp Glu Lys Pro
 275 280 285
 Gln Cys Lys Gly Lys Asp Lys His Cys Gln Arg Pro Arg Gly Ser Gln
 290 295 300
 Ser Lys Ser Arg Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr Met Arg
 305 310 315 320
 Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Tyr Asn Pro
 325 330 335
 Gln Ala Gly Ser Val Thr Thr Ala Thr Ser Leu Asp Phe Pro Ala Leu
 340 345 350

25

Ser Trp Leu Arg Leu Ser Ala Glu Phe Gly Ser Leu Arg Lys Asn Ala
 355 360 365

Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile Tyr Ala
 370 375 380

Leu Asn Gly Arg Ala Leu Ile Gln Val Val Asn Cys Asn Gly Glu Arg
 385 390 395 400

Val Phe Asp Gly Glu Leu Gln Glu Gly Arg Val Leu Ile Val Pro Gln
 405 410 415

Asn Phe Val Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu Tyr Val
 420 425 430

Ser Phe Lys Thr Asn Asp Thr Pro Met Ile Gly Thr Leu Ala Gly Ala
 435 440 445

Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln His Thr Phe
 450 455 460

Asn Leu Lys Ser Gln Gln Ala Arg Gln Ile Lys Asn Asn Asn Pro Phe
 465 470 475 480

Lys Phe Leu Val Pro Pro Gln Glu Ser Gln Lys Arg Ala Val Ala
 485 490 495

<210> 21

<211> 1458

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(1455)

<223> glycinin G2 subunit

<400> 21

atg gcc aag ctt gtt ctt tcc ctt tgt ttc ctt ctt ttc agt ggc tgc	48
Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys	
1 5 10 15	
ttc gct ctg aga gag cag gca cag caa aat gag tgc cag atc caa aag	96
Phe Ala Leu Arg Glu Gln Ala Gln Gln Asn Glu Cys Gln Ile Gln Lys	
20 25 30	
ctg aat gcc ctc aaa ccg gat aac cgt ata gag tcg gaa ggt ggg ttc	144
Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu Gly Gly Phe	
35 40 45	
att gag aca tgg aac cct aac aac aag cca ttc cag tgt gcc ggt gtt	192
Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys Ala Gly Val	
50 55 60	
gcc ctc tct cgc tgc acc ctt aac cgc aat gcc ctt cgt aga cct tcc	240
Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg Arg Pro Ser	
65 70 75 80	
tac acc aac ggt ccc cag gaa atc tac ata caa caa ggt aat ggt att	288
Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly Asn Gly Ile	
85 90 95	
ttt ggc atg ata ttc ccg ggt tgt cct agc act tat caa gag ccg caa	336
Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Tyr Gln Glu Pro Gln	
100 105 110	

26

gaa tct cag caa cga gga cga agc cag agg ccc caa gac cgt cac caa	384
Glu Ser Gln Gln Arg Gly Arg Ser Gln Arg Pro Gln Asp Arg His Gln	
115 120 125	
aag gta cat cgc ttc aga gag ggt gat ttg atc gca gtg cct act ggt	432
Lys Val His Arg Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly	
130 135 140	
gtt gca tgg tgg atg tac aac aat gaa gac act cct gtt gtt gcc gtt	480
Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val	
145 150 155 160	
tct att att gac acc aac agc ttg gag aac cag ctc gac cag atg cct	528
Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp Gln Met Pro	
165 170 175	
agg aga ttc tat ctt gct ggg aac caa gag caa gag ttt cta aaa tat	576
Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Lys Tyr	
180 185 190	
cag cag cag cag caa gga ggt tcc caa agc cag aaa gga aag caa caa	624
Gln Gln Gln Gln Gln Gly Gly Ser Gln Ser Gln Lys Gly Lys Gln Gln	
195 200 205	
gaa gaa gaa aac gaa gga agc aac ata ttg agt ggc ttc gcc cct gaa	672
Glu Glu Glu Asn Glu Gly Ser Asn Ile Leu Ser Gly Phe Ala Pro Glu	
210 215 220	
ttc ttg aaa gaa gcg ttc ggc gtg aac atg cag ata gtg aga aac cta	720
Phe Leu Lys Glu Ala Phe Gly Val Asn Met Gln Ile Val Arg Asn Leu	
225 230 235 240	
caa ggt gag aac gaa gag gag gat agt gga gcc att gtg aca gtg aaa	768
Gln Gly Glu Asn Glu Glu Glu Asp Ser Gly Ala Ile Val Thr Val Lys	
245 250 255	
gga ggt cta aga gtc aca gct cca gcc atg agg aag cca cag caa gaa	816
Gly Gly Leu Arg Val Thr Ala Pro Ala Met Arg Lys Pro Gln Gln Glu	
260 265 270	
gaa gat gat gat gat gag gaa gag cag cca cag tgc gtg gag aca gac	864
Glu Asp Asp Asp Asp Glu Glu Glu Gln Pro Gln Cys Val Glu Thr Asp	
275 280 285	
aaa ggt tgc caa cgc caa agc aaa agg agc aga aat ggc att gat gag	912
Lys Gly Cys Gln Arg Gln Ser Lys Arg Ser Arg Asn Gly Ile Asp Glu	
290 295 300	
acc att tgc aca atg aga ctt cgc caa aac att ggt cag aat tca tca	960
Thr Ile Cys Thr Met Arg Leu Arg Gln Asn Ile Gly Gln Asn Ser Ser	
305 310 315 320	
cct gac atc tac aac cct caa gct ggt agc atc aca acc gcc acc agc	1008
Pro Asp Ile Tyr Asn Pro Gln Ala Gly Ser Ile Thr Thr Ala Thr Ser	
325 330 335	
ctt gac ttc cca gcc ctc tgg ctt ctc aaa ctc agt gcc cag tat gga	1056
Leu Asp Phe Pro Ala Leu Trp Leu Leu Lys Leu Ser Ala Gln Tyr Gly	
340 345 350	
tca ctc cgc aag aat gct atg ttc gtg cca cac tac acc ctg aac gcg	1104
Ser Leu Arg Lys Asn Ala Met Phe Val Pro His Tyr Thr Leu Asn Ala	
355 360 365	
aac agc ata ata tac gca ttg aat ggg cgg gca ttg gta caa gtg gtg	1152
Asn Ser Ile Ile Tyr Ala Leu Asn Gly Arg Ala Leu Val Gln Val Val	
370 375 380	

27

aat tgc aat ggt gag aga gtg ttt gat gga gag ctg caa gag gga ggg 1200
 Asn Cys Asn Gly Glu Arg Val Phe Asp Gly Glu Leu Gln Glu Gly Gly
 385 390 395 400
 gtg ctg atc gtt cca caa aac ttt gcg gtg gct gca aaa tcc cag agc 1248
 Val Leu Ile Val Pro Gln Asn Phe Ala Val Ala Ala Lys Ser Gln Ser
 405 410 415
 gat aac ttt gag tat gtg tca ttc aag acc aat gat aga ccc tcg atc 1296
 Asp Asn Phe Glu Tyr Val Ser Phe Lys Thr Asn Asp Arg Pro Ser Ile
 420 425 430
 gga aac ctt gca ggg gca aac tca ttg ttg aac gca ttg cca gag gaa 1344
 Gly Asn Leu Ala Gly Ala Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu
 435 440 445
 gtg att cag cac act ttt aac cta aag agc cag cag gcc agg cag gtg 1392
 Val Ile Gln His Thr Phe Asn Leu Lys Ser Gln Gln Ala Arg Gln Val
 450 455 460
 aag aac aac aac cct ttc agc ttc ctt gtt cca cct cag gag tct cag 1440
 Lys Asn Asn Asn Pro Phe Ser Phe Leu Val Pro Pro Gln Glu Ser Gln
 465 470 475 480
 agg aga gct gtg gct tag 1458
 Arg Arg Ala Val Ala
 485

<210> 22

<211> 485

<212> PRT

<213> Glycine max

<400> 22

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

28

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

<210> 23

<211> 1689

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(1686)

<223> glycinin A5A4B3 subunits

<400> 23

atg ggg aag ccc ttc act ctc tct ctt tct tcc ctt tgc ttg cta ctc 48

29

Met	Gly	Lys	Pro	Phe	Thr	Leu	Ser	Leu	Ser	Ser	Leu	Cys	Leu	Leu	Leu	
1				5					10					15		
ttg	tcg	agt	gca	tgc	ttt	gct	att	agc	tcc	agc	aag	ctc	aac	gag	tgc	96
Leu	Ser	Ser	Ala	Cys	Phe	Ala	Ile	Ser	Ser	Ser	Lys	Leu	Asn	Glu	Cys	
			20					25					30			
caa	ctc	aac	aac	ctc	aac	gcg	ttg	gaa	ccc	gac	cac	cgc	gtt	gag	tcc	144
Gln	Leu	Asn	Asn	Leu	Asn	Ala	Leu	Glu	Pro	Asp	His	Arg	Val	Glu	Ser	
		35					40					45				
gaa	ggt	ggt	ttg	att	caa	aca	tgg	aac	tct	caa	cac	cct	gag	ctg	aaa	192
Glu	Gly	Gly	Leu	Ile	Gln	Thr	Trp	Asn	Ser	Gln	His	Pro	Glu	Leu	Lys	
	50					55				60						
tgc	gcc	ggt	gtc	act	gtt	tcc	aaa	ctc	acc	ctc	aac	cgc	aat	ggc	ctc	240
Cys	Ala	Gly	Val	Thr	Val	Ser	Lys	Leu	Thr	Leu	Asn	Arg	Asn	Gly	Leu	
65					70					75					80	
cac	tcg	cca	tct	tac	tca	cct	tat	ccc	cgg	atg	atc	atc	atc	gcc	caa	288
His	Ser	Pro	Ser	Tyr	Ser	Pro	Tyr	Pro	Arg	Met	Ile	Ile	Ile	Ala	Gln	
				85				90						95		
ggg	aaa	gga	gca	ctt	gga	gtt	gca	att	cca	gga	tgt	cct	gag	acg	ttt	336
Gly	Lys	Gly	Ala	Leu	Gly	Val	Ala	Ile	Pro	Gly	Cys	Pro	Glu	Thr	Phe	
			100					105					110			
gag	gag	cca	caa	gaa	caa	tca	aac	aga	aga	ggc	tca	agg	tcg	cag	aag	384
Glu	Glu	Pro	Gln	Glu	Gln	Ser	Asn	Arg	Arg	Gly	Ser	Arg	Ser	Gln	Lys	
		115					120					125				
cag	cag	cta	cag	gac	agt	cac	cag	aag	att	cgt	cac	ttc	aat	gaa	gga	432
Gln	Gln	Leu	Gln	Asp	Ser	His	Gln	Lys	Ile	Arg	His	Phe	Asn	Glu	Gly	
	130					135				140						
gac	gta	ctc	gtg	att	cct	cct	agt	gtt	cct	tac	tgg	acc	tat	aac	act	480
Asp	Val	Leu	Val	Ile	Pro	Pro	Ser	Val	Pro	Tyr	Trp	Thr	Tyr	Asn	Thr	
145					150					155					160	
ggc	gat	gaa	cca	gtt	gtt	gcc	atc	agt	ctt	ctt	gac	acc	tct	aac	ttc	528
Gly	Asp	Glu	Pro	Val	Val	Ala	Ile	Ser	Leu	Leu	Asp	Thr	Ser	Asn	Phe	
				165				170						175		
aat	aac	cag	ctt	gat	caa	acc	cct	agg	gta	ttt	tac	ctt	gct	ggg	aac	576
Asn	Asn	Gln	Leu	Asp	Gln	Thr	Pro	Arg	Val	Phe	Tyr	Leu	Ala	Gly	Asn	
		180						185					190			
cca	gat	ata	gag	tac	cca	gag	acc	atg	caa	caa	caa	caa	cag	cag	aaa	624
Pro	Asp	Ile	Glu	Tyr	Pro	Glu	Thr	Met	Gln	Gln	Gln	Gln	Gln	Gln	Lys	
		195					200					205				
agt	cat	ggt	gga	cgc	aag	cag	ggg	caa	cac	cag	cag	gag	gaa	gag	gaa	672
Ser	His	Gly	Gly	Arg	Lys	Gln	Gly	Gln	His	Gln	Gln	Glu	Glu	Glu	Glu	
	210					215				220						
gaa	ggt	ggc	agc	gtg	ctc	agt	ggc	ttc	agc	aaa	cac	ttc	ttg	gca	caa	720
Glu	Gly	Gly	Ser	Val	Leu	Ser	Gly	Phe	Ser	Lys	His	Phe	Leu	Ala	Gln	
225					230					235					240	
tcc	ttc	aac	acc	aac	gag	gac	ata	gct	gag	aaa	ctt	gag	tct	cca	gac	768
Ser	Phe	Asn	Thr	Asn	Glu	Asp	Ile	Ala	Glu	Lys	Leu	Glu	Ser	Pro	Asp	
				245					250					255		
gac	gaa	agg	aag	cag	atc	gtg	aca	gtg	gaa	gga	ggt	ctc	agc	gtt	atc	816
Asp	Glu	Arg	Lys	Gln	Ile	Val	Thr	Val	Glu	Gly	Gly	Leu	Ser	Val	Ile	
			260					265					270			

30

agc ccc aag tgg caa gaa caa caa gat gaa gat gaa gat gaa gac gaa	864
Ser Pro Lys Trp Gln Glu Gln Gln Asp Glu Asp Glu Asp Glu Asp Glu	
275 280 285	
gat gat gaa gat gaa caa att ccc tct cac cct cct cgc cga cca agc	912
Asp Asp Glu Asp Glu Gln Ile Pro Ser His Pro Pro Arg Arg Pro Ser	
290 295 300	
cat gga aag cgt gaa caa gac gag gac gag gac gaa gat gaa gat aaa	960
His Gly Lys Arg Glu Gln Asp Glu Asp Glu Asp Glu Asp Glu Asp Lys	
305 310 315 320	
cct cgt cct agt cga cca agc caa gga aag cgg aac aag aca gga cag	1008
Pro Arg Pro Ser Arg Pro Ser Gln Gly Lys Arg Asn Lys Thr Gly Gln	
325 330 335	
gac gag gac gaa gat gaa gat gaa gat caa cct cgc aag agc cgc gaa	1056
Asp Glu Asp Glu Asp Glu Asp Glu Asp Gln Pro Arg Lys Ser Arg Glu	
340 345 350	
tgg aga tcg aaa aag aca caa ccc aga aga cct aga caa gaa gaa cca	1104
Trp Arg Ser Lys Lys Thr Gln Pro Arg Arg Pro Arg Gln Glu Glu Pro	
355 360 365	
cgt gaa aga gga tgc gag aca aga aac ggg gtt gag gaa aat atc tgc	1152
Arg Glu Arg Gly Cys Glu Thr Arg Asn Gly Val Glu Glu Asn Ile Cys	
370 375 380	
acc ttg aag ctt cac gag aac att gct cgc cct tca cgc gct gac ttc	1200
Thr Leu Lys Leu His Glu Asn Ile Ala Arg Pro Ser Arg Ala Asp Phe	
385 390 395 400	
tac aac cct aaa gct ggt cgc att agt acc ctc aac agc ctc acc ctc	1248
Tyr Asn Pro Lys Ala Gly Arg Ile Ser Thr Leu Asn Ser Leu Thr Leu	
405 410 415	
cca gcc ctc cgc caa ttc caa ctc agt gcc caa tat gtt gtc ctc tac	1296
Pro Ala Leu Arg Gln Phe Gln Leu Ser Ala Gln Tyr Val Val Leu Tyr	
420 425 430	
aag aat gga att tac tct cca cat tgg aat ctg aat gca aac agt gtg	1344
Lys Asn Gly Ile Tyr Ser Pro His Trp Asn Leu Asn Ala Asn Ser Val	
435 440 445	
atc tat gtg act cga gga caa gga aag gtt aga gtt gtg aac tgc caa	1392
Ile Tyr Val Thr Arg Gly Gln Gly Lys Val Arg Val Val Asn Cys Gln	
450 455 460	
ggg aat gca gtg ttc gac ggt gag ctt agg agg gga caa ttg ctg gtg	1440
Gly Asn Ala Val Phe Asp Gly Glu Leu Arg Arg Gly Gln Leu Leu Val	
465 470 475 480	
gta cca cag aac ttc gtg gtg gcg gag caa gcc gga gaa caa gga ttc	1488
Val Pro Gln Asn Phe Val Val Ala Glu Gln Ala Gly Glu Gln Gly Phe	
485 490 495	
gaa tac ata gta ttc aag aca cac cac aac gca gtc act agc tac ttg	1536
Glu Tyr Ile Val Phe Lys Thr His His Asn Ala Val Thr Ser Tyr Leu	
500 505 510	
aag gat gtg ttt agg gca att ccc tca gag gtt ctt gcc cat tct tac	1584
Lys Asp Val Phe Arg Ala Ile Pro Ser Glu Val Leu Ala His Ser Tyr	
515 520 525	
aac ctt cga cag agt caa gtg tct gag ctt aag tat gaa gga aat tgg	1632
Asn Leu Arg Gln Ser Gln Val Ser Glu Leu Lys Tyr Glu Gly Asn Trp	
530 535 540	

31

ggt cct ttg gtc aac cct gag tct caa caa ggc tca ccc cgt gtt aaa 1680
 Gly Pro Leu Val Asn Pro Glu Ser Gln Gln Gly Ser Pro Arg Val Lys
 545 550 555 560

gtc gca taa 1689
 Val Ala

<210> 24

<211> 562

<212> PRT

<213> Glycine max

<400> 24

Met Gly Lys Pro Phe Thr Leu Ser Leu Ser Ser Leu Cys Leu Leu Leu
 1 5 10 15

Leu Ser Ser Ala Cys Phe Ala Ile Ser Ser Ser Lys Leu Asn Glu Cys
 20 25 30

Gln Leu Asn Asn Leu Asn Ala Leu Glu Pro Asp His Arg Val Glu Ser
 35 40 45

Glu Gly Gly Leu Ile Gln Thr Trp Asn Ser Gln His Pro Glu Leu Lys
 50 55 60

Cys Ala Gly Val Thr Val Ser Lys Leu Thr Leu Asn Arg Asn Gly Leu
 65 70 75 80

His Ser Pro Ser Tyr Ser Pro Tyr Pro Arg Met Ile Ile Ile Ala Gln
 85 90 95

Gly Lys Gly Ala Leu Gly Val Ala Ile Pro Gly Cys Pro Glu Thr Phe
 100 105 110

Glu Glu Pro Gln Glu Gln Ser Asn Arg Arg Gly Ser Arg Ser Gln Lys
 115 120 125

Gln Gln Leu Gln Asp Ser His Gln Lys Ile Arg His Phe Asn Glu Gly
 130 135 140

Asp Val Leu Val Ile Pro Pro Ser Val Pro Tyr Trp Thr Tyr Asn Thr
 145 150 155 160

Gly Asp Glu Pro Val Val Ala Ile Ser Leu Leu Asp Thr Ser Asn Phe
 165 170 175

Asn Asn Gln Leu Asp Gln Thr Pro Arg Val Phe Tyr Leu Ala Gly Asn
 180 185 190

Pro Asp Ile Glu Tyr Pro Glu Thr Met Gln Gln Gln Gln Gln Lys
 195 200 205

Ser His Gly Gly Arg Lys Gln Gly Gln His Gln Gln Glu Glu Glu Glu
 210 215 220

Glu Gly Gly Ser Val Leu Ser Gly Phe Ser Lys His Phe Leu Ala Gln
 225 230 235 240

Ser Phe Asn Thr Asn Glu Asp Ile Ala Glu Lys Leu Glu Ser Pro Asp
 245 250 255

Asp Glu Arg Lys Gln Ile Val Thr Val Glu Gly Gly Leu Ser Val Ile
 260 265 270

Ser Pro Lys Trp Gln Glu Gln Gln Asp Glu Asp Glu Asp Glu Asp Glu
 275 280 285

Asp Asp Glu Asp Glu Gln Ile Pro Ser His Pro Pro Arg Arg Pro Ser
 290 295 300

32

His Gly Lys Arg Glu Gln Asp Glu Asp Glu Asp Glu Asp Glu Asp Lys
 305 310 315 320
 Pro Arg Pro Ser Arg Pro Ser Gln Gly Lys Arg Asn Lys Thr Gly Gln
 325 330 335
 Asp Glu Asp Glu Asp Glu Asp Glu Asp Gln Pro Arg Lys Ser Arg Glu
 340 345 350
 Trp Arg Ser Lys Lys Thr Gln Pro Arg Arg Pro Arg Gln Glu Glu Pro
 355 360 365
 Arg Glu Arg Gly Cys Glu Thr Arg Asn Gly Val Glu Glu Asn Ile Cys
 370 375 380
 Thr Leu Lys Leu His Glu Asn Ile Ala Arg Pro Ser Arg Ala Asp Phe
 385 390 395 400
 Tyr Asn Pro Lys Ala Gly Arg Ile Ser Thr Leu Asn Ser Leu Thr Leu
 405 410 415
 Pro Ala Leu Arg Gln Phe Gln Leu Ser Ala Gln Tyr Val Val Leu Tyr
 420 425 430
 Lys Asn Gly Ile Tyr Ser Pro His Trp Asn Leu Asn Ala Asn Ser Val
 435 440 445
 Ile Tyr Val Thr Arg Gly Gln Gly Lys Val Arg Val Val Asn Cys Gln
 450 455 460
 Gly Asn Ala Val Phe Asp Gly Glu Leu Arg Arg Gly Gln Leu Leu Val
 465 470 475 480
 Val Pro Gln Asn Phe Val Val Ala Glu Gln Ala Gly Glu Gln Gly Phe
 485 490 495
 Glu Tyr Ile Val Phe Lys Thr His His Asn Ala Val Thr Ser Tyr Leu
 500 505 510
 Lys Asp Val Phe Arg Ala Ile Pro Ser Glu Val Leu Ala His Ser Tyr
 515 520 525
 Asn Leu Arg Gln Ser Gln Val Ser Glu Leu Lys Tyr Glu Gly Asn Trp
 530 535 540
 Gly Pro Leu Val Asn Pro Glu Ser Gln Gln Gly Ser Pro Arg Val Lys
 545 550 555 560
 Val Ala

<210> 25

<211> 1551

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(1548)

<223> glycinin A3-B4 subunit

<400> 25

atg ggg aag ccc ttc ttc act ctc tct ctt tct tcc ctt tgc ttg cta 48
 Met Gly Lys Pro Phe Phe Thr Leu Ser Leu Ser Ser Leu Cys Leu Leu
 1 5 10 15

ctc ttg tcg agt gca tgc ttt gct att acc tcc agc aag ttc aac gag 96
 Leu Leu Ser Ser Ala Cys Phe Ala Ile Thr Ser Ser Lys Phe Asn Glu
 20 25 30

33

tgc	caa	ctc	aac	aac	ctc	aac	gcg	ttg	gaa	ccc	gac	cac	cg	ggt	gag	144
Cys	Gln	Leu	Asn	Asn	Leu	Asn	Ala	Leu	Glu	Pro	Asp	His	Arg	Val	Glu	
	35						40					45				
tcc	gaa	ggg	ggg	ctt	att	gaa	aca	tgg	aac	tct	caa	cac	cct	gag	ctg	192
Ser	Glu	Gly	Gly	Leu	Ile	Glu	Thr	Trp	Asn	Ser	Gln	His	Pro	Glu	Leu	
	50					55					60					
caa	tgc	gcc	ggg	gtc	act	ggt	tcc	aaa	cg	acc	ctc	aac	cg	aac	ggc	240
Gln	Cys	Ala	Gly	Val	Thr	Val	Ser	Lys	Arg	Thr	Leu	Asn	Arg	Asn	Gly	
	65				70					75					80	
tcc	cac	ttg	cca	tct	tac	tta	cct	tat	ccc	caa	atg	atc	att	gtc	ggt	288
Ser	His	Leu	Pro	Ser	Tyr	Leu	Pro	Tyr	Pro	Gln	Met	Ile	Ile	Val	Val	
				85					90					95		
caa	ggg	aag	gga	gca	att	gga	ttt	gca	ttt	ccg	gga	tgt	ccc	gag	acg	336
Gln	Gly	Lys	Gly	Ala	Ile	Gly	Phe	Ala	Phe	Pro	Gly	Cys	Pro	Glu	Thr	
		100					105					110				
ttt	gag	aag	cca	caa	caa	caa	tca	agc	aga	aga	ggc	tca	agg	tca	cag	384
Phe	Glu	Lys	Pro	Gln	Gln	Gln	Ser	Ser	Arg	Arg	Gly	Ser	Arg	Ser	Gln	
	115					120					125					
cag	caa	cta	caa	gac	agt	cac	cag	aag	att	cgt	cac	ttc	aat	gaa	gga	432
Gln	Gln	Leu	Gln	Asp	Ser	His	Gln	Lys	Ile	Arg	His	Phe	Asn	Glu	Gly	
	130					135					140					
gac	gta	cta	gtg	att	cct	ctt	ggg	ggt	cct	tac	tgg	acc	tat	aac	act	480
Asp	Val	Leu	Val	Ile	Pro	Leu	Gly	Val	Pro	Tyr	Trp	Thr	Tyr	Asn	Thr	
	145				150					155					160	
ggc	gat	gaa	cca	ggt	ggt	gcc	atc	agt	cct	ctt	gac	acc	tcc	aac	ttc	528
Gly	Asp	Glu	Pro	Val	Val	Ala	Ile	Ser	Pro	Leu	Asp	Thr	Ser	Asn	Phe	
				165					170					175		
aac	aat	cag	ctt	gat	caa	aac	ccc	aga	gta	ttt	tac	ctt	gct	ggg	aac	576
Asn	Asn	Gln	Leu	Asp	Gln	Asn	Pro	Arg	Val	Phe	Tyr	Leu	Ala	Gly	Asn	
		180					185					190				
cca	gat	ata	gag	cat	ccc	gag	acc	atg	caa	caa	cag	cag	cag	cag	aag	624
Pro	Asp	Ile	Glu	His	Pro	Glu	Thr	Met	Gln	Gln	Gln	Gln	Gln	Gln	Lys	
	195					200					205					
agt	cat	ggg	gga	cg	aag	cag	ggg	caa	cac	cga	cag	cag	cag	gag	gaa	672
Ser	His	Gly	Gly	Arg	Lys	Gln	Gly	Gln	His	Arg	Gln	Gln	Glu	Glu	Glu	
	210					215					220					
ggg	ggc	agt	gtg	ctc	agt	ggc	ttc	agc	aaa	cat	ttc	tta	gca	caa	tcc	720
Gly	Gly	Ser	Val	Leu	Ser	Gly	Phe	Ser	Lys	His	Phe	Leu	Ala	Gln	Ser	
	225				230					235					240	
ttc	aac	acc	aac	gag	gac	aca	gct	gag	aaa	ctt	cgg	tct	cca	gat	gac	768
Phe	Asn	Thr	Asn	Glu	Asp	Thr	Ala	Glu	Lys	Leu	Arg	Ser	Pro	Asp	Asp	
				245					250					255		
gaa	agg	aag	cag	atc	gtg	aca	gtg	gag	gga	ggc	ctc	agc	ggt	atc	agc	816
Glu	Arg	Lys	Gln	Ile	Val	Thr	Val	Glu	Gly	Gly	Leu	Ser	Val	Ile	Ser	
		260					265					270				
ccc	aag	tgg	caa	gaa	caa	gaa	gac	gaa	gac	gaa	gac	gaa	gac	gaa	gaa	864
Pro	Lys	Trp	Gln	Glu	Gln	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Glu	
		275					280					285				
tat	gga	cgg	acg	ccc	tct	tat	cct	cca	cga	cga	cca	agc	cat	gga	aag	912
Tyr	Gly	Arg	Thr	Pro	Ser	Tyr	Pro	Pro	Arg	Arg	Pro	Ser	His	Gly	Lys	
	290					295					300					

34

cat gaa gat gac gag gac gag gac gaa gaa gaa gat caa cct cgt cct 960
 His Glu Asp Asp Glu Asp Glu Asp Glu Glu Glu Asp Gln Pro Arg Pro
 305 310 315 320
 gat cac cct cca cag cga cca agc agg ccc gaa caa caa gaa cca cgt 1008
 Asp His Pro Pro Gln Arg Pro Ser Arg Pro Glu Gln Gln Glu Pro Arg
 325 330 335
 gga aga gga tgt cag act aga aat ggg gtt gag gaa aat att tgc acc 1056
 Gly Arg Gly Cys Gln Thr Arg Asn Gly Val Glu Glu Asn Ile Cys Thr
 340 345 350
 atg aag ctt cac gag aac att gct cgc cct tca cgt gct gac ttc tac 1104
 Met Lys Leu His Glu Asn Ile Ala Arg Pro Ser Arg Ala Asp Phe Tyr
 355 360 365
 aac cca aaa gct ggt cgc att agc acc ctc aac agt ctc acc ctc cca 1152
 Asn Pro Lys Ala Gly Arg Ile Ser Thr Leu Asn Ser Leu Thr Leu Pro
 370 375 380
 gcc ctc cgc caa ttc gga ctc agt gcc caa tat gtt gtc ctc tac agg 1200
 Ala Leu Arg Gln Phe Gly Leu Ser Ala Gln Tyr Val Val Leu Tyr Arg
 385 390 395 400
 aat gga att tac tct cca gat tgg aac ttg aac gcg aac agt gtg acg 1248
 Asn Gly Ile Tyr Ser Pro Asp Trp Asn Leu Asn Ala Asn Ser Val Thr
 405 410 415
 atg act cga ggg aaa gga aga gtt aga gtg gtg aac tgc caa ggg aat 1296
 Met Thr Arg Gly Lys Gly Arg Val Arg Val Val Asn Cys Gln Gly Asn
 420 425 430
 gca gtg ttc gac ggt gag cta agg agg gga caa ttg cta gtg gtg ccg 1344
 Ala Val Phe Asp Gly Glu Leu Arg Arg Gly Gln Leu Leu Val Val Pro
 435 440 445
 cag aac ccc gcg gtg gct gag caa ggg gga gaa caa gga ttg gaa tat 1392
 Gln Asn Pro Ala Val Ala Glu Gln Gly Gly Glu Gln Gly Leu Glu Tyr
 450 455 460
 gta gtg ttc aag aca cac cac aac gcc gtg agc agc tac att aag gat 1440
 Val Val Phe Lys Thr His His Asn Ala Val Ser Ser Tyr Ile Lys Asp
 465 470 475 480
 gtg ttt agg gta atc cct tcg gag gtt ctt tcc aat tct tac aac ctt 1488
 Val Phe Arg Val Ile Pro Ser Glu Val Leu Ser Asn Ser Tyr Asn Leu
 485 490 495
 ggc cag agt caa gtg cgt cag ctc aag tat caa gga aac tcc ggc cct 1536
 Gly Gln Ser Gln Val Arg Gln Leu Lys Tyr Gln Gly Asn Ser Gly Pro
 500 505 510
 ttg gtc aac cca taa 1551
 Leu Val Asn Pro
 515
 <210> 26
 <211> 516
 <212> PRT
 <213> Glycine max
 <400> 26
 Met Gly Lys Pro Phe Phe Thr Leu Ser Leu Ser Ser Leu Cys Leu Leu
 1 5 10 15
 Leu Leu Ser Ser Ala Cys Phe Ala Ile Thr Ser Ser Lys Phe Asn Glu
 20 25 30

35

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

36

Met Thr Arg Gly Lys Gly Arg Val Arg Val Val Asn Cys Gln Gly Asn
 420 425 430
 Ala Val Phe Asp Gly Glu Leu Arg Arg Gly Gln Leu Leu Val Val Pro
 435 440 445
 Gln Asn Pro Ala Val Ala Glu Gln Gly Gly Glu Gln Gly Leu Glu Tyr
 450 455 460
 Val Val Phe Lys Thr His His Asn Ala Val Ser Ser Tyr Ile Lys Asp
 465 470 475 480
 Val Phe Arg Val Ile Pro Ser Glu Val Leu Ser Asn Ser Tyr Asn Leu
 485 490 495
 Gly Gln Ser Gln Val Arg Gln Leu Lys Tyr Gln Gly Asn Ser Gly Pro
 500 505 510
 Leu Val Asn Pro
 515

<210> 27

<211> 1446

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(1443)

<223> glycinin G3 subunit

<400> 27

atg gct aag ctt gtt ctt tcc ctt tgt ttt ctg ctt ttc agt ggc tgc	48
Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys	
1 5 10 15	
tgc ttc gct ttc agt ttc aga gag cag cca cag caa aac gag tgc cag	96
Cys Phe Ala Phe Ser Phe Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln	
20 25 30	
atc caa cgc ctc aat gcc cta aaa ccg gat aac cgt ata gag tca gaa	144
Ile Gln Arg Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu	
35 40 45	
ggg ggc ttc att gag aca tgg aac cct aac aac aag cca ttc cag tgt	192
Gly Gly Phe Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys	
50 55 60	
gcc ggt gtt gcc ctc tct cgc tgc acc ctc aac cgc aac gcc ctt cgc	240
Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg	
65 70 75 80	
aga cct tcc tac acc aac gct ccc cag gag atc tac atc caa caa ggt	288
Arg Pro Ser Tyr Thr Asn Ala Pro Gln Glu Ile Tyr Ile Gln Gln Gly	
85 90 95	
agt ggt att ttt ggc atg ata ttc ccg ggt tgt cct agc aca ttt gaa	336
Ser Gly Ile Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Phe Glu	
100 105 110	
gag cct caa caa aaa gga caa agc agc agg ccc caa gac cgt cac cag	384
Glu Pro Gln Gln Lys Gly Gln Ser Ser Arg Pro Gln Asp Arg His Gln	
115 120 125	
aag atc tat cac ttc aga gag ggt gat ttg att gca gtg cca acc ggt	432
Lys Ile Tyr His Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly	
130 135 140	

37

ttt gca tac tgg atg tac aac aat gaa gac act cct gtt gtt gcc gtt	480
Phe Ala Tyr Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val	
145 150 155 160	
tct ctt att gac acc aac agc ttc cag aac cag ctc gac cag atg cct	528
Ser Leu Ile Asp Thr Asn Ser Phe Gln Asn Gln Leu Asp Gln Met Pro	
165 170 175	
agg aga ttc tat ctt gct ggg aac caa gag caa gag ttt cta cag tat	576
Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Gln Tyr	
180 185 190	
cag cca cag aag cag caa gga ggt act caa agc cag aaa gga aag cgt	624
Gln Pro Gln Lys Gln Gln Gly Gly Thr Gln Ser Gln Lys Gly Lys Arg	
195 200 205	
cag caa gaa gaa gaa aac gaa gga ggc agc ata ttg agt ggc ttc gcc	672
Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe Ala	
210 215 220	
ccg gaa ttc ttg gaa cat gcg ttc gtc gtg gac agg cag ata gtg aga	720
Pro Glu Phe Leu Glu His Ala Phe Val Val Asp Arg Gln Ile Val Arg	
225 230 235 240	
aag cta caa ggt gag aac gaa gag gaa gag aag ggt gcc att gtg aca	768
Lys Leu Gln Gly Glu Asn Glu Glu Glu Glu Lys Gly Ala Ile Val Thr	
245 250 255	
gtg aaa gga ggt ctc agc gtg ata agc cca ccc acg gaa gag cag caa	816
Val Lys Gly Gly Leu Ser Val Ile Ser Pro Pro Thr Glu Glu Gln Gln	
260 265 270	
caa aga ccc gag gaa gag gag aag cca gat tgt gac gag aaa gac aaa	864
Gln Arg Pro Glu Glu Glu Lys Pro Asp Cys Asp Glu Lys Asp Lys	
275 280 285	
cat tgc caa agc caa agc aga aat ggc att gac gag acc att tgc aca	912
His Cys Gln Ser Gln Ser Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr	
290 295 300	
atg aga ctt cgc cac aac att ggc cag act tca tca cct gac atc ttc	960
Met Arg Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Phe	
305 310 315 320	
aac cct caa gct ggt agc atc aca acc gct acc agc ctc gac ttc cca	1008
Asn Pro Gln Ala Gly Ser Ile Thr Thr Ala Thr Ser Leu Asp Phe Pro	
325 330 335	
gcc ctc tcg tgg ctc aaa ctc agt gcc cag ttt gga tca ctc cgc aag	1056
Ala Leu Ser Trp Leu Lys Leu Ser Ala Gln Phe Gly Ser Leu Arg Lys	
340 345 350	
aat gct atg ttc gtg cca cac tac aac ctg aac gca aac agc ata ata	1104
Asn Ala Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile	
355 360 365	
tac gca ttg aat gga cgg gca ttg gta caa gtg gtg aat tgc aat ggt	1152
Tyr Ala Leu Asn Gly Arg Ala Leu Val Gln Val Val Asn Cys Asn Gly	
370 375 380	
gag aga gtg ttt gat gga gag ctg caa gag gga cag gtg tta att gtg	1200
Glu Arg Val Phe Asp Gly Glu Leu Gln Glu Gly Gln Val Leu Ile Val	
385 390 395 400	
cca caa aac ttt gcg gtg gct gca aga tca cag agc gac aac ttc gag	1248
Pro Gln Asn Phe Ala Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu	
405 410 415	

38

tat gtt tca ttc aag acc aat gat aga ccc tcg atc ggc aac ctt gca 1296
 Tyr Val Ser Phe Lys Thr Asn Asp Arg Pro Ser Ile Gly Asn Leu Ala
 420 425 430

ggc gca aac tca ttg ttg aac gca ttg ccg gag gaa gtg att cag caa 1344
 Gly Ala Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln Gln
 435 440 445

act ttt aac cta agg agg cag cag gcc agg cag gtc aag aac aac aac 1392
 Thr Phe Asn Leu Arg Arg Gln Gln Ala Arg Gln Val Lys Asn Asn Asn
 450 455 460

cct ttc agc ttc ctg gtt cca cct aag gag tct cag agg aga gtt gtg 1440
 Pro Phe Ser Phe Leu Val Pro Pro Lys Glu Ser Gln Arg Arg Val Val
 465 470 475 480

gct tag 1446
 Ala

<210> 28

<211> 481

<212> PRT

<213> Glycine max

<400> 28

Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys
 1 5 10 15

Cys Phe Ala Phe Ser Phe Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln
 20 25 30

Ile Gln Arg Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu
 35 40 45

Gly Gly Phe Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys
 50 55 60

Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg
 65 70 75 80

Arg Pro Ser Tyr Thr Asn Ala Pro Gln Glu Ile Tyr Ile Gln Gln Gly
 85 90 95

Ser Gly Ile Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Phe Glu
 100 105 110

Glu Pro Gln Gln Lys Gly Gln Ser Ser Arg Pro Gln Asp Arg His Gln
 115 120 125

Lys Ile Tyr His Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly
 130 135 140

Phe Ala Tyr Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val
 145 150 155 160

Ser Leu Ile Asp Thr Asn Ser Phe Gln Asn Gln Leu Asp Gln Met Pro
 165 170 175

Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Gln Tyr
 180 185 190

Gln Pro Gln Lys Gln Gln Gly Gly Thr Gln Ser Gln Lys Gly Lys Arg
 195 200 205

Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe Ala
 210 215 220

Pro Glu Phe Leu Glu His Ala Phe Val Val Asp Arg Gln Ile Val Arg
 225 230 235 240

39

```
<210> 29
<211> 1482
<212> DNA
<213> Helianthus annuus
```

```
<220>
<221> CDS
<222> (1)..(1479)
<223> 11S storage protein G3-D1
```

```

<400> 29
atg gca tcc aaa gca act ttg ctc tta gct ttt acc ctt ctc ttt gcc 48
Met Ala Ser Lys Ala Thr Leu Leu Leu Ala Phe Thr Leu Leu Phe Ala
  1                    5                10                15

act tgc att gcc cgc cac cag caa cgg caa cag caa cag aac cag tgc 96
Thr Cys Ile Ala Arg His Gln Gln Arg Gln Gln Gln Gln Asn Gln Cys
                20                25                30

```

40

cag ctt caa aac atc gag gcg ctc gag ccc atc gaa gtt atc caa gct	144
Gln Leu Gln Asn Ile Glu Ala Leu Glu Pro Ile Glu Val Ile Gln Ala	
35 40 45	
gaa gcc ggt gtg acc gaa att tgg gac gcc tat gac caa cag ttc cag	192
Glu Ala Gly Val Thr Glu Ile Trp Asp Ala Tyr Asp Gln Gln Phe Gln	
50 55 60	
tgt gcg tgg tcg att tta ttc gac acc gga ttc aac ctg gtg gcc ttc	240
Cys Ala Trp Ser Ile Leu Phe Asp Thr Gly Phe Asn Leu Val Ala Phe	
65 70 75 80	
tct tgc ctt cct acg tca aca ccc cta ttt tgg cct tcg tcg aga gag	288
Ser Cys Leu Pro Thr Ser Thr Pro Leu Phe Trp Pro Ser Ser Arg Glu	
85 90 95	
ggg gtt ata ttg ccg gga tgc cgc aga acc tat gaa tat tcg cag gag	336
Gly Val Ile Leu Pro Gly Cys Arg Arg Thr Tyr Glu Tyr Ser Gln Glu	
100 105 110	
caa cag ttt tcc ggt gag ggt ggc cgc aga gga gga gga gag ggc aca	384
Gln Gln Phe Ser Gly Glu Gly Gly Arg Arg Gly Gly Gly Glu Gly Thr	
115 120 125	
ttc agg acc gtc atc aga aag tta gag aac tta aag gag ggt gac gtg	432
Phe Arg Thr Val Ile Arg Lys Leu Glu Asn Leu Lys Glu Gly Asp Val	
130 135 140	
gtt gcc atc ccc acc gga aca gct cac tgg ctt cac aac gac ggc aac	480
Val Ala Ile Pro Thr Gly Thr Ala His Trp Leu His Asn Asp Gly Asn	
145 150 155 160	
aca gaa ctt gtg gtc gtc ttc ttg gat act cag aac cat gag aac cag	528
Thr Glu Leu Val Val Val Phe Leu Asp Thr Gln Asn His Glu Asn Gln	
165 170 175	
ctt gac gaa aac caa agg aga ttc ttc tta gcc gga aac cct caa gct	576
Leu Asp Glu Asn Gln Arg Arg Phe Phe Leu Ala Gly Asn Pro Gln Ala	
180 185 190	
caa gct caa agc cag cag caa caa caa aga caa cca cgc caa caa tct	624
Gln Ala Gln Ser Gln Gln Gln Gln Gln Arg Gln Pro Arg Gln Gln Ser	
195 200 205	
cct caa agg caa agg caa agg caa agg caa ggg caa ggt cag aac gcc	672
Pro Gln Arg Gln Arg Gln Arg Gln Arg Gln Gly Gln Gly Gln Asn Ala	
210 215 220	
ggc aac atc ttc aac ggt ttc acc ccc gag ctc att gca caa tca ttc	720
Gly Asn Ile Phe Asn Gly Phe Thr Pro Glu Leu Ile Ala Gln Ser Phe	
225 230 235 240	
aac gtc gac caa gag acc gcc cag aag cta caa gga caa aac gac cag	768
Asn Val Asp Gln Glu Thr Ala Gln Lys Leu Gln Gly Gln Asn Asp Gln	
245 250 255	
aga ggc cac att gtt aat gtc gga caa gac ctt caa ata gtc cgc cca	816
Arg Gly His Ile Val Asn Val Gly Gln Asp Leu Gln Ile Val Arg Pro	
260 265 270	
cca caa gac aga cgc tct cct cgc caa caa caa gag caa gcg acg tct	864
Pro Gln Asp Arg Arg Ser Pro Arg Gln Gln Gln Glu Gln Ala Thr Ser	
275 280 285	
cct cgc caa caa caa gag cag cag caa ggc aga cgt ggc gga tgg agc	912
Pro Arg Gln Gln Gln Glu Gln Gln Gln Gly Arg Arg Gly Gly Trp Ser	
290 295 300	

41

aac ggt gtg gaa gaa acc atc tgc agc atg aag ttc aaa gtg aac att 960
 Asn Gly Val Glu Glu Thr Ile Cys Ser Met Lys Phe Lys Val Asn Ile
 305 310 315 320
 gac aac cct tcc cag gct gac ttt gta aac ccg caa gcc ggc agc att 1008
 Asp Asn Pro Ser Gln Ala Asp Phe Val Asn Pro Gln Ala Gly Ser Ile
 325 330 335
 gca aac ctc aac agc ttc aaa ttc ccc att ctc gag cac ctc cgg ctc 1056
 Ala Asn Leu Asn Ser Phe Lys Phe Pro Ile Leu Glu His Leu Arg Leu
 340 345 350
 agc gtg gaa aga ggc gaa ctc cgt ccg aat gcc atc caa tcc cca cac 1104
 Ser Val Glu Arg Gly Glu Leu Arg Pro Asn Ala Ile Gln Ser Pro His
 355 360 365
 tgg aca atc aac gcc cac aat ctt ctc tac gta acc gag gga gcc ttg 1152
 Trp Thr Ile Asn Ala His Asn Leu Leu Tyr Val Thr Glu Gly Ala Leu
 370 375 380
 agg gta caa atc gtc gac aac caa gga aac tca gtt ttc gac aac gag 1200
 Arg Val Gln Ile Val Asp Asn Gln Gly Asn Ser Val Phe Asp Asn Glu
 385 390 395 400
 ctc cgt gag gga cag gtg gtg gtg atc ccg cag aac ttt gcg gtg atc 1248
 Leu Arg Glu Gly Gln Val Val Val Ile Pro Gln Asn Phe Ala Val Ile
 405 410 415
 aag aga gcc aat gaa caa gga agc agg tgg gtg tct ttc aag act aat 1296
 Lys Arg Ala Asn Glu Gln Gly Ser Arg Trp Val Ser Phe Lys Thr Asn
 420 425 430
 gat aat gcc atg ata gca aac ctt gca ggg cgt gtg tcc gca tca gca 1344
 Asp Asn Ala Met Ile Ala Asn Leu Ala Gly Arg Val Ser Ala Ser Ala
 435 440 445
 gca tcg ccg ttg acg ttg tgg gcg aat cgg tat cag cta tct cga gag 1392
 Ala Ser Pro Leu Thr Leu Trp Ala Asn Arg Tyr Gln Leu Ser Arg Glu
 450 455 460
 gaa gct cag cag ctc aag ttt agc cag agg gag acg gtt ttg ttt gca 1440
 Glu Ala Gln Gln Leu Lys Phe Ser Gln Arg Glu Thr Val Leu Phe Ala
 465 470 475 480
 cca agt ttt tcc agg ggc caa ggg atc agg gct tca cgt taa 1482
 Pro Ser Phe Ser Arg Gly Gln Gly Ile Arg Ala Ser Arg
 485 490

<210> 30

<211> 493

<212> PRT

<213> Helianthus annuus

<400> 30

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

42

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

43

Glu Ala Gln Gln Leu Lys Phe Ser Gln Arg Glu Thr Val Leu Phe Ala
 465 470 475 480
 Pro Ser Phe Ser Arg Gly Gln Gly Ile Arg Ala Ser Arg
 485 490

<210> 31

<211> 537

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (1)..(534)

<223> NAPIN

<400> 31

atg gcg aac aag ctc ttc ctc gtc tgc gca act ctc gcc ttc ttc ttc	48
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe	
1 5 10 15	
ctt ctc acc aat gcc tcc atc tac cgg acg gtc gtc gag ttc gac gaa	96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu	
20 25 30	
gat gat gcc aca gac tca gcc ggc cca ttt agg att cca aaa tgt agg	144
Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg	
35 40 45	
aag gag ttt cag caa gca caa cac cta aga gct tgc cag cag tgg ctc	192
Lys Glu Phe Gln Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu	
50 55 60	
cac aag caa gca atg cag tct ggc ggt ggt cct agc tgg acc ctc gac	240
His Lys Gln Ala Met Gln Ser Gly Gly Gly Pro Ser Trp Thr Leu Asp	
65 70 75 80	
ggt gag ttt gac ttt gaa gac gac atg gag aac ccg cag ggt cca cag	288
Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln	
85 90 95	
cag aga ccg cct cta ctc cag cag tgc tgt aac gag ctc cac cag gaa	336
Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu	
100 105 110	
gag ccc ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt	384
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val	
115 120 125	
aaa caa caa att caa caa cag gga caa cag caa gga aag cag caa atg	432
Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met	
130 135 140	
gtg agc cgt atc tac cag acc gct acg cac tta cct aaa gtt tgc aac	480
Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn	
145 150 155 160	
atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc	528
Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro	
165 170 175	
tcc tac tag	537
Ser Tyr	

<210> 32

<211> 178

<212> PRT

<213> Brassica napus

<400> 32

```

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

Ser Tyr

```

<210> 33

<211> 537

<212> DNA

<213> Brassica juncea

<220>

<221> CDS

<222> (1)..(534)

<223> 2S storage protein

<400> 33

```

atg gcg aac aag ctc ttc ctc gtc tcg gca act ctc gcc ttc ttc ttc      48
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
 1             5             10             15
ctt ctc acc aat gcc tcc atc tac cgg acg gtc gtc gag ttc gac gaa      96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
          20             25             30
gat gat gcc aca gac tca gcc ggc cca ttt agg att cca aaa tgt agg      144
Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
          35             40             45
aag gag ttt cag caa gca caa cac cta aga gtc tgc cag cag tgg ctc      192
Lys Glu Phe Gln Gln Ala Gln His Leu Arg Val Cys Gln Gln Trp Leu
          50             55             60

```

45

```

cac aag caa gca atg cag tct ggc ggt ggt ctc agc tgg acc ctc gac 240
His Lys Gln Ala Met Gln Ser Gly Gly Gly Leu Ser Trp Thr Leu Asp
 65              70              75              80

ggt gag ttt gac ttt gaa gac gac atg gag aac tcg cag ggt cca cag 288
Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Ser Gln Gly Pro Gln
      85              90              95

cag aga ccg cct cta ctc cag cag tgc tgt aac gag ctc cac cag gaa 336
Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
      100              105              110

gag ccc ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt 384
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
      115              120              125

aaa caa caa att caa caa cag gga caa cag caa gga aag cag caa atg 432
Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
      130              135              140

gtg agc cgt atc tac cag acc gct acg cac tta cct aaa gtt tgc aac 480
Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
      145              150              155              160

atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc 528
Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
      165              170              175

tcc tac tag 537
Ser Tyr

<210> 34
<211> 178
<212> PRT
<213> Brassica juncea

<400> 34
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
 1              5              10              15

Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
      20              25              30

Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
      35              40              45

Lys Glu Phe Gln Gln Ala Gln His Leu Arg Val Cys Gln Gln Trp Leu
      50              55              60

His Lys Gln Ala Met Gln Ser Gly Gly Gly Leu Ser Trp Thr Leu Asp
      65              70              75              80

Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Ser Gln Gly Pro Gln
      85              90              95

Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
      100              105              110

Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
      115              120              125

Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
      130              135              140

Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
      145              150              155              160

Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
      165              170              175

```

Ser Tyr

<210> 35

<211> 537

<212> DNA

<213> Brassica oleracea

<220>

<221> CDS

<222> (1)..(534)

<223> 2S storage protein

<400> 35

atg	gcg	aac	aag	ctc	ttc	ctc	gtc	tcg	gca	act	ctc	gcc	ttc	ttc	ttc	48
Met	Ala	Asn	Lys	Leu	Phe	Leu	Val	Ser	Ala	Thr	Leu	Ala	Phe	Phe	Phe	
1				5					10				15			

ctt	ctc	acc	aat	gcc	tcc	atc	tac	cgg	acg	gtg	gtc	gag	ttc	gac	gaa	96
Leu	Leu	Thr	Asn	Ala	Ser	Ile	Tyr	Arg	Thr	Val	Val	Glu	Phe	Asp	Glu	
			20					25					30			

gat	gat	gcc	aca	aac	cca	gcc	ggc	cca	ttt	agg	atc	cca	aaa	tgt	agg	144
Asp	Asp	Ala	Thr	Asn	Pro	Ala	Gly	Pro	Phe	Arg	Ile	Pro	Lys	Cys	Arg	
		35					40					45				

aag	gag	ttt	cag	caa	gca	caa	cac	cta	aga	gct	tgc	cag	cag	tggt	ctc	192
Lys	Glu	Phe	Gln	Gln	Ala	Gln	His	Leu	Arg	Ala	Cys	Gln	Gln	Trp	Leu	
	50					55					60					

cac	aag	caa	gca	atg	cag	tct	ggc	ggt	ggt	cct	agc	tggt	acc	ctc	gac	240
His	Lys	Gln	Ala	Met	Gln	Ser	Gly	Gly	Gly	Pro	Ser	Trp	Thr	Leu	Asp	
	65				70					75					80	

agt	gag	ttt	gac	ttt	gaa	gac	gac	atg	gag	aac	ccg	cag	ggt	cca	cag	288
Ser	Glu	Phe	Asp	Phe	Glu	Asp	Asp	Met	Glu	Asn	Pro	Gln	Gly	Pro	Gln	
			85						90					95		

cag	aga	ccg	cct	cta	ctc	ctg	caa	tgc	tgt	aac	gag	ctg	gac	cag	gaa	336
Gln	Arg	Pro	Pro	Leu	Leu	Leu	Gln	Cys	Cys	Asn	Glu	Leu	Asp	Gln	Glu	
			100					105					110			

gag	ccc	ctt	tgc	ggt	tgc	cca	acc	ttg	aaa	gga	gca	tcc	aaa	gcg	ggt	384
Glu	Pro	Leu	Cys	Val	Cys	Pro	Thr	Leu	Lys	Gly	Ala	Ser	Lys	Ala	Val	
		115					120					125				

aaa	caa	caa	att	caa	caa	cag	gga	caa	cag	caa	gga	aag	cag	caa	atg	432
Lys	Gln	Gln	Ile	Gln	Gln	Gln	Gly	Gln	Gln	Gln	Gly	Lys	Gln	Gln	Met	
	130					135					140					

gtg	agc	cgt	atc	tac	cag	acc	gct	acg	cac	tta	cct	aaa	ggt	tgc	aac	480
Val	Ser	Arg	Ile	Tyr	Gln	Thr	Ala	Thr	His	Leu	Pro	Lys	Val	Cys	Asn	
	145				150				155					160		

atc	ccg	caa	ggt	agc	ggt	tgt	ccc	ttc	cag	aag	acc	atg	cct	ggg	ccc	528
Ile	Pro	Gln	Val	Ser	Val	Cys	Pro	Phe	Gln	Lys	Thr	Met	Pro	Gly	Pro	
			165						170					175		

tcc	tac	tag														537
Ser	Tyr															

<210> 36

<211> 178

<212> PRT

<213> Brassica oleracea

<400> 36

```

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

```

<210> 37

<211> 543

<212> DNA

<213> Brassica napus cv. Topas

<220>

<221> CDS

<222> (1)..(540)

<223> Napin

<400> 37

```

atg gcg aac aag ctc ttc ctc gtc tcg gca act ctt gcc ttc ttc ttc      48
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
  1           5           10           15
ctt ctc acc aac gcc tcc atc tac cgc acc atc gtg gaa gtc gac gaa      96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Ile Val Glu Val Asp Glu
           20           25           30
gat gat gcc aca aac cca gcc ggc cca ttt agg att cca aaa tgt agg      144
Asp Asp Ala Thr Asn Pro Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
           35           40           45
aag gag ttt cag caa gca caa cac ctg aaa gct tgc caa caa tgg ctc      192
Lys Glu Phe Gln Gln Ala Gln His Leu Lys Ala Cys Gln Gln Trp Leu
           50           55           60
cac aag cag gca atg cag tcc ggt agt ggc cca agc tgg acc ctc gac      240
His Lys Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp
           65           70           75           80

```

48

```

ggt gag ttt gat ttt gaa gat gac atg gag aac ccc cag ggc cca caa 288
Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
      85                      90                      95

cag agg ccg cca cta ctc cag cag tgc tgc aac gag ctc cac cag gaa 336
Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
      100                      105                      110

gag cca ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcc gtt 384
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
      115                      120                      125

aaa caa cag gtt cga caa cag caa gga cag cag gga cag cag ctg cag 432
Lys Gln Gln Val Arg Gln Gln Gln Gly Gln Gln Gly Gln Gln Leu Gln
      130                      135                      140

caa gta att agc cgt atc tac cag act gct acg cac tta cct aaa gtt 480
Gln Val Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val
      145                      150                      155                      160

tgc aac atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct 528
Cys Asn Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro
      165                      170                      175

gga ccc tcc tac tag 543
Gly Pro Ser Tyr
      180

```

<210> 38

<211> 180

<212> PRT

<213> Brassica napus cv. Topas

<400> 38

```

Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
  1              5              10              15

Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Ile Val Glu Val Asp Glu
      20              25              30

Asp Asp Ala Thr Asn Pro Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
      35              40              45

Lys Glu Phe Gln Gln Ala Gln His Leu Lys Ala Cys Gln Gln Trp Leu
      50              55              60

His Lys Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp
      65              70              75              80

Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
      85              90              95

Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
      100              105              110

Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
      115              120              125

Lys Gln Gln Val Arg Gln Gln Gln Gly Gln Gln Gly Gln Gln Leu Gln
      130              135              140

Gln Val Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val
      145              150              155              160

Cys Asn Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro
      165              170              175

Gly Pro Ser Tyr
      180

```

```

<210> 39
<211> 435
<212> DNA
<213> Sinapis alba

<220>
<221> CDS
<222> (1)..(432)
<223> coding for partial sin1 storage protein

<400> 39
cca gcc ggc cca ttt ggg att cca aaa tgt agg aag gag ttt caa caa 48
Pro Ala Gly Pro Phe Gly Ile Pro Lys Cys Arg Lys Glu Phe Gln Gln
  1             5             10             15

gca caa cac cta aga gct tgc cag caa tgg ctc cac aag cag gca atg 96
Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu His Lys Gln Ala Met
      20             25             30

cag tct ggt agt ggt cca agc tgg acc ctc gac gat gag ttt gat ttt 144
Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp Asp Glu Phe Asp Phe
      35             40             45

gaa gat gac atg gag aac cca cag gga cca cag cag agg cca cca cta 192
Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln Gln Arg Pro Pro Leu
      50             55             60

ctc cag cag tgc tgc aac gag ctc cac cag gaa gag cca ctt tgc gtt 240
Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu Glu Pro Leu Cys Val
      65             70             75             80

tgc cca acc ttg aaa gga gca tcc aaa gcc gtt aaa cag cag gtt aga 288
Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val Lys Gln Gln Val Arg
      85             90             95

caa cag ctg gag cag cag gga cag cag gga ccg cac ctg cag cat gta 336
Gln Gln Leu Glu Gln Gln Gly Gln Gln Gly Pro His Leu Gln His Val
      100            105            110

att agc cgt atc tac cag acc gct acg cac tta cct aga gtt tgc aac 384
Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Arg Val Cys Asn
      115            120            125

att agg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct gga ccc 432
Ile Arg Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
      130            135            140

tcc 435

<210> 40
<211> 144
<212> PRT
<213> Sinapis alba

<400> 40
Pro Ala Gly Pro Phe Gly Ile Pro Lys Cys Arg Lys Glu Phe Gln Gln
  1             5             10             15

Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu His Lys Gln Ala Met
      20             25             30

Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp Asp Glu Phe Asp Phe
      35             40             45

Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln Gln Arg Pro Pro Leu
      50             55             60

```

50

Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu Glu Pro Leu Cys Val
 65 70 75 80
 Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val Lys Gln Gln Val Arg
 85 90 95
 Gln Gln Leu Glu Gln Gln Gly Gln Gln Gly Pro His Leu Gln His Val
 100 105 110
 Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Arg Val Cys Asn
 115 120 125
 Ile Arg Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
 130 135 140

<210> 41

<211> 468

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(465)

<223> 2S albumine 1

<400> 41

atg acc aag ctt aca att ctc ctc atc gct ctt ctc ttc atc gcc cac 48
 Met Thr Lys Leu Thr Ile Leu Leu Ile Ala Leu Leu Phe Ile Ala His
 1 5 10 15
 acc tgc tgc gcc tcc aaa tgg caa cag cac cag caa gag agc tgc cgc 96
 Thr Cys Cys Ala Ser Lys Trp Gln Gln His Gln Gln Glu Ser Cys Arg
 20 25 30
 gag cag ctc aag ggg atc aac ctc aac ccc tgt gag cac atc atg gag 144
 Glu Gln Leu Lys Gly Ile Asn Leu Asn Pro Cys Glu His Ile Met Glu
 35 40 45
 aag atc caa gct ggc cgc cgc ggc gag gac ggc agc gac gaa gat cac 192
 Lys Ile Gln Ala Gly Arg Arg Gly Glu Asp Gly Ser Asp Glu Asp His
 50 55 60
 att ctc atc agg acc atg ccg gga aga atc aac tac atc agg aag aag 240
 Ile Leu Ile Arg Thr Met Pro Gly Arg Ile Asn Tyr Ile Arg Lys Lys
 65 70 75 80
 gaa gga aaa gaa gaa gaa gaa gaa gga cac atg cag aag tgc tgc agc 288
 Glu Gly Lys Glu Glu Glu Glu Glu Gly His Met Gln Lys Cys Cys Ser
 85 90 95
 gaa atg agc gag ctg aaa agc ccc ata tgc cag tgc aaa gcg cta cag 336
 Glu Met Ser Glu Leu Lys Ser Pro Ile Cys Gln Cys Lys Ala Leu Gln
 100 105 110
 aag ata atg gat aac cag agc gag caa ctg gag ggg aag gag aag aag 384
 Lys Ile Met Asp Asn Gln Ser Glu Gln Leu Glu Gly Lys Glu Lys Lys
 115 120 125
 cag atg gag aga gag ctc atg aac ttg gct att agg tgc agg ttg gga 432
 Gln Met Glu Arg Glu Leu Met Asn Leu Ala Ile Arg Cys Arg Leu Gly
 130 135 140
 ccc atg ata ggg tgc gac ttg tcc tcc gat gac tga 468
 Pro Met Ile Gly Cys Asp Leu Ser Ser Asp Asp
 145 150 155

<210> 42

<211> 155

<212> PRT

<213> Glycine max

<400> 42

```

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

```

<210> 43

<211> 477

<212> DNA

<213> Glycine max

<220>

<221> CDS

<222> (1)..(474)

<223> 2S albumine

<400> 43

```

atg acc aag ttc aca atc ctc ctc atc tct ctt ctc ttc tgc atc gcc      48
Met Thr Lys Phe Thr Ile Leu Leu Ile Ser Leu Leu Phe Cys Ile Ala
 1              5              10              15
cac act tgc agc gcc tcc aaa tgg cag cac cag caa gat agc tgc cgc      96
His Thr Cys Ser Ala Ser Lys Trp Gln His Gln Gln Asp Ser Cys Arg
          20              25              30
aag cag ctc cag ggg gtg aac ctc acg ccc tgc gag aag cac atc atg      144
Lys Gln Leu Gln Gly Val Asn Leu Thr Pro Cys Glu Lys His Ile Met
          35              40              45
gag aag atc caa ggc cgc ggc gat gac gat gat gat gat gac gac gac      192
Glu Lys Ile Gln Gly Arg Gly Asp Asp Asp Asp Asp Asp Asp Asp Asp
          50              55              60
aat cac att ctc agg acc atg cgg gga aga atc aac tac ata agg agg      240
Asn His Ile Leu Arg Thr Met Arg Gly Arg Ile Asn Tyr Ile Arg Arg
          65              70              75              80

```

52

```

aac gaa gga aaa gac gaa gac gaa gaa gaa gaa gga cac atg cag aag      288
Asn Glu Gly Lys Asp Glu Asp Glu Glu Glu Glu Gly His Met Gln Lys
      85                      90                      95

tgc tgc aca gaa atg agc gag ctg aga agc ccc aaa tgc cag tgc aaa      336
Cys Cys Thr Glu Met Ser Glu Leu Arg Ser Pro Lys Cys Gln Cys Lys
      100                      105                      110

gcg ctg cag aag ata atg gag aac cag agc gag gaa ctg gag gag aag      384
Ala Leu Gln Lys Ile Met Glu Asn Gln Ser Glu Glu Leu Glu Glu Lys
      115                      120                      125

cag aag aag aaa atg gag aag gag ctc att aac ttg gct act atg tgc      432
Gln Lys Lys Lys Met Glu Lys Glu Leu Ile Asn Leu Ala Thr Met Cys
      130                      135                      140

agg ttt gga ccc atg atc cag tgc gac ttg tcc tcc gat gac taa      477
Arg Phe Gly Pro Met Ile Gln Cys Asp Leu Ser Ser Asp Asp
      145                      150                      155

<210> 44
<211> 158
<212> PRT
<213> Glycine max

<400> 44
Met Thr Lys Phe Thr Ile Leu Leu Ile Ser Leu Leu Phe Cys Ile Ala
  1                      5                      10                      15
His Thr Cys Ser Ala Ser Lys Trp Gln His Gln Gln Asp Ser Cys Arg
      20                      25                      30
Lys Gln Leu Gln Gly Val Asn Leu Thr Pro Cys Glu Lys His Ile Met
      35                      40                      45
Glu Lys Ile Gln Gly Arg Gly Asp Asp Asp Asp Asp Asp Asp Asp
      50                      55                      60
Asn His Ile Leu Arg Thr Met Arg Gly Arg Ile Asn Tyr Ile Arg Arg
      65                      70                      75                      80
Asn Glu Gly Lys Asp Glu Asp Glu Glu Glu Glu Gly His Met Gln Lys
      85                      90                      95
Cys Cys Thr Glu Met Ser Glu Leu Arg Ser Pro Lys Cys Gln Cys Lys
      100                      105                      110
Ala Leu Gln Lys Ile Met Glu Asn Gln Ser Glu Glu Leu Glu Glu Lys
      115                      120                      125
Gln Lys Lys Lys Met Glu Lys Glu Leu Ile Asn Leu Ala Thr Met Cys
      130                      135                      140
Arg Phe Gly Pro Met Ile Gln Cys Asp Leu Ser Ser Asp Asp
      145                      150                      155

```

```

<210> 45
<211> 537
<212> DNA
<213> Brassica nigra

<220>
<221> CDS
<222> (1)..(534)
<223> 2S storage protein

```

53

<400> 45
 atg gcg aac aag ctc ttc ctc gtc tcg gca act ctc gcc ttc ttc ttc 48
 Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
 1 5 10 15
 ctg ctc acc aat gcc tcc atc tac cgg acg gtc gtc gag ttc gac gaa 96
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
 20 25 30
 gat gat gac aca aac caa gcc gga cca ttt agg att cca aga tgt cga 144
 Asp Asp Asp Thr Asn Gln Ala Gly Pro Phe Arg Ile Pro Arg Cys Arg
 35 40 45
 aag gag ttt cgg caa gca caa cat cta aga gct tgc cag caa tgg ctc 192
 Lys Glu Phe Arg Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu
 50 55 60
 cac agg cag gca atg cag tcc ggt agt ggt cca agc tgg acc ctg gac 240
 His Arg Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp
 65 70 75 80
 ggt gag ttt gac ttt gaa gac gac atg gag aac caa cag ggc cca cag 288
 Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Gln Gln Gly Pro Gln
 85 90 95
 cag agg cca cct cta ctc cag caa tgc tgc aac gag ctc cac cag gaa 336
 Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
 100 105 110
 gag gca ctt tgt gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt 384
 Glu Ala Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
 115 120 125
 aga caa cag gtt cga caa cag gga cac cag cag cag atg cag cat gta 432
 Arg Gln Gln Val Arg Gln Gln Gly His Gln Gln Gln Met Gln His Val
 130 135 140
 att agc cgt atc tac cag acc gct acg cac tta cct aga gtt tgc aac 480
 Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Arg Val Cys Asn
 145 150 155 160
 atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc 528
 Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
 165 170 175
 tcc tac tag 537
 Ser Tyr
 <210> 46
 <211> 178
 <212> PRT
 <213> Brassica nigra
 <400> 46
 Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
 1 5 10 15
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
 20 25 30
 Asp Asp Asp Thr Asn Gln Ala Gly Pro Phe Arg Ile Pro Arg Cys Arg
 35 40 45
 Lys Glu Phe Arg Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu
 50 55 60
 His Arg Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp
 65 70 75 80

[illegible]

```
<210> 47
<211> 435
<212> DNA
<213> Sinapis alba
<220>
<221> CDS
<222> (1)..(432)
<223> sin5 storage protein
```

<400> 47																
cca gcc ggc	cca ttt ggg	att cca aaa tgt	agg aag gag	ttt caa caa	48											
Pro Ala Gly	Pro Phe Gly	Ile Pro Lys Cys	Arg Lys Glu	Phe Gln Gln												
1	5	10		15												
gca caa cac	cta aga gct	tgc cag caa tgg	ctc cac aag	cag gca atg	96											
Ala Gln His	Leu Arg Ala	Cys Gln Gln	Trp Leu His	Lys Gln Ala												
	20	25		30												
cag tct ggt	agt ggt cca	agc tgg acc	ctc gac gat	gag ttt gat	144											
Gln Ser Gly	Ser Gly Pro	Ser Trp Thr	Leu Asp Asp	Glu Phe Asp												
35		40		45												
gaa gac gac	atg gag aac	ccc cag gga	cca cag cag	aag ccg cca	192											
Glu Asp Asp	Met Glu Asn	Pro Gln Gly	Pro Gln Gln	Lys Pro Pro												
50		55		60												
ctc cag caa	tgc tgc aac	gag ctt cac	cag gag gag	cca ctt tgc	240											
Leu Gln Gln	Cys Cys Asn	Glu Leu His	Gln Glu Glu	Pro Leu Cys												
65		70		75	80											
tgc cca act	ttg aaa gga	gct tcc aaa	gcc gtt aaa	caa cag gtt	288											
Cys Pro Thr	Leu Lys Gly	Ala Ser Lys	Ala Val Lys	Gln Gln Val												
	85		90	95												
caa cag ttg	ggg cag cag	gga cag cag	gga ccg cag	gtg cag cat	336											
Gln Gln Leu	Gly Gln Gln	Gly Gln Gln	Gly Pro Gln	Val Gln His												
100		105		110												
att agc cgt	atc tac cag	acc gct acg	cac tta cct	aaa gtt tgc	384											
Ile Ser Arg	Ile Tyr Gln	Thr Ala Thr	His Leu Pro	Lys Val Cys												
115		120		125												
atc ccc caa	gta agc gtt	tgt ccc ttc	aag aag acc	atg cct gga	432											
Ile Pro Gln	Val Ser Val	Cys Pro Phe	Lys Lys Thr	Met Pro Gly												
130		135		140												

tcc

435

<210> 48

<211> 144

<212> PRT

<213> Sinapis alba

<400> 48

Pro Ala Gly Pro Phe Gly Ile Pro Lys Cys Arg Lys Glu Phe Gln Gln
 1 5 10 15

Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu His Lys Gln Ala Met
 20 25 30

Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp Asp Glu Phe Asp Phe
 35 40 45

Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln Gln Lys Pro Pro Leu
 50 55 60

Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu Glu Pro Leu Cys Val
 65 70 75 80

Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val Lys Gln Gln Val Arg
 85 90 95

Gln Gln Leu Gly Gln Gln Gly Gln Gln Gly Pro Gln Val Gln His Val
 100 105 110

Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
 115 120 125

Ile Pro Gln Val Ser Val Cys Pro Phe Lys Lys Thr Met Pro Gly Pro
 130 135 140

<210> 49

<211> 888

<212> DNA

<213> Helianthus annuus

<220>

<221> CDS

<222> (1)..(885)

<223> HaG5 2S albumine

<400> 49

atg gca aag caa ata gtt ctc gca ctc gct ttc gcc gcc ctt gta gcc 48
 Met Ala Lys Gln Ile Val Leu Ala Leu Ala Phe Ala Ala Leu Val Ala
 1 5 10 15

ttt gct acc gcc cac aca acc ata atc acc acc acc atc gaa gac gag 96
 Phe Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Glu Asp Glu
 20 25 30

aac ccg atc tcc gga caa agg caa gtg agc caa cgg ata cag gga caa 144
 Asn Pro Ile Ser Gly Gln Arg Gln Val Ser Gln Arg Ile Gln Gly Gln
 35 40 45

agg ctg aac cag tgt cgc atg ttc ctc cag cag ggt cag aac att cct 192
 Arg Leu Asn Gln Cys Arg Met Phe Leu Gln Gln Gly Gln Asn Ile Pro
 50 55 60

cgc gaa ttc gat aac cct cag atg ggg cgg cag cag gag cag cag ctc 240
 Arg Glu Phe Asp Asn Pro Gln Met Gly Arg Gln Gln Glu Gln Gln Leu
 65 70 75 80

56

cag cag tgt tgt caa gag ctc caa aac atc gaa ggg cag tgc caa tgt	288
Gln Gln Cys Cys Gln Glu Leu Gln Asn Ile Glu Gly Gln Cys Gln Cys	
85 90 95	
gag gcg gtg aag cag gtg ttc cga gaa gcc cag cag caa gta caa cag	336
Glu Ala Val Lys Gln Val Phe Arg Glu Ala Gln Gln Gln Val Gln Gln	
100 105 110	
caa cag gga cgg cag ctt gta ccc ttc cgc ggt tgc cag cag acc caa	384
Gln Gln Gly Arg Gln Leu Val Pro Phe Arg Gly Ser Gln Gln Thr Gln	
115 120 125	
cag ttg aag cag aag gct cag att ctc cct aac gta tgc aac ctt caa	432
Gln Leu Lys Gln Lys Ala Gln Ile Leu Pro Asn Val Cys Asn Leu Gln	
130 135 140	
tca aga cga tgt gaa atc gga acc atc acc acc acc gtc acc gag agc	480
Ser Arg Arg Cys Glu Ile Gly Thr Ile Thr Thr Thr Val Thr Glu Ser	
145 150 155 160	
aat atc gat atc ccc ttc cgt gac agg ccc ttt ggc act gga tca caa	528
Asn Ile Asp Ile Pro Phe Arg Asp Arg Pro Phe Gly Thr Gly Ser Gln	
165 170 175	
cag tgc aga gaa act gaa atc caa cga ccc gtt ggt gaa tgc caa agg	576
Gln Cys Arg Glu Thr Glu Ile Gln Arg Pro Val Gly Glu Cys Gln Arg	
180 185 190	
ttc gtg gag cag caa atg cag cag tct ccg agg tcc act aga cca tac	624
Phe Val Glu Gln Gln Met Gln Gln Ser Pro Arg Ser Thr Arg Pro Tyr	
195 200 205	
caa cag cgg cca gga caa cag cag cag cag cag aga ggg ctc caa caa	672
Gln Gln Arg Pro Gly Gln Gln Gln Gln Gln Gln Arg Gly Leu Gln Gln	
210 215 220	
caa tgc tgc aac gag cta caa aac gtg aag agg gag tgt cat tgc gag	720
Gln Cys Cys Asn Glu Leu Gln Asn Val Lys Arg Glu Cys His Cys Glu	
225 230 235 240	
gca att caa gaa gtg gct agg aga gtg atg agg cag cca cag cag cag	768
Ala Ile Gln Glu Val Ala Arg Arg Val Met Arg Gln Pro Gln Gln Gln	
245 250 255	
cag cag caa cgt cgt ggg cag ttc ggt ggg cag gag atg gaa acc gcg	816
Gln Gln Gln Arg Arg Gly Gln Phe Gly Gly Gln Glu Met Glu Thr Ala	
260 265 270	
agg agg gtg att cag aat ctg ccc aac cag tgc gac ttg gaa gtc cag	864
Arg Arg Val Ile Gln Asn Leu Pro Asn Gln Cys Asp Leu Glu Val Gln	
275 280 285	
caa tgc aca acc tgt acg gga tga	888
Gln Cys Thr Thr Cys Thr Gly	
290 295	
<210> 50	
<211> 295	
<212> PRT	
<213> Helianthus annuus	
<400> 50	
Met Ala Lys Gln Ile Val Leu Ala Leu Ala Phe Ala Ala Leu Val Ala	
1 5 10 15	
Phe Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Glu Asp Glu	
20 25 30	

Asn	Pro	Ile	Ser	Gly	Gln	Arg	Gln	Val	Ser	Gln	Arg	Ile	Gln	Gly	Gln
35												45			
Arg	Leu	Asn	Gln	Cys	Arg	Met	Phe	Leu	Gln	Gln	Gly	Gln	Asn	Ile	Pro
50												60			
Arg	Glu	Phe	Asp	Asn	Pro	Gln	Met	Gly	Arg	Gln	Gln	Glu	Gln	Gln	Leu
65												75			
Gln	Gln	Cys	Cys	Gln	Glu	Leu	Gln	Asn	Ile	Glu	Gly	Gln	Cys	Gln	Cys
				85								95			
Glu	Ala	Val	Lys	Gln	Val	Phe	Arg	Glu	Ala	Gln	Gln	Gln	Val	Gln	Gln
				100								105			
Gln	Gln	Gly	Arg	Gln	Leu	Val	Pro	Phe	Arg	Gly	Ser	Gln	Gln	Thr	Gln
				115								125			
Gln	Leu	Lys	Gln	Lys	Ala	Gln	Ile	Leu	Pro	Asn	Val	Cys	Asn	Leu	Gln
130												140			
Ser	Arg	Arg	Cys	Glu	Ile	Gly	Thr	Ile	Thr	Thr	Thr	Val	Thr	Glu	Ser
145				150								155			
Asn	Ile	Asp	Ile	Pro	Phe	Arg	Asp	Arg	Pro	Phe	Gly	Thr	Gly	Ser	Gln
				165								175			
Gln	Cys	Arg	Glu	Thr	Glu	Ile	Gln	Arg	Pro	Val	Gly	Glu	Cys	Gln	Arg
				180								185			
Phe	Val	Glu	Gln	Gln	Met	Gln	Gln	Ser	Pro	Arg	Ser	Thr	Arg	Pro	Tyr
195												205			
Gln	Gln	Arg	Pro	Gly	Gln	Gln	Gln	Gln	Gln	Gln	Arg	Gly	Leu	Gln	Gln
210				215								220			
Gln	Cys	Cys	Asn	Glu	Leu	Gln	Asn	Val	Lys	Arg	Glu	Cys	His	Cys	Glu
225				230								235			
Ala	Ile	Gln	Glu	Val	Ala	Arg	Arg	Val	Met	Arg	Gln	Pro	Gln	Gln	Gln
				245								255			
Gln	Gln	Gln	Arg	Arg	Gly	Gln	Phe	Gly	Gly	Gln	Glu	Met	Glu	Thr	Ala
				260								265			
Arg	Arg	Val	Ile	Gln	Asn	Leu	Pro	Asn	Gln	Cys	Asp	Leu	Glu	Val	Gln
275												285			
Gln	Cys	Thr	Thr	Cys	Thr	Gly									
290				295											

gct aca gcc cac aca acc atc atc acc acc acc atc gac gac gag aac 97
Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Asp Asp Glu Asn
20 25 30

58

ccg atc tcc gaa caa agg caa tgt tgg caa cag gta cag gga caa agg	145
Pro Ile Ser Glu Gln Arg Gln Cys Trp Gln Gln Val Gln Gly Gln Arg	
35 40 45	
ttg aac cag tgt cgc atg ttc ctc cag caa ggt cag agg ggg cag caa	193
Leu Asn Gln Cys Arg Met Phe Leu Gln Gln Gly Gln Arg Gly Gln Gln	
50 55 60	
cac caa cag caa cag cat cag cag cag gag cag cag ctc ctc cag cag	241
His Gln Gln Gln Gln His Gln Gln Gln Glu Gln Gln Leu Leu Gln Gln	
65 70 75 80	
tgt tgt caa gag ctt caa aac atc gaa gga cag tgc caa tgt gag gcg	289
Cys Cys Gln Glu Leu Gln Asn Ile Glu Gly Gln Cys Gln Cys Glu Ala	
85 90 95	
gtg aag cag gtg gtc cga gat gct cag cga cac gag caa cag cga ccg	337
Val Lys Gln Val Val Arg Asp Ala Gln Arg His Glu Gln Gln Arg Pro	
100 105 110	
cga gtg ccc ttc cag ggt tct cag cag tct caa cag ttg aag cag agg	385
Arg Val Pro Phe Gln Gly Ser Gln Gln Ser Gln Gln Leu Lys Gln Arg	
115 120 125	
gct cag att ctc cct aac gta tgc aac ctt caa tca aga cga tgc gaa	433
Ala Gln Ile Leu Pro Asn Val Cys Asn Leu Gln Ser Arg Arg Cys Glu	
130 135 140	
atc gaa agc gtc agg agt gtt gct gag agc aat ttt gaa atc cca ttt	481
Ile Glu Ser Val Arg Ser Val Ala Glu Ser Asn Phe Glu Ile Pro Phe	
145 150 155 160	
gat atg ccg ttt gat atc cct tgg ccc ttt cgc cca agc tca gag tca	529
Asp Met Pro Phe Asp Ile Pro Trp Pro Phe Arg Pro Ser Ser Glu Ser	
165 170 175	
cag caa tgc aga cag agt gaa atc caa agg cct gtg agt cag tgc caa	577
Gln Gln Cys Arg Gln Ser Glu Ile Gln Arg Pro Val Ser Gln Cys Gln	
180 185 190	
agg tat gtg gag cag caa att cag tcc tcc agg cca tac caa cag agc	625
Arg Tyr Val Glu Gln Gln Ile Gln Ser Ser Arg Pro Tyr Gln Gln Ser	
195 200 205	
ccg tac gac cgg agg caa cag agc cca tac gac cgg agg caa cag agc	673
Pro Tyr Asp Arg Arg Gln Gln Ser Pro Tyr Asp Arg Arg Gln Gln Ser	
210 215 220	
cca tat gaa cag agg caa gga cca tac gaa cag agg cca tac gaa cag	721
Pro Tyr Glu Gln Arg Gln Gly Pro Tyr Glu Gln Arg Pro Tyr Glu Gln	
225 230 235 240	
agg cca tac caa cag cga gga gga cga cag cag gag cag caa ggg ctc	769
Arg Pro Tyr Gln Gln Arg Gly Gly Arg Gln Gln Glu Gln Gln Gly Leu	
245 250 255	
cag caa tgc tgc aac gag ctc caa aac gtg agg agg gag tgt cag tgc	817
Gln Gln Cys Cys Asn Glu Leu Gln Asn Val Arg Arg Glu Cys Gln Cys	
260 265 270	
gag gcg att aag gaa gtg ggc caa aga atg agg cag cag caa caa caa	865
Glu Ala Ile Lys Glu Val Gly Gln Arg Met Arg Gln Gln Gln Gln Gln	
275 280 285	
caa cgt agg cag tat ggt ggg cag cag aca caa act gtg gag aga att	913
Gln Arg Arg Gln Tyr Gly Gly Gln Gln Thr Gln Thr Val Glu Arg Ile	
290 295 300	

59

ctt gag aat ctg cct aac caa tgc gac cta gat gtc cag caa tgc aac 961
 Leu Glu Asn Leu Pro Asn Gln Cys Asp Leu Asp Val Gln Gln Cys Asn
 305 310 315 320

atc ccc tac tga 973
 Ile Pro Tyr

<210> 52

<211> 323

<212> PRT

<213> Helianthus annuus

<400> 52

Ala Lys Ile Thr Leu Leu Leu Leu Ala Leu Ala Ala Leu Val Ala Leu
 1 5 10 15

Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Asp Asp Glu Asn
 20 25 30

Pro Ile Ser Glu Gln Arg Gln Cys Trp Gln Gln Val Gln Gly Gln Arg
 35 40 45

Leu Asn Gln Cys Arg Met Phe Leu Gln Gln Gly Gln Arg Gly Gln Gln
 50 55 60

His Gln Gln Gln Gln His Gln Gln Gln Glu Gln Gln Leu Leu Gln Gln
 65 70 75 80

Cys Cys Gln Glu Leu Gln Asn Ile Glu Gly Gln Cys Gln Cys Glu Ala
 85 90 95

Val Lys Gln Val Val Arg Asp Ala Gln Arg His Glu Gln Gln Arg Pro
 100 105 110

Arg Val Pro Phe Gln Gly Ser Gln Gln Ser Gln Gln Leu Lys Gln Arg
 115 120 125

Ala Gln Ile Leu Pro Asn Val Cys Asn Leu Gln Ser Arg Arg Cys Glu
 130 135 140

Ile Glu Ser Val Arg Ser Val Ala Glu Ser Asn Phe Glu Ile Pro Phe
 145 150 155 160

Asp Met Pro Phe Asp Ile Pro Trp Pro Phe Arg Pro Ser Ser Glu Ser
 165 170 175

Gln Gln Cys Arg Gln Ser Glu Ile Gln Arg Pro Val Ser Gln Cys Gln
 180 185 190

Arg Tyr Val Glu Gln Gln Ile Gln Ser Ser Arg Pro Tyr Gln Gln Ser
 195 200 205

Pro Tyr Asp Arg Arg Gln Gln Ser Pro Tyr Asp Arg Arg Gln Gln Ser
 210 215 220

Pro Tyr Glu Gln Arg Gln Gly Pro Tyr Glu Gln Arg Pro Tyr Glu Gln
 225 230 235 240

Arg Pro Tyr Gln Gln Arg Gly Gly Arg Gln Gln Glu Gln Gln Gly Leu
 245 250 255

Gln Gln Cys Cys Asn Glu Leu Gln Asn Val Arg Arg Glu Cys Gln Cys
 260 265 270

Glu Ala Ile Lys Glu Val Gly Gln Arg Met Arg Gln Gln Gln Gln
 275 280 285

Gln Arg Arg Gln Tyr Gly Gly Gln Gln Thr Gln Thr Val Glu Arg Ile
 290 295 300

60

Leu Glu Asn Leu Pro Asn Gln Cys Asp Leu Asp Val Gln Gln Cys Asn
 305 310 315 320

Ile Pro Tyr

<210> 53

<211> 15

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 53

ttcctcgtct gcgca

15

<210> 54

<211> 17

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 54

tcctcaccaa cgcttcc

17

<210> 55

<211> 18

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 55

cacctaagag cttgccag

18

<210> 56

<211> 16

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 56

gtttgccccca ccttga

16

<210> 57

<211> 22

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 57

atggcgaaca agctcttcct cg

22

<210> 58
<211> 23
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 58
gccttcttct tccttctcac caa 23
<210> 59
<211> 23
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 59
ccaaaatgta ggaaggagtt tca 23
<210> 60
<211> 25
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 60
ctttgcgttt gcccaacctt gaaag 25
<210> 61
<211> 25
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 61
gcaagttagc gtttgtccct tccag 25
<210> 62
<211> 25
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 62
gtcccttcca gaagaccatg cctgg 25
<210> 63
<211> 22
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 63
cacatcatgg agaagatcca ag 22

<210> 64
<211> 29
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 64
agaagaagga cacatgcaga agtgctgca 29

<210> 65
<211> 18
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 65
atgccagtgc aaagcgct 18

<210> 66
<211> 30
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 66
agatcatcac tggatgaatgg tgatcgtgta 30

<210> 67
<211> 31
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 67
ggcaatggaa gcacttagag tgtgctttgt g 31

<210> 68
<211> 19
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 68
gacgagaacc cgatctccg 19

<210> 69
<211> 28
<212> DNA
<213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 69
 tgaaccagtg tcgcatgttc ctccagca 28

 <210> 70
 <211> 23
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 70
 ctccagcagt gttgtcaaga gct 23

 <210> 71
 <211> 30
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 71
 cagtgccaat gtgagggcggg gaagcagggt 30

 <210> 72
 <211> 22
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 72
 ggctcagatt ctcccctaacg ta 22

 <210> 73
 <211> 23
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 73
 tgcaaccttc aatcaagacg atg 23

 <210> 74
 <211> 127
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 74
 atggcdaaca agytsttcct cgtctgcgcr actytgcgyc tstgyttcmt cctcaccaac 60

gcttcctrctc taycgcacyg tygtygartt cgamgaagat gacgccasya accccrtrgg 120
yccaada 127

<210> 75

<211> 39

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 75

aaggagtttc agmaakmmca rcacctaaga gcttgccag 39

<210> 76

<211> 97

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 76

tccagmagtg ctgcargcag cttcgccarg aagagccagw ttgygtttgc cccaccttga 60
racaagctgc carggcmgtt agrytccagg gacarca 97

<210> 77

<211> 59

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 77

caggaaaatt tacmagdcag cyaagyactt gccyaacrtr tgcravatcc mgcaagttg 59

<210> 78

<211> 96

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 78

agccgtatct accagacygc tacgcactta cctaragttt gcaacatyms gcaagtttagc 60
gtttgtccct tccagaagac catgcctggr cctcc 96

<210> 79

<211> 61

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 79

accaggaaga gccmctttgc gtttgcccaa ccttgaaagg agcatccaaa gcsgttaaac 60
a 61

<210> 80

<211> 58

<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 80
cagccggccc atttrggaty ccaaaatgta ggaaggagtt tcarcaagca caacacct 58
<210> 81
<211> 32
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 81
atcgcccaca cytgwgcgc ctccaaatgg ca 32
<210> 82
<211> 44
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 82
gaaatgagcg agctgaraag cccawatgc cagtgcaaag cgct 44
<210> 83
<211> 37
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 83
atcaccacca ccatacgamga cgagaacccg atctccg 37
<210> 84
<211> 43
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein
<400> 84
tacagggaca aaggytgaac cagtgtcgca tgttcctcca gca 43
<210> 85
<211> 45
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 85
caaaacatcg aaggrcagtg ccaatgtgag gcggtgaagc aggtg 45

<210> 86
<211> 52
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 86
caacagttga agcagarggc tcagattctc cctaacgtat gcaaccttca at 52

<210> 87
<211> 17
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 87
gaggctggtc gcatcga 17

<210> 88
<211> 20
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 88
ttccgtgata tgcaccagaa 20

<210> 89
<211> 19
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 89
tggttaacga caacggtga 19

<210> 90
<211> 17
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: sequence
motive from plant storage protein

<400> 90
gaggctggtc gggtcga 17

<210> 91
<211> 20
<212> DNA
<213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 91
 ttccgtgaca tgcaccagaa 20

 <210> 92
 <211> 19
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 92
 tggatgaacga caacggaca 19

 <210> 93
 <211> 35
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 93
 carggbttcc gtgayatgca ccagaargtv garca 35

 <210> 94
 <211> 20
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 94
 attgagacat ggaaccctaa 20

 <210> 95
 <211> 20
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 95
 tggatgtaca acaatgaaga 20

 <210> 96
 <211> 22
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 96
 cactcctggtt gttgccgttt ct 22

<210> 97
<211> 24
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
<400> 97
attctatctt gctggaacc aaga 24
<210> 98
<211> 26
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
<400> 98
gccattgtga cagtgaagg aggtct 26
<210> 99
<211> 26
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
<400> 99
agggaatgca gtgttcgacg gtgagc 26
<210> 100
<211> 20
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
<400> 100
aaacatggaa ctctcaacac 20
<210> 101
<211> 20
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
<400> 101
tggacctata acactggcga 20
<210> 102
<211> 22
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 102
 tgaaccagtt gttgccatca gt 22
 <210> 103
 <211> 24
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
 <400> 103
 attttacctt gctgggaacc caga 24
 <210> 104
 <211> 26
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
 <400> 104
 ggtgagagag tgtttgatgg agagct 26
 <210> 105
 <211> 1114
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: DNA
 construct coding for dsRNA
 <400> 105
 ggccgcgtgt tccatttggc cggaacaac cagcagggag gctttggcgg ttcacagcaa 60
 caacaagaac agaaaaactt gtggagcggg ttcgacgcac aggtcatagc tcaagcattg 120
 aaaattgacg ttcagttggc tcagcagctt cagaaccaac aagacagcag aggaaacatc 180
 gttcgtgtta agggacctt ccaggctcgtg aggccacctc taagacagcc ctacgagagc 240
 gaggagtggg gacaccacg tagcccacag ggcaacggcc ttgaggagac tatctgcagc 300
 atgaggtccc acgagaacat tgacgaccct gctcgtgctg acgtgtacaa gccagccta 360
 ggtcgcgtga ccagcgtcaa cagctatacc ttgcccattc tggagtatgt caggctcagt 420
 gccactcgtg gcgttctcca ggggtgatcc ttctgtaaca tttgacaaaa catgtgaaca 480
 cgtcatccgt catatagaac ttccaatttt aatatgtttt gctaaagaaa aaaaaaggga 540
 ataaatatct atcaaattca tttttaaaac atttgtatac gttcttaaata aatttaggat 600
 atgactaatt tttctttttg gtaaaaatgt taatatctat atttaattta ttaagaaaaa 660
 tgtacttaca ccctggagaa cgccacgagt ggcactgagc ctgacatact ccaagatggg 720
 caaggtatag ctggttgacg tggtcacgag acctaggctg ggcttgtaga cgtcagcacg 780
 agcagggctg tcaatgttct cgtgggacct catgctgcag atagtctcct caaggccgtt 840
 gccctgtggg ctacgtgggt gtctccactc ctcgctctcg tagggctgtc ttagaggtgg 900
 cctcacgacc tggaaagggt ccttaacacg aacgatgttt cctctgctgt cttgttggtt 960
 ctgaagctgc tgagccaact gaacgtcaat tttcaatgct tgagctatga cctgtgcgtc 1020
 gaaccgctc cacaagtttt tctgttcttg ttgttgctgt gaaccgcaa agcctccctg 1080
 ctggttggtt ccggccaaat ggaacacgag gccg 1114
 <210> 106
 <211> 1114
 <212> RNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: RNA sequence
 for forming dsRNA

<400> 106

```

ggccgcgugu uccauuuggc cggaacaac cagcagggag gcuuuggcgg uucacagcaa 60
caacaagaac agaaaaacuu guggagcggg uucgacgcac aggucauagc ucaagcauug 120
aaaauugacg uucaguuggc ucagcagcuu cagaaccaac aagacagcag aggaacauc 180
guucguguua agggaccuuu ccaggucgug aggccaccuc uaagacagcc cuacgagagc 240
gaggagugga gacaccacg uagccacag ggcaacggcc uugaggagac uaucugcagc 300
augagguccc acgagaacau ugacgaccu gcucgugcug acguguacaa gccagccua 360
ggucgcguga ccagcgucua cagcuauacc uugcccaucu uggaguaugu caggcucagu 420
gccacucgug gcuucucca ggguggaucc uucuguaaca uuugacaaa caugugaaca 480
cgucauccgu cauauagaac uucaauuuu aaauuguuuu gcuaaagaaa aaaaaaggga 540
auaaaauauc aucaauuca uuuuuuuuac auuuguauac guucuuuuu auuuuaggau 600
augacuaauu uuucuuuuug guaaaaaugu uaauaucuau auuuuuuuu uuaagaaaa 660
uguacuuaca ccuggagaa cgccacgagu ggcacugagc cugacauacu ccaagauagg 720
caagguauag cuugagcgc uggucacgcg accuaggcug ggcuuuguaca cgucagcacg 780
agcagggucg ucaauguucu cgugggaccu caugcugcag auagucuccu caaggccguu 840
gcccuguggg cuacgugggu gucuccacuc cucgcucugc uagggcuguc uuagaggugg 900
ccucacgacc uggaaagguc ccuaaacacg aacgauguu ccucugcugu cuuguugguu 960
cugaagcugc ugagccaacu gaacgucaau uucaaugcu ugagcuanga ccugugcugc 1020
gaaccgcuc cacaaguuuu ucuguucuu uuguugcugu gaaccgcaa agccuccug 1080
cugguuguuu ccggccaaau ggaacacgcg gccg 1114

```

<210> 107

<211> 1789

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: DNA
construct coding for dsRNA

<400> 107

```

gcggccgcgg atcctcaggg tcttttcttg cccactttct tgaacgccgg caaactcacg 60
tttgttggtc acggaagggg tctaattgga agagtatttc cgggatgcgc cgagacgttc 120
atggagtcac cggatatttg agaaggtcaa ggtcagggtc agagtcaagg gttccgtgac 180
atgcaccaga aagtagagca cctacgggtg ggtgacacca ttgcaacacc atctggtgta 240
gctcaatggt tctacaacaa tggaaatgag cctctcattc ttgttgagc cgcggtatctc 300
gccagcaacc agaaccagct tgaccgcaac cttagaccat ttttgatagc cggaacaac 360
ccacaagggc aggaatggct acaaggccga aagcaacaga agcaaaaca catcttcaat 420
ggcttcgcac ctgagatctt ggctcaagcc ttcaagatca atgtcgagac ggctcagcag 480
ctccagaacc agcaagataa ccgtggcaac atcgtcaagg tcaacggacc tttcggcgtc 540
attaggccac ccttgagacg cggcgaaggc ggccaacaac cacatgaaat agctaattggt 600
ttagaggaga ctttgtgcac catgcgatgc actgaaaacc tcgatgacc gtcggatgct 660
gacgtgtaca agccatcact cggatacatt agcacactta acagctacaa tcttctatc 720
ctcagacttc tccgccttag cgctcttcgt ggctccatcc gtaaaactcg aggtaagctc 780
aacaaatctt tagaaaatta attttatgtg acatatgcaa taatttgatt tggcaagata 840
aactaataga ttttgcgatt tggagtttta aactctaaat aatctaaatc gttttcaatt 900
ggttttaaata tatactttgc atttttaatc gtttttaatt aaaaaatata tatatatata 960
tatatcttgc atttttaatc gttttcaatt taaaaatat cttgcacgca gaacgctgtc 1020
gagttttacg gatggagcca cgaagagcgc taaggcggag aagtctgagg ataggaagat 1080
tgtagctgtt aagtgtgcta atgtatccga gtgatggctt gtacacgtca gcatccgacg 1140
ggctcatcgag gttttcagtg catcgcatgg tgcacaaagt ctctctaaa ccattagcta 1200
tttcatgttg ttgttgccg ccttcgccgc gtctcaaggg tggcctaatt acgccgaaag 1260
gtccgttgac cttgacgatg ttgccacggg tatcttgctg gttctggagc tgctgagccg 1320
tctcgacatt gatcttgaag gcttgagcca agatctcagg tgcgaagcca ttgaagatgt 1380
tgttttgctt ctgttgcttt cggccttgta gccattcctg cccttggtggg ttgtttccgg 1440
ctatcaaaaa tgggtctaagg ttgcggtcaa gctggttctg gttgctggcg agatccgcgg 1500
ctgcaacaag aatgagaggc tcatttccat tggtgtagaa ccattgagct acaccagatg 1560
gtgttgcaat ggtgtcacgg caccgtaggg gctctacttt ctggtgcatg tcacggaacc 1620
cttgactctg accctgacct tgaccttctc caaataccgg tgactccatg aacgtctcgg 1680

```

71

cgcatcccg aataactctt cccattagac cccttccgtg aacaacaaac gtgagtttgc 1740
 cggcggttcaa gaaagtgggc aagaaaagac cctgaggatc cgcgggccgc 1789

<210> 108

<211> 1789

<212> RNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: : RNA
 sequence for forming dsRNA

<400> 108

gcgggccgcgg auccucaggg ucuuuuucug cccacuuucu ugaacgccgg caaacucacg 60
 uuuguuguuc acggaagggg ucuaauggga agaguuauc cggaugcgc cgagacguuc 120
 auggagucac cgguaauugg agaaggucac ggucaggguc agagucagg guuccgugac 180
 augcaccaga aaguagagca ccuacggugc ggugacacca uugcaacacc aucuggugua 240
 gcucaauggu ucuaacaacaa uggaaaugag ccucucuuuc uuguugcagc cgcggaucuc 300
 gccagcaacc agaaccagcu ugaccgcaac cuuagaccac uuuugauagc cggaacaac 360
 ccacaagggc aggaauggcu acaaggccga aagcaacaga agcaaaacaa caucucaau 420
 ggcuucgcac cugagaucuu ggcucaagcc uucaagauc augucgagac ggcucagcag 480
 cuccagaacc agcaagauaa ccguggcaac aucgucaagg ucaacggacc uuucggcguc 540
 auuaggccac ccuugagacg cggcgaaggc ggccaacaac cacaugaaau agcuauuggu 600
 uuagaggaga cuuugugcac caugcgaugc acugaaaacc ucgaugacc gucggagcu 660
 gacguguaca agccaucacu cggauacau agcacacua acagcuaca ucuccuau 720
 cucagacuuc uccgccuuag cgcucuucgu ggcuccauc guaaaacuc agguagcuc 780
 aaaaaucuu uagaaaaua auuuuauug acaaugcaa uauuuugau uggcaagau 840
 aacuaauaga uuugcgauu uggaguuaa aacucuaaa aaucuaaauc guuuucauu 900
 gguuuuuaua uauaucuugc auuuuuuau guuuuuuau aaaaaaua uauauaua 960
 uauaucuugc auuuuuuau guuuucauu uaaaaaua cuugcacgca gaacgcuguc 1020
 gaguuuuacg gauggagcca cgaagagcgc uaaggcggag aagucugagg auaggaagau 1080
 uguagcuguu aagugugcua auguaucgga gugauggcu guacacguca gcauccgacg 1140
 ggucaucgag guuuucagug caucgcaugc ugcacaaagu cuccucuaaa ccauugcua 1200
 uuucaugugg uuguuggccg ccuucgcgc gucucaagg uggccuaug acgccgaaag 1260
 guccguugac cuugacgaug uggccaggu uaucuugcug guucuggagc ugcugagccg 1320
 ucucgacauu gaudiugaag gcuugagcca agaucucagg ugcgaagcca uugaagaugu 1380
 uguuuugcuu cugugcuuu gggccuugua gccauuccug ccuuguggg uuguuuccg 1440
 cuaucaaaa uggucuaagg uggcggucaa gcugguucug guugcuggcg agauccgagg 1500
 cugcaacaag aaugagaggg ucauuuccau uguuguagaa ccauugagcu acaccagaug 1560
 gugugcaau ggugucaccg caccguaggu gcucuaucuu cuggugcaug ucacggaacc 1620
 cuugacucug acccugaccu ugaccuucuc caauuaccg ugacuccaug aacgucucgg 1680
 cgcaucccg aauaacucuu ccauuagac ccuuccgug aacaacaaac gugaguugc 1740
 cggcgaucaa gaaagugggc aagaaaagac ccugaggau cgcgggccgc 1789

<210> 109

<211> 1273

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: DNA
 construct coding for dsRNA

<400> 109

goggccgcgg atccatggct aacaagctct tcctcgtctg cgcaactctc gccctctgct 60
 tcctcctcac caacgcttcc atctaccgca ccgttgctga attcgaagaa gatgacgcca 120
 gcaaccccg aggtccaaga cagagatgcc agaaggagtt tcagcaatca caacacctaa 180
 gagcttgcca gagatggatg agcaagcaaa tgaggcaagg acgtggtggt ggtccttccc 240
 tcgacgatga gttcgatttc gagggccccc agcagggata ccagctactc cagcagtgc 300
 gcaacgagct tcgccaggaa gagccagttt gcgtttgccc caccttgaaa caagctgcca 360
 gggcagttag cctccaggga cagcacggac cattccaatc caggaaaatt taccagtgc 420

```

ctaagtacttt gcctaacatt tgcaagatcc agcaagttgg tgaatgtccc ttccagacca 480
ccatccctttt cttccctcct tactactagg gtactcgagg taagctcaac aaatcttttag 540
aaaatttaatt ttatgtgaca tatgcaataa tttgatttgg caagataaac taatagattt 600
tgcgatttgg agtttttaaac tctaaataat ctaaactcgtt ttcaattggg ttaaataatat 660
atcttgcatt tttaatcgtt tttaatataa aaatatatat atatatatat atcttgcatt 720
tttaactcgtt ttcaatttaa aaaatatctt gcacgcagaa cgctgtcgac taccctagta 780
gtaaggaggg aagaaaggga tgggtggtctg gaaggacat tcaccaactt gctggatctt 840
gcaaatgtta ggcaagtact tagctgactg gtaaattttc ctggattgga atgggtccgtg 900
ctgtccctgg aggctaactg ccctggcagc ttgtttcaag gtgggggcaaa cgcaactgg 960
ctcttcctgg cgaagctcgt tgcagcactg ctggagtagc tggatatccct gctggggggcc 1020
ctcgaaatcg aactcatcgt cgagggaagg accaccacca cgtccttgcc tcatttgctt 1080
gtcatccat ctctggcaag ctcttaggtg ttgtgattgc tgaaactcct tctggcatct 1140
ctgtcttggga cctacggggt tgctggcgtc atcttcttcg aattcgacaa cgggtcggta 1200
gatggaagcg ttggtgagga ggaagcagag ggcgagagtt gcgcagacga ggaagagctt 1260
gtagccatg gat 1273

```

<210> 110

<211> 1273

<212> RNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: : RNA
sequence for forming dsRNA

<400> 110

```

gcggccgcgg auccauggcu aacaagcucu uccucgucug cgcaacucuc gcccucugcu 60
uccuccucac caacgcuucc aucuaccgca ccguugucga auucgaagaa gaugacgcca 120
gcaaccccggu agguccaaga cagagaugcc agaaggaguu ucagcaauca caacaccuaa 180
gagcuugcca gagauggaug agcaagcaaa ugaggcaagg acgugguggu gguccuuccc 240
ucgacgauga guucgauuuc gagggccccc agcagggaua ccagcuacuc cagcagugcu 300
gcaacgagcu ucgccaggaa gagccaguuu gcguuugccc caccuugaaa caagcugcca 360
gggcaguuag ccuccaggga cagcacggac cauuccaau caggaaaauu uaccagucag 420
cuaaguacuu gccuaacauu ugcaagauc agcaaguugg ugaauugccc uuccagacca 480
ccauccuuu cuucccuccu uacuacuagg guacucgagg uaagcucaac aaauuuuag 540
aaaauuauu uuaugugaca uaugcaaua uuugauuugg caagauaaac uaauagauuu 600
ugcgauuugg aguuuuuuaa ucuaauuau cuaaaucguu uucaauuggu uuaauauau 660
aucuugcauu uuuaaucguu uuuaauuaa aaauauauau auauauauau aucuugcauu 720
uuuaaucguu uucaauuuaa aaaauaucuu gcacgcagaa cgcugucgac uaccuagua 780
guaaggaggg aagaaaggga ugguggucug gaagggacau ucaccaacu gcuggaucuu 840
gcaaauguaa ggcaaguacu uagcugacug guaaaauuuc cuggauugga augguccgug 900
cugucccugg aggcuaacug ccuggcagc uuguuucag guggggcaaa cgcaaacugg 960
cucuuccugg cgaagcucgu ugcagcacug cuggaguagc ugguaucucc gcuggggggcc 1020
cucgaaaucg aacucaucgu cgagggaagg accaccacca cguccuugcc ucauuugcuu 1080
gcucauccau cucuggcaag cucuuaggug uugugauugc ugaaacuccu ucuggcaucu 1140
cugucuugga ccuacggggu ugcuggcguc aucuucucg aaucgacaa cggugcgguu 1200
gauggaagcg uuggugagga ggaagcagag ggcgagaguu gcgcagacga ggaagagcuu 1260
guuagccaug gau 1273

```

<210> 111

<211> 1575

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1572)

<223> 12S cruciferin

<400> 111

```

atg gtt aag ctc agc aat ctc ctc gtt gca acc ttc ggg gtt ctc ctc 48

```

73

Met	Val	Lys	Leu	Ser	Asn	Leu	Leu	Val	Ala	Thr	Phe	Gly	Val	Leu	Leu		
1				5					10					15			
gtc	ctt	aac	ggc	tgc	ctt	gcg	agg	cag	tca	ctt	ggg	gtt	cct	cct	cag	96	
Val	Leu	Asn	Gly	Cys	Leu	Ala	Arg	Gln	Ser	Leu	Gly	Val	Pro	Pro	Gln		
		20						25					30				
cta	cag	aac	gag	tgt	aac	ctc	gac	aac	cta	gat	gtt	ctc	caa	gcc	acc	144	
Leu	Gln	Asn	Glu	Cys	Asn	Leu	Asp	Asn	Leu	Asp	Val	Leu	Gln	Ala	Thr		
		35					40					45					
gaa	act	atc	aag	agt	gaa	gcc	ggt	cag	atc	gag	tac	tgg	gac	cac	aac	192	
Glu	Thr	Ile	Lys	Ser	Glu	Ala	Gly	Gln	Ile	Glu	Tyr	Trp	Asp	His	Asn		
	50					55				60							
cac	cct	cag	ctc	cga	tgt	gtt	ggt	gtt	tcc	gtt	gct	cgt	tat	gta	att	240	
His	Pro	Gln	Leu	Arg	Cys	Val	Gly	Val	Ser	Val	Ala	Arg	Tyr	Val	Ile		
	65				70				75						80		
gaa	caa	ggc	ggt	ctt	tac	ttg	ccc	acc	ttc	ttc	act	tcc	cca	aaa	att	288	
Glu	Gln	Gly	Gly	Leu	Tyr	Leu	Pro	Thr	Phe	Phe	Thr	Ser	Pro	Lys	Ile		
				85					90					95			
tcc	tac	gtc	gtt	caa	gga	acg	ggt	atc	agc	gga	aga	gtg	gtc	cct	gga	336	
Ser	Tyr	Val	Val	Gln	Gly	Thr	Gly	Ile	Ser	Gly	Arg	Val	Val	Pro	Gly		
		100						105					110				
tgt	gcc	gag	acc	ttc	atg	gac	tcg	cag	ccg	atg	caa	gga	caa	caa	caa	384	
Cys	Ala	Glu	Thr	Phe	Met	Asp	Ser	Gln	Pro	Met	Gln	Gly	Gln	Gln	Gln		
	115					120						125					
ggc	caa	cca	tgg	caa	gga	cga	cag	gga	caa	caa	ggc	caa	cca	tgg	gaa	432	
Gly	Gln	Pro	Trp	Gln	Gly	Arg	Gln	Gly	Gln	Gln	Gly	Gln	Pro	Trp	Glu		
	130					135					140						
gga	cag	gga	caa	cag	gga	caa	caa	gga	aga	caa	ggc	caa	cca	tgg	gaa	480	
Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Arg	Gln	Gly	Gln	Pro	Trp	Glu		
	145				150					155					160		
gga	cag	gga	caa	cag	gga	caa	caa	gga	cga	cag	gga	caa	caa	ggc	caa	528	
Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Arg	Gln	Gly	Gln	Gln	Gly	Gln		
			165					170						175			
cca	tgg	gaa	gga	cag	gga	cag	cag	gga	caa	caa	ggg	ttc	cgt	gac	atg	576	
Pro	Trp	Glu	Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Phe	Arg	Asp	Met		
		180						185					190				
cac	cag	aag	gtg	gaa	cat	gtg	aga	cgc	gga	gac	gtc	ttt	gcc	aac	act	624	
His	Gln	Lys	Val	Glu	His	Val	Arg	Arg	Gly	Asp	Val	Phe	Ala	Asn	Thr		
		195					200					205					
cca	ggc	tct	gcc	cac	tgg	atc	tac	aac	tca	gga	gaa	cag	cca	ctt	gtc	672	
Pro	Gly	Ser	Ala	His	Trp	Ile	Tyr	Asn	Ser	Gly	Glu	Gln	Pro	Leu	Val		
	210					215						220					
atc	atc	gct	ctt	ctc	gac	atc	gcc	aac	tac	caa	aac	caa	ctc	gac	cgc	720	
Ile	Ile	Ala	Leu	Leu	Asp	Ile	Ala	Asn	Tyr	Gln	Asn	Gln	Leu	Asp	Arg		
	225				230					235				240			
aac	cct	aga	gtg	ttc	cat	ttg	gcc	gga	aac	aac	cag	cag	gga	ggc	ttt	768	
Asn	Pro	Arg	Val	Phe	His	Leu	Ala	Gly	Asn	Asn	Gln	Gln	Gly	Gly	Phe		
				245				250						255			
ggc	ggt	tca	cag	caa	caa	caa	gaa	cag	aaa	aac	ttg	tgg	agc	ggg	ttc	816	
Gly	Gly	Ser	Gln	Gln	Gln	Gln	Glu	Gln	Lys	Asn	Leu	Trp	Ser	Gly	Phe		
			260					265					270				

gac gca cag gtc ata gct caa gca ttg aaa att gac gtt cag ttg gct	864
Asp Ala Gln Val Ile Ala Gln Ala Leu Lys Ile Asp Val Gln Leu Ala	
275 280 285	
cag cag ctt cag aac caa caa gac agc aga gga aac atc gtt cgt gtt	912
Gln Gln Leu Gln Asn Gln Gln Asp Ser Arg Gly Asn Ile Val Arg Val	
290 295 300	
aag gga cct ttc cag gtc gtg agg cca cct cta aga cag ccc tac gag	960
Lys Gly Pro Phe Gln Val Val Arg Pro Pro Leu Arg Gln Pro Tyr Glu	
305 310 315 320	
agc gag gag tgg aga cac cca cgt agc cca cag ggc aac ggc ctt gag	1008
Ser Glu Glu Trp Arg His Pro Arg Ser Pro Gln Gly Asn Gly Leu Glu	
325 330 335	
gag act atc tgc agc atg agg tcc cac gag aac att gac gac cct gct	1056
Glu Thr Ile Cys Ser Met Arg Ser His Glu Asn Ile Asp Asp Pro Ala	
340 345 350	
cgt gct gac gtg tac aag ccc agc cta ggt cgc gtg acc agc gtc aac	1104
Arg Ala Asp Val Tyr Lys Pro Ser Leu Gly Arg Val Thr Ser Val Asn	
355 360 365	
agc tat acc ttg ccc atc ttg gag tat gtc agg ctc agt gcc act cgt	1152
Ser Tyr Thr Leu Pro Ile Leu Glu Tyr Val Arg Leu Ser Ala Thr Arg	
370 375 380	
ggc gtt ctc cag ggt aat gcg atg gtg ctt cct aaa tac aac atg aac	1200
Gly Val Leu Gln Gly Asn Ala Met Val Leu Pro Lys Tyr Asn Met Asn	
385 390 395 400	
gct aac gag atc ttg tac tgc act gga gga caa gga agg atc caa gtg	1248
Ala Asn Glu Ile Leu Tyr Cys Thr Gly Gly Gln Gly Arg Ile Gln Val	
405 410 415	
gtc aac gac aac gga cag aac gtg ttg gac caa cag gtg cag aag gga	1296
Val Asn Asp Asn Gly Gln Asn Val Leu Asp Gln Gln Val Gln Lys Gly	
420 425 430	
cag ctc gtg gtc atc cca caa ggg ttc gca tac gtt gtc cag tcc cac	1344
Gln Leu Val Val Ile Pro Gln Gly Phe Ala Tyr Val Val Gln Ser His	
435 440 445	
gga aac aag ttc gag tgg atc tct ttc aaa act aat gaa aac gca atg	1392
Gly Asn Lys Phe Glu Trp Ile Ser Phe Lys Thr Asn Glu Asn Ala Met	
450 455 460	
atc agc act ttg gcg ggt aga acc tcg ctc ttg agg gca ttg cca ttg	1440
Ile Ser Thr Leu Ala Gly Arg Thr Ser Leu Leu Arg Ala Leu Pro Leu	
465 470 475 480	
gag gtc ata tca aat ggt ttc cag atc tct ccc gag gaa gct agg aag	1488
Glu Val Ile Ser Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Lys	
485 490 495	
atc aag ttc aac aca ctt gag acc act ttg acc cgc gct gcc ggt agg	1536
Ile Lys Phe Asn Thr Leu Glu Thr Thr Leu Thr Arg Ala Ala Gly Arg	
500 505 510	
caa caa caa cag ttg atc gag gag att gtc gag gct taa	1575
Gln Gln Gln Gln Leu Ile Glu Glu Ile Val Glu Ala	
515 520	

<210> 112

<211> 524

75

<212> PRT

<213> Arabidopsis thaliana

<400> 112

```

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

```

76

Arg Ala Asp Val Tyr Lys Pro Ser Leu Gly Arg Val Thr Ser Val Asn
 355 360 365

Ser Tyr Thr Leu Pro Ile Leu Glu Tyr Val Arg Leu Ser Ala Thr Arg
 370 375 380

Gly Val Leu Gln Gly Asn Ala Met Val Leu Pro Lys Tyr Asn Met Asn
 385 390 395 400

Ala Asn Glu Ile Leu Tyr Cys Thr Gly Gly Gln Gly Arg Ile Gln Val
 405 410 415

Val Asn Asp Asn Gly Gln Asn Val Leu Asp Gln Gln Val Gln Lys Gly
 420 425 430

Gln Leu Val Val Ile Pro Gln Gly Phe Ala Tyr Val Val Gln Ser His
 435 440 445

Gly Asn Lys Phe Glu Trp Ile Ser Phe Lys Thr Asn Glu Asn Ala Met
 450 455 460

Ile Ser Thr Leu Ala Gly Arg Thr Ser Leu Leu Arg Ala Leu Pro Leu
 465 470 475 480

Glu Val Ile Ser Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Lys
 485 490 495

Ile Lys Phe Asn Thr Leu Glu Thr Thr Leu Thr Arg Ala Ala Gly Arg
 500 505 510

Gln Gln Gln Gln Leu Ile Glu Glu Ile Val Glu Ala
 515 520

<210> 113

<211> 1419

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1416)

<223> 12S Cra1 storage protein

<400> 113

atg gct cga gtc tct tct ctt ctt tct ttc tgc tta aca ctt ttg atc 48
 Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile
 1 5 10 15

ctt ttc cat ggc tac gcg gct caa cag ggt cag cag ggt cag cag ttt 96
 Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe
 20 25 30

ccg aac gag tgc cag ctc gac cag ctc aat gcg ctc gag ccg tca cac 144
 Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His
 35 40 45

gta ctg aag agc gag gct ggt cgc atc gag gtg tgg gac cac cac gct 192
 Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala
 50 55 60

cct cag ctc cgt tgc tca ggt gtc tcc ttt gca cgt tac atc atc gag 240
 Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu
 65 70 75 80

tct aag ggt ctc tac ttg ccc tct ttc ttt aac acc gcg aag ctc tct 288
 Ser Lys Gly Leu Tyr Leu Pro Ser Phe Phe Asn Thr Ala Lys Leu Ser
 85 90 95

ttc	gtg	gct	aag	gga	cga	ggt	ctt	atg	gga	aaa	gtg	atc	cct	gga	tgc	336
Phe	Val	Ala	Lys	Gly	Arg	Gly	Leu	Met	Gly	Lys	Val	Ile	Pro	Gly	Cys	
			100					105					110			
gcc	gaa	aca	ttc	caa	gac	tca	tca	gag	ttc	caa	cca	cgc	ttc	gaa	ggt	384
Ala	Glu	Thr	Phe	Gln	Asp	Ser	Ser	Glu	Phe	Gln	Pro	Arg	Phe	Glu	Gly	
			115					120					125			
caa	ggt	caa	agc	cag	agg	ttc	cgt	gac	atg	cac	cag	aaa	gtg	gag	cac	432
Gln	Gly	Gln	Ser	Gln	Arg	Phe	Arg	Asp	Met	His	Gln	Lys	Val	Glu	His	
			130					135					140			
att	agg	agc	ggt	gat	acc	att	gcc	aca	aca	ccc	ggt	gta	gca	cag	tgg	480
Ile	Arg	Ser	Gly	Asp	Thr	Ile	Ala	Thr	Thr	Pro	Gly	Val	Ala	Gln	Trp	
			145					150					155			160
ttc	tac	aac	gac	gga	cag	cag	cca	ctt	gtc	atc	gtc	agc	gtc	ttc	gat	528
Phe	Tyr	Asn	Asp	Gly	Gln	Gln	Pro	Leu	Val	Ile	Val	Ser	Val	Phe	Asp	
				165					170						175	
cta	gcc	agt	cac	cag	aac	cag	ctt	gac	cgc	aac	cca	agg	cca	ttt	tac	576
Leu	Ala	Ser	His	Gln	Asn	Gln	Leu	Asp	Arg	Asn	Pro	Arg	Pro	Phe	Tyr	
			180					185						190		
tta	gcc	gga	aac	aac	cca	caa	ggt	caa	gta	tgg	cta	caa	gga	cga	gag	624
Leu	Ala	Gly	Asn	Asn	Pro	Gln	Gly	Gln	Val	Trp	Leu	Gln	Gly	Arg	Glu	
			195					200						205		
caa	cag	cca	cag	aag	aac	att	ttc	aat	gga	ttt	gga	ccc	gag	ggt	att	672
Gln	Gln	Pro	Gln	Lys	Asn	Ile	Phe	Asn	Gly	Phe	Gly	Pro	Glu	Val	Ile	
			210					215					220			
gct	caa	gct	ttg	aag	atc	gat	ctt	cag	aca	gca	cag	caa	ctt	cag	aac	720
Ala	Gln	Ala	Leu	Lys	Ile	Asp	Leu	Gln	Thr	Ala	Gln	Gln	Leu	Gln	Asn	
			225					230					235			240
caa	gat	gac	aac	cgt	gga	aac	att	gtc	cga	gtc	caa	gga	ccg	ttc	ggt	768
Gln	Asp	Asp	Asn	Arg	Gly	Asn	Ile	Val	Arg	Val	Gln	Gly	Pro	Phe	Gly	
				245					250						255	
gtc	att	agg	ccg	cct	ttg	agg	ggc	cag	aga	cct	cag	gag	gag	gaa	gaa	816
Val	Ile	Arg	Pro	Pro	Leu	Arg	Gly	Gln	Arg	Pro	Gln	Glu	Glu	Glu	Glu	
			260					265							270	
gaa	gaa	gga	cga	cat	gga	cga	cac	ggt	aat	ggc	tta	gag	gag	acc	atc	864
Glu	Glu	Gly	Arg	His	Gly	Arg	His	Gly	Asn	Gly	Leu	Glu	Glu	Thr	Ile	
			275					280							285	
tgc	agc	gcc	agg	tgc	acc	gat	aac	ctc	gat	gac	ccg	tct	cgt	gct	gac	912
Cys	Ser	Ala	Arg	Cys	Thr	Asp	Asn	Leu	Asp	Asp	Pro	Ser	Arg	Ala	Asp	
			290					295					300			
gtg	tac	aag	cca	cag	ctc	ggt	tac	atc	agc	act	ctc	aac	agt	tac	gat	960
Val	Tyr	Lys	Pro	Gln	Leu	Gly	Tyr	Ile	Ser	Thr	Leu	Asn	Ser	Tyr	Asp	
			305					310					315			320
ctc	ccc	att	ctt	cgc	ttc	atc	cgt	ctc	tca	gcc	ctc	cgt	gga	tct	atc	1008
Leu	Pro	Ile	Leu	Arg	Phe	Ile	Arg	Leu	Ser	Ala	Leu	Arg	Gly	Ser	Ile	
				325					330						335	
cgt	caa	aac	gca	atg	gtg	ctt	cca	cag	tgg	aac	gca	aac	gcg	aac	gct	1056
Arg	Gln	Asn	Ala	Met	Val	Leu	Pro	Gln	Trp	Asn	Ala	Asn	Ala	Asn	Ala	
			340					345							350	
att	ctt	tac	gag	aca	gac	ggg	gaa	gcc	caa	atc	cag	atc	gta	aac	gac	1104
Ile	Leu	Tyr	Glu	Thr	Asp	Gly	Glu	Ala	Gln	Ile	Gln	Ile	Val	Asn	Asp	
			355					360							365	

78

aat ggt aac aga gtg ttt gac gga caa gtc tct caa gga cag ctc ata 1152
 Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile
 370 375 380

gcc gta cca caa ggt ttc tcg gtg gtg aaa cgc gca aca agc aac cga 1200
 Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg
 385 390 395 400

ttc cag tgg gtt gag ttc aaa aca aac gct aac gcg caa atc aac act 1248
 Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr
 405 410 415

ctg gcg gga cga acc tca gtc ttg aga ggt tta cca ctt gaa gtc ata 1296
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile
 420 425 430

acc aat ggg ttc caa atc tca ccc gaa gaa gca agg agg gtc aag ttc 1344
 Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe
 435 440 445

aac acg ctc gag acc act ttg act cac agc agt ggc cca gct agc tac 1392
 Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr
 450 455 460

gga agg cca aga gtg gct gca gct taa 1419
 Gly Arg Pro Arg Val Ala Ala Ala
 465 470

<210> 114
 <211> 472
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 114

Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile
 1 5 10 15

Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe
 20 25 30

Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His
 35 40 45

Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala
 50 55 60

Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu
 65 70 75 80

Ser Lys Gly Leu Tyr Leu Pro Ser Phe Phe Asn Thr Ala Lys Leu Ser
 85 90 95

Phe Val Ala Lys Gly Arg Gly Leu Met Gly Lys Val Ile Pro Gly Cys
 100 105 110

Ala Glu Thr Phe Gln Asp Ser Ser Glu Phe Gln Pro Arg Phe Glu Gly
 115 120 125

Gln Gly Gln Ser Gln Arg Phe Arg Asp Met His Gln Lys Val Glu His
 130 135 140

Ile Arg Ser Gly Asp Thr Ile Ala Thr Thr Pro Gly Val Ala Gln Trp
 145 150 155 160

Phe Tyr Asn Asp Gly Gln Gln Pro Leu Val Ile Val Ser Val Phe Asp
 165 170 175

Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr
 180 185 190

79

Leu Ala Gly Asn Asn Pro Gln Gly Gln Val Trp Leu Gln Gly Arg Glu
 195 200 205
 Gln Gln Pro Gln Lys Asn Ile Phe Asn Gly Phe Gly Pro Glu Val Ile
 210 215 220
 Ala Gln Ala Leu Lys Ile Asp Leu Gln Thr Ala Gln Gln Leu Gln Asn
 225 230 235 240
 Gln Asp Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe Gly
 245 250 255
 Val Ile Arg Pro Pro Leu Arg Gly Gln Arg Pro Gln Glu Glu Glu Glu
 260 265 270
 Glu Glu Gly Arg His Gly Arg His Gly Asn Gly Leu Glu Glu Thr Ile
 275 280 285
 Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Arg Ala Asp
 290 295 300
 Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp
 305 310 315 320
 Leu Pro Ile Leu Arg Phe Ile Arg Leu Ser Ala Leu Arg Gly Ser Ile
 325 330 335
 Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala
 340 345 350
 Ile Leu Tyr Glu Thr Asp Gly Glu Ala Gln Ile Gln Ile Val Asn Asp
 355 360 365
 Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile
 370 375 380
 Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg
 385 390 395 400
 Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr
 405 410 415
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile
 420 425 430
 Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe
 435 440 445
 Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr
 450 455 460
 Gly Arg Pro Arg Val Ala Ala Ala
 465 470

<210> 115

<211> 1419

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1416)

<223> At5g442120/MLN1_4 storage protein

<400> 115

atg gct cga gtc tct tct ctt ctt tct ttc tgc tta aca ctt ttg atc 48
 Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile
 1 5 10 15

80

ctt ttc cat ggc tac gcg gct caa cag ggt cag cag ggt cag cag ttt	96
Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe	
20 25 30	
ccg aac gag tgc cag ctc gac cag ctc aat gcg ctc gag ccg tca cac	144
Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His	
35 40 45	
gta ctg aag agc gag gct ggt cgc atc gag gtg tgg gac cac cac gct	192
Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala	
50 55 60	
cct cag ctc cgt tgc tca ggt gtc tcc ttt gca cgt tac atc atc gag	240
Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu	
65 70 75 80	
tct aag ggt ctc tac ttg ccc tct ttc ttt aac acc gcg aag ctc tct	288
Ser Lys Gly Leu Tyr Leu Pro Ser Phe Phe Asn Thr Ala Lys Leu Ser	
85 90 95	
ttc gtg gct aag gga cga ggt ctt atg gga aaa gtg atc cct gga tgc	336
Phe Val Ala Lys Gly Arg Gly Leu Met Gly Lys Val Ile Pro Gly Cys	
100 105 110	
gcc gaa aca ttc caa gac tca tca gag ttc caa cca cgc ttc gaa ggt	384
Ala Glu Thr Phe Gln Asp Ser Ser Glu Phe Gln Pro Arg Phe Glu Gly	
115 120 125	
caa ggt caa agc cag agg ttc cgt gac atg cac cag aaa gtg gag cac	432
Gln Gly Gln Ser Gln Arg Phe Arg Asp Met His Gln Lys Val Glu His	
130 135 140	
att agg agc ggt gat acc att gcc aca aca ccc ggt gta gca cag tgg	480
Ile Arg Ser Gly Asp Thr Ile Ala Thr Thr Pro Gly Val Ala Gln Trp	
145 150 155 160	
ttc tac aac gac gga cag gaa cca ctt gtc atc gtc agc gtc ttc gat	528
Phe Tyr Asn Asp Gly Gln Glu Pro Leu Val Ile Val Ser Val Phe Asp	
165 170 175	
cta gcc agt cac cag aac cag ctt gac cgc aac cca agg cca ttt tac	576
Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr	
180 185 190	
tta gcc gga aac aac cca caa ggt caa gta tgg cta caa gga cga gag	624
Leu Ala Gly Asn Asn Pro Gln Gly Gln Val Trp Leu Gln Gly Arg Glu	
195 200 205	
caa cag cca cag aag aac att ttc aat gga ttt gga ccc gag gtt att	672
Gln Gln Pro Gln Lys Asn Ile Phe Asn Gly Phe Gly Pro Glu Val Ile	
210 215 220	
gct caa gct ttg aag atc gat ctt cag aca gca cag caa ctt cag aac	720
Ala Gln Ala Leu Lys Ile Asp Leu Gln Thr Ala Gln Gln Leu Gln Asn	
225 230 235 240	
caa gat gac aac cgt gga aac att gtc cga gtc caa gga ccg ttc ggt	768
Gln Asp Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe Gly	
245 250 255	
gtc att agg ccg cct ttg agg ggc cag aga cct cag gag gag gaa gaa	816
Val Ile Arg Pro Pro Leu Arg Gly Gln Arg Pro Gln Glu Glu Glu Glu	
260 265 270	
gaa gaa gga cga cat gga cga cac ggt aat ggc tta gag gag acc atc	864
Glu Glu Gly Arg His Gly Arg His Gly Asn Gly Leu Glu Glu Thr Ile	
275 280 285	

81

tgc agc gcc agg tgc acc gat aac ctc gat gac ccg tct cgt gct gac 912
 Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Arg Ala Asp
 290 295 300

gtg tac aag cca cag ctc ggt tac atc agc act ctc aac agt tac gat 960
 Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp
 305 310 315 320

ctc ccc atc ctt cgc ttc atc cgt ctc tca gcc ctc cgt gga tct atc 1008
 Leu Pro Ile Leu Arg Phe Ile Arg Leu Ser Ala Leu Arg Gly Ser Ile
 325 330 335

cgt caa aac gca atg gtg ctt cca cag tgg aac gca aac gcg aac gct 1056
 Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala
 340 345 350

att ctt tac gtg aca gac ggg gaa gcc caa atc cag atc gta aac gac 1104
 Ile Leu Tyr Val Thr Asp Gly Glu Ala Gln Ile Gln Ile Val Asn Asp
 355 360 365

aat ggt aac aga gtg ttt gac gga caa gtc tct caa gga cag ctc ata 1152
 Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile
 370 375 380

gcc gta cca caa ggt ttc tcg gtg gtg aaa cgc gca aca agc aac cga 1200
 Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg
 385 390 395 400

ttc cag tgg gtt gag ttc aaa aca aac gct aac gcg caa atc aac act 1248
 Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr
 405 410 415

ctg gcg gga cga acc tca gtc ttg aga ggt tta cca ctt gaa gtc ata 1296
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile
 420 425 430

acc aat ggg ttc caa atc tca ccc gaa gaa gca agg agg gtc aag ttc 1344
 Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe
 435 440 445

aac acg ctc gag acc act ttg act cac agc agt ggc cca gct agc tac 1392
 Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr
 450 455 460

gga agg cca agg gtg gct gca gct taa 1419
 Gly Arg Pro Arg Val Ala Ala Ala
 465 470

<210> 116
 <211> 472
 <212> PRT
 <213> Arabidopsis thaliana
 <400> 116

Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile
 1 5 10 15

Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe
 20 25 30

Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His
 35 40 45

Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala
 50 55 60

Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu
 65 70 75 80

82

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

Gly Arg Pro Arg Val Ala Ala Ala
465 470

<210> 117

<211> 1368

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1365)

<223> 12S Crb storage protein

<400> 117

atg ggt cga gtc tca tct att atc tct ttc tct ttg aca ctc ttg atc	48
Met Gly Arg Val Ser Ser Ile Ile Ser Phe Ser Leu Thr Leu Ile	
1 5 10 15	
ctc ttc aat ggc tac act gcc caa cag tgg ccc aac gag tgc cag ctc	96
Leu Phe Asn Gly Tyr Thr Ala Gln Gln Trp Pro Asn Glu Cys Gln Leu	
20 25 30	
gat caa ctc aat gcg ctc gaa cca tcc caa atc atc aag agc gag ggt	144
Asp Gln Leu Asn Ala Leu Glu Pro Ser Gln Ile Ile Lys Ser Glu Gly	
35 40 45	
ggt cgc atc gag gtc tgg gac cac cat gca ccc cag ctc cgt tgc tcc	192
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	
ggc ttt gcc ttt gag cgt ttc gtc att gag cct cag ggt ctt ttc ttg	240
Gly Phe Ala Phe Glu Arg Phe Val Ile Glu Pro Gln Gly Leu Phe Leu	
65 70 75 80	
ccc act ttc ttg aac gcc ggc aaa ctc acg ttt gtt gtt cac gga agg	288
Pro Thr Phe Leu Asn Ala Gly Lys Leu Thr Phe Val Val His Gly Arg	
85 90 95	
ggt cta atg gga aga gtt att ccg gga tgc gcc gag acg ttc atg gag	336
Gly Leu Met Gly Arg Val Ile Pro Gly Cys Ala Glu Thr Phe Met Glu	
100 105 110	
tca ccg gta ttt gga gaa ggt caa ggt cag ggt cag agt caa ggg ttc	384
Ser Pro Val Phe Gly Glu Gly Gln Gly Gln Gly Gln Ser Gln Gly Phe	
115 120 125	
cgt gac atg cac cag aaa gta gag cac cta cgg tgc ggt gac acc att	432
Arg Asp Met His Gln Lys Val Glu His Leu Arg Cys Gly Asp Thr Ile	
130 135 140	
gca aca cca tct ggt gta gct caa tgg ttc tac aac aat gga aat gag	480
Ala Thr Pro Ser Gly Val Ala Gln Trp Phe Tyr Asn Asn Gly Asn Glu	
145 150 155 160	
cct ctc att ctt gtt gca gcc gcg gat ctc gcc agc aac cag aac cag	528
Pro Leu Ile Leu Val Ala Ala Ala Asp Leu Ala Ser Asn Gln Asn Gln	
165 170 175	
ctt gac cgc aac ctt aga cca ttt ttg ata gcc gga aac aac cca caa	576
Leu Asp Arg Asn Leu Arg Pro Phe Leu Ile Ala Gly Asn Asn Pro Gln	
180 185 190	
ggg cag gaa tgg cta caa ggc cga aag caa cag aag caa aac aac atc	624
Gly Gln Glu Trp Leu Gln Gly Arg Lys Gln Gln Lys Gln Asn Asn Ile	
195 200 205	

ttc aat ggc ttc gca cct gag atc ttg gct caa gcc ttc aag atc aat	672
Phe Asn Gly Phe Ala Pro Glu Ile Leu Ala Gln Ala Phe Lys Ile Asn	
210 215 220	
gtc gag acg gct cag cag ctc cag aac cag caa gat aac cgt ggc aac	720
Val Glu Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg Gly Asn	
225 230 235 240	
atc gtc aag gtc aac gga cct ttc ggc gtc att agg cca ccc ttg aga	768
Ile Val Lys Val Asn Gly Pro Phe Gly Val Ile Arg Pro Pro Leu Arg	
245 250 255	
cgc ggc gaa ggc ggc caa caa cca cat gaa ata gct aat ggt tta gag	816
Arg Gly Glu Gly Gly Gln Gln Pro His Glu Ile Ala Asn Gly Leu Glu	
260 265 270	
gag act ttg tgc acc atg cga tgc act gaa aac ctc gat gac ccg tcg	864
Glu Thr Leu Cys Thr Met Arg Cys Thr Glu Asn Leu Asp Asp Pro Ser	
275 280 285	
gat gct gac gtg tac aag cca tca ctc gga tac att agc aca ctt aac	912
Asp Ala Asp Val Tyr Lys Pro Ser Leu Gly Tyr Ile Ser Thr Leu Asn	
290 295 300	
agc tac aat ctt cct atc ctc aga ctt ctc cgc ctt agc gct ctt cgt	960
Ser Tyr Asn Leu Pro Ile Leu Arg Leu Leu Arg Leu Ser Ala Leu Arg	
305 310 315 320	
ggc tcc atc cgt aaa aac gct atg gtg cta ccg caa tgg aac gta aac	1008
Gly Ser Ile Arg Lys Asn Ala Met Val Leu Pro Gln Trp Asn Val Asn	
325 330 335	
gca aac gcg gca ctc tac gtg aca aac gga aag gct cat ata caa atg	1056
Ala Asn Ala Ala Leu Tyr Val Thr Asn Gly Lys Ala His Ile Gln Met	
340 345 350	
gtg aac gac aac gga gaa aga gtg ttc gac caa gag atc tcc agc gga	1104
Val Asn Asp Asn Gly Glu Arg Val Phe Asp Gln Glu Ile Ser Ser Gly	
355 360 365	
cag tta cta gtc gtg cca caa ggc ttt tcg gtc atg aaa cat cgc ata	1152
Gln Leu Leu Val Val Pro Gln Gly Phe Ser Val Met Lys His Arg Ile	
370 375 380	
ggc gaa cag ttc gag tgg atc gaa ttc aag aca aac gaa aac gca cag	1200
Gly Glu Gln Phe Glu Trp Ile Glu Phe Lys Thr Asn Glu Asn Ala Gln	
385 390 395 400	
gtc aac aca ctc gcg ggc cgt acc tca gtc atg aga ggt ttg ccg ctt	1248
Val Asn Thr Leu Ala Gly Arg Thr Ser Val Met Arg Gly Leu Pro Leu	
405 410 415	
gag gtt ata acc aat ggg tac cag atc tct ccc gaa gaa gct aaa cga	1296
Glu Val Ile Thr Asn Gly Tyr Gln Ile Ser Pro Glu Glu Ala Lys Arg	
420 425 430	
gta aag ttt agc acg att gag acc aca ctg acc cat agc agt cca atg	1344
Val Lys Phe Ser Thr Ile Glu Thr Thr Leu Thr His Ser Ser Pro Met	
435 440 445	
agc tac gga agg cct agg gct tga	1368
Ser Tyr Gly Arg Pro Arg Ala	
450 455	
<210> 118	
<211> 455	

<212> PRT

<213> Arabidopsis thaliana

<400> 118

```

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

```

86

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

<210> 119

<211> 1356

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1353)

<223> putative 12S storage protein

<400> 119

atg cat aag ctt ttg ttt tct ctt ctc tcc gtc gtc tca ctc tca ttt	48
Met His Lys Leu Leu Phe Ser Leu Leu Ser Val Val Ser Leu Ser Phe	
1 5 10 15	
ctc ctc ttc ttc cat ggc gcc gag gca cgc cag cga gag gcg ccg ttt	96
Leu Leu Phe Phe His Gly Ala Glu Ala Arg Gln Arg Glu Ala Pro Phe	
20 25 30	
cca aac gcc tgc cat ttc agc caa atc aac agc ctc gcg ccc gct cag	144
Pro Asn Ala Cys His Phe Ser Gln Ile Asn Ser Leu Ala Pro Ala Gln	
35 40 45	
gcg acg aag ttc gaa gcc ggt cag atg gaa gta tgg gac cac atg agc	192
Ala Thr Lys Phe Glu Ala Gly Gln Met Glu Val Trp Asp His Met Ser	
50 55 60	
cct gag ctc cga tgc gcc ggt gta acg gtg gct cgc atc acc ctt cag	240
Pro Glu Leu Arg Cys Ala Gly Val Thr Val Ala Arg Ile Thr Leu Gln	
65 70 75 80	
ccc aat tcc att ttc ttg ccc gct ttc ttt agc cca cct gcc ctt gct	288
Pro Asn Ser Ile Phe Leu Pro Ala Phe Phe Ser Pro Pro Ala Leu Ala	
85 90 95	
tac gtt gtc caa gga gaa gga gtt atg ggg acg att gct tct ggt tgt	336
Tyr Val Val Gln Gly Glu Gly Val Met Gly Thr Ile Ala Ser Gly Cys	
100 105 110	
cct gag act ttt gca gaa gtt gaa gga tca tca gga aga gga gga gga	384
Pro Glu Thr Phe Ala Glu Val Glu Gly Ser Ser Gly Arg Gly Gly Gly	
115 120 125	
gga gac ccg ggt cga cgt ttt gag gac atg cac cag aag ttg gag aat	432
Gly Asp Pro Gly Arg Arg Phe Glu Asp Met His Gln Lys Leu Glu Asn	
130 135 140	

ttc cgg cga ggg gat gtg ttt gct tcg ctt gcc gga gtt tca cag tgg	480
Phe Arg Arg Gly Asp Val Phe Ala Ser Leu Ala Gly Val Ser Gln Trp	
145 150 155 160	
tggtac aac cgc ggt gat tcc gat gcc gtc att gtc att gtt ctt gat	528
Trp Tyr Asn Arg Gly Asp Ser Asp Ala Val Ile Val Ile Val Leu Asp	
165 170 175	
gtc acc aac aga gaa aac cag ctt gac caa gtc cct agg atg ttc caa	576
Val Thr Asn Arg Glu Asn Gln Leu Asp Gln Val Pro Arg Met Phe Gln	
180 185 190	
cta gcc ggg agc aga acg caa gaa gaa gaa caa cca tta acg tgg cca	624
Leu Ala Gly Ser Arg Thr Gln Glu Glu Glu Gln Pro Leu Thr Trp Pro	
195 200 205	
tca ggc aac aac gct ttc agc ggt ttc gac cca aac ata atc gcg gaa	672
Ser Gly Asn Asn Ala Phe Ser Gly Phe Asp Pro Asn Ile Ile Ala Glu	
210 215 220	
gca ttc aaa atc aac atc gag aca gct aag caa cta caa aac cag aag	720
Ala Phe Lys Ile Asn Ile Glu Thr Ala Lys Gln Leu Gln Asn Gln Lys	
225 230 235 240	
gac aac aga gga aac ata atc cga gca aat ggt cct ctc cat ttc gtc	768
Asp Asn Arg Gly Asn Ile Ile Arg Ala Asn Gly Pro Leu His Phe Val	
245 250 255	
atc cca ccg cct cgt gaa tgg cag caa gat ggc att gct aat ggc atc	816
Ile Pro Pro Pro Arg Glu Trp Gln Gln Asp Gly Ile Ala Asn Gly Ile	
260 265 270	
gaa gag act tat tgc acg gct aag att cat gag aat atc gat gat cca	864
Glu Glu Thr Tyr Cys Thr Ala Lys Ile His Glu Asn Ile Asp Asp Pro	
275 280 285	
gaa cgg tct gac cat ttt agc aca cga gcc gga aga atc agc act ctt	912
Glu Arg Ser Asp His Phe Ser Thr Arg Ala Gly Arg Ile Ser Thr Leu	
290 295 300	
aac agc ctt aat ctc cct gtt cta cgt cta gtc aga ctt aac gcc ctt	960
Asn Ser Leu Asn Leu Pro Val Leu Arg Leu Val Arg Leu Asn Ala Leu	
305 310 315 320	
aga ggt tat ctc tac agc gga gga atg gtg ttg cca caa tgg acg gca	1008
Arg Gly Tyr Leu Tyr Ser Gly Gly Met Val Leu Pro Gln Trp Thr Ala	
325 330 335	
aac gcg cac acg gtg cta tac gtc aca gga ggt caa gcc aag ata caa	1056
Asn Ala His Thr Val Leu Tyr Val Thr Gly Gly Gln Ala Lys Ile Gln	
340 345 350	
gtg gtg gac gac aat ggt cag tcg gtg ttc aat gag caa gtg gga caa	1104
Val Val Asp Asp Asn Gly Gln Ser Val Phe Asn Glu Gln Val Gly Gln	
355 360 365	
ggc caa atc att gtg att cca caa ggc ttt gca gtt tca aaa acg gct	1152
Gly Gln Ile Ile Val Ile Pro Gln Gly Phe Ala Val Ser Lys Thr Ala	
370 375 380	
ggt gaa acg ggt ttc gag tgg ata tca ttc aag aca aac gat aac gct	1200
Gly Glu Thr Gly Phe Glu Trp Ile Ser Phe Lys Thr Asn Asp Asn Ala	
385 390 395 400	
tac att aac aca ctg agc ggc caa aca tcg tac ttg aga gca gtt cca	1248
Tyr Ile Asn Thr Leu Ser Gly Gln Thr Ser Tyr Leu Arg Ala Val Pro	
405 410 415	

88

gtg gat gtg atc aaa gcg tca tat gga gtg aac gag gaa gaa gcc aag 1296
 Val Asp Val Ile Lys Ala Ser Tyr Gly Val Asn Glu Glu Glu Ala Lys
 420 425 430

agg atc aag ttt agt cag caa gag acc atg ttg tct atg aca cca agc 1344
 Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser
 435 440 445

tct tct tct taa 1356
 Ser Ser Ser
 450

<210> 120
 <211> 451
 <212> PRT
 <213> Arabidopsis thaliana

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

cct gag ctc cga tgc gcc ggt gta acg gtg gct cgc atc acc ctt cag 240
Pro Glu Leu Arg Cys Ala Gly Val Thr Val Ala Arg Ile Thr Leu Gln
65 70 75 80

90

ccc aat tcc att ttc ttg ccc gct ttc ttt agc cca cct gcc ctt gct	288
Pro Asn Ser Ile Phe Leu Pro Ala Phe Phe Ser Pro Pro Ala Leu Ala	
85 90 95	
tac gtt gtc caa gga gaa gga gtt atg ggg acg att gct tct ggt tgt	336
Tyr Val Val Gln Gly Glu Gly Val Met Gly Thr Ile Ala Ser Gly Cys	
100 105 110	
cct gag act ttt gca gaa gtt gaa gga tca tca gga aga gga gga gga	384
Pro Glu Thr Phe Ala Glu Val Glu Gly Ser Ser Gly Arg Gly Gly Gly	
115 120 125	
gga gac ccg ggt cga cgt ttt gag gac atg cac cag aag ttg gag aat	432
Gly Asp Pro Gly Arg Arg Phe Glu Asp Met His Gln Lys Leu Glu Asn	
130 135 140	
ttc cgg cga ggg gat gtg ttt gct tcg ctt gcc gga gtt tca cag tgg	480
Phe Arg Arg Gly Asp Val Phe Ala Ser Leu Ala Gly Val Ser Gln Trp	
145 150 155 160	
tgg tac aac cgc ggt gat tcc gat gcc gtc att gtc att gtt ctt gat	528
Trp Tyr Asn Arg Gly Asp Ser Asp Ala Val Ile Val Ile Val Leu Asp	
165 170 175	
gtc acc aac aga gaa aac cag ctt gac caa gtc cct agg atg ttc caa	576
Val Thr Asn Arg Glu Asn Gln Leu Asp Gln Val Pro Arg Met Phe Gln	
180 185 190	
cta gcc ggg agc aga acg caa gaa gaa gaa caa cca tta acg tgg cca	624
Leu Ala Gly Ser Arg Thr Gln Glu Glu Glu Gln Pro Leu Thr Trp Pro	
195 200 205	
tca ggc aac aac gct ttc agc ggt ttc gac cca aac ata atc gcg gaa	672
Ser Gly Asn Asn Ala Phe Ser Gly Phe Asp Pro Asn Ile Ile Ala Glu	
210 215 220	
gca ttc aaa atc aac atc gag aca gct aag caa cta caa aac cag aag	720
Ala Phe Lys Ile Asn Ile Glu Thr Ala Lys Gln Leu Gln Asn Gln Lys	
225 230 235 240	
gac aac aga gga aac ata atc cga gca aat ggt cct ctc cat ttc gtc	768
Asp Asn Arg Gly Asn Ile Ile Arg Ala Asn Gly Pro Leu His Phe Val	
245 250 255	
atc cca ccg cct cgt gaa tgg cag caa gat ggc att gct aat ggc atc	816
Ile Pro Pro Pro Arg Glu Trp Gln Gln Asp Gly Ile Ala Asn Gly Ile	
260 265 270	
gaa gag act tat tgc acg gct aag att cat gag aat atc gat gat cca	864
Glu Glu Thr Tyr Cys Thr Ala Lys Ile His Glu Asn Ile Asp Asp Pro	
275 280 285	
gaa cgg tct gac cat ttt agc aca cga gcc gga aga atc agc act ctt	912
Glu Arg Ser Asp His Phe Ser Thr Arg Ala Gly Arg Ile Ser Thr Leu	
290 295 300	
aac agc ctt aat ctc cct gtt cta cgt cta gtc aga ctt aac gcc ctt	960
Asn Ser Leu Asn Leu Pro Val Leu Arg Leu Val Arg Leu Asn Ala Leu	
305 310 315 320	
aga ggt tat ctc tac agc gga gga atg gtg ttg cca caa tgg acg gca	1008
Arg Gly Tyr Leu Tyr Ser Gly Gly Met Val Leu Pro Gln Trp Thr Ala	
325 330 335	
aac gcg cac acg gtg cta tac gtc aca gga ggt caa gcc aag ata caa	1056
Asn Ala His Thr Val Leu Tyr Val Thr Gly Gly Gln Ala Lys Ile Gln	
340 345 350	

91

gtg gtg gac gac aat ggt cag tcg gtg ttc aat gag caa gtg gga caa 1104
 Val Val Asp Asp Asn Gly Gln Ser Val Phe Asn Glu Gln Val Gly Gln
 355 360 365

ggc caa atc att gtg att cca caa ggc ttt gca gtt tca aaa acg gct 1152
 Gly Gln Ile Ile Val Ile Pro Gln Gly Phe Ala Val Ser Lys Thr Ala
 370 375 380

ggt gaa acg ggt ttc gag tgg ata tca ttc aag aca aac gat aac gct 1200
 Gly Glu Thr Gly Phe Glu Trp Ile Ser Phe Lys Thr Asn Asp Asn Ala
 385 390 395 400

tac att aac aca ctg agc ggc caa aca tcg tac ttg aga gca gtt cca 1248
 Tyr Ile Asn Thr Leu Ser Gly Gln Thr Ser Tyr Leu Arg Ala Val Pro
 405 410 415

gtg gat gtg atc aaa gcg tca tat gga gtg aac gag gaa gaa gcc aag 1296
 Val Asp Val Ile Lys Ala Ser Tyr Gly Val Asn Glu Glu Glu Ala Lys
 420 425 430

agg atc aag ttt agt cag caa gag acc atg ttg tct atg aca cca agc 1344
 Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser
 435 440 445

tct tct tct taa 1356
 Ser Ser Ser
 450

<210> 122
 <211> 451
 <212> PRT
 <213> Arabidopsis thaliana

<400> 122
 Met His Lys Leu Leu Phe Ser Leu Leu Ser Val Val Ser Leu Ser Phe
 1 5 10 15
 Leu Leu Phe Phe His Gly Ala Glu Ala Arg Gln Arg Glu Ala Pro Phe
 20 25 30
 Pro Asn Ala Cys His Phe Ser Gln Ile Asn Ser Leu Ala Pro Ala Gln
 35 40 45
 Ala Thr Lys Phe Glu Ala Gly Gln Met Glu Val Trp Asp His Met Ser
 50 55 60
 Pro Glu Leu Arg Cys Ala Gly Val Thr Val Ala Arg Ile Thr Leu Gln
 65 70 75 80
 Pro Asn Ser Ile Phe Leu Pro Ala Phe Phe Ser Pro Pro Ala Leu Ala
 85 90 95
 Tyr Val Val Gln Gly Glu Gly Val Met Gly Thr Ile Ala Ser Gly Cys
 100 105 110
 Pro Glu Thr Phe Ala Glu Val Glu Gly Ser Ser Gly Arg Gly Gly Gly
 115 120 125
 Gly Asp Pro Gly Arg Arg Phe Glu Asp Met His Gln Lys Leu Glu Asn
 130 135 140
 Phe Arg Arg Gly Asp Val Phe Ala Ser Leu Ala Gly Val Ser Gln Trp
 145 150 155 160
 Trp Tyr Asn Arg Gly Asp Ser Asp Ala Val Ile Val Ile Val Leu Asp
 165 170 175
 Val Thr Asn Arg Glu Asn Gln Leu Asp Gln Val Pro Arg Met Phe Gln
 180 185 190

92

Leu Ala Gly Ser Arg Thr Gln Glu Glu Glu Gln Pro Leu Thr Trp Pro
 195 200 205
 Ser Gly Asn Asn Ala Phe Ser Gly Phe Asp Pro Asn Ile Ile Ala Glu
 210 215 220
 Ala Phe Lys Ile Asn Ile Glu Thr Ala Lys Gln Leu Gln Asn Gln Lys
 225 230 235 240
 Asp Asn Arg Gly Asn Ile Ile Arg Ala Asn Gly Pro Leu His Phe Val
 245 250 255
 Ile Pro Pro Pro Arg Glu Trp Gln Gln Asp Gly Ile Ala Asn Gly Ile
 260 265 270
 Glu Glu Thr Tyr Cys Thr Ala Lys Ile His Glu Asn Ile Asp Asp Pro
 275 280 285
 Glu Arg Ser Asp His Phe Ser Thr Arg Ala Gly Arg Ile Ser Thr Leu
 290 295 300
 Asn Ser Leu Asn Leu Pro Val Leu Arg Leu Val Arg Leu Asn Ala Leu
 305 310 315 320
 Arg Gly Tyr Leu Tyr Ser Gly Gly Met Val Leu Pro Gln Trp Thr Ala
 325 330 335
 Asn Ala His Thr Val Leu Tyr Val Thr Gly Gly Gln Ala Lys Ile Gln
 340 345 350
 Val Val Asp Asp Asn Gly Gln Ser Val Phe Asn Glu Gln Val Gly Gln
 355 360 365
 Gly Gln Ile Ile Val Ile Pro Gln Gly Phe Ala Val Ser Lys Thr Ala
 370 375 380
 Gly Glu Thr Gly Phe Glu Trp Ile Ser Phe Lys Thr Asn Asp Asn Ala
 385 390 395 400
 Tyr Ile Asn Thr Leu Ser Gly Gln Thr Ser Tyr Leu Arg Ala Val Pro
 405 410 415
 Val Asp Val Ile Lys Ala Ser Tyr Gly Val Asn Glu Glu Glu Ala Lys
 420 425 430
 Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser
 435 440 445
 Ser Ser Ser
 450

<210> 123

<211> 17

<212> DNA

<213> Künstliche Sequenz

<220>

 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

<400> 123

aacgagtgcc agctcga

<210> 124

<211> 18

<212> DNA

<213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 124
 ttccgtgaca tgcaccag 18

 <210> 125
 <211> 19
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 125
 cagaaccagc ttgaccgca 19

 <210> 126
 <211> 22
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 126
 tagccggaaa caaccacaa gg 22

 <210> 127
 <211> 17
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 127
 aacctcgatg acccgtc 17

 <210> 128
 <211> 19
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 128
 tgctgacgtg tacaagcca 19

 <210> 129
 <211> 20
 <212> DNA
 <213> Künstliche Sequenz

 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein

 <400> 129
 gacatgcacc agaarktrga 20

<210> 130
 <211> 8
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
 <400> 130
 catgcatg 8
 <210> 131
 <211> 8
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: sequence
 motive from plant storage protein
 <400> 131
 gccacystc 8
 <210> 132
 <211> 867
 <212> DNA
 <213> Arabidopsis thaliana
 <220>
 <221> CDS
 <222> (1)..(864)
 <223> prohibitin 1
 <400> 132
 atg aac aac gtc aaa gtt cca aag ata cca ggt ggt ggt gcc att tcg 48
 Met Asn Asn Val Lys Val Pro Lys Ile Pro Gly Gly Gly Ala Ile Ser
 1 5 10 15
 acg ttg ctt aag gtt ggg att att ggt ggg ctt ggc ctc tat ggt gct 96
 Thr Leu Leu Lys Val Gly Ile Ile Gly Gly Leu Gly Leu Tyr Gly Ala
 20 25 30
 acg cac agt ctc tac aat gtt gaa gga gga cat cga gcc atc atg ttc 144
 Thr His Ser Leu Tyr Asn Val Glu Gly Gly His Arg Ala Ile Met Phe
 35 40 45
 aat cgt tta gtc ggt att aaa gat aag gtt tac cct gag ggt aca cac 192
 Asn Arg Leu Val Gly Ile Lys Asp Lys Val Tyr Pro Glu Gly Thr His
 50 55 60
 ctt atg att cct tgg ttt gaa agg ccg gtc atc tat gac gtt cgt gct 240
 Leu Met Ile Pro Trp Phe Glu Arg Pro Val Ile Tyr Asp Val Arg Ala
 65 70 75 80
 cga cct tac ctt gtt gag agt aca tcc gga agc cgt gat ctt cag atg 288
 Arg Pro Tyr Leu Val Glu Ser Thr Ser Gly Ser Arg Asp Leu Gln Met
 85 90 95
 gtg aaa att ggg ctt agg gtt ctc aca cgt ccc atg gca gac cag tta 336
 Val Lys Ile Gly Leu Arg Val Leu Thr Arg Pro Met Ala Asp Gln Leu
 100 105 110
 cct gaa atc tac aga agc ctt ggt gag aac tac agc gag aga gtt cta 384
 Pro Glu Ile Tyr Arg Ser Leu Gly Glu Asn Tyr Ser Glu Arg Val Leu
 115 120 125

95

cct tct ata atc aac gag act ttg aaa gct gtg gtt gct cag tac aat 432
 Pro Ser Ile Ile Asn Glu Thr Leu Lys Ala Val Val Ala Gln Tyr Asn
 130 135 140

gca agc cag ctt att act cag aga gag gcg gtc agt agg gag atc agg 480
 Ala Ser Gln Leu Ile Thr Gln Arg Glu Ala Val Ser Arg Glu Ile Arg
 145 150 155 160

aag att ctg act gaa cga gca gca aac ttc aat gtt gcg ctt gac gat 528
 Lys Ile Leu Thr Glu Arg Ala Ala Asn Phe Asn Val Ala Leu Asp Asp
 165 170 175

gtg tcc atc aca aac ctg aca ttc ggg aag gag ttc aca gct gcc att 576
 Val Ser Ile Thr Asn Leu Thr Phe Gly Lys Glu Phe Thr Ala Ala Ile
 180 185 190

gaa gca aag cag gtg gcg gct caa gag gct gag cgg gct aag ttc att 624
 Glu Ala Lys Gln Val Ala Ala Gln Glu Ala Glu Arg Ala Lys Phe Ile
 195 200 205

gtt gag aag gcc gaa caa gac aag aga agt gct gtt atc cgc gcc cag 672
 Val Glu Lys Ala Glu Gln Asp Lys Arg Ser Ala Val Ile Arg Ala Gln
 210 215 220

gga gaa gcc aag agt gct cag ctc att ggt caa gca att gca aac aac 720
 Gly Glu Ala Lys Ser Ala Gln Leu Ile Gly Gln Ala Ile Ala Asn Asn
 225 230 235 240

caa gcg ttt atc acg ctc agg aag atc gag gct gca aga gag att gca 768
 Gln Ala Phe Ile Thr Leu Arg Lys Ile Glu Ala Ala Arg Glu Ile Ala
 245 250 255

cag acc ata gca aac tcg gcg aac aag gtt tac ttg agc tca gac gat 816
 Gln Thr Ile Ala Asn Ser Ala Asn Lys Val Tyr Leu Ser Ser Asp Asp
 260 265 270

ctt ttg ctt aac cta caa ggg atg aat ttg gat gtt gat gca aag aac 864
 Leu Leu Leu Asn Leu Gln Gly Met Asn Leu Asp Val Asp Ala Lys Asn
 275 280 285

tag 867

<210> 133
 <211> 288
 <212> PRT
 <213> Arabidopsis thaliana

<400> 133
 Met Asn Asn Val Lys Val Pro Lys Ile Pro Gly Gly Gly Ala Ile Ser
 1 5 10 15
 Thr Leu Leu Lys Val Gly Ile Ile Gly Gly Leu Gly Leu Tyr Gly Ala
 20 25 30
 Thr His Ser Leu Tyr Asn Val Glu Gly Gly His Arg Ala Ile Met Phe
 35 40 45
 Asn Arg Leu Val Gly Ile Lys Asp Lys Val Tyr Pro Glu Gly Thr His
 50 55 60
 Leu Met Ile Pro Trp Phe Glu Arg Pro Val Ile Tyr Asp Val Arg Ala
 65 70 75 80
 Arg Pro Tyr Leu Val Glu Ser Thr Ser Gly Ser Arg Asp Leu Gln Met
 85 90 95
 Val Lys Ile Gly Leu Arg Val Leu Thr Arg Pro Met Ala Asp Gln Leu
 100 105 110

96

Pro Glu Ile Tyr Arg Ser Leu Gly Glu Asn Tyr Ser Glu Arg Val Leu
 115 120 125

Pro Ser Ile Ile Asn Glu Thr Leu Lys Ala Val Val Ala Gln Tyr Asn
 130 135 140

Ala Ser Gln Leu Ile Thr Gln Arg Glu Ala Val Ser Arg Glu Ile Arg
 145 150 155 160

Lys Ile Leu Thr Glu Arg Ala Ala Asn Phe Asn Val Ala Leu Asp Asp
 165 170 175

Val Ser Ile Thr Asn Leu Thr Phe Gly Lys Glu Phe Thr Ala Ala Ile
 180 185 190

Glu Ala Lys Gln Val Ala Ala Gln Glu Ala Glu Arg Ala Lys Phe Ile
 195 200 205

Val Glu Lys Ala Glu Gln Asp Lys Arg Ser Ala Val Ile Arg Ala Gln
 210 215 220

Gly Glu Ala Lys Ser Ala Gln Leu Ile Gly Gln Ala Ile Ala Asn Asn
 225 230 235 240

Gln Ala Phe Ile Thr Leu Arg Lys Ile Glu Ala Ala Arg Glu Ile Ala
 245 250 255

Gln Thr Ile Ala Asn Ser Ala Asn Lys Val Tyr Leu Ser Ser Asp Asp
 260 265 270

Leu Leu Leu Asn Leu Gln Gly Met Asn Leu Asp Val Asp Ala Lys Asn
 275 280 285

<210> 134

<211> 40

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer

<400> 134

ataagaatgc ggccgcgtgt tccatttggc cggaacaac

40

<210> 135

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer

<400> 135

cccggatcct tctgtaacat ttgacaaaac atg

33

<210> 136

<211> 40

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer

<400> 136

ataagaatgc ggccgcgtgt tccatttggc cggaacaac

40

<210> 137
<211> 44
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer
<400> 137
ataagaatgc ggccgcggat ccaccctgga gaacgccacg agtg 44
<210> 138
<211> 45
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer
<400> 138
ataagaatgc ggccgcggat ccctcagggc cttttcttgc ccact 45
<210> 139
<211> 30
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer
<400> 139
ccgctcgagt ttacggatgg agccacgaag 30
<210> 140
<211> 30
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer
<400> 140
ccgctcgagg taagctcaac aaatctttag 30
<210> 141
<211> 31
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer
<400> 141
acgcgtcgac gcgttctgcg tgcaagatat t 31
<210> 142
<211> 46
<212> DNA
<213> Künstliche Sequenz
<220>
<223> Beschreibung der künstlichen Sequenz:
oligonucleotid primer

<400> 142
 ataagaatgc ggccgcggat ccatggctaa caagctcttc ctctgc 46

<210> 143
 <211> 45
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotid primer

<400> 143
 ataagaatgc ggccgcggat ccctagtagt aaggagggaa gaaag 45

<210> 144
 <211> 4954
 <212> DNA
 <213> Künstliche Sequenz

<220>
 <223> Beschreibung der künstlichen Sequenz: DNA
 construct coding for dsRNA for suppression of
 multiple storage proteins

<400> 144
 agcttggtac cgagctcgga tccactagta acggccgcca gtgtgctgga attcgccctt 60
 gcggccgcgt gttccatttg gccggaaca accagcaggg aggctttggc gggtcacagc 120
 aacaacaaga acagaaaaac ttgtggagcg gggtcgacgc acaggtcata gctcaagcat 180
 tgaaaattga cggttcagttg gctcagcagc ttcagaacca acaagacagc agaggaaaca 240
 tcgttcgtgt taagggacct ttccaggctc tgaggccacc tctaagacag ccctacgaga 300
 gcgaggagtg gagacacca cgtagccac agggcaacgg ccttgaggag actatctgca 360
 gcatgaggtc ccacgagaac attgacgacc ctgctcgtgc tgacgtgtac aagcccagcc 420
 taggtcgcgt gaccagcgtc aacagctata ccttgcccat cttggagtat gtcaggctca 480
 gtgccactcg tggcgttctc caggggtggat ccttctgtaa catttgacaa aacatgtgaa 540
 cacgtcatcc gtcatataga acttccaatt ttaatatgtt ttgctaaaga aaaaaaaaag 600
 gaataaatat ctatcaaatt cattttttaa acatttgtat acgttcttaa ataatttagg 660
 atatgactaa tttttctttt tggtaaaaat gttaatatct atatttaatt tattaagaaa 720
 aatgtactta caccctggag aacgccacga gtggcactga gcctgacata ctccaagatg 780
 ggcaaggtat agctgttgac gctggtcacg cgacctaggc tgggcttgta cacgtcagca 840
 cgagcagggt cgtcaatgtt ctgctgggac ctcatgctgc agatagtctc ctcaaggccg 900
 ttgccctgtg ggctacgtgg gtgtctccac tcctcgtctc cgtagggctg tcttagaggt 960
 ggccctcacga cctggaaagg tcccttaaca cgaacgatgt ttctctgct gtcttgttg 1020
 ttctgaagct gctgagccaa ctgaacgtca attttcaatg cttgagctat gacctgtgcg 1080
 tcgaaccgcg tccacaagtt tttctgttct tgttgttgct gtgaaccgcc aaagcctccc 1140
 tgctggttgt ttccggccaa atggaacacg cgccgcaag ggcgaattct cgagatatcc 1200
 atcacactgg cggccgctcg acgtaagctc aacaaatctt tagaaaatta attttatgtg 1260
 acatatgcaa taatttgatt tggcaagata aactaataga ttttgcgatt tggagtttta 1320
 aactctaaat aatctaaatc gttttcaatt ggtttaaata tatactctgc atttttaatc 1380
 gtttttaatt aaaaaatata tatatatata tatatcttgc atttttaatc gttttcaatt 1440
 taaaaaatat cttgcacgca gaacgctctc gagcgccgc ggatcctcag ggtcttttct 1500
 tgcccacttt cttgaacgcc ggcaaacctca cgtttgttgt tcacggaagg ggtctaattg 1560
 gaagagttat tccgggatgc gccgagacgt tcatggagtc accggtatgt ggagaaggctc 1620
 aaggctcagg tcagagtcaa gggttccgtg acatgcacca gaaagtagag cacctacggg 1680
 gcggtgacac cattgcaaca ccatctggtg tagctcaatg gttctacaac aatggaaatg 1740
 agcctctcat tcttgttgca gccgcggatc tcgccagcaa ccagaaccag cttgaccgca 1800
 accttagacc atttttgata gccggaaaca accacaagg gcaggaatgg ctacaaggcc 1860
 gaaagcaaca gaagcaaac aacatcttca atggcttcgc acctgagatc ttgggtcaag 1920
 ctttcaagat caatgtcgag acggctcagc agctccagaa ccagcaagat aaccgtggca 1980
 acatcgtcaa ggtcaacgga cctttcggcg tcattaggcc acccttgaga cgcggcgaag 2040
 gcggccaaca accacatgaa atagctaatt gtttagagga gactttgtgc accatgcgat 2100

```

gcactgaaaa cctcgatgac ccgtcggatg ctgacgtgta caagccatca ctcgataca 2160
ttagcacact taacagctac aatcttcccta tcctcagact tctccgcctt agcgcctctc 2220
gtggctccat ccgtaaaact cgaggtaagc tcaacaaatc tttagaaaat taattttatg 2280
tgacatatgc aataatttga tttggcaaga taaactaata gattttgcga tttggagttt 2340
taaaactctaa ataactctaaa tcgttttcaa ttggttttaa tatatatctt gcatttttaa 2400
tcgttttttaa ttaaaaaata tatatatata tatatatctt gcatttttaa tcgttttcaa 2460
tttaaaaaat atcttgcacg cagaacgctg tcgagtttta cggatggagc cacgaagagc 2520
gctaaggcgg agaagtctga ggataggaag attgtagctg ttaagtgtgc taatgtatcc 2580
gagtgatggc ttgtacacgt cagcatccga cgggtcatcg aggttttcag tgcacgcat 2640
gggtgcacaaa gtctcctcta aaccattagc tatttcatgt ggttggtggc cgccttcgcc 2700
gcgtctcaag ggtggcctaa tgacgccgaa aggtccgttg acctgacga tgttgccacg 2760
gttatcttgc tggttctgga gctgctgagc cgtctcgaca ttgatcttga aggttgagc 2820
caagatctca ggtgcgaagc cattgaagat gttgttttgc ttctgttgct ttcggccttg 2880
tagccattcc tgcccttggt ggttggttcc ggctatcaaa aatggtctaa ggttgcggtc 2940
aagctgggtc tggttgctgg cgagatccgc ggctgcaaca agaagagag gctcatttcc 3000
attgtttag aaccattgag ctacaccaga tgggtgtgca atggtgtcac cgcaccgtag 3060
gtgctctact ttctggtgca tgtcacggaa cccttgactc tgaccctgac cttgaccttc 3120
tccaaatacc ggtgactcca tgaacgtctc ggcgcacccc ggaataactc ttcccattag 3180
acccttccg tgaacaacaa acgtgagttt gccggcggtc aagaaagtgg gcaagaaaag 3240
accctgagga tccgcggccg cgcacgcac tagctcgagg taagctcaac aaatctttag 3300
aaaattaatt ttatgtgaca tatgcaataa tttgatttgg caagataaac taatagattt 3360
tgcgatttgg agtttttaaac tctaaataat ctaaactggt ttcaattggt ttaaataatat 3420
atcttgcatt tttaatcggt tttaattaaa aaatatatat atatatatat atcttgcatt 3480
tttaatcggt ttcaatttaa aaaatatctt gcacgcagaa cgctagggcc gcggcccgcg 3540
atccatggct aacaagctct tcctcgtctg cgcaactctc gccctctgct tcctcctcac 3600
caacgcttcc atctaccgca ccgttgtcga attcgaagaa gatgacgcca gcaaccccg 3660
aggtccaaga cagagatgcc agaaggagtt tcagcaatca caacaccta gagcttgcca 3720
gagatggatg agcaagcaaa tgaggcaagg acgtggtggt ggtccttccc tcgacgatga 3780
gttcgatattc gagggccccc agcagggata ccagctactc cagcagtgtc gcaacgagct 3840
tcgccaggaa gagccagttt gcgtttgccc caccttgaaa caagctgcca gggcagttag 3900
cctccaggga cagcacggac cattccaatc caggaaaatt taccagtcag ctaagtaact 3960
gcctaacatt tgcaagatcc agcaagttgg tgaatgtccc ttccagacca ccatcccttt 4020
cttccctcct tactactagg gtactcgagg taagctcaac aaatctttag aaaattaatt 4080
ttatgtgaca tatgcaataa tttgatttgg caagataaac taatagattt tgcgatttgg 4140
agtttttaaac tctaaataat ctaaactggt ttcaattggt ttaaataatat atcttgcatt 4200
tttaatcggt tttaattaaa aaatatatat atatatatat atcttgcatt tttaatcggt 4260
ttcaatttaa aaaatatctt gcacgcagaa cgctgtcgac taccctagta gtaaggaggg 4320
aagaaagggg tgggtggtctg gaaggacat tcaccaactt gctggatctt gcaaatgtta 4380
ggcaagtact tagctgactg gtaaattttc ctggattgga atggtccgtg ctgtccctgg 4440
aggctaactg ccctggcagc ttgtttcaag gtggggcaaa cgcaaactgg ctcttccctgg 4500
cgaagctcgt tgcagcactg ctggagtagc tggtatccct gctggggggc ctcgaaatcg 4560
aactcatcgt cgagggaagg accaccacca cgtccttgcc tcatttgcct gctcatccat 4620
ctctggcaag ctcttaggtg ttgtgattgc tgaaactcct tctggcatct ctgtcttggg 4680
cctacggggg ttgtggcgct atcttcttcg aattcgacaa cgggtgcggta gatggaagcg 4740
ttggtgagga ggaagcagag ggcgagagtt gcgcagacga ggaagagctt gtagccatg 4800
gatcaattcg ccctatagtg agtcgtatta caattcactg gccgtcggtt tacaacgtcg 4860
tgactgggaa aaccctggcg ttacccaact taatcgctt gcagcacatc ccccttctgc 4920
cagctggcgt aatagcgaag aggccgcac cgat 4954

```

<210> 145

<211> 4954

<212> RNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: RNA coding
for dsRNA for suppression of multiple storage
proteins

<400> 145

```

agcuugguac cgagcucgga uccacuagua acggccgcca gugugcugga auucgccccu 60
gcgggccgcu guuccauuug gccggaaaca accagcaggg aggcuuuggc gguucacagc 120
aacaacaaga acagaaaaac uuguggagcg gguucgacgc acaggucuaa gcucaagcau 180
ugaaaauuga cguucaguug gcucagcagc uucagaacca acaagacagc agaggaaaca 240
ucguucgugu uaagggaccu uuccaggucg ugaggccacc ucuaagacag cccuacgaga 300
gcgaggagug gagacaccca cguagcccac agggcaacgg ccuugaggag acuaucugca 360
gcaugagguc ccacgagaac auugacgacc cugcucgugc ugacguguac aagcccagcc 420
uaggucgcu gaccagcguc aacagcuaua ccuugcccau cuuggaguau gucaggcuca 480
gugccacucg uggcguucuc caggguggau ccuucuguaa cauugacaa aacauguga 540
cacgucaucc gucauauaga acuuccaauu uuaauauguu uugcuaaaga aaaaaaaaag 600
gaauaaaaua cuaucaaaau cauuuuuuuu acauuugua acguucuaa auauuuuagg 660
auaugacuaa uuuuucuuuu ugguaaaaau guuaauaucu auauuuauu uauuaagaaa 720
aaugacuua caccucggag aacgccacga guggcacuga gccugacuaa cuccaagau 780
ggcaaggua agcuguugac gcuggucacg cgaccuaggc ugggcuugua cagcucagca 840
cgagcagggu cgucaaugu cucgugggac cucaugcugc agauagucuc cucaaggccg 900
uugcccugug ggcucgugg gugucuccac uccucgcucu cguagggcug ucuaagggu 960
ggccucacga ccuggaaagg ucccuaaca cgaacgaugu uccucugcu gucuuguug 1020
uucugaagcu gcugagccaa cugaacguc auuuucaaug cuugagcuau gaccugugc 1080
ucgaacccgc uccacaaguu uuucuguucu uguuguugcu gugaaccgcc aaagccucc 1140
ugcugguugu uuccggccaa auggaacacg cggccgcaag ggcgaauuc gcagauauc 1200
aucacacug cgccgcucg acguaagcuc acaaaucuu uagaaaaua auuuuauug 1260
acauaugcaa uaauuugau uggaagaua aacuaauaga uuuugcga uggaguuuu 1320
aacucuaaa auucuaaauc guuuucaauu gguuuuaa uauaucuugc auuuuaa 1380
guuuuuauu aaaaaaua uauauaua uauaucuugc auuuuaa 1440
uaaaaaau cuugcacgca gaacgcucuc gagcggccgc ggauccucag ggucuuuc 1500
ugcccacuu cuugaacgcc ggcaaacuca cguuuguugu ucacggaagg ggucuaaug 1560
gaagaguau uccgggaug gccgagacgu ucauggaguc accgguuuu ggagaagg 1620
aaggucagg ucagagucaa gguuuccgug acaugacca gaaaguagag caccuacgg 1680
gcgguagac cauugcaaca ccaucuggug uagcucaug guucuacaac aauggaaa 1740
agccucuc auuuguugca gccgcggauc ucgccagcaa ccagaaccag cuugaccg 1800
accuagacc auuuuugau gccggaaaca acccacaagg gcaggaauug cuacaagg 1860
gaaagcaaca gaagcaaac acaucuuca auggcucgc accugagau ugggcuaag 1920
ccuucagau caaugucgag acggcucagc agcuccagaa ccagcaagau aaccgugg 1980
acaucgucaa ggucaacgga ccuucggcg ucauaggcc acccuugaga cgcggcga 2040
gcggccaa acccaugaa auagcuauug guuagagga gacuugugc accaugcgau 2100
gcacugaaaa ccucgaugac ccgucggaug cugacgugua caagccaua cucggaua 2160
uuagcacacu uaacagcuac aaucuccua uccucagacu ucuccgccu agcgcucu 2220
guggcuccau ccguaaaacu cgagguuagc ucaacaauc uuuagaaa auuuuuuag 2280
ugacauaugc aaauuuuuga uuuggcaaga uaaacuaa gauuuugcga uuuggagu 2340
uaaacucua auauucuaa ucguuuuc uugguuuaa uauauaucu gcauuuuua 2400
ucguuuuua uuaaaaaa uauauaua uauauaucu gcauuuuua ucguuuuc 2460
uuuaaaaa auucugcacg cagaacgcug ucgaguuaa cggauggagc cacgaagag 2520
gcuaaggcga agaagucuga ggauaggaag auuguagcug uuaagugugc uaaugua 2580
gagugauggc uuguacacgu cagcauccga cgggucauc agguuuucag ugcaucgca 2640
ggugcaciaa gucuccua aaccauuagc uauuucagu gguuguuggc cgccuucg 2700
gcgucuaag gguggccua ugacgccgaa agguccguug accuugacga uguugccag 2760
guuauucugc ugguucugga gcugcugagc cgucucgaca uugaucuuga aggcugag 2820
caagaucuca ggugcgaagc cauugaagau guuguuuugc uucuguugcu uucggccu 2880
uagccauucc ugcccuugug gguuguuucc ggcuaucaaa aauggucuaa gguugcgg 2940
aagcugguuc ugguugcugg cgagaucgc ggcugcaaca agaagagag gcucauuucc 3000
auuguuguag aaccauugag cuacaccaga ugguguugca augguugac cgaccguag 3060
gugcucuacu uucuggugca ugucacggaa ccuugacuc ugaccugac cuugaccu 3120
uccaaauacc ggugacucca ugaacgcuc ggcgcaucc ggaauaacu uucccauag 3180
accccuuccg ugaacaacaa acgugaguuu gccggcguc aagaaagug gcaagaaa 3240
accugagga uccgcggccg cgcaugcauc uagcucgagg uaagcucaac aaaucuuag 3300
aaaauaaau uuaugugaca uaugcaaua uuugauuug caagauaaac uauagauu 3360

```

101

```

ugcgauuugg aguuuuuaac ucuaaaauau cuaaaucguu uucaauuggu uuaaaauauau 3420
aucuugcauu uuuaaauucguu uuuaaaauaaa aaauauauau auauauauau aucuugcauu 3480
uuuaaauucguu uucaauuuuaa aaaauauucuu gcacgcagaa cgcuaagggcc gcgggccgcgg 3540
auccauggcu aacaagcucu uccucgucug cgcaacucuc gccucugcu uccuccucac 3600
caacgcuucc aucuaccgca ccguugucga auucgaagaa gaugacgcca gcaaccccg 3660
agguccaaga cagagaugcc agaaggaguu ucagcaauca caacaccuaa gagcuugcca 3720
gagauggaug agcaagcaaa ugaggcaagg acgugguggu gguccuuccc ucgacgauga 3780
guucgauuuc gagggccccc agcaggggaa ccagcuacuc cagcagugcu gcaacgagcu 3840
ucgccaggaa gagccaguuu gcguuugccc caccuugaaa caagcugcca gggcaguuag 3900
ccuccaggga cagcacggac cauuccaauc caggaaaauu uaccagucag cuaaguacuu 3960
gccuaacauu ugcaagauc agcaaguugg ugaauugccc uuccagacca ccaucccuu 4020
cuucccuccu uacuacuagg guacucgagg uaagcucaac aaauuuuag aaaaauaaau 4080
uuauugaca uauagcaaua uuugauuugg caagauaaac uauagauuu ugcgauuugg 4140
aguuuuuaac ucuaaaauau cuaaaucguu uucaauuggu uuaaaauauau aucuugcauu 4200
uuuaaauucguu uuuaaaauaaa aaauauauau auauauauau aucuugcauu uuuaaauucguu 4260
uucaauuuuaa aaaauaucuu gcacgcagaa cgcugucgac uaccuagua guaaaggagg 4320
aagaaaggga ugguggucug gaagggacau ucaccaacu gcuggaucuu gcaaaugua 4380
ggcaaguacu uagcugacug guaaaauuuc cuggauugga augguccgug cugucccug 4440
aggcuaacug ccuggcagc uuguuucag guggggcaaa cgcaaacugg cucuuccug 4500
cgaagcucgu ugcagcacug cuggaguagc ugguaucuu gcugggggccc cucgaaucg 4560
aacucaucgu cgagggaagg accaccacca cguccuugcc ucauuugcu gcucauccau 4620
cucuggcaag cucuaggug uugugauugc ugaacuccu ucuggcauc cuugcuugga 4680
ccuacggggu ugcuggcguc aucuucucg aaucgacaa cggugcgua gauggaagcg 4740
uuggugagga ggaagcagag ggcgagaguu gcgcagacga ggaagagcu guuagccaug 4800
gaucaauucg ccuauagug agucguauua caauucacug gccgucguu uacaacgucg 4860
ugacugggaa aaccuggcg uuaccaca cuauccguu gcagcacauc ccccuucgc 4920
cagcuggcgu aaugcgaag aggcccgac cgau 4954

```

<210> 146

<211> 4456

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: DNA
construct coding for dsRNA for suppression of
multiple storage proteins

<400> 146

```

agcttggtac cgagctcgga tccactagta acggccgcca gtgtgctgga attcgccctt 60
gcggccgagg atcctcaggg tcttttcttg ccacttttct tgaacgcccg caaactcacg 120
tttgttggtc acggaagggg tctaattgga agagttattc cgggatgcgc cgagacgttc 180
atggagtcac cggatatttg agaaggtcaa ggtcagggtc agagtcaagg gttccgtgac 240
atgcaccaga aagtagagca cctacgggtg ggtgacacca ttgcaacacc atctggtgta 300
gctcaatggt tctacaacaa tggaaatgag cctctcattc ttgttgacgc cgcggatctc 360
gccagcaacc agaaccagct tgaccgcaac cttagaccat ttttgatagc cggaaacaac 420
ccacaagggc aggaatggct acaaggccga aagcaacaga agcaaaaca catcttcaat 480
ggcttcgcac ctgagatctt ggctcaagcc ttcaagatca atgtcgagac ggctcagcag 540
ctccagaacc agcaagataa ccgtggcaac atcgtcaagg tcaacggacc tttcggcgct 600
attaggccac ctttgagacg cggcgaaggc ggcaacaac cacatgaaat agctaattgt 660
ttagaggaga ctttgtgcac catgcgatgc actgaaaacc tcgatgacct gtcggatgct 720
gacgtgtaca agccatcact cggatacatt agcacactta acagctacaa tcttcctatc 780
ctcagacttc tccgccttag cgctcttcgt ggctccatcc gtaaaaggat ccgcggccgc 840
aaggcgcaat tctgcagatc cttcagggtc ttttcttgcc cactttcttg aacgccggca 900
aactcacgtt tgttggtcac ggaaggggtc taatgggaag agttattccg ggatgcgccc 960
agacgttcat ggagtcaccg gtatttgag aaggtcaagg tcagggtcag agtcaagggt 1020
tccgtgacat gcaccagaaa gtagagcacc tacgggtgcg tgacaccatt gcaacaccat 1080
ctggtgtagc tcaatggttc tacaacaatg gaaatgagcc tctcattctt gttgcagccg 1140
cggatctcgc cagcaaccag aaccagcttg accgcaacct tagaccattt ttgatagccg 1200

```

```

gaaacaaccc acaagggcag gaatggctac aaggccgaaa gcaacagaag caaaacaaca 1260
tcttcaatgg cttcgcacct gagatcttgg ctcaagcctt caagatcaat gtcgagacgg 1320
ctcagcagct ccagaaccag caagataacc gtggcaacat cgtcaaggtc aacggacctt 1380
tcggcgatcat tagggcacc ctagagacgg gcgaaggcgg ccaacaacca catgaaatag 1440
ctaattggttt agaggagact ttgtgcacca tgcgatgcac tgaaaacctc gatgacccgt 1500
cggatgctga cgtgtacaag ccatcactcg gatacattag cacacttaac agctacaatc 1560
ttcctatcct cagactcttc cgccttagcg ctcttcgtgg ctccatccgt aaaagatcct 1620
atggctaaca agctcttcct cgtctgcgca actctcgccc tctgcttcct cctcaccaac 1680
gcttccatct accgcaccgt tgtcgaattc gaagaagatg acgccagcaa ccccgtaggt 1740
ccaagacaga gatgccagaa ggagtttcag caatcacaac acctaagagc ttgccagaga 1800
tggatgagca agcaaatgag gcaaggacgt ggtgggtggc cttccctcga cgatgagttc 1860
gatttcgagg gccccagca gggataccag ctactccagc agtgctgcaa cgagcttcgc 1920
caggaagagc cagtttgctg ttgccccacc ttgaaacaag ctgccagggc agttagcctc 1980
cagggacagc acggaccatt ccaatccagg aaaattttacc agtcagctaa gtacttgctc 2040
aacatttgca agatccagca agttgggtgaa tgtcccttcc agaccaccat ccctttcttc 2100
cctccttact actagggtag atatccatca cactggcggc cgctcgacgt aagctcaaca 2160
aatctttaga aaattaattt tatgtgacat atgcaataat ttgatttggc aagataaact 2220
aatagatttt gcgatttggg gttttaaact ctaaataatc taaatcgttt tcaattgggt 2280
taaataatata tcttgcatct ttaatcgttt ttaattaaaa aatatatata tatatatata 2340
tcttgcatct ttaatcgttt tcaattttaa aaatatcttg cacgcagaac gcctcgacta 2400
ccctagtagt aaggagggaa gaaagggatg gtggtctgga agggacattc accaacttgc 2460
tggatcttgc aaatgcttag caagtactta gctgactggt aaattttcct ggattggaat 2520
ggtccgtgct gtccctggag gctaactgcc ctggcagctt gtttcaaggt ggggcaaacg 2580
caaactggct cttcctggcg aagctcgttg cagcactgct ggagtagctg gtatccctgc 2640
tggggggcct cgaaatcgaa ctcatcgctg aggggaaggac caccaccacg tccttgctc 2700
atttgcttgc tcatccatct ctggcaagct cttaggtggt gtgattgctg aaactcctc 2760
tggcatctct gtcttggaac tacggggttg ctggcgctcat cttcttcgaa ttcgacaacg 2820
gtgcggtaga tgggaagcgtt ggtgaggagg aagcagaggc cgagagttgc gcagacgagg 2880
aagagcttgt tagccatagg atcttttacg gatggagcca cgaagagcgc taaggcggag 2940
aagtctgagg ataggaagat tgtagctggt aagtgtgcta atgtatccga gtgatggctt 3000
gtacacgtca gcatccgacg ggtcatcgag gttttcagtg catcgcatgg tgcacaaagt 3060
ctcctctaaa ccattagcta tttcatgttg ttgttgggcg ccttcgccgc gtctcaaggg 3120
tggcctaata acgccgaaag gtccgttgac cttgacgatg ttgccacggt tatcttgctg 3180
gttctggagc tgctgagccg tctcgacatt gatcttgaag gcttgagcca agatctcagg 3240
tgcaagacca ttgaagatgt tgttttgctt ctggtgcttt cggccttgta gccattcctg 3300
cccttggtgg ttgtttccgg ctatcaaaaa tgggtctaagg ttgcggtcaa gctggttctg 3360
gttctggcg agatccgcgg ctgcaacaag aatgagaggc tcatttccat tgtttagtaa 3420
ccattgagct acacgagatg gtgttgcaat ggtgtcaccg caccgtaggt gctctacttt 3480
ctggtgcatg tcacggaacc cttgactctg accctgacct tgaccttctc caaataccgg 3540
tgactccatg aacgtctcgg cgcacccgg aataactctt cccattagac cccttcctg 3600
aacaacaaac gtgagtttgc cggcgttcaa gaaagtgggc aagaâaagac cctgaaggat 3660
ctgcagaatt cgcccttgcg gccgcggtac cttttacgga tggagccacg aagagcgcta 3720
aggcggagaa gtctgaggat aggaagattg tagctgttaa gtgtgctaat gtatccgagt 3780
gatggcttgt acacgtcagc atccgacggg tcatcgaggt tttcagtgca tcgcatggtg 3840
cacaaagtct cctctaaacc attagctatt tcatgtggtt gttggccgct ttcgcccgct 3900
ctcaagggtg gcctaataac gccgaaaggt ccgttgacct tgacgatgtt gccacgggta 3960
tcttgctggt tctggagctg ctgagccgtc tcgacattga tcttgaaggc ttgagccaag 4020
atctcagggt cgaagccatt gaagatgttg ttttgcttct gttgctttcg gccttgtagc 4080
cattcctgcc cttgtgggtt gtttcggct atcaaaaatg gtctaagggt gcggtcaagc 4140
tggttctggt tgctggcgag atccgcggct gcaacaagaa tgagaggctc atttccattg 4200
ttgtagaacc attgagctac accagatggt gttgcaatgg tgtcaccgca ccgtaggtgc 4260
tctactttct ggtgcatgtc acggaaccct tgactctgac cctgaccttg accttctcca 4320
aataccggtg actccatgaa cgtctcggcg catcccgga taactcttcc cattagaccc 4380
cttccgtgaa caacaaacgt gagtttgccg gcgttcaaga aagtgggcaa gaaaagaccc 4440
tgactcgagc atgcat 4456

```

<210> 147

<211> 4456

<212> RNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: RNA coding
for dsRNA for suppression of multiple storage
proteins

<400> 147

```

agcuugguac cgagcucgga uccacuagua acggccgcca gugugcugga auucgccccu 60
gcggccgcgg auccucaggg ucuuuuuucg cccacuuucu ugaacgccgg caaacucacg 120
uuuguuguuc acggaagggg ucuaauggga agaguuauuc cgggaugcgc cgagacguuc 180
auggagucac cgguauuugg agaaggucac ggucaggguu agagucaagg guuccgugac 240
augcaccaga aaguagagca ccuacggguc ggugacacca uugcaacacc aucuggugua 300
gcuaauggu ucuacaacaa uggaaaugag ccucucauuc uuguugcagc cgcggauuc 360
gccagcaacc agaaccagcu ugaccgcaac cuuagaccu uuuugauagc cggaaacaac 420
ccacaagggc aggaauuggcu acaaggccga aagcaacaga agcaaaacaa caucuucuu 480
ggcuucgcac cugagaucuu ggcucaagcc uucaagauc augucgagac ggcucagcag 540
cuccagaacc agcaagauaa ccguggcaac aucguicaagg ucaacggacc uuucggcguc 600
auuaggccac ccuugagacg cggcggaagg ggccaacaac cacaugaaau agcuauuggu 660
uuagaggaga cuuugugcag caugcgauuc acugaaaacc ucgaugaccc gucggauuc 720
gacguguaca agccaucacu cggauacauu agcacacuua acagcuacaa ucuuccuau 780
cucagacuuc uccgccuuag cgcucuucgu ggcuccaucc guaaaaggau ccgcggccgc 840
aagggcgaa ucuugcagau cuucaggguc uuucuuugc cacuuucuu aacgccggca 900
aacucacguu uguuguucac ggaagggguc uaaugggaag aguauuuccg ggaugcgccg 960
agacguucau ggagucaccg guauuuggag aaggucaagg ucagggucag agucaagggu 1020
uccgugacau gcaccagaaa guagagcacc uacggugcgg ugacaccau gcaacaccau 1080
cugguguaag ucauugguuc uacaacaau gaaugagacc ucucuuucuu guugcagccg 1140
cggauucgc cagcaaccag aaccagcuug accgcaaccu uagaccuuu uugauagccg 1200
gaaacaaccc acaagggcag gaauggcuac aaggccgaaa gcaacagaag caaaacaaca 1260
ucuucaauug cuucgcaccu gagaucuugg cucaagccuu caagaucuu gucagagcgg 1320
cucagcagcu ccagaaccag caagauaacc guggcaacau cgucaagguc aacggaccuu 1380
ucggcgucac uaggccaccc uugagacgcg gcgaaggcgg ccaacaacca caugaaauag 1440
cuauugguu agaggagacu uugugcacca ugcgaugcac ugaaaaccuc gaugacccgu 1500
cggauugcga cguguacaag ccaucacucg gauacauuag cacacuuaac agcuacaau 1560
uuccuauccu cagacuucuc cgccuuagcg cucuucgugg cuccauccgu aaaagaucuu 1620
auggcuaaca agcucuuccu cgucugcgca acucucgccc ucugcuuccu ccucaccaac 1680
gcuuccauuc accgcaccgu ugucgaauuc gaagaagaug acgccagcaa ccccguaagg 1740
ccaagacaga gaugccagaa ggaguuucag caaacaac accuaagagc uugccagaga 1800
uggaugagca agcaauugag gcaaggacgu gguggugugc cuuccucga cgaugaguuc 1860
gauuucgagg gccccagca gggauaccag cuacuccagc agugcugcaa cgagcuucgc 1920
caggaagagc caguugcgu uugccccacc uugaaacaag cugccagggc aguagccuc 1980
cagggacagc acggaccuu ccaauccagg aaaauuuacc agucagcuua guacuugccu 2040
aacauuugca agauccagca aguuggugaa ugucuccuuc agaccaccau cccuuucuu 2100
ccuccuuacu acuagguuag auauccauca cacuggcggc cgcucgacgu aagcucaaca 2160
aaucuuuaga aaauuaauu uaugugacau augcaauau uugauuuggc aagauaaacu 2220
aaauaguuuu gcgauuugga guuuuaaacu cuaaaauaau uaaucguuu ucaauugguu 2280
uaauauauua ucuugcauuu uuaaucguuu uuaauuaaaa aaauauaua uauauauaua 2340
ucuugcauuu uuaaucguuu ucauuuaaaa aaauaucuug cacgcagaac gccucgacua 2400
ccuaguagu aaggagggaa gaaagggau guggucugga agggacauuc accaacuugc 2460
uggaucuugc aaauuuagg caaguacuu gcugacuggu aaauuuuccu ggauuggau 2520
gguccgugcu guccugggag gcuaacugc cuggcagcu guuucaggu ggggcaaacg 2580
caaacuggcu cuuccuggcg aagcucguug cagcacugcu ggaguagcug guauccugc 2640
ugggggcccu cgaaaucgaa cucaucguc aggggaaggac caccaccacg uccuugccuc 2700
auuugcuugc ucauccauuc cuggcaagcu cuuagguguu gugauugcug aaacuccuuc 2760
uggcaucucu gucuuggacc uacggggguu cuggcgucuu cuucucgaa uucgacaacg 2820
gugcgguaga uggaagcguu ggugaggagg aagcagaggg cgagaguugc gcagacgagg 2880
aagagcuugu uagccauagg aucuuuucg gauggagcca cgaagagcgc uaaggcggag 2940

```

104

```

aagucugagg auaggaagau uguagcuguu aagugugcua auguauccga gugauggcuu 3000
guacacguca gcauccgacg ggucaucgag guuuucagug caucgcaugg ugcacaaagu 3060
cuccucuaaa ccauuagcua uucaugugg uuguuggccg ccuucgccgc gucucaaggg 3120
uggccuaaag acgccgaaag guccguugac cuugacgaug uugccacggu uaucuugcug 3180
guucuggagc ugcugagccg ucucgacauu gaudiugaag gcuugagcca agaucucagg 3240
ugcgaagcca uugaagaugu uguuuugcuu cuguugcuuu cggccuugua gccauuccug 3300
cccuuguggg uuguuuccgg cuaucaaaaa uggucuaaagg uugcggucua gcugguucug 3360
guugcuggcg agauccgagg cugcaacaag aaugagaggc ucauuuccau uguuguagaa 3420
ccaugagcu acaccgaug gugugcaau ggugucaccg caccguaggu gcucuacuuu 3480
cuggugcaug ucacggaacc cuugacucug acccugaccu ugaccuucuc caauuaccgg 3540
ugacuccaug aacgucucgg cgcaucccgg aaauaacucuu cccauuagac ccuuccgug 3600
aacaacaaac gugaguugc cggcguaaa gaaagugggc aagaaaagac ccugaaggau 3660
cugcagaauu cggccuugcg gccgcggau cuuuuacgga uggagccacg aagagcgua 3720
aggcggagaa gucugaggau aggaagaug uagcuguuaa gugugcuaau guauccgagu 3780
gauggcuugu acacgucagc auccgacggg ucaucgaggu uuucagugca ucgcauggug 3840
cacaagucu ccucuaaacc auuagcuau ucaugugguu guuggccgccc uucgccgcu 3900
cucaaggug gccuaaugac gccgaaaggu ccguugaccu ugacgauguu gccacgguua 3960
ucuugcuggu ucuggagcug cugagccguc ugcacauuga ucuugaaggc uugagccaag 4020
aucucaggug cgaagccauu gaagauguu uuuugcuucu guugcuuucg gccuuguagc 4080
cauuccugcc cuuguggguu guuuccggcu aucaaaaug gucuuagggu gcggucaagc 4140
ugguucuggu ugcuggcgag auccgcggcu gcaacaagaa ugagaggcuc auuuccauug 4200
uuuguagaacc auugagcuac accagauggu guugcaaugg ugucaccgca ccguaggugc 4260
ucuacuucuu ggugcauguc acggaacccu ugacucugac ccugaccuug accuucucca 4320
aauaccggug acuccaugaa cgucucggcg caucccgga uaacucuucc cauugaccc 4380
cuuccgugaa caacaaacgu gaguugccg gcguucaaga aagugggcaa gaaaagacc 4440
ugacucgagc augcau 4456

```

<210> 148

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 148

aaaaggcctg tgttccattt ggccggaac aac

33

<210> 149

<211> 31

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 149

aaagatatca ccctggagaa cgccacgagt g

31

<210> 150

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:
oligonucleotide primer

<400> 150

aaaaggccta tggctaacaa gctcttcctc gtc

33

<210> 151
 <211> 32
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 151
 aaagatatcc tagtagtaag gaggggaagaa ag 32
 <210> 152
 <211> 32
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 152
 ccgctcgagc tcaggggtctt ttcttgccca ct 32
 <210> 153
 <211> 32
 <212> DNA
 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz:
 oligonucleotide primer
 <400> 153
 ccggctcgacc tagtagtaag gaggggaagaa ag 32
 <210> 154
 <211> 1500
 <212> DNA
 <213> Arabidopsis thaliana
 <220>
 <221> CDS
 <222> (1)..(1497)
 <223> globuline-like protein
 <400> 154
 atg act aga ttt gcg gta ttg cca ctc tct gtt ctt ctt ctc gtt ctc 48
 Met Thr Arg Phe Ala Val Leu Pro Leu Ser Val Leu Leu Leu Val Leu
 1 5 10 15
 ttg ttc ctc tgc act gag tcg ttg gct aag tcg gag gag tct gaa gaa 96
 Leu Phe Leu Cys Thr Glu Ser Leu Ala Lys Ser Glu Glu Ser Glu Glu
 20 25 30
 tac gac gtc gct gtg cct tca tgt tgc ggg ttt tcg tct cct ctt ctg 144
 Tyr Asp Val Ala Val Pro Ser Cys Cys Gly Phe Ser Ser Pro Leu Leu
 35 40 45
 att aag aaa gat caa tgg aaa cca atc ttc gag acg aag ttt gga cag 192
 Ile Lys Lys Asp Gln Trp Lys Pro Ile Phe Glu Thr Lys Phe Gly Gln
 50 55 60
 atc tca acc gtt caa atc ggc aat gga tgc ggt gga atg gga cct tac 240
 Ile Ser Thr Val Gln Ile Gly Asn Gly Cys Gly Gly Met Gly Pro Tyr
 65 70 75 80

106

aag ata cat tcc ata acg ttg gag cca aac aca att ttg ctc cct ctt	288
Lys Ile His Ser Ile Thr Leu Glu Pro Asn Thr Ile Leu Leu Pro Leu	
85 90 95	
ctt ctt cat tcc gac atg gtc ttc ttc gtt gac tct gga agt ggg att	336
Leu Leu His Ser Asp Met Val Phe Phe Val Asp Ser Gly Ser Gly Ile	
100 105 110	
ctg aat tgg gtt gac gag gaa gcg aag agt acc gag atc aga cta gga	384
Leu Asn Trp Val Asp Glu Glu Ala Lys Ser Thr Glu Ile Arg Leu Gly	
115 120 125	
gac gtt tac agg cta cgt ccc ggt tct gtt ttc tac tta cag agc aaa	432
Asp Val Tyr Arg Leu Arg Pro Gly Ser Val Phe Tyr Leu Gln Ser Lys	
130 135 140	
ccg gat cct tgc ttt ggt gcc tat tct agt atc aca gat cta atg ttt	480
Pro Asp Pro Cys Phe Gly Ala Tyr Ser Ser Ile Thr Asp Leu Met Phe	
145 150 155 160	
ggt ttt gat gag aca att ctc cag tca gct ttt ggg gtt cct gag ggg	528
Gly Phe Asp Glu Thr Ile Leu Gln Ser Ala Phe Gly Val Pro Glu Gly	
165 170 175	
att ata gag ctg atg agg aac cgt acg aag cca cca ttg atc gtg agt	576
Ile Ile Glu Leu Met Arg Asn Arg Thr Lys Pro Pro Leu Ile Val Ser	
180 185 190	
gag acg ctt tgc aca cca ggt gtg gcc aac act tgg cag ctc caa ccg	624
Glu Thr Leu Cys Thr Pro Gly Val Ala Asn Thr Trp Gln Leu Gln Pro	
195 200 205	
cga tta cta aag ctc ttt gct gga agt gca gac ttg gtg gat aac aag	672
Arg Leu Leu Lys Leu Phe Ala Gly Ser Ala Asp Leu Val Asp Asn Lys	
210 215 220	
aag aag aaa gag aag aaa gag aag aag gag aag gtg aag aaa gcg aag	720
Lys Lys Lys Glu Lys Lys Glu Lys Lys Glu Lys Val Lys Lys Ala Lys	
225 230 235 240	
aca ttc aat gtt ttc gaa tct gaa ccg gat ttc gag agc ccc tac ggt	768
Thr Phe Asn Val Phe Glu Ser Glu Pro Asp Phe Glu Ser Pro Tyr Gly	
245 250 255	
cgt act ata acg att aac agg aag gat ctg aaa gtt tta aaa ggc tca	816
Arg Thr Ile Thr Ile Asn Arg Lys Asp Leu Lys Val Leu Lys Gly Ser	
260 265 270	
atg gtt gga gtt tcc atg gtg aat ctc act caa gga tct atg atg gga	864
Met Val Gly Val Ser Met Val Asn Leu Thr Gln Gly Ser Met Met Gly	
275 280 285	
ccg cac tgg aat cca tgg gct tgc gag att tct att gta ttg aaa gga	912
Pro His Trp Asn Pro Trp Ala Cys Glu Ile Ser Ile Val Leu Lys Gly	
290 295 300	
gca gga atg gtt agg gtt ctt agg tct tct ata tca tca aac aca tca	960
Ala Gly Met Val Arg Val Leu Arg Ser Ser Ile Ser Ser Asn Thr Ser	
305 310 315 320	
tca gag tgt aag aac gtg agg ttt aaa gta gag gaa gga gat att ttc	1008
Ser Glu Cys Lys Asn Val Arg Phe Lys Val Glu Glu Gly Asp Ile Phe	
325 330 335	
gca gtt cca cgg tta cat cca atg gct caa atg tct ttc aac aat gac	1056
Ala Val Pro Arg Leu His Pro Met Ala Gln Met Ser Phe Asn Asn Asp	
340 345 350	

107

tcg tta gtg ttc gtt ggg ttt act act tca gct aag aat aac gag ccg 1104
 Ser Leu Val Phe Val Gly Phe Thr Thr Ser Ala Lys Asn Asn Glu Pro
 355 360 365

cag ttc tta gcc ggg gag gac tcg gct ttg cgg atg ctt gac cgg caa 1152
 Gln Phe Leu Ala Gly Glu Asp Ser Ala Leu Arg Met Leu Asp Arg Gln
 370 375 380

gta ttg gct gca tcg ctt aat gtg agt agt gtg acg att gat gga ttg 1200
 Val Leu Ala Ala Ser Leu Asn Val Ser Ser Val Thr Ile Asp Gly Leu
 385 390 395 400

ttg gga gct cag aag gaa gct gtt atc ttg gaa tgt cat tct tgt gcg 1248
 Leu Gly Ala Gln Lys Glu Ala Val Ile Leu Glu Cys His Ser Cys Ala
 405 410 415

gaa gga gag ata gag aag ctt aag gtg gag ata gag agg aag aaa ata 1296
 Glu Gly Glu Ile Glu Lys Leu Lys Val Glu Ile Glu Arg Lys Lys Ile
 420 425 430

gat gat gag agg aag agg aga cat gat gaa agg aag aaa gaa gaa gaa 1344
 Asp Asp Glu Arg Lys Arg Arg His Asp Glu Arg Lys Lys Glu Glu Glu
 435 440 445

gag gcg aag aga gaa gag gaa gag agg agg aaa cgt gaa gaa gaa gaa 1392
 Glu Ala Lys Arg Glu Glu Glu Glu Arg Arg Lys Arg Glu Glu Glu Glu
 450 455 460

gag aag aag cgg tgg cca cct caa caa cca cca caa gaa gaa gaa ctt 1440
 Glu Lys Lys Arg Trp Pro Pro Gln Gln Pro Pro Gln Glu Glu Glu Leu
 465 470 475 480

agg gaa cgg caa cta ccg atg gag aaa gaa tgg gaa atg gaa ggt gaa 1488
 Arg Glu Arg Gln Leu Pro Met Glu Lys Glu Trp Glu Met Glu Gly Glu
 485 490 495

gag gag agt taa 1500
 Glu Glu Ser

<210> 155
 <211> 499
 <212> PRT
 <213> Arabidopsis thaliana

<400> 155
 Met Thr Arg Phe Ala Val Leu Pro Leu Ser Val Leu Leu Leu Val Leu
 1 5 10 15
 Leu Phe Leu Cys Thr Glu Ser Leu Ala Lys Ser Glu Glu Ser Glu Glu
 20 25 30
 Tyr Asp Val Ala Val Pro Ser Cys Cys Gly Phe Ser Ser Pro Leu Leu
 35 40 45
 Ile Lys Lys Asp Gln Trp Lys Pro Ile Phe Glu Thr Lys Phe Gly Gln
 50 55 60
 Ile Ser Thr Val Gln Ile Gly Asn Gly Cys Gly Gly Met Gly Pro Tyr
 65 70 75 80
 Lys Ile His Ser Ile Thr Leu Glu Pro Asn Thr Ile Leu Leu Pro Leu
 85 90 95
 Leu Leu His Ser Asp Met Val Phe Phe Val Asp Ser Gly Ser Gly Ile
 100 105 110
 Leu Asn Trp Val Asp Glu Glu Ala Lys Ser Thr Glu Ile Arg Leu Gly
 115 120 125

108

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

<210> 156
 <211> 1284
 <212> DNA
 <213> Glycine max

<220>
 <221> CDS
 <222> (1)..(1281)
 <223> 7S seed globuline

<400> 156
 atg gct tcc atc ctc cac tac ttt tta gcc ctc tct ctt tct tgc tct 48
 Met Ala Ser Ile Leu His Tyr Phe Leu Ala Leu Ser Leu Ser Cys Ser
 1 5 10 15
 ttt ctt ttc ttc tta tcc gac tca gtc acc cct aca aaa cca ata aac 96
 Phe Leu Phe Phe Leu Ser Asp Ser Val Thr Pro Thr Lys Pro Ile Asn
 20 25 30
 ctt gtt gtt cta ccc gtt caa aat gat ggt tcc aca ggg ctc cat tcg 144
 Leu Val Val Leu Pro Val Gln Asn Asp Gly Ser Thr Gly Leu His Ser
 35 40 45
 gcc aac ctc caa aaa aga acc cct cta atg caa gta cca gtc ctg gtg 192
 Ala Asn Leu Gln Lys Arg Thr Pro Leu Met Gln Val Pro Val Leu Val
 50 55 60
 gac ctc aat gga aat cac ttg tgg gtt aac tgt gag cag cag tac tca 240
 Asp Leu Asn Gly Asn His Leu Trp Val Asn Cys Glu Gln Gln Tyr Ser
 65 70 75 80
 tcc aaa acg tac caa gca ccc ttc tgc cac tcc acc caa tgc tct aga 288
 Ser Lys Thr Tyr Gln Ala Pro Phe Cys His Ser Thr Gln Cys Ser Arg
 85 90 95
 gcc aac acc cac caa tgc ctc agt tgc ccc gcg gca tca agg cca ggg 336
 Ala Asn Thr His Gln Cys Leu Ser Cys Pro Ala Ala Ser Arg Pro Gly
 100 105 110
 tgc cac aaa aac acg tgt ggc ctc atg tcc act aat ccc atc acc caa 384
 Cys His Lys Asn Thr Cys Gly Leu Met Ser Thr Asn Pro Ile Thr Gln
 115 120 125
 caa acc ggt tta ggt gaa cta gga gaa gac gtt ctt gca atc cac gcc 432
 Gln Thr Gly Leu Gly Glu Leu Gly Glu Asp Val Leu Ala Ile His Ala
 130 135 140
 aca caa ggg tcg acc caa caa ctt ggc cca ttg gtc aca gtc cca caa 480
 Thr Gln Gly Ser Thr Gln Gln Leu Gly Pro Leu Val Thr Val Pro Gln
 145 150 155 160
 ttc ctc ttt tct tgt gca cct tcc ttc ctt gtt caa aag ggt ctt cct 528
 Phe Leu Phe Ser Cys Ala Pro Ser Phe Leu Val Gln Lys Gly Leu Pro
 165 170 175
 aga aac act caa ggt gtg gct ggg tta ggc cat gca cca att tct ctt 576
 Arg Asn Thr Gln Gly Val Ala Gly Leu Gly His Ala Pro Ile Ser Leu
 180 185 190
 cca aac caa ctc gct tcc cac ttt ggc cta caa cgc caa ttc acc act 624
 Pro Asn Gln Leu Ala Ser His Phe Gly Leu Gln Arg Gln Phe Thr Thr
 195 200 205
 tgc ctt tct cgc tac cct act tca aag ggt gct ata ata ttc ggg gat 672
 Cys Leu Ser Arg Tyr Pro Thr Ser Lys Gly Ala Ile Ile Phe Gly Asp
 210 215 220

110

gca cct aac aac atg cga cag ttt caa aac caa gat att ttc cac gat 720
 Ala Pro Asn Asn Met Arg Gln Phe Gln Asn Gln Asp Ile Phe His Asp
 225 230 235 240
 ttg gcc ttc acc cca tta acc atc acc ctg cag gga gag tac aac gtg 768
 Leu Ala Phe Thr Pro Leu Thr Ile Thr Leu Gln Gly Glu Tyr Asn Val
 245 250 255
 aga gtc aac tca ata aga atc aac cag cac agt gtg ttc cca ctg aac 816
 Arg Val Asn Ser Ile Arg Ile Asn Gln His Ser Val Phe Pro Leu Asn
 260 265 270
 aag ata tca tcc acc atc gta ggg tcg acc tct gga gga acc atg att 864
 Lys Ile Ser Ser Thr Ile Val Gly Ser Thr Ser Gly Gly Thr Met Ile
 275 280 285
 agc acc tca act cct cac atg gtt ctc cag caa tcc gtg tac cag gct 912
 Ser Thr Ser Thr Pro His Met Val Leu Gln Gln Ser Val Tyr Gln Ala
 290 295 300
 ttc act cag gtg ttt gct cag cag cta cca aag caa gca cag gtg aaa 960
 Phe Thr Gln Val Phe Ala Gln Gln Leu Pro Lys Gln Ala Gln Val Lys
 305 310 315 320
 tct gtg gca cca ttt ggg tta tgc ttc aac tcc aac aag atc aat gca 1008
 Ser Val Ala Pro Phe Gly Leu Cys Phe Asn Ser Asn Lys Ile Asn Ala
 325 330 335
 tat cct agc gtg gac ctt gtg atg gac aag ccc aat ggt cct gtt tgg 1056
 Tyr Pro Ser Val Asp Leu Val Met Asp Lys Pro Asn Gly Pro Val Trp
 340 345 350
 aga atc tct ggt gag gac ttg atg gtg cag gca caa cct ggg gtc acg 1104
 Arg Ile Ser Gly Glu Asp Leu Met Val Gln Ala Gln Pro Gly Val Thr
 355 360 365
 tgt ttg ggt gtt atg aat gga gga atg caa cct aga gct gaa att acc 1152
 Cys Leu Gly Val Met Asn Gly Gly Met Gln Pro Arg Ala Glu Ile Thr
 370 375 380
 tta ggg gca cgt cag ttg gaa gag aac ctg gtg gtg ttc gat ctt gca 1200
 Leu Gly Ala Arg Gln Leu Glu Glu Asn Leu Val Val Phe Asp Leu Ala
 385 390 395 400
 agg tca agg gtt ggg ttt agc acc tca tca ctg cac tcg cat gga gtc 1248
 Arg Ser Arg Val Gly Phe Ser Thr Ser Ser Leu His Ser His Gly Val
 405 410 415
 aaa tgt gct gac ctc ttc aac ttt gcc aat gca tga 1284
 Lys Cys Ala Asp Leu Phe Asn Phe Ala Asn Ala
 420 425
 <210> 157
 <211> 427
 <212> PRT
 <213> Glycine max
 <400> 157
 Met Ala Ser Ile Leu His Tyr Phe Leu Ala Leu Ser Leu Ser Cys Ser
 1 5 10 15
 Phe Leu Phe Phe Leu Ser Asp Ser Val Thr Pro Thr Lys Pro Ile Asn
 20 25 30
 Leu Val Val Leu Pro Val Gln Asn Asp Gly Ser Thr Gly Leu His Ser
 35 40 45

111

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

<210> 158

<211> 814

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(720)

<223> 19kd zein

<400> 158

atg gca gcc aag att ttt gcc ctc ctt gcc ctc ctt gct ctt tca gca	48
Met Ala Ala Lys Ile Phe Ala Leu Leu Ala Leu Leu Ala Leu Ser Ala	
1 5 10 15	
aac gtt gct acc gcg act att att cca caa tgc tca caa caa tac ctc	96
Asn Val Ala Thr Ala Thr Ile Ile Pro Gln Cys Ser Gln Gln Tyr Leu	
20 25 30	
tct ccg gtg aca gcc gcg aga ttt gaa tac cca act ata caa tcc tac	144
Ser Pro Val Thr Ala Ala Arg Phe Glu Tyr Pro Thr Ile Gln Ser Tyr	
35 40 45	
agg cta caa cag gcc atc gca gca agc atc tta cgg tcg tta gca ttg	192
Arg Leu Gln Gln Ala Ile Ala Ala Ser Ile Leu Arg Ser Leu Ala Leu	
50 55 60	
act gtc caa caa cca tat gcc cta ttg caa caa cca tcc tta gtg aat	240
Thr Val Gln Gln Pro Tyr Ala Leu Leu Gln Gln Pro Ser Leu Val Asn	
65 70 75 80	
cta tat ctc caa aga atc gta gca caa caa cta caa caa caa ttg ctt	288
Leu Tyr Leu Gln Arg Ile Val Ala Gln Gln Leu Gln Gln Gln Leu Leu	
85 90 95	
cca aca atc aat gaa gta gtt gca gcg aac ctt gat gct tac ctc cag	336
Pro Thr Ile Asn Glu Val Val Ala Ala Asn Leu Asp Ala Tyr Leu Gln	
100 105 110	
caa caa caa ttt ctt cca ttc aat caa cta gct ggg gtg aac cct gct	384
Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Gly Val Asn Pro Ala	
115 120 125	
gct tac ttg cag gca caa cag cta cta cca ttc aac caa ctt gtc agg	432
Ala Tyr Leu Gln Ala Gln Gln Leu Leu Pro Phe Asn Gln Leu Val Arg	
130 135 140	
agc cct gct gcc ttc tta ctg cag caa cag ttg ttg cca ttc cat cta	480
Ser Pro Ala Ala Phe Leu Leu Gln Gln Gln Leu Leu Pro Phe His Leu	
145 150 155 160	
caa gtt gtg gca aac att gct gct ttc ttg caa caa caa caa ttg ctg	528
Gln Val Val Ala Asn Ile Ala Ala Phe Leu Gln Gln Gln Gln Leu Leu	
165 170 175	
cca ttt tac cca cag gtt gtg gga aac att aac gcc ttc ttg caa cag	576
Pro Phe Tyr Pro Gln Val Val Gly Asn Ile Asn Ala Phe Leu Gln Gln	
180 185 190	
caa cag ttg ctg cca ttc tac cca cag gat gtg gca aac aat gtc gcc	624
Gln Gln Leu Leu Pro Phe Tyr Pro Gln Asp Val Ala Asn Asn Val Ala	
195 200 205	
ttc tta caa caa caa caa ttg ctg cca ttt agc caa ctt gct ttg acg	672
Phe Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Leu Thr	
210 215 220	

113

aat cct acc acc tta ttg cag cag ccc acc att ggt ggt gcc atc ttc 720
 Asn Pro Thr Thr Leu Leu Gln Gln Pro Thr Ile Gly Gly Ala Ile Phe
 225 230 235 240

tagatttttt atgctttata ctgtaataat aaagttctca tactgatatg tgcaacttct 780
 cagtaataaa agattagaga tctatatatttt atta 814

<210> 159

<211> 240

<212> PRT

<213> Zea mays

<400> 159

Met Ala Ala Lys Ile Phe Ala Leu Leu Ala Leu Leu Ala Leu Ser Ala
 1 5 10 15

Asn Val Ala Thr Ala Thr Ile Ile Pro Gln Cys Ser Gln Gln Tyr Leu
 20 25 30

Ser Pro Val Thr Ala Ala Arg Phe Glu Tyr Pro Thr Ile Gln Ser Tyr
 35 40 45

Arg Leu Gln Gln Ala Ile Ala Ala Ser Ile Leu Arg Ser Leu Ala Leu
 50 55 60

Thr Val Gln Gln Pro Tyr Ala Leu Leu Gln Gln Pro Ser Leu Val Asn
 65 70 75 80

Leu Tyr Leu Gln Arg Ile Val Ala Gln Gln Leu Gln Gln Gln Leu Leu
 85 90 95

Pro Thr Ile Asn Glu Val Val Ala Ala Asn Leu Asp Ala Tyr Leu Gln
 100 105 110

Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Gly Val Asn Pro Ala
 115 120 125

Ala Tyr Leu Gln Ala Gln Gln Leu Leu Pro Phe Asn Gln Leu Val Arg
 130 135 140

Ser Pro Ala Ala Phe Leu Leu Gln Gln Gln Leu Leu Pro Phe His Leu
 145 150 155 160

Gln Val Val Ala Asn Ile Ala Ala Phe Leu Gln Gln Gln Gln Leu Leu
 165 170 175

Pro Phe Tyr Pro Gln Val Val Gly Asn Ile Asn Ala Phe Leu Gln Gln
 180 185 190

Gln Gln Leu Leu Pro Phe Tyr Pro Gln Asp Val Ala Asn Asn Val Ala
 195 200 205

Phe Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Leu Thr
 210 215 220

Asn Pro Thr Thr Leu Leu Gln Gln Pro Thr Ile Gly Gly Ala Ile Phe
 225 230 235 240

<210> 160

<211> 705

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(702)

<223> 19 kd zein B1

114

<400> 160
 atg gca gcc aaa ata ttt tgc ctc ctt atg ctc ctt ggt ctt tct gca 48
 Met Ala Ala Lys Ile Phe Cys Leu Leu Met Leu Leu Gly Leu Ser Ala
 1 5 10 15
 agt gct gct acg gcg acc att ttc cca caa tgc tca caa gct cct ata 96
 Ser Ala Ala Thr Ala Thr Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile
 20 25 30
 gct tcc ctt ctt ccc ccg tac ctc tca cca gcg gtg tct tcg gta tgt 144
 Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Val Ser Ser Val Cys
 35 40 45
 gaa aac cca att ctt caa ccc tat agg atc caa cag gca atc gca gct 192
 Glu Asn Pro Ile Leu Gln Pro Tyr Arg Ile Gln Gln Ala Ile Ala Ala
 50 55 60
 ggc atc tta cct tta tca ccc ttg ttc ctc caa caa tca tca gcc cta 240
 Gly Ile Leu Pro Leu Ser Pro Leu Phe Leu Gln Gln Ser Ser Ala Leu
 65 70 75 80
 tta cag cag tta cct ttg gtg cat tta ttg gca caa aac atc agg gca 288
 Leu Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala
 85 90 95
 caa caa cta caa caa ctt gtg cta gca aac ctt gct gcc tac tct cag 336
 Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln
 100 105 110
 caa caa cag ttt ctt cca ttc aac caa cta gct gca ttg aac tct gct 384
 Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn Ser Ala
 115 120 125
 tct tat ttg caa caa caa caa cta cca ttc agc cag cta tct gct gcc 432
 Ser Tyr Leu Gln Gln Gln Gln Leu Pro Phe Ser Gln Leu Ser Ala Ala
 130 135 140
 tac ccc cag caa ttt ctt cca ttc aac caa ctg aca gct ttg aac tct 480
 Tyr Pro Gln Gln Phe Leu Pro Phe Asn Gln Leu Thr Ala Leu Asn Ser
 145 150 155 160
 cct gct tat tta cag cag caa caa cta cta cca ttc agc cag cta gct 528
 Pro Ala Tyr Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala
 165 170 175
 ggt gtg agc cct gct acc ttc ttg aca caa cca caa ttg ttg ccg ttc 576
 Gly Val Ser Pro Ala Thr Phe Leu Thr Gln Pro Gln Leu Leu Pro Phe
 180 185 190
 tac cag cac gct gcg cct aac gct ggc acc ctc tta caa ctg caa caa 624
 Tyr Gln His Ala Ala Pro Asn Ala Gly Thr Leu Leu Gln Leu Gln Gln
 195 200 205
 ttg ctg cca ttc aac caa ctt gct ttg aca aac cca aca gca ttc tac 672
 Leu Leu Pro Phe Asn Gln Leu Ala Leu Thr Asn Pro Thr Ala Phe Tyr
 210 215 220
 caa caa ccc atc att ggt ggt gcc ctc ttt tag 705
 Gln Gln Pro Ile Ile Gly Gly Ala Leu Phe
 225 230
 <210> 161
 <211> 234
 <212> PRT
 <213> Zea mays
 <400> 161
 Met Ala Ala Lys Ile Phe Cys Leu Leu Met Leu Leu Gly Leu Ser Ala
 1 5 10 15

115

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

<210> 162

<211> 804

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(801)

<223> 19 kd zein B2

<400> 162

atg gca gcc aaa ata ttt tgc ctc att atg ctc ctt ggt ctt tct gca 48
Met Ala Ala Lys Ile Phe Cys Leu Ile Met Leu Leu Gly Leu Ser Ala
1 5 10 15
agt gct gct acg gcg agc att ttc ccg caa tgc tca caa gct cct ata 96
Ser Ala Ala Thr Ala Ser Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile
20 25 30
gct tcc ctt ctt ccc cca tac ctc tca cca gcg atg tct tca gta tgt 144
Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Met Ser Ser Val Cys
35 40 45
gaa aat cca att ctt cta ccc tac agg atc caa cag gca atc gca gca 192
Glu Asn Pro Ile Leu Leu Pro Tyr Arg Ile Gln Gln Ala Ile Ala Ala
50 55 60

116

```

ggc atc tta cct tta tca ccc ttg ttc ctc caa caa tca tca gcc cta 240
Gly Ile Leu Pro Leu Ser Pro Leu Phe Leu Gln Gln Ser Ser Ala Leu
65 70 75 80

tta cag cag tta cct ttg gtg cat tta ttg gca caa aac atc agg gca 288
Leu Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala
85 90 95

caa caa cta caa caa ctc gtg cta gca aac ctt gct gcc tac tct cag 336
Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln
100 105 110

caa cag cag tta cct ttg gtg cat ttg ttg gca caa aac atc agg gca 384
Gln Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala
115 120 125

caa caa cta caa caa ctc gtg cta gca aac ctt gct gcc tac tct cag 432
Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln
130 135 140

caa caa cag ttt ctg cca ttc aac caa cta gct gca ttg aac tct gct 480
Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn Ser Ala
145 150 155 160

gct tat ttg cag caa caa caa cta cta cca ttc agc cag cta gct gct 528
Ala Tyr Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Ala
165 170 175

gcc tac ccc cgg caa ttt ctt cca ttc aac caa ctg gca gca ttg aac 576
Ala Tyr Pro Arg Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn
180 185 190

tct cat gct tat gta caa caa caa caa cta cta cca ttc agc cag cta 624
Ser His Ala Tyr Val Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu
195 200 205

gct gct gtg agc cct gct gcc ttc ttg aca cag caa cat ttg ttg ccg 672
Ala Ala Val Ser Pro Ala Ala Phe Leu Thr Gln Gln His Leu Leu Pro
210 215 220

ttc tac ctg cac act gcg cct aac gtt ggc acc ctc tta caa ctg caa 720
Phe Tyr Leu His Thr Ala Pro Asn Val Gly Thr Leu Leu Gln Leu Gln
225 230 235 240

caa ttg ctg cca ttc gac caa ctt gct ttg aca aac cca gca gtg ttc 768
Gln Leu Leu Pro Phe Asp Gln Leu Ala Leu Thr Asn Pro Ala Val Phe
245 250 255

tac caa caa ccc atc att ggt ggt gcc ctc ttt tag 804
Tyr Gln Gln Pro Ile Ile Gly Gly Ala Leu Phe
260 265

```

<210> 163

<211> 267

<212> PRT

<213> Zea mays

<400> 163

```

Met Ala Ala Lys Ile Phe Cys Leu Ile Met Leu Leu Gly Leu Ser Ala
1 5 10 15

Ser Ala Ala Thr Ala Ser Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile
20 25 30

Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Met Ser Ser Val Cys
35 40 45

```

117

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

<210> 164

<211> 801

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(798)

<223> 22kd alpha-zein

<400> 164

atg gct acc aag ata tta gcc ctc ctt gcg ctc ctt tcc ctt tca gtg 48
 Met Ala Thr Lys Ile Leu Ala Leu Leu Ala Leu Leu Ser Leu Ser Val
 1 5 10 15
 agc gca aca act gca ttc att att cca caa tgc tca ctt gct cct aat 96
 Ser Ala Thr Thr Ala Phe Ile Ile Pro Gln Cys Ser Leu Ala Pro Asn
 20 25 30
 gcc att att cca cag ttc ctc cca tca gtt aca tca atg ggc atc gaa 144
 Ala Ile Ile Pro Gln Phe Leu Pro Ser Val Thr Ser Met Gly Ile Glu
 35 40 45
 cac cct att gtg caa gcc tat agg cta caa caa gcg ctt gcg gcg agc 192
 His Pro Ile Val Gln Ala Tyr Arg Leu Gln Gln Ala Leu Ala Ala Ser
 50 55 60

118

gtc tta caa caa ccg ttt gcc caa tta caa caa caa tcc ttg gca cat 240
 Val Leu Gln Gln Pro Phe Ala Gln Leu Gln Gln Gln Ser Leu Ala His
 65 70 75 80
 cta acc ata caa acc atc gca aca caa cta gag caa cag ttt gtg ccc 288
 Leu Thr Ile Gln Thr Ile Ala Thr Gln Leu Glu Gln Gln Phe Val Pro
 85 90 95
 gca ttg agc caa cta gcc gcg gtg aac cct gtc tcc tac ttg caa cag 336
 Ala Leu Ser Gln Leu Ala Ala Val Asn Pro Val Ser Tyr Leu Gln Gln
 100 105 110
 caa atg ctt gca tcc aac cca ctt gct ctg gcg aac aca gcc gca tac 384
 Gln Met Leu Ala Ser Asn Pro Leu Ala Leu Ala Asn Thr Ala Ala Tyr
 115 120 125
 cag caa caa cta cag ttg caa cag ttt cta cca gct ctt agt caa cta 432
 Gln Gln Gln Leu Gln Leu Gln Gln Phe Leu Pro Ala Leu Ser Gln Leu
 130 135 140
 gcc agg gtg aac cct gcc aca tac ctg caa cag caa caa ctg ctt tca 480
 Ala Arg Val Asn Pro Ala Thr Tyr Leu Gln Gln Gln Leu Leu Ser
 145 150 155 160
 tct agc cca ctc gct gtg ggc aat gcg gct aca tac ctg caa cag cag 528
 Ser Ser Pro Leu Ala Val Gly Asn Ala Ala Thr Tyr Leu Gln Gln Gln
 165 170 175
 ctg cta caa cag atc gta ccg gct ctt agt cag cta gtt gtg gcg aac 576
 Leu Leu Gln Gln Ile Val Pro Ala Leu Ser Gln Leu Val Val Ala Asn
 180 185 190
 cct act gcc tac tta caa cag ctt ctt cca ttc aac caa cta gat gtg 624
 Pro Thr Ala Tyr Leu Gln Gln Leu Leu Pro Phe Asn Gln Leu Asp Val
 195 200 205
 gca aac tct gct gcg tac cta caa cag cgg cag caa cta ctt aat cca 672
 Ala Asn Ser Ala Ala Tyr Leu Gln Gln Arg Gln Gln Leu Leu Asn Pro
 210 215 220
 ctt gca gcg gct aac cca ttg gtc gcc gcc ttc ctg caa cag caa caa 720
 Leu Ala Ala Ala Asn Pro Leu Val Ala Ala Phe Leu Gln Gln Gln Gln
 225 230 235 240
 ttt ctg cca tac aac caa atc tct ttg atg aac ctt gcc ttg tca agg 768
 Phe Leu Pro Tyr Asn Gln Ile Ser Leu Met Asn Leu Ala Leu Ser Arg
 245 250 255
 cag caa ccg atc gtt gga ggt gcc atc ttt tag 801
 Gln Gln Pro Ile Val Gly Gly Ala Ile Phe
 260 265

<210> 165

<211> 266

<212> PRT

<213> Zea mays

<400> 165

Met Ala Thr Lys Ile Leu Ala Leu Leu Ala Leu Leu Ser Leu Ser Val
 1 5 10 15

Ser Ala Thr Thr Ala Phe Ile Ile Pro Gln Cys Ser Leu Ala Pro Asn
 20 25 30

Ala Ile Ile Pro Gln Phe Leu Pro Ser Val Thr Ser Met Gly Ile Glu
 35 40 45

119

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

<210> 166

<211> 471

<212> DNA

<213> Oryza sativa

<220>

<221> CDS

<222> (1)..(468)

<223> prolamin

<400> 166

atg aag atc att ttc gta ttt gct ctc ctt gct att gtt gca tgc aac 48
 Met Lys Ile Ile Phe Val Phe Ala Leu Leu Ala Ile Val Ala Cys Asn
 1 5 10 15
 gct tct gca cgg ttt gat gct ctt agt caa agt tat aga caa tat caa 96
 Ala Ser Ala Arg Phe Asp Ala Leu Ser Gln Ser Tyr Arg Gln Tyr Gln
 20 25 30
 cta caa tcg cat ctc ctg cta cag caa caa gtg ctc agc cca tgc agt 144
 Leu Gln Ser His Leu Leu Leu Gln Gln Gln Val Leu Ser Pro Cys Ser
 35 40 45
 gag ttc gta agg caa cag cat agc ata gtg gca acc ccc ttc tgg caa 192
 Glu Phe Val Arg Gln Gln His Ser Ile Val Ala Thr Pro Phe Trp Gln
 50 55 60

120

cca gct acg ttt caa ttg ata aac aac caa gtc atg cag caa cag tgt 240
 Pro Ala Thr Phe Gln Leu Ile Asn Asn Gln Val Met Gln Gln Gln Cys
 65 70 75 80

tgc caa cag ctc agg ctg gta gcg caa caa tct cac tac cag gcc att 288
 Cys Gln Gln Leu Arg Leu Val Ala Gln Gln Ser His Tyr Gln Ala Ile
 85 90 95

agt agc gtt cag gcg att gtg cag caa cta cag ctg cag cag gtc ggt 336
 Ser Ser Val Gln Ala Ile Val Gln Gln Leu Gln Leu Gln Gln Val Gly
 100 105 110

gtt gtc tac ttt gat cag act caa gct caa gct caa gct ttg ctg gcc 384
 Val Val Tyr Phe Asp Gln Thr Gln Ala Gln Ala Gln Ala Leu Leu Ala
 115 120 125

tta aac ttg cca tcc ata tgt ggt atc tat cct aac tac tac att gct 432
 Leu Asn Leu Pro Ser Ile Cys Gly Ile Tyr Pro Asn Tyr Tyr Ile Ala
 130 135 140

ccg agg agc att ccc acc gtt ggt ggt gtc tgg tac tga 471
 Pro Arg Ser Ile Pro Thr Val Gly Gly Val Trp Tyr
 145 150 155

<210> 167

<211> 156

<212> PRT

<213> Oryza sativa

<400> 167

Met Lys Ile Ile Phe Val Phe Ala Leu Leu Ala Ile Val Ala Cys Asn
 1 5 10 15

Ala Ser Ala Arg Phe Asp Ala Leu Ser Gln Ser Tyr Arg Gln Tyr Gln
 20 25 30

Leu Gln Ser His Leu Leu Leu Gln Gln Gln Val Leu Ser Pro Cys Ser
 35 40 45

Glu Phe Val Arg Gln Gln His Ser Ile Val Ala Thr Pro Phe Trp Gln
 50 55 60

Pro Ala Thr Phe Gln Leu Ile Asn Asn Gln Val Met Gln Gln Gln Cys
 65 70 75 80

Cys Gln Gln Leu Arg Leu Val Ala Gln Gln Ser His Tyr Gln Ala Ile
 85 90 95

Ser Ser Val Gln Ala Ile Val Gln Gln Leu Gln Leu Gln Gln Val Gly
 100 105 110

Val Val Tyr Phe Asp Gln Thr Gln Ala Gln Ala Gln Ala Leu Leu Ala
 115 120 125

Leu Asn Leu Pro Ser Ile Cys Gly Ile Tyr Pro Asn Tyr Tyr Ile Ala
 130 135 140

Pro Arg Ser Ile Pro Thr Val Gly Gly Val Trp Tyr
 145 150 155

<210> 168

<211> 645

<212> DNA

<213> Avena sativa

<220>

<221> CDS

121

<222> (1)..(642)

<223> avenin

<400> 168

```

atg aag atc ttc ttc ttc tta gct ctc ctt gct ctg gta gtg agc gcc 48
Met Lys Ile Phe Phe Phe Leu Ala Leu Leu Ala Leu Val Val Ser Ala
  1             5             10             15

acc ttt gca caa tat gca gaa tct gac ggt agt tat gag gaa gtg gag 96
Thr Phe Ala Gln Tyr Ala Glu Ser Asp Gly Ser Tyr Glu Glu Val Glu
             20             25             30

ggt tct cat gat cga tgc caa caa cat cag atg aag ctg gac tct tgc 144
Gly Ser His Asp Arg Cys Gln Gln His Gln Met Lys Leu Asp Ser Cys
             35             40             45

aga gag tac gtg gcg gag cgg tgc aca acg atg aga gat ttt ccg atc 192
Arg Glu Tyr Val Ala Glu Arg Cys Thr Thr Met Arg Asp Phe Pro Ile
             50             55             60

acc tgg cca tgg aaa tgg tgg aag ggt ggt tgc gag gag ctc cgc aat 240
Thr Trp Pro Trp Lys Trp Trp Lys Gly Gly Cys Glu Glu Leu Arg Asn
             65             70             75             80

gag tgc tgc caa ctg ttg ggc cag atg cca tcg gag tgt cgc tgt gat 288
Glu Cys Cys Gln Leu Leu Gly Gln Met Pro Ser Glu Cys Arg Cys Asp
             85             90             95

gcg att tgg aga tca atc cag cgc gag ctt ggt ggc ttc ttt gga act 336
Ala Ile Trp Arg Ser Ile Gln Arg Glu Leu Gly Gly Phe Phe Gly Thr
             100             105             110

caa caa ggt ctg ata ggg aaa agg ttg aag ata gcc aag agt ttg ccc 384
Gln Gln Gly Leu Ile Gly Lys Arg Leu Lys Ile Ala Lys Ser Leu Pro
             115             120             125

acg cag tca aca tgg gcc ctg agt gca ata tcc cca aac tcc atg gtt 432
Thr Gln Ser Thr Trp Ala Leu Ser Ala Ile Ser Pro Asn Ser Met Val
             130             135             140

agc cac att gct gga aag agc tcc att ctt cgt gcc ttg ccc gtg gat 480
Ser His Ile Ala Gly Lys Ser Ser Ile Leu Arg Ala Leu Pro Val Asp
             145             150             155             160

gtc ctc gcc aat gca tac cgc att tcc agg caa gaa gcc cga aac ctc 528
Val Leu Ala Asn Ala Tyr Arg Ile Ser Arg Gln Glu Ala Arg Asn Leu
             165             170             175

aaa aac aac agg gga caa gag tct ggt gta ttc act cca aaa ttt acc 576
Lys Asn Asn Arg Gly Gln Glu Ser Gly Val Phe Thr Pro Lys Phe Thr
             180             185             190

caa acg agc ttc caa cct tat cca gag ggc gag gat gag tca tct ttg 624
Gln Thr Ser Phe Gln Pro Tyr Pro Glu Gly Glu Asp Glu Ser Ser Leu
             195             200             205

att aat aag gca tca gag taa 645
Ile Asn Lys Ala Ser Glu
             210

```

<210> 169

<211> 214

<212> PRT

<213> Avena sativa

<400> 169

```

Met Lys Ile Phe Phe Phe Leu Ala Leu Leu Ala Leu Val Val Ser Ala
  1             5             10             15

```

122

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

<210> 170

<211> 1044

<212> DNA

<213> Hordeum vulgare

<220>

<221> CDS

<222> (1)..(1041)

<223> c-hordein

<220>

<221> misc_feature

<222> (481)..(482)

<223> /transl_except=(pos:481..483,aa:OTHER)

<400> 170

atg aag acg ttc ctc acc ttt gtc ctc ctt gcc atg gcg atg agc atc 48
 Met Lys Thr Phe Leu Thr Phe Val Leu Leu Ala Met Ala Met Ser Ile
 1 5 10 15
 gtc act acc gct agg cag cta aac cct agc cac caa gag ttg caa tca 96
 Val Thr Thr Ala Arg Gln Leu Asn Pro Ser His Gln Glu Leu Gln Ser
 20 25 30
 cca caa caa cca ttt ctg aaa caa caa tca tat ctg caa caa cca tat 144
 Pro Gln Gln Pro Phe Leu Lys Gln Gln Ser Tyr Leu Gln Gln Pro Tyr
 35 40 45

123

cca caa caa cca tat cta ccg cag caa cca ttc ccc aca ccc caa caa	192
Pro Gln Gln Pro Tyr Leu Pro Gln Gln Pro Phe Pro Thr Pro Gln Gln	
50 55 60	
ttt ttc ccc tat cta cca cag caa aca ttt ccc cca tcc caa caa cca	240
Phe Phe Pro Tyr Leu Pro Gln Gln Thr Phe Pro Pro Ser Gln Gln Pro	
65 70 75 80	
aac ccc cta caa cca caa caa cca ttc ccc ctg caa ccc caa cca cca	288
Asn Pro Leu Gln Pro Gln Gln Pro Phe Pro Leu Gln Pro Gln Pro Pro	
85 90 95	
caa caa cct ttt cct cag ccc caa caa cca aat ccc cag caa cca caa	336
Gln Gln Pro Phe Pro Gln Pro Gln Gln Pro Asn Pro Gln Gln Pro Gln	
100 105 110	
caa cct ttc ccc cgg caa cca caa caa ata gta ccc cag caa cca caa	384
Gln Pro Phe Pro Arg Gln Pro Gln Gln Ile Val Pro Gln Gln Pro Gln	
115 120 125	
caa cca ttc cct cag caa cca caa caa cct ttt cct cag ccc caa caa	432
Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Phe Pro Gln Pro Gln Gln	
130 135 140	
cca ttc tct tgg caa cca caa caa cca ttt ctc cag ccc cta caa cta	480
Pro Phe Ser Trp Gln Pro Gln Gln Pro Phe Leu Gln Pro Leu Gln Leu	
145 150 155 160	
tag ccc ctg caa gca caa caa cca ttc ccc ttg caa cct caa cta cca	528
Pro Leu Gln Ala Gln Gln Pro Phe Pro Leu Gln Pro Gln Leu Pro	
165 170 175	
ttt ccg caa ccc caa caa cca att gga cag caa cca aaa caa cca ctc	576
Phe Pro Gln Pro Gln Gln Pro Ile Gly Gln Gln Pro Lys Gln Pro Leu	
180 185 190	
ctg cag caa cca caa caa aca att ccc cag caa cca caa caa cca ttc	624
Leu Gln Gln Pro Gln Gln Thr Ile Pro Gln Gln Pro Gln Gln Pro Phe	
195 200 205	
ccc ctg cag ccg caa caa cca ttc ccc caa caa cca caa caa cca ctt	672
Pro Leu Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Leu	
210 215 220	
ccc caa caa ccc caa caa ata att tcc cag caa ccc caa caa cca ttc	720
Pro Gln Gln Pro Gln Gln Ile Ile Ser Gln Gln Pro Gln Gln Pro Phe	
225 230 235 240	
cct cta caa cct caa caa cca ttc ccc caa ccc caa cca ttc ccc cag	768
Pro Leu Gln Pro Gln Gln Pro Phe Pro Gln Pro Gln Pro Phe Pro Gln	
245 250 255	
gag caa ccc caa caa gca ttc ccc cta caa ccg caa caa cca ttc ccc	816
Glu Gln Pro Gln Gln Ala Phe Pro Leu Gln Pro Gln Gln Pro Phe Pro	
260 265 270	
gag gaa tca gaa caa ata att acc caa caa cca ttc cct cta caa cca	864
Glu Glu Ser Glu Gln Ile Ile Thr Gln Gln Pro Phe Pro Leu Gln Pro	
275 280 285	
caa caa ctg ttc ccc cag caa cca caa caa cca ctt ccc cag ccc caa	912
Gln Gln Leu Phe Pro Gln Gln Pro Gln Gln Pro Leu Pro Gln Pro Gln	
290 295 300	
caa cca ttc cgc caa cta cca aaa tat ata att ccc cag caa cct caa	960
Gln Pro Phe Arg Gln Leu Pro Lys Tyr Ile Ile Pro Gln Gln Pro Gln	
305 310 315 320	

caa cca ttc ctt ctg	caa cca cac caa cct cag caa cct tat gca caa	1008
Gln Pro Phe Leu Leu	Gln Pro His Gln Pro Gln Gln Pro Tyr Ala Gln	
	325 330 335	
caa gac atc tgg agt gat ata gcc ctc ttg ggc taa		1044
Gln Asp Ile Trp Ser Asp Ile Ala Leu Leu Gly		
	340 345	

```
<210> 171
<211> 160
<212> PRT
<213> Hordeum vulgare
```

<400> 171															
Met	Lys	Thr	Phe	Leu	Thr	Phe	Val	Leu	Leu	Ala	Met	Ala	Met	Ser	Ile
1				5					10					15	
Val	Thr	Thr	Ala	Arg	Gln	Leu	Asn	Pro	Ser	His	Gln	Glu	Leu	Gln	Ser
			20					25					30		
Pro	Gln	Gln	Pro	Phe	Leu	Lys	Gln	Gln	Ser	Tyr	Leu	Gln	Gln	Pro	Tyr
		35					40					45			
Pro	Gln	Gln	Pro	Tyr	Leu	Pro	Gln	Gln	Pro	Phe	Pro	Thr	Pro	Gln	Gln
	50					55					60				
Phe	Phe	Pro	Tyr	Leu	Pro	Gln	Gln	Thr	Phe	Pro	Pro	Ser	Gln	Gln	Pro
65				70					75						80
Asn	Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	Pro	Pro
				85					90					95	
Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Gln	Pro	Asn	Pro	Gln	Gln	Pro	Gln
			100					105					110		
Gln	Pro	Phe	Pro	Arg	Gln	Pro	Gln	Gln	Ile	Val	Pro	Gln	Gln	Pro	Gln
		115					120					125			
Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Gln
	130					135					140				
Pro	Phe	Ser	Trp	Gln	Pro	Gln	Gln	Pro	Phe	Leu	Gln	Pro	Leu	Gln	Leu
145					150					155					160

```
<210> 172
<211> 186
<212> PRT
<213> Hordeum vulgare
```

<400> 172																
Pro	Leu	Gln	Ala	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	Leu	Pro	Phe	
1				5					10					15		
Pro	Gln	Pro	Gln	Gln	Pro	Ile	Gly	Gln	Gln	Pro	Lys	Gln	Pro	Leu	Leu	
			20					25					30			
Gln	Gln	Pro	Gln	Gln	Thr	Ile	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	
			35				40					45				
Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	Pro	
	50					55					60					
Gln	Gln	Pro	Gln	Gln	Ile	Ile	Ser	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	
65					70					75					80	
Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Pro	Phe	Pro	Gln	Glu	
				85					90					95		
Gln	Pro	Gln	Gln	Ala	Phe	Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Glu	
			100					105					110			
Glu	Ser	Glu	Gln	Ile	Ile	Thr	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	
		115					120					125				
Gln	Leu	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	Pro	Gln	Pro	Gln	Gln	
	130					135					140					
Pro	Phe	Arg	Gln	Leu	Pro	Lys	Tyr	Ile	Ile	Pro	Gln	Gln	Pro	Gln	Gln	
145					150					155					160	

125

Pro Phe Leu Leu Gln Pro His Gln Pro Gln Gln Pro Tyr Ala Gln Gln
 165 170 175
 Asp Ile Trp Ser Asp Ile Ala Leu Leu Gly
 180 185

<210> 173

<211> 924

<212> DNA

<213> Triticum aestivum

<220>

<221> CDS

<222> (1)..(921)

<223> glutenin-1D1

<400> 173

atg aag acc ttc ctc gtc ttt gcc ctc ctc gcc gtt gcg gcg aca agt	48
Met Lys Thr Phe Leu Val Phe Ala Leu Leu Ala Val Ala Ala Thr Ser	
1 5 10 15	
gca att gcg cag atg gag act aga tgc atc cct ggt ttg gag aga cca	96
Ala Ile Ala Gln Met Glu Thr Arg Cys Ile Pro Gly Leu Glu Arg Pro	
20 25 30	
tgg cag cag caa cca tta cca cca caa cag aca ttt cca caa caa cca	144
Trp Gln Gln Gln Pro Leu Pro Pro Gln Gln Thr Phe Pro Gln Gln Pro	
35 40 45	
cta ttt tca caa caa caa caa caa caa cta ttt cct caa caa cca tca	192
Leu Phe Ser Gln Gln Gln Gln Gln Gln Leu Phe Pro Gln Gln Pro Ser	
50 55 60	
ttt tcg cag caa caa cca cca ttt tgg cag caa caa cca cca ttt tct	240
Phe Ser Gln Gln Gln Pro Phe Trp Gln Gln Gln Pro Pro Phe Ser	
65 70 75 80	
cag caa caa cca att cta cca cag caa cca cca ttt tcg cag caa caa	288
Gln Gln Gln Pro Ile Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln	
85 90 95	
caa cta gtt cta ccg caa caa cca cca ttt tca cag caa caa caa cca	336
Gln Leu Val Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln Gln Pro	
100 105 110	
gtt tta cct cca caa caa tca cct ttt cca caa caa caa caa caa cac	384
Val Leu Pro Pro Gln Gln Ser Pro Phe Pro Gln Gln Gln Gln Gln His	
115 120 125	
caa cag ctg gtg caa caa caa atc cct gtt gtt cag cca tcc att ttg	432
Gln Gln Leu Val Gln Gln Gln Ile Pro Val Val Gln Pro Ser Ile Leu	
130 135 140	
cag cag cta aac cca tgc aag gta ttc ctc cag cag cag tgc agc cct	480
Gln Gln Leu Asn Pro Cys Lys Val Phe Leu Gln Gln Gln Cys Ser Pro	
145 150 155 160	
gtg gca atg cca caa cgt ctt gct agg tcg caa atg ttg cag cag agc	528
Val Ala Met Pro Gln Arg Leu Ala Arg Ser Gln Met Leu Gln Gln Ser	
165 170 175	
agt tgc cat gtg atg caa caa caa tgt tgc cag cag ttg ccg caa atc	576
Ser Cys His Val Met Gln Gln Gln Cys Cys Gln Gln Leu Pro Gln Ile	
180 185 190	
ccc cag caa tcc cgc tat gag gca atc cgt gct atc atc tac tcc atc	624
Pro Gln Gln Ser Arg Tyr Glu Ala Ile Arg Ala Ile Ile Tyr Ser Ile	
195 200 205	

126

atc ctg caa gaa caa caa cag gtt cag ggt tcc atc caa tct cag cag 672
 Ile Leu Gln Glu Gln Gln Gln Val Gln Gly Ser Ile Gln Ser Gln Gln
 210 215 220
 cag caa ccc caa cag ttg ggc caa tgt gtt tcc caa ccc caa cag cag 720
 Gln Gln Pro Gln Gln Leu Gly Gln Cys Val Ser Gln Pro Gln Gln Gln
 225 230 235 240
 tcg cag cag caa ctc ggg caa caa cct caa caa caa caa ttg gca cag 768
 Ser Gln Gln Gln Leu Gly Gln Gln Pro Gln Gln Gln Gln Leu Ala Gln
 245 250 255
 ggt acc ttt ttg cag cca cac cag ata gct cag ctt gag gtg atg act 816
 Gly Thr Phe Leu Gln Pro His Gln Ile Ala Gln Leu Glu Val Met Thr
 260 265 270
 tcc att gcg ctc cgt atc ctg cca acg atg tgc agt gtt aat gtg ccg 864
 Ser Ile Ala Leu Arg Ile Leu Pro Thr Met Cys Ser Val Asn Val Pro
 275 280 285
 ttg tac aga acc acc act agt gtg cca ttc ggc gtt ggc acc gga gtt 912
 Leu Tyr Arg Thr Thr Thr Ser Val Pro Phe Gly Val Gly Thr Gly Val
 290 295 300
 ggt gcc tac tga 924
 Gly Ala Tyr
 305
 <210> 174
 <211> 307
 <212> PRT
 <213> Triticum aestivum
 <400> 174
 Met Lys Thr Phe Leu Val Phe Ala Leu Leu Ala Val Ala Ala Thr Ser
 1 5 10 15
 Ala Ile Ala Gln Met Glu Thr Arg Cys Ile Pro Gly Leu Glu Arg Pro
 20 25 30
 Trp Gln Gln Gln Pro Leu Pro Pro Gln Gln Thr Phe Pro Gln Gln Pro
 35 40 45
 Leu Phe Ser Gln Gln Gln Gln Gln Gln Leu Phe Pro Gln Gln Pro Ser
 50 55 60
 Phe Ser Gln Gln Gln Pro Pro Phe Trp Gln Gln Gln Pro Pro Phe Ser
 65 70 75 80
 Gln Gln Gln Pro Ile Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln
 85 90 95
 Gln Leu Val Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln Gln Pro
 100 105 110
 Val Leu Pro Pro Gln Gln Ser Pro Phe Pro Gln Gln Gln Gln Gln His
 115 120 125
 Gln Gln Leu Val Gln Gln Gln Ile Pro Val Val Gln Pro Ser Ile Leu
 130 135 140
 Gln Gln Leu Asn Pro Cys Lys Val Phe Leu Gln Gln Gln Cys Ser Pro
 145 150 155 160
 Val Ala Met Pro Gln Arg Leu Ala Arg Ser Gln Met Leu Gln Gln Ser
 165 170 175
 Ser Cys His Val Met Gln Gln Gln Cys Cys Gln Gln Leu Pro Gln Ile
 180 185 190

127

Pro Gln Gln Ser Arg Tyr Glu Ala Ile Arg Ala Ile Ile Tyr Ser Ile
 195 200 205
 Ile Leu Gln Glu Gln Gln Gln Val Gln Gly Ser Ile Gln Ser Gln Gln
 210 215 220
 Gln Gln Pro Gln Gln Leu Gly Gln Cys Val Ser Gln Pro Gln Gln Gln
 225 230 235 240
 Ser Gln Gln Gln Leu Gly Gln Gln Pro Gln Gln Gln Gln Leu Ala Gln
 245 250 255
 Gly Thr Phe Leu Gln Pro His Gln Ile Ala Gln Leu Glu Val Met Thr
 260 265 270
 Ser Ile Ala Leu Arg Ile Leu Pro Thr Met Cys Ser Val Asn Val Pro
 275 280 285
 Leu Tyr Arg Thr Thr Thr Ser Val Pro Phe Gly Val Gly Thr Gly Val
 290 295 300
 Gly Ala Tyr
 305

<210> 175

<211> 8324

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: construct
 coding for artificial fusion protein of A.thaliana
 Accase and plastidic transitpeptide of tobacco
 transketolase

<220>

<221> CDS

<222> (1154)..(7999)

<223> coding for artificial fusion protein of A.thaliana
 Accase and plastidic transitpeptide of tobacco
 transketolase

<400> 175

aagcttttctt catcggtgat tgattccttt aaagacttat gtttcttattc ttgcttctga 60
 ggcaagtatt cagttaccag ttaccactta tattctggac tttctgactg catcctcatt 120
 tttccaacat ttttaatttc actattggct gaatgcttct tctttgagga agaaacaatt 180
 cagatggcag aaatgtatca accaatgcat atatacaaat gtacctcttg ttctcaaaac 240
 atctatcgga tggttccatt tgctttgtca tccaattagt gactacttta tattattcac 300
 tcctctttat tactattttc atgcgagggt gccatgtaca ttatatttgt aaggattgac 360
 gctattgagc gtttttcttc aattttcttt atttttagaca tgggtatgaa atgtgtgtta 420
 gagttgggtt gaatgagata tacgttcaag tgaagtggca taccgttctc gagtaaggat 480
 gacctaccca ttcttgagac aaatgttaca ttttagtatc agagtaaaat gtgtacctat 540
 aactcaaatt cgattgacat gtatccattc aacataaaat taaaccagcc tgcacctgca 600
 tccacatttc aagtattttc aaaccgttcg gctcctatcc accgggtgta acaagacgga 660
 ttccgaattt ggaagatttt gactcaaatt cccaatttat attgaccgtg actaaatcaa 720
 ctttaacttc tataattctg attaagctcc caatttatat tccaacggc actacctcca 780

128

aaatttatag actctcatcc ccttttaaac caacttagta aacgtttttt tttttaattt 840
 tatgaagtta agttttttacc ttgttttttaa aaagaatcgt tcataagatg ccatgccaga 900
 acattagcta cacgttacac atagcatgca gccgcggaga attgtttttc ttcgccactt 960
 gtcactccct tcaaacacct aagagcttct ctctcacagc acacacatac aatcacatgc 1020
 gtgcatgcat tattacacgt gatcgccatg caaatctcct ttatagccta taaattaact 1080
 catccgcttc actcttttact caaaccaaaa ctcatcaata caaacaagat taaaaacata 1140
 cacgagtggg acc atg gcg tct tct tct tct ctc act ctc tct caa gct 1189
 Met Ala Ser Ser Ser Ser Leu Thr Leu Ser Gln Ala
 1 5 10
 atc ctc tct cgt tct gtc cct cgc cat ggc tct gcc tct tct tct caa 1237
 Ile Leu Ser Arg Ser Val Pro Arg His Gly Ser Ala Ser Ser Ser Gln
 15 20 25
 ctt tcc cct tct tct ctc act ttt tcc ggc ctt aaa tcc aat ccc aat 1285
 Leu Ser Pro Ser Ser Leu Thr Phe Ser Gly Leu Lys Ser Asn Pro Asn
 30 35 40
 atc acc acc tcc cgc cgc cgt act cct tcc tcc gcc gcc gcc gcc gcc 1333
 Ile Thr Thr Ser Arg Arg Arg Thr Pro Ser Ser Ala Ala Ala Ala Ala
 45 50 55 60
 gtc gta agg tca ccg gcg att cgt gcc tca gct gca acc gaa acc ata 1381
 Val Val Arg Ser Pro Ala Ile Arg Ala Ser Ala Ala Thr Glu Thr Ile
 65 70 75
 gag aaa act gag act gcg tgg atc ccg ggt cga caa act agt gct ggc 1429
 Glu Lys Thr Glu Thr Ala Trp Ile Pro Gly Arg Gln Thr Ser Ala Gly
 80 85 90
 tcg gtt aac ggg aat cat agt gct gta gga cct ggt ata aat tat gag 1477
 Ser Val Asn Gly Asn His Ser Ala Val Gly Pro Gly Ile Asn Tyr Glu
 95 100 105
 acg gtg tct caa gtg gat gag ttc tgt aaa gca ctt aga ggg aaa agg 1525
 Thr Val Ser Gln Val Asp Glu Phe Cys Lys Ala Leu Arg Gly Lys Arg
 110 115 120
 ccg atc cat agt att ttg ata gct aac aat gga atg gcg gct gtg aag 1573
 Pro Ile His Ser Ile Leu Ile Ala Asn Asn Gly Met Ala Ala Val Lys
 125 130 135 140
 ttt ata cgt agt gtc aga aca tgg gct tat gaa aca ttt ngg acg gaa 1621
 Phe Ile Arg Ser Val Arg Thr Trp Ala Tyr Glu Thr Phe Xaa Thr Glu
 145 150 155
 aaa gcc ata ttg ttg gtg ggg atg gca acc cct gaa gac atg cgg atc 1669
 Lys Ala Ile Leu Leu Val Gly Met Ala Thr Pro Glu Asp Met Arg Ile
 160 165 170
 aat gcg gag cat atc aga atc gct gat cag ttt gtt gag gtt ccc gga 1717
 Asn Ala Glu His Ile Arg Ile Ala Asp Gln Phe Val Glu Val Pro Gly
 175 180 185
 gga acc aac aat aac aat tat gct aac gtt cag ctg att gtg gag atg 1765
 Gly Thr Asn Asn Asn Asn Tyr Ala Asn Val Gln Leu Ile Val Glu Met
 190 195 200
 gct gaa gta aca cgc gtg gat gca gtt tgg cct ggt tgg ggt can gca 1813
 Ala Glu Val Thr Arg Val Asp Ala Val Trp Pro Gly Trp Gly Thr Ala
 205 210 215 220
 tct gaa aac ccc gaa tta cct gat gcc cta gat gca aaa gga atc ata 1861
 Ser Glu Asn Pro Glu Leu Pro Asp Ala Leu Asp Ala Lys Gly Ile Ile
 225 230 235

129

ttt ctt ggt cct cca gca tct tca atg gca gca ctg gga gat aag att	1909
Phe Leu Gly Pro Pro Ala Ser Ser Met Ala Ala Leu Gly Asp Lys Ile	
240 245 250	
ggt tct tcg ttg att gca caa gct gct gat gta ccc act ctg cca tgg	1957
Gly Ser Ser Leu Ile Ala Gln Ala Ala Asp Val Pro Thr Leu Pro Trp	
255 260 265	
agt ggt tcc cat gtt aaa ata cct cct aat agc aac ttg gta acc atc	2005
Ser Gly Ser His Val Lys Ile Pro Pro Asn Ser Asn Leu Val Thr Ile	
270 275 280	
cca gag gag atc tac cgg caa gca tgt gtc tac aca act gaa gaa gcg	2053
Pro Glu Glu Ile Tyr Arg Gln Ala Cys Val Tyr Thr Thr Glu Glu Ala	
285 290 295 300	
att gct agc tgt caa gtt gtc ggt tac cca gca atg atc aaa gca tcg	2101
Ile Ala Ser Cys Gln Val Val Gly Tyr Pro Ala Met Ile Lys Ala Ser	
305 310 315	
tgg ggt ggt ggt ggt aaa gga atc agg aag gtt cat aat gat gat gag	2149
Trp Gly Gly Gly Gly Lys Gly Ile Arg Lys Val His Asn Asp Asp Glu	
320 325 330	
gtt agg gct cta ttc aag caa gtt cag ggt gag gtc cca ggc tca cca	2197
Val Arg Ala Leu Phe Lys Gln Val Gln Gly Glu Val Pro Gly Ser Pro	
335 340 345	
ata ttc ata atg aag gtt gcg tca cag agt cgg cat cta gag gtc cag	2245
Ile Phe Ile Met Lys Val Ala Ser Gln Ser Arg His Leu Glu Val Gln	
350 355 360	
ctg ctc tgt gac aag cat gga aat gtt tca gct ctg cat agc cgt gat	2293
Leu Leu Cys Asp Lys His Gly Asn Val Ser Ala Leu His Ser Arg Asp	
365 370 375 380	
tgt agc gtc cag aga aga cat caa aag atc ata gag gag ggt cca att	2341
Cys Ser Val Gln Arg Arg His Gln Lys Ile Ile Glu Glu Gly Pro Ile	
385 390 395	
act gtg gct ccg cca gaa gct gtc aag aaa ctt gaa caa gca gct aga	2389
Thr Val Ala Pro Pro Glu Ala Val Lys Lys Leu Glu Gln Ala Ala Arg	
400 405 410	
agg ttg gct aag agt gtt aac tat gtt gga gct gct act gtt gag tat	2437
Arg Leu Ala Lys Ser Val Asn Tyr Val Gly Ala Ala Thr Val Glu Tyr	
415 420 425	
ctc tac agt atg gac act ggg gag tac tac ttc tta gag ctt aac cct	2485
Leu Tyr Ser Met Asp Thr Gly Glu Tyr Tyr Phe Leu Glu Leu Asn Pro	
430 435 440	
cgc tta cag gtt gag cat cct gtc act gag tgg att gcc gag ata aat	2533
Arg Leu Gln Val Glu His Pro Val Thr Glu Trp Ile Ala Glu Ile Asn	
445 450 455 460	
ctt cct gct gcc caa gtt gct gtg ggg atg gga att cct ctc tgg caa	2581
Leu Pro Ala Ala Gln Val Ala Val Gly Met Gly Ile Pro Leu Trp Gln	
465 470 475	
atc cct gag ata aga cgg ttc tat gga ata gaa cat ggt gga ggt tat	2629
Ile Pro Glu Ile Arg Arg Phe Tyr Gly Ile Glu His Gly Gly Gly Tyr	
480 485 490	
gat tct tgg cga aaa aca tct gtt gta gcc ttc cct ttt gat ttt gat	2677
Asp Ser Trp Arg Lys Thr Ser Val Ala Phe Pro Phe Asp Phe Asp	
495 500 505	

130

aaa gct caa tct ata agg cca aaa ggt cat tgt gtg gct gta cgt gtg	2725
Lys Ala Gln Ser Ile Arg Pro Lys Gly His Cys Val Ala Val Arg Val	
510 515 520	
aca agt gag gat cct gat gac ggg ttc aaa cca acc agc ggt aga gtt	2773
Thr Ser Glu Asp Pro Asp Asp Gly Phe Lys Pro Thr Ser Gly Arg Val	
525 530 535 540	
cag gag ttg agt ttt aag agc aag cca aat gtg tgg gcg tac ttc tct	2821
Gln Glu Leu Ser Phe Lys Ser Lys Pro Asn Val Trp Ala Tyr Phe Ser	
545 550 555	
gtc aag tct ggt gga ggc atc cac gag ttc tcg gat tcc cag ttt gga	2869
Val Lys Ser Gly Gly Gly Ile His Glu Phe Ser Asp Ser Gln Phe Gly	
560 565 570	
cat gtt ttt gca ttt ggg gaa tcc aga gcc ctg gcg ata gcg aat atg	2917
His Val Phe Ala Phe Gly Glu Ser Arg Ala Leu Ala Ile Ala Asn Met	
575 580 585	
gtt ctt ggg cta aaa gaa att cag atc cgt gga gaa att agg act aac	2965
Val Leu Gly Leu Lys Glu Ile Gln Ile Arg Gly Glu Ile Arg Thr Asn	
590 595 600	
gtt gac tac acg atc gac ctt tta cat gct tct gat tac cgt gat aac	3013
Val Asp Tyr Thr Ile Asp Leu Leu His Ala Ser Asp Tyr Arg Asp Asn	
605 610 615 620	
aaa att cac act ggt tgg ttg gat agt agg att gct atg cgg gtc aga	3061
Lys Ile His Thr Gly Trp Leu Asp Ser Arg Ile Ala Met Arg Val Arg	
625 630 635	
gct gag agg cct cca tgg tat ctc tct gtt gtc ggc gga gct ctc tat	3109
Ala Glu Arg Pro Pro Trp Tyr Leu Ser Val Val Gly Gly Ala Leu Tyr	
640 645 650	
aaa gca tca gcg acc agt gct gct gtg gtt tca gat tac gtt ggt tat	3157
Lys Ala Ser Ala Thr Ser Ala Ala Val Val Ser Asp Tyr Val Gly Tyr	
655 660 665	
ctg gag aag ggg caa atc cct cca aag cat ata tct ctt gta cat tct	3205
Leu Glu Lys Gly Gln Ile Pro Pro Lys His Ile Ser Leu Val His Ser	
670 675 680	
caa gtg tct ctg aat att gaa gga agt aaa tat acg att gat gta gtc	3253
Gln Val Ser Leu Asn Ile Glu Gly Ser Lys Tyr Thr Ile Asp Val Val	
685 690 695 700	
cgg ggt gga tca gga acc tac agg cta aga atg aac aag tca gaa gtg	3301
Arg Gly Gly Ser Gly Thr Tyr Arg Leu Arg Met Asn Lys Ser Glu Val	
705 710 715	
gta gca gaa ata cac act cta cgt gat gga ggt ctg ttg atg cag ttg	3349
Val Ala Glu Ile His Thr Leu Arg Asp Gly Gly Leu Leu Met Gln Leu	
720 725 730	
gat ggc aaa agc cat gtg ata tat gca gag gaa gaa gct gca gga act	3397
Asp Gly Lys Ser His Val Ile Tyr Ala Glu Glu Glu Ala Ala Gly Thr	
735 740 745	
cgt ctt ctc att gat gga aga act tgt ttg cta cag aat gac cac gat	3445
Arg Leu Leu Ile Asp Gly Arg Thr Cys Leu Leu Gln Asn Asp His Asp	
750 755 760	
cca tca aag tta atg gct gag aca ccg tgc aag ttg atg agg tat ttg	3493
Pro Ser Lys Leu Met Ala Glu Thr Pro Cys Lys Leu Met Arg Tyr Leu	
765 770 775 780	

131

att tcc gac aac agc aat att gac gct gat acg cct tat gcc gaa gtt	3541
Ile Ser Asp Asn Ser Asn Ile Asp Ala Asp Thr Pro Tyr Ala Glu Val	
785 790 795	
gag gtc atg aag atg tgc atg cca ctt ctt tca cct gct tca gga gtt	3589
Glu Val Met Lys Met Cys Met Pro Leu Leu Ser Pro Ala Ser Gly Val	
800 805 810	
atc cat ttt aaa atg tct gaa gga caa gcc atg cag gct ggt gaa ctt	3637
Ile His Phe Lys Met Ser Glu Gly Gln Ala Met Gln Ala Gly Glu Leu	
815 820 825	
ata gcc aat ctt gat ctt gat gat cct tct gct gta aga aag gcc gaa	3685
Ile Ala Asn Leu Asp Leu Asp Asp Pro Ser Ala Val Arg Lys Ala Glu	
830 835 840	
ccc ttc cat gga agt ttc cca aga tta ggg ctt cca act gca ata tcc	3733
Pro Phe His Gly Ser Phe Pro Arg Leu Gly Leu Pro Thr Ala Ile Ser	
845 850 855 860	
ggt aga gtt cat cag aga tgt gcc gca aca tta aat gct gca cgc atg	3781
Gly Arg Val His Gln Arg Cys Ala Ala Thr Leu Asn Ala Ala Arg Met	
865 870 875	
att ctt gct ggc tat gag cat aaa gta gat gag gtt gtt caa gac tta	3829
Ile Leu Ala Gly Tyr Glu His Lys Val Asp Glu Val Val Gln Asp Leu	
880 885 890	
ctt aat tgc ctt gat agc cct gaa ctc cca ttt ctt cag tgg caa gag	3877
Leu Asn Cys Leu Asp Ser Pro Glu Leu Pro Phe Leu Gln Trp Gln Glu	
895 900 905	
tgc ttt gca gtt ctg gcg aca cga cta cct aaa aat ctc agg aac atg	3925
Cys Phe Ala Val Leu Ala Thr Arg Leu Pro Lys Asn Leu Arg Asn Met	
910 915 920	
cta gaa tca aag tat agg gaa ttt gag agt att tcc aga aac tct ttg	3973
Leu Glu Ser Lys Tyr Arg Glu Phe Glu Ser Ile Ser Arg Asn Ser Leu	
925 930 935 940	
acc acc gat ttc cct gcc aaa ctt tta aaa ggc att ctt gag gca cat	4021
Thr Thr Asp Phe Pro Ala Lys Leu Leu Lys Gly Ile Leu Glu Ala His	
945 950 955	
tta tct tct tgt gat gag aaa gag aga ggt gcc ctt gaa agg ctc att	4069
Leu Ser Ser Cys Asp Glu Lys Glu Arg Gly Ala Leu Glu Arg Leu Ile	
960 965 970	
gaa cca ttg atg agc ctt gca aaa tct tat gaa ggt ggt aga gaa agt	4117
Glu Pro Leu Met Ser Leu Ala Lys Ser Tyr Glu Gly Gly Arg Glu Ser	
975 980 985	
cat gcc cgt gtt att gtt cat tct ctc ttt gaa gaa tat cta tca gta	4165
His Ala Arg Val Ile Val His Ser Leu Phe Glu Glu Tyr Leu Ser Val	
990 995 1000	
gaa gaa tta ttc aat gat aac atg ctg gct gat gtt ata gaa cgc atg	4213
Glu Glu Leu Phe Asn Asp Asn Met Leu Ala Asp Val Ile Glu Arg Met	
1005 1010 1015 1020	
cgt cag cta tac aag aaa gat ctg ttg aaa att gtg gat ata gtg ctc	4261
Arg Gln Leu Tyr Lys Lys Asp Leu Leu Lys Ile Val Asp Ile Val Leu	
1025 1030 1035	
tca cac cag ggc ata aaa aac aaa aac aaa ctc gtt ctc cgg ctc atg	4309
Ser His Gln Gly Ile Lys Asn Lys Asn Lys Leu Val Leu Arg Leu Met	
1040 1045 1050	

132

gag cag ctt gtt tac cct aat cct gct gct tac aga gat aaa ctt att 4357
 Glu Gln Leu Val Tyr Pro Asn Pro Ala Ala Tyr Arg Asp Lys Leu Ile
 1055 1060 1065

cga ttc tca aca ctt aac cat act aac tac tct gag ttg gcg ctc aag 4405
 Arg Phe Ser Thr Leu Asn His Thr Asn Tyr Ser Glu Leu Ala Leu Lys
 1070 1075 1080

gcg agt caa tta ctt gaa cag acc aaa cta agt gag ctt cgt tca aac 4453
 Ala Ser Gln Leu Leu Glu Gln Thr Lys Leu Ser Glu Leu Arg Ser Asn
 1085 1090 1095 1100

att gct aga agc ctt tca gag tta gaa atg ttt aca gag gac gga gaa 4501
 Ile Ala Arg Ser Leu Ser Glu Leu Glu Met Phe Thr Glu Asp Gly Glu
 1105 1110 1115

aat atg gat act ccc aag agg aaa agt gcc att aat gaa aga ata gaa 4549
 Asn Met Asp Thr Pro Lys Arg Lys Ser Ala Ile Asn Glu Arg Ile Glu
 1120 1125 1130

gat ctt gta agc gca tct tta gct gtt gaa gac gct ctc gtg gga cta 4597
 Asp Leu Val Ser Ala Ser Leu Ala Val Glu Asp Ala Leu Val Gly Leu
 1135 1140 1145

ttt gac cat agc gat cac aca ctt caa aga cgg gtt gtt gag act tat 4645
 Phe Asp His Ser Asp His Thr Leu Gln Arg Arg Val Val Glu Thr Tyr
 1150 1155 1160

att cgc aga tta tac cag ccc tac gtc gtt aaa gat agc gtg agg atg 4693
 Ile Arg Arg Leu Tyr Gln Pro Tyr Val Val Lys Asp Ser Val Arg Met
 1165 1170 1175 1180

cag tgg cac cgt tct ggt ctt ctt gct tcc tgg gag ttc cta gag gag 4741
 Gln Trp His Arg Ser Gly Leu Leu Ala Ser Trp Glu Phe Leu Glu Glu
 1185 1190 1195

cat atg gaa aga aaa aac att ggc tta gac gat ccc gac aca cct gaa 4789
 His Met Glu Arg Lys Asn Ile Gly Leu Asp Asp Pro Asp Thr Pro Glu
 1200 1205 1210

aaa gga ttg gtt gag aag cgt agt aag aga aaa tgg ggg gct atg gtt 4837
 Lys Gly Leu Val Glu Lys Arg Ser Lys Arg Lys Trp Gly Ala Met Val
 1215 1220 1225

ata atc aaa tct ttg cag ttt ctt cca agt ata ata agt gca gca ttg 4885
 Ile Ile Lys Ser Leu Gln Phe Leu Pro Ser Ile Ile Ser Ala Ala Leu
 1230 1235 1240

aga gaa aca aag cac aac gac tat gaa act gcc gga gct cct tta tct 4933
 Arg Glu Thr Lys His Asn Asp Tyr Glu Thr Ala Gly Ala Pro Leu Ser
 1245 1250 1255 1260

ggc aat atg atg cac att gct att gtg ggc atc aac aac cag atg agt 4981
 Gly Asn Met Met His Ile Ala Ile Val Gly Ile Asn Asn Gln Met Ser
 1265 1270 1275

ctg ctt cag gac agt ggg gat gaa gac caa gct cag gaa aga gta aac 5029
 Leu Leu Gln Asp Ser Gly Asp Glu Asp Gln Ala Gln Glu Arg Val Asn
 1280 1285 1290

aag ttg gcc aaa att ctt aaa gag gaa gaa gtg agt tca agc ctc tgt 5077
 Lys Leu Ala Lys Ile Leu Lys Glu Glu Glu Val Ser Ser Ser Leu Cys
 1295 1300 1305

tct gcc ggt gtt ggt gta atc agc tgt ata att cag cga gat gaa gga 5125
 Ser Ala Gly Val Gly Val Ile Ser Cys Ile Ile Gln Arg Asp Glu Gly
 1310 1315 1320

cga	aca	ccc	atg	aga	cat	tct	ttc	cat	tgg	tcg	ttg	gag	aaa	cag	tat	5173
Arg	Thr	Pro	Met	Arg	His	Ser	Phe	His	Trp	Ser	Leu	Glu	Lys	Gln	Tyr	
1325					1330					1335					1340	
tat	gta	gaa	gag	ccg	ttg	ctg	cgt	cat	ctt	gaa	cct	cct	ctg	tcc	att	5221
Tyr	Val	Glu	Glu	Pro	Leu	Leu	Arg	His	Leu	Glu	Pro	Pro	Leu	Ser	Ile	
				1345					1350						1355	
tac	ctt	gag	ttg	gat	aag	ctg	aaa	gga	tac	tca	aat	ata	caa	tat	acg	5269
Tyr	Leu	Glu	Leu	Asp	Lys	Leu	Lys	Gly	Tyr	Ser	Asn	Ile	Gln	Tyr	Thr	
				1360				1365						1370		
cct	tct	cga	gat	cgt	caa	tgg	cat	ctg	tat	act	ggt	aca	gac	aag	cca	5317
Pro	Ser	Arg	Asp	Arg	Gln	Trp	His	Leu	Tyr	Thr	Val	Thr	Asp	Lys	Pro	
		1375					1380					1385				
gtg	cca	atc	aag	agg	atg	ttc	ctg	aga	tct	ctt	ggt	cga	cag	gct	aca	5365
Val	Pro	Ile	Lys	Arg	Met	Phe	Leu	Arg	Ser	Leu	Val	Arg	Gln	Ala	Thr	
	1390					1395					1400					
atg	aac	gat	gga	ttt	ata	ttg	cag	caa	ggg	cag	gat	aag	cag	ctt	agc	5413
Met	Asn	Asp	Gly	Phe	Ile	Leu	Gln	Gln	Gly	Gln	Asp	Lys	Gln	Leu	Ser	
1405				1410					1415						1420	
caa	aca	ctg	atc	tcc	atg	gcg	ttt	acg	tcg	aaa	tgt	ggt	ctg	agg	tct	5461
Gln	Thr	Leu	Ile	Ser	Met	Ala	Phe	Thr	Ser	Lys	Cys	Val	Leu	Arg	Ser	
				1425				1430						1435		
ttg	atg	gat	gcc	atg	gag	gaa	ctg	gaa	ctg	aat	gcc	cat	aat	gct	gca	5509
Leu	Met	Asp	Ala	Met	Glu	Glu	Leu	Glu	Leu	Asn	Ala	His	Asn	Ala	Ala	
			1440				1445					1450				
atg	aaa	cca	gat	cac	gca	cat	atg	ttt	ctt	tgc	ata	ttg	cgt	gag	cag	5557
Met	Lys	Pro	Asp	His	Ala	His	Met	Phe	Leu	Cys	Ile	Leu	Arg	Glu	Gln	
	1455					1460						1465				
cag	ata	gat	gat	ctt	gtg	cct	ttc	ccc	agg	aga	ggt	gaa	gtg	aat	gcg	5605
Gln	Ile	Asp	Asp	Leu	Val	Pro	Phe	Pro	Arg	Arg	Val	Glu	Val	Asn	Ala	
	1470				1475					1480						
gag	gat	gaa	gaa	act	aca	ggt	gaa	atg	atc	tta	gaa	gaa	gca	gca	cga	5653
Glu	Asp	Glu	Glu	Thr	Thr	Val	Glu	Met	Ile	Leu	Glu	Glu	Ala	Ala	Arg	
1485				1490					1495						1500	
gag	ata	cat	aga	tct	ggt	gga	gtg	aga	atg	cat	agg	ttg	ggc	gtg	tgc	5701
Glu	Ile	His	Arg	Ser	Val	Gly	Val	Arg	Met	His	Arg	Leu	Gly	Val	Cys	
				1505				1510						1515		
gag	tgg	gaa	gtg	cgg	ctg	tgg	ttg	gtg	tcc	tct	gga	ctg	gca	tgt	ggt	5749
Glu	Trp	Glu	Val	Arg	Leu	Trp	Leu	Val	Ser</							

134

agg agt aac act act tat tgc tat gac ttc ccg ttg gca ttt ggg aca	5989
Arg Ser Asn Thr Thr Tyr Cys Tyr Asp Phe Pro Leu Ala Phe Gly Thr	
1600 1605 1610	
gcc ttg gaa ctg ttg tgg gca tca caa cac cca gga gtt aag aaa cca	6037
Ala Leu Glu Leu Leu Trp Ala Ser Gln His Pro Gly Val Lys Lys Pro	
1615 1620 1625	
tat aag gat act ctg atc aat gtt aaa gag ctt gta ttc tca aaa cca	6085
Tyr Lys Asp Thr Leu Ile Asn Val Lys Glu Leu Val Phe Ser Lys Pro	
1630 1635 1640	
gaa ggt tct tcg ggt aca tct cta gat ctg gtt gaa aga cca ccc ggt	6133
Glu Gly Ser Ser Gly Thr Ser Leu Asp Leu Val Glu Arg Pro Pro Gly	
1645 1650 1655 1660	
ctc aac gac ttt gga atg gtt gcc tgg tgc cta gat atg tcg acc cca	6181
Leu Asn Asp Phe Gly Met Val Ala Trp Cys Leu Asp Met Ser Thr Pro	
1665 1670 1675	
gag ttt cct atg ggg cgg aaa ctt ctc gtg att gcg aat gat gtc acc	6229
Glu Phe Pro Met Gly Arg Lys Leu Leu Val Ile Ala Asn Asp Val Thr	
1680 1685 1690	
ttc aaa gct ggt tct ttt ggt cct aga gag gac gcg ttt ttc ctt gct	6277
Phe Lys Ala Gly Ser Phe Gly Pro Arg Glu Asp Ala Phe Phe Leu Ala	
1695 1700 1705	
gtt act gaa ctc gct tgt gcc aag aag ctt ccc ttg att tac ttg gca	6325
Val Thr Glu Leu Ala Cys Ala Lys Lys Leu Pro Leu Ile Tyr Leu Ala	
1710 1715 1720	
gca aat tct ggt gcc cga ctt ggg gtt gct gaa gaa gtc aaa gcc tgc	6373
Ala Asn Ser Gly Ala Arg Leu Gly Val Ala Glu Glu Val Lys Ala Cys	
1725 1730 1735 1740	
ttc aaa gtt gga tgg tcg gat gaa att tcc cct gag aat ggt ttt cag	6421
Phe Lys Val Gly Trp Ser Asp Glu Ile Ser Pro Glu Asn Gly Phe Gln	
1745 1750 1755	
tat ata tac cta agc cct gaa gac cac gaa agg att gga tca tct gtc	6469
Tyr Ile Tyr Leu Ser Pro Glu Asp His Glu Arg Ile Gly Ser Ser Val	
1760 1765 1770	
att gcc cat gaa gta aag ctc tct agt ggg gaa act agg tgg gtg att	6517
Ile Ala His Glu Val Lys Leu Ser Ser Gly Glu Thr Arg Trp Val Ile	
1775 1780 1785	
gat acg atc gtt ggc aaa gaa gat ggt att ggt gta gag aac tta aca	6565
Asp Thr Ile Val Gly Lys Glu Asp Gly Ile Gly Val Glu Asn Leu Thr	
1790 1795 1800	
gga agt ggg gcc ata gcg ggt gct tac tca aag gca tac aat gaa act	6613
Gly Ser Gly Ala Ile Ala Gly Ala Tyr Ser Lys Ala Tyr Asn Glu Thr	
1805 1810 1815 1820	
ttt act tta acc ttt gtt agt gga aga acg gtt gga att ggt gct tat	6661
Phe Thr Leu Thr Phe Val Ser Gly Arg Thr Val Gly Ile Gly Ala Tyr	
1825 1830 1835	
ctt gcc cgc cta ggt atg cgg tgc ata cag aga ctt gat cag ccg atc	6709
Leu Ala Arg Leu Gly Met Arg Cys Ile Gln Arg Leu Asp Gln Pro Ile	
1840 1845 1850	
atc ttg act ggc ttc tct aca ctc aac aag tta ctt ggg cgt gag gtc	6757
Ile Leu Thr Gly Phe Ser Thr Leu Asn Lys Leu Leu Gly Arg Glu Val	
1855 1860 1865	

135

tat agc tct cac atg caa ctg ggt ggc ccg aaa atc atg ggc aca aat	6805
Tyr Ser Ser His Met Gln Leu Gly Gly Pro Lys Ile Met Gly Thr Asn	
1870 1875 1880	
ggt gtt gtt cat ctt aca gtc tca gat gat ctt gaa ggc gta tca gca	6853
Gly Val Val His Leu Thr Val Ser Asp Asp Leu Glu Gly Val Ser Ala	
1885 1890 1895 1900	
att ctc aac tgg ctc agc tac att cct gct tac gtg ggt ggt cct ctt	6901
Ile Leu Asn Trp Leu Ser Tyr Ile Pro Ala Tyr Val Gly Gly Pro Leu	
1905 1910 1915	
cct gtt ctt gcc cct tta gat cca ccg gag aga att gtg gag tat gtc	6949
Pro Val Leu Ala Pro Leu Asp Pro Pro Glu Arg Ile Val Glu Tyr Val	
1920 1925 1930	
cca gag aac tct tgc gac cca cga gcg gct ata gct ggg gtc aaa gac	6997
Pro Glu Asn Ser Cys Asp Pro Arg Ala Ala Ile Ala Gly Val Lys Asp	
1935 1940 1945	
aat acc ggt aaa tgg ctt gga ggt atc ttt gat aaa aat agt ttc att	7045
Asn Thr Gly Lys Trp Leu Gly Gly Ile Phe Asp Lys Asn Ser Phe Ile	
1950 1955 1960	
gag act ctt gaa ggc tgg gca agg acg gta gtg act ggt aga gcc aag	7093
Glu Thr Leu Glu Gly Trp Ala Arg Thr Val Val Thr Gly Arg Ala Lys	
1965 1970 1975 1980	
ctc ggg gga ata ccc gtt gga gtt gtt gca gtt gag aca cag act gtc	7141
Leu Gly Gly Ile Pro Val Gly Val Val Ala Val Glu Thr Gln Thr Val	
1985 1990 1995	
atg cag atc atc cca gcc gat cct gga cag ctt gac tct cat gaa aga	7189
Met Gln Ile Ile Pro Ala Asp Pro Gly Gln Leu Asp Ser His Glu Arg	
2000 2005 2010	
gtg gtt ccg caa gca ggg caa gtc tgg ttt cct gat tca gcg gcc aag	7237
Val Val Pro Gln Ala Gly Gln Val Trp Phe Pro Asp Ser Ala Ala Lys	
2015 2020 2025	
act gct caa gcg ctt atg gat ttc aac cgg gaa gag ctt cca ttg ttt	7285
Thr Ala Gln Ala Leu Met Asp Phe Asn Arg Glu Glu Leu Pro Leu Phe	
2030 2035 2040	
atc cta gcg aac tgg aga ggg ttt tca ggt ggg cag aga gat ctt ttc	7333
Ile Leu Ala Asn Trp Arg Gly Phe Ser Gly Gly Gln Arg Asp Leu Phe	
2045 2050 2055 2060	
gaa gga ata ctt cag gca ggt tca act ata gta gaa aat ctg aga acc	7381
Glu Gly Ile Leu Gln Ala Gly Ser Thr Ile Val Glu Asn Leu Arg Thr	
2065 2070 2075	
tat cgt cag cca gtg ttt gtg tac atc cca atg atg gga gag ctg cgc	7429
Tyr Arg Gln Pro Val Phe Val Tyr Ile Pro Met Met Gly Glu Leu Arg	
2080 2085 2090	
ggt gga gcg tgg gtt gtt gtt gac agc cag ata aat tcg gat tat gtt	7477
Gly Gly Ala Trp Val Val Val Asp Ser Gln Ile Asn Ser Asp Tyr Val	
2095 2100 2105	
gaa atg tat gct gat gaa aca gct cgt gga aat gtg ctt gag cca gaa	7525
Glu Met Tyr Ala Asp Glu Thr Ala Arg Gly Asn Val Leu Glu Pro Glu	
2110 2115 2120	
ggg aca ata gag ata aaa ttt aga aca aaa gag cta tta gag tgc atg	7573
Gly Thr Ile Glu Ile Lys Phe Arg Thr Lys Glu Leu Leu Glu Cys Met	
2125 2130 2135 2140	

136

gga agg ttg gac cag aag cta atc agt ctg aaa gca aaa ctg caa gat 7621
 Gly Arg Leu Asp Gln Lys Leu Ile Ser Leu Lys Ala Lys Leu Gln Asp
 2145 2150 2155

gcc aag caa agc gag gcc tat gca aac atc gag ctt ctc cag caa cag 7669
 Ala Lys Gln Ser Glu Ala Tyr Ala Asn Ile Glu Leu Leu Gln Gln Gln
 2160 2165 2170

att aaa gcc cga gag aaa cag ctt tta cca gtt tat atc caa atc gcc 7717
 Ile Lys Ala Arg Glu Lys Gln Leu Leu Pro Val Tyr Ile Gln Ile Ala
 2175 2180 2185

acc aaa ttt gca gaa ctt cat gac act tcc atg aga atg gct gca aag 7765
 Thr Lys Phe Ala Glu Leu His Asp Thr Ser Met Arg Met Ala Ala Lys
 2190 2195 2200

gga gtg atc aaa agt gtt gtg gaa tgg agc ggc tcg cgg tcc ttc ttc 7813
 Gly Val Ile Lys Ser Val Val Glu Trp Ser Gly Ser Arg Ser Phe Phe
 2205 2210 2215 2220

tac aaa aag ctc aat agg aga atc gct gag agc tct ctt gtg aaa aac 7861
 Tyr Lys Lys Leu Asn Arg Arg Ile Ala Glu Ser Ser Leu Val Lys Asn
 2225 2230 2235

gta aga gaa gca tct gga gac aac tta gca tat aaa tct tca atg cgt 7909
 Val Arg Glu Ala Ser Gly Asp Asn Leu Ala Tyr Lys Ser Ser Met Arg
 2240 2245 2250

ctg att cag gat tgg ttc tgc aac tct gat att gca aag ggg aaa gaa 7957
 Leu Ile Gln Asp Trp Phe Cys Asn Ser Asp Ile Ala Lys Gly Lys Glu
 2255 2260 2265

gaa gct tgg aca gac gac caa gtg ttc ttt aca tgg aag gac 7999
 Glu Ala Trp Thr Asp Asp Gln Val Phe Phe Thr Trp Lys Asp
 2270 2275 2280

aatgtagta actacgagtt gaagctgagc gagttgagag cgcagaaact actgaaccaa 8059
 cttgcagaga ttgggaattc ctcagatttg caagctctgc cacaaggact tgctaattctt 8119
 ctaaacaagg tggagccgctc gaaaagagaa gagctggtgg ctgctaatacg aaagggtcttg 8179
 ggtagactga tatcgaagac tttagcttct aatccaagaa agatggacat ttaaagtttg 8239
 cttgtgtcca tttggaccat cttccttata tttgttggtc acagttgtaa atgttggtgt 8299
 agctttgtca tttcgtgctg gccgc 8324

<210> 176
 <211> 2282
 <212> PRT
 <213> Künstliche Sequenz
 <223> Beschreibung der künstlichen Sequenz: construct
 coding for artificial fusion protein of A.thaliana
 Accase and plastidic transitpeptide of tobacco
 transketolase

<400> 176
 Met Ala Ser Ser Ser Ser Leu Thr Leu Ser Gln Ala Ile Leu Ser Arg
 1 5 10 15
 Ser Val Pro Arg His Gly Ser Ala Ser Ser Ser Gln Leu Ser Pro Ser
 20 25 30
 Ser Leu Thr Phe Ser Gly Leu Lys Ser Asn Pro Asn Ile Thr Thr Ser
 35 40 45
 Arg Arg Arg Thr Pro Ser Ser Ala Ala Ala Ala Val Val Arg Ser
 50 55 60

137

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

138

Glu His Pro Val Thr Glu Trp Ile Ala Glu Ile Asn Leu Pro Ala Ala
 450 455 460
 Gln Val Ala Val Gly Met Gly Ile Pro Leu Trp Gln Ile Pro Glu Ile
 465 470 475 480
 Arg Arg Phe Tyr Gly Ile Glu His Gly Gly Gly Tyr Asp Ser Trp Arg
 485 490 495
 Lys Thr Ser Val Val Ala Phe Pro Phe Asp Phe Asp Lys Ala Gln Ser
 500 505 510
 Ile Arg Pro Lys Gly His Cys Val Ala Val Arg Val Thr Ser Glu Asp
 515 520 525
 Pro Asp Asp Gly Phe Lys Pro Thr Ser Gly Arg Val Gln Glu Leu Ser
 530 535 540
 Phe Lys Ser Lys Pro Asn Val Trp Ala Tyr Phe Ser Val Lys Ser Gly
 545 550 555 560
 Gly Gly Ile His Glu Phe Ser Asp Ser Gln Phe Gly His Val Phe Ala
 565 570 575
 Phe Gly Glu Ser Arg Ala Leu Ala Ile Ala Asn Met Val Leu Gly Leu
 580 585 590
 Lys Glu Ile Gln Ile Arg Gly Glu Ile Arg Thr Asn Val Asp Tyr Thr
 595 600 605
 Ile Asp Leu Leu His Ala Ser Asp Tyr Arg Asp Asn Lys Ile His Thr
 610 615 620
 Gly Trp Leu Asp Ser Arg Ile Ala Met Arg Val Arg Ala Glu Arg Pro
 625 630 635 640
 Pro Trp Tyr Leu Ser Val Val Gly Gly Ala Leu Tyr Lys Ala Ser Ala
 645 650 655
 Thr Ser Ala Ala Val Val Ser Asp Tyr Val Gly Tyr Leu Glu Lys Gly
 660 665 670
 Gln Ile Pro Pro Lys His Ile Ser Leu Val His Ser Gln Val Ser Leu
 675 680 685
 Asn Ile Glu Gly Ser Lys Tyr Thr Ile Asp Val Val Arg Gly Gly Ser
 690 695 700
 Gly Thr Tyr Arg Leu Arg Met Asn Lys Ser Glu Val Val Ala Glu Ile
 705 710 715 720
 His Thr Leu Arg Asp Gly Gly Leu Leu Met Gln Leu Asp Gly Lys Ser
 725 730 735
 His Val Ile Tyr Ala Glu Glu Glu Ala Ala Gly Thr Arg Leu Leu Ile
 740 745 750
 Asp Gly Arg Thr Cys Leu Leu Gln Asn Asp His Asp Pro Ser Lys Leu
 755 760 765
 Met Ala Glu Thr Pro Cys Lys Leu Met Arg Tyr Leu Ile Ser Asp Asn
 770 775 780
 Ser Asn Ile Asp Ala Asp Thr Pro Tyr Ala Glu Val Glu Val Met Lys
 785 790 795 800
 Met Cys Met Pro Leu Leu Ser Pro Ala Ser Gly Val Ile His Phe Lys
 805 810 815
 Met Ser Glu Gly Gln Ala Met Gln Ala Gly Glu Leu Ile Ala Asn Leu
 820 825 830

139

Asp Leu Asp Asp Pro Ser Ala Val Arg Lys Ala Glu Pro Phe His Gly
 835 840 845
 Ser Phe Pro Arg Leu Gly Leu Pro Thr Ala Ile Ser Gly Arg Val His
 850 855 860
 Gln Arg Cys Ala Ala Thr Leu Asn Ala Ala Arg Met Ile Leu Ala Gly
 865 870 875 880
 Tyr Glu His Lys Val Asp Glu Val Val Gln Asp Leu Leu Asn Cys Leu
 885 890 895
 Asp Ser Pro Glu Leu Pro Phe Leu Gln Trp Gln Glu Cys Phe Ala Val
 900 905 910
 Leu Ala Thr Arg Leu Pro Lys Asn Leu Arg Asn Met Leu Glu Ser Lys
 915 920 925
 Tyr Arg Glu Phe Glu Ser Ile Ser Arg Asn Ser Leu Thr Thr Asp Phe
 930 935 940
 Pro Ala Lys Leu Leu Lys Gly Ile Leu Glu Ala His Leu Ser Ser Cys
 945 950 955 960
 Asp Glu Lys Glu Arg Gly Ala Leu Glu Arg Leu Ile Glu Pro Leu Met
 965 970 975
 Ser Leu Ala Lys Ser Tyr Glu Gly Gly Arg Glu Ser His Ala Arg Val
 980 985 990
 Ile Val His Ser Leu Phe Glu Glu Tyr Leu Ser Val Glu Glu Leu Phe
 995 1000 1005
 Asn Asp Asn Met Leu Ala Asp Val Ile Glu Arg Met Arg Gln Leu Tyr
 1010 1015 1020
 Lys Lys Asp Leu Leu Lys Ile Val Asp Ile Val Leu Ser His Gln Gly
 1025 1030 1035 1040
 Ile Lys Asn Lys Asn Lys Leu Val Leu Arg Leu Met Glu Gln Leu Val
 1045 1050 1055
 Tyr Pro Asn Pro Ala Ala Tyr Arg Asp Lys Leu Ile Arg Phe Ser Thr
 1060 1065 1070
 Leu Asn His Thr Asn Tyr Ser Glu Leu Ala Leu Lys Ala Ser Gln Leu
 1075 1080 1085
 Leu Glu Gln Thr Lys Leu Ser Glu Leu Arg Ser Asn Ile Ala Arg Ser
 1090 1095 1100
 Leu Ser Glu Leu Glu Met Phe Thr Glu Asp Gly Glu Asn Met Asp Thr
 1105 1110 1115 1120
 Pro Lys Arg Lys Ser Ala Ile Asn Glu Arg Ile Glu Asp Leu Val Ser
 1125 1130 1135
 Ala Ser Leu Ala Val Glu Asp Ala Leu Val Gly Leu Phe Asp His Ser
 1140 1145 1150
 Asp His Thr Leu Gln Arg Arg Val Val Glu Thr Tyr Ile Arg Arg Leu
 1155 1160 1165
 Tyr Gln Pro Tyr Val Val Lys Asp Ser Val Arg Met Gln Trp His Arg
 1170 1175 1180
 Ser Gly Leu Leu Ala Ser Trp Glu Phe Leu Glu Glu His Met Glu Arg
 1185 1190 1195 1200
 Lys Asn Ile Gly Leu Asp Asp Pro Asp Thr Pro Glu Lys Gly Leu Val
 1205 1210 1215

140

Glu Lys Arg Ser Lys Arg Lys Trp Gly Ala Met Val Ile Ile Lys Ser
 1220 1225 1230
 Leu Gln Phe Leu Pro Ser Ile Ile Ser Ala Ala Leu Arg Glu Thr Lys
 1235 1240 1245
 His Asn Asp Tyr Glu Thr Ala Gly Ala Pro Leu Ser Gly Asn Met Met
 1250 1255 1260
 His Ile Ala Ile Val Gly Ile Asn Asn Gln Met Ser Leu Leu Gln Asp
 265 1270 1275 1280
 Ser Gly Asp Glu Asp Gln Ala Gln Glu Arg Val Asn Lys Leu Ala Lys
 1285 1290 1295
 Ile Leu Lys Glu Glu Glu Val Ser Ser Ser Leu Cys Ser Ala Gly Val
 1300 1305 1310
 Gly Val Ile Ser Cys Ile Ile Gln Arg Asp Glu Gly Arg Thr Pro Met
 1315 1320 1325
 Arg His Ser Phe His Trp Ser Leu Glu Lys Gln Tyr Tyr Val Glu Glu
 1330 1335 1340
 Pro Leu Leu Arg His Leu Glu Pro Pro Leu Ser Ile Tyr Leu Glu Leu
 345 1350 1355 1360
 Asp Lys Leu Lys Gly Tyr Ser Asn Ile Gln Tyr Thr Pro Ser Arg Asp
 1365 1370 1375
 Arg Gln Trp His Leu Tyr Thr Val Thr Asp Lys Pro Val Pro Ile Lys
 1380 1385 1390
 Arg Met Phe Leu Arg Ser Leu Val Arg Gln Ala Thr Met Asn Asp Gly
 1395 1400 1405
 Phe Ile Leu Gln Gln Gly Gln Asp Lys Gln Leu Ser Gln Thr Leu Ile
 1410 1415 1420
 Ser Met Ala Phe Thr Ser Lys Cys Val Leu Arg Ser Leu Met Asp Ala
 425 1430 1435 1440
 Met Glu Glu Leu Glu Leu Asn Ala His Asn Ala Ala Met Lys Pro Asp
 1445 1450 1455
 His Ala His Met Phe Leu Cys Ile Leu Arg Glu Gln Gln Ile Asp Asp
 1460 1465 1470
 Leu Val Pro Phe Pro Arg Arg Val Glu Val Asn Ala Glu Asp Glu Glu
 1475 1480 1485
 Thr Thr Val Glu Met Ile Leu Glu Glu Ala Ala Arg Glu Ile His Arg
 1490 1495 1500
 Ser Val Gly Val Arg Met His Arg Leu Gly Val Cys Glu Trp Glu Val
 505 1510 1515 1520
 Arg Leu Trp Leu Val Ser Ser Gly Leu Ala Cys Gly Ala Trp Arg Val
 1525 1530 1535
 Val Val Ala Asn Val Thr Gly Arg Thr Cys Thr Val His Ile Tyr Arg
 1540 1545 1550
 Glu Val Glu Thr Pro Gly Arg Asn Ser Leu Phe Tyr His Ser Ile Thr
 1555 1560 1565
 Lys Lys Gly Pro Leu His Glu Thr Pro Ile Ser Asp Gln Tyr Lys Pro
 1570 1575 1580
 Leu Gly Tyr Leu Asp Arg Gln Arg Leu Ala Ala Arg Arg Ser Asn Thr
 585 1590 1595 1600

141

Thr Tyr Cys Tyr Asp Phe Pro Leu Ala Phe Gly Thr Ala Leu Glu Leu
 1605 1610 1615
 Leu Trp Ala Ser Gln His Pro Gly Val Lys Lys Pro Tyr Lys Asp Thr
 1620 1625 1630
 Leu Ile Asn Val Lys Glu Leu Val Phe Ser Lys Pro Glu Gly Ser Ser
 1635 1640 1645
 Gly Thr Ser Leu Asp Leu Val Glu Arg Pro Pro Gly Leu Asn Asp Phe
 1650 1655 1660
 Gly Met Val Ala Trp Cys Leu Asp Met Ser Thr Pro Glu Phe Pro Met
 665 1670 1675 1680
 Gly Arg Lys Leu Leu Val Ile Ala Asn Asp Val Thr Phe Lys Ala Gly
 1685 1690 1695
 Ser Phe Gly Pro Arg Glu Asp Ala Phe Phe Leu Ala Val Thr Glu Leu
 1700 1705 1710
 Ala Cys Ala Lys Lys Leu Pro Leu Ile Tyr Leu Ala Ala Asn Ser Gly
 1715 1720 1725
 Ala Arg Leu Gly Val Ala Glu Glu Val Lys Ala Cys Phe Lys Val Gly
 1730 1735 1740
 Trp Ser Asp Glu Ile Ser Pro Glu Asn Gly Phe Gln Tyr Ile Tyr Leu
 745 1750 1755 1760
 Ser Pro Glu Asp His Glu Arg Ile Gly Ser Ser Val Ile Ala His Glu
 1765 1770 1775
 Val Lys Leu Ser Ser Gly Glu Thr Arg Trp Val Ile Asp Thr Ile Val
 1780 1785 1790
 Gly Lys Glu Asp Gly Ile Gly Val Glu Asn Leu Thr Gly Ser Gly Ala
 1795 1800 1805
 Ile Ala Gly Ala Tyr Ser Lys Ala Tyr Asn Glu Thr Phe Thr Leu Thr
 1810 1815 1820
 Phe Val Ser Gly Arg Thr Val Gly Ile Gly Ala Tyr Leu Ala Arg Leu
 825 1830 1835 1840
 Gly Met Arg Cys Ile Gln Arg Leu Asp Gln Pro Ile Ile Leu Thr Gly
 1845 1850 1855
 Phe Ser Thr Leu Asn Lys Leu Leu Gly Arg Glu Val Tyr Ser Ser His
 1860 1865 1870
 Met Gln Leu Gly Gly Pro Lys Ile Met Gly Thr Asn Gly Val Val His
 1875 1880 1885
 Leu Thr Val Ser Asp Asp Leu Glu Gly Val Ser Ala Ile Leu Asn Trp
 1890 1895 1900
 Leu Ser Tyr Ile Pro Ala Tyr Val Gly Gly Pro Leu Pro Val Leu Ala
 905 1910 1915 1920
 Pro Leu Asp Pro Pro Glu Arg Ile Val Glu Tyr Val Pro Glu Asn Ser
 1925 1930 1935
 Cys Asp Pro Arg Ala Ala Ile Ala Gly Val Lys Asp Asn Thr Gly Lys
 1940 1945 1950
 Trp Leu Gly Gly Ile Phe Asp Lys Asn Ser Phe Ile Glu Thr Leu Glu
 1955 1960 1965
 Gly Trp Ala Arg Thr Val Val Thr Gly Arg Ala Lys Leu Gly Gly Ile
 1970 1975 1980

142

Pro Val Gly Val Val Ala Val Glu Thr Gln Thr Val Met Gln Ile Ile
 985 1990 1995 2000
 Pro Ala Asp Pro Gly Gln Leu Asp Ser His Glu Arg Val Val Pro Gln
 2005 2010 2015
 Ala Gly Gln Val Trp Phe Pro Asp Ser Ala Ala Lys Thr Ala Gln Ala
 2020 2025 2030
 Leu Met Asp Phe Asn Arg Glu Glu Leu Pro Leu Phe Ile Leu Ala Asn
 2035 2040 2045
 Trp Arg Gly Phe Ser Gly Gly Gln Arg Asp Leu Phe Glu Gly Ile Leu
 2050 2055 2060
 Gln Ala Gly Ser Thr Ile Val Glu Asn Leu Arg Thr Tyr Arg Gln Pro
 065 2070 2075 2080
 Val Phe Val Tyr Ile Pro Met Met Gly Glu Leu Arg Gly Gly Ala Trp
 2085 2090 2095
 Val Val Val Asp Ser Gln Ile Asn Ser Asp Tyr Val Glu Met Tyr Ala
 2100 2105 2110
 Asp Glu Thr Ala Arg Gly Asn Val Leu Glu Pro Glu Gly Thr Ile Glu
 2115 2120 2125
 Ile Lys Phe Arg Thr Lys Glu Leu Leu Glu Cys Met Gly Arg Leu Asp
 2130 2135 2140
 Gln Lys Leu Ile Ser Leu Lys Ala Lys Leu Gln Asp Ala Lys Gln Ser
 145 2150 2155 2160
 Glu Ala Tyr Ala Asn Ile Glu Leu Leu Gln Gln Gln Ile Lys Ala Arg
 2165 2170 2175
 Glu Lys Gln Leu Leu Pro Val Tyr Ile Gln Ile Ala Thr Lys Phe Ala
 2180 2185 2190
 Glu Leu His Asp Thr Ser Met Arg Met Ala Ala Lys Gly Val Ile Lys
 2195 2200 2205
 Ser Val Val Glu Trp Ser Gly Ser Arg Ser Phe Phe Tyr Lys Lys Leu
 2210 2215 2220
 Asn Arg Arg Ile Ala Glu Ser Ser Leu Val Lys Asn Val Arg Glu Ala
 225 2230 2235 2240
 Ser Gly Asp Asn Leu Ala Tyr Lys Ser Ser Met Arg Leu Ile Gln Asp
 2245 2250 2255
 Trp Phe Cys Asn Ser Asp Ile Ala Lys Gly Lys Glu Glu Ala Trp Thr
 2260 2265 2270
 Asp Asp Gln Val Phe Phe Thr Trp Lys Asp
 2275 2280

<210> 177

<211> 7029

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: construct
 coding for artificial fusion protein of A.thaliana
 Accase and plastidic transitpeptide of tobacco
 transketolase

<400> 177

```

atggcgtctt cttcttctct cactctctct caagctatcc tctctcgttc tgtccctcgc 60
catggctctg cctcttcttc tcaacttttc ccttcttctc tcactttttc cggccttaaa 120
tccaatccca atatcaccac ctcccgccgc cgtactcctt cctccgccgc cgccgccgcc 180
gtcgtaaagg caccggcgat tctgtcctca gctgcaaccg aaaccataga gaaaactgag 240
actgcgtgga tcccggtcgc acaaactagt gctggctcgg ttaacgggaa tcatagtgtc 300
gtaggacctg gtataaatta tgagacggtg tctcaagtgg atgagttctg taaagcactt 360
agagggaaaa ggccgatcca tagtattttg atagctaaca atggaatggc ggctgtgaag 420
tttatacgta gtgtcagaac atgggcttat gaaacatttg gtacggaaaa agccatattg 480
ttgggtggga tggcaacccc tgaagacatg cggatcaatg cggagcatat cagaatcgct 540
gatcagtttg ttgaggttcc cggaggaacc aacaataaca attatgctaa cgttcagctg 600
attgtggaga tggctgaagt aacacgcgtg gatgcagttt ggctgggttg gggctcatgca 660
tctgaaaacc ccgaattacc tgatgcccta gatgcaaaag gaatcatatt tcttggctct 720
ccagcatctt caatggcagc actgggagat aagattgggt cttcgttgat tgcacaagct 780
gctgatgtac ccactctgcc atggagtggg tcccatgtta aaatacctcc taatagcaac 840
ttggtaacca tcccagagga gatctaccgg caagcatgtg tctacacaac tgaagaagcg 900
attgctagct gtcaagttgt cggttaccca gcaatgatca aagcatcgtg ggggtgggtg 960
ggtaaaggaa tcaggaaggt tcataatgat gatgaggtta gggctctatt caagcaagtt 1020
caggggtgagg tcccaggctc accaatattc ataatgaagg ttgcgtcaca gagtccgcat 1080
ctagaggctc agctgctctg tgacaagcat ggaaatgttt cagctctgca tagccgtgat 1140
tgtagcgtcc agagaagaca tcaaaagatc atagaggagg gtccaattac tgtggctccg 1200
ccagaagctg tcaagaaact tgaacaagca gctagaaggt tggctaagag tgttaactat 1260
gttgagctg ctactgttga gtatctctac agtatggaca ctggggagta ctacttctta 1320
gagcttaacc ctcgcttaca ggttgagcat cctgtcactg agtggattgc cgagataaat 1380
cttctgctg cccaagttgc tgtggggatg ggaattcctc tctggcaaat ccctgagata 1440
agacggttct atggaataga acatggtgga ggttatgatt cttggcgaaa aacatctgtt 1500
gtagccttcc cttttgattt tgataaagct caatctataa ggccaaaagg tcattgtgtg 1560
gctgtacgtg tgacaagtga ggatcctgat gacgggttca aaccaaccag cggtagagtt 1620
caggagtga gttttaagag caagccaaat gtgtgggcgt acttctctgt caagtctggt 1680
ggaggcatcc acgagttctc ggattcccag tttggacatg tttttgcatt tggggaatcc 1740
agagccctgg cgatagcgaa tatggttctt gggctaaaag aaattcagat ccgtggagaa 1800
attaggacta acgttgacta cacgatcgac cttttacatg cttctgatta ccgtgataac 1860
aaaattcaca ctggttggtt ggatagtagg attgctatgc gggtcagagc tgagaggcct 1920
ccatggtatc tctctgttgt cggcggagct ctctataaag catcagcgac cagtgtgtct 1980
gtggtttcag attacgttgg ttatctggag aaggggcaaa tccctccaaa gcatatatct 2040
cttgtagatt ctcaagtgtc tctgaatatt gaagggaagta aatatacgat tgatgtagtc 2100
cggggtggat caggaaccta caggctaaga atgaacaagt cagaagtggg agcagaaata 2160
cacactctac gtgatggagg tctgttgatg cagttggatg gcaaaagcca tgtgatatat 2220
gcagaggaag aagctgcagg aactcgtctt ctcatgtgat gaagaacttg tttgctacag 2280
aatgaccacg atccatcaaa gttaatggct gagacaccgt gcaagttgat gaggtatttg 2340
atttccgaca acagcaatat tgacgtgat acgccttatg ccgaagttga ggtcatgaag 2400
atgtgcatgc cacttctttc acctgcttca ggagttatcc attttaaaat gtctgaagga 2460
caagccatgc aggtcgttga acttatagcc aatcttgatc ttgatgatcc ttctgtgtga 2520
agaaaggccg aacccttcca tgggaagttt ccaagattag ggcttccaac tgcaatatcc 2580
ggtagagttc atcagagatg tgccgcaaca ttaaagtctg cacgcatgat tcttgtgtgc 2640
tatgagcata aagtagatga ggttggtcaa gacttactta attgccttga tagccctgaa 2700
ctccatttcc ttcagtggca agagtgtctt gcagttctgg cgacacgact acctaaaaat 2760
ctcaggaaca tgctagaatc aaagtatagg gaatttgaga gtatttccag aaactctttg 2820
accaccgatt tccctgccaa acttttaaaa ggcattcctt aggcacattt atcttcttgt 2880
gatgagaaag agagaggtgc ccttgaaagg ctcatgaac cattgatgag ccttgcaaaa 2940
tcttatgaag gtggtagaga aagtcatgcc cgtgttattg ttcattctct ctttgaagaa 3000
tatctatcag tagaagaatt attcaatgat aacatgctgg ctgatgttat agaacgcagt 3060
cgtcagctat acaagaaaga tctgttgaaa attgtggata tagtgctctc acaccagggc 3120
ataaaaaaca aaaacaaact cgttctccgg ctcatggagc agcttgttta ccctaatact 3180
gctgcttaca gagataaact tattcgattc tcaacactta accatactaa ctactctgag 3240
ttggcgctca aggcgagtca attacttgaa cagaccaaac taagttagct tcgttcaaac 3300
attgctagaa gcctttcaga gttagaaatg tttacagagg acggagaaaa tatggatact 3360

```

acatcgcac	gagcgcac	gtactcggat	ggaagccggt	cttgtecatc	aggatgatct	3840
ggacgaagag	catcaggggc	tcgcgccagc	cgaactgttc	gccaggctca	aggcgcgcac	3900
gcccgcgcgc	gaggatctcg	tcgtgaccca	tggcgatgcc	tgcttgccga	atatcatggt	3960
ggaaaatggc	cgcttttctg	gattcatcga	ctgtggcccg	ctgggtgtgg	cggaccgcta	4020
tcaggacata	gcgttggtta	cccgtgatat	tgctgaagag	cttggcggcg	aatgggctga	4080
ccgcttcctc	gtgcttttac	gtatcgccgc	tcccgattcg	cagcgcacac	ccttctatcg	4140
ccttcttgac	gagttcttct	gagcgggacc	caagctctag	atcttgctgc	gttcggatat	4200
tttcgtggag	ttcccgcac	agaccgggat	gatccccgat	cgttcaaaca	tttggcaata	4260
aagtttctta	agattgaatc	ctgttgcccg	tcttgcatg	attatcatat	aatttctggt	4320
gaattacgtt	aagcatgtaa	taattaacat	gtaatgcatg	acgttattta	tgagatgggt	4380
ttttatgatt	agagtcgccg	aattatacat	ttaatacgcg	atagaaaaca	aaatatagcg	4440
cgcaaactag	gataaattat	cgcgcgcggt	gtcatctatg	ttactagatc	gggcctcctg	4500
tcaagctctg	cttggttaata	attgtcatta	gattgttttt	atgcatagat	gcactcgaaa	4560
tcagccaatt	ttagacaagt	atcaaaccga	tgtaatttca	gtacattaaa	gacgtccgca	4620
atgtgttatt	aagttgtcta	agcgtcaatt	tgtttacacc	acaatatatc	ctgccaccag	4680
ccagccaaca	gtcctccgac	cggcagctcg	gcacaaaatc	accacgcgtt	accaccacgc	4740
cggccggccg	catggtgttg	accgtgttcg	ccggcattgc	cgagttcgag	cgttccctaa	4800
tcacgcaccg	caccgggagc	gggcgcgagg	ccgccaaggc	ccgaggcggtg	aagtttggtc	4860
cccgccttac	cctcaccctg	gcacagatcg	cgcacgcccc	cgagctgac	gaccaggaag	4920
gccgcaccgt	gaaagaggcg	gctgcaactg	ttggcgtgca	tcgctcgacc	ctgtaccgcg	4980
cacttgagcg	cagcgaggaa	gtgacgcccc	ccgaggccag	gcggcgcggt	gccttccgtg	5040
aggacgcatt	gaccgaggcc	gacgccttgg	cggccgcgca	gaatgaacgc	caagaggaac	5100
aagcatgaaa	ccgcaccagg	acggccaggga	cgaaccgttt	ttcattaccg	aagagatcga	5160
ggcggagatg	atcgcggccg	ggtacgtgtt	cgagccgccc	gcgcacgtct	caaccgtgcg	5220
gctgcatgaa	atcctggccg	gtttgtctga	tgccaagctg	gcggcctggc	cggccagctt	5280
ggccgctgaa	gaaaccgagc	gccgcctgtc	aaaaaggtga	tgtgtatttg	agtaaaacag	5340
cttgcgatcat	gcggtcgctg	cgtatatgat	gcgatgagta	aataaaca	tacgcaaggg	5400
gaacgcacga	aggttatcgc	tgtacttaac	cagaaaggcg	ggtcaggcaa	gacgaccatc	5460
gcaacccatc	tagcccgccg	cctgcaactc	gccggggccg	atgttctgtt	agtcgattcc	5520
gatccccagg	gcagtgcctg	cgattgggag	gccgtgcggg	aagatcaacc	gctaaccgtt	5580
gtcggcatcg	accgcccgc	gattgaccgc	gacgtgaagg	ccatcgcccg	gcgcgacttc	5640
gtagtgatcg	acggagcgcc	ccaggcgccg	gacttggtcg	tgtccgcgat	caaggcagcc	5700
gacttcgtgc	tgattccggt	gcagccaagc	ccttacgaca	tatgggccac	cgccgacctg	5760
gtggagctgg	ttaagcagcg	cattgaggtc	acggatggaa	ggctacaagc	ggcctttgtc	5820
gtgtcgcggg	cgatcaaagg	cacgcgcac	ggcgggtgag	ttgccgaggc	gctggccggg	5880
tacgagctgc	ccattcttga	gtcccgtatc	acgcagcgcg	tgagctaccc	aggcactgcc	5940
gccgcgggca	caaccgttct	tgaatcagaa	cccgaggcg	acgtgcccgc	cgagggtccag	6000
gcgctggccg	ctgaaattaa	atcaaaactc	atlttgagtta	atgaggtaaa	gagaaaatga	6060
gcaaaagcac	aaacacgcta	agtgcgggcc	gtccgagcgc	acgcagcagc	aaggctgcaa	6120
cgttggccag	cctggcagac	acgccagcca	tgaagcgggt	caactttcag	ttgccggcgg	6180
aggatcacac	caagctgaag	atgtacgcgg	tacgccaagg	caagaccatt	accgagctgc	6240
tatctgaata	catcgcgag	ctaccagagt	aaatgagcaa	atgaataaat	gagtagatga	6300
atlttagcgg	ctaaaggagg	cggcatggaa	aatcaagaac	aaccaggcac	cgacgcctgt	6360
gaatgcccc	tgtgtggagg	aacgggcggg	tggccaggcg	taagcggctg	ggttgtctgc	6420
cggccctgca	atggcactgg	aacccccaa	cccagggaat	cggcgtgagc	ggtcgcaaac	6480
catccggccc	ggtacaaatc	ggcgcgccgc	tgggtgatga	cctggtggag	aagttgaagg	6540
ccgcgcaggc	cgcccagcgg	caacgcacgc	aggcagaagc	acgccccggt	gaatcggtgc	6600
aagcggccgc	tgatcgaatc	cgcaagaat	cccggcaacc	gccggcagcc	ggtgcgccgt	6660
cgattaggaa	gccgccccag	ggcgacgagc	aaccagattt	tttcgttccg	atgctctatg	6720
acgtgggcac	ccgcgatagt	cgcagcatca	tggacgtggc	cgttttccgt	ctgtcgaaag	6780
gtgaccgacg	agctggcgag	gtgatccgct	acgagcttcc	agacgggcac	gtagagggtt	6840
ccgcagggcc	ggccggcatg	gccagtgtgt	gggattacga	cctggtactg	atggcggttt	6900
cccatctaac	cgaatccatg	aaccgatata	gggaaggga	gggagacaag	cccggccgcg	6960
tgttccgtcc	acacgttgcg	gacgtactca	agttctgccc	gcgagccgat	ggcggaaagc	7020
agaaagacga	cctggtagaa	acctgcattc	ggttaaaccac	cacgcacgtt	gccatgcagc	7080
gtacgaagaa	ggccaagaac	ggccgcctgg	tgacgggtatc	cgagggtgaa	gccttgatta	7140
gccgctacaa	gatcgtaaa	agcgaaaccg	ggcggccgga	gtacatcgag	atcgagctag	7200

ctgattggat	gtaccgcgag	atcacagaag	gcaagaaccc	ggacgtgctg	acggttcacc	7260
ccgattactt	tttgatcgat	cccggcatcg	gccgttttct	ctaccgcctg	gcacgccgcg	7320
ccgcaggcaa	ggcagaagcc	agatggttgt	tcaagacgat	ctacgaacgc	agtggcagcg	7380
ccggagagtt	caagaagtcc	tgtttcaccg	tgcgcaagct	gatcgggtca	aatgacctgc	7440
cggagtacga	tttgaaggag	gaggcggggc	aggctggccc	gacccctagtc	atgcgctacc	7500
gcaacctgat	cgagggcgaa	gcatccgccg	gttcctaattg	tacggagcag	atgctagggc	7560
aaattgccct	agcaggggaa	aaaggctcgaa	aaggctctctt	tcctgtggat	agcacgtaca	7620
ttgggaaccc	aaagccgtac	attgggaacc	ggaacccgta	cattgggaac	ccaaagccgt	7680
acattgggaa	ccggtcacac	atgtaagtga	ctgatataaa	agagaaaaaa	ggcgattttt	7740
ccgcctaaaa	ctctttaaaa	cttattaaaa	ctcttaaaac	ccgcctggcc	tgtgcataac	7800
tgtctggcca	gcgcacagcc	gaagagctgc	aaaaagcgcc	tacccttcgg	tcgctgcgct	7860
ccctacgccc	cgccgcttcg	cgtcggccta	tcgcgccgcg	tggccgctca	aaaatggctg	7920
gcctacggcc	aggcaatcta	ccagggcgcg	gacaagccgc	gccgtcgcca	ctcgaccgcc	7980
ggcgcccaca	tcaaggcacc	ctgcctcgcg	cgtttcgggtg	atgacgggtga	aaacctctga	8040
cacatgcagc	tcccggagac	ggtcacagct	tgtctgtaag	cggatgccgg	gagcagacaa	8100
gcccgtcagg	gcgcgtcagc	gggtgttggc	gggtgtcggg	gcgcagccat	gaccagtcga	8160
cgtacgcgata	gcggagtgtg	tactggctta	actatgcggc	atcagagcag	attgtactga	8220
gagtgcacca	tatgcggtgt	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	8280
ggcgctcttc	cgcttcctcg	ctcactgact	cgctgcgctc	ggcgttcgg	ctgcggcgag	8340
cggtatcagc	tcactcaaag	gcggtaatat	ggttatccac	agaatcaggg	gataacgcag	8400
gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	8460
tggcgttttt	ccataggctc	cgccccctcg	acgagcatca	caaaaatcga	cgctcaagtc	8520
agaggtggcg	aaacccgaca	ggactataaa	gataaccaggc	gtttccccct	ggaagctccc	8580
tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	8640
cggaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	8700
ttcgctccaa	gctgggctgt	gtgcacgaac	cccccgttca	gcccgaccgc	tgcgccttat	8760
ccggtaaacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	8820
ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	8880
ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	8940
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaacaaacc	accgctggta	9000
gcgggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	9060
atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgtaagggga	9120
ttttggatcat	gcatgatata	tctcccaatt	tgtgtagggc	ttattatgca	cgcttaaaaa	9180
taataaaaagc	agacttgacc	tgatagtttg	gctgtgagca	attatgtgct	tagtgcattct	9240
aacgcttgag	ttaagccgcg	ccgcgaagcg	gcgtcggctt	gaacgaattt	ctagctagac	9300
attattttgcc	gactaccttg	gtgatctcgc	ctttcacgta	gtggacaaat	tcttccaact	9360
gatctgcgcg	cagggccaaag	cgatcttctt	cttgtccaag	ataagcctgt	ctagcttcaa	9420
gtatgacggg	ctgatactgg	gccggcaggc	gctccattgc	ccagtcggca	gcgacatcct	9480
tcggcgcgat	tttgccgggtt	actgcgctgt	accaaattgcg	ggacaacgta	agcactacat	9540
ttcgctcctc	gccagcccag	tcgggcccgcg	agttccatag	cgtaagggtt	tcatttagcg	9600
cctcaaatag	atcctgttca	ggaaccggat	caaagagttc	ctccgcccgt	ggacctacca	9660
aggcaacgct	atgtttctctt	gctttttgtca	gcaagatagc	cagatcaatg	tcgatcgtgg	9720
ctggctcgaa	gatacctgca	agaatgtcat	tgcgctgccca	ttctccaaat	tgcagttcgc	9780
gcttagctgg	ataacgccac	ggaatgatgt	cgctcgtgcac	aacaatgggtg	acttctacag	9840
cgcgagagaat	ctcgctctct	ccaggggaag	ccgaagtttc	caaaagggtcg	ttgatcaaag	9900
ctcgcccgct	tgtttcatca	agccttacgg	tcaccgtaac	cagcaaatca	atatcactgt	9960
gtggcttcag	gccgccatcc	actgcggagc	cgtacaaatg	tacggccagc	aacgtcgggt	10020
cgagatggcg	ctcgatgacg	ccaactacct	ctgatagttg	agtcgatact	tcggcgatca	10080
ccgcttcccc	catgatgttt	aactttgttt	tagggcgact	gccctgctgc	gtaacatcgt	10140
tgctgctcca	taacatcaaa	catcgacca	cggcgtaacg	cgcttgctgc	ttggatgccc	10200
gaggcataga	ctgtacccca	aaaaaacagt	cataacaagc	catgaaaacc	gccactgcg	10259

cccaagagga	aaagtgccat	taatgaaaga	atagaagatc	ttgtaagcgc	atcttttagct	3420
gttgaagacg	ctctcgtggg	actatttgac	catagcgatc	acacacttca	aagacgggtt	3480
gttgagactt	atattcgcag	attataccag	ccctacgtcg	ttaaagatag	cgtgaggatg	3540
cagtggcacc	gttctggtct	tcttgcttcc	tgggagttcc	tagaggagca	tatggaaaga	3600
aaaaacattg	gcttagacga	tcccgacaca	cctgaaaaag	gattgggttg	gaagcgtagt	3660
aagagaaaat	ggggggctat	ggttataatc	aaatctttgc	agtttcttcc	aagtataata	3720
agtgcagcat	tgagagaaac	aaagcacaac	gactatgaaa	ctgccggagc	tcctttatct	3780
ggcaatatga	tgacatttgc	tattgtgggc	atcaacaacc	agatgagtct	gcttcaggac	3840
agtggggatg	aagaccaagc	tcaggaaaga	gtaaacaagt	tggccaaaat	tcttaaagag	3900
gaagaagtga	gttcaagcct	ctgttctgcc	ggtgttggtg	taatcagctg	tataattcag	3960
cgagatgaag	gacgaacacc	catgagacat	tctttccatt	ggtcgttgga	gaaacagtat	4020
tatgtagaag	agccgttgct	gcgtcatctt	gaacctcttc	tgtccattta	ccttgagttg	4080
gataagctga	aaggatactc	aaatatacaa	tatacgcttt	ctcgagatcg	tcaatggcat	4140
ctgtatactg	ttacagacaa	gccagtgcc	atcaagagga	tgttcctgag	atctcttggt	4200
cgacaggcta	caatgaacga	tggatttata	ttgcagcaag	ggcaggataa	gcagcttagc	4260
caaacactga	tctccatggc	gtttacgtcg	aatgtgttcc	tgaggtcttt	gatggatgcc	4320
atggagggaac	tggaaactgaa	tgcccataat	gctgcaatga	aaccagatca	cgcacatatg	4380
tttcttttga	tattgcgtga	gcagcagata	gatgatcttg	tgcctttccc	caggagagtt	4440
gaagtgaatg	cggaggatga	agaaactaca	gttgaaatga	tcttagaaga	agcagcacga	4500
gagatacata	gatctgttgg	agtgagaatg	cataggtttg	gcgtgtgcga	gtgggaagtg	4560
cggctgtggg	tgggtgcctc	tggactggca	tgtggtgctt	ggagggttgt	ggttgcaaac	4620
gtgacaggcc	gtacatgcac	tgtccacata	taccgagaag	ttgaaactcc	tggaaagaaac	4680
agttttattct	accactcaat	aaccaagaag	ggacctttgc	atgaaacacc	aatcagtgat	4740
caatataagc	ccctgggata	tctcgacagc	caacgtttag	cagcaaggag	gagtaacact	4800
acttatttgc	atgacttccc	gttggcattt	gggacagcct	tggaaactgtt	gtgggcatca	4860
caacaccag	gagttaagaa	accatataag	gatactctga	tcaatgttaa	agagcttgta	4920
ttctcaaaac	cagaaggttc	ttcgggtaca	tctctagatc	tggttgaaag	accacccggt	4980
ctcaacgact	ttggaatggg	tgcctgggtc	ctagatatgt	cgaccccgag	gtttcctatg	5040
gggcggaac	ttctcgtgat	tgcgaatgat	gtcaccttca	aagctgggtc	ttttggtcct	5100
agagaggacg	cgtttttccc	tgctgttact	gaactcgctt	gtgccaagaa	gcttcccttg	5160
atttacttgg	cagcaaattc	tgggtgcccg	cttggggttg	ctgaagaagt	caaagcctgc	5220
ttcaaagtgg	gatggtcgga	tgaattttcc	cctgagaatg	gttttcagta	tatatacct	5280
agccctgaag	accacgaaag	gattggatca	tctgtcattg	cccatgaagt	aaagctctct	5340
agtggggaaa	ctaggtgggt	gattgatagc	atcgltggca	aagaagatgg	tattggtgta	5400
gagaacttaa	caggaagtgg	ggccatagcg	ggtgcttact	caaaggcata	caatgaaact	5460
tttactttta	cctttgttag	tgggaagaacg	gttgggaattg	gtgcttatct	tgcccgccct	5520
ggtatgcggg	gcatacagag	acttgatcag	ccgatcatct	tgactggctt	ctctacactc	5580
aacaagttac	ttgggcgtga	ggtctatagc	tctcacatgc	aactgggtgg	cccgaaaatc	5640
atgggcacaa	atgggtgttg	tcatcttaca	gtctcagatg	atcttgaagg	cgtatcagca	5700
attctcaact	ggctcagcta	cattcctgct	tacgtgggtg	gtcctcttcc	tgttcttgcc	5760
ccttttagatc	caccggagag	aattgtggag	tatgtcccag	agaactcttg	cgacccacga	5820
gcggctatag	ctgggggtcaa	agacaatacc	ggtaaatggc	ttggagggtat	ccttgataaa	5880
aatagtttca	ttgagactct	tgaaggctgg	gcaaggacgg	tagtgactgg	tagagccaag	5940
ctcgggggaa	taccggttgg	agttgttgca	gttgagacac	agactgtcat	gcagatcatc	6000
ccagccgatc	ctggacagct	tgactctcat	gaaagagtgg	ttccgcaagc	agggcaagtc	6060
tggtttccctg	attcagcggc	caagactgct	caagcgctta	tggatttcaa	ccgggaagag	6120
cttccattgt	ttatcctagc	gaactggaga	gggttttccag	gtgggcagag	agatcttttc	6180
gaagggaatac	ttcaggcagg	ttcaactata	gtagaaaatc	tgagaaccta	tgcgtcagcca	6240
gtgtttgtgt	acatcccaat	gatgggagag	ctgcgcggtg	gagcgtgggt	tgttgttgac	6300
agccagataa	attcggatta	tgttgaaatg	tatgctgatg	aaacagctcg	tggaaatgtg	6360
cttgagccag	aagggacaat	agagataaaa	tttagaacia	aagagctatt	agagtgcattg	6420
ggaagggttg	accgaagct	aatcagtctg	aaagcaaaac	tgcaagatgc	caagcaaaagc	6480
gaggccatag	caaacatcga	gcttctccag	caacagatta	aagcccgaga	gaaacagctt	6540
ttaccagttt	atatccaaat	cgccacccaa	tttgagaac	ttcatgacac	ttccatgaga	6600
atggctgcaa	agggagtgat	caaaagtgtt	gtggaatgga	gcggctcgcg	gtccttcttc	6660
tacaaaaagc	tcaataggag	aatcgctgag	agctctcttg	tgaaaaacgt	aagagaagca	6720
tctggagaca	acttagcata	taaatcttca	atgcgtctga	ttcaggattg	gttctgcaac	6780

145

```
tctgatattg caaaggggaa agaagaagct tggacagacg accaagtgtt ctttacatgg 6840
aaggacaatg ttagtaacta cgagttgaag ctgagcgagt tgagagcgca gaaactactg 6900
aaccaacttg cagagattgg gaattcctca gatttgcaag ctctgccaca aggacttgct 6960
aatcttctaa acaaggtgga gccgtcgaaa agagaagagc tgggtggctgc taatcgaaag 7020
gtcttgggt
```

<210> 178

<211> 8482

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: binary
expression vector

<400> 178

```
ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttattctaataa aaacgctctt ttctcttagg tttaccgcgc aatatatcct gtcaaacact 120
gatagttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aatcaccact ttgtacaaga aagctgggtc tagatgacgg acaatcagta 240
aattgaacgg agaataattat tcaaaaaaat acgatagtaa cgggtgatat attcattaga 300
atgaaccgaa accggcggtg aggatctgag ctacacatgc tcagggtttt tacaacgtgc 360
acaacagaat tgaaagcaaa tatcatgcga tcataggcgt ctcgcatatc tcattaaagc 420
aggaggcctt ctagactgca ggcggccgcc caccgcggtg ggctggctat gaagaaatta 480
taatcgtgta aaacttagtg agtgtgtatg aatgaaagta ttgcaaaatc ctcatatat 540
agactacatg cataactagt tgcattgtaa tttgtagttt tcttcattat tgcattcctc 600
aagtggatgt catgggtttt cacatggcct ccatgcaaat catttccaaa atatttttaa 660
actttccaca gggcatccat gcatgcacct caaaacttgt gtgtggtaac attgttgtct 720
tgaaaaatta ctaaaccctt tgtccacgtg acgttcatgc acctcaaatc ttgtgtggta 780
ccattattat cctcaagaat tattgaatgt ttggtgtata tgccatccat gcagcattgc 840
aacaattaaa tctccaaacc ttgtggtacc atattcactc actttaattc tcctatagta 900
gaaatattag caaatattta catttccagt tgattagtat atgtatttag aagacaaaaa 960
taatttagaa tcaattaatc aacttgcaaa ttgctaagtg ttggcaaacg ttagcataaa 1020
agggtgtata aatttagtac caaatataaa aatttatcgc aaatcaaata cataacacac 1080
atagtaaaac aaaaacaaat tacaagggtt tagacgttta gtggcaatgt gtaaatttgc 1140
tcgactgaat tggttccttt aagcctgctt ttttgtacaa acttgtgata attcactggc 1200
cgtcgtttta caacgactca ggatcctgtc aaacactgat agtttaaact gaaggcggga 1260
aacgacaatc tgatcatgag cggagaatta agggagtcac gttatgacct ccgccgatga 1320
cgccgggacaa gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc 1380
agccgcgggt ttctggagtt taatgagcta agcacaatag tcagaaacca ttattgcgcg 1440
ttcaaaagtc gcctaaggtc actatcagct agcaaatatt tcttgtcaaa aatgctccac 1500
tgacgttcca taaattcccc tcggtatcca attagagtct catattcact ctcaatccaa 1560
ataatctgca ccggatctgg atcgtttcgc atgattgaac aagatggatt gcacgcaggt 1620
tctccggccg cttgggtgga gaggtattc ggctatgact gggcacaaca gacaatcggc 1680
tgctctgatg ccgcgctggt ccggtgtcga gcgcaggggc gcccggttct ttttgtcaag 1740
accgacctgt ccggtgccct gaatgaactg caggacgagg cagcgcggct atcgtggctg 1800
gccacgacgg gcgttccttg cgcagctgtg ctgcagcttg tcaactgaagc gggaaggga 1860
tggctgctat tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc 1920
gagaaagtat ccatcatggc tgatgcaatg cggcggctgc atacgcttga tccggctacc 1980
tgcccatctg accaccaagc gaaacatcgc atcgagcgag cacgtactcg gatggaagcc 2040
ggtcttgctg atcaggatga tctggacgaa gagcatcagg ggctcgcgcc agccgaactg 2100
ttcgccaggc tcaaggcgcg catgcccgac ggcgaggatc tcgtcgtgac ccatggcgat 2160
gcctgcttgc cgaatatcat ggtggaaaat ggcgcgtttt ctggattcat cgactgtggc 2220
cggctgggtg tggcggaacc ctatcaggac atagcgttgg ctaccctga tattgtgaa 2280
gagcttggcg gcgaatgggc tgaccgcttc ctctgtcttt acggtatcgc cgctcccgat 2340
tcgcagcgca tcgccttcta tcgccttctt gacgagttct tctgagcggg acccaagctc 2400
tagatcttgc tgcgttcgga tattttcgtg gagttccgcg cacagaccgg gatgatcccc 2460
gatcgttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cggctcttgcg 2520
atgattatca tataatttct gttgaattac gtttaagcatg taataattaa catgtaatgc 2580
```

atgacgttat	ttatgagatg	ggttttttatg	attagagtcc	cgcaattata	cattttaatac	2640
gcgataaaaa	acaaaatata	gcgcgcgaaac	taggataaat	tatcgcgcg	ggtgtcatct	2700
atgttactag	atcgggcctc	ctgtcaagct	ctgcttggtg	ataattgtca	ttagattggt	2760
tttatgcata	gatgcactcg	aaatcagcca	atttttagaca	agtatcaaac	ggatgttaat	2820
tcagtacatt	aaagacgtcc	gcaatgtgtt	attaagtgtg	ctaagcgtca	atgtgtttac	2880
accacaatat	atcctgccac	cagccagcca	acagctcccc	gaccggcagc	tcggcacaaa	2940
atcaccacgc	gttaccacca	cgccggcccg	ccgcatgggtg	ttgaccgtgt	tcgcccgcac	3000
tgccgagttc	gagcgttccc	taatcatcga	ccgcaccccg	agcgggccc	aggccgcca	3060
ggccccgagg	gtgaagtttg	gccccgcgcc	tacctcacc	ccggcacaga	tcgcgcacgc	3120
ccgcgagctg	atcgaccagg	aaggccgcac	cgtgaaagag	gcggctgcac	tgcttggcgt	3180
gcacgcctcg	accctgtacc	gcgcacttga	gcgcagcgag	gaagtgaacg	ccaccgaggc	3240
caggcggcgc	ggtgccttcc	gtgaggacgc	attgaccgag	gccgacgccc	tgccggccgc	3300
cgagaatgaa	cgccaagagg	aacaagcatg	aaaccgcacc	aggacggcca	ggacgaaccg	3360
tttttcatta	ccgaagagat	cgaggcggag	atgatcgcg	ccgggtacgt	gttcgagccg	3420
cccgcgcacg	tctcaaccgt	gcggctgcat	gaaatcctgg	ccggtttgtc	tgatgccaa	3480
ctggcggcct	ggccggccag	cttgcccgct	gaagaaaccg	agcgcgcgc	tctaaaaagg	3540
tgatgtgtat	ttgagtaaaa	cagcttgctg	catgcggctg	ctgcgtatat	gatgcgatga	3600
gtaataaaac	aaatacgcga	ggggaacgca	tgaaggttat	cgctgtactt	aaccagaaag	3660
gcgggtcagg	caagacgacc	atcgcaaccc	atctagcccc	cgccctgcaa	ctcgccgggg	3720
ccgatgttct	gttagtcgat	tccgatcccc	agggcagtg	ccgcgattgg	gcggccgtgc	3780
gggaagatca	accgctaacc	gttgtcggca	tcgaccgccc	gacgattgac	cgcgacgtga	3840
aggccatcgg	ccggcgcgac	tccgtagtga	tcgaccgagc	gcccaggcg	gcggacttgg	3900
ctgtgtccgc	gatcaaggca	gccgacttcg	tgctgattcc	ggtgcagcca	agcccttacg	3960
acatatgggc	caccgccgac	ctggtggagc	tggttaagca	gcgcattgag	gtcacggatg	4020
gaaggctaca	agcggccttt	gtcgtgtcgc	gggcatcaaa	aggcacgcgc	atcggcgggtg	4080
agggtgcccga	ggcgttggcc	gggtacgagc	tgcccattct	tgagtcccgt	atcacgcagc	4140
gcgtgagcta	cccaggcact	gccgcgcgcg	gcacaaccgt	tcttgaatca	gaaccgagg	4200
gcgacgctgc	ccgcgagggtc	caggcgcgtg	ccgctgaaat	taaatcaaaa	ctcatttgag	4260
ttaatgaggt	aaagagaaaa	tgagcaaaa	cacaaacacg	ctaagtgcgc	gccgtccgag	4320
cgcacgcagc	agcaaggctg	caacgttggc	cagcctggca	gacacgccag	ccatgaagcg	4380
ggtcaacttt	cagttgcgcg	cggaggatca	caccaagctg	aagatgtacg	cggtagccca	4440
aggcaagacc	attaccgagc	tgctatctga	atacatcgcg	cagctaccag	agtaaatgag	4500
caaatgaata	aatgagtaga	tgaatttttag	cggctaaagg	aggcggcatg	gaaaatcaag	4560
aacaaccagg	caccgacgcc	gtggaatgcc	ccatgtgtgg	aggaaacgggc	ggttggccag	4620
gcgtaagcgg	ctgggttgtc	tgccggccct	gcaatggcac	tggaaccccc	aagcccagg	4680
aatcggcgtg	agcggctcga	aaccatccgg	cccgggtaca	atcggcgcgc	cgctgggtga	4740
tgacctgggtg	gagaagttga	aggcdgcgca	ggccgcccag	cggcaacgca	tcgaggcaga	4800
agcacgcccc	ggtgaatcgt	ggcaagcggc	cgctgatcga	atccgcaaag	aatcccggca	4860
accgcccggca	gccggtgcgc	cgctgattag	gaagcgcgcc	aagggcgacg	agcaaccaga	4920
ttttttcgtt	ccgatgctct	atgacgtggg	caccgcgat	agtcgcagca	tcattggacgt	4980
ggccgttttc	cgtctgtcga	agcgtgaccg	acgagctggc	gaggtgatcc	gctacgagct	5040
tccagacggg	cacgtagagg	tttccgcagg	gccggccggc	atggccagtg	tgtgggatta	5100
cgacctggta	ctgatggcgg	tttcccatct	aaccgaatcc	atgaaccgat	accgggaagg	5160
gaagggagac	aagcccggcc	gcgtgttccg	tccacacgtt	gcggacgtac	tcaagttctg	5220
ccggcgagcc	gatggcggaa	agcagaaa	cgacctggta	gaaacctgca	ttcgggttaa	5280
caccacgcac	gttgccatgc	agcgtacgaa	gaaggccaag	aacggccgcc	tggtgacggg	5340
atccgagggg	gaagccttga	ttagccgcta	caagatcgta	aagagcgaaa	ccgggcggcc	5400
ggagtacatc	gagatcgagc	tagctgattg	gatgtaccgc	gagatcacag	aaggcaagaa	5460
cccggacgtg	ctgacggttc	accccgatta	ctttttgatc	gatcccggca	tcggccgttt	5520
tctctaccgc	ctggcacgcc	gcgccgcagg	caaggcagaa	gccagatggg	tgttcaagac	5580
gatctacgaa	cgcagtggca	gcgccggaga	gttcaagaag	ttctgtttca	ccgtgcgcaa	5640
gctgatcggg	tcaaatgacc	tgccggagta	cgatttgaag	gaggaggcgg	ggcaggctgg	5700
cccgatccta	gtcatgcgct	accgcaacct	gatcgagggc	gaagcatccg	ccggttccta	5760
atgtacggag	cagatgctag	ggcaaattgc	cctagcaggg	gaaaaagggtc	gaaaagggtc	5820
ctttcctgtg	gatagcacgt	acattgggaa	cccaaagccg	tacattggga	accggaaccc	5880
gtacattggg	aacccaaagc	cgtacattgg	gaaccgggtca	cacatgtaag	tgactgatat	5940
aaaagagaaa	aaaggcgatt	tttccgccta	aaactcttta	aaacttatta	aaactcttaa	6000

```

aaccgcgctg gcctgtgcat aactgtcttg ccagcgcaca gccgaagagc tgcaaaaagc 6060
gcctaccctt cggctcgtgc gctccctacg cccgcgcgt tcgctcggc ctatcgcggc 6120
cgctggccgc tcaaaaatgg ctggcctaag gccaggcaat ctaccagggc gcggaacaagc 6180
cgcgccgtcg ccaactcgacc gccggcgccc acatcaaggc accctgcctc gcgcggttcg 6240
gtgatgacgg tgaaaacctc tgacacatgc agctcccgga gacggtcaca gcttgtctgt 6300
aagcggatgc cgggagcaga caagcccgtc agggcgcgctc agcgggtgtt ggcgggtgtc 6360
ggggcgcgag catgaccagc tcacgtagcg atagcggagt gtatactggc ttaactatgc 6420
ggcatcagag cagattgtac tgagagtgc ccatatgcgg tgtgaaatac cgcacagatg 6480
cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg 6540
ctcggtcgtt cggctcgggc gagcggatat agctcactca aaggcggtaa tacggttatc 6600
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 6660
gaaccgtaaa aaggcccgct tgctggcggt tttccatagg ctccgcccc ctgacgagca 6720
tcacaaaaat cgacgctcaa gtcagagggt gcgaaacccg acaggactat aaagatacca 6780
ggcggtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 6840
atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag 6900
gtatctcagt tcgggtgtag tcgttcgctc caagctgggc tgtgtgcacg aacccccctg 6960
tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 7020
cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 7080
cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt 7140
tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 7200
cggcaaaaca accaccgctg gtagecgggtg tttttttgtt tgcaagcagc agattacgcg 7260
cagaaaaaaa ggatctcaag aagatccttt gatctttct acgggggtctg acgctcagtg 7320
gaacgaaaac tcacgttaag ggatttttgt catgcatgat atatctccca atttgtgtag 7380
ggcttattat gcacgcttaa aaataataaa agcagacttg acctgatagt ttggctgtga 7440
gcaattatgt gcttagtgca tctaacgctt gagttaagcc gcgccgcgaa gcgcgctcgg 7500
cttgaacgaa tttctagcta gacattattt gccgactacc ttggtgatct cgcctttcac 7560
gtagtggaca aattcttcca actgatctgc gcgcgaggcc aagcgatctt cttcttgtcc 7620
aagataagcc tgtctagctt caagtatgac gggctgatac tgggcccggca ggcgctccat 7680
tgcccagtcg gcagcgacat ccttcggcgc gatcttgccg gttactgcgc cagtcgggcg gcgagttcca 7800
gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg gcgagttcca 7860
tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg gatcaaagag 7920
ttcctccgcc gctggacctt ccaaggcaac gctatgttct cttgcttttg tcagcaagat 7980
agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt cattgcgctg 8040
ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga tgtcgtcgtg 8100
cacaacaatg gtgacttcta cagcgcggag aatctcgtc tctccagggg aagccgaagt 8160
ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta cggtcaccgt 8220
aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg agccgtacaa 8280
atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta cctctgatag 8340
ttgagtcgat acttcggcga tcaccgcttc ccccatgatg tttaactttg ttttagggcg 8400
actgccttgc tgcgtaacat cgttgctgct ccataacatc aaacatcgac ccacggcgta 8460
acgcgcttgc tgcttggatg cccgaggcat agactgtacc ccaaaaaaac agtcataaca 8482
agccatgaaa accgccactg cg

```

<210> 179

<211> 10259

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: : binary
expression vector

<400> 179

```

ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttattctaata aaacgctctt ttctcttagg ttaccgcgc aatatatcct gtcaaact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttggcg cgccgagctc gtcgacggac aatcagtaaa tgtaacggag 240
aatattattc ataaaaatac gatagtaacg ggtgatatat tcattagaat gaaccgaaac 300
cggcggtaag gatctgagct acacatgctc aggtttttta caacgtgcac aacagaattg 360

```

aaagcaaata	tcattgcatc	atagggcgtct	cgcataatctc	attaaagcag	gccatggccc	420
gggatcgatg	ccggcagatc	tgctggctat	gaagaaatta	taatcgtgta	aaacttagtg	480
agtgtgtatg	aatgaaagta	ttgcaaaatc	ctcattatat	agactacatg	cataactagt	540
tgcatgtaaa	tttgtagttt	tcttcattat	tgcatcctcc	aagtggatgt	catggtttta	600
cacatggctt	ccatgcaaat	catttccaaa	atatttttaa	actttccaca	gggcatccat	660
gcatgcacct	caaaacttgt	gtgtggtaac	attgttgtct	tgaaaaatta	ctaaaccttt	720
tgtccacgtg	acgttcatgc	acctcaaatc	tttgtgtgta	ccattattat	cctcaagaat	780
tattgaatgt	ttgggtgata	tgccatccat	gcagcattgc	aacaattaaa	tctccaaacc	840
ttgtggtacc	atattcactc	actttaattc	tcctatagta	gaaatattag	caaataattta	900
catttccagt	tgattagtat	atgtattttag	aagacaaaaa	taatttagaa	tcaattaatc	960
aacttgcaaa	ttgctaagtg	ttggcaaacg	ttagcataaa	aggtgttata	aatttagtac	1020
caaataataa	aatttatcgc	aaatcaaaata	catacacac	atagtaaaac	aaaaacaaat	1080
tacaagggtt	tagacgttta	gtggcaatgt	gtaaatttgc	tcgacggaca	atcagtaaata	1140
tgaacggaga	atattattca	taaaaatacg	atagtaacgg	gtgatataat	cattagaatg	1200
aaccgaaacc	ggcggtaagg	atctgagcta	cacatgctca	ggttttttac	aacgtgcaca	1260
acagaattga	aagcaaatat	catgcatcga	taggcgtctc	gcataatctc	ttaaagcagg	1320
gttaacgcta	gcgggcccga	tatcggatcc	gctggctatg	aagaaattat	aatcgtgtaa	1380
aacttagtga	gtgtgtatga	atgaaagtat	tgcaaaatcc	tcattatata	gactacatgc	1440
ataactagtt	gcatgtaaat	ttgtagtttt	cttcattatt	gcacccctca	agtggatgtc	1500
atggttttac	acatggcttc	catgcaaatc	atttccaaaa	tattttttaa	ctttccacag	1560
ggcatccatg	catgcacctc	aaaacttgtg	tgtggtaaca	ttgttgtctt	gaaaaattac	1620
taaacctttt	gtccacgtga	cgttcatgca	cctcaaactc	tgtgtggtac	cattattatc	1680
ctcaagaatt	attgaatgtt	tggtgtatat	gccatccatg	cagcattgca	acaattaaat	1740
ctccaaacct	tgtggtacca	tattcactca	ctttaattct	cctatagtag	aaatattagc	1800
aaatattttac	atttccagtt	gattagtata	tgtattttaga	agacaaaaat	aatttagaat	1860
caattaatca	acttgcaaat	tgctaagtgt	tggcaaacgt	tagcataaaa	gggtgttataa	1920
atttagtacc	aaatataaaa	atttatcgca	aatcaaatac	ataacacaca	tagtaaaaaa	1980
aaaacaaatt	acaagggttt	agacgttttag	tggcaatgtg	taaattttgt	cgacggacaa	2040
tcagtaaatt	gaacggagaa	tattattcat	aaaaatacga	tagtaacggg	tgatataattc	2100
attagaatga	accgaaaccg	gcggtaagga	tctgagctac	acatgctcag	gtttttttaca	2160
acgtgcacaa	cagaattgaa	agcaaatatc	atgcgatcat	aggcgtctcg	cataatctcat	2220
taaagcagga	ggccttctag	actgcaggcg	gccgccacc	gcgggtgggt	ggctatgaag	2280
aaattataat	cgtgtaaaaa	ttagtgtgtg	tgtatgaatg	aaagtattgc	aaaatcctca	2340
ttatatagac	tacatgcata	actagttgca	tgtaaatttg	tagttttctt	cattattgca	2400
tcctccaagt	ggatgtcatg	gttttacaca	tggcttccat	gcaaatcatt	tccaaaatat	2460
ttttaaactt	tccacagggc	atccatgcat	gcacctcaaa	acttgttgtg	ggtaacattg	2520
ttgtcttgaa	aaattactaa	accttttgtc	cacgtgacgt	tcatgcacct	caaatcttgt	2580
gtggtaccat	tattatcctc	aagaattatt	gaatgttttg	tgtatatgcc	atccatgcag	2640
cattgcaaca	attaaatctc	caaaccttgt	ggtaccatat	tcactcactt	taattctcct	2700
atagtagaaa	tatttagcaa	tattttacatt	tccagttgat	tagtatatgt	atttagaaga	2760
caaaaataat	ttagaatcaa	ttaatcaact	tgcaaatgtc	taagtgttgg	caaacgttag	2820
cataaaaagg	gttataaaat	tagtaccaaa	tataaaaatt	tatcgcaaat	caaatacata	2880
acacacatag	taaaacaaaa	acaaattaca	agggtttaga	cgtttagtgg	caatgtgtaa	2940
atttgctcga	ggagctcggc	gcgccgaatt	cactggccgt	cgtttttaca	cgactcagga	3000
tcctgtcaaa	cactgatagt	ttaaactgaa	ggcgggaaac	gacaatctga	tcattgagcgg	3060
agaattaagg	gagtcacgtt	atgacccccg	ccgatgacgc	gggacaagcc	gtttttacgtt	3120
tggaaactgac	agaaccgcaa	cgttgaaagga	gccactcagc	cgcggttttc	tggagttaa	3180
tgagctaagc	acatacgtca	gaaaccatta	ttgcgcgttc	aaaagtgcgc	taaggctcact	3240
atcagctagc	aaatattttc	tgtcaaaaaa	gctccactga	cgttccataa	attcccctcg	3300
gtatccaatt	agagtctcat	attcactctc	aatccaaata	atctgcaccg	gatctggatc	3360
gtttcgcatg	attgaacaag	atggattgca	cgcaggttct	ccggccgctt	gggtggagag	3420
gctattcggc	tatgactggg	cacaacagac	aatcggtgc	tctgatgccg	ccgtgttccg	3480
gctgtcagcg	caggggccc	cggttctttt	tgtcaagacc	gacctgtccg	gtgccctgaa	3540
tgaactgcag	gacgaggcag	cgcggtctatc	gtggctggcc	acgacggggc	ttccttgccg	3600
agctgtgctc	gacgttgtca	ctgaagcggg	aaggactgg	ctgctattgg	gcgaagtgcc	3660
ggggcaggat	ctcctgtcat	ctcaccttgc	tcctgccgag	aaagtatcca	tcattggctga	3720
tgcaatgcgg	cggctgcata	cgtttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	3780