



(22) Date de dépôt/Filing Date: 1997/05/01
(41) Mise à la disp. pub./Open to Public Insp.: 1997/11/06
(45) Date de délivrance/Issue Date: 2010/11/16
(62) Demande originale/Original Application: 2 252 412
(30) Priorité/Priority: 1996/05/01 (US US 60/016,345)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01),
A01H 1/04 (2006.01), *A01H 5/00* (2006.01),
C12N 15/12 (2006.01), *C12N 15/65* (2006.01),
C12N 5/10 (2006.01), *G01N 33/58* (2006.01)

(72) Inventeurs/Inventors:
GORDON-KAMM, WILLIAM, US;
PIERCE, DOROTHY A., US;
BOWEN, BENJAMIN, US;
BIDNEY, DENNIS, US;
ROSS, MARGIT, US;
...

(73) Propriétaire/Owner:
PIONEER HI-BRED INTERNATIONAL, INC., US

(74) Agent: TORYS LLP

(54) Titre : UTILISATION DE LA PROTEINE FLUORESCENTE VERTE COMME MARQUEUR DE CRIBLAGE POUR LA TRANSFORMATION DE PLANTES
(54) Title: USE OF THE GREEN FLUORESCENT PROTEIN AS A SCREENABLE MARKER FOR PLANT TRANSFORMATION

```

1  ATGTCCAAGGGCGAGGAGCTCTTCACCGGCGTGGTGCCCATCCTCGTGAGCTCGACGGC
   -----+-----+-----+-----+-----+-----+ 60
   TACAGGTTCCCGCTCCTCGAGAAGTGGCCGCACCACGGGTAGGAGCACCTCGAGCTGCCG
   M S K G E E L P T G V V P I L V E L D G
61  GACGTGAACGGCCACAAGTTCTCCGTGTCCGGCGAGGGCGAGGGCGACGCCACCTACGGC
   -----+-----+-----+-----+-----+ 120
   CTGCACTTGCCGGTGTTCAAGAGGCACAGGCCGCTCCCGCTCCCGCTGCGGTGGATGCCG
   D V N G H K F S V S G E G E G D A T Y G
121 AAGCTCACCTCAAGTTCATCTGCACCACCGGCAAGCTCCCCGTGCCCTGGCCCCACCTC
   -----+-----+-----+-----+-----+ 180
   TTCGAGTGGGAGTTCAAGTAGACGTGGTGGCCGTTTCGAGGGGCACGGGACCGGGTGGGAG
   K L T L K F I C T T G K L P V P W P T L
181 GTGACCACCTTCTCCTACGGCGTGCAGTGCTTCTCCAGGTACCCCGACCATGAAGCAG
   -----+-----+-----+-----+-----+ 240
   CACTGGTGGAAGAGGATGCCGCACGTCACGAAGAGGTCCATGGGGCTGGTGTACTTCGTC
   V T T F S Y G V Q C F S R Y P D H M K Q
241 CACGACTTCTTCAAGTCAGCCATGCCCGAGGGCTACGTGCAGGAGAGGACCATCTTCTTC
   -----+-----+-----+-----+-----+ 300
   GTGCTGAAGAAGTTCAGTCGGTACGGGCTCCCGATGCACGTCCTCTCCTGGTAGAAGAAG
   H D F F K S A M P E G Y V Q E R T I F F
301 AAGGACGACGGCAACTACAAGACCAGGGCCGAGGTGAAGTTCGAAGGGGACACCTCGTC
   -----+-----+-----+-----+-----+ 360
   TTCCTGCTGCCGTTGATGTTCTGGTCCCGGCTCCACTTCAAGCTTCCGCTGTGGGAGCAC
   K D D G N Y K T R A E V K F E G D T L V
```

(57) Abrégé/Abstract:

A method for the production of transgenic plants is provided in which a vector carrying a gene encoding the green fluorescent protein is introduced into cells, the cells are screened for the protein and transformed cells are selected and regenerated. The



(72) **Inventeurs(suite)/Inventors(continued):** SCELONGE, CHRISTOPHER, US; MILLER, MIKE, US; SANDAHL, GARY, US; WANG, LIJUAN, US

(57) **Abrégé(suite)/Abstract(continued):**

cellular toxicity of the green fluorescent protein is circumvented by regulating expression of the gene encoding the protein or directing the protein to a subcellular compartment where it is not toxic to the cell. DNA constructs are provided for cell transformation in which the expression of a gene encoding the green fluorescent protein is placed under the control of an inducible promoter. In addition, DNA constructs are provided in which a nucleotide sequence encoding the green fluorescent protein is operably linked to a signal sequence which directs the expressed protein to a subcellular compartment where the protein is not toxic to the cell. Oxidative stress to plant cells transformed with GFP also can be ameliorated by transforming cells with an expression vector comprising genes encoding GFP and an oxygen scavenger enzyme such as superoxide dismutase. The toxicity of GFP in transformed plants can be eliminated by excising the screenable marker gene following detection of transformed cells or sectors. The FLP/FRT system is used in conjunction with GFP as a visible marker for transformation and FRT excision. A nucleotide sequence optimized for expression of the green fluorescent protein in plants is also provided.

**CORRECTED
VERSION***

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, 15/82, C07K 14/435 A01H 5/00		A3	(11) International Publication Number: WO 97/41228
			(43) International Publication Date: 6 November 1997 (06.11.97)
(21) International Application Number: PCT/US97/07688			
(22) International Filing Date: 1 May 1997 (01.05.97)			
(30) Priority Data: 60/016,345 1 May 1996 (01.05.96) US			
(60) Parent Application or Grant (63) Related by Continuation US 60/016,345 (CIP) Filed on 1 May 1996 (01.05.96)			
(71) Applicant (for all designated States except US): PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; Darwin Building, 7100 N.W. 62nd Avenue, P.O. Box 1000, Johnston, IA 50131 (US).		(74) Agents: BENT, Stephen, A. et al.; Foley & Lardner, Suite 500, 3000 K Street, N.W., Washington, DC 20007-5109 (US).	
(72) Inventors; and (75) Inventors/Applicants (for US only): GORDON-KAMM, William [US/US]; 3916 67th Street, Urbandale, IA 50322 (US). PIERCE, Dorothy, A. [US/US]; 9124 Hammontree Drive, Urbandale, IA 50322 (US). BOWEN, Benjamin [GB/US]; 3008 36th Street, Des Moines, IA 50310 (US). BIDNEY, Dennis [US/US]; 8385 Plum Drive, Urbandale, IA 50322 (US). ROSS, Margit [US/US]; 5534 Kensington		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
		(88) Date of publication of the international search report: 11 December 1997 (11.12.97)	
(54) Title: USE OF THE GREEN FLUORESCENT PROTEIN AS A SCREENABLE MARKER FOR PLANT TRANSFORMATION			
(57) Abstract A method for the production of transgenic plants is provided in which a vector carrying a gene encoding the green fluorescent protein is introduced into cells, the cells are screened for the protein and transformed cells are selected and regenerated. The cellular toxicity of the green fluorescent protein is circumvented by regulating expression of the gene encoding the protein or directing the protein to a subcellular compartment where it is not toxic to the cell. DNA constructs are provided for cell transformation in which the expression of a gene encoding the green fluorescent protein is placed under the control of an inducible promoter. In addition, DNA constructs are provided in which a nucleotide sequence encoding the green fluorescent protein is operably linked to a signal sequence which directs the expressed protein to a subcellular compartment where the protein is not toxic to the cell. Oxidative stress to plant cells transformed with GFP also can be ameliorated by transforming cells with an expression vector comprising genes encoding GFP and an oxygen scavenger enzyme such as superoxide dismutase. The toxicity of GFP in transformed plants can be eliminated by excising the screenable marker gene following detection of transformed cells or sectors. The FLP/FRT system is used in conjunction with GFP as a visible marker for transformation and FRT excision. A nucleotide sequence optimized for expression of the green fluorescent protein in plants is also provided.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

**USE OF THE GREEN FLUORESCENT PROTEIN AS A SCREENABLE
MARKER FOR PLANT TRANSFORMATION**

BACKGROUND OF THE INVENTION

1. Field of the Invention

5 The present invention relates to a method of plant transformation in which a DNA construct carrying a gene encoding the green fluorescent protein (GFP) is introduced into plant cells which are then screened for the presence of the protein and transformed cells are regenerated into transgenic plants. In particular, the present invention provides methods for circumventing the cellular toxicity of the GFP by regulating expression of the gene encoding the protein or directing the protein to a subcellular compartment where it is not toxic to the cell. The present invention provides DNA constructs for cell transformation in which expression of a gene encoding the GFP is placed under the control of an inducible, constitutive or tissue-specific promoter. In addition, DNA constructs are provided in which a nucleotide sequence encoding the GFP is operably linked to a signal or targeting sequence which directs the expressed protein to a subcellular compartment where the protein is not toxic to the cell. Moreover, the present invention provides a nucleotide sequence encoding GFP that is optimized for expression of the GFP gene in

10

15

20

25

plants and to GFP-encoding nucleotide sequences that code for light-shifted versions of GFP. The present invention also provides a method for selecting plant cells transformed with a gene encoding a screenable marker flanked on the 5-prime and 3-prime ends with a recombinase-specific target sequence, and introducing a gene encoding a site specific recombinase into the transformed plant cells and selecting transformed plant cells that no longer express the screenable marker. In addition, the present invention provides a method of reducing GFP toxicity by transforming plant cells with a gene encoding the GFP together with a gene encoding an oxygen scavenger such as superoxidase dismutase.

2. Background

Expression vectors include at least one genetic marker that allows transformed cells to be either recovered by negative selection, i.e. inhibiting growth of cells that do not contain the selectable marker gene, or by screening for product encoded by the genetic marker. Many of the commonly used selectable marker genes for plant transformation were isolated from bacteria and code for enzymes that metabolically detoxify a selective chemical agent which may be an antibiotic or a herbicide. Other selectable marker genes encode an altered target which is insensitive to the inhibitor.

The most commonly used selectable marker gene for plant transformation is the neomycin phosphotransferase II (*nptII*) gene, isolated from Tn5, which when placed under the control of plant regulatory signals confers resistance to kanamycin. Fraley et al., *Proc. Natl. Acad. Sci. U.S.A.*, 80: 4803 (1983). Another commonly used selectable marker gene is the hygromycin phosphotransferase gene which confers resistance to the antibiotic hygromycin. Vanden Elzen et al., *Plant Mol. Biol.*, 5: 299 (1985).

Additional selectable marker genes of bacterial origin that confer resistance to antibiotics include gentamycin acetyl transferase, streptomycin phosphotransferase, aminoglycoside-3'-adenyl transferase, the bleomycin resistance determinant. Hayford et al., *Plant Physiol.* 86: 1216 (1988), Jones et al., *Mol. Gen. Genet.*, 210: 86 (1987), Svab et al., *Plant Mol. Biol.* 14: 197 (1990), Hille et al., *Plant Mol. Biol.* 7: 171 (1986). Other selectable marker genes confer resistance to herbicides such as glyphosate, glufosinate or broxynil. Comai et al., *Nature* 317: 741-744 (1985), Gordon-Kamm et al., *Plant Cell* 2: 603-618 (1990) and Stalker et al., *Science* 242: 419-423 (1988).

Other selectable marker genes for plant transformation are not of bacterial origin. These genes include, for example, mouse dihydrofolate reductase, plant 5-enolpyruvylshikimate-3-phosphate synthase and plant acetolactate synthase. Eichholtz et al., *Somatic Cell Mol. Genet.* 13: 67 (1987), Shah et al., *Science* 233: 478 (1986), Charest et al., *Plant Cell Rep.* 8: 643 (1990).

Although many of these markers have been used for selecting transformed plant tissue, these selection systems involving toxic chemical agents can have disadvantages or limitations. One disadvantage is that it may be difficult to recover normal, viable transformed plants directly from chemical selection. Everett et al., *Bio/Technology* 5: 1201-1204 (1987). Another disadvantage is that not all selectable marker systems work for all tissues, in all plant species, due in part to differences in sensitivity of a particular tissue or plant species to the selective agent. The success of any given marker for transformation of a given plant species is not easily predicted. Moreover, potential regulatory issues surrounding the use of antibiotic resistance genes and the use of herbicide resistance genes for plant species capable of outcrossing with weedy species are additional disadvantages of these markers.

Another class of marker genes for plant transformation require screening of presumptively transformed plant cells rather than direct genetic selection of transformed cells for resistance to a toxic substance such as an antibiotic. These genes are particularly useful to quantify or visualize the spatial pattern of expression of a gene in specific tissues and are frequently referred to as reporter genes because they can be fused to a gene or gene regulatory sequence for the investigation of gene expression. Commonly used genes for screening presumptively transformed cells include β -glucuronidase (GUS), β -galactosidase, luciferase, and chloramphenicol acetyltransferase. Jefferson, R.A., *Plant Mol. Biol. Rep.* 5: 387 (1987), Teeri et al., *EMBO J.* 8: 343 (1989), Koncz et al., *Proc. Natl. Acad. Sci. U.S.A.* 84: 131 (1987), De Block et al., *EMBO J.* 3: 1681 (1984). Another approach to the identification of relatively rare transformation events has been use of a gene that encodes a dominant constitutive regulator of the *Zea mays* anthocyanin pigmentation pathway. Ludwig et al., *Science* 247: 449 (1990).

Although chemical selection of plant cells transformed with selectable marker genes has been successful with plant species and varieties that are easily cultured in vitro, the choice of selectable marker systems that have been shown to be successful for cereals and many other agronomically important plant species is very limited. In general, plant species that tend toward organogenesis and/or shoot propagation have been difficult to transform by means of chemical selection. The success rate with these plant species continues to improve, however, as evidenced by recent advances in Type I selection of maize inbreds and small grain cereals such as barley. Koziel et al., *Bio/Technology* 11: 194-200 (1993) and Mendel et al. In: *Transformation of Plants and Soil Microorganisms*, Wang et al. eds., Cambridge Press (1995).

Likewise, there has been little success in using visual screening methods for primary identification of transformed cells. The GUS gene was used to investigate germline transmission. McCabe et al., *Plant Physiol.*, 87(3): 671 (1988) and McCabe et al., *Plant Cell Tissue Organ Cult.* 33(3): 227 (1993). Histochemical staining for GUS activity was used to locate transgenic sectors in cotton and soybean transformants that ultimately produced transgenic seeds. Since histochemical analysis for GUS activity requires destruction of portions of the presumptively transformed plant tissue, this method is labor intensive and impractical for routine production of transgenic plants. This method is particularly unsuitable for plant species such as maize and other cereals in which transformants are recovered, even under optimum conditions, at low frequency. Recovery of transformed progeny was reported once in barley using GUS expression as a screening tool, but the method was found to be very labor-intensive. Ritala et al., *Plant Mol. Biol.* 24: 317-325 (1994). There have been no reports at all of success with GUS or other screenable markers with maize.

More recently, in vivo methods for visualizing GUS activity that do not require destruction of plant tissue have been made available. Molecular Probes Publication 2908, *Imagene Green™*, p. 1-4 (1993) and Naleway et al., *J. Cell Biol.* 115: 151a (1991). However, these in vivo methods for visualizing GUS activity have not proven useful for recovery of transformed cells because of low sensitivity and high fluorescent backgrounds.

Despite the fact that luciferase genes have been available for many years, this strategy for visualizing transformed cells has also not been successfully adapted for routine recovery of plant transformants. The luciferase-based screening methods are limited by the fact that most of these systems require the presence of luciferin, a compatible luciferase and an exogenously supplied cofactor. In the absence of the substrate,

enzyme or cofactor, the system does not bioluminesce. Cells transformed with a luciferase gene must have cell walls and plasma membranes that are permeable, or rendered permeable, to a compatible luciferin in order to detect bioluminescence. For example, tobacco plants regenerated from cells transformed with a firefly luciferase gene and exposed to a liquid medium containing firefly luciferin exhibited bioluminescence primarily along their major veins. Ow et al., *Science* 234: 856 (1986). Accordingly, a screening method has not been successfully developed for routine plant transformation that does not involve chemical selection or assays, often labor-intensive, that require the sacrifice or destruction of tissue samples for analysis.

A gene encoding GFP has been utilized as a marker for gene expression in prokaryotic and eukaryotic cells. Chalfie et al., *Science* 263: 802 (1994). Many cnidarians utilize GFPs as energy-transfer acceptors in bioluminescence. A gene encoding GFP isolated from a cnidarian and expressed in a heterologous prokaryotic or eukaryotic host produces a protein capable of fluorescence. A cDNA encoding the *Aequorea victoria* GFP produced a fluorescent product when expressed in *Escherichia coli* or *Caenorhabditis elegans* cells. Green fluorescence was detected in transformed cells upon illumination with a long-wave ultraviolet (UV) source without having to supply substrates or cofactors. In addition, fluorescence was stable for at least 10 min when illuminated with 450 to 490 nm light. Transformation of plant cells with a gene encoding GFP and detection of fluorescence has been reported. Haseloff et al., *TIG* 11: 328-329 (1995). Transformed *Arabidopsis* cells could be regenerated into whole plants. However, the regenerated plants expressing GFP exhibited signs of mild to moderate toxicity in the light compared to plants not expressing GFP. The strongest GFP expressors proved more difficult to regenerate. Likewise, in a recent report describing GFP expression in

kanamycin-selected tobacco transformants, Chiu et al., *Current Biol.* 6: 325-330 (1996) note that high GFP expression levels inhibited regeneration of transgenic plants.

5 A need therefore exists for a cell transformation method which does not rely, or does not rely solely on, selection of cells carrying a gene that confers resistance to a toxic substance. A need exists for a method for efficiently and easily identifying transformed
10 plant cells with a visual screenable marker. A need also exists for a method of cell transformation that does not require destruction of presumptively transformed tissue to assay for the presence of a selectable marker gene. A need exists for a method for cell transformation that
15 combines a selectable marker gene and a screenable marker gene. In addition, a need exists for a method of cell transformation which does not require exogenous supply of a substrate or cofactor for detection of the polypeptide encoded by a selectable marker gene. Yet another need
20 exists for a method of cell transformation that circumvents the cellular toxicity of the GFP.

SUMMARY OF THE INVENTION

25 Accordingly, it is an object of the present invention to provide a method for cell transformation which does not depend on selection of cells carrying a gene that confers resistance to a toxic substance.

30 It is another object of the present invention to provide a method for cell transformation which combines the selection of cells carrying a gene that confers resistance to a toxic substance with screening cells for the presence of a substance that renders transformed cells identifiable.

35 It is a further object of the present invention to provide a method for cell transformation which does not require destruction of presumptively transformed tissue to assay for the presence of a selectable marker gene.

Yet another object of the present invention is to provide a method for cell transformation which does not require that an exogenous substrate or cofactor be provided to assay for a polypeptide encoded by a selectable marker gene.

It is another object of the present invention to provide a method for cell transformation that circumvents the cellular toxicity of the GFP.

These and other objects are achieved, in accordance with one embodiment of the present invention, by providing an isolated DNA molecule comprising a nucleotide sequence encoding the GFP operably linked to an inducible promoter. The inducible promoter can be selected from the group consisting of the estrogen-inducible promoter, the estradiol-inducible promoter, the ACE1 promoter, the IN2 promoter and the tetracycline repressor promoter.

Also provided is an isolated DNA molecule comprising as nucleotide sequence encoding the GFP wherein the nucleotide sequence is operably linked to a targeting sequence for subcellular localization which directs a protein to a subcellular compartment.

An isolated DNA molecule is provided comprising a nucleotide sequence selected from the group consisting of (a) SEQ ID NO: 1; (b) a nucleotide sequence that has substantial sequence similarity with SEQ ID NO: 1; and (c) a functional fragment of (a) or (b), wherein said DNA molecule encodes a GFP. The nucleotide sequence encoding a GFP may be operably linked to a nucleotide sequence encoding a targeting sequence for subcellular localization which directs a protein to a subcellular compartment. The nucleotide sequence encoding a GFP and a targeting sequence may further comprise a promoter operably linked to said nucleotide sequence.

Also provided are expression vectors comprising DNA molecules encoding a GFP, optionally operably linked to an inducible promoter and/or targeting sequence. The expression vector of the instant invention may carry a

nucleotide sequence encoding a foreign protein operably linked to a second promoter.

Also provided is a method of using the expression vectors of the instant invention to produce a transformed plant, comprising the steps of introducing an expression carrying a gene encoding GFP into regenerable plant cells and selecting cells containing the GFP for regeneration. The method of the instant invention may include the step of inducing GFP expression where the nucleotide sequence encoding GFP is operably linked to an inducible promoter. The regenerable plant cells utilized in the method of the instant invention are selected from the group consisting of *Zea*, *Brassica* and *Helianthus* cells.

Also provided are transgenic plants expressing the isolated DNA molecule encoding GFP. In addition, transgenic plants comprising a vector carrying a nucleotide sequence encoding GFP are provided.

A method for producing a transgenic plant is provided comprising the steps of (a) constructing an expression vector comprising (i) a first promoter which is an inducible promoter, operably linked to a nucleotide sequence encoding a GFP, and (ii) a second promoter operably linked to a foreign gene; (b) introducing said expression vector into regenerable plant cells; (c) inducing expression of the gene encoding the GFP and selecting transformed plant cells containing said protein; and (d) regenerating transformed plants from said selected transformed plant cells. The inducible promoter may be selected from the group including the estrogen-inducible promoter, the estradiol-inducible promoter, the ACE1 promoter, the IN2 promoter and the tetracycline repressor promoter.

Also provided is a method for producing a transgenic plant, comprising the steps: (a) constructing an expression vector comprising (i) a first promoter operably linked to a nucleotide sequence encoding a sequence for subcellular localization which directs a protein to a subcellular compartment which is operably

linked to a nucleotide sequence encoding a GFP, and (ii) a second promoter operably linked to a foreign gene; (b) introducing said expression vector into regenerable plant cells; (c) selecting transformed plant cells containing said GFP; and (d) regenerating transformed plants from said selected transformed plant cells. The targeting sequence for subcellular localization directs the GFP to the mitochondria, chloroplasts, peroxisomes, vacuole, endoplasmic reticulum, cell wall or for secretion generally into the apoplast. This method may include a nucleotide sequence encoding a GFP selected from the group consisting of: (a) SEQ ID NO: 1; (b) a nucleotide sequence that has substantial sequence similarity with SEQ ID NO: 1; and (c) a functional fragment of (a) or (b).

Also provided is a method for producing a transgenic plant, comprising the steps: (a) constructing an expression vector comprising (i) a first promoter operably linked to a nucleotide sequence encoding a screenable marker flanked on the 5-prime and 3-prime ends with a recombinase-specific target sequence, and (ii) a second promoter operably linked to a foreign gene; (b) introducing said expression vector into regenerable plant cells; (c) selecting transformed plant cells containing said screenable marker; (d) transforming the plant cells with a second expression vector containing a gene encoding a site-specific recombinase; (e) selecting plant cells that no longer express the screenable marker; (f) regenerating transformed plants from said selected transformed plant cells; and (g) isolating said foreign protein. The site-specific recombinase may be selected from the group consisting of the FLP/FRT, Ac/DS and cre/lox systems. In addition, the screenable marker may be the GFP.

An isolated DNA molecule comprising a nucleotide sequence encoding the GFP fused in frame to a nucleotide sequence encoding superoxide dismutase is also provided. Cells transformed with the DNA molecule contain a fusion

-11-

protein that (i) produces green fluorescence in the presence of UV to blue light and (ii) displays superoxide dismutase activity.

Also provided is a method for producing a transgenic plant comprising: (a) constructing an expression vector comprising (i) a first promoter operably linked to a nucleotide sequence encoding a GFP, and (ii) a second promoter operably linked to a gene encoding an enzyme that is an oxygen scavenger, and (iii) a third promoter operably linked to a foreign gene; (b) introducing said expression vector into regenerable plant cells; (c) selecting transformed plant cells containing said GFP; and (d) regenerating transformed plants from said selected transformed plant cells. The oxygen scavenger enzyme may be superoxide dismutase.

Alternatively, a method for producing a transgenic plant is provided comprising: (a) constructing an expression vector comprising (i) a first promoter operably linked to a nucleotide sequence encoding a fusion protein comprising the GFP and an oxygen scavenger enzyme fused in frame, and (ii) a second promoter operably linked to a foreign gene; (b) introducing said expression vector into regenerable plant cells; (c) selecting transformed plant cells containing said the GFP and oxygen scavenger enzyme activity; and (d) regenerating transformed plants from said selected transformed plant cells. The oxygen scavenger enzyme may be superoxide dismutase.

11a

An aspect of the invention is to provide a method for producing a transgenic plant, wherein said transgenic plant is produced by positive selection, by screening
5 for a transformed plant cell without selection of a cell carrying a gene that confers resistance to a toxic substance, comprising the steps: (a) providing an expression vector comprising (i) a first promoter operably linked to a nucleotide sequence encoding a green fluorescent protein flanked on the 5-prime and 3-prime ends with a recombinase-specific target sequence, and (ii) a second promoter operably linked to
10 a foreign gene; (b) introducing said expression vector into regenerable plant cells; (c) screening for a transformed plant cell containing said green fluorescent protein by positive selection, wherein said screening does not depend on negative selection of a cell carrying a gene that confers resistance to a toxic substance; (d) transforming the plant cell identified by said screening of step (c), or a second plant cell derived
15 from said plant cell, with a second expression vector containing a gene encoding a site-specific recombinase; (e) screening for a plant cell that no longer expresses the green fluorescent protein; and (f) regenerating a transformed plant from a cell identified by said screening of step (e). The site-specific recombinase can be FLP/FRT, Ac/DS or cre/lox. The plant can be a monocot. The plant can be maize.
20 The plant can be a dicot. The plant can be sunflower. The green fluorescent protein (GFP) can be GFPm as set forth in SEQ ID NO: 1. The site-specific recombinase can be FLP/FRT. The screening in step (c) can be done by illumination with UV-blue light.

Another aspect of the invention is to provide a use of a first expression
25 vector and a second expression vector to transform a plant cell, wherein the first expression vector, comprises: (i) a first promoter operably linked to a nucleotide sequence encoding a green fluorescent protein flanked on the 5-prime and 3-prime ends with a recombinase-specific target sequence, and (ii) a second promoter operably linked to a foreign gene, and the second expression vector comprises: (i) a
30 gene encoding a site-specific recombinase wherein the transgenic plant is produced by positive selection, by screening for a transformed plant cell without selection of a cell carrying a gene that confers resistance to a toxic

11b

substance. The site-specific recombinase can be FLP/FRT, Ac/DS or cre/lox. The plant can be a monocot. The plant can be maize. The plant can be a dicot. The plant can be sunflower. The green fluorescent protein (GFP) can be GFPm as set forth in
5 SEQ ID NO: 1. The site-specific recombinase can be FLP/FRT. The screening can be done by illumination with UV-blue light.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 presents the nucleotide sequence [SEQ ID NO: 1] encoding a GFP
10 with its corresponding amino acid sequence [SEQ ID NO: 2].

Figure 2 shows the Plasmid Map of PHP167.

Figure 3 shows the Plasmid Map of PHP762.

Figure 4 shows the Plasmid Map of PHP7921.

Figure 5 shows the Plasmid Map of PHP8144.

DETAILED DESCRIPTION

1. Definitions

5 In the description that follows, a number of terms are used extensively. The following definitions are provided to facilitate understanding of the invention.

10 A structural gene is a DNA sequence that is transcribed into messenger RNA (mRNA) which is then translated into a sequence of amino acids characteristic of a specific polypeptide.

15 A promoter is a DNA sequence that directs the transcription of a structural gene. Typically, a promoter is located in the 5-prime region of a gene, proximal to the transcriptional start site of a structural gene. If a promoter is an inducible promoter, then the rate of transcription increases in response to an inducing agent. In contrast, the rate of transcription is not regulated by an inducing agent if the promoter is a constitutive promoter. For example, a promoter may be regulated in a tissue-specific or tissue-preferred manner such that it is only active in transcribing the associated coding region in a specific tissue type(s) such as leaves, roots or meristem.

25 A cytoplasmically localized protein is a protein coded for by a gene which does not include any specific signal for targeting of the protein into any subcellular organelle or compartment or any signals for secretion of the protein. For example, the GFP structural gene operably linked 5-prime to a promoter and to an appropriate 3-prime sequence encodes for a protein compartmentalized in the cytoplasm.

30 A signal or targeting sequence is a structural peptide domain required for targeting of a given polypeptide to a subcellular organelle, subcellular compartment or secretion from the cell. As used herein, the phrase sequence for subcellular localization is

intended to refer collectively to any form of signal, targeting or retention sequence as defined herein. Signal sequences for a polypeptide directed to the chloroplast or mitochondrion are largely localized to the amino-terminus of the polypeptide. Signal sequences for a polypeptide directed to the glyoxysomes and peroxisomes are largely localized to the carboxy-terminal domains of the polypeptide. Accordingly, targeted transport of the GFP protein is accomplished by means of chemically joining in proper reading frame, or operably linking, the nucleotide sequence encoding a signal sequence to the 5-prime and/or 3-prime region of the GFP structural gene.

A mitochondrial targeting sequence facilitates the transport of a protein to a mitochondrial compartment. Typically, the mitochondrial targeting sequence is located at the amino-terminus of a polypeptide.

A secretion targeting sequence targets a polypeptide for export into the extracellular space through the ER. For example, operably linking a nucleotide sequence encoding the barley alpha amylase 1 (BAA) secretory targeting sequence to the 5-prime end of a structural gene targets the encoded protein for export into the extracellular space.

A cell wall targeting sequence targets a polypeptide for export from the cell but the polypeptide is specifically localized to the cell wall. For example, cell wall localization of a polypeptide is accomplished by operably linking a nucleotide sequence encoding BAA 5-prime, and operably linking a nucleotide sequence encoding a portion of the maize hydroxyproline-rich glycoprotein 3-prime to a gene encoding the polypeptide. Steifel et al., *Plant Cell* 2: 785-793 (1990).

A vacuolar signal sequence facilitates the transport of a protein to the vacuole. For example, vacuolar targeting is accomplished by fusing the BAA secretory signal sequence at the amino-terminus of the protein and a sequence encoding a vacuolar signal sequence to the carboxy-terminus. Transport of a polypeptide to the

vacuole is therefore accomplished by means operably linking a nucleotide sequence encoding BAA 5-prime, and a nucleotide sequence encoding a vacuolar signal sequence 3-prime to a gene encoding a polypeptide. Alternatively, vacuolar targeting is accomplished by constructing a nucleotide sequence comprising in the 5-prime to 3-prime direction nucleotide sequences encoding a vacuole signal sequence, BAA and a polypeptide.

An endoplasmic reticulum retention sequence targets a polypeptide for localization in the lumen of the endoplasmic reticulum. For example, a polypeptide is targeted for retention in the endoplasmic reticulum through the addition of the BAA sequence on the amino-terminus and an endoplasmic reticulum signal sequence on the carboxy-terminus of the polypeptide.

A nuclear targeting sequence facilitates transport of a polypeptide to the nucleus. Typically, the nuclear signal sequence is located at the amino-terminus of a polypeptide. In order to retain the nuclear targeted protein in the nucleus, it may be necessary to increase the molecular weight of the protein by means of fusing an unrelated protein to the carboxy-terminus of the targeted protein. For example, GFP was retained in the nucleus by operably linking a nucleotide sequence encoding a nuclear signal sequence 5-prime and a nucleotide sequence encoding maize acetolactate synthase 3-prime to a gene encoding a polypeptide.

A peroxisomal targeting sequence facilitates the transport of a polypeptide into the peroxisome. Typically, the peroxisomal signal sequence is a tripeptide located at the carboxy-terminus of a polypeptide.

A chloroplast targeting sequence facilitates the transport of a nuclear encoded protein to a chloroplast compartment. Typically, the chloroplast signal sequence is located at the amino-terminus of a polypeptide. Accordingly, transport of a polypeptide to a chloroplast compartment is accomplished by means of operably linking

the nucleotide sequence encoding a chloroplast signal sequence to the 5-prime region of a gene encoding a polypeptide.

5 An isolated DNA molecule is a fragment of DNA that is not integrated in the genomic DNA of an organism.

10 Complementary DNA (cDNA) is a single-stranded DNA molecule that is formed from an mRNA template by the enzyme reverse transcriptase. Typically, a primer complementary to portions of mRNA is employed for the initiation of reverse transcription. Those skilled in the art also use the term "cDNA" to refer to a double-stranded DNA molecule consisting of such a single-stranded DNA molecule and its complementary DNA strand.

15 The term expression refers to the biosynthesis of a gene product. For example, in the case of a structural gene, expression involves transcription of the structural gene into mRNA and the translation of mRNA into one or more polypeptides.

20 A cloning vector is a DNA molecule, such as a plasmid, cosmid, or bacteriophage, that has the capability of replicating autonomously in a host cell. Cloning vectors typically contain one or a small number of restriction endonuclease recognition sites at which foreign DNA sequences can be inserted in a determinable fashion without loss of an essential biological function of the vector, as well as a marker gene that is suitable for use in the identification and selection of cells transformed with the cloning vector. Marker genes typically include, for example, genes that provide tetracycline resistance, ampicillin resistance or
25 30 kanamycin resistance.

35 An expression vector is a DNA molecule comprising a gene that is expressed in a host cell. Typically, gene expression is placed under the control of certain regulatory elements, including constitutive or inducible promoters, tissue-specific regulatory elements, and enhancers. Such a gene is said to be operably linked to the regulatory elements.

Transformation includes introduction of genetic material into plant cells resulting in chromosomal integration and stable heritability through meiosis. Transformation also includes introduction of genetic material into plant cells in the form of plant viral vectors involving epichromosomal replication and gene expression which may exhibit variable properties with respect to meiotic stability.

A foreign gene refers in the present description to a DNA sequence that is operably linked to at least one heterologous regulatory element.

A recombinant host may be any prokaryotic or eukaryotic cell that contains either a cloning vector or expression vector. This term also includes those prokaryotic or eukaryotic cells that have been genetically engineered to contain the cloned gene(s) in the chromosome or genome of the host cell.

A transgenic plant is a plant having one or more plant cells that contain an expression vector.

A green fluorescent protein (GFP), or a functional fragment of a GFP, is capable of producing a green fluorescence. GFP absorbs in the UV to blue range with a peak at 395 nm and emits in the green with a peak at 510 nm. The "red-shifted" version of GFP is a modified version of GFP which absorbs between 480 to 490 nm and emits in the green with a peak at 510nm. Red-shifted GFP is abbreviated as GFP_r. The "blue fluorescent protein" is a modified version of GFP which absorbs around 380nm and emits in the blue with a peak at around 445nm. Blue fluorescent protein is abbreviated as BFP. GFP_m is a nucleotide sequence coding for GFP in which the DNA sequence has been modified based on codon preference in maize (Fig. 1).

Two nucleic acid molecules are considered to have a substantial sequence similarity if their nucleotide sequences share a similarity of at least 50%. Sequence similarity determinations can be performed, for example, using the FASTA program (Genetics Computer Group;

Madison, WI). Alternatively, sequence similarity determinations can be performed using BLAST (Basic Local Alignment Search Tool) of the Experimental GENIFO(R) BLAST Network Service. See Altschul et al., *J. Mol. Biol.* 215: 403 (1990). Also, see Pasternak et al., "Sequence Similarity Searches, Multiple Sequence Alignments, and Molecular Tree Building," in *METHODS IN PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY*, Glick et al. (eds.), pages 251-267 (CRC Press 1993).

2. Reducing GFP Toxicity in Transformed Plants

Oxidative stress associated with GFP fluorescence has been suggested to cause cellular toxicity. Haseloff et al., *TIG* 11: 328-329 (1995). If GFP is to be useful as a screenable marker for plant transformation, as well as a reporter gene for detection of gene expression, strategies for reducing GFP toxicity must be available. The gene encoding GFP can be operably linked to an inducible promoter so that GFP can be transiently expressed and transformed cells identified. Alternatively, the gene encoding GFP can be operably linked to a signal sequence for targeting to an organelle or subcellular compartment or for secretion to the apoplast where GFP is not toxic to the cell.

Oxidative stress can be ameliorated by transforming cells containing GFP with an enzyme that serves as an oxygen scavenger. Such enzymes are well known in the art. For example, genes encoding the enzyme superoxide dismutase may protect against the primary oxidative stresses associated with GFP fluorescence. Balzan et al., *Proc. Nat Acad. Sci.* 92: 4219-4223 (1995). Other enzymes involved in the oxidative stress response, such as ascorbate peroxidase or glutathione S-transferase, could help mitigate secondary effects in the cell. Conklin, *Plant Physiol.* 109: 203-212 (1995).

More specifically, oxidative stress can be ameliorated by transforming cells with a DNA construct

carrying genes encoding GFP and SOD. Alternatively, the gene encoding GFP can be fused in frame to the gene encoding SOD. Cells transformed with this DNA construct produce a fusion protein that cause the cells to fluoresce in the presence of UV-blue light and express SOD activity.

The toxicity of GFP in transformed plants can be eliminated by excising the screenable marker gene following detection of transformed cells or sectors. The FLP/FRT system is used in conjunction with GFPm as a visible viable marker for FRT excision. The FLP/FRT system has been demonstrated in maize suspension cells using GUS expression as an indicator of FRT excision. Lysnik et al., NAR 21: 969-975 (1993). For example, plant cells are bombarded with a DNA construct containing the GFP gene flanked by FRT sequences as well as a foreign or agronomic gene of interest. The GFP gene may be operably linked to a constitutive promoter or an inducible promoter. In addition, the GFP gene may be operably linked to a signal sequence. Stable transformants are detected by means of screening for GFP.

Transgenic callus pieces are spread on medium and bombarded a second time with a FLP recombinase construct. Callus is monitored periodically under UV to blue illumination to detect cells that no longer express GFP. Callus pieces that no longer express GFP are regenerated and analyzed for expression of the foreign or agronomic gene. Agronomically useful transgenic plants are thereby produced that do not contain a marker gene.

The sensitivity of the screening method can be further increased by placing two markers genes between the FRT sequences. For example, plant cells are bombarded with a DNA construct containing the PAT and GFP genes flanked by FRT sequences and a foreign or agronomic gene. Stable transformants are recovered on bialaphos-containing medium and positive GFP-expression is confirmed. The transformed callus is then bombarded with FLP. The callus is then grown for 2-6 weeks with no

selection until clear GFP-null sectors can be identified. These sectors can be transferred onto bialaphos/chlorophenol red multiwell test plates to confirm bialaphos sensitivity (i.e. within 3-5 days). Callus pieces that no longer express GFP or PAT are regenerated and analyzed for expression of the foreign or agronomic gene. This permits recovery of agronomically useful transformants without any marker genes in the final product.

Likewise, the Ac/Ds system of maize can also be used in transgenic plants to excise the screenable marker gene that is transformed together with a foreign or agronomic gene. Mobilization of Ac and/or Ds has been demonstrated in diverse plants such as tomato, tobacco and Arabidopsis. Yoder et al., In Tomato Technology, Alan R. Liss, Inc. pp 189-198 (1987); Yoder et al., U.S. Patent No. 5,225,341; Baker et al., *EMBO J* 6: 1547-1554 (1987) and Lawson et al. *Mol. Gen. Genet.*, 245: 608-615 (1994). Likewise, the cre/lox recombinase system from bacteriophage P1 could also be used in conjunction with GFP. Excision of transgenes in plants using the cre/lox system was first demonstrated in tobacco. Odell et al., *Mol. Gen. Genet.*, 223: 369-378 (1990) and Dale and Ow, *Proc. Natl. Acad. Sci. USA*, 88: 10558-10562 (1991). Similar to the FLP and Ac systems described above, GFP expression provides an efficient, easily scorable phenotype for monitoring excision.

3. Isolation of Nucleotide Sequences Encoding GFP

Peptide sequences from purified GFPs are used to design degenerate oligonucleotide primers for polymerase chain reactions and gene cloning. Protein extracts can be prepared from cnidarian cells by standard methods known to the art. See, for example, Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Press, (1988). In a preferred embodiment, cnidarian cells are extracted into a buffer, and the extracts

separated into membrane and soluble fractions. Each fraction is tested for green fluorescence activity in the presence of blue light.

Fractions which contain GFP activity are then purified further to determine which components are responsible for the activity. Purification of the active fractions can be carried out by methods known in the art. See, for example, PROTEIN PURIFICATION METHODS - A PRACTICAL APPROACH, Harris et al., Eds. (IRL Press, Oxford, 1989). Extracts prepared as described above are purified by sequential size exclusion chromatography, isoelectric focusing, HPLC size exclusion chromatography, and chromatography on an affinity column. Fractions which display GFP activity can be analyzed further by SDS-PAGE analysis to determine the approximate molecular mass of the active component.

Purified GFPs prepared by the methods described above can be sequenced using methods well known in the art, for example using a gas phase peptide sequencer (Applied Biosystems, Foster City, CA). To determine as much of the peptide sequence as possible, it is preferred that the proteins of the present invention be cleaved into smaller fragments more suitable for gas-phase sequence analysis. This can be achieved by treatment of a purified GFP with selective peptidases, and in a particularly preferred embodiment, with endoproteinase lys-C (Boehringer). The fragments so produced can be separated by reversed-phase HPLC chromatography.

The peptide sequences of the proteins determined as above can be used to determine the DNA sequence encoding the protein. Methods for carrying out this determination are well known in the art. See, for example Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 1989).

In a preferred embodiment of the present invention, the peptide sequences are used to design degenerate oligonucleotide primers for polymerase chain reactions.

Each degenerate primer set will preferably contain every possible DNA sequence encoding the corresponding peptide sequences. Primer sets are prepared in both the sense and antisense orientation. Suitable oligonucleotide primers can be synthesized using commercial synthesizers, such as those supplied by Applied Biosystems (Foster City, CA). In a particularly preferred embodiment, the primers include additional nucleotide sequences containing restriction endonuclease cleavage sites. The presence of such sites allows for the directional cloning of PCR products into suitable cloning vectors after treatment with an appropriate restriction enzyme. See Finney, "Molecular Cloning of PCR Products" in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel et al. Eds. (John Wiley & Sons, New York, 1987) p. 15.7.1.

Template DNA for the PCR can be prepared from appropriate cnidarian cells or tissues using methods well known in the art. See Sambrook et al., *supra*. In a preferred embodiment, host cells are crushed under liquid nitrogen and mRNA is extracted using a commercially available kit (Pharmacia, Piscataway, NJ).

The mRNA preparation can then be used as a template for cDNA synthesis using poly(dT) or random hexamer primers by standard techniques. See Sambrook et al., *supra*. In a particularly preferred embodiment, cDNA synthesis is carried out using a commercially available kit (Pharmacia).

The cDNA can then be used directly for PCR using the method of Saiki et al., *Science* 239: 487 (1988). The cDNA also is used to prepare a cDNA library by standard methods. See Sambrook et al., *supra*. In a particularly preferred embodiment, the cDNA is packaged into bacteriophage particles using a commercially available kit (Promega, Madison, WI). The packaged cDNA is then transfected into *E. coli* to produce a cDNA library.

In an alternative preferred embodiment, genomic DNA from cnidarian cells or tissue can be used as the template DNA for the PCR. Genomic DNA can be prepared by

standard methods and PCR can then be used to prepare double stranded DNA molecules to probe the cDNA library and the genomic DNA for the gene(s) encoding GFP. In a preferred embodiment, degenerate primers are prepared corresponding to the termini of the longest peptide sequence determined by peptide sequencing. In a particularly preferred embodiment, primers are used in a PCR with first strand cDNA as template to amplify the DNA encoding the peptide. PCR is carried out under standard conditions. See Sakai et al., supra.

PCR amplification products are analyzed by polyacrylamide gel electrophoresis using standard methods. If an amplification product of the expected size (based on the peptide sequence) is found, the product is digested with appropriate restriction enzymes, ligated into a cloning vector and cloned by standard methods. See Sambrook et al, supra. In a preferred embodiment, clones are sequenced to verify that sequences according to the expected peptide sequence are present.

Once the DNA sequence encoding the peptide is known, it can be used to prepare non-degenerate primers corresponding to that sequence, again containing restriction enzyme recognition sequences to aid in cloning of DNA products. These primers are used in combination with degenerate primers corresponding to other peptide sequences to generate PCR amplification products which can be cloned and then analyzed as above. By these means, fragments of the gene sequence of the protein can be determined. Alternative methods for carrying out this PCR analysis include use of the 5' or 3' RACE methods using commercially available kits, such as those manufactured by Life Technologies (Gaithersburg, MD) or Clontech (Palo Alto, CA). Primers for this method are selected according to the manufacturer's directions.

Gene fragments prepared as above are excised from the cloning vector by restriction enzyme digestion, labeled with ^{32}P by conventional methods and used as probes to identify the complete gene encoding the

protein from within a cDNA or genomic library. In a preferred embodiment, the probe is chosen such that it is long enough to ensure hybridization specificity, while remaining short enough to allow reasonable rates of hybridization to the target gene.

A cnidarian genomic DNA library can be prepared by means well-known in the art. See, for example, Slightom et al. "Construction of λ Clone Banks," in METHODS IN PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY, Glick et al. (eds.), pages 121-146 (CRC Press, 1993). Genomic DNA can be isolated from cnidarian tissue, for example, by lysing plant tissue with the detergent Sarkosyl™, digesting the lysate with proteinase K, clearing insoluble debris from the lysate by centrifugation, precipitating nucleic acid from the lysate using isopropanol, and purifying resuspended DNA on a cesium chloride density gradient. See, for example, Ausubel et al. (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, pages 2.3.1-2.3.3 (Wiley Interscience 1990) ["Ausubel"].

DNA fragments that are suitable for the production of a genomic library can be obtained by the random shearing of genomic DNA or by the partial digestion of genomic DNA with restriction endonucleases. See, for example, Ausubel at pages 5.3.2-5.4.4, and Slightom et al., supra. Genomic DNA fragments can be inserted into a vector, such as a bacteriophage or cosmid vector, in accordance with conventional techniques, such as the use of restriction enzyme digestion to provide appropriate termini, the use of alkaline phosphatase treatment to avoid undesirable joining of DNA molecules, and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Slightom et al., supra, and are well-known in the art. Also see Ausubel at pages 3.0.5-3.17.5.

Screening of the cDNA or genomic library is carried out by conventional methods. See Sambrook et al, supra. Clones which hybridize to the probe are purified and their sequences determined. To facilitate sequencing,

5 nested deletions in the clones can be created using
standard protocols, or by commercially available kits
such as Erase-a-baseTM (Promega, Madison, WI) or The
Deletion Factory (Life Technologies, Gaithersburg, MD),
10 following the manufacturer's instructions. The sequences
obtained are analyzed for the presence of open reading
frames by conventional methods and to check if the entire
gene sequence has been found. In a preferred embodiment,
cDNA libraries are prepared by both random hexamer and
poly (dT) priming from samples, and are used to maximize
15 the chances of finding the complete coding sequence of
the desired gene.

Once the entire coding sequence of the gene for GFP
has been determined, the gene can be inserted into an
15 appropriate expression system. The gene can be expressed
in any number of different recombinant DNA expression
systems to generate large amounts of protein, which can
then be purified.

20 4. Chemical Synthesis of Genes Encoding GFP

Once the amino acid sequence of a GFP is known, a
gene encoding this GFP can be synthesized. In addition,
the nucleotide sequence of a synthetic gene encoding GFP
25 can be optimized for expression in plants by modifying
the codon usage to include plant preferred codons. See,
for example, Murray et al., *NAR* 17: 477 (1989). Even
more specifically, the nucleotide sequence of a synthetic
gene encoding GFP can be optimized for expression in
30 monocotyledonous or dicotyledonous plants. See, for
example, Campbell et al., *Plant Physiol.* 92: 1 (1990).

Genes encoding GFP can be obtained, for example, by
synthesizing the genes with mutually priming long
oligonucleotides. See, for example, Ausubel at pages
35 8.2.8 to 8.2.13. Also, see Wosnick et al., *Gene* 60:115
(1987). Moreover, current techniques using the
polymerase chain reaction provide the ability to
synthesize genes as large as 1.8 kilobases in length.

Adang et al., *Plant Molec. Biol.* 21: 1131 (1993); Bambot et al., *PCR Methods and Applications* 2: 266 (1993).

5. Identification of Functional Fragments of the GFP

5
DNA clones can be analyzed using a variety of techniques such as restriction analysis, Southern analysis, primer extension analysis, and DNA sequence
10 analysis. Primer extension analysis or S1 nuclease protection analysis, for example, can be used to localize the putative start site of transcription of the cloned gene. Ausubel at pages 4.8.1-4.8.5; Walmsley et al., "Quantitative and Qualitative Analysis of Exogenous Gene
15 Expression by the S1 Nuclease Protection Assay," in METHODS IN MOLECULAR BIOLOGY, VOL. 7: GENE TRANSFER AND EXPRESSION PROTOCOLS, Murray (ed.), pages 271-281 (Humana Press Inc. 1991). Functional fragments of the GFP protein are identified by production of green
20 fluorescence in the presence of blue UV light.

The general approach of such functional analysis involves subcloning DNA fragments of a genomic clone, cDNA clone or synthetic gene sequence into an expression vector, introducing the expression vector into a
25 heterologous host, and screening to detect green fluorescence in the presence of UV to blue light. Methods for generating fragments of a cDNA or genomic clone are well-known. Variants of an isolated DNA encoding GFP can be produced by deleting, adding and/or
30 substituting nucleotides for the isolated nucleotides, for example, the nucleotide sequence of SEQ ID NO: 1. Such variants can be obtained, for example, by oligonucleotide-directed mutagenesis, linker-scanning mutagenesis, mutagenesis using the polymerase chain reaction, and the like. See, for example, Ausubel, pages
35 8.0.3-8.5.9. Also see generally, McPherson (ed.), DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, (IRL Press 1991). Thus, the present invention also encompasses DNA molecules comprising nucleotide sequences that have

substantial sequence similarity with SEQ ID NO: 1 and encode a GFP.

6. Methods for Plant Transformation

5 The GFP-based screening method for plant transformation of the instant application can be used in conjunction with any method of plant transformation and regeneration. Numerous methods for plant transformation have been developed, including biological and physical, plant transformation protocols. See, for example, Miki et al., "Procedures for Introducing Foreign DNA into Plants" in *Methods in Plant Molecular Biology and Biotechnology*, Glick, B.R. and Thompson, J.E. Eds. (CRC Press, Inc., Boca Raton, 1993) pages 67-88. In addition, expression vectors and in vitro culture methods for plant cell or tissue transformation and regeneration of plants are available. See, for example, Gruber et al., "Vectors for Plant Transformation" in *Methods in Plant Molecular Biology and Biotechnology*, Glick, B.R. and Thompson, J.E. Eds. (CRC Press, Inc., Boca Raton, 1993) pages 89-119.

A. Agrobacterium-mediated Transformation

25 The most widely utilized method for introducing an expression vector into plants is based on the natural transformation system of *Agrobacterium*. See, for example, Horsch et al., *Science* 227:1229 (1985). *A. tumefaciens* and *A. rhizogenes* are plant pathogenic soil bacteria which genetically transform plant cells. The Ti and Ri plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of the plant. See, for example, Kado, C.I., *Crit. Rev. Plant. Sci.* 10: 1 (1991). Descriptions of *Agrobacterium* vector systems and methods for *Agrobacterium*-mediated gene transfer are provided by Gruber et al., *supra*, Miki et al., *supra*, and Moloney et al., *Plant Cell Reports* 8: 238 (1989).

B. Direct Gene Transfer

Despite the fact the host range for *Agrobacterium*-mediated transformation is broad, some major cereal crop species and gymnosperms have generally been recalcitrant to this mode of gene transfer, even though some success has recently been achieved in rice. Hiei et al., *The Plant Journal* 6: 271-282 (1994). Several methods of plant transformation, collectively referred to as direct gene transfer, have been developed as an alternative to *Agrobacterium*-mediated transformation.

A generally applicable method of plant transformation is microprojectile-mediated transformation wherein DNA is carried on the surface of microprojectiles measuring 1 to 4 μm . The expression vector is introduced into plant tissues with a biolistic device that accelerates the microprojectiles to speeds of 300 to 600 m/s which is sufficient to penetrate plant cell walls and membranes. Sanford et al., *Part. Sci. Technol.* 5: 27 (1987), Sanford, J.C., *Trends Biotech.* 6: 299 (1988) Sanford, J.C., *Physiol. Plant* 79: 206 (1990), Klein et al., *Biotechnology* 10: 268 (1992).

Another method for physical delivery of DNA to plants is sonication of target cells. Zhang et al., *Bio/Technology* 9: 996 (1991). Alternatively, liposome or spheroplast fusion have been used to introduce expression vectors into plants. Deshayes et al., *EMBO J.*, 4: 2731 (1985), Christou et al., *Proc Natl. Acad. Sci. U.S.A.* 84: 3962 (1987). Direct uptake of DNA into protoplasts using CaCl_2 precipitation, polyvinyl alcohol or poly-L-ornithine have also been reported. Hain et al., *Mol. Gen. Genet.* 199: 161 (1985) and Draper et al., *Plant Cell Physiol.* 23: 451 (1982). Electroporation of protoplasts and whole cells and tissues have also been described. Donn et al., In Abstracts of VIIth International Congress on Plant Cell and Tissue Culture IAPTC, A2-38, p 53 (1990); D'Halluin et al., *Plant Cell* 4: 1495-1505 (1992) and Spencer et al., *Plant Mol. Biol.* 24: 51-61 (1994).

A preferred method is microprojectile-mediated bombardment of immature embryos. The embryos can be bombarded on the embryo axis side to target the meristem at a very early stage of development or bombarded on the scutellar side to target cells that typically form callus and somatic embryos. Targeting of the scutellum using projectile bombardment is well known to those in the art of cereal tissue culture. Klein et al., *Bio/Technol.*, 6: 559-563 (1988); Kartha et al., *Plant Cell Rep.* 8: 429-432 (1989); Sautter et al., *Bio/Technol.*, 9: 1080-1085 (1991); Chibbar et al., *Genome*, 34: 435-460 (1991). The scutellar origin of regenerable callus from cereals is well known. Green et al., *Crop Sci.*, 15: 417-421 (1975); Lu et al., *TAG* 62: 109-112 (1982); and Thomas and Scott, *J. Plant Physiol.* 121: 159-169 (1985). Targeting the scutellum and then using chemical selection to recover transgenic plants is well established in cereals. D/Halluin et al., *Plant Cell* 4: 1495-1505 (1992); Perl et al., *MGG* 235: 279-284 (1992); Cristou et al., *Bio/Technol.* 9: 957-962 (1991). This literature reports DNA targeting of the scutellum and recovery of transgenic callus, plants and progeny based on a chemical selection regime. None of these references teach successful plant transformation wherein transformed cells are: visualized with a screenable marker such as GUS.

A preferred transformation method involves bombardment of the scutellar surface of immature embryos to introduce a fGP expression cassette and any other cotransformed genes. Embryos can be pretreated for 1 to 48 hours with a high osmoticum medium or left on a high-osmoticum medium for 24-48 hours after bombardment to improve cell survival and transformation frequencies. Immature embryos are then cultured on typical callus-inducing medium with no selective agent. At each subculture transfer, i.e., every two weeks, the culture is monitored using UV-blue light for GFP fluorescence. Fluorescing calli are separated from non-fluorescing

callus, and grown to the point where plants can be regenerated through standard media progressions.

Many exemplary variations, in terms of target cells and tissue culture routes are known in the art that could be used with GFP in this fashion. Some of these variations include introduction of GFP expression into protoplasts, suspension cells, type II callus, and type I callus, as in Morocz et al., *Theor. Appl. Genet.* 80: 721-726 (1990), Gordon-Kamm et al., *Plant Cell* 2: 603-618 (1990), Fromm et al., *Bio/Technol.* 8: 833-839 (1990) and D'Halluin et al., *Plant Cell* 4: 1495-1505 (1992). Microspores and microspore-derived embryogenic callus also represent a feasible alternative target/culturing route for using GFP screening to recover transformants. See Sun et al., *Plant Cell Rep.* 8: 313-316 (1989) and Mitchell et al., *J. Plant Physiol.* 37: 530-536 (1991). Furthermore, it is expected that the method of the present invention is also be applicable to any dicot species that is known to have a callus stage, for example, tobacco, canola or soybean.

In the case of the maize meristem target, the method entails selective enlargement of transgenic sectors, toward genetic homogeneity, in cell layers that contribute to germline transmission. This method is described in U.S. Patent No: 5,736,369,

and comprises the steps of (A) introducing foreign DNA into a meristem that is not enclosed by sheathing leaves; (B) inducing reorganization of the meristem to increase transgenic sector size, whereby the likelihood that a transgenic sector will contribute to germline transmission is increased; and (C) exposing the meristem to conditions under which it differentiates to form a plantlet, wherein the plantlet contains the transgenic sector or is homogeneously transformed by the foreign DNA, such that the plantlet can be grown into a transformed cereal plant that will transmit the foreign DNA to progeny. The foreign DNA can be introduced into

a plurality of meristems, at least some of which differentiate in step (C) to form a plurality of plantlets.

5 In one preferred embodiment, reorganization is effected through at least one manipulation selected from the group consisting of (i) imposition of a nonlethal selective pressure on the meristems, (ii) mechanically-induced meristem reorganization, and (iii) hormonally-induced shoot multiplication. In another preferred
10 embodiment the conditions in step (C) are such that the meristems undergo maturation and plant differentiation to form shoot apices, and the method further comprises effecting reorganization of meristem tissue in the shoot apices to enlarge transformed sectors or to produce periclinal L2 chimeras. The reorganization in this
15 regard can be effected, for example, by exposing the shoot apices to nonlethal selection pressure such that transformed cells have a competitive growth advantage over nontransformed cells in the shoot apices, and the proportion of transformed cells in the shoot apices is
20 increased. In yet another preferred embodiment, the method further comprises a step before step (B), e.g., before step (A), of wounding the apical dome selectively. A method of the present invention also can comprise the further steps of (i) dissecting out an axillary bud from
25 above the base of a leaf of a plantlet when a chimeric sector is observed in a substantial portion of the leaf, and then (ii) germinating the axillary bud into a whole plant or subjecting the axillary bud to shoot multiplication.
30

Likewise, the introduction of GFP-encoding DNA into dicot meristems provides an efficient method of mapping transgenic sectors and following these up into a germline and into progeny. Using mature imbibed seeds as the
35 starting material, the meristem is exposed by removal of the cotyledons, and DNA is introduced into meristem cells. A preferred method for introducing DNA into meristem cells is *Agrobacterium*-mediated delivery, but

other DNA delivery methods can be used with this explant such as bombardment with DNA-coated particles, or silicon carbide fiber-mediated delivery. Kaeppler et al., *Plant Cell Rep.* 9: 415-418 (1990). After DNA introduction, the meristem explants are cultured.

Plants can be manipulated, for example, by removal of the apical meristem, to stimulate axillary or secondary buds which can exhibit larger transgenic sectors relative to the primary shoot. Flowers above transgenic shoots are pollinated and the progeny are analyzed for transgene presence and expression. A variety of starting explants can regenerate shoots in sunflower, and thus represent alternative targets for GFP-encoding DNA delivery and transmission to progeny. These include the seedling meristem (as above), also the seedling hypocotyl, the mature cotyledon, the immature cotyledon, zygotic immature embryos, somatic embryos, and primary leaflets. See for example, respectively, Greco et al., *Plant Sci. Lett.* 36: 73-77 (1984); Krauter et al., *Helia* 14: 117-122 (1991); Power *Am. J. Bot.* 74: 497-503 (1987); Krauter et al., *Theor. Appl. Genet.* 82: 521-525 (1991); Finer, *Plant Cell Rep.* 6: 372-374 (1987), and Greco et al., *Plant Sci. Lett.* 36: 73-77 (1984).

In similar fashion, GFP expression could be used in other dicot species, for example, soybean, to map sectors and/or identify transformants and follow the transgenes through to progeny. A GFP-encoding cassette and other cotransformed genes can be introduced into soybean cotyledonary node cells using *Agrobacterium*. The explants are cultured on Gamborg (B5) medium with or without kanamycin selection. With no selection, GFP expression is used to identify transgenic chimeric or homogeneously transformed plantlets, which are rooted, grown to maturity and pollinated in the greenhouse. Progeny seed are collected for analysis. The use of cotyledonary nodes and *Agrobacterium* is presented by way of example, but other target cells and culturing methods could be used with GFP for soybean. In addition to the

cotyledonary node cells, mature cotyledons and immature cotyledons have both been used for transformation. See Hinchee et al., *Bio/Technol.* 6: 915-922 (1988) and Parrott et al., *Plant Cell Rep.* 7: 615-617 (1989).
5 Meristems from immature soybean seed embryonic axes have also been used for bombardment-mediated soybean transformation. See Christou et al., *Plant Physiol.* 87: 671-674 (1988) and McCabe et al., *Bio/Technol.* 6: 923-926 (1988).

10

7. Promoters

A. Inducible Promoters

15

An inducible promoter is operably linked to a nucleotide sequence encoding a GFP. Optionally, the inducible promoter is operably linked to a nucleotide sequence encoding a signal sequence which is operably linked to a nucleotide sequence encoding GFP. With an inducible promoter the rate of transcription increases in response to an inducing agent.

20

25

30

35

Any inducible promoter can be used in the instant invention. See Ward et al. *Plant Mol. Biol.* 22: 361-366 (1993). Exemplary inducible promoters include that from the ACE1 system which responds to copper (Mett et al. *PNAS* 90: 4567-4571 (1993)); In2 gene from maize which responds to benzenesulfonamide herbicide safeners (Hershey et al., *Mol. Gen. Genetics* 227: 229-237 (1991) and Gatz et al., *Mol. Gen. Genetics* 243: 32-38 (1994)) or Tet repressor from Tn10 (Gatz et al., *Mol. Gen. Genet.* 227: 229-237 (1991)). A particularly preferred inducible promoter is a promoter that responds to an inducing agent to which plants do not normally respond. An exemplary inducible promoter is the inducible promoter from a steroid hormone gene the transcriptional activity of which is induced by a glucocorticosteroid hormone. Schena et al., *Proc. Natl. Acad. Sci. U.S.A.* 88: 10421 (1991).

The expression vector comprises an inducible promoter operably linked to a nucleotide sequence encoding GFP. The expression vector is introduced into plant cells and presumptively transformed cells are exposed to an inducer of the inducible promoter. The cells are screened for the presence of GFP protein by means of illuminating the cells with UV to blue light and screening for the presence of green fluorescence.

B. Tissue-specific or Tissue Preferred Promoters

A tissue-specific promoter is operably linked to a nucleotide sequence encoding a GFP. Optionally, the tissue-specific promoter is operably linked to a nucleotide sequence encoding a signal sequence which is operably linked to a nucleotide sequence encoding GFP. Plants transformed with a gene encoding GFP operably linked to a tissue-specific promoter produce the GFP protein exclusively, or preferentially, in a specific tissue.

Any tissue-specific or tissue-preferred promoter can be utilized in the instant invention. Exemplary tissue-specific or tissue-preferred promoters include a root-preferred promoter such as that from the phaseolin gene (Murai et al., *Science* 23: 476-482 (1983) and Sengupta-Gopalan et al., *Proc. Natl. Acad. Sci. USA* 82: 3320-3324 (1985)); a leaf-specific and light-induced promoter such as that from *cab* or *rubisco* (Simpson et al., *EMBO J.* 4(11): 2723-2729 (1985) and Timko et al., *Nature* 318: 579-582 (1985)); an anther-specific promoter such as that from *LAT52* (Twell et al., *Mol. Gen. Genet.* 217: 240-245 (1989)); a pollen-specific promoter such as that from *Zml3* (Guerrero et al., *Mol. Gen. Genet.* 224: 161-168 (1993)) or a microspore-preferred promoter such as that from *apg* (Twell et al., *Sex. Plant Reprod.* 6: 217-224 (1993)).

The expression vector comprises a tissue-specific or tissue-preferred promoter operably linked to a nucleotide sequence encoding GFP. The expression vector is

introduced into plant cells. The cells are screened for the presence of GFP protein by means of illuminating the cells with UV to blue light and screening for the presence of green fluorescence.

5

C. Constitutive Promoters

10

A constitutive promoter is operably linked to a nucleotide sequence encoding a GFP or the constitutive promoter is operably linked to a nucleotide sequence encoding a signal sequence which is operably linked to a nucleotide sequence encoding GFP.

15

20

Many different constitutive promoters can be utilized in the instant invention. Exemplary constitutive promoters include the promoters from plant viruses such as the 35S promoter from CaMV (Odell et al., *Nature* 313: 810-812 (1985) and the promoters from such genes as rice actin (McElroy et al., *Plant Cell* 2: 163-171 (1990)); ubiquitin (Christensen et al., *Plant Mol. Biol.* 12: 619-632 (1989) and Christensen et al., *Plant Mol. Biol.* 18: 675-689 (1992)); pEMU (Last et al., *Theor. Appl. Genet.* 81: 581-588 (1991)); MAS (Velten et al., *EMBO J.* 3: 2723-2730 (1984)) and maize H3 histone (Lepetit et al., *Mol. Gen. Genet.* 231: 276-285 (1992) and Atanassova et al., *Plant Journal* 2(3): 291-300 (1992)).

25

The ALS promoter, a XbaI/NcoI fragment 5-prime to the *Brassica napus* ALS3 structural gene (or a nucleotide sequence that has substantial sequence similarity to said XbaI/NcoI fragment), represents a particularly useful constitutive promoter, Pioneer Hi-Bred

30

International US Patent No: 5,659,026.

35

The expression vector comprises a constitutive promoter operably linked to a nucleotide sequence encoding GFP. The expression vector is introduced into plant cells and presumptively transformed cells are screened for the presence of GFP protein by means of illuminating the cells with UV to blue light and screening for the presence of green fluorescence.

8. Signal Sequences For Targeting Proteins to Subcellular Compartments

Transport of GFP to a subcellular compartment such as the chloroplast, vacuole, peroxisome, glyoxysome, cell wall or mitochondrion, or for secretion into the apoplast, is accomplished by means of operably linking the nucleotide sequence encoding a signal sequence to the 5-prime and/or 3-prime region of a gene encoding GFP. Targeting sequences at the 5-prime and/or 3-prime end of the structural gene may determine, during protein synthesis and processing, where the encoded protein is ultimately compartmentalized. The presence of a signal sequence directs a polypeptide to either an intracellular organelle or subcellular compartment or for secretion to the apoplast.

In addition to the targeting sequences described above, the addition of amino acids to the encoded protein (fusing GFP to either another protein, a protein fragment, or a peptide) can further influence the fate of GFP. For example, a nuclear localization signal (NLS) alone is not sufficient to accumulate GFP in the nucleus. The GFP protein is sufficiently small that it freely diffuses from the nucleus into the cytoplasm, probably through the nuclear pores. The fusion of GFP to a sufficiently large protein permits accumulation in the nucleus if the GFP protein also contains an NLS at the amino-terminus. GFP can also be targeted to the nucleus by fusing GFP to proteins that are normally targeted to the nucleus.

The following exemplary signal sequences were utilized for targeting of GFP. However, the invention encompasses use of any signal sequence, or combination of signal sequences, to facilitate transport of GFP to a subcellular compartment or the apoplast where the protein is not toxic to the cell but can be detected.

More specifically, GFP is expressed in plant cells transformed with a expression vector carrying a

nucleotide sequence encoding a signal sequence(s), which directs a protein to a subcellular compartment or to the apoplast, operably linked to a gene encoding GFP. This GFP gene construct can be expressed from a constitutive, tissue-specific, tissue-preferred or inducible promoter. GFP is transported to a subcellular compartment or the apoplast where it is not toxic to the plant cell. Cells illuminated with UV to blue light can be detected because they emit green fluorescence. The transformed cells are selected for regeneration. If the expression vector also carries a gene encoding a foreign protein or agronomically useful protein, the regenerated plant can be used for production of this protein.

A. PHP8080 - Ubi::CTP-GFPm

A chloroplast signal sequence from the maize cab-m7 has been characterized and the nucleotide sequence is shown below. Becker et al., *Plant Mol. Biol.* 20: 49 (1992). The signal sequence includes 21 amino acids of the cab-m7 protein that extends beyond the processing site (marked by *).

```

ATG GCG TCC TCC ACG ATG GCC CTC TCC TCC ACC GCC TTC GCC
M  A  S  S  T  M  A  L  S  S  T  A  F  A
-
GGC AAG GCC GTG AAC GTG CCG TCG TCG TCC GCC TTC GAG
G  K  A  V  N  V  P  S  S  S  A  F  E
-
GCC CGC GTG ACC ATG AGG AAG ACG GCG GCG AAG GCC AAG
A  R  V  T  *M  R  K  T  A  A  K  A  K  K
-
CCA GCT GCG GCG TCC GGG AGC CCG TGG TAC GGC CCC ATG
P  A  A  A  S  G  S  P  W  Y  G  P  M

```

B. PHP8087 - CPN60::GFPm

The mitochondrial signal sequence from the maize chaperonin 60 gene is shown below. Close, P.S., Master's Theses, Iowa State University (1993). The signal sequence includes intron 1 of cpn60II (marked by -) and

includes 2 amino acids from the cpn60II protein beyond the processing site (marked by *).

5 ATG TAC CGC GCG GCC GCT AGC CTC GCC TCC AAG GCG CGG CAA GCC GGG AGC
M Y R A A A S L A S K A R Q A G S
AGC TCC GCC GCT CGC CAG-gtgagagcagactcgtgtttatacgcgtgtgatgggtctgatg
S S A A R Q
10 gaggatccccgccctcagatttggtaatgttggcctgggggttctagtagttctgcgtgcagggctgtg
ggtttgcgtgcatgggtgctgtttatttttggtggcgatctgccggaatctgtagttcgctcgcgcaaaa
tctaagctagctcgctaattggcgtagtggcggtgggggtttcacattaatctacgggtggtgaactcgctc
actaccgtcctccagttagctgttagacaccgaatacactgattggcagttgagaaacattgatctg
atccagcagaaatcgatgtcttgtgaaattcggtattttattgtcgtgtaaccttggggcatggcag
15 tctctaattgatcagcactcacctctgttgtgtgtgatgctttatag-GTT GGA AGC AGG CTT
V G S R L
GCC TGG AGC AGG AAC TAT GCT GCC ATG
A W S R N Y *A A M

20 C. PHP8144 - Ubi::BAA-GFPm - Cloned sequence
includes 1 amino acid beyond processing site
(marked by *, lower case bases show linker
sequence).

25 The barley alpha amylase 1 signal
sequence is shown below. Knox, C., et al.,
"Structure and Organization of Two Divergent
Alpha-Amylase Genes From Barley", Plant Mol.
Biol. 9:3-17 (1987). The cloned sequence
30 includes sequence only up to processing site.

ATG GCC AAC AAG CAC CTG AGC CTC TCC CTC TTC CTC GTG CTC CTC GGC CTC TCC
M A N K H L S L S L F L V L L G L S
GCC TCC CTC GCC TCC GGA tcc ATG
35 A S L A S *G S M

D. PHP8757 - Ubi::BAA-GFPm-BLVT - Signal
sequence starts at *, lower case bases show
linker sequence.

40 The barley lectin vacuole signal sequence
is shown below. Lerner et al., Plant Physiol.
91: 124-129 (1989).

AAG atc tac GTG TTC GCC GAG GCC ATC GCC GCC AAC TCC ACC CTC GTG GCC GAG
45 K I Y *V F A E A I A A N S T L V A E

E. PHP8758 - Ubi::BAA-GFPm-HDEL - Signal sequence starts at *, lower case bases show linker sequence.

The HDEL endoplasmic reticulum retention signal from maize b-70 is shown below. Fontes et al., Plant Cell 3: 483-496 (1991).

AAG atc tac GAC GGC GGC GTG GAC GAC GAC CAC GAC GAG CTC
K I Y *D G G V D D D H D E L

F. PHP8759 - Ubi::BAA-SVT-GFPm - Marked by ^, lower case bases show linker sequence.

The sweet potato sporamin vacuole signal sequence is shown below. Matsuoka et al., Proc. Natl. Acad. Sci. 88: 834 (1991). The sequence shown includes 4 additional sporamin amino acids beyond the processing site (marked by *).

ATG GCC AAC AAG CAC CTG AGC CTC TCC CTC TTC CTC GTG CTC CTC GGC CTC TCC
M A N K H L S L S L F L V L L G L S
GCC TCC CTC GCC TCC GGA CAC AGC AGG TTC AAC CCC ATC AGG CTG CCC ACC ACC
A S L A S G ^H S R F N P I R L P T T
CAC GAG CCC GCC AGC AGC GAG ACC gga tcc ATG
H E P A *S S E T G S M

G. PHP8882 - Ubi::GFPm-SKL - (marked by *)

The consensus peroxisome signal sequence designated SLK is shown below. Gould et al., J. Cell Biol 108: 1657 (1989).

AAG TCG AAG CTT
K *S K L

H. Ubi::GR-GFPm

The pea glutathione reductase mitochondrial and chloroplast signal sequence is shown below. Creissen et al., Plant J. 2: 129 (1991). The sequence shown includes 20 additional amino acids beyond the putative processing site (marked by *).

ATG AAC CAG GCG ATG GCC ACC CCG CTG TCC CTG TCC TGC TGC TCC CCG ACC CTG
 M N Q A M A T P L S L S C C S P T L
 ACC AGG TCC ACC CTG TTC TTC ACC AAG ACC TTC CCG TTC TCC CGC TCC TTC TCC
 T R S T L F F T K T F P F S R S F S
 ACC CCG CTG CCG CTG TCC ACC AAG ACC CTG ATC TCC CTG TCC CCG CCG CAC AGG
 T P L P L S T K T L I S L S P P H R
 ACC TTC GCC GTG AGG GCT GAG TCC CAG AAC GGC GCG GAC CCG GCC AGG CAG TAC
 E S Q N G A D P A R Q Y D F D L F T
 GAC TTC GAC CTG TTC ACC ATC GGC
 I G T F A V *R A

I. PHP9053 - Ubi::NLS-GFPm-MALS

The nuclear localization signal from simian virus 40 (SV40) was fused to the N terminus of GFPm. In order to retain the protein in the nucleus the molecular weight of NLS-GFPm was increased by making a carboxy terminal addition of the large unrelated protein maize acetolactate synthase (ALS) (lower case bases show linker sequence). Kalderon, D., Robers, B., Richardson, W., and Smith A., "A short amino acid sequence able to specify nuclear location", *Cell* 39: 499-509 (1984).

CCG CCC AAG AAG AAG CGC AAG GTG ccc ATG ...
 P P K K K R K V P M
 AAG atc cac ATG
 K I H M

J. Ubi::GFPm-HRGP

A carboxy terminal fusion of a portion of the maize HRGP coding sequence to GFPm was used to anchor GFP to the cell wall. Sequence encoding amino acids 177 to 328 was used. Stiefel, V., Ruiz-Avila, L., Raz R., Valles M., Gomez J., Pages M., Martinez-Izquierdo J., Ludevid M., Landale J., Nelson T., and Puigdomenech P., "Expression of a maize cell wall hydroxyproline-rich glycoprotein gene in early leaf and root vascular differentiation", *Plant Cell* 2: 785-793 (1990).

TAC ACT CCA AGC CCT AAG CCA CCG GCT ACC AAG CCT CCC ACG
 Y T P S P K P P A T K P P T

5 CCC AAG CCG ACC CCG CCA ACG TAC ACC CCT TCG CCA AAG CCT
 P K P T P P T Y T P S P K P
 CCG ACA CCC AAG CCG ACC CCG CCG ACG TAC ACC CCT TCT CCC
 P T P K P T P P T Y T P S P
 AAG CCT CCG ACG CCC AAG CCG ACC CCG CCG ACG TAC ACT CCA
 K P P T P K P T P P T Y T P
 AGC CCC AAG CCT CCC ACA CAC CCG ACG CCC AAG CCG ACC CCA
 S P K P P T H P T P K P T P
 10 CCG ACG TAC ACC CCT TCC CCA AAG CCT CCG ACG CCC AAG CCG
 P T Y T P S P K P P T P K P
 ACC CCA CCG ACG TAC ACC CCT TCC CCA AAG CCT CCG ACA CCC
 T P P T Y T P S P K P P T P
 AAG CCG ACC CCA CCG ACG TAC ACC CCT TCC CCA AAG CCT CCG
 K P T P P T Y T P S P K P P
 15 ACA CCC AAG CCG ACC CCA CCG ACG TAC ACT CCC ACA CCG AAG
 T P K P T P P T Y T P T P K
 CCG CCG GCC ACC AAG CCG CCC ACC TAC ACT CCG ACG CCG CCG
 P P A T K P P T Y T P T P P
 20 GTG TCT CAC ACC CCC AGC CCG CCG CCA CCA TAC TAC
 V S H T P S P P P P Y Y

K. Ubi::GFPm-SOD

Superoxide dismutase coding sequence fused to
 GFPm. Alcendor D., Chapman G., and Beaman B.,
 25 "Isolation, sequencing and expression of the
 superoxide dismutase-encoding gene (sod) of *Nocardia
 asteroides* strain GUH-2", *Gene* 164: 143-147 (1995).

30 GTG GCT GAG TAC ACG CTG CCG GAT CTG GAT TAC GAC TAC AGC
 V A E Y T L P D L D Y D Y S
 GCC TGG AAC CCC ACA TCT CCG GGC AGA TCA ACG AGC TGA CAC
 A L E P H I S G Q I N E L H
 CAT TCC AAG CAC CAC GCC GCC TAC GTC GCC GGT GCC AAC ACG
 H S K H H A A Y V A G A N T
 35 GCA CTG GAG AAG CTG GAA GCC GCC CGT GAG GCC GGC GAT CAC
 A L E K L E A A R E A G D H
 AGC GCG ATC TTC CTG CAC GAG AAG AAC CTC GCG TTC CAC CTC
 S A I F L H E K N L A F H L
 GGC GGA CAC GTC AAC CAC TCC ATC TGG TGG AAG AAC CTG TCC
 40 G G H V N H S I W W K N L S
 CCC AAC GGT GGC GAC AAG CCG GTC GGC GAG CTG GCC GCG GCC
 P N G G D K P V G E L A A A
 ATC GAC GAC CAG TTC GGT TCG TTC GAC AAG TTC CGC GCG CAG
 I D D Q F G S F D K F R A Q
 45 TTC ACC GCC GCG GCC AAC GGC CTG CAG GGC TCG GGC TGG GCG
 F T A A A N G L Q G S G W A
 GTG CTC GGT TAC GAC ACC CTC GGC CAG AAG CTG CTG ACC TTC
 V L G Y D T L G Q K L L T F
 CAG CTC TAC GAC CAG CAG GCC AAC GTG CCG CTG GGC ATC ATC
 50 Q L Y D Q Q A N V P L G I I
 CCG CTG CTC CAG GTC GAC ATG TGG GAG CAC GCC TTC TAC CTG
 P L L Q V D M W E H A F Y L
 CAG TAC AAG AAC GTC AAG GCC GAC TAC GTG ACC GCG TTC TGG
 Q Y K N V K A D Y V T A F W

AAC GTC GTC AAC TGG GCC GAC GTG CAG GAC CGC TTC GGC AAG
 N V V N W A D V Q D R F G K
 GCC GTC AAC CAG GGC AAG GGC CTT ATC TTC GGG
 A V N Q G K G L I F G

5

9. Detection of GFP in Transformed Plant Cells

10

GFP is detected in transformed plant cells by conventional spectrophometric methods. The transformed cells or tissue are screened for the presence of GFP protein by means of the illuminating the cells with UV to blue light and screening for the presence of green fluorescence.

15

20

25

30

35

Compound and dissecting microscopes were fitted with appropriate filter combinations for fluorescent protein excitation. Illumination with UV-blue light (exciting around the absorption maximum of 395 nm or around the minor peak at approximately 475 nm for GFP; around 480-490 nm for the red-shifted version and around 380 nm for the blue fluorescent protein) is required for visualization. A hand-held lamp for benchtop work also permits good visualization. Cut-off filters or bandpass filters between the fluorescing tissue and the viewer (i.e. around the intermediate objective or the eyepieces of the microscope, or hand-held in front of the eyes if working on the benchtop) greatly reduced background autofluorescence from the tissue. For GFP and the red-shifted GFP this cut-off or bandpass filter permitted light between 500-550 nm to reach the viewer. For the blue fluorescent protein, this filter permitted the passage of 420 to 450 nm light. Useful filters and wavelengths are not restricted to those described above, and further generic description of the optical characteristics of this system are available. Heim et al., *Currently Biology* 6: 178-182 (1996).

10. Foreign Protein Genes and Agronomic Genes

With transgenic plants according to the present invention, a foreign protein can be produced in commercial quantities. Thus, the selection and propagation techniques described above yield a plurality of transgenic plants which are harvested in a conventional manner, and a foreign protein then is extracted from a tissue of interest or from total biomass. Protein extraction from plant biomass can be accomplished by known methods which are discussed, for example, by Heney and Orr, *Anal. Biochem.* 114: 92-6 (1981).

According to a preferred embodiment, the transgenic plant provided for commercial production of foreign protein is maize. In another preferred embodiment, the biomass of interest is seed. For the relatively small number of transgenic plants that show higher levels of expression, a genetic map can be generated, primarily via conventional RFLP and PCR analysis, which identifies the approximate chromosomal location of the integrated DNA molecule. For exemplary methodologies in this regard, see Glick and Thompson, *METHODS IN PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY* 269-284 (CRC Press, 1993). Map information concerning chromosomal location is useful for proprietary protection of a subject transgenic plant. If unauthorized propagation is undertaken and crosses made with other germplasm, the map of the integration region can be compared to similar maps for suspect plants, to determine if the latter have a common parentage with the subject plant. Map comparisons would involve hybridizations, RFLP, PCR and sequencing, all of which are conventional techniques.

Likewise, by means of the present invention, agronomic genes can be expressed in transformed plants. More particularly, plants can be genetically engineered to express various phenotypes of agronomic interest. The

genes implicated in this regard include, but are not limited to, those categorized below.

1. Genes That Confer Resistance To Pests or Disease And That Encode:

5

10

15

20

(A) Plant disease resistance genes. Plant defenses are often activated by specific interaction between the product of a disease resistance gene (R) in the plant and the product of a corresponding avirulence (Avr) gene in the pathogen. A plant variety can be transformed with cloned resistance gene to engineer plants that are resistant to specific pathogen strains. See, for example Jones et al., *Science* 266: 789 (1994) (cloning of the tomato Cf-9 gene for resistance to *Cladosporium fulvum*); Martin et al., *Science* 262: 1432 (1993) (tomato Pto gene for resistance to *Pseudomonas syringae* pv. tomato encodes a protein kinase); Mindrinos et al., *Cell* 78: 1089 (1994) (*Arabidopsis* RSP2 gene for resistance to *Pseudomonas syringae*).

25

30

(B) A *Bacillus thuringiensis* protein, a derivative thereof or a synthetic polypeptide modeled thereon. See, for example, Geiser et al., *Gene* 48: 109 (1986), who disclose the cloning and nucleotide sequence of a Bt δ -endotoxin gene. Moreover, DNA molecules encoding δ -endotoxin genes can be purchased from American Type Culture Collection (Rockville, MD), under ATCC accession Nos. 40098, 67136, 31995 and 31998.

35

(C) A lectin. See, for example, the disclosure by Van Damme et al., *Plant Molec. Biol.* 24: 825 (1994), who disclose the nucleotide sequences of several *Clivia miniata* mannose-binding lectin genes.

(D) A vitamin-binding protein such as avidin.

The application teaches the use of avidin and avidin homologues as larvicides against insect pests.

(E) An enzyme inhibitor, for example, a protease inhibitor or an amylase inhibitor. See, for example, Abe et al., *J. Biol. Chem.* 262: 16793 (1987) (nucleotide sequence of rice cysteine proteinase inhibitor), Huub et al., *Plant Molec. Biol.* 21: 985 (1993) (nucleotide sequence of cDNA encoding tobacco proteinase inhibitor I), and Sumitani et al., *Biosci. Biotech. Biochem.* 57: 1243 (1993) (nucleotide sequence of *Streptomyces nitrosporeus* α -amylase inhibitor).

(F) An insect-specific hormone or pheromone such as an ecdysteroid and juvenile hormone, a variant thereof, a mimetic based thereon, or an antagonist or agonist thereof. See, for example, the disclosure by Hammock et al., *Nature* 344: 458 (1990), of baculovirus expression of cloned juvenile hormone esterase, an inactivator of juvenile hormone.

(G) An insect-specific peptide or neuropeptide which, upon expression, disrupts the physiology of the affected pest. For example, see the disclosures of Regan, *J. Biol. Chem.* 269: 9 (1994) (expression cloning yields DNA coding for insect diuretic hormone receptor), and Pratt et al., *Biochem. Biophys. Res. Comm.* 163: 1243 (1989) (an allostatin is identified in *Diploptera punctata*). See also U.S. patent No. 5,266,317 to Tomalski et al., who disclose genes encoding insect-specific, paralytic neurotoxins.

(H) An insect-specific venom produced in nature by a snake, a wasp, etc. For example, see Pang et al., *Gene* 116: 165 (1992), for disclosure of heterologous expression in plants of a gene coding for a scorpion insectotoxic peptide.

(I) An enzyme responsible for an hyperaccumulation of a monerpene, a sesquiterpene, a steroid, hydroxamic acid, a phenylpropanoid derivative or another non-protein molecule with insecticidal activity.

5

(J) An enzyme involved in the modification, including the post-translational modification, of a biologically active molecule; for example, a glycolytic enzyme, a proteolytic enzyme, a lipolytic enzyme, a
10 nuclease, a cyclase, a transaminase, an esterase, a hydrolase, a phosphatase, a kinase, a phosphorylase, a polymerase, an elastase, a chitinase and a glucanase, whether natural or synthetic. See PCT application WO 93/02197 in the name of Scott et al., which discloses the
15 nucleotide sequence of a callase gene. DNA molecules which contain chitinase-encoding sequences can be obtained, for example, from the ATCC under accession Nos. 39637 and 67152. See also Kramer et al., *Insect Biochem. Molec. Biol.* 23: 691 (1993), who teach the nucleotide
20 sequence of a cDNA encoding tobacco hookworm chitinase, and Kawalleck et al., *Plant Molec. Biol.* 21: 673 (1993), who provide the nucleotide sequence of the parsley ubi4-2 polyubiquitin gene.

25

(K) A molecule that stimulates signal transduction. For example, see the disclosure by Botella et al., *Plant Molec. Biol.* 24: 757 (1994), of nucleotide sequences for mung bean calmodulin cDNA clones, and Griess et al., *Plant Physiol.* 104: 1467 (1994), who
30 provide the nucleotide sequence of a maize calmodulin cDNA clone.

35

(L) A hydrophobic moment peptide. See U.S. patent No: 5,580,852 (disclosure of peptide derivatives of Tachyplesin which inhibit fungal plant pathogens) and U.S. Patent No: 5,607,914 (teaches synthetic antimicrobial peptides that confer disease resistance).

5 (M) A membrane permease, a channel former or a channel blocker. For example, see the disclosure by Jaynes et al., Plant Sci. 89: 43 (1993), of heterologous expression of a cecropin- β lytic peptide analog to render transgenic tobacco plants resistant to *Pseudomonas solanacearum*.

10

(N) A viral-invasive protein or a complex toxin derived therefrom. For example, the accumulation of viral coat proteins in transformed plant cells imparts resistance to viral infection and/or disease development effected by the virus from which the coat protein gene is derived, as well as by related viruses. See Beachy et al., Ann. Rev. Phytopathol. 28: 451 (1990). Coat protein-mediated resistance has been conferred upon transformed plants against alfalfa mosaic virus, cucumber mosaic virus, tobacco streak virus, potato virus X, potato virus Y, tobacco etch virus, tobacco rattle virus and tobacco mosaic virus. Id.

20

(O) An insect-specific antibody or an immunotoxin derived therefrom. Thus, an antibody targeted to a critical metabolic function in the insect gut would inactivate an affected enzyme, killing the insect. Cf. Taylor et al., Abstract #497, SEVENTH INT'L SYMPOSIUM ON MOLECULAR PLANT-MICROBE INTERACTIONS (1994) (enzymatic inactivation in transgenic tobacco via production of single-chain antibody fragments).

25

30

(P) A virus-specific antibody. See, for example, Tavladoraki et al., Nature 366: 469 (1993), who show that transgenic plants expressing recombinant antibody genes are protected from virus attack.

35

(Q) A developmental-arrestive protein produced in nature by a pathogen or a parasite. Thus, fungal endo α -1,4-D-polygalacturonases facilitate fungal colonization and plant nutrient release by solubilizing plant cell wall homo- α -1,4-D-galacturonase. See Lamb et al., *Bio/Technology* 10: 1436 (1992). The cloning and characterization of a gene which encodes a bean endopolygalacturonase-inhibiting protein is described by Toubart et al., *Plant J.* 2: 367 (1992).

(R) A developmental-arrestive protein produced in nature by a plant. For example, Logemann et al., *Bio/Technology* 10: 305 (1992), have shown that transgenic plants expressing the barley ribosome-inactivating gene have an increased resistance to fungal disease.

2. Genes That Confer Resistance To A Herbicide, For Example:

(A) A herbicide that inhibits the growing point or meristem, such as an imidazalinone or a sulfonylurea. Exemplary genes in this category code for mutant ALS and AHAS enzyme as described, for example, by Lee et al., *EMBO J.* 7: 1241 (1988), and Miki et al., *Theor. Appl. Genet.* 80: 449 (1990), respectively.

(B) Glyphosate (resistance imparted by mutant EPSP synthase and *aroA* genes, respectively) and other phosphono compounds such as glufosinate (PAT and *bar* genes), and pyridinoxy or phenoxy proprionic acids and cyclohexones (ACCase inhibitor-encoding genes). See, for example, U.S. patent No. 4,940,835 to Shah et al., which discloses the nucleotide sequence of a form of EPSP which can confer glyphosate resistance. A DNA molecule encoding a mutant *aroA* gene can be obtained under ATCC accession No. 39256, and the nucleotide sequence of the mutant gene is disclosed in U.S. patent No. 4,769,061 to Comai. European patent application No. 0 333 033 to

Kumada et al. and U.S. patent No. 4,975,374 to Goodman et al. disclose nucleotide sequences of glutamine synthetase genes which confer resistance to herbicides such as L-phosphinothricin. The nucleotide sequence of a phosphinothricin-acetyl-transferase gene is provided in European application No. 0 242 246 to Leemans et al. De Greef et al., *Bio/Technology* 7: 61 (1989), describe the production of transgenic plants that express chimeric bar genes coding for phosphinothricin acetyl transferase activity. Exemplary of genes conferring resistance to phenoxy proprionic acids and cyclohexones, such as sethoxydim and haloxyfop, are the Acc1-S1, Acc1-S2 and Acc1-S3 genes described by Marshall et al., *Theor. Appl. Genet.* 83: 435 (1992).

(C) A herbicide that inhibits photosynthesis, such as a triazine (*psbA* and *gs+* genes) and a benzonitrile (nitrilase gene). Przibilla et al., *Plant Cell* 3: 169 (1991), describe the transformation of *Chlamydomonas* with plasmids encoding mutant *psbA* genes. Nucleotide sequences for nitrilase genes are disclosed in U.S. patent No. 4,810,648 to Stalker, and DNA molecules containing these genes are available under ATCC accession Nos. 53435, 67441 and 67442. Cloning and expression of DNA coding for a glutathione S-transferase is described by Hayes et al., *Biochem. J.* 285: 173 (1992).

3. Genes That Confer Or Contribute To A Value-Added Trait, Such As:

(A) Modified fatty acid metabolism, for example, by transforming a plant with an antisense gene of stearoyl-ACP desaturase to increase stearic acid content of the plant. See Knultzon et al., *Proc. Nat'l Acad Sci. USA* 89: 2624 (1992).

(B) Decreased phytate content

(1) Introduction of a phytase-encoding gene would enhance breakdown of phytate, adding more free phosphate to the transformed plant. For example, see Van Hartingsveldt et al., *Gene* 127: 87 (1993), for a disclosure of the nucleotide sequence of an *Aspergillus niger* phytase gene.

(2) A gene could be introduced that reduces phytate content. In maize, this, for example, could be accomplished, by cloning and then re-introducing DNA associated with the single allele which is responsible for maize mutants characterized by low levels of phytic acid. See Raboy et al., *Maydica* 35: 383 (1990).

(C) Modified carbohydrate composition effected, for example, by transforming plants with a gene coding for an enzyme that alters the branching pattern of starch. See Shiroza et al., *J. Bacteriol.* 170: 810 (1988) (nucleotide sequence of *Streptococcus mutans* fructosyltransferase gene), Steinmetz et al., *Mol. Gen. Genet.* 200: 220 (1985) (nucleotide sequence of *Bacillus subtilis* levansucrase gene), Pen et al., *Bio/Technology* 10: 292 (1992) (production of transgenic plants that express *Bacillus licheniformis* α -amylase), Elliot et al., *Plant Molec. Biol.* 21: 515 (1993) (nucleotide sequences of tomato invertase genes), Søgaard et al., *J. Biol. Chem.* 268: 22480 (1993) (site-directed mutagenesis of barley amylase gene), and Fisher et al., *Plant Physiol.* 102: 1045 (1993) (maize endosperm starch branching enzyme II).

Example 1**Construction of Genes and Vectors**

5 The GFPm nucleotide sequence was derived from
a back translation of the protein sequence using maize
preferred codons. Sequence analysis was performed using
the Wisconsin Sequence Analysis Package from Genetics
Computer Group, Madison, WI. The nucleotide sequence was
10 assembled from a series of synthetic oligonucleotides.
Cloning sites include a 5-prime flanking BamHI
restriction site, an AflIII site at the start codon, a 3-
prime flanking HpaI site or a BglII site converting the
stop codon to an isoleucine (Fig. 1).

15 Amino terminal fusions to GFPm were created using
synthetic oligonucleotides encoding the signal sequence
flanked by a 5-prime BamHI or BglII restriction site and
3-prime NcoI site. The synthetic signal sequences were
cloned into PHP7921 (Fig. 4) using the 5-prime flanking
20 BamHI and AflIII restriction sites of GFPm to create in-
frame fusions. Sequences cloned in this manner include
the maize chloroplast signal sequence, barley alpha-
amylase signal sequence, pea glutathione reductase signal
sequence, and SV40 nuclear localization signal. The
25 sporamin vacuole signal sequence was synthesized with a
5-prime BspEI and 3-prime BamHI for insertion between the
BAA pre-sequence and GFPm coding sequence of pHP8144
(Fig. 5). The CPN60 promoter site, created by site-
directed mutagenesis, was used with the GFPm AflIII site
30 to make the fusion sequence. Each of plasmids PHP7921
and PHP8144 are based on pUC18.

35 Carboxy terminal sequences were synthesized with
a 5-prime BglII and 3-prime HpaI site, and cloned into
pHP7921 or pHP8144 using the 3-prime BflII and HpaI sites
of GFPm. The barley lectin vacuole target, HDEL ER
retention signal, SKL peroxisome signal, ALS coding
sequence and HRGP coding sequence fusions were made using
this strategy.

TABLE 1
PLASMIDS

PHP Number	Target Crop	Promoter	Targeting				Other 5' Sequences	Structural Gene	3' Sequence	Comments
			5' Sequence	3' Sequence	Localized To:					
7956	maize	4xERE	--	--	Cytoplasm	Omega Prime (O') Adh1 intron I	GFPm	pinII	Estrogen-inducible expression	
7667	maize	ubiquitin	--	--	Cytoplasm	--	GFP(native)	nos	Jellyfish structural gene	
7921	maize	ubiquitin	--	--	Cytoplasm	--	GFPm	pinII	Cytoplasmic expression	
8489	maize	ubiquitin	--	--	Cytoplasm	--	GFPm	pinII	MAR flanking cassette	
8490	maize	ubiquitin	--	--	Cytoplasm	--	GFPm-S65T	pinII	MAR flanking cassette, red-shift	
8080	maize	ubiquitin	CTP	--	chloroplast	--	GFPm	pinII	Chloroplast signal sequence, cab-m7 from maize	

8087	maize	CPN60	CPN60	--	mitochondria a	--	GFPm	pinII	Targeted to mitochondria, signal from maize chaparonin 60
8144	maize	ubiquitin	BAA	--	secreted	--	GFPm	pinII	Secreted GFP; signal from barley alpha amylase 1.
8757	maize	ubiquitin	BAA	BLVT	vacuole	--	GFPm	pinII	Barley lectin vacuole signal seq. on carboxy end of BAA- GFPm
8758	maize	ubiquitin	BAA	HDEL	endoplasmic reticulum	--	GFPm	pinII	HDEL endoplasmic reticulum signal from maize b-70
8759	maize	ubiquitin	BAA	--	vacuole	SVT	GFPm	pinII	Sweet potato sporamin vacuole signal sequence

WO 97/41228

PCT/US97/07688

- 53 -

8882	maize	ubiquitin	--	SKL	peroxisome	--	GFPm	pinII	Consensus peroxisomal signal sequence at carboxy terminus
8986	maize	ubiquitin	T7 TAG-NLS-	--	nucleus	--	GFPm	pinII	Nuclear localization sequence
9053	maize	ubiquitin	T7 TAG-NLS-	--	nucleus	--	MALS-GFPm	pinII	Nuclear localization sequence; protein fusion to prevent escape
9057	maize	ubiquitin	--	--	nucleus	--	MALS-GFPm	pinII	Fusion protein to test effect on GFP as reporter
9058	maize	ubiquitin	--	--	cytoplasm	--	PATm-GFPm	pinII	PAT-GFP fusion to test efficacy

8744	maize	ubiquitin	--	--	nucleus	--	GFPm- ZmRAD51	pinII	Protein fusion: for studies on Rad51 expression/ function
8081	maize	ubiquitin	--	--	cytoplasm	FRT	GFPm	pinII	To verify that FRT sequence does not impair expression
8674	maize	2x358	--	--	cytoplasm	O' Adh1 intron	FRT-bar- FRT-GFPm	pinII	Select on bialaphos, then excise bar and activate GFP
8911	maize	ubiquitin /2x358	--	adh1 intron I	cytoplasm	Ds	GFPm bar	pinII pinII	ubi-Ds-2x35S: :bar::pin-Ds- GFPm::pinII
8088	maize	ubiquitin	--	--	cytoplasm	--	GFPm-S65T	pinII	Red-shifted GFPm
8138	maize	ubiquitin	--	--	cytoplasm	--	GFPm-Red2	pinII	Red-shifted version #2 F64M, S65G, Q69L

WO 97/41228

PCT/US97/07688

-55-

8175	maize	ubiquitin	--	--	cytoplasm	--	Blue FP	pinII	GFPm-modified Y66H
8007	maize	2x35s ubiquitin					adh1 intronI::bar ubi intron::FLP	pinII pinII	Tandem vector for FLP excision of FRT to activate GFPm
8011	sunflo wer	2x35s ALS	--	--	cytoplasm	O'	GFPm NPT-II	pinII pinII	Binary
8491	sunflo wer	2x35s ALS	--	--	cytoplasm	O'	GFPm-S65T NPT-II	pinII pinII	Binary with red-shifted GFP

The ubiquitin promoter is described in Christensen et al., *Plant Mol. Biol.* 18: 675-689 (1992). The 4xERE promoter is described in Klein-Hitpass et al., *Cell* 46: 1053-1061 (1986). The CPN60; 2x35S and ALS promoters are described in Close P. Ph.D. Dissertation, Iowa State University; Gardner et al., *Nucl. Acid Res.* 9: 2871-2888 (1981) and U.S. Patent No: 5,659,026, respectively. The Omega prime (O') 5-prime sequence is described by Gallie et al., *Nucl. Acids Res.* 15: 3257-3273 (1987) while the Adh 1 intron I is described by Dennis et al., *Nucl. Acids Res.* 12: 3983-3990 (1984). The 3-prime sequence pinII is described in An et al., *Plant Cell* 1: 115-122 (1989). Structural genes used in the construction of the vectors include the GFP gene (native jellyfish) described by Prasher et al., *Gene* 111: 229-233 (1992); the maize ALS gene described by Fromm et al., *Bio/Technol.* 8: 883 (1990) and the maize HPRG described by Steifel et al., *Plant Mol. Biol.* 11: 483-493 (1988). The FLP/FRT system contained in plasmids PHP8674 and PHP8007 is described in O'Gorman et al., *Science* 251: 1351-1355 (1991) and Jayaram M., *Proc. Natl. Acad. Sci. USA* 82: 5875-5879 (1985). The Ac/Ds system is described by Muller-Neumann et al., *Mol. Gen. Genet.* 198: 19-24 (1984).

For *Agrobacterium*-mediated transformation of sunflower, the backbone of PHP8011 is a bin 19-based vector. Beven, *Nucl. Acids Res.* 12: 8711-8721 (1984). The NPT-II cassette was removed from bin 19 and the following cloned between the left and right border sequences; the double 35S promoter, omega', the GFPm coding sequence and the pinII 3' region-(all described above) were cloned downstream of the left border sequence, followed in a head-to-tail orientation by the BnALS3 promoter, the NPT-II coding sequence and the pinII 3' region. Beven, "Binary *Agrobacterium* Vectors for Plant Transformation", *Nucl. Acids Res.* 12: 8711-8721 (1984).

GFP_r, BFP and GFPs (soluble GFP) were constructed by means of oligonucleotide-directed mutagenesis of the GFP_m nucleotide sequence shown in Figure 1 as described by Ausubel, *supra*. With regard to GFP_r, TCC was changed to ACG resulting in the S65T substitution in the GFP amino acid sequence. With regard to BFP, TAC was changed to CAC, resulting in the Y66H substitution in the GFP amino acid sequence.

With regard to GFPs, the following sequence changes in GFP_m were made. TCC was changed to AGC, resulting in a F99S substitution in the GFP amino acid sequence. The oligonucleotide used to introduce this mutation also caused a silent mutation at R96 (AGC changed to CGG), thereby adding a BsrBI restriction site. ATG was changed to ACC, resulting in an M153T substitution in the GFP amino acid sequence. GTG was changed to GCC, resulting in a V163A substitution in the GFP amino acid sequence. A single oligonucleotide was used to change the amino acids at 153 and 163. This oligonucleotide also caused a silent mutation at A154 (GCC changed to GCG), which added a SacII restriction site. As reported by Cramer *et al.*, *supra*, these three amino acid changes did not alter the absorption/emission characteristics of GFP. These three amino acid substitutions, however, increased the solubility of GFP. Maize cells transformed with the GFPs gene fluoresce in blue to UV light based on transient assays.

Example 2**Maize Transformation with GFP Screening or Bialaphos Selection**

Maize Hi-II immature embryos were co-transformed via microprojectile bombardment with a mixture of PHP7814 (ubi::PAT::pinII) and a plasmid comprising the GFP gene. Plasmid PHP7814 carries the PAT gene encoding resistance to bialaphos operably linked to the ubiquitin promoter and the terminator from the pinII gene. Standard CaCl₂/spermidine precipitation of DNA onto tungsten microprojectiles was used to prepare the DNA-particles. Prior to bombardment, the embryos of approximately 1.5-2.0 mm in length were cultured on 560P medium including N6 salts, Erikkson's vitamins, 0.69 mg/L proline, 2 mg/L 2,4-D, and 3% sucrose. After 4-5 days of incubation in the dark at 28°C, embryos were removed from 560P medium and cultured, coleorhizal end up, on 560L medium which is equivalent to 560P except that it contains 12% sucrose. Embryos were allowed to acclimate to this medium for 3 h prior to transformation.

Embryos were transformed using the PDS-1000 Helium Gun from Bio-Rad at one shot per sample using 650 PSI rupture disks. On average approximately 0.0067 µg of DNA was delivered per shot. For each treatment (pat gene with bialaphos selection or GFP gene with screening for GFP), 447 embryos were used. For the bialaphos-selection treatment; following bombardment, the embryos were maintained on 560L medium for 48 h before transfer to 560R selection medium containing N6 salts, Erikkson's vitamins, 2 mg/L 2,4-D, 3% sucrose and 3mg/L bialaphos. Plates were maintained at 28°C in the dark and were observed for colony recovery. Transfer of colonies to fresh medium occurred every two weeks. Transgenic colony recovery under bialaphos

selection was accomplished by picking the healthy, growing transgenic calli from amongst the inhibited background of non-transgenic tissue as the material was progressively transferred to fresh media. For the GFP screening treatment, 560P media was used after bombardment and the two-day recovery on 560L. Calli were transferred to fresh non-selective 560P medium every two weeks. Screening for GFP expression was carried out at each transfer using a Xenon and/or Mercury light source with the appropriate filters for GFP visualization.

Table 2

Recovery of Transgenic Events in Maize Hi-II Callus
Using
Either GFP Visualization or Bialaphos Selection

	Number of Events Recovered	
Selection Method	Initial Events	Events Capable of Regeneration
Bialaphos selection	16	12
Screened using GFP	9	8

Transgenic events were recovered from each treatment. The largest number of transgenic events were consistently obtained by means of bialaphos selection. However, the frequency of recovery of transgenic events with GFP screening was approximately one-half the rate of recovery with bialaphos selection which was surprisingly high considering that GFP screening does not rely on selection of transgenic events in the presence of a cellular toxin.

Once GFP expressing colonies were identified they were monitored regularly for new growth and expression using the Xenon light source. Plant cells containing GFP were regenerated by transferring the callus to 288 medium containing MS salts, 1 mg/L IAA,

0.5 mg/L zeatin and 4% sucrose. The callus was placed in the light. As plantlets developed they were transferred to tubes containing 272K, hormone-free MS medium and 3% sucrose. The percentage of green
5 fluorescent colonies that regenerated into whole plants was determined.

Leaf samples were collected from TO transgenic plants from events recovered by means of bialaphos selection, or for events recovered by means of GFP
10 visualization (non-selected). Genomic DNA was extracted from all leaf samples and Southern analysis was performed using the PAT and GFP genes as hybridization probes. Hybridization for the PAT gene and GFP gene were performed sequentially on the same
15 blots, by stripping and reprobing for the other transgene. Bands hybridizing for the PAT gene and for the GFP gene were observed in all 8 GFP-screened events tested and in all 12 of the bialaphos selected events tested. Copy numbers and integration patterns for both
20 treatments appeared similar, and were consistent with previous experiments by the inventors using bialaphos selection of callus. Copy numbers ranged from low (single or two copies for each gene) to numerous (i.e., >5 copies per gene). Integration events were
25 predominantly at single integration sites but in some cases integration at multiple locations in the genome was observed.

Bialaphos is the most efficient chemical selective agent for maize. Consequently, the high
30 frequency with which transgenic events were obtained with GFP screening, compared to bialaphos selection, was unexpected. An advantage of GFP screening over bialaphos selection is that non-transformed tissue is not killed. Accordingly, transgenic sectors containing
35 the GFP gene develop in a healthier environment for growth and development.

Example 3**The GFP Marker in Maize Cytoplasm, Chloroplast and Mitochondria**

5 Maize A188 x B73-derived cells were transformed with plasmids PHP8080, PHP7921 and PHP8087. Each of these plasmids is described in Example 1 and Table 1. Plasmid PHP8080 carries the GFPm gene operably linked to the ubiquitin promoter and a chloroplast signal
10 sequence. GFP is directed to the chloroplasts of cells transformed with PHP8080. Plasmid PHP7921 carries the GFPm gene operably linked to ubiquitin promoter. GFP remains in the cytoplasm of cells transformed with pPGP7921. Plasmid PHP8087 carries the GFPm gene
15 operably lined to the CPN60 promoter and CPN60 signal sequence. GFP is directed to mitochondria of cells transformed with pPGP8087.

The cells were transformed by culturing maize embryos approximately 1.5 to 2 mm in length onto 560P medium containing N6 salts, Eriksson's vitamins, 0.69
20 g/L Proline, 2 mg/L 2,4-D and 3% sucrose. After 4-5 days of incubation in the dark at 28°C, embryos were removed from 560P medium and cultured, coleorhizal end up, onto 560L medium which is equivalent to 560P but
25 contains 12% sucrose. Embryos were allowed to acclimate to this medium for 3 h prior to transformation. Embryos were transformed using the PDS-1000 Helium Gun from Bio-Rad at one shot per sample using 650PSI rupture disks. DNA delivered per shot
30 averaged at 0.0667 µg. Following bombardment, all embryos were maintained on 560L medium for 48 hours before transfer to 3% sucrose and 3 mg/L bialaphos. Plates were maintained at 28°C in the dark and were observed for colony recovery with transfers to fresh
35 medium occurring every two weeks. Transgenic colony recovery was noted initially as growing callus tissue with a healthy phenotype on selection. Screening for

GFP expression was completed with each transfer using a Xenon light source with UV filter attachments.

GFP localization was confirmed by analysis of TO callus. Samples of calli from the targeted GFP transformations were fixed in FAA and then examined using an inverted microscopes with UV filters to visualize GFP. For each targeted construct GFP expression was localized to the specified organelle.

Once GFP expressing colonies were identified they were monitored regularly for new growth and expression using the Xenon light source. Plant cells containing GFP were regenerated by transferring the callus to 288 medium containing MS salts, 1 mg/L IAA, 0.5 mg/L zeatin and 4% sucrose. The callus was placed in the light. As plantlets developed they were transferred to tubes containing 272K, hormone-free MS medium and 3% sucrose. The percentage of green fluorescent colonies that regenerated into whole plants was determined.

Regeneration of plants from callus transformed with PHP8080, PHP7921, and PHP8087 appeared to vary between constructs; plants were successfully regenerated from independent transformed calli at respective frequencies of 35, 54 and 86%.

Transformation was confirmed by Southern blot analyses using both the BAR and GFPm genes as hybridization probes. GFP localization in various subcellular compartments was confirmed in stable transgenic maize cells using epifluorescent microscopy and image enhancement software.

Table 3

**Regeneration Frequency Of Maize Transformed With GFP
Directed to Different Subcellular Compartments**

Chloroplast					
Exp #	Total Colonies	Green	Regenerate	Percentage	Ave. Plant Event
..4.01	32	23	4	17	6
..4.03	0	-	-	-	-
..4.06	13	11	8	73	3
Totals	45	34	12	35*	
Cytoplasm					
Exp #	Total Colonies	Green	Regenerate	Percentage	Ave. Plants Event
..4.01	7	6	2	33	4
..4.03	7	5	4	80	5
..4.06	8	7	6	86	4
..4.08	9	9	3	33	2
..4.11	3	3	1	33	4
Totals	34	30	16	54*	
Mitochondrial					
Exp #	Total Colonies	Green	Regenerate	Percentage	Ave. Plants Event
..4.03	5	3	1	33	9
..4.06	13	11	11	100	4
Totals	18	14	12	86*	
* This is not a mean regeneration frequency, but instead is calculated from the totals for all treatments.					

Transformed plants containing GFP in the mitochondrial subcellular compartment grew more vigorously than plants containing GFP in the cytoplasm or chloroplast. The transformed plants containing GFP in the mitochondria were equivalent to the controls in terms of growth rate and appearance.

In contrast, transformed plants containing GFP in the chloroplast were difficult to regenerate. This callus produced a preponderance of roots with few shoots. Almost all the plants containing GFP in the chloroplast that were regenerated and reached maturity clearly exhibited stress with narrow and somewhat flaccid leaves, reduced stature, yellow leaves (presumably caused by reduced chlorophyll levels) and male sterility.

Germination of T1 embryos have been tested in the following manner. Aliquots of seed from a number of independent transformants were imbibed and the embryos were excised and placed on hormone-free MS medium with 3% sucrose for germination (this in vitro test made it easier to visualize the seedling). All embryos were germinated in the dark and then scored for visual GFP expression using UV-blue excitation under the dissecting microscope.

For each event tested aliquots of 5 or 10 seeds/event were imbibed, excised and scored for germination and fluorescence. The germination frequency for excised progeny produced on plants transformed with PHP7921 and PHP8080 was good. The germination frequency for excised progeny produced on plants transformed with PHP8087 was good for 2 events (10/10), low for one event (2/10). No detectable germination was observed for 2 events. This lowered germination frequency could be due to mitochondrially-expressed GFP expression per se, but it could also be the result of event-to-event variation sometimes observed in transgenic lines.

Germination of a single event from plants transformed with PHP7921 was tested differently. Seed were screened using a UV-blue light source, and 10 seed with GFP-expressing seed (embryos and endosperm) were selected. Five of these seed were planted in soil, and in the other 5 the embryos were excised and germinated on medium. None of the seed in soil germinated, while

all excised embryos germinated, grew and the seedlings continued to fluoresce. For events where germination appears to be negatively impacted, this could be due to increased GFP concentrations during seed maturation and dry-down. High GFP concentrations and aggregation of the protein has been found to be toxic in *E. coli*. Cramer et al., *Nature Biotechnol.* 14: 315-319 (1995).

Table 4
T1 Progeny Analysis
Plant Germination and GFPm Expression

Plasmid	Event Tested	# of T1 Seed Tested	# of T1 Seed Germinated Per Event	# of T1 Plants Expressing GFPm
PHP8087: CPN60::mit-GFPm-PinII	1	10	0	0
PHP8087: CPN60::mit-GFPm-PinII	2	10	0	0
PHP8087: CPN60::mit-GFPm-PinII	3	10	10	1
PHP8087: CPN60::mit-GFPm-PinII	4	10	2	0
PHP8087: CPN60::mit-GFPm-PinII	5	10	10	2
PHP7921: Ubi::Ubi intron-GFPm-PinII	1	10	10	4
PHP8080: Ubi::Ubi intron-CTP-GFPm-PinII	1	5	5	0
PHP8080: Ubi::Ubi intron-CTP-GFPm-PinII	2	5	4	0
PHP8080: Ubi::Ubi intron-CTP-GFPm-PinII	3	5	4	0
PHP4810: Ubi::Ubi intron-BAR-PinII	1	5	5	0
*PHP7921: Ubi::Ubi intron-GFPm-PinII	2	10 5 germinated in vitro and 5 germinated in soil	5 Germination occurred only from in vitro cultures	5 Seed planted in soil also were GFPm+

* For all seed except those designated in the PHP7921 transformation, embryos were excised from imbibed seed and germinated on MS medium containing no hormones and 3% sucrose. All embryos were germinated in the dark to allow for visualization of GFPm expression. One event from PHP7921 was tested by planting one half of the seed in soil and one half with embryos excised from endosperm with culture in vitro. All seed planted in soil did not germinate. The germination rate for embryos excised from the endosperm and cultured in vitro was 100%. All plants expressed GFPm.

Example 4**Sunflower Transformation with GFP Screening, GUS Screening and Kanamycin Selection**

5

10

15

20

25

Experiments were undertaken to test GFP, GUS and NPTII as marker genes for sunflower transformation. Two different kinds of meristem systems (split meristem; intact meristem) and two different kinds of particle systems (FG, focusing gun; He, helium gun) were used. The screenable markers GFP and GUS were tested with the intact meristem system under conditions involving no chemical selection but rather selection of transformants based on visual (GFP) or assayable (GUS) expression in transformed sectors. The NPTII marker gene was tested with the split meristem method using kanamycin for chemical selection of transformants and assays for NPTII expression to identify transformed sectors. In some cases, the experiments were carried only so far as identifying transformed sectors while other experiments were carried out all the way through to test for seed transmission of the marker gene to progeny. Transmission to progeny was confirmed by NPTII ELISA.

30

35

Regardless of the exact meristem system or type of particle gun used, the basic transformation protocol involved a combination of wounding by particle bombardment, followed by use of *Agrobacterium* for DNA delivery, as described by Bidney et al., *Plant Mol. Biol.* 18: 301-313 (1992). The plasmids used for plant transformation in these experiments were PHP167 containing GUS as the screenable marker, PHP8011 containing GFP as the screenable marker as well as the NPTII gene, and PHP762 containing NPTII as the selectable marker. Both PHP167 and PHP762 are binary vectors derived from pGA473. See An et al., *EMBO J.* 4: 277-284 (1985). The binary backbone for PHP8011 is bin 19. See Bevan, *NAR* 12: 8711-8721 (1984).

Procedures for preparation of *Agrobacterium* and preparation of particles for all of the experiments are described in Bidney et al., *supra*. The Pioneer sunflower line SMF3 was used in all experiments. See Burrus et al., *Plant Cell Rep.* 10: 161-166 (1991) for information concerning SMF3. The *Agrobacterium* strain EHA101 is described in Bidney et al. *supra* and strain EHA105 is described in Hood et al., *Transgen. Res.* 2: 208-218 (1993). Procedures for use of the helium gun, intact meristem preparation, tissue culture and co-cultivation conditions, and transgenic plant recovery for the intact meristem experiments are also described in Bidney et al., *supra*. Procedures for the use of the focusing gun, i.e., as in the intact meristem experiments, are described in Sautter et al., *Bio/Technol.* 9: 1080-1085 (1991). Procedures for use of the helium gun, split meristem preparation, tissue culture, co-cultivation, kanamycin selection and transgenic plant recovery for the split meristem experiments are described in Malone-Schoneberg et al., *Plant Sci.* 103: 199-207 (1994).

Parameters and general use of the focusing gun were similar to those 1.4 μm described in Sautter et al., *supra*. Gold particles measuring at a concentration of $0.5 \times 10^6/\mu\text{l}$ and 80 bars of nitrogen were typically used although different sized particles (0.25 to 2.5 μm) and different nitrogen pressures (40-160 bar) were tested. For the Dupont PDS-1000 helium driven particle bombardment device, the meristems were bombarded twice at the highest shelf with 600 psi rupture discs and a vacuum of 26 inches Hg and 1.8 μm tungsten particles. The intact meristem transformation method involved imbibing seed for 24 hours in the dark, removing the cotyledons and root radical, followed by culturing of the meristem explants. Twenty-four hours later, the primary leaves were removed to expose the apical meristem. The explants were placed apical dome side up and bombarded twice with particles, followed by

co-cultivation with *Agrobacterium*. To start the co-cultivation for intact meristems a 0.5 μ l droplet of OD⁶⁰⁰ 4.0 *Agrobacterium* was placed on the meristem. After a 3-day co-cultivation the meristems were transferred to culture medium with 250 μ g/ml cefotaxime (+ 100 mg/l kanamycin for the NPTII selection experiments). Selection has also been done using 50 gm/l kanamycin. The split meristem method involved imbibing seed, breaking of the cotyledons to produce a clean fracture at the plane of the embryonic axis, excising the root tip and then bisecting the explants longitudinally between the primordial leaves. The two halves were placed cut surface up on the medium then bombarded twice with particles, followed by co-cultivation with *Agrobacterium*. For split meristems, after bombardment the meristems were placed in an OD⁶⁰⁰ 0.6 *Agrobacterium* suspension for 30 minutes. They were then removed from the suspension onto solid culture medium for 3-day co-cultivation. After this period, the meristems were transferred to fresh medium with 250 μ g/ml cefotaxime (+ 100mg/l kanamycin for selection). Selection has also been done using 50 gm/l kanamycin.

Table 5 summarizes the results from experiments looking at the frequency of transformed sectors in primary shoots and secondary branches. Table 6 summarizes the results from experiments carried through to progeny for analysis of seed transmission. Transformants were identified in either seed or plants by assaying for marker gene expression. For GUS expression, activity was determined using either histochemical staining or the fluorometric assay, see Jefferson et al., *EMBO J.* 6: 3901-3907 (1987). Inheritance of transgene expression was assessed using an NPTII ELISA.

TABLE 5
Overall Transformed Sector Frequencies
for Sunflower Transformation

MARKER	PLASMID PHP #	GUN	MERISTEM PREPARATION	AVE. % SECTORS	REPS (N)
Screening:					
GUS	167	FG	Intact	22.0	15
GUS	167	FG	Intact	15.0	3
GUS	167	He	Intact	13.6	12
GFP	8011	He	Intact	7.5	1
Selection:					
NPTII/Kan	762	FG	Intact	12.7	3

TABLE 6

Overall Sector and Seed Transmission
for Sunflower Transformation

MARKER	PLASMID PHP #	GUN	MERISTEM PREPAR- ATION	AVE. % SECTORS	REPS (N)	SEED TRANS- MISSION
Screening:						
GUS	167	FG	Intact	15.0	3	0/8
GFP	8011	He	Intact	7.5	1	9/10
Selection:						
NPTII/KAN	762	FG	Intact	12.7	4	0/2
NPTII/KAN	8011	He	Split	1.8	4	6/8

For evaluation of initial sector frequencies, GUS seemed to be the most effective marker gene (Tables 5 and 6). However, based on screening alone (no selection) the observed GUS sectors did not subsequently contribute to the germline and thus did not transmit to progeny (Table 6). It was unlikely that this was due to the use of the focusing gun because, both quantitatively and qualitatively, mapping of GUS sectors after use of each type of gun was comparable. GFP was significantly better than NPTII.

(7.5% vs. 1.8%, see Table 6) for both sector frequency and seed transmission when compared using the helium gun. Progeny from GFP-expressing (no chemical selection but instead visually selecting using GFP) TO plants were analyzed for NPTII expression (Table 6) Progeny from 10 plants were tested and 9 showed inheritance of NPTII activity. From kanamycin-selected TO plants, 6 out of 8 tested showed transmission of NPTII expression to the progeny. Not only has GFP "visual selection" worked unexpectedly well for recovering the primary transformants (i.e. compare Tables 3 and 4), but cotransformation and inheritance of transgene expression was also very efficient.

Table 7: Progeny Analysis of Transformation Experiments of SMF3
with GFP (PHP8011: CaMV-GFPm, ALS-NPTII)

Treatment (Comments)	SID	No. of Seeds Screened	TO NPTII Map Range	Maximum NPTII	Transmissi on Efficiency
Topdown Non-selected (from 1 experiment, 120 explants in total)	281976	246	12.6-->30	9.9	9/10
	281977	139	16.3-->	11.9	
	281978	127	>30	>30	
	281979	175	9-27	6	
	281980	74	>30	4	
	281983	15	3-21.3	6.5	
	319454	124	4.1-18.7	5.9	
	319453	118	++	8.6	
	281982	47	>30	6.2	
	281981	3	>30	0	
Split selected (from 4 experiments, 955 explants in total)	248112	6	++	2.3	6/8
	248113	24	>30	7.8	
	248111	41	>30	>30	
	248107	23	17	2.7	
	248114	67	>30	1.2	
	248110	8	4.1	0	
	248109	6	10.7	0	
	248115	86	++	5.5	

Parallel experiments comparing GFP screening, GUS screening and NPTII/kanamycin selection can be performed using intact meristems, the *Agrobacterium* strain EHA105 and the helium gun, in side-by-side comparisons. Plasmid PHP8011 can be used for GFP/NPTII comparisons. An identical plasmid substituting the GUS structural gene for GFP can be used for GUS/NPTII comparisons, and both plasmids can be used in side-by-side experiments for GUS/GFP comparisons. Sector frequency in primary shoots and secondary shoots can be

assayed and calculated as described above. Inheritance and expression of transgenes can also be analyzed in these experiments.

Example 5**GFP As A FLP-Mediated Excision Marker In Maize**

5 Immature embryos of the Hi-II genotype were
bombarde d with PHP8674 which contains a GFPm reporter
gene that is activated upon FLP recombinase-mediated
FRT excision. Plasmid PHP8674 also carries a
10 constitutively expressed gene for bialaphos resistance.
Stably transformed callus was selected on bialaphos and
examined under UV-blue light for GFP expression. No
GFP expression was observed in these primary
transformants. Pieces of callus transformed with
PHP8674 were then bombarded with PHP8007. Plasmid
15 PHP8007 contains the FLP recombinase gene operably
linked to the ubiquitin promoter. Callus sectors
expressing GFP were first observed ten days after the
second bombardment. These GFP-expressing sectors
continued to grow and express the transgene after one
20 month of culture. As a negative control the FRT/GFPm
stably transformed callus was also bombarded with a
non-FLP containing plasmid. This treatment did not
produce GFP expressing sectors.

 As a second control for this experiment the
25 consecutive treatments were done in opposite order.
Hi-II immature embryos were bombarded with PHP8007
which carries the FLP recombinase gene and a bialaphos
resistance gene. Stable transformants were recovered
following selection on medium containing bialaphos.
30 These calli were re-bombarded with PHP8674. GFP
expression was observed in the callus but this faded
over 5-10 days until it could no longer be detected.
The stably transformed callus expresses FLP
recombinase. Introduction of PHP8674 likely provided
35 the substrate for excision (the FRT sites along with
spacer DNA between the sites) which activated the GFP
cassette while still on the plasmid. This "activated

WO 97/41228

PCT/US97/07688

-75-

expression" was only transient and disappeared as the plasmid was degraded.

Example 6**GFP Targeting to the Nucleus as a Fusion Protein**

5 Plasmid PHP9053 was constructed in which the GFP
protein was fused to the ALS enzyme (designated MALS
for maize-ALS). The ALS gene, which encodes a protein
normally found in the cytoplasm, is typically used to
confer herbicide resistance. Fang et al., *Plant Mol.*
10 *Biol.* 18: 1185-1187 (1992). A nuclear localization
signal from simian virus 40 was fused to the N terminus
of GFPm as described in Example 1 (Kalderon et al.,
Cell 39: 499-509 (1984) with the ALS enzyme fused to
the C terminus. The sequence encoding this nuclear
15 targeted fusion protein (in construct PHP9053) was
introduced into maize cells using particle bombardment
as described previously. Once gene expression was
confirmed through GFP fluorescence under the
microscope, tissue was fixed as described earlier and
20 examined to determine subcellular localization of GFP.
Fluorescence was localized to the nucleus. Plant
regeneration and assessment of toxicity for this
experiment has not yet been done.

 GFP was targeted to the nucleus by fusing the
25 protein to another polypeptide that is normally
compartmentalized in the nucleus. For example, the
GFPZm-RAD51 fusion found in plasmid PHP8744 contains
RAD51, a protein required for mitotic and meiotic
recombination and double-stranded break repair, fused
30 to GFP. The gene encoding RAD51 was cloned from yeast.
Haaf et al., *Proc. Nat Acad. Sci.* 92: 2298-2302 (1995).
Degenerate primers based on the nucleotide sequence of
the RAD51 gene cloned from yeast were used to identify
and isolate clones containing the RAD51 gene from
35 maize. The carboxy terminus of GFPm was fused to the
amino terminus of ZmRAD51a. Maize cells were
transformed with the GFPm-ZmRAD51 fusion and
fluorescence was localized in the nucleus. Plant

regeneration and assessment of toxicity for this experiment has not yet been done.

5 Although the foregoing refers to particular preferred embodiments, it will be understood that the present invention is not so limited. It will occur to those of ordinary skill in the art that various modifications may be made to the disclosed embodiments and that such modifications are intended to be within the scope of the present invention, which is defined by
10 the following claims.

 All publications and patent applications mentioned in this specification are indicative of the level of skill of those in the art to which the
15 invention pertains.

20

-78-

SEQUENCE LISTING

GENERAL INFORMATION:

APPLICANT: PIONEER HI-BRED INTERNATIONAL, INC.

TITLE OF INVENTION: USE OF THE GREEN FLUORESCENT PROTEIN AS A
SCREENABLE MARKER FOR PLANT TRANSFORMATION

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS: 800 Capital Square, 400 Locust Street, Des
Moines, IA 50309, U.S.A.

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: CA 2,502,657

FILING DATE: 01-MAY-1997

CLASSIFICATION: C12N - 15/12

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/016,345

FILING DATE: 01-MAY-1996

PATENT AGENT INFORMATION:

NAME: TORYS LLP

REFERENCE NUMBER: 31539-2196

INFORMATION FOR SEQ ID NO.:1:

SEQUENCE CHARACTERISTICS:

LENGTH: 717 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..714

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 1..714

SEQUENCE DESCRIPTION: SEQ ID NO.:1:

ATG	TCC	AAG	GGC	GAG	GAG	CTC	TTC	ACC	GGC	GTG	GTG	CCC	ATC	CTC	GTG	48
Met	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	
1				5					10					15		
GAG	CTC	GAC	GGC	GAC	GTG	AAC	GGC	CAC	AAG	TTC	TCC	GTG	TCC	GGC	GAG	96

-79-

Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	
			20					25					30			
GGC	GAG	GGC	GAC	GCC	ACC	TAC	GGC	AAG	CTC	ACC	CTC	AAG	TTC	ATC	TGC	144
Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	
		35					40					45				
ACC	ACC	GGC	AAG	CTC	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	TTC	192
Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Phe	
	50					55					60					
TCC	TAC	GGC	GTG	CAG	TGC	TTC	TCC	AGG	TAC	CCC	GAC	CAC	ATG	AAG	CAG	240
Ser	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	
65					70					75					80	
CAC	GAC	TTC	TTC	AAG	TCA	GCC	ATG	CCC	GAG	GGC	TAC	GTG	CAG	GAG	AGG	288
His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	
				85					90					95		
ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	AGG	GCC	GAG	GTG	336
Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	
			100					105					110			
AAG	TTC	GAA	GGC	GAC	ACC	CTC	GTG	AAC	AGG	ATT	GAG	CTC	AAG	GGC	ATC	384
Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	
		115					120					125				
GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTC	GGC	CAC	AAG	CTC	GAG	TAC	AAC	432
Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	
	130					135					140					
TAC	AAC	TCC	CAC	AAC	GTG	TAC	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	GGC	480
Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly	
145					150					155					160	
ATC	AAG	GTG	AAC	TTC	AAG	ATC	AGG	CAC	AAC	ATC	GAG	GAC	GGC	TCA	GTG	528
Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	
				165					170					175		
CAG	CTC	GCT	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	CCC	576
Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	
			180					185					190			
GTG	CTC	CTC	CCC	GAC	AAC	CAC	TAC	CTC	TCC	ACC	CAG	TCC	GCC	CTC	TCC	624
Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	
		195					200					205				
AAG	GAC	CCC	AAC	GAG	AAG	AGG	GAC	CAC	ATG	GTG	CTC	CTC	GAG	TTC	GTG	672
Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val	
	210					215					220					
ACC	GCT	GCT	GGC	ATC	ACC	CAC	GGC	ATG	GAC	GAG	CTC	TAC	AAG			714
Thr	Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp	Glu	Leu	Tyr	Lys			
225					230					235						

TGA

717

INFORMATION FOR SEQ ID NO.:2:

SEQUENCE CHARACTERISTICS:

LENGTH: 238 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:2:

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

INFORMATION FOR SEQ ID NO.:3:

SEQUENCE CHARACTERISTICS:

LENGTH: 159 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE :

NAME/KEY: CDS
LOCATION: 1..159

FEATURE :

NAME/KEY: mat_peptide
LOCATION: 1..159

SEQUENCE DESCRIPTION: SEQ ID NO.:3:

[illegible]

INFORMATION FOR SEQ ID NO.:4:

SEQUENCE CHARACTERISTICS:

LENGTH: 53 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:4:

Met Ala Ser Ser Thr Met Ala Leu Ser Ser Thr Ala Phe Ala Gly Lys
1 5 10 15

Ala Val Asn Val Pro Ser Ser Ser Ala Phe Glu Ala Arg Val Thr Met

-82-

```

                20                25                30
Arg Lys Thr Ala Ala Lys Ala Lys Pro Ala Ala Ala Ser Gly Ser Pro
      35                40                45

Trp Tyr Gly Pro Met
      50

```

INFORMATION FOR SEQ ID NO.:5:

SEQUENCE CHARACTERISTICS:

LENGTH: 537 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
 LOCATION: join(1..69, 496..537)

FEATURE:

NAME/KEY: mat_peptide
 LOCATION: join(1..69, 496..537)

SEQUENCE DESCRIPTION: SEQ ID NO.:5:

```

ATG TAC CGC GCG GCC GCT AGC CTC GCC TCC AAG GCG CGG CAA GCC GGG      48
Met Tyr Arg Ala Ala Ala Ser Leu Ala Ser Lys Ala Arg Gln Ala Gly
  1                5                10                15

AGC AGC TCC GCC GCT CGC CAG GTGAGAGCAG ACTCGTGTTT ATACGCGTGT      99
Ser Ser Ser Ala Ala Arg Gln
      20

GATGGGTCTG ATGGAGGATC CCGCCCTCAG ATTTGGTAAT GTTGGCCTGG GGTTCTAGTA  159

GTTCTGCGTG CAGGGCTGTG GGTTTGCTGC ATGGTGCTGT TTATTTTGGT GGCGATCTGC  219

CGGAATCTGT AGTTCGCTCG CGCAAAATCT AAGCTAGCTC GCTAATGGCG TACTGGCGTG  279

GGGTTTCACA TTAATCTACG GTGGTGAAC TCGTCACTACC GTCCTCCAGT TAGCTGTTAG  339

ACACCGAATA CACTGATTGG CAGTTGAGAA ACATTGATCT GATCCAGCAG AAATCGATGT  399

CTTGTGAAAT TCGTTATTTT ATTGTCGTGT AACCTTGGGG CATGGCAGTC TCTAATTGAT  459

CACGCACTCA CCTCTGTTGT GTGTGATGCT TTATAG GTT GGA AGC AGG CTT GCC      513
                        Val Gly Ser Arg Leu Ala
                        25

TGG AGC AGG AAC TAT GCT GCC ATG      537
Trp Ser Arg Asn Tyr Ala Ala Met
  30                35

```

-83-

INFORMATION FOR SEQ ID NO.:6:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:6:

Met	Tyr	Arg	Ala	Ala	Ala	Ser	Leu	Ala	Ser	Lys	Ala	Arg	Gln	Ala	Gly
1				5					10					15	
Ser	Ser	Ser	Ala	Ala	Arg	Gln	Val	Gly	Ser	Arg	Leu	Ala	Trp	Ser	Arg
			20					25					30		
Asn	Tyr	Ala	Ala	Met											
			35												

INFORMATION FOR SEQ ID NO.:7:

SEQUENCE CHARACTERISTICS:

LENGTH: 72 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
 LOCATION: 1..72

FEATURE:

NAME/KEY: mat_peptide
 LOCATION: 1..72

SEQUENCE DESCRIPTION: SEQ ID NO.:7:

ATG	GCC	AAC	AAG	CAC	CTG	AGC	CTC	TCC	CTC	TTC	CTC	GTG	CTC	CTC	GGC	48
Met	Ala	Asn	Lys	His	Leu	Ser	Leu	Ser	Leu	Phe	Leu	Val	Leu	Leu	Gly	
1				5					10					15		
CTC	TCC	GCC	TCC	CTC	GCC	TCC	ATG									72
Leu	Ser	Ala	Ser	Leu	Ala	Ser	Met									
				20												

INFORMATION FOR SEQ ID NO.:8:

SEQUENCE CHARACTERISTICS:

LENGTH: 24 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

-84-

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:8:

Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15
Leu Ser Ala Ser Leu Ala Ser Met
20

INFORMATION FOR SEQ ID NO.:9:

SEQUENCE CHARACTERISTICS:
LENGTH: 45 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
LOCATION: 1..45

FEATURE:

. NAME/KEY: mat_peptide
LOCATION: 1..45

SEQUENCE DESCRIPTION: SEQ ID NO.:9:

GTG TTC GCC GAG GCC ATC GCC GCC AAC TCC ACC CTC GTG GCC GAG 45
Val Phe Ala Glu Ala Ile Ala Ala Asn Ser Thr Leu Val Ala Glu
1 5 10 15

INFORMATION FOR SEQ ID NO.:10:

SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:10:

Val Phe Ala Glu Ala Ile Ala Ala Asn Ser Thr Leu Val Ala Glu
1 5 10 15

INFORMATION FOR SEQ ID NO.:11:

SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid

-85-

STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
LOCATION: 1..33

FEATURE:

NAME/KEY: mat_peptide
LOCATION: 1..33

SEQUENCE DESCRIPTION: SEQ ID NO.:11:

GAC GGC GGC GTG GAC GAC GAC CAC GAC GAG CTC
Asp Gly Gly Val Asp Asp Asp His Asp Glu Leu
1 5 10

33

INFORMATION FOR SEQ ID NO.:12:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:12:

Asp Gly Gly Val Asp Asp Asp His Asp Glu Leu
1 5 10

INFORMATION FOR SEQ ID NO.:13:

SEQUENCE CHARACTERISTICS:

LENGTH: 141 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
LOCATION: 1..141

FEATURE:

NAME/KEY: mat_peptide
LOCATION: 1..141

SEQUENCE DESCRIPTION: SEQ ID NO.:13:

ATG GCC AAC AAG CAC CTG AGC CTC TCC CTC TTC CTC GTG CTC CTC GGC
Met Ala Asn Lys His Leu Ser Leu Ser Leu Phe Leu Val Leu Leu Gly
1 5 10 15

48

-86-

CTC TCC GCC TCC CTC GCC TCC GGA CAC AGC AGG TTC AAC CCC ATC AGG 96
 Leu Ser Ala Ser Leu Ala Ser Gly His Ser Arg Phe Asn Pro Ile Arg
 20 25 30

CTG CCC ACC ACC CAC GAG CCC GCC AGC AGC GAG ACC GGA TCC ATG 141
 Leu Pro Thr Thr His Glu Pro Ala Ser Ser Glu Thr Gly Ser Met
 35 40 45

INFORMATION FOR SEQ ID NO.:14:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:14:

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

Leu Ser Ala Ser Leu Ala Ser Gly His Ser Arg Phe Asn Pro Ile Arg
 20 25 30

Leu Pro Thr Thr His Glu Pro Ala Ser Ser Glu Thr Gly Ser Met
 35 40 45

INFORMATION FOR SEQ ID NO.:15:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 1..12

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 1..12

SEQUENCE DESCRIPTION: SEQ ID NO.:15:

AAG TCG AAG CTT
 Lys Ser Lys Leu
 1

12

INFORMATION FOR SEQ ID NO.:16:

-87-

SEQUENCE CHARACTERISTICS:
 LENGTH: 4 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:16:

Lys Ser Lys Leu
 1

INFORMATION FOR SEQ ID NO.:17:

SEQUENCE CHARACTERISTICS:
 LENGTH: 240 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..240

FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 1..240

SEQUENCE DESCRIPTION: SEQ ID NO.:17:

ATG AAC CAG GCG ATG GCC ACC CCG CTG TCC CTG TCC TGC TGC TCC CCG	48
Met Asn Gln Ala Met Ala Thr Pro Leu Ser Leu Ser Cys Cys Ser Pro	
1 5 10 15	
ACC CTG ACC AGG TCC ACC CTG TTC TTC ACC AAG ACC TTC CCG TTC TCC	96
Thr Leu Thr Arg Ser Thr Leu Phe Phe Thr Lys Thr Phe Pro Phe Ser	
20 25 30	
CGC TCC TTC TCC ACC CCG CTG CCG CTG TCC ACC AAG ACC CTG ATC TCC	144
Arg Ser Phe Ser Thr Pro Leu Pro Leu Ser Thr Lys Thr Leu Ile Ser	
35 40 45	
CTG TCC CCG CCG CAC AGG ACC TTC GCC GTG AGG GCT GAG TCC CAG AAC	192
Leu Ser Pro Pro His Arg Thr Phe Ala Val Arg Ala Glu Ser Gln Asn	
50 55 60	
GGC GCG GAC CCG GCC AGG CAG TAC GAC TTC GAC CTG TTC ACC ATC GGC	240
Gly Ala Asp Pro Ala Arg Gln Tyr Asp Phe Asp Leu Phe Thr Ile Gly	
65 70 75 80	

INFORMATION FOR SEQ ID NO.:18:

SEQUENCE CHARACTERISTICS:

-88-

LENGTH: 80 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:18:

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

INFORMATION FOR SEQ ID NO.:19:

SEQUENCE CHARACTERISTICS:
 LENGTH: 42 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
 LOCATION: 1..42

FEATURE:

NAME/KEY: mat_peptide
 LOCATION: 1..42

SEQUENCE DESCRIPTION: SEQ ID NO.:19:

CCG	CCC	AAG	AAG	AAG	CGC	AAG	GTG	CCC	ATG	AAG	ATC	CAC	ATG
Pro	Pro	Lys	Lys	Lys	Arg	Lys	Val	Pro	Met	Lys	Ile	His	Met
1				5					10				

42

INFORMATION FOR SEQ ID NO.:20:

SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein

-89-

SEQUENCE DESCRIPTION: SEQ ID NO.:20:

Pro Pro Lys Lys Lys Arg Lys Val Pro Met Lys Ile His Met
 1 5 10

INFORMATION FOR SEQ ID NO.:21:

SEQUENCE CHARACTERISTICS:

LENGTH: 456 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
 LOCATION: 1..456

FEATURE:

NAME/KEY: mat_peptide
 LOCATION: 1..456

SEQUENCE DESCRIPTION: SEQ ID NO.:21:

TAC	ACT	CCA	AGC	CCT	AAG	CCA	CCG	GCT	ACC	AAG	CCT	CCC	ACG	CCC	AAG	48
Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Ala	Thr	Lys	Pro	Pro	Thr	Pro	Lys	
1				5				10						15		
CCG	ACC	CCG	CCA	ACG	TAC	ACC	CCT	TCG	CCA	AAG	CCT	CCG	ACA	CCC	AAG	96
Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Thr	Pro	Lys	
			20					25					30			
CCG	ACC	CCG	CCG	ACG	TAC	ACC	CCT	TCT	CCC	AAG	CCT	CCG	ACG	CCC	AAG	144
Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Thr	Pro	Lys	
		35					40					45				
CCG	ACC	CCG	CCG	ACG	TAC	ACT	CCA	AGC	CCC	AAG	CCT	CCC	ACA	CAC	CCG	192
Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Thr	His	Pro	
	50					55					60					
ACG	CCC	AAG	CCG	ACC	CCA	CCG	ACG	TAC	ACC	CCT	TCC	CCA	AAG	CCT	CCG	240
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	
65					70					75					80	
ACG	CCC	AAG	CCG	ACC	CCA	CCG	ACG	TAC	ACC	CCT	TCC	CCA	AAG	CCT	CCG	288
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	
				85					90					95		
ACA	CCC	AAG	CCG	ACC	CCA	CCG	ACG	TAC	ACC	CCT	TCC	CCA	AAG	CCT	CCG	336
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	
			100					105					110			
ACA	CCC	AAG	CCG	ACC	CCA	CCG	ACG	TAC	ACT	CCC	ACA	CCG	AAG	CCG	CCG	384
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Thr	Pro	Lys	Pro	Pro	

-90-

115	120	125	
GCC ACC AAG CCG CCC ACC TAC ACT CCG ACG CCG CCG GTG TCT CAC ACC			432
Ala Thr Lys Pro Pro Thr Tyr Thr Pro Thr Pro Pro Val Ser His Thr			
130	135	140	
CCC AGC CCG CCG CCA CCA TAC TAC			456
Pro Ser Pro Pro Pro Pro Tyr Tyr			
145	150		

INFORMATION FOR SEQ ID NO.:22:

SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:22:

Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Ala	Thr	Lys	Pro	Pro	Thr	Pro	Lys
1				5					10					15	
Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Thr	Pro	Lys
			20					25					30		
Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Thr	Pro	Lys
			35				40					45			
Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro	Thr	His	Pro
	50					55					60				
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro
65					70					75					80
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro
				85					90					95	
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Ser	Pro	Lys	Pro	Pro
			100					105					110		
Thr	Pro	Lys	Pro	Thr	Pro	Pro	Thr	Tyr	Thr	Pro	Thr	Pro	Lys	Pro	Pro
			115				120					125			
Ala	Thr	Lys	Pro	Pro	Thr	Tyr	Thr	Pro	Thr	Pro	Pro	Val	Ser	His	Thr
	130					135					140				
Pro	Ser	Pro	Pro	Pro	Pro	Tyr	Tyr								
145						150									

INFORMATION FOR SEQ ID NO.:23:

-91-

SEQUENCE CHARACTERISTICS:

LENGTH: 621 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
 LOCATION: 1..621

FEATURE:

NAME/KEY: mat_peptide
 LOCATION: 1..621

SEQUENCE DESCRIPTION: SEQ ID NO.:23:

GTG	GCT	GAG	TAC	ACG	CTG	CCG	GAT	CTG	GAT	TAC	GAC	TAC	AGC	GCC	CTG	48
Val	Ala	Glu	Tyr	Thr	Leu	Pro	Asp	Leu	Asp	Tyr	Asp	Tyr	Ser	Ala	Leu	
1				5				10						15		
GAA	CCC	CAC	ATC	TCC	GGG	CAG	ATC	AAC	GAG	CTG	CAC	CAT	TCC	AAG	CAC	96
Glu	Pro	His	Ile	Ser	Gly	Gln	Ile	Asn	Glu	Leu	His	His	Ser	Lys	His	
			20					25					30			
CAC	GCC	GCC	TAC	GTC	GCC	GGT	GCC	AAC	ACG	GCA	CTG	GAG	AAG	CTG	GAA	144
His	Ala	Ala	Tyr	Val	Ala	Gly	Ala	Asn	Thr	Ala	Leu	Glu	Lys	Leu	Glu	
			35				40					45				
GCC	GCC	CGT	GAG	GCC	GGC	GAT	CAC	AGC	GCG	ATC	TTC	CTG	CAC	GAG	AAG	192
Ala	Ala	Arg	Glu	Ala	Gly	Asp	His	Ser	Ala	Ile	Phe	Leu	His	Glu	Lys	
	50					55					60					
AAC	CTC	GCG	TTC	CAC	CTC	GGC	GGA	CAC	GTC	AAC	CAC	TCC	ATC	TGG	TGG	240
Asn	Leu	Ala	Phe	His	Leu	Gly	Gly	His	Val	Asn	His	Ser	Ile	Trp	Trp	
65					70				75					80		
AAG	AAC	CTG	TCC	CCC	AAC	GGT	GGC	GAC	AAG	CCG	GTC	GGC	GAG	CTG	GCC	288
Lys	Asn	Leu	Ser	Pro	Asn	Gly	Gly	Asp	Lys	Pro	Val	Gly	Glu	Leu	Ala	
				85				90						95		
GCG	GCC	ATC	GAC	GAC	CAG	TTC	GGT	TCG	TTC	GAC	AAG	TTC	CGC	GCG	CAG	336
Ala	Ala	Ile	Asp	Asp	Gln	Phe	Gly	Ser	Phe	Asp	Lys	Phe	Arg	Ala	Gln	
			100					105					110			
TTC	ACC	GCC	GCG	GCC	AAC	GGC	CTG	CAG	GGC	TCG	GGC	TGG	GCG	GTG	CTC	384
Phe	Thr	Ala	Ala	Ala	Asn	Gly	Leu	Gln	Gly	Ser	Gly	Trp	Ala	Val	Leu	
		115					120					125				
GGT	TAC	GAC	ACC	CTC	GGC	CAG	AAG	CTG	CTG	ACC	TTC	CAG	CTC	TAC	GAC	432
Gly	Tyr	Asp	Thr	Leu	Gly	Gln	Lys	Leu	Leu	Thr	Phe	Gln	Leu	Tyr	Asp	
	130					135				140						
CAG	CAG	GCC	AAC	GTG	CCG	CTG	GGC	ATC	ATC	CCG	CTG	CTC	CAG	GTC	GAC	480
Gln	Gln	Ala	Asn	Val	Pro	Leu	Gly	Ile	Ile	Pro	Leu	Leu	Gln	Val	Asp	
145					150					155					160	

-92-

ATG	TGG	GAG	CAC	GCC	TTC	TAC	CTG	CAG	TAC	AAG	AAC	GTC	AAG	GCC	GAC	528
Met	Trp	Glu	His	Ala	Phe	Tyr	Leu	Gln	Tyr	Lys	Asn	Val	Lys	Ala	Asp	
			.	165					170					175		
TAC	GTG	ACC	GCG	TTC	TGG	AAC	GTC	GTC	AAC	TGG	GCC	GAC	GTG	CAG	GAC	576
Tyr	Val	Thr	Ala	Phe	Trp	Asn	Val	Val	Asn	Trp	Ala	Asp	Val	Gln	Asp	
			180					185					190			
CGC	TTC	GGC	AAG	GCC	GTC	AAC	CAG	GGC	AAG	GGC	CTT	ATC	TTC	GGG		621
Arg	Phe	Gly	Lys	Ala	Val	Asn	Gln	Gly	Lys	Gly	Leu	Ile	Phe	Gly		
		195					200					205				

INFORMATION FOR SEQ ID NO.:24:

SEQUENCE CHARACTERISTICS:

LENGTH: 207 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO.:24:

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

-93-

Met	Trp	Glu	His	Ala	Phe	Tyr	Leu	Gln	Tyr	Lys	Asn	Val	Lys	Ala	Asp
				165					170					175	
Tyr	Val	Thr	Ala	Phe	Trp	Asn	Val	Val	Asn	Trp	Ala	Asp	Val	Gln	Asp
			180					185					190		
Arg	Phe	Gly	Lys	Ala	Val	Asn	Gln	Gly	Lys	Gly	Leu	Ile	Phe	Gly	
		195					200					205			

-94-

What is claimed is:

1. A method for producing a transgenic plant, wherein said transgenic plant is produced by positive selection, by screening for a transformed plant cell without selection of a cell carrying a gene that confers resistance to a toxic substance, comprising the steps:

- (a) providing an expression vector comprising
 - (i) a first promoter operably linked to a nucleotide sequence encoding a green fluorescent protein flanked on the 5-prime and 3-prime ends with a recombinase-specific target sequence, and
 - (ii) a second promoter operably linked to a foreign gene;
- (b) introducing said expression vector into regenerable plant cells;
- (c) screening for a transformed plant cell containing said green fluorescent protein by positive selection, wherein said screening does not depend on negative selection of a cell carrying a gene that confers resistance to a toxic substance;
- (d) transforming the plant cell identified by said screening of step (c), or a second plant cell derived from said plant cell, with a second expression vector containing a gene encoding a site-specific recombinase;
- (e) screening for a plant cell that no longer expresses the green fluorescent protein; and
- (f) regenerating a transformed plant from a cell identified by said screening of step (e).

2. The method according to claim 1, wherein said site-specific recombinase is FLP/FRT, Ac/DS or cre/lox.

3. The method of any one of claims 1 or 2 wherein the plant is a monocot.

4. The method of claim 3 wherein the plant is maize.

5. The method of any one of claims 1 or 2 wherein the plant is a dicot.

-95-

6. The method of claim 5 wherein the plant is sunflower.
7. The method of any one of claims 1-6 wherein the green fluorescent protein (GFP) is GFPm as set forth in SEQ ID NO: 1.
8. The method of any one of claims 1-7 wherein the site-specific recombinase is FLP/FRT.
9. The method of any one of claims 1-8 wherein the screening in step (c) is done by illumination with UV-blue light.
10. Use of a first expression vector and a second expression vector to transform a plant cell, wherein the first expression vector, comprises:
 - (i) a first promoter operably linked to a nucleotide sequence encoding a green fluorescent protein flanked on the 5-prime and 3-prime ends with a recombinase-specific target sequence, and
 - (ii) a second promoter operably linked to a foreign gene,and the second expression vector comprises:
 - (i) a gene encoding a site-specific recombinasewherein the transgenic plant is produced by positive selection, by screening for a transformed plant cell without selection of a cell carrying a gene that confers resistance to a toxic substance.
11. The use according to claim 10, wherein said site-specific recombinase is FLP/FRT, Ac/DS or cre/lox.
12. The use of any one of claims 10 or 11 wherein the plant is a monocot.
13. The use of claim 12 wherein the plant is maize.
14. The use of any one of claims 10 or 11 wherein the plant is a dicot.
15. The use of claim 14 wherein the plant is sunflower.

-96-

16. The use of any one of claims 10-15 wherein the green fluorescent protein (GFP) is GFPm as set forth in SEQ ID NO: 1.

17. The use of any one of claims 10-16 wherein the site-specific recombinase is FLP/FRT.

18. The use of any one of claims 10-17 wherein the screening is done by illumination with UV-blue light.

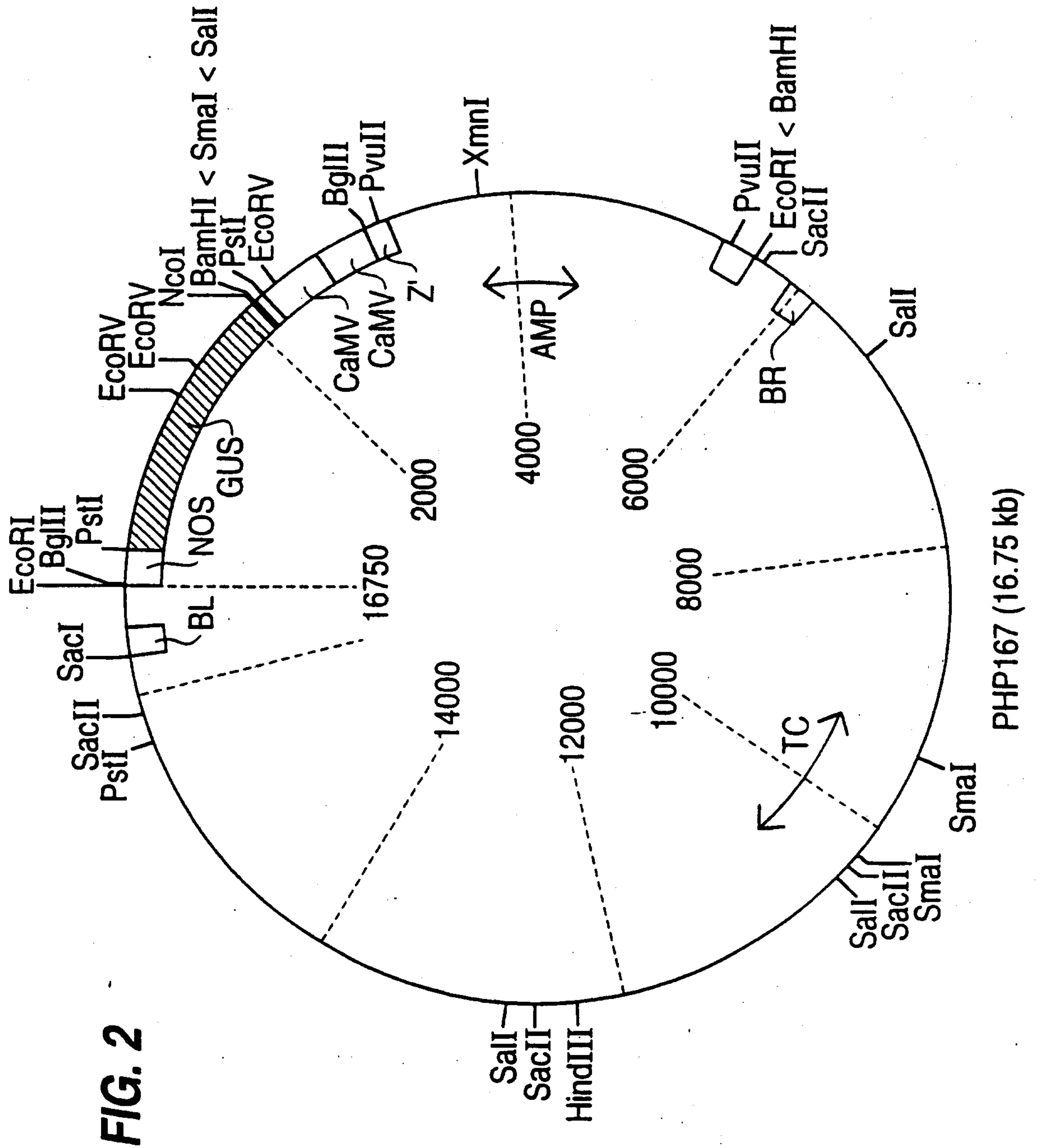
FIG. 1A

1 ATGTCCAAGGGCGAGGAGCTCTTACCGGGGTGGTGCCCATCTCTGGAGCTCGACGGC + 60
-----+-----+-----+-----+-----+-----+
TACAGTTCCCGCTCCTCGAGAAGTGCGCCGACCAACGGGTAGGAGCACCTCGAGCTGCCG
M S K G E E L P T G V V P I L V E L D G
61 GACGTGAACGGCCACAAGTTCTCCGTGTCCGGCGAGGGCGAGGGCGACCCACCTACGGC + 120
-----+-----+-----+-----+-----+-----+
CTGCACTTGCCGGTGTTCAGAGGACAGGCCGCTCCCGCTCCCGCTGCGGTGATGCCG
D V N G H K F S V S G E G E G D A T Y G
121 AAGCTCACCCTCAAGTTCACTGCACCACGGCAAGTCCCGTGCCCTGCCACCCTC + 180
-----+-----+-----+-----+-----+-----+
TTCGAGTGGAGTTCAAGTAGACGTGGTGGCCGTTCGAGGGGCACGGGACCGGGTGGGAG
K L T L K F I C T T G K L P V P W P T L
181 GTGACCACCTTCTCCTACGGCGTGCGAGTGCTTCTCCAGGTACCCCGACCATGAAGCAG + 240
-----+-----+-----+-----+-----+-----+
CACTGGTGAAGAGGATGCCCGCACGTACGAAGAGGTCCATGGGGCTGGTACTTCGTC
V T T F S Y G V Q C F S R Y P D H M K Q
241 CACGACTTCTCAAGTCAGCCATGCCCGAGGGCTACGTGCAGGAGGACCATCTCTTC + 300
-----+-----+-----+-----+-----+-----+
GTGCTGAAGAAGTTCAAGTCGGTACGGGCTCCCGATGCACGTCTCTCTGGTAGAAGAAG
H D F F K S A M P E G Y V Q E R T I F P
301 AAGGACGCGCAACTACAAGACCAGGGCCGAGGTGAAGTTCGAAGGCGACACCCTCGTG + 360
-----+-----+-----+-----+-----+-----+
TTCCTGCTGCCGTTGATGTTCTGTGTCGGCTCCACTTCAAGCTTCCGCTGTGGGAGCAC
K D D G N Y K T R A E V K F E G D T L V

2/6

FIG. 1B

361 AACAGGATTGAGCTCAAGGGCATCGACTTCAAGGAGGACGCAACATCCTCGGCCACAAG 420
-----+-----+-----+-----+-----+
TTGTCCTAACTCGAGTTCCCGTAGCTGAAGTTCTCCTCGCGTTGTAGGAGCCGGTGTTCTC
N R I E L K G I D F K E D G N I L G H K
CTCGAGTACAACCTCCACAACGTGTACATCATGGCCGACAAGCAGAAGAACGGC 480
-----+-----+-----+-----+-----+
GAGCTCATGTTGATGTTGAGGGTGTGCACATGTAGTACCGGCTGTCTCTTCTTGCCCG
421
L E Y N Y N S H N V Y I M A D K Q K N G
ATCAAGGTGAACCTTCAAGATCAGGCACAACATCGAGGACGGCTCAGTGCAGCTCGCTGAC 540
-----+-----+-----+-----+-----+
TAGTTCCACTTGAAGTTCTAGTCCGTGTGTAGCTCCTGCCGAGTCACGTCGAGCGACTG
I K V N F K I R H N I E D G S V Q L A D
CACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTCTCCTCCCGACAACCACTAC 600
541 -----+-----+-----+-----+-----+
GTGATGGTCTGTTGTGGGGTAGCCGCTGCCGGGCACGAGGAGGGGCTGTTGGTGATG
H Y Q Q N T P I G D G P V L L P D N H Y
CTCTCCACCAGTCCGCCCTCTCCAAGGACCCCAACGAGAAGAGGACCAATGGTGCTC 660
601 -----+-----+-----+-----+-----+
GAGAGTGGTCAAGCGCGGAGAGGTTCTCTGGGGTGTCTTCTCCTCGGTGTACCACGAG
L S T Q S A L S K D P N E K R D H M V L
CTCGAGTTCGTGACCGCTGCTGGCATCACCCACGGCATGGACGAGCTCTACAAGTGA 717
661 -----+-----+-----+-----+-----+
GAGCTCAAGCACTGGCGACGACCGTAGTGGGTGCCGTACCTGCTCGAGATGTTCACT
L E F V T A A G I T H G M D E L Y K *



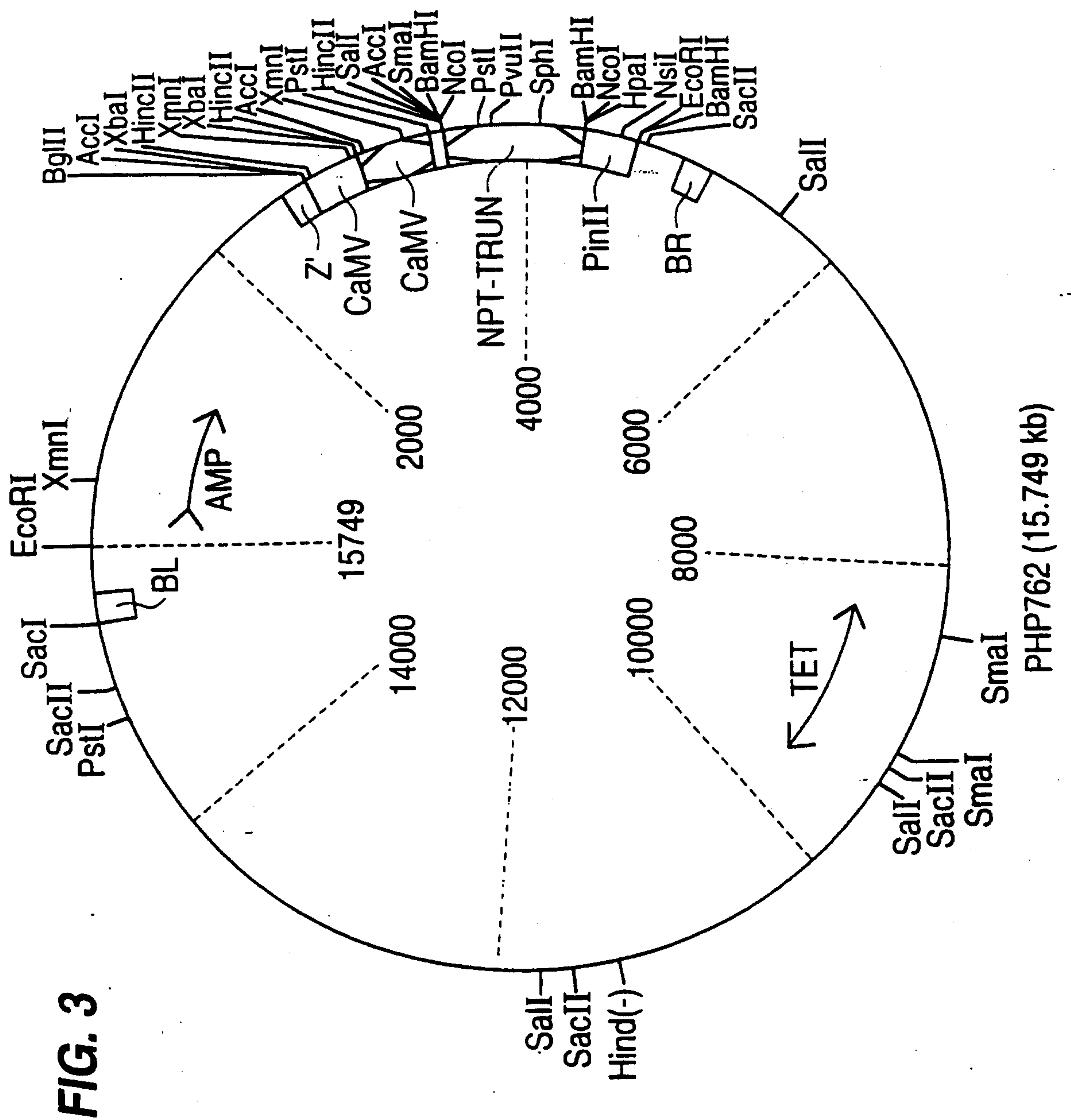
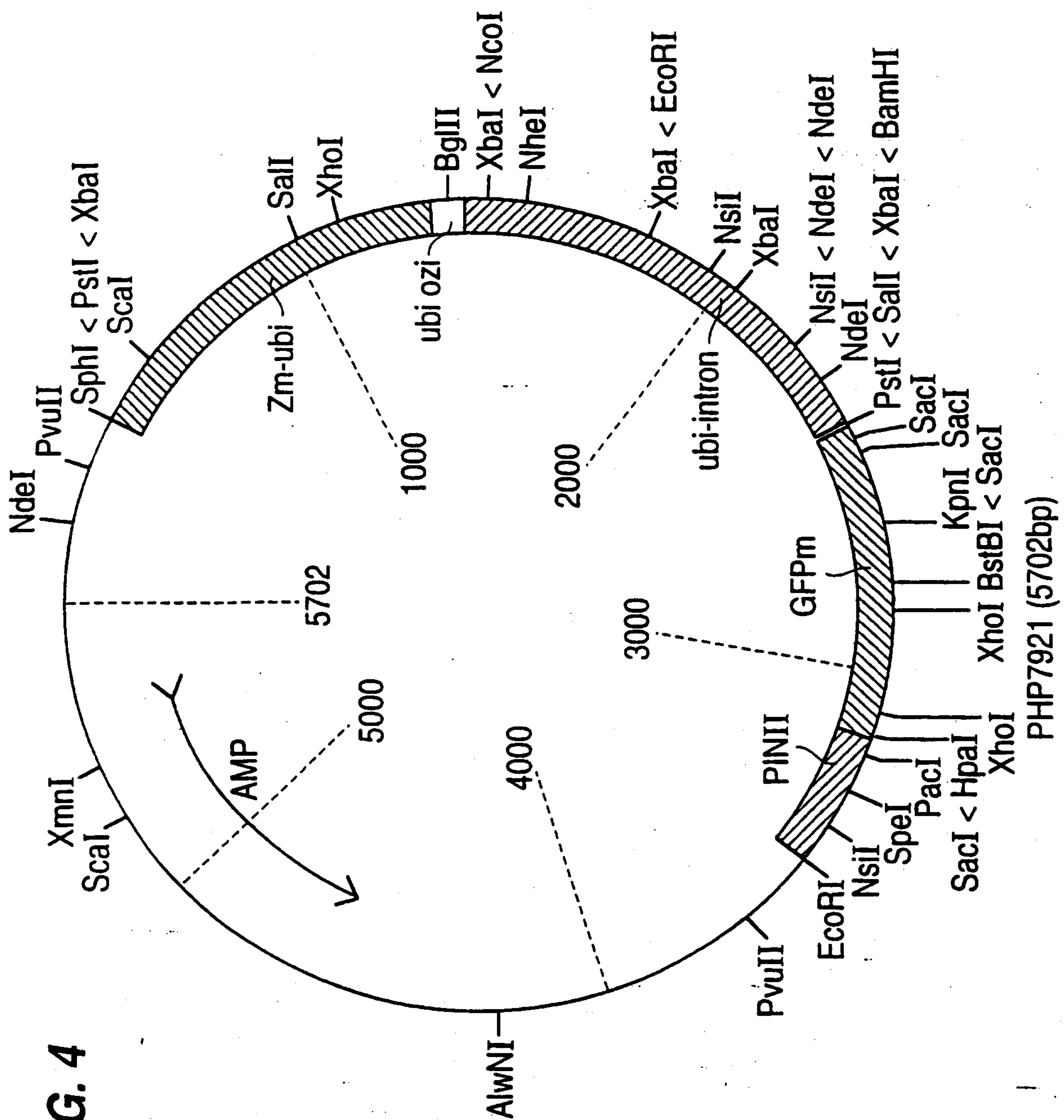
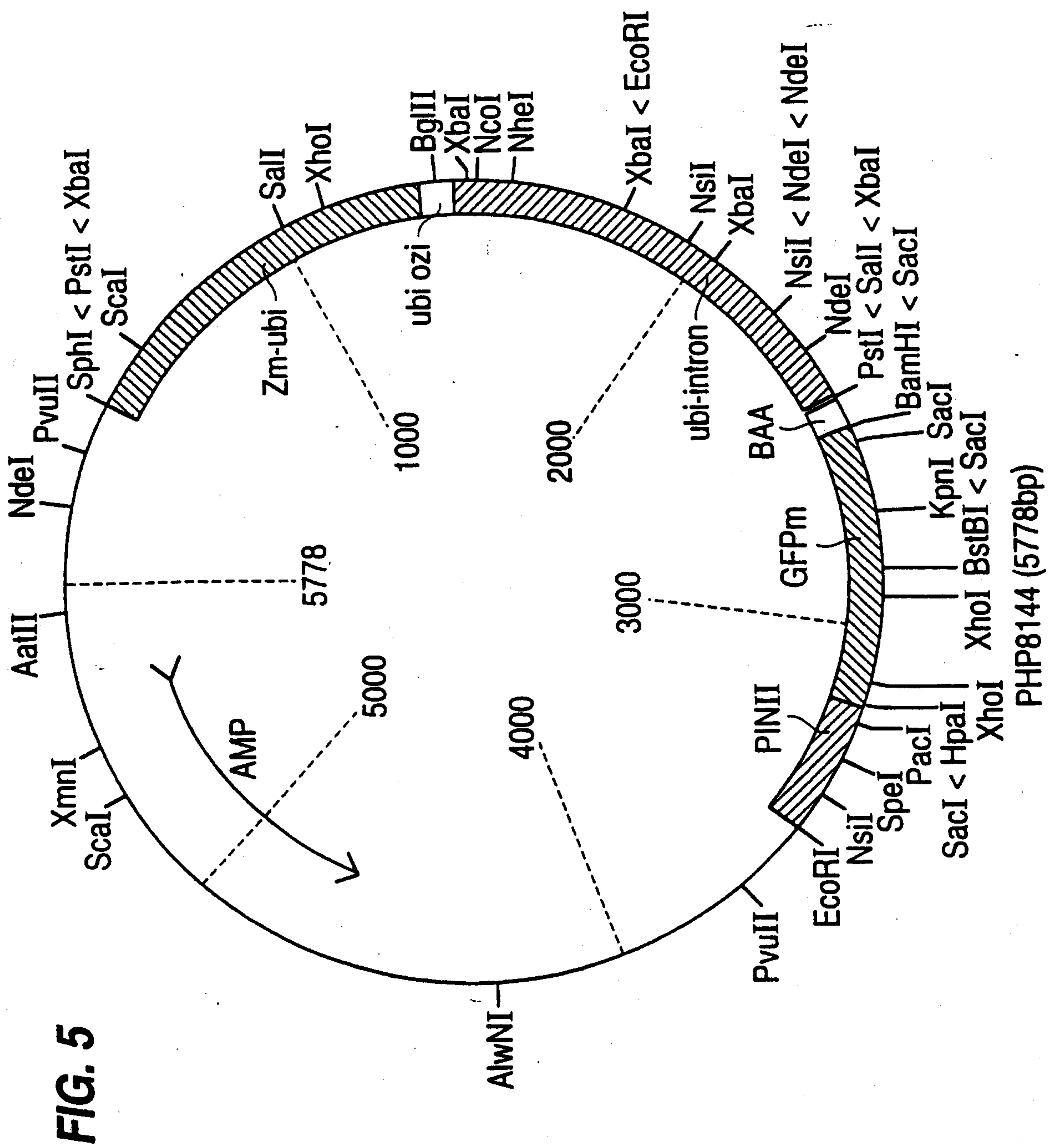


FIG. 3

FIG. 4





ATGTCCAAGGGCGAGGAGCTCTTCAACCGGCGTGGTGCCCATCCTCGTGGAGCTCGACGGC
1 -----+-----+-----+-----+-----+-----+-----+ 60
TACAGGTTCCCGCTCCTCGAGAAGTGGCCGCACCACGGGTAGGAGCACCTCGAGCTGCCG
M S K G E E L P T G V V P I L V E L D G
GACGTGAACGGCCACAAGTTCTCCGTGTCCGGCGAGGGCGAGGGCGACGCCACCTACGGC
61 -----+-----+-----+-----+-----+-----+ 120
CTGCACTTGCCGGTGTTCAAGAGGCACAGGCCGCTCCCGCTCCCGCTGCGGTGGATGCCG
D V N G H K F S V S G E G E G D A T Y G
AAGCTCACCTCAAGTTCATCTGCACCACCGGCAAGCTCCCCGTGCCCTGGCCCCACCCTC
121 -----+-----+-----+-----+-----+-----+ 180
TTCGAGTGGGAGTTCAAGTAGACGTGGTGGCCGTTCGAGGGGGCACGGGACCGGGTGGGAG
K L T L K F I C T T G K L P V P W P T L
GTGACCACCTTCTCCTACGGCGTGCAAGTGCTTCTCCAGGTACCCCGACCATGAAGCAG
181 -----+-----+-----+-----+-----+-----+ 240
CACTGGTGGGAAGAGGATGCCGCACGTCACGAAGAGGTCCATGGGGCTGGTGTACTTCGTC
V T T F S Y G V Q C F S R Y P D H M K Q
CACGACTTCTTCAAGTCAGCCATGCCCCGAGGGCTACGTGCAGGAGAGGACCATCTTCTTC
241 -----+-----+-----+-----+-----+-----+ 300
GTGCTGAAGAAGTTCAGTCGGTACGGGCTCCCGATGCACGTCCTCTCCTGGTAGAAGAAG
H D F F K S A M P E G Y V Q E R T I F F
AAGGACGACGGCAACTACAAGACCAGGGCCGAGGTGAAGTTCGAAGGCGACACCCTCGTG
301 -----+-----+-----+-----+-----+-----+ 360
TTCCTGCTGCCGTTGATGTTCTGGTCCCGGCTCCACTTCAAGCTTCCGCTGTGGGAGCAC
K D D G N Y K T R A E V K F E G D T L V