(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 March 2006 (30.03.2006)

PCT

(10) International Publication Number
WO 2006/032426 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/EP2005/010058

(22) International Filing Date:
17 September 2005 (17.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/612432 23 September 2004 (23.09.2004) US

(71) Applicant *(for all designated States except US)*: **BASF PLANT SCIENCE GMBH** [DE/DE]; 67056 Ludwigshafen (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SANCHEZ-FERNANDEZ, Rocio** [ES/DE]; Markt 9, 06484 Quedlinburg (DE). **BIESGEN, Christian** [DE/DE]; Rebenstr. 10, 06484 Quedlinburg (DE). **LEPS, Michael** [DE/DE]; Juri-Gagarin-Str. 12, 38820 Halberstadt (DE). **BROWN, Jeffrey A.** [US/US]; 1005 Modest Way, Apex, NC 27502 (US).

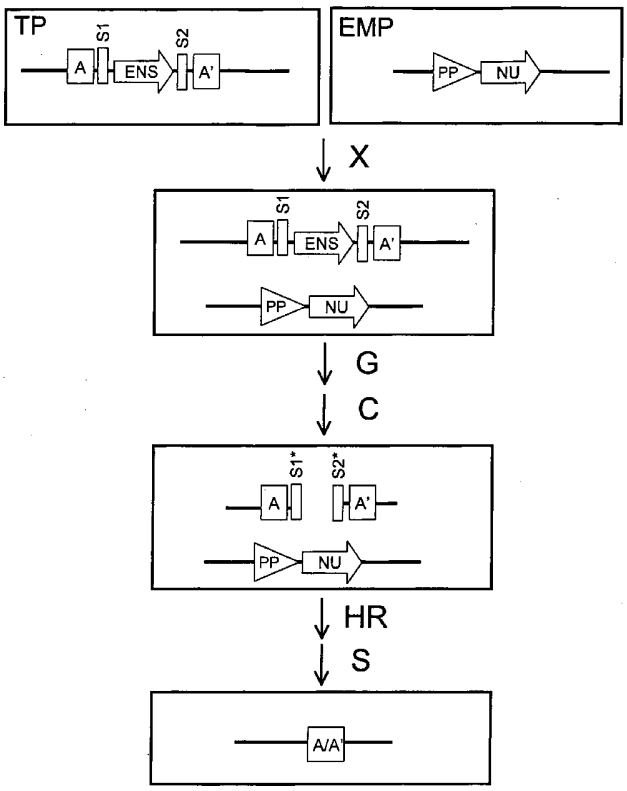
(74) Agent: **POPP, Andreas**; c/o BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,

[Continued on next page]

(54) Title: RECOMBINATION CASSETTES AND METHODS FOR SEQUENCE EXCISION IN PLANTS



(57) Abstract: Summary The invention relates to improved recombination systems and methods for eliminating maker sequences from the genome of plants. Particulary the invention is based on use of an expression cassette comprising the parsley ubiquitin promoter, and operably linked thereto a nucleic acid sequence coding for a sequence specific DNA-endonuclease.

WO 2006/032426 A2

RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

RECOMBINATION CASSETTES AND METHODS FOR SEQUENCE EXCISION IN PLANTS

FIELD OF THE INVENTION

The invention relates to improved recombination systems and methods for eliminating maker sequences from the genome of plants.

BACKGROUND OF THE INVENTION

An aim of plant biotechnology is the generation of plants with advantageous novel characteristics, for example for increasing agricultural productivity, improving the quality in foodstuffs or for the production of certain chemicals or pharmaceuticals (Dunwell JM (2000) J Exp Bot 51:487-96). Transformation of plants typically involves the introduction of a gene of interest ("trait gene") and a marker sequence (for example a selectable marker such as a herbicide resiatance gene) into the organism. The marker sequence is useful during the transformation process to select for, and identify, transformed organisms, but typically provides no useful function once the transformed organism has been identified and contributes substantially to the lack of acceptance of these "gene food" products among consumers. In consequence, there are multiple attempts to develop techniques by means of which marker sequences can be excised from plant genome (Ow DW and Medberry SL (1995) Crit Rev in Plant Sci 14:239-261).

The person skilled in the art is familiar with a variety of systems for the site-directed removal of recombinantly introduced nucleic acid sequences. They are based on the use of sequence specific recombinases and two recognition sequences of said recombinases which flank the sequence to be removed. The effect of the recombinase on this construct brings about the excision of the flanked sequence, one of the recognition sequences remaining in the genome of the organism. Various sequence-specific recombination systems are described, such as the Cre/lox system of the bacteriophage P1 (Dale EC and Ow DW (1991) Proc Natl Acad Sci USA 88:10558-10562; Russell SH *et al.* (1992) Mol Gen Genet 234: 49-59; Osborne BI *et al.* (1995) Plant J. 7, 687-701), the yeast FLP/FRT system (Kilby NJ *et al.* (1995) Plant J 8:637-652; Lyznik LA *et al.* (1996) Nucleic Acids Res 24:3784-3789), the Mu phage Gin recombinase, the E. coli Pin recombinase or the R/RS system of the plasmid pSR1 (Onouchi H *et al.*(1995) Mol Gen Genet 247:653-660; Sugita K *et al.* (2000) Plant J. 22:461-469). A disadvantage of the sequence-specific recombination systems is the reversibility of the reaction, that is to say an equilibrium exists between excision and integration of the marker sequence in question. This frequently brings about unwanted mutations by multiple consequtive insertions and excisions. This not only applies to the Cre-lox system, but also to the other sequence-specific recombinases (see above). A further disadvantage is the fact that one of the recognition sequences of the recombinase remains in the genome, which is thus modified: The remaining recognition sequence excludes a further use of the recombination system, for example for a second genetic modification, since interactions with the subsequently introduced recognition sequences cannot be ruled out. Substantial chromosomal rearrangements or deletions may result.

Zubko *et al.* describe a system for the deletion of nucleic acid sequences from the tobacco genome, where the sequence to be deleted is flanked by two 352 bp attP recognition sequences from the bacteriophage Lambda. Deletion of the flanked region takes

place independently of the expression of helper proteins in two out of eleven transgenic tobacco lines by spontaneous intrachromosomal recombination between the attP recognition regions. The disadvantage of this method is that recombination, or deletion, cannot be induced specifically at a particular point in time, but takes place spontaneously. The fact that the method worked only in a small number of lines suggests that the integration locus in the cases in question tends to be unstable (Puchta H (2000) Trends in Plant Sci 5:273-274).

WO 02/29071 discloses a method for conditional excision of transgenic sequences from the genome of a transgenic organism. Excision occurs directly by action of an enzyme (e.g., a recombinase or a endonuclease) but not via homologous recombination of flanking sequences. The recombination mechanism mediated by recombinases differs from the mechanism leading to homologous recombination between homologous sequences. It is the purpose of the method to prevent occurrence of the transgenic sequence in the agricultural product but to have it remaining in other plant parts. In consequence, promoters employed here are inducible, seed or fruit specific promoters.

Self-excising constructs based on a site-specific recombinase are described in WO97/037012 and WO02/10415. Here also no homologous recombination but recombinase mediated recombination occurs and the recombinase recognition sequence remains in the genome making further applications of the system impossible (as described above as a general disadvantage for recombinase systems).

Several constitutive promoters in plants are known. Most of them are derived from viral or bacterial sources such as the nopaline synthase (nos) promoter (Shaw *et al.* (1984) Nucleic Acids Res. 12 (20) : 7831-7846), the mannopine synthase (mas) promoter (Comai *et al.* (1990) Plant Mol Biol 15(3):373-381), or the octopine synthase (ocs) promoter (Leisner and Gelvin (1988) Proc Natl Acad Sci USA 85 (5) :2553-2557) from *Agrobacterium* tumefaciens or the CaMV35S promote from the Cauliflower Mosaic Virus (US 5,352, 605). The latter was most frequently used in constitutive expression of transgenes in plants (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). However, the CaMV 35S promoter demonstrates variability not only in different plant species but also in different plant tissues (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). An additional disadvantage is an interference of the transcription regulating activity of the 35S promoter with wild-type CaMV virus (Al-Kaff *et al.* (2000) Nature Biotechnology 18 :995-99). Another viral promoter for constitutive expression is the Sugarcane bacilliform badnavirus (ScBV) promoter (Schenk *et al.* (1999) Plant Mol Biol 39 (6) :1221-1230).

Several plant constitutive promoters are described such as the ubiquitin promoter from *Arabidopsis* thaliana (Callis *et al.* (1990) J Biol Chem 265:12486- 12493; Holtorf S *et al.* (1995) Plant Mol Biol 29:637-747), which - however - is reported to be unable to regulate expression of selection markers (WO03102198), or two maize ubiquitin promoter (Ubi-1 and Ubi-2; US 5,510,474; US 6,020, 190; US 6,054574), which beside a constitutive expression profile demonstrate a heat-shock induction (Christensen *et al.* (1992)

Plant. Mol. Biol. 18(4):675-689). A comparison of specificity and expression level of the CaMV 35S, the barley thionine promoter, and the *Arabidopsis* ubiquitin promoter based on stably transformed *Arabidopsis* plants demonstrates a high expression rate for the CaMV 35S promoter, while the thionine promoter was inactive in most lines and the ubi1 promoter from Arabisopsis resulted only in moderate expression activity (Holtorf *et al.* (1995) Plant Mol Biol 29 (4):637-6469) .

While the maize Ubi-1 promoter demonstrates acceptable expression activity in maize and other monocotyledonous plants, expression is low (10%) in dicotyledonous tobacco plants in comparison to the 35S CaMV promoter, which makes the promoter unsuitable for most applications in dicots. Ubiquitines are ubiquitous proteins found in all eukaryotes analyzed so far. The genes for parsley (*Petroselinum crispum*) are described (Kawalleck *et al.* (1993) Plant Mol Biol 21;673-684. Furthermore the promoter of the parsley ubiquitin gene was analyzed and described as a constitutive promoter (WO 03/102198). Other constitutive promoters are the rice atin 1(Actl) promoter (McElroy *et al.* (1991) Mol Gen Genet 231:150-1609), and the S-adenosyl-L-methionine synthetase promoter (WO 00/37662). The latter is however dependant on the methionine concentration.

WO 03/004659 describes a recombination system based on homologous recombination between two homologous sequences induced by action of a sequence specific double-strand break inducing enzyme, preferably a meganuclease (homing-endonuclease). Although general statements are made about the preferable use of *inter alia* homing-endonucleases and the potential use of *inter alia* tissue specific promoters, there is no specific teaching suggesting the specific combination of features of the invention disclosed herein. European Patent Applications Appl. No. 03028884.9 and 03028885.6 describe various combination of homing endonucleases with promoters having activity in reproductive tissues.

Although these inventions solve some problems, still the extent of excision is low and the generation of homogenous, non-chimeric plants (i.e., plants in which the sequence was deleted from all cells) is time and labor intensive. The reason is mainly an non-homogeneous or insufficient expression of the endonuclease, which is needed for induction of site-specific double-strand breaks to induce homologous recombination between directed repeats flanking the sequences to be deleted. For example for the strong 35S CaMV recombination could only be observed in less than 10% of the plant cells. Such insufficient or non-homogeneous expression results in plants which are mosaic or chimeric plants (i.e., plants which comprise both cells which have undergone recombination and sequence excision and cells which have not). This requires additional plant generations (either by sexual or asexual propagation). The related efforts highly depend on the frequency of cells which have undergone homologous recombination.

It is an object of the present invention to develop systems and methods which enable the easy-to-use, highly-efficient, predictable elimination of sequences, preferably marker sequences, from the genome of a plant and allow the repeated, successive application to the same organism. This has been achieved by the present invention.

SUMMARY OF THE INVENTION

Accordingly a first embodiment of the invention relates to a method for producing a transgenic plant comprising:

i) crossing a first transgenic plant comprising in its genome a DNA construct comprising

a1) at least one recognition sequence of at least 10 base pairs for the site-directed induction of DNA double-strand breaks by a sequence specific DNA-endonuclease and

b1) a nucleic acid sequence to be excised,

wherein said elements a1) and b1) and optionally further elements are flanked by homology sequences A and A', having sufficient length and sufficient homology in order to ensure homologous recombination between A and A', and having an orientation which – upon recombination between A and A' – will lead to an excision of said elements a1) and b1), and

c1) at least one additional sequence conferring to said plant an agronomically valuable trait, wherein said sequence is not localized between the homology sequences A and A' and would not be excised from the genome upon recombination between A and A'

with a second transgenic plant comprising in its genome an expression cassette comprising

a2) the parsley ubiquitin promoter, and operably linked thereto

b2) a nucleic acid sequence coding for a sequence specific DNA-endonuclease having a sequence specificity for said recognition sequence a1),

ii) generating descendants (F1) following this crossing, and – optionally - sexually or asexually generating further descendants, and

iii) isolating descendants which have undergone recombination between the homology sequences A and A' and which do not comprise in their genome said elements a1) and b1) but comprise sequence c1).

Preferably the element b1) is an expression cassette for a marker sequence, more preferably selected from the group consisting of negative selection marker, counter selection marker, positive selection marker, and reporter genes.

In an preferred embodiment, the method further comprises the step of segregating the expression cassette for the endonuclease from the sequence c1) for the agronomically valuable trait and isolating plants comprising sequence c1) but not said expression cassette for the endonuclease.

Preferably the parsley ubiquitin promoter comprises a sequence described by SEQ ID NO: 8 or 15 or a functional equivalent or functional equivalent fragment thereof.

Preferably the orientation of the homology sequences is in the form of direct repeats, which are flanking elements a1) and b1) and optionally further elements.

Preferably the recombination mechanism between A and A' is homologous recombination.

The sequence specific DNA-endonuclease is preferably a homing endonuclease, more preferably selected from the group consisting of I-SceI, I-CpaI, I-CpaII, I-CreI and I-ChuI.

In an preferred embodiment the construct employed in the method of the invention comprises two recognition sequences a1) which are localized between the homology sequences A and A' and are flanking element b1) and optionally further elements in a way that cleavage at this two recognition sequences excises said element b1). The homology sequences A and A' are preferably part of the expression cassette comprised in the DNA construct.

In an preferred embodiment of the invention, the method is employed to generate marker-free plants, thus preferably the resulting plant is selection marker-free.

Another embodiment of the invention relates to a transgenic expression cassette comprising a sequence coding for a sequence specific DNA-endonuclease operably linked to the parsely ubiquitin promoter. The endonuclease is preferably a homing endonuclease, more preferably selected from the group consisting of I-SceI, I-CpaI, I-CpaII, I-CreI and I-ChuI. The parsley ubiquitin promoter preferably comprises a sequence described by SEQ ID NO: 8 or 15 or a functional equivalent or functional equivalent fragment thereof. Other embodiments of the invention relate to transgenic vectors comprising a expression cassette of the invention, and transgenic cells or non-human organisms, preferably plant or plant cells, comprising a expression cassette or a vector of the invention. Preferably the expression cassette is comprised in the genome of the plant or plant cell.

GENENERAL DEFINITIONS

The teachings, methods, sequences etc. employed and described in the international patent application WO 03/004659 are hereby incorporated by reference.

" Agronomically valuable trait" includes any phenotype in a plant organism that is useful or advantageous for food production or food products, including plant parts and plant products. Non-food agricultural products such as paper, etc. are also included. A partial list of agronomically valuable traits includes pest resistance, vigor, development time (time to harvest), enhanced nutrient content, novel growth patterns, flavors or colors, salt, heat, drought and cold tolerance, and the like. Preferably, agronomically valuable traits do not include marker sequences (e. g., selectable marker such as herbicide or antibiotic resistance genes used only to facilitate detection or selection of transformed cells), hormone biosynthesis genes leading to the production of a plant hormone (e.g., auxins, gibberellins, cytokinins, abscisic acid and ethylene that are used only for selection), or reporter genes (e.g. luciferase, glucuronidase, chloramphenicol acetyl transferase (CAT, etc.). Such agronomically valuable important traits may include improvement of pest resistance (e.g., Melchers *et al.* (2000) Curr Opin Plant Biol 3(2):147-52), vigor, development time (time to harvest), enhanced nutrient content, novel growth patterns, flavors or colors, salt, heat, drought, and cold tolerance (e.g., Sakamoto *et al.*

(2000) J Exp Bot 51(342):81-8; Saijo *et al.* (2000) Plant J 23(3): 319-327; Yeo *et al.*(2000) Mol Cells 10(3):263-8; Cushman *et al.* (2000) Curr Opin Plant Biol 3(2):117-24), and the like. Those of skill will recognize that there are numerous polynucleotides from which to choose to confer these and other agronomically valuable traits.

The phrase "nucleic acid sequence" refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5'- to the 3'-end. It includes chromosomal DNA, self-replicating plasmids, infectious polymers of DNA or RNA and DNA or RNA that performs a primarily structural role. A "polynucleotide construct" refers to a nucleic acid at least partly created by recombinant methods.

The term "promoter" refers to regions or sequences located upstream and/or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription.

A polynucleotide sequence is "heterologous to" an organism or a second polynucleotide sequence if it originates from a foreign species, or, if from the same species, is modified from its original form. For example, a promoter operably linked to a heterologous coding sequence refers to a coding sequence from a species different from that from which the promoter was derived, or, if from the same species, a coding sequence which is not naturally associated with the promoter (e. g. a genetically engineered coding sequence or an allele from a different ecotype or variety).

"Transgene", "transgenic" or "recombinant" refers to a polynucleotide manipulated by man or a copy or complement of a polynucleotide manipulated by man. For instance, a transgenic expression cassette comprising a promoter operably linked to a second polynucleotide may include a promoter that is heterologous to the second polynucleotide as the result of manipulation by man (e.g., by methods described in Sambrook *et al.*, Molecular Cloning-A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989) or Current Protocols in Molecular Biology Volumes 1-3, John Wiley & Sons, Inc. (1994-1998)) of an isolated nucleic acid comprising the expression cassette. In another example, a recombinant expression cassette may comprise polynucleotides combined in such a way that the polynucleotides are extremely unlikely to be found in nature. For instance, restriction sites or plasmid vector sequences manipulated by man may flank or separate the promoter from the second polynucleotide. One of skill will recognize that polynucleotides can be manipulated in many ways and are not limited to the examples above.

A polynucleotide "exogenous to" an individual organism is a polynucleotide which is introduced into the organism by any means other than by a sexual cross.

The term "expression cassette" – for example when referring to the expression cassette for the sequence specific DNA-endonuclease – means those constructions in which the DNA to be expressed is linked operably to at least one genetic control element which enables or regulates its expression (i.e. transcription and / or translation). Here, expression may be for example stable or transient, constitutive or inducible.

The terms "operable linkage" or "operably linked" are generally understood as meaning an arrangement in which a genetic control sequence is capable of exerting its function with regard to a nucleic acid sequence, for example while encoding a sequence specific DNA-endonuclease. Function, in this context, may mean for example control of the expression, i.e. transcription and/or translation, of the nucleic acid sequence, for example one encoding a sequence specific DNA-endonuclease. Control, in this context, encompasses for example initiating, increasing, governing or suppressing the expression, i.e. transcription and, if appropriate, translation. Controlling, in turn, may be, for example, tissue- and / or time-specific. It may also be inducible, for example by certain chemicals, stress, pathogens and the like. Preferably, operable linkage is understood as meaning for example the sequential arrangement of a promoter, of the nucleic acid sequence to be expressed - for example one encoding a sequence specific DNA-endonuclease – and, if appropriate, further regulatory elements such as, for example, a terminator, in such a way that each of the regulatory elements can fulfil its function when the nucleic acid sequence - for example one encoding a sequence specific DNA-endonuclease – is expressed. An operably linkage does not necessarily require a direct linkage in the chemical sense. Genetic control sequences such as, for example, enhancer sequences are also capable of exerting their function on the target sequence from positions located at a distance or indeed other DNA molecules. Preferred arrangements are those in which the nucleic acid sequence to be expressed - for example one encoding a sequence specific DNA-endonuclease – is positioned after a sequence acting as promoter so that the two sequences are linked covalently to one another. The distance between the promoter sequence and the nucleic acid sequence – for example one encoding a sequence specific DNA-endonuclease – is preferably less than 200 base pairs, especially preferably less than 100 base pairs, very especially preferably less than 50 base pairs. The skilled worker is familiar with a variety of ways in order to obtain such an expression cassette. References for customary recombination and cloning techniques as given below. However, an expression cassette may also be constructed in such a way that the nucleic acid sequence to be expressed (for example one encoding a marker sequence, an agronomically valuable trait, or a sequence specific endonuclease) is brought under the control of an endogenous genetic control element, for example an endogenous promoter, for example by means of homologous recombination or else by random insertion. Such constructs are likewise understood as being expression cassettes for the purposes of the invention.

A "genetically-modified organism" or "GMO" refers to any organism that comprises transgene DNA. Exemplary organisms include plants, animals and microorganisms.

Homology between two nucleic acid sequences is understood as meaning the identity of the nucleic acid sequence over in each case the entire sequence length which is calculated by alignment with the aid of the program algorithm GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), setting the following parameters:

Gap Weight: 12	Length Weight: 4
Average Match: 2,912	Average Mismatch:-2,003

"Genome" or "genomic DNA" is conferring to the heritable genetic information of a host organism. Said genomic DNA comprises the DNA of the nucleus (also referred to as chromosomal DNA) but also the DNA of the plastids (e.g., chloroplasts) and other cellular organelles (e.g., mitochondria). Preferably the terms genome or genomic DNA is referring to the chromosomal DNA of the nucleus.

DETAILED DESCRIPTION OF THE INVENTION

Accordingly, a first embodiment of the invention relates to a method for producing a transgenic plant comprising:

1) crossing a first transgenic plant comprising in its genome a DNA construct comprising

a1) at least one recognition sequence of at least 10 base pairs for the site-directed induction of DNA double-strand breaks by a sequence specific DNA-endonuclease and

b1) a nucleic acid sequence to be excised,

wherein said elements a1) and b) and optionally further elements are flanked by homology sequences A and A', having sufficient length and sufficient homology in order to ensure homologous recombination between A and A', and having an orientation which – upon recombination between A and A' – will lead to an excision of said elements a1) and b), and

c1) at least one additional sequence conferring to said plant an agronomically valuable trait, wherein said sequence is not localized between the homology sequences A and A' and would not be excised from the genome upon recombination between A and A'

with a second transgenic plant comprising in its genome an expression cassette comprising

a2) the parsley ubiquitin promoter, and operably linked thereto

b2) a nucleic acid sequence coding for a sequence specific DNA-endonuclease having a sequence specificity for said recognition sequence a1),

2) generating descendants (F1) following this crossing, and – optionally - sexually or asexually generating further descendants, and

3) isolating descendants which have undergone recombination between the homology sequences A and A' and which do not comprise in their genome said elements a1) and b1) but comprise sequence c1).

In an preferred embodiment the element b1) is an expression cassette for a marker sequence. More preferably, this expression cassette enables the expression of a sequence allowing selection of transformed plant material, wherein the DNA sequence encoding said selectable sequence is operably linked with a promoter functional in plants. However, excision may also be advantageous in other circumstance and for other non-marker sequences, for example, in cases of hybrid technology or trait containment.

In a preferred embodiment the orientation of the homology sequences in the DNA construct of the invention is in the form of directed repeats, which are flanking elements a1) and b1) and optionally further elements.

In another preferred embodiment the expression cassette for the endonuclease is segregated from the sequences c1) for the agronomically valuable trait by e.g., conventional breeding techniques. Plants are isolated which comprise sequence c1) but not said expression cassette for the endonuclease. In a preferred embodiment the resulting plant is marker free or selection marker free.

In another preferred embodiment the DNA construct of the invention, comprises two recognition sequences a1). It is especially preferred that these two recognition sequences are flanking the marker sequence (and optionally further elements) in a way that a cleavage at this two sides excises the marker sequence (and optionally further elements).

Other embodiments of the invention are related to vector comprising said DNA construct, and transgenic plants comprising said vector or said DNA construct.

The present invention enables sequences (such as marker sequences e.g., genes for resistance to antibiotics or herbicides) to be deleted from the genome (e.g., chromosomal DNA) of a plant organism in an accurately predictable manner with high efficiency.

Within the method of the invention it is an essential feature that two plants are crossed, each of these comprising a specific DNA construct:

i) a first plant (hereinafter the "endonuclease master plant") comprising a expression cassette for expression of a sequence specific DNA-endonuclease (hereinafter the "endonuclease expression cassette"). Expression here is under control of the parsley ubiquitin promoter as specified in more detail below.

ii) a second plant (hereinafter the "trait plant") comprising a recombination cassette for excision (hereinafter the "excision cassette") of a sequence to be deleted (e.g., a marker sequences) and further comprising – optionally – sequences (e.g., an expression cassette) for an agronomically valuable trait.

The sequence to be eliminated (e.g., the marker sequence) is flanked by homology sequences A and A' having sufficient length and sufficient homology in order to ensure homologous recombination between A and A', and having an orientation which – upon recombination between A and A' – will lead to an excision of said sequence (e.g., the marker sequence) from the genome. Efficiency and accuracy of homologous recombination between A and A' is mediated by action of a sequence specific DNA-endonuclease, which is able to cleave at a recognition site between the two homology sequences, inducing a double-strand break, and in consequence, triggering said homologous recombination between A and A'. By this homologous recombination also the recognition sequence for the sequence specific DNA-endonuclease is excised likewise, which allows the method of the invention to be used repeatedly for further controlled genetic modifications. The sequences which are deleted are those located between the homology sequences A and A'. In contrast to systems such as, for example, the cre/lox

or the FRT/FLP system, one is not bound to specific sequences when performing recombination. The skilled worker knows that any sequence can undergo homologous recombination with another sequence provided that sufficient length and homology exist.

It is another inventive feature of the present invention, that expression of the sequence specific DNA-endonuclease is mediated by the parsley ubiquitin promoter. The method of the invention has at least four advantageous effects:

1) The use of parsley ubiquitin promoters for the endonuclease expression cassette surprisingly outmatches all other constitutive promoters tested so far including the gold-standard CaMV 35S promoter. The performance of the parsley ubiquitin promoter is by far better than the performance of any other promoter tested under equivalent conditions (see comparison examples). The parsley ubiquitin promoter seems to regulate efficient transcription and expression of the endonucleases in tissues and at times which are essential to allow induced homologous recombination. In consequence, many more cells of the F1 generation of a cross between an endonuclease master plant and a trait plant contain the respective recombination event. Therefore, the isolation of plants having the recombination event present in all cells is highly facilitated. Accordingly, the specific use of the parsley ubiquitin promoter solves problems which still adhere to constitutive promoters such as 35S CaMV and others.
2) Physical separation of the expression cassette for the endonuclease and the excision cassette (comprising its recognition sequences) by employing separate plants prevents premature excision that may occur in a co-transformation approach with both constructs and which may negatively affect the transformation / selection efficiency.
3) The method of the invention reduces multiple insertion (e.g., of a T-DNA) in one genomic location to a single insertion event by excision of the redundant copies (Fig. 10) in addition to the excision of the sequence to be eliminated (e.g., the selection marker).
4) The fact that the endonuclease is expressed from a construct in a separate plant allows for generation and use of a master plant. This means, that this master plant for the endonuclease can be crossed with various plants comprising different recombination cassettes (for the introduction of different agronomically valuable traits). This makes the method time and work efficient since only one of the two plants employed need to be generated for a new approach. Moreover, such endonuclease master plant could be in elite germplasm. Thus, upon crossing to the trait plant one could already initiate the first cross to breed the agronomical valuable trait into elite germplasm.

1. Sequence specific DNA endonuclease

"Sequence specific DNA-endonuclease" generally refers to all those enzymes which are capable of generating double-strand breaks in double stranded DNA in a sequence-specific manner at one or more recognition sequences. Said DNA cleavage may result in blunt ends, or so called "sticky" ends of the DNA (having a 5'- or 3'-overhang). The cleavage site may be localized within or outside the recognition sequence. Various kinds of endonucleases can be employed. Endonucleases can be, for

example, of the Class II or Class IIs type. Class IIs R-M restriction endonucleases catalyze the DNA cleavage at sequences other than the recognition sequence, i.e. they cleave at a DNA sequence at a particular number of nucleotides away from the recognition sequence (Szybalski *et al.* (1991) Gene 100:13-26). The following may be mentioned by way of example, but not by limitation:

1. Restriction endonucleases (e.g., type II or IIs), preferably homing endonucleases as described in detail herein below.
2. Chimeric or synthetic nucleases as described in detail herein below.

Unlike recombinases, restriction enzymes typically do not ligate DNA, but only cleave DNA. Restriction enzymes are described, for instance, in the New England Biolabs online catalog (www.neb.com), Promega online catalog (www.promega.com) and Rao *et al.* (2000) Prog Nucleic Acid Res Mol Biol 64:1-63. Within this invention "ligation" of the DNA ends resulting from the cleavage by the endonuclease is realized by fusion by homologous recombination of the homology sequences. The enzymes facilitating homologous recombination are naturally provided by the plant.

Preferably, the endonuclease is chosen in a way that its corresponding recognition sequences are rarely, if ever, found in the unmodified genome of the target plant organism. Ideally, the only copy (or copies) of the recognition sequence in the genome is (or are) the one(s) introduced by the DNA construct of the invention, thereby eliminating the chance that other DNA in the genome is excised or rearranged when the sequence-specific endonuclease is expressed.

One criterion for selecting a suitable endonuclease is the length of its corresponding recognition sequence. Said recognition sequence has an appropriate length to allow for rare cleavage, more preferably cleavage only at the recognition sequence(s) comprised in the DNA construct of the invention. One factor determining the minimum length of said recognition sequence is – from a statistical point of view – the size of the genome of the host organism. In an preferred embodiment the recognition sequence has a length of at least 10 base pairs, preferably at least 14 base pairs, more preferably at least 16 base pairs, especially preferably at least 18 base pairs, most preferably at least 20 base pairs.

A restriction enzyme that cleaves a 10 base pair recognition sequence is described in Huang B *et al.* (1996) J Protein Chem 15(5):481-9 .

Suitable enzymes are not only natural enzymes, but also synthetic enzymes. Preferred enzymes are all those sequence specific DNA-endonucleases whose recognition sequence is known and which can either be obtained in the form of their proteins (for example by purification) or expressed using their nucleic acid sequence. This is why homing endonucleases are very especially preferred (Review: (Belfort M and Roberts RJ (1997) Nucleic Acids Res 25: 3379-3388; Jasin M (1996) Trends Genet. 12:224-228; Internet: http://rebase.neb.com/rebase/rebase.homing.html). Owing to their long recognition sequences, they have no, or only a few, further recognition sequences in the chromosomal DNA of eukaryotic organisms in most cases.

The sequences encoding for such homing endonucleases can be isolated for example from the chloroplast genome of Chlamydomonas (Turmel M *et al.* (1993) J Mol Biol 232: 446-467). They are small (18 to 26 kD) and their open reading frames (ORF) have a "codon usage" which is suitable directly for nuclear expression in eukaryotes (Monnat RJ Jr *et al.* (1999) Biochem Biophys Res Com 255:88-93). Homing endonucleases which are very especially preferably isolated are the homing endonucleases I-SceI (WO96/14408), I-SceII (Sarguiel B *et al.* (1990) Nucleic Acids Res 18:5659-5665), I-SceIII (Sarguiel B *et al.* (1991) Mol Gen Genet. 255:340-341), I-CeuI (Marshall (1991) Gene 104:241-245), I-CreI (Wang J *et al.* (1997) Nucleic Acids Res 25: 3767-3776), I-ChuI (Cote V *et al.* (1993) Gene 129:69-76), I-TevI (Chu *et al.* (1990) Proc Natl Acad Sci USA 87:3574-3578; Bell-Pedersen *et al.* (1990) Nucleic Acids Res 18:3763-3770), I-TevII (Bell-Pedersen *et al.* (1990) Nucleic Acids Res 18:3763-3770), I-TevIII (Eddy *et al.* (1991) Genes Dev. 5:1032-1041), Endo SceI (Kawasaki *et al.* (1991) J Biol Chem 266:5342-5347), I-CpaI (Turmel M *et al.* (1995a) Nucleic Acids Res 23:2519-2525) and I-CpaII (Turmel M *et al.* (1995b) Mol. Biol. Evol. 12, 533-545).

Further homing endonucleases are detailed in the abovementioned Internet website, and examples which may be mentioned are homing endonucleases such as F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-AmaI, I-AniI, I-CeuI, I-CeuAIIP, I-ChuI, I-CmoeI, I-CpaI, I-CpaII, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-CsmI, I-CvuI, I-CvuAIP, I-DdiI, I-DdiII, I-DirI, I-DmoI, I-HmuI, I-HmuII, I-HspNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-PpoI, I-SPBetaIP, I-ScaI, I-SceI, I-SceII, I-SceIII , I-SceIV, I-SceV, I-SceVI, I-SceVII, I-SexIP, I-SneIP, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiS3bP, I-TdeIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPA1P, I-UarHGPA13P, I-VinIP, I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIP, PI-PfuI, PI-PfuII, PI-PkoI, PI-PkoII, PI-PspI, PI-Rma43812IP, PI-SPBetaIP, PI-SceI, PI-TfuI, PI-TfuII, PI-ThyI, PI-TliI, PI-TliII, H-DreI, I-BasI, I-BmoI, I-PogI, I-TwoI, PI-MgaI, PI-PabI, PI-PabII.

Preferred in this context are the homing endonucleases whose gene sequences are already known, such as, for example, F-SceI, I-CeuI, I-ChuI, I-DmoI, I-CpaI, I-CpaII, I-CreI, I-CsmI, F-TevI, F-TevII, I-TevI, I-TevII, I-AniI, I-CvuI, I-DdiI, I-HmuI, I-HmuII, I-LlaI, I-NanI, I-MsoI, I-NitI, I-NjaI, I-PakI, I-PorI, I-PpoI, I-ScaI, I-Ssp6803I, PI-PkoI, PI-PkoII, PI-PspI, PI-TfuI, PI-TliI. Especially preferred are commercially available homing endonucleases such as I-CeuI, I-SceI, I-DmoI, I-PpoI, PI-PspI or PI-SceI. Endonucleases with particularly long recognition sequences, and which therefore only rarely (if ever) cleave within a genome include: I-CeuI (26 bp recognition sequence), PI-PspI (30 bp recognition sequence), PI-SceI (39 bp recognition sequence), I-SceI (18 bp recognition sequence) and I-PpoI (15 bp recognition sequence).

The enzymes can be isolated from their organisms of origin in the manner with which the skilled worker is familiar, and/or their coding nucleic acid sequence can be cloned. The sequences of various enzymes are deposited in GenBank.

Very especially preferred are the homing endonucleases I-SceI, I-CpaI, I-CpaII, I-CreI and I-ChuI. Sequences encoding said nucleases are known in the art and – for exam-

ple – specified in WO 03/004659 (e.g., as SEQ ID NO: 2, 4, 6, 8, and 10 of WO 03/004659 hereby incorporated by reference).

In an preferred embodiment, the sequences encoding said homing endonucleases can be modified by insertion of an intron sequence. This prevents expression of a functional enzyme in procaryotic host organisms and thereby facilitates cloning and transformations procedures (e.g., based on E. coli or *Agrobacterium*). In plant organisms, expression of a functional enzyme is realized, since plants are able to recognize and "splice" out introns. Preferably, introns are inserted in the homing endonucleases mentioned as preferred above (e.g., into I-SceI or I-CreI). Another preferred embodiment of the invention is related to a intron-comprising I-Sce-I sequence and its use in methods of the invention (more preferably a sequence as described by SEQ ID NO: 14).

In some aspects of the invention, molecular evolution can be employed to create an improved endonuclease. Polynucleotides encoding a candidate endonuclease enzyme can, for example, be modulated with DNA shuffling protocols. DNA shuffling is a process of recursive recombination and mutation, performed by random fragmentation of a pool of related genes, followed by reassembly of the fragments by a polymerase chain reaction-like process. See, e.g., Stemmer (1994) Proc Natl Acad Sci USA 91:10747-10751; Stemmer (1994) Nature 370:389-391; and US 5,605,793, US 5,837,458, US 5,830,721 and US 5, 811,238.

Other synthetic sequence specific DNA-endonucleases which may be mentioned by way of example are chimeric nucleases which are composed of an unspecific nuclease domain and a sequence-specific DNA binding domain consisting of zinc fingers (Bibikova M *et al.* (2001) Mol Cell Biol. 21:289-297). These DNA-binding zinc finger domains can be adapted to suit any DNA sequence. Suitable methods for preparing suitable zink finger domaines are described and known to the skilled worker (Beerli RR *et al.*, Proc Natl Acad Sci U S A. 2000; 97 (4):1495-1500; Beerli RR, *et al.*, J Biol Chem 2000; 275(42):32617-32627; Segal DJ and Barbas CF 3rd., Curr Opin Chem Biol 2000; 4(1):34-39; Kang JS and Kim JS, J Biol Chem 2000; 275(12):8742-8748; Beerli RR *et al.*, Proc Natl Acad Sci USA 1998; 95(25):14628-14633; Kim JS *et al.*, Proc Natl Acad Sci USA 1997; 94(8):3616-3620; Klug A, J Mol Biol 1999; 293(2):215-218; Tsai SY *et al.*, Adv Drug Deliv Rev 1998;30(1-3):23-31; Mapp AK *et al.*, Proc Natl Acad Sci USA 2000; 97(8):3930-3935; Sharrocks AD *et al.*, Int J Biochem Cell Biol 1997; 29(12):1371-1387; Zhang L *et al.*, J Biol Chem 2000; 275(43):33850-33860).

The sequence specific DNA-endonuclease may be expressed as a fusion protein with a nuclear localization sequence (NLS). This NLS sequence enables facilitated transport into the nucleus and increases the efficacy of the recombination system. A variety of NLS sequences are known in the art (Jicks GR and Raikhel NV (1995) Annu Rev Cell Biol 11:155-188; WO 03/004659). Preferred for plant organisms is, for example, the NLS sequence of the SV40 large antigen. However, owing to the small size of many sequence specific DNA-endonucleases (such as, for example, the homing endonucleases), a NLS sequence is not necessarily required. These enzymes are capable of passing through the nuclear pores even without an additional NLS.

In a further preferred embodiment, the activity of the sequence specific DNA-endonuclease can be induced. Suitable methods have been described for sequence-specific recombinases (Angrand PO *et al.* (1998) Nucl. Acids Res. 26(13):3263-3269; Logie C and Stewart AF (1995) Proc Natl Acad Sci USA 92(13):5940-5944; Imai T *et al.* (2001) Proc Natl Acad Sci USA 98(1):224-228). These methods employ fusion proteins of the sequence specific DNA-endonuclease and the ligand binding domain for steroid hormone receptor (for example the human androgen receptor, or mutated variants of the human estrogen receptor as described therein). Induction may be effected with ligands such as, for example, estradiol, dexamethasone, 4-hydroxytamoxifen or raloxifen.

Some sequence specific DNA-endonucleases enzymes are active as dimers (homo- or heterodimers; I-CreI forms a homodimer; I-SecIV forms a heterodimer) (Wernette CM (1998) Biochemical & Biophysical Research Communications 248(1):127-333)). Dimerization can be designed as an inducible feature, for example by exchanging the natural dimerization domains for the binding domain of a low-molecular-weight ligand. Addition of a dimeric ligand then brings about dimerization of the fusion protein. Corresponding inducible dimerization methods, and the preparation of the dimeric ligands, have been described (Amara JF *et al.* (1997) Proc Natl Acad Sci USA 94(20): 10618-1623; Muthuswamy SK *et al.* (1999) Mol Cell Biol 19(10):6845-685; Schultz LW and Clardy J (1998) Bioorg Med Chem Lett. 8(1):1-6; Keenan T *et al.* (1998) Bioorg Med Chem. 6(8):1309-1335).

2. Recognition sequences for sequence specific DNA endonuclease

"Recognition sequence" refers to a DNA sequence that is recognized by a sequence-specific DNA endonuclease of the invention. The recognition sequence will typically be at least 10 base pairs long, is more usually 10 to 30 base pairs long, and in most embodiments, is less than 50 base pairs long.

"Recognition sequence" generally refers to those sequences which, under the conditions in a plant cell used within this invention, enable the recognition and cleavage by the sequence specific DNA-endonuclease. The recognition sequences for the respective sequence specific DNA-endonucleases are mentioned in Table 1 hereinbelow by way of example, but not by limitation.

Table 1: Recognition sequences and organisms of origin of sequence specific DNA-endonuclease ("^" indicates the cleavage site of the sequence specific DNA-endonuclease within a recognition sequence).

Nuclease	**Organism of origin**	**Recognition sequence**
I-AniI	*Aspergillus nidulans*	5'-TTGAGGAGGTT^TCTCTGTAAATAANNNNNNNNNNNNNNN 3'-AACTCCTCCAAAGAGACATTTATTNNNNNNNNNNNNNNN^
I-DdiI	*Dictyostelium discoideumAX3*	5'-TTTTTTGGTCATCCAGAAGTATAT 3'-AAAAAACCAG^TAGGTCTTCATATA
I-CvuI	*Chlorella vulgaris*	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-CsmI	*Chlamydomonas smithii*	5'-GTACTAGCATGGGGTCAAATGTCTTTCTGG

Nuclease	Organism of origin	Recognition sequence
I-Cmoel	*Chlamydomonas-moewusii*	5'-TCGTAGCAGCT^CACGGTT 3'-AGCATCG^TCGAGTGCCAA
I-Crel	*Chlamydomonas reinhardtii*	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-Chul	*Chlamydomonas humicola*	5'-GAAGGTTTGGCACCTCG^ATGTCGGCTCATC 3'-CTTCCAAACCGTG^GAGCTACAGCCGAGTAG
I-Cpal	*Chlamydomonas pallidostigmatica*	5'-CGATCCTAAGGTAGCGAA^ATTCA 3'-GCTAGGATTCCATC^GCTTTAAGT
I-Cpall	*Chlamydomonas pallidostigmatica*	5'-CCCGGCTAACTC^TGTGCCAG 3'-GGGCCGAT^TGAGACACGGTC
I-Ceul	*Chlamydomonas eugametos*	5'-CGTAACTATAACGGTCCTAA^GGTAGCGAA 3'-GCATTGATATTGCCAG^GATTCCATCGCTT
I-Dmol	*Desulfuro-coccus mobilis*	5'-ATGCCTTGCCGGGTAA^GTTCCGGCGCGCAT 3'-TACGGAACGGCC^CATTCAAGGCCGCGCGTA
I-Scel	*Saccharomyces cerevisiae*	5'-AGTTACGCTAGGGATAA^CAGGGTAATATAG 3'-TCAATGCGATCCC^TATTGTCCCATTATATC 5'-TAGGGATAA^CAGGGTAAT 3'-ATCCC^TATTGTCCCATTA ("Core"-Sequence)
I-SceII	*S.cerevisiae*	5'-TTTTGATTCTTTGGTCACCC^TGAAGTATA 3'-AAAACTAAGAAACCAG^TGGGACTTCATAT
I-SceIII	*S.cerevisiae*	5'-ATTGGAGGTTTTGGTAAC^TATTTATTACC 3'-TAACCTCCAAAACC^ATTGATAAATAATGG
I-SceIV	S.cerevisiae	5'-TCTTTTCTCTTGATTA^GCCCTAATCTACG 3'-AGAAAAGAGAAC^TAATCGGGATTAGATGC
I-SceV	*S.cerevisiae*	5'-AATAATTTTCT^TCTTAGTAATGCC 3'-TTATTAAAAGAAGAATCATTA^CGG
I-SceVI	*S.cerevisiae*	5'-GTTATTTAATG^TTTTAGTAGTTGG 3'-CAATAAATTACAAAATCATCA^ACC
I-SceVII	*S.cerevisiae*	5'-TGTCACATTGAGGTGCACTAGTTATTAC
PI-Scel	*S.cerevisiae*	5'-ATCTATGTCGGGTGC^GGAGAAAGAGGTAAT 3'-TAGATACAGCC^CACGCCTCTTTCTCCATTA
F-Scel	*S.cerevisiae*	5'-GATGCTGTAGGC^ATAGGCTTGGTT 3'-CTACGACA^TCCGTATCCGAACCAA
F-SceII	*S.cerevisiae*	5'-CTTTCCGCAACA^GTAAAATT 3'-GAAAGGCG^TTGTCATTTTAA
I-Hmul	*Bacillus subtilis* bacteriophage SPO1	5'-AGTAATGAGCCTAACGCTCAGCAA 3'-TCATTACTCGGATTGC^GAGTCGTT
I-Hmull	Bacillus subtilis bacteriophage SP82	5'-AGTAATGAGCCTAACGCTCAACAANNNNNNNNNNNNNNNN-NNNNNNNNNNNNNNNNNNNNN
I-Llal	*Lactococcus lactis*	5'-CACATCCATAAC^CATATCATTTTT 3'-GTGTAGGTATTGGTATAGTAA^AAA

Nuclease	Organism of origin	Recognition sequence
I-MsoI	*Monomastix species*	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-NanI	*Naegleria andersoni*	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
I-NitI	*Naegleria italica*	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
I-NjaI	*Naegleria jamieso-ni*	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
I-PakI	*Pseudendoclonium akinetum*	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-PorI	*Pyrobaculum organotrophum*	5'-GCGAGCCCGTAAGGGT^GTGTACGGG 3'-CGCTCGGGCATT^CCCACACATGCCC
I-PpoI	*Physarum polycephalum*	5'-TAACTATGACTCTCTTAA^GGTAGCCAAAT 3'-ATTGATACTGAGAG^AATTCCATCGGTTTA
I-ScaI	*Saccharomyces capensis*	5'-TGTCACATTGAGGTGCACT^AGTTATTAC 3'-ACAGTGTAACTCCAC^GTGATCAATAATG
I-Ssp6803I	*Synechocystis* species	5'-GTCGGGCT^CATAACCCGAA 3'-CAGCCCGAGTA^TTGGGCTT
PI-PfuI	*Pyrococcus furiosus Vc1*	5'-GAAGATGGGAGGAGGG^ACCGGACTCAACTT 3'-CTTCTACCCTCC^TCCCTGGCCTGAGTTGAA
PI-PfuII	*Pyrococcus furiosus Vc1*	5'-ACGAATCCATGTGGAGA^AGAGCCTCTATA 3'-TGCTTAGGTACAC^CTCTTCTCGGAGATAT
PI-PkoI	*Pyrococcus koda-karaensis KOD1*	5'-GATTTTAGAT^CCCTGTACC 3'-CTAAAA^TCTAGGGACATGG
PI-PkoII	*Pyrococcus koda-karaensis KOD1*	5'-CAGTACTACG^GTTAC 3'-GTCATG^ATGCCAATG
PI-PspI	*Pyrococcus sp.*	5'-AAAATCCTGGCAAACAGCTATTAT^GGGTAT 3'-TTTTAGGACCGTTTGTCGAT^AATACCCATA
PI-TfuI	*Thermococcus fumicolans ST557*	5'-TAGATTTTAGGT^CGCTATATCCTTCC 3'-ATCTAAAA^TCCAGCGATATAGGAAGG
PI-TfuII	*Thermococcus fumicolans ST557*	5'-TAYGCNGAYACN^GACGGYTTYT 3'-ATRCGNCT^RTGNCTGCCRAARA
PI-ThyI	*Thermococcus hydrothermalis*	5'-TAYGCNGAYACN^GACGGYTTYT 3'-ATRCGNCT^RTGNCTGCCRAARA
PI-TliI	*Thermococcus litoralis*	5'-TAYGCNGAYACNGACGG^YTTYT 3'-ATRCGNCTRTGNC^TGCCRAARA
PI-TliII	*Thermococcus litoralis*	5'-AAATTGCTTGCAAACAGCTATTACGGCTAT
I-TevI	Bacteriophage T4	5'-AGTGGTATCAAC^GCTCAGTAGATG 3'-TCACCATAGT^TGCGAGTCATCTAC
I-TevII	Bacteriophage T4	5'-GCTTATGAGTATGAAGTGAACACGT^TATTC 3'-CGAATACTCATACTTCACTTGTG^CAATAAG

Nuclease	Organism of origin	Recognition sequence
F-TevI	Bacteriophage T4	5'-GAAACACAAGA^AATGTTTAGTAAANNNNNNNNNNNNN 3'-CTTTGTGTTCTTTACAAATCATTTNNNNNNNNNNNNN^
F-TevII	Bacteriophage T4	5'-TTTAATCCTCGCTTC^AGATATGGCAACTG 3'-AAATTAGGAGCGA^AGTCTATACCGTTGAC
H-DreI	*E. coli pI-DreI*	5'-CAAAACGTCGTAA^GTTCCGGCGCG 3'-GTTTTGCAG^CATTCAAGGCCGCGC
I-BasI	*Bacillus thuringiensis phage Bastille*	5' AGTAATGAGCCTAACGCTCAGCAA 3'- TCATTACGAGTCGAACTCGGATTG
I-BmoI	*Bacillus mojavensis* s87-18	5'-GAGTAAGAGCCCG^TAGTAATGACATGGC 3'-CTCATTCTCG^GGCATCATTACTGTACCG
I-PogI	*Pyrobaculum oguniense*	5'-CTTCAGTAT^GCCCCGAAAC 3'-GAAGT^CATACGGGGCTTTG
I-TwoI	*Staphylococcus aureus phage Twort*	5'-TCTTGCACCTACACAATCCA 3'-AGAACGTGGATGTGTTAGGT
PI-MgaI	*Mycobacterium gastri*	5'-CGTAGCTGCCCAGTATGAGTCA 3'-GCATCGACGGGTCATACTCAGT
PI-PabI	*Pyrococcus abyssi*	5'-GGGGGCAGCCAGTGGTCCCGTT 3'-CCCCCGTCGGTCACCAGGGCAA
PI-PabII	*Pyrococcus abyssi*	5'-ACCCCTGTGGAGAGGAGCCCCTC 3'-TGGGGACACCTCTCCTCGGGGAG

Also encompassed are minor deviations (degenerations) of the recognition sequence which still enable recognition and cleavage by the sequence specific DNA-endonuclease in question. Such deviations - also in connection with different framework conditions such as, for example, calcium or magnesium concentration – have been described (Argast GM *et al.* (1998) J Mol Biol 280: 345-353). Also encompassed are core sequences of these recognition sequences and minor deviations (degenerations) in there. It is known that the inner portions of the recognition sequences suffice for an induced double-strand break and that the outer ones are not absolutely relevant, but can codetermine the cleavage efficacy. Thus, for example, an 18 bp core sequence can be defined for I-SceI.

3. Promoters of the Invention

Various promoters for expression in plants and plant cells can be employed in the invention. A first promoter – the parsley ubiquitin promoter - regulates the expression of the sequence-specific endonuclease. Other promoters may regulate the expression of the selection marker or the agronomically valuable trait.

3.1 Parsley ubiquitin promoter

Expression of the polynucleotide encoding a sequence-specific DNA endonuclease is controlled by a parsley ubiquitin promoter. The term "parsley ubiquitin promoter" mean the transcription regulating region of the ubiquitin gene from parsley (*Petroselinum crispum*), preferably the promoter sequences disclosed and claimed in international patent application WO 03/102198, hereby incorporated entirely by reference. More

preferably, a parsley ubiquitin promoter is described by the nucleic acid sequence of SEQ ID NO: 8 or 15, and functional equivalents and functional equivalent fragments thereof.

Functional equivalents means transcription regulating sequences derived from a sequence as described by or obtainable from SEQ ID NO: 8 or 15 for example by substitution, insertion or deletion of one or more nucleotides which have a identity of at least 30 %, preferably at least 50 % or 70 %, more preferably at least 90 %, most preferably at least 95 %, and demonstrate substantially the same transcription regulating properties than the parsley ubiquitin promoter as described by SEQ ID NO: 8 or 15. Functional equivalents may be obtained synthetically or from orthologous genes of other organisms by – for example – homology-based database screening or hybridization-based library screening.

Functionally equivalent fragments of a parsley ubiquitin promoter as described by SEQ ID NO: 8 or 15 can be obtained – for example – by deleting non-essential sequences without substantially modifying its transcription regulating properties. It is well known in the art that not all sequences in a promoter region are required for transcription regulation but that the essential regions are restricted to limited portions thereof (so called promoter elements). Functionally equivalent fragments of a promoter sequence can be obtained by deleting non-essential sequences (e.g., of a promoter sequence as described by SEQ ID NO: 8 or 15). Such a functionally equivalent fragment consists of at least 50, preferably at least 100, more preferably at least 150, most preferably at least 200 consecutive base pairs of a promoter as described by SEQ ID NO: 8 or 15 and has substantially the same promoter activity as the promoter described by SEQ ID NO: 8 or 15. Narrowing of a promoter sequence to specific, essential regulatory regions or elements can be facilitated by using computer algorithms for the prediction of promoter elements. In most promoters the essential regulatory regions are characterized by a clustering of promoter elements. A promoter element analysis can be done by computer programs like e.g., PLACE ("Plant Cis-acting Regulatory DNA Elements"; Higo K *et al.* (1999) Nucleic Acids Res 27:1, 297-300) or by using the BIOBASE database "Transfac" (Biologische Datenbanken GmbH, Braunschweig).

A promoter activity of a functional equivalent or equivalent fragment is regarded substantially the same if transcription of a specific nucleic acid sequence under transcriptional control of such sequences does not derivate more than 50%, preferably more than 40%, more preferably more than 30% or 20%, most preferably more than 10% from a comparison value obtained under same conditions using the promoter sequence as described by SEQ ID NO: 8 or 15. The level of expression may be higher or lower than the standard value. Preferably the transcription level is assessed by expression of nucleic acids encoding for readily quantifiable proteins such as reporter proteins (e.g., green fluorescence protein (GFP); Chui *et al.* (1996) Curr Biol 6 : 325-330; Leffel SM *et al.* (1997) Biotechniques. 23(5):912-8), chloramphenicol transferase, luciferase (Millar *et al.* (1992) Plant Mol Biol Rep 10 :324-414), β-galactosidase, or – preferably- β-glucuronidase (Jefferson *et al.* (1987) EMBO J. 6:3901-3907).

A functional equivalent preferably comprises one or more of the promoter elements identified in the parsley ubiquitin promoter presumably constituting its essential parts

for transcription regulation, such elements may be identified by computer algorithms such as PLACE (Higo *et al.* (1999) Nucl. Acid. Res., Vol. 27, No.1, 297-300) and may be selected from the group consisting of:

a) a putative heat shock inducible element (= HSE) at a position equivalent to position base 534-547 of SEQ ID NO: 8,

b) two CAAACAC-elements at a position equivalent to position base 264 to 270, and 716 (complementary strand) of SEQ ID NO: 8 (Stalberg K *et al.* (1996) Planta 199:515-519)

c) two AACAAAC-elements at a position equivalent to position base 140 to 146, and 461 (complementary strand) (Wu C *et al.* (2000) Plant J 23 : 415-421)

d) a TATA-Box (TATATATA) at a position equivalent to position base 291 to 297 of SEQ ID NO: 8 and close to the expected transcription start at position 237 (Joshi CP (1987) Nucleic Acids Res 15(16):6643-53)

e) all together 4 ACGTA-boxes (at a position equivalent to position base 214,674, 692, and 880, respectively)

f) abscisic acid responsive element (at a position equivalent to position base 227 of SEQ ID NO: 8) (Hattori T *et al.* (2002) Plant Cell Physiol 43 : 136-140) ;

g) several amylase-boxes at a position equivalent to position base 139, 421 (complementary strand), 462 (complementary strand), 789 (complementary strand), and 871 (complementary strand), respectively, of SEQ ID NO: 8 (Huang N *et al.* (1990) Plant Mol Biol 14 :655-668)

h) a CACGTG motif at a position equivalent to position base 565 of SEQ ID NO: 8 (Menkens AE (1995) Trends in Biochemistry 20:506-510)

i) altogether 16 GATA boxes (Gilmartin PM *et al.* (1990) Plant Cell 2:369-378)

j) 5 GT1 consensus binding sites (GRWAAW) at a position equivalent to position base 395, and on the complementary strand at position 52, 387, 504, and 647, respectively, of SEQ ID NO: 8 (Villain P *et al.* (1996) J Biol Chem 271:32593-32598)

k) a Ibox (GATAAG) at a position equivalent to position base 474 of SEQ ID NO: 8 (complementary strand) (Rose A et al. (1999) Plant J 20:641-652)

l) a LTRE (low-temperature-responsive element) CCGAAA at a position equivalent to position base 632 of SEQ ID NO: 8 (Dunn MA et al. (1998) Plant Mol Biol 38:551-564)

m) several binding sites for various classes of myb transcription factors (Jin H et al. (1999) Plant Mol Biol 41(5):577-85)

n) several W-box binding sites at a position equivalent to position base 549, 61, 550, and 919, respectively, of SEQ ID NO: 8, which are bound by WRKY transcription factors (Eulgem T et al. (2000) Trends Plant Sci 5:199-206).

Preferably the equivalent promoter comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 or all of the above mentioned elements. Preferably the promoter comprises at least the elements a, b, c, and d.

Sequence comparison between the parsley ubiquitin promoter (PcUbi4-2) and the maize ubiquitin promoter demonstrates a very low identify of only 26 % which is non-significant (Gap opening penalty 15, Gap extension penalty 6,66) (Altschul et al. (1990)

J Mol Biol 215:403-410, Altschul *et al.* (1997) Nucl. Acid Res 25:3389-3402). A BLAST search with one of sequences in the GenBank database would not identify the other promoter. The homology between the coding regions for the ubiquitins from maize and *P. crispum* is on nucleic acid level as high as 66,1%.

Accordingly, another subject matter of the invention relates to a transgenic expression cassettes comprising a sequence coding for a homing endonuclease operably linked to a parsley ubiquitin promoter as defined above.

Other embodiments of the invention are related to a transgenic vector comprising said expression cassette, and transgenic plants or plant cells comprising in their genome, preferably in their nuclear, chromosomal DNA, said expression cassette or said vector. Enclosed are also cells, cell cultures, tissues, parts or propagation material – such as, for example, in the case of plant organisms leaves, roots, seeds, fruit, pollen and the like – derived from said transgenic plants.

Obviously, also the promoter controlling expression of the agronomically valuable trait or marker sequence may be a parsley ubiquitin promoter.

3.2 Promoter for general use

Promoters for the expression of the marker sequence or the agronomically valuable trait can be selected from all promoter having activity in plants or parts thereof. These promoters are selected for the tissues or cells where expression of the marker sequence and/or trait gene is desired. A number of exemplary promoters are described below. The following promoters, however, are only provided as examples and are not intended to limit the invention. Those of skill in the art will recognize that other promoters with desired expression patterns are well known or can be selected with routine molecular techniques.

A promoter can be derived from a gene that is under investigation, or can be a heterologous promoter that is obtained from a different gene, or from a different species. Suitable promoters can be derived from plants or plant pathogens like e.g., plant viruses. Where expression of a gene in all tissues of a transgenic plant or other organism is desired, one can use a "constitutive" promoter, which is generally active under most environmental conditions and states of development or cell differentiation (Benfey *et al.* (1989) EMBO J. 8:2195-2202). The promoter controlling expression of the trait gene and/or marker sequence can be constitutive. Suitable constitutive promoters for use in plants include, for example, the cauliflower mosaic virus (CaMV) 35S transcription initiation region (Franck *et al.* (1980) Cell 21:285-294; Odell *et al.* (1985) Nature 313:810-812; Shewmaker *et al.* (1985) Virology 140:281-288; Gardner *et al.* 1986, Plant Mol. Biol. 6, 221-228), the 19S transcription initiation region (US 5,352,605 and WO 84/02913), and region VI promoters, the 1'-or 2'-promoter derived from T-DNA of *Agrobacterium* tumefaciens, and other promoters active in plant cells that are known to those of skill in the art. Other suitable promoters include the full-length transcript promoter from Figwort mosaic virus, actin promoters, histone promoters, tubulin promoters, or the mannopine synthase promoter (MAS). Other constitutive plant promoters include various ubiquitin or polyubiquitin promoters derived from, inter alia, *Arabidopsis* (Sun and Callis (1997) Plant J 11(5): 1017-1027), the mas, Mac or DoubleMac promot-

ers (US 5,106,739; Comai *et al.* (1990) Plant Mol Biol 15:373-381), the ubiquitin promoter (Holtorf S *et al.* (1995) Plant Mol Biol 29:637-649) and other transcription initiation regions from various plant genes known to those of skill in the art. Useful promoters for plants also include those obtained from Ti-or Ri-plasmids, from plant cells, plant viruses or other organisms whose promoters are found to be functional in plants. Bacterial promoters that function in plants, and thus are suitable for use in the methods of the invention include the octopine synthetase promoter, the nopaline synthase promoter, and the mannopine synthetase promoter. Suitable endogenous plant promoters include the ribulose-1,6-biphosphate (RUBP) carboxylase small subunit (ssu) promoter, the α-conglycinin promoter, the phaseolin promoter, the ADH promoter, and heat-shock promoters.

Of course, promoters can regulate expression all of the time in only one or some tissues. Alternatively, a promoter can regulate expression in all tissues but only at a specific developmental time point.

One can use a promoter that directs expression of a gene of interest in a specific tissue or is otherwise under more precise environmental or developmental control. Examples of environmental conditions that may affect transcription by inducible promoters include pathogen attack, anaerobic conditions, ethylene or the presence of light. Promoters under developmental control include promoters that initiate transcription only in certain tissues or organs, such as leaves, roots, fruit, seeds, or flowers, or parts thereof. The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

Examples of tissue-specific plant promoters under developmental control include promoters that initiate transcription only in certain organs or tissues, such as fruits, seeds, flowers, anthers, ovaries, pollen, the meristem, flowers, leaves, stems, roots and seeds. The tissue-specific ES promoter from tomato is particularly useful for directing gene expression so that a desired gene product is located in fruits. See, e. g., Lincoln *et al.* (1988) Proc Natl Acad Sci USA 84:2793-2797; Deikman *et al.* (1988) EMBO J 7:3315-3320 ; Deikman *et al.* (1992) Plant Physiol 100:2013-2017. Other suitable seed specific promoters include those derived from the following genes: MAC1 from maize (Sheridan *et al.* (1996) Genetics 142:1009-1020, Cat3 from maize (GenBank No. L05934, Ableretal. (1993) Plant Mol Biol 22:10131-1038, the gene encoding oleosin 18kD from maize (GenBank No. J05212, Lee *et al.* (1994) Plant Mol Biol 26:1981-1987), viviparous-1 from *Arabidopsis* (Genbank No. U93215), the gene encoding oleosin from *Arabidopsis* (Genbank No. Z17657), Atmycl from *Arabidopsis* (Urao *et al.* (1996) Plant Mol Biol 32:571-576, the 2s seed storage protein gene family from *Arabidopsis* (Conceicao *et al.* (1994) Plant 5:493-505) the gene encoding oleosin 20kD from Brassica napus (GenBank No. M63985), napin from Brassica napus (GenBank No. J02798, Josefsson *et al.* (1987) J. Biol. Chem. 262:12196-12201), the napin gene family (e.g., from Brassica napus ; Sjodahl *et al.* (1995) Planta 197:264-271, US 5,608,152; Stalberg K, *et al.* (1996) L. Planta 199: 515-519), the gene encoding the 2S storage protein from Brassica napus (Dasgupta *et al.* (1993) Gene 133: 301-302), the genes encoding oleosin A (Genbank No. U09118) and oleosin B (Genbank No. U09119) from soybean, the gene encoding low molecular weight sulphur rich protein from soybean (Choi *et al.* (1995) Mol Gen Genet 246:266-268), the phaseolin gene (US 5,504,200,

Bustos MM *et al.*, Plant Cell. 1989;1(9):839-53), the 2S albumin gene (Joseffson LG *et al.*(1987) J Biol Chem 262: 12196-12201), the legumin gene (Shirsat A *et al.* (1989) Mol Gen Genet. 215(2):326-331), the USP (unknown seed protein) gene (Bäumlein H *et al.* (1991) Mol Gen Genetics 225(3):459-67), the sucrose binding protein gene (WO 00/26388), the legumin B4 gene (LeB4; Bäumlein H *et al.* (1991) Mol Gen Genet 225:121-128; Baeumlein *et al.* (1992) Plant J 2(2):233-239; Fiedler U *et al.* (1995) Biotechnology (NY) 13(10):1090-1093), the Ins *Arabidopsis* oleosin gene (WO9845461), the Brassica Bce4 gene (WO 91/13980), genes encoding the "high-molecular-weight glutenin" (HMWG), gliadin, branching enzyme, ADP-glucose pyrophosphatase (AGPase) or starch synthase. Furthermore preferred promoters are those which enable seed-specific expression in monocots such as maize, barley, wheat, rye, rice and the like. Promoters which may advantageously be employed are the promoter of the lpt2 or lpt1 gene (WO 95/15389, WO 95/23230) or the promoters described in WO 99/16890 (promoters of the hordein gene, the glutelin gene, the oryzin gene, the prolamine gene, the gliadin gene, the zein gene, the kasirin gene or the secalin gene).

Further suitable promoters are, for example, specific promoters for tubers, storage roots or roots such as, for example, the class I patatin promoter (B33), the potato cathepsin D inhibitor promoter, the starch synthase (GBSS1) promoter or the sporamin promoter, and fruit-specific promoters such as, for example, the tomato fruit-specific promoter(EP-A 409 625).

Promoters which are furthermore suitable are those which ensure leaf-specific expression. Promoters which may be mentioned are the potato cytosolic FBPase promoter (WO 98/18940), the Rubisco (ribulose-1,5-bisphosphate carboxylase) SSU (small subunit) promoter or the potato ST-LSI promoter (Stockhaus *et al.* (1989) EMBO J 8(9):2445-2451). Other preferred promoters are those which govern expression in seeds and plant embryos.

Further suitable promoters are, for example, fruit-maturation-specific promoters such as, for example, the tomato fruit-maturation-specific promoter (WO 94/21794), flower-specific promoters such as, for example, the phytoene synthase promoter (WO 92/16635) or the promoter of the P-rr gene (WO 98/22593) or another node-specific promoter as described in EP-A 249676 may be used advantageously. The promoter may also be a pith-specific promoter, such as the promoter isolated from a plant TrpA gene as described in W0 93/07278.

A development-regulated promoter is, inter alia, described by Baerson *et al.* (Baerson SR, Lamppa GK (1993) Plant Mol Biol 22(2):255-67).

Other preferred promoters are promoters induced by biotic or abiotic stress, such as, for example, the pathogen-inducible promoter of the PRP1 gene (Ward *et al.*, Plant Mol Biol 1993, 22: 361-366), the tomato heat-inducible hsp80 promoter (US 5,187,267), the potato chill-inducible alpha-amylase promoter (WO 96/12814) or the wound-induced pinII promoter (EP375091).

Promoters may also encompass further promoters, promoter elements or minimal promoters capable of modifying the expression-specific characteristics. Thus, for example,

the tissue-specific expression may take place in addition as a function of certain stress factors, owing to genetic control sequences. Such elements are, for example, described for water stress, abscisic acid (Lam E and Chua NH (1991) J Biol Chem 266(26):17131 -17135) and heat stress (Schoffl F *et al.* (1989) Molecular & General Genetics 217(2-3):246-53).

4. The homology sequences

Referring to the homology sequences (e.g., A, A') "sufficient length" preferably refers to sequences with a length of at least 20 base pairs, preferably at least 50 base pairs, especially preferably at least 100 base pairs, very especially preferably at least 300 base pairs, most preferably at least 500 base pairs.

Referring to the homology sequences (e.g., A, A'), "sufficient homology" preferably refers to sequences with at least 70%, preferably 80%, by preference at least 90%, especially preferably at least 95%, very especially preferably at least 99%, most preferably 100%, homology within these homology sequences over a length of at least 20 base pairs, preferably at least 50 base pairs, especially preferably at least 100 base pairs, very especially preferably at least 300 base pairs, most preferably at least 500 base pairs.

The homology sequences A and A' are preferably organized in the form of a direct repeat. The term "direct repeat" means a subsequent localization of two sequences on the same strand of a DNA molecule in the same orientation, wherein these two sequences fulfill the above given requirements for homologous recombination between said two sequences.

In an preferred embodiment, the homology sequences may be a duplication of a sequence having additional use within the DNA construct. For example, the homology sequences may be two transcription terminator sequences. One of these terminator sequences may be operably linked to the agronomically valuable trait, while the other may be linked to the marker sequence, which is localized in 3'-direction of the trait gene. Recombination between the two terminator sequences will excise the marker sequence but will reconstitute the terminator of the trait gene (see Fig. 4).

In another example, the homology sequences may be two promoter sequences. One of these promoter sequences may be operably linked to the agronomically valuable trait, while the other may be linked to the marker sequence, which is localized in 5'-direction of the trait gene. Recombination between the two promoter sequences will excise the marker sequence but will reconstitute the promoter of the trait gene (see Fig. 3).

The person skilled in the art will know that the homology sequences do not need to be restricted to a single functional element (e.g. promoter or terminator), but may comprise or extent to other sequences (e.g. being part of the coding region of the trait gene and the respective terminator sequence of said trait gene (see Fig. 5).

5. Additional elements in the DNA construct

The DNA construct may – beside the various promoter sequences – comprise additional genetic control sequences. The term "genetic control sequences" is to be under-

stood in the broad sense and refers to all those sequences which affect the making or function of the DNA construct to the invention or an expression cassette comprised therein. Preferably, a expression cassettes according to the invention encompass 5'-upstream of the respective nucleic acid sequence to be expressed a promoter and 3'-downstream a terminator sequence as additional genetic control sequence, and, if appropriate, further customary regulatory elements, in each case in operable linkage with the nucleic acid sequence to be expressed.

Genetic control sequences are described, for example, in "Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990)" or "Gruber and Crosby, in: Methods in Plant Molecular Biology and Biotechnolgy, CRC Press, Boca Raton, Florida, eds.: Glick and Thompson, Chapter 7, 89-108" and the references cited therein.

Examples of such control sequences are sequences to which inductors or repressors bind and thus regulate the expression of the nucleic acid. Genetic control sequences furthermore also encompass the the 5'-untranslated region, introns or the noncoding 3'-region of genes. It has been demonstrated that they may play a significant role in the regulation of gene expression. Thus, it has been demonstrated that 5'-untranslated sequences are capable of enhancing the transient expression of heterologous genes. Furthermore, they may promote tissue specificity (Rouster J *et al.* (1998) Plant J 15:435-440.). Conversely, the 5'-untranslated region of the opaque-2 gene suppresses expression. Deletion of the region in question leads to an increased gene activity (Lohmer S *et al.* (1993) Plant Cell 5:65-73). Genetic control sequences may also encompass ribosome binding sequences for initiating translation. This is preferred in particular when the nucleic acid sequence to be expressed does not provide suitable sequences or when they are not compatible with the expression system.

The expression cassette can advantageously comprise one or more of what are known as enhancer sequences in operable linkage with the promoter, which enable the increased transgenic expression of the nucleic acid sequence. Additional advantageous sequences, such as further regulatory elements or terminators, may also be inserted at the 3' end of the nucleic acid sequences to be expressed recombinantly. One or more copies of the nucleic acid sequences to be expressed recombinantly may be present in the gene construct. Genetic control sequences are furthermore understood as meaning sequences which encode fusion proteins consisting of a signal peptide sequence.

Polyadenylation signals which are suitable as genetic control sequences are plant polyadenylation signals, preferably those which correspond essentially to T-DNA polyadenylation signals from *Agrobacterium tumefaciens*. Examples of particularly suitable terminator sequences are the OCS (octopine synthase) terminator and the NOS (nopaline synthase) terminator. Also preferably are those taken from viral sequences, e.g. the 35S terminator.

The DNA constructs according to the invention and any vectors derived therefrom may comprise further functional elements. The term "further functional elements" is to be understood in the broad sense. It preferably refers to all those elements which affect the generation, multiplication, function, use or value of said DNA construct or vectors

comprising said DNA construct, or cells or organisms comprising the before mentioned. These further functional elements may include but shall not be limited to:

i) Origins of replication which ensure replication of the expression cassettes or vectors according to the invention in, for example, E. coli. Examples which may be mentioned are ORI (origin of DNA replication), the pBR322 ori or the P15A ori (Sambrook *et al.*: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

ii) Multiple cloning sites (MCS) to enable and facilitate the insertion of one or more nucleic acid sequences.

iii) Sequences which make possible homologous recombination or insertion into the genome of a host organism.

iv) Elements, for example border sequences, which make possible the *Agrobacterium*-mediated transfer in plant cells for the transfer and integration into the plant genome, such as, for example, the right or left border of the T-DNA or the vir region.

6. The Marker Sequence

The term "marker sequence" is to be understood in the broad sense to include all nucleotide sequences (and/or polypeptide sequences translated therefrom) which facilitate detection, identification, or selection of transformed cells, tissues or organism (e.g., plants). The terms "sequence allowing selection of a transformed plant material", "selection marker" or "selection marker gene" or "selection marker protein" or "marker" have essentially the same meaning.

Markers may include (but are not limited to) selectable marker and screenable marker. A selectable marker confers to the cell or organism a phenotype resulting in a growth or viability difference. The selectable marker may interact with a selection agent (such as a herbicide or antibiotic or pro-drug) to bring about this phenotype. A screenable marker confers to the cell or organism a readily detectable phenotype, preferably a visibly detectable phenotype such a color or staining. The screenable marker may interact with a screening agent (such as a dye) to bring about this phenotype.

Selectable marker (or selectable marker sequences) comprise but are not limited to

a) negative selection marker, which confer a resistance against toxic (in case of plants phytotoxic) agent such as an antibiotic, herbicides or other biocides,

b) counter selection marker, which confer a sensitivity against certain chemical compounds (e.g., by converting a non-toxic compound into a toxic compound), and

c) positive selection marker, which confer a growth advantage (e.g., by expression of key elements of the cytokinin or hormone biosynthesis leading to the production of a plant hormone e. g., auxins, gibberllins, cytokinins, abscisic acid and ethylene; Ebinuma H *et al.* (2000) Proc Natl Acad Sci USA 94:2117-2121).

When using negative selection markers, only plants are selected which comprise said negative selection marker. When using counter selection marker, only plants are selected which lack said counter-selection marker. Counter-selection marker may be em-

ployed to verify successful excision of a sequence (comprising said counter-selection marker) from a genome. Screenable marker sequences include but are not limited to reporter genes (e. g. luciferase, glucuronidase, chloramphenicol acetyl transferase (CAT, etc.). Preferred marker sequences include but shall not be limited to:

i) Negative selection marker

As a rule, negative selection markers are useful for selecting cells which have successfully undergone transformation. The negative selection marker, which has been introduced with the DNA construct of the invention, may confer resistance to a biocide or phytotoxic agent (for example a herbicide such as phosphinothricin, glyphosate or bromoxynil), a metabolism inhibitor such as 2-deoxyglucose-6-phosphate (WO 98/45456) or an antibiotic such as, for example, tetracyclin, ampicillin, kanamycin, G 418, neomycin, bleomycin or hygromycin to the cells which have successfully undergone transformation. The negative selection marker permits the selection of the transformed cells from untransformed cells (McCormick *et al.* (1986) Plant Cell Reports 5:81-84). Negative selevtion marker in a vector of the invention may be employed to conferr resistance in more than one organism. For example a vector of the invention may comprise a selection marker for amplification in bacteria (such as E.coli or Agrobacterium) and plants. Examples of selectable markers for E. coli include: genes specifying resistance to antibiotics, i.e., ampicillin, tetracycline, kanamycin, erythromycin, or genes conferring other types of selectable enzymatic activities such as galactosidase, or the lactose operon. Suitable selectable markers for use in mammalian cells include, for example, the dihydrofolate reductase gene (DHFR), the thymidine kinase gene (TK), or prokaryotic genes conferring drug resistance, gpt (xanthine-guanine phosphoribosyltransferase, which can be selected for with mycophenolic acid; neo (neomycin phosphotransferase), which can be selected for with G418, hygromycin, or puromycin; and DHFR (dihydrofolate reductase), which can be selected for with methotrexate (Mulligan & Berg (1981) Proc Natl Acad Sci USA 78:2072; Southern & Berg (1982) J Mol Appl Genet 1: 327). Selection markers for plant cells often confer resistance to a biocide or an antibiotic, such as, for example, kanamycin, G 418, bleomycin, hygromycin, or chloramphenicol, or herbicide resistance, such as resistance to chlorsulfuron or Basta.

Especially preferred negative selection markers are those which confer resistance to herbicides. Examples of negative selection markers are:

- DNA sequences which encode phosphinothricin acetyltransferases (PAT), which acetylates the free amino group of the glutamine synthase inhibitor phosphinothricin (PPT) and thus brings about detoxification of PPT (de Block *et al.* (1987) EMBO J 6:2513-2518) (also referred to as Bialophos® resistence gene bar; EP 242236),
- 5-enolpyruvylshikimate-3-phosphate synthase genes (EPSP synthase genes), which confer resistence to Glyphosate® (N-(phosphonomethyl)glycine),
- the gox gene, which encodes the Glyphosate®-degrading enzyme Glyphosate oxidoreductase,
- the deh gene (encoding a dehalogenase which inactivates Dalapon®),
- acetolactate synthases which confer resistance to sulfonylurea and imidazolinone,
- bxn genes which encode Bromoxynil®-degrading nitrilase enzymes,

- the kanamycin, or G418, resistence gene (NPTII). The NPTII gene encodes a neomycin phosphotransferase which reduces the inhibitory effect of kanamycin, neomycin, G418 and paromomycin owing to a phosphorylation reaction (Beck et al (1982) Gene 19: 327),
- the DOG^R1 gene. The DOG^R1 gene has been isolated from the yeast Saccharomyces cerevisiae (EP 0 807 836). It encodes a 2-deoxyglucose-6-phosphate phosphatase which confers resistence to 2-DOG (Randez-Gil *et al.* (1995) Yeast 11:1233-1240).
- the hyg gene, which codes for the enzyme hygromycin phosphotransferase and confers resistance to the antibiotic hygromycin (Gritz and Davies (1983) Gene 25: 179);
- especially preferred are negative selection markers that confer resistance against the toxic effects imposed by D-amino acids like e.g., D-alanine and D-serine (WO 03/060133; Erikson 2004). Especially preferred as negative selection marker in this contest are the *dao*l gene (EC: 1.4. 3.3 : GenBank Acc.-No.: U60066) from the yeast *Rhodotorula gracilis* (*Rhodosporidium toruloides*) and the *E. coli* gene *dsd*A (D-serine dehydratase (D-serine deaminase) (EC: 4.3. 1.18; GenBank Acc.-No.: J01603).

ii) Positive selection marker

Positive selection marker comprise but are not limited to growth stimulating selection marker.asGenes like isopentenyltransferase from *Agrobacterium* tumefaciens (strain:PO22; Genbank Acc.-No.: AB025109) may – as a key enzyme of the cytokinin biosynthesis – facilitate regeneration of transformed plants (e.g., by selection on cytokinin-free medium). Corresponding selection methods are described (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). Additional positive selection markers, which confer a growth advantage to a transformed plant in comparison with a non-transformed one, are described e.g., in EP-A 0 601 092. Growth stimulation selection markers may include (but shall not be limited to) β-Glucuronidase (in combination with e.g., a cytokinin glucuronide), mannose-6-phosphate isomerase (in combination with mannose), UDP-galactose-4-epimerase (in combination with e.g., galactose), wherein mannose-6-phosphate isomerase in combination with mannose is especially preferred.

iii) Counter selection markers

Counter-selection markerenable the selection of organisms with successfully deleted sequences (Koprek T *et al.* (1999) Plant J 19(6):719-726). TK thymidine kinase (TK) and diphtheria toxin A fragment (DT-A), codA gene encoding a cytosine deaminase (Gleve AP *et al.* (1999) Plant Mol Biol 40(2):223-35; Pereat RI *et al.* (1993) Plant Mol Biol 23(4):793-799; Stougaard J (1993) Plant J 3:755-761), the cytochrome P450 gene (Koprek *et al.* (1999) Plant J 16:719-726), genes encoding a haloalkane dehalogenase (Naested H (1999) Plant J 18:571-576), the iaaH gene (Sundaresan V *et al.* (1995) Genes & Development 9:1797-1810), the tms2 gene (Fedoroff NV & Smith DL (1993) Plant J 3:273- 289), and D-amino acid oxidases causing toxic effects by conversion of D-amino acids (WO 03/ 060133).

In a preferred embodiment the excision cassette includes at least one of said counter-selection markers to distinguish plant cells or plants with successfully excised sequences from plant which still contain these. In a more preferred embodiment the excision cassette of the invention comprises a dual-function marker i.e. a marker with can be employed as both a negative and a counter selection marker depending on the substrate employed in the selection scheme. An example for a dual-function marker is the *dao*l gene (EC: 1.4. 3.3 : GenBank Acc.-No.: U60066) from the yeast *Rhodotorula gracilis,* which can be empo,oyed as negative selection marker with D.-amino acids such as D-alanine and D-serine, and as counter-selection marker with D-amino acids such as D-isoleucine and D-valine (see European Patent Appl. No.: 04006358.8)

iv) Screenable marker (reporter genes)

Screenable marker (such as reporter genes) encode readily quantifiable or detectable proteins and which, via intrinsic color or enzyme activity, ensure the assessment of the transformation efficacy or of the location or timing of expression. Especially preferred are genes encoding reporter proteins (see also Schenborn E, Groskreutz D. (1999) Mol Biotechnol 13(1):29-44) such as

- "green fluorescence protein" (GFP) (Chui WL *et al.* (1996) Curr Biol 6:325-330; Leffel SM *et al.* (1997) Biotechniques 23(5):912-8; Sheen *et al.* (1995) Plant J 8(5):777-784; Haseloff *et al.* (1997) Proc Natl Acad Sci 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).
- Chloramphenicol transferase,
- luciferase (Millar *et al.* (1992) Plant Mol Biol Rep 10:324-414; Ow *et al.* (1986) Science 234:856-859) permits selection by detection of bioluminescence,
- β–galactosidase, encodes an enzyme for which a variety of chromogenic substrates are available,
- ß-glucuronidase (GUS) (Jefferson *et al.* (1987) EMBO J 6:3901-3907) or the uidA gene, which encodes an enzyme for a variety of chromogenic substrates,
- R locus gene product: protein which regulates the production of anthocyanin pigments (red coloration) in plant tissue and thus makes possible the direct analysis of the promoter activity without the addition of additional adjuvants or chromogenic substrates (Dellaporta *et al.* (1988) In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, 11:263-282,),
- ß-lactamase (Sutcliffe (1978) Proc Natl Acad Sci USA 75:3737-3741), enzyme for a variety of chromogenic substrates (for example PADAC, a chromogenic cephalosporin),
- xylE gene product (Zukowsky *et al.* (1983) Proc Natl Acad Sci USA 80:1101-1105), catechol dioxygenase capable of converting chromogenic catechols,
- α-amylase (Ikuta *et al.* (1990) Bio/technol. 8:241-242),
- tyrosinase (Katz *et al.*(1983) J Gene Microbiol 129:2703-2714), enzyme which oxidizes tyrosine to give DOPA and dopaquinone which subsequently form melanine, which is readily detectable,
- aequorin (Prasher *et al.*(1985) Biochem Biophys Res Commun 126(3):1259-1268), can be used in the calcium-sensitive bioluminescence detection.

7. Target Organisms

The methods of the invention are useful for obtaining marker-free plants, or cells, parts, tissues, harvested material derived therefrom.

The term "plant" includes whole plants, shoot vegetative organs/structures (e. g. leaves, stems and tubers), roots, flowers and floral organs/structures (e. g. bracts, sepals, petals, stamens, carpels, anthers and ovules), seeds (including embryo, endosperm, and seed coat) and fruits (the mature ovary), plant tissues (e. g. vascular tissue, ground tissue, and the like) and cells (e. g. guard cells, egg cells, trichomes and the like), and progeny of same. The class of plants that can be used in the method of the invention is generally as broad as the class of higher and lower plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), gymnosperms, ferns, and multicellular algae. It includes plants of a variety of ploidy levels, including aneuploid, polyploid, diploid, haploid and hemizygous.

Included within the scope of the invention are all genera and species of higher and lower plants of the plant kingdom. Included are furthermore the mature plants, seed, shoots and seedlings, and parts, propagation material (for example seeds and fruit) and cultures, for example cell cultures, derived therefrom.

Preferred are plants and plant materials of the following plant families: Amaranthaceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Tetragoniaceae.

Annual, perennial, monocotyledonous and dicotyledonous plants are preferred host organisms for the generation of transgenic plants. The use of the recombination system, or method according to the invention is furthermore advantageous in all ornamental plants, useful or ornamental trees, flowers, cut flowers, shrubs or turf. Said plant may include – but shall not be limited to - bryophytes such as, for example, Hepaticae (hepaticas) and Musci (mosses); pteridophytes such as ferns, horsetail and clubmosses; gymnosperms such as conifers, cycads, ginkgo and Gnetaeae; algae such as Chlorophyceae, Phaeophpyceae, Rhodophyceae, Myxophyceae, Xanthophyceae, Bacillariophyceae (diatoms) and Euglenophyceae.

Plants for the purposes of the invention may comprise the families of the Rosaceae such as rose, Ericaceae such as rhododendrons and azaleas, Euphorbiaceae such as poinsettias and croton, Caryophyllaceae such as pinks, Solanaceae such as petunias, Gesneriaceae such as African violet, Balsaminaceae such as touch-me-not, Orchidaceae such as orchids, Iridaceae such as gladioli, iris, freesia and crocus, Compositae such as marigold, Geraniaceae such as geraniums, Liliaceae such as drachaena, Moraceae such as ficus, Araceae such as philodendron and many others.

The transgenic plants according to the invention are furthermore selected in particular from among dicotyledonous crop plants such as, for example, from the families of the Leguminosae such as pea, alfalfa and soybean; Solanaceae such as tobacco and and many others; the family of the Umbelliferae, particularly the genus Daucus (very particularly the species carota (carrot)) and Apium (very particularly the species

graveolens dulce (celery)) and many others; the family of the Solanaceae, particularly the genus Lycopersicon, very particularly the species esculentum (tomato) and the genus Solanum, very particularly the species tuberosum (potato) and melongena (aubergine) and many others; and the genus Capsicum, very particularly the species annum (pepper) and many others; the family of the Leguminosae, particularly the genus Glycine, very particularly the species max (soybean) and many others; and the family of the Cruciferae, particularly the genus Brassica, very particularly the species napus (oilseed rape), campestris (beet), oleracea cv Tastie (cabbage), oleracea cv Snowball Y (cauliflower) and oleracea cv Emperor (broccoli); and the genus *Arabidopsis*, very particularly the species thaliana and many others; the family of the Compositae, particularly the genus Lactuca, very particularly the species sativa (lettuce) and many others.

The transgenic plants according to the invention are selected in particular among monocotyledonous crop plants, such as, for example, cereals such as wheat, barley, sorghum and millet, rye, triticale, maize, rice or oats, and sugar cane. Especially preferred are *Arabidopsis* thaliana, Nicotiana tabacum, oilseed rape, soybean, corn (maize), wheat, linseed, potato and tagetes.

Plant organisms are furthermore, for the purposes of the invention, other organisms which are capable of photosynthetic activity, such as, for example, algae or cyanobacteria, and also mosses. Preferred algae are green algae, such as, for example, algae of the genus Haematococcus, Phaedactylum tricornatum, Volvox or Dunaliella.

Genetically modified plants according to the invention which can be consumed by humans or animals can also be used as food or feedstuffs, for example directly or following processing known in the art.

8. Generation of the plants for the method of the invention

Within the method of the invention it as an essential feature that two plants are crossed each of these comprising a specific DNA construct:

i) a first plant (the "endonuclease master plant") comprising an expression cassette for expression of sequence specific DNA-endonuclease (the "endonuclease expression cassette"). Expression here is under the control of a parsley ubiquitin promoter as specified above.

ii) a second plant (the "trait plant") comprising a recombination cassette for excision of a marker sequence and further comprising – optionally – an expression cassette for an agronomically valuable trait.

The individual features and preferred embodiments for the elements of said expression constructs or recombination cassettes are explained above in detail. The generation of the endonuclease master plant and the trait plant can be done by any of the multiple methods known in the art. The following procedures are only given by way of example.

8.1 Construction of Polynucleotide Constructs

Typically, DNA constructs (e.g., for an expression or recombination cassette) to be introduced into plants or plant cells are prepared using transgene expression tech-

niques. Recombinant expression techniques involve the construction of recombinant nucleic acids and the expression of genes in transfected cells. Molecular cloning techniques to achieve these ends are known in the art. A wide variety of cloning and in vitro amplification methods suitable for the construction of recombinant nucleic acids are well-known to persons of skill. Examples of these techniques and instructions sufficient to direct persons of skill through many cloning exercises are found in Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol.152, Academic Press, hic., San Diego, CA (Berger) ; Current Protocols in Molecular Biology, F. M. Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1998 Supplement), T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989), in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984). Preferably, the DNA constructs employed in the invention are generated by joining the abovementioned essential constituents of the DNA construct together in the abovementioned sequence using the recombination and cloning techniques with which the skilled worker is familiar.

Generally, a gene to be expressed will be present in an expression cassette, meaning that the gene is operably linked to expression control signals, e. g., promoters and terminators, that are functional in the host cell of interest. The genes that encode the sequence-specific DNA cleaving enzyme and, optionally, the selectable marker, will also be under the control of such signals that are functional in the host cell. Control of expression is most easily achieved by selection of a promoter. The transcription terminator is not generally as critical and a variety of known elements may be used so long as they are recognized by the cell. The invention contemplates polynucleotides operably linked to a promoter in the sense or antisense orientation.

The construction of polynucleotide constructs generally requires the use of vectors able to replicate in bacteria. A plethora of kits are commercially available for the purification of plasmids from bacteria. For their proper use, follow the manufacturer's instructions (see, for example, EasyPrep™, FlexiPrep™, both from Pharmacia Biotech; StrataClean™, from Stratagene; and, QIAexpress™ Expression System, Qiagen). The isolated and purified plasmids can then be further manipulated to produce other plasmids, used to transfect cells or incorporated into *Agrobacterium* tumefaciens to infect and transform plants. Where *Agrobacterium* is the means of transformation, shuttle vectors are constructed.

However, the skilled worker is aware that he may also obtain the DNA construct according to the invention in other ways. Thus, the host organism may already comprise one or more of the essential components of a DNA construct. A DNA construct is then generated by introducing one further, or more, essential components of said DNA construct in the correct position relative to the existing components in said organism. Thus, for example, the starting organism may already comprise one of the homology sequences (e.g., A or A'). If the organism already comprises a homology sequence A, introducing a DNA construct comprising all other elements (beside A) into the genomic DNA in proximity to the already existing homology sequence A gives rise to a DNA construct according to the invention.

Furthermore, the skilled worker is familiar with various ways in which the DNA construct according to the invention may be introduced into the genome of a host cell or organism. In this context, the insertion may be directed (i.e. taking place at a defined insertion site) or undirected (i.e. taking place randomly). Suitable techniques are known to the skilled worker and described by way of example herein below.

8.2 Methods for Introducing Constructs into Target Cells

A DNA construct employed in the invention may advantageously be introduced into cells using vectors into which said DNA construct is inserted. Examples of vectors may be plasmids, cosmids, phages, viruses, retroviruses or agrobacteria. In an advantageous embodiment, the expression cassette is introduced by means of plasmid vectors. Preferred vectors are those which enable the stable integration of the expression cassette into the host genome.

A DNA construct can be introduced into the target plant cells and/or organisms by any of the several means known to those of skill in the art, a procedure which is termed transformation (see also Keown *et al.* (1990) Meth Enzymol 185:527-537). For instance, the DNA constructs can be introduced into cells, either in culture or in the organs of a plant by a variety of conventional techniques. For example, the DNA constructs can be introduced directly to plant cells using ballistic methods, such as DNA particle bombardment, or the DNA construct can be introduced using techniques such as electroporation and microinjection of cell. Particle-mediated transformation techniques (also known as "biolistics") are described in, e.g., Klein *et al.* (1987) Nature 327:70-73; Vasil V *et al.* (1993) BiolTechnol 11:1553-1558; and Becker D *et al.* (1994) Plant J 5:299-307. These methods involve penetration of cells by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface. The biolistic PDS-1000 Gene Gun (Biorad, Hercules, CA) uses helium pressure to accelerate DNA-coated gold or tungsten microcarriers toward target cells. The process is applicable to a wide range of tissues and cells from organisms, including plants. Other transformation methods are also known to those of skill in the art.

Microinjection techniques are known in the art and are well described in the scientific and patent literature. Also, the cell can be permeabilized chemically, for example using polyethylene glycol, so that the DNA can enter the cell by diffusion. The DNA can also be introduced by protoplast fusion with other DNA-containing units such as minicells, cells, lysosomes or liposomes. The introduction of DNA constructs using polyethylene glycol (PEG) precipitation is described in Paszkowski *et al.* (1984) EMBO J 3:2717. Liposome-based gene delivery is e.g., described in WO 93/24640; Mannino and Gould-Fogerite (1988) BioTechniques 6(7):682-691; US 5,279,833; WO 91/06309; and Felgner *et al.* (1987) Proc Natl Acad Sci USA 84:7413-7414).

Another suitable method of introducing DNA is electroporation, where the cells are permeabilized reversibly by an electrical pulse. Electroporation techniques are described in Fromm *et al.* (1985) Proc Natl Acad Sci USA 82:5824. PEG-mediated transformation and electroporation of plant protoplasts are also discussed in Lazzeri P (1995) Methods Mol Biol 49:95-106. Preferred general methods which may be mentioned are the calcium-phosphate-mediated transfection, the DEAE-dextran-mediated transfection, the cationic lipid-mediated transfection, electroporation, transduction and

infection. Such methods are known to the skilled worker and described, for example, in Davis *et al.*, Basic Methods In Molecular Biology (1986). For a review of gene transfer methods for plant and cell cultures, see, Fisk *et al.* (1993) Scientia Horticulturae 55:5-36 and Potrykus (1990) CIBA Found Symp 154:198.

Methods are known for introduction and expression of heterologous genes in both monocot and dicot plants. See, e.g., US 5,633,446, US 5,317,096, US 5,689,052, US 5,159,135, and US 5,679,558; Weising *et al.* (1988) Ann. Rev. Genet. 22: 421-477. Transformation of monocots in particular can use various techniques including electroporation (e.g., Shimamoto *et al.* (1992) Nature 338:274-276; biolistics (e.g., EP-A1 270,356); and *Agrobacterium* (e.g., Bytebier *et al.* (1987) Proc Natl Acad Sci USA 84:5345-5349).

In plants, methods for transforming and regenerating plants from plant tissues or plant cells with which the skilled worker is familiar are exploited for transient or stable transformation. Suitable methods are especially protoplast transformation by means of polyethylene-glycol-induced DNA uptake, biolistic methods such as the gene gun ("particle bombardment" method), electroporation, the incubation of dry embryos in DNA-containing solution, sonication and microinjection, and the transformation of intact cells or tissues by micro- or macroinjection into tissues or embryos, tissue electroporation, or vacuum infiltration of seeds. In the case of injection or electroporation of DNA into plant cells, the plasmid used does not need to meet any particular requirement. Simple plasmids such as those of the pUC series may be used. If intact plants are to be regenerated from the transformed cells, the presence of an additional selectable marker gene on the plasmid is useful.

In addition to these "direct" transformation techniques, transformation can also be carried out by bacterial infection by means of *Agrobacterium* tumefaciens or *Agrobacterium* rhizogenes. These strains contain a plasmid (Ti or Ri plasmid). Part of this plasmid, termed T-DNA (transferred DNA), is transferred to the plant following Agrobacterium infection and integrated into the genome of the plant cell.

For *Agrobacterium*-mediated transformation of plants, a DNA construct of the invention may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium* tumefaciens host vector. The virulence functions of the A. tumefaciens host will direct the insertion of a transgene and adjacent marker gene(s) (if present) into the plant cell DNA when the cell is infected by the bacteria. *Agrobacterium* tumefaciens-mediated transformation techniques are well described in the scientific literature. See, for example, Horsch *et al.* (1984) Science 233:496-498, Fraley *et al.* (1983) Proc Natl Acad Sci USA 80:4803-4807, Hooykaas (1989) Plant Mol Biol 13:327-336, Horsch RB (1986) Proc Natl Acad Sci USA 83(8):2571-2575), Bevans *et al.* (1983) Nature 304:184-187, Bechtold *et al.* (1993) Comptes Rendus De L'Academie Des Sciences Serie III-Sciences De La Vie-Life Sciences 316:1194-1199, Valvekens *et al.* (1988) Proc Natl Acad Sci USA 85:5536-5540.

A DNA construct of the invention is preferably integrated into specific plasmids, either into a shuttle, or intermediate, vector or into a binary vector). If, for example, a Ti or Ri plasmid is to be used for the transformation, at least the right border, but in most cases

the right and the left border, of the Ti or Ri plasmid T-DNA is linked with the expression cassette to be introduced as a flanking region. Binary vectors are preferably used. Binary vectors are capable of replication both in E. coli and in *Agrobacterium*. As a rule, they contain a selection marker gene and a linker or polylinker flanked by the right or left T-DNA flanking sequence. They can be transformed directly into *Agrobacterium* (Holsters *et al.* (1978) Mol Gen Genet 163:181-187). The selection marker gene permits the selection of transformed agrobacteria and is, for example, the nptII gene, which imparts resistance to kanamycin. The *Agrobacterium*, which acts as host organism in this case, should already contain a plasmid with the vir region. The latter is required for transferring the T-DNA to the plant cell. An *Agrobacterium* thus transformed can be used for transforming plant cells.

Many strains of *Agrobacterium* tumefaciens are capable of transferring genetic material - for example a DNA constructs according to the invention -, such as, for example, the strains EHA101(pEHA101) (Hood EE *et al.* (1996) J Bacteriol 168(3):1291-1301), EHA105(pEHA105) (Hood *et al.* 1993, Transgenic Research 2, 208-218), LBA4404(pAL4404) (Hoekema *et al.* (1983) Nature 303:179-181), C58C1(pMP90) (Koncz and Schell (1986) Mol Gen Genet 204,383-396) and C58C1(pGV2260) (Deblaere *et al.* (1985) Nucl Acids Res. 13, 4777-4788).

The agrobacterial strain employed for the transformation comprises, in addition to its disarmed Ti plasmid, a binary plasmid with the T-DNA to be transferred, which, as a rule, comprises a gene for the selection of the transformed cells and the gene to be transferred. Both genes must be equipped with transcriptional and translational initiation and termination signals. The binary plasmid can be transferred into the agrobacterial strain for example by electroporation or other transformation methods (Mozo & Hooykaas (1991) Plant Mol Biol 16:917-918). Coculture of the plant explants with the agrobacterial strain is usually performed for two to three days.

A variety of vectors could, or can, be used. In principle, one differentiates between those vectors which can be employed for the *Agrobacterium*-mediated transformation or agroinfection, i.e. which comprise a DNA construct of the invention within a T-DNA, which indeed permits stable integration of the T-DNA into the plant genome. Moreover, border-sequence-free vectors may be employed, which can be transformed into the plant cells for example by particle bombardment, where they can lead both to transient and to stable expression.

The use of T-DNA for the transformation of plant cells has been studied and described intensively (EP-A1 120 516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B. V., Alblasserdam, Chapter V; Fraley *et al.* (1985) Crit Rev Plant Sci 4:1-45 and An *et al.* (1985) EMBO J 4:277-287). Various binary vectors are known, some of which are commercially available such as, for example, pBIN19 (Clontech Laboratories, Inc. USA).

To transfer the DNA to the plant cell, plant explants are cocultured with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. Starting from infected plant material (for example leaf, root or stalk sections, but also protoplasts or suspensions of plant cells), intact plants can be regenerated using a suitable medium which may contain, for ex-

ample, antibiotics or biocides for selecting transformed cells. The plants obtained can then be screened for the presence of the DNA introduced, in this case a DNA construct according to the invention. As soon as the DNA has integrated into the host genome, the genotype in question is, as a rule, stable and the insertion in question is also found in the subsequent generations. As a rule, the expression cassette integrated contains a selection marker which confers a resistance to a biocide (for example a herbicide) or an antibiotic such as kanamycin, G 418, bleomycin, hygromycin or phosphinotricin and the like to the transformed plant. The selection marker permits the selection of transformed cells (McCormick *et al.*, Plant Cell Reports 5 (1986), 81-84). The plants obtained can be cultured and hybridized in the customary fashion. Two or more generations should be grown in order to ensure that the genomic integration is stable and hereditary.

The abovementioned methods are described, for example, in B. Jenes *et al.*, Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by SD Kung and R Wu, Academic Press (1993), 128-143 and in Potrykus (1991) Annu Rev Plant Physiol Plant Molec Biol 42:205-225). The construct to be expressed is preferably cloned into a vector which is suitable for the transformation of *Agrobacterium* tumefaciens, for example pBin19 (Bevan *et al.* (1984) Nucl Acids Res 12:8711).

The DNA construct of the invention can be used to confer desired traits on essentially any plant. One of skill will recognize that after DNA construct is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

8.3 Regeneration of Transgenic Plants

Transformed cells, i.e. those which comprise the DNA integrated into the DNA of the host cell, can be selected from untransformed cells if a selectable marker is part of the DNA introduced. A marker can be, for example, any gene which is capable of conferring a resistance to antibiotics or herbicides (for examples see above). Transformed cells which express such a marker gene are capable of surviving in the presence of concentrations of a suitable antibiotic or herbicide which kill an untransformed wild type. As soon as a transformed plant cell has been generated, an intact plant can be obtained using methods known to the skilled worker. For example, callus cultures are used as starting material. The formation of shoot and root can be induced in this as yet undifferentiated cell biomass in the known fashion. The shoots obtained can be planted and cultured.

Transformed plant cells, derived by any of the above transformation techniques, can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker that has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al.*, Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, pp. 124176, Macmillian Publishing Company, New York (1983); and in Binding, Regeneration of Plants, Plant Protoplasts, pp. 21-73, CRC Press, Boca Raton, (1985). Regeneration can also

be obtained from plant callus, explants, somatic embryos (Dandekar *et al.* (1989) J Tissue Cult Meth 12:145; McGranahan *et al.* (1990) Plant Cell Rep 8:512), organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al.* (1987) Ann Rev Plant Physiol 38:467-486.

9. Crossing of the trait plant and the endonuclease master plant and generation of marker-free descendants

After transformation, selection and regeneration of transgenic trait plants and endonuclease master plants, these plants can be further bred (e.g. to homozygous plants by selfing or crossing into elite germplasm). The trait plants and/or endonuclease master plants to be employed in the method of the invention may comprise one or more copies of the respective DNA construct introduced into their genome and may be homozygous or heterozygous with respect to said recombination cassette or endonuclease expression cassette, respectively.

Selfing and/or crossing of the trait plant with the endonuclease marker plant can be done by any procedure known in the art. To reduce the possibility of self-pollination during crossing, the flowers from the female parent may be used before the anthers begin to shed pollen onto the stigma. For the male parent, an open flower that is visibly shedding pollen should be chosen. The appearance of flowers at the appropriate developmental stage varies among plant species, cultivars and growth conditions.

After the fertilization process, F1 seeds are harvested, germinated and grown into mature plants. Plants from which the selectable marker of the "trait construct" is removed from all cells may be obtained in the first generation (F1) or in the second (F2) or later (Fn) generation. Isolation and identification of descendants which underwent an excision process can be done at any stage of plant development. Methods for said identification are well known in the art and may comprise – for example – PCR analysis, Northern blot, Southern blot, or phenotypic screening (e.g., for an negative selection marker; compare examples 8c and 8d).

Descendants of the F1 plants may be obtained by sexual propagation as outlined above. They may also be obtained by asexual propagation. For the latter tissue culture procedures may be applied. Asexual propagation is especially preferred in cases, where plants are not yet homogenously consisting of cells which all have undergone successful sequence excision. For asexual propagation tissues from plants (which may be chimeric for the recombination event, i.e. the excision did not take place in all cells) are used as explants to regenerate new plants. Since these new plants may be regenerated from a single cell, all cells of this asexually obtained descendant are identical regarding the excision event to the original cell. By selecting for plants which originate from a cell, in which the desired recombination event (e.g. excision of a selectable marker) occurred, plants may be obtained which have the recombination event present in all cells. A very efficient method to regenerate whole plants from single cells may be applied by making a slurry of the explant (see example 8c where a detailed method for efficient regeneration of rapeseed plants is disclosed as an example).

Descendants may comprise the construct encoding the sequence specific endonuclease (optionally together with a selectable marker on the same construct). These cas-

settes are preferably removed by segregation in the progeny (sexual propagation). This might, for example, be achieved by selfing or crossing to a non-transgenic wildtype plant.

Descendants may comprise one or more copies of the agronomically valuable trait gene. Preferably, descendants are isolated which only comprise one copy of said trait gene. It is another inventive feature of the present invention that multiple insertion (e.g., of a T-DNA) in one genomic location will be reduced to a single insertion event by excision of the redundant copies together with the marker gene of the remaining copy (Fig. 10).

In a preferred embodiment the transgenic plant made by the process of the invention is marker-free. The terms "marker-free" or "selection marker free" as used herein with respect to a cell or an organisms are intended to mean a cell or an organism which is not able to express a functional selection marker protein (encoded by expression cassette b1; as defined above) which was inserted into said cell or organism in combination with the gene encoding for the agronomically valuable trait. The sequence encoding said selection marker protein may be absent in part or –preferably – entirely. Furthermore the promoter operably linked thereto may be dysfunctional by being absent in part or entirely.

The resulting plant may however comprise other sequences which may function as a selection marker. For example the plant may comprise as a agronomically valuable trait a herbicide resistance conferring gene or the endonuclease expression cassette may be linked to a selection marker. However, it is most preferred that the resulting plant does not comprise any selection marker.

10. Combination with other recombination enhancing techniques

In a further preferred embodiment, the efficacy of the recombination system is increased by combination with systems which promote homologous recombination. Such systems are described and encompass, for example, the expression of proteins such as RecA or the treatment with PARP inhibitors. It has been demonstrated that the intrachromosomal homologous recombination in tobacco plants can be increased by using PARP inhibitors (Puchta H *et al.* (1995) Plant J. 7:203-210). Using these inhibitors, the homologous recombination rate in the recombination cassette after induction of the sequence-specific DNA double-strand break, and thus the efficacy of the deletion of the transgene sequences, can be increased further. Various PARP inhibitors may be employed for this purpose. Preferably encompassed are inhibitors such as 3-aminobenzamide, 8-hydroxy-2-methylquinazolin-4-one (NU1025), 1,11b-dihydro-(2H)benzopyrano(4,3,2-de)isoquinolin-3-one (GPI 6150), 5-aminoisoquino-linone, 3,4-dihydro-5-(4-(1-piperidinyl)butoxy)-1(2H)-isoquinolinone, or the compounds described in WO 00/26192, WO 00/29384, WO 00/32579, WO 00/64878, WO 00/68206, WO 00/67734, WO 01/23386 and WO 01/23390.

In addition, it was possible to increase the frequency of various homologous recombination reactions in plants by expressing the E. coli RecA gene (Reiss B *et al.* (1996) Proc Natl Acad Sci USA 93(7):3094-3098). Also, the presence of the protein shifts the ratio between homologous and illegitimate DSB repair in favor of homologous repair

(Reiss B *et al.* (2000) Proc Natl Acad Sci USA 97(7):3358-3363). Reference may also be made to the methods described in WO 97/08331 for increasing the homologous recombination in plants. A further increase in the efficacy of the recombination system might be achieved by the simultaneous expression of the RecA gene or other genes which increase the homologous recombination efficacy (Shalev G *et al.* (1999) Proc Natl Acad Sci USA 96(13):7398-402). The above-stated systems for promoting homologous recombination can also be advantageously employed in cases where the recombination construct is to be introduced in a site-directed fashion into the genome of a eukaryotic organism by means of homologous recombination.

11. Preferred combinations

11.1 Basic Principle (see Fig. 1 and 2)

In a preferred embodiment the DNA construct in the first plant (the trait plant, TP) comprises an expression cassette for a negative selections marker (ENS) under control of a promoter and a transcription terminator. In addition, the DNA construct comprises two (or more) recognition sequences (S1 and S2) flanking the selection marker expression cassette in a way that cleavage at this two recognition sequences excises said cassette. These two recognition sequences are flanked by the two homology sequences A and A'. Preferably, the distance between a homology sequences (A and A', respectively) and a recognition sequence (e.g., S1 or S2) is less than 100 base pairs, preferably less than 50 base pairs, more preferably less than 25 base pairs, most preferably the recognition sequence is attached directly to the end of the homology sequence. Furthermore the DNA construct in the second plant (the endonuclease master plant, EMP) comprises a second expression cassette for a sequence specific endonuclease (EE) under control of a parsley ubiquitin promoter and a transcription terminator.

After crossing (X) of the two plants (symbolized by the boxes), the resulting descendant is grown (G) and optionally further propagated into following generation(s). Cleavage (C) at the recognition sequences (S1 and S2) is induced when the parsley ubiquitin promoter EP becomes active and causes expression of the sequence specific endonuclease. By generation of the double-strand breaks homologous recombination (HR) is induced between the homology sequences A and A'. The endonuclease may be separated from the trait by further crossing and segregation (S).

In an preferred embodiment the DNA construct in the trait plant comprises a second expression cassette (e.g., encoding an agronomically valuable trait) outside of the region flanked by the homology sequences, which is therefore not excised from the genome by the homologous recombination reaction (see. Fig. 2). Furthermore the DNA construct may comprise a fourth expression cassette for a counter-selection marker, preferably localized between the homology sequences A and A' (Fig. 9).

11.2 Variations where the homology sequences are part of the expression cassettes (see Fig. 3 to 5)

In another preferred embodiment the homology sequences (A and A') are part of the expression cassettes of the DNA construct in the trait plant. Any part of the expression cassettes may be suitable to function as a homology sequence. Preferably, the homology sequences are identical with the promoter regions (Fig. 3) or the terminator region (Fig. 4). Homologous recombination between these functional elements preferably re-

constitutes the expression cassette for the agronomically valuable trait but excises the expression cassettes for the selection marker. It is possible that the homology region does not only comprises promoter or terminator regions but extents into the coding region (e.g., of the gene encoding the agronomically valuable trait (Fig. 5).

11.3 DNA constructs comprising only one recognition / cleavage site (Fig. 6 - 8)

The DNA construct of the invention may comprise only one recognition / cleavage sequence for the sequence specific endonuclease. Preferably this sequence is localized close to one of the homology sequences and is within the region flanked by said homology sequences. Various locations are possible (Fig. 6-8).

12. Sequences

1. SEQ ID NO: 1 Nucleic acid sequence coding for Construct I

Features of the T-DNA:

Position 6 to 151		right border
Position 152 to161		N, region encoding different expression cassettes for the sequence specific DNA-endonuclease
Position 1681 to 198	(compl.)	nosP::nptII::nosT cassette (with nptII ORF from 1324 to 546, compl.)
Position 1694 to 1908		left border

Rest of the plasmid is pSUN backbone (including aadA bacterial resistance gene encoded at position 7279 to 6488, complementary)

2. SEQ ID NO: 2 Nucleic acid sequence coding for insert of pCB603-100

Features

Position 249 to 43	(compl.)	ocs terminator
Position 1010 to 303	(compl.)	ORF encoding I-SceI
Position 2790 to 1014	(compl.)	fragment of promoter AtSERK1 (incl.5'-UTR)

3. SEQ ID NO: 3 Nucleic acid sequence coding for insert of pCB622-3

Features

Position 249 to 43	(compl.)	ocs terminator
Position 1011 to 304	(compl.)	ORF encoding I-SceI
Position 2168 to 1014	(compl.)	fragment of promoter Atcyc1 (incl.5'-UTR)

4. SEQ ID NO: 4 Nucleic acid sequence coding for insert of pCB653-37

Features

Position 249 to 43	(compl.)	ocs terminator
Position 1011 to 304	(compl.)	ORF encoding I-SceI
Position 2201 to 1013	(compl.)	fragment of promoter erecta (incl. 5'-UTR)

5. SEQ ID NO: 5 Nucleic acid sequence coding for insert of pCB652-124

Features

Position 249 to 43	(compl.)	ocs terminator
Position 1011 to 304	(compl.)	ORF encoding I-SceI
Position 2956 to 1059	(compl.)	fragment of promoter invGF (incl. 5'-UTR)

6. SEQ ID NO: 6 Nucleic acid sequence coding for insert of pCB632-17

Features		
Position 90 to 1407		STPT promoter
Position 1491 to 2198		ORF encoding I-SceI
Position 2282 to 2516		CatDpA terminator

7. SEQ ID NO: 7 Amino acid sequence coding for I-SceI

8. SEQ ID NO: 8 Nucleic acid sequence coding for parsley (*Petroselinum crispum*) ubiquitin promoter

9. SEQ ID NO: 9 Nucleic acid sequence coding for binary vector pCB666-3

Features of T-DNA		
Position 3636 to 3850		Left Border
Position 3630 to 2313	(compl)	STPT promoter from *Arabidopsis*
Position 2289 to 277	(compl)	AHAS resistance gene from *Arabidopsis;* encodes the S653N mutation conferring resistance towards imidazoline herbicides
Position 260 to 8	(compl)	nos terminator
Position 12807 to 11606	(compl)	sequence derived from *Arabidopsis* downstream of AHAS coding region
Position 11582 to 10600	(compl)	Parsley ubiquitin promoter / 5'UTR with intron
Position 10582 to 9875	(compl)	sequence encoding I-SceI
Position 9599 to 9851	(compl)	nos terminator
Position 9395 to 9540		Right Border

10. SEQ ID NO: 10 Nucleic acid sequence coding for binary vector pCB657-41

Features of T-DNA		
Position 6976to 7192		Left Border
Position 6904 to 5587	(compl)	STPT promoter from *Arabidopsis*
Position 5533 to 3536	(compl)	GUS gene; PIV2 intron at 5148 to 4960 (compl)
Position 3461 to 3257	(compl)	35S terminator
Position 3182 to 3199		I-SceI site
Position 3200 to 3229		I-CreI site
Position 3167 to 1356	(compl)	A. thaliana nitrilase 1 promoter
Position 1312 to 509	(compl)	nptII gene
Position 462 to 445	(compl)	I-SceI site
Position 444 to 415	(compl)	I-CreI site
Position 38 to 183		Right Border

11. SEQ ID NO: 11 Nucleic acid sequence coding for binary vector JB010qcz

Features of T-DNA		
Position 363 to 149	(compl)	Left Border
Position 372 to 6088		AHAS expression cassette[1]
Position 6106 to 7088		PcUbi promoter (including an intron in the 5'UTR)
Position 7106 to 7813		I-SceI coding region
Position 7837 to 8089		nos terminator
Position 8148 to 8293		Right Border

[1]The AHAS ORF (position 2855 to 4867) from *Arabidopsis* encodes the S653N mutation conferring resistance towards imidazoline herbicides

12. SEQ ID NO:12 Nucleic acid sequence coding for binary vector pCB583-40

Features of T-DNA		
Position 8 to 153		Right Border
Position 419 to 244	(compl)	nos terminator
Position 3144 to 525	(compl)	overlapping, non-functional helves of GUS gene with an I-SceI and I-CreI site inbetween
Position 3620 to 3251	(compl)	35S promoter
Position 4041 to 3786	(compl)	nos terminator
Position 4663 to 4112	(compl)	pat gene conferring BASTA herbicide resistance
Position 4999 to 4669	(compl)	nos promoter
Position 5015 to 5228		Left Border

13. SEQ ID NO: 13 Nucleic acid sequence coding for binary vector pRS8

Features of T-DNA		
Position 8 to 153		Right Border
Position 490 to 252	(compl)	ocs terminator
Position 1208 to 501		I-SceI CDS
Position 1809 to 1280	(compl)	35S promoter
Position 2111 to 1856	(compl)	nos terminator
Position 2982 to 2204	(compl)	nptII gene conferring kanamycin resistance
Position 3339 to 3003	(compl)	nos promoter
Position 3566 to 3352		Left Border

14. SEQ ID NO: 14 I-SceI coding region interrupted by an intron

I-SceI coding sequence comprising the potato PIV2 intron at position 565 to 753

15. SEQ ID NO: 15 Nucleic acid sequence coding for parsley (*Petroselinum crispum*) ubiquitin promoter (alternative form)

13. Figures

The following abbreviations apply to the figures in general:

A:	Homology sequence A
A':	Homology sequence A'
A/A':	Sequence as the result of homologous recombination between A and A'
C:	Sequence specific cleavage
CS:	Counter selection marker
E:	Sequence encoding sequence specific DNA-endonuclease
EE:	Complete expression cassette for endonuclease
EN:	Complete expression cassette for further nucleic acid sequence (coding for e.g., agronomically valuable trait)
PP:	Parsley ubiquitin promoter
ENS:	Complete expression cassette for negative selection marker
I:	Insertion into the genome (e.g., chromosomal DNA)
G	Growing of plants and – optionally – generation of subsequent generation
HR:	Homologous recombination
N:	Further nucleic acid sequence (coding for e.g., agronomically valuable trait)
NS:	Negative selection marker
NU:	Endonuclease
P_n:	Promoter
S_n:	Recognition sequence for the site-directed induction of DNA double-strand breaks (e.g., S1: First recognition sequence). The recognition sequences may be different (e.g., functioning for different endonucleases) or –preferably - identical (but only placed in different locations).
S_n*:	Part of recognition sequence S_n remaining after cleavage
T_n:	Terminator sequence
RB/LB:	Right/left border of *Agrobacterium* T-DNA
X:	Crossing of plants
S:	Segregation by crossing or selfing, may be monitored by e.g. PCR or Southern

Fig. 1: Basic Principle

The boxes represent the individual plants (EMP: Endonuclease master plant; TP trait plant). The DNA construct in the trait plant comprises:

- An expression cassette for a negative selection marker (ENS) under control of a first promoter and a transcription terminator,
- Two recognition sequences (S1 and S2) for the sequence specific endonuclease expressed by the second expression cassette, flanking the two expression cassettes in a way that cleavage at this two recognition sequences excises said cassettes, and
- Two homology sequences A and A' flanking the two recognition sequences.

The DNA construct in the endonuclease master plant comprises:

- An expression cassette for a sequence specific endonuclease (NU) under control of an parsley ubiquitin promoter and a transcription terminator.

After crossing (X) of the two plants (symbolized by the boxes), the resulting descendants are grown (G) and optionally further propagated into following generation(s). Cleavage (C) at the recognition sequences (S1 and S2) is induced when the parsley ubiquitin promoter (PP) becomes active and causes expression of the sequence specific endonuclease (NU). By generation of the double-strand breaks, homologous recombination (HR) is induced between the homology sequences A and A'. The endonuclease may be separated from the trait by further crossing and segregation (S).

Fig. 2: Introduction of an agronomically valuable trait

The DNA construct in the trait plant further comprises a second expression cassette (e.g., encoding an agronomically valuable trait (EN) under control of a promoter and a terminator) outside of the domain flanked by the homology sequences. (In this case the DNA construct was introduced into the chromosomal DNA by *Agrobacterium* mediated transformation. Therefore the inserted elements are flanked by right (RB) and left border (LB) of *Agrobacterium* T-DNA). Crossing (X) , growing (G) occurs as described above. Cleavage (C) and the subsequent homologous recombination (HR) excises the expression cassettes for the selection marker. However, the expression cassette for the agronomically valuable trait is not excised but remains in the chromosomal DNA. The endonuclease may be separated from the trait by further crossing and segregation (S).

Fig. 3 Use of promoter sequences as homology sequences

(Only the cleavage and homologous recombination part of the method are shown)

The homology sequences (A and A') are the promoters of the expression cassettes of the DNA construct (P1=A; P1=A'). Cleavage (C) and subsequent homologous recombination (HR) between these promoters reconstitutes the expression cassette for the agronomically valuable trait (N) but excises the expression cassettes for the negative selection marker (NS). In the present example DNA introduction was realized by *Agrobacterium* transformation and the inserted sequence is flanked by *Agrobacterium* left/right borders (other ways of introduction e.g., by particle bombardment are possible and would not require these borders).

Fig. 4 Use of terminator sequences as homology sequences

(Only the cleavage and homologous recombination part of the method are shown)

The homology sequences (A and A') are the terminators of the expression cassettes of the DNA construct (T1=A; T1=A'). Cleavage (C) and subsequent homologous recombination (HR) between these terminators reconstitutes the expression cassette for the agronomically valuable trait (N) but excises the expression cassettes for the negative selection marker (NS). In the present example DNA introduction was realized by *Agrobacterium* transformation and the inserted sequence is flanked by *Agrobacterium* left/right borders (other ways of introduction e.g., by particle bombardment are possible and would not require these borders).

Fig. 5 Use of part of the excision cassettes as homology sequences

(Only the cleavage and homologous recombination part of the method are shown)

The homology sequences (A and A') are the part of the excision cassettes of the DNA construct (indicated by black bars below; A; A'). Cleavage (C) and subsequent homologous recombination (HR) between these terminators reconstitutes the expression cassette for the agronomically valuable trait (N) but excises the expression cassettes for the negative selection marker (NS). In the present example DNA introduction was realized by *Agrobacterium* transformation and the inserted sequence is flanked by *Agrobacterium* left/right borders (other ways of introduction e.g., by particle bombardment are possible and would not require these borders).

Fig.6-8 DNA constructs with one recognition sequence
(Only the cleavage and homologous recombination part of the method are shown)

The DNA construct of the invention may comprise only one recognition / cleavage sequence for the sequence specific endonuclease. Preferably this sequence is localized close to one of the homology sequences and is within the region flanked by said homology sequences. Various locations are possible (Fig. 6-8).

Fig. 9: DNA construct comprising a counter-selection marker
(Only the cleavage and homologous recombination part of the method are shown)

The DNA construct in the trait plant may comprise:

- A first expression cassette for a negative selection marker (NS) under control of a promoter (P1) and a transcription terminator (T1),
- A second expression cassette for a counter-selection marker (CS) under control of a promoter (P2) and a transcription terminator (T2),
- Two recognition sequences (S1 and S2) for the sequence-specific endonuclease encoded by the second expression cassette, flanking the two expression cassettes in a way that cleavage at this two recognition sequences excises said cassettes,
- Two homology sequences A and A' flanking the two recognition sequences, and
- A third expression cassette for an agronomically valuable trait (localized outside of the region flanked by the homology sequences A and A') under control of a promoter (P3) and a transcription terminator (T4)

After crossing with the endonuclease master plant (as described for Fig. 1) and growing of the descendants, cleavage (C) at the recognition sequences (S1 and S2) is induced when the parsley ubiquitin promoter becomes active and causes expression of the sequence specific endonuclease. By generation of the double-strand breaks homologous recombination (HR) is induced between the homology sequences A and A'.

Fig. 10 Application of the method of the invention to simplify transformation events.
(Only the cleavage and homologous recombination part of the method are shown)

It is another inventive feature of the present invention that multiple insertion (e.g., of a T-DNA) in one genomic location will be reduced to a single insertion event by excision of the redundant copies. In the depicted case two copies of a T-DNA have inserted into the genome (box 1 and box 2). Both comprise the expression cassette for a negative selection marker (ENS) and an expression cassette for an agronomically valuable trait (EN). Cleavage and subsequent homologous recombination deletes the marker and the oblivious copy and the resulting event becomes undistinguishable from a single insertion event of which the selectable marker has been eliminated by homologous recombination (see Fig. 2 in comparison).

Fig. 11-12: Use of endogenous promoters

One or more expression cassette of the DNA construct may constituted after insertion into the genome by inserting the nucleic acid to be expressed (e.g., the negative selection marker (Fig. 11) or the gene encoding the agronomically valuable trait (Fig. 12)) under control of an endogenous promoter.

Fig. 13: Schematic representation of the pGU-sce-US construct before (A) and after (B) intrachromosomal homologous recombination between the identical homology sequences A and A'. The original construct (A) did not produce active GUS protein, while the arrangement obtained after intrachromosomal homologous recombination (hereinafter called ICHR) (B) restored an intact GUS gene and therefore produced active GUS protein. The cells / plants in which ICHR had taken place were identified by histochemical GUS staining.

Fig. 14: Leaves from plants (A / B) containing the pGU-Sce-US reporter construct. The only difference between these two plants is that plant (B) contained in addition an I-SceI expression cassette (pRS8).

Fig. 15-26: Vector maps

Vector backbone elements:

aadA:	prokaryotic selectable marker conferring spectinomycin resistance
ColE1:	origin of replication, e.g. for E. coli
repA/pVS1:	elements for replication, e.g. in *Agrobacterium*

T-DNA elements:

LB:	Left border
RB:	Right border
sTPT:	constitutive plant promoter
GUS(int):	uidA gene encoding GUS with an intron
GUS:	uidA gene encoding GUS
35SpA:	terminator sequence derived from CaMV 35S RNA encoding gene (duplicated, serves as homology region A and A', respectively)
Nit1-P:	constitutive plant promoter
nptII:	eukaryotic selectable marker conferring kanamycin resistance
nosT:	terminator sequence derived from nos gene from *Agrobacterium*
USP-P:	USP promoter active in immature embyos
invGF promoter:	invGF promoter active in pollen
Perecta	Erecta promoter
Prom AtSERK1	AtSERK1 promoterAtcyc1A Prom Atcyc1A promoter
I-SceI:	gene encoding the homing endonuclease I-SceI
I-SceI RC	recognition sequence for I-SceI homing endonuclease
I-CreI RC	recognition sequence for I-SceI homing endonuclease
I-CreI Exon1/2:	Two artifical exons of the I-CreI gene
T-CatD	terminator sequence
35SpA	terminator sequence
nosT	terminator sequence
NNNNNNNNNN	region encoding different expression cassettes for the sequence specific DNA-endonuclease (the number of Ns is only symbolic, the insert at this place can have any length)

Fig. 27 Seedling and leafs from seedlings obtained from crosses between GU-US reporter lines (construct pCB583-40) and sTPT::I-SceI (construct pCB632-17; based on the constitutive sTPT promoter; panel "A") and PcUbi::I-SceI (construct JB010cqz, based on the parsley ubiquitin promoter; panel "B"), respectively, after histochemical GUS staining. A significant more intense staining (indicated by dark areas) can be observed for the parsley based promoter construct in panel "B" indicating that more cells harbor the recombination event, i.e. the functional GUS gene in this particular example.

Fig. 28 Schematic representation of the procedure to obtain full blue plants (equivalent to full marker-free plants) from plants with blue spots (chimeric plants)

Fig. 29 Regenerants obtained from cross C18-5. Plants 4 and 6 are white, while plant 10 is totally blue. The rest of the plants have many blue spots, in some cases very small.

Fig. 30-I PCR analysis of some of the plants regenerated from cross C18-5. Expected results in white ("A"), blue ("B"), and spotted plants ("C").

Fig. 30-II PCR analysis of some of the plants regenerated from cross C18-5. PCR results of six plants, before (Panel "A") and after (Panel "B") digestion of the fragment with I-SceI

Fig. 31-I Southern blot of several the plants regenerated from cross C18-5, hybridized with a GUS probe. Schematic drawing for expected results with and without ICHR.
Upper Panel ("A"); genomic DNA digested with *Eco*RI + *Not*I.
Lower Panel ("B"): genomic DNA digested with *Bam*HI + *Pvu*I.

Fig. 31-II Southern blot of several the plants regenerated from cross C18-5, hybridized with a GUS probe. Southern results of six regenerated plants and a wild type control.
Left(Panel A): genomic DNA digested with *Eco*RI + *Not*I.
Right (Panel B): genomic DNA digested with *Bam*HI + *Pvu*I.

Fig. 32 Northern blot analysis of some of the plants regenerated from cross C18-5.
A: Ethidium bromide stained agarose gel to be blotted (10 µg total RNA)
B: Northern blot, probed with GUS probe
Note that due to the recombination event, the transcript produced by plant number 10 is shorter than the transcript produced by any of the other plants

Fig. 33: A: Principle of Southern hybridisation to distinguish between recombination cassette before (A-1) and after (A-2) homologous recombination (HR) occurred.

B: Southern blot of leaf material from F2 plants originating from crosses between *Arabidopsis* plants harbouring a single copy of the T-DNA from the construct to monitor ICHR (pCB583-40) and *Arabidopsis* plants harbouring PcUbi::I-SceI (JB010cqz). DNA was extracted from plant leaves, digested with SacI, separated on an agarose gel, transferred onto nylon membrane and hybridised with a radioactive labelled GUS probe. The result was analysed with the help of a phosphoimager.

Fig 34. Leaves from plants (A / B) after histochemical GUS staining, which originated from crosses of I-SceI expressing plants and plants containing the pGU-Sce-US reporter construct. The only difference between these two plants is that in plant B I-SceI expression is under control of PcUbi promoter, while I-SceI in plant A is under control of 35S promoter.

Examples

General methods:

The chemical synthesis of oligonucleotides can be effected for example in the known manner using the phosphoamidite method (Voet, Voet, 2nd edition, Wiley Press New York, pages 896-897). The cloning steps carried out for the purposes of the present invention, such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, the transfer of nucleic acids to nitrocellulose and nylon membranes, the linkage of DNA fragments, the transformation of E. coli cells, bacterial cultures, the propagation of phages and the sequence analysis of recombinant DNA are carried out as described by Sambrook *et al.* (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6. Recombinant DNA molecules were sequenced using an ALF Express laser fluorescence DNA sequencer (Pharmacia, Upsala, Sweden) following the method of Sanger (Sanger *et al.*, Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

Example 1: Plant transformation

Example 1a: Transformation of *Arabidopsis thaliana*

A. thaliana plants were grown in soil until they flowered. *Agrobacterium tumefaciens* (strain C58C1 (pMP90)) transformed with the construct of interest was grown in 500 mL in liquid YEB medium (5 g/L Beef extract, 1 g/L Yeast Extract (Duchefa), 5 g/L Peptone (Duchefa), 5 g/L sucrose (Duchefa), 0,49 g/L $MgSO_4$ (Merck)) until the culture reached an OD_{600} 0.8-1.0. The bacterial cells were harvested by centrifugation (15 minutes, 5,000 rpm) and resuspended in 500 mL infiltration solution (5% sucrose, 0.05% SILWET L-77 (distributed by Lehle seeds, Cat.No. VIS-02)). Flowering plants were dipped for 10-20 seconds into the *Agrobacterium* solution. Afterwards the plants were kept in the greenhouse until seeds could be harvested. Transgenic seeds were selected by plating surface sterilized seeds on growth medium A (4.4g/L MS salts (Sigma-Aldrich), 0.5g/L MES (Duchefa); 8g/L Plant Agar (Duchefa)) supplemented with 50 mg/L kanamycin for plants carrying the nptII resistance marker, 100 nM Bimazethapyr for plants carrying a mutated AHAS gene and 10 mg/L Phosphinotricin for plants carrying the pat gene, respectively. Surviving plants were transferred to soil and grown in the greenhouse.

Example 1b: *Agrobacterium*-mediated transformation of *Brassica napus*

Agrobacterium tumefaciens strain GV3101 transformed with the plasmid of interest was grown in 50 mL YEB medium (see Example 4a) at 28°C overnight. The *Agrobacterium* solution is mixed with liquid co-cultivation medium (double concentrated MSB5 salts (Duchefa), 30 g/L sucrose (Duchefa), 3.75 mg/l BAP (6-benzylamino purine, Duchefa), 0.5 g/l MES (Duchefa), 0.5 mg/l GA3 (Gibberellic Acid, Duchefa); pH5.2) until OD_{600} of 0.5 is reached. Petiols of 4 days old seedlings of *Brassica napus* cv. Westar grown on growth medium B (MSB5 salts (Duchefa), 3% sucrose (Duchefa), 0.8% oxoidagar (Oxoid GmbH); pH 5,8) are cut. Petiols are dipped for 2-3 seconds in the *Agrobacterium* solution and afterwards put into solid medium for co-cultivation (co-cultivation medium supplemented with 1.6% Oxoidagar). The co-cultivation lasts 3 days (at 24°C and ~50 µMol/m^2s light intensity). Afterwards petiols are transferred to co-cultivation medium supplemented with 18 mg/L kanamycin (Duchefa) and 300 mg/L Timetin (Duchefa) and incubated for four weeks at 24°C. This step is repeated until shoots appear. Shoots are transferred to A6 medium (MS salts (Sigma Aldrich), 20 g/L sucrose,

100 mg/L myo-inositol (Duchefa), 40 mg/L adenine sulfate (Sigma Aldrich), 500 mg/L MES, 0.0025 mg/L BAP (Sigma), 5 g/L oxoidagar (Oxoid GmbH), 150 mg/L timetin (Duchefa), 15 mg/L kanamycin (Sigma), 0.1 mg/L IBA (indol butyric acid, Duchefa); pH 5,8) until they elongated. Elongated shoots are cultivated in A7 medium (A6 medium without BAP) for rooting. Rooted plants are transferred to soil and grown in the greenhouse.

Example 2: Constructs harbouring sequence specific DNA-endonuclease expression cassettes

Example 2a: Basic construct

In this example we present the general outline of a binary vector, named "Construct I" suitable for plant transformation. This general outline of the binary vector comprises a T-DNA with a nos-promoter::nptII::nos-terminator cassette, which confers kanamycin resistance when integrated into the plant genome. SEQ ID NO: 1 shows a sequence stretch of "NNNNNNNNNN". This is meant to be a placeholder for different expression cassettes for the sequence specific DNA-endonuclease. The sequence of the latter is given in the following examples.

Example 2b: Comparison Constructs

2b.1 Ovule primordia, egg cell, zygote and early embryo-specific promoter fused to endonuclease I-SceI

In Plant Phys. 127: 803-816 (2001) Hecht *et al.* described the promoter of the *Arabidopsis* gene AtSERK1 (Somatic Embryogenesis Receptor-Like Kinase 1; At1G71830). The respective promoter fragment was fused to I-SceI. The resulting plasmid was called pCB603-100. The sequence of the construct is identical to the sequence of construct I, whereas the sequence "NNNNNNNNNN" was replaced by the sequence described by SEQ ID NO: 2.

2b.2 Zygote, early embryo and meristematic cells specific promoter fused to endonuclease I-SceI

In Plant Cell 6: 1763-1774 (1994) Ferreira *et al.* described the promoter of the *Arabidopsis* gene Atcyc1A (Cyclin cyc1 gene, type cyclin B; At4g37490). The respective promoter fragment was fused to I-SceI. The resulting plasmid was called pCB622-3. The sequence of the construct is identical to the sequence of construct I, whereas the sequence "NNNNNNNNNN" was replaced by the sequence described by SEQ ID NO: 3.

2b.3 Shoot and flower meristems specific promoter fused to endonuclease I-SceI

The promoter of gene erecta (Acc. No. D83257) from *Arabidopsis* was described to be active in meristematic cells (Yokoyama *et al.*, 1998, Plant J. 15: 301-310). The erecta promoter was fused to I-SceI resulting in plasmid pCB653-37. The sequence of the construct is identical to the sequence of construct I, whereas the sequence "NNNNNNNNNN" was replaced by the sequence described by SEQ ID NO: 4.

2b.4 Pollen-specific promoter fused to endonuclease I-SceI

The promoter of the potato invGF gene is active in pollen of potato (Plant Mol. Biol., 1999, 41: 741-751; EMBL Acc No. AJ133765) and rapeseed. The respective promoter

fragment has fused to I-SceI. The resulting plasmid was called pCB652-124. The sequence of the construct is identical to the sequence of construct I, whereas the sequence "NNNNNNNNNN" is replaced by the sequence described by SEQ ID NO: 5.

2b.5 Constitutive sTPT promoter fused to endonuclease I-SceI

The sTPT promoter from *Arabidopsis* (i.e. TPT promoter truncated version, WO 03/006660; SEQ ID NO: 27 cited therein) is comparable to the well known 35S promoter. The sTPT promoter was fused to I-SceI. The sequence of the construct pCB632-17 is identical to the sequence of construct I, whereas the sequence "NNNNNNNNNN" is replaced by the sequence described by SEQ ID NO: 6.

2b.6 Constitutive CaMV 35S promoter fused to the endonuclease I-SceI

The 35S promoter from Cauliflower mosaic virus (CaMV) is a strong, well known promoter. The 35S promoter was fused to I-SceI. The resulting construct was named pRS8 (SEQ ID NO: 13).

Example 2c: Parsley promoter fused to endonuclease I-SceI

The PcUbi promoter from parsley (WO 03/102198) is a strong constitutive promoter. The PcUbi promoter was fused to I-SceI. This resulted in two constructs named JB010qcz (SEQ ID NO: 11) and pCB666-3 (SEQ ID NO: 9).

Example 3: Constructs used to monitor intrachromosomal homologous recombination (ICHR) and marker excision

Example 3a: Construct to monitor ICHR by restoring GUS activity

In construct pCB583-40 (SEQ ID NO: 12) the T-DNA comprises the 35S CaMV constitutive promoter, a partial uidA (GUS) gene (called "GU"), an I-SceI recognition sequence and another partial uidA gene (called "US") as well as ocs terminator.

The partially overlapping halves of the GUS gene (GU and US) are non-functional, but as a result of ICHR a functional GUS gene will be restored. This can be monitored by hostochemical GUS staining (Jefferson 1985)

Example 3b: Construct to demonstrate marker excision

In construct pCB657-41 (SEQ ID NO: 10) the nit1P::nptII selectable marker expression cassette is surrounded by recognition sites for I-SceI and I-CreI and a direct repeat of the 35S terminator sequence. The 35S terminator sequence also functions as terminator for the nit1P::nptII expression cassette. The GUS gene functions as a reporter gene under control of the constitutive sTPT promoter (see above). The GUS gene as well as the sTPT promoter can easily be replaced by any other promoter::gene of interest cassette.

Thus, upon induction of ICHR by double-strand breaks, recombination between the duplicated terminator sequences may occur and lead to the loss of the nitP::nptII selectable marker expression cassette (i.e. marker excision), while the GUS expression cassette stays in the genome.

Example 4: Transformation of sequence-specific DNA endonuclease encoding constructs into *Arabidopsis thaliana*

Plasmids pCB603-100, pCB622-3, pCB653-37, pCB632-17, pCB652-124, pRS8, JB010qcz were transformed into *Arabidopsis* according to the protocol described in Example 1a. Selected transgenic lines (T1 generation) were grown in the greenhouse and some flowers were used for crossings (see below).

Example 5: Transformation of constructs to monitor marker excision into *Arabidopsis thaliana*

Plasmid pCB583-40 as well as pCB657-41 were transformed into *Arabidopsis* according to the protocol described in Example 1a. Selected transgenic lines (T1 generation) had been grown in the greenhouse and seeds had been harvested. T2 seeds had been grown *in vitro* on growth medium A (see Example 1a) supplemented with the respective selective agent (10 mg/L Phosphinotricin and 50 mg/L kanamycin, respectively). Individual, resistant plants from lines showing a 3:1 segregation have been transferred to soil and grown in the greenhouse.

Example 6: Transformation of sequence-specific DNA endonuclease encoding constructs into *Brassica napus*

Plasmid pCB632-17, pRS8 and pCB666-3 are transformed into rapeseed according to the protocol described in Example 1b. Selected transgenic lines (T1 generation) are grown in the greenhouse.

Example 7: Transformation of constructs to monitor marker excision into *Brassica napus*

Plasmid pCB657-41 is transformed into rapeseed according to the protocol described in Example 1b. Selected transgenic lines (T1 generation) are grown in the greenhouse. For comparison a construct pGU-Sce-US (identical to pCB583-40, but with nptII instead of pat selectable marker) has been transformed into rapeseed according to the protocol described in Example 1b. Selected transgenic lines (T1 generation) are grown in the greenhouse.

Example 8: Induction of ICHR by crossing sequence-specific DNA endonuclease expressing lines and lines harboring constructs to monitor ICHR and marker excision

Example 8a: Monitoring ICHR in *Arabidopsis*

Transgenic lines of *Arabidopsis* harboring the T-DNA of construct pCB583-40 have been crossed with lines of *Arabidopsis* harboring the T-DNA of constructs pCB603-100, pCB622-3, pCB653-37, pCB632-17, pCB652-124, pRS8, JB010qcz, respectively. F1 seeds of the crosses have been harvested. The seeds have been surface sterilized and grown on medium A supplemented with the respective antibiotics and/or herbicides. 3-4 old seedlings have been harvested and were used for histochemical GUS staining. The result is summarized in Table 2 and illustrated in Figure 27 for one particular example.

Table 2: Amount of blue areas as an indicator of tissues / parts of tissues in which ICHR occurred in crosses

Promoter driving I-SceI expression	Construct	Relative amount of tissues in which ICHR occurred
Control (no I-SceI crossed into reporter lines)	/	- (spontaneous frequency of ICHR in seedlings)
erecta	pCB653-37	-
AtCyc1	pCB622-3	+
AtSERK1	pCB603-100	-
invGF	pCB652-124	-
35S	pRS8	+++
sTPT	pCB666-3	++++
PcUbi	JB010cqz	++++++++

Some seedlings have been transferred to soil and seeds have been harvested. 3-4 weeks old F2 seedlings have been analysed by histochemical GUS staining. Only in crosses harbouring the GU-US reporter (pCB583-40) as well as the PcUbi::I-SceI construct (JB010cqz) completely blue plants have been detected. This indicates that almost all or all cells of the respective plants harboured the ICHR event. This was further confirmed by Southern hybridisation (Fig. 33). While F2 plants 6, 3 and13 from crosses K165-10, K170-10 and K170-10, respectively, show only the recombined event, other plants are chimeric and show the hybridisation band indicative for both, recombined and non-recombined events. This correlated with the histochmical GUS staining: K165-10-6, K170-10-3 and K170-10-13 were completely blue after histochemical GUS staining, while K170-4-3, K161-6-3, K161-6-7, for example, showed only blue sectors and blue spots. The conclusion that completely blue (recombined) F2 plants have been obtained was further confirmed by analysing the F3 descendants of the respective F2 plants. As expected, F3 plants were completely blue or completely white (as a result of segregation of the GUS gene). F3 descendants of K170-10-3, for example, were completely blue or completely white after histochemical GUS staining. Completely blue and completely white F3 plants appeared in a 3:1 ratio, indicating that the respective F2 plant was heterozygous for the GUS gene and the adjacent recombination cassette. F3 descendants of K170-10-13, for example, were all completely blue indicating that the respective F2 plant was homozygous for the GUS gene and the adjacent recombination cassette. Thus expression of I-SceI via parsley ubiquitin (PcUbi) promoter was so efficient that recombination had been induced on the maternal and paternal chromosome.

Therefore, the cassette PcUbi::I-SceI appears to be suitable for obtaining fully marker-free plants in the F2 of a respective cross.

Example 8b: Demonstrating marker excision in *Arabidopsis*

Transgenic lines of *Arabidopsis* harbouring a single integration of the T-DNA of construct pCB657-41 are crossed with lines of *Arabidopsis* harbouring the T-DNA of construct JB010qcz. F1 seeds of the crosses are harvested. The seeds are surface sterilized and grown on medium A supplemented with 100 nM Bimazethapyr. After transfer to soil F2 seeds are harvested. The F2 seeds are sawn and analysed by PCR. Plants which do not show a PCR fragment for the nptII selectable marker, but show a PCR

fragment for the GUS gene still present in the T-DNA of construct pCB657-41 after the ICHR event occurred, are free of the selectable marker encoded in the T-DNA of construct pCB657-41 due to the I-SceI induced induction of ICHR. This is confirmed by Southern Analysis. In the F3 generation plants will be selected, which lost the T-DNA of construct JB010qcz due to segregation. The respective plants are free of any selectable marker and do not obtain I-SceI, but only the gene of interest (i.e. GUS in this particular example).

The following Table 3 exemplifies such analysis for a particular cross (CR-CB613) between *Arabidopsis* plants harbouring a single integration of the T-DNA of construct pCB657-41 and lines of *Arabidopsis* harbouring the T-DNA of construct JB010qcz (PcUbi::I-SceI). Shown are results of a histochemical GUS staining of 30 F2 plants (named C24-CR-CB1055-PL-1 to 30) originating from line 3 of cross CR-CB613. The histochemical GUS staining is indicative for the presence of the GUS gene from the T-DNA of construct pCB657-41. A PCR to detect the nptII gene was conducted on GUS positive plants. The absence of nptII indicates that marker excision occurred and the T-DNA of pCB657-41 integrated into the respective *Arabidopsis* plants was recombined (nd – not determined). C24-CR-CB1055-PL-6, C24-CR-CB1055-PL-7, C24-CR-CB1055-PL-14 and C24-CR-CB1055-PL-20 may have lost the nitP::nptII selectable marker cassette from the T-DNA of pCB657-41 due to ICHR induced by I-SceI expressed under control of PcUbi promoter.

Table 3: Analysis of F2 *Arabidopsis* plants in order to identify selectable marker-free plants

Plant name	GUS staining	PCR nptII
C24-CR-CB1055-PL-1	-	nd
C24-CR-CB1055-PL-2	+	+
C24-CR-CB1055-PL-3	+	+
C24-CR-CB1055-PL-4	+	+
C24-CR-CB1055-PL-5	+	+
C24-CR-CB1055-PL-6	+	-
C24-CR-CB1055-PL-7	+	-
C24-CR-CB1055-PL-8	+	+
C24-CR-CB1055-PL-9	-	-
C24-CR-CB1055-PL-10	-	nd
C24-CR-CB1055-PL-11	+	+
C24-CR-CB1055-PL-12	+	+
C24-CR-CB1055-PL-13	+	+
C24-CR-CB1055-PL-14	+	-
C24-CR-CB1055-PL-15	+	+
C24-CR-CB1055-PL-16	+	+
C24-CR-CB1055-PL-17	+	+
C24-CR-CB1055-PL-18	+	+

Plant name	GUS staining	PCR nptII
C24-CR-CB1055-PL-19	+	+
C24-CR-CB1055-PL-20	+	-
C24-CR-CB1055-PL-21	+	+
C24-CR-CB1055-PL-22	+	+
C24-CR-CB1055-PL-23	+	+
C24-CR-CB1055-PL-24	-	nd
C24-CR-CB1055-PL-25	-	nd
C24-CR-CB1055-PL-26	+	+
C24-CR-CB1055-PL-27	+	+
C24-CR-CB1055-PL-28	+	+
C24-CR-CB1055-PL-29	-	nd
C24-CR-CB1055-PL-30	-	nd

Example 8c: Monitoring ICHR in rapeseed

Intrachromosomal homologous recombination (ICHR) is the mechanism underlying the marker excision concept described in this invention. This comparison example illustrates the results obtained when using ICHR in combination with the expression of an endonuclease under the control of a constitutive promoter. The example uses two constructs: pRS8 and pGU-sce-US, both binary vectors with nptII as a selection marker.

The T-DNA from pRS8 (SEQ ID NO.: 13), which is defined by its left and right borders (LB and RB, respectively), contains the following elements (from LB to RB): nptII expression cassette (nos promoter – nptII coding sequence – nos terminator); I-SceI expression cassette (CaMV 35S promoter – I-SceI coding sequence - ocs terminator). The T-DNA described is located on a plasmid (pSUN derivative) that contains origins for the propagation in *E. coli* as well as in *Agrobacterium* and an aadA expression cassette (conferring spectinomycin and streptomycin resistance) to select for transgenic bacteria cells.

The T-DNA from pGU-sce-US (Fig.13), which is defined by its left and right borders (LB and RB, respectively), contains the following elements (from LB to RB): nptII expression cassette (nos promoter – nptII coding sequence – nos terminator); constitutive promoter – partial uidA (GUS) gene (called "GU") – I-SceI recognition sequence – partial uidA gene (called "US") – ocs terminator. The partially overlapping halves of the GUS gene (GU and US) are non-functional, but as a result of ICHR a functional GUS gene will be restored. The T-DNA described is located on a plasmid (pSUN derivative) that contains origins for the propagation in *E. coli* as well as in *Agrobacterium* and an aadA expression cassette (conferring spectinomycin and streptomycin resistance) to select for transgenic bacteria cells.

The constructs pRS8 and pGU-sce-US, respectively, were separately introduced into B. napus *via Agrobacterium*-mediated transformation using the procedure described in Example 1b. Transgenic lines containing the separate constructs were selected in kanamycin-containing media and confirmed by molecular analysis (genomic PCR and genomic Southern blots). Several independent lines containing the pRS8 T-DNA and

the pGU-sce-US T-DNA, respectively, were isolated. These lines contained between 1 and 5 copies of the respective T-DNA, as determined by Southern blot.

T0 pRS8 and pGU-sce-US transgenic rapeseed plants (heterozygous for the respective transgenes) were crossed. F1 lines of these crosses were analyzed by genomic PCR and Southern blot in order to identify plants containing both transgenes. These plants were used for histochemical GUS staining and compared to siblings containing only the GU-sce-US transgene in order to determine the effect of double-strand breaks on ICHR.

Results: A dramatic increase in the frequency of ICHR due to the expression of the I-SceI can be observed, as shown by the number of blue spots per leaf, which are originated via ICHR between the duplicated parts of the GUS gene present in the GU-sce-US construct. Although the introduction of double-strand breaks by the expression of the I-SceI nuclease caused a very significant increase on the frequency of ICHR, so far all F1 plants analyzed were chimeric and no F1 plant having a entire blue staining was observed (Fig. 14).

Non-chimeric plants (exhibiting a complete blue staining) can be obtained by regenerating new plants from the marker free (or blue in this example) sectors on the chimeric plants (e.g., by inducing somatic embryogenesis and/or organogenesis / shoot generation). (The tissue culture process is summarized in Figure 28).

One particular pRS8 X pGU-sce-US cross, called C18, produced 9 F_1 plants that were called C18-1, C18-2, C18-3, C18-4, C18-5, C18-6, C18-7, C18-8, and C18-9, respectively. Three of these plants had many blues spots (C18-5, C18-7 and C18-9). Around 30 sterile F_2 seeds from plant C18-5 were germinated on MSB5 medium (4.4 g/l MS medium with B5 vitamins, 0.5 g/l MES, 3% sucrose, 0.8% oxoid agar; pH 5.8) and incubated at 21°C, 16h light (40-50µE/m^2s) / 8h dark for three weeks. The plants were then cut approx. 1 cm below the cotyledons and the top was transferred to fresh MSB5 medium for ~3 more weeks in the same conditions. When the leaves were 1-3 cm, one leaf per plant was used for GUS staining, and only the plants showing many spots per leaf were kept. The rest of the leaves of the positive plants were harvested, pooled, immersed in disruption medium (4.4 g/l MS medium with B5 vitamins, 0.5 g/l MES, 13% sucrose, 3,75 mg/l BAP, pH 5.8) and disrupted using a waring blender (3-5 pulses of 3-5 seconds). 100 mg aliquots of the leaf slurry obtained were plated on osmotic rafts over liquid regeneration medium (4.4 g/l MS medium with B5 vitamins, 0.5 g/l MES, 3% sucrose, 3 mg/l $AgNO_3$, 5 mg/l BAP, 5 mg/l NAA; pH 5.8), and subcultured every ~10 days until calli (first) and shoots (later) appeared. The shoot-forming explants were transferred to shoot development medium (4.4 g/l MS medium with B5 vitamins, 0.5 g/l MES, 1% sucrose, 3 mg/l BAP, 1 mg/l zeatin; pH 5.8), and subcultured every ~10 days until shoots reached a size of 1.5 –2 cm. Then they were transferred to Magenta boxes with shoot elongation medium (4.4 g/l MS medium with B5 vitamins, 0.5 g/l MES, 1% sucrose, 0.6% oxoid agar; pH 5.8) and subcultured every 15-20 days. When the shoots were big enough, all callus tissue was discarded by cutting and the rest of the explant was transferred to fresh medium for rooting. In total we regenerated 25 plants (Figure 29 shows regenerants 1 to 10). Of these, 20 had many spots, 3 were white (numbers 4, 6 and 26), and 2 were totally blue (numbers 10 and 15). The blue plants were regener-

ated from cells in which I-SceI cut between "GU" and "US" and the double strand break was repaired by intrachromosomal homologous recombination. The white plants were regenerated from cells in which I-SceI cut between "GU" and "US" and the double strand break was repaired by non homologous end joining (illegitimate recombination). The plants with spots were regenerated from cells in which I-SceI did not cut (or it did cut and the break was repaired, restoring the I-SceI recognition sequence). This was confirmed by the molecular analysis of several of these plants (PCR, Southern blot, and Northern blot). We performed PCRs (Fig. 30-I/II) with primers that amplify a band of 1433 bp on the unrecombined GU-US substrate. If recombination has taken place, we expect a fragment of 782 bp, which is what we obtained with regenerated plant number 10 (full blue plant). The white plants should produce a band of ~1400 bp, which is what we obtained with regenerated plants number 4 and 6 (white plants). The plants with spots must show both the 1433 and the 782 bp bands, and as expected regenerated plants number 1, 5 and 8 did so. These results were confirmed by digesting the PCR fragments with I-SceI. The enzyme only cut the 1433 bp band produced by the plants with spots, which are the only ones that still contained an I-SceI recognition sequence. In addition we performed a Southern blot analysis, digesting genomic DNA from regenerated plants with EcoRI + NotI or BamHI + PvuI (Figure 31-I/II). With the first double digestion and using a GUS probe, we obtained two bands of 1.3 and 2.3 kb in all plants except number 10, which gave only one band of 3 Kb. With the second double digestion and the same probe, we obtained one band of 2.6 Kb in all plants except number 10, which gave one band of 1.9 Kb after. These results confirmed that plant number 10 was indeed totally blue (equivalent to full marker excision). In addition Northern blot analysis (Figure 32) using a GUS probe showed that plant number 10 produced a transcript that was ~750 bp shorter than the transcript produced by the other plants, as expected after ICHR. Taken together, the results showed in figures 3, 4 and 5 unequivocally prove that we can obtain fully recombined plants (i.e. full blue plants) after regeneration from plants in which recombination was partial (i.e. plants with blue spots).

Rapeseed lines harbouring the T-DNA of pCB666-3, i.e. the I-SceI under control of the PcUbi promoter have been crossed to rapeseed plants harbouring the T-DNA of pGU-sce-US. Figure 34 B shows a leaf of such a cross after histochemical GUS staining in comparison to a cross of RS8 (35S::I-SceI) and pGU-sce-US plants (Figure 34 A). PcUbi driving I-SceI is especially good in creating big sectors in which recombination – as monitored by restoration of a functional GUS gene – occurs. As a result when PcUbi promoter was used, a much bigger area of the leaf comprises cells with the recombined event. In addition, these results demonstrate that PcUbi promoter driving I-SceI expression is superior over other promoter::I-SceI combinations including the gold-standard 35S promoter not only in the model *Arabidopsis* but also in other species such as the important crop rapeseed.

The frequency of obtaining completely blue plants with the method described above is much higher when the pGU-sce-US harbouring lines are crossed with rapeseed lines harbouring the T-DNA of pCB666-3, i.e. the I-SceI under control of the PcUbi. This demonstrates that the PcUbi promoter is much better suited for the purpose described in this invention.

Example 8d: Demonstrating marker excision in rapeseed

T1 plants harboring the T-DNA of construct pCB657-41 are crossed with lines harboring the plasmids pCB666-3 and pCB632-17, respectively. Seeds of the crosses are harvested and are germinated. The F1 seedlings of the crosses are used for regenerating new, completely marker free plants (regarding the selectable marker encoded between the duplication of the 35S terminator in pCB657-41; the GUS gene in this construct in this particular example is not used as a selectable marker. The selectable marker present in the T-DNA comprising the I-SceI – T-DNA from construct pCB663-3 and pCB632-17, respectively – is not intended to be deleted by this process. This marker is being segregated from the remaining GUS gene in the next generation; for the detailed protocol of regenerating new rapeseed plants from the F1 seedlings see above).

Alternatively to the regeneration protocol marker free plants are obtained by the following procedure. F2 plants are analyzed by PCR for the presence of I-SceI and the GUS reporter gene as well as for the absence of the nptII selectable marker cassette from the T-DNA of pCB657-41. In the F3 progeny then seedlings can be identified in which the T-DNA comprising I-SceI is segregated.

CLAIMS

What is claimed is:

1. A method for producing a transgenic plant comprising:

   i) crossing a first transgenic plant comprising in its genome a DNA construct comprising

      a1) at least one recognition sequence of at least 10 base pairs for the site-directed induction of DNA double-strand breaks by a sequence specific DNA-endonuclease and

      b1) a nucleic acid sequence to be excised,

      wherein said elements a1) and b1) and optionally further elements are flanked by homology sequences A and A', having sufficient length and sufficient homology in order to ensure homologous recombination between A and A', and having an orientation which – upon recombination between A and A' – will lead to an excision of said elements a1) and b1), and

      c1) at least one additional sequence conferring to said plant an agronomically valuable trait, wherein said sequence is not localized between the homology sequences A and A' and would not be excised from the genome upon recombination between A and A'

      with a second transgenic plant comprising in its genome an expression cassette comprising

      a2) the parsley ubiquitin promoter, and operably linked thereto

      b2) a nucleic acid sequence coding for a sequence specific DNA-endonuclease having a sequence specificity for said recognition sequence a1),

   ii) generating descendants (F1) following this crossing, and – optionally - sexually or asexually generating further descendants, and

   iii) isolating descendants which have undergone recombination between the homology sequences A and A' and which do not comprise in their genome said elements a1) and b1) but comprise sequence c1).

2. The method of Claim 1, wherein the element b1) is an expression cassette for a marker sequence.

3. The method of Claim 2, wherein the marker sequence is selected from the group consisting of negative selection marker, counter selection marker, positive selection marker, and reporter genes.

4. The method of any of Claim 1 to 3, wherein the method further comprises the step of segregating the expression cassette for the endonuclease from the sequence c1) for the agronomically valuable trait and isolating plants comprising sequence c1) but not said expression cassette for the endonuclease.

5. The method of any of Claim 1 to 4, wherein the parsley ubiquitin promoter comprises a sequence described by SEQ ID NO: 8 or 15 or a functional equivalent or functional equivalent fragment thereof.

6. The method of any of Claim 1 to 5, wherein the orientation of the homology sequences is in the form of direct repeats, which are flanking elements a1) and b1) and optionally further elements.

7. The method of any of Claim 1 to 6, wherein the sequence specific DNA-endonuclease is a homing endonuclease.

8. The method of any of Claim 1 to 7, wherein sequence specific DNA-endonuclease is a homing endonuclease selected from the group consisting of I-SceI, I-CpaI, I-CpaII, I-CreI and I-ChuI.

9. The method of Claim 8, wherein the sequence encoding the endonuclease is comprising an intron.

10. The method of any of Claim 1 to 9, wherein said construct comprises two recognition sequences a1) which are localized between the homology sequences A and A' and are flanking element b1) and optionally further elements in a way that cleavage at this two recognition sequences excises said element b1).

11. The method of any of Claim 1 to 10, wherein the homology sequences A and A' are part of the expression cassette comprised in the DNA construct.

12. The method of any Claim 1 or 11, wherein the resulting plant is marker-free.

13. A transgenic expression cassette comprising a sequence coding for a sequence specific DNA-endonuclease operably linked to the parsely ubiquitin promoter.

14. The transgenic expression cassette of Claim 13, wherein sequence specific DNA-endonuclease is a homing endonuclease selected from the group consisting of I-SceI, I-CpaI, I-CpaII, I-CreI and I-ChuI.

15. The transgenic expression cassette of Claim 14, wherein the sequence encoding the homing endonuclease is comprising an intron.

16. The transgenic expression cassette of any of Claim 13 to 15, wherein the parsley ubiquitin promoter comprises a sequence described by SEQ ID NO: 8 or 15 or a functional equivalent or functional equivalent fragment thereof.

17. A transgenic vector comprising a expression cassette of any of Claim 13 to 16.

18. A transgenic cell or nun-human organism comprising a expression cassette of any of Claim 13 to 16 or a vector of Claim 17.

19. A transgenic plant or plant cells comprising in the genome a expression cassette of any of Claim 13 to 16 or a vector of Claim 17.

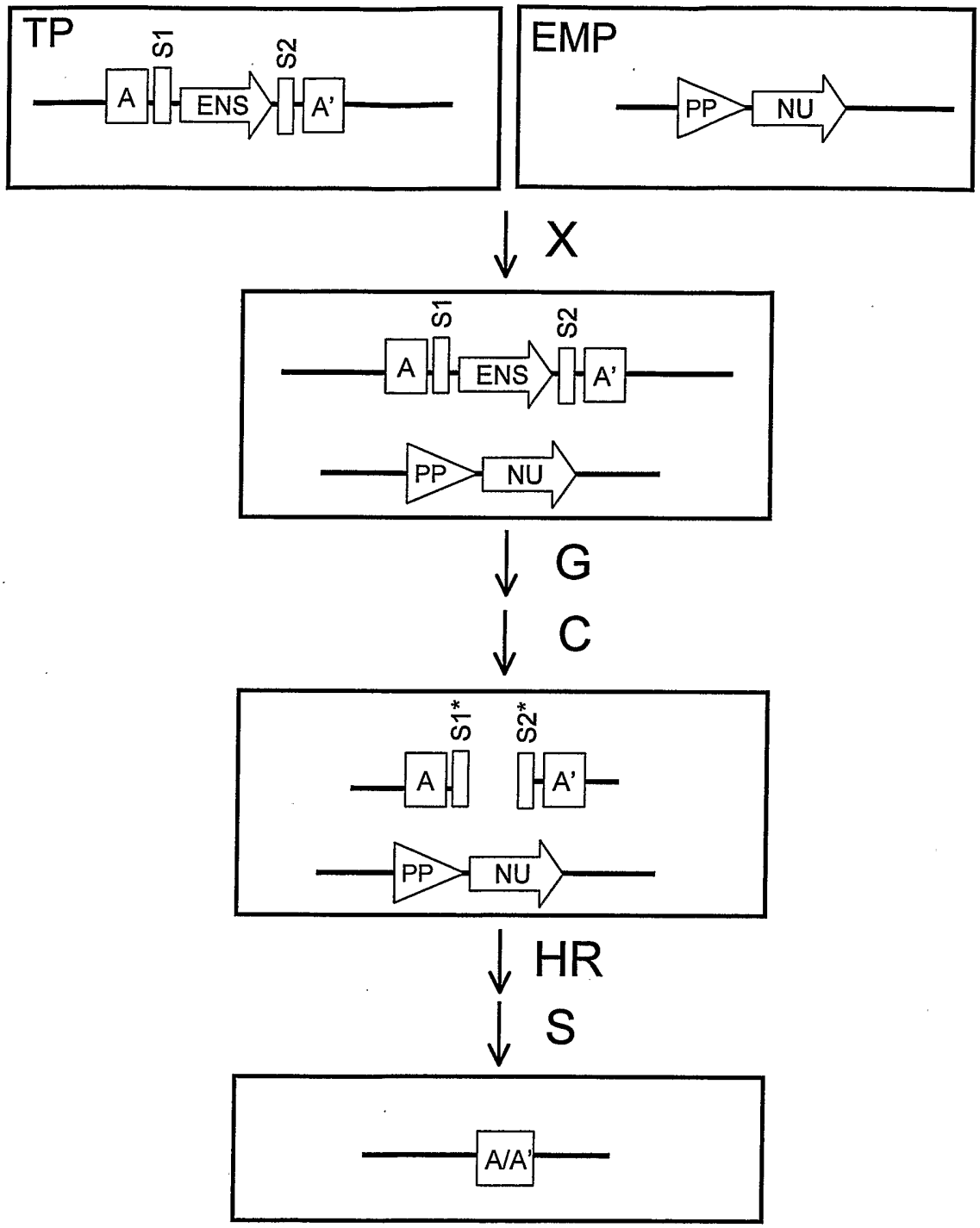


Fig. 1

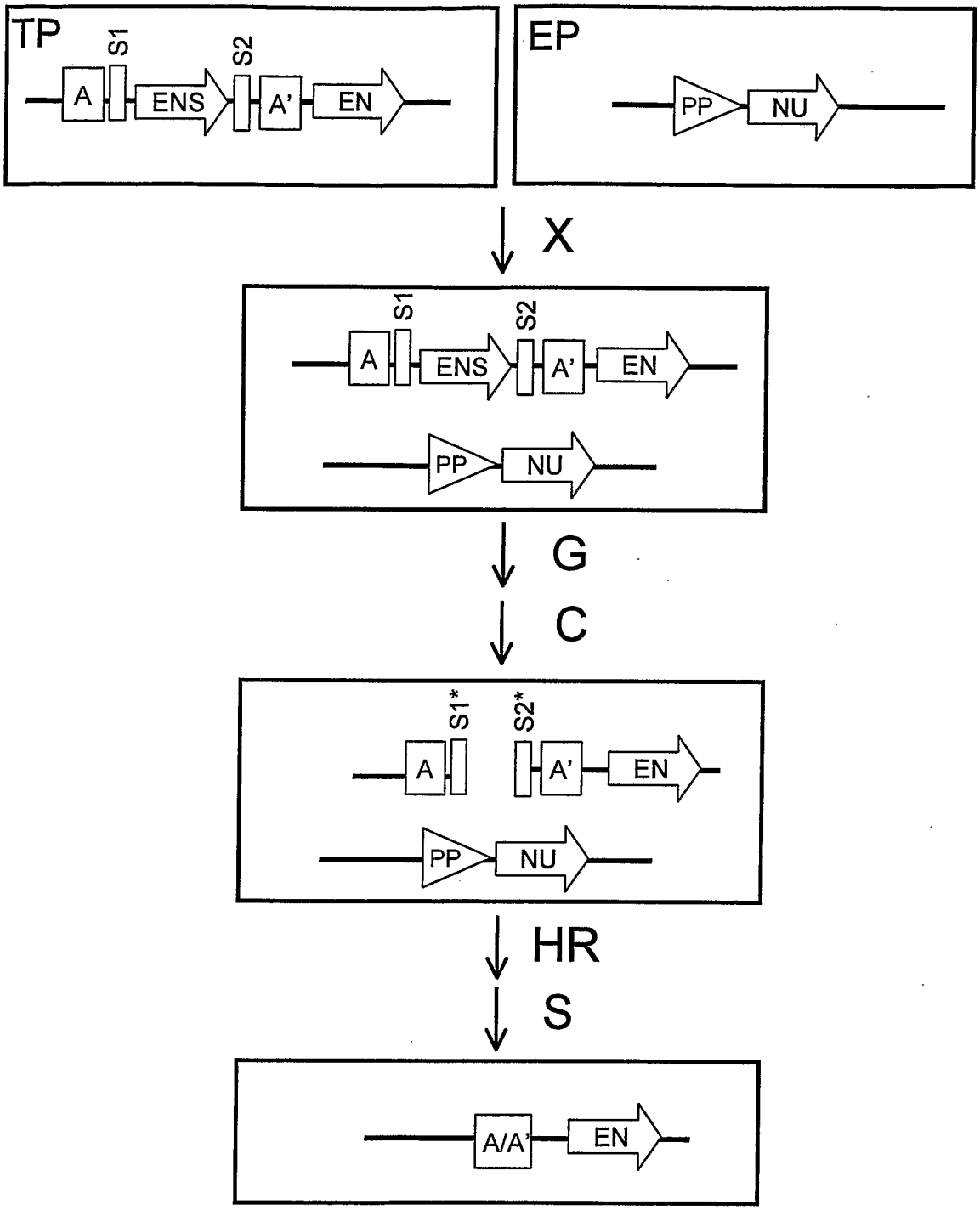

TP
S1
S2
A
ENS
A'
EN
EP
PP
NU
X
S1
S2
A
ENS
A'
EN
PP
NU
G
C
S1*
S2*
A
A'
EN
PP
NU
HR
S
A/A'
EN

Fig. 2

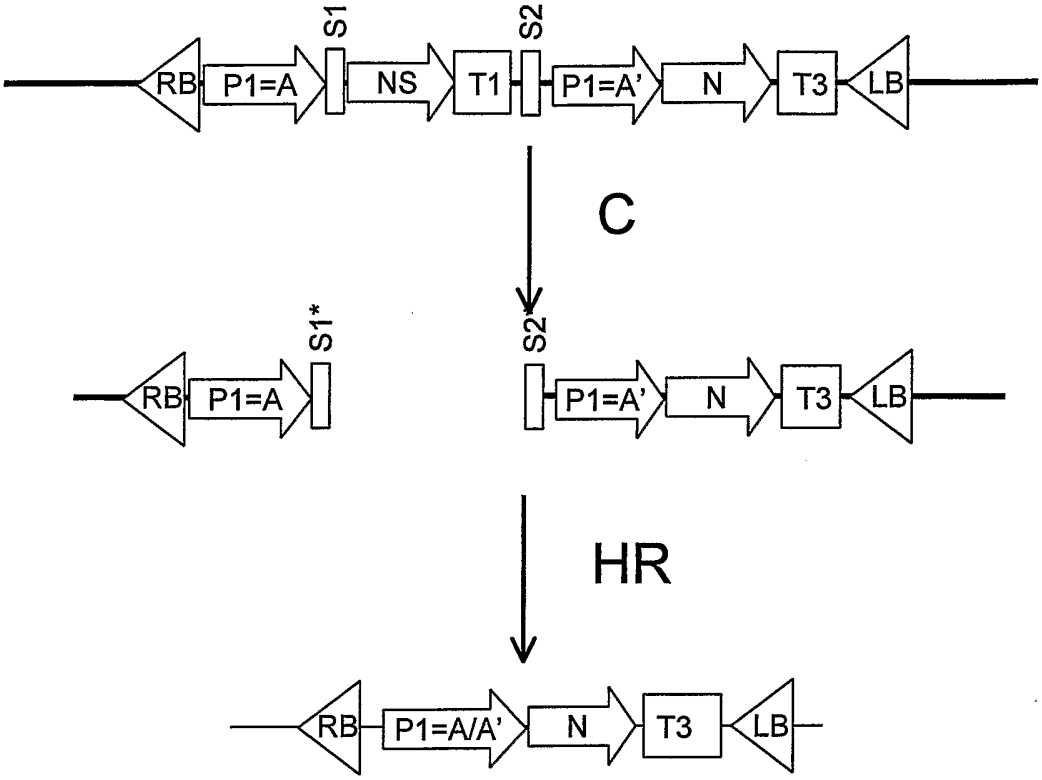

S1
S2
RB
P1=A
NS
T1
P1=A'
N
T3
LB
C
S1*
S2
RB
P1=A
P1=A'
N
T3
LB
HR
RB
P1=A/A'
N
T3
LB

Fig. 3

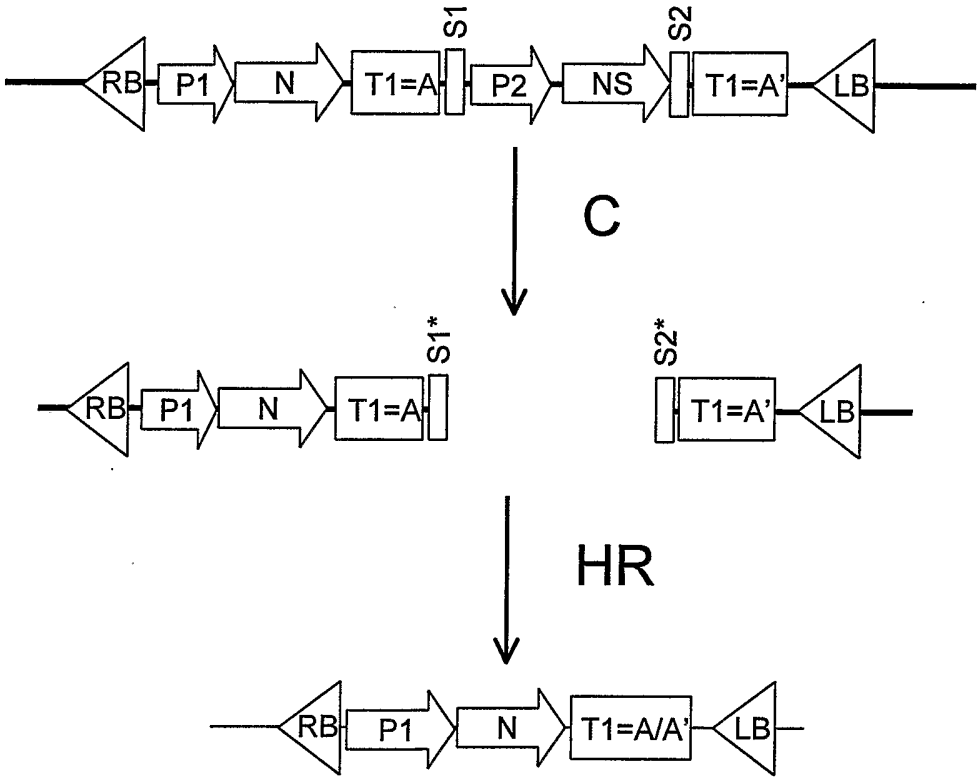


Fig. 4

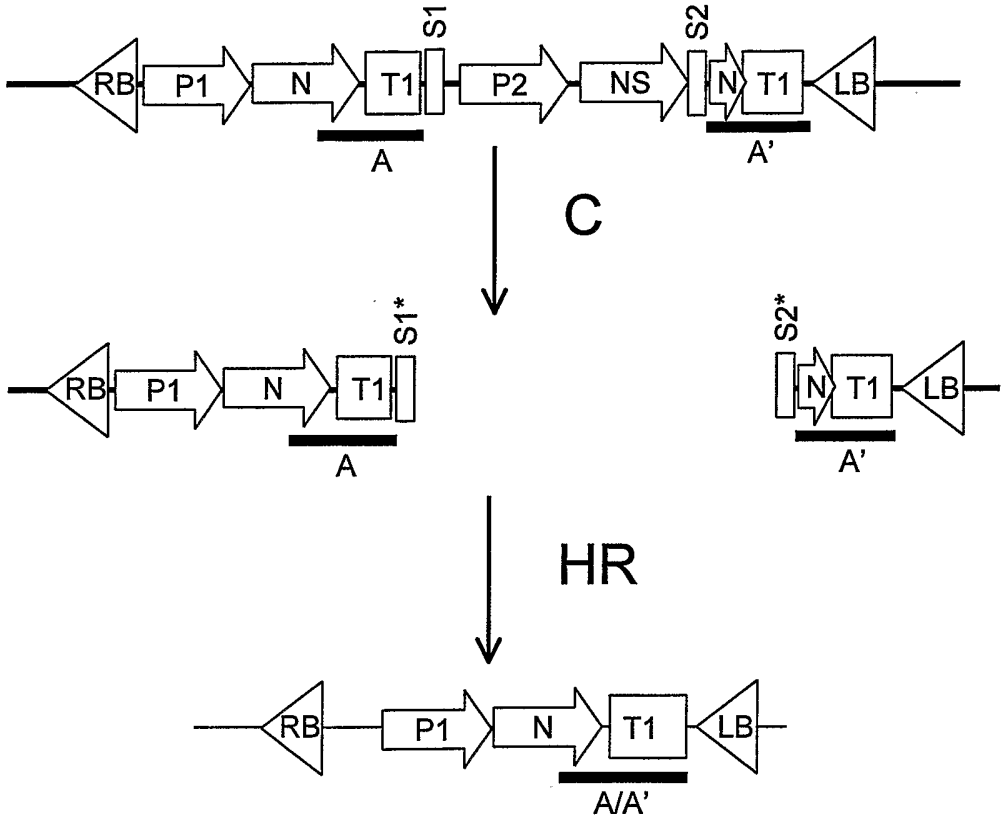


Fig. 5

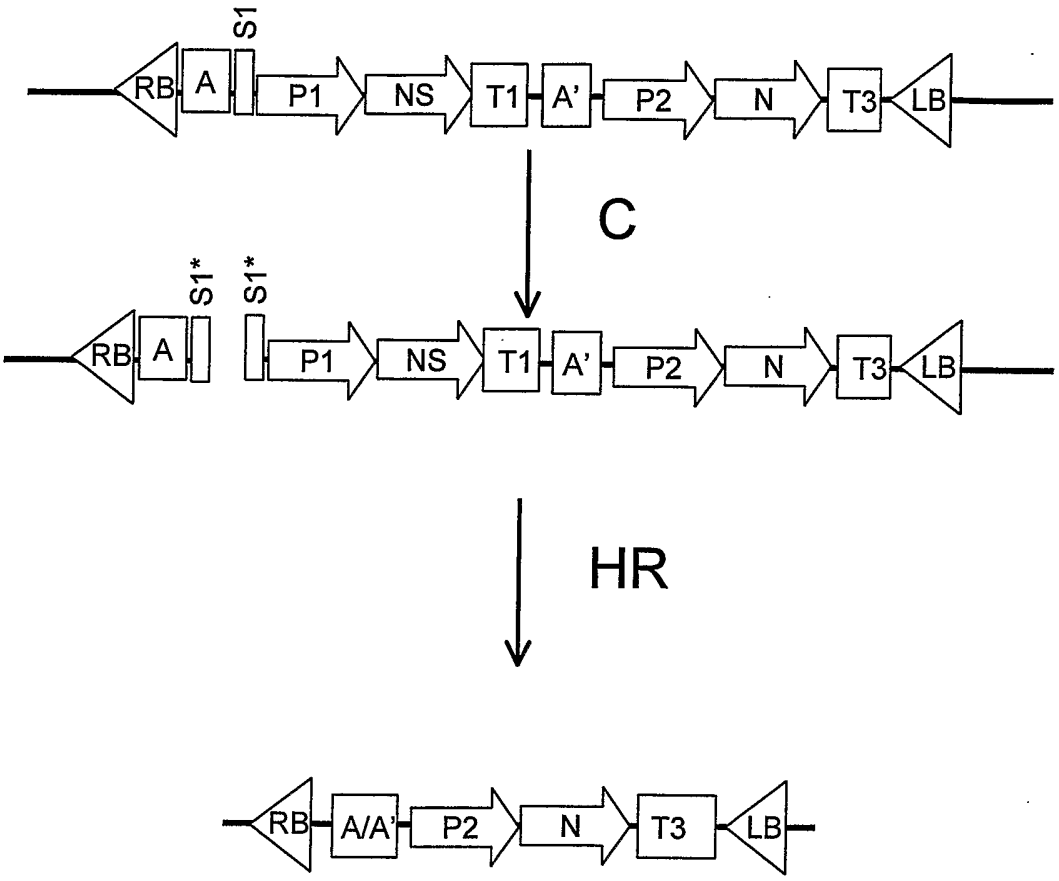

S1
RB
A
P1
NS
T1
A'
P2
N
T3
LB
C
S1*
S1*
RB
A
P1
NS
T1
A'
P2
N
T3
LB
HR
RB
A/A'
P2
N
T3
LB

Fig. 6

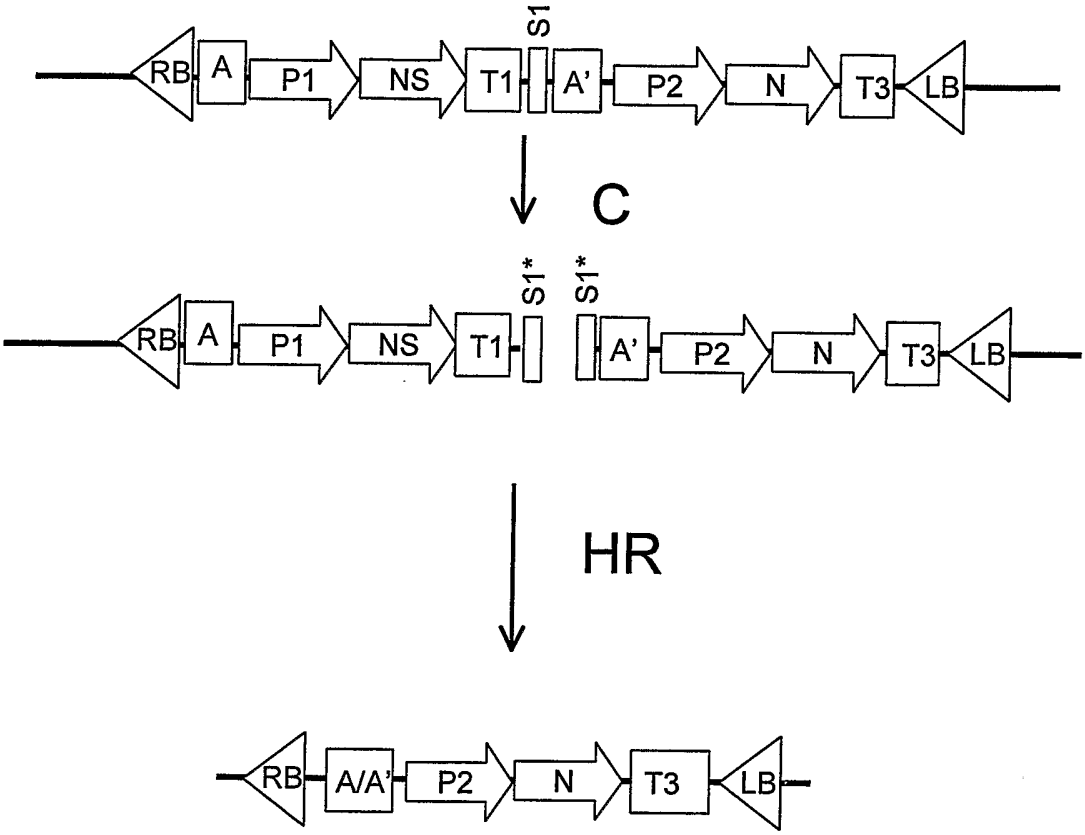

S1
RB
A
P1
NS
T1
A'
P2
N
T3
LB
C
S1*
S1*
RB
A
P1
NS
T1
A'
P2
N
T3
LB
HR
RB
A/A'
P2
N
T3
LB

Fig. 7

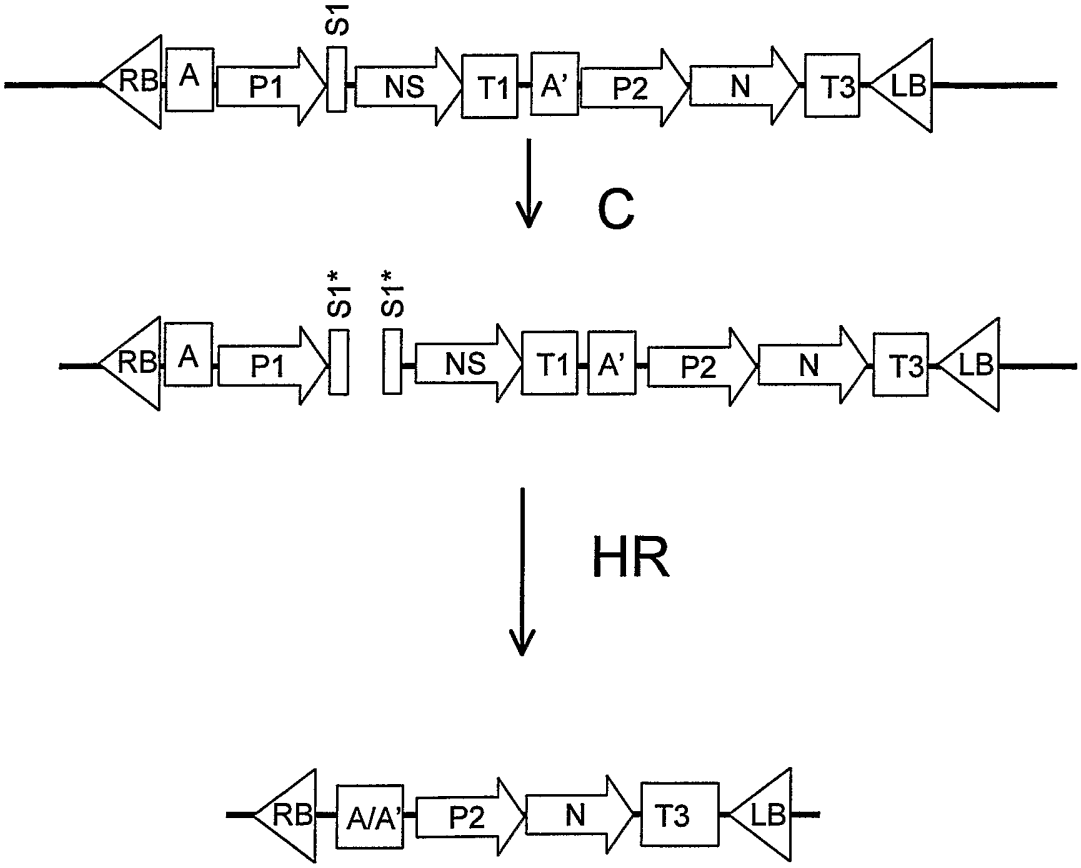


Fig. 8

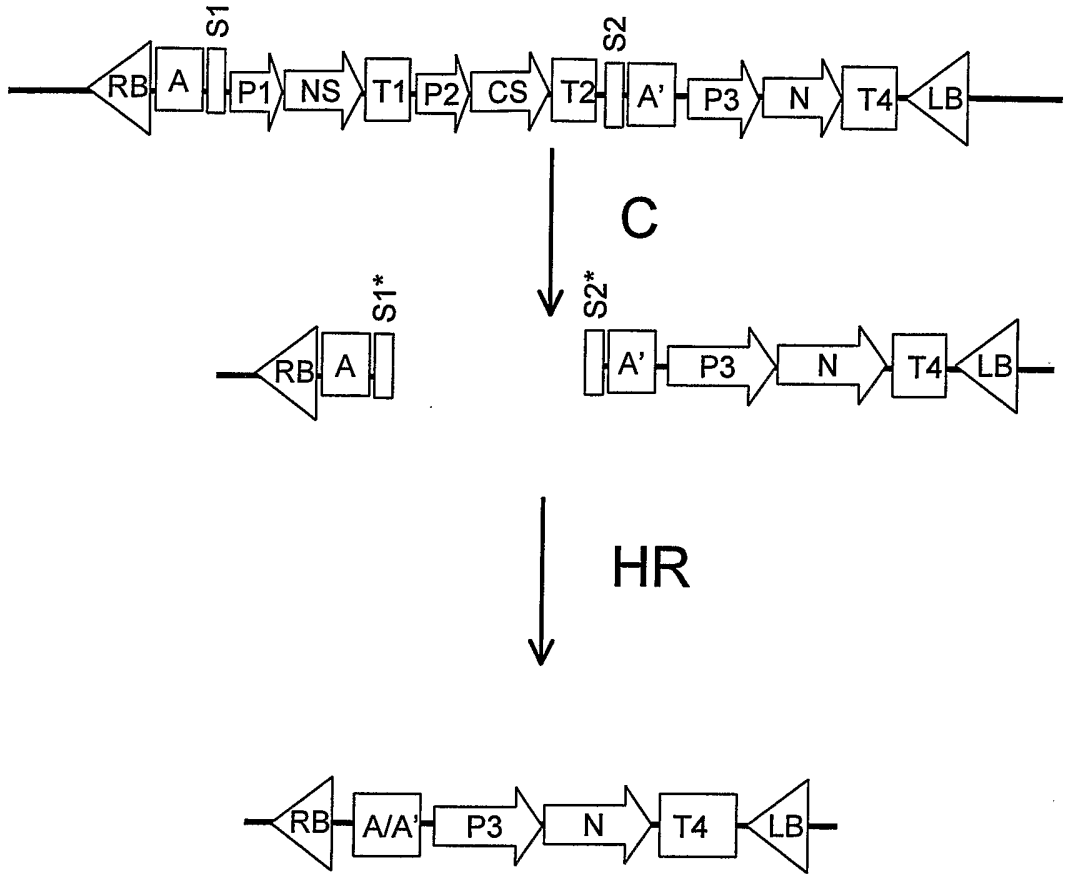

S1
S2
RB
A
P1
NS
T1
P2
CS
T2
A'
P3
N
T4
LB
C
S1*
S2*
RB
A
A'
P3
N
T4
LB
HR
RB
A/A'
P3
N
T4
LB

Fig. 9

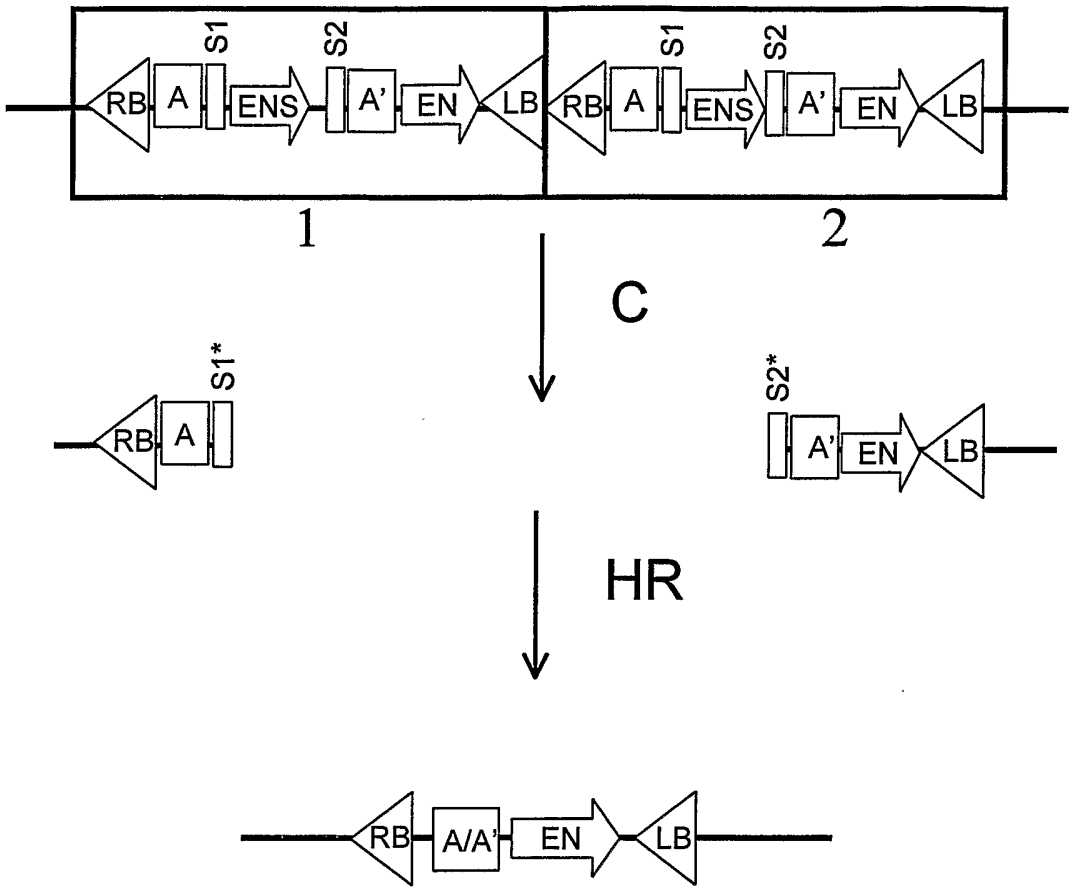


Fig. 10

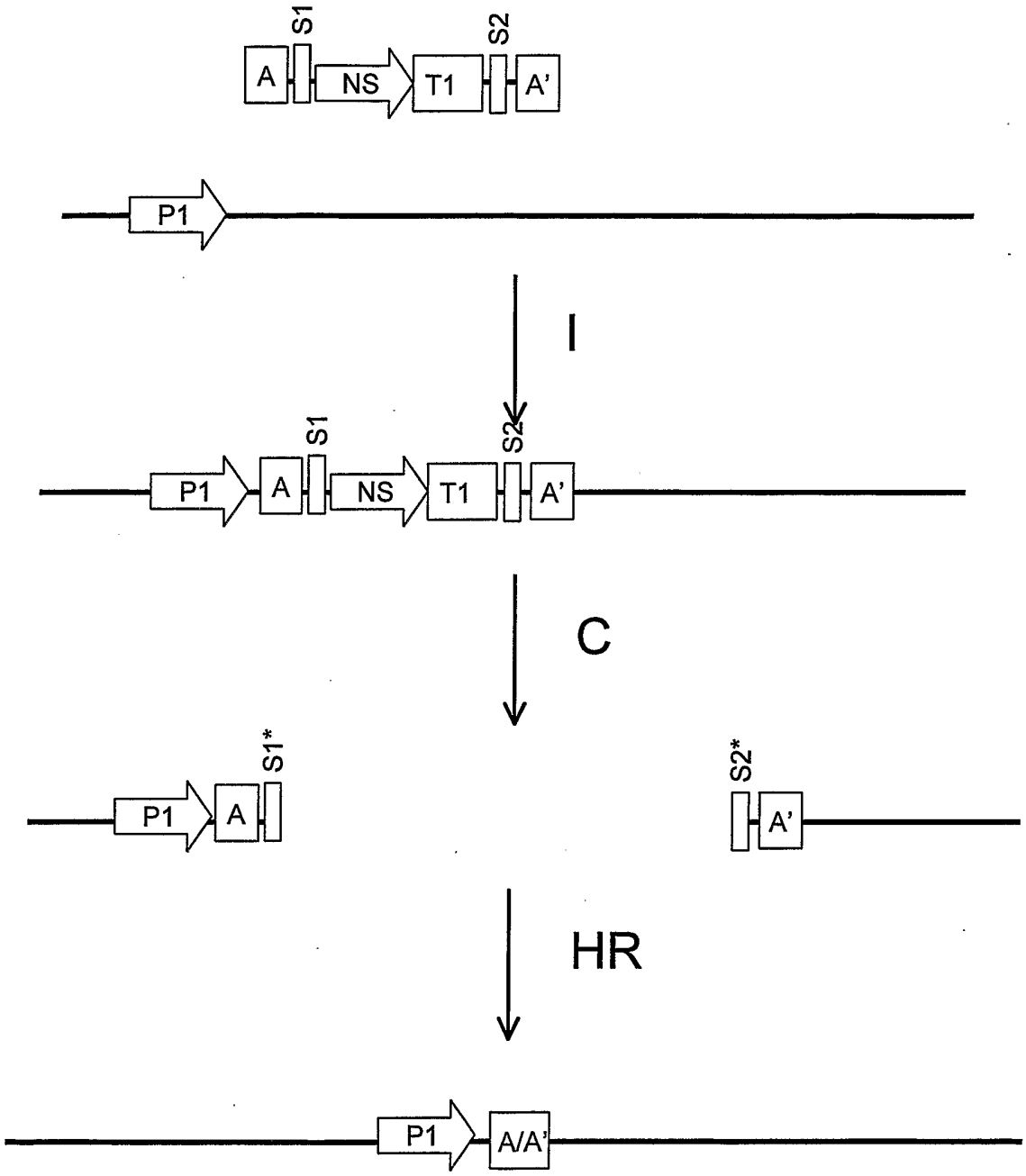


Fig. 11

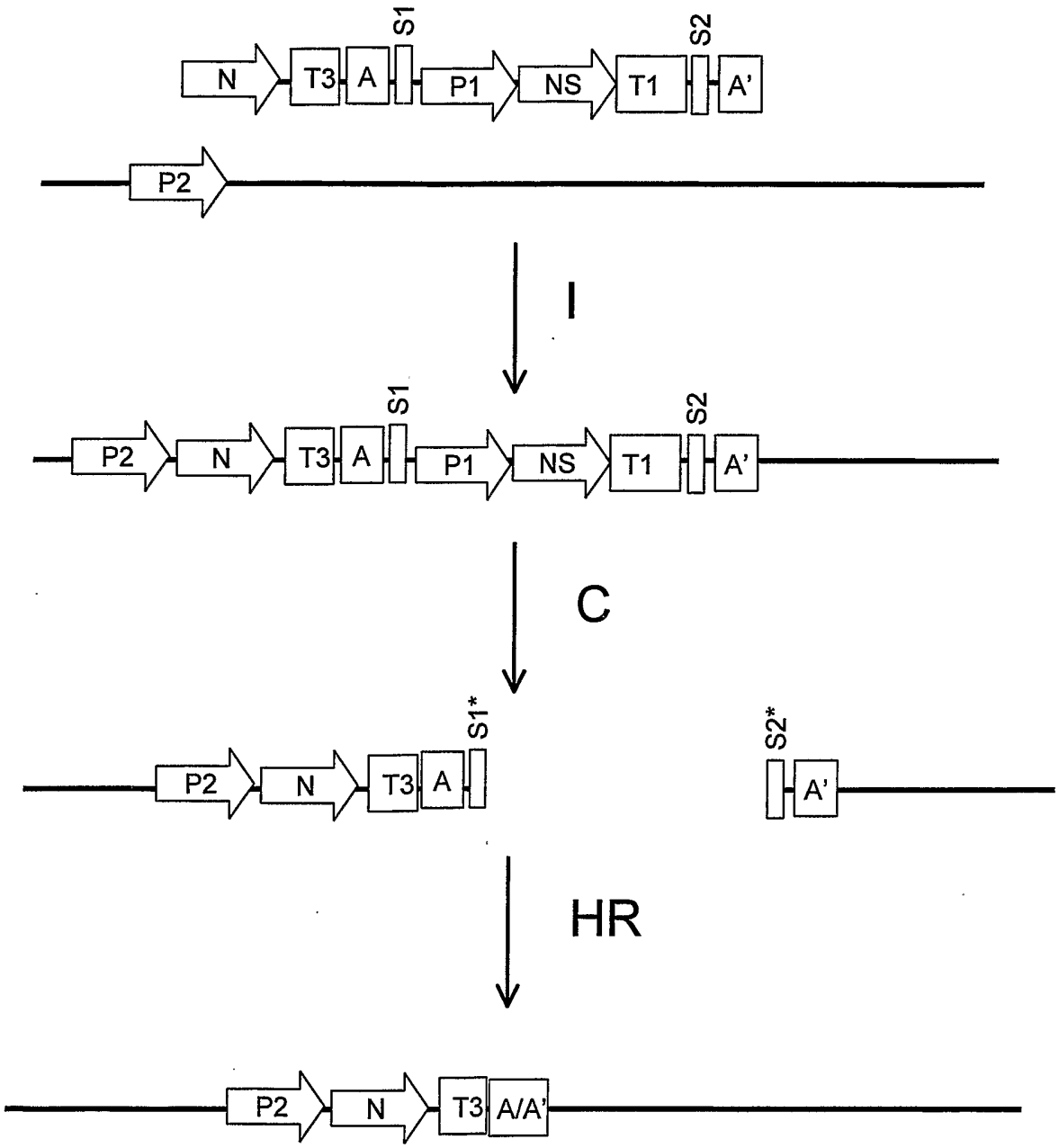

N
T3
A
S1
P1
NS
T1
S2
A'
P2
I
P2
N
T3
A
S1
P1
NS
T1
S2
A'
C
P2
N
T3
A
S1*
S2*
A'
HR
P2
N
T3
A/A'

Fig. 12

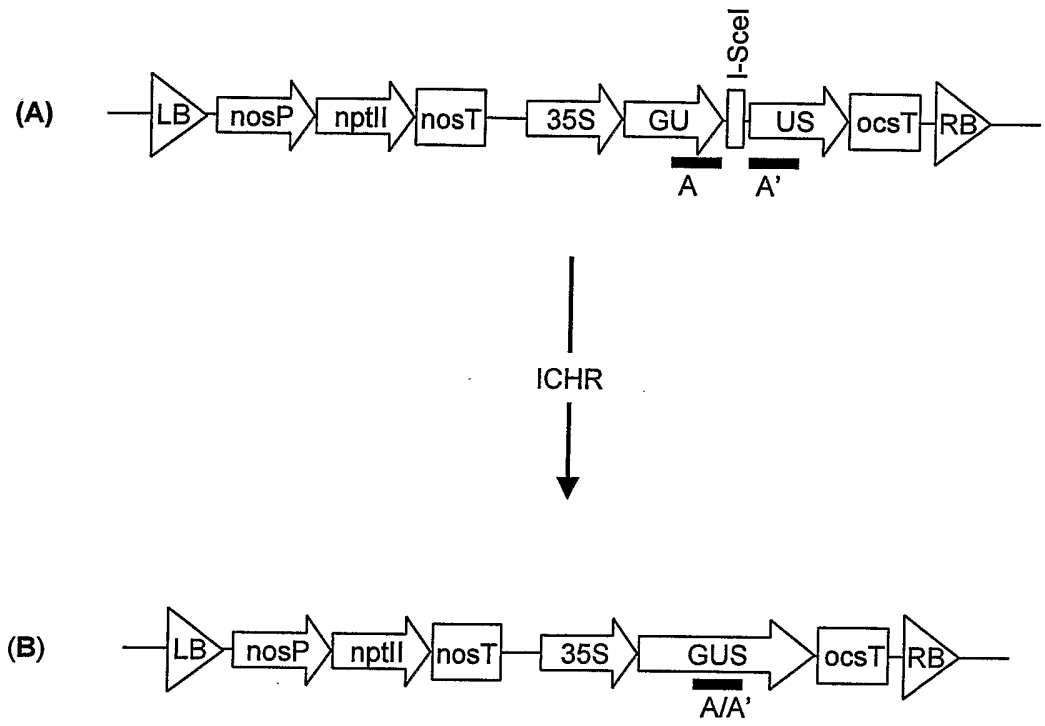

(A)
LB
nosP
nptII
nosT
35S
GU
I-SceI
US
ocsT
RB
A
A'
ICHR
(B)
LB
nosP
nptII
nosT
35S
GUS
ocsT
RB
A/A'

Fig. 13

A B

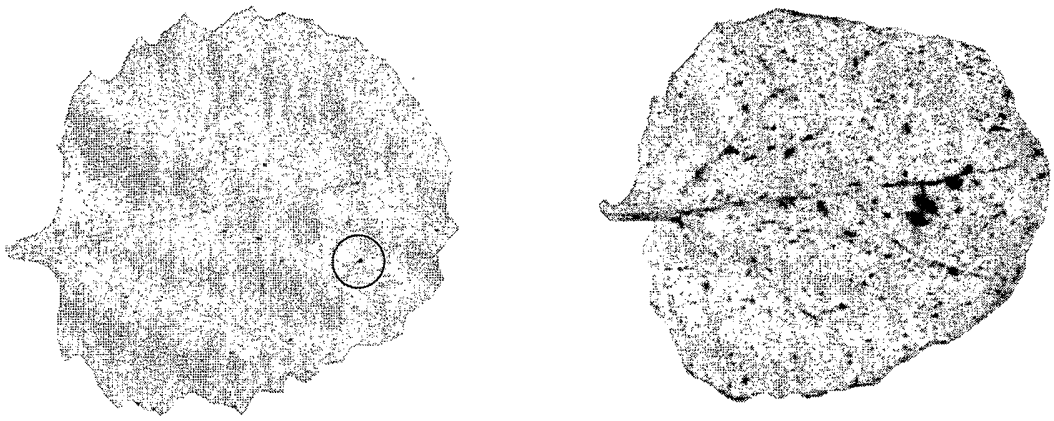


Fig. 14

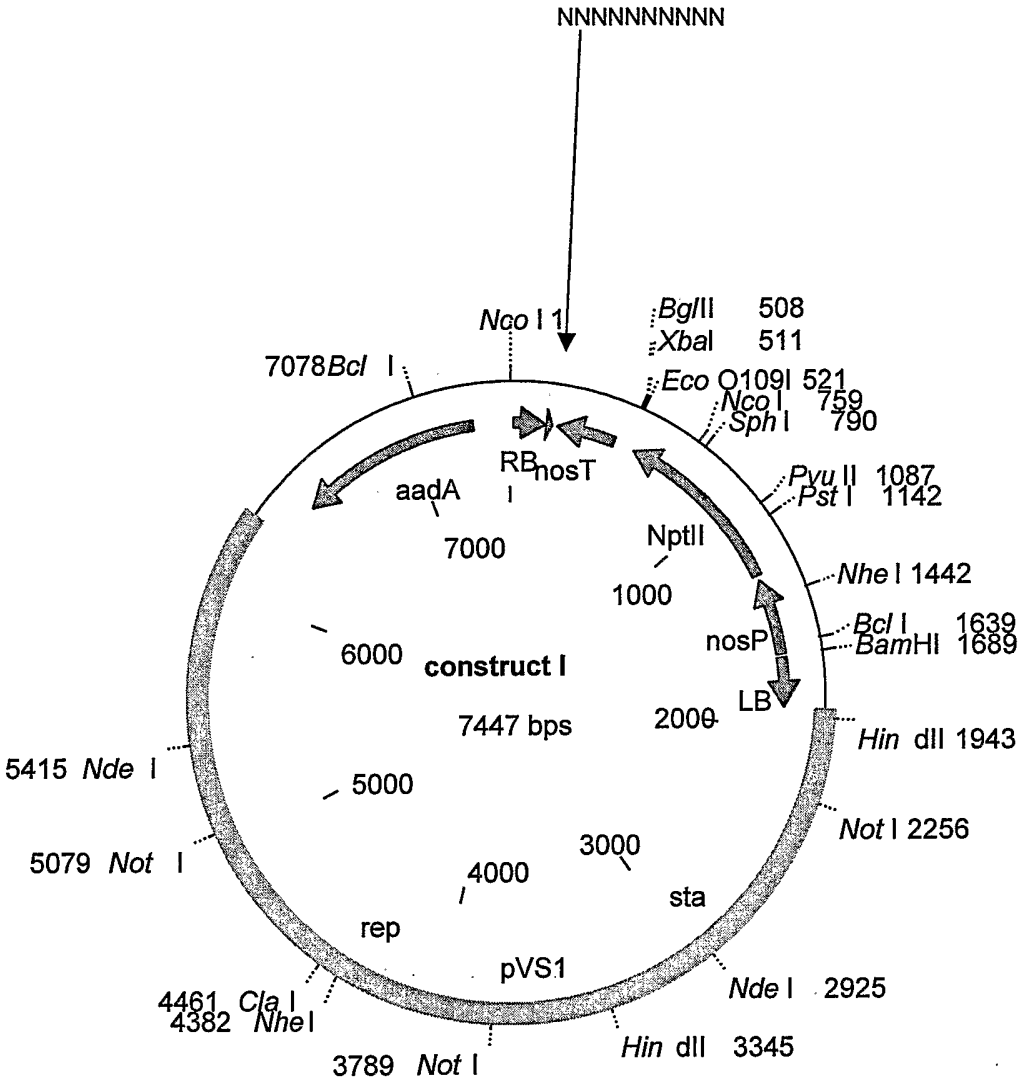

NNNNNNNNNN
Nco I 1
BglII 508
XbaI 511
Eco O109I 521
NcoI 759
SphI 790
PvuII 1087
PstI 1142
NheI 1442
BclI 1639
BamHI 1689
Hin dII 1943
Not I 2256
Nde I 2925
Hin dII 3345
Not I 3789
Cla I 4461
Nhe I 4382
Nde I 5415
Not I 5079
7078 Bcl I
RB
nosT
aadA
7000
NptII
1000
nosP
LB
2000
3000
sta
4000
rep
pVS1
5000
6000
construct I
7447 bps

Fig. 15

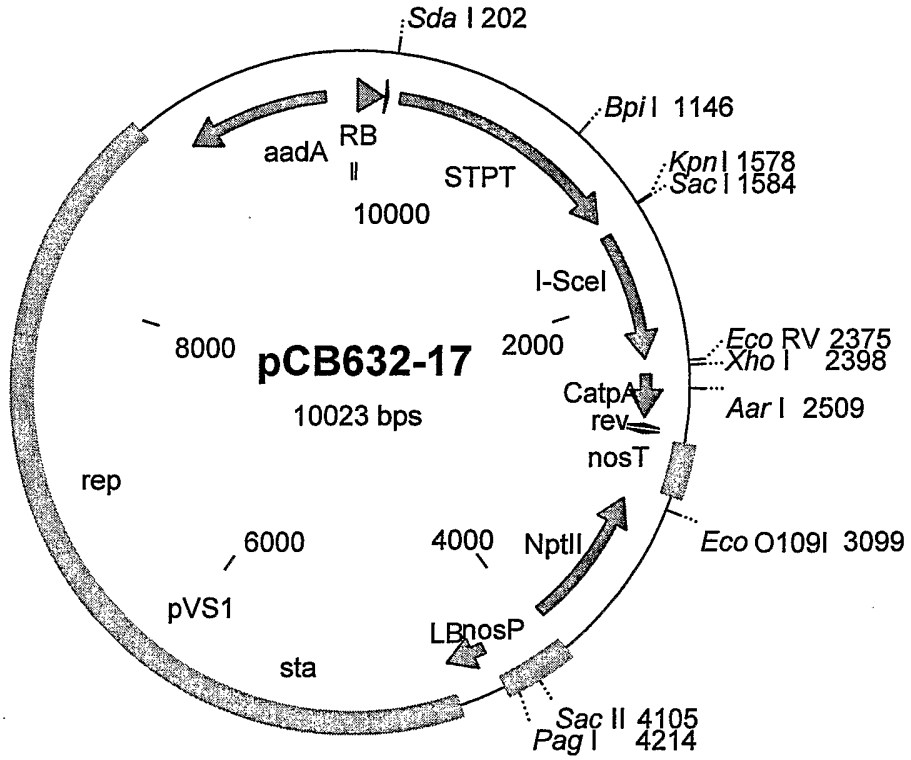

Sda I 202
Bpi I 1146
Kpn I 1578
Sac I 1584
aadA
RB
STPT
10000
I-SceI
8000
pCB632-17
2000
10023 bps
Eco RV 2375
Xho I 2398
Aar I 2509
CatpA
rev
nosT
rep
6000
4000
NptII
Eco O109I 3099
pVS1
LBnosP
sta
Sac II 4105
Pag I 4214

Fig. 16

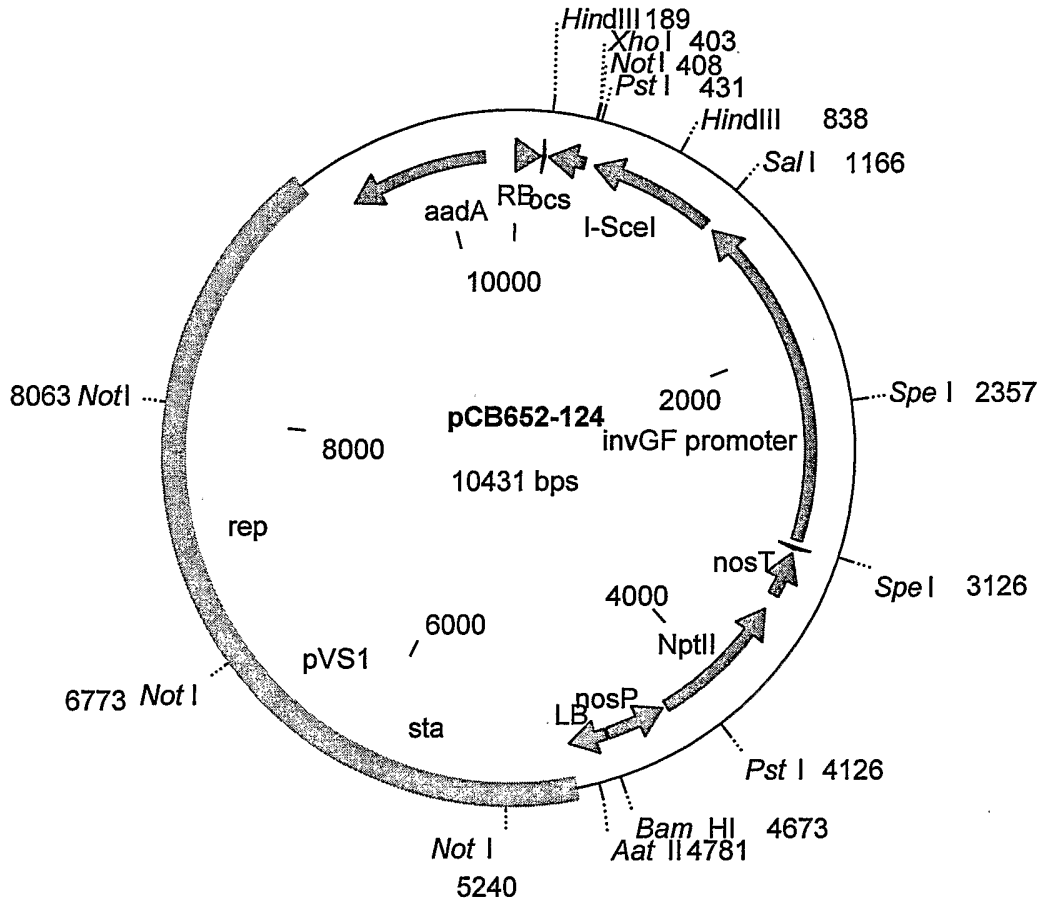

HindIII 189
Xho I 403
Not I 408
Pst I 431
HindIII 838
Sal I 1166
RB
bcs
aadA
I-SceI
10000
2000
Spe I 2357
8063 Not I
pCB652-124
invGF promoter
8000
10431 bps
rep
nosT
Spe I 3126
4000
6000
NptII
pVS1
nosP
LB
6773 Not I
sta
Pst I 4126
Bam HI 4673
Not I
Aat II 4781
5240

Fig. 17

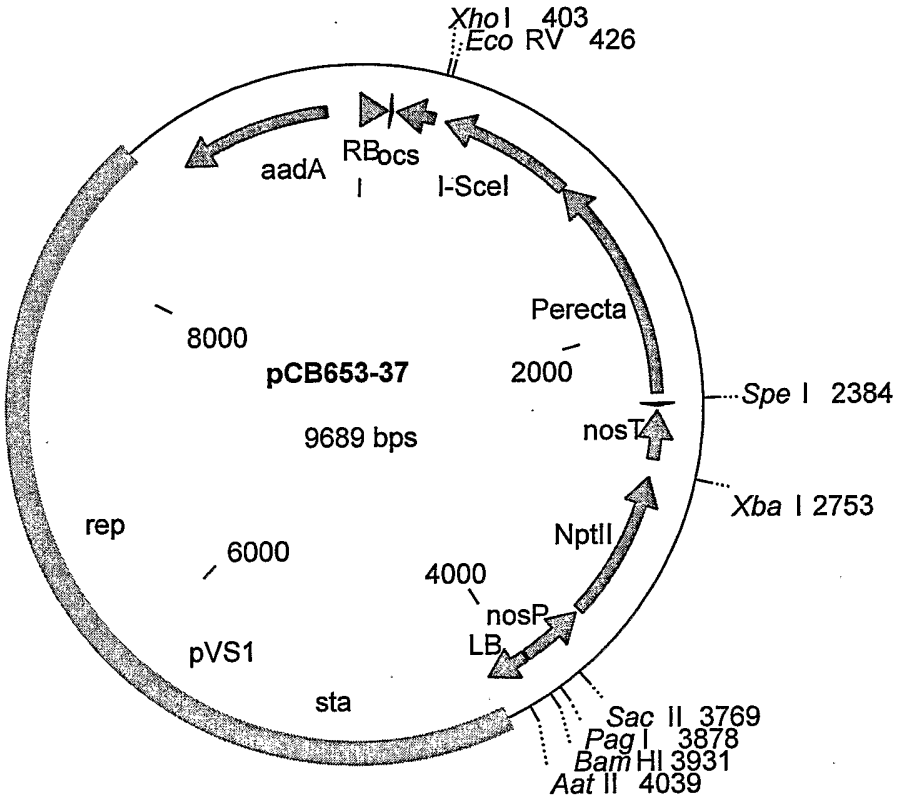

Xho I 403
Eco RV 426
RB
ocs
I
aadA
I-SceI
Perecta
8000
pCB653-37
2000
9689 bps
Spe I 2384
nosT
Xba I 2753
rep
NptII
6000
4000
nosP
LB
pVS1
sta
Sac II 3769
Pag I 3878
Bam HI 3931
Aat II 4039

Fig. 18

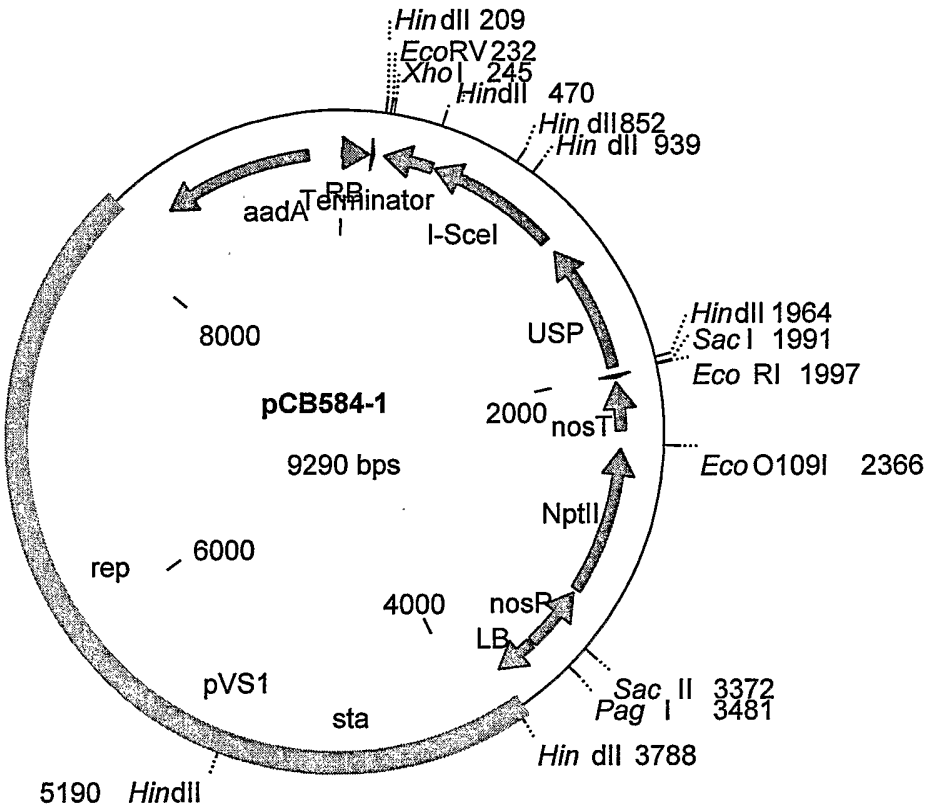

HindIII 209
EcoRV 232
XhoI 245
HindIII 470
HindIII 852
HindIII 939
aadA
RB
Terminator
I-SceI
8000
USP
HindIII 1964
SacI 1991
EcoRI 1997
pCB584-1
2000
nosT
9290 bps
EcoO109I 2366
NptII
6000
rep
4000
nosP
LB
pVS1
sta
SacII 3372
PagI 3481
HindIII 3788
5190 HindIII

Fig. 19

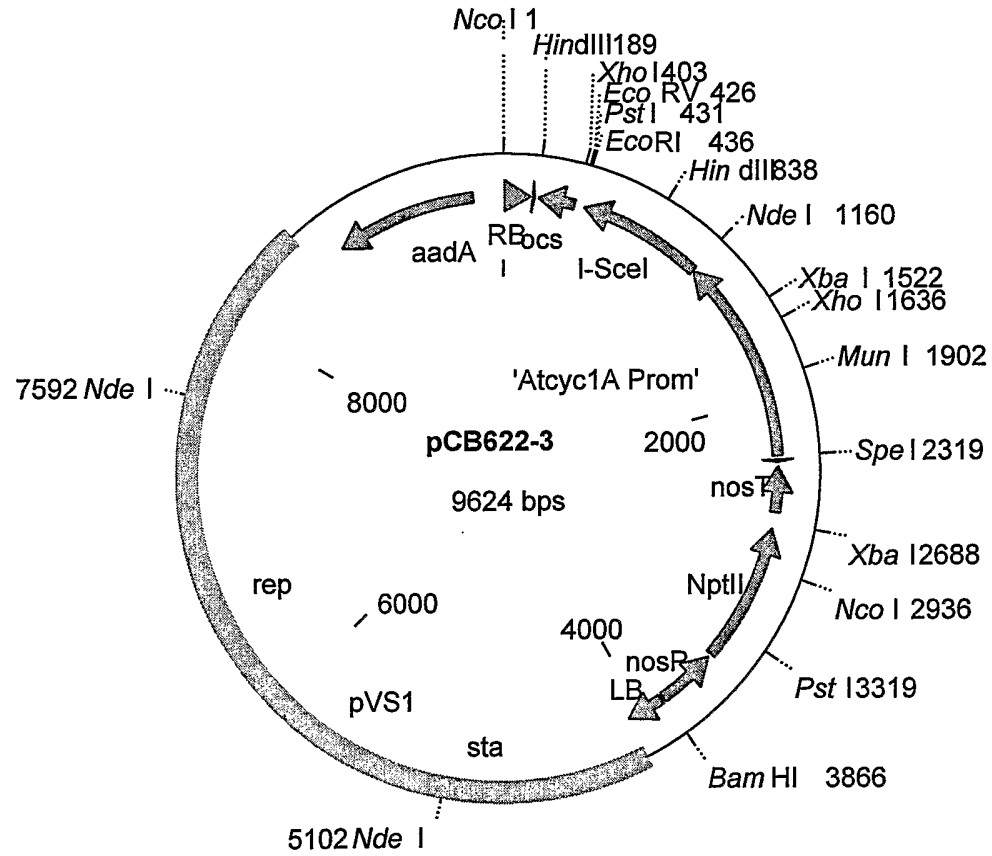

Nco I 1
HindIII 189
Xho I 403
Eco RV 426
Pst I 431
EcoRI 436
Hin dIII 838
Nde I 1160
Xba I 1522
Xho I 1636
Mun I 1902
Spe I 2319
Xba I 2688
Nco I 2936
Pst I 3319
Bam HI 3866
5102 Nde I
7592 Nde I
aadA
RB
ocs
I-SceI
'Atcyc1A Prom'
8000
2000
pCB622-3
9624 bps
nosT
NptII
nosP
LB
4000
6000
rep
pVS1
sta

Fig. 20

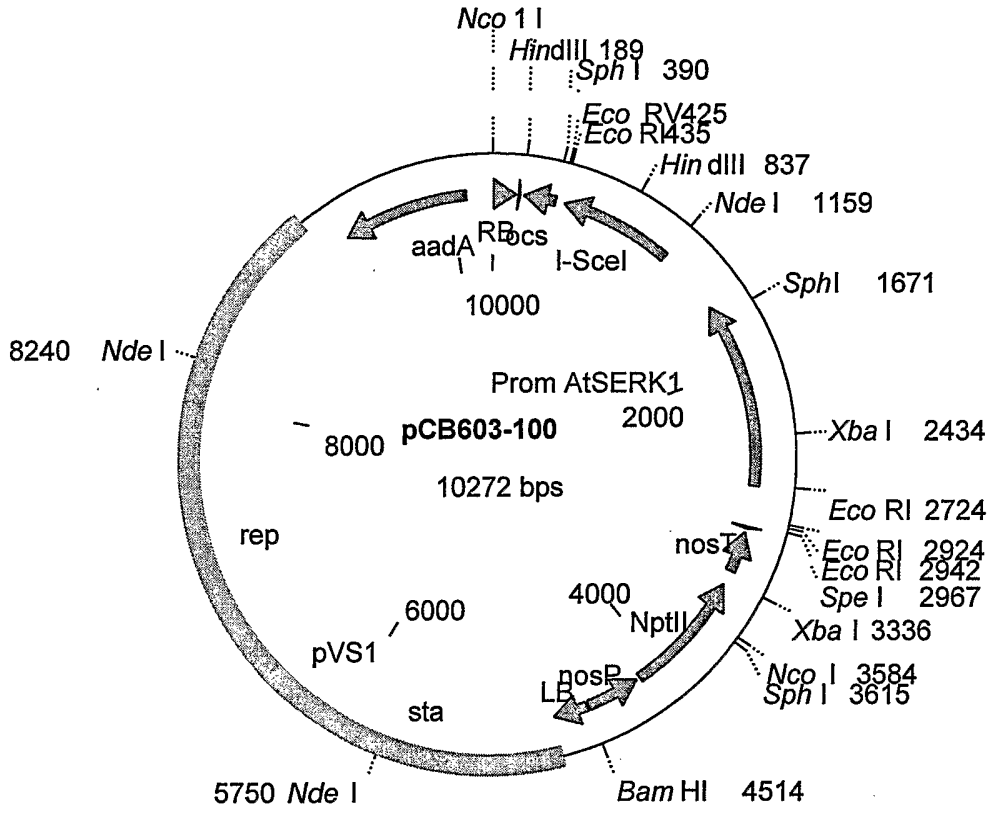

Nco 1 I
HindIII 189
Sph I 390
Eco RV425
Eco RI435
Hin dIII 837
Nde I 1159
SphI 1671
Xba I 2434
Eco RI 2724
Eco RI 2924
Eco RI 2942
Spe I 2967
Xba I 3336
Nco I 3584
Sph I 3615
Bam HI 4514
5750 Nde I
sta
pVS1
6000
rep
8000
8240 Nde I
aadA
RB
ocs
I-SceI
10000
Prom AtSERK1
2000
pCB603-100
10272 bps
nosT
4000
NptII
nosP
LB

Fig. 21

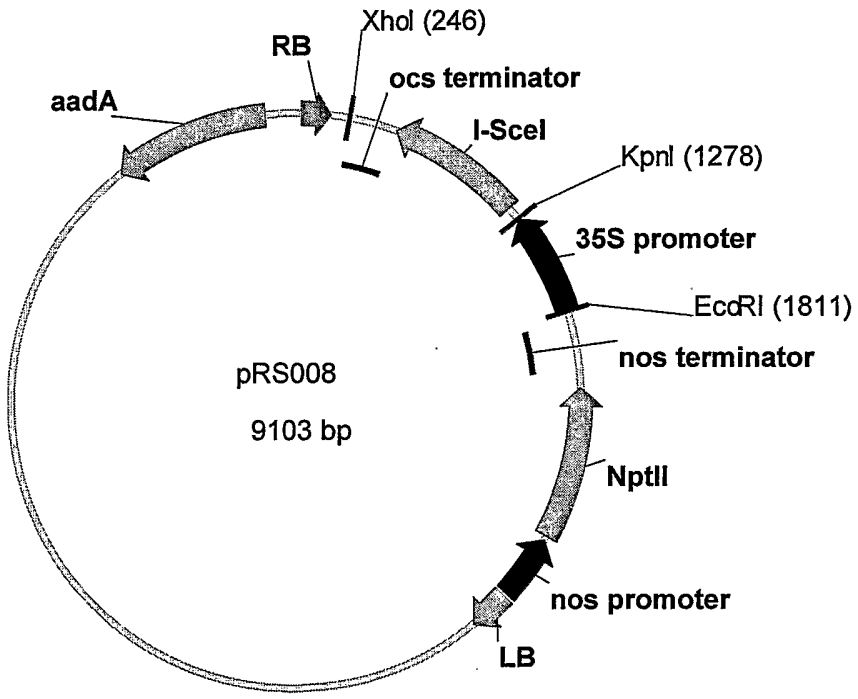

XhoI (246)
RB
ocs terminator
aadA
I-SceI
KpnI (1278)
35S promoter
EcoRI (1811)
nos terminator
pRS008
9103 bp
NptII
nos promoter
LB

Fig. 22

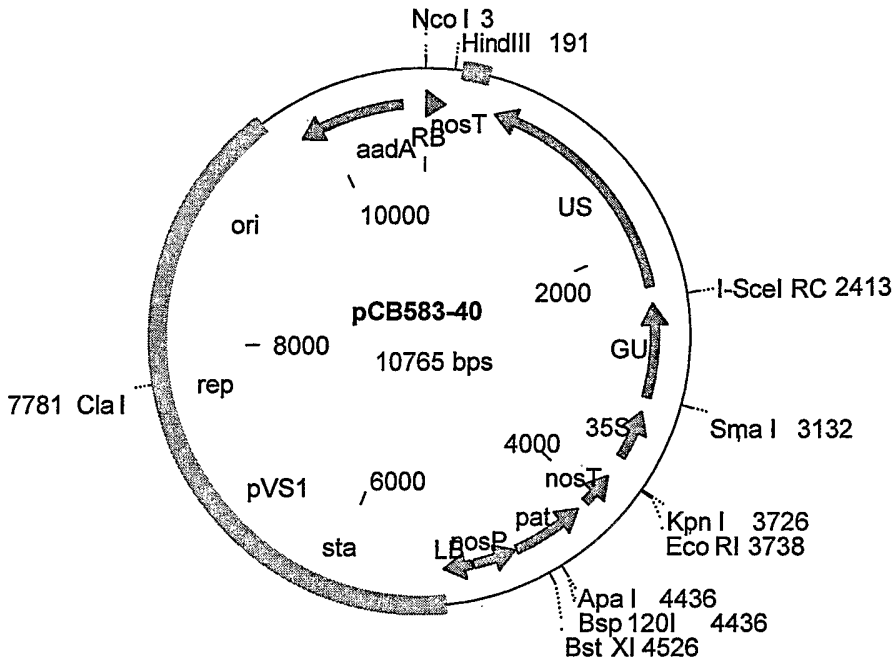

Nco I 3
HindIII 191
aadA
RB
nosT
10000
US
ori
2000
I-SceI RC 2413
pCB583-40
8000
10765 bps
GU
rep
7781 Cla I
35S
Sma I 3132
4000
nosT
pVS1
6000
pat
Kpn I 3726
Eco RI 3738
sta
LB
nosP
Apa I 4436
Bsp 120I 4436
Bst XI 4526

Fig. 23

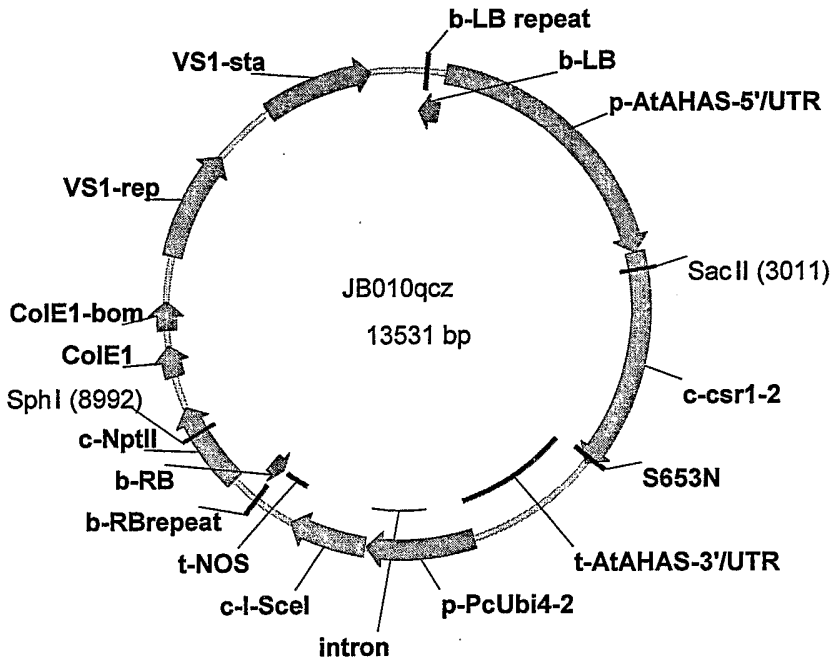

b-LB repeat
b-LB
p-AtAHAS-5'/UTR
SacII (3011)
c-csr1-2
S653N
t-AtAHAS-3'/UTR
p-PcUbi4-2
intron
c-I-SceI
t-NOS
b-RBrepeat
b-RB
c-NptII
SphI (8992)
ColE1
ColE1-bom
VS1-rep
VS1-sta
JB010qcz
13531 bp

Fig. 24

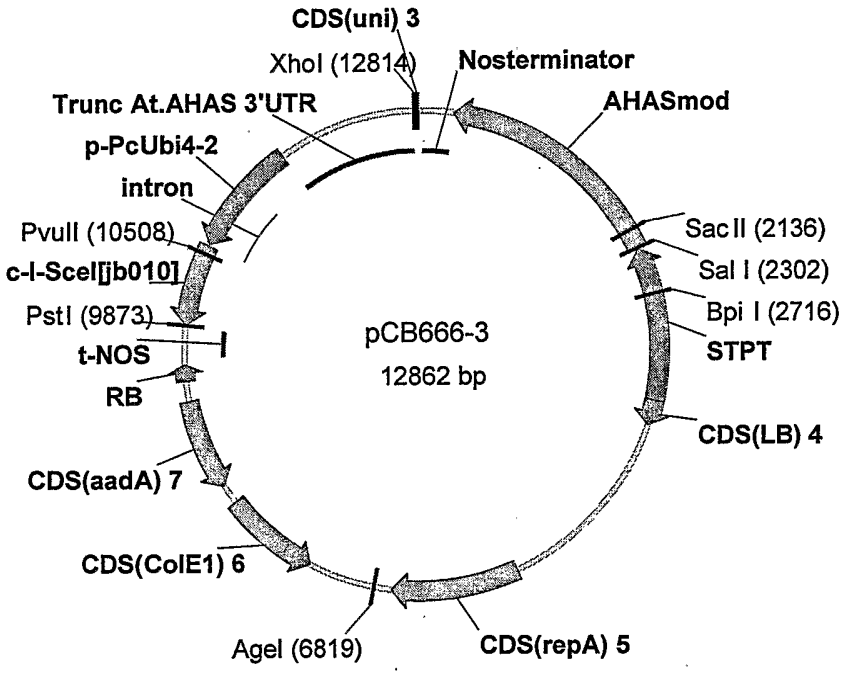

CDS(uni) 3
XhoI (12814)
Nosterminator
Trunc At.AHAS 3'UTR
AHASmod
p-PcUbi4-2
intron
PvuII (10508)
SacII (2136)
c-I-SceI[jb0010]
SalI (2302)
PstI (9873)
BpiI (2716)
t-NOS
pCB666-3
12862 bp
STPT
RB
CDS(LB) 4
CDS(aadA) 7
CDS(ColE1) 6
AgeI (6819)
CDS(repA) 5

Fig. 25

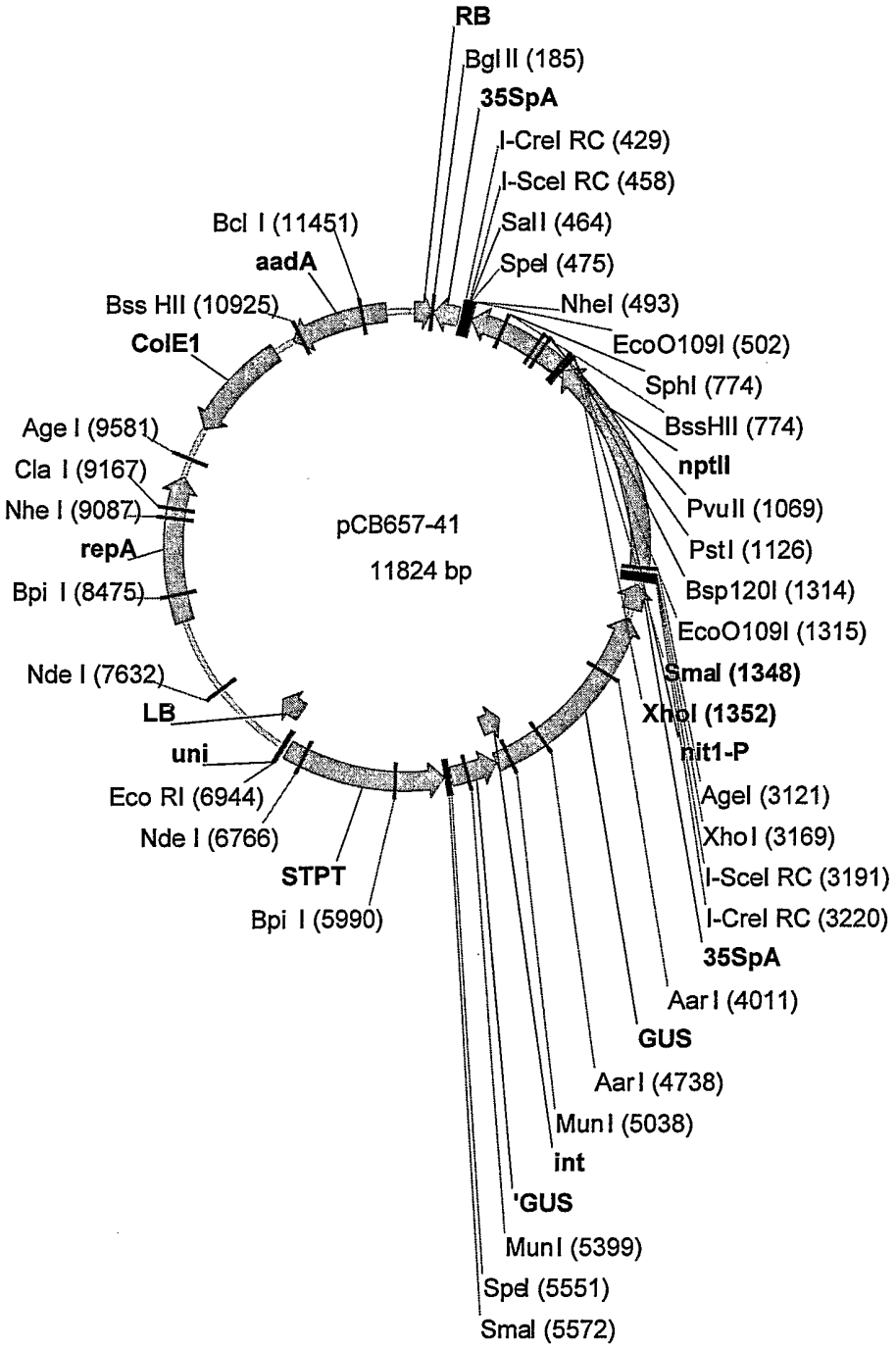
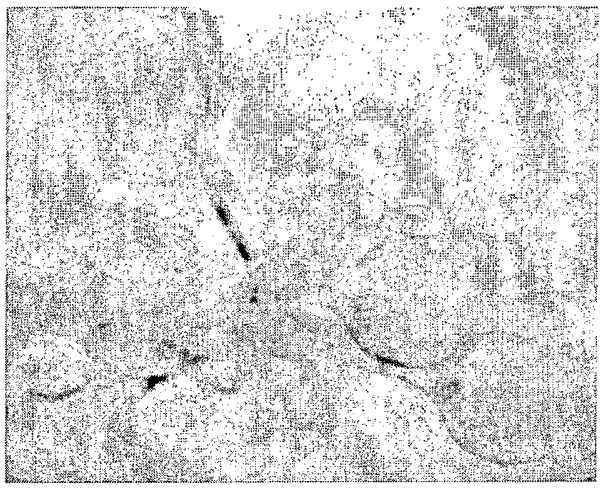

RB
BglII (185)
35SpA
I-CreI RC (429)
I-SceI RC (458)
SalI (464)
SpeI (475)
NheI (493)
EcoO109I (502)
SphI (774)
BssHII (774)
nptII
PvuII (1069)
PstI (1126)
Bsp120I (1314)
EcoO109I (1315)
SmaI (1348)
XhoI (1352)
nit1-P
AgeI (3121)
XhoI (3169)
I-SceI RC (3191)
I-CreI RC (3220)
35SpA
AarI (4011)
GUS
AarI (4738)
MunI (5038)
int
'GUS
MunI (5399)
SpeI (5551)
SmaI (5572)
Bpi I (5990)
STPT
Nde I (6766)
Eco RI (6944)
uni
LB
Nde I (7632)
Bpi I (8475)
repA
Nhe I (9087)
Cla I (9167)
Age I (9581)
ColE1
Bss HII (10925)
aadA
BclI (11451)
pCB657-41
11824 bp

Fig. 26

A



B



Fig. 27

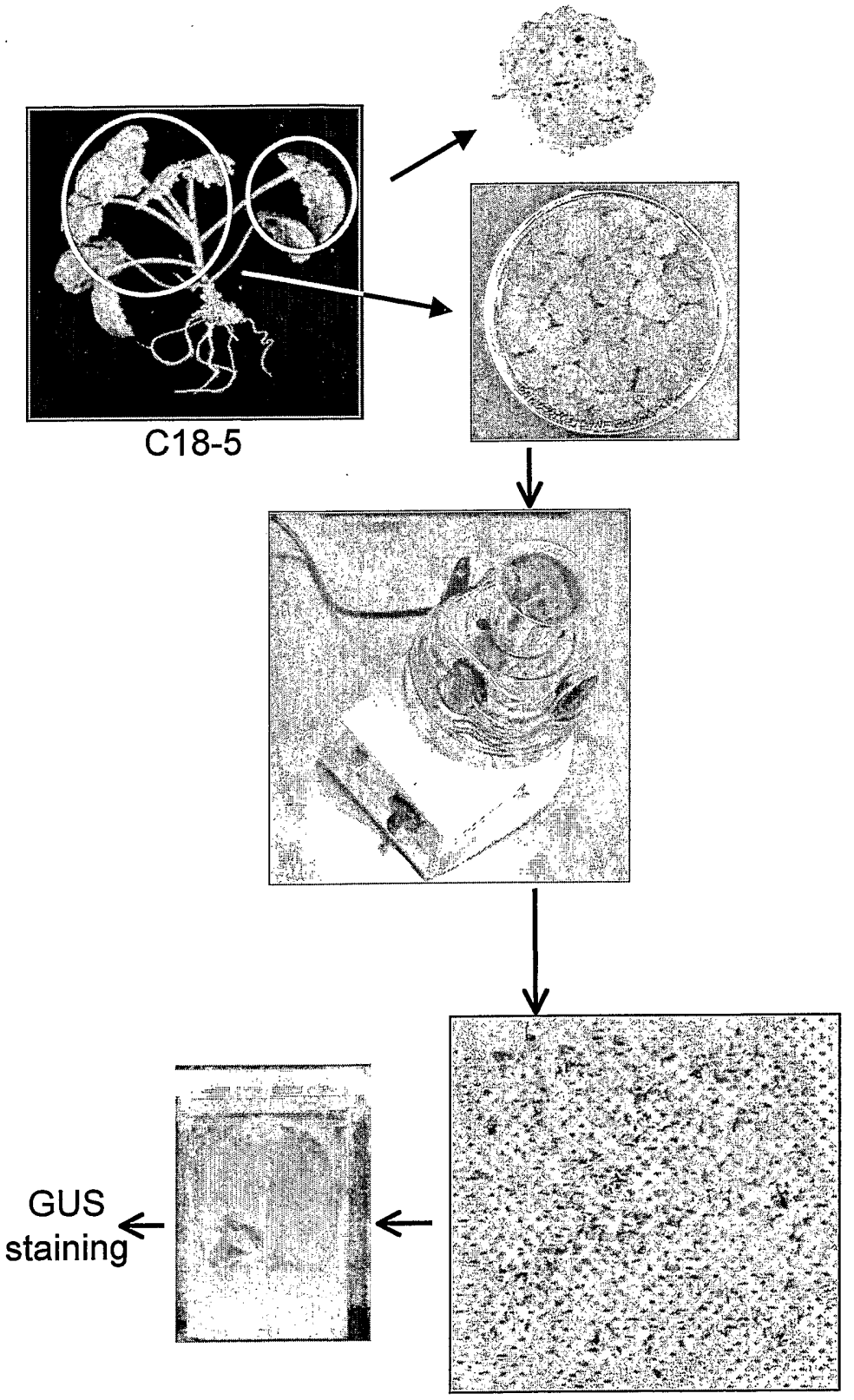


Fig. 28

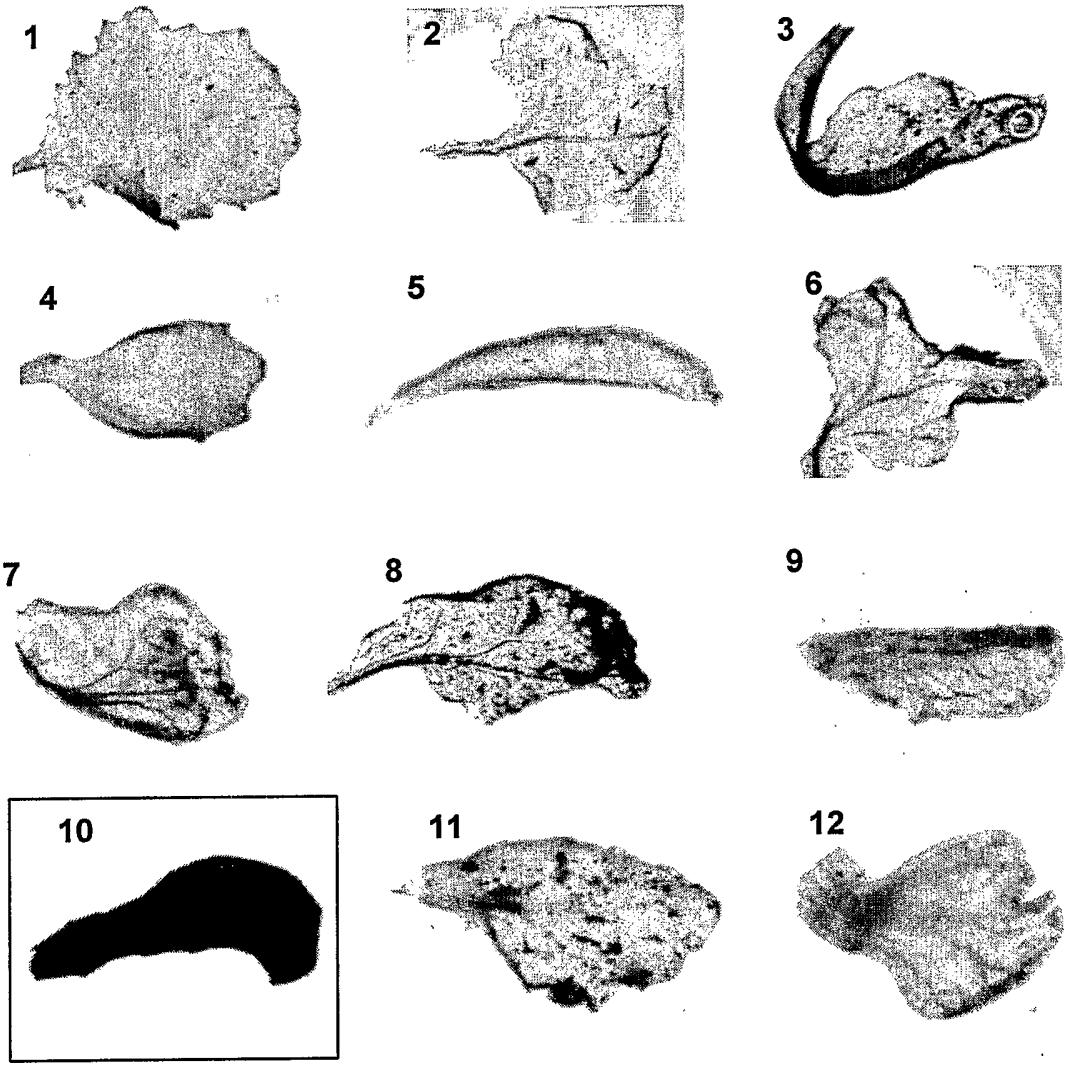


Fig. 29

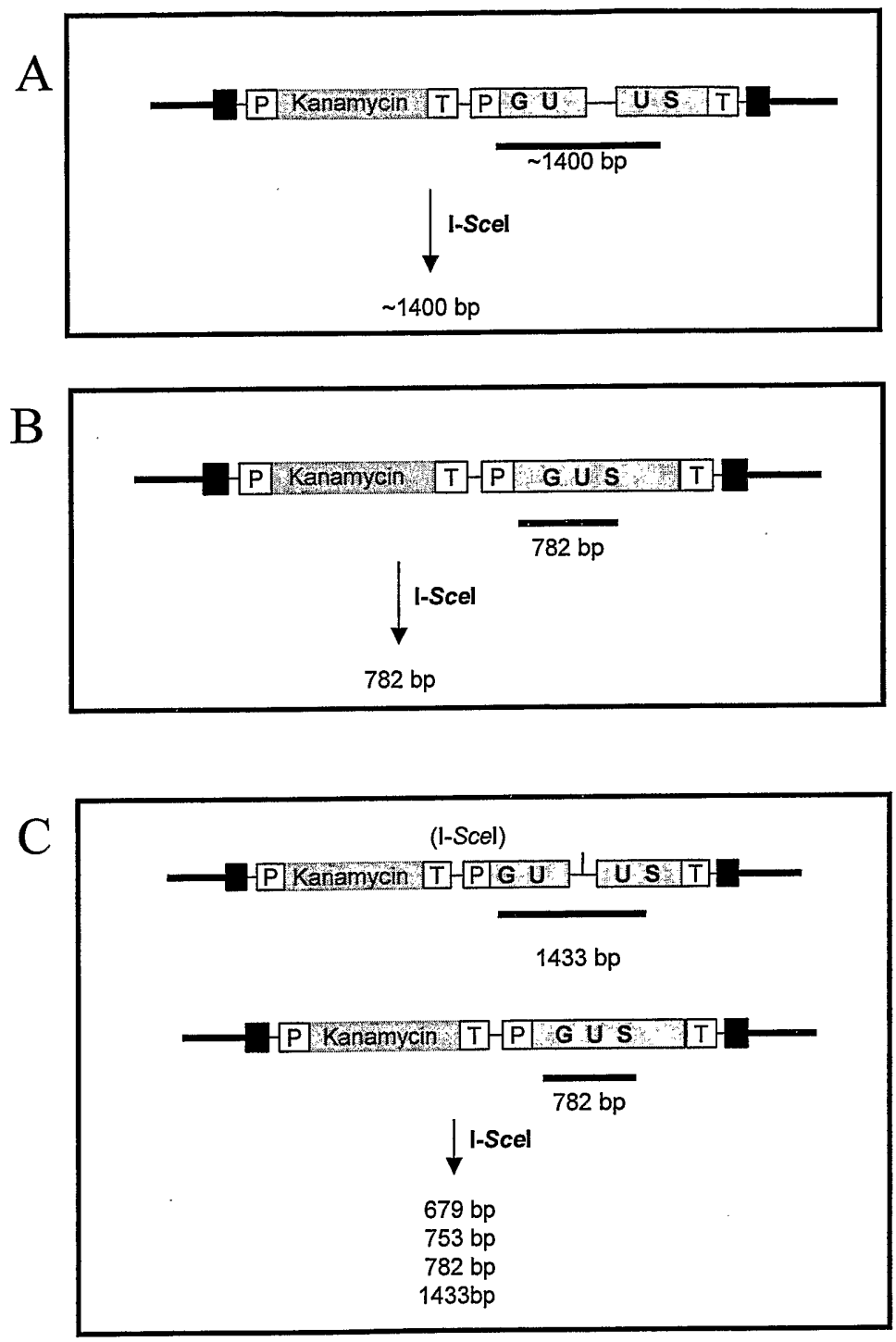

A
P
Kanamycin
T
P
G U
U S
T
~1400 bp
I-SceI
~1400 bp
B
P
Kanamycin
T
P
G U S
T
782 bp
I-SceI
782 bp
C
(I-SceI)
P
Kanamycin
T
P
G U
U S
T
1433 bp
P
Kanamycin
T
P
G U S
T
782 bp
I-SceI
679 bp
753 bp
782 bp
1433bp

Fig. 30-I

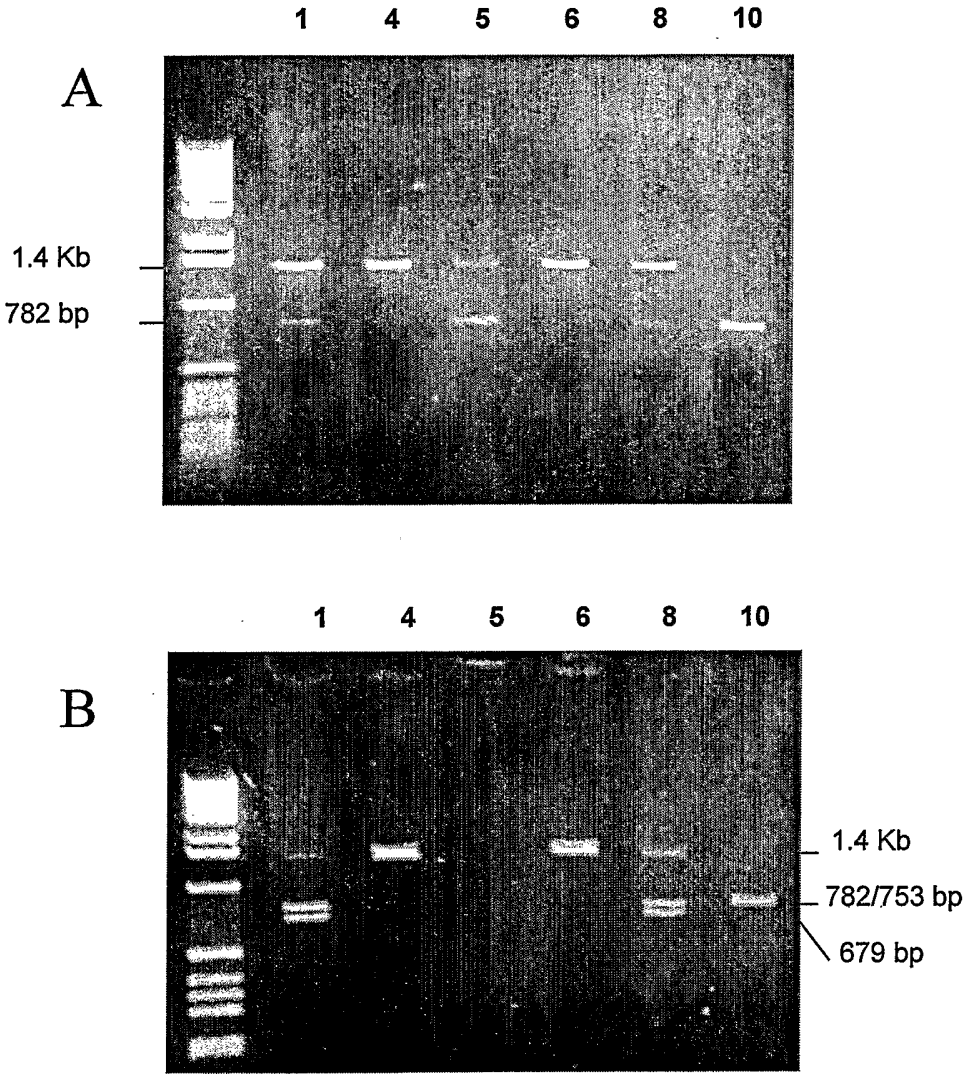
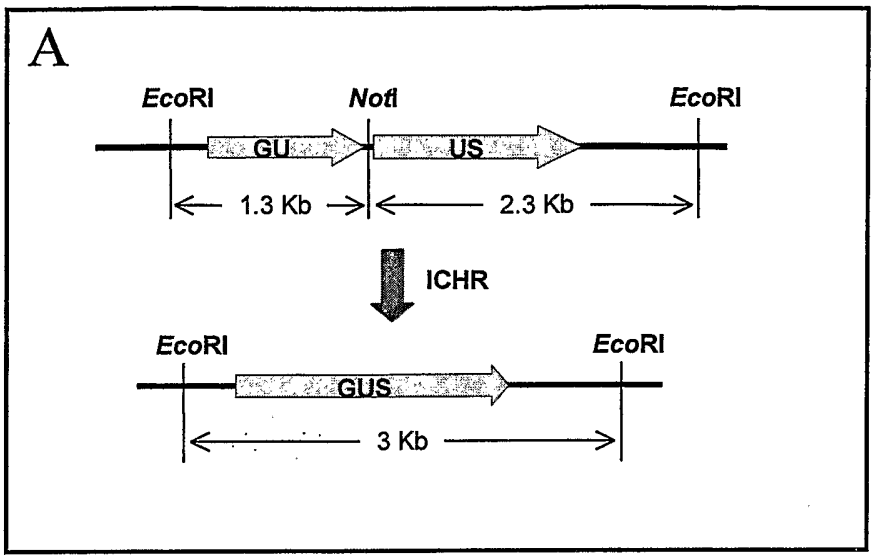


Fig. 30-II


A
EcoRI
NotI
EcoRI
GU
US
1.3 Kb
2.3 Kb
ICHR
EcoRI
EcoRI
GUS
3 Kb

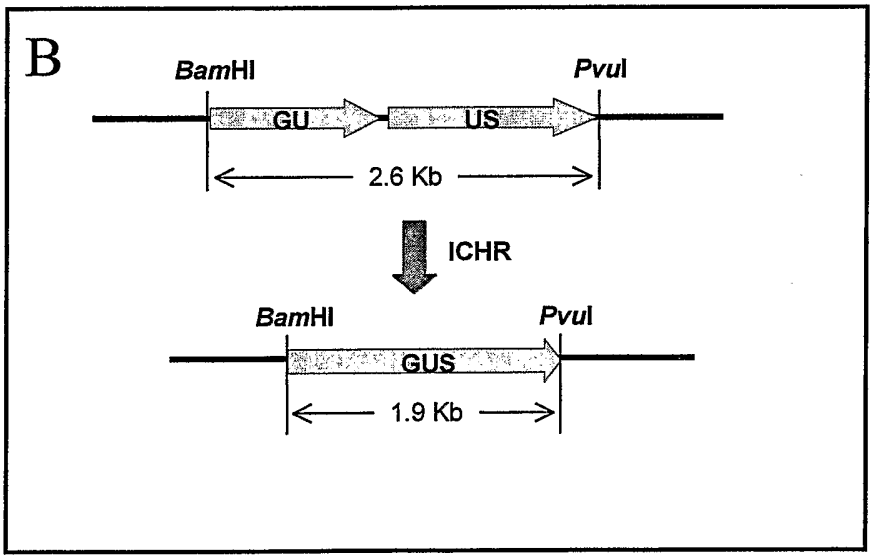

B
BamHI
PvuI
GU
US
2.6 Kb
ICHR
BamHI
PvuI
GUS
1.9 Kb

Fig. 31-I

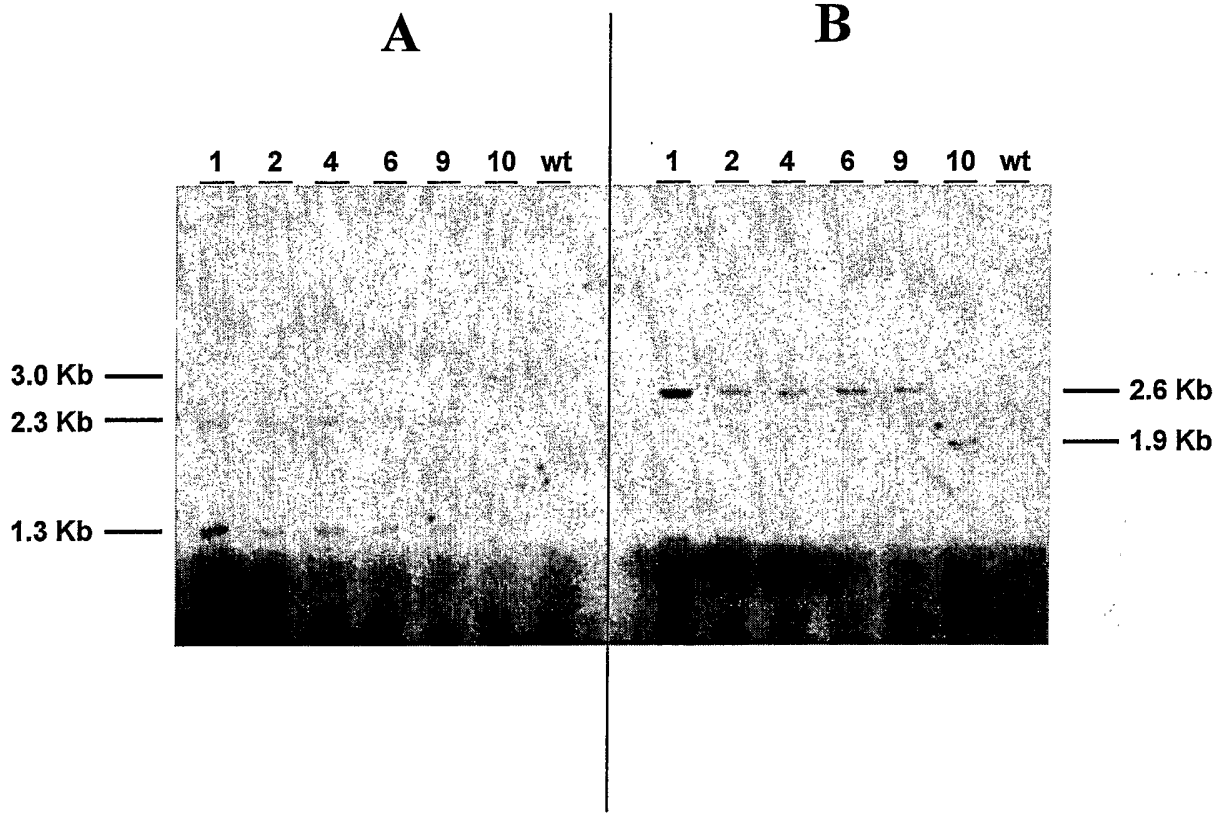


Fig. 31-II

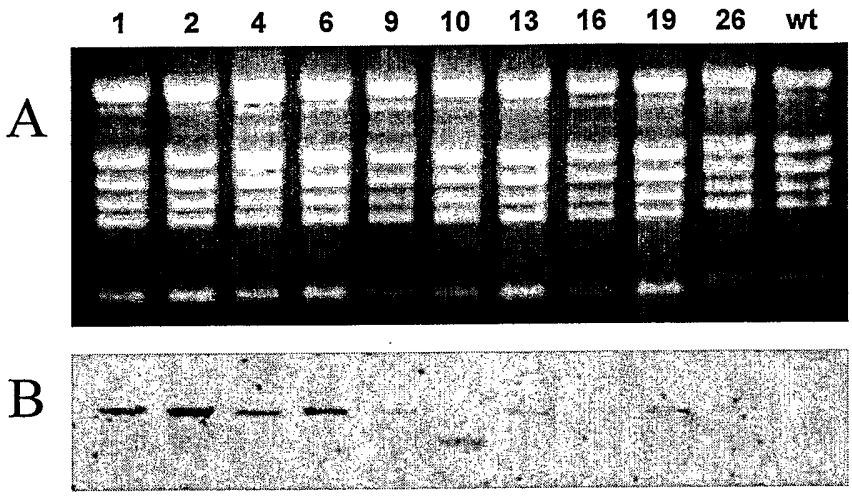


Fig. 32

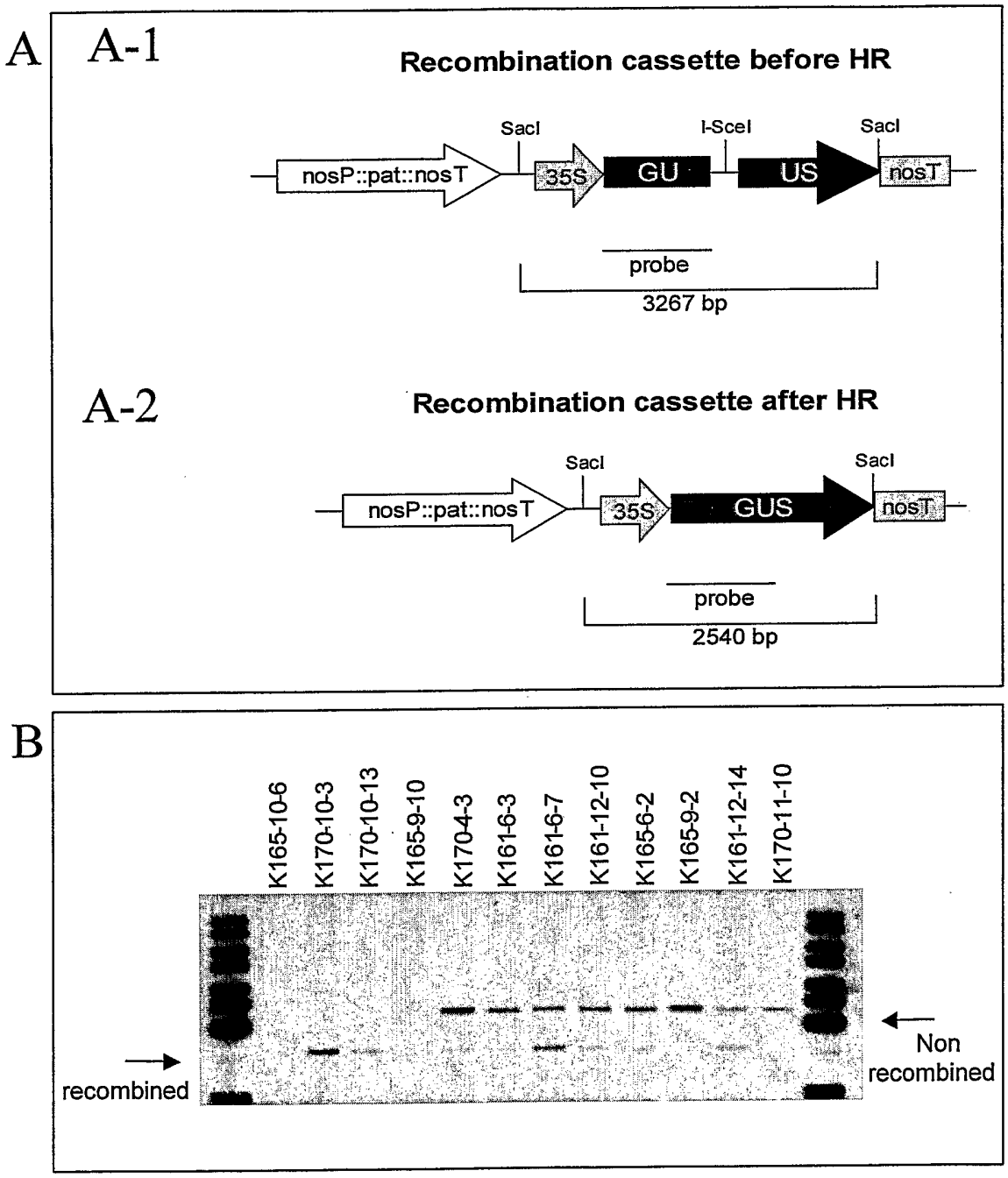
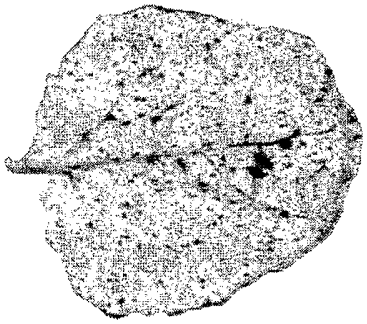

A
A-1
Recombination cassette before HR
SacI
I-SceI
SacI
nosP::pat::nosT
35S
GU
US
nosT
probe
3267 bp
A-2
Recombination cassette after HR
SacI
SacI
nosP::pat::nosT
35S
GUS
nosT
probe
2540 bp
B
K165-10-6
K170-10-3
K170-10-13
K165-9-10
K1704-3
K161-6-3
K161-6-7
K161-12-10
K165-6-2
K165-9-2
K161-12-14
K170-11-10
recombined
Non recombined

Fig. 33

A



B



Fig. 34

SEQUENCE LISTING

<110> BASF Plant Science GmbH
<120> RECOMBINATION CASSETTES AND METHODS FOR SEQUENCE EXCISION IN PLANTS
<130> AE20040748 / PF 55912
<160> 15
<170> PatentIn version 3.3

<210> 1
<211> 7447
<212> DNA
<213> Artificial

<220>
<223> artificial DNA construct (Construct I)

<220>
<221> misc_feature
<222> (6)..(151)
<223> T-DNA Right Border

<220>
<221> misc_feature
<222> (152)..(161)
<223> symbolizing insertion site for various endonuclease expression cassettes

<220>
<221> misc_feature
<222> (1681)..(198)
<223> complementary: nos:P::notII::nosT expression cassettes with nptII ORF from 1324 to 546

<220>
<221> misc_feature
<222> (1694)..(1908)
<223> T-DNA Left Border

<400> 1
ccatggacat acaaatggac gaacggataa acctttcac gccctttaa atatccgatt 60
attctaataa acgctctttt ctcttaggtt tacccgccaa tatatcctgt caaacactga 120
tagtttaaac tgaaggcggg aaacgacaat cnnnnnnnnn ngccgtcgtt ttacaacgac 180
tcagagcttg acaggaggcc cgatctagta acatagatga caccgcgcgc gataatttat 240
cctagtttgc gcgctatatt ttgttttcta tcgcgtatta aatgtataat tgcgggactc 300

taatcataaa aacccatctc ataaataacg tcatgcatta catgttaatt attacatgct 360
taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca ggattcaatc 420
ttaagaaact ttattgccaa atgtttgaac gatcggggat catccgggtc tgtggcggga 480
actccacgaa aatatccgaa cgcagcaaga tctagagctt gggtcccgct cagaagaact 540
cgtcaagaag gcgatagaag gcgatgcgct gcgaatcggg agcggcgata ccgtaaagca 600
cgaggaagcg gtcagcccat tcgccgccaa gctcttcagc aatatcacgg gtagccaacg 660
ctatgtcctg atagcggtcc gccacaccca gccggccaca gtcgatgaat ccagaaaagc 720
ggccattttc caccatgata ttcggcaagc aggcatcgcc atgggtcacg acgagatcct 780
cgccgtcggg catgcgcgcc ttgagcctgg cgaacagttc ggctggcgcg agcccctgat 840
gctcttcgtc cagatcatcc tgatcgacaa gaccggcttc catccgagta cgtgctcgct 900
cgatgcgatg tttcgcttgg tggtcgaatg ggcaggtagc cggatcaagc gtatgcagcc 960
gccgcattgc atcagccatg atggatactt tctcggcagg agcaaggtga gatgacagga 1020
gatcctgccc cggcacttcg cccaatagca gccagtccct tcccgcttca gtgacaacgt 1080
cgagcacagc tgcgcaagga acgcccgtcg tggccagcca cgatagccgc gctgcctcgt 1140
cctgcagttc attcagggca ccggacaggt cggtcttgac aaaaagaacc gggcgccccc 1200
gcgctgacag ccggaacacg gcggcatcag agcagccgat tgtctgttgt gcccagtcat 1260
agccgaatag cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa 1320
tcatgcgaaa cgatccagat ccggtgcaga ttatttggat tgagagtgaa tatgagactc 1380
taattggata ccgaggggaa tttatggaac gtcagtggag cattttgac aagaaatatt 1440
tgctagctga tagtgacctt aggcgacttt tgaacgcgca ataatggttt ctgacgtatg 1500
tgcttagctc attaaactcc agaaacccgc ggctgagtgg ctccttcaac gttgcggttc 1560
tgtcagttcc aaacgtaaaa cggcttgtcc cgcgtcatcg gcgggggtca taacgtgact 1620
cccttaattc tccgctcatg atcagattgt cgtttcccgc cttcagttta aactatcagt 1680
gtttgacagg atcctgcttg gtaataattg tcattagatt gtttttatgc atagatgcac 1740
tcgaaatcag ccaattttag acaagtatca aacggatgtt aattcagtac attaaagacg 1800
tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa tatatcctgc 1860
caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca cgcgttacca 1920
ccacgccggc cggccgcatg gtgttgaccg tgttcgccgg cattgccgag ttcgagcgtt 1980
ccctaatcat cgaccgcacc cggagcgggc gcgaggccgc caaggcccga ggcgtgaagt 2040
ttggcccccg ccctaccctc accccggcac agatcgcgca cgcccgcgag ctgatcgacc 2100
aggaaggccg caccgtgaaa gaggcggctg cactgcttgg cgtgcatcgc tcgaccctgt 2160
accgcgcact tgagcgcagc gaggaagtga cgcccaccga ggccaggcgg cgcggtgcct 2220
tccgtgagga cgcattgacc gaggccgacg ccctggcggc cgccgagaat gaacgccaag 2280
aggaacaagc atgaaaccgc accaggacgg ccaggacgaa ccgtttttca ttaccgaaga 2340
gatcgaggcg gagatgatcg cggccgggta cgtgttcgag ccgcccgcgc acgtctcaac 2400
cgtgcggctg catgaaatcc tggccggttt gtctgatgcc aagctggcgg cctggccggc 2460
cagcttggcc gctgaagaaa ccgagcgccg ccgtctaaaa aggtgatgtg tatttgagta 2520
aaacagcttg cgtcatgcgg tcgctgcgta tatgatgcga tgagtaaata aacaaatacg 2580
caaggggaac gcatgaaggt tatcgctgta cttaaccaga aaggcgggtc aggcaagacg 2640
accatcgcaa cccatctagc ccgcgccctg caactcgccg gggccgatgt tctgttagtc 2700
gattccgatc cccagggcag tgcccgcgat tgggcggccg tgcgggaaga tcaaccgcta 2760
accgttgtcg gcatcgaccg cccgacgatt gaccgcgacg tgaaggccat cggccggcgc 2820
gacttcgtag tgatcgacgg agcgccccag gcggcggact tggctgtgtc cgcgatcaag 2880
gcagccgact tcgtgctgat tccggtgcag ccaagccctt acgacatatg ggccaccgcc 2940
gacctggtgg agctggttaa gcagcgcatt gaggtcacgg atggaaggct acaagcggcc 3000

tttgtcgtgt cgcgggcgat caaaggcacg cgcatcggcg gtgaggttgc cgaggcgctg 3060
gccgggtacg agctgcccat tcttgagtcc cgtatcacgc agcgcgtgag ctacccaggc 3120
actgccgccg ccggcacaac cgttcttgaa tcagaacccg agggcgacgc tgcccgcgag 3180
gtccaggcgc tggccgctga aattaaatca aaactcattt gagttaatga ggtaaagaga 3240
aaatgagcaa aagcacaaac acgctaagtg ccggccgtcc gagcgcacgc agcagcaagg 3300
ctgcaacgtt ggccagcctg gcagacacgc cagccatgaa gcgggtcaac tttcagttgc 3360
cggcggagga tcacaccaag ctgaagatgt acgcggtacg ccaaggcaag accattaccg 3420
agctgctatc tgaatacatc gcgcagctac cagagtaaat gagcaaatga ataaatgagt 3480
agatgaattt tagcggctaa aggaggcggc atggaaaatc aagaacaacc aggcaccgac 3540
gccgtggaat gccccatgtg tggaggaacg ggcggttggc caggcgtaag cggctgggtt 3600
gtctgccggc cctgcaatgg cactggaacc cccaagcccg aggaatcggc gtgagcggtc 3660
gcaaaccatc cggcccggta caaatcggcg cggcgctggg tgatgacctg gtggagaagt 3720
tgaaggccgc gcaggccgcc cagcggcaac gcatcgaggc agaagcacgc cccggtgaat 3780
cgtggcaagc ggccgctgat cgaatccgca aagaatcccg gcaaccgccg gcagccggtg 3840
cgccgtcgat taggaagccg cccaagggcg acgagcaacc agatttttc gttccgatgc 3900
tctatgacgt gggcacccgc gatagtcgca gcatcatgga cgtggccgtt ttccgtctgt 3960
cgaagcgtga ccgacgagct ggcgaggtga tccgctacga gcttccagac gggcacgtag 4020
aggtttccgc agggccggcc ggcatggcca gtgtgtggga ttacgacctg gtactgatgg 4080
cggtttccca tctaaccgaa tccatgaacc gataccggga agggaaggga gacaagcccg 4140
gccgcgtgtt ccgtccacac gttgcggacg tactcaagtt ctgccggcga gccgatggcg 4200
gaaagcagaa agacgacctg gtagaaacct gcattcggtt aaacaccacg cacgttgcca 4260
tgcagcgtac gaagaaggcc aagaacggcc gcctggtgac ggtatccgag ggtgaagcct 4320
tgattagccg ctacaagatc gtaaagagcg aaaccgggcg gccggagtac atcgagatcg 4380
agctagctga ttggatgtac cgcgagatca cagaaggcaa gaacccggac gtgctgacgg 4440
ttcaccccga ttacttttttg atcgatcccg gcatcggccg ttttctctac cgcctggcac 4500
gccgcgccgc aggcaaggca gaagccagat ggttgttcaa gacgatctac gaacgcagtg 4560
gcagcgccgg agagttcaag aagttctgtt tcaccgtgcg caagctgatc gggtcaaatg 4620
acctgccgga gtacgatttg aaggaggagg cggggcaggc tggcccgatc ctagtcatgc 4680
gctaccgcaa cctgatcgag ggcgaagcat ccgccggttc ctaatgtacg gagcagatgc 4740
tagggcaaat tgccctagca ggggaaaaag gtcgaaaagg tctctttcct gtggatagca 4800
cgtacattgg gaacccaaag ccgtacattg ggaaccggaa cccgtacatt gggaacccaa 4860
agccgtacat tgggaaccgg tcacacatgt aagtgactga tataaaagag aaaaaaggcg 4920
atttttccgc ctaaaactct ttaaaactta ttaaaactct taaaacccgc ctggcctgtg 4980
cataactgtc tggccagcgc acagccgaag agctgcaaaa agcgcctacc cttcggtcgc 5040
tgcgctccct acgccccgcc gcttcgcgtc ggcctatcgc ggccgctggc cgctcaaaaa 5100
tggctggcct acggccaggc aatctaccag ggcgcggaca agccgcgccg tcgccactcg 5160
accgccggcg cccacatcaa ggcaccctgc ctcgcgcgtt tcggtgatga cggtgaaaac 5220
ctctgacaca tgcagctccc ggagacggtc acagcttgtc tgtaagcgga tgccgggagc 5280
agacaagccc gtcagggcgc gtcagcgggt gttggcgggt gtcggggcgc agccatgacc 5340
cagtcacgta gcgatagcgg agtgtatact ggcttaacta tgcggcatca gagcagattg 5400
tactgagagt gcaccatatg cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc 5460
gcatcaggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 5520
ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 5580
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 5640
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 5700

caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa 5760
gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 5820
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 5880
aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 5940
ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 6000
cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 6060
tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 6120
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 6180
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 6240
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 6300
aagggatttt ggtcatgcat gatatatctc ccaatttgtg tagggcttat tatgcacgct 6360
taaaaataat aaaagcagac ttgacctgat agtttggctg tgagcaatta tgtgcttagt 6420
gcatctaacg cttgagttaa gccgcgccgc gaagcggcgt cggcttgaac gaatttctag 6480
ctagacatta tttgccgact accttggtga tctcgccttt cacgtagtgg acaaattctt 6540
ccaactgatc tgcgcgcgag gccaagcgat cttcttcttg tccaagataa gcctgtctag 6600
cttcaagtat gacgggctga tactgggccg gcaggcgctc cattgcccag tcggcagcga 6660
catccttcgg cgcgattttg ccggttactg cgctgtacca aatgcgggac aacgtaagca 6720
ctacatttcg ctcatcgcca gcccagtcgg gcggcgagtt ccatagcgtt aaggtttcat 6780
ttagcgcctc aaatagatcc tgttcaggaa ccggatcaaa gagttcctcc gccgctggac 6840
ctaccaaggc aacgctatgt tctcttgctt ttgtcagcaa gatagccaga tcaatgtcga 6900
tcgtggctgg ctcgaagata cctgcaagaa tgtcattgcg ctgccattct ccaaattgca 6960
gttcgcgctt agctggataa cgccacggaa tgatgtcgtc gtgcacaaca atggtgactt 7020
ctacagcgcg gagaatctcg ctctctccag gggaagccga agtttccaaa aggtcgttga 7080
tcaaagctcg ccgcgttgtt tcatcaagcc ttacggtcac cgtaaccagc aaatcaatat 7140
cactgtgtgg cttcaggccg ccatccactg cggagccgta caaatgtacg gccagcaacg 7200
tcggttcgag atggcgctcg atgacgccaa ctacctctga tagttgagtc gatacttcgg 7260
cgatcaccgc ttcccccatg atgtttaact ttgttttagg gcgactgccc tgctgcgtaa 7320
catcgttgct gctccataac atcaaacatc gacccacggc gtaacgcgct tgctgcttgg 7380
atgcccgagg catagactgt accccaaaaa aacagtcata acaagccatg aaaaccgcca 7440
ctgcgtt 7447

<210> 2
<211> 2835
<212> DNA
<213> artificial DNA construct (insert of pCB603-100)

<220>
<223> artificial DNA construct (insert of pCB603-100)

<220>
<221> terminator
<222> (43)..(249)
<223> complentary: OCS terminator

<220>

<221> misc_feature
<222> (303)..(1010)
<223> complementary: sequence encoding for ORF of I-SceI

<220>
<221> promoter
<222> (1014)..(2790)
<223> complementary: fragment of AtSERK1 promoter incl. 5'-UTR

<400> 2

agatctagta ggaaacagct atgaccatga ttacgccaag cttggacaat cagtaaattg 60
aacggagaat attattcata aaaatacgat agtaacgggt gatatattca ttagaatgaa 120
ccgaaaccgg cggtaaggat ctgagctaca catgctcagg ttttttacaa cgtgcacaac 180
agaattgaaa gcaaatatca tgcgatcata ggcgtctcgc atatctcatt aaagcagggc 240
atgcctgcag gtcgagcggc cgccagtgtg atggatatct gcagaattcg ccctttttaga 300
tcttatttca ggaaagtttc ggaggagata gtgttcggca gtttgtacat catctgcggg 360
atcaggtacg gtttgatcag gttgtagaag atcaggtaag acatagaatc gatgtagatg 420
atcggtttgt ttttgttgat ttttacgtaa cagttcagtt ggaatttgtt acgcagaccc 480
ttaaccaggt attctacttc ttcgaaagtg aaagactggg tgttcagtac gatcgatttg 540
ttggtagagt ttttgttgta atcccattta ccaccatcat ccatgaacca gtatgccaga 600
gacatcgggg tcaggtagtt ttcaaccagg ttgttcggga tggtttttttt gttgttaacg 660
atgaacaggc tagccagttt gttgaaagct tggtgtttga aagtctgggc gccccaggtg 720
attaccaggt tacccaggtg gttaacacgt tctttttttgt gcggcgggga cagtacccac 780
tgatcgtaca gcagacatac gtggtccatg tatgctttgt ttttccactc gaactgcata 840
cagtaggttt taccttcatc acgagaacgg atgtaagcat cacccaggat cagaccgata 900
cctgcttcga actgttcgat gttcagttcg atcagctggg atttgtattc tttcagcagt 960
ttagagttcg gacccaggtt cattacctgg ttttttttga tgtttttcat atgaaacaac 1020
aatgctaaat ttcgaatctt gatttgatct cttcagatcc gagaaaaccc taacttccga 1080
aaatcaccac cattaacact ttacagagac aacctcaaca caaactcact catttttttcg 1140
ataacccttt taaaaaccca ttagcattaa aacaaacaaa gagtaaaaaa taaaatcaaa 1200
atcctaattc tccggcgttc tcgtcgtcgt ctttcctcaa tcaaaccaca ctctctctca 1260
cacacacact gtaaaacaaa agttgaataa agatgaaact tttttttttta aacaaaaaaa 1320
taagagcttt ttggggaaaa tctacgtgaa gtaataataa caaaagagct ttttattaag 1380
ccgcaacaac aagcctgcca gagaaactat gattagtgtg ttgttgttgt tcctctcttt 1440
ttcttcttat gaataaaata gctttttcgt tgatatatta tttaatattt ttaaagatct 1500
caacgggctt tatttatggg catgcagtgg tgtagaggag gaagacgagg aatataagct 1560
gcccgtaacg gtgctgctgc ccgaactgac cggtcaaaaa aacgggttcg ggggagagaa 1620
gagatggttg aaagtgacat cgctttgggt ttggggtaaa gtagttgaaa gctcttaagg 1680
ctatgcaccc gtttctcctg gtcaaacctc tctctgttgg gtccctttttt tctttaagta 1740
gaaaaacctt attctaaatt tttaatgcaa gtaagattcg taatcccata tttatgaaag 1800
taaagttaaa atgttaaaga ggttcactat catgtgagca atctcatacc aaaaatgtga 1860
ttttttttttt acaatggttg tgaattttgt gattgataaa ccatgcaaca ataaagatat 1920
tagtgtttgc tgatattctt tctaaaccca actttggcag agatgagatt gaaagatttg 1980
agttttgtca acaaaaaatt tggtggaaga aactgtagag acggcaaaga caaaatttgg 2040
ccaagttcca ctgtggaagt agtatggcgt gcccgatatg actttactac tgtttcatca 2100

caataccatt aaacagatac ttgtcttatc atttaaaaaa aaaaaaaaac tgcccataca 2160
atttgagtag aggtaaaaag taaaataaaa agcagattac atttgttacc ctttaaatat 2220
cacacattat gcttacccca tgtggtggat ggactcttta ttttcattat gatcaaaagg 2280
tttctagaac aaaatttaat taaagtaaaa aagaaacagt catcttcaga aaatatttca 2340
agataatgct tacaattatt tatttagata ttgtgagtgg ctaaatcagc tgccaatgag 2400
tgaatttaca aacttttta taatccccaa catttttat attcatgata aaaatatgct 2460
tcaaattcaa aatttcacaa atcagaaata attctaaaaa caataagagc taccttgtta 2520
tgaaagagaa acgaacttta agaagatttt atcaaaatta aatgaaacag tagaattcaa 2580
atttgcgaca aaaaatacgc ctcctaaata caacgggata atatcctagt gcaagtggcc 2640
accgagcgtt tcaactttca aatccaggaa taaaggtaag aaaaagaaca aaacaagaga 2700
aaacacagaa tttacagagg aaagatacca tgcaaagaaa aagaagacat ggacgcaaac 2760
gccataattt gggaattcta tattaagggc gaattccagc acactggcgg ccgttactag 2820
tggatcaatt cactg 2835

<210> 3
<211> 2187
<212> DNA
<213> artificial DNA construct (insert of pCB622-3)

<220>
<223> artificial DNA construct (insert of pCB622-3)

<220>
<221> terminator
<222> (43)..(249)
<223> complementary: OCS terminator

<220>
<221> misc_feature
<222> (304)..(1011)
<223> complementary: sequence encoding ORF of I-SceI

<220>
<221> promoter
<222> (304)..(1011)
<223> complementary: fragment of Atcyc1 promoter incl. 5'-UTR

<400> 3
agatctagta ggaaacagct atgaccatga ttacgccaag cttggacaat cagtaaattg 60
aacggagaat attattcata aaaatacgat agtaacgggt gatatattca ttagaatgaa 120
ccgaaaccgg cggtaaggat ctgagctaca catgctcagg ttttttacaa cgtgcacaac 180
agaattgaaa gcaaatatca tgcgatcata ggcgtctcgc atatctcatt aaagcagggc 240
atgcctgcat gctcgagcgg ccgccagtgt gatggatatc tgcagaattc gccctttag 300
atcttatttc aggaaagttt cggaggagat agtgttcggc agtttgtaca tcatctgcgg 360
gatcaggtac ggtttgatca ggttgtagaa gatcaggtaa gacatagaat cgatgtagat 420

gatcggtttg tttttgttga tttttacgta acagttcagt tggaatttgt tacgcagacc 480
cttaaccagg tattctactt cttcgaaagt gaaagactgg gtgttcagta cgatcgattt 540
gttggtagag tttttgttgt aatcccattt accaccatca tccatgaacc agtatgccag 600
agacatcggg gtcaggtagt tttcaaccag gttgttcggg atggtttttt tgttgttaac 660
gatgaacagg ctagccagtt tgttgaaagc ttggtgtttg aaagtctggg cgccccaggt 720
gattaccagg ttacccaggt ggttaacacg ttctttttttg tgcggcgggg acagtaccca 780
ctgatcgtac agcagacata cgtggtccat gtatgctttg tttttccact cgaactgcat 840
acagtaggtt ttaccttcat cacgagaacg gatgtaagca tcacccagga tcagaccgat 900
acctgcttcg aactgttcga tgttcagttc gatcagctgg gatttgtatt ctttcagcag 960
tttagagttc ggacccaggt tcattacctg gtttttttttg atgtttttca tatgagtgtt 1020
ctcttctctt tctctcagac taaaatctca ggtttgtagt agaaaaagtc taagtagtgg 1080
gggcgtcttc tatataggag gcttcttgag gaagatgagc cgttttataa gaaataaaaa 1140
caaaattaaa gaaaataaat gaaccccaac ggtctgattg tttttgaaaa attagaggtc 1200
gtgggcctgt tacgggcctt acggctgtcg agacttacaa aatcgggctt tttcactatt 1260
gggattttga agtcacaacg gcccacttat tctttttatt gactgtgata taggaaaaaa 1320
gaaaagaaaa aaagatttgt gtgaattaga ttgctagtta aatccggtta tctagatgtt 1380
gatttctgtt tggtttacat ggttttgaaa taaaaataga tttttgttcg gtttgcatta 1440
gaattagtac ggttttgaga aaaaaaaaca aaagaaaaag gtcactcgag gctgttttgg 1500
aaacaaagcg agtagatgta aaatatgcat gattttgtgg gtttacccct acagaaatcg 1560
acgacacgcg ctcaacgcga ggttggtgcg agtgggcgtg atctggagcg cgtgacgcga 1620
agttggtgcg agtgggcgtt ttctggtgcg ttctgattgg ctaaaaaaat gaaaatagta 1680
gtgatgctcc tcctctatta ttttgttttt cgtcattctt cagttatgaa agttgtgtgg 1740
ttttaagatg caattgagtt tattctgagg ttttttggtt attttgaaat cattgattaa 1800
tgtcacaaaa aaaaagtaga aactagattt tgagatatta aaattttcat tttcattttg 1860
ttctgcaatc attgaaaact agattcagct gagccattga agaaatcaaa gtgagttgaa 1920
tatttttcta aaaatgtgaa aattggtaaa agttggtcct cattttgtta gtttaatgtt 1980
ttcttgactt taggagatga actttcaact gcattgcaga tcccatcaca cattaatctc 2040
actctcttct gcacattgca tggtgtggtg tagatatggt tcccacacaa tttgtcatat 2100
tcatttacca caaagttcat gaatcatgtt tgaagtgaat ccttttttctt tcaaatttag 2160
tcatctcact agtggatcaa ttcactg 2187

<210> 4
<211> 2252
<212> DNA
<213> artificial DNA construct (insert of pCB653-37)

<220>
<223> artificial DNA construct (insert of pCB653-37)

<220>
<221> terminator
<222> (43)..(249)
<223> complementary: OCS terminator

<220>

<221> misc_feature
<222> (304)..(1011)
<223> complementary: sequence encoding ORF of I-SceI

<220>
<221> promoter
<222> (1013)..(2201)
<223> complementary: fragment of erecta promoter incl. 5'-UTR

<400> 4

agatctagta ggaaacagct atgaccatga ttacgccaag cttggacaat cagtaaattg 60
aacggagaat attattcata aaaatacgat agtaacgggt gatatattca ttagaatgaa 120
ccgaaaccgg cggtaaggat ctgagctaca catgctcagg ttttttacaa cgtgcacaac 180
agaattgaaa gcaaatatca tgcgatcata ggcgtctcgc atatctcatt aaagcagggc 240
atgcctgcat gctcgagcgg ccgccagtgt gatggatatc tgcagaattc gccctttttag 300
atcttatttc aggaaagttt cggaggagat agtgttcggc agtttgtaca tcatctgcgg 360
gatcaggtac ggtttgatca ggttgtagaa gatcaggtaa gacatagaat cgatgtagat 420
gatcggtttg tttttgttga tttttacgta acagttcagt tggaatttgt tacgcagacc 480
cttaaccagg tattctactt cttcgaaagt gaaagactgg gtgttcagta cgatcgattt 540
gttggtagag tttttgttgt aatcccattt accaccatca tccatgaacc agtatgccag 600
agacatcggg gtcaggtagt tttcaaccag gttgttcggg atggtttttt tgttgttaac 660
gatgaacagg ctagccagtt tgttgaaagc ttggtgtttg aaagtctggg cgccccaggt 720
gattaccagg ttacccaggt ggttaacacg ttctttttttg tgcggcgggg acagtaccca 780
ctgatcgtac agcagacata cgtggtccat gtatgctttg tttttccact cgaactgcat 840
acagtaggtt ttaccttcat cacgagaacg gatgtaagca tcacccagga tcagaccgat 900
acctgcttcg aactgttcga tgttcagttc gatcagctgg gatttgtatt ctttcagcag 960
tttagagttc ggacccaggt tcattacctg gttttttttg atgttttttca tatgcacaca 1020
cagtcttaaa acgactgcgt ttttagatat actttaaaag ctttcaagtc tccatgaaga 1080
agaaacagag gatagagaag aagaagagaa ccgtattatg taccgacatt agatttttttt 1140
tctttctttt attatgattt ctacagaaat ttgtttttgt tttttcactt taaccgagaa 1200
aaaatgattt ttatcgttcc atcacatcaa attctggcgg aaatttacag tggtttcatt 1260
gagtgctcag acgacagaaa tacagagtaa gatagagatg gagagagaaa gagagagagt 1320
ctttaaagag aggagattgt acaaatgtag ggacctgttg tgtaagagag agagatgaag 1380
ggaagaagaa ttatttgggg aggataaagt taaaaacccc atcatctttt atctctctgc 1440
cagagttctg aagctttgca gaagagagag atagagtgag tgagagggag tgattttgtt 1500
tctctctctt gaagggtaaa acggtcataa aatgttgaag aagaaggaga agaagaacat 1560
gttaatggca gtggtgaata aaggaggtcg ctttcgttgc ccaatactta cgagtctttg 1620
tctctctgtg gcttgtttca cagatgaccc atatcaataa atccatctta tgggtacttt 1680
tgtaattatg tatttttata tgattttagt gtaatttttt tgtttgctta acatgttctt 1740
cttcatcttc ttacttcaat tactaattta aattcaggag tttacaaaaa gtttggttca 1800
gttttctctt ctttaattat tagtatagtt gttgtgtttt tgtcaacgtt ttttagtata 1860
gtgaacgtgt gatttctaag tgtttttgtt tattactaaa tttagaaata tgcatgccaa 1920
tgatttttgg tttagcatca acacaaaact taataaaata tgattggttt taggatgtta 1980
accggtttgg ttggctccgg tcggtttgtc acaaaattaa tattttgaag ctttgagaac 2040
gaagtggact aaaactaaaa gggtcacagg cccaacaaaa gaaggaagaa cttaaaaaaa 2100

ataataattc cttttcattt taacttctct ttttttttt gttaaatcat ttatttgttt 2160
caataaaaat acacataata ttatcttttc tcacatagta taagggcgaa ttccagcaca 2220
ctggcggccg ttactagtgg atcaattcac tg 2252

<210> 5
<211> 2994
<212> DNA
<213> artificial DNA construct (insert of pCB652-124)

<220>
<223> artificial DNA construct (insert of pCB652-124)

<220>
<221> terminator
<222> (43)..(249)
<223> complementary: OCS terminator

<220>
<221> misc_feature
<222> (304)..(1011)
<223> complementary: sequence encoding ORF of I-SceI

<220>
<221> promoter
<222> (1059)..(2956)
<223> complementary: fragment of promoter invGF (incl. 5'-UTR)

<400> 5
agatctagta ggaaacagct atgaccatga ttacgccaag cttggacaat cagtaaattg 60
aacggagaat attattcata aaaatacgat agtaacgggt gatatattca ttagaatgaa 120
ccgaaaccgg cggtaaggat ctgagctaca catgctcagg ttttttacaa cgtgcacaac 180
agaattgaaa gcaaatatca tgcgatcata ggcgtctcgc atatctcatt aaagcagggc 240
atgcctgcat gctcgagcgg ccgccagtgt gatggatatc tgcagaattc gccctttag 300
atcttatttc aggaaagttt cggaggagat agtgttcggc agtttgtaca tcatctgcgg 360
gatcaggtac ggtttgatca ggttgtagaa gatcaggtaa gacatagaat cgatgtagat 420
gatcggtttg tttttgttga tttttacgta acagttcagt tggaatttgt tacgcagacc 480
cttaaccagg tattctactt cttcgaaagt gaaagactgg gtgttcagta cgatcgattt 540
gttggtagag tttttgttgt aatcccattt accaccatca tccatgaacc agtatgccag 600
agacatcggg gtcaggtagt tttcaaccag gttgttcggg atggtttttt tgttgttaac 660
gatgaacagg ctagccagtt tgttgaaagc ttggtgtttg aaagtctggg cgccccaggt 720
gattaccagg ttacccaggt ggttaacacg ttctttttttg tgcggcgggg acagtaccca 780
ctgatcgtac agcagacata cgtggtccat gtatgctttg tttttccact cgaactgcat 840
acagtaggtt ttaccttcat cacgagaacg gatgtaagca tcacccagga tcagaccgat 900
acctgcttcg aactgttcga tgttcagttc gatcagctgg gatttgtatt ctttcagcag 960
tttagagttc ggacccaggt tcattacctg gttttttttg atgtttttca tatggtcgac 1020

taaagggcga attccagcac actggcggcc gttactagag ttttttttt ttcaataatt 1080
gggctttgac agatagtggt aattaatcta aataattttt ttttattttt tttggatagg 1140
gagaaagaaa cttgggaatg aagaataaat gaaaattatg aagagtattt atagtgaaaa 1200
aaatgggtag agtttgtttt tgagtgaaag atggagagtt aatttaggat tttttttaaa 1260
aataaattaa atggattgga tgattatttt ttatttgttt ggataatatt tatctaaaat 1320
tttggataat ttgttggtta ttgttacatt tcatgagtct gttttactaa taatttttg 1380
ttagaacaag tcttttgcca tttaatttgc tattccgttc tttaatgttt tgacttttga 1440
gtgtgttaaa aaaaaaggaa ttcgcgaaaa ataatgcaaa atttgggatt ttccaattta 1500
cataaataag ataaaggggt aaatttgacc ttaaataaaa aatgaaaaag aaaagaaaat 1560
atgatcacaa ataatgttta gtggcgacta aatataatcg ccacaaaagg ttatcaccgg 1620
tcgtggcgac ttttttgagt cgccacttga aatgttgttc tttgttacga ttgaggaagt 1680
agtctccacg aaggcgtaac aaaaatatga tttcttgtta ccttgaaaat tataagtttt 1740
ttccacatta tttatatatg gcgtgaaatt tggtcatata tactatatcc ctagggcgaa 1800
aagctaaaag aatagtaaga aacgatattg tgcatctctt aaaattaaaa attaactatc 1860
aagtgttttt gaacagattt gattatgaat ggaagtatca acttcatcaa gttactctat 1920
tttagccatc tttattcatt tgtgattttt ttttccgag gctcatcact cctcttcctt 1980
ctttgctttt atagttcgtc attaattaag tattcactat tcttatttgt aatattatcc 2040
tagagtgtgt tataaattat acaataatta taacttactt agttctttcg tcttctcata 2100
ttttacatgc ttgttaatag ggatttggtg cctcaataga gttagtaagg taaaaatttg 2160
tccttccgct tacttgcaga ggcggatcta caagtatctt gagaaactag taaaaataga 2220
gaatataaaa aagagtgcac ctttagaagg cattgatcta gtacgttgtt ggaagtaagg 2280
gagtccaccc ctacgaccta gagtcgattt cggccaacaa cccatttttt actcatttta 2340
ttacaagtta aagagtttat ttgctattaa atatgacatt attcgttagc cgtcattaat 2400
atttttacta atgcttaaat taaaataata ttgatgcacc caaactcttc aaatcctgaa 2460
ttcgcctcta cttacttgta ctcttctcaa agaacttctt attatatttg aataaaataa 2520
gtcacataac aacttgccta atattaccaa atattagttc tacataaaac cacacattta 2580
ttgtgtaaca aaacaaaaag gaaagctaac aaatttcttt cttagtatat atatggtact 2640
agacaagtcc atatacacac tctttgcctt ctaaattcac caaaataaat taagagacaa 2700
ttatagagga ttcttgagga catcatctac tgttaaattt gacacaattt ttgtctcttt 2760
agacatatat ggttgtagac ataattcctt tttaaaaaaa aatatttagt gcatcttagg 2820
tacatccatg ctccaagcat taagagtctc aattgtgatt gtctcagatc cattattgaa 2880
gacaaataaa tgtgcattat tatgaatcgc taacgttgga tacaccctcg atgttatgca 2940
tgtaagggcg aattccagca cactggcggc cgttactagt ggatcaattc actg 2994

<210> 6
<211> 2586
<212> DNA
<213> artificial DNA construct (insert of pCB632-17)

<220>
<223> artificial DNA construct (insert of pCB632-17)

<220>
<221> promoter
<222> (90)..(1407)

<223> sTPT promoter

<220>
<221> CDS
<222> (1491)..(2198)
<223> coding for I-SceI

<220>
<221> terminator
<222> (2282)..(2516)
<223> CatDpA terminator

<400> 6

```
agatctagta ggaaacagct atgaccatga ttacgccaag cttgcatgcc tgcaggtcga     60
ctctagagga tccccaagtc gacggatcct gatagcttat actcaaattc aacaagttat    120
atataaatgt atagatacta caatatcatt aacaaaagtc accttaaata aatacacata    180
tcttttatgt tctctattgt tttgcgtacg ctaacacaat ttctcatatg caaaaggatg    240
aatgagtaac aaattacctc ataagaacaa tcatctttgc ttacatacta atacaataat    300
cactcaatca accaataaca tcaatcacat aggtttacat acaataatca ctcaatcaac    360
ttcataagaa gaatcatgtt tacttaattc atcaattatc cccaaaaaca ccactattaa    420
gtataaacta taacatattt gtagtgatgg gtcaacattt ttatcatatt taaactcggg    480
ttccctcaaa tcgagaaata gtgaacatgt aatattaatt ttaaatcgca attacagaaa    540
ttaattgaat ttggtcaaat ggacagaatt ttatagattg ggtggaacta gaaaaaaaaa    600
aaaaaagagt atagggtgaa ttgagtacat gaaagtacat ggtaatccta gttaaacgca    660
taatacatgt gggttcattt gtatttttt gtaacttacg agtaaactgg ctacaacaaa    720
aaaaattaga agatttttt gttttgtaga aaaccctaat tttagttata gttgtataac    780
tttgataaaa ttataaaatt gtattacgaa aaaagtaata agatattcaa aaaagcctag    840
aataacgtat atgactatga gcatgaaact gcaagtcaaa tgctgacaga caaccataaa    900
caaaagaaat taaatagaga tacctttaaa ataagtaaaa tttcatttat aaaaaatcta    960
ctttcttgtg aatctgtcac gttcaataat ttgaagacca ctcaacatac aaggtaaata   1020
atgaaaaata aaatctacca aaatttcaat cattattatc ttccaaaaaa acaaaattat   1080
acagatgatg atggtgatat ggaacttcga ttggctaata ttcactgtgt ctctaaaaac   1140
catccactta tcaagataag atggacccta cactcatcca atctaaacca gtatctcaag   1200
attcttatct aattacatca ttctctaccg ttagatgaaa ttgaccatta accctaccat   1260
aactccatac accgcgagat actggattaa ccaaatcgag atcatcgtag ccgtccgatc   1320
aacaagtacc atctcttgaa atactcgaaa tcctcataag tccgtccctc tttgctctca   1380
ctatcaaaac tctgaatttc gatttcatct agagtcgacg tctgggtacc gagctcggat   1440
ccactagtaa cggccgccag tgtgctggaa ttcgcccttt agtcgaccat atg aaa       1496
                                                        Met Lys
                                                        1
aac atc aaa aaa aac cag gta atg aac ctg ggt ccg aac tct aaa ctg      1544
Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser Lys Leu
        5                  10                  15
ctg aaa gaa tac aaa tcc cag ctg atc gaa ctg aac atc gaa cag ttc      1592
Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu Gln Phe
```

```
    20                  25                  30
gaa gca ggt atc ggt ctg atc ctg ggt gat gct tac atc cgt tct cgt     1640
Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg Ser Arg
35                  40                  45                  50
gat gaa ggt aaa acc tac tgt atg cag ttc gag tgg aaa aac aaa gca     1688
Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn Lys Ala
                55                  60                  65
tac atg gac cac gta tgt ctg ctg tac gat cag tgg gta ctg tcc ccg     1736
Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu Ser Pro
            70                  75                  80
ccg cac aaa aaa gaa cgt gtt aac cac ctg ggt aac ctg gta atc acc     1784
Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val Ile Thr
        85                  90                  95
tgg ggc gcc cag act ttc aaa cac caa gct ttc aac aaa ctg gct agc     1832
Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu Ala Ser
    100                 105                 110
ctg ttc atc gtt aac aac aaa aaa acc atc ccg aac aac ctg gtt gaa     1880
Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn Asn Leu Val Glu
115                 120                 125                 130
aac tac ctg acc ccg atg tct ctg gca tac tgg ttc atg gat gat ggt     1928
Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp Asp Gly
                135                 140                 145
ggt aaa tgg gat tac aac aaa aac tct acc aac aaa tcg atc gta ctg     1976
Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys Ser Ile Val Leu
            150                 155                 160
aac acc cag tct ttc act ttc gaa gaa gta gaa tac ctg gtt aag ggt     2024
Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr Leu Val Lys Gly
        165                 170                 175
ctg cgt aac aaa ttc caa ctg aac tgt tac gta aaa atc aac aaa aac     2072
Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile Asn Lys Asn
    180                 185                 190
aaa ccg atc atc tac atc gat tct atg tct tac ctg atc ttc tac aac     2120
Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe Tyr Asn
195                 200                 205                 210
ctg atc aaa ccg tac ctg atc ccg cag atg atg tac aaa ctg ccg aac     2168
Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys Leu Pro Asn
                215                 220                 225
act atc tcc tcc gaa act ttc ctg aaa taa gatctaaaag ggcgaattct      2218
Thr Ile Ser Ser Glu Thr Phe Leu Lys
            230                 235
gcagatatcc atcacactgg cggccgctcg agccgaatca ctagtctaga gtcgaccgta   2278
ccccctgcag atagactata ctatgtttta gcctgcctgc tggctagcta ctatgttatg   2338
ttatgttgta aaataaacac ctgctaaggt atatctatct atattttagc atggctttct   2398
caataaattg tctttcctta tcgtttacta tcttatacct aataatgaaa taataatatc   2458
acatatgagg aacggggcag gtttaggcat atatatacga gtgtagggcg gagtgggggg   2518
```

gatcggcatg caagctagct tggcgtaatc atggtcatag ctgtttccta ctagatctga 2578
attcactg 2586

<210> 7
<211> 235
<212> PRT
<213> artificial DNA construct (insert of pCB632-17)

<220>
<223> Synthetic Construct

<400> 7
Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser
1 5 10 15

Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu
20 25 30

Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg
35 40 45

Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn
50 55 60

Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu
65 70 75 80

Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val
85 90 95

Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu
100 105 110

Ala Ser Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn Asn Leu
115 120 125

Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp
130 135 140

Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys Ser Ile
145 150 155 160

Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr Leu Val
165 170 175

Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile Asn
180 185 190

Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe
195 200 205

Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys Leu
210 215 220

Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
225 230 235

<210> 8
<211> 982
<212> DNA
<213> Petroselinum crispum

<400> 8
gaattcgaat ccaaaaatta cggatatgaa tataggcata tccgtatccg aattatccgt 60
ttgacagcta gcaacgattg tacaattgct tctttaaaaa aggaagaaag aaagaaagaa 120
aagaatcaac atcagcgtta acaaacggcc ccgttacggc ccaaacggtc atatagagta 180
acggcgttaa gcgttgaaag actcctatcg aaatacgtaa ccgcaaacgt gtcatagtca 240
gatcccctct tccttcaccg cctcaaacac aaaaataatc ttctacagcc tatatataca 300
acccccccct ctatctctcc tttctcacaa ttcatcatct ttctttctct acccccaatt 360
ttaagaaatc ctctcttctc ctcttcattt tcaaggtaaa tctctctctc tctctctctc 420
tctgttattc cttgttttaa ttaggtatgt attattgcta gtttgttaat ctgcttatct 480
tatgtatgcc ttatgtgaat atctttatct tgttcatctc atccgtttag aagctataaa 540
tttgttgatt tgactgtgta tctacacgtg gttatgttta tatctaatca gatatgaatt 600
tcttcatatt gttgcgtttg tgtgtaccaa tccgaaatcg ttgatttttt tcatttaatc 660
gtgtagctaa ttgtacgtat acatatggat ctacgtatca attgttcatc tgtttgtgtt 720
tgtatgtata cagatctgaa aacatcactt ctctcatctg attgtgttgt tacatacata 780
gatatagatc tgttatatca tttttttatt aattgtgtat atatatatgt gcatagatct 840
ggattacatg attgtgatta tttacatgat tttgttattt acgtatgtat atatgtagat 900
ctggactttt tggagttgtt gacttgattg tatttgtgtg tgtatatgtg tgttctgatc 960
ttgatatgtt atgtatgtgc ag 982

<210> 9
<211> 12862
<212> DNA
<213> artificial DNA construct (binary vector pCB666-3)

<220>
<223> artificial DNA construct (binary vector pCB666-3)

<220>

<221> terminator
<222> (8)..(260)
<223> complement: nos terminator

<220>
<221> misc_feature
<222> (277)..(2289)
<223> complement: AHAS herbicide resistance gene

<220>
<221> promoter
<222> (2313)..(3630)
<223> complement: sTPT promoter

<220>
<221> misc_feature
<222> (3636)..(3850)
<223> T-DNA left border

<220>
<221> misc_feature
<222> (9395)..(9540)
<223> T-DNA right border

<220>
<221> terminator
<222> (9599)..(9851)
<223> complement: nos terminator

<220>
<221> misc_feature
<222> (9875)..(10582)
<223> complement: encoding I-SceI homing endonuclease

<220>
<221> promoter
<222> (10600)..(11582)
<223> complement: parsley ubiquitin promoter

<400> 9
aattcccgat ctagtaacat agatgacacc gcgcgcgata atttatccta gtttgcgcgc 60
tatattttgt tttctatcgc gtattaaatg tataattgcg ggactctaat cataaaaacc 120
catctcataa ataacgtcat gcattacatg ttaattatta catgcttaac gtaattcaac 180
agaaattata tgataatcat cgcaagaccg gcaacaggat tcaatcttaa gaaactttat 240
tgccaaatgt ttgaacgatc ggggaaattc gagctctcag tatttaatcc ggccatctcc 300

ttccgttatg acatcgttga aagtgccacc attcgggatc atcggcaaca catgttcttg 360
gtgcggacaa atcacatcca acaggtaagg tcctggtgta tccagcattg tctgaatagc 420
ttctcggaga tctgctttct ttgtcaccct cgccgctgga atcccgcaag ctgctgcaaa 480
cagcaacatg ttcgggaata tctcgtcctc ctgagccgga tctccgagaa atgtgtgtgc 540
tcggttagct ttgtagaacc gatcttccca ttgcataacc atgccaagat gctggttgtt 600
taataaaagt accttcactg gaagattctc tacacgaata gtggctagct cttgcacatt 660
cattataaaa cttccatctc cgtcaatatc cacaactatc gcatcagggt tagcaacaga 720
cgctccaatc gcagcaggaa gtccaaatcc catagctcca aggcctcctg atgatagcca 780
ctgccttggt ttcttgtaat tgtagaactg cgccgcccac atttgatgtt gcccgacacc 840
agtacttatt atggctttтc catcagtcaa ctcatcaagg accttaatcg catactgtgg 900
aggaatagct tccccaaacg tcttaaagct caacggaaac ttctgtttct gtacgttcaa 960
ctcattcctc caaactccaa aatcaagttt aagctcctcc gctcggttct caagaacctt 1020
attcatccct tgcaaagcca gcttaacatc accacacaca gacacatgag gagtcttatt 1080
cttcccaatc tcagccgagt caatatcaat atgaacaatc ttagccctac tagcaaaagc 1140
ctcaagttta cccgtgacac gatcatcaaa ccttacccca aacgccaaca acaaatcact 1200
atgctccaca gcgtaatttg catacacagt cccatgcatt ccaagcatat gtaacgacaa 1260
ctcatcatca caaggataag atcccagccc catcaacgta ctcgcaacag ggattcccgt 1320
aagctcaaca aacctaccca attcatcgct agagttcaaa caaccaccac caacatacaa 1380
cacaggcttc ttagactcag aaatcaacct aacaatctgc tccaaatgag aatcttccgg 1440
aggtttaggc atcctagaca tataaccagg taatctcata gcctgttccc aattaggaat 1500
cgcaagctgt tgttgaatat ctttaggaac atcaaccaaa acaggtccag gtctaccaga 1560
agtagctaaa aagaaagcct cctcaataat cctaggaata tcttcaacat ccatcacaag 1620
atagttatgc ttcgtaatcg aacgcgttac ctcaacaatc ggagtctctt gaaacgcatc 1680
tgtaccaatc atacgacgag ggacttgtcc tgtgattgct acaagaggaa cactatctaa 1740
caacgcatcg gctaatccgc taacgagatt tgtagctccg ggacctgaag tggctataca 1800
gatacctggt ttacctgagg atcgagcgta tccttctgct gcgaatacac ctccttgttc 1860
gtgacgagga aggacgttac ggattgagga agagcgggtt aaggcttggt gaatctccat 1920
tgatgcacct ccagggtaag cgaatacggt ttctacgcct tgacgttcta aagcctcgac 1980
gagaatatca gcgcctttgc ggggttgatc tggagcgaat cgggagatga atgtttcggg 2040
tttggtaggt ttggttggag agggagtggt tgtgacattg gtggttgtgt tgagcacggc 2100
ggagatggag gagggagagc tggatttgat accgcggcgg cgggaggagg aggatgattt 2160
gttggggttt agggagaatg ggagggagaa tctggagatt ggtaatggtg atttggagga 2220
ggaaggagat ggtttggtgg agaaggagat cgaagaagat gttgttgttg ttgttgttgc 2280
cgccgccatg gtacccagac gtcgactcta gatgaaatcg aaattcagag ttttgatagt 2340
gagagcaaag agggacggac ttatgaggat ttcgagtatt tcaagagatg gtacttgttg 2400
atcggacggc tacgatgatc tcgatttggt taatccagta tctcgcggtg tatggagtta 2460
tggtagggtt aatggtcaat ttcatctaac ggtagagaat gatgtaatta gataagaatc 2520
ttgagatact ggtttagatt ggatgagtgt agggtccatc ttatcttgat aagtggatgg 2580
tttttagaga cacagtgaat attagccaat cgaagttcca tatcaccatc atcatctgta 2640
taattttgtt tttttggaag ataataatga ttgaaatttt ggtagatttt atttttcatt 2700
atttaccttg tatgttgagt ggtcttcaaa ttattgaacg tgacagattc acaagaaagt 2760
agatttttta taaatgaaat tttacttatt ttaaaggtat ctctatttaa tttcttttgt 2820
ttatggttgt ctgtcagcat ttgacttgca gtttcatgct catagtcata tacgttattc 2880
taggcttttt tgaatatctt attacttttt tcgtaataca attttataat tttatcaaag 2940
ttatacaact ataactaaaa ttagggtttt ctacaaaaca aaaaaatctt ctaatttttt 3000

ttgttgtagc cagtttactc gtaagttaca aaaaaataca aatgaaccca catgtattat 3060
gcgtttaact aggattacca tgtactttca tgtactcaat tcaccctata ctctttttt 3120
ttttttttct agttccaccc aatctataaa attctgtcca tttgaccaaa ttcaattaat 3180
ttctgtaatt gcgatttaaa attaatatta catgttcact atttctcgat ttgagggaac 3240
ccgagtttaa atatgataaa aatgttgacc catcactaca aatatgttat agtttatact 3300
taatagtggt gtttttgggg ataattgatg aattaagtaa acatgattct tcttatgaag 3360
ttgattgagt gattattgta tgtaaaccta tgtgattgat gttattggtt gattgagtga 3420
ttattgtatt agtatgtaag caaagatgat tgttcttatg aggtaatttg ttactcattc 3480
atccttttgc atatgagaaa ttgtgttagc gtacgcaaaa caatagagaa cataaaagat 3540
atgtgtattt atttaaggtg acttttgtta atgatattgt agtatctata catttatata 3600
taacttgttg aatttgagta taagctatca ggatcctgct tggtaataat tgtcattaga 3660
ttgtttttat gcatagatgc actcgaaatc agccaatttt agacaagtat caaacggatg 3720
ttaattcagt acattaaaga cgtccgcaat gtgttattaa gttgtctaag cgtcaatttg 3780
tttacaccac aatatatcct gccaccagcc agccaacagc tccccgaccg gcagctcggc 3840
acaaaatcac cacgcgttac caccacgccg gccggccgca tggtgttgac cgtgttcgcc 3900
ggcattgccg agttcgagcg ttccctaatc atcgaccgca cccggagcgg gcgcgaggcc 3960
gccaaggccc gaggcgtgaa gtttggcccc cgccctaccc tcaccccggc acagatcgcg 4020
cacgcccgcg agctgatcga ccaggaaggc cgcaccgtga aagaggcggc tgcactgctt 4080
ggcgtgcatc gctcgaccct gtaccgcgca cttgagcgca gcgaggaagt gacgcccacc 4140
gaggccaggc ggcgcggtgc cttccgtgag gacgcattga ccgaggccga cgccctggcg 4200
gccgccgaga atgaacgcca agaggaacaa gcatgaaacc gcaccaggac ggccaggacg 4260
aaccgttttt cattaccgaa gagatcgagg cggagatgat cgcggccggg tacgtgttcg 4320
agccgcccgc gcacgtctca accgtgcggc tgcatgaaat cctggccggt ttgtctgatg 4380
ccaagctggc ggcctggccg gccagcttgg ccgctgaaga aaccgagcgc cgccgtctaa 4440
aaaggtgatg tgtatttgag taaaacagct tgcgtcatgc ggtcgctgcg tatatgatgc 4500
gatgagtaaa taaacaaata cgcaagggga acgcatgaag gttatcgctg tacttaacca 4560
gaaaggcggg tcaggcaaga cgaccatcgc aacccatcta gcccgcgccc tgcaactcgc 4620
cggggccgat gttctgttag tcgattccga tccccagggc agtgcccgcg attgggcggc 4680
cgtgcgggaa gatcaaccgc taaccgttgt cggcatcgac cgcccgacga ttgaccgcga 4740
cgtgaaggcc atcggccggc gcgacttcgt agtgatcgac ggagcgcccc aggcggcgga 4800
cttggctgtg tccgcgatca aggcagccga cttcgtgctg attccggtgc agccaagccc 4860
ttacgacata tgggccaccg ccgacctggt ggagctggtt aagcagcgca ttgaggtcac 4920
ggatggaagg ctacaagcgg cctttgtcgt gtcgcgggcg atcaaaggca cgcgcatcgg 4980
cggtgaggtt gccgaggcgc tggccgggta cgagctgccc attcttgagt cccgtatcac 5040
gcagcgcgtg agctacccag gcactgccgc cgccggcaca accgttcttg aatcagaacc 5100
cgagggcgac gctgcccgcg aggtccaggc gctggccgct gaaattaaat caaaactcat 5160
ttgagttaat gaggtaaaga gaaaatgagc aaaagcacaa acacgctaag tgccggccgt 5220
ccgagcgcac gcagcagcaa ggctgcaacg ttggccagcc tggcagacac gccagccatg 5280
aagcgggtca actttcagtt gccggcggag gatcacacca agctgaagat gtacgcggta 5340
cgccaaggca agaccattac cgagctgcta tctgaataca tcgcgcagct accagagtaa 5400
atgagcaaat gaataaatga gtagatgaat tttagcggct aaaggaggcg gcatggaaaa 5460
tcaagaacaa ccaggcaccg acgccgtgga atgccccatg tgtggaggaa cgggcggttg 5520
gccaggcgta agcggctggg ttgtctgccg gccctgcaat ggcactggaa cccccaagcc 5580
cgaggaatcg gcgtgagcgg tcgcaaacca tccggcccgg tacaaatcgg cgcggcgctg 5640
ggtgatgacc tggtggagaa gttgaaggcc gcgcaggccg cccagcggca acgcatcgag 5700

gcagaagcac gccccggtga atcgtggcaa gcggccgctg atcgaatccg caaagaatcc 5760
cggcaaccgc cggcagccgg tgcgccgtcg attaggaagc cgcccaaggg cgacgagcaa 5820
ccagattttt tcgttccgat gctctatgac gtgggcaccc gcgatagtcg cagcatcatg 5880
gacgtggccg ttttccgtct gtcgaagcgt gaccgacgag ctggcgaggt gatccgctac 5940
gagcttccag acgggcacgt agaggtttcc gcagggccgg ccggcatggc cagtgtgtgg 6000
gattacgacc tggtactgat ggcggtttcc catctaaccg aatccatgaa ccgataccgg 6060
gaagggaagg gagacaagcc cggccgcgtg ttccgtccac acgttgcgga cgtactcaag 6120
ttctgccggc gagccgatgg cggaaagcag aaagacgacc tggtagaaac ctgcattcgg 6180
ttaaacacca cgcacgttgc catgcagcgt acgaagaagg ccaagaacgg ccgcctggtg 6240
acggtatccg agggtgaagc cttgattagc cgctacaaga tcgtaaagag cgaaaccggg 6300
cggccggagt acatcgagat cgagctagct gattggatgt accgcgagat cacagaaggc 6360
aagaacccgg acgtgctgac ggttcacccc gattactttt tgatcgatcc cggcatcggc 6420
cgttttctct accgcctggc acgccgcgcc gcaggcaagg cagaagccag atggttgttc 6480
aagacgatct acgaacgcag tggcagcgcc ggagagttca agaagttctg tttcaccgtg 6540
cgcaagctga tcgggtcaaa tgacctgccg gagtacgatt tgaaggagga ggcggggcag 6600
gctggcccga tcctagtcat gcgctaccgc aacctgatcg agggcgaagc atccgccggt 6660
tcctaatgta cggagcagat gctagggcaa attgccctag caggggaaaa aggtcgaaaa 6720
ggtctctttc ctgtggatag cacgtacatt gggaacccaa agccgtacat tgggaaccgg 6780
aacccgtaca ttgggaaccc aaagccgtac attgggaacc ggtcacacat gtaagtgact 6840
gatataaaag agaaaaaagg cgatttttcc gcctaaaact ctttaaaact tattaaaact 6900
cttaaaaccc gcctggcctg tgcataactg tctggccagc gcacagccga agagctgcaa 6960
aaagcgccta cccttcggtc gctgcgctcc ctacgccccg ccgcttcgcg tcggcctatc 7020
gcggccgctg gccgctcaaa aatggctggc ctacggccag gcaatctacc agggcgcgga 7080
caagccgcgc cgtcgccact cgaccgccgg cgcccacatc aaggcaccct gcctcgcgcg 7140
tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg 7200
tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg 7260
gtgtcggggc gcagccatga cccagtcacg tagcgatagc ggagtgtata ctggcttaac 7320
tatgcggcat cagagcagat tgtactgaga gtgcaccata tgcggtgtga aataccgcac 7380
agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg cttcctcgct cactgactcg 7440
ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg 7500
ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag 7560
gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac 7620
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga 7680
taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt 7740
accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc 7800
tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc 7860
cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta 7920
agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat 7980
gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca 8040
gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct 8100
tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt 8160
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct 8220
cagtggaacg aaaactcacg ttaagggatt ttggtcatgc atgatatatc tcccaatttg 8280
tgtagggctt attatgcacg cttaaaaata ataaaagcag acttgacctg atagtttggc 8340
tgtgagcaat tatgtgctta gtgcatctaa cgcttgagtt aagccgcgcc gcgaagcggc 8400

gtcggcttga acgaatttct agctagacat tatttgccga ctaccttggt gatctcgcct 8460
ttcacgtagt ggacaaattc ttccaactga tctgcgcgcg aggccaagcg atcttcttct 8520
tgtccaagat aagcctgtct agcttcaagt atgacgggct gatactgggc cggcaggcgc 8580
tccattgccc agtcggcagc gacatccttc ggcgcgattt tgccggttac tgcgctgtac 8640
caaatgcggg acaacgtaag cactacattt cgctcatcgc cagcccagtc gggcggcgag 8700
ttccatagcg ttaaggtttc atttagcgcc tcaaatagat cctgttcagg aaccggatca 8760
aagagttcct ccgccgctgg acctaccaag gcaacgctat gttctcttgc ttttgtcagc 8820
aagatagcca gatcaatgtc gatcgtggct ggctcgaaga tacctgcaag aatgtcattg 8880
cgctgccatt ctccaaattg cagttcgcgc ttagctggat aacgccacgg aatgatgtcg 8940
tcgtgcacaa caatggtgac ttctacagcg cggagaatct cgctctctcc aggggaagcc 9000
gaagtttcca aaaggtcgtt gatcaaagct cgccgcgttg tttcatcaag ccttacggtc 9060
accgtaacca gcaaatcaat atcactgtgt ggcttcaggc cgccatccac tgcggagccg 9120
tacaaatgta cggccagcaa cgtcggttcg agatggcgct cgatgacgcc aactacctct 9180
gatagttgag tcgatacttc ggcgatcacc gcttccccca tgatgtttaa ctttgtttta 9240
gggcgactgc cctgctgcgt aacatcgttg ctgctccata acatcaaaca tcgacccacg 9300
gcgtaacgcg cttgctgctt ggatgcccga ggcatagact gtaccccaaa aaaacagtca 9360
taacaagcca tgaaaaccgc cactgcgttc catggacata caaatggacg aacggataaa 9420
ccttttcacg ccctttttaaa tatccgatta ttctaataaa cgctcttttc tcttaggttt 9480
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcggga aacgacaatc 9540
agatctagta ggaaacagct atgaccatga ttacgccaag ctgcgatcgc caattcccga 9600
tctagtaaca tagatgacac cgcgcgcgat aatttatcct agtttgcgcg ctatattttg 9660
ttttctatcg cgtattaaat gtataattgc gggactctaa tcataaaaac ccatctcata 9720
aataacgtca tgcattacat gttaattatt acatgcttaa cgtaattcaa cagaaattat 9780
atgataatca tcgcaagacc ggcaacagga ttcaatctta agaaacttta ttgccaaatg 9840
tttgaacgat cggggaaatt cgagctcctg caggttattt caggaaagtt tcggaggaga 9900
tagtgttcgg cagtttgtac atcatctgcg ggatcaggta cggtttgatc aggttgtaga 9960
agatcaggta agacatagaa tcgatgtaga tgatcggttt gtttttgttg atttttacgt 10020
aacagttcag ttggaatttg ttacgcagac ccttaaccag gtattctact tcttcgaaag 10080
tgaaagactg ggtgttcagt acgatcgatt tgttggtaga gtttttgttg taatcccatt 10140
taccaccatc atccatgaac cagtatgcca gagacatcgg ggtcaggtag ttttcaacca 10200
ggttgttcgg gatggttttt ttgttgttaa cgatgaacag gctagccagt ttgttgaaag 10260
cttggtgttt gaaggtctgg gcgccccagg tgattaccag gttacccagg tggttaacac 10320
gttctttttt gtgcggcggg gacagtaccc actgatcgta cagcagacat acgtggtcca 10380
tgtatgcttt gtttttccac tcgaactgca tacagtaggt tttaccttca tcacgagaac 10440
ggatgtaagc atcacccagg atcagaccga tacctgcttc gaactgttcg atgttcagtt 10500
cgatcagctg ggatttgtat tctttcagca gtttagagtt cggacccagg ttcattacct 10560
ggtttttttt gatgtttttc atggcgcgcc ggatccgggc tgcacataca taacatatca 10620
agatcagaac acacatatac acacacaaat acaatcaagt caacaactcc aaaaagtcca 10680
gatctacata tatacatacg taaataacaa aatcatgtaa ataatcacaa tcatgtaatc 10740
cagatctatg cacatatata tatacacaat taataaaaaa aatgatataa cagatctata 10800
tctatgtatg taacaacaca atcagatgag agaagtgatg ttttcagatc tgtatacata 10860
caaacacaaa cagatgaaca attgatacgt agatccatat gtatacgtac aattagctac 10920
acgattaaat gaaaaaaatc aacgatttcg gattggtaca cacaaacgca acaatatgaa 10980
gaaattcata tctgattaga tataaacata accacgtgta gatacacagt caaatcaaca 11040
aatttatagc ttctaaacgg atgagatgaa caagataaag atattcacat aaggcataca 11100

taagataagc agattaacaa actagcaata atacatacct aattaaaaca aggaataaca 11160
gagagagaga gagagagaga gatttacctt gaaaatgaag aggagaagag aggatttctt 11220
aaaattgggg gtagagaaag aaagatgatg aattgtgaga aaggagagat agaaggggag 11280
gttgtatata taggctgtag aagattattt ttgtgtttga ggcggtgaag gaagagggga 11340
tctgactatg acacgtttgc ggttacgtat ttcgatagga gtctttcaac gcttaacgcc 11400
gttactctat atgaccgttt gggccgtaac ggggccgttt gttaacgctg atgttgattc 11460
ttttctttct ttctttcttc ctttttaaa gaagcaattg tacaatcgtt gctagctgtc 11520
aaacggataa ttcggatacg gatatgccta tattcatatc cgtaattttt ggattcgaat 11580
tctctagtca agcttacggt ctagatattg tttttgtttc acataaatgt cgttttggat 11640
tattcatgta atattttaaa ctaaagtaca atttttgact actttagttt actagttaag 11700
cttttatttt tttgactaac cattgaatga tgaagagatc aacgcatcat atttacaact 11760
tacatagtct tttggaagtg taaattgcta atactaccta aaatatatct ataattaact 11820
aatatttttt cgtcaattat aatagatcaa ttaaaaggct atcaaaagga aaaaaatgaa 11880
atccacatcc tgccatcata acctcatgct ggaaaaagaa atgaaaaaat ataaaaaatt 11940
tcttttgttt attaaattta caactttaat actagtttct tttctatttt ttaaaagctt 12000
ttgtcactta cttaaaaaaa aaaaactttt tgaaatattc ctacttccaa tgtctgatta 12060
gtgcttctgg atttcctttt tggatcatgt gaatcctaaa tcagaaaaat tcatataata 12120
cccaattcag tatattttca tacttcaatt tacaagagtt ctctatgttt ttagcttctt 12180
tcttttaagc caaatgtttt aagcatcttt tatacattaa aataatttag tgttgagttg 12240
agatttttt ttttttttt tggatttact tgttcaaaat ctgaaaaaat gtttacagaa 12300
ggttaaaatg aaccaaaagg catatcaagc tagattttga attaccctat ttcatcgtat 12360
acacaaaact gataatgtgg acacagttga ttttacttct cgatgacatc gtagttttat 12420
ttaatttgga aaccacggcc catatgagca catttcaatt aaaaaccaat ggtaagagca 12480
ttttccatgc aagattcgag agatattaac ccagtgactg ttaaaacagc ttagaaccct 12540
aataacgaat ttcaattact caatttacca ttcgcatttc gcaataacca aactgagcca 12600
gtcacaagga gtaaaccgaa ccggattatt tatttataaa atgaaagaaa ggaaaccaaa 12660
caacaacagc agtagtagtc tgacgtaaac caaaaagcag gcagatcaac aactaaaaga 12720
aactcaaatt accaaaacaa acaggaaatt gcaaactaag ttttttttacc atatgcatac 12780
aaagaccata aaaggttctg ataatcaccg ggctcgaggt accgagctcg aattcactgg 12840
ccgtcgtttt acaacgactc ag 12862

<210> 10
<211> 11824
<212> DNA
<213> artificial DNA construct (binary vector pCB657-41)

<220>
<223> artificial DNA construct (binary vector pCB657-41)

<220>
<221> misc_feature
<222> (38)..(183)
<223> T-DNA right border

<220>

<221> misc_feature
<222> (415)..(444)
<223> complement: I-CreI site

<220>
<221> misc_feature
<222> (445)..(462)
<223> complement: I-SceI site

<220>
<221> misc_feature
<222> (509)..(1312)
<223> complement: nptII gene

<220>
<221> promoter
<222> (1356)..(3167)
<223> complement: A.thaliana nitrilase-1 promoter

<220>
<221> misc_feature
<222> (3182)..(3199)
<223> I-SceI site

<220>
<221> misc_feature
<222> (3200)..(3229)
<223> I-CreI site

<220>
<221> terminator
<222> (3257)..(3461)
<223> complement: 35S terminator

<220>
<221> misc_feature
<222> (3536)..(5533)
<223> complement:GUS gene

<220>
<221> Intron
<222> (4960)..(5148)
<223> complement: PIV2 intron

<220>

<221> promoter
<222> (5587)..(6904)
<223> complement: sTPT promoter

<220>
<221> misc_feature
<222> (6976)..(7192)
<223> T-DNA left border

<400> 10
aatattcaaa caaacacata cagcgcgact tatcatggac atacaaatgg acgaacggat 60
aaaccttttc acgccctttt aaatatccga ttattctaat aaacgctctt ttctcttagg 120
tttacccgcc aatatatcct gtcaaacact gatagtttaa actgaaggcg ggaaacgaca 180
atcagatctg gtacccggtc actggatttt ggttttagga attagaaatt ttattgatag 240
aagtatttta caaatacaaa tacatactaa gggtttctta tatgctcaac acatgagcga 300
aaccctataa gaaccctaat tcccttatct gggaactact cacacattat tctggagaaa 360
aatagagaga gatagatttg tagagagaga ctggtgattt ttgcgccggg taccccaaac 420
tgtctcacga cgttttgaac ccagattacc ctgttatccc tagtcgactc tagactagtg 480
gatccgtcga agctagcttg ggtcccgctc agaagaactc gtcaagaagg cgatagaagg 540
cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg tcagcccatt 600
cgccgccaag ctcttcagca atatcacggg tagccaacgc tatgtcctga tagcggtccg 660
ccacacccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc accatgatat 720
tcggcaagca ggcatcgcca tgggtcacga cgagatcctc gccgtcgggc atgcgcgcct 780
tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc agatcatcct 840
gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt ttcgcttggt 900
ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca tcagccatga 960
tggatacttt ctcggcagga gcaaggtgag atgacaggag atcctgcccc ggcacttcgc 1020
ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct gcgcaaggaa 1080
cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca ttcagggcac 1140
cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc cggaacacgg 1200
cggcatcaga gcagccgatt gtctgttgtg cccagtcata gccgaatagc ctctccaccc 1260
aagcggccgg agaacctgcg tgcaatccat cttgttcaat ccaagctccc atgggccctc 1320
gactagagtc gagatccgat atcgcccggg ctcgagtctt tgttttttac tttggttcat 1380
gacactcaga gacttgagag aagcaatata tagacttttt tttgtttttt ttttgtggtc 1440
acgtttattt tcctattgga gacggtaacg aagatcgaac ctgtggtgga aatgaaacaa 1500
ggtgggacta gcccacgtgg tttcttttct ctgcattgat ttgtttttgt ttttttgta 1560
aagttcacat caaacctact aataattgag aagaaaaata aaatctattg attgattaaa 1620
ccagccgatg ctttatgtct gaatataaaa aagaagtgaa aaccccgttt aagaattaca 1680
acggtggttt acaaagtatt tggacacaat aaatccaaac gaaataaaac aaaatggaga 1740
actaccaaat aaaaaacaaa taaaaaactt aaaagaattt attccatttt ttttcccgta 1800
gaatttattc ttttatggat tccttaaatc catatttgat gcattttgat tcctcataat 1860
aggtaataat atatactatg ttatagatat gtttctaatt cgtattaacc tacctttttt 1920
tggtcgtacg attctaccta ataatattga acggaattga tgttttggac cacttagaaa 1980
gtatttttt tttggtttgt cttagctgta tttcattaaa tataaattta aataagaaat 2040
gtcataaata aaatttgacg tatagatttt ttaaatccat tttatgttat ttaatatttg 2100

aaatgtgagt ttggctccta tttaatctta ggatgggtta atactaagtt ttccttaatg 2160
aattatctca gagaaactgg attaaataaa ctaaaaaata gatcaatgtg ttttggtccg 2220
gtcaaatatc tttggattta ctattattgg cgaaaagaaa gtctcatata gtaaatcata 2280
ttcctacaag agaaatcaaa atttttgaat taacatggat tgtatagttt cttatataac 2340
caattagttc gcatcaagaa aaccaaaccc caattaataa tcaaacgggc ttggtaggaa 2400
tatttcattg cagctttcag ataaaagaaa aaaacacaca ctcaagtctt ttatttcatc 2460
tttcttactt gcaggaactc aaattccact ttgccacttt tctttacaaa taaacacaaa 2520
ttgtcaatga aacgaaatag tctttttatg caaacactgt ttgtcttttt tcgatcacgt 2580
ttctgattgt gacagccatc catatatata gggaatgtaa aacaacaaca tgtgaagtca 2640
catatacgta atggtttagc atagcttcta ttttcgttgt caatattagt cattccaaaa 2700
cattttaag aaaaataaat taatatatgt atattcttgg aactaatgta tgtggaaata 2760
cagtaactta attattaaac attctaaatg caaatatgca aagaaaaaaa agaaaagaac 2820
acaactgaaa tcaaagccag attcataata attggctaca tggttgtaga atgtagggta 2880
acacaacatc cagaattgaa cactcaaatt ggatgataga tggataatct ttagatacaa 2940
gagaattggt tctcttccat tattaacgaa aataaagaaa aaaagtttag cataaaagtt 3000
tgaaactcaa cataacattt tgaacttgac tccttcatag gagtgacatg aactgacgaa 3060
tcacaaccga ttacttgttt gagtcatctt ccgctttctc caccttcgaa atgaatgtga 3120
ccggtttctt cgggtgctca tttacggtca agtgtaaaac atctggtctc gaggtacctg 3180
gtagggataa cagggtaatc tgggttcaaa acgtcgtgag acagtttggt gcaggtcgaa 3240
attcgagctc ggtacccggt cactggattt tggttttagg aattagaaat tttattgata 3300
gaagtatttt acaaatacaa atacatacta agggtttctt atatgctcaa cacatgagcg 3360
aaaccctata agaaccctaa ttcccttatc tgggaactac tcacacatta ttctggagaa 3420
aaatagagag agatagattt gtagagagag actggtgatt tttgcgccgg gtaccgagct 3480
cggtagcaat tcccgaggct gtagccgacg atggtgcgcc aggagagttg ttgattcatt 3540
gtttgcctcc ctgctgcggt ttttcaccga agttcatgcc agtccagcgt ttttgcagca 3600
gaaaagccgc cgacttcggt ttgcggtcgc gagtgaagat ccctttcttg ttaccgccaa 3660
cgcgcaatat gccttgcgag gtcgcaaaat cggcgaaatt ccatacctgt tcaccgacga 3720
cggcgctgac gcgatcaaag acgcggtgat acatatccag ccatgcacac tgatactctt 3780
cactccacat gtcggtgtac attgagtgca gcccggctaa cgtatccacg ccgtattcgg 3840
tgatgataat cggctgatgc agtttctcct gccaggccag aagttctttt tccagtacct 3900
tctctgccgt ttccaaatcg ccgctttgga cataccatcc gtaataacgg ttcaggcaca 3960
gcacatcaaa gagatcgctg atggtatcgg tgtgagcgtc gcagaacatt acattgacgc 4020
aggtgatcgg acgcgtcggg tcgagtttac gcgttgcttc cgccagtggc gaaatattcc 4080
cgtgcacttg cggacgggta tccggttcgt tggcaatact ccacatcacc acgcttgggt 4140
ggtttttgtc acgcgctatc agctctttaa tcgcctgtaa gtgcgcttgc tgagtttccc 4200
cgttgactgc ctcttcgctg tacagttctt tcggcttgtt gcccgcttcg aaaccaatgc 4260
ctaaagagag gttaaagccg acagcagcag tttcatcaat caccacgatg ccatgttcat 4320
ctgcccagtc gagcatctct tcagcgtaag ggtaatgcga ggtacggtag gagttggccc 4380
caatccagtc cattaatgcg tggtcgtgca ccatcagcac gttatcgaat cctttgccac 4440
gtaagtccgc atcttcatga cgaccaaagc cagtaaagta gaacggtttg tggttaatca 4500
ggaactgttc gcccttcact gccactgacc ggatgccgac gcgaagcggg tagatatcac 4560
actctgtctg gcttttggct gtgacgcaca gttcatagag ataaccttca cccggttgcc 4620
agaggtgcgg attcaccact tgcaaagtcc cgctagtgcc ttgtccagtt gcaaccacct 4680
gttgatccgc atcacgcagt tcaacgctga catcaccatt ggccaccacc tgccagtcaa 4740
cagacgcgtg gttacagtct tgcgcgacat gcgtcaccac ggtgatatcg tccacccagg 4800

tgttcggcgt ggtgtagagc attacgctgc gatggattcc ggcatagtta aagaaatcat 4860
ggaagtaaga ctgctttttc ttgccgtttt cgtcggtaat caccattccc ggcgggatag 4920
tctgccagtt cagttcgttg ttcacacaaa cggtgatacc tgcacatcaa caaattttgg 4980
tcatatatta gaaaagttat aaattaaaat atacacactt ataaactaca gaaaagcaat 5040
tgctatatac tacattcttt tattttgaaa aaaatatttg aaatattata ttactactaa 5100
ttaatgataa ttattatata tatatcaaag gtagaagcag aaacttacgt acacttttcc 5160
cggcaataac atacggcgtg acatcggctt caaatggcgt atagccgccc tgatgctcca 5220
tcacttcctg attattgacc cacactttgc cgtaatgagt gaccgcatcg aaacgcagca 5280
cgatacgctg gcctgcccaa cctttcggta taaagacttc gcgctgatac cagacgttgc 5340
ccgcataatt acgaatatct gcatcggcga actgatcgtt aaaactgcct ggcacagcaa 5400
ttgcccggct ttcttgtaac gcgctttccc accaacgctg accaattcca cagttttcgc 5460
gatccagact gaatgcccac aggccgtcga gtttttttgat ttcacgggtt ggggtttcta 5520
caggacgtac catggtcgat cgactctaga ctagtggatc cgatatcgcc cgggctcgac 5580
tctagatgaa atcgaaattc agagttttga tagtgagagc aaagagggac ggacttatga 5640
ggatttcgag tatttcaaga gatggtactt gttgatcgga cggctacgat gatctcgatt 5700
tggttaatcc agtatctcgc ggtgtatgga gttatggtag ggttaatggt caatttcatc 5760
taacggtaga gaatgatgta attagataag aatcttgaga tactggttta gattggatga 5820
gtgtagggtc catcttatct tgataagtgg atggttttta gagacacagt gaatattagc 5880
caatcgaagt tccatatcac catcatcatc tgtataattt tgtttttttg gaagataata 5940
atgattgaaa ttttggtaga ttttatttt cattatttac cttgtatgtt gagtggtctt 6000
caaattattg aacgtgacag attcacaaga aagtagattt tttataaatg aaattttact 6060
tattttaaag gtatctctat ttaatttctt ttgtttatgg ttgtctgtca gcatttgact 6120
tgcagtttca tgctcatagt catatacgtt attctaggct tttttgaata tcttattact 6180
tttttcgtaa tacaatttta taattttatc aaagttatac aactataact aaaattaggg 6240
ttttctacaa aacaaaaaaa tcttctaatt ttttttgttg tagccagttt actcgtaagt 6300
tacaaaaaaa tacaaatgaa cccacatgta ttatgcgttt aactaggatt accatgtact 6360
ttcatgtact caattcaccc tatactcttt ttttttttt ttctagttcc acccaatcta 6420
taaaattctg tccatttgac caaattcaat taatttctgt aattgcgatt taaaattaat 6480
attacatgtt cactatttct cgatttgagg gaacccgagt ttaaatatga taaaaatgtt 6540
gacccatcac tacaaatatg ttatagttta tacttaatag tggtgttttt ggggataatt 6600
gatgaattaa gtaaacatga ttcttcttat gaagttgatt gagtgattat tgtatgtaaa 6660
cctatgtgat tgatgttatt ggttgattga gtgattattg tattagtatg taagcaaaga 6720
tgattgttct tatgaggtaa tttgttactc attcatcctt ttgcatatga gaaattgtgt 6780
tagcgtacgc aaaacaatag agaacataaa agatatgtgt atttatttaa ggtgactttt 6840
gttaatgata ttgtagtatc tatacattta tatataactt gttgaatttg agtataagct 6900
atcaggatcc gggggatcct ctagagtcga ggtaccgagc tcgaattcac tggccgtcgt 6960
tttacaacga ctcagtactg cttggtaata attgtcatta gattgttttt atgcatagat 7020
gcactcgaaa tcagccaatt ttagacaagt atcaaacgga tgttaattca gtacattaaa 7080
gacgtccgca atgtgttatt aagttgtcta agcgtcaatt tgtttacacc acaatatatc 7140
ctgccaccag ccagccaaca gctccccgac cggcagctcg gcacaaaatc accacgcgtc 7200
taaaaaggtg atgtgtattt gagtaaaaca gcttgcgtca tgcggtcgct gcgtatatga 7260
tgcgatgagt aaataaacaa atacgcaagg ggaacgcatg aaggttatcg ctgtacttaa 7320
ccagaaaggc gggtcaggca agacgaccat cgcaacccat ctagcccgcg ccctgcaact 7380
cgccggggcc gatgttctgt tagtcgattc cgatccccag ggcagtgccc gcgattgggc 7440
ggccgtgcgg gaagatcaac cgctaaccgt tgtcggcatc gaccgcccga cgattgaccg 7500

cgacgtgaag gccatcggcc ggcgcgactt cgtagtgatc gacggagcgc cccaggcggc 7560
ggacttggct gtgtccgcga tcaaggcagc cgacttcgtg ctgattccgg tgcagccaag 7620
cccttacgac atatgggcca ccgccgacct ggtggagctg gttaagcagc gcattgaggt 7680
cacggatgga aggctacaag cggcctttgt cgtgtcgcgg gcgatcaaag gcacgcgcat 7740
cggcggtgag gttgccgagg cgctggccgg gtacgagctg cccattcttg agtcccgtat 7800
cacgcagcgc gtgagctacc caggcactgc cgccgccggc acaaccgttc ttgaatcaga 7860
acccgagggc gacgctgccc gcgaggtcca ggcgctggcc gctgaaatta aatcaaaact 7920
catttgagtt aatgaggtaa agagaaaatg agcaaaagca caaacacgct aagtgccggc 7980
cgtccgagcg cacgcagcag caaggctgca acgttggcca gcctggcaga cacgccagcc 8040
atgaagcggg tcaactttca gttgccggcg gaggatcaca ccaagctgaa gatgtacgcg 8100
gtacgccaag gcaagaccat taccgagctg ctatctgaat acatcgcgca gctaccagag 8160
taaatgagca aatgaataaa tgagtagatg aattttagcg gctaaaggag gcggcatgga 8220
aaatcaagaa caaccaggca ccgacgccgt ggaatgcccc atgtgtggag gaacgggcgg 8280
ttggccaggc gtaagcggct gggttgtctg ccggccctgc aatggcactg gaacccccaa 8340
gcccgaggaa tcggcgtgag cggtcgcaaa ccatccggcc cggtacaaat cggcgcggcg 8400
ctgggtgatg acctggtgga gaagttgaag gccgcgcagg ccgcccagcg gcaacgcatc 8460
gaggcagaag acgccccggt gaatcgtggc aaggggccgc tgatcgaatc cgcaaagaat 8520
cccggcaacc gccggcagcc ggtgcgccgt cgattaggaa gccgcccaag ggcgacgagc 8580
aaccagattt tttcgttccg atgctctatg acgtgggcac ccgcgatagt cgcagcatca 8640
tggacgtggc cgttttccgt ctgtcgaagc gtgaccgacg agctggcgag gtgatccgct 8700
acgagcttcc agacgggcac gtagaggttt ccgcagggcc ggccggcatg gccagtgtgt 8760
gggattacga cctggtactg atggcggttt cccatctaac cgaatccatg aaccgatacc 8820
gggaagggaa gggagacaag cccggccgcg tgttccgtcc acacgttgcg gacgtactca 8880
agttctgccg gcgagccgat ggcggaaagc agaaagacga cctggtagaa acctgcattc 8940
ggttaaacac cacgcacgtt gccatgcagc gtacgaagaa ggccaagaac ggccgcctgg 9000
tgacggtatc cgagggtgaa gccttgatta gccgctacaa gatcgtaaag agcgaaaccg 9060
ggcggccgga gtacatcgag atcgagctag ctgattggat gtaccgcgag atcacagaag 9120
gcaagaaccc ggacgtgctg acggttcacc ccgattactt tttgatcgat cccggcatcg 9180
gccgttttct ctaccgcctg gcacgccgcg ccgcaggcaa ggcagaagcc agatggttgt 9240
tcaagacgat ctacgaacgc agtggcagcg ccggagagtt caagaagttc tgtttcaccg 9300
tgcgcaagct gatcgggtca aatgacctgc cggagtacga tttgaaggag gaggcggggc 9360
aggctggccc gatcctagtc atgcgctacc gcaacctgat cgagggcgaa gcatccgccg 9420
gttcctaatg tacggagcag atgctagggc aaattgccct agcaggggaa aaaggtcgaa 9480
aaggtctctt tcctgtggat agcacgtaca ttgggaaccc aaagccgtac attgggaacc 9540
ggaacccgta cattgggaac ccaaagccgt acattgggaa ccggtcacac atgtaagtga 9600
ctgatataaa agagaaaaaa ggcgattttt ccgcctaaaa ctctttaaaa cttattaaaa 9660
ctcttaaaac ccgcctggcc tgtgcataac tgtctggcca gcgcacagcc gaagagctgc 9720
aaaaagcgcc taccttcgg tcgctgcgct ccctacgccc cgccgcttcg cgtcggccta 9780
tcgcggccta tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg 9840
cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 9900
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 9960
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 10020
gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 10080
aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 10140
gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 10200

ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 10260
cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 10320
ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 10380
actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 10440
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 10500
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 10560
ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 10620
cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 10680
ttggtcatgc atgatatatc tcccaatttg tgtagggctt attatgcacg cttaaaaata 10740
ataaaagcag acttgacctg atagtttggc tgtgagcaat tatgtgctta gtgcatctaa 10800
cgcttgagtt aagccgcgcc gcgaagcggc gtcggcttga acgaatttct agctagacat 10860
tatttgccga ctaccttggt gatctcgcct ttcacgtagt ggacaaattc ttccaactga 10920
tctgcgcgcg aggccaagcg atcttcttct tgtccaagat aagcctgtct agcttcaagt 10980
atgacgggct gatactgggc cggcaggcgc tccattgccc agtcggcagc gacatccttc 11040
ggcgcgattt tgccggttac tgcgctgtac caaatgcggg acaacgtaag cactacattt 11100
cgctcatcgc cagcccagtc gggcggcgag ttccatagcg ttaaggtttc atttagcgcc 11160
tcaaatagat cctgttcagg aaccggatca aagagttcct ccgccgctgg acctaccaag 11220
gcaacgctat gttctcttgc ttttgtcagc aagatagcca gatcaatgtc gatcgtggct 11280
ggctcgaaga tacctgcaag aatgtcattg cgctgccatt ctccaaattg cagttcgcgc 11340
ttagctggat aacgccacgg aatgatgtcg tcgtgcacaa caatggtgac ttctacagcg 11400
cggagaatct cgctctctcc aggggaagcc gaagtttcca aaaggtcgtt gatcaaagct 11460
cgccgcgttg tttcatcaag ccttacggtc accgtaacca gcaaatcaat atcactgtgt 11520
ggcttcaggc cgccatccac tgcggagccg tacaaatgta cggccagcaa cgtcggttcg 11580
agatggcgct cgatgacgcc aactacctct gatagttgag tcgatacttc ggcgatcacc 11640
gcttccccca tgatgtttaa ctttgtttta gggcgactgc cctgctgcgt aacatcgttg 11700
ctgctccata acatcaaaca tcgacccacg gcgtaacgcg cttgctgctt ggatgcccga 11760
ggcatagact gtaccccaaa aaaacagtca taacaagcca tgaaaaccgc cactgcgttc 11820
catg 11824

<210> 11
<211> 13531
<212> DNA
<213> artificial DNA construct (binary vector JB010qcz)

<220>
<223> artificial DNA construct (binary vector JB010qcz)

<220>
<221> misc_feature
<222> (149)..(363)
<223> complement: T-DNA left border

<220>
<221> misc_feature
<222> (372)..(6088)

<223> AHAS herbicide resistance gene expression cassette

<220>
<221> promoter
<222> (6106)..(7088)
<223> parsley ubiquitin promoter (incl. intron in 5' UTR)

<220>
<221> misc_feature
<222> (7106)..(7813)
<223> coding region for I-SceI

<220>
<221> terminator
<222> (7837)..(8089)
<223> nos terminator

<220>
<221> misc_feature
<222> (8148)..(8293)
<223> T-DNA right border

<400> 11
gtagggcggg ggccaaactt cacgcctcgg gccttggcgg cctcgcgccc gctccgggtg 60
cggtcgatga ttagggaacg ctcgaactcg gcaatgccgg cgaacacggt caacaccatg 120
cggccggccg gcgtggtggt aacgcgtggt gattttgtgc cgagctgccg gtcggggagc 180
tgttggctgg ctggtggcag gatatattgt ggtgtaaaca aattgacgct tagacaactt 240
aataacacat tgcggacgtc tttaatgtac tgaattaaca tccgtttgat acttgtctaa 300
aattggctga tttcgagtgc atctatgcat aaaaacaatc taatgacaat tattaccaag 360
caggatccca gtctagatta tgtatttcca actttcatta acaatataat cgcatataaa 420
tgaaaaatcg tttccaggat aatattttga tgaaatctca tattattgtt cgtactcgga 480
ttgatgttga aggcttgaag cgcttcaaat tatagaccag attatttaag tttttctttt 540
gtttactcca tatcaatttg atccattata ctacctaaga aaatttaggt aacatagaat 600
tatttattgt tatagtaaaa aaaaggaaaa ccacaaaaat aatctacttt tacgtatata 660
ctattttcat gacataagta attaagttgt acaacttttt tttaatgaaa agagagagta 720
aatttatcat gttcatgtgt agttacctcg tgaataaccg acggttatat agacgcctaa 780
catgaattgt tcagttgaag acagttcaaa acatgtgttt cactctaaaa tcctcaacaa 840
aaaaaaagtg ttaaaatttg taaacctctt tcaagcaaaa aaagaaaaag tgttagaatc 900
ccaagattct ttcataatcc ggaatcttgg ctgaaaacgt ataaaagaga ttgacgtagt 960
aacaaggagt cttggtatgc ttccatgctt tttatccttt tttgtcatgg aaccatgatt 1020
tggttaccat ttattatgta accgaaattt tcattgtaat aatgaatatt taaattttta 1080
gcaaaaaaaa acaaaaaaaa acaaggagtc ttgtcttcgt tctcaaattt cagagctctt 1140
gcacttttca agagttttac tttgatgagt gagacatttg tcttttttagt gtttattttc 1200
taaacttaaa atagtagcat caacatcact caattataat tcttaagatg ttgtagaaaa 1260
atattttata gatggaaagt aatcgatatt aagacaaata agaaaccaaa ccggactttg 1320

tgttcagacc gaatcaaatc tgaattggag aaattatggt ggaggcgaaa gtcaacggaa 1380
ctaaagtata aaaccaaatg tcaaaaataa aacccaattt tcatccttaa acgaacctgc 1440
tgaaacccta atttcgatta ccaattccga tctaaaaaga agtcatggaa gccattgatt 1500
ccgcaatcga tcctctcaga gatttcgcta agagcagtgt tcgtctcgtc cagcgctgtc 1560
acaaacccga tcgcaagggt aacgcctttt ctcaaaaaaa tctcatttcc gatttttgat 1620
ctgtagatta gggttttctg aaattttgat atcatttgta attgaattgg ttatcagaat 1680
tcacgaaagt agctgtgcgt acggcgattg gatttgtggt gatgggattc gttggattct 1740
tcgtgaagct cgttttcatc ccaatcaaca acatcatcgt tggatcttct tagtgtagta 1800
ctttctttac gaggtaattg atctcgcatt atatatctac attttggtta tgttacttga 1860
catatagtca ttgattcaat agttctgtta attcctttaa agatcatttt gactagacca 1920
cattcttggt tcattcctca ataatttgta atcatattgg tggatataga agtagattgg 1980
ttatagatca gatagtggaa gactttagga tgaatttcag ctagtttttt tttttggctt 2040
attgtctcaa aagattagtg ctttgctgtc tccattgctt ctgctatcga cacgcttctg 2100
tctccttgta tctttattat atctattcgt cccatgagtt ttgtttgttc tgtattcgtt 2160
cgctctggtg tcatggatgg agtctctgtt ccatgtttct gtaatgcatg ttgggttgtt 2220
tcatgcaaga aatgctgaga taaacactca tttgtgaaag tttctaaact ctgaatcgcg 2280
ctacaggcaa tgctccgagg agtaggagga gaagaacgaa ccaaacgaca ttatcagccc 2340
tttgaggaag ctcttagttt tgttattgtt tttgtagcca aattctccat tcttattcca 2400
ttttcactta tctcttgttc cttatagacc ttataagttt tttattcatg tatacaaatt 2460
atattgtcat caagaagtat ctttaaaatc taaatctcaa atcaccagga ctatgttttt 2520
gtccaattcg tggaaccaac ttgcagcttg tatccattct cttaaccaat aaaaaaagaa 2580
agaaagatca atttgataaa tttctcagcc acaaattcta catttaggtt ttagcatatc 2640
gaaggctcaa tcacaaatac aatagataga ctagagattc cagcgtcacg tgagttttat 2700
ctataaataa aggaccaaaa atcaaatccc gagggcattt tcgtaatcca acataaaacc 2760
cttaaacttc aagtctcatt tttaaacaaa tcatgttcac aagtctcttc ttcttctctg 2820
tttctctatc tcttgctcat ctttctcctg aaccatggcg gcggcaacaa caacaacaac 2880
aacatcttct tcgatctcct tctccaccaa accatctcct tcctcctcca aatcaccatt 2940
accaatctcc agattctccc tcccattctc cctaaacccc aacaaatcat cctcctcctc 3000
ccgccgccgc ggtatcaaat ccagctctcc ctcctccatc tccgccgtgc tcaacacaac 3060
caccaatgtc acaaccactc cctctccaac caaacctacc aaacccgaaa cattcatctc 3120
ccgattcgct ccagatcaac cccgcaaagg cgctgatatc ctcgtcgaag ctttagaacg 3180
tcaaggcgta gaaaccgtat tcgcttaccc tggaggtgca tcaatggaga ttcaccaagc 3240
cttaacccgc tcttcctcaa tccgtaacgt ccttcctcgt cacgaacaag gaggtgtatt 3300
cgcagcagaa ggatacgctc gatcctcagg taaaccaggt atctgtatag ccacttcagg 3360
tcccggagct acaaatctcg ttagcggatt agccgatgcg ttgttagata gtgttcctct 3420
tgtagcaatc acaggacaag tccctcgtcg tatgattggt acagatgcgt ttcaagagac 3480
tccgattgtt gaggtaacgc gttcgattac gaagcataac tatcttgtga tggatgttga 3540
agatatccct aggattattg aggaagcttt ctttttagct acttctggta gacctggacc 3600
tgttttggtt gatgttccta aagatattca acaacagctt gcgattccta attgggaaca 3660
ggctatgaga ttacctggtt atatgtctag gatgcctaaa cctccggaag attctcattt 3720
ggagcagatt gttaggttga tttctgagtc taagaagcct gtgttgtatg ttggtggtgg 3780
ttgtttgaat tctagcgatg aattgggtag gtttgttgag cttacgggga tccctgttgc 3840
gagtacgttg atggggctgg gatcttatcc ttgtgatgat gagttgtcgt tacatatgct 3900
tggaatgcat gggactgtgt atgcaaatta cgctgtggag catagtgatt tgttgttggc 3960
gtttggggta aggtttgatg atcgtgtcac gggtaagctt gaggcttttg ctagtagggc 4020

taagattgtt catattgata ttgactcggc tgagattggg aagaataaga ctcctcatgt 4080
gtctgtgtgt ggtgatgtta agctggcttt gcaagggatg aataaggttc ttgagaaccg 4140
agcggaggag cttaagcttg attttggagt ttggaggaat gagttgaacg tacagaaaca 4200
gaagtttccg ttgagcttta agacgtttgg ggaagctatt cctccacagt atgcgattaa 4260
ggtccttgat gagttgactg atggaaaagc cataataagt actggtgtcg ggcaacatca 4320
aatgtgggcg gcgcagttct acaattacaa gaaaccaagg cagtggctat catcaggagg 4380
ccttggagct atgggatttg gacttcctgc tgcgattgga gcgtctgttg ctaaccctga 4440
tgcgatagtt gtggatattg acggagatgg aagctttata atgaatgtgc aagagctagc 4500
cactattcgt gtagagaatc ttccagtgaa ggtactttta ttaaacaacc agcatcttgg 4560
catggttatg caatgggaag atcggttcta caaagctaac cgagctcaca catttctcgg 4620
ggatccggct caggaggacg agatattccc gaacatgttg ctgtttgcag cagcttgcgg 4680
gattccagcg gcgagggtga caaagaaagc agatctccga gaagctattc agacaatgct 4740
ggatacacca ggaccttacc tgttggatgt gatttgtccg caccaagaac atgtgttgcc 4800
gatgatcccg aatggtggca ctttcaacga tgtcataacg gaaggagatg gccggattaa 4860
atactgagag atgaaaccgg tgattatcag aacctttttat ggtctttgta tgcatatggt 4920
aaaaaaactt agtttgcaat ttcctgtttg ttttggtaat ttgagtttct tttagttgtt 4980
gatctgcctg ctttttggtt tacgtcagac tactactgct gttgttgttt ggtttccttt 5040
ctttcatttt ataaataaat aatccggttc ggtttactcc ttgtgactgg ctcagtttgg 5100
ttattgcgaa atgcgaatgg taaattgagt aattgaaatt cgttattagg gttctaagct 5160
gttttaacag tcactgggtt aatatctctc gaatcttgca tggaaaatgc tcttaccatt 5220
ggtttttaat tgaaatgtgc tcatatgggc cgtggtttcc aaattaaata aaactacgat 5280
gtcatcgaga agtaaaatca actgtgtcca cattatcagt tttgtgtata cgatgaaata 5340
gggtaattca aaatctagct tgatatgcct tttggttcat tttaaccttc tgtaaacatt 5400
ttttcagatt ttgaacaagt aaatccaaaa aaaaaaaaaa aaaatctcaa ctcaacacta 5460
aattatttta atgtataaaa gatgcttaaa acatttggct taaaagaaag aagctaaaaa 5520
catagagaac tcttgtaaat tgaagtatga aaatatactg aattgggtat tatatgaatt 5580
tttctgattt aggattcaca tgatccaaaa aggaaatcca gaagcactaa tcagacattg 5640
gaagtaggaa tatttcaaaa agtttttttt ttttaagtaa gtgacaaaag cttttaaaaa 5700
atagaaaaga aactagtatt aaagttgtaa atttaataaa caaaagaaat tttttatatt 5760
ttttcatttc tttttccagc atgaggttat gatggcagga tgtggatttc atttttttcc 5820
ttttgatagc cttttaattg atctattata attgacgaaa aaatattagt taattataga 5880
tatattttag gtagtattag caatttacac ttccaaaaga ctatgtaagt tgtaaatatg 5940
atgcgttgat ctcttcatca ttcaatggtt agtcaaaaaa ataaaagctt aactagtaaa 6000
ctaaagtagt caaaaattgt actttagttt aaaatattac atgaataatc caaaacgaca 6060
tttatgtgaa acaaaaacaa tatctagacc gtaagcttga ctagagaatt cgaatccaaa 6120
aattacggat atgaatatag gcatatccgt atccgaatta tccgtttgac agctagcaac 6180
gattgtacaa ttgcttcttt aaaaaaggaa gaaagaaaga aagaaaagaa tcaacatcag 6240
cgttaacaaa cggccccgtt acggcccaaa cggtcatata gagtaacggc gttaagcgtt 6300
gaaagactcc tatcgaaata cgtaaccgca aacgtgtcat agtcagatcc cctcttcctt 6360
caccgcctca aacacaaaaa taatcttcta cagcctatat atacaacccc cccttctatc 6420
tctcctttct cacaattcat catctttctt tctctacccc caattttaag aaatcctctc 6480
ttctcctctt cattttcaag gtaaatctct ctctctctct ctctctctgt tattccttgt 6540
tttaattagg tatgtattat tgctagtttg ttaatctgct tatcttatgt atgccttatg 6600
tgaatatctt tatcttgttc atctcatccg tttagaagct ataaatttgt tgatttgact 6660
gtgtatctac acgtggttat gtttatatct aatcagatat gaatttcttc atattgttgc 6720

gtttgtgtgt accaatccga aatcgttgat ttttttcatt taatcgtgta gctaattgta 6780
cgtatacata tggatctacg tatcaattgt tcatctgttt gtgtttgtat gtatacagat 6840
ctgaaaacat cacttctctc atctgattgt gttgttacat acatagatat agatctgtta 6900
tatcattttt tttattaatt gtgtatatat atatgtgcat agatctggat tacatgattg 6960
tgattattta catgattttg ttatttacgt atgtatatat gtagatctgg acttttgga 7020
gttgttgact tgattgtatt tgtgtgtgta tatgtgtgtt ctgatcttga tatgttatgt 7080
atgtgcagcc cggatccggc gcgccatgaa aaacatcaaa aaaaaccagg taatgaacct 7140
gggtccgaac tctaaactgc tgaaagaata caaatcccag ctgatcgaac tgaacatcga 7200
acagttcgaa gcaggtatcg gtctgatcct gggtgatgct tacatccgtt ctcgtgatga 7260
aggtaaaacc tactgtatgc agttcgagtg gaaaaacaaa gcatacatgg accacgtatg 7320
tctgctgtac gatcagtggg tactgtcccc gccgcacaaa aaagaacgtg ttaaccacct 7380
gggtaacctg gtaatcacct ggggcgccca gaccttcaaa caccaagctt tcaacaaact 7440
ggctagcctg ttcatcgtta acaacaaaaa aaccatcccg aacaacctgg ttgaaaacta 7500
cctgaccccg atgtctctgg catactggtt catggatgat ggtggtaaat gggattacaa 7560
caaaaactct accaacaaat cgatcgtact gaacacccag tctttcactt tcgaagaagt 7620
agaatacctg gttaagggtc tgcgtaacaa attccaactg aactgttacg taaaaatcaa 7680
caaaaacaaa ccgatcatct acatcgattc tatgtcttac ctgatcttct acaacctgat 7740
caaaccgtac ctgatcccgc agatgatgta caaactgccg aacactatct cctccgaaac 7800
tttcctgaaa taacctgcag gagctcgaat ttccccgatc gttcaaacat ttggcaataa 7860
agtttcttaa gattgaatcc tgttgccggt cttgcgatga ttatcatata atttctgttg 7920
aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt 7980
tttatgatta gagtcccgca attatacatt taatacgcga tagaaaacaa aatatagcgc 8040
gcaaactagg ataaattatc gcgcgcggtg tcatctatgt tactagatcg ggaattggcg 8100
atcgcagctt ggcgtaatca tggtcatagc tgtttcctac tagatctgat tgtcgtttcc 8160
cgccttcagt ttaaactatc agtgtttgac aggatatatt ggcgggtaaa cctaagagaa 8220
aagagcgttt attagaataa tcggatattt aaaagggcgt gaaaaggttt atccgttcgt 8280
ccatttgtat gtccatgtgt tttatggaca gcaagcgaac cggaattgcc agctggggcg 8340
ccctctggta aggttgggaa gccctgcaaa gtaaactgga tggctttctt gccgccaagg 8400
atctgatggc gcaggggatc aagatctgat caagagacag gatgaggatc gtttcgcatg 8460
attgaacaag atggattgca cgcaggttct ccggccgctt gggtggagag gctattcggc 8520
tatgactggg cacaacagac aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg 8580
caggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag 8640
gacgaggcag cgcggctatc gtggctggcc acgacgggcg ttccttgcgc agctgtgctc 8700
gacgttgtca ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc ggggcaggat 8760
ctcctgtcat cccaccttgc tcctgccgag aaagtatcca tcatggctga tgcaatgcgg 8820
cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcatc 8880
gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct ggacgaagag 8940
catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat gcccgacggc 9000
gaggatctcg tcgtgaccca tggcgatgcc tgcttgccga atatcatggt ggaaaatggc 9060
cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata 9120
gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga ccgcttcctc 9180
gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg ccttcttgac 9240
gagttcttct gaattgaaaa aggaagaatg catgaccaaa atcccttaac gtgagtttc 9300
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt 9360
tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt 9420

gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat 9480
accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc 9540
accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca gtggcgataa 9600
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg 9660
ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag 9720
atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa aggcggacag 9780
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa 9840
cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt 9900
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg 9960
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc 10020
tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac 10080
cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ctgatgcggt attttctcct 10140
tacgcatctg tgcggtattt cacaccgcat atggtgcact ctcagtacaa tctgctctga 10200
tgccgcatag ttaagccagt atacactccg ctatcgctac gtgactgggt catggctgcg 10260
ccccgacacc cgccaacacc cgctgacgcg ccctgacggg cttgtctgct cccggcatcc 10320
gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagaggtt ttcaccgtca 10380
tcaccgaaac gcgcgaggca gggtgccttg atgtgggcgc cggcggtcga gtggcgacgg 10440
cgcggcttgt ccgcgccctg gtagattgcc tggccgtagg ccagccattt ttgagcggcc 10500
agcggccgcg ataggccgac gcgaagcggc ggggcgtagg gagcgcagcg accgaagggt 10560
aggcgctttt tgcagctctt cggctgtgcg ctggccagac agttatgcac aggccaggcg 10620
ggttttaaga gttttaataa gttttaaaga gttttaggcg gaaaaatcgc ctttttttctc 10680
ttttatatca gtcacttaca tgtgtgaccg gttcccaatg tacggctttg ggttcccaat 10740
gtacgggttc cggttcccaa tgtacggctt tgggttccca atgtacgtgc tatccacagg 10800
aaagagacct tttcgacctt tttcccctgc tagggcaatt tgccctagca tctgctccgt 10860
acattaggaa ccggcggatg cttcgccctc gatcaggttg cggtagcgca tgactaggat 10920
cgggccagcc tgccccgcct cctccttcaa atcgtactcc ggcaggtcat ttgacccgat 10980
cagcttgcgc acggtgaaac agaacttctt gaactctccg gcgctgccac tgcgttcgta 11040
gatcgtcttg aacaaccatc tggcttctgc cttgcctgcg gcgcggcgtg ccaggcggta 11100
gagaaaacgg ccgatgccgg gatcgatcaa aaagtaatcg gggtgaaccg tcagcacgtc 11160
cgggttcttg ccttctgtga tctcgcggta catccaatca gctagctcga tctcgatgta 11220
ctccggccgc ccggtttcgc tctttacgat cttgtagcgg ctaatcaagg cttcaccctc 11280
ggataccgtc accaggcggc cgttcttggc cttcttcgta cgctgcatgg caacgtgcgt 11340
ggtgtttaac cgaatgcagg tttctaccag gtcgtctttc tgctttccgc catcggctcg 11400
ccggcagaac ttgagtacgt ccgcaacgtg tggacggaac acgcggccgg gcttgtctcc 11460
cttcccttcc cggtatcggt tcatggattc ggttagatgg gaaaccgcca tcagtaccag 11520
gtcgtaatcc cacacactgg ccatgccggc cggccctgcg gaaacctcta cgtgcccgtc 11580
tggaagctcg tagcggatca cctcgccagc tcgtcggtca cgcttcgaca gacggaaaac 11640
ggccacgtcc atgatgctgc gactatcgcg ggtgcccacg tcatagagca tcggaacgaa 11700
aaaatctggt tgctcgtcgc ccttgggcgg cttcctaatc gacggcgcac cggctgccgg 11760
cggttgccgg gattctttgc ggattcgatc agcggccgct tgccacgatt caccggggcg 11820
tgcttctgcc tcgatgcgtt gccgctgggc ggcctgcgcg gccttcaact tctccaccag 11880
gtcatcaccc agcgccgcgc cgatttgtac cgggccggat ggtttgcgac cgctcacgcc 11940
gattcctcgg gcttgggggt tccagtgcca ttgcagggcc ggcagacaac ccagccgctt 12000
acgcctggcc aaccgcccgt tcctccacac atggggcatt ccacggcgtc ggtgcctggt 12060
tgttcttgat tttccatgcc gcctccttta gccgctaaaa ttcatctact catttattca 12120

tttgctcatt tactctggta gctgcgcgat gtattcagat agcagctcgg taatggtctt 12180
gccttggcgt accgcgtaca tcttcagctt ggtgtgatcc tccgccggca actgaaagtt 12240
gacccgcttc atggctggcg tgtctgccag gctggccaac gttgcagcct tgctgctgcg 12300
tgcgctcgga cggccggcac ttagcgtgtt tgtgcttttg ctcattttct ctttacctca 12360
ttaactcaaa tgagttttga tttaatttca gcggccagcg cctggacctc gcgggcagcg 12420
tcgccctcgg gttctgattc aagaacggtt gtgccggcgg cggcagtgcc tgggtagctc 12480
acgcgctgcg tgatacggga ctcaagaatg ggcagctcgt acccggccag cgcctcggca 12540
acctcaccgc cgatgcgcgt gcctttgatc gcccgcgaca cgacaaaggc cgcttgtagc 12600
cttccatccg tgacctcaat gcgctgctta accagctcca ccaggtcggc ggtggcccat 12660
atgtcgtaag ggcttggctg caccggaatc agcacgaagt cggctgcctt gatcgcggac 12720
acagccaagt ccgccgcctg gggcgctccg tcgatcacta cgaagtcgcg ccggccgatg 12780
gccttcacgt cgcggtcaat cgtcgggcgg tcgatgccga caacggttag cggttgatct 12840
tcccgcacgg ccgcccaatc gcgggcactg ccctggggat cggaatcgac taacagaaca 12900
tcggccccgg cgagttgcag ggcgcgggct agatgggttg cgatggtcgt cttgcctgac 12960
ccgcctttct ggttaagtac agcgataacc ttcatgcgtt ccccttgcgt atttgtttat 13020
ttactcatcg catcatatac gcagcgaccg catgacgcaa gctgttttac tcaaatacac 13080
atcacctttt tagacggcgg cgctcggttt cttcagcggc caagctggcc ggccaggccg 13140
ccagcttggc atcagacaaa ccggccagga tttcatgcag ccgcacggtt gagacgtgcg 13200
cgggcggctc gaacacgtac ccggccgcga tcatctccgc ctcgatctct tcggtaatga 13260
aaaacggttc gtcctggccg tcctggtgcg gtttcatgct tgttcctctt ggcgttcatt 13320
ctcggcggcc gccagggcgt cggcctcggt caatgcgtcc tcacggaagg caccgcgccg 13380
cctggcctcg gtgggcgtca cttcctcgct gcgctcaagt gcgcggtaca gggtcgagcg 13440
atgcacgcca agcagtgcag ccgcctcttt cacggtgcgg ccttcctggt cgatcagctc 13500
gcgggcgtgc gcgatctgtg ccggggtgag g 13531

<210> 12
<211> 10765
<212> DNA
<213> artificial DNA construct (binary vector pCB583-40)

<220>
<223> artificial DNA construct (binary vector pCB583-40)

<220>
<221> misc_feature
<222> (8)..(153)
<223> T-DNA right border

<220>
<221> terminator
<222> (244)..(419)
<223> complement: nos terminator

<220>
<221> misc_feature

<222> (525)..(3144)
<223> overlapping, non-functional helves of GUS gene with a I-SceI and I-CreI site in-between

<220>
<221> promoter
<222> (3251)..(3620)
<223> complement: 35S CaMV pomoter

<220>
<221> terminator
<222> (3786)..(4041)
<223> complement: nos terminator

<220>
<221> misc_feature
<222> (4112)..(4663)
<223> complement: pat gene coding for phosphinotricin resistance

<220>
<221> promoter
<222> (4669)..(4999)
<223> complement: nos pomoter

<220>
<221> misc_feature
<222> (5015)..(5228)
<223> T-DNA left border

<400> 12
ttccatggac atacaaatgg acgaacggat aaacctttc acgccctttt aaatatccga 60
ttattctaat aaacgctctt ttctcttagg tttacccgcc aatatatcct gtcaaacact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttattc cgatctagta acatagatga caccgcgcgc gataatttat 240
cctagtttgc gcgctatatt ttgttttcta tcgcgtatta aatgtataat tgcgggactc 300
taatcataaa aacccatctc ataaataacg tcatgcatta catgttaatt attacatgct 360
taacgtaatt caacagaaat tatatgataa tcatcgcaag tccggcaaca ggattcaatc 420
ttaagaaact ttattgccaa atgtttgaac gatcggggaa attcgagctc ggtagcaatt 480
cccgaggctg tagccgacga tggtgcgcca ggagagttgt tgattcattg tttgcctccc 540
tgctgcggtt tttcaccgaa gttcatgcca gtccagcgtt tttgcagcag aaaagccgcc 600
gacttcggtt tgcggtcgcg agtgaagatc cctttcttgt taccgccaac gcgcaatatg 660
ccttgcgagg tcgcaaaatc ggcgaaattc catacctgtt caccgacgac ggcgctgacg 720
cgatcaaaga cgcggtgata catatccagc catgcacact gatactcttc actccacatg 780
tcggtgtaca ttgagtgcag cccggctaac gtatccacgc cgtattcggt gatgataatc 840
ggctgatgca gtttctcctg ccaggccaga agttcttttt ccagtacctt ctctgccgct 900

tccaaatcgc cgctttggac ataccatccg taataacggt tcaggcacag cacatcaaag 960
agatcgctga tggtatcggt gtgagcgtcg cagaacatta cattgacgca ggtgatcgga 1020
cgcgtcgggt cgagtttacg cgttgcttcc gccagtggcg cgaaatattc ccgtgcacct 1080
tgcggacggg tatccggttc gttggcaata ctccacatca ccacgcttgg gtggtttttg 1140
tcacgcgcta tcagctcttt aatcgcctgt aagtgcgctt gctgagtttc cccgttgact 1200
gcctcttcgc tgtacagttc tttcggcttg ttgcccgctt cgaaaccaat gcctaaagag 1260
aggttaaagc cgacagcagc agtttcatca atcaccacga tgccatgttc atctgcccag 1320
tcgagcatct cttcagcgta agggtaatgc gaggtacggt aggagttggc cccaatccag 1380
tccattaatg cgtggtcgtg caccatcagc acgttatcga atcctttgcc acgcaagtcc 1440
gcatcttcat gacgaccaaa gccagtaaag tagaacggtt tgtggttaat caggaactgt 1500
tcgcccttca ctgccactga ccggatgccg acgcgaagcg ggtagatatc acactctgtc 1560
tggcttttgg ctgtgacgca cagttcatag agataacctt cacccggttg ccagaggtgc 1620
ggattcacca cttgcaaagt cccgctagtg ccttgtccag ttgcaaccac ctgttgatcc 1680
gcatcacgca gttcaacgct gacatcacca ttggccacca cctgccagtc aacagacgcg 1740
tggttacagt cttgcgcgac atgcgtcacc acggtgatat cgtccaccca ggtgttcggc 1800
gtggtgtaga gcattacgct gcgatggatt ccggcatagt taaagaaatc atggaagtaa 1860
gactgctttt tcttgccgtt ttcgtcggta atcaccattc ccggcgggat agtctgccag 1920
ttcagttcgt tgttcacaca aacggtgata cgtacacttt tcccggcaat aacatacggc 1980
gtgacatcgg cttcaaatgg cgtatagccg ccctgatgct ccatcacttc ctgattattg 2040
acccacactt tgccgtaatg agtgaccgca tcgaaacgca gcacgatacg ctggcctgcc 2100
caacctttcg gtataaagac ttcgcgctga taccagacgt tgcccgcata attacgaata 2160
tctgcatcgg cgaactgatc gttaaaactg cctggcacag caattgcccg gctttcttgt 2220
aacgcgcttt cccaccaacg ctgatcaatt ccacagtttt cgcggtccag actgaatgcc 2280
cacaggccgt cgagtttttt gatttcacgg gttggggttt ctacaggact ctagctggca 2340
cagagttacc gggtgaattt cgctacctta ggatcgtaac caatatgtct cacggcgttt 2400
tcggactaga ctattaccct gttatcccta ggcggccgct ggcaccacct gccagtcaac 2460
agacgcgtgg ttacagtctt gcgcgacatg cgtcaccacg gtgatatcgt ccacccaggt 2520
gttcggcgtg gtgtagagca ttacgctgcg atggattccg gcatagttaa agaaatcatg 2580
gaagtaagac tgcttttttct tgccgtttc gtcggtaatc accattcccg gcgggatagt 2640
ctgccagttc agttcgttgt tcacacaaac ggtgatacgt acacttttcc cggcaataac 2700
atacggcgtg acatcggctt caaatggcgt atagccgccc tgatgctcca tcacttcctg 2760
attattgacc cacactttgc cgtaatgagt gaccgcatcg aaacgcagca cgatacgctg 2820
gcctgcccaa cctttcggta taaagacttc gcgctgatac cagacgttgc ccgcataatt 2880
acgaatatct gcatcggcga actgatcgtt aaaactgcct ggcacagcaa ttgcccggct 2940
ttcttgtaac gcgctttccc accaacgctg atcaattcca cagttttcgc gatccagact 3000
gaatgcccac aggccgtcga gtttttgat ttcacgggtt ggggtttcta caggacggat 3060
gtagatcgaa ttgggattgg tgacgttcgt cacctgctta gtctgagtct gagtcttcag 3120
aagtagatga tcccggggat ccatcacttt cttctgtaga ggtttcttct tcctcttctt 3180
tgtattctag gtacgtaccc tgtcctctcc aaatgaaatg aacttcctta tatagaggaa 3240
gggtcttgcg aaggatagtg ggattgtgcg tcatccctta cgtcagtgga gatatcacat 3300
caatccactt gctttgaaga cgtggttgga acgtcttctt tttccacgat gctcctcgtg 3360
ggtgggggtc catctttggg accactgtcg gcagaggcat cttcaacgat ggcctttcct 3420
ttatcgcaat gatggcattt gtaggagcca ccttcctttt ccactatctt cacaataaag 3480
tgacagatag ctgggcaatg gaatccgagg aggtttccgg atattaccct ttgttgaaaa 3540
gtctcaattg ccctttggtc ttctgagact gtatctttga tatttttgga gtagacaagc 3600

gtgtcgtgct ccaccatgtt gacgaagatt ttcttcttgt cattgagtcg taagagactc 3660
tgtatgaact gttcgccagt ctttacggcg agttctgtta ggtcctctat ttgaatcttt 3720
gactcggtac cgagctcgaa ttcactggcc gtcgttttac aacgactcag ccagcttgac 3780
aggaggcccg atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgcgc 3840
gctatatttt gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa 3900
cccatctcat aaataacgtc atgcattaca tgttaattat tacatgctta acgtaattca 3960
acagaaatta tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt 4020
attgccaaat gtttgaacga tcggggatca tccgggtctg tggcgggaac tccacgaaaa 4080
tatccgaacg cagcaagatc ggtcgatcga ctcagatctg ggtaactggc ctaactggcc 4140
ttggaggagc tggcaactca aaatcccttt gccaaaaacc aacatcatgc catccaccat 4200
gcttgtatcc agccgcgcgc aatgtacccc gcgctgtgta tcccaaagcc tcatgcaacc 4260
taacagatgg atcgtttgga aggcctataa cagcaaccac agacttaaaa ccttgcgcct 4320
ccatagactt aagcaaatgt gtgtacaatg tagatcctag gcccaacctt tgatgcctat 4380
gtgacacgta aacagtactc tcaactgtcc aatcgtaagc gttcctagcc ttccagggcc 4440
cagcgtaagc aataccagcc acaacaccct caacctcagc aaccaaccaa gggtatctat 4500
cttgcaacct ctctaggtca tcaatccact cttgtggtgt ttgtggctct gtcctaaagt 4560
tcactgtaga cgtctcaatg taatggttaa cgatgtcaca aaccgcggcc atatcggctg 4620
ctgtagctgg cctaatctca actggtctcc tctccggaga catgtcgaga ttatttggat 4680
tgagagtgaa tatgagactc taattggata ccgaggggaa tttatggaac gtcagtggag 4740
catttttgac aagaaatatt tgctagctga tagtgacctt aggcgacttt tgaacgcgca 4800
ataatggttt ctgacgtatg tgcttagctc attaaactcc agaaacccgc ggctgagtgg 4860
ctccttcaac gttgcggttc tgtcagttcc aaacgtaaaa cggcttgtcc cgcgtcatcg 4920
gcgggggtca taacgtgact cccttaattc tccgctcatg atcagattgt cgtttcccgc 4980
cttcagttta aactatcagt gtttgacagg atcctgcttg gtaataattg tcattagatt 5040
gtttttatgc atagatgcac tcgaaatcag ccaatttttag acaagtatca aacggatgtt 5100
aattcagtac attaaagacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt 5160
tacaccacaa tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac 5220
aaaatcacca cgcgttacca ccacgccggc cggccgcatg gtgttgaccg tgttcgccgg 5280
cattgccgag ttcgagcgtt ccctaatcat cgaccgcacc cggagcgggc gcgaggccgc 5340
caaggcccga ggcgtgaagt ttggcccccg ccctaccctc accccggcac agatcgcgca 5400
cgcccgcgag ctgatcgacc aggaaggccg caccgtgaaa gaggcggctg cactgcttgg 5460
cgtgcatcgc tcgaccctgt accgcgcact tgagcgcagc gaggaagtga cgcccaccga 5520
ggccaggcgg cgcggtgcct tccgtgagga cgcattgacc gaggccgacg ccctggcggc 5580
cgccgagaat gaacgccaag aggaacaagc atgaaaccgc accaggacgg ccaggacgaa 5640
ccgtttttca ttaccgaaga gatcgaggcg gagatgatcg cggccgggta cgtgttcgag 5700
ccgcccgcgc acgtctcaac cgtgcggctg catgaaatcc tggccggttt gtctgatgcc 5760
aagctggcgg cctggccggc cagcttggcc gctgaagaaa ccgagcgccg ccgtctaaaa 5820
aggtgatgtg tatttgagta aaacagcttg cgtcatgcgg tcgctgcgta tatgatgcga 5880
tgagtaaata aacaaatacg caaggggaac gcatgaaggt tatcgctgta cttaaccaga 5940
aaggcgggtc aggcaagacg accatcgcaa cccatctagc ccgcgccctg caactcgccg 6000
gggccgatgt tctgttagtc gattccgatc cccagggcag tgcccgcgat tgggcggccg 6060
tgcgggaaga tcaaccgcta accgttgtcg gcatcgaccg cccgacgatt gaccgcgacg 6120
tgaaggccat cggccggcgc gacttcgtag tgatcgacgg agcgccccag gcggcggact 6180
tggctgtgtc cgcgatcaag gcagccgact tcgtgctgat tccggtgcag ccaagccctt 6240
acgacatatg ggccaccgcc gacctggtgg agctggttaa gcagcgcatt gaggtcacgg 6300

atggaaggct acaagcggcc tttgtcgtgt cgcgggcgat caaaggcacg cgcatcggcg 6360
gtgaggttgc cgaggcgctg gccgggtacg agctgcccat tcttgagtcc cgtatcacgc 6420
agcgcgtgag ctacccaggc actgccgccg ccggcacaac cgttcttgaa tcagaacccg 6480
agggcgacgc tgcccgcgag gtccaggcgc tggccgctga attaaatca aaactcattt 6540
gagttaatga ggtaaagaga aaatgagcaa aagcacaaac acgctaagtg ccggccgtcc 6600
gagcgcacgc agcagcaagg ctgcaacgtt ggccagcctg gcagacacgc cagccatgaa 6660
gcgggtcaac tttcagttgc cggcggagga tcacaccaag ctgaagatgt acgcggtacg 6720
ccaaggcaag accattaccg agctgctatc tgaatacatc gcgcagctac cagagtaaat 6780
gagcaaatga ataaatgagt agatgaattt tagcggctaa aggaggcggc atggaaaatc 6840
aagaacaacc aggcaccgac gccgtggaat gccccatgtg tggaggaacg ggcggttggc 6900
caggcgtaag cggctgggtt gtctgccggc cctgcaatgg cactggaacc cccaagcccg 6960
aggaatcggc gtgagcggtc gcaaaccatc cggcccggta caaatcggcg cggcgctggg 7020
tgatgacctg gtggagaagt tgaaggccgc gcaggccgcc cagcggcaac gcatcgaggc 7080
agaagcacgc cccggtgaat cgtggcaagc ggccgctgat cgaatccgca aagaatcccg 7140
gcaaccgccg gcagccggtg cgccgtcgat taggaagccg cccaagggcg acgagcaacc 7200
agatttttc gttccgatgc tctatgacgt gggcacccgc gatagtcgca gcatcatgga 7260
cgtggccgtt ttccgtctgt cgaagcgtga ccgacgagct ggcgaggtga tccgctacga 7320
gcttccagac gggcacgtag aggtttccgc agggccggcc ggcatggcca gtgtgtggga 7380
ttacgacctg gtactgatgg cggtttccca tctaaccgaa tccatgaacc gataccggga 7440
agggaaggga gacaagcccg gccgcgtgtt ccgtccacac gttgcggacg tactcaagtt 7500
ctgccggcga gccgatggcg gaaagcagaa agacgacctg gtagaaacct gcattcggtt 7560
aaacaccacg cacgttgcca tgcagcgtac gaagaaggcc aagaacggcc gcctggtgac 7620
ggtatccgag ggtgaagcct tgattagccg ctacaagatc gtaaagagcg aaaccgggcg 7680
gccggagtac atcgagatcg agctagctga ttggatgtac cgcgagatca cagaaggcaa 7740
gaacccggac gtgctgacgg ttcaccccga ttacttttg atcgatcccg gcatcggccg 7800
ttttctctac cgcctggcac gccgcgccgc aggcaaggca gaagccagat ggttgttcaa 7860
gacgatctac gaacgcagtg gcagcgccgg agagttcaag aagttctgtt tcaccgtgcg 7920
caagctgatc gggtcaaatg acctgccgga gtacgatttg aaggaggagg cggggcaggc 7980
tggcccgatc ctagtcatgc gctaccgcaa cctgatcgag ggcgaagcat ccgccggttc 8040
ctaatgtacg gagcagatgc tagggcaaat tgccctagca ggggaaaaag gtcgaaaagg 8100
tctctttcct gtggatagca cgtacattgg gaacccaaag ccgtacattg ggaaccggaa 8160
cccgtacatt gggaacccaa agccgtacat tgggaaccgg tcacacatgt aagtgactga 8220
tataaaagag aaaaaaggcg atttttccgc ctaaaactct ttaaaactta ttaaaactct 8280
taaaacccgc ctggcctgtg cataactgtc tggccagcgc acagccgaag agctgcaaaa 8340
agcgcctacc cttcggtcgc tgcgctccct acgccccgcc gcttcgcgtc ggcctatcgc 8400
ggccgctggc cgctcaaaaa tggctggcct acggccaggc aatctaccag ggcgcggaca 8460
agccgcgccg tcgccactcg accgccggcg cccacatcaa ggcaccctgc ctcgcgcgtt 8520
tcggtgatga cggtgaaaac ctctgacaca tgcagctccc ggagacggtc acagcttgtc 8580
tgtaagcgga tgccgggagc agacaagccc gtcagggcgc gtcagcgggt gttggcgggt 8640
gtcggggcgc agccatgacc cagtcacgta gcgatagcgg agtgtatact ggcttaacta 8700
tgcggcatca gagcagattg tactgagagt gcaccatatg cggtgtgaaa taccgcacag 8760
atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct tcctcgctca ctgactcgct 8820
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 8880
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 8940
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 9000

gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 9060
ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 9120
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 9180
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 9240
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 9300
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 9360
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt 9420
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 9480
atccggcaaa caaaccaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac 9540
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 9600
gtggaacgaa aactcacgtt aagggatttt ggtcatgcat gatatatctc ccaatttgtg 9660
tagggcttat tatgcacgct taaaaataat aaaagcagac ttgacctgat agtttggctg 9720
tgagcaatta tgtgcttagt gcatctaacg cttgagttaa gccgcgccgc gaagcggcgt 9780
cggcttgaac gaatttctag ctagacatta tttgccgact accttggtga tctcgccttt 9840
cacgtagtgg acaaattctt ccaactgatc tgcgcgcgag gccaagcgat cttcttcttg 9900
tccaagataa gcctgtctag cttcaagtat gacgggctga tactgggccg gcaggcgctc 9960
cattgcccag tcggcagcga catccttcgg cgcgattttg ccggttactg cgctgtacca 10020
aatgcgggac aacgtaagca ctacatttcg ctcatcgcca gcccagtcgg gcggcgagtt 10080
ccatagcgtt aaggtttcat ttagcgcctc aaatagatcc tgttcaggaa ccggatcaaa 10140
gagttcctcc gccgctggac ctaccaaggc aacgctatgt tctcttgctt ttgtcagcaa 10200
gatagccaga tcaatgtcga tcgtggctgg ctcgaagata cctgcaagaa tgtcattgcg 10260
ctgccattct ccaaattgca gttcgcgctt agctggataa cgccacggaa tgatgtcgtc 10320
gtgcacaaca atggtgactt ctacagcgcg gagaatctcg ctctctccag gggaagccga 10380
agtttccaaa aggtcgttga tcaaagctcg ccgcgttgtt tcatcaagcc ttacggtcac 10440
cgtaaccagc aaatcaatat cactgtgtgg cttcaggccg ccatccactg cggagccgta 10500
caaatgtacg gccagcaacg tcggttcgag atggcgctcg atgacgccaa ctacctctga 10560
tagttgagtc gatacttcgg cgatcaccgc ttcccccatg atgtttaact ttgttttagg 10620
gcgactgccc tgctgcgtaa catcgttgct gctccataac atcaaacatc gacccacggc 10680
gtaacgcgct tgctgcttgg atgcccgagg catagactgt accccaaaaa aacagtcata 10740
acaagccatg aaaaccgcca ctgcg 10765

<210> 13
<211> 9103
<212> DNA
<213> artificial DNA construct (binary vector pRS8)

<220>
<223> artificial DNA construct (binary vector pRS8)

<220>
<221> misc_feature
<222> (8)..(153)
<223> T-DNA right border

<220>

<221> terminator
<222> (252)..(490)
<223> complement: ocs terminator

<220>
<221> misc_feature
<222> (501)..(1208)
<223> complement: coding for I-SceI

<220>
<221> promoter
<222> (1280)..(1809)
<223> complement: 35S CaMV promoter

<220>
<221> terminator
<222> (1856)..(2111)
<223> complement: nos terminator

<220>
<221> misc_feature
<222> (2204)..(2982)
<223> complement: nos gene (kanamycin resistance)

<220>
<221> promoter
<222> (3003)..(3339)
<223> complement: nos promoter

<220>
<221> misc_feature
<222> (3352)..(3566)
<223> T-DNA left border

<400> 13
ttccatggac atacaaatgg acgaacggat aaacctttc acgccctttt aaatatccga 60
ttattctaat aaacgctctt ttctcttagg tttacccgcc aatatatcct gtcaaacact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttgcat gcctgcaggt cgactctaga ctagtggatc cgatatcgcc 240
cgggctcgag agtgccaagc ttggacaatc agtaaattga acggagaata ttattcataa 300
aaatacgata gtaacgggtg atatattcat tagaatgaac cgaaaccggc ggtaaggatc 360
tgagctacac atgctcaggt tttttacaac gtgcacaaca gaattgaaag caaatatcat 420
gcgatcatag gcgtctcgca tatctcatta aagcagggca tgcctgcagg tcgactctag 480
aggatccagt actcgactta ttatttcagg aaagtttcgg aggagatagt gttcggcagt 540
ttgtacatca tctgcgggat caggtacggt ttgatcaggt tgtagaagat caggtaagac 600

atagaatcga tgtagatgat cggtttgttt ttgttgattt ttacgtaaca gttcagttgg 660
aatttgttac gcagaccctt aaccaggtat tctacttctt cgaaagtgaa agactgggtg 720
ttcagtacga tcgatttgtt ggtagagttt ttgttgtaat cccatttacc accatcatcc 780
atgaaccagt atgccagaga catcggggtc aggtagtttt caaccaggtt gttcgggatg 840
gttttttttgt tgttaacgat gaacaggcta gccagtttgt tgaaagcttg gtgtttgaaa 900
gtctgggcgc cccaggtgat taccaggtta cccaggtggt taacacgttc ttttttgtgc 960
ggcggggaca gtacccactg atcgtacagc agacatacgt ggtccatgta tgctttgttt 1020
ttccactcga actgcataca gtaggtttta ccttcatcac gagaacggat gtaagcatca 1080
cccaggatca gaccgatacc tgcttcgaac tgttcgatgt tcagttcgat cagctgggat 1140
ttgtattctt tcagcagttt agagttcgga cccaggttca ttacctggtt ttttttgatg 1200
tttttcatca cttcttctgt agaggtttct tcttcctctt ctttgtattc taggtacagt 1260
actggatccc cgggtaccct gtcctctcca aatgaaatga acttccttat atagaggaag 1320
ggtcttgcga aggatagtgg gattgtgcgt catcccttac gtcagtggag atatcacatc 1380
aatccacttg ctttgaagac gtggttggaa cgtcttcttt ttccacgatg ctcctcgtgg 1440
gtgggggtcc atctttggga ccactgtcgg cagaggcatc ttcaacgatg gcctttcctt 1500
tatcgcaatg atggcatttg taggagccac cttccttttc cactatcttc acaataaagt 1560
gacagatagc tgggcaatgg aatccgagga ggtttccgga tattaccctt tgttgaaaag 1620
tctcaattgc cctttggtct tctgagactg tatctttgat atttttggag tagacaagcg 1680
tgtcgtgctc caccatgttg acgaagattt tcttcttgtc attgagtcgt aagagactct 1740
gtatgaactg ttcgccagtc tttacggcga gttctgttag gtcctctatt tgaatctttg 1800
actccatggg aattcactgg ccgtcgtttt acaacgactc agagcttgac aggaggcccg 1860
atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgcgc gctatatttt 1920
gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa cccatctcat 1980
aaataacgtc atgcattaca tgttaattat tacatgctta acgtaattca acagaaatta 2040
tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt attgccaaat 2100
gtttgaacga tcggggatca tccgggtctg tggcgggaac tccacgaaaa tatccgaacg 2160
cagcaagatc tagagcttgg gtcccgctca gaagaactcg tcaagaaggc gatagaaggc 2220
gatgcgctgc gaatcgggag cggcgatacc gtaaagcacg aggaagcggt cagcccattc 2280
gccgccaagc tcttcagcaa tatcacgggt agccaacgct atgtcctgat agcggtccgc 2340
cacacccagc cggccacagt cgatgaatcc agaaaagcgg ccattttcca ccatgatatt 2400
cggcaagcag gcatcgccat gggtcacgac gagatcctcg ccgtcgggca tgcgcgcctt 2460
gagcctggcg aacagttcgg ctggcgcgag cccctgatgc tcttcgtcca gatcatcctg 2520
atcgacaaga ccggcttcca tccgagtacg tgctcgctcg atgcgatgtt tcgcttggtg 2580
gtcgaatggg caggtagccg gatcaagcgt atgcagccgc cgcattgcat cagccatgat 2640
ggatactttc tcggcaggag caaggtgaga tgacaggaga tcctgccccg gcacttcgcc 2700
caatagcagc cagtcccttc ccgcttcagt gacaacgtcg agcacagctg cgcaaggaac 2760
gcccgtcgtg gccagccacg atagccgcgc tgcctcgtcc tgcagttcat tcagggcacc 2820
ggacaggtcg gtcttgacaa aaagaaccgg gcgcccctgc gctgacagcc ggaacacggc 2880
ggcatcagag cagccgattg tctgttgtgc ccagtcatag ccgaatagcc tctccaccca 2940
agcggccgga gaacctgcgt gcaatccatc ttgttcaatc atgcgaaacg atccagatcc 3000
ggtgcagatt atttggattg agagtgaata tgagactcta attggatacc gaggggaatt 3060
tatggaacgt cagtggagca tttttgacaa gaaatatttg ctagctgata gtgaccttag 3120
gcgacttttg aacgcgcaat aatggtttct gacgtatgtg cttagctcat taaactccag 3180
aaacccgcgg ctgagtggct ccttcaacgt tgcggttctg tcagttccaa acgtaaaacg 3240
gcttgtcccg cgtcatcggc gggggtcata acgtgactcc cttaattctc cgctcatgat 3300

cagattgtcg tttcccgcct tcagtttaaa ctatcagtgt ttgacaggat cctgcttggt 3360
aataattgtc attagattgt ttttatgcat agatgcactc gaaatcagcc aattttagac 3420
aagtatcaaa cggatgttaa ttcagtacat taaagacgtc cgcaatgtgt tattaagttg 3480
tctaagcgtc aatttgttta caccacaata tatcctgcca ccagccagcc aacagctccc 3540
cgaccggcag ctcggcacaa aatcaccacg cgttaccacc acgccggccg gccgcatggt 3600
gttgaccgtg ttcgccggca ttgccgagtt cgagcgttcc ctaatcatcg accgcacccg 3660
gagcgggcgc gaggccgcca aggcccgagg cgtgaagttt ggcccccgcc ctaccctcac 3720
cccggcacag atcgcgcacg cccgcgagct gatcgaccag gaaggccgca ccgtgaaaga 3780
ggcggctgca ctgcttggcg tgcatcgctc gaccctgtac cgcgcacttg agcgcagcga 3840
ggaagtgacg cccaccgagg ccaggcggcg cggtgccttc cgtgaggacg cattgaccga 3900
ggccgacgcc ctggcggccg ccgagaatga acgccaagag gaacaagcat gaaaccgcac 3960
caggacggcc aggacgaacc gtttttcatt accgaagaga tcgaggcgga gatgatcgcg 4020
gccgggtacg tgttcgagcc gcccgcgcac gtctcaaccg tgcggctgca tgaaatcctg 4080
gccggtttgt ctgatgccaa gctggcggcc tggccggcca gcttggccgc tgaagaaacc 4140
gagcgccgcc gtctaaaaag gtgatgtgta tttgagtaaa acagcttgcg tcatgcggtc 4200
gctgcgtata tgatgcgatg agtaaataaa caaatacgca aggggaacgc atgaaggtta 4260
tcgctgtact taaccagaaa ggcgggtcag gcaagacgac catcgcaacc catctagccc 4320
gcgccctgca actcgccggg gccgatgttc tgttagtcga ttccgatccc cagggcagtg 4380
cccgcgattg ggcggccgtg cgggaagatc aaccgctaac cgttgtcggc atcgaccgcc 4440
cgacgattga ccgcgacgtg aaggccatcg gccggcgcga cttcgtagtg atcgacggag 4500
cgccccaggc ggcggacttg gctgtgtccg cgatcaaggc agccgacttc gtgctgattc 4560
cggtgcagcc aagcccttac gacatatggg ccaccgccga cctggtggag ctggttaagc 4620
agcgcattga ggtcacggat ggaaggctac aagcggcctt tgtcgtgtcg cgggcgatca 4680
aaggcacgcg catcggcggt gaggttgccg aggcgctggc cgggtacgag ctgcccattc 4740
ttgagtcccg tatcacgcag cgcgtgagct acccaggcac tgccgccgcc ggcacaaccg 4800
ttcttgaatc agaacccgag ggcgacgctg cccgcgaggt ccaggcgctg gccgctgaaa 4860
ttaaatcaaa actcatttga gttaatgagg taaagagaaa atgagcaaaa gcacaaacac 4920
gctaagtgcc ggccgtccga gcgcacgcag cagcaaggct gcaacgttgg ccagcctggc 4980
agacacgcca gccatgaagc gggtcaactt tcagttgccg gcggaggatc acaccaagct 5040
gaagatgtac gcggtacgcc aaggcaagac cattaccgag ctgctatctg aatacatcgc 5100
gcagctacca gagtaaatga gcaaatgaat aaatgagtag atgaatttta gcggctaaag 5160
gaggcggcat ggaaaatcaa gaacaaccag gcaccgacgc cgtggaatgc cccatgtgtg 5220
gaggaacggg cggttggcca ggcgtaagcg gctgggttgt ctgccggccc tgcaatggca 5280
ctggaacccc caagcccgag gaatcggcgt gagcggtcgc aaaccatccg gcccggtaca 5340
aatcggcgcg gcgctgggtg atgacctggt ggagaagttg aaggccgcgc aggccgccca 5400
gcggcaacgc atcgaggcag aagcacgccc cggtgaatcg tggcaagcgg ccgctgatcg 5460
aatccgcaaa gaatcccggc aaccgccggc agccggtgcg ccgtcgatta ggaagccgcc 5520
caagggcgac gagcaaccag atttttttcgt tccgatgctc tatgacgtgg gcacccgcga 5580
tagtcgcagc atcatggacg tggccgtttt ccgtctgtcg aagcgtgacc gacgagctgg 5640
cgaggtgatc cgctacgagc ttccagacgg gcacgtagag gtttccgcag ggccggccgg 5700
catggccagt gtgtgggatt acgacctggt actgatggcg gtttcccatc taaccgaatc 5760
catgaaccga taccgggaag ggaagggaga caagcccggc cgcgtgttcc gtccacacgt 5820
tgcggacgta ctcaagttct gccggcgagc cgatggcgga aagcagaaag acgacctggt 5880
agaaacctgc attcggttaa acaccacgca cgttgccatg cagcgtacga agaaggccaa 5940
gaacggccgc ctggtgacgg tatccgaggg tgaagccttg attagccgct acaagatcgt 6000

aaagagcgaa accgggcggc cggagtacat cgagatcgag ctagctgatt ggatgtaccg 6060
cgagatcaca gaaggcaaga acccggacgt gctgacggtt caccccgatt acttttgat 6120
cgatcccggc atcggccgtt ttctctaccg cctggcacgc cgcgccgcag gcaaggcaga 6180
agccagatgg ttgttcaaga cgatctacga acgcagtggc agcgccggag agttcaagaa 6240
gttctgtttc accgtgcgca agctgatcgg gtcaaatgac ctgccggagt acgatttgaa 6300
ggaggaggcg gggcaggctg gcccgatcct agtcatgcgc taccgcaacc tgatcgaggg 6360
cgaagcatcc gccggttcct aatgtacgga gcagatgcta gggcaaattg ccctagcagg 6420
ggaaaaaggt cgaaaaggtc tctttcctgt ggatagcacg tacattggga acccaaagcc 6480
gtacattggg aaccggaacc cgtacattgg gaacccaaag ccgtacattg ggaaccggtc 6540
acacatgtaa gtgactgata taaaagagaa aaaaggcgat ttttccgcct aaaactcttt 6600
aaaacttatt aaaactctta aaacccgcct ggcctgtgca taactgtctg gccagcgcac 6660
agccgaagag ctgcaaaaag cgcctaccct tcggtcgctg cgctccctac gccccgccgc 6720
ttcgcgtcgg cctatcgcgg ccgctggccg ctcaaaaatg gctggcctac ggccaggcaa 6780
tctaccaggg cgcggacaag ccgcgccgtc gccactcgac cgccggcgcc cacatcaagg 6840
caccctgcct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg 6900
agacggtcac agcttgtctg taagcggatg ccgggagcag acaagcccgt cagggcgcgt 6960
cagcgggtgt tggcgggtgt cggggcgcag ccatgaccca gtcacgtagc gatagcggag 7020
tgtatactgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatgcg 7080
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgct cttccgcttc 7140
ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc 7200
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 7260
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 7320
gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 7380
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 7440
tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 7500
ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 7560
ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct 7620
tgagtccaac ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat 7680
tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg 7740
ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 7800
aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt 7860
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 7920
tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgcatga 7980
tatatctccc aatttgtgta gggcttatta tgcacgctta aaaataataa aagcagactt 8040
gacctgatag tttggctgtg agcaattatg tgcttagtgc atctaacgct tgagttaagc 8100
cgcgccgcga agcggcgtcg gcttgaacga atttctagct agacattatt tgccgactac 8160
cttggtgatc tcgcctttca cgtagtggac aaattcttcc aactgatctg cgcgcgaggc 8220
caagcgatct tcttcttgtc caagataagc ctgtctagct tcaagtatga cgggctgata 8280
ctgggccggc aggcgctcca ttgcccagtc ggcagcgaca tccttcggcg cgattttgcc 8340
ggttactgcg ctgtaccaaa tgcgggacaa cgtaagcact acatttcgct catcgccagc 8400
ccagtcgggc ggcgagttcc atagcgttaa ggtttcattt agcgcctcaa atagatcctg 8460
ttcaggaacc ggatcaaaga gttcctccgc cgctggacct accaaggcaa cgctatgttc 8520
tcttgctttt gtcagcaaga tagccagatc aatgtcgatc gtggctggct cgaagatacc 8580
tgcaagaatg tcattgcgct gccattctcc aaattgcagt tcgcgcttag ctggataacg 8640
ccacggaatg atgtcgtcgt gcacaacaat ggtgacttct acagcgcgga gaatctcgct 8700

ctctccaggg gaagccgaag tttccaaaag gtcgttgatc aaagctcgcc gcgttgtttc 8760
atcaagcctt acggtcaccg taaccagcaa atcaatatca ctgtgtggct tcaggccgcc 8820
atccactgcg gagccgtaca aatgtacggc cagcaacgtc ggttcgagat ggcgctcgat 8880
gacgccaact acctctgata gttgagtcga tacttcggcg atcaccgctt cccccatgat 8940
gtttaacttt gttttagggc gactgccctg ctgcgtaaca tcgttgctgc tccataacat 9000
caaacatcga cccacggcgt aacgcgcttg ctgcttggat gcccgaggca tagactgtac 9060
cccaaaaaaa cagtcataac aagccatgaa aaccgccact gcg 9103

<210> 14
<211> 897
<212> DNA
<213> Artificial

<220>
<223> artificial sequence coding for I-SceI with inserted intron

<220>
<221> Intron
<222> (565)..(753)
<223> PIV2 intron

<220>
<221> Intron
<222> (565)..(753)
<223> PIV2 intron

<400> 14
atgaaaaaca tcaaaaaaaa ccaggtaatg aacctgggtc cgaactctaa actgctgaaa 60
gaatacaaat cccagctgat cgaactgaac atcgaacagt tcgaagcagg tatcggtctg 120
atcctgggtg atgcttacat ccgttctcgt gatgaaggta aaacctactg tatgcagttc 180
gagtggaaaa acaaagcata catggaccac gtatgtctgc tgtacgatca gtgggtactg 240
tccccgccgc acaaaaaaga acgtgttaac cacctgggta acctggtaat cacctggggc 300
gcccagactt tcaaacacca agctttcaac aaactggcta gcctgttcat cgttaacaac 360
aaaaaaacca tcccgaacaa cctggttgaa aactacctga ccccgatgtc tctggcatac 420
tggttcatgg atgatggtgg taaatgggat tacaacaaaa actctaccaa caaatcgatc 480
gtactgaaca cccagtcttt cactttcgaa gaagtagaat acctggttaa gggtctgcgt 540
aacaaattcc aactgaactg ttacgtaagt ttctgcttct acctttgata tatatataat 600
aattatcatt aattagtagt aatataatat ttcaaatatt tttttcaaaa taaaagaatg 660
tagtatatag caattgcttt tctgtagttt ataagtgtgt atattttaat ttataacttt 720
tctaatatat gaccaaaatt tgttgatgtg caggtaaaaa tcaacaaaaa caaaccgatc 780
atctacatcg attctatgtc ttacctgatc ttctacaacc tgatcaaacc gtacctgatc 840
ccgcagatga tgtacaaact gccgaacact atctcctccg aaactttcct gaaataa 897

<210> 15
<211> 983

<212> DNA
<213> Petroselinum crispum

<220>
<221> promoter
<222> (1)..(983)
<223> parsley ubiquitin promoter (alternative sequence)

<400> 15

gaattcgaat ccaaaaatta cggatatgaa tataggcata tccgtatccg aattatccgt 60
ttgacagcta gcaacgattg tacaattgct tctttaaaaa aggaagaaag aaagaaagaa 120
aagaatcaac atcagcgtta acaaacggcc ccgttacggc ccaaacggtc atatagagta 180
acggcgttaa gcgttgaaag actcctatcg aaatacgtaa ccgcaaacgt gtcatagtca 240
gatcccctct tccttcaccg cctcaaacac aaaaataatc ttctacagcc tatatataca 300
acccccccct ctatctctcc tttctcacaa ttcatcatct ttctttctct acccccaatt 360
ttaagaaatc ctctcttctc ctcttcattt tcaaggtaaa tctctctctc tctctctctc 420
tctgttattc cttgttttaa ttaggtatgt attattgcta gtttgttaat ctgcttatct 480
tatgtatgcc ttatgtgaat atctttatct tgttcatctc atccgtttag aagctataaa 540
tttgttgatt tgactgtgta tctacacgtg gttatgttta tatctaatca gatatgaatt 600
tcttcatatt gttgcgtttg tgtgtaccaa tccgaaatcg ttgatttttt tcatttaatc 660
gtgtagctaa ttgtacgtat acatatggat ctacgtatca attgttcatc tgtttgtgtt 720
tgtatgtata cagatctgaa aacatcactt ctctcatctg attgtgttgt tacatacata 780
gatatagatc tgttatatca tttttttttat taattgtgta tatatatatg tgcatagatc 840
tggattacat gattgtgatt atttacatga ttttgttatt tacgtatgta tatatgtaga 900
tctggacttt ttggagttgt tgacttgatt gtatttgtgt gtgtatatgt gtgttctgat 960
cttgatatgt tatgtatgtg cag 983