

(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  
30. Januar 2003 (30.01.2003)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 03/008596 A2

(51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/82, (74) Anwalt: WERNER, Frank; BASF Aktiengesellschaft,  
A01H 5/00, A01K 67/027 67056 Ludwigshafen (DE).

(21) Internationales Aktenzeichen: PCT/EP02/07485 (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,

(22) Internationales Anmeldedatum: 5. Juli 2002 (05.07.2002)  
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, NO, NZ, OM, PH, PL, PT, RO, RU,  
SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG,  
US, UZ, VN, YU, ZA, ZM, ZW.

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität: 101 33 398.6 13. Juli 2001 (13.07.2001) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von  
US): SUNGENE GMBH & CO. KGaA [DE/DE]; 06466  
Gatersleben (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): HEIM, Ute [DE/DE];  
Wasserstr. 1, 06466 Gatersleben (DE). HILLEBRAND,  
Helke [DE/DE]; B2, 11, 68159 Mannheim (DE). KUNZE,  
Irene [DE/DE]; Mühlenweg 11, 06466 Gatersleben (DE).  
HERBERS, Karin [DE/DE]; Am Hange 6, 06484  
Quedlinburg (DE).

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: EXPRESSION CASSETTES FOR TRANSGENICALLY EXPRESSING SELECTION MARKERS

(54) Bezeichnung: EXPRESSIONSKASSETTEN ZUR TRANSGENEN EXPRESSION VON SELEKTIONSMARKERN

(57) Abstract: The invention relates to expression cassettes and vectors, which contain selection marker genes while controlling the plant constitutive nitrilase 1 (NIT1) promoter from Arabidopsis thaliana, and to the use of these expression constructs or vectors for selecting transgenic organisms, preferably plants. The invention also relates to transgenic plants that are transformed by using these expression cassettes or vectors, to cultures derived therefrom, parts or multiplication products and to the use thereof for producing foodstuffs, feedstuffs, seeds, pharmaceuticals or fine chemicals.

(57) Zusammenfassung: Die Erfindung betrifft Expressionskassetten und Vektoren, die Selektionsmarker-Gene unter Kontrolle des pflanzlichen konstitutiven Nitrilase 1 (NIT1) Promotors aus Arabidopsis thaliana enthalten, sowie die Verwendung dieser Expressionskonstrukte oder Vektoren zur Selektion transgener Organismen, bevorzugt Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder Vermehrungsgut, sowie die Verwendung derselben zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien.



WO 03/008596 A2

Expressionskassetten zur transgenen Expression von Selektionsmarkern

## 5 Beschreibung

Die Erfindung betrifft Expressionskassetten und Vektoren, die Selektionsmarker-Gene unter Kontrolle des pflanzlichen konstitutiven Nitrilase 1 (NIT1) Promotors aus Arabidopsis thaliana enthalten, sowie die Verwendung dieser Expressionskonstrukte oder Vektoren zur Selektion transgener Organismen, bevorzugt Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder Vermehrungsgut, sowie die Verwendung derselben zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien.

Ziel biotechnologischer Arbeiten an Pflanzen ist die Herstellung von Pflanzen mit verbesserten Eigenschaften zum Beispiel zur Steigerung der landwirtschaftlichen Produktivität. Die Herstellung transgener Pflanzen ist eine grundlegende Technik der Pflanzenbiotechnologie und damit eine unerlässliche Voraussetzung für die pflanzliche Grundlagenforschung, sowie für die Herstellung von Pflanzen mit verbesserten, neuen Eigenschaften für die Landwirtschaft, zur Qualitätssteigerung bei Nahrungsmitteln oder zur Produktion bestimmter Chemikalien oder Pharmazeutika (Dunwell JM, J Exp Bot. 2000; 51 Spec No:487-96).

Eine Grundvoraussetzung für die transgene Expression bestimmter Gene in Pflanzen ist - aufgrund der geringen Effizienz der Transformation von Pflanzen - die Möglichkeit, transformierte von nicht transformierten Pflanzen zu unterscheiden. Dazu stehen dem Fachmann verschiedene sogenannte Selektionsmarker zur Verfügung. Die Expression dieser Gene verleiht einer entsprechend transformierten Pflanze eine Resistenz gegen Antibiotika (wie z.B. Kanamycin) oder Herbizide (wie z.B. Phosphinothricin oder Glyphosat). Eine wesentliche Voraussetzung, um diese Selektionsmarker-Gene zu exprimieren, ist die Bereitstellung pflanzen-spezifischer Promotoren.

Verschiedene pflanzliche Promotoren sind bekannt. Dabei kann zwischen konstitutiven Promotoren, die eine lokal und zeitlich nur wenig beschränkte Expression in verschiedenen Teilen einer Pflanze ermöglichen, und spezifischen Promotoren, die eine Expression nur in bestimmten Teilen oder Zellen einer Pflanze (z.B. Wurzel, Samen, Pollen, Blättern etc.) oder nur zu bestimmten Zeitpunkten der Entwicklung gestatten, unterschieden

## 2

werden. Konstitutive Promotoren sind für die Expression von Selektionsmarkern bevorzugt. Sie ermöglichen eine effiziente, frühe Selektion, meist schon im Zellkulturstadium (beispielsweise in einer Kalluskultur) was mit gewebe- oder entwicklungsabhängigen Promotoren nicht gewährleistet werden kann.

Konstitutive, in Pflanzen funktionelle Promotoren sind bislang relativ selten beschrieben. Zu nennen sind der TR-Doppelpromotor, die Promotoren der vakuolären ATPase Untereinheiten, der Ppcl Promotor aus *Mesembryanthemum crystallinum* (Cushman et al. (1993) *Plant Mol Biol* 21:561-566) oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991).

Die derzeit in Pflanzen überwiegend verwendeten konstitutiven Promotoren sind fast ausschließlich virale oder aus *Agrobacterium* isolierte Promotoren. Dabei handelt es sich im einzelnen um den Nopalinsynthase (*nos*) Promotor (Shaw et al. (1984) *Nucleic Acids Res.* 12(20):7831-7846), den Mannopinsynthase (*mas*) Promotor (Comai et al. (1990) *Plant Mol Biol* 15 (3):373-381) und den Octopinsynthase (*ocs*) Promotor (Leisner and Gelvin (1988) *Proc Natl Acad Sci USA* 85(5):2553-2557) aus *Agrobacterium tumefaciens* oder der CaMV35S Promotor aus dem Blumenkohl-Mosaikvirus. Letzterer ist der am häufigsten verwendete Promotor in Expressionssystemen mit ubiquitärer und andauernder Expression (Odell et al. (1985) *Nature* 313:810-812; Battraw and Hall (1990) *Plant Mol Biol* 15:527-538; Benfey et al. (1990) *EMBO J* 9(69):1677-1684; US 5,612,472). Der häufig als konstitutiver Promoter angewandte CaMV 35S Promotor zeigt jedoch Variationen in der Aktivität in unterschiedlichen Pflanzen und in unterschiedlichen Geweben derselben Pflanze (Atanassova et al. (1998) *Plant Mol Biol* 37:275-85; Battraw and Hall (1990) *Plant Mol Biol* 15:527-538; Holtorf et al. (1995) *Plant Mol Biol* 29:637-646; Jefferson et al. (1987) *EMBO J* 6:3901-3907). Ein weiterer Nachteil des 35S Promotors ist, dass bei einer Infektion mit dem Blumenkohlmosaikvirus und seinen typischen pathogenen Varianten die Expression des Transgens verändert wird. So sind Pflanzen, die das BAR Gen unter der Kontrolle des 35S Promotors exprimieren nicht mehr resistent nach Infektion mit dem Virus, der in der Natur typischerweise vorkommt (Al-Kaff et al. (2000) *Natur Biotechnology* 18:995-99).

Als Alternative zum CaMV 35S Promotor aus dem Bereich viraler Promotoren wurde der Sugarcane bacilliform badnavirus (ScBV) beschrieben (Schenk et al. (1999) *Plant Mol Biol* 39(6):1221-1230), der ein dem CaMV ähnliches Expressionsmuster vermittelt. Die Aktivität des ScBV Promotors wurde in transienten Expressionsanalysen mit verschiedenen zweikeimblättrigen Pflanzen, darunter

## 3

*Nicotiana tabacum* und *N. benthamiana*, Sonnenblume und Raps, und einkeimblättrigen Pflanzen, hier anhand von Banane, Mais und Hirse, analysiert. In den transienten Analysen in Mais war das ScBV Promotor vermittelte Expressionsniveau vergleichbar mit dem  
5 des Ubiquitin-Promotors aus Mais (s. unten). Weiterhin wurde die ScBV Promotor-vermittelte Expressionsrate in transgenen Bananen- und Tabakpflanzen getestet und zeigte in beiden Pflanzenspezies im wesentlichen konstitutive Expression.

10 Als gängige Promotoren zur Expression von Selektionsmarkern in Pflanzen werden vor allem der *nos*-Promotor, aber auch der *mas*- und *ocs*-Promotor verwendet, die allesamt aus *Agrobacterium*-Stämmen isoliert wurden.

15 Der Verwendung viraler Sequenzen wird seitens des Verbrauchers oft mit großem Vorbehalt begegnet. Diese Bedenken werden nicht zuletzt durch Untersuchungen genährt, die die Sicherheit des 35S CaMV Promotors - aufgrund eines möglichen horizontalen Gen-  
transfers bedingt durch einen Rekombinations-"Hot Spot" - in

20 Frage stellen (Ho MW et al. (1999) *Microbial Ecology in Health and Disease* 11:194-197; Cummins J et al. (2000) *Nature Biotechnology* 18:363). Ein Ziel künftiger biotechnologischer Arbeiten an Pflanzen ist daher der Ersatz viraler genetischer Elemente durch regulatorische pflanzliche Elemente, um so nah  
25 wie möglich am pflanzlichen System zu bleiben.

Aufgrund der herrschenden Bedenken hinsichtlich viraler Promotoren gibt es umfangreiche Bemühungen, diese durch pflanzliche Promotoren zu ersetzen. Ein den viralen Elementen  
30 vergleichbarer Promotor pflanzlicher Herkunft ist bislang jedoch noch nicht beschrieben worden.

Das vom Mais-Ubiquitin Promotor vermittelte Expressionsmuster wurde für die Ubi-1 and Ubi-2 Promotoren aus Mais beschrieben  
35 (Christensen et al. (1992) *Plant Mol Biol* 18(4):675-689). Während der Ubi-1 Promotor in Mais und anderen monokotylen Pflanzen gute Expressionsaktivität aufweist, zeigt er in dikotylen Tabakpflanzen nur 10 % der Aktivität, die in vergleichbaren Experimenten mit dem viralen 35S Promotor erzielt worden war.  
40 Ferner wurde gezeigt, dass der Mais Ubi-1 Promotor für die Überexpression von Genen in monokotylen Pflanzensystemen geeignet ist und darüber hinaus auch hinreichend stark ist, um eine Herbizid-Resistenz durch Expression von Selektionsmarkern zu vermitteln (Christensen and Quail (1996) *Transgenic Res* 5(3):213-218). Für  
45 dikotyle Expressionssysteme erwies sich der Ubi-1 Promotor als ungeeignet.

## 4

Ein Vergleich der Organspezifität und Stärke verschiedener konstitutiver Promotoren wurde von Holtorf (Holtorf et al. (1995) Plant Mol Biol 29(4):637-646) anhand stabil transformierter *Arabidopsis*-Pflanzen durchgeführt. Die Studie umfasste u.a.

5 den CaMV35S-Promotor, den blattspezifischen Thionin-Promotor aus Gerste und den *Arabidopsis* Ubiquitin-Promotor (UBQ1). Die höchste Expressionsrate zeigte der CaMV35S-Promotor. Anhand der Verwendung eines zusätzlichen translationalen Enhancers (TMV omega element) konnte die Expressionsrate des Promotors bei un-

10 veränderter Organspezifität um das Zwei- bis Dreifache gesteigert werden. Der blattspezifische Thionin-Promotor aus Gerste war in der Mehrzahl der transformierten Linien inaktiv, während der UBQ1 Promotor aus *Arabidopsis* mittlere Expressionsraten ergab. Entgegen den Befunden von Callis und Holtorf (Callis et al. (1990)

15 J Biol Chem 265:186- 12493; Holtorf S et al. (1995) Plant Mol Biol 29:637-747), ergaben eigene Untersuchungen, dass der *Arabidopsis* Ubiquitin Promotor für die Expression von Selektionsmarkergenen ungeeignet ist und seine allgemeine Anwendbarkeit aus diesem Grund in Frage gestellt werden muss (s. Vergleichs-

20 beispiel 1 und 3).

McElroy und Mitarbeiter berichteten von einem auf dem Reis Actin-1 (Act1) Promotor basierenden Konstrukt zur Transformation

25 monokotyle Pflanzen (McElroy et al. (1991) Mol Gen Genet 231:150-160). Insgesamt wurde aus den zuvor beschriebenen Untersuchungen geschlussfolgert, dass die auf dem Act1-Promotor basierenden Expressionsvektoren dazu geeignet sind, eine hinreichend starke und konstitutive Expression von Fremd-DNA in transformierten Zellen monokotyle Pflanzen zu steuern.

30 Als konstitutiver Promotor ist ferner der Promotor einer S-Adenosyl-L-Methionin Synthetase beschrieben (WO 00/37662). Nachteilig ist hier vor allem eine Abhängigkeit der Expressionsstärke von der Methioninkonzentration (siehe WO 00/37662;

35 Fig. 7).

WO 99/31258 beschreibt chimäre, konstitutive pflanzliche Promotoren, die sich aus verschiedenen Elementen verschiedener Promotoren mit komplementären Expressionsmustern zusammensetzen,

40 so dass die Kombination einzelner Gewebespezifitäten additiv zu einem konstitutiven Expressionsmuster führt.

Die Nitrilasen aus *Arabidopsis thaliana* sind in einer kleinen Genfamilie organisiert, von denen das Nitrilase 1 (Nit1) Gen

45 die vorherrschende Isoform darstellt. Nitrilasen sind Schlüsselenzyme der Auxinbiosynthese und katalysieren die Konversion von Indol-3-acetonitril zu Indol-3-essigsäure (Bartel und Fink (1994)

Proc Natl Acad Sci USA 91:6649-6653; Bartling et al. (1994) Proc Natl Acad Sci USA 91:6021-6025). Expressionsanalysen zeigten, dass die Nit1-mRNA vornehmlich in den Blättern der Rosetten, aber auch in den Wurzeln, Blüten, Sprossachsen und Schoten vor-

5 kommt. Eine Promotoranalyse mit einem Nit1-Promotor/GUS-Reporter Fusionskonstrukt zeigte in Tabak eine mit dem Hormonstoffwechsel koordinierte Expression (Hillebrand et al. (1996) Gene 170:197-200). Es wurde eine Blaufärbung in den Blattknospen und bei weiterer Blattentwicklung in den Spitzen der jungen Blättern

10 gezeigt. Blätter von 10 cm Länge zeigen eine homogene Verteilung des Farbstoffs. Schnitte durch Sprossachsen zeigten nur eine Blaufärbung, wenn sie sich in der Nähe sich entwickelnder Blätter und Knospen befanden. Dieses entwicklungsabhängige Muster zeigt sich auch in Wurzeln, wo nur die Wurzelspitzen und junge Wurzeln

15 gefärbt waren. Andere Arbeiten zeigten eine Expression der NIT1 mRNA in Blättern und Stengeln der reifen Pflanze sowie in der Wurzel älterer Pflanzen bzw. Wurzelspitze jüngerer Pflanzen. Die Expression korreliert mit der Lokalisation der Auxin-Biosynthese (Bartel B et al. (1994) Proc. Natl Acad Sci USA 91:6649-6653).

20 Aufgrund dieses gewebe- und entwicklungsabhängigen Expressionsmusters wäre der Promotor schlecht geeignet für Expression von Selektionsmarkern. Hier ist eine Selektion in einem sehr frühen Stadium wünschenswert. Vorteilhaft ist ferner eine effiziente Selektion von dedifferenzierten Zellen (beispielsweise in ver-

25 schiedenen Kallus-Phasen) aus Gewebekultur und weiteren Entwicklungsstadien, die für eine Gewebekultur geeignet sind. Zudem ist eine Expression in möglichst allen Gewebeteilen erforderlich, um eine effiziente Selektion zu gewährleisten.

30 Die im Stand der Technik beschriebenen, sogenannten "konstitutiven" Promotoren weisen einen oder mehrere der nachfolgenden Nachteile auf:

1. Mangelhafte Homogenität der Expression:
- 35 Die bekannten, sogenannten "konstitutiven" Promotoren zeigen vielfach eine unterschiedliche Expressionshöhe je nach Gewebe- oder Zelltyp. Zudem ist die Expressionseigenschaft oft stark von dem Insertionsort des Wirtsgenoms abhängig. Dies hat zur Folge, dass die durch heterologe Expression zu
- 40 erzielenden Effekte nicht in dem gleichen Ausmaß homogen in der Pflanze erreicht werden können. Es kann zu Unter- oder Überdosierungen kommen. Dies kann sich nachteilig auf das Pflanzenwachstum oder den Pflanzenwert auswirken.

## 6

2. Mangelhaftes Zeitprofil:  
Die im Stand der Technik bekannten sogenannten "konstitutiven" Promotoren zeigen oft einen zu späten Beginn ihrer Aktivität und/oder ein zu frühes Abflachen der Aktivität.  
5 Damit können zum Beispiel in der frühen Phase der Embryogenese keine wünschenswerten Effekte erzielt werden, was gerade hier - aufgrund der Empfindlichkeit des Embryos gegen biotische und abiotische Stressfaktoren - vorteilhaft wäre.
- 10 3. Mangelnde Anwendbarkeit auf viele Pflanzenarten:  
Die im Stand der Technik beschriebenen sogenannten "konstitutiven" Promotoren sind oft nicht in allen Arten in gleicher Weise aktiv.
- 15 4. Beim Vorliegen mehrerer Expressionskassetten mit jeweils dem gleichen "konstitutiven" Promotor in einem Organismus kann es zu Wechselwirkungen zwischen denselben und sogar zum Abschalten ("Gene Silencing") einzelner Expressionskassetten kommen (Mette et al. (1999) EMBO J. 18:241-248).
- 20 5. Promotoren viralen Ursprungs können durch Virusinfektionen der transgenen Pflanze beeinflusst werden und die gewünschte Eigenschaft dann nicht mehr ausprägen (Al-Kaff et al. (2000) Natur Biotechnology 18:995-99).
- 25 6. Die öffentliche Akzeptanz gegenüber der Verwendung von Promotoren und Elementen aus pflanzlichen Systemen ist höher als bei viralen Systemen.
- 30 7. Die Zahl der für die Expression von Selektionsmarkern in Pflanzen geeigneten Promotoren ist gering und sie sind gewöhnlicher Weise viralen oder bakteriellen Ursprungs.
- 35 8. Pollen/Antheren-Expression: Viele der genannten Promotoren (wie beispielsweise 35S CaMV) zeigen eine starke Aktivität im Pollen oder den Staubbeuteln (Antheren). Dies kann unvorteilhafte Auswirkungen auf die Umwelt haben. So hatte eine unspezifische Expression von *Bacillus thuringiensis* Endotoxinen nicht nur den gewünschten Effekt auf Fraßinsekten durch Expression in der Wurzel, sondern auch infolge einer  
40 Expression im Pollen bedeutende Schäden im Bestand des Monarchfalters zur Folge, der den Pollen als Hauptnahrungsquelle nutzt (Losey JE et al. (1999) Nature 399, 214).

## 7

Ein idealer konstitutiver Promotor zur Expression von Selektionsmarkern sollte möglichst zahlreiche der nachfolgenden Eigenschaften haben:

- 5 a) Möglichst homogene lokale und zeitliche Genexpression d.h. eine Expression in möglichst vielen Zelltypen oder Geweben eines Organismus während der verschiedenen Phasen des Entwicklungszyklus. Vorteilhaft ist ferner eine effiziente Selektion von dedifferenzierten Zellen (beispielsweise in  
10 verschiedenen Kallus-Phasen) aus Gewebekultur und weiteren Entwicklungsstadien, die für eine Gewebekultur geeignet sind.
- b) Möglichst breite Anwendbarkeit auf verschiedene Pflanzenarten bzw. Anwendbarkeit auf Arten, in denen mit den bisher  
15 bekannten "konstitutiven" Promotoren keine Expression erreicht werden kann.
- c) Um eine Kombination von mehreren Transgenen in einer Pflanze zu realisieren, ist es wünschenswert, mehrere  
20 Transformationen hintereinander zu schalten oder ein Konstrukt mit mehreren Promotor-Kassetten zu verwenden, ohne dass durch die mehrmalige Verwendung identischer regulatorischer Sequenzen Silencing-Effekte erzeugt werden...
- 25 d) Einen pflanzlichen Ursprung zur Vermeidung von Akzeptanzproblemen beim Konsumenten und eventuellen zukünftigen Zulassungsproblemen.
- e) Nebenaktivitäten eines Promotors in den Antheren/Pollen sind  
30 nicht wünschenswert, zum Beispiel um Umweltschäden zu vermeiden (s.o.).

Die der vorliegenden Erfindung zugrunde liegende Aufgabe bestand also in der Bereitstellung von pflanzlichen regulatorischen  
35 Sequenzen, die möglichst viele der oben genannten Eigenschaften erfüllen, vor allem eine ubiquitäre und entwicklungsunabhängige (konstitutive) Expression eines nachgeschalteten Selektionsmarker-Gens vermitteln, die ausreicht, um im Rahmen der Transformation von Pflanzen Selektionsmarkergene hinreichend hoch  
40 zu exprimieren, um die transformierten Zellen aufgrund des per Transformation erzielten Selektionsvorteils von den nicht transformierten Zellen unterscheiden zu können. Trotz verschiedener pflanzlicher Promotoren für die eine konstitutive Expression zumindest in einzelnen Arten beansprucht wird, wurde bislang  
45 kein Promotor mit den oben aufgezählten gewünschten Eigen-

## 8

schaften beschrieben. Es bestand daher die Aufgabe, entsprechende Promotoren zu identifizieren.

Diese Aufgabe wurde durch Bereitstellung von Expressions-  
5 kassetten, in denen Selektionsmarker unter der Kontrolle des Nitrilase 1 (infolge NIT1) Promotors aus *Arabidopsis thaliana* stehen, gelöst.

Die Sequenz des NIT1 Promotors entspricht größtenteils der  
10 Sequenz von Position 3 bis 1879 der Sequenz des NIT1 Gens in der Datenbank unter Genbank Acc.-No: X86454 bzw. einer in der GenBank unter der Acc.-No. Y07648 hinterlegten Sequenz (Version Y07648.2 vom 20.12.1999; Basenpaar 2456 bis 4340; das Gen beginnend ab bp 4345 ist mit "nitrilase 1" annotiert).  
15 Unter Verwendung von Glucuronidase-Reporter-Versuchen wurde eine Promotoraktivität - vergleichbar intensiv wie bei entsprechenden mit dem 35S Promotor/GUS-Konstrukten transformierten Pflanzen - in zahlreichen Geweben detektiert. Im Gegensatz zum 35S Promotor konnte keine Blaufärbung in den Antheren/Pollen nachgewiesen  
20 werden. Entgegen dem aus den Literaturbefunden abgeleiteten Vorbehalt gegen eine Eignung des Promotors zur effektiven Expression von Selektionsmarkern (zum Beispiel aufgrund der demonstrierten gewebe- und entwicklungsabhängigen Expression), konnte eine hocheffiziente Selektion durch Kombination mit  
25 beispielsweise dem Phosphinothricin-Resistenzgen (bar/pat) demonstriert werden.

Weiterhin wurden im Rahmen der genannten Arbeiten der *Arabidopsis thaliana* Ubiquitin-Promotor (Holtorf et al. (1995) Plant Mol  
30 Biol 29:637-646) und der Squalen Synthase-Promotor (Kribii et al. (1997) Eur J Biochem 249:61-69) untersucht, die aber beide überraschenderweise nicht geeignet waren, Selektionsmarkergen-Expression zu vermitteln, obwohl die Literaturdaten der Ubiquitin-Promotoren aus Monokotylen (siehe oben) hatten vermuten  
35 lassen, dass insbesondere der Ubiquitin-Promotor einer dikotylen Pflanze als Promotor eines Selektionsmarkersystems hätte funktionieren müssen (siehe Vergleichsbeispiele 1 und 3). Ähnliches gilt für den Squalensynthase-Promotor, dessen Charakterisierung hätte erwarten lassen, dass ausreichend hohe Expressions-  
40 raten für die erfolgreiche Steuerung eines Selektionsmarkergens erzielt werden können (Del Arco and Boronat (1999) 4th European Symposium on Plant Isoprenoids, 21.-23.4.1999, Barcelona, Spanien) (siehe Vergleichsbeispiele 2 und 3).

## 9

Ein erster Gegenstand der Erfindung betrifft daher transgene Expressionskassetten zur Expression von Nukleinsäuresequenzen kodierend für einen Selektionsmarker enthaltend in 5'-3' Richtung

- 5 a) einen Promotor gemäß SEQ ID NO: 1 oder ein funktionelles Äquivalent oder äquivalentes Fragment desselben, das im wesentlichen die gleiche Promotoraktivität wie der Promotor gemäß SEQ ID NO: 1 besitzt, und
- 10 b) eine transgen zu exprimierende Nukleinsäuresequenz kodierend für einen Selektionsmarker, und
- c) eine in pflanzlichen Zellen oder pflanzlichen Organismen funktionelle Terminatorsequenz,

15

wobei a) oder b) funktionell miteinander so verknüpft sind, dass eine transgene Expression der Nukleinsäuresequenz kodierend für einen Selektionsmarker in einer pflanzlichen Zelle oder einem pflanzlichen Organismus ermöglicht wird.

20

Ferner betrifft die Erfindung Verfahren zur Selektion transformierter pflanzlicher Zellen oder pflanzlicher Organismen, wobei eine transgene Expressionskassette enthaltend in 5'-3' Richtung

25

- a) einen Promotor gemäß SEQ ID NO: 1, oder ein funktionelles Äquivalent oder äquivalentes Fragment desselben, das im wesentlichen die gleiche Promotoraktivität wie der Promotor gemäß SEQ ID NO: 1 besitzt, und

30

- b) eine transgen zu exprimierende Nukleinsäuresequenz kodierend für einen Selektionsmarker, und

- c) eine in Pflanzen funktionelle Terminatorsequenz,

35

wobei a) oder b) funktionell miteinander so verknüpft sind, dass eine transgene Expression der Nukleinsäuresequenz kodierend für einen Selektionsmarker in einer pflanzlichen Zelle oder pflanzlichen Organismen ermöglicht wird, in besagte pflanzliche Zelle oder besagten pflanzlichen Organismus eingebracht wird, und die Expression des Selektionsmarkers erfolgt und eine Selektion der transformierten pflanzlichen Zellen oder pflanzlichen Organismen ausgeübt wird.

45

## 10

"Expression" umfasst die Transkription und/oder die Translation der transkribierten RNA der transgen zu exprimierenden Nukleinsäuresequenz kodierend für einen Selektionsmarker in ein korrespondierendes Polypeptid.

5

"Transgen" meint in Bezug auf eine Expressionskassette zur transgenen Expression von Nukleinsäuren kodierend für einen Selektionsmarker oder die Verknüpfung von NIT1-Promotor und Nukleinsäuren kodierend für einen Selektionsmarker alle solche

10 durch gentechnische Methoden zustande gekommene Konstruktionen oder Verfahren, in denen sich entweder

a) der NIT1 Promotor gemäß SEQ ID NO: 1 oder ein funktionelles Äquivalent oder äquivalente Fragment derselben, oder

15

b) die zu exprimierende Nukleinsäuresequenz kodierend für einen Selektionsmarker, oder

c) (a) und (b)

20

sich nicht in ihrer natürlichen, genetischen Umgebung (d.h. an ihrem natürlichen chromosomalen Locus) befinden oder durch gentechnische Methoden modifiziert wurden, wobei die Modifikation beispielhaft eine Substitution, Addition, Deletion, Inversion

25 oder Insertion eines oder mehrerer Nukleotidreste sein kann.

Die erfindungsgemäßen Expressionskassetten, von ihnen abgeleitete Vektoren oder die erfindungsgemäßen Verfahren können funktionelle Äquivalente zu den unter SEQ ID NO: 1 beschriebenen NIT1

30 Promotorsequenz umfassen. Funktionell äquivalente Sequenzen umfassen auch all die Sequenzen, die von dem komplementären Gegenstrang den durch SEQ ID NO:1 definierten Sequenz abgeleitet sind und im wesentlichen die gleiche Promotoraktivität aufweisen.

35 Funktionelle Äquivalente in bezug auf den NIT1 Promotor meint insbesondere natürliche oder künstliche Mutationen der unter SEQ ID NO: 1 beschriebenen NIT1 Promotorsequenz sowie seine Homologen aus anderen Pflanzengattungen und -arten, welche weiterhin im wesentlichen die gleiche Promotoraktivität aufweisen.

40

Eine Promotoraktivität wird im wesentlichen als gleich bezeichnet, wenn die Transkription eines bestimmten zu exprimierenden Gens unter Kontrolle eines bestimmten von SEQ ID NO: 1 abgeleiteten Promotors unter ansonsten unveränderten Bedingungen eine

45 Lokalisation innerhalb der Pflanze aufweist, die zu mindestens 50 %, bevorzugt mindestens 70 %, besonders bevorzugt mindestens 90 % ganz besonders bevorzugt mindestens 95 % deckungsgleich

## 11

ist mit einer Vergleichsexpression erhalten unter Verwendung eines des durch SEQ ID NO: 1 beschriebenen NIT1 Promotors. Dabei kann die Expressionshöhe sowohl nach unten als auch nach oben im Vergleich zu einem Vergleichswert abweichen. Bevorzugt sind dabei solche Sequenzen, deren Expressionshöhe, gemessen anhand der transkribierten mRNA oder dem infolge translatierten Protein, unter ansonsten unveränderten Bedingungen quantitativ um nicht mehr als 50 %, bevorzugt 25 %, besonders bevorzugt 10 % von einem Vergleichswert erhalten mit dem durch SEQ ID NO: 1 beschriebenen NIT1 Promotor unterscheidet. Besonders bevorzugt sind solche Sequenzen, deren Expressionshöhe, gemessen anhand der transkribierten mRNA oder dem infolge translatierten Protein, unter ansonsten unveränderten Bedingungen quantitativ um mehr als 50 %, bevorzugt 100 %, besonders bevorzugt 500 %, ganz besonders bevorzugt 1000 % einen Vergleichswert erhalten mit dem durch SEQ ID NO:1 beschriebenen NIT1 Promotor übersteigt. Bevorzugt ist als Vergleichswert die Expressionshöhe der natürliche mRNA des jeweiligen Gens oder des natürlichen Genproduktes. Bevorzugt ist ferner als Vergleichswert die Expressionshöhe erhalten mit einer beliebigen, aber bestimmten Nukleinsäuresequenz, bevorzugt solchen Nukleinsäuresequenzen, die für leicht quantifizierbare Proteine kodieren. Ganz besonders bevorzugt sind dabei Reporter-Proteine (Schenborn E; Groskreutz D (1999) Mol Biotechnol. 13(1):29-44) wie das "green fluorescence protein" (GFP) (Chui WL et al. (1996) Curr Biol 6:325-330; Leffel SM et al. (1997) Bio-techniques 23(5):912-918), die Chloramphenicoltransferase, eine Luziferase (Millar et al. (1992) Plant Mol Biol Rep 10:324-414) oder die  $\beta$ -Galactosidase, ganz besonders bevorzugt ist die  $\beta$ -Glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).

Ansonsten unveränderte Bedingungen bedeutet, dass die Expression, die durch eine der zu vergleichenden Expressionskassetten initiiert wird, nicht durch Kombination mit zusätzlichen genetischen Kontrollsequenzen, zum Beispiel Enhancer-Sequenzen, modifiziert wird. Unveränderte Bedingungen bedeutet ferner, dass alle Rahmenbedingungen wie beispielsweise Pflanzenart, Entwicklungsstadium der Pflanzen, Zuchtbedingungen, Assaybedingungen (wie Puffer, Temperatur, Substrate etc.) zwischen den zu vergleichenden Expressionen identisch gehalten werden.

Mutationen umfassen Substitutionen, Additionen, Deletionen, Inversionen oder Insertionen eines oder mehrerer Nukleotide. Somit werden beispielsweise auch solche Nukleotidsequenzen durch die vorliegende Erfindung mit umfasst, welche man durch Modifikation des NIT1 Promotors gemäß SEQ ID NO: 1 erhält. Ziel einer solchen Modifikation kann die weitere Eingrenzung der darin enthaltenen Sequenz oder z.B. auch die Einfügung weiterer

## 12

Restriktionsenzymchnittstellen, die Entfernung überflüssiger DNA oder das Hinzufügen weiterer Sequenzen, zum Beispiel weiterer regulatorischer Sequenzen, sein.

- 5 Wo Insertionen, Deletionen oder Substitutionen, wie z.B. Transitionen und Transversionen, in Frage kommen, können an sich bekannte Techniken, wie in vitro-Mutagenese, "primer repair", Restriktion oder Ligation verwendet werden. Durch Manipulationen, wie z.B. Restriktion, "chewing-back" oder Auffüllen von Über-
- 10 hängen für "blunt ends" können komplementäre Enden der Fragmente für die Ligation zur Verfügung gestellt werden. Zu analogen Ergebnissen kann man auch unter Verwendung der Polymerasekettenreaktion (PCR) unter Verwendung spezifischer Oligonukleotid-Primer kommen.

## 15

- Unter Homologie zwischen zwei Nukleinsäuren 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) unter Einstellung
- 20 folgender Parameter berechnet wird:

Gap Weight: 12

Length Weight: 4

25 Average Match: 2,912

Average Mismatch:-2,003

- Beispielhaft wird unter einer Sequenz, die eine Homologie von mindestens 50 % auf Nukleinsäurebasis mit der Sequenz gemäß SEQ ID NO. 1 aufweist, eine Sequenz verstanden, die bei einem
- 30 Vergleich mit der Sequenz SEQ ID NO. 1 nach obigem Programmalgorithmus mit obigem Parametersatz eine Homologie von mindestens 50 % aufweist.

- Funktionelle Homologe zu dem NIT1 Promotor gemäß SEQ ID NO: 1
- 35 abgeleitet zum Beispiel durch Substitution, Insertion oder Deletion von Nukleotiden umfasst bevorzugt solche Sequenzen, die eine Homologie aufweisen von mindestens 50 %, bevorzugt 70 %, vorzugsweise mindestens 80 %, besonders bevorzugt mindestens 90 %, ganz besonders bevorzugt mindestens 95 %, am meisten
- 40 bevorzugt 99 % über eine Länge von von mindestens 100 Basenpaaren, bevorzugt mindestens 200 Basenpaaren, besonders bevorzugt von mindestens 300 Basenpaaren, ganz besonders bevorzugt von mindestens 400 Basenpaaren, am meistens bevorzugt von mindestens 500 Basenpaaren.

## 45

## 13

Weitere Beispiele für die in den erfindungsgemäßen Expressionskassetten oder Vektoren zum Einsatz kommenden Promotorsequenzen lassen sich beispielsweise in verschiedenen Organismen, deren genomische Sequenz bekannt ist, wie beispielsweise aus Arabidopsis thaliana, Brassica napus, Nicotiana tabacum, Solanum tuberosum, Helianthium anuus, Linum sativum durch Homologievergleiche in Datenbanken leicht auffinden. Für die Suche kann sowohl die Promotorsequenz des Nit1 Promotors gemäß SEQ ID NO:1 verwendet werden, also auch die kodierende Sequenz des NIT1 Gens (GenBank Acc.-No.: X86454).

Funktionelle Äquivalente meint ferner DNA Sequenzen, die unter Standardbedingungen mit der Nukleinsäuresequenz kodierend für den NIT1 Promotor gemäß SEQ ID NO:1 oder der zu ihr komplementären Nukleinsäuresequenzen hybridisieren und die im wesentlichen gleichen Eigenschaften haben. Standardhybridisierungsbedingungen ist breit zu verstehen und meint somit 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, S. 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, pH7,0).

Darüber hinaus 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 Waschschrilles sind infolge gegeben:

(1) Hybridisierungsbedingungen mit zum Beispiel

a) 4X SSC bei 65°C, oder

b) 6X SSC, 0,5 % SDS, 100 µg/ml denaturierte, fragmentierte Lachssperma-DNA bei 65°C, oder

## 14

- c) 4X SSC, 50 % Formamid, bei 42°C, oder
- d) 6X SSC, 0,5 % SDS, 100 µg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder
- 5 e) 2X oder 4X SSC bei 50°C (schwach stringente Bedingung), oder
- f) 2X oder 4X SSC, 30 bis 40 % Formamid bei 42°C (schwach stringente Bedingung).
- 10 g) 6x SSC bei 45°C, oder,
- h) 50 % Formamid, 4x SSC bei 42°C, oder
- 15 i) 50 % (vol/vol) Formamid, 0,1 % Rinderserumalbumin, 0,1 % Ficoll, 0,1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6,5, 750 mM NaCl, 75 mM Natriumcitrate bei 42°C, oder
- 20 j) 0,05 M Natriumphosphatpuffer pH 7,0, 2 mM EDTA, 1 % BSA und 7 % SDS.
- (2) Waschschritte mit zum Beispiel
- 25 a) 0,1X SSC bei 65°C, oder
- b) 0,1X SSC, 0,5 % SDS bei 68°C, oder
- 30 c) 0,1X SSC, 0,5 % SDS, 50 % Formamid bei 42°C, oder
- d) 0,2X SSC, 0,1 % SDS bei 42°C, oder
- e) 2X SSC bei 65°C (schwach stringente Bedingung).
- 35 f) 40 mM Natriumphosphatpuffer pH 7,0, 1 % SDS, 2mM EDTA.

Verfahren zur Herstellung erfindungsgemäßer funktioneller Äquivalente umfasst bevorzugt die Einführung von Mutationen

40 in den NIT1 Promotor gemäß SEQ ID NO: 1. Eine Mutagenese kann ungerichtet ("random") erfolgen, wobei die mutagenisierten Sequenzen anschließend bezüglich ihrer Eigenschaften nach einer "trial-by-error" Prozedur durchmustert werden. Besonders vorteilhafte Selektionskriterien umfassen beispielsweise eine erhöhte

45 Resistenz gegenüber einem Antibiotikum, Herbizid oder Biozid, gegen das der zu exprimierende Selektionsmarker Resistenz verleiht, oder die Höhe der resultierenden Expression der ein-

## 15

geführten Nukleinsäuresequenz kodierend für den Selektionsmarker. Letztere kann auf Transkriptionsebene (beispielsweise durch Northern-Blot) oder Translationsebene (beispielsweise durch Immuno/Western-Blot) in der dem Fachmann geläufigen Weise  
5 analysiert werden.

Alternativ können nicht-essentielle Sequenzen des erfindungs-  
gemäßen Promotors deletiert werden ohne die genannten Eigenschaf-  
ten signifikant zu beeinträchtigen. Derartige Deletionsvarianten  
10 stellen funktionell äquivalente Teile des Promotors beschrieben  
durch SEQ ID NO: 1 dar.

Die Eingrenzung der Promotorsequenz auf bestimmte, essentielle  
regulatorische Regionen kann auch mit Hilfe von Suchroutine  
15 zur Suche von Promotorelementen vorgenommen werden. Oft sind  
in den für die Promotoraktivität relevanten Regionen bestimmte  
Promotorelemente gehäuft vorhanden. Diese Analyse kann beispiele-  
weise mit Computerprogrammen wie dem Programm PLACE ("Plant Cis-  
acting Regulatory DNA Elements"; Higo K et al. (1999) Nucleic  
20 Acids Research 27(1):297-300) oder der BIOBASE Datenbank "Trans-  
fac" (Biologische Datenbanken GmbH, Braunschweig) vorgenommen  
werden.

Auf diese Weise, wie auch beispielsweise durch sukzessive Verkür-  
25 zung und Analyse von Promotorsequenzen, kann der Fachmann zu  
äquivalenten Fragmenten des Promotor gemäß SEQ ID NO: 1 oder ei-  
nes funktionellen Äquivalentes desselben gelangen. Derartige Se-  
quenzen sind aufgrund ihrer reduzierten Sequenzlänge leichter zu  
handhaben und daher bevorzugt. Beispielhaft für derartige ver-  
30 kürzte Sequenzen sind die durch SEQ ID NO: 18 und 19 beschriebe-  
nen Sequenzen. Diese zeichnen sich durch im wesentlichen gleiche  
Promotoreigenschaften wie der Promotor gemäß SEQ ID NO: 1 aus.

Verfahren zur Mutagenisierung von Nukleinsäuresequenzen sind dem  
35 Fachmann bekannt und schließen beispielhaft die Verwendung von  
Oligonukleotiden mit einer oder mehr Mutationen im Vergleich  
zu der zu mutierenden Region ein (z.B. im Rahmen einer "Site-  
specific mutagenesis"). Typischerweise kommen Primer mit ungefähr  
15 bis ungefähr 75 Nukleotiden oder mehr zum Einsatz, wobei  
40 bevorzugt ca. 10 bis ca. 25 oder mehr Nukleotidreste an beiden  
Seiten der zu verändernden Sequenz lokalisiert sind. Details  
und Durchführung besagter Mutageneseverfahren sind dem Fachmann  
geläufig (Kunkel et al. (1987) Methods Enzymol, 154:367-382;  
Tomic et al. (1990) Nucl Acids Res 12:1656; Upender, Raj, Weir  
45 (1995) Biotechniques 18(1):29-30; US 4,237,224). Eine Mutagenese  
kann auch durch Behandlung von beispielsweise Vektoren, die

## 16

eine der erfindungsgemäßen Nukleinsäuresequenzen enthalten, mit mutagenisierenden Agentien wie Hydroxylamin realisiert werden.

- Die in den erfindungsgemäßen Expressionskassetten enthaltenen
- 5 transgen zu exprimierenden Nukleinsäuresequenzen kodierend für Selektionsmarker können mit weiteren genetischen Kontrollsequenzen neben einem der erfindungsgemäßen Promotor oder seinem Äquivalent oder äquivalenten Teil funktionell verknüpft sein.
- 10 Unter einer funktionellen Verknüpfung versteht man zum Beispiel die sequentielle Anordnung des Promotors, der transgen zu exprimierenden Nukleinsäuresequenz kodierend für einen Selektionsmarker und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator derart, dass jedes der regulativen Elemente
- 15 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
- 20 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 Promotor fungierenden Sequenz
- 25 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.
- 30 Die Herstellung einer funktionellen Verknüpfung 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
- 35 Spring Harbor Laboratory, Cold Spring Harbor, NY sowie in Silhavy TJ, Berman ML und Enquist LW(1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY und in Ausubel FM et al.(1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience be-
- 40 schrieben 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.

## 17

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  
5 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  
10 Nukleinsäuresequenz einen der erfindungsgemäßen Promotoren 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.

15

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

Es können ferner weitere Promotoren funktionell mit der zu exprimierenden Nukleinsäuresequenz verknüpft sein, die eine Expression in weiteren Pflanzengeweben oder in anderen  
30 Organismen, wie zum Beispiel *E.coli* Bakterien ermöglichen.

Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Region, Introns oder die nichtkodierende 3'-Region von Genen, bevorzugt Nit-Gens. Es ist gezeigt worden, dass diese eine  
35 signifikante Funktionen bei der Regulation der Genexpression spielen können. So wurde gezeigt, dass 5'-untranslatierte Sequenzen die transiente Expression heterologer Gene verstärken können. Sie können ferner die Gewebespezifität fördern (Rouster J et al.(1998) Plant J. 15:435-440). Umgekehrt unterdrückt die  
40 5'-untranslatierte Region des opaque-2 Gens die Expression. Eine Deletion der entsprechenden Region führt zu einer Erhöhung der Genaktivität (Lohmer S et al. (1993) Plant Cell 5:65-73). Die unter SEQ ID NO: 1 angegebene Nukleinsäuresequenz enthält den Abschnitt des NIT1-Gens, der den Promotor und die 5'-untrans-  
45 latierte Region bis vor das ATG-Startcodon des NIT1 Proteins aus *Arabidopsis thaliana* repräsentiert.

## 18

- McElroy und Mitarbeiter (McElroy et al. (1991) Mol Gen Genet 231(1):150-160) berichteten von einem auf dem Reis Actin 1 (Act1) Promotor basierenden Konstrukt zur Transformation monokotyle Pflanz. In transgenen Reiszellen führte die Verwendung des
- 5 Act1-Introns in Kombination mit dem 35S-Promotor zu einer gegenüber dem isolierten 35S-Promotor um das Zehnfache gesteigerten Expressionsrate. Eine Optimierung der Sequenzumgebung der Translationsinitiationsstelle des Reporter-gens (GUS)
- 10 in transformierten Reiszellen. Eine Kombination der optimierten Translations-Initiationsstelle und des Act1-Introns resultierte in einer 40-fachen Steigerung der GUS-Expression durch den CaMV35S-Promotor in transformierten Reiszellen; ähnliche Ergebnisse wurden anhand von transformierten Maiszellen erzielt.
- 15 Insgesamt wurde aus den zuvor beschriebenen Untersuchungen geschlossen, dass die auf dem Act1-Promotor basierenden Expressionsvektoren dazu geeignet sind, eine hinreichend starke und konstitutive Expression von Fremd-DNA in transformierten Zellen monokotyle Pflanz zu steuern.
- 20 Die Expressionskassette kann vorteilhafterweise eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte transgene Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der transgen zu
- 25 exprimierenden Nukleinsäuresequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden, wie weitere regulatorische Elemente oder Terminatoren. Die transgen zu exprimierenden Nukleinsäuresequenzen können in einer oder mehreren Kopien in der Expressionskassette enthalten sein.
- 30 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
- 35 der natürliche Promotor eines bestimmten Gens gegen einen der erfindungsgemäßen Promotor 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.
- 40 14(4):381-92). Hier werden bestimmte flankierende Sequenzen dem Zielgen angefügt (lox-Sequenzen), die später eine Entfernung mittels der cre-Rekombinase ermöglichen. Die Rekombinase-Technik kann auch zur Erhöhung der Effizienz der homologen Rekombination verwendet werden. Entsprechende Verfahren sind dem Fachmann
- 45 bekannt (US 6,110,736). Eine Vielzahl von sequenzspezifischen Rekombinationssystemen kann verwendet werden, beispielhaft sind das Cre/lox-System des Bacteriophagen P1, das FLP/FRT System der

## 19

Hefe, die Gin Recombinase des Mu Phagen, die Pin Recombinase 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 5 pflanzlichen Systemen angewendet (Odell et al. (1990) Mol Gen Genet 223:369-378)

Als Kontrollsequenzen geeignete Polyadenylierungssignale sind in Pflanzen funktionelle Polyadenylierungssignale sowie - vor- 10 zugsweise - solche, die im wesentlichen T-DNA Polyadenylierungssignalen aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACHS entsprechen (Gielen et al. (1984) EMBO J 3:835ff) oder funktionelle Äquivalente davon.

15

In einer besonders bevorzugten Ausführungsform enthält die Expressionskassette eine in Pflanzen funktionelle Terminatorsequenz. In Pflanzen funktionelle Terminatorsequenzen meint allgemein solche Sequenzen, die in der Lage sind, in Pflanzen den 20 Abbruch der Transkription einer DNA-Sequenz zu bewirken. Beispiele für geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopalin-Synthase)-Terminator. Besonders bevorzugt sind jedoch pflanzliche Terminatorsequenzen. Pflanzliche Terminatorsequenzen meint allgemein solche Sequenzen, 25 die Bestandteil eines natürlichen pflanzlichen Gens sind. Besonders bevorzugt sind dabei der Terminators des Cathepsin D Inhibitor Gens aus Kartoffel (GenBank Acc. No.: X74985; Terminator: SEQ ID NO: 16) oder des Terminators der Speicherproteingens VfLE1B3 (GenBank Acc. No.: Z26489; Terminator: SEQ ID NO: 17) aus 30 der Ackerbohne. Diese Terminatoren sind den im Stand der Technik beschriebenen viralen oder T-DNA Terminatoren mindestens gleichwertig. Das Plasmid pSUN5NPTIICat (SEQ ID NO: 15) enthält den pflanzlichen Terminator des Cathepsin D Inhibitor Gens aus Kartoffel.

35

Selektionsmarker umfasst sowohl positive Selektionsmarker, die eine Resistenz gegen ein Antibiotikum, Herbizid oder Biozid verleihen, als auch negative Selektionsmarker, die eine Sensitivität gegen eben diese verleihen, als auch Marker die 40 dem transformierten Organismus einen Wachstumsvorteil gewähren (beispielsweise durch Expression von Schlüsselgenen der Cytokinbiosynthese; Ebinuma H et al. (2000) Proc Natl Acad Sci USA 94:2117-2121). Bei der positiven Selektion gedeihen nur die Organismen, die den entsprechenden Selektionsmarker exprimieren, 45 während bei der negativen Selektion eben diese eingehen. Bei der Herstellung transgener Pflanzen ist die Verwendung eines positiven Selektionsmarkers bevorzugt. Bevorzugt ist ferner

## 20

die Verwendung von Selektionsmarkern, die Wachstumsvorteile verleihen. Negative Selektionsmarker können vorteilhaft verwendet werden, wenn es darum geht, bestimmte Gene oder Genomabschnitte aus einem Organismus zu entfernen (beispielsweise im Rahmen eines Kreuzungsprozesses).

Der mit der Expressionskassette eingebrachte selektionierbare Marker verleiht den erfolgreich rekombinierten oder transformierten Zellen eine Resistenz gegen ein Biozid (zum Beispiel ein Herbizid wie Phosphinothricin, Glyphosat, Imidazolinon- oder Sulfonylharnstoff-Herbizide oder Bromoxynil), einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456) oder ein Antibiotikum, wie zum Beispiel Kanamycin, G 418, Bleomycin, Hygromycin. Der Selektionsmarker erlaubt die Selektion der transformierten Zellen von untransformierten (McCormick et al., Plant Cell Reports 5 (1986), 81-84). Besonders bevorzugte Selektionsmarker sind solche die eine Resistenz gegen Herbizide verleihen. Dem Fachmann sind zahlreiche derartiger Selektionsmarker und die für diese kodierenden Sequenzen bekannt. Nachfolgend seien beispielhaft jedoch nicht einschränkend zu nennen:

i) Positive Selektionsmarker:

Der mit der Expressionskassette eingebrachte selektionierbare Marker verleiht den erfolgreich rekombinierten oder transformierten Zellen eine Resistenz gegen ein Biozid (zum Beispiel ein Herbizid wie Phosphinothricin, Glyphosat oder Bromoxynil), einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456) oder ein Antibiotikum, wie zum Beispiel Tetracycline, Ampicillin, Kanamycin, G 418, Neomycin, Bleomycin oder Hygromycin. Der Selektionsmarker erlaubt die Selektion der transformierten Zellen von untransformierten (McCormick et al., Plant Cell Reports 5 (1986), 81-84). Besonders bevorzugte Selektionsmarker sind solche, die eine Resistenz gegen Herbizide verleihen. Beispielhaft als Selektionsmarker seien genannt:

- DNA Sequenzen, die für Phosphinothricinacetyltransferasen (PAT) kodieren, welche die freie Aminogruppe des Glutaminsynthaseinhibitors Phosphinothricin (PPT) acetylieren und damit eine Detoxifizierung des PPT erreichen (de Block et al. 1987, EMBO J. 6, 2513-2518) (auch Bialophos® resistenzgen (bar) genannt). Das bar Gen kodierend für eine Phosphinothricinacetyltransferase (PAT) kann aus beispielsweise Streptomyces hygroscopicus oder S. viridochromogenes isoliert werden. Entsprechende Sequenzen sind dem Fachmann bekannt (aus Streptomyces hygroscopicus GenBank Acc.-No.: X17220 und X05822, aus Streptomyces viridochromogenes GenBank

## 21

Acc.-No.: M 22827 und X65195; US 5,489,520). Ferner sind synthetische Gene beispielsweise für die Expression in Plastiden beschrieben. Ein synthetisches Pat Gen ist beschrieben in Becker et al. (1994) *The Plant J.* 5:299-307.

5 Ganz besonders bevorzugt ist ebenfalls die Expression des Polypeptides gemäß SEQ ID NO: 3 beispielsweise kodiert durch eine Nukleinsäuresequenz gemäß SEQ ID NO: 2. Die Gene verleihen Resistenz gegen das Herbizid Bialaphos oder Glufosinat und sind vielbenutzer Marker in transgenen Pflanzen (Vickers

10 JE et al. (1996). *Plant Mol Biol Reporter* 14:363-368; Thompson CJ et al. (1987) *EMBO J* 6:2519-2523).

- 5-Enolpyruvylshikimat-3-phosphatsynthasegene (EPSP Synthasegene), die eine Resistenz gegen Glyphosat (N-(phosphonomethyl)glycin) verleihen. Das unselektive Herbizid Glyphosat hat die 5-Enolpyruvyl-3-phosphoshikimatsynthase (EPSPS) als

15 molekulares Target. Diese hat eine Schlüsselfunktion in der Biosynthese aromatischer Aminosäuren in Mikroben und Pflanzen, jedoch nicht in Säugern (Steinrucken HC et al. (1980) *Biochem. Biophys. Res. Commun.* 94:1207-1212; Levin JG und Sprinson DB (1964) *J. Biol. Chem.* 239: 1142-1150; Cole DJ

20 (1985) Mode of action of glyphosate a literature analysis, p. 48-74. In: Grossbard E und Atkinson D (eds.). *The herbicide glyphosate*. Butterworths, Boston.). Glyphosat-tolerante EPSPS Varianten werden bevorzugt als Selektionsmarker

25 verwendet (Padgett SR et al. (1996). *New weed control opportunities: development of soybeans with a Roundup Ready™ gene*. In: *Herbicide Resistant Crops* (Duke, S.O., ed.), pp. 53-84. CRC Press, Boca Raton, FL; Saroha MK und Malik VS (1998) *J Plant Biochemistry and Biotechnology* 7:65-72). Das

30 EPSPS Gen des *Agrobacterium* sp. strain CP4 hat eine natürliche Toleranz gegen Glyphosat, die auf entsprechende transgene Pflanzen transferiert werden kann. Das CP4 EPSPS Gen wurde aus *Agrobacterium* sp. strain CP4 kloniert (Padgett SR et al. (1995) *Crop Science* 35(5):1451-1461). Sequenzen

35 von 5-Enolpyruvylshikimate-3-phosphate-synthasen, die Glyphosat-tolerant sind, wie beispielsweise beschrieben in US 5,510,471; US 5,776,760; US 5,864,425; US 5,633,435; US 5,627,061; US 5,463,175; EP 0 218 571, sind sowohl in den Patenten beschrieben als auch in der GenBank hinterlegt. Weitere Sequenzen sind beschrieben unter GenBank

40 Accession X63374. Ferner ist das *aroA* Gen bevorzugt (M10947 *S. typhimurium aroA* locus 5-enolpyruvylshikimate-3-phosphate synthase (*aroA* protein) gene).

45 - das für das Glyphosat® degradierende Enzyme kodierende *gox* Gen (Glyphosatoxidoreduktase). GOX (beispielsweise die Glyphosatoxidoreduktase aus *Achromobacter* sp.) katalysiert

## 22

- die Spaltung einer C-N Bindung im Glyphosat, welches so zu Aminomethylphosphonsäure (AMPA) und Glyoxylat umgesetzt wird. GOX kann dadurch eine Resistenz gegen Glyphosat vermitteln (Padgett SR et al. (1996) J Nutr. 1996 Mar;126(3):702-16; Shah D et al. (1986) Science 233: 478-481).
- 5
- das deh Gen (kodierend für eine Dehalogenase, die Dalapon® inaktiviert), (GenBank Acc.-No.: AX022822, AX022820 sowie WO99/27116)
- 10
- bxn Gene, die für Bromoxynil degradierende Nitrilaseenzyme kodieren. Beispielsweise die Nitrilase aus Klebsiella ozananae. Sequenzen sind in der Genbank beispielsweise unter den Acc.-No: E01313 (DNA encoding bromoxynil specific nitrilase) und J03196 (K. pneumoniae bromoxynil-specific nitrilase (bxn) gene, complete cds) zu finden.
- 15
- Neomycinphosphotransferasen verleihen eine Resistenz gegen Antibiotika (Aminoglykoside) wie Neomycin, G418, Hygromycin, Paromomycin oder Kanamycin, indem sie durch eine Phosphorylierungsreaktion deren inhibierende Wirkung reduzieren. Besonders bevorzugt ist das nptII Gen. Sequenzen können aus der GenBank erhalten werden (AF080390 Minitransposon mTn5-GNm; AF080389 Minitransposon mTn5-Nm, complete sequence). Zudem ist das Gen bereits Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden (AF234316 pCAMBIA-2301; AF234315 pCAMBIA-2300, AF234314 pCAMBIA-2201). Das NPTII Gen kodiert für eine Aminoglycosid-3'O-phosphotransferase aus E.coli, Tn5 (GenBank Acc.-No: U00004 Position 1401-2300; Beck et al. (1982) Gene 19 327-336).
- 20
- 25
- 30
- das DOG<sup>R1</sup>-Gen. Das Gen DOG<sup>R1</sup> wurde aus der Hefe Saccharomyces cerevisiae isoliert (EP 0 807 836). Es codiert für eine 2-Desoxyglukose-6-phosphat Phosphatase, die Resistenz gegenüber 2-DOG verleiht (Randez-Gil et al. 1995, Yeast 11, 1233-1240; Sanz et al. (1994) Yeast 10:1195- 1202, Sequenz: GenBank Acc.-No.: NC001140 chromosom VIII, Saccharomyces cervisiae Position 194799-194056).
- 35
- 40
- Sulfonylurea- und Imidazolinon inaktivierende Acetolactat-synthasen, die eine Resistenz gegen Imidazolinon/Sulfonylurea-Herbizide verleihen. Beispielhaft seien für Imidazolinon-Herbizide die Wirkstoffe Imazamethabenz-methyl, Imazzamox, Imazapyr, Imazaquin, Imazethapyr zu nennen. Für Sulfonylharn-
- 45

## 23

- stoff-Herbizide seien beispielhaft Amidosulfuron, Azimsulfuron, Chlorimuronethyl, Chlorsulfuron, Cinosulfuron, Imazosulfuron, Oxasulfuron, Prosulfuron, Rimsulfuron, Sulfo-
- 5 Wirkstoffe der genannten Klassen bekannt. Geeignet sind Nukleinsäuresequenzen wie beispielsweise die unter der Gen-Bank Acc-No.: X51514 abgelegte Sequenz für das Arabidopsis thaliana Csr 1.2 Gen (EC 4.1.3.18) (Sathasivan K et al. (1990) Nucleic Acids Res. 18(8):2188). Acetolactatesynthasen,
- 10 die eine Resistenz gegen Imidazolinon-Herbizide verleihen, sind ferner beschrieben unter den GenBank Acc.-No.:
- a) AB049823 Oryza sativa ALS mRNA for acetolactate synthase, complete cds, herbicide resistant biotype
  - 15 b) AF094326 Bassia scoparia herbicide resistant acetolactate synthase precursor (ALS) gene, complete cds
  - c) X07645 Tobacco acetolactate synthase gene, ALS SuRB (EC 4.1.3.18)
  - 20 d) X07644 Tobacco acetolactate synthase gene, ALS SuRA (EC 4.1.3.18)
  - e) A19547 Synthetic nucleotide mutant acetolactate synthase
  - f) A19546 Synthetic nucleotide mutant acetolactate synthase
  - g) A19545 Synthetic nucleotide mutant acetolactate synthase
  - h) I05376 Sequence 5 from Patent EP 0257993
  - 25 i) I05373 Sequence 2 from Patent EP 0257993
  - j) AL133315
- Bevorzugt ist die Expression einer Acetolactatsynthase gemäß SEQ ID NO: 5, beispielsweise kodiert durch eine Nukleinsäure-
- 30 sequenz gemäß SEQ ID NO: 4.
- Hygromycinphosphotransferasen (X74325 P. pseudomallei gene for hygromycin phosphotransferase) die eine Resistenz gegen das Antibiotikum Hygromycin verleihen. Das Gen ist Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden (AF294981 pINDEX4; AF234301 pCAMBIA-1380; AF234300 pCAMBIA-1304; AF234299 pCAM-
  - 40 BIA-1303; AF234298 pCAMBIA-1302; AF354046 pCAMBIA-1305.; AF354045 pCAMBIA-1305.1)
  - Resistenzgene gegen
  - 45 a) Chloramphenicol (Chloramphenicolacetyltransferase),

## 24

- b) Tetracyclin, verschiedene Resistenzgene sind beschrieben z.B. X65876 *S. ordonez* genes class D tetA and tetR for tetracycline resistance and repressor proteins X51366 *Bacillus cereus* plasmid pBC16 tetracycline resistance gene. Zudem ist das Gen bereits Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden
- 5
- c) Streptomycin, verschiedene Resistenzgene sind beschrieben z.B. mit der GenBank Acc.-No.:AJ278607 *Corynebacterium acetoacidophilum* ant gene for streptomycin adenyl-yl-transferase.
- 10
- d) Zeocin, das entsprechende Resistenzgen ist Bestandteil zahlreicher Klonierungsvektoren (z.B. L36849 Cloning vector pZEO) und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden.
- 15
- e) Ampicillin ( $\beta$ -Lactamase Gen; Datta N, Richmond MH. (1966) *Biochem J.* 98(1):204-9; Heffron F et al (1975) *J. Bacteriol*-122: 250-256; das Amp Gen wurde zuerst zur Herstellung des *E. coli* Vektors pBR322 kloniert; Bolivar F et al. (1977) *Gene* 2:95-114). Die Sequenz ist Bestandteil zahlreicher Klonierungsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden.
- 20
- 25
- 30
- Gene wie die Isopentenyltransferase aus *Agrobacterium tumefaciens* (strain:PO22) (Genbank Acc.-No.: AB025109). Das *ipt* Gen ist ein Schlüsselenzym der Cytokin-Biosynthese. Seine Überexpression erleichtert die Regeneration von Pflanzen (z.B. Selektion auf Cytokin-freiem Medium). Das Verfahren zur Nutzung des *ipt* Gens ist beschrieben (Ebinuma H et al. (2000) *Proc Natl Acad Sci USA* 94:2117-2121; Ebinuma, H et al. (2000) Selection of Marker-free transgenic plants using the oncogenes (*ipt*, *rol* A, B, C) of *Agrobacterium* as selectable markers, In *Molecular Biology of Woody Plants*. Kluwer Academic Publishers).
- 35
- 40

Verschiedene weitere positive Selektionsmarker, die den transformierten Pflanzen einen Wachstumsvorteil gegenüber nicht-transformierten verleihen, sowie Verfahren zu ihrer Verwendung sind u.a. beschrieben in EP-A 0 601 092. Beispielhaft sind zu nennen  $\beta$ -Glucuronidase (in Verbindung mit z.B. Cyto-

45

## 25

kininglucuronid), Mannose-6-phosphat-Isomerase (in Verbindung mit Mannose), UDP-Galaktose-4-Epimerase (in Verbindung mit z.B. Galactose), wobei Mannose-6-phosphat-Isomerase in Verbindung mit Mannose besonders bevorzugt ist.

5

ii) Negative Selektionsmarker ermöglichen beispielsweise die Selektion von Organismen mit erfolgreich deletierten Sequenzen, die das Markergen umfassen (Koprek T et al. (1999) The Plant Journal 19(6):719-726). Bei der negativen Selektion wird beispielsweise durch den in die Pflanze eingebrachten negativen Selektionsmarker eine Verbindung, die ansonsten für die Pflanze keine nachteilige Wirkung hat, in eine Verbindung mit nachteiliger Wirkung umgesetzt. Ferner sind Gene geeignet, die per se eine nachteilige Wirkung haben, wie zum Beispiel TK thymidine kinase (TK) und Diphtheria Toxin A Fragment (DT-A), das codA Genprodukt kodierend für eine Cytosindeaminase (Gleave AP et al. (1999) Plant Mol Biol 40(2):223-35; Perera RJ et al. (1993) Plant Mol Biol 23(4): 793-799; Stougaard J (1993) Plant J 3:755-761), das Cytochrom P450 Gen (Koprek et al. (1999) Plant J. 16:719-726), Gene kodierend für eine Haloalkan Dehalogenase (Naested H (1999) Plant J. 18:571-576), das iaaH Gen (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) oder das tms2 Gen (Fedoroff NV & Smith DL 1993, Plant J 3: 273-289).

10

15

20

25

Die jeweils für die Selektion verwendeten Konzentrationen der Antibiotika, Herbizide, Biozide oder Toxine müssen an die jeweiligen Testbedingungen bzw. Organismen angepasst werden. Beispielhaft seien für Pflanzen zu nennen Kanamycin (Km) 50 mg/l, Hygromycin B 40 mg/l, Phosphinothricin (Ppt) 6 mg/l.

30

Ferner können funktionelle Analoga der genannten Nukleinsäuren kodierend für Selektionsmarker exprimiert werden. Funktionelle Analoga meint hier all die Sequenzen, die im wesentlichen die gleiche Funktion haben d.h. zu einer Selektion transformierter Organismen befähigt sind. Dabei kann das funktionelle Analogon sich in anderen Merkmalen durchaus unterscheiden. Es kann zum Beispiel eine höhere oder niedrigere Aktivität haben oder auch über weitere Funktionalitäten verfügen.

35

Funktionelle Analoga meint ferner Sequenzen, die für Fusionsproteine kodieren bestehend aus einem der bevorzugten Selektionsmarker und anderen Proteinen zum Beispiel einem weiteren bevorzugten Selektionsmarker oder aber auch einer Signalpeptidsequenz.

40

45

## 26

Die Expression des Selektionsmarkers ist in jedem gewünschten Zellkompartiment, wie z.B. dem Endomembransystem, der Vakuole und den Chloroplasten möglich. Durch Nutzung des sekretorischen Weges sind gewünschte Glykosylierungsreaktionen, besondere Faltungen  
5 u.ä. möglich. Auch die Sekretion des Zielproteins zur Zelloberfläche bzw. die Sezernierung ins Kulturmedium, beispielsweise bei Nutzung suspensionskultivierter Zellen oder Protoplasten ist möglich. Die dafür notwendigen Targetsequenzen können sowohl  
10 in einzelnen Vektorvariationen berücksichtigt werden als auch durch Verwendung einer geeigneten Klonierungsstrategie gemeinsam mit dem zu klonierenden Zielgen in den Vektor mit eingebracht werden. Als Targetsequenzen können sowohl Gen eigene, sofern vorhanden, oder heterologe Sequenzen genutzt werden. Zusätzliche, heterologe zur funktionellen Verknüpfung bevorzugte  
15 aber nicht darauf beschränkte Sequenzen sind weitere Targeting-Sequenzen zur Gewährleistung der subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endoplasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten; sowie Translationsverstärker wie die  
20 5'-Leadersequenz aus dem Tabak-Mosaik-Virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711) und dergleichen. Das Verfahren, an sich nicht in den Plastiden lokalisierte Proteine, gezielt in die Plastiden zu transportieren, ist beschrieben (Klosgen RB und Weil JH (1991) Mol Gen Genet 225(2):297-304; Van Breusegem  
25 F et al. (1998) Plant Mol Biol 38(3):491-496). Bevorzugte Sequenzen sind:

- a) kleine Untereinheit (SSU) der Ribulosebisphosphatcarboxylase (Rubisco ssu) aus Erbse, Mais, Sonnenblume  
30
- b) Transitpeptide abgeleitet von Genen der pflanzlichen Fettsäurebiosynthese wie das Transitpeptid des plastidären "Acyl Carrier Protein" (ACP), die Stearyl-ACP-Desaturase,  $\beta$ -Ketoacyl-ACP Synthase oder die Acyl-ACP-Thioesterase.  
35
- c) das Transitpeptid für GBSSI ("Granule Bound Starch Synthase I")  
40
- d) LHCP II Gene.

Die Zielsequenzen können mit anderen, von dem Transitpeptid kodierenden Sequenzen verschiedenen, Targeting-Sequenzen verknüpft sein, um eine subzellulären Lokalisation im Apoplasten, in der Vakuole, in Plastiden, im Mitochondrium, im Endo-  
45 plasmatischen Retikulum (ER), im Zellkern, in Ölkörperchen oder anderen Kompartimenten zu gewährleisten. Ferner können Translationsverstärker wie die 5'-Leadersequenz aus dem Tabak-

Mosaik-Virus (Gallie et al. (1987) Nucl Acids Res 15:8693-8711) und dergleichen zum Einsatz kommen.

Die erfindungsgemäßen Expressionskassetten und die von ihnen  
 5 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 oder von  
 diesen abgeleitete Vektoren oder Organismen haben. Beispielhaft  
 10 aber nicht einschränkend seien zu nennen:

- a) Reporter gene, die für leicht quantifizierbare Proteine  
 kodieren und über Eigenfarbe oder Enzymaktivität eine  
 Bewertung der Transformationseffizienz, des Expressionsortes  
 15 oder -zeitpunktes gewährleisten. Ganz besonders bevorzugt  
 sind dabei Gene kodierend für Reporter-Proteine (siehe auch  
 Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44)  
 wie
- 20 - "green fluorescence protein" (GFP) (Chui WL et al., Curr  
 Biol 1996, 6:325-330; Leffel SM et al., Biotechniques.  
 23(5):912-8, 1997; Sheen et al.(1995) Plant Journal  
 8(5):777-784; Haseloff et al.(1997) Proc Natl Acad Sci  
 25 USA 94(6):2122-2127; Reichel et al.(1996) Proc Natl Acad  
 Sci USA 93(12):5888-5893; Tian et al. (1997) Plant Cell  
 Rep 16:267-271; WO 97/41228).
- Chloramphenicoltransferase (Fromm et al. (1985) Proc.  
 Natl. Acad. Sci. USA 82:5824-5828),
- 30 - Luziferase (Millar et al., Plant Mol Biol Rep 1992  
 10:324-414; Ow et al. (1986) Science, 234:856-859);  
 erlaubt Biolumineszenzdetektion.
- 35 -  $\beta$ -Galactosidase, kodiert für ein Enzym für das ver-  
 schiedene chromogene Substrate zur Verfügung stehen.
- $\beta$ -Glucuronidase (GUS) (Jefferson et al., EMBO J. 1987, 6,  
 3901-3907) oder das uidA Gen, das ein Enzym für ver-  
 40 schiedene chromogene Substrate kodiert.
- R-Locus Genprodukt:Protein, das die Produktion von  
 Anthocyaninpigmenten (rote Färbung) in pflanzlichen,  
 Gewebe reguliert und so eine direkte Analyse der  
 45 Promotoraktivität ohne Zugabe zusätzlicher Hilfsstoffe  
 oder chromogener Substrate ermöglicht (Dellaporta et al.,  
 In: Chromosome Structure and Function: Impact of New

## 28

Concepts, 18th Stadler Genetics Symposium, 11:263-282, 1988).

- 5           -     $\beta$ -Lactamase (Sutcliffe (1978) Proc Natl Acad Sci USA 75:3737-3741), Enzym für verschiedene chromogene Substrate (z.B. PADAC, eine chromogenes Cephalosporin).
- 10           -    xylE Genprodukt (Zukowsky et al. (1983) Proc Natl Acad Sci USA 80:1101-1105), Catecholdioxygenase, die chromogene Catechole umsetzen kann.
- Alpha-Amylase (Ikuta et al. (1990) Bio/technol. 8:241-242).
- 15           -    Tyrosinase (Katz et al. (1983) J Gen Microbiol 129:2703-2714), Enzym, das Tyrosin zu DOPA und Dopaquinon oxidiert, die infolge das leicht nachweisbare Melanin bilden.
- 20           -    Aequorin (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), kann in der Calcium-sensitiven Biolumineszenzdetektion verwendet werden.
- 25           b)    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, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
- 30           c)    Elemente zum Beispiel "Bordersequenzen", die einen Agrobakterien-vermittelte Transfer in Pflanzenzellen für die Übertragung und Integration ins Pflanzengenom ermöglichen, wie zum Beispiel die rechte oder linke Begrenzung der T-DNA oder die vir-Region.
- 35           d)    Multiple Klonierungsregionen (MCS) erlauben und erleichtern die Insertion eines oder mehrerer Nukleinsäuresequenzen.
- 40           Dem Fachmann sind verschiedene Wege bekannt, um zu einer erfindungsgemäßen Expressionskassette zu gelangen. Die Herstellung einer erfindungsgemäßen Expressionskassette erfolgt beispielsweise durch Fusion des Nit1 Promotors gemäß SEQ-ID NO: 1
- 45           (oder eines funktionellen Äquivalentes oder funktionell äquivalenten Teils) mit einer zu exprimierenden Nukleotidsequenz kodierend für einen Selektionsmarker, gegebenenfalls einer für

## 29

ein Transitpeptid kodierenden Sequenz, vorzugsweise ein chloroplastenspezifisches Transitpeptid, welches vorzugsweise zwischen dem Promotor und der jeweiligen Nukleotidsequenz angeordnet ist, sowie einem Terminator- oder Polyadenylierungssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, 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 sowie in Silhavy TJ, Berman ML und Enquist LW (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY und in Ausubel FM et al. (1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience beschrieben sind.

Unter einer Expressionskassette sind aber auch solche Konstruktionen zu verstehen, bei denen eine transgen zu exprimierende Nukleinsäuresequenz kodierend für einen Selektionsmarker zum Beispiel durch eine homologe Rekombination hinter den endogenen, natürlichen NIT1-Promotor (oder eins seiner funktionellen Äquivalente beispielsweise aus anderen Pflanzenarten) platziert wird, wodurch man eine erfindungsgemäße Expressionskassette erhält, die die Expression der transgen zu exprimierenden Nukleinsäuresequenz kodierend für einen Selektionsmarker in der erfindungsgemäßen Weise steuert.

Erfindungsgemäß sind ferner Vektoren, die die oben beschriebenen Expressionskassetten enthalten. Vektoren können beispielhaft Plasmide, Cosmide, Phagen, Viren oder auch Agrobakterien sein.

Ein anderer Gegenstand der Erfindung betrifft transgene Organismen, transformiert mit wenigstens einer erfindungsgemäßen Expressionskassette oder einem erfindungsgemäßen Vektor, sowie Zellen, Zellkulturen, Gewebe, Teile - wie zum Beispiel bei pflanzlichen Organismen Blätter, Wurzeln usw.- oder Vermehrungsgut abgeleitet von solchen Organismen.

Unter Organismus, Ausgangs- oder Wirtsorganismen werden prokaryotische oder eukaryotische Organismen, wie beispielsweise Mikroorganismen oder pflanzliche Organismen verstanden. Bevorzugte Mikroorganismen sind Bakterien, Hefen, Algen oder Pilze.

Bevorzugte Bakterien sind Bakterien der Gattung Escherichia, Erwinia, Agrobacterium, Flavobacterium, Alcaligenes oder Cyanobakterien zum Beispiel der Gattung Synechocystis.

Bevorzugt sind vor allem Mikroorganismen, welche zur Infektion von Pflanzen und damit zur Übertragung der erfindungsgemäßen Kassetten befähigt sind. Bevorzugte Mikroorganismen sind solche

## 30

aus der Gattung Agrobakterium und insbesondere der Art Agrobakterium tumefaciens.

Bevorzugte Hefen sind Candida, Saccharomyces, Hansenula oder  
5 Pichia.

Bevorzugte Pilze sind Aspergillus, Trichoderma, Ashbya, Neurospora, Fusarium, Beauveria oder weitere in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) auf Seite 15, Tabelle 6  
10 beschriebene Pilze.

Als transgene Organismen bevorzugte Wirts- oder Ausgangsorganismen sind vor allem Pflanzen. Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Eingeschlossen sind ferner die  
15 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.

20

Einjährige, mehrjährige, monocotyledone und dicotyledone Pflanzen sind bevorzugte Wirtsorganismen für die Herstellung transgener Pflanzen. Die Expression von Genen ist ferner vorteilhaft bei allen Schmuckpflanzen, Nutz- oder Zierbäumen, Blumen, Schnitt-  
25 blumen, Sträuchern oder Rasen. Beispielhaft 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; Algen wie Chlorophyceae, Phaeophyceae, Rhodophyceae, Myxophyceae, Xanthophyceae,  
30 Bacillariophyceae (Diatomeen) und Euglenophyceae.

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

40 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.

45

## 31

Bevorzugte dikotyle Pflanzen sind insbesondere ausgewählt aus den dikotylen Kulturpflanzen, wie zum Beispiel

- 5 - Asteraceae wie Sonnenblume, Tagetes oder Calendula und andere mehr,
- Compositae, besonders die Gattung Lactuca, ganz besonders die Art sativa (Salat) und andere mehr,
- 10 - 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
- 15 oder Canola und andere mehr,
- Cucurbitaceae wie Melone, Kürbis oder Zucchini und andere mehr,
- Leguminosae besonders die Gattung Glycine, ganz besonders die
- 20 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 (Kaffestrauch) und
- 25 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)
- 30 sowie Tabak oder Paprika und andere mehr,
- Sterculiaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Theobroma cacao (Kakaostrauch) und andere mehr,
- 35 - Theaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Camellia sinensis oder Thea sinensis (Teestrauch) und andere mehr,
- Umbelliferae, besonders die Gattung Daucus (ganz besonders
- 40 die Art carota (Karrotte)) und Apium (ganz besonders die Art graveolens dulce (Selarie)) und andere mehr; und die Gattung Capsicum, ganz besonders die Art annum (Pfeffer) und andere mehr,

## 32

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

5. Umfasst sind ferner Schmuckpflanzen, Nutz- oder Zierbäumen, Blumen, Schnittblumen, Sträuchern oder Rasen. Beispielhaft aber nicht einschränkend seien zu nennen Angiospermen, Bryophyten wie zum Beispiel Hepaticae (Leberblümchen) und Musci (Moose); Pteridophyten wie Farne, Schachtelhalm und Lycopoden; Gymno-
- 10 spermen wie Koniferen, Cycaden, Ginkgo und Gnetales, die Familien der Rosaceae wie Rose, Ericaceae wie Rhododendrons und Azaleen, Euphorbiaceae wie Weihnachtssterne und Kroton, Caryophyllaceae wie Nelken, Solanaceae wie Petunien, Gesneriaceae wie das Usambaraveilchen, Balsaminaceae wie das Springkraut, Orchidaceae
- 15 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.

- 20 Am meisten bevorzugt sind *Arabidopsis thaliana*, *Nicotiana tabacum*, *Tagetes erecta*, *Calendula officinalis* und *Brassica napus* sowie alle Gattungen und Arten, die als Nahrungs- oder Futtermittel zum Einsatz kommen, wie die beschriebenen Getreidearten, oder sich zur Herstellung von Ölen eignen, wie Ölsaaten (wie zum
- 25 Beispiel Raps), Nussarten, Soja, Sonnenblume, Kürbis und Erdnuss.

Pflanzliche Organismen im Sinne der Erfindung sind weiterhin weitere photosynthetisch aktive befähigte Organismen, wie zum Beispiel Algen oder Cyanobakterien, sowie Moose.

- 30 Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* oder *Dunaliella*.
- 35 Die Herstellung eines transformierten Organismus oder einer transformierten Zelle erfordert, dass die entsprechende DNA in die entsprechende Wirtszelle eingebracht wird. Für diesen Vorgang, der als Transformation bezeichnet wird, steht eine Vielzahl von Methoden zur Verfügung (siehe auch Keown et al. (1990)
- 40 *Methods in Enzymology* 185:527-537). So kann die DNA 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
- 45 kann. Die DNA kann auch durch Protoplastenfusion mit anderen DNA-enthaltenden Einheiten wie Minicells, Zellen, Lysosomen oder Liposomen erfolgen. Elektroporation ist eine weitere geeignete

## 33

Methode zur Einführung von DNA, bei der die Zellen reversibel durch einen elektrischen Impuls permeabilisiert werden.

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. Diese Stämme enthalten ein Plasmid (Ti bzw. Ri Plasmid), das auf die Pflanze nach *Agrobacterium*-Infektion übertragen wird. Ein Teil dieses Plasmons, genannt T-DNA (transferred DNA), wird in das Genom der Pflanzenzelle integriert.

Die *Agrobacterium*-vermittelte Transformation ist am besten für dicotyledone, diploide Pflanzenzellen geeignet, wohingegen die direkten Transformationstechniken sich für jeden Zelltyp eignen.

25

Die Einführung einer erfindungsgemäßen Expressionskassette in Zellen, bevorzugt in pflanzliche Zellen, kann vorteilhaft unter Verwendung von Vektoren realisiert werden.

In einer vorteilhaften Ausführungsform wird die Einführung der Expressionskassette mittels Plasmidvektoren realisiert. Bevorzugt sind solche Vektoren, die eine stabile Integration der Expressionskassette in das Wirtsgenom ermöglichen.

Im Falle von Injektion oder Elektroporation von DNA 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 erforderlich, das sich auf dem Plasmid ein zusätzliches selektionierbares Markergen befindet.

Transformationstechniken sind für verschiedene monokotyle und dikotyle pflanzliche Organismen beschrieben. Ferner stehen verschiedene mögliche Plasmidvektoren für die Einführung fremder Gene in Pflanzen zur Verfügung, die in der Regel einen Replikationsursprung für eine Vermehrung in *E. coli* und ein Markergen

## 34

für eine Selektion transformierter Bakterien enthalten. Beispiele sind pBR322, pUC Reihe, M13mp Reihe, pACYC184 etc.

Die Expressionskassette kann in den Vektor über eine geeignete  
 5 Restriktionsschnittstelle eingeführt werden. Das entstandene  
 Plasmid wird zunächst in E.coli eingeführt. Korrekt trans-  
 formierte E.coli werden selektioniert, gezüchtet und das rekombi-  
 nante Plasmid mit dem Fachmann geläufigen Methoden gewonnen.  
 Restriktionsanalyse und Sequenzierung können dazu dienen, den  
 10 Klonierungsschritt zu überprüfen.

Transformierte Zellen d.h. solche, die die eingeführte DNA  
 integriert in die DNA der Wirtszelle enthalten, können von  
 untransformierten selektioniert werden, wenn ein selektionier-  
 15 barer Marker Bestandteil der eingeführten DNA ist. Als Marker  
 kann beispielhaft jedes Gen fungieren, das eine Resistenz gegen  
 Antibiotika oder Herbizide zu verleihen vermag. Transformierte  
 Zellen, die ein solches Markergen exprimieren, sind in der Lage,  
 in der Gegenwart von Konzentrationen eines entsprechenden Anti-  
 20 biotikums oder Herbizides zu überleben, die einen untrans-  
 formierten Wildtyp abtöten. Beispiele sind das bar Gen, das  
 Resistenz gegen das Herbizid Phosphinothricin verleiht (Rathore  
 KS et al. (1993) Plant Mol Biol 21(5):871-884), das nptII Gen,  
 das Resistenz gegen Kanamycin verleiht, das hpt Gen,  
 25 das Resistenz gegen Hygromycin verleiht, oder das EPSP-Gen,  
 das Resistenz gegen das Herbizid Glyphosat verleiht.

Je nach Methode der DNA-Einführung können weitere Gene auf dem  
 Vektorplasmid erforderlich sein. Werden Agrobacteria verwendet,  
 30 so ist die Expressionskassette in spezielle Plasmide zu inte-  
 griieren, entweder in einen Zwischenvektor (englisch: shuttle  
 or intermediate vector) oder einen binären Vektor. Wenn zum Bei-  
 spiel ein Ti oder Ri Plasmid zur Transformation verwendet werden  
 soll, ist zumindest die rechte Begrenzung, meistens jedoch die  
 35 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 repli-  
 zieren. Sie enthalten in der Regel ein Selektionsmarkergen und  
 40 einen Linker oder Polylinker flankiert von der rechten und linken  
 T-DNA Begrenzungssequenz. Sie können direkt in Agrobacterium  
 transformiert werden (Holsters et al., Mol. Gen. Genet. 163  
 (1978), 181-187). Das Selektionsmarkergen erlaubt eine Selektion  
 transformierter Agrobacteria und ist zum Beispiel das nptII Gen,  
 45 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 Über-

## 35

tragung der T-DNA auf die pflanzliche Zelle erforderlich. Ein so transformiertes Agrobacterium kann zur Transformation pflanzlicher Zellen verwendet werden.

- 5 Die Verwendung von T-DNA zur Transformation pflanzlicher Zellen ist intensiv untersucht und beschrieben (EP 120516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasterdam, Chapter V; Fraley et al., Crit. Rev. Plant. Sci., 4:1-46 and An et al. (1985) EMBO J 4:277-287). Verschiedene  
10 binäre Vektoren sind bekannt und teilweise kommerziell erhältlich wie zum Beispiel pBIN19 (Clontech Laboratories, Inc. USA).

- Für den Transfer der DNA in die pflanzliche Zelle werden pflanzliche Explantate mit Agrobacterium tumefaciens oder Agrobacterium  
15 rhizogenes kokultiviert. Ausgehend von infiziertem Pflanzenmaterial (z.B. Blatt-, Wurzel- oder Stengelteile, aber auch Protoplasten oder Suspensionen von Pflanzenzellen) können ganze Pflanzen unter Verwendung eines geeigneten Mediums, das zum  
Beispiel Antibiotika oder Biozide zur Selektion transformierter  
20 Zellen enthalten kann, regeneriert werden. Die erhaltenen Pflanzen können dann auf die Präsenz der eingeführten DNA, hier der erfindungsgemäßen Expressionskassette, durchmustert werden. Sobald die DNA in das Wirtsgenom integriert ist, ist der entsprechende Genotyp in der Regel stabil und die entsprechende  
25 Insertion wird auch in den Nachfolgenerationen wiedergefunden. In der Regel enthält die integrierte Expressionskassette einen Selektionsmarker, der der transformierten Pflanze eine Resistenz gegen ein Biozid (zum Beispiel ein Herbizid), einen Metabolismusinhibitor wie 2-DOG oder ein Antibiotikum wie Kanamycin, G 418,  
30 Bleomycin, Hygromycin oder Phosphinothricin etc. verleiht. Der Selektionsmarker erlaubt die Selektion von transformierten Zellen von untransformierten (McCormick et al., Plant Cell Reports 5 (1986), 81-84). Die erhaltenen Pflanzen können in üblicher Weise gezüchtet und gekreuzt werden. Zwei oder mehr Generationen soll-  
35 ten kultiviert werden, um sicherzustellen, dass die genomische Integration stabil und vererblich ist.

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

## 36

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 Zell-

5 massen kann die Bildung von Spross und Wurzel in bekannter Weise induziert werden. Die erhaltenen Sprösslinge können ausgepflanzt und gezüchtet werden.

Die Wirksamkeit der Expression der transgen exprimierten Nuklein-

10 säuren kann beispielsweise *in vitro* durch Sprossmeristemvermehrung unter Verwendung einer der oben beschriebenen Selektionsmethoden ermittelt werden.

Erfindungsgemäß sind ferner von den oben beschriebenen transgenen

15 Organismen abgeleitete Zellen, Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln, Blätter etc.-, und transgenes Vermehrungsgut wie Saaten oder Früchte.

Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch

20 veränderte Pflanzen können auch beispielsweise direkt oder nach an sich bekannter Aufbereitung als Nahrungsmittel oder Futtermittel verwendet werden.

Ein weiterer Gegenstand der Erfindung betrifft die Verwendung der

25 oben beschriebenen 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

30 oder Feinchemikalien.

Bevorzugt ist ferner ein Verfahren zur rekombinanten Herstellung von Pharmazeutika oder Feinchemikalien in Wirtsorganismen wobei ein Wirtsorganismus mit einer der oben beschriebenen Expressions-

35 kassetten transformiert wird und diese Expressionskassette ein oder mehrere Strukturgene enthält; die für die gewünschte Feinchemikalie kodieren oder die Biosynthese der gewünschten Feinchemikalie katalysieren, der transformierte Wirtsorganismus gezüchtet wird und die gewünschte Feinchemikalie aus dem Züchtungs-

40 medium isoliert wird. Dieses Verfahren ist für Feinchemikalien wie Enzyme, Vitamine, Aminosäuren, Zucker, Fettsäuren, natürliche und synthetische Geschmacks-, Aroma- und Farbstoffe breit anwendbar. Besonders bevorzugt ist die Produktion von Tocopherolen und Tocotrienolen sowie Carotinoiden. Die Züchtung der trans-

45 formierten Wirtsorganismen sowie die Isolierung aus den Wirtsorganismen bzw. aus dem Züchtungsmedium erfolgt mit dem Fachmann bekannten Verfahren. Die Produktion von Pharmazeutika, wie zum

## 37

Beispiel Antikörpern oder Vakzinen ist beschrieben bei Hood EE, Jilka JM (1999) Curr Opin Biotechnol 10(4):382-6; Ma JK, Vine ND (1999) Curr Top Microbiol Immunol 236:275-92.

## 5 Sequenzen

1. SEQ ID NO: 1  
Promotor and 5'-untranslatierte Region des Arabidopsis thaliana NIT1-Promotors.
- 10 2. SEQ ID NO: 2  
Nukleinsäure kodierend für eine Phosphinothricinacetyltransferase.
- 15 3. SEQ ID NO: 3  
Aminosäuresequenz kodierend für eine Phosphinothricinacetyltransferase.
4. SEQ ID NO: 4  
20 Nukleinsäure kodierend für eine Acetolactatsynthase.
5. SEQ ID NO: 5  
Aminosäuresequenz kodierend für eine Acetolactatsynthase.
- 25 6. SEQ ID NO: 6 - Oligonukleotid-Primer nit3  
5'-CCGGATCCACTCGAGTCTTTGTTTTTTACTTTGG-3'
7. SEQ ID NO: 7 - Oligonukleotid-Primer nit5  
5'-CATCAAGATCTTGGTGATGTAGCAA-3'
- 30 8. SEQ ID NO: 8 - Oligonukleotid-Primer ubi5  
5'-CCAAACCATGGTAAGTTTGTCTAAAGCTTA-3'
9. SEQ ID NO: 9 - Oligonukleotid-Primer ubi3  
35 5'-CGGATCCTTTTGTGTTTCGTCTTCTCTCACG-3'
10. SEQ ID NO: 10 - Oligonukleotid-Primer sqs5  
5'-GTCTAGAGGCAAACCACCGAGTGTT-3'
- 40 11. SEQ ID NO: 11 - Oligonukleotid-Primer sqs3  
5'-CGGTACCTGTTTCCAGAAAATTTTGATTCAG-3'
12. SEQ ID NO: 12  
Binärplasmid pSUN3 (Sungene GmbH & Co KGaA)

13. SEQ ID NO: 13  
Binärplasmid pSUN14 (Sungene GmbH & Co KGaA)
14. SEQ ID NO: 14  
5 Binärplasmid pSUN3PatNos (Sungene GmbH & Co KGaA)
15. SEQ ID NO:15  
Binärplasmid pSun5NptIICat (SunGene GmbH & Co KGaA)
- 10 16. SEQ ID NO: 16  
Nukleinsäuresequenz des Terminators des Cathepsin D Inhibitor  
Gens aus Kartoffel (GenBank Acc. No.: X74985)
- 15 17. SEQ ID NO: 17  
Nukleinsäuresequenz des Terminators der Speicherproteingens  
VfLE1B3 aus der Ackerbohne (GenBank Acc. No.: Z26489)
18. SEQ ID NO: 18  
20 Verkürzter Promotor (1478 bp) des *Arabidopsis thaliana*  
NIT1-Gens
19. SEQ ID NO: 19  
25 Verkürzter Promotor (1330 bp) des *Arabidopsis thaliana*  
NIT1-Gens

#### Abbildungen

- 30 Fig.1: Vergleich der Expressionsstärke des isolierten NIT1-  
Promotors aus *Arabidopsis thaliana* mit dem 35S Promotor.  
Die Tabaklinien UH76 und UH77 enthalten das GUS Gen unter  
der Kontrolle des 35S Promotors. Die Linie UH127 ent-  
hält das GUS-Gen unter Kontrolle des NIT1-Promotors aus  
35 *Arabidopsis thaliana*. WT stellt den untransformierten  
Tabak-Wildtyp dar.

#### Beispiele

- 40 Allgemeine Methoden:

Die chemische Synthese von Oligonukleotiden kann beispielsweise,  
in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet,  
2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die  
45 im Rahmen der vorliegenden Erfindung durchgeführten Klonierungs-  
schritte wie z.B. Restriktionsspaltungen, Agarosegelelektro-  
phorese, Reinigung von DNA-Fragmenten, Transfer von Nuklein-

## 39

säuren 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).

10

Beispiel 1: Isolierung von genomischer DNA aus *Arabidopsis thaliana* (CTAB-Methode)

Zur Isolierung genomischer DNA aus *Arabidopsis thaliana* 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

## 40

DNA mit 2,5 Volumina eiskaltem Ethanol vermischt und für 1h bei -20°C inkubiert. Alternativ kann der Ansatz mit 0.6 Volumina Iso-propanol 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.

## 15 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)
- 2 % (w/v) CTAB (4,0 g)

Jeweils vor der Verwendung werden frisch zugesetzt:

2 % β-Mercaptoethanol (20 µl für 1 ml Lösung I).

## 25 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 µl 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.

## 45 Beispiel 2: Transformation von Tabak, Raps und Kartoffeln

## 41

- Die Transformation von Tabak erfolgte über Infektion mit *Agrobacterium tumefaciens*. gemäß der von Horsch entwickelten Methode (Horsch et al. (1985) Science 227: 1229-1231). Alle zur Transformation verwendeten Konstrukte wurden anhand der Gefrier/
- 5 Tau-Methode (wiederholtes Auftauen und Einfrieren) in *Agrobacterium tumefaciens* transformiert. Die das gewünschte Konstrukt enthaltenden *Agrobacterium*-Kolonien wurden auf Mannitol/Glutamat-Medium mit 50 µg/ml Kanamycin, 50 µg/ml Ampicilin und 25 µg/ml Rifampicin selektioniert.
- 10 Zur Transformation von Tabakpflanzen (*Nicotiana tabacum* L. cv. Samsun NN) wurden 10 ml einer unter Selektion gewachsenen Übernachtskultur von *Agrobacterium tumefaciens* abzentrifugiert, der Überstand verworfen, und die Bakterien im gleichen Volumen Anti-
- 15 biotika-freien Mediums resuspendiert. In einer sterilen Petrischale wurden Blattscheiben steriler Pflanzen (Durchmesser ca. 1 cm) in dieser Bakterienlösung gebadet. Anschließend wurden die Blattscheiben in Petrischalen auf MS-Medium (Murashige und Skoog (1962) Physiol Plant 15:473ff.) mit 2 % Saccharose und 0,8 %
- 20 Bacto-Agar ausgelegt. Nach 2tägiger Inkubation im Dunkeln bei 25°C wurden sie auf MS-Medium mit 100 mg/l Kanamycin, 500 mg/l Claforan, 1 mg/l Benzylaminopurin (BAP), 0,2 mg/l Naphtylessigsäure (NAA), 1,6 % Glukose und 0,8 % Bacto-Agar übertragen und die Kultivierung (16 Stunden Licht/8 Stunden Dunkelheit) fort-
- 25 gesetzt. Wachsende Sprosse wurden auf hormonfreies MS-Medium mit 2 % Saccharose, 250 mg/l Claforan und 0,8 % Bacto-Agar überführt.

- Die Transformation von Raps erfolgte mittels der Petiolentransformation nach Moloney et al. (Moloney MM et al. (1989) Plant
- 30 Cell Reports 8:238-242).

- Die Transformation von Kartoffel erfolgte mittels der Methode nach Kunze et al. (Kunze I et al. (2001) Molecular Breeding 7: 221-227).
- 35

Beispiel 3: Untersuchungen zur Eignung des Nitrilase 1 (NIT1) Promotors

- 40 a) Klonierung des NIT1 Promotor aus *Arabidopsis thaliana*

Der putative Nitrilase1 Promotor wurde mittels PCR aus genomischer *Arabidopsis thaliana* DNA mit den Primern nit3 und nit5 amplifiziert.

- 45 nit3 (SEQ ID NO: 6):  
5'-CCGGATCCACTCGAGTCTTTGTTTTTTACTTTGG-3'

nit5 (SEQ ID NO: 7):

5'-CATCAAGATCTTGGTGATGTAGCAA-3'

5 Reaktionsansatz:

	37,5 µl	H <sub>2</sub> O
	5 µl	10X Reaktionspuffer ("genomic PCR")
	4 µl	dNTP mix (jeweils 2,5 mM)
10	2,2 µl	25 mM Mg(OAc) <sub>2</sub> (Endkonzentration 1,1 mM)
	1 µl	Primer nit3 (10 µM)
	1 µl	Primer nit5 (10 µM)
	0,5 µl	Pfu-turbo Polymerase Mix
	1 µl	Genomische <i>Arabidopsis</i> DNA (ca. 250 ng)

15

PCR-Bedingungen:

- 1 Zyklus mit 5 min. bei 95°C
- 25 Zyklen mit 94°C für 30 sec, 50°C für 60 sec und 72°C für 1 min.
- 20 1 Zyklus mit 50°C für 60 sec, 72°C für 10 min., anschließend Kühlung auf 4°C bis zur Weiterverarbeitung.

Das entstandene PCR Fragment wurde in das *Sma*I geschnittene Plasmid pUC18 kloniert (pNit22) und mittels Sequenzanalyse verifiziert.

25

b) Konstruktion der NIT1-Promotor-Expressionskassetten

30 Ausgangsplasmid für Analysen mit dem Glucuronidasegen ist das Plasmid pGUSINT37. Zur Herstellung des Plasmids pGUSINT37 wurde die 35S-GUS-Intron Kasette aus dem Binärplasmid p35SGUSINT (G. Vancanneyt et al. (1989) Mol. Gen. Genet. 220: 245-50) durch Verdau mit *Pst*I herausgeschnitten. Die Enden des Fragmentes wurden durch Klenow-"Fill-In" geglättet und das Fragment in den *Sma*I geschnittenen Vektor pUC18 kloniert. Dieses Konstrukt trägt die Bezeichnung pGUSINT37.

35

Für die Fusion mit dem GUS Gen wurde das *Bam*HI Fragment aus dem pNit22 Vektor in das *Bam*HI geschnittene, dephosphorylierte Plasmid pGUSINT37 kloniert. Das erhaltene Konstrukt trägt die Bezeichnung pNit1GUSINT.

40

Das *Eco*RI/*Sal*I Fragment aus dem Vektor pNit1GUSINT, das die Nit1P/GUS Kasette enthält, wurde in das Binärplasmid pSUN3 (Sungene GmbH & Co KGaA; SEQ ID NO: 12) gespalten mit *Eco*RI/*Sal*I, integriert. Das resultierende Plasmid pSUN3Nit1GUS wurde in Tabak

45

## 43

und in Raps transformiert. Die erzeugten Tabakpflanzen wurden UH127, die erzeugten Rapspflanzen BN127 genannt.

Für die Vermittlung von Resistenz gegen Phosphinothricin wurde  
5 der NIT1 Promotor als *Bam*HI Fragment vor das Phosphinothricin-  
Resistenzgen in das *Bam*HI geschnittene, dephosphorylierte Plasmid  
pSUN3PatNos (Sungene GmbH & Co KGaA, SEQ ID NO: 14) kloniert. Das  
resultierende Plasmid mit der Bezeichnung pSUN3Nit1Pat wurde in  
Tabak transformiert.

10

Für die Vermittlung von Resistenz gegen Kanamycin wurde der NIT1  
Promotor als *Xho*I Fragment vor das NptII-Resistenzgen in das *Xho*I  
geschnittene, dephosphorylierte Plasmid pSUN5NptIICat (Sungene  
GmbH & Co KGaA, SEQ ID NO: 15) kloniert. Das resultierende  
15 Plasmid mit der Bezeichnung pS5NitNptII wurde in Tabak, Raps  
und in Kartoffel transformiert.

c) Ergebnisse der GUS Analyse der transgenen Tabak- und Raps-  
20 pflanzen

Die mit dem Konstrukt pSUN3Nit1GUS transformierten Tabak- und  
Rapspflanzen (Linie UH127 bzw. BN127) wurden sowohl histochemisch  
als auch quantitativ auf die Expression des GUS Gens unter der  
25 Kontrolle des NIT1 Promotors analysiert.

Eine Blaufärbung, vergleichbar intensiv wie bei entsprechenden  
mit dem 35S Promotor/GUS-Konstrukten transformierten Pflanzen,  
wurde in den ‚sink‘ und ‚source‘ Blättern, Knospen und Kelch-  
30 blättern detektiert. In den Spitzen der Blütenblätter wurde  
gelegentlich eine schwache Färbung gefunden. Im Gegensatz zum  
35S Promotor konnte keine Blaufärbung in den Antheren (Staub-  
beutel) nachgewiesen werden. Nur vereinzelt zeigten beim Tabak  
Pollenkörner eine schwache Blaufärbung. Die Wurzeln der *in vitro*  
35 Pflanzen zeigten ebenfalls eine Blaufärbung.

Bei der quantitativen Analyse wurde die Aktivität des NIT1  
Promotors mit der Aktivität des 35S Promotors verglichen.  
Parallel zu den mit pSUN3Nit1GUS transformierten Pflanzen  
40 (UH127) wurden Tabakpflanzen der Linien UH76 und UH77 ange-  
zogen, die das GUS Gen unter der Kontrolle des 35S Promotors  
enthielten. Die Glucuronidaseaktivität wurden nach Extraktion  
von Blattmaterial durch den Umsatz von 4-Methylumbelliferyl-  
β-D-glucuronid quantitativ bestimmt (Jefferson et al. (1987)  
EMBO J 6:3901-3907). Die Analysen zeigten, dass die Stärke des  
45 isolierten Promotors des Gens der Nitrilase 1 aus *Arabidopsis*  
*thaliana* mit dem 35S Promotor vergleichbar ist (Fig.1).

## 44

- d) Ergebnisse der Analyse der Phosphinothricinresistenz der transgenen Tabakpflanzen

Für die Untersuchung der NIT1-Promotor gestützten Vermittlung  
5 von Resistenz gegen Phosphinothricin wurde der Nit1-Promotor mit dem bar Gen (UH126) kombiniert und in den Vektor pSUN3, der das nptII Gen unter Kontrolle des nos Promotors enthält, kloniert.

- 10 Als Vergleichskontrollen dienten das Plasmid pSUN14 (SEQ ID NO: 13), das das pat Gen unter der Kontrolle des nosP als selektiven Marker enthält, sowie das Konstrukt UH39, das das bar Gen unter Kontrolle des 35S Promotors im Vektor pPZP200 trägt.  
15 Alle drei Konstrukte wurden in den Agrobakterium tumefaciens Stamm EHA101[pEHA101] transformiert.

Die resultierenden Stämme sind für die Transformation, wie unter Beispiel 1 beschrieben, eingesetzt worden. Die selektive  
20 Regeneration wurde in Gegenwart von 5 mg/l Phosphinothricin erreicht. Etwa 4 Wochen nach Versuchsstart wurde die Regenerations-effizienz erfasst (Tab. 1).

Tabelle 1:

25

Konstrukt	Anzahl getesteter Explantate	Anzahl Sprossknospen bildender Explantate	Regenerations-effizienz (%)
30 UH39 (pPZP200-35S-bar)	130	126	97
UH126 (pSUN3-Nit-Prom-bar)	97	92	95
UH131 (pSUN14-nosP-bar)	95	71	75

35

Sprosse wurden in Gegenwart von 5 mg/l PPT bewurzelt (Tab. 2)

Tabelle 2:

40

Konstrukt	Anzahl getesteter Sprosse	Anzahl bewurzelter Sprosse	Effizienz der Bewurzelung (%)
45 UH39 (pPZP200-35S-bar)	68	61	90

45

UH126 (pSUN3-Nit-Prom-bar)	23	22	96
UH131 (pSUN14-nosP-bar)	50	48	96

5

Das mit dem Nitrilasepromotor ausgerüstete bar Gen lieferte nach Transformation Regenerationseffizienzen, die mit denen des 35S-Promotorkonstruktes vergleichbar sind.

- 10 Um die Selektion während der Keimung der Samen zu prüfen, wurden reife Samen der transgenen Tabaklinien (UH126) auf MS Medium mit 10mg/l Phosphinothricin ausgelegt. Dabei zeigte sich, dass die Nachkommen, die die Nit/bar Kassette enthielten, keimen und sich zu normalen Pflanzen entwickeln konnten. Somit ist der Nitrilase1
- 15 Promoter auch während der Keimungsphase aktiv und kann auch hier selektive Marker exprimieren.

Die Ergebnisse demonstrieren, dass die isolierte Nukleinsäuresequenz die gewünschten vorteilhaften Promotoreigenschaften

20 besitzt, d.h. sie zeigt eine Promotoraktivität, die geeignet ist Selektionsmarker effektiv zu exprimieren und weist keine Aktivität im Pollen auf.

- e) Ergebnisse der Analyse der Kanamycinresistenz der transgenen
- 25 Tabak-, Raps und Kartoffelpflanzen

Für die Untersuchung der NIT1-Promotor gestützten Vermittlung von Resistenz gegen Kanamycin wurde der Nit1-Promotor mit dem NptII

30 Gen kombiniert. Das resultierende Konstrukt pS5NitNptII wurde in die Agrobakterium tumefaciens Stämme GV3101[mp90] für die Transformation in Tabak und Raps und in C58C1[Gv2260] für die Transformation in Kartoffel transformiert.

- 35 Die resultierenden Stämme sind für die Transformation, wie unter Beispiel 2 beschrieben, eingesetzt worden. Die selektive Regeneration wurde in Gegenwart von 100 mg/l (18 mg/l Raps und 50mg/l Kartoffel) Kanamycin erreicht. Etwa 4 Wochen nach Versuchsstart wurde die Regenerationseffizienz erfasst. Nach erfolgter Sprossknospenbildung bzw. der Bewurzelung der Sprosse auf kanamycin-
- 40 haltigen Medium wurden die resistenten Transformanten gezählt und deren Zahlen verglichen. Eine PCR anhand genomischer DNA der regenerierten Pflanzen, mit der das NptII Gen nachgewiesen wurde, zeigte den hohen Anteil der transgenen Pflanzen.

45

	Nos-P	Nit1-P
Sprossbildung	100 %	88 %

46

Bewurzelte Pflanzen	80 %	66 %
Transgene Pflanzen (PCR)	92 %	92 %

- 5 Bei der Kartoffeltransformation ergab eine Analyse der Sprossknospenbildung, dass bei der Verwendung des nos-Promotors bzw. des Nit1-Promotors zur Expression des NptII Gens jeweils 85 % der Explantate Sprossknospen bilden. Die Analyse der transgenen Pflanzen mittels PCR auf Integration des NptII Gens zeigte bei 10 96 % der nosP-NptII Pflanzen und 100 % der Nit1-P-NptII Pflanzen die Transgenizität.

- Das mit dem Nitrilasepromotor ausgerüstete NptII Gen lieferte 15 nach Transformation Regenerationseffizienzen, die mit denen von NOS-Promotorkonstrukten vergleichbar sind. Die Ergebnisse demonstrieren, dass die isolierte Nukleinsäuresequenz die gewünschten vorteilhaften Promotoreigenschaften besitzt, d.h. sie zeigt eine Promotoraktivität, die geeignet ist Selektionsmarker effektiv 20 zu exprimieren.

Vergleichsbeispiel 1: Untersuchungen zur Eignung des Ubiquitin-Promotors

- 25 a) Klonierung des Ubiquitin Promotor aus *Arabidopsis thaliana*

Der Ubiquitin Promotor wurde mittels PCR aus genomischer *Arabidopsis thaliana* DNA mit den Primern ubi5 und ubi3 amplifiziert.

- 30 ubi5 (SEQ ID NO: 8) :  
5'-CCAAACCATGGTAAGTTTGTCTAAAGCTTA-3'

ubi3 (SEQ ID NO: 9):  
5'-CGGATCCTTTTGTGTTTCGTCTTCTCTCACG-3'

- 35 Reaktionsansatz:

- |    |         |  |
|----|---------|--|
|    | 37,5 µl | H <sub>2</sub> O                                     |
|    | 5 µl    | 10X Reaktionspuffer ("genomic PCR")                  |
| 40 | 4 µl    | dNTP mix (jeweils 2,5 mM)                            |
|    | 2,2 µl  | 25 mM Mg(OAc) <sub>2</sub> (Endkonzentration 1,1 mM) |
|    | 1 µl    | Primer ubi3 (10 µM)                                  |
|    | 1 µl    | Primer ubi5 (10 µM)                                  |
|    | 0,5 µl  | Pfu-turbo polymerase mix                             |
| 45 | 1 µl    | Genomische <i>Arabidopsis</i> DNA (ca. 250 ng)       |

## 47

PCR-Bedingungen:

- 1 Zyklus mit 5 min. bei 94°C  
 25 Zyklen mit 94°C für 30 sec, 52°C für 1 min. und 72°C für  
 5 1 min.  
 1 Zyklus mit 52°C für 1 min. und 72°C für 10 min.,  
 anschließend Kühlung auf 4°C bis zur Weiterverarbeitung.

Das entstandene PCR Fragment wurde als *HindIII/BamHI* Fragment  
 10 in das mit *HindIII/BamHI* geschnittene Plasmid pGUSINT37 kloniert  
 (pUBI42GUS) und mittels Sequenzanalyse verifiziert.

b) Klonierung des Ubiquitin Promotors vor das PAT-Gen

15 Für die Untersuchung der Ubiquitin-Promotor gestützten  
 Vermittlung von Resistenz gegen Phosphinothricin wurde  
 der Ubiquitin -Promotor als *BamHI/HindIII* Fragment vor  
 das Phosphinothricin Resistenzgen in das *BamHI/HindIII*  
 geschnittene Plasmid pSUN3PatNos kloniert. Das resultierende  
 20 Plasmid pSUN3UBIPat wurde unter Verwendung des *Agrobacterium*  
*tumefaciens* Stamm EHA101 zur Tabaktransformation verwendet.  
 Die selektive Regeneration der Tabakpflänzchen erfolgte  
 einerseits auf Phosphinothricin (5 mg/l) und andererseits  
 zur Kontrolle auf Kanamycin (100 mg/l).

25

c) Ergebnisse der Analyse der Phosphinothricinresistenz der  
 transgenen Tabakpflanzen

Im Gegensatz zur Selektion auf Kanamycin, die normal verlief,  
 30 konnten unter Selektion auf Phosphinothricin keine Kalli oder  
 Sprosse erhalten werden. So ist der Ubiquitin Promotor nicht  
 geeignet für die Expression eines selektiven Markers für  
 den *Agrobacterium tumefaciens*-vermittelten Gentransfer mit  
 anschließender Regeneration von Geweben.

35

Vergleichsbeispiel 2: Untersuchungen zur Eignung des Squalen  
 Synthase (SQS) Promotors

a) Klonierung des Squalen Synthase (SQS) Promotors aus  
 40 *Arabidopsis thaliana*

Der Squalen Synthase Promotor wurde mittels PCR aus genomischer  
*Arabidopsis thaliana* DNA mit den Primern sqs5 und sqs3 ampli-  
 fiziert.

45

sqs5 (SEQ ID NO: 10): 5'-GTCTAGAGGCAAACCACCGAGTGT-3'

## 48

sqs3 (SEQ ID NO: 11): 5'-CGGTACCTGTTTCCAGAAAATTTTGATTCAG-3'

## Reaktionsansatz

5	37,5 µl	H <sub>2</sub> O
	5 µl	10X Reaktionspuffer ("genomic PCR")
	4 µl	dNTP mix (jeweils 2,5 mM)
	2,2 µl	25 mM Mg(OAc) <sub>2</sub> (Endkonzentration 1,1 mM)
	1 µl	Primer sqs3 (10 µM) (10 µM)
10	1 µl	Primer sqs5 (10 µM)
	0,5 µl	Pfu-turbo polymerase mix
	1 µl	Genomische <i>Arabidopsis</i> DNA (ca. 250 ng)

## PCR-Bedingungen:

- 15
- 1 Zyklus mit 5 min. bei 94°C
  - 25 Zyklen mit 94°C für 30 sec, 52°C für 1 min. und 72°C für 1 min.
  - 1 Zyklus mit 52°C für 1 min. und 72°C für 10 min.,
- 20       anschliessend Kühlung auf 4°C bis zur Weiterverarbeitung.

Das entstandene PCR Fragment wurde als *Xba*II/*Bam*HI Fragment in das mit *Xba*II/*Bam*HI geschnittene Plasmid pGUSINT37 kloniert (pSQSPGUS) und mittels Sequenzanalyse verifiziert.

25

b) Klonierung des Squalen Synthase Promotors vor das PAT-Gen

- Für die Untersuchung der Squalen Synthase -Promotor gestützten Vermittlung von Resistenz gegen Phosphinothricin wurde der
- 30 Squalen Synthase-Promotor als *Bam*HI/*Sal*I Fragment vor das Phosphinothricin Resistenzgen in das *Bam*HI/*Sal*I geschnittene Plasmid pSUN3PatNos kloniert. Das resultierende Plasmid pSUN3SQSPat wurde unter Verwendung des *Agrobacterium tumefaciens* Stamm EHA101 zur Tabaktransformation verwendet. Die selektive
- 35 Regeneration der Tabakpflänzchen erfolgte einerseits auf Phosphinothricin (5mg/l) und andererseits zur Kontrolle auf Kanamycin (100mg/l).

- c) Ergebnisse der Analyse der Phosphinothricinresistenz der
- 40 transgenen Tabakpflanzen

Im Gegensatz zur Selektion auf Kanamycin, die normal verlief, konnten unter Selektion auf Phosphinothricin keine Kalli oder Sprosse erhalten werden. So ist der Squalensynthase-Promotor

45 nicht geeignet für die Expression eines selektiven Markers für

## 49

den *Agrobacterium tumefaciens*-vermittelten Gentransfer mit anschließender Regeneration von Geweben.

- 5 Vergleichsbeispiel 3: Testung der Promotoraktivität des Ubiquitin- und Squalen Synthase-Promotors mittels Partikelkanone

Um transient die Aktivität des Ubiquitin- und des Squalen Synthase-Promotors zu testen, wurden sterile Tabakblätter mit  
10 Plasmid-DNA der Plasmide pUBI42GUS, pSQSPGUS und pGUSINT37 mit der Partikelkanone "Biolistics" von BioRad beschossen. Dabei wurden Microcarrier (25 µg Gold, Hereus 0,3 bis 3 µm) mit 10 µg Plasmid DNA, 2,5 M CaCl<sub>2</sub>, und 0,1 M Spermidine behandelt, mit  
15 Alkohol gewaschen und bei einem Vakuum von 26 inch und einem Druck von 1100 psi auf die Blätter, die auf MS-Medium lagen, geschossen. Die Explantate wurden anschließend für 24 h in MS-Medium mit 2 % Sucrose inkubiert und dann mit X-gluc histochemisch gefärbt.

20 Der Ubiquitin-Promotor und der Squalen Synthase-Promotor zeigten hierbei im Gegensatz zum Vergleichskonstrukt pGUSINT37, wo das GUS Gen unter der Kontrolle des 35S Promotors exprimiert wurde, nur sehr wenige und sehr schwache Punkte mit GUS-Färbung.  
Dies zeigt, dass die Aktivität des Ubiquitin- und des Squalen  
25 Synthase-Promotors deutlich schwächer ist, als die des CaMV35S-Promotors.

30

35

40

45

## Patentansprüche

1. Transgene Expressionskassette zur Expression von Nukleinsäuresequenzen kodierend für Selektionsmarker enthaltend in 5'-3' Richtung
- 5
- a) einen Promotor gemäß SEQ ID NO: 1, oder ein funktionelles Äquivalent oder äquivalentes Fragment desselben, das im wesentlichen die gleiche Promotoraktivität wie der Promotor gemäß SEQ ID NO: 1 besitzt, und
- 10
- b) eine transgen zu exprimierende Nukleinsäuresequenz kodierend für einen Selektionsmarker, und
- 15
- c) eine in pflanzlichen Zellen oder pflanzlichen Organismen funktionelle Terminatorsequenz,
- wobei a) oder b) funktionell miteinander so verknüpft sind, dass eine transgene Expression der Nukleinsäuresequenz kodierend für einen Selektionsmarker in einer pflanzlichen Zelle oder einem pflanzlichen Organismus ermöglicht wird.
- 20
2. Transgene Expressionskassette nach Anspruch 1, dadurch gekennzeichnet, dass das äquivalente Fragment durch eine Sequenz mit der SEQ ID NO: 18 oder 19 beschrieben wird.
- 25
3. Transgene Expressionskassette nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass der Selektionsmarker ausgewählt ist aus der Gruppe bestehend aus positiven Selektionsmarkern, negativen Selektionsmarkern und Faktoren die einen Wachstumsvorteil gewähren.
- 30
4. Transgene Expressionskassette nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass der Selektionsmarker ausgewählt ist aus der Gruppe bestehend aus Proteinen die eine Resistenz gegen Antibiotika, Metabolismus-Inhibitoren, Herbizide oder Biozide verleihen.
- 35

40

45

## 51

5. Transgene Expressionskassette nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass der Selektionsmarker ausgewählt ist aus der Gruppe bestehend aus Proteinen die eine Resistenz verleihen gegen Phosphinothricin, Glyphosat, Bromoxynil, Dalapon, 2-Desoxyglucose-6-phosphat, Tetracycline, Ampicillin, Kanamycin, G 418, Neomycin, Paromomycin, Bleomycin, Zeocin, Hygromycin, Chloramphenicol, Sulfonylharnstoff-Herbizide, Imidazolinon-Herbizide.
- 10 6. Transgene Expressionskassette nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass der Selektionsmarker ausgewählt ist aus der Gruppe bestehend aus Phosphinothricinacetyltransferasen, 5-Enolpyruvylshikimat-3-phosphatsynthasen, Glyphosatoxidoreduktasen, Dehalogenase, 15 Nitrilasen, Neomycinphosphotransferasen, DOGR1-Genen, Acetolactatsynthasen, Hygromycinphosphotransferasen, Chloramphenicolacetyltransferasen, Streptomycinadenylyltransferasen,  $\beta$ -Lactamasen, tetA Genen, tetR Genen, Isopentenyltransferasen, Thymidinkinase, Diphtheriatoxin A, 20 Cytosindeaminase (codA), Cytochrom P450, Haloalkandehalogenasen, iaaH Gene, tms2 Gene,  $\beta$ -Glucuronidasen, Mannose-6-phosphat-Isomerasen, UDP-Galaktose-4-Epimerasen..
7. Transgene Expressionskassette nach einem der Ansprüche 1 bis 25 6, dadurch gekennzeichnet, dass der Selektionsmarker kodiert wird durch Nukleinsäuresequenzen
- i) beschrieben durch SEQ ID NO: 2 oder 3, oder
- 30 ii) beschrieben durch oder enthalten in den Sequenzen beschrieben durch die GenBank Acc.-No.: X17220, X05822, M22827, X65195, AJ028212, X17220, X05822 M22827, X65195, AJ028212, X63374, M10947, AX022822, AX022820, E01313, J03196, AF080390, AF234316, AF080389, AF234315, AF234314, 35 U00004, NC001140, X51514, AB049823, AF094326, X07645, X07644, A19547, A19546, A19545, I05376, I05373, X74325, AF294981, AF234301, AF234300, AF234299, AF234298, AF354046, AF354045, X65876, X51366, AJ278607, L36849, AB025109, AL133315.
- 40 8. Transgene Expressionskassette nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass
- a) die zu exprimierende Nukleinsäuresequenz kodierend für 45 einen Selektionsmarker mit weiteren genetischen Kontrollsequenzen funktionell verknüpft ist, oder

- b) die Expressionskassette zusätzliche Funktionselemente enthält, oder
- c) a) und b) gegeben sind.
- 5
9. Transgene Expressionskassette nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, dass die Terminatorsequenz ausgewählt ist aus der Gruppe bestehend aus dem OCS (Octopin-Synthase)-Terminator, dem NOS (Nopalin-Synthase)-Terminator, dem Terminator des Cathepsin D Inhibitor Gens aus Kartoffel und dem Terminator der Speicherproteingens VFLE1B3 aus der Ackerbohne.
- 10
10. Vektoren enthaltend eine Expressionskassette gemäß einem der Ansprüche 1 bis 9.
- 15
11. Verfahren zur Selektion transformierter pflanzlicher Zellen oder pflanzlicher Organismen, dadurch gekennzeichnet, dass eine transgene Expressionskassette enthaltend in
- 20
- 5'-3' Richtung
- a) einen Promotor gemäß SEQ ID NO: 1 oder ein funktionelles Äquivalent oder äquivalentes Fragment desselben, das im wesentlichen die gleiche Promotoraktivität wie der Promotor gemäß SEQ ID NO: 1 besitzt, und
- 25
- b) eine transgen zu exprimierende Nukleinsäuresequenz kodierend für einen Selektionsmarker, und
- 30
- c) eine in Pflanzen funktionelle Terminatorsequenz,
- wobei a) oder b) funktionell miteinander so verknüpft sind, dass eine transgene Expression der Nukleinsäuresequenz kodierend für einen Selektionsmarker in einer pflanzlichen Zelle oder pflanzlichen Organismen ermöglicht wird,
- 35
- in besagte pflanzliche Zelle oder besagten pflanzlichen Organismus eingebracht wird, die Expression des Selektionsmarkers erfolgt und eine Selektion der transformierten pflanzlichen Zellen oder pflanzlichen Organismen ausgeübt wird.
- 40
12. Verfahren nach Anspruch 11, dadurch gekennzeichnet, dass das äquivalente Fragment durch eine Sequenz mit der SEQ ID NO: 18 oder 19 beschrieben wird.
- 45

## 53

13. Verfahren nach Anspruch 11 oder 12, dadurch gekennzeichnet, dass der Selektionsmarker ausgewählt ist aus den in den Ansprüchen 2, 3, 4, 5 oder 6 beschriebenen Gruppen von Selektionsmarkern.
- 5
14. Transgener Organismus transformiert mit einer Expressionskassette gemäß den Ansprüchen 1 bis 9 oder einem Vektor gemäß Anspruch 10.
- 10
15. Transgener Organismus nach Anspruch 14 ausgewählt aus der Gruppe bestehend aus Bakterien, Hefen, Pilzen, tierischen und pflanzlichen Organismen
16. Transgener Organismus nach einem der Ansprüche 14 oder 15
- 15 ausgewählt aus der Gruppe bestehend aus Arabidopsis, Tomate, Tabak, Kartoffeln, Mais, Raps, Weizen, Gerste, Sonnenblumen, Hirse, Rübe, Roggen, Hafer, Zuckerrübe, Bohngewächse und Soja.
- 20
17. Zellkulturen, Teile oder transgenes Vermehrungsgut abgeleitet von einem transgenen Organismus nach einem der Ansprüche 14 bis 16.
18. Verwendung eines transgenen Organismus nach einem der
- 25 Ansprüche 14 bis 16 oder von diesem abgeleitete Zellkulturen, Teile oder transgenes Vermehrungsgut nach Anspruch 17 zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien.

30

35

40

45

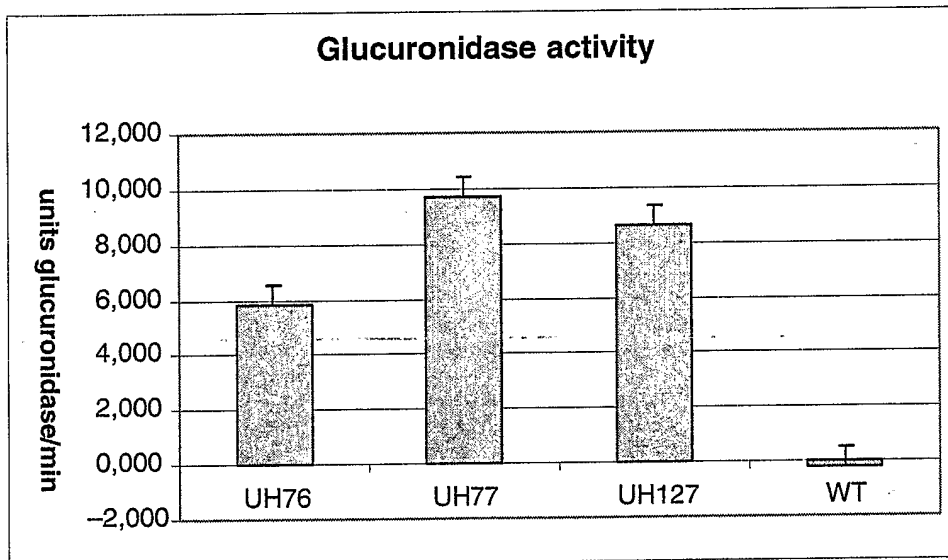


Fig. 1

## SEQUENZPROTOKOLL

<110> SunGene GmbH & Co KGaA

<120> Expressionskassetten zur transgenen Expression von Selektionsmarkern

<130> NAE 2568/01 AT

<140>

<141>

<160> 19

<170> PatentIn Ver. 2.1

<210> 1

<211> 1883

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> promoter

<222> (1)..(1883)

<400> 1

```

catcaagatc ttggtgatgt agcaagagct aagttgtact tcgattcgggt tggacattac 60
tcgagaccag atgttttaca cttgaccgta aatgagcacc cgaagaaacc ggtcacattc 120
atctcgaagg tggagaaaagc ggaagatgac tcaaaacaagt aatcggttgt gattcgtcag 180
ttcatgtcac tcctatgaag gagtcaagtt caaaaatgta tgttgagttt caaactttta 240
tgctaaactt tttttcttta ttttcgtaa taatggaaga gaaccaattc tcttgtatct 300
aaagattatc catctatcat ccaatttgag tgttcaattc tggatgttgt gttaccctac 360
attctacaac catgtagcca attattatga atctggcttt gatttcagtt gtgttctttt 420
cttttttttc tttgcatatt tgcatttaga atgtttaata attaagttac tgtatttcca 480
catacattag ttccaagaat atacatata taatttattt ttcttaaaaa tgttttggaa 540
tgactaatat tgacaacgaa aatagaagct atgctaaacc attacgtata tgtgacttca 600
catgttggtg ttttacattc cctatatata tggatggctg tcacaatcag aaacgtgatc 660
gaaaaaagac aaacagtgtt tgcataaaaa gactatttcg tttcattgac aatttgtgtt 720
tatttghtaa gaaaagtggc aaagtggaat ttgagttcct gcaagtaaga aagatgaaat 780
aaaagacttg agtgtgtgtt tttttctttt atctgaaagc tgcaatgaaa tattcctacc 840
aagcccgttt gattattaat tggggtttgg ttttcttgat gcgaactaat tggttatata 900
agaaactata caatccatgt taattcaaaa attttgattt ctctttagg aatatgattt 960
actatatgag actttctttt cgccaataat agtaaattcca aagatatttg accggaccaa 1020
aacacattga tctatttttt agtttattta atccagtttc tctgagataa ttcatthaagg 1080
aaaacttagt attaaccat cctaagatta aataggagcc aaactcacat ttcaaatatt 1140
aaataacata aatggattt aaaaaatcta tacgtcaaat tttatttatg acatttctta 1200
tttaaattta tatttaatga aatacagcta agacaaacca aaaaaaaaaat actttctaag 1260
tgggtccaaa catcaattcc gttcaatatt attaggtaga atcgtacgac caaaaaaagg 1320
taggttaata cgaattagaa acatatctat aacatagtat atattattac ctattatgag 1380
gaatcaaaat gcatcaaata tggatttaag gaatccataa aagaataaat tctacgggaa 1440
aaaaaatgga ataaattctt ttaagttttt tatttgtttt ttatttggta gttctccatt 1500
ttgttttatt tcgtttggat ttattgtgtc caaatacttt gtaaaccacc gttgtaattc 1560
ttaaacgggg ttttacttcc ttttttata tccagacataa agcatcggct ggtttaatca 1620
atcaatagat tttatttttc ttctcaatta ttagtaggtt tgatgtgaac ttacaaaaa 1680
aaacaaaaac aatcaatgc agagaaaaga aaccacgtgg gctagtcca cttgtttca 1740
tttccaccac aggttcgate ttcgttaccg tctccaatag gaaaataaac gtgaccacaa 1800
aaaaaaaaaca aaaaaaagtc tatatatgtc ttctctcaag tctctgagtg tcatgaacca 1860
aagtaaaaaa caaagactcg agt 1883

```

<210> 2

<211> 552

<212> DNA

<213> Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz: codon adapted sequence coding for bar gene

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(549)

&lt;223&gt; phosphinotricin-N-acetyltransferase (PAT)

&lt;400&gt; 2

```

atg tct ccg gag agg aga cca gtt gag att agg cca gct aca gca gcc 48
Met Ser Pro Glu Arg Arg Pro Val Glu Ile Arg Pro Ala Thr Ala Ala
  1                5                10                15

gat atg gcc gcg gtt tgt gac atc gtt aac cat tac att gag acg tct 96
Asp Met Ala Ala Val Cys Asp Ile Val Asn His Tyr Ile Glu Thr Ser
                20                25                30

aca gtg aac ttt agg aca gag cca caa aca cca caa gag tgg att gat 144
Thr Val Asn Phe Arg Thr Glu Pro Gln Thr Pro Gln Glu Trp Ile Asp
                35                40                45

gac cta gag agg ttg caa gat aga tac cct tgg ttg gtt gct gag gtt 192
Asp Leu Glu Arg Leu Gln Asp Arg Tyr Pro Trp Leu Val Ala Glu Val
                50                55                60

gag ggt gtt gtg gct ggt att gct tac gct ggg ccc tgg aag gct agg 240
Glu Gly Val Val Ala Gly Ile Ala Tyr Ala Gly Pro Trp Lys Ala Arg
                65                70                75                80

aac gct tac gat tgg aca gtt gag agt act gtt tac gtg tca cat agg 288
Asn Ala Tyr Asp Trp Thr Val Glu Ser Thr Val Tyr Val Ser His Arg
                85                90                95

cat caa agg ttg ggc cta gga tct aca ttg tac aca cat ttg ctt aag 336
His Gln Arg Leu Gly Leu Gly Ser Thr Leu Tyr Thr His Leu Leu Lys
                100                105                110

tct atg gag gcg caa ggt ttt aag tct gtg gtt gct gtt ata ggc ctt 384
Ser Met Glu Ala Gln Gly Phe Lys Ser Val Val Ala Val Ile Gly Leu
                115                120                125

cca aac gat cca tct gtt agg ttg cat gag gct ttg gga tac aca gcg 432
Pro Asn Asp Pro Ser Val Arg Leu His Glu Ala Leu Gly Tyr Thr Ala
                130                135                140

cgg ggt aca ttg cgc gcg gct gga tac aag cat ggt gga tgg cat gat 480
Arg Gly Thr Leu Arg Ala Ala Gly Tyr Lys His Gly Gly Trp His Asp
                145                150                155                160

gtt ggt ttt tgg caa agg gat ttt gag ttg cca gct cct cca agg cca 528
Val Gly Phe Trp Gln Arg Asp Phe Glu Leu Pro Ala Pro Pro Arg Pro
                165                170                175

gtt agg cca gtt acc cag atc tga 552
Val Arg Pro Val Thr Gln Ile
                180

```

&lt;210&gt; 3

&lt;211&gt; 183

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;223&gt; Beschreibung der künstlichen Sequenz: codon adapted sequence coding for bar gene

&lt;400&gt; 3

```

Met Ser Pro Glu Arg Arg Pro Val Glu Ile Arg Pro Ala Thr Ala Ala
  1                5                10                15

```

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

<210> 4  
 <211> 2013  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (1)..(2010)  
 <223> acetolactate synthase

<400> 4  
 atg gcg gcg gca aca aca aca aca aca tct tct tcg atc tcc ttc 48  
 Met Ala Ala Ala Thr Thr Thr Thr Thr Ser Ser Ser Ile Ser Phe  
 1 5 10 15  
 tcc acc aaa cca tct cct tcc tcc tcc aaa tca cca tta cca atc tcc 96  
 Ser Thr Lys Pro Ser Pro Ser Ser Ser Lys Ser Pro Leu Pro Ile Ser  
 20 25 30  
 aga ttc tcc ctc cca ttc tcc cta aac ccc aac aaa tca tcc tcc tcc 144  
 Arg Phe Ser Leu Pro Phe Ser Leu Asn Pro Asn Lys Ser Ser Ser Ser  
 35 40 45  
 tcc cgc cgc cgc ggt atc aaa tcc agc tct ccc tcc tcc atc tcc gcc 192  
 Ser Arg Arg Arg Gly Ile Lys Ser Ser Ser Pro Ser Ser Ile Ser Ala  
 50 55 60  
 gtg ctc aac aca acc acc aat gtc aca acc act ccc tct cca acc aaa 240  
 Val Leu Asn Thr Thr Thr Asn Val Thr Thr Thr Pro Ser Pro Thr Lys  
 65 70 75 80  
 cct acc aaa ccc gaa aca ttc atc tcc cga ttc gct cca gat caa ccc 288  
 Pro Thr Lys Pro Glu Thr Phe Ile Ser Arg Phe Ala Pro Asp Gln Pro  
 85 90 95

cgc aaa ggc gct gat atc ctc gtc gaa gct tta gaa cgt caa ggc gta 336  
 Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg Gln Gly Val  
 100 105 110

gaa acc gta ttc gct tac cct gga ggt gca tca atg gag att cac caa 384  
 Glu Thr Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu Ile His Gln  
 115 120 125

gcc tta acc cgc tct tcc tca atc cgt aac gtc ctt cct cgt cac gaa 432  
 Ala Leu Thr Arg Ser Ser Ser Ile Arg Asn Val Leu Pro Arg His Glu  
 130 135 140

caa gga ggt gta ttc gca gca gaa gga tac gct cga tcc tca ggt aaa 480  
 Gln Gly Gly Val Phe Ala Ala Glu Gly Tyr Ala Arg Ser Ser Gly Lys  
 145 150 155 160

cca ggt atc tgt ata gcc act tca ggt ccc gga gct aca aat ctc gtt 528  
 Pro Gly Ile Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu Val  
 165 170 175

agc gga tta gcc gat gcg ttg tta gat agt gtt cct ctt gta gca atc 576  
 Ser Gly Leu Ala Asp Ala Leu Leu Asp Ser Val Pro Leu Val Ala Ile  
 180 185 190

aca gga caa gtc cct cgt cgt atg att ggt aca gat gcg ttt caa gag 624  
 Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala Phe Gln Glu  
 195 200 205

act ccg att gtt gag gta acg cgt tcg att acg aag cat aac tat ctt 672  
 Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His Asn Tyr Leu  
 210 215 220

gtg atg gat gtt gaa gat atc cct agg att att gag gaa gct ttc ttt 720  
 Val Met Asp Val Glu Asp Ile Pro Arg Ile Ile Glu Glu Ala Phe Phe  
 225 230 235 240

tta gct act tct ggt aga cct gga cct gtt ttg gtt gat gtt cct aaa 768  
 Leu Ala Thr Ser Gly Arg Pro Gly Pro Val Leu Val Asp Val Pro Lys  
 245 250 255

gat att caa caa cag ctt gcg att cct aat tgg gaa cag gct atg aga 816  
 Asp Ile Gln Gln Gln Leu Ala Ile Pro Asn Trp Glu Gln Ala Met Arg  
 260 265 270

tta cct ggt tat atg tct agg atg cct aaa cct ccg gaa gat tct cat 864  
 Leu Pro Gly Tyr Met Ser Arg Met Pro Lys Pro Pro Glu Asp Ser His  
 275 280 285

ttg gag cag att gtt agg ttg att tct gag tct aag aag cct gtg ttg 912  
 Leu Glu Gln Ile Val Arg Leu Ile Ser Glu Ser Lys Lys Pro Val Leu  
 290 295 300

tat gtt ggt ggt ggt tgt ttg aat tct agc gat gaa ttg ggt agg ttt 960  
 Tyr Val Gly Gly Gly Cys Leu Asn Ser Ser Asp Glu Leu Gly Arg Phe  
 305 310 315 320

gtt gag ctt acg ggg atc cct gtt gcg agt acg ttg atg ggg ctg gga 1008  
 Val Glu Leu Thr Gly Ile Pro Val Ala Ser Thr Leu Met Gly Leu Gly  
 325 330 335

tct tat cct tgt gat gat gag ttg tcg tta cat atg ctt gga atg cat 1056  
 Ser Tyr Pro Cys Asp Asp Glu Leu Ser Leu His Met Leu Gly Met His  
 340 345 350

ggg act gtg tat gca aat tac gct gtg gag cat agt gat ttg ttg ttg 1104  
 Gly Thr Val Tyr Ala Asn Tyr Ala Val Glu His Ser Asp Leu Leu Leu  
 355 360 365

gcg	ttt	ggg	gta	agg	ttt	gat	gat	cgt	gtc	acg	ggt	aag	ctt	gag	gct	1152
Ala	Phe	Gly	Val	Arg	Phe	Asp	Asp	Arg	Val	Thr	Gly	Lys	Leu	Glu	Ala	
	370					375					380					
ttt	gct	agt	agg	gct	aag	att	gtt	cat	att	gat	att	gac	tcg	gct	gag	1200
Phe	Ala	Ser	Arg	Ala	Lys	Ile	Val	His	Ile	Asp	Ile	Asp	Ser	Ala	Glu	
385					390					395					400	
att	ggg	aag	aat	aag	act	cct	cat	gtg	tct	gtg	tgt	ggt	gat	gtt	aag	1248
Ile	Gly	Lys	Asn	Lys	Thr	Pro	His	Val	Ser	Val	Cys	Gly	Asp	Val	Lys	
				405					410					415		
ctg	gct	ttg	caa	ggg	atg	aat	aag	gtt	ctt	gag	aac	cga	gcg	gag	gag	1296
Leu	Ala	Leu	Gln	Gly	Met	Asn	Lys	Val	Leu	Glu	Asn	Arg	Ala	Glu	Glu	
			420					425					430			
ctt	aag	ctt	gat	ttt	gga	gtt	tgg	agg	aat	gag	ttg	aac	gta	cag	aaa	1344
Leu	Lys	Leu	Asp	Phe	Gly	Val	Trp	Arg	Asn	Glu	Leu	Asn	Val	Gln	Lys	
		435					440						445			
cag	aag	ttt	ccg	ttg	agc	ttt	aag	acg	ttt	ggg	gaa	gct	att	cct	cca	1392
Gln	Lys	Phe	Pro	Leu	Ser	Phe	Lys	Thr	Phe	Gly	Glu	Ala	Ile	Pro	Pro	
	450					455					460					
cag	tat	gcg	att	aag	gtc	ctt	gat	gag	ttg	act	gat	gga	aaa	gcc	ata	1440
Gln	Tyr	Ala	Ile	Lys	Val	Leu	Asp	Glu	Leu	Thr	Asp	Gly	Lys	Ala	Ile	
465					470					475					480	
ata	agt	act	ggt	gtc	ggg	caa	cat	caa	atg	tgg	gcg	gcg	cag	ttc	tac	1488
Ile	Ser	Thr	Gly	Val	Gly	Gln	His	Gln	Met	Trp	Ala	Ala	Gln	Phe	Tyr	
				485					490					495		
aat	tac	aag	aaa	cca	agg	cag	tgg	cta	tca	tca	gga	ggc	ctt	gga	gct	1536
Asn	Tyr	Lys	Lys	Pro	Arg	Gln	Trp	Leu	Ser	Ser	Gly	Gly	Leu	Gly	Ala	
			500					505					510			
atg	gga	ttt	gga	ctt	cct	gct	gcg	att	gga	gcg	tct	gtt	gct	aac	cct	1584
Met	Gly	Phe	Gly	Leu	Pro	Ala	Ala	Ile	Gly	Ala	Ser	Val	Ala	Asn	Pro	
		515					520						525			
gat	gcg	ata	gtt	gtg	gat	att	gac	gga	gat	gga	agc	ttt	ata	atg	aat	1632
Asp	Ala	Ile	Val	Val	Asp	Ile	Asp	Gly	Asp	Gly	Ser	Phe	Ile	Met	Asn	
	530					535					540					
gtg	caa	gag	cta	gcc	act	att	cgt	gta	gag	aat	ctt	cca	gtg	aag	gta	1680
Val	Gln	Glu	Leu	Ala	Thr	Ile	Arg	Val	Glu	Asn	Leu	Pro	Val	Lys	Val	
545					550					555					560	
ctt	tta	tta	aac	aac	cag	cat	ctt	ggc	atg	gtt	atg	caa	tgg	gaa	gat	1728
Leu	Leu	Leu	Asn	Asn	Gln	His	Leu	Gly	Met	Val	Met	Gln	Trp	Glu	Asp	
				565					570					575		
cgg	ttc	tac	aaa	gct	aac	cga	gct	cac	aca	ttt	ctc	ggg	gat	ccg	gct	1776
Arg	Phe	Tyr	Lys	Ala	Asn	Arg	Ala	His	Thr	Phe	Leu	Gly	Asp	Pro	Ala	
			580					585					590			
cag	gag	gac	gag	ata	ttc	ccg	aac	atg	ttg	ctg	ttt	gca	gca	gct	tgc	1824
Gln	Glu	Asp	Glu	Ile	Phe	Pro	Asn	Met	Leu	Leu	Phe	Ala	Ala	Ala	Cys	
		595					600					605				
ggg	att	cca	gcg	gcg	agg	gtg	aca	aag	aaa	gca	gat	ctc	cga	gaa	gct	1872
Gly	Ile	Pro	Ala	Ala	Arg	Val	Thr	Lys	Lys	Ala	Asp	Leu	Arg	Glu	Ala	
	610					615					620					
att	cag	aca	atg	ctg	gat	aca	cca	gga	cct	tac	ctg	ttg	gat	gtg	att	1920
Ile	Gln	Thr	Met	Leu	Asp	Thr	Pro	Gly	Pro	Tyr	Leu	Leu	Asp	Val	Ile	
625					630					635					640	

tgt ccg cac caa gaa cat gtg ttg ccg atg atc ccg aat ggt ggc act 1968  
 Cys Pro His Gln Glu His Val Leu Pro Met Ile Pro Asn Gly Gly Thr  
 645 650 655

ttc aac gat gtc ata acg gaa gga gat ggc cgg att aaa tac tga 2013  
 Phe Asn Asp Val Ile Thr Glu Gly Asp Gly Arg Ile Lys Tyr  
 660 665 670

<210> 5  
 <211> 670  
 <212> PRT  
 <213> Arabidopsis thaliana

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

Tyr Val Gly Gly Gly Cys Leu Asn Ser Ser Asp Glu Leu Gly Arg Phe  
 305 310 315 320  
 Val Glu Leu Thr Gly Ile Pro Val Ala Ser Thr Leu Met Gly Leu Gly  
 325 330 335  
 Ser Tyr Pro Cys Asp Asp Glu Leu Ser Leu His Met Leu Gly Met His  
 340 345 350  
 Gly Thr Val Tyr Ala Asn Tyr Ala Val Glu His Ser Asp Leu Leu Leu  
 355 360 365  
 Ala Phe Gly Val Arg Phe Asp Asp Arg Val Thr Gly Lys Leu Glu Ala  
 370 375 380  
 Phe Ala Ser Arg Ala Lys Ile Val His Ile Asp Ile Asp Ser Ala Glu  
 385 390 395 400  
 Ile Gly Lys Asn Lys Thr Pro His Val Ser Val Cys Gly Asp Val Lys  
 405 410 415  
 Leu Ala Leu Gln Gly Met Asn Lys Val Leu Glu Asn Arg Ala Glu Glu  
 420 425 430  
 Leu Lys Leu Asp Phe Gly Val Trp Arg Asn Glu Leu Asn Val Gln Lys  
 435 440 445  
 Gln Lys Phe Pro Leu Ser Phe Lys Thr Phe Gly Glu Ala Ile Pro Pro  
 450 455 460  
 Gln Tyr Ala Ile Lys Val Leu Asp Glu Leu Thr Asp Gly Lys Ala Ile  
 465 470 475 480  
 Ile Ser Thr Gly Val Gly Gln His Gln Met Trp Ala Ala Gln Phe Tyr  
 485 490 495  
 Asn Tyr Lys Lys Pro Arg Gln Trp Leu Ser Ser Gly Gly Leu Gly Ala  
 500 505 510  
 Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ser Val Ala Asn Pro  
 515 520 525  
 Asp Ala Ile Val Val Asp Ile Asp Gly Asp Gly Ser Phe Ile Met Asn  
 530 535 540  
 Val Gln Glu Leu Ala Thr Ile Arg Val Glu Asn Leu Pro Val Lys Val  
 545 550 555 560  
 Leu Leu Leu Asn Asn Gln His Leu Gly Met Val Met Gln Trp Glu Asp  
 565 570 575  
 Arg Phe Tyr Lys Ala Asn Arg Ala His Thr Phe Leu Gly Asp Pro Ala  
 580 585 590  
 Gln Glu Asp Glu Ile Phe Pro Asn Met Leu Leu Phe Ala Ala Ala Cys  
 595 600 605  
 Gly Ile Pro Ala Ala Arg Val Thr Lys Lys Ala Asp Leu Arg Glu Ala  
 610 615 620  
 Ile Gln Thr Met Leu Asp Thr Pro Gly Pro Tyr Leu Leu Asp Val Ile  
 625 630 635 640  
 Cys Pro His Gln Glu His Val Leu Pro Met Ile Pro Asn Gly Gly Thr  
 645 650 655  
 Phe Asn Asp Val Ile Thr Glu Gly Asp Gly Arg Ile Lys Tyr  
 660 665 670

<210> 6  
 <211> 34  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotide primer  
 <400> 6  
 ccggatccac tcgagtccttt gttttttact ttgg 34  
 <210> 7  
 <211> 25  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotide primer  
 <400> 7  
 catcaagatc ttggtgatgt agcaa 25  
 <210> 8  
 <211> 30  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotide primer  
 <400> 8  
 ccaaaccatg gtaagtttgt ctaaagctta 30  
 <210> 9  
 <211> 31  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotide primer  
 <400> 9  
 cggatccttt tgtgtttcgt cttctctcac g 31  
 <210> 10  
 <211> 25  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotide primer  
 <400> 10  
 gtctagaggc aaaccaccga gtggt 25  
 <210> 11  
 <211> 31  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotide primer

<400> 11  
 cggtacctgt ttccagaaaa ttttgattca g

31

<210> 12  
 <211> 7554  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz: circular  
 binary vector

<400> 12  
 ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60  
 ttattctaataa aaacgctctt ttctcttagg tttaccgcgc aatataatcct gtcaaact 120  
 gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180  
 tgattacgcc aagcttgcac gcctgcaggc cgactctaga ctagtggatc cgatcgcgc 240  
 cgggctcgag gtaccgagct cgaattcact ggccgctcgt ttacaacgac tcagagcttg 300  
 acaggaggcc cgatctagta acatagatga caccgcgcgc gataatttat cctagtttgc 360  
 gcgctatatt ttgttttcta tcgcgtatta aatgtataat tgcgggactc taatcataaa 420  
 aacctatctc ataaataacg tcatgcatta catgttaatt attacatgct taacgtaatt 480  
 caacagaaat tatatgataa tcatcgcaag accggcaaca ggattcaatc ttaagaaact 540  
 ttattgcaa atgtttgaac gatcggggat catccgggctc tgtggcggga actccacgaa 600  
 aatatccgaa cgcagcaaga tctaagcttg ggtcccgcctc agaagaactc gtcaagaagg 660  
 cgatagaagg cgatgcgctg cgaatcggga gcggcgatc cgtaaagcac gaggaagcgg 720  
 tcagcccatt cgcgcgaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 780  
 tagcggctcg ccacaccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc 840  
 accatgatata tcggcaagca ggcatcgcca tgggtcacga cgagatcctc gccgctcggc 900  
 atgcgcgcct tgagcctggc gaacagttcg gctggcgcga gccctgatg ctcttcgctc 960  
 agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgtc gatgcgatgt 1020  
 ttcgcttggt ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcatgca 1080  
 tcagccatga tggatacttt ctccggcagga gcaaggtgag atgacaggag atcctgcccc 1140  
 ggcacttcgc ccaatagcag ccagtcctt cccgcttcag tgacaacgctc gagcacagct 1200  
 gcgcaaggaa cgcccgctcg ggccagccac gatagccgcg ctgcctcgtc ctgcagttca 1260  
 ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgccccctg cgctgacagc 1320  
 cggaacacgg cggcatcaga gcagccgatt gtctgtttgt cccagtcata gccgaatagc 1380  
 ctctccacc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac 1440  
 gatccagatc cggcgcagat tttttggatt gagagtgaat atgagactct aattggatac 1500  
 cgaggggaat ttatggaacg tcagtgagc atttttgaca agaaatattt gctagctgat 1560  
 agtgacctta ggagactttt gaacgcgcaa taatggtttc tgacgtatgt gcttagctca 1620  
 ttaactcca gaaaccgcg gctgagtgcc tcttcaacg ttgcggttct gtcagttcca 1680  
 aacgtaaaac ggcttgctcc gcgtcatcgg cgggggtcat aacgtgactc ccttaattct 1740  
 ccgctcatga tcagattgtc gtttcccgc ttcagtttaa actatcagtg tttgacagga 1800  
 tctgcttg taataattgt cattagattg tttttatgca tagatgcact cgaatcagc 1860  
 caattttaga caagtatcaa acggatgtta attcagtaca ttaaagacgt ccgcaatgtg 1920  
 ttattaagtt gtctaagcgt caatttgttt acaccacaat atatacctgcc accagccagc 1980  
 caacagctcc ccgaccggca gctcggcaca aatcaccac gcgttaccac cagccggcc 2040  
 ggccgcatgg tgttgaccgt gttcggcggc attgcccagc tcgagcgttc cctaatactc 2100  
 gaccgaccc ggagcggggc cgaggccgc aaggcccagc gcgtgaagtt tggccccgc 2160  
 cctaccctca ccccggcaca gatcgcgcac gcccgcgagc tgatcgacca ggaaggccgc 2220  
 accgtgaaag aggcggctgc actgcttggc gtgcatcgct cgaccctgta ccgcgactt 2280  
 gagcgcagc aggaagtgc gccaccgag gccaggcggc gcggtgcctt ccgtgaggac 2340  
 gcattgaccg aggcgcagc cctggcggc gccgagaatg aacgccaaga ggaacaagca 2400  
 tgaaccgca ccaggacggc caggacgaac cgtttttcat taccgaagag atcgaggcgg 2460  
 agatgatcgc ggccgggtac gtgttcgagc cgcgcgcga cgtctcaacc gtgcgctgc 2520  
 atgaaatcct ggccggtttg tctgatgcca agctggcggc ctggccggcc agcttgggccg 2580  
 ctgaagaaac cgagcgcgc cgtctaaaaa ggtgatgtgt atttgagtaa aacagcttgc 2640  
 gtcagcgggt cgctgcgtat atgatgcgat gagtaaataa acaaatcgc aaggggaacg 2700  
 catgaaggtt atcgcgtgac ttaaccagaa aggcgggtca ggcaagacga ccatcgcaac 2760

ccatctagcc cgcgcctgc aactcgccgg ggccgatggt ctgttagtgc attccgatcc 2820  
 ccagggcagt gcccgcgatt gggcgccgt gcgggaagat caaccgctaa ccgttgctcg 2880  
 catcgaccgc ccgacgattg accgcgacgt gaaggccatc ggccggcgcg acttcgtagt 2940  
 gatcgacgga gcgccccagg cggcggactt ggctgtgtcc gcgatcaagg cagccgactt 3000  
 cgtgctgatt ccggtgcagc caagccctta cgacatatgg gccaccgccc acctggtgga 3060  
 gctggttaag cagcgcattg aggtcacgga tggagggcta caagcggcct ttgtcgtgtc 3120  
 gcgggcgac aaaggcacgc gcatcgccgg tgaggttgcc gaggcgctgg ccgggtacga 3180  
 gctgcccatt cttgagtccc gtatcacgca gcgcgtgagc taccaggca ctgccgccc 3240  
 cggcacaacc gttcttgaat cagaaccgga gggcgacgct gcccgcgagg tccaggcgt 3300  
 ggccgctgaa attaatcaa aactcatttg agttaatgag gtaaagagaa aatgagcaaa 3360  
 agcacaaca cgctaagtgc cggccgtccg agcgcacgca gcagcaaggc tgcaacgttg 3420  
 gccagcctgg cagacacgcc agccatgaag cgggtcaact ttcagttgcc ggcggaggat 3480  
 cacaccaagc tgaagatgta cgcggtacgc caaggcaaga ccattaccga gctgctatct 3540  
 gaatacatcg cgcagctacc agagtaaatg agcaaatgaa taaatgagta gatgaatttt 3600  
 agcggctaaa ggaggcggca tggaaaatca agaacaacca ggcaccgacg ccgtggaatg 3660  
 ccccatgtgt ggaggaacgg gcggttgcc aggcgtaagc ggctgggttg tctgccggcc 3720  
 ctgcaatggc actggaacc ccaagcccga ggaatcggcg tgagcggctc caaacatcc 3780  
 ggcccggtac aatcggcgc ggcgctgggt gatgacctgg tggagaagtt gaaggccgcg 3840  
 caggccccc agcggcaacg catcgaggca gaagcacgcc ccggtgaatc gtggcaacgc 3900  
 gccgctgac gaatccgcaa agaatcccgg caaccgcccg cagccggtgc gccgtcgatt 3960  
 aggaagccgc ccaaggcga agagcaacca gattttttcg ttccgatgct ctatgacgtg 4020  
 ggcaccgcg atagtgcag catcatggac gtggccgttt tccgtctgtc gaagcgtgac 4080  
 cgacgagctg gcgaggtgat ccgctacgag cttccagacg ggcacgtaga ggtttccgca 4140  
 gggccggccc gcatggccag tgtgtgggat tacgacctgg tactgatggc ggtttcccat 4200  
 ctaaccgaat ccatgaaccg ataccgggaa gggaaaggag acaagcccgg ccgctgttc 4260  
 cgtccacacg ttgctgacgt actcaagttc tgccggcgag ccgatggcgg aaagcagaaa 4320  
 gacgacctgg tagaaacctg cattcggtta aacaccacgc acgttgccat gcagcgtacg 4380  
 aagaaggcca agaacggccg cctggtgacg gtatccgagg gtgaagcctt gattagccgc 4440  
 tacaagatcg taaagagcga aaccgggccc ccggagtaca tcgagatcga gctagctgat 4500  
 tggatgtacc gcgagatcac agaaggcaag aaccggacg tgctgacggt tcaccccgat 4560  
 tactttttga tccatcccgg catcggccgt tttctctacc gcctggcacg ccgcgccgca 4620  
 ggcaaggcag aagccagatg gttgttcaag acgatctacg aacgcagtgg cagcgcggga 4680  
 gagttcaaga agttctgttt caccgtgccc aagctgatcg ggtcaaatga cctgccggag 4740  
 tacgatattga aggaggaggc ggggcaggct ggcccgatcc tagtcatgcg ctaccgcaac 4800  
 ctgatcagag gcgaagcadc ccgcggttcc taatgtacgg agcagatgct agggcaaat 4860  
 gccctagcag gggaaaagg tcgaaaagg ctctttctc tggatagcac gtacattggg 4920  
 aacccaaagc cgtacattgg gaaccggaac ccgtacattg ggaacccaaa gccgtacatt 4980  
 gggaaaccgt cacacatgta agtgactgat ataaaagaga aaaaaggcga tttttccgcc 5040  
 taaaactctt taaaacttat taaaactctt aaaaccgcc tggcctgtgc ataactgtct 5100  
 ggccagcgc caagcgaaga gctgcaaaaa gcgcctacc ttcggtcgtc gcgctcccta 5160  
 cgccccgccc cttcgcgtcg gcctatcgc gccgctggcc gctcaaaaat ggctggccta 5220  
 cggccaggca atctaccagg gcgcggacaa gccgcgccgt cgcactcga ccgccggcgc 5280  
 ccacatcaag gcaccctgcc tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat 5340  
 gcagctccc gagacgggtca cagcttgtct gtaagcggat gccgggagca gacaagcccg 5400  
 tcagggcgcg tcagcgggtg ttggcgggtg tcggggcgca gccatgacc agtcacgtag 5460  
 cgatagcgg gtgtatactg gcttaactat gcggcatcag agcagattgt actgagagtg 5520  
 caccatagc ggtgtgaaat accgcacaga tgcgtaagga gaaaatacc catcaggcgc 5580  
 tcttccgctt cctcgtcac tgactcgtc gcctcggctc ttcggtcgtc gcgagcggta 5640  
 tcagctcact caaaggcgg aatacggtta tccacagaa caggggataa cgcaggaaag 5700  
 aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggcgc gttgctggcg 5760  
 tttttccata ggctccgccc ccctgacgag catcacaata atcgacgctc aagtcagagg 5820  
 tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 5880  
 cgctctcctg ttccgacct gccgcttacc ggatacctgt ccgccttct cccttcggga 5940  
 agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcgggtgta ggtcgttcgc 6000  
 tccaagctgg gctgtgtgca cgaaccccc gttcagccc accgctgcgc cttatccggt 6060  
 aactatcgtc ttgagtccaa cccgtaaga cacgacttat cgcactggc agcagccact 6120  
 ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg 6180

```

cctaactacg gctacactag aaggacagta tttggatct gcgctctgct gaagccagtt 6240
accttcgtaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg 6300
ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 6360
ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgta agggattttg 6420
gtcatgcatg atatatctcc caatttgtgt agggcttatt atgcacgctt aaaaataata 6480
aaagcagact tgacctgata gtttggctgt gagcaattat gtgcttagtg catctaacgc 6540
ttgagttaag ccgcgccgcg aagcggcgctc ggcttgaacg aatttctagc tagacattat 6600
ttgcccacta ccttgggtgat ctgcgctttc acgtagtggg caaattcttc caactgatct 6660
gcgcgcgagg ccaagcgatc ttcttcttgt ccaagataag cctgtctagc ttcaagtatg 6720
acgggctgat actgggcccgg caggcgctcc attgccagc cggcagcgac atccttcggc 6780
gcgatttttg cggttactgc gctgtaccaa atgcgggaca acgtaagcac tacatttcgc 6840
tcatcgccag cccagtcggg cggcgagttc catagcgta aggtttcatt tagcgcctca 6900
aatagatcct gttcaggaac cggatcaaag agttcctccg ccgctggacc taccaaggca 6960
acgctatggt ctcttgcttt tgtcagcaag atagccagat caatgtcgat cgtggctggc 7020
tcgaagatac ctgcaagaat gtcattgcgc tgccattctc caaattgcag ttcgcgctta 7080
gctggataac gccacggaat gatgtcgtcg tgcacaacaa tggtgacttc tacagcgcgg 7140
agaatctcgc tctctccagg ggaagccgaa gtttccaaaa ggctcgttgat caaagctcgc 7200
cgcgttgttt catcaagcct tacggtcacc gtaaccagca aatcaatc actgtgtggc 7260
ttcaggccgc catccactgc ggagcgcgtac aaatgtacgg ccagcaacgt cggttcgaga 7320
tgccgctcga tgacgccaac tacctctgat agttgagtcg atacttcggc gatcaccgct 7380
tccccatga tgtttaactt tgttttaggg cgactgccct gctgcgtaac atcgttgctg 7440
ctccataaca tcaaacatcg acccagggcg taacgcgctt gctgcttggg tgcccggagg 7500
atagactgta ccccaaaaa acagtcataa caagccatga aaaccgccac tgcg 7554

```

<210> 13

<211> 6080

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: circular  
binary vector

<400> 13

```

ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttattctaataaacgctctt ttctcttagg tttaccggcc aatatacctt gtcaaacact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttgcat gcctgcaggt cgactctaga ctagtggatc cgatatcgcc 240
cgggctcgag gtaccgagct cgaattcact ggcgctcggt ttacaacgac tcaggatcct 300
gtcaaacact gatagtttaa actgaaggcg ggaaacgaca atctgatcat gagcggagaa 360
ttaagggagt cacgttatga cccccgcga tgacgcggga caagccggtt tacgtttggg 420
actgacagaa ccgcaacggt gaaggagcca ctacgcccgc ggtttctgga gtttaatgag 480
ctaagcacat acgtcagaaa ccattattgc gcgttcaaaa gtcgcctaag gtcactatca 540
gctagcaaat atttcttgct aaaaatgctc cactgacggt ccataaattc ccctcgggat 600
ccaattagag tctcatattc actctcaatc caaataatct cgacatgtct ccggagagga 660
gaccagttga gattaggcca gctacagcag ccgatatggc cgcggtttgt gacatcgta 720
accattacat tgagacgtct acagtgaact ttaggacaga gccacaaaca ccacaagagt 780
ggattgatga cctagagagg ttgcaagata gatacccttg gttggttgcg gaggttgagg 840
gtggttgggc tggattgct tacgctgggc cctggaaggc taggaacgct tacgattgga 900
cagttgagag tactgtttac gtgtcacata ggcatcaaag gttgggccta ggatctacat 960
tgtacacaca tttgcttaag tctatggagg cgcaagggtt taagtctgtg gttgctgta 1020
taggccttcc aaacgatcca tctgtaggt tgcatgaggc tttgggatac acagcgcggg 1080
gtacattgcg cgcggctgga tacaagcatg gtggatggca tgatggtggg ttttggcaaa 1140
gggattttga gttgccagct cctccaaggc cagttaggcc agttaccag atctgagtcg 1200
atcgaccgat cttgctgcgt tcggataatt tcgtagggtt cccgccacag aaccggatga 1260
tccccgatcg ttcaaacatt tggcaataaa gtttcttaag attgaatcct gttgcccgtc 1320
ttgcatgat tatcatataa tttctgttga attacgttaa gcatgtaata attaacatgt 1380
aatgcatgac gttatattatg agatgggttt ttatgattag agtcccgcaa ttatacattt 1440
aatacgcgat agaaaacaaa atatagcgcg caaactagga taaattatcg cgcgcggtgt 1500

```

catctatgtt actagatcgg gcctcctgtc aagctggctg cttggtaata attgtcatta 1560  
gattgttttt atgcatagat gcaactcgaat tcagccaatt ttagacaagt atcaaacgga 1620  
tgtaattca gtacattaaa gacgtccgca atgtgttatt aagttgtcta agcgtcaatt 1680  
tgtttacacc acaatatatc ctgccaccag ccagccaaca gctccccgac cggcagctcg 1740  
gcacaaaatc accacgcgcg tctaaaaagg tgatgtgtat ttgagtaaaa cagcttgcgt 1800  
catgcggtcg ctgcttatat gatgctgatg gtaaataaac aaatacgcaa ggggaacgca 1860  
tgaaggttat cgctgtactt aaccagaaaag gcggtcagg caagacgacc atcgcaacc 1920  
atctagcccg cgccctgcaa ctgcgcccgg ccatgttct gttagtcgat tccgatcccc 1980  
agggcagtg cgcgattgg gcggcctg gcggaagatca accgtaacc gttgtcggca 2040  
tcgaccgcc gacgattgac cgcgacgtg aggccatcgg ccggcgcgac ttcgtagtga 2100  
tcgacggagc gccccaggcg gcggacttgg ctgtgtccgc gatcaaggca gccgacttcg 2160  
tgctgattcc ggtgcagcca agcccttacc acatatgggc caccgccgac ctggtggagc 2220  
tggtaagca ggcgattgag gtcacggatg gaaggctaca agcggccttt gtcgtgtcgc 2280  
gggcgatcaa aggcacgcgc atcggcggg aggttgccga ggcgctggcc gggtagcagc 2340  
tgcccattct tgagtcccgt atcagcgagc gcgtgagcta cccaggcact gccgcccgcg 2400  
gcacaaccgt tcttgaatca gaaccgagg gcgacgctg ccgagaggtc caggcgtg 2460  
ccgctgaaat taaatcaaaa ctcatgtgag ttaatgagg aaagagaaaa tgagcaaaaag 2520  
cacaacacg ctaagtgcg gccgtccgag cgcacgcagc agcaaggctg caacgttggc 2580  
cagcctggca gacacgccag ccatgaagcg ggtcaacttt cagttgccgg cggaggatca 2640  
caccaagctg aagatgtacg cggtagcca aggcaagacc attaccgagc tgctatctga 2700  
atacatcgcg cagctaccag agtaaagag caaatgaata aatgagtaga tgaattttag 2760  
cggctaaagg aggcggcatg gaaaatcaag aacaaccagg caccgacgcc gtggaatgcc 2820  
ccatgtgtgg aggaacgggc ggttggccag gcgtaagcgg ctgggttgtc tgccggccct 2880  
gcaatggcac tggaaacccc aagcccagg aatcggcgtg acggtcgcaa accatccggc 2940  
ccggtacaaa tcggcgcggc gctgggtgat gacctggtg agaagttgaa ggcgcgcag 3000  
gccgcccagc ggcaacgcat cgaggcagaa gcacgcccc gtgaatcgtg gcaaccggcc 3060  
gctgatcgaa tccgcaaaga atcccggcaa ccgcccggc cgggtgcgcc gtcgattagg 3120  
aagccgccca agggcgacga gcaaccagat ttttctgtc cgatgctcta tgacgtgggc 3180  
accgcgata gtcgcagcat catggacgtg gccgttttcc gtctgtcgaa gcgtgaccga 3240  
cgagctggcg aggtgatccg ctacgagctt ccagacgggc acgtagagg ttccgcaggg 3300  
ccggccggca tggccagtg gtgggattac gacctggtac tgatggcggt tcccattcta 3360  
accgaatcca tgaaccgata ccgggaagg aagggagaca agcccggccg cgtgttccgt 3420  
ccacacgttg cggacgtact caagttctgc cggcgagcc atggcggaaa gcagaaagac 3480  
gacctggtag aaacctgcat tcggttaaac accacgcagc ttgccatgca gcgtacgaag 3540  
aaggccaaga acggccgct ggtgacggta tccgagggtg aagccttgat tagccgctac 3600  
aagatcgtaa agagcgaac cgggcggccg gactacatcg agatcgagct agctgattgg 3660  
atgtaccgag agatcacaga aggcaagaac ccggacgtg tgacggttca cccgattac 3720  
ttttgatcg atcccggcat cggccgttt ctctaccgcc tggcacgccc cgccgagggc 3780  
aaggcagaag ccagatggtt gttcaagac atctacgaac gcagtggcag cgccggagag 3840  
ttcaagaagt tctgtttcac cgtgcgcaag ctgatcgggt caaatgacct gccggagtac 3900  
gatttgaagg aggaggcggg gcaggctggc ccgatcctag tcatgctgta ccgcaacctg 3960  
atcgaggggc aagcatccgc cggttcctaa tgtacggagc agatgctagg gcaattgcc 4020  
ctagcagggg aaaaaggctc aaaaggctct tttcctgtgg atagcacgta cattgggaac 4080  
ccaaagccgt acattgggaa ccggaaccgc tacattggga acccaagcc gtacattggg 4140  
aaccggtcac acatgtaagt gactgatata aaagagaaaa aaggcgattt tccgcctaa 4200  
aactctttaa aacttattaa aactcttaa acccgcctgg cctgtgcata actgtctggc 4260  
cagcgcacag ccgaagagct gcaaaaagcg cctaccctt ggtcgtcg ctccctacgc 4320  
cccgcgctt cgcgctcgcc tatcgcggcc tatgcggtg gaaataccgc acagatgcgt 4380  
aaggagaaaa taccgcatca ggcgctctt cgttctctc ctactgact cgctgcgctc 4440  
ggtcgttcgg ctgcgcgag cggtatcagc tcaactcaaag gcgtaatac ggttatccac 4500  
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 4560  
ccgtaaaaag gccgcgttgc tggcgtttt ccataggctc cgccccctg acgagcatca 4620  
caaaaatcga cgctcaagtc agaggtggcg aaaccgcaca ggactataaa gataccaggc 4680  
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 4740  
cctgtccgcc tttctccctt cgggaagcgt ggcgcttct catagctcac gctgtaggta 4800  
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 4860  
gccccaccgc tgcgcttat ccggtaacta tcgtcttgag tccaaccgg taagacacga 4920

```

cttatcgcca ctggcagcag ccaactggtaa caggattagc agagcgaggt atgtaggcgg 4980
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg 5040
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gtttagcggcg tcggcttgaa 5100
cgaatttcta gctagacatt atttgccgac taccttgggtg atctcgcctt tcacgtagtg 5160
gacaaattct tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata 5220
agcctgtcta gcttcaagta tgacgggctg atactgggcc ggcaggcgtt ccattgcca 5280
gtcggcagcg acatccttcg gcgcgatctt gccggttact gcgctgtacc aaatgcggga 5340
caacgtaagc actacatttc gctcatcgcc agcccagtcg ggcggcgagt tccatagcgt 5400
taaggtttca ttttagcgct caaatagatc ctgttcagga accggatcaa agagtccctc 5460
cgccgctgga cctaccaagg caacgctatg ttctcttgct tttgtcagca agatagccag 5520
atcaatgtcg atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc 5580
tccaaattgc agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac 5640
aatggtgact tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttccaa 5700
aaggtcgttg atcaaagctc gccgctgtgt ttcatcaagc cttacggta ccgtaaccag 5760
caaatcaata tcaactgtgt gcttcaggcc gccatccact gcggagccgt acaaatgtac 5820
ggccagcaac gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt 5880
cgatacttcg gcgatcaccg cttcccccat gatgtttaac tttgttttag ggcgactgcc 5940
ctgctgcgta acatcgttgc tgctccataa catcaaacat cgaccacgg cgtaacgcgc 6000
ttgctgcttg gatgcccag gcatagactg taccocaaaa aaacatgtca taacaagaag 6060
ccatgaaaac cgccactgcg                                     6080

```

<210> 14

<211> 8441

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: circular  
binary vector

<400> 14

```

ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttatttctaat aaacgctctt ttctcttagg ttaccgcccc aatatatcct gtcaaact 120
gatagttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaca gctatgacca 180
tgattacgcc aagcttgcac gccagcttga caggaggccc gatctagtaa catagatgac 240
accgcgcgcg ataatttctc ctagtcttgc cgctatattt tgttttctat cgcgtattaa 300
atgtataatt gcgggactct aatcataaaa accatctca taaataacgt catgcattac 360
atgttaatta ttacatgctt aacgtaattc aacagaaatt atatgataat catcgcaaga 420
ccggcaacag gattcaatct taagaaactt tattgccaaa tgtttgaacg atcggggatc 480
atccgggtct gtggcgggaa ctccacgaaa atatccgaac gcagcaagat cggctgatcg 540
actcagatct gggtaactgg cctaactggc cttggaggag ctggcaactc aaaatccctt 600
tgccaaaaac caacatcatg ccatccacca tgcttgatc cagccgcgcg caatgtacc 660
cgcgctgtgt atcccaaagc ctcatgcaac ctaacagatg gatcgtttg aaggcctata 720
acagcaacca cagacttaaa accttgcgcc tccatagact taagcaaag tgtgtacaat 780
gtagatccta ggccaaact ttgatgccta tgtgacacgt aaacagtact ctcaactgtc 840
caatcgtaag cgttcctagc cttccagggc ccagcgtaag caataccagc cacaacacc 900
tcaacctcag caaccaacca aggtatctc tcttgaacc tctctaggtc atcaatccac 960
tcttgtggtg tttgtggctc tgcctaaag ttcactgtag acgtctcaat gtaatggtta 1020
acgatgtcac aaaccgcggc catatcggct gctgtagctg gcctaactc aactggtctc 1080
ctctccggag acatgtcgac tctagactag tggatccgat atcgcccggg ctcgaggtag 1140
cgagctcgaa tcaactggcc gtcgttttac aacgactcag agcttgacag gaggcccgat 1200
ctagtaacat agatgacacc gcgcgcgata atttatccta gtttgcgccc tatattttgt 1260
tttctatcgc gtattaaatg tataattgcy ggactctaat cataaaaacc catctcataa 1320
ataacgtcat gcattacatg ttaattatta catgcttaac gtaattcaac agaaattata 1380
tgataatcat cgcaagaccg gcaacaggat tcaatcttaa gaaactttat tgccaaatgt 1440
ttgaacgatc ggggatcatc cgggtctgtg gcgggaaactc cacgaaaata tccgaacgca 1500
gcaagatcta gagcttgggt cccgctcaga agaactcgtc aagaaggcga tagaaggcga 1560
tgcgctgcga atcgggagcg gcgataccgt aaagcacgag gaagcgggtca gccattcgc 1620
cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggctccgcca 1680

```

caccagccg gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg 1740  
gcaagcagcc atcgccatgg gtcacgacga gatcctcgcc gtcgggcatg cgcgccctga 1800  
gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcatcctgat 1860  
cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggt 1920  
cgaatgggca ggtagccgga tcaagcgtat gcagccgccc cattgcatca gccatgatgg 1980  
atactttctc ggaggagca aggtgagatg acaggagatc ctgccccggc acttcgcca 2040  
atagcagcca gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc 2100  
ccgtcgtggc cagccacgat agccgcgctg cctcgctcctg cagttcattc agggcaccgg 2160  
acaggtcggc cttgacaaaa agaaccgggc gccctcgcc tgacagccgg aacacggcgg 2220  
catcagagca gccgattgtc tgttggtccc agtcatagcc gaatagcctc tccaccaag 2280  
cggccggaga acctgctgct aatccatctt gttcaatcat gcgaaacgat ccagatccgg 2340  
tgcagattat ttggattgag agtgaatatg agactctaag tggataccga ggggaattta 2400  
tggaacgtca gtggagcatt tttgacaaga aatatttgcg agctgatagt gaccttaggc 2460  
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520  
accgcggctc gagtggctcc ttcaacgttg cggttctgct agttccaaac gtaaaacggc 2580  
ttgtcccgcg tcatcggcgg gggcataaac gtgactccct taattctccg ctcatgatca 2640  
gattgtcgtt tcccgccttc agtttaaact atcagtgttt gacaggatcc tgcttggtaa 2700  
taattgtcat tagattgttt ttatgcatag atgcactcga aatcagccaa ttttagacaa 2760  
gtatcaaacg gatgttaatt cagtacatta aagacgtccg caatgtgtta ttaagttgtc 2820  
taagcgtcaa tttgtttaca ccacaatata tcttgccacc agccagccaa cagctccccg 2880  
accggcagct cggcacaaaa tcaccacgcg ttaccaccac gccggcggc cgcatgggtg 2940  
tgaccgtggt cgccggcatt gccgagttcg agcgttccct aatcatcgac cgcaccggga 3000  
gcgggcgcga ggccgccaag gcccgaggcg tgaagtttgg cccccgccct accctcacc 3060  
cggcacagat cgcgcacgcc cgcgagctga tgcaccagga aggcgcgacc gtgaaagagg 3120  
cggctgcact gcttggcgtg catcgctcga cctgtaccg cgcacttgag cgcagcgagg 3180  
aagtgacgcc caccgagggc aggcggcgcg gtgccttccg tgaggacgca ttgaccgagg 3240  
ccgacgccct ggccggccgc gagaatgaac gccaaagagga acaagcatga aaccgacca 3300  
ggacggccag gacgaaccgt ttttcattac cgaagagatc gaggcggaga tgatcgcggc 3360  
cgggtacgtg ttcgagccgc ccgcgcacgt ctcaaccgtg cggctgcatg aaatcctggc 3420  
cggtttgtct gatgccaaagc tggcggcctg gccggccagc ttggccgctg aagaaaccga 3480  
gcgcccggcgt ctaaaaaggt gatgtgtatt tgagtaaaac agcttgcgct atgcggctgc 3540  
tgcgtatatg atgcgatgag taaataaaca aatacgcaag gggaaacgat gaaggttatc 3600  
gctgtactta accagaaagg cgggtcaggc aagacgacca tgcgaaccca tctagcccg 3660  
gccctgcaac tcgcccgggc cgatgttctg ttagtcgatt ccgatcccca gggcagtgcc 3720  
cgcgattggg cggccgtgcg ggaagatcaa ccgctaaccg ttgtcggcat cgaccgcccg 3780  
acgattgacc gcgacgtgaa ggccatcggc cggcgcgact tctgtagtgat cgacggagcg 3840  
ccccaggcgg cggacttggc tgtgtccgcg atcaaggcag ccgacttctg gctgattccg 3900  
gtgcagccaa gcccttacga catatgggccc accgccgacc tggtgagct ggttaagcag 3960  
cgcattgagg tcacggatgg aaggctaaa gggccctttg tctgttcgcg ggcgatcaaa 4020  
ggcacgcgca tcggcgggtga ggttgccgag gcgctggccg ggtacgagct gccattctt 4080  
gagtcgccga tcacgcagcg cgtgagctac ccaggcactg ccgcccggc cacaaccgtt 4140  
cttgaatcag aaccggaggc cgacgctgcc cgcgaggtcc aggcgctggc cgctgaaatt 4200  
aaatcaaac tcatttgagt taatgaggtg aagagaaaat gagcaaaagc acaaacacgc 4260  
taagtgcggc ccgtccgagc gcacgcagca gcaaggctgc aacgttggc agcctggcag 4320  
acacgccagc catgaagcgg gtcaactttc agttgccggc ggaggatcac accaagctga 4380  
agatgtacgc ggtacgcca ggcaagacca ttaccgagct gctatctgaa tacatcgcgc 4440  
agctaccaga gtaaatgagc aatgaataa atgagtagat gaattttagc ggctaaagga 4500  
ggcggcatgg aaaatcaaga acaaccaggc accgacgccc tggaatgccc catgtgtgga 4560  
ggaacgggcg gttggccagg cgtgagctac ccaggcactg ccgcccggc caatggcact 4620  
ggaacccccca agcccgagga atcggcgtga gcggtcgcaa accatccggc ccggtacaaa 4680  
tcggcgcggc gctgggtgat gacctggtgg agaagttgaa ggccgcgag gcccccagc 4740  
ggcaacgcat cgaggcagaa gcacgccccg gtgaatcgtg gcaagcggc gctgatcgaa 4800  
tccgcaaaga atcccggcaa ccgcccggc cgggtgcgccc gtcgattagg aagccgcca 4860  
agggcgacga gcaaccagat tttttcgttc cgtgtctcta tgacgtgggc acccgcgata 4920  
gtcgcagcat catggacgtg gcggttttcc gtctgtcgaa gcgtgaccga cgagctggcg 4980  
aggtgatccg ctacgagctt ccagacgggc acgtagaggt ttccgcaggg ccggccggca 5040  
tggccagtggt gtgggattac gacctggtac tgatggcggt tccccatcta accgaatcca 5100

tgaaccgata	ccgggaaggg	aagggagaca	agccccggccg	ctgtttccgt	ccacacgttg	5160
cggacgtact	caagttctgc	cggcgagccg	atggcggaaa	gcagaaagac	gacctggtag	5220
aaacctgcat	tccgttaaac	accacgcacg	ttgccatgca	gcgtacgaag	aaggccaaga	5280
acggccgcct	ggtgacggta	tccgaggggtg	aagccttgat	tagccgctac	aagatcgtaa	5340
agagcgaaac	cgggcggccg	gagtacatcg	agatcgagct	agctgattgg	atgtaccgcg	5400
agatcacaga	aggcaagaac	cgggacgtgc	tgacggttca	ccccgattac	ttttgatcg	5460
atcccggcat	cggccgtttt	ctctaccgcc	tggcacgcgc	cgccgcaggg	aaggcagaag	5520
ccagatgggt	gttcaagacg	atctacgaac	gcagtggcag	cgccggagag	ttcaagaagt	5580
tctgtttcac	ctgtgcgaag	ctgatcgggt	caaatgacct	gccggagtac	gatttgaagg	5640
aggaggcggg	gcaggctggc	ccgatcctag	tcatgcgcta	ccgcaacctg	atcgagggcg	5700
aagcatccgc	cggttcctaa	tgtacggagc	agatgctagg	gcaaattgcc	ctagcagggg	5760
aaaaaggtcg	aaaaggtctc	tttctgtgg	atagcacgta	cattgggaac	ccaaagccgt	5820
acattgggaa	ccggaaccgg	tacattggga	acccaaagcc	gtacattggg	aaccggtcac	5880
acatgtaagt	gactgatata	aaagagaaaa	aaggcgattt	ttccgcctaa	aactctttaa	5940
aacttattaa	aactcttaaa	acccgcctgg	cctgtgcata	actgtctggc	cagcgcacag	6000
ccgaagagct	gcaaaaagcg	cctacccttc	ggtcgtcgcg	ctccctacgc	cccgcgcctt	6060
cgcgtcggcc	tatcgcggcc	gctggccgct	caaaaatggc	tggcctacgg	ccaggcaatc	6120
taccagggcg	cggacaagcc	gcgcctcgc	cactcgaccg	ccggcgccca	catcaaggca	6180
ccctgcctcg	cgcttttcgg	tgatgacggt	gaaaacctct	gacacatgca	gctcccggag	6240
acggtcacag	cttgtctgta	agcggatgcc	gggagcagac	aagcccgtca	gggcgcgtca	6300
gcgggtggtg	gcgggtgctg	gggagcagcc	atgacccagt	cacgtagcga	tagcggagtg	6360
tatactggct	taactatgcg	gcatcagagc	agattgtact	gagagtgcac	catatgcggt	6420
gtgaaatacc	gcacagatgc	gtaaggagaa	aataccgcat	caggcgctct	tccgcttctt	6480
cgctcactga	ctcgtcgcgc	tccgtcgttc	ggctgocggc	agcggtatca	gctcactcaa	6540
aggcggtaat	acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	6600
aaggccagca	aaaggccagg	aaccgtaaaa	aggcgcgctt	gctggcgttt	ttccataggg	6660
tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaacccga	6720
caggactata	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	6780
cgaccctgcc	gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	6840
ctcatagctc	acgctgtagg	tatctcagtt	cggtgtaggt	cgctcgtctc	aagctgggct	6900
gtgtgcacga	acccccgctt	cagcccagcc	gctgcgcctt	atccggtaac	tatcgtcttg	6960
agtccaaccc	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	7020
gcagagcgag	gtatgtaggc	ggtgctacag	agttcttgaa	tggtggcct	aactacggct	7080
acactagaag	gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	7140
gagttggtag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	7200
gcaagcagca	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	7260
cggggtctga	cgctcagtgg	aacgaaaact	cacgttaagg	gattttggtc	atgcatgata	7320
tatctcccaa	tttgtgtagg	gcttattatg	cacgcttaaa	aataataaaa	gcagacttga	7380
cctgatagtt	tggctgtgag	caattatgtg	cttagtgcat	ctaacgcttg	agttaagccg	7440
cgccgcgaag	cggcgtcggc	ttgaacgaat	ttctagctag	acattatttg	ccgactacct	7500
tggatgatctc	gcctttcacg	tagtggacaa	attcttccaa	ctgatctgcg	cgcgaggcca	7560
agcgatcttc	ttctgttcca	agataagcct	gtctagcttc	aagtatgacg	ggctgatact	7620
gggcccggcag	gcgctccatt	gcccagtcgg	cagcgcacatc	cttcggcgcg	atcttgccgg	7680
ttactgcgct	gtaccaaag	cgggacaacg	taagcactac	atctcgtca	tcgccagccc	7740
agtcggggcg	cgagttccat	agcgttaagg	tttcatttag	cgctcaaat	agatcctggt	7800
caggaaccgg	atcaaagagt	tctcccgccg	ctggacctac	caaggcaacg	ctatgttctc	7860
ttgcttttgt	cagcaagata	gccagatcaa	tgctgatcgt	ggctggctcg	aagataacctg	7920
caagaatgtc	attgcgctgc	cattctccaa	attgcagttc	gcgcttagct	ggataacgcc	7980
acggaatgat	gtcgtcgtgc	acaacaatgg	tgacttctac	agcgcggaga	atctcgtctt	8040
ctccagggga	agccgaagtt	tccaaaaggt	cgttgatcaa	agctcgcgcg	gttgtttcat	8100
caagccttac	ggtcaccgta	accagcaaat	caatatcact	gtgtggcttc	aggccgccat	8160
ccactgcgga	gccgtacaaa	tgtacggcca	gcaacgtcgg	ttcgagatgg	cgctcgatga	8220
cgccaactac	ctctgatagt	tgagtcgata	cttcggcgat	caccgcttcc	cccatgatgt	8280
ttaactttgt	tttagggcga	ctgccctgct	gcgtaacatc	gcttgcgtctc	cataacatca	8340
aacatcgacc	cacggcgtaa	cgcgcttgct	gcttggatgc	ccgaggcata	gactgtacct	8400
caaaaaaaca	gtcataacaa	gccatgaaaa	ccgccactgc	g		8441

<210> 15  
 <211> 8384  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz: circular  
 binary expression vector pSUN5NPTIICAT

<400> 15  
 ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60  
 ttattctaataaacgctctt ttctcttagg tttaccggcc aatataatcct gtcaaacact 120  
 gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180  
 tgattacgcc aagcttgcac gccgatcccc cccactccgc cctacactcg tataatatatg 240  
 cctaaacctg ccccgttcct catatgtgat attattatth cattattagg tataagatag 300  
 taaacgataa ggaaagacaa tttattgaga aagccatgct aaaatataga tagatatacc 360  
 ttagcagggtg tttatthttac aacataacat aacatagtag ctagccagca ggcaggctaa 420  
 aacatagtat agtctatctg caggggggtac ggtcgcactct agactagtgg atccgctcga 480  
 gctagcttgg gtcccgcctca gaagaactcg tcaagaaggc gatagaaggc gatcgcctgc 540  
 gaatcgggag cggcgatacc gtaaagcacg aggaagcggg cagcccattc gccgccaagc 600  
 tcttcagcaa tatcacgggt agccaacgct atgtcctgat agcgggtccgc cacaccagc 660  
 cggccacagt cgatgaatcc agaaaagcgg ccatthttcca ccatgatatt cggcaagcag 720  
 gcatcgccat gggtcacgac gagatcctcg ccgtcgggca tgcgcgcctt gaggcctggcg 780  
 aacagttcgg ctggcgcgag cccctgatgc tcttcgtcca gatcatcctg atcgacaaga 840  
 ccggcttcca tccgagtacg tgctcgtctg atgcgatgth tgcgttgggtg gtcgaatggg 900  
 caggtagccg gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatacttc 960  
 tcggcaggag caaggtgaga tgacaggaga tcttgccccg gcaacttcgca caatagcagc 1020  
 cagtccttc ccgcttcagt gacaacgctg agcacagctg cgcaaggaac gcccgctcgtg 1080  
 gccagccacg atagccgcgc tgccctgctc tgcaattcat tcagggcacc ggacaggctc 1140  
 gtcttgacaa aaagaaccgg ggcgccctgc gctgacagcc ggaacacggc ggcacagag 1200  
 cagccgattg tctgttgtgc ccagtcatag ccgaatagcc tctccacca agcggccgga 1260  
 gaacctgcgt gcaatccatc ttgttcaatc caagctccca tgggcccctcg actagagtcg 1320  
 agatccgata tcgcccgggc tcgaggtacc gagctcgaat tcaactggccg tcgthttaca 1380  
 acgactcagc cagcttgaca ggaggcccga tctagtaaca tagatgacac cgcgcgcgat 1440  
 aatthtatcct agthttgcgc ctatathttg thttctatcg cgtatthaat gtataattgc 1500  
 gggactctaa tcataaaaac ccatctcata aataacgtca tgcattacat gthattatt 1560  
 acatgcttaa cgtaattcaa cagaaattat atgataatca tcgcaagacc ggcaacagga 1620  
 ttcaatctta agaaacttha ttgccaaatg thtgaacgat cggggatcat cgggtctgt 1680  
 ggcgggaact ccacgaaaat atccgaacgc agcaagatcg gtcgatcgac tcagatctgg 1740  
 gtaactggcc taactggcct tggaggagct ggcaactcaa aatccctthg ccaaaaacca 1800  
 acatcatgcc atccaccatg cthgtatcca gccgcgcgca atgtaccccg cgtgtgtat 1860  
 cccaaagcct catgcaacct aacagatgga tgcthttgga ggcctataac agcaaccaca 1920  
 gactthaaaac cthgcccctc catagactta agcaaatgtg tgtacaatgt agatcctagg 1980  
 cccaacctth gatgcctatg tgacacgtaa acagttactc caactgtcca atcgtaagcg 2040  
 thcctagcct tccagggccc agcgtaaagca ataccagcca caacaccctc aacctcagca 2100  
 accaaccaag ggtatctatc thgcaacctc tctaggatcat caatccactc thgtgggtgt 2160  
 tgtggctctg tcctaaagth cactgtagac gthctcaatgt aatggthaac gatgtcaca 2220  
 accgcggcca tatcggtcgc tgtagctggc ctaatctcaa ctggctcctc ctccggagac 2280  
 atgtcgagat thttgggatt gagagtgaat atgagactct aatgggatac cgaggggaa 2340  
 thtatggaacg tcagtgagc atthttgaca agaaatath gctagctgat agtgacctta 2400  
 ggcgacttht gaacgcgcaa thaatggtht tgacgtatgt gcttagctca thaaactcca 2460  
 gaaacccgcg gctgagtggtc thcttcaacg thgctgtctc gtcagthcca aacgthaaac 2520  
 ggtthgtccc gcgtcatcgg cgggggtcat aacgtgactc cthattctc ccgctcatga 2580  
 tcagattgtc gthttcccgc thcagthtaa actatcagtg thtgacagga thctgcttg 2640  
 thataattgt cattagattg ththttatgca tagatgcact cgaaatcagc caatthttaga 2700  
 caagtatcaa acggatgth atthcagthca thaaagacgt ccgcaatgtg ththattagth 2760  
 gthtaagcgt caatthgtth acaccacaat atatcctgcc accagccagc caacagctcc 2820  
 ccgaccggca gthcggcaca aatthaccac gctthaccac cagccggcc ggcgcgatg 2880

tgttgaccgt	gttcgccggc	attgccgagt	tcgagcgttc	cctaatac	gaccgcaccc	2940
ggagcgggcg	cgaggccgcc	aaggcccgag	gcgtgaagtt	tggccccgc	cctaccctca	3000
ccccggcaca	gatcgcgcac	gcccgcgagc	tgatcgacca	ggaaggccgc	accgtgaaag	3060
aggcggctgc	actgcttggc	gtgcatcgct	cgaccctgta	ccgcgcactt	gagcgcagcg	3120
aggaagtgac	gcccaccgag	gccaggcggc	gcggtgcctt	ccgtgaggac	gcattgaccg	3180
aggccgacgc	cctggcggcc	gccgagaatg	aacgccaaaga	ggaacaagca	tgaaaccgca	3240
ccaggacggc	caggacgaac	cgtttttcat	taccgaagag	atcgaggcgg	agatgatcgc	3300
ggccgggtac	gtgttcgagc	cgcccgcgca	cgtctcaacc	gtgcccgtgc	atgaaatcct	3360
ggccggtttg	tctgatgcc	agctggcggc	ctggccggcc	agcttggccg	ctgaagaaac	3420
cgagcggcgc	cgtctaaaaa	ggtgatgtgt	atltgagtaa	aacagcttgc	gtcatgcggt	3480
cgctgcgtat	atgatgcgat	gagtaaataa	acaatacgc	aaggggaacg	catgaagggt	3540
atcgctgtac	ttaaccagaa	aggcgggtca	ggcaagacga	ccatcgcaac	ccatctagcc	3600
cgccccctgc	aactcgccgg	ggccgatggt	ctgttagtcg	attccgatcc	ccagggcagt	3660
gcccgcgatt	gggcccggct	gcccgaagat	caaccgctaa	ccgttgtcgg	catcgaccgc	3720
ccgacgattg	accgcgacgt	gaaggccatc	ggcccggcgcg	acttcgtagt	gatcgacgga	3780
gcgcccaggg	cggcggactt	ggctgtgtcc	gcgatcaagg	cagccgactt	cgtgctgatt	3840
ccggtgcagc	caagccctta	cgacatatgg	gccaccgccc	acctggtgga	gctggttaa	3900
cagcgcattg	aggtcacgga	tgggaaggcta	caagcggcct	ttgtcgtgtc	gcccggcgtc	3960
aaaggcacgc	gcatcggcgg	tgaggttggc	gaggcgtctg	ccgggtacga	gctgccatt	4020
cttgagtccc	gtatcacgca	gcgcgtgagc	taccaggca	ctgccgcgc	cggcacaacc	4080
gttcttgaat	cagaaccgga	gggcgacgct	gcccgcgagg	tccaggcgt	ggccgctgaa	4140
attaatcaa	aactcatttg	agttaatgag	gtaaagagaa	aatgagcaaa	agcacaaca	4200
cgctaagtgc	cgcccgctcc	agcgcacgca	gcagcaaggc	tgcaacgttg	gccagcctgg	4260
cagacacgcc	agccatgaag	cggttcaact	ttcagttgcc	ggcggaggat	cacaccaagc	4320
tgaagatgta	cgcggtacgc	caaggcaaga	ccattaccga	gctgctatct	gaatacatcg	4380
cgcagctacc	agagtaaatg	agcaaatgaa	taaatgagta	gatgaatltt	agcggctaaa	4440
ggagggcgga	tgaaaaatca	agaacaacca	ggcaccgacg	ccgtggaatg	ccccatgtgt	4500
ggaggaacgg	gcggttggcc	aggcgtaaag	ggctgggttg	tctgccggcc	ctgcaatggc	4560
actggaaccc	ccaagcccga	ggaatcggcg	tgagcggctg	caaaccatcc	ggcccggtac	4620
aaatcggcgc	ggcgttgggt	gatgacctgg	tggagaagtt	gaaggccgcg	caggccgccc	4680
agcggcaacg	catcgaggca	gaagcacgcc	ccggtgaatc	gtggcaagcg	gcccgtgatc	4740
gaatccgcaa	agaatcccgg	caaccgcccg	cagccggtgc	gccgtcgatt	aggaagccgc	4800
ccaagggcga	cgagcaacca	gattttttcg	ttccgatgct	ctatgacgtg	ggcaccgccc	4860
atagtcgcag	catcatggac	gtggccggtt	tccgtctgtc	gaagcgtgac	cgacgagctg	4920
gcgaggtgat	ccgctacgag	cttccagacg	ggcacgtaga	ggtttccgca	ggcccggccc	4980
gcatggccag	tgtgtgggat	tacgacctgg	tactgatggc	ggtttcccat	ctaaccgaat	5040
ccatgaaccg	ataccgggaa	gggaagggag	acaagcccgg	ccgcgtgttc	cgccacacg	5100
ttgcggacgt	actcaagttc	tgcccggcag	ccgatggcgg	aaagcagaaa	gacgacctgg	5160
tagaaacctg	cattcggtta	aacaccacgc	acgttgccat	gcagcgtacg	agaaggcca	5220
agaacggccc	cctggtgacg	gtatccgagg	gtgaagcctt	gattagccgc	tacaagatcg	5280
taaagagcga	aaccgggccc	ccggagtaca	tcgagatcga	gctagctgat	tggatgtacc	5340
gcgagatcac	agaaggcaag	aaccggacg	tgctgacggg	tcaccccgat	tactttttga	5400
tcgatcccgg	catcggccgt	tttctctacc	gcctggcacg	ccgcgccgca	ggcaaggcag	5460
aagccagatg	gttgttcaag	acgatctacg	aacgcagtgg	cagcgcggga	gagttcaaga	5520
agttctgttt	caccgtgcgc	aagctgatcg	ggtcaaatga	cctgcccggg	tacgatttga	5580
aggaggaggc	ggggcaggct	ggcccgatcc	tagtcatgcg	ctaccgcaac	ctgatcgagg	5640
gcgaagcatc	cgccggttcc	taatgtacgg	agcagatgct	agggcaaat	gccctagcag	5700
gggaaaaagg	tcgaaaagg	ctctttcctg	tggatagcac	gtacattggg	aaccacaagc	5760
cgtagacattg	gaaccggaac	ccgtacattg	ggaacccaaa	gccgtacatt	gggaaccggt	5820
cacacatgta	agtgactgat	ataaaaagaga	aaaaaggcga	tttttccgcc	taaaactcct	5880
taaaacttat	taaaactcct	aaaaccggcc	tggcctgtgc	ataactgtct	ggccagcgc	5940
cagccgaaga	gctgcaaaaa	gcgcctacc	ttcggctcgt	gcgctcccta	cgccccgccc	6000
cttcgcgtcg	gcctatcgcg	gccgctggcc	gctcaaaaa	ggctggccta	cgccaggcca	6060
atctaccagg	gcgcggacaa	gccgcgccc	cgccactcga	ccgcggcgc	ccacatcaag	6120
gcaccctgcc	tcgcgcggtt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctccc	6180
gagacgggtca	cagcttgtct	gtaagcggat	gcccgggagca	gacaagccc	tcagggcgcg	6240
tcagcgggtg	ttggcgggtg	tcggggcgca	gccatgacc	agtcacgtag	cgatagcgg	6300

```

gtgtatactg gcttaactat gcggcacag agcagattgt actgagagtg caccatattg 6360
gggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcagggcg tcttccgctt 6420
cctcgctcac tgactcgctg cgctcggtcg ttcggctgcy gcgagcggta tcagctcact 6480
caaaggcggg aatacgggta tccacagaat caggggataa cgcaggaaag aacatgtgag 6540
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 6600
ggctccgccc ccctgacgag catcacaata atcgacgctc aagtcagagg tggcgaacc 6660
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgctg cgctctcctg 6720
ttccgaccct gccgcttacc ggatacctgt ccgctttct cccttcggga agcgtggcg 6780
tttctcatag ctacgctgt aggtatctca gttcggtgta ggtcgttcg tccaagctgg 6840
gctgtgtgca cgaaccccc gttcagcccg accgctgcy cttatccgg aactatcgtc 6900
ttgagtccaa cccggtaaga cagacttat cgccactggc agcagccact ggtaacagga 6960
ttagcagagc gaggtatgta ggcgggtgta cagagttctt gaagtggagg cctaactacg 7020
gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 7080
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggttttttg 7140
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatcttt 7200
ctacgggggtc tgacgctcag tggaaacgaaa actcacgta agggattttg gtcatgcatg 7260
atatatctcc caatttgtgt agggcttatt atgcacgctt aaaaataata aaagcagact 7320
tgacctgata gtttggctgt gagcaattat gtgcttagtg catctaaccg ttgagttaag 7380
ccgcgccgcy aagcggcgctc ggcttgaacg aatttctagc tagacattat ttgccgacta 7440
ccttgggtgat ctgcctttc acgtagtgga caaattcttc caactgatct gcgcgcgagg 7500
ccaagcgatc ttcttcttgt ccaagataag cctgtctagc ttcaagtatg acgggctgat 7560
actgggcccg caggcgctcc attgccagc cggcagcgc atccttcggc gcgattttgc 7620
cggttactgc gctgtaccaa atgcccggaca acgtaagcac tacatttcgc tcatcgccag 7680
cccagtcggg cggcgagttc catagcgtta aggtttcatt tagcgcctca aatagatcct 7740
gttcaggaac cggatcaaag agttcctccg ccgctggacc taccaaggca acgctatggt 7800
ctcttgcttt tgtcagcaag atagccagat caatgtcgat cgtggctggc tcgaagatac 7860
ctgcaagaat gtcattgcgc tgccattctc caaattgcag ttcgcgctta gctggataac 7920
gccacggaat gatgtcgtcg tgcacaacaa tgggtgacttc tacagcgcgg agaatctcgc 7980
tctctccagg ggaagccgaa gtttccaaaa ggtcgttgat caaagctcgc cgcgttggtt 8040
catcaagcct tacggtcacc gtaaccagca aatcaatatc actgtgtggc ttcaggccgc 8100
catccactgc ggagccgtac aaatgtacgg ccagcaacgt cggttcgaga tggcgcctcga 8160
tgacgccaac tacctctgat agttgagtcg atacttcggc gatcaccgct tccccatga 8220
tgtttaactt tgtttttaggg cgactgcctt gctgcgtaac atcgttgctg ctccataaca 8280
tcaaacatcg acccagggcg taacgcgctt gctgcttggg tggccgaggc atagactgta 8340
ccccaaaaaa acagtcataa caagccatga aaaccgccac tgcy 8384

```

<210> 16

<211> 234

<212> DNA

<213> Solanum tuberosum

<220>

<221> terminator

<222> (1)..(234)

<223> terminators sequence of the Cathepsin D Inhibitor gene from potato

<400> 16

```

cctgcagata gactatacta tgtttttagcc tgcctgctgg ctagctacta tgttatgtta 60
tgttgtaaaa taaacacctg ctaaggtata tctatctata ttttagcatg gctttctcaa 120
taaattgtct ttccttatcg ttactatct tatacctaata atgaaataa taatatcaca 180
tatgaggaac ggggcagggt taggcatata tatacagagtg tagggcggag tggg 234

```

<210> 17

<211> 298

<212> DNA

<213> Vicia faba

<220>

<221> terminator

<222> (1)..(298)

<223> terminator of storage protein gene Vf1E1B3 from  
Vicia faba

<400> 17

```
gatcctgcaa tagaatgttg aggtgaccac tttctgtaat aaaataatta taaaataaat 60
ttagaattgc tgtagtcaag aacatcagtt ctaaaatatt aataaagtta tggccttttg 120
acatatgtgt ttcgataaaa aaatcaaaaat aaattgagat ttattcgaaa tacaatgaaa 180
gtttgcagat atgagatatg tttctacaaa ataataactt aaaactcaac tatatgctaa 240
tgtttttctt ggtgtgtttc atagaaaatt gtatccgttt cttagaaaat gctcgttaa 298
```

<210> 18

<211> 1478

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> promoter

<222> (1)..(1478)

<223> truncated version of nit1 promoter

<400> 18

```
cagttgtggt ctttctttt ttttctttgc atatttgcatt ttagaatggt taataattaa 60
gttactgtat ttccacatac attagttcca agaataatac tatattaatt tatttttctt 120
aaaaatgttt tggaaatgact aatattgaca acgaaaatag aagctatgct aaaccattac 180
gtatatgtga cttcacatgt tgttgtttta cattccctat atatatggat ggctgtcaca 240
atcagaaaacg tgatcgaaaa aagacaaaaca gtgtttgcatt aaaaagacta tttcgtttca 300
ttgacaattt gtgtttattt gtaaagaaaa gtggcaaagt ggaatttgag ttcctgcaag 360
taagaaagat gaaataaaag acttgagtgt gtgtttttt cttttatctg aaagctgcaa 420
tgaaatattc ctaccaagcc cgtttgatta ttaattgggg tttggttttc ttgatgcgaa 480
ctaattgggt atataagaaa ctatacaatc catgttaatt caaaaatttt gatttctctt 540
gtaggaatat gatttactat atgagacttt ctttctgcca ataataagtaa atccaaagat 600
atgtgaccgg accaaaacac attgatctat tttttagttt atttaatacca gtttctctga 660
gataattcat taaggaaaac ttagtattaa cccatcctaa gattaaatag gagccaaact 720
cacatttcaa atattaata acataaaatg gatttaaaaa atctatacgt caaattttat 780
ttatgacatt tcttatttaa atttatattt aatgaaatac agctaagaca aaccaaaaaa 840
aaaatacttt ctaagtgggt caaaacatca attccgttca atattattag gtagaatcgt 900
acgaccaaaa aaaggtagggt taatacgaat tagaaacata tctataacat agtatatatt 960
attacctatt atgaggaatc aaaatgcatt aaatatggat ttaaggaatc cataaaagaa 1020
taaattctac gggaaaaaaa atggaataaa ttcttttaag ttttttattt gttttttatt 1080
tggtagttct ccattttggt ttatttctgt tggatttatt gtgtccaaat actttgtaaa 1140
ccaccgttgt aattcttaaa cggggttttc acttctttt tatattcaga cataaagcat 1200
cggctgggtt aatcaatcaa tagattttat ttttctctc aattattagt aggtttgatg 1260
tgaactttac aaaaaaaca aaaacaaatc aatgcagaga aaagaaacca cgtgggctag 1320
tcccaccttg tttcatttcc accacagggt cgatctcctg taccgtctcc aataggaaaa 1380
taaacgtgac cacaaaaaaa aaacaaaaaa aagtctatat attgcttctc tcaagtctct 1440
gagtgtcatg aaccaaagta aaaaacaaag actcgagt 1478
```

<210> 19

<211> 1330

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> promoter

<222> (1)..(1330)

<223> truncated version of nit1 promoter

<400> 19

```
caacgaaaat agaagctatg ctaaaccatt acgtatatgt gacttcacat gttgttgttt 60
tacattccct atatatatgg atggctgtca caatcagaaa cgtgatcgaa aaaagacaaa 120
cagtgtttgc ataaaaagac tatttctgtt cattgacaat ttgtgtttat ttgtaaagaa 180
aagtggcaaa gtggaatttg agttcctgca agtaagaaag atgaaataaa agacttgagt 240
```

gtgtggtttt	ttcttttata	tgaaagctgc	aatgaaatat	tcctaccaag	cccgtttgat	300
tattaattgg	ggtttgggtt	tcttgatgcg	aactaattgg	ttatataaga	aactatacaa	360
tccatggtta	ttcaaaaatt	ttgattttct	ttgtaggaat	atgatttact	atatgagact	420
ttcttttcgc	caataatagt	aatccaaaag	atatttgacc	ggacccaaaac	acattgatct	480
atTTTTtagt	ttatttaatc	cagtttctct	gagataattc	attaaggaaa	acttagtatt	540
aaccatcct	aagattaat	aggagccaaa	ctcacatttc	aaatattaaa	taacataaaa	600
tggatttaaa	aaatctatac	gtcaaatttt	atTTatgaca	tttcttattt	aaatttata	660
ttaatgaaat	acagctaaga	caaaccaaaa	aaaaaatact	ttctaagtgg	tccaaaacat	720
caattccgtt	caatattatt	aggtagaatc	gtacgaccaa	aaaaaggtag	gttaatacga	780
attagaaaca	tatctataac	atagtatata	ttattaccta	ttatgaggaa	tcaaaatgca	840
tcaaatatgg	atTTaaggaa	tccataaaaag	aataaattct	acgggaaaaa	aaatggaata	900
aattctttta	agttttttat	ttgTTTTtta	tttggtagtt	ctccattttg	ttttatttcg	960
tttggattta	ttgtgtccaa	atactttgta	aaccaccgtt	gtaattctta	aacgggggtt	1020
tcacttcttt	tttatattca	gacataaagc	atcggtggt	ttaatcaatc	aatagatttt	1080
atTTTTcttc	tcaattatta	gtaggtttga	tgtgaacttt	acaaaaaaaa	caaaaacaaa	1140
tcaatgcaga	gaaaagaaac	cacgtgggct	agtcccacct	tgtttcattt	ccaccacagg	1200
ttcgatcttc	gttaccgtct	ccaataggaa	aataaacgtg	accacaaaaa	aaaaacaaaa	1260
aaaagtctat	atattgcttc	tctcaagtct	ctgagtgtca	tgaaccaaag	taaaaaacaa	1320
agactcgagt						1330