superior plants for several generations to produce a series of inbred lines which, although different from each other, each breed true and are highly uniform; (3) crossing the selected inbred lines with unrelated inbred lines to produce hybrid progeny (F_1). During the inbreeding process, the vigor of the lines decreases. Vigor is restored in the F_1. Once the best hybrid is identified, the hybrid seed can be reproduced indefinitely as long as the homogeneity of the inbred parent is maintained.

A single cross hybrid is produced when two inbred lines are crossed to produce the F_1 progeny. A double cross hybrid is produced from four inbred lines crossed in pairs (A x B and C x D) and then the two F_1 hybrids are crossed again. Because the vigor exhibited by F_1 hybrids is lost in the next generation, seed from hybrid varieties is not used for planting stock.

Hybrid corn seed can be produced by manual detasseling. Alternative strips of two inbred varieties of corn are planted in a field, and the pollen-bearing tassels are removed from one of the inbreds. Providing that there is sufficient isolation from sources of foreign corn pollen, the ears of the detassels inbred (female) will be fertilized only by the other inbred (male), and the resulting seed is therefore hybrid and will form hybrid plants.

The laborious detasseling process can be avoided by using cytoplasmic male sterile (CMS) inbreds. Plants of a CMS inbred are fertilized with pollen from another inbred that is not male sterile. Pollen from the second inbred can contribute genes that make the hybrid plants male fertile. Such breeding methods are well known to the art. See, for example, Hallauer et al., In: Corn and Corn Improvement, pp. 463-564 (1988).

Herbicide resistant or tolerant plant variants can be obtained by several methods including, but not limited to, spontaneous variation and direct mutant selection in cultures, direct or indirect mutagenesis procedures on tissue cultures of all cell types, seeds or plants, and mutation of cloned ACCase genes by methods such as site directed mutagenesis. For example, the preparation of herbicide tolerant maize from herbicide tolerant cell lines is described in U.S. Patent No. 5,162,602, issued November 10, 1992, the disclosure of which is

incorporated by reference herein, and in Examples I-III. Briefly, partially differentiated cell cultures are grown and subcultured with continuous exposures to low herbicide levels. Herbicide concentrations are then gradually increased over several subculture intervals. Maize cells or tissues growing in the presence of normally toxic herbicide levels are repeatedly subcultured in the presence of the herbicide and characterized. Stability of the herbicide tolerance trait of the cultured cells may be evaluated by growing the selected cell lines in the absence of herbicides for various periods of time and then analyzing growth after exposing the tissue to normally toxic amounts of herbicide. Mature maize plants are then obtained from maize cell lines that are known to express the trait. Thus, this method is useful to isolate heterozygous dominant, or partially dominant, herbicide resistant mutants.

If possible, the regenerated plants are self-pollinated. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important inbred lines. Conversely, pollen from plants of these inbred lines is used to pollinate regenerated plants. The genetics of the trait are then characterized by evaluating the segregation of the trait in the first and later generation progeny. Stable inheritance of the herbicide tolerance trait is achieved if the plants maintain herbicide tolerance for at least about three to six generations.

Seed from maize plants regenerated from transformed tissue cultures is grown in the field and self-pollinated to generate true breeding plants. Progenies from these plants become true breeding lines which are evaluated for herbicide tolerance in the field under a range of environmental conditions. Herbicide tolerance must be sufficient to protect the monocot plants at the maximum labeled delivery rate under field conditions which cause herbicides to be most active. Appropriate herbicide concentrations and methods of application are those which are known and have been developed for the cyclohexanedione and/or aryloxyphenoxypropanoic acid herbicides disclosed herein.

The commercial value of herbicide resistant corn is greatest if many different hybrid combinations are available for sale. The farmer typically grows

more than one kind of hybrid based on such differences as maturity, standability or other agronomic traits. Additionally, hybrids adapted to one part of the corn belt are not adapted to another part because of differences in such traits as maturity, disease, and insect resistance. Because of this, it is necessary to breed herbicide tolerance into a large number of parental lines so that many hybrid combinations can be produced.

To introduce a particular herbicide resistance allele into an inbred background, a conversion process (backcrossing) is carried out by crossing the original herbicide resistant line to normal elite lines and crossing the progeny back to the normal parent. The progeny from this cross will segregate such that some plants carry the gene responsible for tolerance whereas some do not. Plants carrying such genes will be crossed again to the normal parent resulting in progeny which segregate for herbicide resistance and sensitivity once more. This is repeated until the original normal parent has been converted to an herbicide resistant line, yet possesses all other important attributes as originally found in the normal parent. A separate backcrossing program is implemented for every elite line that is to be converted to herbicide resistant line.

Subsequent to the backcrossing, the new resistant lines and the appropriate combinations of lines which make good commercial hybrids are evaluated for herbicide resistance as well as a battery of important agronomic traits. Resistant lines and hybrids are produced which are true to type of the original normal lines and hybrids. This requires evaluation under a range of environmental conditions where the lines or hybrids will generally be grown commercially. Parental lines of hybrids that perform satisfactorily are increased and used for hybrid production using standard hybrid seed corn production practices.

The plants of the invention are expected to be useful for a variety of commercial and research purposes. The plants can be created for use in traditional agriculture to possess traits beneficial to the consumer of the grain harvested from the plant (e.g., improved nutritive content in human food or animal feed). In such uses, the plants are generally grown for the use of their

grain in human or animal foods. However, other parts of the plants, including stalks, husks, vegetative parts, and the like, may also have utility, including use as part of animal silage or for ornamental purposes.

The plants may also find use in the commercial manufacture of proteins or other molecules, where the molecule of interest is extracted or purified from plant parts, seeds, and the like. Cells or tissue from the plants may also be cultured, grown *in vitro*, or fermented to manufacture such molecules.

The plants may also be used in commercial breeding programs, or may be crossed or bred to plants of related crop species.

The invention will be further described by the following examples.

EXAMPLE I

Identification of Herbicide Mechanism and Site of Action

The objective of this Example was to identify the mechanism whereby sethoxydim and/or haloxyfop inhibit fatty acid synthesis in maize. The results, reported in J. D. Burton et al., Biochem. Biophys. Res. Comm., 148, 1039 (November 13, 1987), show that both sethoxydim and haloxyfop inhibit acetyl-coenzyme A carboxylase (ACCase) (EC 6.4.1.2) in maize chloroplasts.

A. Chemicals

Buffers and cofactors were purchased from Sigma Chemical Company (St. Louis, Missouri); [2-^{14}C]acetate was purchased from Research Products International; [2-^{14}C]pyruvate and [^{14}C]$NaHCO_3$ were purchased from New England Nuclear; and [2-^{14}C]malonyl coenzyme A was purchased from Amersham. Sethoxydim was a gift from BASF (Parsippany, New Jersey), and haloxyfop was provided by Dow Chemical USA (Midland, Michigan).

B. Plant Growth Conditions

Corn (Z. mays L., 'B37 x Oh43') seeds were germinated in darkness for 96 hours in vermiculite in an incubation chamber maintained at 30°C, 80% RH. Seedlings were then transferred to a growth chamber with a 16 hour light (25°C) and an 8 hour dark (20°C) cycle, 90% relative humidity (RH). After greening 48 hours, seedlings were returned to the dark incubation chamber for 12 hours to

deplete chloroplast starch reserves. Seedlings were harvested 6 days after planting. Pea (P. sativum L., 'PI 9901-C') seedlings were grown in vermiculite in a growth chamber with a 16 hour light (21°C) and 8 hour dark (16°C) cycle, 80% RH. Peas were harvested 10 to 13 days after planting. Black Mexican Sweet (BMS) corn suspension cultures were maintained in a supplemented Murashige-Skoog (MS) medium (C. E. Green, Hort. Sci., 12, 7-10 (1977)), and subcultured weekly by 20-fold dilution of the suspension culture into fresh medium.

C. Chloroplast Isolation

Chloroplasts from corn and pea seedlings were isolated at 4°C (K. Cline et al., J. Biol. Chem., 260, 3691-3696 (1985)). Seedlings (50 g of shoots) were homogenized in 200 ml buffer A (50 mM HEPES-NaOH pH 7.5, 330 mM sorbitol, 0.1% w/v BSA, 1 mM $MgCl_2$, 1 mM $MnCl_2$, 2 mM EDTA, 5 mM isoascorbate, 1.3 mM glutathione) in an omnimixer (five, 3-second bursts at full speed). The homogenate was filtered through six layers of cheesecloth and two layers of miracloth, and then centrifuged at 3000 g for 3 minutes with hand-braking. The pellet was gently resuspended in buffer A and layered onto a preformed linear Percoll gradient (50 mM HEPES-NaOH pH 7.5, 330 mM sorbitol, 1.9 mM isoascorbate, 1.08 mM glutathione, 0.1% w/v BSA, 50% Percoll) which was centrifuged at 3000 g for 20 minutes in a Sorvall HB-4 rotor. The lower band in the gradient, containing intact chloroplasts, was washed twice by gently resuspending it in 20 ml of buffer B (50 mM HEPES-NaOH, pH 7.5, and 330 mM sorbitol) followed by repelleting (3000 g, 5 minutes). The final pellet, consisting of intact chloroplasts, was resuspended in 2 to 3 ml of buffer B and stored on ice in the dark until use.

D. Fatty Acid Synthesis

[^{14}C]acetate and [^{14}C]pyruvate were used as precursors to measure fatty acid biosynthesis in isolated chloroplasts (B. Liedvogel et al., Planta, 169, 481-489 (1986)). [^{14}C]acetate incorporation was assayed in a 0.5 ml-volume containing: 50 mM HEPES-NaOH (pH 7.5), 330 mM sorbitol, 5 mM KH_2PO_4,

10 mM $NaHCO_3$, 1 mM $MgCl_2$, 1 mM ATP, 0.1 mM CoA, 0.15 mM [^{14}C]acetate (3.33 mCi/mmol), and chloroplasts (20 to 50 µg chlorophyll). [^{14}C]pyruvate incorporation into fatty acids was assayed in the same medium except that it included 2 mM TPP, 1 mM NAD^+, 0.15 mM [^{14}C]-pyruvate (1.33 mCi/mmol), but no acetate. Assay suspensions were illuminated with 1400 µE/m^2·second PAR at 25°C. Assays were initiated by the addition of the labelled substrate and stopped by the addition of 0.5 ml of 40% KOH. To determine the incorporation of radiolabel into a non-polar (fatty acid) fraction, each treatment was saponified at 90°C for 30 minutes in capped vials (P. B. Hoj et al., Carlsberg Res. Commun., 47, 119-141 (1982)). The vials were acidified with 0.5 ml 40% H_2SO_4, and carrier fatty acids (20 µg each of C 14:0, C 16:0, and C 18:0) were added. The assay mixture was extracted twice with 4 ml hexane. The extracts were combined, dried under N_2, and redissolved in 0.3 ml hexane. Aliquots (50 µl) were counted for radioactivity by liquid scintillation spectrometry.

Incorporation of [^{14}C]malonyl-Coenzyme A into fatty acids (P. B. Hoj et al., supra; and J. B. Ohlrogge et al., Proc. Natl. Acad. Sci. USA, 76, 1194-1198 (1979)) was assayed using cell-free preparations from BMS tissue culture. Cells harvested during logarithmic growth phase were frozen in liquid nitrogen, ground with a mortar and pestle, and thawed in a medium containing: 0.1 M HEPES-KOH, pH 7.5; 0.3 M glycerol, and 5 mM DTT (buffer:tissue, 2:1, v/w). The homogenate was centrifuged at 12,000 g for 20 minutes. The supernatant was filtered through miracloth and centrifuged (125,000 g) for 60 minutes and then filtered through miracloth and assayed. Assays were conducted at 25°C in a 0.4 ml volume containing: 1.0 mM ATP, 0.32 mM NADPH, 0.38 mM NADH, 25 µM CoA, 10 µM acetyl-CoA, 25 µg acyl-carrier protein, and 12 µM malonyl-CoA (11.54 µCi/µmol). Reactions were initiated by addition of [^{14}C]malonyl CoA and stopped by addition of 0.4 ml 40% KOH. Label incorporation into fatty acids was determined as above. Chlorophyll (D. I. Arnon, Plant Physiol., 24, 1-15 (1949)) and protein (P. K. Smith et al., Anal. Biochem., 150, 76-85 (1985)) were determined as described therein.

E. **Acetyl-Coenzyme A Carboxylase (ACCase) Activity**

Maize chloroplasts, isolated as described above, were suspended in buffer C (0.1 M Tricine-KOH, pH 8.0; 0.3 M glycerol, and 1 mM DTT) and homogenized in a glass tissue homogenizer. The disrupted chloroplast fraction was centrifuged at 16,000 g for 15 minutes. The supernatant was desalted on a Sephadex G-25 column (1.5 x 5 cm equilibrated with 0.1 M Tricine-KOH, pH 8.0; and 0.3 M glycerol) and assayed directly. ACCase activity (B. J. Nikolau et al., Arch. Biochem. Biophys., 211, 605-612 (1981)) was assayed at 30°C in a 0.2 ml volume which contained 1 mM ATP, 3 mM acetyl coenzyme A, 2.5 mM $MgCl_2$, 50 mM KCl, 0.5 mM DTT, and 15 mM [^{14}C]$NaHCO_3$ (0.17 mCi/mmol). Reactions were initiated by addition of acetyl coenzyme A and stopped by addition of 25 µl of 12 N HCl. Product formation was determined by the radioactivity found in an acid stable fraction by liquid scintillation spectrometry. Enzyme activity was linear for 15 minutes.

F. **Results**

To probe for the site of herbicidal activity of sethoxydim and haloxyfop, labelled acetate, pyruvate, and malonyl-CoA were used individually as precursors for fatty acid synthesis. Isolated chloroplasts from corn seedlings incorporated [^{14}C]acetate and [^{14}C]pyruvate into a non-polar fraction (fatty acids). Acetate incorporation was linear for 30 min after a 5 min lag period, and dependent upon the addition of free acetyl coenzyme A. Addition of either 10 µM sethoxydim or 1 µM haloxyfop inhibited [^{14}C]acetate incorporation into fatty acids by 90% and 89%, respectively, as shown in Table I, below. Sethoxydim

(10 μM) and haloxyfop (1 μM) also inhibited the incorporation of [^{14}C]pyruvate into fatty acids by 98% and 99%, respectively.

TABLE I

Inhibition of [^{14}C]acetate and [^{14}C]pyruvate Incorporation into Fatty Acids in Corn Seedling Chloroplasts by Sethoxydim (10 μM) and Haloxyfop (1 μM), 10 minute assay time

	Acetate	Pyruvate
	---Activity (nmol/mg chl·min)---	
Control	4.4 ± 0.4[1]	10.8 ± 2.3
	----------% Inhibition----------	
Sethoxydim	90 ± 2.5	98 ± 1.1
Haloxyfop	89 ± 3.1	99 ± 0.3

[1]Results are expressed as mean of two experiments ± standard error.

The effect of 10 μM sethoxydim and 1 μM haloxyfop on [^{14}C]malonyl-CoA incorporation into fatty acids was determined using cell-free extracts from corn suspension cultures. Neither sethoxydim (10 μM) nor haloxyfop (1 μM) inhibited fatty acid synthetase activity. Thus, both herbicides inhibited fatty acid synthesis in intact chloroplasts from corn seedlings with either acetate or pyruvate as a precursor, but did not inhibit incorporation of malonyl-CoA into fatty acids. This suggests that ACCase which catalyzes the formation of malonyl-CoA is the site of action of these herbicides.

EXAMPLE II

Selection and Characterization of Herbicide-tolerant Cell Lines

A selection protocol to identify and isolate herbicide-tolerant maize cells was developed to minimize the adverse effects of high herbicide concentrations on somatic embryo development and plant regeneration capacity. The procedure involved exposing tissue to gradually increasing concentrations of herbicide beginning with a sethoxydim concentration representing 1/20th of lethal dose and doubling the herbicide concentration at approximately two-week intervals until the lethal dose (10 μM sethoxydim) was reached. In this way, the herbicide

was allowed to take effect slowly with continuous selection pressure, thus permitting herbicide-tolerant cells to accumulate over time while not affecting the potential for plant regeneration.

A. **Selection of a Sethoxydim-Tolerant Cell Line**

Many selections were carried out utilizing the selection protocol described in the preceding paragraph. The selection of one such sethoxydim-tolerant cell line that was identified and characterized is described below in detail.

Approximately 100 grams of vigorously growing, regenerable, friable, embryogenic maize callus tissue established from an F_1 immature embryo resulting from the cross A188 x B73 were transferred to agar-solidified maintenance medium (Armstrong and Green, Planta, 164, 207 (1985)) in petri plates containing 0.5 μM sethoxydim (BASF) (Parsippany, New Jersey). This callus line was designated 2167-9/2160-154. Forty plates were prepared and five clumps of callus tissue weighing about 0.5 grams each were placed on each plate. The 0.5 μM sethoxydim concentration was chosen from growth inhibition studies to provide less than 10-20% growth inhibition during the first two weeks of herbicide exposure. After 14 days, 0.25-0.5 gram pieces of tissue showing vigorous growth rate and retention of embryogenic morphology (i.e., presence of somatic embryos) were subcultured on fresh medium containing 1.0 μM sethoxydim. Eighty plates containing five pieces of tissue per plate were prepared. For each subsequent transfer, all callus tissue showing growth and somatic embryo forming ability was placed on fresh media containing a two-fold increased sethoxydim concentration. Therefore, callus was transferred at two-week intervals to petri plates containing 0.5, 1.0, 2.0, 5.0 and 10.0 μM sethoxydim. During the course of the selection process, the total number of lines decreased as the herbicide-mediated growth inhibition became more intense. Cell lines exhibiting growth on 10 μM sethoxydim were designated as herbicide-tolerant and given an identification number. Two sethoxydim-tolerant lines were recovered that exhibited uninhibited growth at 10 μM sethoxydim. These lines were designated 2167-9/2160-154 S-1 and 2167-9/2160-154 S-2.

B. **Characterization of Herbicide-Tolerant Maize Cell Line 2167-9/2160-154 S-2**

Tolerant cell line 2167-9/2160-154 S-2 ("S-2") was characterized to evaluate: (1) the magnitude of sethoxydim tolerance; (2) cross-tolerance of haloxyfop; and (3) the biochemical basis for the tolerance. Callus tissue from S-2 that had been maintained on 10 μM sethoxydim was transferred to media containing up to 100 μM sethoxydim. One-half gram of S-2 tissue was plated on a 7 cm filter paper as a lawn overlaying 50 ml agar-solidified culture medium containing 0, 0.5, 1.0, 2.0, 5.0, 10.0, 50.0 and 100 μM sethoxydim, and cultured for two weeks. Control cell line 2167-9/2160-154 was plated similarly on medium containing the same levels of sethoxydim. The control cell line growth after two weeks was inhibited 50% at 1 μM sethoxydim. Growth of S-2 was not inhibited at 100 μM sethoxydim, indicating that S-2 was at least 100-fold more tolerant than the control callus line.

Growth of S-2 was inhibited with 0.65 μM haloxyfop, whereas the control cell line was inhibited 50% with 0.02 μM, indicating approximately a 30-fold increase in tolerance.

C. **Acetyl-Coenzyme A Carboxylase (ACCase) Activity of Maize Cell Line S-2**

Assays were conducted to determine if ACCase extracted from cell line S-2 was altered with respect to herbicide activity. ACCase activity of control tissue was 50% inhibited either by 1.5 μM sethoxydim, or by 0.25 μM haloxyfop. ACCase activity of S-2 tissue was inhibited 50% either by 70 μM sethoxydim, or by 1.8 μM haloxyfop, indicating at least 40-fold and 7-fold decreases in herbicide sensitivity on concentration basis, respectively.

EXAMPLE III
Plant Regeneration and Production of Herbicide-Tolerant Seed

A. **Plant Regeneration Protocol**

Sixteen ca. 150 mg clumps of S-2 callus were transferred per 25 x 100 mm petri plate containing agar-solidified N6 basal salts and 6% sucrose and incubated 7-14 days in low light (20 $\mu E\ m^{-2}\ s^{-1}$). Several plates containing callus on plant regeneration medium were prepared. Callus was transferred to agar-

solidified Murashige-Skoog (MS) medium without hormones and incubated in high intensity light (200 μE m-2 s^{-1}) for shoot elongation. Developing plants (1-3 cm long) were isolated from the callus surface and transferred to magenta boxes containing agar-solidified MS salts, 2% sucrose with no hormones for two weeks of further growth. When plants reached the 2-3 leaf stage, they were transplanted to peat pots containing potting soil, and were incubated in the growth room until growing stably. Surviving plants were transferred to soil in 4" diameter plastic pots and grown in the greenhouse.

B. Expression of Herbicide Tolerance in Plants Regenerated from S-2 Callus Tissue

Groups of eight control (2167-9/2160-154 unselected) and eight S-2 plants were sprayed with either 0.0, 0.01, 0.05, 0.11, 0.22 or 0.44 kg/ha sethoxydim to determine whole plant sethoxydim-tolerance of greenhouse-grown plants. Control plants were killed by 0.05 kg/ha or more sethoxydim. Plants regenerated from the S-2 cell line survived the 0.44 kg/ha sethoxydim treatment, indicating that S-2 plants exhibit at least 20-fold more tolerance of sethoxydim than control. Shoot height of regenerated S-2 plants was only slightly reduced 14 days after treatment with 0.44 kg/ha sethoxydim.

C. Seed Production from S-2 Plants

Plants surviving sethoxydim treatments of up to 0.44 kg/ha were transplanted to the genetics plot on the University of Minnesota campus, St. Paul, Minnesota. Additional S-2 plants were transplanted to the field that had not been sprayed. Sixty-five 2167-9/2160-154 control plants and ninety-five S-2 plants were grown to maturity in the field. Plants were either self-pollinated or cross-pollinated to inbred maize lines A188, A619, A641, A661, A665, B37, B73, R806, and W153R. Control seed were produced by selfing 2167-9/2160-154 regenerated plants, or by crossing them with the inbreds listed above.

D. Expression of Herbicide Tolerance in Progeny of Regenerated Plants

Seeds obtained by the crossing procedure described above were viable and germinated normally. Seeds from thirty S-2 selfed plants and fifteen 2167-9/2160-154 control plants were planted in 25 x 50 cm trays of soil (28 seeds from each plant in one tray) and grown in the greenhouse. Seedlings at the 3-4

leaf stage were treated with 0.1, 0.44, and 1.1 kg/ha sethoxydim and evaluated for visual herbicide damage and shoot height. Based on visual rating of herbicide damage two weeks after treatment, selfed progeny of S-2 plants segregated approximately 1:2:1 for healthy, uninjured plants: to plants showing partial injury: to dead plants, respectively, at 0.44 and 1.1 kg/ha sethoxydim treatments. All control progeny of 2167-9/2160-154 control plants were killed by 0.1 kg/ha and greater levels of sethoxydim. These results demonstrate partially dominant expression of sethoxydim tolerance indicating that sethoxydim tolerance in S-2 plants is a heritable trait. Similar tests were conducted on progeny of S-2 plants crossed to the other inbreds. In all cases, these test cross progeny treated with 0.44 kg/ha sethoxydim segregated 1:1 for growing shoots versus dead shoots whether S-2 plants were used as male or female parents. These results confirm that sethoxydim tolerance is controlled by a single partially dominant nuclear gene. In all cases, control plants crossed to the other inbreds were killed and therefore sethoxydim-sensitive.

E. Method for Obtaining Uniform Herbicide-Tolerant Seed

Progeny of S-2 plants surviving sethoxydim treatments of 0.44 and 1.1 kg/ha and showing no herbicide injury were transferred to the greenhouse and grown to maturity. These plants may be selfed and their progeny evaluated for sethoxydim and haloxyfop tolerance to identify pure breeding herbicide-tolerant maize lines.

Progeny of S-2 plants crossed to inbred lines and exhibiting sethoxydim tolerance may be recurrently backcrossed to the same inbreds. Progeny of each cross may be screened for sethoxydim-tolerance, and tolerant plants grown to maturity and again crossed to the recurrent parent. After six or seven cycles of backcrossing, sethoxydim-tolerant plants may be selfed and progeny screened for tolerance to produce homozygous sethoxydim tolerant maize inbreds.

EXAMPLE IV

Selection of Additional Herbicide-Tolerant Maize Cell Lines

One primarily sethoxydim-tolerant maize cell line, 2167-9/2160-154 S-1, and two haloxyfop-tolerant maize cell lines, 2167-9/2160-154 H-1 and 2167-9/2160-154 H-2, were selected and characterized as follows:

A. **Selection of Maize Cell Line 2167-9/2160-154 S-1**

Maize cell line 2167-9/2160-154 S-1 was selected from maize cell culture using the protocol described in detail above for the selection of Line 2167-9/2160-154 S-2. Approximately 70 plants were regenerated from Line 2167-9/2160-154 S-1, and either self-pollinated or cross-pollinated to the inbred maize lines A188, A619, A641, A661, A665, B37, B73, R806, and W153R.

B. **Selection of Maize Cell Line 2167-9/2160-154 H-1**

Line 2167-9/2160-154 H-1 was selected from maize cell culture using a similar protocol described in detail above except maize callus tissue was selected using the herbicide haloxyfop. Maize callus tissue was initially plated on 0.01 μM haloxyfop. At two-week intervals, surviving tissue was subcultured onto 0.05, 0.10 and 0.20 μM haloxyfop. Approximately 50 plants were regenerated from Line 2167-9/2160-154 H-1, and were self-pollinated.

C. **Selection of Maize Cell Line 2167-9/2160-154 H-2**

Line 2167-9/2160-154 H-2 was selected from maize cell culture using a similar protocol described in detail for line 2167-9/2160-154 H-1. No plants have been successfully regenerated from this line.

D. **Characterization of Lines 2167-9/2160-154 S-1, H-1 and H-2**

The tolerant callus cultures were characterized to determine the magnitude of sethoxydim and haloxyfop tolerance. Callus tissue from these lines was evaluated in experiments as described above in the characterization of line 2167-9/2160-154 S-2. Table II summarizes the results of these studies. Line 2167-9/2160-154 S-1 and Line 2167-9/2160-154 H-2 showed a four-fold increase in haloxyfop tolerance, while Line 2167-9/2160-154 H-1 exhibited approximately a 60-fold increase in haloxyfop tolerance. Neither haloxyfop selected line showed a significant degree of sethoxydim tolerance, while the sethoxydim selected line S-1 exhibited approximately a 100-fold increase in sethoxydim tolerance.

TABLE II

Herbicide Tolerance of Cell Lines S-1, H-1 and H-2

Cell Line	Herbicide Haloxyfop	Herbicide Sethoxydim
2167-9/2160-154 S-1	4[1]	100
2167-9/2160-154 H-1	61	0
2167-9/2160-154 H-2	4	0

[1]The numbers represent the fold increase in herbicide concentration that results in a 50% reduction in growth of the selected cell lines compared to the unselected control cell line 2167-9/2160-154.

E. **Herbicide Inhibition of Acetyl Coenzyme A Carboxylase of Maize Cell Lines S-1, H-1 and H-2**

Acetyl Coenzyme A Carboxylase (ACCase) was extracted from cell lines S-1, H-1 and H-2 and assayed as described in detail for maize cell line S-2. above. Table III below summarizes the results of these studies. The ACCase from line S-1 was more tolerant of both sethoxydim and haloxyfop, while the ACCase from line H-1 was more tolerant of haloxyfop, but not of sethoxydim. The ACCase from line H-2 showed no difference from the unselected parent line 2167-9/2160-154 in sensitivity to either herbicide.

However, cell line H-2 exhibited approximately a five-fold higher level of ACCase activity as compared to the unselected parent line 2167-9/2160-154. Thus, selection for sethoxydim or haloxyfop tolerance resulted in a less sensitive

ACCase in cell line S-1 and H-1, as well as a higher level of ACCase activity in cell line H-2.

TABLE III

Herbicide Inhibition of ACCase of Maize Cell Lines S-1, H-1 and H-2

Cell Line	Herbicide Haloxyfop	Herbicide Sethoxydim
2167-9/2160-154 S-1	3	4
2167-9/2160-154 H-1	7	0
2167-9/2160-154 H-2	0	0

[1]The numbers represent the fold increase in herbicide concentration that inhibits ACCase activity of the selected cell lines by 50% compared to the unselected parent cell line 2167-9/2160-154.

F. Plant Regeneration and Production of Seed

Cell lines derived by the procedures described above which exhibit herbicide tolerance are put through a plant regeneration protocol to obtain mature plants and seed expressing the resistance trait. The plant regeneration protocol allows the development of somatic embryos and the subsequent growth of roots and shoots.

Mature plants are then obtained from cell lines that are known to express the trait. If possible, the regenerated plants are self pollinated. In addition, pollen obtained from the regenerated plants is crossed to seed grown plants of agronomically important inbred lines. In some cases, pollen from plants of these inbred lines is used to pollinate regenerated plants. The trait is genetically characterized by evaluating the segregation of the trait in first and later generation progeny. The heritability and expression in plants of traits selected in tissue culture are of particular importance if the traits are to be commercially useful.

G. **Deposit of Seeds**

Seeds from representative S-2 plants (Ex. III (B)) and H-1 plants (Ex. IV(B)) have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 USA on October 25, 1988 and assigned accession numbers ATCC 40507, and ATCC 40508, respectively.

EXAMPLE V

Allelic Analysis of Herbicide Tolerant Maize Lines

To determine whether the mutations in herbicide tolerant maize plants derived from herbicide tolerant cell lines were allelic, plants derived from the resistant cell lines were crossed. The results, reported in Marshall et al. (Theor. Appl. Genet., 83, 435 (1992)), show that of the five mutations tested, all were allelic.

Materials and Methods

Foliar applications of the herbicides were made to greenhouse- or field-grown seedlings at the three- to five-leaf stage. Herbicides were applied in a spray volume of 187 l/ha that includes 2.3 l/ha crop oil concentrate to enhance uptake. Sethoxydim and haloxyfop were applied using commercial formulations (Poast from BASF Corp and Verdict from Dow Elanco, respectively). Visual observations were recorded 14-21 days after herbicide treatments.

For each mutant, two homozygous tolerant R_4 families were derived from two different regenerated plants and two heterozygous F_1 families were obtained from crosses of R_3 homozygous tolerant lines with the susceptible inbreds A188 and A619. Sethoxydim and haloxyfop each were applied at rates ranging from 0.0034 to 7.0 kg/ha. For each genotype a five-rate doubling series was used to determine herbicide lethality. Because of the large differences between wildtype and mutants in herbicide susceptibility, the ranges of herbicide application rates were varied. Two five plant/pot replicates of each family were evaluated for each rate. Herbicide injury symptoms were scored visually 21 days after herbicide treatment. The lethal rate was considered to be the lowest application rate resulting in death of all seedlings in the four replications of each

homozygous tolerant genotype or in the eight replications of the heterozygous combinations.

All F_1 combinations (no reciprocals) between S1, S2, S3, H1 and H2 homozygotes were made, and these F_1s were test crossed to susceptible inbred lines. The testcross progeny were grown in the greenhouse and treated with the herbicides. In Experiment 1, progeny from testcrosses with each of the ten F_1 combinations were treated with 0.05 kg/ha sethoxydim for combinations involving only S lines or with 0.009 kg/ha haloxyfop for combinations involving either of the H lines. In Experiment 2, a selected subset of the F_1 combinations was tested that used testcross families that were separately derived from those used in Experiment 1. Testcross progeny of S lines were treated with 0.03 kg/ha sethoxydim, and H1 and H2 progeny were treated with 0.009 kg/ha haloxyfop. In each experiment, known susceptible and heterozygous plants were included as controls.

For each mutant, homozygous tolerant and homozygous susceptible families derived from R_0 (heterozygous) plants were classified by evaluating plant responses to the herbicides. Enzyme extracts from these families were assayed for ACCase activity (% of no herbicide control) at herbicide concentrations near or greater than the susceptible wildtype I_{50} value. Activities in tolerant and susceptible extracts included in the same experiments were compared. ACCase activity was considered tolerant if when assayed in the presence of herbicide, it was at least 20% greater than the inhibited wildtype ACCase activity. For each of the mutant lines, tolerant lines tracing back to at least two different R_0 plants were evaluated.

Results

Homozygous seedlings of S2 and S3 exhibited only slightly bleached leaves and a slight reduction in plant height, but were not killed at 7 kg/ha sethoxydim, which is at least a 127-fold increase over the lethal rate for wildtype plants (Table IV). S2 and S3 lines also exhibited cross tolerance to haloxyfop. Homozygous S1 seedlings exhibited herbicide tolerance similar to that of S2 and S3. Heterozygotes derived from crosses of the S lines with the wildtype

exhibited a tolerance to both herbicides that was intermediate between wildtype and homozygous tolerant seedlings.

In contrast to the S lines, a second distinctive phenotype was represented by H1 homozygous seedlings that exhibited little or no sethoxydim tolerance but significant haloxyfop tolerance. Heterozygotes with the H1 lines exhibited haloxyfop tolerance that was similar to that of the homozygous H1 seedlings. H2 seedlings represented the third phenotype, which exhibited about 16-fold increases in tolerance to sethoxydim and haloxyfop compared to the wildtype. The H2 heterozygotes exhibited a tolerance to both herbicides that was intermediate between wildtype and homozygous tolerant seedlings.

Table IV

Genotype	Lethal rates (kg/ha)	
	Sethoxydim	Haloxyfop
Wildtype[a]	0.055	0.014
Homozygous tolerant[b]		
S2	> 7.0	0.22
S3	> 7.0	0.22
H1	0.055	0.11
H2	0.88	0.22
Heterozygous tolerant[c]		
Inbred x S2	≥ 1.8	0.11
Inbred x S3	≥ 1.8	0.11
Inbred x H1	0.055	0.11
Inbred x H2	0.22	0.11

[a] Average response of A188 and A619

[b] Average response of two families per mutant line

[c] Average response of A188- and A619-derived F_1 families

The two families of each mutant, derived from different R_0 plants and representing independent derivations from the parental A188 x B73 cross. had

similar responses to the herbicides. Herbicide lethality of heterozygotes of each mutant from the A188 cross was similar to that of the A619 cross. This suggested that background effects did not play a large role in modifying herbicide tolerance.

To determine allelic relationships among the mutations, F_1s made between the different homozygous mutants were crossed to herbicide-susceptible inbreds. The testcross progeny were treated with rates of herbicides intended to kill wildtype plants but allow heterozygous plants to survive. If the mutant alleles from the different mutant lines were at the same locus, then all testcross progeny would have had one or the other mutant allele contributed by the F_1 parent and would have been tolerant to the herbicide rates used. No susceptible plants were found in a total of 260 testcross progeny from crosses among S1, S2 and S3. Interpretation of these results was straightforward because all S line heterozygous plants included as controls in these tests survived the herbicide treatments as expected, and all inbred wildtype plants died when treated with sethoxydim. Additionally, no haloxyfop-susceptible plants were found in a total of 190 testcross progeny from H1 combinations with S1, S2 or S3, indicating that the H1 mutation is allelic to S1, S2 and S3 mutations.

Testcrosses involving H1 and H2 alleles were more difficult to interpret. The heterozygous controls for H1 and H2 were injured, and some were killed with the haloxyfop rate (0.009 kg/ha) that killed most, but not all, wildtype control plants. Four susceptible plants were found in a total of 310 testcross progeny from combinations of H2 and the S lines. This frequency was clearly different from the 77 susceptible progeny expected if the H2 allele was at an independent locus. It seems most likely that these deaths occurred due to variability in response of the H2 heterozygous plants, but the possibility that the haloxyfop-susceptible plants represented recombination between two closely linked loci cannot be ruled out.

The 5 haloxyfop-susceptible plants among 102 testcross progeny from the H1 and H2 combination also represented a clear difference from the 25 susceptible expected for independent loci, and again may be explained by a

variable response of H2 or H1 heterozygous plants. In the testcrosses involving H2, all of the plants that died were from crosses to the inbred A188. A188 was slightly more susceptible than other inbreds, which may have resulted in a slightly lower tolerance in heterozygotes derived from A188. The cause of rare deaths cannot be conclusively interpreted. However, the most likely interpretation is that the S1, S2, S3 and H1 mutations are allelic and that H2 either is allelic or linked (≤ 4.4 map units, if all 9 plants that died were due to recombination).

Summary

Herbicide tolerance for five maize (*Zea mays* L.) mutants regenerated from tissue cultures selected for tolerance to the ACCase-inhibiting herbicides, sethoxydim and haloxyfop was determined. Herbicide tolerance in each mutant was inherited as a partially dominant, nuclear mutation. Allelism tests indicated that the five mutations were allelic. Three distinguishable herbicide tolerance phenotypes were differentiated among the five mutants. Seedling tolerance to herbicide treatments cosegregated with reduced inhibition of seedling leaf ACCase activity by sethoxydim and haloxyfop demonstrating that alterations of ACCase conferred herbicide tolerance. Five putative new alleles of the maize ACCase structural gene (*Acc1*) were identified based on their differential response to sethoxydim and haloxyfop. The group represented by *Acc1-S1*, *Acc1-S2* and *Acc1-S3* alleles, which had similar phenotypes, exhibited tolerance to high rates of sethoxydim and haloxyfop. The *Acc1-H1* allele lacked sethoxydim tolerance but was tolerant to haloxyfop, whereas the *Acc1-H2* allele had intermediate tolerance to sethoxydim but was tolerant to haloxyfop. Differences in tolerance to the two herbicides among mutants homozygous for different *Acc1* alleles suggested that sites on ACCase that interact with the different herbicides do not completely overlap.

EXAMPLE VI

Identification and Cloning of Maize cDNA and Genomic Clones Encoding ACCase

Maize cDNA clones encoding a portion of the ACCase gene were identified by screening a DNA library generated from maize. The cDNA clones

were used to identify the sequence of the ACCase gene and to identify the genomic DNA fragments encoding the gene or genes for ACCase.

A λ gt11 cDNA library from maize inbred A188 seedlings was prepared by standard method for oligo-dT priming, as described for pea cDNA. (Gantt and Key, Eur. J. Biochem., 166:119-125 (1987). Plaque lifts of the maize cDNA library were screened with maize ACCase antiserum (Egli et al., Plant Physiol., 101, 499 (1993)) to identify plaques expressing ACCase-like proteins, as described by Sambrook et al., cited supra. (1989). The initial screen of 800,000 plaques yielded 120 positives. Rescreening and plaque purification reduced the number of positives to 14. All 14 clones bound ACCase antibodies that, when eluted from plaque lifts (J. Hammarback et al., J. Biol. Chem., 265:12763 (1990)), recognized a 227-kD biotinylated polypeptide on SDS-PAGE western blots of embryo and leaf crude extracts. The strongest western blot reaction was obtained with cDNA clone #15-14. The six best clones were digested with EcoRI to excise maize cDNA inserts. Total insert sizes ranged from 1.2 to 5.1 kb indicating the clones most likely did not contain the full coding sequences for the mature 219-kD and 227-kD ACCase polypeptides (minimum estimates of 6.1 and 6.3 kb, respectively).

Clone #15-14 contained three EcoRI fragments of 2.0, 1.2 and 0.23 kb Southern blots showed that the 1.2 and 2.0-kb fragments of clone #15-14 each hybridized to different fragments in the other five clones, with the exception of clone #4-4 which only contained a 1.2-kb fragment. The six maize cDNA clones contained EcoRI fragments that hybridized to a large transcript (ca. 8.3 kb) on Northern blots of total RNA from maize leaves, embryos and endosperm. BMS cell culture RNA also contained a 7.8 kb transcript. The relative abundance of the 7.8-kb transcript in embryos was higher than the other sources which is consistent with their ACCase activity.

The three EcoRI fragments were subcloned from cDNA clone #15-14 into BlueScript vector and sequenced by the dideoxy chain termination method (Sequenase 2.0 USB) initially using T3 and T7 primers and then oligonucleotide primers based on insert sequence. A clone #16-6 was also sequenced in a similar

manner. Clone #16-6 included three EcoRI fragments of 3.1 kb, 1.2 kb, and 0.23 kb and had additional sequence located upstream from that of clone #15-14. After comparing the sequence and determining that the sequence was the same, the additional 1.2 kb sequence at the 5' end was sequenced.

Clone #18-5 was sequenced in a similar manner. Clone #18-5 included 3.9 kb, 1.2 kb, and 0.23 kb EcoRI fragments and contains an additional 1.9 kb 5' sequence upstream from clone #15-14. Subclone #18-5I (3.9 kb EcoRI fragment) has been deposited with the American Type Culture Collection and given Accession No. 69236.

GenBank, PIR-29, and Swiss-Prot 19 data banks have been searched for amino acid homology with the corresponding amino acid sequences of the three subclones of clone #18-5. Peptide sequences corresponding to the maize cDNA subclones had higher similarity to chicken, rat, yeast and other plant and animal ACCases than to any other peptide sequence in the data banks. A comparison of the relative organization of the 3.9, 1.2 and 0.23-kb EcoRI fragments of clone #18-5, their co-linearity and extent of amino acid identity with chicken ACCase cDNA sequence shows that the maize clone #18-5 has a large region near the 3' end with high amino acid identity (40 to 61%) to chicken ACCase, a longer region with 23% identity in the middle of the 3.9-kb sequence, and a short region with 52% identify near the 5' of the 3.9 kb sequence.

Portions of the sequence of the #18-5I subclone have been identified as encoding domains of the ACCase enzyme of functional significance. Those functional regions include a fragment that spans the presumed transcarboxylase

active site in the enzyme having the following presumed sequence SEQ ID NO: 1:

1112-856

GTT CCT GCA AAC ATT GGT GGA CCT CTT CCT ATT ACC AAA CCT CTG GAC
CCT CCA GAC AGA CCT GTT GCT TAC ATC CCT GAG AAC ACA TGC GAT CCA
CGT GCA GCT ATC TGT GGT GTA GAT GAC AGC CAA GGG AAA TGG TTG GGT
GGT ATG TTT GAC AAA GAC AGC TTT GTG GAG ACA TTT GAA GGA TGG GCA
AAA ACA GTG GTT ACT GGC AGA GCA AAG CTT GGA GGA ATT CCT GTG GGC
GTC ATA GCT GTG GAG ACA

This functional domain is contained in the sequence 1112 to 856 base pair from the 3' stop codon or carboxy terminus region of the ACCase coding sequence of maize. This transcarboxylase active sequence is also present in clone #15-14.

Another functional region that has been identified spans the 12 base pair sequence encoding the biotin binding site having the following peptide sequence SEQ ID NO: 2:

5' GTT ATG AAG ATG 3'
Val Met Lys Met

The biotin binding site is encoded approximately 30% in from the 5' (N-terminus) end of rat, chicken and yeast ACCase genes. These functional domains are useful in mapping and further identifying other cDNA and/or genomic fragments encoding ACCase genes.

The cDNA clones encoding portions of the acetyl CoA carboxylase genes are useful to identify the sequence of the gene or genes and are useful as probes to locate the genomic copies of the gene or genes. Because the ACCase antibodies used to screen the λ gt11 library recognize both the 219 and 227 kD ACCase polypeptides, it has not been determined which polypeptide is encoded by these less than full length clones. It is likely that the majority of the clones encode the 227 kD polypeptide since that polypeptide is more abundant in the leaf tissue source of the DNA library and the antibodies have a higher affinity for the 227 kD ACCase polypeptide.

The maize genome was analyzed to identify copy number and location of the genomic copies of ACCase gene or genes. Four distinct types of maize ACCase genomic clones have been identified, termed A1, A2, B1 and B2 (see below).

To obtain genomic copies of ACCase genes, a maize B73 genomic library (Clontech, Palo Alto, CA) was screened with the 2 kb subclone from #15-14 and several clones of about 15 kb were identified as having homology to the ACCase cDNA. Restriction mapping and partial sequence analysis revealed two types of genomic clones (Type A and Type B) that differed in restriction sites and in their position relative to the ACCase partial cDNA sequence.

The 2.5 kb EcoRI-SalI fragment (#16) from the Type A genomic clone and the 3.0 kb EcoRI-EcoRI fragment (#34) from the Type B genomic clone were shown to hybridize to the 3.9 kb probe from #18-5 and were subcloned into the Bluescript vector and sequenced. Approximately 1.5 kb of DNA sequence from the genomic Type A 2.5 kb fragment were 100% identical to coding sequence from the 3.9 kb cDNA subclone #18-5I described hereinabove; the remaining sequence exhibited no identity with the cDNA clone and presumably represents a noncoding intron sequence. A 350 nucleotide sequence derived from the genomic Type B 3.0 kb fragment was about 95% identical to the cDNA clone indicating that its coding sequence differs from that of genomic Type A. These results also indicate that the maize genome encodes at least two different genes encoding a polypeptide having acetyl CoA carboxylase activity.

To identify and clone the remainder of the gene representing the amino-terminus of maize ACCase, additional regions from the Type A genomic clone have been subcloned and sequenced. To synthesize the remaining coding region between the end of the cDNA clone #18-5 and the start of transcription, two oligonucleotide primers were synthesized. Primer 1 is complementary to the DNA sequence: (SEQ ID NO:3)

5' GCCAGATTCC ACCAAAGCAT ATATCC 3'

near the 5' end of cDNA subclone #18-5I and was be used as a primer for synthesis of cDNA molecules from maize seedling, leaf or embryo RNA. Several independent clones were sequenced and their sequences compared to the known sequence of the Type A genomic clone to determine the exact coding sequence corresponding to that maize gene for ACCase. A similar strategy can be used to obtain the complete coding sequence for genomic Type B ACCase.

The remaining cDNA sequence was obtained by three successive rounds of RT-PCR using oligonucleotide primers based on genomic apparent exon (5') and known cDNA (3') sequences. The primers used to amplify nucleotides 1-240 of the cDNA were 28sst-a5+ (SEQ ID NO:4) and 28sst-6at3+ (SEQ ID NO:5), nucleotides 217-610 of the cDNA were 28sst-5+ (SEQ ID NO:6) and 28-2t3+ (SEQ ID NO:7), and nucleotides 537-2094 of the cDNA were ACCPCR5' (SEQ ID NO:8) and I55- (SEQ ID NO:3) (Table V). PCR products corresponding to nucleotides 1-240, 217-610, and 537-2094 of the final sequence were cloned into PCR-script (Stratagene).

TABLE V

cDNA Position	5′ primer designation	5′ primer sequence	3′ primer designation	3′ primer sequence
nt-1-240	28sst-a5+	GGTCTTCAATTGTGCTGTCTGG (SEQ ID NO:4)	28sst-6at3+	CCTTGACGAACAGACTGGCTGTGC (SEQ ID NO:5)
nt 217-610	28sst-5+	CACAGCCAGTCTGTTCGTCAAGG (SEQ ID NO:6)	28-2t3+	CCTCTACGTAATTGGTCAGC (SEQ ID NO:7)
nt 537-2094	ACCPCR5′	CATAGCTATGGCAACTCCGG (SEQ ID NO:8)	I55-	GGATATATGCTTTGGTGGAATCTGGC (SEQ ID NO:3)

The original 5.4-kb cDNA clone #18-5 and PCR products from at least three individual PCR per oligonucleotide pair were sequenced in both directions by the dideoxy chain-termination method, using either Sequenase II (U.S. Biochemicals) or ABI 373 (Applied Biosystems, Inc.) protocols. No sequence differences were found in regions of clone overlaps. The complete sequence of the cDNA of maize ACCase (nucleotides 1-7470; SEQ ID NO:9) and its corresponding amino acid sequence (amino acids 1-2325; SEQ ID NO:10) are shown in Figures 3 and 4. The 7470 bp cDNA includes a 459 nucleotide 3' untranslated region and 36 nucleotides of 5' untranslated sequences.

The first Met codon in the cDNA (nucleotides 37-39) was identified as the start codon based on its similarity to consensus initiation sequences (Kozak, J. Cell. Biol., 108, 229 (1989); Lutcke et al., Embo. J., 6, 43 (1987)). An in-frame stop was found in the genomic sequence 6 nucleotides upstream of the sequenced cDNA, and RT-PCR analysis of this region suggested that the in frame stop codon was also present in the cDNA. The 3' end of the coding sequence was defined by two stop codons present in the large open reading frame after nucleotide 7011. The translated coding sequence predicted a polypeptide of 2325 amino acids (257 kD; SEQ ID NO:10) which was 79 to 81% identical to the multifunctional (MF) ACCases from alfalfa (Shorrosh et al., Proc. Nat'l. Acad. Sci., 91, 4323 (1994)) and wheat (Gornicki et al., Proc. Nat'l Acad. Sci., 91, 6860 (1994)), and to a 118-amino acid predicted polypeptide of a rice expressed sequence tag (Genbank accession # D39099, T. Sasaki), but only 53 to 55% identical to ACCase from other eukaryotes.

In a pileup alignment of plant ACCases (Genetics Computer Group, Madison, Wisconsin), Met 1 of both maize and *Brassica napus* ACCases was located about 130 amino acids upstream of the conserved sequence VDEFCKALGG, compared to only 25 amino acids upstream for other plant ACCases. The predicted 2325 amino acids of maize ACCase contains a biotinylation site at position 806, within the conserved VMKM motif (Ton et al., Eur. J. Biochem., 215, 687 (1993)). The arrangement and amino acid sequence of binding sites (Shorrosh et al., Proc. Nat'l. Acad. Sci., 91, 4323 (1994)) for

ATP (amino acids 318-333), biotin (amino acids 799-811; biotin at 806), acetyl-CoA (amino acids 1952-1961), and carboxybiotin (amino acids 1662-1711) were highly conserved among all MF ACCases.

EXAMPLE VII

Characterization of Maize ACCase Genomic Clones

The initial restriction fragment length polymorphism (RFLP) analysis of EcoRI-digested total DNA from three maize inbred lines showed one band when probed with the 2 kb subclone from #15-14 (internal to gene) and two bands when probed with the 1.2 kb subclone (near the 3' end of the gene). Fragments homologous to the 2 kb probe were monomorphic and the more intense of the two bands hybridizing with the 1.2 kb probe was dimorphic. As discussed above, these results support the view that maize contains at least two distinguishable ACCase genes and that they may be quite similar for much of the coding region. Additional genomic Southern blots of a set of recombinant inbred lines were used to map polymorphisms for the ACCase probes to maize chromosomes. One polymorphism was mapped to the short arm of chromosome 2; other polymorphisms were not evident in these initial tests to identify a chromosomal location for other maize ACCase genes.

The isolation and restriction mapping of additional genomic clones from a B73 genomic library (Clontech) resulted in the identification of four different types of clones termed A1, A2, B1 and B2 which had 96% nucleotide sequence identity. Types A and B correspond to previously published pA3 and pA4 cDNAs (Ashton et al., Plant Mol. Biol., 24, 35 (1994)) and differ from pA3 and pA4 by ~ 4% in their coding sequences.

Type A and B genomic clones have linear sequence homology except for an insertion in an intron of the Type B genes about 1400 bp 3′ of the A1(SEQ ID NO:9) translation start site. Analysis of the insert boundaries revealed a 3-bp target site duplication and a 6-bp direct repeat, and further sequence analysis showed the presence of two new and unique LINE elements (Long Interspersed Nuclear Elements) in B1 and B2. Mammalian LINE elements are highly abundant (10^4 to 10^5 copies), 6 to 7 kb long, and have frequent 5′-end deletions

and an A-rich 3′ terminus. They are flanked by short direct repeats, and contain two ORFs, one encoding a reverse transcriptase. Three LINE elements (Cin4, 50-100 copies in maize; del2, 250,000 copies in lily; BNR1, 2-5% of genome in sugarbeet) have been described in plants (Leeton et al., Mol. Gen. Geneti., 237, 97 (1993); Schmidt et al., Chromo. Res., 3, 335 (1995); Schwarz-Sommer et al., EMBO J., 6, 3873 (1987)). Maize ACCase B1 has one unique LINE element and B2 has two. The two B2 LINE elements were characterized by differences in their reverse transcriptase sequence. The B genomic clone inserts have characteristic LINE features including cysteine motifs and a possible polyA tail, and high abundance. The LINE insert also has been found in an intron of the maize *Shrunken-2* gene (Hannah et al., Plant Physiol., 98, 1214 (1992)).

The nucleotide sequence (3544 nucleotides) of a Type A1 ACCase genomic clone is shown in Figure 5 (SEQ ID NO:11).

The nucleotide sequence of another Type A clone is shown in Figure 6 (SEQ ID NO:12). The sequence is all 5′ untranslated sequence and contains two non-identical 7 nucleotide inserts, a CTP (i.e., it can be amplified with primers 28sst-97F and 28sst-6t3+) and at least the first 1/3 of an ACCase coding sequence. The CTP of SEQ ID NO:12 is identical to the CTP of A1 clones.

The partial nucleotide sequence of five Type A2 clones is shown in Figure 7 (SEQ ID NOs 13, 14, 15, 16 and 17, respectively).

A limited sequence comparison of SEQ ID NO:12 and A2 clones showed that SEQ ID NO:12 and A2 clones were more closely related than SEQ ID NO:12 and A1 clones.

The partial nucleotide sequence (231, 207 and 180 nucleotides) of three Type B clones is shown in Figure 8 (SEQ ID NOs 18, 19 and 20, respectively).

The cDNAs corresponding to genomic clones A2, B1 and B2 are cloned and sequenced in a manner similar to that described above. The derived amino acid sequences are aligned with known ACCase sequences. If putative CTP sequences are identified, functionality is tested as described below. Also if the tissue specificity and developmental timing of expression differ for different ACCase genes, the sequences of the promoter regions of the corresponding

genomic clones are compared. Gene-specific probes for specific ACCase genes can provide more information on their roles in lipid synthesis (plastid and cytoplasmic isoforms), secondary metabolism (cytoplasmic isoforms), and herbicide resistance (likely plastid isoforms).

Only one plastidic ACCase polypeptide was identified by SDS-PAGE of maize leaf extracts, although 2-D gel analyses might provide evidence for a second, highly similar isoform. Of the two ACCase isoforms, only ACCase I shows altered herbicide inhibition in *Acc1-S2* mutants, and most of the ACCase activity in leaves and developing embryos is herbicide-resistant and thus attributed to the *Acc1-S2* gene product.

Although a 3′ ACCase probe has been mapped both to chromosome 10 near *Acc2-S5* and to chromosome 2, the conserved sequence of ACCase genes and lack of polymorphism in multiple bands complicates identification of genes encoded at these loci. The Type A1 ACCase gene is probably located on chromosome 2, since (I) 5′ untranslated and chloroplast transit peptide probes from Type A1 hybridize to two bands (dark and light) in maize inbreds, and (ii) analysis of maize-oat addition lines carrying maize chromosomes 2 through 9 indicates the dark band is on chromosome 2 and the light band is on chromosome 1 or 10.

Type B ACCase genes are likely to encode cytosolic isoforms. Given that cytosolic malonyl-CoA is a precursor in the synthesis of many secondary metabolites including flavonoids (e.g. maysin, a corn silk component associated with corn earworm resistance), these cytosolic ACCases can have agronomic utility.

Northern blot analysis of total maize RNA with an ACCase probe (nucleotides 3400-5932) showed a single 8.3 kilobase band. To determine whether the expression of ACCase RNAs was developmentally regulated, blots of total RNA from 16 to 42 DAP (days after pollination) embryos were probed with an ACCase cDNA fragment. Transcript abundance peaked about 23 DAP and the steady state pattern was similar to *in vitro* ACCase enzyme activities and protein measured from developing embryos. Type A- and B- specific ^{32}P-CTP-

labeled antisense transcripts were 780 nt long (662 nt of ACCase sequence + 118 nt of vector/promoter sequence) and were identical except for 15 base mismatches scattered along their length. Each antisense transcript was hybridized to total RNA from embryos at 16, 20, 23, and 42 DAP and digested with RNAse A/TI mixture to yield a 662-base fragment specific to the probe used. The results showed that the Type A transcript was more abundant than Type B at all tested stages, and that only Type A remained high in older embryos. Types A and B had similar expression patterns and peaked around 20-23 DAP. The ratio of Type A:B mRNA in leaves was about 2:1, similar to its relative abundance in cDNA expression libraries.

EXAMPLE VIII

Identification of a Second Herbicide Resistance Locus on Chromosome 10

To determine if sethoxydim and haloxyfop mutations segregated, all possible pairwise combinations of crosses were made between the sethoxydim (designated S1, S2, S3, S4, S5) or haloxyfop (designated H1, H2) tolerant mutant lines. The resulting F1 plants were test-crossed to wild-type susceptible plants. The testcross progeny were grown in the greenhouse and sprayed with Poast at a rate of 0.4 pounds/acre. If the two parents of the original cross had tolerance mutations in the same gene, then all testcross plants were expected to be heterozygous for a tolerance gene and would not segregate for wild-type susceptible plants. If the two parents had tolerance mutations in two different genes, then the testcross plants were expected to segregate for wild-type susceptible plants. Table VI below summarizes the results from these testcross progeny.

Table VI

	PARENT 1					
PARENT 2	S2	S3	S4	S5	H1	H2
S1	Res	Res	Res	Seg	Res	Res
S2		Res	Res	Seg	Res	Res
S3			Res	Seg	Res	Res
S4				Seg	Res	Res
S5						Seg
H1						Res

RES = all resistant testcross progeny

SEG = segregation for wildtype susceptible progeny at an approximate frequency of 25%

Subsequent to these analyses, another *Acc* resistance allele, S6, was identified. S6 is allelic to S5. S5 is a mutant *Acc* allele, which maps to chromosome 10, identified by Van Dee (an allele referred to as "M" in Van Dee, M.S. Thesis, University of Minnesota, 1994)).

These testcross progeny data clearly showed that the mutant line designated S5 was not allelic to any of the other sethoxydim-tolerant mutants. The tolerance gene, *Acc2*, was mapped to chromosome 10, flanked by umc155 and umc146 by a distance of 5.9 and 4.1 centimorgans (cM), respectively, by analyzing standard RFLP markers in an F2 population segregating for tolerance and susceptibility (VanDee, M.S. Thesis, University of Minnesota (1994)). The *Acc2* locus is about 10 map units from *gl*, which is a "golden plant" phenotypic marker on the long arm of chromosome 10. The *Acc2* mutant is designated *Acc2-S5*.

The testcross data also showed that all the mutant lines other than S5 have mutations in the same gene (*Acc1*) because their testcross progeny did not segregate. Thus, tissue culture selection for sethoxydim or haloxyfop resistance

resulted in 7 independently isolated mutations representing 2 different, unlinked genes.

As described above, genomic clones representing at least 4 different ACCase genes have been isolated from a genomic library of the inbred line B73. The complete coding sequence for one gene (A1) was determined and, where determined, the coding sequences for the other clones were highly identical to that of A1. A1 has a functional chloroplast transit peptide sequence as expected for an ACCase that is localized in plastids. Another genomic clone (designated 5A) from the 5′ end of the gene also contained a transit peptide sequence and other 5′ sequences that differed slightly from A1. These results indicated that corn has at least 2 genes for plastidic ACCase. The 5′ sequence differences between A1 and 5A (Type A2) genomic clones are useful in designing PCR primers that would be specific for either A1 or 5A (Type A2) genes. Restriction site differences in the upstream 5′ regions for Type A1 and A2 genes may also be useful for RFLP mapping with gene-specific ACCase probes.

A 3′ Type A1 ACCase cDNA probe mapped to chromosome 2 (Egli et al., Maize Genetics Newsletter, 68, 92 (1994)) and to chromosome 10 (Caffrey et al., Maize Gen. Coop., 69, 3 (1995)). Two 5′ Type A1 cDNA probes which span the transit peptide mapped to chromosome 2 in the same location as the 3′ probe (see maize genetic map, 1996 version, Maize Genomic Database). Similar results were observed when the 5' or 3' probe was used to analyze total genomic DNA from some inbred lines relative to oat-corn additional lines. Both probes hybridized to two different fragments in total maize DNA but to only one fragment in oat-corn lines containing maize chromosome 2 and chromosome 3, 4, 5, 6, 7, 8 or 9 (Rines et al., In: Modification of Gene Expression and NonMendelian Inheritance, pp. 235-251 (1995)). Thus, A1 and A2, described hereinabove, appear to encode plastic ACCases that correspond to the *Acc1* and *Acc2* loci, respectively.

To identify segregating *Acc1-S3* tolerance genes, a pair of PCR primers (28sst-97F,CCTTTTTATGGCACTGTGCG, SEQ ID NO:21) and 28sst-6t3+, CATCGTAGCCTATATGAGGACG, SEQ ID NO:22) were identified that

amplify the 5′ end of the A1 gene sequence from B73, which segregates with the resistance trait, but not from A188 or A641. These PCR primers are located in non-coding regions of A1 that span the chloroplast transit peptide. A control reaction employed a nearby 5′ primer (28sst-a5+, SEQ ID NO:4) and 28sst-6t3+. The amplification reactions comprised 1 µl sense primer (from a 2.5 µM 10X stock solution), 1 µl anti-sense primer (from a 2.5 µM stock solution), 0.5-1.0 µl genomic DNA (1 µg/µl), 2.5 µl 10X *Taq* buffer (Promega), 2.5 µl $MgCl_2$ (Promega, stock solution is 25 mM), 0.5 µl of a dNTP stock solution (Promega, 10 mM stock solution), 0.25 µl *Taq* polymerase, and 24.75 µl water. DMSO at 5% can be added to the amplification reaction to improve specificity.

The following cycling parameters were used : 94°C for 2 minutes for 1 cycle; 94°C for 1 minute for 35 cycles; 54° for 1 minute and 74°C for 70 seconds, and then 4°C. Amplified products were analyzed by gel electrophoresis.

A nearby 5′ primer (28sst-a5+, SEQ ID NO:4), when employed with 28sst-6t3+, amplified all genotypes and functioned as a positive control. In an F2 family segregating for the S3 source of sethoxydim tolerance, amplification of the A1 PCR product (528 bp which includes nucleotides corresponding to nucleotides 1-238 of SEQ ID NO:11, 29 bp of noncoding 5' sequence and 261 bp of 3' intronic sequence) was always associated with sethoxydim tolerance (29 tolerant plants were PCR+). The B73 PCR product was not detected in DNA from 15 of 17 susceptible plants (died after spraying), but was in 2 susceptible plants. If these plants died only from herbicide exposure and not from an unrelated cause, then the results indicate that the A1 PCR marker maps approximately 12 cM from the resistance locus. These results showed that the B73 chromosome was the donor of the *Acc1-S3* tolerance gene in the selected tissue cultures and suggests that the *Acc1* gene is at least closely linked to the A1 genomic clone.

RFLP mapping was also employed to determine whether the *Acc1* gene cosegregates with the herbicide tolerance locus on the short arm of chromosome 2. A probe, a 1.2 kb 3' fragment of *Acc1* (22J), was labeled and used to probe

EcoRI digested genomic DNA isolated from eighty progeny of a cross between S3 x A641. DNA from all herbicide resistant progeny had the B73 *Acc1* polymorphism while DNA from all herbicide sensitive progeny did not show the polymorphism. These results provide further support that the *Acc1* gene is, or is very closely linked to, the locus conferring herbicide tolerance. Moreover, a monomorphic hybridizing fragment was observed in this mapping population which may represent the *Acc2* locus on the long arm of chromosome 10.

EXAMPLE IX

Methods to Prepare a Two-Gene Heterozygous Herbicide Tolerant Hybrid

A preferred embodiment of the invention is an herbicide resistant inbred double homozygous maize plant (e.g., Inbred A (*Acc1-S3/Acc1-S3;Acc2-S5/Acc2-S5*)) that can be crossed to any normal susceptible inbred line. The resulting F1 contains one dominant allele for herbicide tolerance from each tolerance gene. The F1, which is grown by farmers, has herbicide tolerance equivalent to that given by single homozygous hybrids.

One method to prepare a two-gene heterozygous herbicide tolerant hybrid is to incorporate both tolerance genes at the end of inbred parent development. For example, cross Inbred A separately to S2 and to S5 homozygous mutant lines as shown:

Inbred A x S2	Inbred A x S5
+/+; +/+ x *Acc1-S2/Acc1-S2*	*+/+; +/+* x *Acc2-S5/Acc2-S5*

This results in two genotypes each heterozygous for a different herbicide tolerance allele.

Acc1-S2/+; +/+	*+/+; Acc2-S5/+*

These genotypes are then backcrossed to Inbred A for the desired number of generations to recover the Inbred A parent. The susceptible plants segregating in each backcross generation can be eliminated by spraying with Poast. Segregating backcross progeny include:

Inbred A x *Acc1-S2/+; +/+*	Inbred A x *+/+; Acc2-S5/+*
+/+; +/+ x *Acc1-S2/+; +/+*	*+/+; +/+* x *+/+; Acc2-S5/+*

Progeny

+/+; +/+ (discard) +/+; +/+ (discard)

Acc1-S2/+; +/+ (repeat cross) +/+; *Acc2-S5*/+ (repeat cross)

At end of backcrossing process, self pollinate to recover Inbred A with the homozygous mutant gene.

Inbred A (*Acc1-S2/Acc1-S2*; +/+) Inbred A (+/+; *Acc2-S5/Acc2-S5*)

The two mutant versions of Inbred A are then crossed to produce double heterozygous mutant version of Inbred A:

Inbred A (*Acc1-S2*/+; *Acc2-S5*/+)

Self pollinate to produce F2 generation segregating for the following genotypes:

1/16 *Acc1-S2/Acc1-S2; Acc2-S5/Acc2-S5* identify homozygote

2/16 *Acc1-S2/Acc1-S2; Acc2-S5*/+

1/16 *Acc1-S2/Acc1-S2*; +/+; +/+

2/16 *Acc1-S2*/+; *Acc2-S5/Acc2-S5*

4/16 *Acc1-S2*/+; *Acc2-S5*/+

2/16 *Acc1-S2*/+; +/+; +/+

1/16 +/+; +/+; *Acc2-S5/Acc2-S5*

2/16 +/+; +/+; *Acc2-S5*/+

1/16 +/+; +/+; +/+; +/+

One sixteenth of the F2 plants are double homozygotes in Inbred A background. These F2 plants can be identified by:

a) Crossing F2 plants to susceptible plants and testing the progeny for sethoxydim tolerance. Plants homozygous for both genes produce nonsegregating double heterozygous testcross progeny that do not exhibit herbicide damage symptoms typical of single heterozygous mutant testcross genotypes.

b) Using RFLP molecular markers to identify chromosome regions that flank the *Acc1* and *Acc2* loci. Preferred RFLP markers flank and are tightly linked to, the ACCase coding sequences of *Acc1* or *Acc2*. For instance, RFLP markers umc131 and umc2, or umc131 and uox, for the *Acc1* region on chromosome 2 and umc155 and umc146, or umc 155 and ncsu2, for

the *Acc2* region on chromosome 10 can be used to identify F2 plants homozygous for both sets of flanking polymorphisms from the tolerant mutant donor parents. It is also envisioned that other more closely, or distantly, linked flanking RFLP markers or genotype-specific ACCase probes can be used to identify the desired progeny.

c) Using PCR amplification which employs genotype-specific primers to detect the presence of a particular allele of either *Acc1* or *Acc2*. For example, a primer that hybridizes to a region in the 5' untranslated portion of *Acc1* and another primer that hybridizes to a region in the first intron of *Acc1* are useful in an amplification reaction to detect the presence or absence of the fragment in progeny plants.

Another method to prepare a two-gene heterozygous herbicide tolerant hybrid is to incorporate both tolerance genes at the start of inbred parent development. For example, the two sources of tolerance (i.e., S2 and S5) are crossed to produce a double heterozygous mutant F1 genotype. This genotype is the equivalent of the Inbred A genotype indicated above (*Acc1-S2/+;Accs-S5/+*) except that the tolerance genes are not yet in the Inbred A background. The double heterozygous mutant F1 genotype is crossed to Inbred A and the resulting progeny plants analyzed for presence of the RFLP flanking markers for both *Acc1* and *Acc2* tolerance genes. Plants containing both sets of markers are used for the next cycle of backcrossing to Inbred A. The backcrossing and RFLP analysis are repeated for appropriate number of cycles to recover the Inbred A genotype. The final steps in developing the double homozygous mutant Inbred A genotype are the same as the final two steps in the incorporation of both tolerance genes at the end of inbred parent development (see above). The crossing of the inbred plant to a wild-type (susceptible) plant then results in a plant which is a hybrid double heterozygote.

A third method to prepare a two-gene heterozygous herbicide tolerant hybrid is to incorporate one tolerance gene into one inbred line (e.g., (Inbred A (*Acc1-S3/Acc1-S3*;+/+) and incorporate another tolerance gene into another

inbred line (e.g., Inbred B (+/+;*Acc2-S5/Acc2-S5*). These two inbred lines are then crossed to produce a two-gene heterozygous herbicide tolerant hybrid.

EXAMPLE X

Field Tests of Herbicide Tolerant Double Heterozygotes

A double heterozygous hybrid was produced by crossing the following parent lines each homozygous for one of the sethoxydim tolerance mutations: *Acc1-S2/Acc1-S2*; +/+ x +/+; *Acc2-S5/Acc2-S5*, where + represents the normal alleles for sethoxydim susceptibility. The resultant F1 (*Acc1-S2/+; Acc2-S5/+*) contained one copy of each sethoxydim-tolerant allele. Similar F1's were produced between S3 and S5 mutant lines.

Several double heterozygous F1 hybrids were tested for tolerance to sethoxydim (Poast) applications in the field in 1996. Poast was applied at the 6 to 8 leaf stage at a rate of 0.8 pounds/acre. Six days after application, visual ratings of the double heterozygous F1 hybrids unexpectedly indicated either no herbicide injury symptoms or slight symptoms on a few plants (< 10%) as evidenced by transitory yellowing of leaves emerging from the whorl of the main stalk or from side tillers. These symptoms disappeared thirteen days after herbicide application. Similar results were obtained with the single mutant homozygous parent lines S2, S3 and S5. In contrast, single heterozygous F1 hybrids, obtained by crossing S2, S3 and S5 homozygous parent lines with normal susceptible lines, exhibited herbicide injury symptoms in most of the plants. These symptoms, especially yellowing of the tillers, persisted through flowering. Plants of normal susceptible lines were killed by this application.

Twenty-three days after the first application of herbicide, when tassels were emerging from most plants, another application of herbicide was performed (1.6 pounds Poast/acre). The double heterozygous F1 plants and the single homozygous parent lines showed little visible herbicide injury after the second treatment of the same plants. These herbicide applications showed that the herbicide tolerance of the double heterozygous F1 hybrids was equivalent to that of the single mutant homozygous parent lines. The response to herbicide treatment of hybrids with one mutant allele from each of the chromosome 2 and

the chromosome 10 genes (i.e., *Acc1-S2/+; Acc2-S5/+*) was indistinguishable from the response of plants containing two mutant alleles from either gene (i.e., *Acc1-S2/Acc1-S2; +/+* or *+/+; Acc2-S5/Acc2-S5*).

EXAMPLE XI

Greenhouse Trials of Herbicide Tolerant Double Heterozygotes

Two greenhouse trials were conducted to assess relative heterozygous sethoxydim resistant corn safety to 1/8X, 1/4X, 1/2X, 1X, 2X, and 4X use rates of POAST PLUS® in comparison to homozygous sethoxydim resistant corn and to sethoxydim susceptible corn (commercial Pioneer hybrid). Tolerance to sethoxydim was compared between double heterozygotes that contained the chromosome 2 resistance gene plus the chromosome 10 resistance gene, e.g., (*Acc1-S3/+; Acc2-S5/+*) to heterozygotes with only the chromosome 2 gene, e.g., (*Acc1-S3/+; +/+*) or the chromosome 10 gene (*+/+; Acc2-S5/+*).

The following plant genotypes and growth stages were employed in the greenhouse trials: 1) commercial homozygous sethoxydim resistant corn hybrid (*Zea mays*, var. DK 592 $_{SR}$), growth stage-14 (4 leaves expanded); 2) commercial non-sethoxydim resistant corn hybrid (*Zea mays*, var. PI 3140), growth stage-14, 3) heterozygous sethoxydim resistant corn (*Zea mays*), growth stage-14. Sethoxydim (POAST PLUS, 120 g/l) was applied at 0.02 (1/8X), 0.04 1/4X), 0.09 (1/2X), 0.18 (1X), 0.36 (2X), and 0.71 (4X) lb aI/A (active ingredient per acre). The spray volume was 20 GPA (gallons per acre) and an adjuvant, ATPLUS® 411 F, was added at 1.25% v/v COC (crop oil concentrate).

Metro 360 soiless potting media was used. Osmocote 14-14-14 was applied at planting. STEM (Soluble Trace Element Mixture) and Peters 20-20-20 was applied within 3 days of germination and again 1 day after treatment at normal greenhouse use rates. Plants were watered overtop prior to application then subirrigated as needed. Plants were not exposed to UV light after application to maximize sethoxydim activity and potential injury. Corn plants were evaluated for percent injury at 7-8 and 14-15 days after treatment (DAT).

Results

GH Trial 067/96: Germination rates for the heterozygous sethoxydim resistant corn hybrids varied from 3 to 94 percent. Heterozygotes with the following line designations had germination rates of 80% or better and were included in GH trial 067/96: 4163/4142-2, 4142-2/4163, 4142-8/4158, 4158/4142-20, 4143-3/4160, 4143-1/4160. Heterozygotes with the following line designations did not have sufficient germination to test, most likely due to seed age (the seed was several years old): 4145-19/4160, 4144-12/4160, and 4142-7/4163. The heterozygotes in this trial all contained both chromosome 2 plus chromosome 10 resistance genes (homozygous chromosome 2 plants were crossed with homozygous chromosome 10 plants resulting in the same heterozygous genotype as results from a two gene homozygote crossed to a wild type). Homozygous mutant sources used in the crosses were S2 and S3 for the chromosome 2 gene location, and S5 for the chromosome 10 gene location.

Four of the six 2 + 10 double heterozygotes tested demonstrated less then 5% injury to the 2X rate (0.367 lb al/A) of POAST PLUS (see Table VIII). Only 9% initial injury 7 DAT was observed to the 4X rate of POAST PLUS with one of the 2 + 10 heterozygotes (4163/4142-2). This degree of tolerance was almost as high as the chromosome 2 homozygous commercial hybrid DK 592sr. The susceptible Pioneer Hybrid 3140 was severely injured (22%) at the 1/8X use rate (0.02 lb ai/A) of POAST PLUS.

GH Trial 126/96: Chromosome 2 and chromosome 10 single heterozygotes were employed in a second comparative trial testing the best 2 +10 double heterozygotes from trial 067/96 plus three new double heterozygotes produced in 1995, to these single heterozygotes. The chromosome 10 single heterozygotes demonstrated slight tolerance to sethoxydim in comparison to the susceptible Pioneer hybrid, through the plants were still injured 15 to 19% at 0.02 lb al/A, 8 DAT (see Table IX). The chromosome 2 single heterozygotes demonstrated more tolerance than the chromosome 10 single heterozygotes but still demonstrated a range of 17 to 48% injury at 0.09 to 0.18 lb al/A. All of the 2 +10 double heterozygotes demonstrated excellent tolerance (0% injury) at 0.36 lb al/A (2X rate) of POAST PLUS. Five of the seven double heterozygotes

demonstrated less than 10% injury to POAST PLUS at 0.76 lb ai/A. This was better tolerance than was observed with the chromosome 10 single homozygous line, and almost equivalent to the tolerance observed with the chromosome 2 single homozygous commercial hybrid, DK 592sr.

Thus, double chromosome 2 plus chromosome 10 heterozygous resistant corn demonstrated greatly improved tolerance to sethoxydim as compared to chromosome 2 heterozygous material.

Table VIII

GH Trial 067/96 (Spray Volume = 20 GPA)

Growth Stage		% CORN INJURY							
		GS 14 (4 Leaves Expanded at Application)							
Hybrids		2 +10 4163/4142-2		2 + 10 4142-2/4163		2 +10 4142-8/4158		2 +10 4158/4142-20	
Treatment in lb al/A		7 DAT	14 DAT	7 DAT	14 DAT	7 DAT	14 DAT	7 DAT	14 DAT
ATPLUS 411 F 1.25% v/v + POAST PLUS	0.02	0	0	0	0	0	0	2	0
	0.04	0	0	0	0	3	1	0	0
	0.09	0	0	0	0	3	0	3	0
	0.18	0	3	3	0	5	2	0	0
	0.36	0	2	3	2	3	2	0	0
	0.71	9	4	15	9	15	5	13	7
Hybrids		2 homo DK 592_{SR}		Susceptible Pl 3140		2 +10 4143-3/4160		2 + 10 4143-1/4160	
		7 DAT	14 DAT	7 DAT	14 DAT	7 DAT	14 DAT	7 DAT	14 DAT
ATPLUS 411 F 1.25% v/v + POAST PLUS	0.02	0	0	22	14	0	0	0	0
	0.04	0	0	30	78	0	0	2	0
	0.09	0	0	43	88	10	8	18	7
	0.18	0	0	85	99	45	60	20	23
	0.36	0	0	83	100				
	0.71	4	0	93	100				

Table IX

GH Trial 126/96

Growth Stage		% CORN INJURY													
		GS 14 (4 Leaves Expanded at Application)													
Hybrids		2 (single het) A 156/6200-6		10 (single het) A 188/6202-7		10 (single het) A 188/6202-g		2 (single het) A 188/6202-1		2 + 10 het 4142-5/4158		2 + 10 het 4158/4142-20		2 + 10 het 4163/4142-2	
Treatment in lb aI/A		8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT
ATPLUS 411 F 1.25% v/v + POAST PLUS	0.02	3	0	19	13	15	9	0	0	0	0	0	0	0	0
	0.04	18	9	43	45	33	38	3	3	0	0	0	0	0	0
	0.09	48	38	58	60	50	60	18	15	0	0	5	0	5	0
	0.18	28	17	68	58	63	58	17	8	0	0	0	0	0	0
	0.36	50	45	73	78	73	80	48	28	0	0	0	0	0	0
	0.71	63	53	78	95	78	100	65	40	7	7	0	3	5	4
Hybrids		2 + 10 het 4142-2/4165		2 + 10 het 6245		2 +10 het 6224		2 + 10 het 6223		10 homo 6202-9		2 homo DK 592$_{SR}$		Susceptible PI 3140	
		8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT	8 DAT	15 DAT
ATPLUS 411 F 1.25% v/v + POAST PLUS	0.02	0	0	0	0	0	0	0	0	0	0	0	0	73	45
	0.04	0	0	0	0	0	0	0	0	5	3	0	0	65	70
	0.09	0	0	5	0	0	0	0	0	9	4	0	0	78	93
	0.18	0	0	0	0	0	0	0	0	5	0	0	0	88	100
	0.36	0	0	0	0	0	0	0	0	13	7	0	0	99	100
	0.71	17	5	10	7	7	2	5	3	45	40	0	0	100	100

Spray Volume = 20 GPA

The invention is not limited to the exact details shown and described, for it should be understood that many variations and modifications may be made while remaining within the spirit and scope of the invention defined by the claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Regents of the University of Minnesota

(ii) TITLE OF THE INVENTION: GENE COMBINATIONS FOR HERBICIDE TOLERANCE IN CORN

(iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
(B) STREET: P.O. Box 2938
(C) CITY: Minneapolis
(D) STATE: MN
(E) COUNTRY: USA
(F) ZIP: 55402

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: Unknown
(B) FILING DATE: 29-AUG-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US96/04625
(B) FILING DATE: 04-APR-1996

(A) APPLICATION NUMBER: 08/679,826
(B) FILING DATE: 30-AUG-1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Woessner, Warren D
(B) REGISTRATION NUMBER: 30,440
(C) REFERENCE/DOCKET NUMBER: 600.370WO1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 612-359-3260
(B) TELEFAX: 612-359-3263
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 258 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTTCCTGCAA ACATTGGTGG ACCTCTTCCT ATTACCAAAC CTCTGGACCC TCCAGACAGA 60
CCTGTTGCTT ACATCCCTGA GAACACATGC GATCCACGTG CAGCTATCTG TGGTGTAGAT 120
GACAGCCAAG GGAAATGGTT GGGTGGTATG TTTGACAAAG ACAGCTTTGT GGAGACATTT 180
GAAGGATGGG CAAAAACAGT GGTTACTGGC AGAGCAAAGC TTGGAGGAAT TCCTGTGGGC 240
GTCATAGCTG TGGAGACA 258

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Val Met Lys Met
1

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGATATATGC TTTGGTGGAA TCTGGC 26

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGTCTTCAAT TGTGCTGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCTTGACGAA CAGACTGGCT GTGC 24

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CACAGCCAGT CTGTTCGTCA AGG 23

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCTCTACGTA ATTGGTCAGC 20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CATAGCTATG GCAACTCCGG 20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7470 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGTCTTCAAT TGTGCTGTCT GGGCCACGGA ACGACAATGT CACAGCTTGG ATTAGCCGCA 60
GCTGCCTCAA AGGCCTTGCC ACTACTCCCT AATCGCCAGA GAAGTTCAGC TGGGACTACA 120
TTCTCATCAT CTTCATTATC GAGGCCCTTA AACAGAAGGA AAAGCCATAC TCGTTCACTC 180
CGTGATGGCG GAGATGGGGT ATCAGATGCC AAAAAGCACA GCCAGTCTGT TCGTCAAGGT 240
CTTGCTGGCA TTATCGACCT CCCAAGTGAG GCACCTTCCG AAGTGGATAT TTCACATGGA 300
TCTGAGGATC CTAGGGGGCC AACAGATTCT TATCAAATGA ATGGGATTAT CAATGAAACA 360
CATAATGGAA GACATGCCTC AGTGTCCAAG GTTGTTGAAT TTTGTGCGGC ACTAGGTGGC 420
AAAACACCAA TTCACAGTAT ATTAGTGGCC AACAATGGAA TGGCAGCAGC AAAATTTATG 480
AGGAGTGTCC GGACATGGGC TAATGATACT TTTGGATCTG AGAAGGCAAT TCAACTCATA 540
GCTATGGCAA CTCCGGAAGA CATGAGGATA AATGCAGAAC ACATTAGAAT TGCTGACCAA 600
TTCGTAGAGG TGCCTGGTGG AACAAACAAT AATAACTACG CCAATGTTCA ACTCATAGTG 660
GGGATGGCAC AAAAACTAGG TGTTTCTGCT GTTTGGCCTG GTTGGGGTCA TGCTTCTGAG 720
AATCCTGAAC TGCCAGATGC ATTGACCGCA AAAGGGATCG TTTTTCTTGG CCCACCTGCA 780
TCATCAATGA ATGCTTTGGG AGATAAGGTC GGCTCAGCTC TCATTGCTCA AGCAGCCGGG 840
GTCCCAACTC TTGCTTGGAG TGGATCACAT GTTGAAGTTC CATTAGAGTG CTGCTTAGAC 900
GCGATACCTG AGGAGATGTA TAGAAAAGCT TGCGTTACTA CCACAGAGGA AGCAGTTGCA 960
AGTTGTCAAG TGGTTGGTTA TCCTGCCATG ATTAAGGCAT CCTGGGGAGG TGGTGGTAAA 1020
GGAATAAGAA AGGTTCATAA TGATGATGAG GTTAGAGCGC TGTTTAAGCA AGTACAAGGT 1080
GAAGTCCCTG GCTCCCCAAT ATTTGTCATG AGGCTTGCAT CCCAGAGTCG GCATCTTGAA 1140
GTTCAGTTGC TTTGTGATCA ATATGGTAAT GTAGCAGCAC TTCACAGTCG TGATTGCAGT 1200
GTGCAACGGC GACACCAGAA GATTATTGAA GAAGGTCCAG TTACTGTTGC TCCTCGTGAG 1260
ACAGTTAAAG CACTTGAGCA GGCAGCAAGG AGGCTTGCTA AGGCTGTGGG TTATGTTGGT 1320
GCTGCTACTG TTGAGTATCT TTACAGCATG GAAACTGGAG ACTACTATTT TCTGGAACTT 1380
AATCCCCGAC TACAGGTTGA GCATCCAGTC ACTGAGTGGA TAGCTGAAGT GAATCTGCCT 1440
GCAGCTCAAG TTGCTGTTGG AATGGGCATA CCTCTTTGGC AGATTCCAGA AATCAGACGT 1500
TTCTATGGAA TGGACTATGG AGGAGGGTAT GACATTTGGA GGAAAACAGC AGCTCTTGCT 1560
ACACCATTTA ATTTTGATGA AGTAGATTCT CAATGGCCAA AGGGCCATTG TGTAGCAGTT 1620
AGAATTACTA GTGAGGACCC AGATGATGGT TTCAAACCTA CTGGTGGGAA AGTGAAGGAG 1680
ATAAGTTTTA AAAGCAAGCC TAATGTTTGG GCCTACTTCT CAGTAAAGTC TGGTGGAGGC 1740
ATTCATGAAT TTGCTGATTC TCAGTTTGGA CATGCTTTTG CATATGGACT CTCTAGACCA 1800
GCAGCTATAA CAAACATGTC TCTTGCATTA AAAGAGATTC AGATTCGTGG AGAAATTCAT 1860
TCAAATGTTG ATTACACAGT TGACCTCTTA AACGCTTCAG ACTTCAGAGA AAACAAGATC 1920
CACACTGGTT GGCTGGATAC AAGAATAGCT ATGCGTGTTC AAGCTGAGAG GCCCCCATGG 1980
TATATCTCAG TGGTTGGAGG TGCTTTATAT AAAACAGTAA CCACCAATGC AGCCACTGTT 2040
TCTGAATATG TTAGTTATCT CACCAAGGGC CATATTCCAC CAAAGCATAT ATCCCTTGTC 2100
AATTCTACAG TTAATTTGAA TATAGAAGGG AGCAAATACA CAATTGAAAC TGTAAGGACT 2160
GGACATGGTA GCTACAGGTT GAGAATGAAT GATTCAACAG TTGAAGCGAA TGTACAATCT 2220
TTATGTGATG GTGGCCTCTT AATGCAGTTG GATGGAAACA GCCATGTAAT TTATGCAGAA 2280
GAAGAAGCTG GTGGTACACG GCTTCAGATT GATGGAAAGA CATGTTTATT GCAGAATGAC 2340
CATGATCCAT CGAAGTTATT AGCTGAGACA CCCTGCAAAC TTCTTCGTTT CTTGGTTGCT 2400
GATGGTGCTC ATGTTGATGC GGATGTACCA TACGCGGAAG TTGAGGTTAT GAAGATGTGC 2460

ATGCCTCTCT TGTCACCTGC TTCTGGTGTC ATTCATTGTA TGATGTCTGA GGGCCAGGCA 2520
TTGCAGGCTG GTGATCTTAT AGCAAGGTTG GATCTTGATG ACCCTTCTGC TGTGAAAAGA 2580
GCTGAGCCAT TTGATGGAAT ATTTCCACAA ATGGAGCTCC CTGTTGCTGT CTCTAGTCAA 2640
GTACACAAAA GATATGCTGC AAGTTTGAAT GCTGCTCGAA TGGTCCTTGC AGGATATGAG 2700
CACAATATTA ATGAAGTCGT TCAAGATTTG GTATGCTGCC TGGACAACCC TGAGCTTCCT 2760
TTCCTACAGT GGGATGAACT TATGTCTGTT CTAGCAACGA GGCTTCCAAG AAATCTCAAG 2820
AGTGAGTTAG AGGATAAATA CAAGGAATAC AAGTTGAATT TTTACCATGG AAAAAACGAG 2880
GACTTTCCAT CCAAGTTGCT AAGAGACATC ATTGAGGAAA ATCTTTCTTA TGGTTCAGAG 2940
AAGGAAAAGG CTACAAATGA GAGGCTTGTT GAGCCTCTTA TGAACCTACT GAAGTCATAT 3000
GAGGGTGGGA GAGAGAGCCA TGCACATTTT GTTGTCAAGT CTCTTTTCGA GGAGTATCTT 3060
ACAGTGGAAG AACTTTTTAG TGATGGCATT CAGTCTGACG TGATTGAAAC ATTGCGGCAT 3120
CAGCACAGTA AAGACCTGCA GAAGGTTGTA GACATTGTGT TGTCTCACCA GGGTGTGAGG 3180
AACAAAGCTA AGCTTGTAAC GGCACTTATG GAAAAGCTGG TTTATCCAAA TCCTGGTGGT 3240
TACAGGGATC TGTTAGTTCG CTTTTCTTCC CTCAATCATA AAAGATATTA TAAGTTGGCC 3300
CTTAAAGCAA GTGAACTTCT TGAACAAACC AAACTAAGTG AACTCCGTGC AAGCGTTGCA 3360
AGAAGCCTTT CGGATCTGGG GATGCATAAG GGAGAAATGA GTATTAAGGA TAACATGGAA 3420
GATTTAGTCT CTGCCCCATT ACCTGTTGAA GATGCTCTGA TTTCTTTGTT TGATTACAGT 3480
GATCGAACTG TTCAGCAGAA AGTGATTGAG ACATACATAT CACGATTGTA CCAGCCTCAT 3540
CTTGTAAAGG ATAGCATCCA AATGAAATTC AAGGAATCTG GTGCTATTAC TTTTTGGGAA 3600
TTTTATGAAG GGCATGTTGA TACTAGAAAT GGACATGGGG CTATTATTGG TGGGAAGCGA 3660
TGGGGTGCCA TGGTCGTTCT CAAATCACTT GAATCTGCGT CAACAGCCAT TGTGGCTGCA 3720
TTAAAGGATT CGGCACAGTT CAACAGCTCT GAGGGCAACA TGATGCACAT TGCATTATTG 3780
AGTGCTGAAA ATGAAAGTAA TATAAGTGGA ATAAGCAGTG ATGATCAAGC TCAACATAAG 3840
ATGGAAAAGC TTAGCAAGAT ACTGAAGGAT ACTAGCGTTG CAAGTGATCT CCAAGCTGCT 3900
GGTTTGAAGG TTATAAGTTG CATTGTTCAA AGAGATGAAG CTCGCATGCC AATGCGCCAC 3960
ACATTCCTCT GGTTGGATGA CAAGAGTTGT TATGAAGAAG AGCAGATTCT CCGGCATGTG 4020
GAGCCTCCCC TCTCTACACT TCTTGAATTG GATAAGTTGA AGGTGAAAGG ATACAATGAA 4080
ATGAAGTATA CTCCTTCGCG TGACCGCCAA TGGCATATCT ACACACTAAG AAATACTGAA 4140
AACCCCAAAA TGTTGCATAG GGTGTTTTTC CGAACTATTG TCAGGCAACC CAATGCAGGC 4200
AACAAGTTTA GATCGGCTCA GATCAGCGAC GCTGAGGTAG GATGTCCCGA AGAATCTCTT 4260
TCATTTACAT CAAATAGCAT CTTAAGATCA TTGATGACTG CTATTGAAGA ATTAGAGCTT 4320
CATGCAATTA GGACAGGTGA TTCTCACATG TATTTGTGCA TACTGAAAGA GCAAAAGCTT 4380
CTTGACCTCA TTCCATTTTC AGGGAGTACA ATTGTTGATG TTGGCCAAGA TGAAGCTACC 4440
GCTTGTTCAC TTTTAAAATC AATGGCTTTG AAGATACATG AGCTTGTTGG TGCAAGGATG 4500
CATCATCTGT CTGTATGCCA GTGGGAGGTG AAACTCAAGT TGGACTGTGA TGGCCCTGCA 4560
AGTGGTACCT GGAGAGTTGT AACTACAAAT GTTACTGGTC ACACCTGCAC CATTGATATA 4620
TACCGAGAAG TGGAGGAAAT AGAATCACAG AAGTTAGTGT ACCATTCAGC CAGTTCGTCA 4680
GCTGGACCAT TGCATGGTGT TGCACTGAAT AATCCATATC AACCTTTGAG TGTGATTGAT 4740
CTAAAGCGCT GCTCTGCTAG GAACAACAGA ACAACATATT GCTATGATTT TCCGCTGGCC 4800
TTTGAAACTG CACTGCAGAA GTCATGGCAG TCCAATGGCT CTACTGTTTC TGAAGGCAAT 4860
GAAAATAGTA AATCCTACGT GAAGGCAACT GAGCTAGTGT TTGCTGAAAA ACATGGGTCC 4920
TGGGGCACTC CTATAATTCC GATGGAACGC CCTGCTGGGC TCAACGACAT TGGTATGGTC 4980
GCTTGGATCA TGGAGATGTC AACACCTGAA TTTCCCAATG GCAGGCAGAT TATTGTTGTA 5040
GCAAATGATA TCACTTTCAG AGCTGGATCA TTTGGCCCAA GGGAAGATGC ATTTTTTGAA 5100
ACTGTCACTA ACCTGGCTTG CGAAAGGAAA CTTCCTCTTA TATACTTGGC AGCAAACTCT 5160
GGTGCTAGGA TTGGCATAGC TGATGAAGTA AAATCTTGCT TCCGTGTTGG ATGGTCTGAC 5220
GAAGGCAGTC CTGAACGAGG GTTTCAGTAC ATCTATCTGA CTGAAGAAGA CTATGCTCGC 5280
ATTAGCTCTT CTGTTATAGC ACATAAGCTG GAGCTAGATA GTGGTGAAAT TAGGTGGATT 5340
ATTGACTCTG TTGTGGGCAA GGAGGATGGG CTTGGTGTCG AGAACATACA TGGAAGTGCT 5400
GCTATTGCCA GTGCTTATTC TAGGGCATAT GAGGAGACAT TTACACTTAC ATTTGTGACT 5460
GGGCGGACTG TAGGAATAGG AGCTTATCTT GCTCGACTTG GTATACGGTG CATACAGCGT 5520
CTTGACCAGC CTATTATTTT AACAGGGTTT TCTGCCCTGA ACAAGCTCCT TGGGCGGGAA 5580
GTGTACAGCT CCCACATGCA GCTTGGTGGT CCTAAGATCA TGGCGACCAA TGGTGTTGTC 5640
CACCTCACTG TTCCAGATGT CCTTGAAGGT GTTTCCAATA TATTGAGGTG GCTCAGCTAT 5700
GTTCCTGCAA ACATTGGTGG ACCTCTTCCT ATTACCAAAC CTCTGGACCC TCCAGACAGA 5760

```
CCTGTTGCTT ACATCCCTGA GAACACATGC GATCCACGTG CAGCTATCTG TGGTGTAGAT   5820
GACAGCCAAG GGAAATGGTT GGGTGGTATG TTTGACAAAG ACAGCTTTGT GGAGACATTT   5880
GAAGGATGGG CAAAAACAGT GGTTACTGGC AGAGCAAAGC TTGGAGGAAT TCCTGTGGGC   5940
GTCATAGCTG TGGAGACACA GACCATGATG CAGATCATCC CTGCTGATCC AGGTCAGCTT   6000
GATTCCCATG AGCGATCTGT CCCTCGTGCT GGACAAGTGT GGTTCCCAGA TTCTGCAACC   6060
AAGACCGCTC AGGCATTATT AGACTTCAAC CGTGAAGGAT TGCCTCTGTT CATCCTGGCT   6120
AATTGGAGAG GCTTCTCTGG TGGACAAAGA GATCTCTTTG AAGGAATTCT TCAGGCTGGG   6180
TCAACAATTG TCGAGAACCT TAGGACATAT AATCAGCCTG CTTTTGTGTA CATTCCTATG   6240
GCTGGAGAGC TTCGTGGAGG AGCTTGGGTT GTGGTCGATA GCAAAATAAA TCCAGACCGC   6300
ATTGAGTGTT ATGCTGAAAG GACTGCCAAA GGTAATGTTC TCGAACCTCA AGGGTTAATT   6360
GAAATCAAGT TCAGGTCAGA GGAACTCCAA GACTGTATGG GTAGGCTTGA CCCAGAGTTG   6420
ATAAATCTGA AAGCAAAACT CCAAGATGTA AATCATGGAA ATGGAAGTCT ACCAGACATA   6480
GAAGGGATTC GGAAGAGTAT AGAAGCACGT ACGAAACAGT TGCTGCCTTT ATATACCCAG   6540
ATTGCAATAC GGTTTGCTGA ATTGCATGAT ACTTCCCTAA GAATGGCAGC TAAAGGTGTG   6600
ATTAAGAAAG TTGTAGACTG GGAAGAATCA CGCTCGTTCT TCTATAAAAG GCTACGGAGG   6660
AGGATCGCAG AAGATGTTCT TGCAAAAGAA ATAAGGCAGA TAGTCGGTGA TAAATTTACG   6720
CACCAATTAG CAATGGAGCT CATCAAGGAA TGGTACCTTG CTTCTCAGGC CACAACAGGA   6780
AGCACTGGAT GGGATGACGA TGATGCTTTT GTTGCCTGGA AGGACAGTCC TGAAAACTAC   6840
AAGGGGCATA TCCAAAAGCT TAGGGCTCAA AAAGTGTCTC ATTCGCTCTC TGATCTTGCT   6900
GACTCCAGTT CAGATCTGCA AGCATTCTCG CAGGGTCTTT CTACGCTATT AGATAAGATG   6960
GATCCCTCTC AGAGAGCGAA GTTTGTTCAG GAAGTCAAGA AGGTCCTTGA TTGATGATAC   7020
CAACACATCC AACACAATGT GTGCATGTCA CATCTTTTTG TTCTAGTACA TACATAGAAG   7080
GATATTGCTT GGTCTTGATT GATCATGTCT GATTTAAGTC GACTATTATT TCTTGGAATT   7140
TTCTTTTGGA CCTGGTGCTA TGGTTGATGG ATGTATATTG GATATGTGCG TTCTGCCAGG   7200
TGTAAGCACA AAGGTTTAGA CARAMMRARA RCAAGAGCGA GTGAACCTGT TCTGGTTTTG   7260
CAGTGGTTCA GTAAGGCAGA AAGTTGTTAA ACCGTAGTTC TGAGATGTAT TACCAGTGNC   7320
GCCATGCTGT ACTTTTAGGG TGTATAATGC GGATACAAAT AAACAATTTA GCGGTTCATT   7380
AAAGTTTGAA CTCAAATAAC ATGTTCTTTG TAAGCATATG TACCGTACCT CTACGTGAAA   7440
TAAAGTTGTT GAATTAGCAT TCGAAAAAAA                                    7470
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2325 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

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

```
                100                 105                 110
His Ala Ser Val Ser Lys Val Val Glu Phe Cys Ala Ala Leu Gly Gly
        115                 120                 125
Lys Thr Pro Ile His Ser Ile Leu Val Ala Asn Asn Gly Met Ala Ala
    130                 135                 140
Ala Lys Phe Met Arg Ser Val Arg Thr Trp Ala Asn Asp Thr Phe Gly
145                 150                 155                 160
Ser Glu Lys Ala Ile Gln Leu Ile Ala Met Ala Thr Pro Glu Asp Met
                165                 170                 175
Arg Ile Asn Ala Glu His Ile Arg Ile Ala Asp Gln Phe Val Glu Val
            180                 185                 190
Pro Gly Gly Thr Asn Asn Asn Asn Tyr Ala Asn Val Gln Leu Ile Val
        195                 200                 205
Gly Met Ala Gln Lys Leu Gly Val Ser Ala Val Trp Pro Gly Trp Gly
    210                 215                 220
His Ala Ser Glu Asn Pro Glu Leu Pro Asp Ala Leu Thr Ala Lys Gly
225                 230                 235                 240
Ile Val Phe Leu Gly Pro Pro Ala Ser Ser Met Asn Ala Leu Gly Asp
                245                 250                 255
Lys Val Gly Ser Ala Leu Ile Ala Gln Ala Ala Gly Val Pro Thr Leu
            260                 265                 270
Ala Trp Ser Gly Ser His Val Glu Val Pro Leu Glu Cys Cys Leu Asp
        275                 280                 285
Ala Ile Pro Glu Glu Met Tyr Arg Lys Ala Cys Val Thr Thr Thr Glu
    290                 295                 300
Glu Ala Val Ala Ser Cys Gln Val Val Gly Tyr Pro Ala Met Ile Lys
305                 310                 315                 320
Ala Ser Trp Gly Gly Gly Gly Lys Gly Ile Arg Lys Val His Asn Asp
                325                 330                 335
Asp Glu Val Arg Ala Leu Phe Lys Gln Val Gln Gly Glu Val Pro Gly
            340                 345                 350
Ser Pro Ile Phe Val Met Arg Leu Ala Ser Gln Ser Arg His Leu Glu
        355                 360                 365
Val Gln Leu Leu Cys Asp Gln Tyr Gly Asn Val Ala Ala Leu His Ser
    370                 375                 380
Arg Asp Cys Ser Val Gln Arg Arg His Gln Lys Ile Ile Glu Glu Gly
385                 390                 395                 400
Pro Val Thr Val Ala Pro Arg Glu Thr Val Lys Ala Leu Glu Gln Ala
                405                 410                 415
Ala Arg Arg Leu Ala Lys Ala Val Gly Tyr Val Gly Ala Ala Thr Val
            420                 425                 430
Glu Tyr Leu Tyr Ser Met Glu Thr Gly Asp Tyr Tyr Phe Leu Glu Leu
        435                 440                 445
Asn Pro Arg Leu Gln Val Glu His Pro Val Thr Glu Trp Ile Ala Glu
    450                 455                 460
Val Asn Leu Pro Ala Ala Gln Val Ala Val Gly Met Gly Ile Pro Leu
465                 470                 475                 480
Trp Gln Ile Pro Glu Ile Arg Arg Phe Tyr Gly Met Asp Tyr Gly Gly
                485                 490                 495
Gly Tyr Asp Ile Trp Arg Lys Thr Ala Ala Leu Ala Thr Pro Phe Asn
            500                 505                 510
Phe Asp Glu Val Asp Ser Gln Trp Pro Lys Gly His Cys Val Ala Val
        515                 520                 525
Arg Ile Thr Ser Glu Asp Pro Asp Asp Gly Phe Lys Pro Thr Gly Gly
    530                 535                 540
```

```
Lys Val Lys Glu Ile Ser Phe Lys Ser Lys Pro Asn Val Trp Ala Tyr
545                 550                 555                 560
Phe Ser Val Lys Ser Gly Gly Gly Ile His Glu Phe Ala Asp Ser Gln
                565                 570                 575
Phe Gly His Ala Phe Ala Tyr Gly Leu Ser Arg Pro Ala Ala Ile Thr
            580                 585                 590
Asn Met Ser Leu Ala Leu Lys Glu Ile Gln Ile Arg Gly Glu Ile His
        595                 600                 605
Ser Asn Val Asp Tyr Thr Val Asp Leu Leu Asn Ala Ser Asp Phe Arg
    610                 615                 620
Glu Asn Lys Ile His Thr Gly Trp Leu Asp Thr Arg Ile Ala Met Arg
625                 630                 635                 640
Val Gln Ala Glu Arg Pro Pro Trp Tyr Ile Ser Val Val Gly Gly Ala
                645                 650                 655
Leu Tyr Lys Thr Val Thr Thr Asn Ala Ala Thr Val Ser Glu Tyr Val
            660                 665                 670
Ser Tyr Leu Thr Lys Gly His Ile Pro Pro Lys His Ile Ser Leu Val
        675                 680                 685
Asn Ser Thr Val Asn Leu Asn Ile Glu Gly Ser Lys Tyr Thr Ile Glu
    690                 695                 700
Thr Val Arg Thr Gly His Gly Ser Tyr Arg Leu Arg Met Asn Asp Ser
705                 710                 715                 720
Thr Val Glu Ala Asn Val Gln Ser Leu Cys Asp Gly Gly Leu Leu Met
                725                 730                 735
Gln Leu Asp Gly Asn Ser His Val Ile Tyr Ala Glu Glu Glu Ala Gly
            740                 745                 750
Gly Thr Arg Leu Gln Ile Asp Gly Lys Thr Cys Leu Leu Gln Asn Asp
        755                 760                 765
His Asp Pro Ser Lys Leu Leu Ala Glu Thr Pro Cys Lys Leu Leu Arg
    770                 775                 780
Phe Leu Val Ala Asp Gly Ala His Val Asp Ala Asp Val Pro Tyr Ala
785                 790                 795                 800
Glu Val Glu Val Met Lys Met Cys Met Pro Leu Leu Ser Pro Ala Ser
                805                 810                 815
Gly Val Ile His Cys Met Met Ser Glu Gly Gln Ala Leu Gln Ala Gly
            820                 825                 830
Asp Leu Ile Ala Arg Leu Asp Leu Asp Asp Pro Ser Ala Val Lys Arg
        835                 840                 845
Ala Glu Pro Phe Asp Gly Ile Phe Pro Gln Met Glu Leu Pro Val Ala
    850                 855                 860
Val Ser Ser Gln Val His Lys Arg Tyr Ala Ala Ser Leu Asn Ala Ala
865                 870                 875                 880
Arg Met Val Leu Ala Gly Tyr Glu His Asn Ile Asn Glu Val Val Gln
                885                 890                 895
Asp Leu Val Cys Cys Leu Asp Asn Pro Glu Leu Pro Phe Leu Gln Trp
            900                 905                 910
Asp Glu Leu Met Ser Val Leu Ala Thr Arg Leu Pro Arg Asn Leu Lys
        915                 920                 925
Ser Glu Leu Glu Asp Lys Tyr Lys Glu Tyr Lys Leu Asn Phe Tyr His
    930                 935                 940
Gly Lys Asn Glu Asp Phe Pro Ser Lys Leu Leu Arg Asp Ile Ile Glu
945                 950                 955                 960
Glu Asn Leu Ser Tyr Gly Ser Glu Lys Glu Lys Ala Thr Asn Glu Arg
                965                 970                 975
Leu Val Glu Pro Leu Met Asn Leu Leu Lys Ser Tyr Glu Gly Gly Arg
```

```
            980                 985                 990
Glu Ser His Ala His Phe Val Val Lys Ser Leu Phe Glu Glu Tyr Leu
        995                1000                1005
Thr Val Glu Glu Leu Phe Ser Asp Gly Ile Gln Ser Asp Val Ile Glu
   1010                1015                1020
Thr Leu Arg His Gln His Ser Lys Asp Leu Gln Lys Val Val Asp Ile
025                1030                1035                1040
Val Leu Ser His Gln Gly Val Arg Asn Lys Ala Lys Leu Val Thr Ala
               1045                1050                1055
Leu Met Glu Lys Leu Val Tyr Pro Asn Pro Gly Gly Tyr Arg Asp Leu
           1060                1065                1070
Leu Val Arg Phe Ser Ser Leu Asn His Lys Arg Tyr Tyr Lys Leu Ala
       1075                1080                1085
Leu Lys Ala Ser Glu Leu Leu Glu Gln Thr Lys Leu Ser Glu Leu Arg
   1090                1095                1100
Ala Ser Val Ala Arg Ser Leu Ser Asp Leu Gly Met His Lys Gly Glu
105                1110                1115                1120
Met Ser Ile Lys Asp Asn Met Glu Asp Leu Val Ser Ala Pro Leu Pro
               1125                1130                1135
Val Glu Asp Ala Leu Ile Ser Leu Phe Asp Tyr Ser Asp Arg Thr Val
           1140                1145                1150
Gln Gln Lys Val Ile Glu Thr Tyr Ile Ser Arg Leu Tyr Gln Pro His
       1155                1160                1165
Leu Val Lys Asp Ser Ile Gln Met Lys Phe Lys Glu Ser Gly Ala Ile
   1170                1175                1180
Thr Phe Trp Glu Phe Tyr Glu Gly His Val Asp Thr Arg Asn Gly His
185                1190                1195                1200
Gly Ala Ile Ile Gly Gly Lys Arg Trp Gly Ala Met Val Val Leu Lys
               1205                1210                1215
Ser Leu Glu Ser Ala Ser Thr Ala Ile Val Ala Ala Leu Lys Asp Ser
           1220                1225                1230
Ala Gln Phe Asn Ser Ser Glu Gly Asn Met Met His Ile Ala Leu Leu
       1235                1240                1245
Ser Ala Glu Asn Glu Ser Asn Ile Ser Gly Ile Ser Ser Asp Asp Gln
   1250                1255                1260
Ala Gln His Lys Met Glu Lys Leu Ser Lys Ile Leu Lys Asp Thr Ser
265                1270                1275                1280
Val Ala Ser Asp Leu Gln Ala Ala Gly Leu Lys Val Ile Ser Cys Ile
               1285                1290                1295
Val Gln Arg Asp Glu Ala Arg Met Pro Met Arg His Thr Phe Leu Trp
           1300                1305                1310
Leu Asp Asp Lys Ser Cys Tyr Glu Glu Glu Gln Ile Leu Arg His Val
       1315                1320                1325
Glu Pro Pro Leu Ser Thr Leu Leu Glu Leu Asp Lys Leu Lys Val Lys
   1330                1335                1340
Gly Tyr Asn Glu Met Lys Tyr Thr Pro Ser Arg Asp Arg Gln Trp His
345                1350                1355                1360
Ile Tyr Thr Leu Arg Asn Thr Glu Asn Pro Lys Met Leu His Arg Val
               1365                1370                1375
Phe Phe Arg Thr Ile Val Arg Gln Pro Asn Ala Gly Asn Lys Phe Arg
           1380                1385                1390
Ser Ala Gln Ile Ser Asp Ala Glu Val Gly Cys Pro Glu Glu Ser Leu
       1395                1400                1405
Ser Phe Thr Ser Asn Ser Ile Leu Arg Ser Leu Met Thr Ala Ile Glu
   1410                1415                1420
```

```
Glu Leu Glu Leu His Ala Ile Arg Thr Gly His Ser His Met Tyr Leu
425                 1430                1435                1440
Cys Ile Leu Lys Glu Gln Lys Leu Leu Asp Leu Ile Pro Phe Ser Gly
                1445                1450                1455
Ser Thr Ile Val Asp Val Gly Gln Asp Glu Ala Thr Ala Cys Ser Leu
            1460                1465                1470
Leu Ser Lys Met Ala Leu Lys Ile His Glu Leu Val Gly Ala Arg Met
        1475                1480                1485
His His Leu Ser Val Cys Gln Trp Glu Val Lys Leu Lys Leu Asp Cys
    1490                1495                1500
Asp Gly Pro Ala Ser Gly Thr Trp Arg Val Val Thr Thr Asn Val Thr
505                 1510                1515                1520
Gly His Thr Cys Thr Ile Asp Ile Tyr Arg Glu Val Glu Glu Ile Glu
                1525                1530                1535
Ser Gln Lys Leu Val Tyr His Ser Ala Ser Ser Ser Ala Gly Pro Leu
            1540                1545                1550
His Gly Val Ala Leu Asn Asn Pro Tyr Gln Pro Leu Ser Val Ile Asp
        1555                1560                1565
Leu Lys Arg Cys Ser Ala Arg Asn Asn Arg Thr Thr Tyr Cys Tyr Asp
    1570                1575                1580
Phe Pro Leu Ala Phe Glu Thr Ala Leu Gln Lys Ser Trp Gln Ser Asn
585                 1590                1595                1600
Gly Ser Thr Val Ser Glu Gly Asn Glu Asn Ser Lys Ser Tyr Val Lys
                1605                1610                1615
Ala Thr Glu Leu Val Phe Ala Glu Lys His Gly Ser Trp Gly Thr Pro
            1620                1625                1630
Ile Ile Pro Met Glu Arg Pro Ala Gly Leu Asn Asp Ile Gly Met Val
        1635                1640                1645
Ala Trp Ile Met Glu Met Ser Thr Pro Glu Phe Pro Asn Gly Arg Gln
    1650                1655                1660
Ile Ile Val Val Ala Asn Asp Ile Thr Phe Arg Ala Gly Ser Phe Gly
665                 1670                1675                1680
Pro Arg Glu Asp Ala Phe Phe Glu Thr Val Thr Asn Leu Ala Cys Glu
                1685                1690                1695
Arg Lys Leu Pro Leu Ile Tyr Leu Ala Ala Asn Ser Gly Ala Arg Ile
            1700                1705                1710
Gly Ile Ala Asp Glu Val Lys Ser Cys Phe Arg Val Gly Trp Ser Asp
        1715                1720                1725
Glu Gly Ser Pro Glu Arg Gly Phe Gln Tyr Ile Tyr Leu Thr Glu Glu
    1730                1735                1740
Asp Tyr Ala Arg Ile Ser Ser Ser Val Ile Ala His Lys Leu Glu Leu
745                 1750                1755                1760
Asp Ser Gly Glu Ile Arg Trp Ile Ile Asp Ser Val Val Gly Lys Glu
                1765                1770                1775
Asp Gly Leu Gly Val Glu Asn Ile His Gly Ser Ala Ala Ile Ala Ser
            1780                1785                1790
Ala Tyr Ser Arg Ala Tyr Glu Glu Thr Phe Thr Leu Thr Phe Val Thr
        1795                1800                1805
Gly Arg Thr Val Gly Ile Gly Ala Tyr Leu Ala Arg Leu Gly Ile Arg
    1810                1815                1820
Cys Ile Gln Arg Leu Asp Gln Pro Ile Ile Leu Thr Gly Phe Ser Ala
825                 1830                1835                1840
Leu Asn Lys Leu Leu Gly Arg Glu Val Tyr Ser Ser His Met Gln Leu
                1845                1850                1855
Gly Gly Pro Lys Ile Met Ala Thr Asn Gly Val Val His Leu Thr Val
```

```
          1860                1865                1870
Pro Asp Val Leu Glu Gly Val Ser Asn Ile Leu Arg Trp Leu Ser Tyr
       1875                1880                1885
Val Pro Ala Asn Ile Gly Gly Pro Leu Pro Ile Thr Lys Pro Leu Asp
   1890                1895                1900
Pro Pro Asp Arg Pro Val Ala Tyr Ile Pro Glu Asn Thr Cys Asp Pro
905                1910                1915                1920
Arg Ala Ala Ile Cys Gly Val Asp Asp Ser Gln Gly Lys Trp Leu Gly
               1925                1930                1935
Gly Met Phe Asp Lys Asp Ser Phe Val Glu Thr Phe Glu Gly Trp Ala
           1940                1945                1950
Lys Thr Val Val Thr Gly Arg Ala Lys Leu Gly Gly Ile Pro Val Gly
       1955                1960                1965
Val Ile Ala Val Glu Thr Gln Thr Met Met Gln Ile Ile Pro Ala Asp
   1970                1975                1980
Pro Gly Gln Leu Asp Ser His Glu Arg Ser Val Pro Arg Ala Gly Gln
985                1990                1995                2000
Val Trp Phe Pro Asp Ser Ala Thr Lys Thr Ala Gln Ala Leu Leu Asp
               2005                2010                2015
Phe Asn Arg Glu Gly Leu Pro Leu Phe Ile Leu Ala Asn Trp Arg Gly
           2020                2025                2030
Phe Ser Gly Gly Gln Arg Asp Leu Phe Glu Gly Ile Leu Gln Ala Gly
       2035                2040                2045
Ser Thr Ile Val Glu Asn Leu Arg Thr Tyr Asn Gln Pro Ala Phe Val
   2050                2055                2060
Tyr Ile Pro Met Ala Gly Glu Leu Arg Gly Gly Ala Trp Val Val Val
065                2070                2075                2080
Asp Ser Lys Ile Asn Pro Asp Arg Ile Glu Cys Tyr Ala Glu Arg Thr
               2085                2090                2095
Ala Lys Gly Asn Val Leu Glu Pro Gln Gly Leu Ile Glu Ile Lys Phe
           2100                2105                2110
Arg Ser Glu Glu Leu Gln Asp Cys Met Gly Arg Leu Asp Pro Glu Leu
       2115                2120                2125
Ile Asn Leu Lys Ala Lys Leu Gln Asp Val Asn His Gly Asn Gly Ser
   2130                2135                2140
Leu Pro Asp Ile Glu Gly Ile Arg Lys Ser Ile Glu Ala Arg Thr Lys
145                2150                2155                2160
Gln Leu Leu Pro Leu Tyr Thr Gln Ile Ala Ile Arg Phe Ala Glu Leu
               2165                2170                2175
His Asp Thr Ser Leu Arg Met Ala Ala Lys Gly Val Ile Lys Lys Val
           2180                2185                2190
Val Asp Trp Glu Glu Ser Arg Ser Phe Phe Tyr Lys Arg Leu Arg Arg
       2195                2200                2205
Arg Ile Ala Glu Asp Val Leu Ala Lys Glu Ile Arg Gln Ile Val Gly
   2210                2215                2220
Asp Lys Phe Thr His Gln Leu Ala Met Glu Leu Ile Lys Glu Trp Tyr
225                2230                2235                2240
Leu Ala Ser Gln Ala Thr Thr Gly Ser Thr Gly Trp Asp Asp Asp Asp
               2245                2250                2255
Ala Phe Val Ala Trp Lys Asp Ser Pro Glu Asn Tyr Lys Gly His Ile
           2260                2265                2270
Gln Lys Leu Arg Ala Gln Lys Val Ser His Ser Leu Ser Asp Leu Ala
       2275                2280                2285
Asp Ser Ser Ser Asp Leu Gln Ala Phe Ser Gln Gly Leu Ser Thr Leu
   2290                2295                2300
```

Leu Asp Lys Met Asp Pro Ser Gln Arg Ala Lys Phe Val Gln Glu Val
305 2310 2315 2320
Lys Lys Val Leu Asp
2325

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3544 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAGCTTGGTA TGGATTCGTC AGCGCCAAGC CGGGGTTTTG CATGCGCCCG ACTGGAARCS 60
GAATTCCGTG AGCCCTGTAC RRCAATGGCA ACCCCASGGT TACTGGGGTG GCTGAATGGT 120
CTCSGCTTAC GCAATTGTTT GTGGCAGCWG CGTGGGCTAA ATGTARGTTG TCTCTTGTTG 180
CACTGCARGA TGGATGGGTA GCCTCTGGGC CGCCTCTGCT ARTGTCTARC GTTTGCTGAC 240
TGTGGTTTAN TCAGGGATGC CCATGCCCAT GCTAGATTGA TAGGTGCCAT TCTAATGGTA 300
GGTGGCGGTA AGGTTTATTA AGCTGTAGTA TCAGTAGGTA ACCTCATGAA TCAGGGTTTA 360
AGCACACCTT TTCCTTTGTG TGGGTGCATA AGGAATGCAC TTGGCTTCGT TCCCTGATAG 420
TCTTTGCTCA TGTGTCATTC TACCAAGTGG GTTACTGTAA CATTGCACTC TATGATGGTT 480
GGTGGTTGTG CATCTTTTTG CTTCCCCTGG TTGTCTAATA CCTGCATGTA ACTGATGACC 540
TTCTTTTATG TATCATATAG ATTACATCCT TTTGTTGTAC ATCTCAATTC TGAAAAAACA 600
ATGTTTTGCA TTCTTAGCGC TCTGTGCACA AGGAAAAGGA GGTTTTACCT GCAACTTTTT 660
TTTTCGAGAA AAAACAAACC TTTCTGAAAG GCAGTGATCA TTTAGTATAA AGAAAATTTG 720
ATTTACTTTC TTCAGAGAGA ATATKCCAAR CAAACAATTT TCTTACTGTC TGAGCCACGA 780
AATTTGATCT TGATCTTACT TTCACAAGCC ACATGAAGCC TTATCATCGC TCTGATAAAA 840
AAGCCAAATA GGTGATTCAT AGAATGAGAG AAAGAACCTG TTGCCATTTG GGGRCCTTGT 900
TGTGTACTCA TTATCCCCCC TGCTCAGGTT GAGGTTTTCC TTGCCACTGC CACCCCTTGG 960
CCCCTTCTTA TACAACCATC TCCATTGAAA AAGATTTTGC ACTACATTTG GGCTCGTATG 1020
ACAAAAAAGG AAAATAAAAC TAAACAGCAG AAACATAGTA TAATTATAGG TAAAAGGTTC 1080
TGGCAAGTTT GAGTGGAAGA GACCTTTGTA TATTTGGACA TATTTCACTA GTAAATAGTT 1140
TTCTAAAATC TTCATGAATG GTGGCCAATA AACTTGATAA GATCTCAACA TGGCAGGTTC 1200
CTTCMAAATG AGAGGAAAAC TGGAAACATC ACAAATATTT TTTAGCGAGT GGCCTATAAA 1260
TTATAATGTT GCTTTCATTT CTTTGATATT CAAAACTTCC TAAGAGTATT CTGCTAGAGC 1320
TCTGATGGTG TCTTTTGCCT CTGTCAGATT TTCCAGGAGT TTTCTTCCCT TTTTATGGCA 1380
CTGTGCGTTT GAGAAGGTCT TCAATTGTGC TGTCTGGGCC ACGGAACGAC AATGTCACAG 1440
CTTGGATTAG CCGCAGCTGC CTCAAAGGCC TTGCCACTAC TCCCTAATCG CCAGAGAAGT 1500
TCAGCTGGGA CTACATTCTC ATCATCTTCA TTATCGAGGC CCTTAAACAG AAGGAAAAGC 1560
CATACTCGTT CACTCCGTGA TGGCGGAGAT GGGGTATCAG ATGCCAAAAA GCACAGCCAG 1620
TCTGTTCGTC AAGGTACTGT GAATATCTTT TGATACAAGC TAAAATTTTG CTACAGAATA 1680
TATATTAAAG AGTTCTTTCT TGGCTGGTGT TGTTTATTTN GTTTTAANCA TGCGAAAGGG 1740
CCTCTAGTTG AGTTGGTTAG GTGGCCTGAA TACCACTCCT TAAGGTCTTG AGTTTGNNAT 1800
TTTCCGTCGG AGCGAATTTT AGGCTAGGGT TACCCCCCCA CCCCCACCCG AATCTGCACA 1860
GTCCGGTCGT GGTCGTCCTC ATATAGGCTA CGATGTCATT GTGTATCGGC GGGCCAGGGG 1920
TTTAAGAGTT TTCTTGACCT TTGTTAGAAG ATCTTAATAA TACAATGTCC AAGGGCTGTC 1980
TTACCCTGTA GGTCGAGTTT TTAGTTGTTT TAACATGGTA ATGTTTGAAG CCTCATTCTA 2040
GGTACCAATA TAGATATGCT CACTGCTCAG TTTCAAATGT TTGTCTGCAT GTAGGTCTTG 2100
CTGGCATTAT CGACCTCCCA AGTGAGGCAC CTTCCGAAGT GGATATTTCA CAGTAAGGAC 2160
TACAATATTT TGCGTACGTT TGNTTTTGGA AAAAGAAAAT ATTCTCAGCT TATTTATACT 2220
AGCTTCGCTA ANTACTGAAA NNTGCTGTCT TAATGTCCTG GTNGCTGTAT GCTCAATCTT 2280

```
TCATAGTAAA TGCTGCAAAA TATGTGATGT AACTGTTGCA ACACAGCCAG GGACCTGTTA   2340
TTTAGAGCAT GGTGAATGCT CTGGTTCAGT TATATGATGT AGTTATAGCT CATGTTGAAG   2400
AATTAGTTGC AGTGTTTGCT GGACAATGGT CACTTATTAT AAATCATATC TGCATACACA   2460
TTTGTGACTT CTGTTGCTGT AAATGCCCGC ATTTTTTGAG AAAAATTTAA ATGCTTGGCC   2520
TAAATTGGAC ATATATGATA GACACCAAGC TGATTTGAAC TTTGTTTATT TTTGACATCC   2580
ATGCATATTG TCAGTGTTGT GAAAACAATA CTAATCCTTT TTTTTTGTCT TTTTCCAGTG   2640
GATCTGAGGA TCCTAGGGGG CCAACAGATT CTTATCAAAT GAATGGGATT ATCAATGAAA   2700
CACATAATGG AAGACATGCC TCAGTGTCCA AGGTTGTTGA ATTTTGTGCG GCACTAGGTG   2760
GCAAAACACC AATTCACAGT ATATTAGTGG CCAACAATGG AATGGCAGCA GCAAAATTTA   2820
TGAGGAGTGT CCGGACATGG GCTAATGATA CTTTTGGATC TGAGAAGGCA ATTCAACTCA   2880
TAGCTATGGC AACTCCGGAA GACATGAGGA TAAATGCAGA ACACATTAGA ATTGCTGACC   2940
AATTNCGTAG AGGTGCCTGG TGGAACAAAC AATAATAACT ACGNCCAATG TTCAACTCAT   3000
AGTGGAGGTT AGCCTTGCTA ATCTGTTAGT TTACTACTGG TCTGCTGTTT CCTTTATTTG   3060
TTGTATAATG ATTGACATAT TTAAGTAGAG AAATTTATAT TTCTCCTCTG CTGTTGTGGA   3120
AGTCCAATTG TCATCATTAA CTGTGAAATA TTGCAGATGG CACAAAAACT AGGTGTTTCT   3180
GCTGTTTGGC CTGGTTGGGG TCATGCTTCT GAGAATCCTG AACTGCCAGA TGCATTGACC   3240
GCAAAAGGGA TCGTTTTTCT TGGCCCACCT GCATCATCAA TGAATGCTTT GGGAGATAAG   3300
GTCGGCTCAG CTCTCATTGC TCAAGCAGCC GGGGTCCCAA CTCTNTGCTT GGAGTGGATC   3360
ACATGTGAGT CTCACTCTTT GATTACTATC CGCCTGTCTC ATTGCTCTCT CTTTCATATT   3420
CTAATGACAC TAAATTTAGG TTGAAGTTCC ATTAGAGTGC TGCTTAGACG CGATACCTGA   3480
GGAGATGTAT AGAAAAGCTT ATCGATACCK TCGACCTCGA GGGGGGGCCC GGTACYAGCT   3540
GSTG                                                                3544
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2166 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
GAATTCCGTG AGCCCTGTAC GGCAATGGCA ACCCCAGGGT TACTGGGGTG GCTGAATGGT     60
CTCGGCTTAC GCAATTGTTT GTGGCAGCTG CGTGGGCTAA ATGTAGGTTG TCTCTTGTTG    120
CACTGCAGGA TGGATGGGTA GCCTCTGGGC CGCCTCTGCT AGTGTCTAGC GTTGCTGACT    180
GTGGTTTATT CAGGGATGCC ATGCCCATGC TAGATTGATA GGTCATAGGT GCCATTCTAA    240
TGGTAGGTGG CGGTAAGGTT TATTAAGCTG TCGTATCAGT AGGTAACCTC ATGAATCAGG    300
GTTTAAGCCC ACCTTCTCCT TTGTGTGGGT GCATAAGGAA TGCACTTGGC TTCGTTCCCT    360
GCTAGTCTTT GCTCATGTGT CATTCTACCA AGTGGGTTAC TGTAACATTG CACTCTATGA    420
TGGTTGGTGG TTGTGCATCT TTTTGCTTCC CCTGGTTGTC TAATACCTGC ATGTAACTGA    480
TGACCTTCTT TTATGTATCA TATAGATTAC ATCCTTTTGT TGTACATCTC AATTCTGAAA    540
AACAATGTTT TGCATTCTTA GCGCTCTGTG CACAAGGAAA AGGAGGTTTT ACCTGCAACT    600
TTTTTTTTCG AGAAAAAACA AACCTTTCTG AAAGGCAGTG ATCATTTAGT ATAAAGAAAA    660
TTTGATTTAC TTTCTTCAGA ARAGAATATT CCAAACAAAC AATTTTCTTA CAGTCTGAGC    720
CACGAAATTT GATCTTGATC TTACTTTCAC AAGCCACATG AAGCCTTATC ATCGCTCTGA    780
TAAAAAAACC AAATAGGTGA TTCATAGAAT GAGAAAAAGA ACCTGTTGCC ATTTGGGGAC    840
CTTGTTGTGT ACTCATTATC CCCCCTGCTC AGGTTGAGGT TTNCCTTGCC ACTGCCACCC    900
CTTGGCCCCT TCTTATACAA CCATCTCCAT TGAAAAAGAT TTTGCACTAC ATTTGGGCTT    960
CGTATAACAA AAAAGGAAAA TAAAACTAAA CAGCAGAAAC ATAGTATAAT TATAGGTAAA   1020
AGGTTCTGGC AAGTTTGAGT GGTAGAGACC TTTGTATATT TGGACATATT TCACTAGTAA   1080
ATAGTTTTCT AAAATGTTCA TGAATGGTGG CCAATAAACT TGATAAGATC TCAACATGGC   1140
AGGTTCCTTC AAAATGAGAG GAAAACTGGA AACATCACAA ATATTTTTTA GCGAGTGGCC   1200
```

TATAAATTAT AATGTTGCTT TCATTTCTTT GATATTCAAA ACTTCCTAAG AGTATTCTGC 1260
TAGAGCTTCT GATGGTGTCT TTTGCCTCTG TCAGATTTTC CAGGAGNTTT TCTTCCCTTT 1320
TTATGGCACT GTGGTTTGAG AAGGTTTCAA TTGTGCTGTC TGGGCCACGG AACGACAATG 1380
TCACAGCTTG GATTAGCCGC AGCTGCCTCA AAGGCCTTGC CACTACTCCC TAATCGCCAG 1440
AGAAGTTCAG CTGGGACTAC ATTCTCATCA TCTTCATTAT CGAGGCCCTT AAACAGAAGG 1500
AAAAGCCGTA CTCGTTCACT CCGTGATGGC GGAGATGGGG TATCAGATGC CAAAAAGCAC 1560
AGCCAGTCTG TTCGTCAAGG TACTGTGAAT ATCTTTTGAT ACAAGCTAAA ATTTTGCTAC 1620
AGAATATATA TTAAAGAGTT CTTTCTTGGC TGGTGTTGTT TATTTGTTTT AACATGCGAA 1680
AGGGCCTCTA GTTGAGTTGG TTAGGTGGCC TGAGTACCAC TCCTTAAGGT CTTGAGTTTG 1740
ATTTTCCGTC AGAGCGAATT TTAGGCTAGG GTTACCCCCC ACCCCCCCAC CCCTACCCGA 1800
ATCTGCACAG TCCGGTCGTG GTCGTCCTCA TATAGGCTAC GATGTCATTG TGTATCGGCG 1860
GGCCAGGGGT TTAAGAGTTT TCTTGACCTT TGTGAGAAGA TCTTAATAAT ACAATGTCCA 1920
AGAGCTGTCT TACCCTGTAG GTCRAGTTTT TTAGTTGTTT TAACATGGTT ATGTTTGAAG 1980
CCTCATTCTA GGTACCAATA TARATATGCT CACTGCTCAG TTTCMAATGT TTGTCTGCAT 2040
KTAGGTCTTG CTGGCATTAT CGACCTCCCA AGTGAGGCAC CTTCCGAAGT GGATATTTCA 2100
CAGTAAGGAN TACAGTATTT TGCGTACGTT TGTTTTGGAA AAAGAAATAT TCTCAGCTTA 2160
TTTAAT 2166

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 484 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAATTCCTGT GGGTGTCATA GCTGTGGAGA CACAGACCAT GATGCAGATC CTGTAATACT 60
TCAAAATCTT AAGCCACAAA ACTTGATTAA TTGTTAGCAC AGTAATTTGC CAAGTGGCTA 120
GAGAAGGATC TCAACACAAC ACAATAACCA AGAGATATCA ATCACAGAGA TGGCACGGTG 180
GTTATCCCGT GGTTCGGCCA AGACCAACGC TTGCCTACTC CACGTTGTGG CGTCCCAACG 240
GACGAGGGTT GCAATCAACC CCTCTCAAGC GGTCCAAAGA CCAACTTGAA TACCACGGTG 300
TTGCTTTGCT TTTCTTAATC CCACTTGCGA GGAATCTCCA CAGCTTGGAG CCTCTCGCCC 360
TTTCAAAAGA TTTCACAAAG AATCACGGAG CAAGGGAAGG ATCAACAACT CACACACGAC 420
ACAAAGATCA CAGTGAATAC GCACACATAA AACCAAGACT TGAGCTCAAG TGACTAGCAC 480
ACTT 484

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 531 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGGAAGTGT GTATTGCCAG TGCTTATTYT RGGGATATGA GGGAATTWAM ATTACATTTG 60
TGACTGGGCG GACTGTAGGA TAGGAGTTAT CTTGYTCGAT TGGTATACGG TGCATACAGS 120
KYTTGACCAG CTATTATTTT AACAGGGTTT TCTGCCCTGA ACAAGTCCTT GGGCGGGAAG 180

```
TGTACAGCTC CCACATGCAG CTTGGTGGTC CTAAGATCAT GGCGACCAAT GGTGTTGTCC    240
ACCTCACTGT TCCAGATGAC CTTGAAGGTG TTTCCAATAT ATTGAGGTGG CTCAGCTATG    300
TTCCTGCAAA CATTGGTGGA CCTCTTCCTA TTACCAAACC TCTGGACCCT CCAGACAGAC    360
CTGTTGCTTA CATCCCTGAG AACACATGCG ATCCACGTGC AGCTATCTGT GGTGTAGATG    420
ACAGCCAAGG GAAATGGTTG GGTGGTATGT TTGACAAAGA CAGCTTTGTG GAGACATTTG    480
AAGGATGGGC AAAAACAGTG GTTACTGGCA GAGCAAAGCT TGGAGGAATT C             531
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 882 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
CTCCCAATAT TGTCATGAGG CTTGCATCCC AGGTTAGTTT TTTTTCCTTT CTGAAATTTA     60
TATTCCATAC CTTTTCACCT TTAGTTATCC TTGTATTTTC TGGAAGCTTC ATCTGATGCA    120
TTATTGACAA ATGCACTAAT GGTCATCATA TTTGSAKAGW YAASATATKT MTCTTAATTG    180
ATGGKWACTM TTGAMWATGA SRATGSTKRA GCAKRTRRYK WAYASTTTTT TAATAAAAAA    240
ACATGCATTT CTAGGAGTTG GACTAAGCTT TTCTTAGTAT GAAGTGCCAT GTTTTACATG    300
GTCCATTTGT GTCAATTTAC AGTCGGTATC ATGGAAAGGT TGTCATAATG GCTGGAGAGA    360
AACAACACAT CTTGTTTCTC AACACTTGTG GGAGAAGATG TTTTACCTTT TTTCCTAAAA    420
TTACTTTTTG TACTAAATTG TATAAKTTTT CCAATATTCT CCATGATTAT TGAACTCTGC    480
TGTGTTCAAA CAGCCAAAAC ATGTTTCCAT ACTTTACACC TTTATTTTTT AGATGGAAGC    540
CTGGAATTGT GCTCTGTTAT CTGTAGTCAT GCATTATATT TGATCTTAAA TCYTATTCTC    600
TATTGTAGAR TCSGCATCTT GAAGTTCAGT TGCTTTGTGA TCAATATGGT AATGTAGCAG    660
CACTTCACAG TCGTGATTGC AGTGTGCAAC GGCGACACCA GAAGGTCTGC CCCTCACCCA    720
CCCAGCCATA AACACGAAGT TTATAGAACC ATGTATTTTG TTATGCAATA TATTTCTCAA    780
TTGTAGCTCC ATTCACATTT TGCTACAACA GATTATTGAA GAAGGTCCAG TTACTGTTGC    840
TNCCTCGTGA GACAGTTAAA GCACTTGAGC AGCAGCAAGG AG                       882
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 867 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
GTCGACCTGC AGGTCAACGG ATCCTAGGGG GCCAACAGAT TCTTATCAAA TGAATGGGAT     60
TATCAATGAA ACACATAATG GAAGACATGC CTCAGTGTCC AAGGTTGTTG AATTTTGTGC    120
GGCACTAGGT GGCAAAACAC CAATTCACAG TATATTAGTG GCCAACAATG GAATGGCAGC    180
AGCAAAATTT ATGAGGAGTG TCCGGACATG GGCTAATGAT ACTTTTGGAT CTGAGAAGGC    240
AATTCAACTC ATAGCTATGG CAACTCCGGA AGACATGAGG ATAAATGCAG AACACATTAG    300
AATTGCTGAC CAATTCGTAG ARGTGCCTGG TGGAACAAAC AATAATAACT ACGCCAATGT    360
TCAACTCATA GTGGAGGTTA GCCTTGCTAA TCTGTTAGTT TACTACTGGT CTGCTGTTTC    420
CTTTATTTGT TGTATAATGA TTGACATATT TAAGTAGAGA AATTTATATT TCTCCTCTGC    480
```

```
TGTTGTGGAA GTCCAATTGT CACCATTAAC TGTGAAATAT TGCAGATGGC ACAAAAACTA    540
GGTGTTTCTG CTGTTTGGCC TGGTTGGGGT CATGCTTCTG AGAATCCTGA ACTGCCAGAT    600
GCATTGACCG CAAAAGGGAT CGTTTTTCTT GGCCCACCTG CATCATCAAT GAATGCTTTG    660
GGAGATAAGK TCGGCTCAGC TCTCATTGCT CAAGCAGCCG GGGTCCCAAC TCTTGCTTGG    720
AGTGGATCAC ATGTGAGTCT CACTCTTTGA TTACTATCCG CCTGTCTCAT TGCTCTCTCT    780
TTCATATTCT AATGACACTA AATTTAGGTT GAAGTTCCAT TAGAGTGCTG CTTAGACGCG    840
ATACCTGAGG AGATGTATAG AAAAGCT                                        867
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
GAATAATCTG CCTGCAGCTC AAGTTGCTGT TGGAATGGGC ATACCTCTTT GGCAGATTCC     60
AGGTAATTAC CAATTTACCA ACTTATTTAG TTCCTTATTG TTTTATTCTC TAATTTTCTA    120
CTTATGTAGA AATCAGACGT TTCTATGGAA TGGACTATGG AGGAGGGTAT GACATTTGGA    180
GGAAAACAGC AGCTCTTGCT ACACCATTTA ATTTTGATGA AGTAGATTCT CAATGGCCAA    240
AGGGCCATTG TGTAGCAGTT AGAATTACTA GTGAGGACCC AGATGATGGT TTCAAACCTA    300
CTGGTGGGAA AGTGAAGGTA AGTTTTCTAG ATGACATGTA TTATATATCG TTCAAAGAGA    360
TTAAGTTTGG TTAAATGACT AGGTCTTGAT TTTTTATCTT TCAGGAGATA AGTTTTAAAA    420
GCAAGCCTAA TGTTTGGGCC TACTTCTCAG TAAAGGTAAC TTGTTAACTT TAGTACGCTG    480
TCACATTATT CTTCGTTGTG AAAATAATTT GAACGGTTCT CTTTGTATTT TAACCATCCA    540
TCGTCTCATT TAGCAGAGCA CACAAATATT TGCACTGACC CCCCTCCCCT TATCTGCTTT    600
CAGTCTGGTG GAGGCATTCA TGAATTTGCT GATTCTCAGT TCGGTATGTG TAAACCAAGA    660
GTATTCTTTG TAATTTATAT TGGTCCTCAA TTTTGAAATA TTGTCTTTCC GTTACAGGAC    720
ADG                                                                  723
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 231 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
AATTCCTGTG GGTGTTATAG CTGTGGAGAC ACAGACCATG ATGCAGCTCA TCCCTGCTGA     60
TCCAGGTCAA CTTGATTCCC ATGAGCGATG TGTTCCTCGG GCTGGACAAG TGTGGTTCCC    120
AGATNCTGCA ACCAAGACAG CTCAGGCATT ATTAGACTTC AACCGTGAAG GATTGCCTCT    180
GTTCATCCTG GCTAACTGGA GAGGCTTCTC TGGGGGACAG AGAGATCTCT T              231
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 207 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AATTCATGCA TCTTAATAAA CACAGTTGGC CCTTAAAGCA AGTGAACTTC TTGAACAAAC 60
CAAACTAAGT GAACTCTGTT CCAGCATTGC AAGAAGCCTT TCAGATCTGG GGATGCATAA 120
GGGAGAAATG ACTATTAAGG ATAGCATGGA AGATTTAGTC TCTGNCCCAT TGCCTGTTGA 180
AGATGCTCTT ATTTCTTTGT TTGATTA 207

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 180 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATAGACCTGT CGCATACATC CCTGAGAACA CATGCGATCC GCGTGCAGCC ATCCGTGGNG 60
TAGATGACAG CCAAGGGAAA TGGTTGGGTG GTATGTTTGA CAAAGACAGC TTTGTGGAGA 120
CATTTGAAGG ATGGGCAAAA ACAGTGGTTA CTGGTAGAGC AAAGCTTGGA GGAAGGAATT 180

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCTTTTTATG GCACTGTGCG 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CATCGTAGCC TATATGAGGA CG 22

WHAT IS CLAIMED IS:

1. A method of imparting cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance to a corn plant comprising:

crossing a first corn plant with a second corn plant so as to yield a progeny plant, wherein the first plant is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, wherein the second plant is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and wherein the progeny plant is heterozygous for the *Acc1* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and heterozygous for the *Acc2* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

2. A method of imparting cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance to a corn plant comprising:

(a) self pollinating a corn plant which comprises (i) an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and (ii) an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, so as to yield a progeny plant; and

(b) identifying a progeny plant that is homozygous for the allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and is homozygous for the allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

3. A method of preparing a plant which is a double heterozygote for alleles of *Acc1* and *Acc2* which impart cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, comprising:

crossing a first corn plant with a second corn plant so as to yield a progeny plant, wherein the first plant is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, wherein the second plant is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and wherein the progeny plant is heterozygous for the *Acc1* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and heterozygous for the *Acc2* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

4. A method of preparing a plant which is a double homozygote for alleles of *Acc1* and *Acc2* which impart cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance comprising:

(a) self pollinating a corn plant which comprises (i) an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and (ii) an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, so as to yield a progeny plant; and

(b) identifying a progeny plant that is homozygous for the allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and is homozygous for the allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

5. The method of claim 2 or 4 wherein an inbred plant is self pollinated in step (a).

6. The method of claim 1 or 3 wherein the first and second plant are substantially isogenic.

7. The method of claim 1 or 3 further comprising crossing the progeny plant to an inbred plant so as to yield a further progeny plant.

8. The method of claim 7 wherein the further progeny plant is crossed to an inbred so as to yield a progeny plant that is substantially isogenic with the inbred plant.

9. The method of claim 2 or 4 wherein the progeny plant is identified by RFLP analysis.

10. The method of claim 2 or 4 wherein the *Acc1* alleles of the progeny plant are identified by an amplification reaction employing a polymerase chain reaction.

11. The method of claim 1, 2, 3, or 4 wherein the progeny plant is tolerant to levels of a cyclohexanedione or an aryloxyphenoxypropanoic acid herbicide which inhibit a corresponding susceptible plant.

12. The method of claim 1, 2, 3, or 4 wherein the herbicide is selected from the group consisting of sethoxydim, haloxyfop, and mixtures thereof.

13. The method of claim 1, 2, 3, or 4 wherein the *Acc1* allele imparts cyclohexanedione herbicide tolerance.

14. The method of claim 13 wherein the *Acc1* allele is selected from the group consisting of *Acc1-S1*, *Acc1-S2*, *Acc1-S3* and *Acc1-S4*.

15. The method of claim 1, 2, 3, or 4 wherein the *Acc2* allele imparts cyclohexanedione herbicide tolerance.

16. The method of claim 17 wherein the *Acc2* allele is selected from the group consisting of *Acc2-S5* and *Acc2-S6*.

17. The method of claim 1, 2, 3, or 4 comprising obtaining seed from said progeny and obtaining further progeny plants.

18. The method of claim 17 wherein the progeny obtained are crossed back to a corn plant, to obtain further progeny.

19. The method of claim 18 wherein seeds are obtained from further said progeny plants and plants are recovered from said seed.

20. The method of claim 17 wherein said further progeny are crossed back to a corn plant, and progeny are obtained.

21. A progeny plant prepared by the method of claim 1, 2, 3, or 4.

22. The progeny plant of claim 21 which is an inbred plant.

23. A method of obtaining a hybrid herbicide resistant plant which is a double heterozygote for *Acc1* and *Acc2* comprising crossing the progeny plant of claim 22 with a member of another inbred corn line so as to yield a hybrid corn plant.

24. A seed of the plant of claim 21.

25. An inbred cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerant *Zea mays* plant, the genome of which is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

26. An inbred cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerant *Zea mays* plant, the genome of which is heterozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and is heterozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

27. The plant of claim 25 or 26 wherein the plant is resistant or tolerant to an amount of herbicide that is toxic to a corresponding susceptible *Zea mays* plant.

28. The plant of claim 25 or 26 which is resistant or tolerant to an herbicide is selected from the group consisting of sethoxydim, haloxyfop, and mixtures thereof.

29. The plant of claim 25 or 26 wherein the *Acc1* allele is selected from the group consisting of *Acc1-S1, Acc1-S2, Acc1-S3* and *Acc1-S4*.

30. The plant of claim 25 or 26 wherein the *Acc2* allele is selected from the group consisting of *Acc2-S5* and *Acc2-S6*.

31. Seed derived from the plant of claim 25 or 26.

32. A progeny plant derived from the plant of claim 25 or 26.

33. A hybrid plant derived from the plant of claim 25 or 26.

34. A seed of the plant of claim 32.

35. A method to prepare an herbicide resistant or tolerant corn plant, comprising: crossing a first corn plant to a second corn plant so as to yield progeny plant, wherein the first corn plant comprises at least one herbicide resistant allele and the second plant comprises at least one herbicide resistant allele which is not allelic to the herbicide resistant allele present in the first plant.

36. A method of imparting cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance to a hybrid corn plant comprising:
crossing a first inbred corn plant with a second inbred corn plant so as to yield a progeny hybrid plant, wherein the first plant is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, wherein the second plant is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and wherein the progeny plant is heterozygous for the *Acc1* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and heterozygous for the *Acc2* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

37. A method of preparing a hybrid plant which is a double heterozygote for alleles of *Acc1* and *Acc2* which impart cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, comprising:
crossing a first inbred corn plant with a second inbred corn plant so as to yield a progeny hybrid plant, wherein the first plant is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, wherein the second

plant is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and wherein the progeny plant is heterozygous for the *Acc1* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance and heterozygous for the *Acc2* allele which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

38. The method of claim 35, 36 or 37 wherein the progeny plant is tolerant to levels of a cyclohexanedione or an aryloxyphenoxypropanoic acid herbicide which inhibit a corresponding susceptible plant.

39. The method of claim 35, 36 or 37 wherein the herbicide is selected from the group consisting of sethoxydim, haloxyfop, and mixtures thereof.

40. The method of claim 36 or 37 wherein the *Acc1* allele imparts cyclohexanedione herbicide tolerance.

41. The method of claim 35 wherein the herbicide resistant allele in the first plant is selected from the group consisting of *Acc1-S1, Acc1-S2, Acc1-S3* and *Acc1-S4*.

42. The method of claim 35 wherein the herbicide resistant allele in the second plant is selected from the group consisting of *Acc2-S5* and *Acc2-S6*.

43. The method of claim 36 or 37 wherein the *Acc1* allele imparts cyclohexanedione herbicide tolerance.

44. The method of claim 43 wherein the *Acc1* allele is selected from the group consisting of *Acc1-S1, Acc1-S2, Acc1-S3* and *Acc1-S4*.

45. The method of claim 36 or 37 wherein the *Acc2* allele imparts cyclohexanedione herbicide tolerance.

46. The method of claim 45 wherein the *Acc2* allele is selected from the group consisting of *Acc2-S5* and *Acc2-S6*.

47. An hybrid cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerant *Zea mays* plant, the genome of which is homozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and is homozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

48. An hybrid cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerant *Zea mays* plant, the genome of which is heterozygous for an allele of *Acc1* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance, and is heterozygous for an allele of *Acc2* which imparts cyclohexanedione or aryloxyphenoxypropanoic acid herbicide tolerance.

49. The plant of claim 47 or 48 wherein the plant is resistant or tolerant to an amount of herbicide that is toxic to a corresponding susceptible *Zea mays* plant.

50. The plant of claim 47 or 48 which is resistant or tolerant to an herbicide is selected from the group consisting of sethoxydim, haloxyfop, and mixtures thereof.

51. The plant of claim 47 or 48 wherein the *Acc1* allele is selected from the group consisting of *Acc1-S1*, *Acc1-S2*, *Acc1-S3* and *Acc1-S4*.

52. The plant of claim 47 or 48 wherein the *Acc2* allele is selected from the group consisting of *Acc2-S5* and *Acc2-S6*.

53. A method of imparting tolerance to a corn plant to an agent which inhibits acetyl CoA carboxylase, comprising:

crossing a first corn plant with a second corn plant so as to yield a progeny plant, wherein the first plant is homozygous for an allele of *Acc1* which imparts tolerance to the agent, wherein the second plant is homozygous for an allele of *Acc2* which imparts tolerance to the agent, wherein the progeny plant is heterozygous for the *Acc1* allele which imparts tolerance to the agent and heterozygous for the *Acc2* allele which imparts tolerance to the agent, and wherein the agent is selected from the group consisting of 3-(2,4-dichlorophenyl)-perhydroindolizine-2,4-dione, 3-isopropyl-6-(N-[2,2-dimethylpropyl]-acetamido-1,3,5-triazine-2,4-(1H,3H)dione, soraphen A, and structural analogs thereof.

54. A method of imparting tolerance to a corn plant to an agent which inhibits acetyl CoA carboxylase comprising:

(a) self pollinating a corn plant which comprises (i) an allele of *Acc1* which imparts tolerance to the agent and (ii) an allele of *Acc2* which imparts tolerance to the agent, so as to yield a progeny plant, wherein the agent is selected from the group consisting of 3-(2,4-dichlorophenyl)-perhydroindolizine-2,4-dione, 3-isopropyl-6-(N-[2,2-dimethylpropyl]-acetamido-1,3,5-triazine-2,4-(1H,3H)dione, soraphen A, and structural analogs thereof; and

(b) identifying a progeny plant that is homozygous for the allele of *Acc1* which imparts tolerance to the agent and is

homozygous for the allele of *Acc2* which imparts tolerance to the agent.

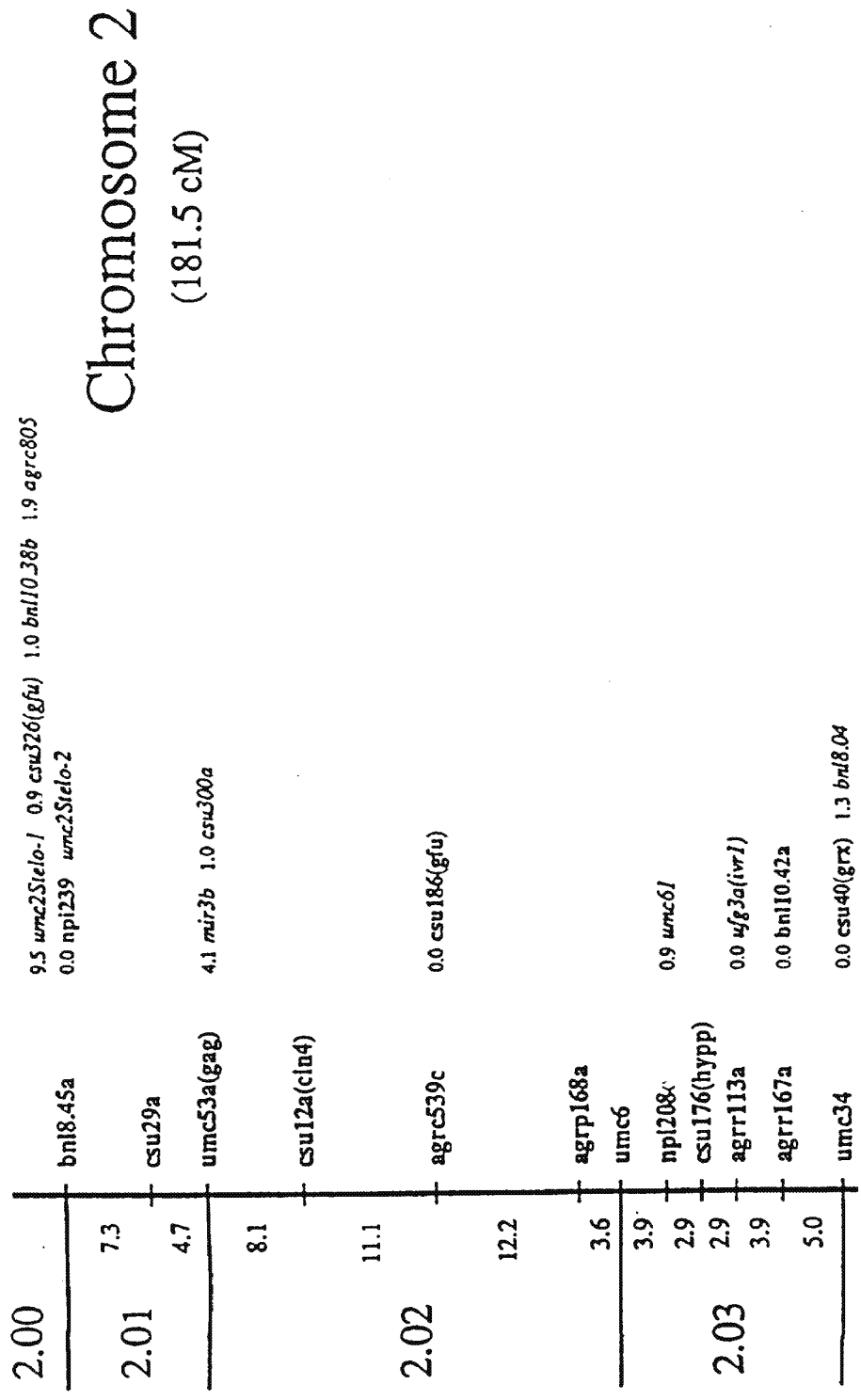


FIG. 1A

SUBSTITUTE SHEET (RULE 26)

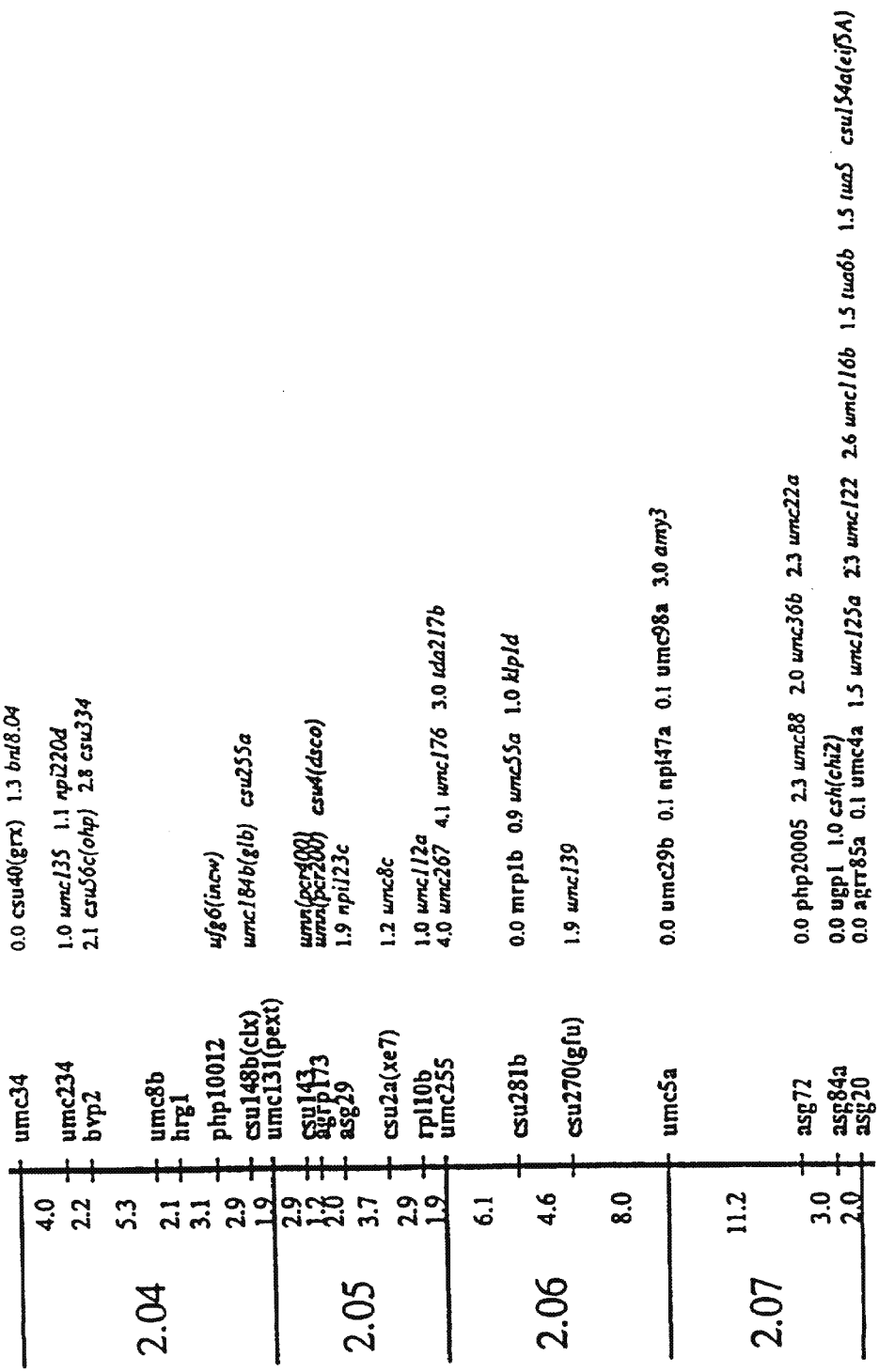


FIG. 1B

SUBSTITUTE SHEET (RULE 28)

COR-AAD1-00001763

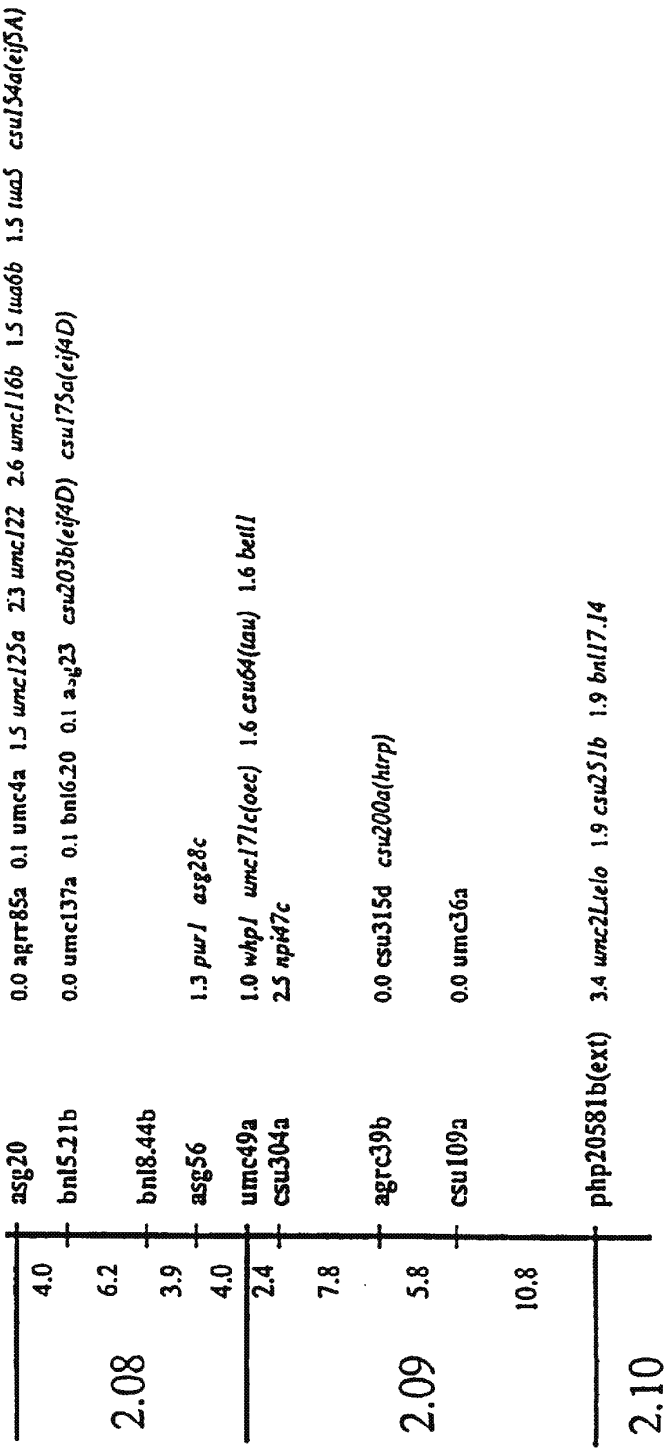


FIG. 1C

SUBSTITUTE SHEET (RULE 26)

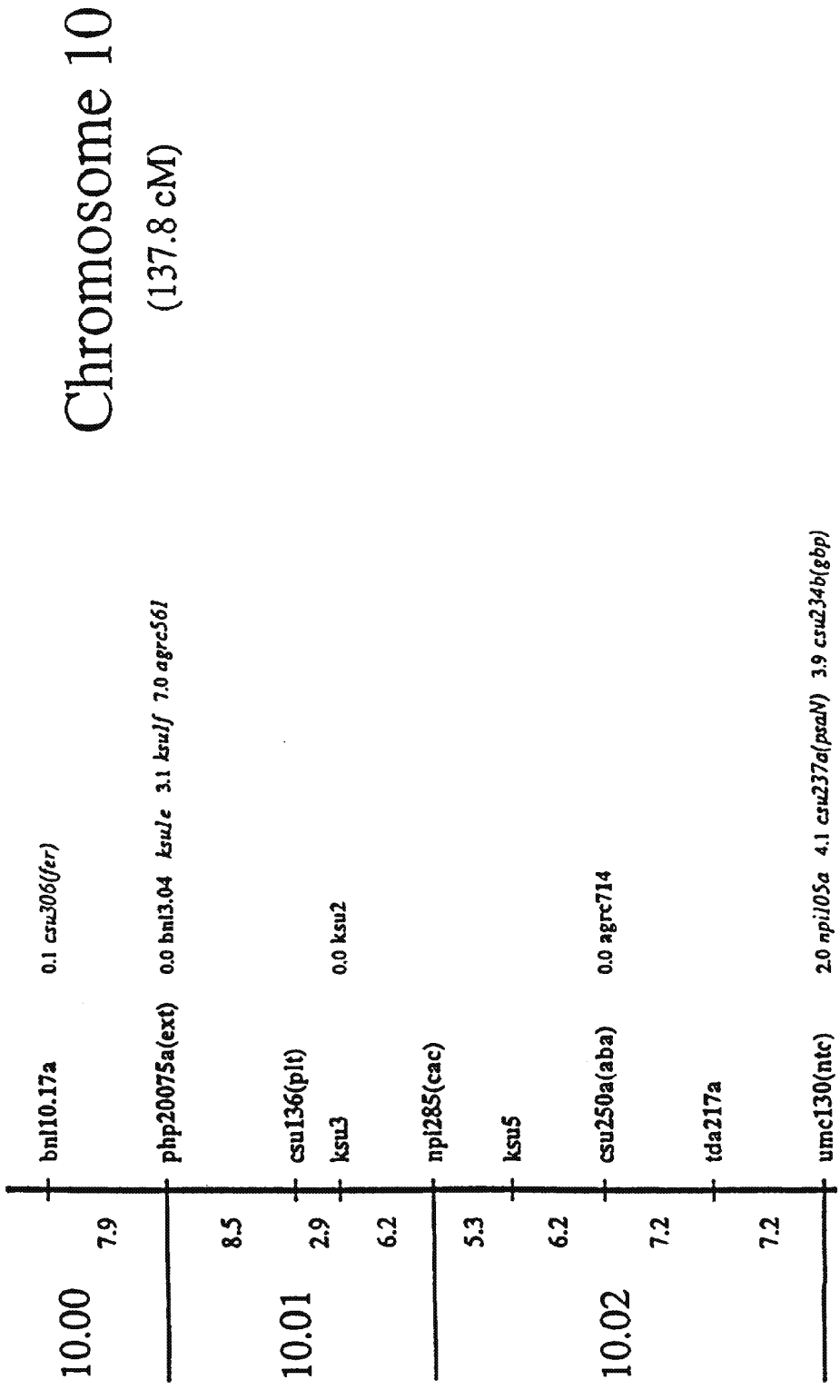


FIG. 2A

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001765

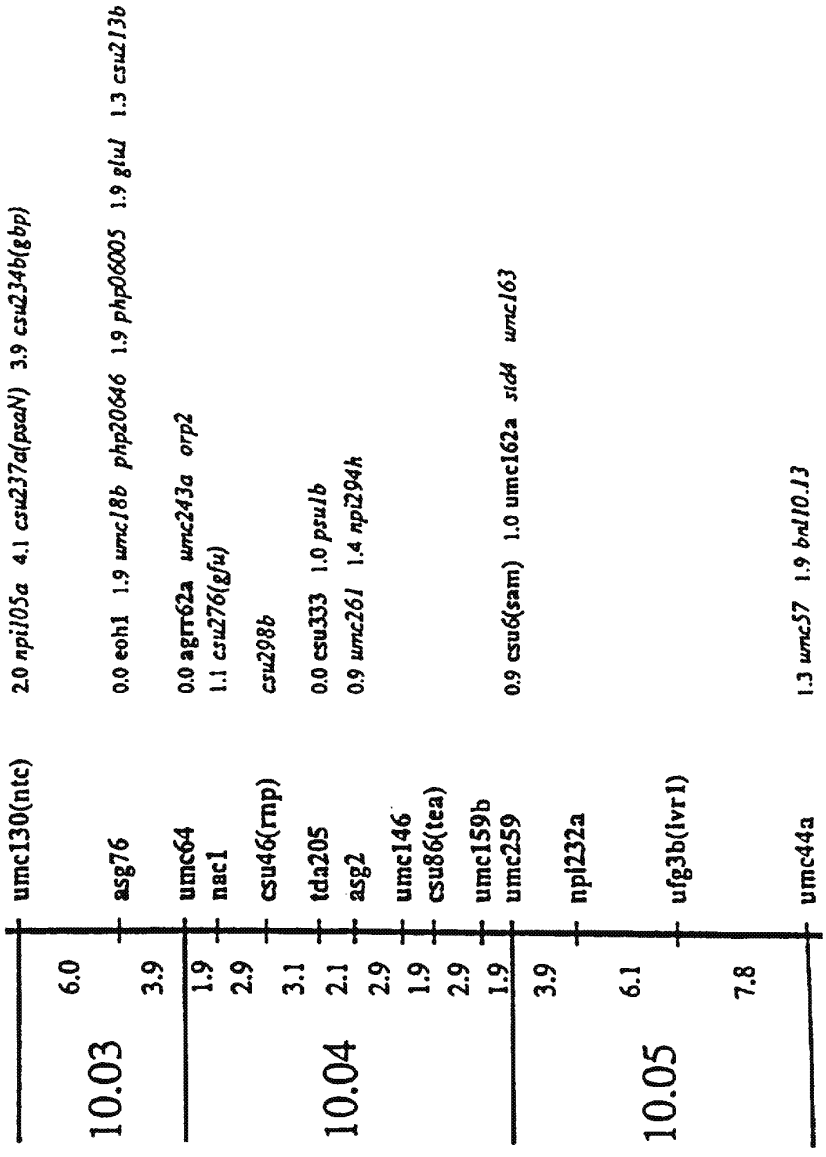


FIG. 2B

SUBSTITUTE SHEET (RULE 26)

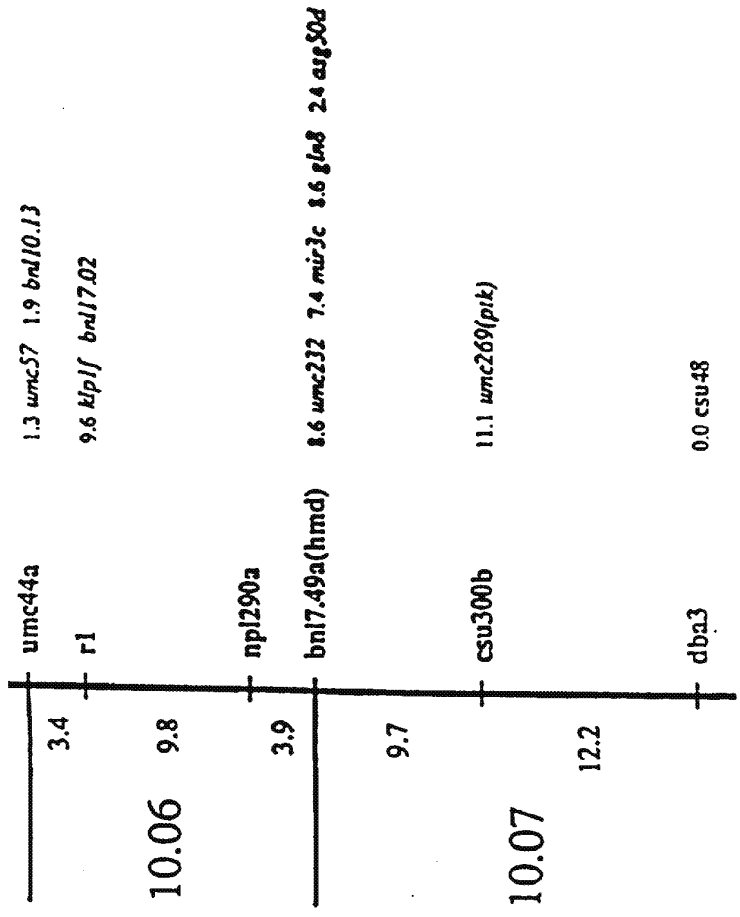


FIG. 2C

COR-AAD1-00001767

```
   1  GGTCTTCAAT TGTGCTGTCT GGGCCACGGA ACGACAATGT CACAGCTTGG
  51  ATTAGCCGCA GCTGCCTCAA AGGCCTTGCC ACTACTCCCT AATCGCCAGA
 101  GAAGTTCAGC TGGGACTACA TTCTCATCAT CTTCATTATC GAGGCCCTTA
 151  AACAGAAGGA AAAGCCATAC TCGTTCACTC CGTGATGGCG GAGATGGGGT
 201  ATCAGATGCC AAAAAGCACA GCCAGTCTGT TCGTCAAGGT CTTGCTGGCA
 251  TTATCGACCT CCCAAGTGAG GCACCTTCCG AAGTGGATAT TTCACATGGA
 301  TCTGAGGATC CTAGGGGGCC AACAGATTCT TATCAAATGA ATGGGATTAT
 351  CAATGAAACA CATAATGGAA GACATGCCTC AGTGTCCAAG GTTGTTGAAT
 401  TTTGTGCGGC ACTAGGTGGC AAAACACCAA TTCACAGTAT ATTAGTGGCC
 451  AACAATGGAA TGGCAGCAGC AAAATTTATG AGGAGTGTCC GGACATGGGC
 501  TAATGATACT TTTGGATCTG AGAAGGCAAT TCAACTCATA GCTATGGCAA
 551  CTCCGGAAGA CATGAGGATA AATGCAGAAC ACATTAGAAT TGCTGACCAA
 601  TTCGTAGAGG TGCCTGGTGG AACAAACAAT AATAACTACG CCAATGTTCA
 651  ACTCATAGTG GGGATGGCAC AAAAACTAGG TGTTTCTGCT GTTTGGCCTG
 701  GTTGGGGTCA TGCTTCTGAG AATCCTGAAC TGCCAGATGC ATTGACCGCA
 751  AAAGGGATCG TTTTTCTTGG CCCACCTGCA TCATCAATGA ATGCTTTGGG
 801  AGATAAGGTC GGCTCAGCTC TCATTGCTCA AGCAGCCGGG GTCCCAACTC
 851  TTGCTTGGAG TGGATCACAT GTTGAAGTTC CATTAGAGTG CTGCTTAGAC
 901  GCGATACCTG AGGAGATGTA TAGAAAAGCT TGCGTTACTA CCACAGAGGA
 951  AGCAGTTGCA AGTTGTCAAG TGGTTGGTTA TCCTGCCATG ATTAAGGCAT
1001  CCTGGGGAGG TGGTGGTAAA GGAATAAGAA AGGTTCATAA TGATGATGAG
1051  GTTAGAGCGC TGTTTAAGCA AGTACAAGGT GAAGTCCCTG GCTCCCCAAT
1101  ATTTGTCATG AGGCTTGCAT CCCAGAGTCG GCATCTTGAA GTTCAGTTGC
1151  TTTGTGATCA ATATGGTAAT GTAGCAGCAC TTCACAGTCG TGATTGCAGT
1201  GTGCAACGGC GACACCAGAA GATTATTGAA GAAGGTCCAG TTACTGTTGC
1251  TCCTCGTGAG ACAGTTAAAG CACTTGAGCA GGCAGCAAGG AGGCTTGCTA
1301  AGGCTGTGGG TTATGTTGGT GCTGCTACTG TTGAGTATCT TTACAGCATG
1351  GAAACTGGAG ACTACTATTT TCTGGAACTT AATCCCCGAC TACAGGTTGA
1401  GCATCCAGTC ACTGAGTGGA TAGCTGAAGT GAATCTGCCT GCAGCTCAAG
```

FIG. 3A 7/25

SUBSTITUTE SHEET (RULE 26)

```
1451  TTGCTGTTGG AATGGGCATA CCTCTTTGGC AGATTCCAGA AATCAGACGT
1501  TTCTATGGAA TGGACTATGG AGGAGGGTAT GACATTTGGA GGAAAACAGC
1551  AGCTCTTGCT ACACCATTTA ATTTTGATGA AGTAGATTCT CAATGGCCAA
1601  AGGGCCATTG TGTAGCAGTT AGAATTACTA GTGAGGACCC AGATGATGGT
1651  TTCAAACCTA CTGGTGGGAA AGTGAAGGAG ATAAGTTTTA AAAGCAAGCC
1701  TAATGTTTGG GCCTACTTCT CAGTAAAGTC TGGTGGAGGC ATTCATGAAT
1751  TTGCTGATTC TCAGTTTGGA CATGCTTTTG CATATGGACT CTCTAGACCA
1801  GCAGCTATAA CAAACATGTC TCTTGCATTA AAAGAGATTC AGATTCGTGG
1851  AGAAATTCAT TCAAATGTTG ATTACACAGT TGACCTCTTA AACGCTTCAG
1901  ACTTCAGAGA AAACAAGATC CACACTGGTT GGCTGGATAC AAGAATAGCT
1951  ATGCGTGTTC AAGCTGAGAG GCCCCCATGG TATATCTCAG TGGTTGGAGG
2001  TGCTTTATAT AAAACAGTAA CCACCAATGC AGCCACTGTT TCTGAATATG
2051  TTAGTTATCT CACCAAGGGC CATATTCCAC CAAAGCATAT ATCCCTTGTC
2101  AATTCTACAG TTAATTTGAA TATAGAAGGG AGCAAATACA CAATTGAAAC
2151  TGTAAGGACT GGACATGGTA GCTACAGGTT GAGAATGAAT GATTCAACAG
2201  TTGAAGCGAA TGTACAATCT TTATGTGATG GTGGCCTCTT AATGCAGTTG
2251  GATGGAAACA GCCATGTAAT TTATGCAGAA GAAGAAGCTG GTGGTACACG
2301  GCTTCAGATT GATGGAAAGA CATGTTTATT GCAGAATGAC CATGATCCAT
2351  CGAAGTTATT AGCTGAGACA CCCTGCAAAC TTCTTCGTTT CTTGGTTGCT
2401  GATGGTGCTC ATGTTGATGC GGATGTACCA TACGCGGAAG TTGAGGTTAT
2451  GAAGATGTGC ATGCCTCTCT TGTCACCTGC TTCTGGTGTC ATTCATTGTA
2501  TGATGTCTGA GGGCCAGGCA TTGCAGGCTG GTGATCTTAT AGCAAGGTTG
2551  GATCTTGATG ACCCTTCTGC TGTGAAAAGA GCTGAGCCAT TTGATGGAAT
2601  ATTTCCACAA ATGGAGCTCC CTGTTGCTGT CTCTAGTCAA GTACACAAAA
2651  GATATGCTGC AAGTTTGAAT GCTGCTCGAA TGGTCCTTGC AGGATATGAG
2701  CACAATATTA ATGAAGTCGT TCAAGATTTG GTATGCTGCC TGGACAACCC
2751  TGAGCTTCCT TTCCTACAGT GGGATGAACT TATGTCTGTT CTAGCAACGA
2801  GGCTTCCAAG AAATCTCAAG AGTGAGTTAG AGGATAAATA CAAGGAATAC
2851  AAGTTGAATT TTTACCATGG AAAAAACGAG GACTTTCCAT CCAAGTTGCT
2901  AAGAGACATC ATTGAGGAAA ATCTTTCTTA TGGTTCAGAG AAGGAAAAGG
```

FIG. 3B 8/25

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001769

```
2951  CTACAAATGA GAGGCTTGTT GAGCCTCTTA TGAACCTACT GAAGTCATAT
3001  GAGGGTGGGA GAGAGAGCCA TGCACATTTT GTTGTCAAGT CTCTTTTCGA
3051  GGAGTATCTT ACAGTGGAAG AACTTTTTAG TGATGGCATT CAGTCTGACG
3101  TGATTGAAAC ATTGCGGCAT CAGCACAGTA AAGACCTGCA GAAGGTTGTA
3151  GACATTGTGT TGTCTCACCA GGGTGTGAGG AACAAAGCTA AGCTTGTAAC
3201  GGCACTTATG GAAAAGCTGG TTTATCCAAA TCCTGGTGGT TACAGGGATC
3251  TGTTAGTTCG CTTTTCTTCC CTCAATCATA AAAGATATTA TAAGTTGGCC
3301  CTTAAAGCAA GTGAACTTCT TGAACAAACC AAACTAAGTG AACTCCGTGC
3351  AAGCGTTGCA AGAAGCCTTT CGGATCTGGG GATGCATAAG GGAGAAATGA
3401  GTATTAAGGA TAACATGGAA GATTTAGTCT CTGCCCCATT ACCTGTTGAA
3451  GATGCTCTGA TTTCTTTGTT TGATTACAGT GATCGAACTG TTCAGCAGAA
3501  AGTGATTGAG ACATACATAT CACGATTGTA CCAGCCTCAT CTTGTAAAGG
3551  ATAGCATCCA AATGAAATTC AAGGAATCTG GTGCTATTAC TTTTTGGGAA
3601  TTTTATGAAG GGCATGTTGA TACTAGAAAT GGACATGGGG CTATTATTGG
3651  TGGGAAGCGA TGGGGTGCCA TGGTCGTTCT CAAATCACTT GAATCTGCGT
3701  CAACAGCCAT TGTGGCTGCA TTAAAGGATT CGGCACAGTT CAACAGCTCT
3751  GAGGGCAACA TGATGCACAT TGCATTATTG AGTGCTGAAA ATGAAAGTAA
3801  TATAAGTGGA ATAAGCAGTG ATGATCAAGC TCAACATAAG ATGGAAAAGC
3851  TTAGCAAGAT ACTGAAGGAT ACTAGCGTTG CAAGTGATCT CCAAGCTGCT
3901  GGTTTGAAGG TTATAAGTTG CATTGTTCAA AGAGATGAAG CTCGCATGCC
3951  AATGCGCCAC ACATTCCTCT GGTTGGATGA CAAGAGTTGT TATGAAGAAG
4001  AGCAGATTCT CCGGCATGTG GAGCCTCCCC TCTCTACACT TCTTGAATTG
4051  GATAAGTTGA AGGTGAAAGG ATACAATGAA ATGAAGTATA CTCCTTCGCG
4101  TGACCGCCAA TGGCATATCT ACACACTAAG AAATACTGAA AACCCCAAAA
4151  TGTTGCATAG GGTGTTTTTC CGAACTATTG TCAGGCAACC CAATGCAGGC
4201  AACAAGTTTA GATCGGCTCA GATCAGCGAC GCTGAGGTAG GATGTCCCGA
4251  AGAATCTCTT TCATTTACAT CAAATAGCAT CTTAAGATCA TTGATGACTG
4301  CTATTGAAGA ATTAGAGCTT CATGCAATTA GGACAGGTGA TTCTCACATG
4351  TATTTGTGCA TACTGAAAGA GCAAAAGCTT CTTGACCTCA TTCCATTTTC
4401  AGGGAGTACA ATTGTTGATG TTGGCCAAGA TGAAGCTACC GCTTGTTCAC
```

FIG. 3C 9/25

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001770

```
4451  TTTTAAAATC AATGGCTTTG AAGATACATG AGCTTGTTGG TGCAAGGATG
4501  CATCATCTGT CTGTATGCCA GTGGGAGGTG AAACTCAAGT TGGACTGTGA
4551  TGGCCCTGCA AGTGGTACCT GGAGAGTTGT AACTACAAAT GTTACTGGTC
4601  ACACCTGCAC CATTGATATA TACCGAGAAG TGGAGGAAAT AGAATCACAG
4651  AAGTTAGTGT ACCATTCAGC CAGTTCGTCA GCTGGACCAT TGCATGGTGT
4701  TGCACTGAAT AATCCATATC AACCTTTGAG TGTGATTGAT CTAAAGCGCT
4751  GCTCTGCTAG GAACAACAGA ACAACATATT GCTATGATTT TCCGCTGGCC
4801  TTTGAAACTG CACTGCAGAA GTCATGGCAG TCCAATGGCT CTACTGTTTC
4851  TGAAGGCAAT GAAAATAGTA AATCCTACGT GAAGGCAACT GAGCTAGTGT
4901  TTGCTGAAAA ACATGGGTCC TGGGGCACTC CTATAATTCC GATGGAACGC
4951  CCTGCTGGGC TCAACGACAT TGGTATGGTC GCTTGGATCA TGGAGATGTC
5001  AACACCTGAA TTTCCCAATG GCAGGCAGAT TATTGTTGTA GCAAATGATA
5051  TCACTTTCAG AGCTGGATCA TTTGGCCCAA GGGAAGATGC ATTTTTTGAA
5101  ACTGTCACTA ACCTGGCTTG CGAAAGGAAA CTTCCTCTTA TATACTTGGC
5151  AGCAAACTCT GGTGCTAGGA TTGGCATAGC TGATGAAGTA AAATCTTGCT
5201  TCCGTGTTGG ATGGTCTGAC GAAGGCAGTC CTGAACGAGG GTTTCAGTAC
5251  ATCTATCTGA CTGAAGAAGA CTATGCTCGC ATTAGCTCTT CTGTTATAGC
5301  ACATAAGCTG GAGCTAGATA GTGGTGAAAT TAGGTGGATT ATTGACTCTG
5351  TTGTGGGCAA GGAGGATGGG CTTGGTGTCG AGAACATACA TGGAAGTGCT
5401  GCTATTGCCA GTGCTTATTC TAGGGCATAT GAGGAGACAT TTACACTTAC
5451  ATTTGTGACT GGGCGGACTG TAGGAATAGG AGCTTATCTT GCTCGACTTG
5501  GTATACGGTG CATACAGCGT CTTGACCAGC CTATTATTTT AACAGGGTTT
5551  TCTGCCCTGA ACAAGCTCCT TGGGCGGGAA GTGTACAGCT CCCACATGCA
5601  GCTTGGTGGT CCTAAGATCA TGGCGACCAA TGGTGTTGTC CACCTCACTG
5651  TTCCAGATGT CCTTGAAGGT GTTTCCAATA TATTGAGGTG GCTCAGCTAT
5701  GTTCCTGCAA ACATTGGTGG ACCTCTTCCT ATTACCAAAC CTCTGGACCC
5751  TCCAGACAGA CCTGTTGCTT ACATCCCTGA GAACACATGC GATCCACGTG
5801  CAGCTATCTG TGGTGTAGAT GACAGCCAAG GGAAATGGTT GGGTGGTATG
5851  TTTGACAAAG ACAGCTTTGT GGAGACATTT GAAGGATGGG CAAAAACAGT
5901  GGTTACTGGC AGAGCAAAGC TTGGAGGAAT TCCTGTGGGC GTCATAGCTG
```

FIG. 3D 10/25

SUBSTITUTE SHEET (RULE 26)

```
5951  TGGAGACACA GACCATGATG CAGATCATCC CTGCTGATC  AGGTCAGCTT
6001  GATTCCCATG AGCGATCTGT CCCTCGTGCT GGACAAGTGT GGTTCCCAGA
6051  TTCTGCAACC AAGACCGCTC AGGCATTATT AGACTTCAAC CGTGAAGGAT
6101  TGCCTCTGTT CATCCTGGCT AATTGGAGAG GCTTCTCTGG TGGACAAAGA
6151  GATCTCTTTG AAGGAATTCT TCAGGCTGGG TCAACAATTG TCGAGAACCT
6201  TAGGACATAT AATCAGCCTG CTTTTGTGTA CATTCCTATG GCTGGAGAGC
6251  TTCGTGGAGG AGCTTGGGTT GTGGTCGATA GCAAAATAAA TCCAGACCGC
6301  ATTGAGTGTT ATGCTGAAAG GACTGCCAAA GGTAATGTTC TCGAACCTCA
6351  AGGGTTAATT GAAATCAAGT TCAGGTCAGA GGAACTCCAA GACTGTATGG
6401  GTAGGCTTGA CCCAGAGTTG ATAAATCTGA AAGCAAAACT CCAAGATGTA
6451  AATCATGGAA ATGGAAGTCT ACCAGACATA GAAGGGATTC GGAAGAGTAT
6501  AGAAGCACGT ACGAAACAGT TGCTGCCTTT ATATACCCAG ATTGCAATAC
6551  GGTTTGCTGA ATTGCATGAT ACTTCCCTAA GAATGGCAGC TAAAGGTGTG
6601  ATTAAGAAAG TTGTAGACTG GGAAGAATCA CGCTCGTTCT TCTATAAAAG
6651  GCTACGGAGG AGGATCGCAG AAGATGTTCT TGCAAAAGAA ATAAGGCAGA
6701  TAGTCGGTGA TAAATTTACG CACCAATTAG CAATGGAGCT CATCAAGGAA
6751  TGGTACCTTG CTTCTCAGGC CACAACAGGA AGCACTGGAT GGGATGACGA
6801  TGATGCTTTT GTTGCCTGGA AGGACAGTCC TGAAAACTAC AAGGGGCATA
6851  TCCAAAAGCT TAGGGCTCAA AAAGTGTCTC ATTCGCTCTC TGATCTTGCT
6901  GACTCCAGTT CAGATCTGCA AGCATTCTCG CAGGGTCTTT CTACGCTATT
6951  AGATAAGATG GATCCCTCTC AGAGAGCGAA GTTTGTTCAG GAAGTCAAGA
7001  AGGTCCTTGA TTGATGATAC CAACACATCC AACACAATGT GTGCATGTCA
7051  CATCTTTTTG TTCTAGTACA TACATAGAAG GATATTGCTT GGTCTTGATT
7101  GATCATGTCT GATTTAAGTC GACTATTATT TCTTGGAATT TTCTTTTGGA
7151  CCTGGTGCTA TGGTTGATGG ATGTATATTG GATATGTGCG TTCTGCCAGG
7201  TGTAAGCACA AAGGTTTAGA CARAMMRARA RCAAGAGCGA GTGAACCTGT
7251  TCTGGTTTTG CAGTGGTTCA GTAAGGCAGA AAGTTGTTAA ACCGTAGTTC
7301  TGAGATGTAT TACCAGTGNC GCCATGCTGT ACTTTTAGGG TGTATAATGC
7351  GGATACAAAT AAACAATTTA GCGGTTCATT AAAGTTTGAA CTCAAATAAC
7401  ATGTTCTTTG TAAGCATATG TACCGTACCT CTACGTGAAA TAAAGTTGTT
7451  GAATTAGCAT TCGAAAAAAA
```

FIG. 3E

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001772

```
   1  MSQLGLAAAA SKALPLLPNR QRSSAGTTFS SSSLSRPLNR RKSHTRSLRD
  51  GGDGVSDAKK HSQSVRQGLA GIIDLPSEAP SEVDISHGSE DPRGPTDSYQ
 101  MNGIINETHN GRHASVSKVV EFCAALGGKT PIHSILVANN GMAAAKFMRS
 151  VRTWANDTFG SEKAIQLIAM ATPEDMRINA EHIRIADQFV EVPGGTNNNN
 201  YANVQLIVGM AQKLGVSAVW PGWGHASENP ELPDALTAKG IVFLGPPASS
 251  MNALGDKVGS ALIAQAAGVP TLAWSGSHVE VPLECCLDAI PEEMYRKACV
 301  TTTEEAVASC QVVGYPAMIK ASWGGGGKGI RKVHNDDEVR ALFKQVQGEV
 351  PGSPIFVMRL ASQSRHLEVQ LLCDQYGNVA ALHSRDCSVQ RRHQKIIEEG
 401  PVTVAPRETV KALEQAARRL AKAVGYVGAA TVEYLYSMET GDYYFLELNP
 451  RLQVEHPVTE WIAEVNLPAA QVAVGMGIPL WQIPEIRRFY GMDYGGGYDI
 501  WRKTAALATP FNFDEVDSQW PKGHCVAVRI TSEDPDDGFK PTGGKVKEIS
 551  FKSKPNVWAY FSVKSGGGIH EFADSQFGHA FAYGLSRPAA ITNMSLALKE
 601  IQIRGEIHSN VDYTVDLLNA SDFRENKIHT GWLDTRIAMR VQAERPPWYI
 651  SVVGGALYKT VTTNAATVSE YVSYLTKGHI PPKHISLVNS TVNLNIEGSK
 701  YTIETVRTGH GSYRLRMNDS TVEANVQSLC DGGLLMQLDG NSHVIYAEEE
 751  AGGTRLQIDG KTCLLQNDHD PSKLLAETPC KLLRFLVADG AHVDADVPYA
 801  EVEVMKMCMP LLSPASGVIH CMMSEGQALQ AGDLIARLDL DDPSAVKRAE
 851  PFDGIFPQME LPVAVSSQVH KRYAASLNAA RMVLAGYEHN INEVVQDLVC
 901  CLDNPELPFL QWDELMSVLA TRLPRNLKSE LEDKYKEYKL NFYHGKNEDF
 951  PSKLLRDIIE ENLSYGSEKE KATNERLVEP LMNLLKSYEG GRESHAHFVV
1001  KSLFEEYLTV EELFSDGIQS DVIETLRHQH SKDLQKVVDI VLSHQGVRNK
1051  AKLVTALMEK LVYPNPGGYR DLLVRFSSLN HKRYYKLALK ASELLEQTKL
1101  SELRASVARS LSDLGMHKGE MSIKDNMEDL VSAPLPVEDA LISLFDYSDR
1151  TVQQKVIETY ISRLYQPHLV KDSIQMKFKE SGAITFWEFY EGHVDTRNGH
1201  GAIIGGKRWG AMVVLKSLES ASTAIVAALK DSAQFNSSEG NMMHIALLSA
1251  ENESNISGIS SDDQAQHKME KLSKILKDTS VASDLQAAGL KVISCIVQRD
1301  EARMPMRHTF LWLDDKSCYE EEQILRHVEP PLSTLLELDK LKVKGYNEMK
```

FIG. 4A 12/25

SUBSTITUTE SHEET (RULE 26)

```
1351  YTPSRDRQWH IYTLRNTENP KMLHRVFFRT IVRQPNAGNK FRSAQISDAE
1401  VGCPEESLSF TSNSILRSLM TAIEELELHA IRTGHSHMYL CILKEQKLLD
1451  LIPFSGSTIV DVGQDEATAC SLLKSMALKI HELVGARMHH LSVCQWEVKL
1501  KLDCDGPASG TWRVVTTNVT GHTCTIDIYR EVEEIESQKL VYHSASSSAG
1551  PLHGVALNNP YQPLSVIDLK RCSARNNRTT YCYDFPLAFE TALQKSWQSN
1601  GSTVSEGNEN SKSYVKATEL VFAEKHGSWG TPIIPMERPA GLNDIGMVAW
1651  IMEMSTPEFP NGRQIIVVAN DITFRAGSFG PREDAFFETV TNLACERKLP
1701  LIYLAANSGA RIGIADEVKS CFRVGWSDEG SPERGFQYIY LTEEDYARIS
1751  SSVIAHKLEL DSGEIRWIID SVVGKEDGLG VENIHGSAAI ASAYSRAYEE
1801  TFTLTFVTGR TVGIGAYLAR LGIRCIQRLD QPIILTGFSA LNKLLGREVY
1851  SSHMQLGGPK IMATNGVVHL TVPDVLEGVS NILRWLSYVP ANIGGPLPIT
1901  KPLDPPDRPV AYIPENTCDP RAAICGVDDS QGKWLGGMFD KDSFVETFEG
1951  WAKTVVTGRA KLGGIPVGVI AVETQTMMQI IPADPGQLDS HERSVPRAGQ
2001  VWFPDSATKT AQALLDFNRE GLPLFILANW RGFSGGQRDL FEGILQAGST
2051  IVENLRTYNQ PAFVYIPMAG ELRGGAWVVV DSKINPDRIE CYAERTAKGN
2101  VLEPQGLIEI KFRSEELQDC MGRLDPELIN LKAKLQDVNH GNGSLPDIEG
2151  IRKSIEARTK QLLPLYTQIA IRFAELHDTS LRMAAKGVIK KVVDWEESRS
2201  FFYKRLRRRI AEDVLAKEIR QIVGDKFTHQ LAMELIKEWY LASQATTGST
2251  GWDDDDAFVA WKDSPENYKG HIQKLRAQKV SHSLSDLADS SSDLQAFSQG
2301  LSTLLDKMDP SQRAKFVQEV KKVLD
```

FIG. 4B

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001774

```
   1  AAGCTTGGTA TGGATTCgTC AGCGCCAAgC CGGGGTTTTG CATGCGCCCg
  51  ACTGGaArCs GAATTCCgTg AgCCCtGTaC rrCaATGGCA ACCCCAsGGT
 101  TACTggGGTG GCTGAATGGT CTCsGCTTAC GCAATTGTTT GTGGCAgCwG
 151  CGTGGGCTAA ATGTArGTTG TCTCTTGTTG CACTGCArGA TGGATGGGTA
 201  gCCTCTGGGC CGCCTCTGCT ArTGTCTArC GtTTGCTGAC TGTGGTTTA.
 251  tCAgGGATGc CCATgcCCAT GcTAgATTGA tAgGTgCCAt TCtAATgGTA
 301  gGTGgcgGTA AGGTTTATTA AgCTGtAgtA TCagTAgGTA ACCTCATGAA
 351  tCAgGGTTTA aGCACACCtT TTCCTTTGTg TgGGTgCATA AgGAAtGCAC
 401  TTGGCtTCgT TCCCtGAtAg TCtTTGctCA TgTgTCATTC TACCAAgTgG
 451  GTTACtgTAA CAtTGCACTC TATGATGGTT GGTGGTtGTG CATCTTTtTG
 501  CTtCCCCTGG tTGTCTAATA CCTGCATGTa ACTGATGACC ttCtTTTATG
 551  TATCATATAg ATTACATCCT TTTGTTGTAC ATCTCAATTC TGAAAAAACA
 601  ATGTTTTGCA TTCTTAGCGc TCTGTGCaCA AgGaAAagGa gGTTTTACCT
 651  gCAAcTtTTT TTTTCgAGAA AAAACAAACC TTTCTGAAag GCAGTGATCA
 701  TTTAGtATAA AGAAAATTTG ATTTACTTTC TTCAGAGAGA AtATkCCAAr
 751  CAAACAATTT TCTTACTGTC TGAGCCACGA AATTTGATCT TGATCtTACT
 801  TTCACAAGCC ACATGAAGCC tTATCATCGC TcTGATAAaA AAgCcaaaTa
 851  GGtGAttcAt aGaAtGaGag aAaGAacCTg TTgCCaTTTG GGGrCCtTGT
 901  TGTGTACTCA ttAtCCCCCC TGCtCAGGTT GaGGtTTTcC TTGccaCTGC
 951  CACCCCtTGG CCCCTtCTTA tACAAcCATC TCcAtTGaaA aAGAtTTTgC
1001  aCtACAtTTG GGCtcGtATG aCaaaaaAGG aAAAtaAAaC TaaaCAGCAG
1051  AAACATAGTA TaATTATAgG TAAAAGGTTc TGGCAAGTTT GAGTGGaAGA
```

FIG. 5A 14/25

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001775

```
1101  GACCTTTGTA TATTTGGACA TATTTCACTA GTAAATAGTT TTCTAAAATc
1151  TTCATGAATG GTGGCCAATA AACTTGATAA GATCTCAACA TGGCAGGTTC
1201  CTTCmAAATG AGAGGAAAAC TGGAAACATC ACAAATATTT TTTAGCGAGT
1251  GGCCTATAAA TTATAATGTT GCTttcAttt CttTGAtaTt caAaACtTCc
1301  taAgAGtatt ctgcTAGAGC TCTGATGGTG TCTTTTGCCT CTGTCAGATT
1351  TTCCAGGaGT TTTCTTCCCT TTTtATGGca CTGTGCGtTT GaGaaGgTct
1401  TCAATTgTgC TGtCTGGgCC acGGaaCGAC AAtgtcAcAg ctTGGATTAg
1451  ccgcAgctgC ctCAaaGGcc TTgcCaCTAC tCccTAATCG ccAGAGaAGt
1501  TCAgCTGGGa CtAcATTCTc ATCAtcttCA ttAtcGAGGC CCtTAaAcAG
1551  aagGaAAagC CaTAcTCGtT CACtCcGTGA TGgCGgAgaT GGgGTATcAG
1601  ATgCCaAAAA GCACAgCCAG TCTGTTcGTC aAGGTactGt GAATATCTtt
1651  tgatACAAgc taaAATtttg ctaCAgaAtA tatatTaAAG AGttCttTCt
1701  TGGctgGtgt tGtttAttT. GttttaA.cA tgcgaaagGg CctctAgttg
1751  agttggttAg gtggCCTGAa tACCACTCCt TAagGTCTTG AgttTG..at
1801  TTtCCgtCGG AgCGAaTttt AGGCTAgGGt TACCcCccCA CCcCCAcCcG
1851  AAtCTGCACA gtccggtcgt ggtcgtcctc atataggcta cGatgtcatt
1901  GtgTATCGGC GGgcCAggGG TTTAAGAgTT TTCTTGAcCT TTGTTAGAAg
1951  ATCTTAATAA TACAATgTCC AAGGGCTGTC TTACCCTGTA GGTCGAGTTT
2001  TTAGTTGTTT TAACATGGTA ATGTTTGAAG CCTCATTCTA GgtaCCAATA
2051  TAGATATGCt cactgctcag tttcaaatgt ttgtctgcat gtaggtcttg
2101  ctGGcattat CGaCctcCca AGtgaggCaC CttcCGaaGT GGAtAttTCa
2151  CAGtaAgGac taCAatattt TGCGtacgTT TG.TTtTGGa aaaaGaAaAT
```

FIG. 5B 15/25

SUBSTITUTE SHEET (RULE 26)

```
2201  AtTCtCAgct tAtttatact AgCttcgCtA A.TacTGaaA ..TgCtGtCt
2251  TaATGtCCtG GT.GCTGtAT GCTCaAtCTT TCATAGtAAA TgCTgCaaaa
2301  taTGTGAtGT AaCTGTTGCA ACACAGCCAG GGaCCTGTTA TTTAGAGCAT
2351  GGTGAATGcT CTGGTTCAGt tAtATGATGT AGTTATAGCT CATGtTGaAG
2401  aAtTAGttGc AgtGttTGCt ggacaatggt cacttatTat aaatcataTc
2451  tgcatacaCa TTtgTGacTt CTgttgctGt AAAtgCCCgc atTttTTGAG
2501  aaaAaTTTaa ATGctTGgcC taaaTTGGac ATAtATGAtA GACaccaAgC
2551  TGatTTGaAC TtTGttTatt tttgacatcC atgCAtAtTg tcAgtgTtGT
2601  GAAAacaata cTAatCcttT tTtttTGtCt ttTtCcaGTg gAtctgaGGA
2651  TCcTAGGGGg CCAACAGATT CTTATCAAAt GaatGGGaTT ATCaAtGaAA
2701  CaCaTaaTgg aagAcatgCc TCaGTGTCCa aGGttGTtga aTTTTgtgcg
2751  gcaCtAGgtg gcaaAaCACC AATTCACaGT ATATTAGTGG CCaACAATGG
2801  aATgGCAgca gcaaAattta TGagGaGTGT cCgGACATgG GCtAaTgATA
2851  ctTTtgGAtC TGagaAgGCA AttCAActCA TAgCTATgGC AACTCCGgaA
2901  gACATGAgGa TAaAtgCAGa ACACAttAGA ATTGCTGACC AATT.CGTAG
2951  AGGTGCCTGG TGGAACaAAC AATAATAACT ACG.CCAATG TTCAACTCAT
3001  AGTGGAGGTT AGCCTTGCTA ATCTGTTAGT TTACTACTGG TCTGCTGTTT
3051  CCTTTATTTG TTGTATAATG ATTGACATAT TTAAGTAGAG AAATTTATAT
3101  TTCTCCTCTG CTGTTGTGGA AGTCCAATTG TCATCATTAA CTGTGAAATA
3151  TTGCAGATGG CACAAAAACT AGGTGTTTCT GCTGtTTGGC CTGGTTGGGG
3201  TCATGCTTCT GAGAATCCTG AACTGCCAGA TGCATTGACC GCAAAAGGGA
```

FIG. 5C 16/25

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001777

```
3251  TCGTTTTTCT TGGCCCACCT GcATCATCAA TGAATGCTTT GGGAGaTAAG
3301  gTCgGCTCAG CTCTCaTTgC TCAAgCagcC GGggtCCCaA CTCT.TgCTT
3351  gGAgTgGATC ACATGTGAgT CTCACTCTTT gAtTACTAtC CgCCTGTCTC
3401  AtTgCTCTCt CtttCAtATT CTAATgACaC taAATtTAGG TTGaaGTtcC
3451  AtTAGagtgc TGCTtAGacG cGATAcCTGA GGAGATGtAt AGAaAAgcTT
3501  ATCGATaCCk TcgACctCGA GGGgGGgccc gGTaCyAgCT GsTG
```

FIG. 5D

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001778

```
  1  GAATTCCGTG AGCCCTGTAC GGCAATGGCA ACCCCAGGGT TACTGGGGTG
 51  GCTGAATGGT CTCGGCTTAC GCAATTGTTT GTGGCAGCTG CGTGGGCTAA
101  ATGTAGGTTG TCTCTTGTTG CACTGCAGGA TGGATGGGTA GCCTCTGGGC
151  CGCCTCTGCT AGTGTCTAGC GTTGCTGACT GTGGTTTATT CAGGGATGCC
201  ATGcCCATGC TAGATTGATA GGTCATAGGT GCCATTCTAA TGGTAGGTGG
251  CGGTAAGGTT TATTAAGCTG TCGTATCAGT AGGTAACCTC ATGAATCAGG
301  GTTTAAGCCC ACCTTCTCCT TTGTGTGGGT GCATAAGGAA TGCACTTGGC
351  TTCGTTCCCT GCTAGTCTTT GCTCATGTGT CATTCTACCA AGTGGGTTAC
401  TGTAACATTG CACTCTATGA TGGTTGGTGG TTGTGCATCT TTTTGCTTCC
451  CCTGGTTGTC TAATACCTGc ATGTAAcTGA TGACcTTCTT TTATGTATCA
501  TATAGATtAc ATcCTTTTGT TGtACATcTC AATTCTGAAA AaCAATGTTT
551  TGCATTCTTA GCGcTCTGTG CACAAGGAAA AGGAgGTTTT ACCTGCAAcT
601  TTTTTTTTCg AgAAAAAACA AACCTTTCTG AAAgGCAGTG ATCATTTAgT
651  ATAAAgAAAA TTTGAtTTAC TTtCTtCAga ArAgAATAtT CCAAACAAAC
701  aATTTTCTTA CAgTcTGAGC CACGAAATTT GATCTTGATC TTACtTTCaC
751  AAGCCACATG AAGCCTTATC ATCGCTCTGA TAAAAAAACC AAATAGGTGA
801  TTCATAGAAT GAGAAAAAGA ACCTGTTGCC ATTTGGGGAC CTTGTTGTGT
851  ACTCATTATC CCCCCTGCTC AGGTTGAGGT TT.CCTTGCC ACTGCCACCC
901  CTTGGCCCCT TCTTATACAA CCATCTCCAT TGAAAAAGAT TTTGCACTAC
951  ATTTGGGcTT cGTATaACaA AAAAGGaAAA TaAAAcTAAA CAGCAGaAAC
```

FIG. 6A 18/25

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001779

```
1001  ATAGTATAAT TATAGgTAAA AGGTTcTGGC AAGTTTGAGT GGTAGAGACC
1051  TTTGTATATT TGGACATATT TCACTAGTAA ATAGTTTTCT AAAATgTTCA
1101  TGAATGGTGG CCAATAAACT TGATAAGATC TCAACATGGC AGGTTCCTTC
1151  AAAATGAGAG GaAAACTGGA AACATCACAA ATATTTTTTA GCGAGTGGCC
1201  TATAAATTAT AATGTTGCTT TCATTTCTTT GATATTCAAA AcTTCCTAAG
1251  AGTATTcTGC TAGAGCtTCT GATGGTGTCT TTtGCcTCTG TCAGATTTTC
1301  CAGGaG.TTT TcTTCCCTTT TTATGGCACT GTGGTTTGAG AAGGTTTCAA
1351  TTGTGCTGTC TGGGCCACGG AACGACAATG TCACAGCTTG GATTAGCCGC
1401  AGCTGCCTCA AAGGCCTTGC CACTACTCCC TAATCGCCAG AGAAGTTCAG
1451  CTGGGACTAC ATTCTCATCA TcTTCATTAT CGAGGCCCTT AAACAGAAGG
1501  AAAAGCCGTA CTCGTTCACT CCGTGATGGC GGAGATGGGG TATCAGATGC
1551  CAAAAAGCAC AGCCAGTCTG TTcGTCAAGG TACTGTGAAT AtCTTTTGAT
1601  ACAAGCTAAA ATTTTGCTAC AGAATATATA TTAAAGAGTT CTTTCTTGGc
1651  TGGtGTTGTT TATTTGTTTT AACATGCGAA AgGGCcTCTA gTTGAGTTGG
1701  TTAGGTGGCC TGAGTACCAc TCCTTAAGGT CTTGAGTTTG ATTTTCCGTC
1751  AgAGCGAATT TTAGGcTAGG GTtaCCCCCC ACCCCCCCAC CCCTACCCGA
1801  ATcTGCACAg TCCGGTCGTG GTCGTCCTCA TATAgGCTAC gATGTCATTG
1851  TGTATCGGCG GGCCAGGGGT TTAAgAgTTT tCTTGACCTT TGTGAgAAgA
1901  TcTTAATAAT ACAATGTCCA AgAgCTGTCt TACCCTGTAG GTCrAgTTTT
1951  TTAgTTGTTT TAACATGGTT ATgTTTGAAg CCTCAtTCtA gGTAcCAATA
2001  TArATATgCt CaCTGCTCAg TTTCmAATGt TTGTCTGCAT kTAgGTCtTG
```

FIG. 6B 19/25

SUBSTITUTE SHEET (RULE 26)

2051 CTGGCATTAT CGACCTCCCA AGTGAGGCAC CTTCCGAAGT GGATATTTCA

2101 CAGTAAGGAn TACAGTATTT TGCGTACGTT TGTTTTGGAA AAAGAAATAT

2151 TCTCAGCTTA TTTAAT

FIG. 6C

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001781

```
  1  GAATTCCTGT GGGTGTCATA GCTGTGGAGA CACAGACCAT GATGCAGATC
 51  CTGTAATACT TCAAAATCTT AAGCCACAAA ACTTGATTAA TTGTTAGCAC
101  AGTAATTTGC CAAGTGGCTA GAGAAgGATC TCAACACAAC ACAATAACCA
151  AGAGATATCA ATCACAGAGA TGGCACGGTG GTTATCCCGT GGTTCGGCCA
201  AGACCAACGC TTGCCTACTC CACGTTGTGG CGTCCCAACG GACGAGGGTT
251  GCAATCAACC CCTCTCAAGC GGTCCAAAGA CCAACTTGAA TACCACGGTG
301  TTGCTTTGCT TTTCTTAATC CCACTTGCGA GGAATCTCCA CAGCTTGGAG
351  CCTCTCGCCC TTTCAAAAGA TTTCACAAAG AATCACGGAg CAAGGGAAGG
401  ATCAACAACT CACACACgAC ACAAAGATCA CAGTGAATAC GCACACATAA
451  AACCAAGACT TGAgCTCAAg TGACTAGCAC ACTT
```

FIG. 7A

```
  1  ATGGAaGTGT GTATTGCCAG TGcTTaTTyT rGGGATATGA GGGAATTwAm
 51  ATTAcATTTG TGAcTGGGCG GACTGTAGGA TAGGAGTTAT CTTGyTcGAT
101  TGGTATACGG TGCATACAGs kyTTGACCAG CTATTATTTT AACAGGgTTT
151  TCTGCCCTGA ACAAGTCCTT GGGCGGGAAG TGTACAGcTC CCACATGCAG
201  CTTGGTGGTC CTAAGATCAT GGCGACCAAT GGTGTTGTCC ACCTCACTGT
251  TCCAGATGAC CTTGAAGGTG TTTCCAATAT ATTGAGGTGG CTCAGCTATG
301  TTCCTGCAAA CATTGGTGGA CCTcTTCCTA TTACCAAACC TCTGGACCCT
351  CCAGACAGAC CTGTTGCTTA CATCCCTGAG AACACATGCG ATCCACGTGC
401  AGCTATCTGT GGTGTAGATG ACAGCCAAGG GAAATGGTTG GGTGGTATGT
451  TTGACAAAGA CAGCTTTGTG GAGACATTTG AAGGATGGGC AAAAACAGTG
501  GTTACTGGCA GAGCAAAGCT TGGAGGAATT C
```

FIG. 7B

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001782

```
  1  CTCCCAATAT TGTCATGAGG CTTGCATCCC AGGTTAGTTT TTTTTCCTTT
 51  CTGAAATTTA TATTCCATAC CTTTTCACCT TTAGTTATCC TTGTATTTTC
101  TGGAAGCTTC ATCTGATGCA TTATTGACAA ATGCACTAAT GGTCATCATA
151  TTTGsAkAgw yAAsATAtkT mTCTTaattg aTGGkwACtm TTgamwATGA
201  srATgsTkrA GCAkrTrryk wAyAstTTTT TaATAAAAAA AcatGCaTTT
251  cTAgGAGtTG GAcTaAGctT TTCTTAGtAT GAaGtGcCAT GtTTTAcATg
301  GTCCATTTGt GTCAATTTAC AGTCgGTATC ATGGaAAGGt TGTCATAATG
351  GcTGGAGAga AACAACAcAT CTTGtTTCTC AACAcTTGTg GGAGAAGAtG
401  TTTTAcCTTT TTTcCTAAAA TTACTTTTTG TACTAAATTG TATAAkTTTT
451  cCAATATTCT cCATGATTAT TGAACTCTGC TGTGtTCAAA CAGCCAAAAC
501  ATGTTTcCAT ACTTTACACC TTTATTTTTT AGATGGAAgC CTGGAATTGT
551  GCTCTGTTAT CTGTAgTCAT GCAtTATAtT TGATCTTAAA tCyTAtTCTC
601  TATTGTAGAr TCsGCAtCTT GAAgTTCAGT TGCtTTGTGA TCAATATGGT
651  AATGTAGCAG CACTTCACAG TCGTGATTGC AGTGTGCAAC GGCGACACCA
701  GAAGGTCTGC CCCTCACCCA CCCAGCCATA AACACGAAGT TTATAGAACC
751  ATGTATTTTG TTATGCAATA TATTTCTCAA TTGTAGCTCC ATTCACATTT
801  TGCTACAACA GATTATTGAA GAAGGTCCAG TTACTGTTGC TnCCTCGTGA
851  GACAGTTAAA GCACTTGAGC AGCAGCAAGG AG
```

FIG. 7C

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001783

```
  1  GTcGACCTGC AGGTCAACGG ATCCTAgGGG GCCAACAGAT TcTTATcaaA
 51  TgAATGggat TATCaATGAA AcacATaATG GaAGACAtGC cTCAGtGTCC
101  AAgGTTGtTG aATTTTGTGC GGCAcTAgGT GGCaAAACAC CaATTCACAG
151  TATATTAGTG GCCAACAATG GaATGGCAgC AGCaAAAtTT ATGagGaGTG
201  TCCGGACATG GGCTAATGAT AcTTTTGGAT CTGAGAAgGC AATTCAACTC
251  ATAGCTATGG CAACTCCGGA AGACATGAGG aTAAATGCAG AACACATTAG
301  AATTGCTGAC CAATTCGTAg ArgTGcctgG TGgaacaaAC AAtAATAAct
351  AcgCCAaTGT TCAAcTcATA gTGGAgGTTA GCCTTGcTAA TCTGTTAgTT
401  TACTACTGgT CtGCtGtTTC CtTTAtTTGt tGTaTAATGA ttGACaTATT
451  taAgTagAgA AAtTTATAtT TCtCctCtgC tGTTGTGgAa gTCCAatTGT
501  CaCCAtTAAC tGTgaAATAt TGCagATgGc aCaAAAACtA gGtGtTTcTg
551  CTgTTTGGCC TgGTTGGGGT CATGCTTCTG AgAAtCcTGA ACTGCCAgAT
601  GCATTGACCG CAAAAgGGaT CgTTTTTCTT GGcCcacCTG CATCAtCAAT
651  GaATGCtTTG GGAGATAAGk TCgGcTCAgC TCTCATTGCT CAAGCAgCCG
701  GGGtCCCaAC tCtTGCTTGG AgTgGATCAC aTGTGAgTCt CACtCtTTGA
751  tTACTAtCCG cCTGTCtCAt TGCtCTcTCt TTCATATTCT AATGACACTA
801  AATTTAGGTT GAAGTTCCAT TAGAGTGCTG CTTAGACGCG ATACCTGAGG
851  AGATGTATAG AAAAGCT
```

FIG. 7D

SUBSTITUTE SHEET (RULE 26)

COR-AAD1-00001784

```
  1  GAATAATCTG CCTGCAGCTC AAGTTGCTGT TGGAATGGGC ATACCTCTTT
 51  GGCAGATTCC AGGTAATTAC CAATTTACCA ACTTATTTAG TTCCTTATTG
101  TTTTATTCTC TAATTTTCTA CTTATGTAgA AATCAGAcGT TtCTATGGAa
151  TGgAcTATGG AGGagGGTAt GAcATTTGGa GGAAAaCAGC AGcTCTTGCT
201  ACACCATTtA ATTTTGATGA AGTAgaTTcT CAATGGCCAA AGGGCCATTG
251  TGTAgCAGTT AGAATTAcTA GTGAGGACCC AGATGATGGT TTCAAACcTA
301  CTGGTGGGAA AGTGAAGGTA AGTTTTCTAG ATGACATGTA TTATATATCG
351  TTCAAAgAgA TTAAGTTTGG TTAAATGAcT AGGTCTTGAT TTTTTATCTT
401  TCAGGAGATA AGTTTTAAAA GCAAGCCTAA TGTTTGGGCC TaCTTCTCAG
451  TAAAGGTaAC TTGTTAACTT TAGTACGCTG TCACATTATt ctTcgTTGTG
501  AAAATAAtTT GAACGGTtCT CTTTGTATTT TaACCAtCCA tCgTcTCATT
551  TAgCAgAgCA CACAAATATT tGCACTGACC CCCcTcCCCt TATCtGctTT
601  CAgTCTGGTG GAgGCATtCA tGAATTtGCT GATTCTCAGT TCGGTATGTG
651  TAAACCAAGA GTATTCTTTG TAATTTATAT TGGTCCTCAA TTTTGAAATA
701  TTGTCTTTCC GTTACAGGAC AdG
```

FIG. 7E

COR-AAD1-00001785

```
  1  AATTCCTGTG GGTGTTATAG CTGTGGAGAC ACAGACCATG ATGCAGCTCA
 51  TCCCTGCTGA TCCAGGTCAA CTTGATTCCC ATGAGCGATG TGTTCCTCGG
101  GCTGGACAAG TGTGGTTCCC AGATNCTGCA ACCAAGACAG CTCAGGCATT
151  ATTAGACTTC AACCGTGAAG GATTGCCTCT GTTCATCCTG GCTAACTGGA
201  GAGGCTtCTC TGGGGgACAG AGAGATCTCT T
```

FIG. 8A

```
  1  AATTCATGCA TCTTAATAAA CACAGTTGGC CCTTAAAGCA AGTGAACTTC
 51  TTGAACAAAC CAAACTAAGT GAACTCTGTT CCAGCATTGC AAGAAGCCTT
101  TCAGATCTGG GGATGCATAA GGGAGAAATG ACTATTAAGG ATAGCATGGA
151  AGATTTAGTC TCTGNCCCAT TGCCTGTTGA AGATGCTCTT ATTTCTTTGT
201  TTGATTA
```

FIG. 8B

```
  1  ATAGACCTGT CGCATACATC CCTGAGAACA CATGCGATCC GCGTGCAGCC
 51  ATCCGTGGnG TAGATGACAG CCAAGGGAAA TGGTTGGGTG GTATGTTTGA
101  CAAAGACAGC TTTGTGGAGA CATTTGAAGG ATGGGCAAAA ACAGTGGTTA
151  CTGGTAGAGC AAAGCTTGGA GGAAGGAATT
```

FIG. 8C

COR-AAD1-00001786

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/US 97/15344

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 //A01H1/00,A01H5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MARSHALL L C ET AL: "ALLELIC MUTATIONS IN ACETYL-COENZYME A CARBOXYLASE CONFER HERBICIDE TOLERANCE IN MAIZE" THEORETICAL AND APPLIED GENETICS, vol. 83, no. 4, 1992, BERLIN DE, pages 435-442, XP002051060 cited in the application see the whole document ---	1,13,14, 29,41, 44,51
A	VAN DEE K.L. ET AL.: "RFLP mapping of Acc1 in Zea mays L." AGRONOMY ABSTRACTS, 1992, MADISON US, page 198 XP002051061 cited in the application see abstract --- -/--	1,15,16, 30,40, 42,45, 46,52

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

2

Date of the actual completion of the international search	Date of mailing of the international search report
29 December 1997	14. 01. 98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	De Kok, A

Form PCT/ISA/210 (second sheet) (July 1992)

COR-AAD1-00001787

INTERNATIONAL SEARCH REPORT

Intern. al Application No
PCT/US 97/15344

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CAFFREY J J ET AL: "GENETIC MAPPING OF TWO ACETYL-COA CARBOXYLASE GENES" MAIZE GENETICS COOPERATION NEWSLETTER, vol. 69, 1995, COLUMBIA US, pages 3-4, XP002051062 cited in the application see the whole document	1,13,14
A	US 5 162 602 A (SOMERS DAVID A ET AL) 10 November 1992 cited in the application see the whole document	1-8, 17-34, 36,37, 47,48, 53,54
A	EGLI M A ET AL: "A MAIZE ACETYL-COENZYME A CARBOXYLASE CDNA SEQUENCE 1" PLANT PHYSIOLOGY, vol. 108, no. 3, July 1995, US, page 1299/1300 XP002013205	1,9,10
A	US 5 498 544 A (GENGENBACH BURLE G ET AL) 12 March 1996 see abstract see column 17 - column 38	1
A	DE 42 34 131 A (MAX PLANCK GESELLSCHAFT) 21 April 1994 see abstract	1
A	DATABASE WPI Section Ch, Week 9307 Derwent Publications Ltd., London, GB; Class C06, AN 93-057074 XP002051063 & SU 1 720 594 A (MOLD VEGETABLES GROWING RES INST) , 23 March 1992 see abstract	1,3,35
A	VAHLENSIECK H F ET AL: "IDENTIFICATION OF THE YEAST ACC1 GENE PRODUCT (ACETYL-COA CARBOXYLASE) AS THE TARGET OF THE POLYKETIDE FUNGICIDE SORAPHEN A" CURRENT GENETICS, vol. 25, no. 2, 1 February 1994, BERLIN DE, pages 95-100, XP000196332 see abstract	53,54
A	WO 89 07647 A (PIONEER HI BRED INTERNATIONAL) 24 August 1989 see the whole document	1,9

2

COR-AAD1-00001788

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern. al Application No
PCT/US 97/15344

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5162602 A	10-11-92	US 5498544 A	12-03-96
		US 5428001 A	27-06-95
		US 5290696 A	01-03-94
US 5498544 A	12-03-96	US 5290696 A	01-03-94
		US 5162602 A	10-11-92
		US 5428001 A	27-06-95
DE 4234131 A	21-04-94	AU 671669 B	05-09-96
		AU 4872093 A	28-04-94
		CA 2108112 A	10-04-94
		EP 0616035 A	21-09-94
		HU 68236 A	28-06-95
		JP 6197651 A	19-07-94
		MX 9306300 A	29-04-94
		US 5689045 A	18-11-97
		ZA 9307202 A	20-04-94
WO 8907647 A	24-08-89	AU 631562 B	03-12-92
		AU 4030289 A	06-09-89
		EP 0402401 A	19-12-90

COR-AAD1-00001789

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/82, 15/53, 15/54, 15/55, 15/60, A01H 5/00

A2

(11) International Publication Number: **WO 98/20144**

(43) International Publication Date: 14 May 1998 (14.05.98)

(21) International Application Number: PCT/GB97/02996

(22) International Filing Date: 31 October 1997 (31.10.97)

(30) Priority Data:

9623248.3	7 November 1996 (07.11.96)	GB
9625957.7	13 December 1996 (13.12.96)	GB
9703855.8	25 February 1997 (25.02.97)	GB

(71) Applicant *(for all designated States except US)*: ZENECA LIMITED [GB/GB]; 15 Stanhope Gate, London W1Y 6LN (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: THOMPSON, Paul, Anthony [GB/GB]; Zeneca Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG42 6ET (GB). KNIGHT, Mary, Elizabeth [GB/GB]; Zeneca Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG42 6ET (GB). JEPSON, Ian [GB/GB]; Zeneca Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG42 6ET (GB). THOMAS, Paul, Graham [GB/GB]; Zeneca Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG42 6ET (GB). HAWKES, Timothy, Robert [GB/GB]; Zeneca Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG42 6ET (GB).

(74) Agents: GAAL, Jozsef, Christopher et al.; Zeneca Agrochemicals, Intellectual Property Dept., Jealott's Hill Research Station, P.O. Box 3538, Bracknell, Berkshire RG42 6YA (GB).

(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, ID, 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, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), 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

Without international search report and to be republished upon receipt of that report.

(54) Title: HERBICIDE RESISTANT PLANTS

(57) Abstract

The present invention provides, *inter alia*, a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5–enol–pyruvyl–3–phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); and (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT).

COR-AAD1-00001790

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

COR-AAD1-00001791

HERBICIDE RESISTANT PLANTS

The present invention relates to recombinant DNA technology, and in particular to the production of transgenic plants which exhibit substantial resistance or substantial tolerance to herbicides when compared with non transgenic like plants.

Plants which are substantially "tolerant" to a herbicide when they are subjected to it provide a dose/response curve which is shifted to the right when compared with that provided by similarly subjected non tolerant like plants. Such dose/response curves have "dose" plotted on the x-axis and "percentage kill", "herbicidal effect" etc. plotted on the y-axis. Tolerant plants will require more herbicide than non tolerant like plants in order to produce a given herbicidal effect. Plants which are substantially "resistant" to the herbicide exhibit few, if any, necrotic, lytic, chlorotic or other lesions when subjected to the herbicide at concentrations and rates which are typically employed by the agrochemical community to kill weeds in the field. Plants which are resistant to a herbicide are also tolerant of the herbicide. The terms "resistant" and "tolerant" are to be construed as "tolerant and/or resistant" within the context of the present application.

According to the present invention there is provided a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); and (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT).

In a preferred embodiment of the invention the regions comprised by the polynucleotide are each under expression control of a plant operable promoter and terminator. Such promoters and terminators are well known to the skilled man who will choose them according to his particular needs. For example, suitable promoters include the 35S CaMV or FMV promoters, and the arabidopsis and maize ubiquitin promoters. Preferably, the promoters are constitutive. This avoids any need for external induction and

means that the plant is permanently tolerant of or resistant to each corresponding herbicide. DNA encoding the herbicide resistance genes may also be included in a plant transformation vector under the control of an inducible promoter, to give inducible herbicide resistance in the transgenic plants. Such promoters include the chemically-inducible known GST-27 promoter by which resistance may be switched on by application of a suitable inducer (such as a chemical safener). In certain circumstances, the ability to express or to increase herbicide resistance only when required may be advantageous. For example, during rotation of crops, individuals of the first crop species may grow the following year in the field to be cultivated with a second crop species. A herbicide may be used to destroy these un-induced and still susceptible "volunteer" plants. Induction of herbicide resistance gene expression only when herbicide resistance is required (that is, just before application of a herbicide) may also be metabolically more efficient in some circumstances as the plant then produces resistance polypeptides only when required. Suitable inducible promoters further include the tetracycline-inducible promoter, the lac bacterial repressor/operator system, the glucocorticoid receptor, together with dexamethasone, copper and salicylic acid-inducible promoters, promoters based on the ecdysone receptor, as described in International Patent Application No. PCT/GB96/01195, and the so-called Alc promoter, as described in International Patent Publication No. WO93/21334.

In a particularly preferred embodiment of the invention, at least one of the regions comprised by the polypeptide provides for resistance to a pre-emergence herbicide and at least one of the regions provides for resistance to a post emergence herbicide. Whilst the skilled man does not need a definition of pre-emergence and post emergence, by "pre-emergence" is meant applied before the germinating seed emerges above the soil surface, ie before any plant material is visible above the ground. Post emergence means applied after the seedling is visible above the surface of the soil.

The pre-emergence herbicide may be selected from the group consisting of a dinitroaniline herbicide, bromacil, flupoxam, picloram, fluorochloridone, tetrazolinones including N-carbamoyltetrazolinones such as those described in EP-A-612,735, sulcatrione, norflurazone, RP201772, atrazine or another triazine, iminothiadozole, diflufenicon, sulfonyl urea, imidazolinone, thiocarbamate, triazine, uracil, urea, triketone, isoxazole, acetanilide,

oxadiazole, the phosphosulfonate herbicides described in EP-A-511,826, triazinone, sulfonanilide, amide, oxyacetamides such as fluthiamide, anilide and triazolinone type herbicide. Examples of triketone herbicides include 2-(2-Nitro-4-trifluoromethylbenzoyl)-cyclohexane-1,3-dione

2-(2-Chloro-4-methanesulphonylbenzoyl)-cyclohexane-1,3-dione,

2-(2-2 Nitro-4-methanesulphonylbenzoyl)-cyclohexane-1,3-dione,

[5-cyclopropyl-4-(2-methylsulphonyl-4-trifluoromethylbenzoyl)isoxazole, etc.

For the avoidance of doubt, by "triketone herbicide" is meant any compound capable of inhibiting a 4-hydroxyphenyl pyruvate (or pyruvic acid) dioxygenase (HPPD). Within the context of the present invention the terms 4-hydroxy phenyl pyruvate (or pyruvic acid) dioxygenase (4-HPPD) and p-hydroxy phenyl pyruvate (or pyruvic acid) dioxygenase (p-OHPP) are synonymous.

The post-emergence herbicide may be selected from the group consisting of glyphosate and salts thereof, glufosinate, diphenyl ether, asulam, bentazon, bialaphos, bromacil, sethoxydim or another cyclohexanedione, dicamba, fosamine, flupoxam, phenoxy propionate, quizalofop or another aryloxy-phenoxypropanoate, picloram, fluormetron, atrazine or another triazine, metribuzin, chlorimuron, chlorsulfuron, flumetsulam, halosulfuron, sulfometron, imazaquin, imazethapyr, isoxaben, imazamox, metosulam, pyrithrobac, rimsulfuron, bensulfuron, nicosulfuron, fomesafen, fluroglycofen, KIH9201, ET751, carfentrazone, ZA1296, ICIA0051, RP201772, flurochloridone, norflurazon, paraquat, diquat, bromoxynil and fenoxaprop. Particularly preferred combinations of these herbicides to which the polynucleotide of the invention is capable of conferring resistance (or to which the plants of the invention are resistant or tolerant) are: (i) glyphosate and diphenyl ether or acetanalide type herbicides: (ii) glyphosate and/or glufosinate and anilide and/or triazolinone type herbicides; (iii) triketones and glyphosate and/or glufosinate; (iv) glyphosate and/or glufosinate and triketones and anilide type herbicides; (v) glyphosate and/or glufosinate and a PDS inhibitor (such as the compounds of formulas I-III depicted below).

The proteins encoded by the said regions of the polynucleotide may be selected from the group consisting of glyphosate oxido-reductase (GOX), 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS), phosphinothricin acetyl transferase (PAT),

COR-AAD1-00001794

hydroxyphenyl pyruvate dioxygenase (HPPD), glutathione S transferase (GST), cytochrome P450, Acetyl-COA carboxylase (ACC), Acetolactate synthase (ALS), protoporphyrinogen oxidase (protox), dihydropteroate synthase, polyamine transport proteins, superoxide dismutase (SOD), bromoxynil nitrilase (BNX), phytoene desaturase (PDS), the product of the *tfd*A gene obtainable from *Alcaligenes eutrophus,* and mutagenised or otherwise modified variants of the said proteins. The product of the said *tfd*A gene is a dioxygenase which is capable of oxidising phenoxycarboxylic acids, such as 2,4-D to the corresponding phenol. The EPSPS enzyme may be a so called class II EPSPS, as described in European Patent No. 546,090. Alternatively, and/or additionally, it may be mutated so as to comprise amino acid substitutions at certain positions which are known to result in enhanced resistance to glyphosate (and agriculturally acceptable salts thereof). For example, the EPSPS may have at least the residues Thr, Pro, Gly and Ala at positions corresponding to 174, 178, 173 and 264 with respect to the EPSPS depicted in SEQ ID No. 9 alerted as follows:

(i) Thr 174 - Ile

(ii) Pro 178 - Ser

(iii) Gly 173 - Ala

(iv) Ala 264 - Thr

wherein (i) Thr 174 occurs within a sequence comprising contiguously Ala -Gly-Thr-Ala-Met; (ii) Pro 178 occurs within a sequence comprising contiguously Met-Arg-Pro-Leu-Thr; (iii) Gly 173 occurs within a sequence comprising contiguously Asn-Ala-Gly-Thr-Ala; and (iv) Ala 264 occurs within a sequence comprising contiguously Pro-Leu-Ala-Leu-Gly. Additionally, the terminal Gly residue within the sequence motif Glu-Arg-Pro-AA1-AA2-Leu-Val-AA3-AAA4-Leu-AA5-AA6-AA7-Gly- in a region of the EPSPS enzyme corresponding to that spanning positions 192 to 232 in SEQ ID No. 9 may be replaced by either an Asp or Asn residue.

In one embodiment of the polynucleotide, the region encoding the HPPD enzyme has the sequence depicted in SEQ ID Nos. 1 or 3, or alternatively is complementary to one which when incubated at a temperature of between 60 and 65°C in 0.3 strength citrate buffered saline containing 0.1% SDS followed by rinsing at the same temperature with 0.3 strength citrate buffered saline containing 0.1% SDS still hybridises with the sequence depicted in SEQ ID No. 1 or 3 respectively.

When the test and inventive sequences are double stranded the nucleic acid constituting the test sequence preferably has a TM within 15°C of that of the said SEQ ID No. 1 sequence. In the case that the test and SEQ ID No. 1 sequences (or test and SEQ ID No. 3 sequences) are mixed together and are denatured simultaneously, the TM values of the sequences are preferably within 5°C of each other. More preferably the hybridisation is performed under relatively stringent conditions, with either the test or inventive sequences preferably being supported. Thus either a denatured test or inventive sequence is preferably first bound to a support and hybridisation is effected for a specified period of time at a temperature of between 60 and 65°C in 0.3 strength citrate buffered saline containing 0.1% SDS followed by rinsing of the support at the same temperature but with 0.1 strength citrate buffered saline. Where the hybridisation involves a fragment of the inventive sequence, the hybridisation conditions may be less stringent, as will be obvious to the skilled man.

When the polynucleotide comprises an HPPD gene capable of conferring resistance to triketone herbicides, plant material transformed therewith may be subjected to a triketone herbicide and visually selected on the basis of a colour difference between the transformed and non transformed material when subjected to the said herbicide. Thus the non-transformed material may become and stay white when subjected to the selection procedure, whereas the transformed material may become white but later turn green, or may remain green, likewise, when subjected to the said selection procedure.

A further embodiment of the polynucleotide of the invention includes a further region encoding a protein capable of providing the plant with resistance or tolerance to insects, desiccation and/or fungal, bacterial or viral infections. The proteins encoded by such regions are known to the skilled man and include the delta endotoxin from *Bacillus thuringiensis* and the coat proteins from viruses, for example.

The polynucleotide may comprise sequences 5' of and contiguous with the said regions, which sequences encode (i) a peptide which is capable of targeting the translation products of the regions to plastids such as chloroplasts, mitochondria, other organelles or plant cell walls; and/or (ii) non-translated translational enhancing sequences. Suitable targeting sequences encode chloroplast transit peptides, particularly in the case that the herbicide resistance conferring region immediately down-stream of it is an EPSPS or GOX enzyme. Translational expression of the protein encoding sequences contained within the

polynucleotide may be relatively enhanced by including known non translatable translational enhancing sequences 5' of the said protein encoding regions. The skilled man is very familiar with such enhancing sequences, which include the TMV-derived sequences known as omega, and omega prime, as well as other sequences derivable, *inter alia*, from the regions 5' of other viral coat protein encoding sequences, such as that of the Tobacco Etch virus. It may be desirable, having regard to the expression of nucleotide sequences *in planta*, to modify the sequences encoding known proteins capable of conferring resistance to herbicides. Accordingly the invention also includes a polynucleotide as indicated above, but which is modified in that mRNA instability motifs and/or fortuitous splice regions are removed, or plant preferred codons are used so that expression of the thus modified polynucleotide in a plant yields substantially similar protein having a substantially similar activity/function to that obtained by expression of the unmodified polynucleotide in the organism in which the protein encoding regions of the unmodified polynucleotide are endogenous, with the *proviso* that if the thus modified polynucleotide comprises plant preferred codons, the degree of identity between the protein encoding regions within the modified polynucleotide and like protein encoding regions endogenously contained within the said plant and encoding substantially the same protein is less than about 70%.

The invention further includes a vector comprising the said polynucleotide.

The invention still further provides plants which comprise at least two nucleotide sequences encoding proteins capable of conferring resistance to at least two herbicides and which have been regenerated from material which has been transformed with the polynucleotide or vector of the invention. Transformation techniques are well known and include particle mediated biolistic transformation, *Agrobacterium*-mediated transformation, protoplast transformation (optionally in the presence of polyethylene glycols); sonication of plant tissues, cells or protoplasts in a medium comprising the polynucleotide; micro-insertion of the polynucleotide into totipotent plant material (optionally employing the known silicon carbide "whiskers" technique), electroporation and the like. The transformed inventive plants include small grain cereals, oil seed crops, fibre plants, fruit, vegetables, plantation crops and trees. Particularly preferred such plants include soybean, cotton, tobacco, sugarbeet, oilseed rape, canola, flax, sunflower, potato, tomato, alfalfa, lettuce, maize, wheat, sorghum, rye, bananas, barley, oat, turf grass, forage grass, sugar cane, pea, field bean,

COR-AAD1-00001797

rice, pine, poplar, apple, grape, citrus or nut plants and the progeny, seeds and parts of such plants.

The invention still further provides plant material which comprises nucleic acid sequences comprising regions encoding at least two proteins capable of conferring upon the material resistance to at least two herbicides, with the *provisos* that the material that the material does not contain a polynucleotide which encodes a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the material does not contain a polynucleotide which comprises only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); (iii) that the material does not contain a polynucleotide which comprises only regions encoding GST and phosphinothricin acetyl transferase (PAT); and (iv), that when the plant from which the material is derived is sugar beet, the herbicide resistance or tolerance conferring genes which it comprises are not solely EPSPS and PAT.

The material may be regenerated into morphologically normal fertile whole plants, by means known to the skilled man. In a preferred embodiment of the material, at least one of the regions encodes a protein capable of conferring resistance to a pre-emergence type herbicide, and at least one of the regions encodes a protein capable of providing resistance to a post emergence type herbicide. Such protein encoding regions and herbicides have been discussed above. The skilled man will recognise that multiple herbicide resistance conferring regions may be present in plants (or parts thereof) as a consequence of the crossing of a first plant comprising a polynucleotide encoding a first protein capable of conferring resistance to a first herbicide with a second plant which comprises a polynucleotide encoding a second protein capable of conferring resistance to a second herbicide (see the experimental part of the application). Preferred combinations of herbicide resistance conferring genes are (i) an HPPD gene and an EPSPS or GOX gene; (ii) an HPPD gene and a PAT gene; (iii) a GST gene and an EPSPS/GOX gene; (iv) an EPSPS/GOX gene and a PAT gene; (iv) an HPPD gene, a GOX and/or EPSPS gene, and a PAT gene; (v) an ACC'ase gene and a PAT and/or EPSPS gene; (vi) a PDS gene and a PAT and/or EPSPS and/or GOX gene; (vii)), the *tfd*A gene obtainable from *Alcaligenes eutrophus* and an EPSPS and/or GOX and/or PAT and/or PDS gene. In addition each of these combinations

may have one or more of the herbicide genes replaced by a SOD, protox and/or ALS gene. Such plants are referred to in this application as plants of the invention.

The invention also includes a method of selectively controlling weeds in a field comprising weeds and crop plants, wherein the crop plants comprise (i) a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT); and (iv), that when the crop plant is sugar beet, the herbicide resistance or tolerance conferring genes which it comprises are not solely EPSPS and PAT; or (ii) a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a polynucleotide comprising a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT); and (iv), that when the crop plant is sugar beet, the herbicide resistance or tolerance conferring genes which it comprises are not solely EPSPS and PAT, the method comprising application to the field of at least one of the said herbicides in an amount sufficient to control the weeds without substantially affecting the crop plants. The herbicide resistance conferring genes may be present on separate polynucleotides within the plant. In a preferred method the plant contains genes encoding an EPSPS and/or GOX enzyme and an HPPD enzyme, the method comprising application to the field of glyphosate and a triketone herbicide in an amount sufficient to control the weeds without substantially affecting the crop plants. In a further embodiment of the method, the plant contains genes encoding an

EPSPS and/or GOX enzyme and a phosphinothricin acetyl transferase, the method comprising application to the field of glyphosate and glufosinate. In a further embodiment of the method, the plant contains genes encoding an EPSPS and/or GOX enzyme and a phosphinothricin acetyl transferase and an HPPD enzyme, the method comprising application to the field of glyphosate and glufosinate and a triketone herbicide. In a further embodiment of the method, the plant contains genes encoding an EPSPS and/or GOX enzyme and/or a phosphinothricin acetyl transferase and a glutathione S transferase, the method comprising application to the field of glyphosate and/or glufosinate and an anilide herbicide such as acetochlor, for example. In a further embodiment of the method, the plant contains genes encoding an ACC'ase and a PAT and/or EPSPS enzyme, the method comprising application to the field of a fluazifop type herbicide and glufosinate and/or glyphosate. In a still further embodiment of the method, the plant contains genes encoding the product of the *tfd*A gene (optionally codon optimised) obtainable from *Alcaligenes eutrophus* and an EPSPS and/or GOX and/or PAT and/or PDS enzyme, the method comprising application to the field of 2,4 D and glyphosate and/or glufosinate and/or a herbicidal inhibitor of phytoene desaturase. In addition each of these combinations may have one or more of the herbicide genes replaced by a SOD, protox and/or ALS gene.

In a particularly preferred embodiment of this inventive method, a pesticidally effective amount of one or more of an insecticide, fungicide, bacteriocide, nematicide and anti-viral is applied to the field either prior to or after application to the field of one or more herbicides.

The present invention further provides a method of producing plants which are substantially tolerant or substantially resistant to two or more herbicides, comprising the steps of:

(i) transforming plant material with the polynucleotide or vector of the invention;

(ii) selecting the thus transformed material; and

(iii) regenerating the thus selected material into morphologically normal fertile whole plants.

The plants of the invention may optionally be obtained by a process which involves transformation of a first plant material with a first herbicide resistance conferring sequence,

COR-AAD1-00001800

and transformation of a second plant material with a second herbicide resistance conferring sequence, regeneration of the thus transformed material into fertile whole plants and cross pollination of the plants to result in progeny which comprises both the said first and second herbicide resistance genes. Optionally the first and/or second material may have been prior transformed with polynucleotides comprising regions encoding one or more of a herbicide resistance conferring protein, an insecticidal protein, an anti-fungal protein, an anti-viral protein, and/or a protein capable of conferring upon a plant improved desiccation tolerance.

The invention still further provides the use of the polynucleotide or vector of the invention in the production of plant tissues and/or morphologically normal fertile whole plants (i) which are substantially tolerant or substantially resistant to two or more herbicides.

The invention still further provides the use of the polynucleotide or vector of the invention in the production of a herbicidal target for the high throughput *in vitro* screening of potential herbicides. The protein encoding regions of the polynucleotide may be heterologously expressed in *E. coli* or yeast.

The invention still further includes plant tissue transformed with a polynucleotide comprising the sequence depicted in SEQ ID No. 1 and encoding a dioxygenase. This may be the only herbicide resistance conferring gene within the material. The material may be regenerated into morphologically normal fertile plants using known means. In a particularly preferred embodiment of the transformed tissue, the polynucleotide which encodes a protein having a substantially similar activity to that encoded by SEQ ID No. 1, is complementary to one which when incubated at a temperature of between 60 and 65°C in 0.3 strength citrate buffered saline containing 0.1% SDS followed by rinsing at the same temperature with 0.3 strength citrate buffered saline containing 0.1% SDS still hybridises with the sequence depicted in SEQ ID No. 1.

The invention will be further apparent from the following description taken in conjunction with the associated figures and sequence listings.

SEQ ID No. 1 shows a DNA sequence, isolated from *Synechocystis sp*, which encodes *a*n enzyme (depicted as SEQ ID No. 2) having the activity of a p-hydroxyphenyl pyruvic acid dioxygenase.

SEQ ID No. 3 shows a DNA sequence, isolated from *Pseudomonas* spp. 87/79, in which nucleotides 1217 to 2290 encode an enzyme (depicted as SEQ ID No. 4) having the activity of a p-hydroxyphenyl pyruvic acid dioxygenase.

SEQ ID Nos. 5 and 6 depict one form of the minimally redundant synthetic PCR primers (see reference to HPPD-P4 and HPPD-REV1 below) which were used to isolate SEQ ID No 3 from the bacterial genome.

SEQ ID Nos. 7 and 8 are also synthetic PCR primers which were used to modify the SEQ ID No. 3 sequence so that it could be incorporated into the desired plant transformation vector.

SEQ ID No. 9 shows the amino acid sequence of an EPSPS enzyme (including chloroplast signal peptide) from petunia.

SEQ ID Nos. 10-32 are PCR primers or poly-linkers which are inserted into restricted plasmids to enable the production of constructs comprising multiple genes capable of conferring resistance to herbicides.

Figure 1 shows a schematic diagram of the clone comprising the sequence depicted in SEQ ID No. 3, in which three open reading frames are identified: the first starting at nucleotide 15 and ending at nucleotide 968; the second starting at nucleotide 215 and ending at nucleotide 1066 and the third starting at nucleotide 1217 and ending at nucleotide 2290 in SEQ ID No. 3. The Figure also shows the restriction sites contained within the sequence which are engineered by use of the primers designated as SEQ ID Nos. 7 and 8. Figure 2 schematically depicts the production of a 4-HPPD containing plant expression cassette in which the PCR edited DNA fragment of Figure 1 is restricted with the enzymes *Nco*1 and *Kpn*1, then ligated into a vector (pMJB1) also restricted with *Nco*1 and *Kpn*1. Figure 3 is a schematic representation showing how the plant transformation binary vector pBin 19 is engineered to contain the 4-HPPD expression cassette of Figure 2.

Figure 4 shows a schematic diagram of the clone comprising the sequence depicted in SEQ ID No. 1. Figure 5 schematically depicts the production of a 4-HPPD containing plant expression cassette in which a PCR edited DNA fragment of Figure 4 is restricted with the enzymes *Nco*1 and *Kpn*1, then ligated into a vector (pMJB1) also restricted with *Nco*1 and *Kpn*1.

COR-AAD1-00001802

Figure 6 shows schematically the construction of a plasmid vector, used in *Agrobacterium* transformation and also includes maps of plasmids pJR1Ri and pGST-27Bin;

Figure 7 shows GST activity in transformed tobacco subjected to four herbicides

Figure 8 is a graph comparing damage to wild type plants and a GST-27 line following metolachlor treatment at 1400 g/ha for 3 weeks;

Figure 9 is a map of the plasmid pDV3-puc;

Figure 10 is a map of the plasmid pDV6-Bin;

Figure 11 is a map of plasmid pUB-1 containing the Ubiquitin promoter fragment PCRed from maize, a 2Kb fragment is cloned into pUC 19 and the junctions are sequenced to confirm the presence of the Ubiquitin promoter;

Figure 12 is a map of plasmid pIE98;

Figure 13 is a map of plasmid pIGPAT;

Figure 14 is a map of plasmid pCAT10;

Figure 15 is a map of plasmid pCAT11;

Figure 16 is a map of plasmid pPG6;

Figure 17 depicts part of the pMV1 plasmid.

EXAMPLE 1

Cloning of the 4-HPPD gene from *Pseudomonas* spp, transformation of the gene into plant material and the production of triketone herbicide resistant plants.

The amino acid sequence of 4-HPPD purified from *Pseudomonas fluorescens* PJ-874, grown on tyrosine as the sole carbon source is known. (Ruetschi *et al.*, Eur. J. Biochem 1992 **202**(2):459-466). Using this sequence minimally redundant PCR primers are designed with which to amplify a large but incomplete segment of the 4-HPPD gene from genomic DNA from a different bacterial strain (*Pseudomonas fluorescens* strain 87-79). The skilled man recognises what is meant by the term "minimally redundant primers", the redundancy being represented by squared brackets in the sequences depicted below. One example of each of the respective primers (corresponding to a 5' and 3' location within the HPPD gene) is given in each of SEQ ID Nos. 3 and 4.

Primer 1 (SEQ ID No. 5) which is a 17mer is designed from a knowledge of the sequence of amino acids 4-9 of the published protein sequence (see above) and Primer 2 (SEQ ID No. 6), likewise a 17mer, is designed from a knowledge of residues 334 to 339.

Primer 1 (HPPD-P4) has the sequence 5'TA[T/C] GA[G/A] AA[T/C] CC[T/C/G/A] ATG GG and primer 2 (HPPD-REV1) has the sequence 5'GC[T/C] TT[G/A] AA[G/A] TT[T/C/G/A] CC[T/C] TC. 100 ng genomic of DNA from Pseudomonas 87-79 was prepared using standard protocols and mixed with 100 pmol of each primer. The mixture is PCR amplified (35 cycles) using a Taq polymerase and other standard reagents under the following DNA synthesis and dissociation conditions:

94°C x 1.5 min

55°C x 2 min

74°C x 3 min

The amplified fragment comprises a region containing 3 codons from the 5' end, and about 30 codons from the 3' end of the coding region of the 4HPPD gene. The PCR product is blunt end cloned in the housekeeping vector pGEM3Z-f(+) using standard procedures.

Partial sequencing confirms that the cloned PCR fragment is 4-HPPD specific. The derived amino sequence contains several discrepancies compared with sequence published in respect of the *Pseudomonas fluorescens* PJ-874 enzyme. This partial fragment of the 4-HPPD gene gives negative hybridisation signals in genomic Southern blots on plant DNA under low stringency hybridisation/wash conditions. A 900 bp *Eco*R1 / *Eco*R1 fragment is excised from the centre of the previously cloned partial gene to use as a probe. Southern blots, using a variety of enzymes to restrict the genomic DNA, are hybridised with the radiolabelled fragment.

*Bcl*1 restricted DNA gives a single positive band of approx. 2.5 kb which is sufficient to contain the entire gene plus flanking regions of untranslated DNA. Genomic DNA is restricted with *Bcl*1 and electrophoresed on a preparative agarose gel. The region of digested DNA containing fragments in the size range 2 - 3 Kb is cut out and the DNA electro-eluted. The recovered DNA is cloned into the *Bam*H1 (which is compatible with *Bcl* 1) site of pUC18. Colony blots are probed with the 900 bp fragment and 12 positives are isolated. Minipreps are made from these, and cut with *Eco*R1 to look for the diagnostic 900 bp band. Of 12 colonies, 7 formed a brown pigment when grown overnight in LB to make the

minipreps, 5 of these are positive for the 900 bp band, the other 5 minipreps are negative and do not produce the brown pigment. The formation of the "brown pigment" is associated with the heterologous expression of a 4-HPPD gene.

Restriction analysis shows that the cloned insert was 2.5 kb long with about 1.2 kb DNA upstream of the 4-HPPD gene and 400 bp downstream. The ends of the gene are sequenced using appropriate primers and primers from pUC18. Such sequencing proves the gene to be intact and present in both orientations with regard to the pUC18 polylinker site.

SDS-PAGE on bacterial cell lysates shows that a new protein is present with a size of 40 kDa, which is correct for a 4-HPPD. A large band is present in extracts from cells having the gene inserted in a first orientation such that the gene is expressed from the plac promoter in the vector. No 40 kD band is obviously visible when the lysate is obtained from the cells in which the gene is in the opposite orientation, although both clones produced the brown pigment suggesting the presence of the active protein in both cell types. The 40 kDa recombinant protein is present in the soluble rather than the insoluble protein fraction. The clone in which the gene is in the second orientation is subjected to automated DNA sequence analysis to reveal the sequence depicted in SEQ ID No. 3. This sequence is edited to introduce several unique restriction sites to facilitate its assembly into a vector suitable for plant transformation work. The editing oligonucleotides, which are depicted in SEQ ID Nos. 7 and 8, are primer 3 (HPPDSYN1) 5'-
GTTAGGTACCAGTCTAGACTGACCATGGCCGACCAATACGAAAACC -3' and primer 4 (HPPDSYN2)
5'TAGCGGTACCTGATCACCCGGGTTATTAGTCGGTGGTCAGTAC-3'.

Expression of the *Pseudomonas* 4-HPPD gene in transgenic tobacco

The PCR edited DNA fragment is restricted with the enzymes *Nco*1 and *Kpn*1, then ligated into a vector (pMJB1) also restricted with *Nco*1 and *Kpn*1. pMJB1 is a pUC19 derived plasmid which contains the double CaMV35S promoter; a TMV omega enhancer and the NOS transcription terminator. A schematic representation of the resulting plasmid is shown in Figure 2. All of the DNA manipulations use standard protocols known to the man skilled in the art of plant molecular biology.

Bulk DNA is isolated and the 4-HPPD expression cassette (i.e. from the 2x35S to the nos 3' terminator), excised by partial restriction *Eco*R1 and then subjected to complete restriction with *Hin*d3. This is to avoid cutting at an *Eco*R1 site within the 4-HPPD gene. Following preparative agarose gel electrophoresis, the required DNA fragment is recovered by electro-elution.

The 4-HPPD expression cassette is then ligated in to the binary vector pBin19 restricted with *Hin*d3 and *Eco*R1. The structure of the resulting plasmid is shown schematically in Figure 3.

DNA is isolated and used to transform *Agrobacterium tumefaciens* LBA4404 to kanamycin resistance again using standard procedures. Leaf discs/slices of *Nicotiana plumbaginifolia* var Samsun are subjected to *Agrobacterium*-mediated transformation using standard procedures. Transformed shoots are regenerated from kanamycin resistant callus. Shoots are rooted on MS agar containing kanamycin. Surviving rooted explants are re-rooted to provide about 80 kanamycin resistant transformed tobacco plants. The presence of the 4-HPPD gene (using pre-existing EDIT primers) is verified by PCR. About 60 plants are PCR positive.

Explants (i.e. a leaf plus short segment of stem containing the axillary bud) are placed into MS agar (+ 3% sucrose) containing various concentrations of ZA1206 (a triketone herbicide) from 0.02 to 2 ppm. Untransformed tobacco explants are fully bleached at 0.02 ppm. They do not recover following prolonged exposure to the herbicide. In these particular experiments, only the shoot which develops from the bud is bleached, the leaf on the explanted tissue remains green.

About 30 of the PCR+ve transformed plants tolerated 0.1 ppm of ZA1296 (about 5x the level which causes symptoms on wild-type tobacco) with no indication of bleaching. They root normally and are phenotypically indistinguishable from untransformed plants. A sub-set of the transformants was tolerant to 0.2 ppm and a few transformants tolerate concentrations of up to 0.5 - 1 ppm. Again these plants look normal and root well in the presence of herbicide. Some of the transformed plants can be initially bleached when subjected to the herbicide at the said higher concentrations, but on prolonged exposure they progressively "green up" and "recover".

A subset of the said herbicide resistant transgenic plants are treated with the known herbicide Isoxaflutole [5-cyclopropyl-4-(2-methylsulphonyl-4-trifluoromethylbenzoyl)isoxazole or RPA 210772]. Such plants are even more resistant to this herbicide than they are to that designated as ZA1296 thus clearly indicating that the plants are cross resistant to multiple classes of 4-HPPD inhibitor.

EXAMPLE 2

Cloning of the 4-HPPD gene from *Synechocystis sp* into plant material and regeneration of the material to yield triketone herbicide resistant plants.

The genome of *Synechocystis sp*, PCC6803 has been sequenced. In order to introduce unique restriction sites to facilitate its assembly into a vector suitable for plant transformation work 100 ng of genomic DNA from *Synechocystis sp.* is prepared using standard protocols and mixed with 100 pmol of two primers suitable for the PCR amplification (35 cycles) of the sequence specified in SEQ ID No. 1, using a thermostable DNA polymerase preferably with proof reading activity and other standard reagents under appropriate DNA synthesis and dissociation conditions, the following being typical:

94°C x 1.5 min

55°C x 2 min

74°C x 3 min

The amplified fragment comprises a region containing the coding region of the 4-HPPD gene. The PCR product is blunt end cloned in a standard housekeeping vector, such as, for example, pGEM3Z-f(+) using standard procedures.

Automated DNA sequence analysis confirms that the cloned PCR product is 4-HPPD specific. Some of the transformed colonies harbouring the cloned 4-HPPD gene form a brown pigment when grown overnight in LB. The formation of the "brown pigment" is associated with the heterologous expression of a 4-HPPD gene (Denoya *et al* 1994 J. Bacteriol. 176:5312-5319).

SDS-PAGE on bacterial cell lysates shows that they contain a new protein having the expected molecular weight for the 4-HPPD gene product. In a preferred embodiment the recombinant protein is either present in the soluble rather than the insoluble protein fraction,

or else is manipulated to be so present. The clone is preferably subjected to automated DNA sequence analysis to confirm the absence of PCR derived artefacts.

Heterologous expression of the *Synechocystis sp*. PCC6803 4-HPPD gene in *E. coli*

The PCR edited DNA fragment is restricted with suitable enzymes such as *Nco*1 and *Kpn*1, for example then ligated into an *E coli* expression vector (such as the known pET series) appropriately restricted. All of the DNA manipulations use standard protocols known to the man skilled in the art of molecular biology.

Suitable host strains such as BL21(DE3) or other DE3 lysogens harbouring the said vector express quantities of HPPD enzyme sufficient to provide for their use in high through put screening to identify alternative 4-HPPD inhibitors. HPPD purified from the said transformed host strain may be used in the provision of antisera for the analysis of plants transformed with a polynucleotide encoding 4-HPPD.

Heterologous expression of the *Synechocystis sp*. PCC6803 4-HPPD gene in transgenic plants

The PCR edited DNA fragment is restricted with suitable enzymes such as *Nco*1 and *Kpn*1, for example then ligated into a suitable house keeping vector, such as pMJB1, to generate an expression cassette which contains an appropriate plant operable promoter and terminator. pMJB1 is a pUC19 derived plasmid which contains the double CaMV35S promoter; a TMV omega enhancer and the nos transcription terminator. A schematic representation of the resulting plasmid is shown in Figure 4.

The 4-HPPD expression cassette is then ligated in to the binary vector pBin19 restricted with *Hin*d3 and *Eco*R1. The structure of the resulting plasmid is shown schematically in Figure 5.

DNA is isolated and used to transform *Agrobacterium tumefaciens* LBA4404 to kanamycin resistance again using standard procedures. Potato and tomato tissue is subjected to *Agrobacterium*-mediated transformation using standard procedures. Transformed shoots are regenerated from kanamycin resistant callus. Shoots are rooted on MS agar containing kanamycin. Surviving rooted explants are re-rooted to provide about 80 kanamycin resistant transformed tobacco plants. The presence of the 4-HPPD gene (using pre-existing EDIT

COR-AAD1-00001808

primers) is verified by PCR. A substantial number of PCR positive plants are selected for further analysis.

Explants (i.e. a leaf plus short segment of stem containing the axillary bud) are placed into MS agar (+ 3% sucrose) containing various concentrations of ZA1206 (a triketone herbicide) from 0.02 to 2 ppm. Untransformed explants are fully bleached at 0.02 ppm. They do not recover following prolonged exposure to the herbicide. In these particular experiments, only the shoot which develops from the bud is bleached, the leaf on the explanted tissue remains green.

About 30 of the PCR+ve transformed plants tolerated 0.1 ppm of ZA1296 (about 5x the level which causes symptoms on wild-type tobacco) with no indication of bleaching. They root normally and are phenotypically indistinguishable from untransformed plants. A sub-set of the transformants is tolerant to 0.2 ppm and a few transformants tolerate concentrations of up to 0.5 - 1 ppm. Again these plants look normal and root well in the presence of herbicide. Some of the transformed plants can be initially bleached when subjected to the herbicide at the said higher concentrations, but on prolonged exposure they progressively "green up" and "recover".

A subset of the said herbicide resistant transgenic plants are treated with the known herbicide Isoxaflutole [5-cyclopropyl-4-(2-methylsulphonyl-4-trifluoromethylbenzoyl)isoxazole or RPA 210772]. Such plants are resistant to this herbicide and that designated as ZA1296 thus clearly indicating that the plants are cross resistant to multiple classes of 4-HPPD inhibitor.

EXAMPLE 3

Cloning of the GST gene into plant material and the generation of plants resistant to anilide and diphenyl ether type herbicides.

Plants Stocks of *Nicotiana tabacum* cv Samsum are kept on Musharige and Skoog medium (MS medium: MS salts (4.6 g/l) supplemented with 3% sucrose and 0.8% Bactoagar , pH 5.9). These plants, explants for the rooting assay and the seeds for the germination tests are grown in culture room at 25°C with 16 hours of light. When grown in

COR-AAD1-00001809

the glasshouse, the plants are transferred into compost (John Innes compost number 3, Minster Brand products).

Bacterial strains *Escherichia coli*, strain DH5 (GIBCO BRL), is: F^-, 80d*lacZ* M15, (*lacZ*YAargF)U169, *deo*R, *rec*A1, *end*A1, *hsd*R17(r_K^-, m_K^+), *sup*E44, $^-$*thi-gyr*A96*rel*A1. *Agrobacterium tumefaciens*, strain LBA 4404, is used to transform tobacco leaves.

Plasmids. DNA of GST-27 is inserted in the 2.961 kb pBluescript® II SK (+/-) phagemid designated pIJ21-3A (Jepson *et al* 1994). pJR1Ri is a 12.6 kb plasmid. The pJR1Ri plasmid contains a bacterial kanamycin resistance marker (KAN). It possesses the 2 repetitive sequences of 25 bp: the right (RB) and the left (LB) borders. The T-DNA contains a kanamycin resistance marker gene driven by NOS promoter. The GST-27 protein encoding sequence is expressed under the control of the CaMV 35S promoter.

Size markers. A 1kb DNA ladder is used as a DNA size marker (Bethesda Research Laboratories Life Technologies, Inc) when digestions and PCR (polymerase chain reaction) products are checked on an agarose gel. The Rainbow protein molecular weight markers (Amersham) are loaded on polyacrylamide gels for the Western analyses, as is known to the skilled man.

Chemicals. The active ingredients acetochlor, alachlor and metolachlor are produced at ZENECA Agrochemicals (UK), Jealott's Hill Research Station. The technical ingredients are formulated in ethanol and used in the HPLC assay, the rooting assay and the germination test (see below).

Plasmid construction. The plasmid pIJ21-3A containing the DNA gene of GST-27 is digested by the restriction enzyme *Eco*RI (Pharmacia) in 1 x Tris acetate (TA) buffer. Digestions are checked on a 0.8% agarose gel. *Eco*RI digested fragments are ligated into the *Sma* 1 (Pharmacia) site of pJR1Ri (Figure 6) after filling the protruding ends with the Klenow DNA polymerase (Pharmacia). The calf-alkaline-phosphatase (C.A.P.) enzyme prevents the self-ligation of pJR1Ri before the ligation of the GST gene. Competent *E. coli* cells (DH5) are transformed with the plasmid by a heat shock method. They are grown on L-agar and kanamycin plates. Positive colonies are checked by PCR or by hybridization overnight at 42°C with labelled probes (α -^{32}P dNTP). The melting temperature (Tm) of

the probes is defined by adding 2°C for each A or T and 4°C for each G or C. The reaction is performed at the lowest Tm-5 C with the *Taq* polymerase (Ampli-Taq DNA polymerase, Perkin Elmer Cetus) according to the manufacturer's protocols. PCR conditions are set up for 35 cycles as following: denaturation of DNA at 94°C for 48 seconds, annealing at the lowest Tm for 1 minute and extension at 72°C for 2.5 minutes. Prior to the first cycle, the reaction starts at 85°C.

Eight positive colonies are chosen and grown at 37°C on an overnight shaking L-broth and kanamycin culture. DNA from these cell culture is extracted and then purified from an ultracentrifugation at 50,000 rpm in a CsCl gradient.

The orientation of the insert into pJR1Ri is checked by sequencing the region between the 35S promoter and the GST gene, according to the Sanger method, using the Sequenase® (version 2.0, United States Biochemical corporation) following the manufacturer's protocols. The resultant plasmid (pGST-27Bin) (Figure 6) is introduced into *Agrobacterium tumefaciens* strain LBA4404, using the freeze thaw method described by Hostlers et al 1987.

Leaf transformation by *Agrobacterium*. The transformation of pGST-27Bin into tobacco is performed according to the method described by Bevan 1984. 3-4 weeks old sterile culture of tobacco (*Nicotiana tabacum* cv Samsum), grown on MS, are used for the transformation. The leaves are incubated on NBM medium (MS medium supplemented with 1 mg/l 6-benzylamino purine (6-BAP), 0.1 mg/l naphthalene acetic acid (NAA)) and kanamycin for 1 day. This medium enables the growth of shoot from leaf. One day later, the edges of the leaves are cut off and leaves cut into pieces. They are then co-incubated with the transformed *Agrobacterium* cells, containing the pJR1RI plasmid with the insert (pGST-27Bin), suspension (strain LBA 4404) for 20 minutes. The pieces are returned to the plates containing the NBM medium afterwards. After 2 days, explants are transferred to culture pots containing the NBM medium supplemented with carbenicillin (500 mg/l) and kanamycin (100 mg/l). Five weeks later, 1 shoot per leaf disc is transferred on NBM medium supplemented with carbenicillin (200 mg/l) and kanamycin (100 mg/l). After 2-3 weeks, shoots with roots are transferred to fresh medium. 2 cuttings from each shoot are transferred to separate pots. One is kept as a tissue culture stock, the other one is transferred

to soil for growth in the glasshouse after rooting. 42 independent transformants carrying the GST-27 construct are transferred to the glasshouse.

Leaf DNA extraction for PCR reactions. The presence of the transgene in the putative transformants is verified by PCR. Leaf samples are taken from 3-4 weeks old plants grown in sterile conditions. Leaf discs of about 5 mm in diameter are ground for 30 seconds in 200 µl of extraction buffer (0.5% sodium dodecyl sulfate (SDS), 250 mM NaCl, 100 mM Tris HCl.(pH 8). The samples are centrifuged for 5 minutes at 13,000 rpm and afterwards 150 µl of isopropanol is added to the same volume of the top layer. The samples are left on ice for 10 minutes, centrifuged for 10 minutes at 13,000 rpm and left to dry. Then they are resuspended in 100µl of deionised water, 15µl of which is used for the PCR reaction. PCR is performed using the conditions described by Jepson *et al.* (1991). Plants transformed with GST-27 DNA are analysed with the primer GST II/7 (AACAAGGTGGCGCAGTT) (SEQ ID No. 10) specific to the 3' region of GST-27 region and NOS 3 (CATCGCAAGACCGGCAACAG) (SEQ ID No. 11) specific to the NOS terminator. 39 of the 42 primary transformants provide a 310 bp fragment by PCR.

Western blot analysis. To verify the heterologous expression of GST-27 in tobacco Western blot analysis is performed. 120 mg of leaf from 3-4 weeks old plants grown in sterile conditions are ground at 4°C in 0.06 g of polyvinylpoly-pyrolidone (PVPP) to adsorb phenolic compounds and in 0.5 ml of extraction buffer (1 M Tris HCl, 0.5 M EDTA (ethylenediamine-tetraacetate), 5 mM DTT (dithiothreitol), pH 7.8). An additional 200 µl of extraction buffer is then added. The samples are mixed and then centrifuged for 15 minutes at 4°C. The supernatant is removed, the concentration of protein being estimated by Bradford assay using BSA as the standard. The samples are kept at -70°C until required.

Samples of 5µg of protein with 33% (v/v) Laemmli dye (97.5% Laemmli buffer (62.5 mM Tris HCl, 10% w/v sucrose, 2% w/v SDS, pH 6.8), 1.5% pyronin y and 1% -mercaptoethanol) are loaded on a SDS-polyacrylamide gel (17.7% 30:0.174 acrylamide:bisacrylamide), after 2 minutes boiling. Protein extracts are separated electrophoretically in the following buffer (14.4% w/v glycine, 1% w/v SDS, 3% w/v Tris Base). Then they are transferred onto nitro-cellulose (Hybond-C , Amersham) using an

electroblotting procedure (Biorad unit) in the following blotting buffer (14.4% w/v glycine, 3% w/v Tris Base, 0.2% w/v SDS, 20% v/v methanol) at 40 mV overnight.

Equal loadings of proteins are checked by staining the freshly blotted nitrocellulose in 0.05% CPTS (copper phtalocyanine tetrasulfonic acid, tetrasodium salt) and 12 mM HCl. Then the blots are destained by 2-3 rinses in 12 mM HCl solution and the excess of dye removed by 0.5 M $NaHCO_3$ solution for 5-10 minutes followed by rinses in deionised water. Filters are blocked for 1 hour with TBS-Tween (2.42% w/v Tris HCl, 8% w/v NaCl, 5% Tween 20 (polyxyethylene sorbitan monolaureate), pH 7.6) containing 5% w/v BSA. Then they are washed for 20 minutes in TBS-Tween supplemented with 2% w/v BSA. Indirect immunodetections are performed with a 1:2000 dilution of a sheep GST-27 antiserum as first antibody and with a 1:1000 dilution of a rabbit anti-sheep antiserum as second antibody, associated with the horseradish peroxidase (HRP). Any excess of antiserum is washed with TBS-Tween supplemented with 2% w/v BSA. ECL (enhanced chemiluminescence) detection is performed using the protocols described by Amersham. Any background is eliminated by additional washes of the membranes in the solution mentioned above.

An estimation of the level of expression of the GST gene is performed on the LKB 2222-020 Ultroscan XL laser densitometer (Pharmacia). Western analysis reveals 8 of the PCR positive primary transformants show no detectable GST-27 expression. The remaining 31 show expression levels which vary from barely detectable to high levels equating to 1% of total soluble protein as determined from signals detected with pure maize GST II samples.

Southern blot analysis. The pattern of integration of transgenes is verified by Southern blot analysis. 2.5 g of fresh tobacco leaf taken from plants grown in glasshouse, placed into a plastic bag containing 0.75 ml of extraction buffer (0.35 M sorbitol, 0.1 M Tris HCl, 0.005 M EDTA, 0.02 M sodium meta bisulphite, pH 7.5), are crushed by passing through the rollers of a "Pasta machine". Crushed extracts are then centrifuged for 5 minutes at 6000 rpm at room temperature. After discarding the supernatant, the pellet is resuspended in 300µl extraction buffer and 300µl nuclei lysis buffer (2% w/v CTAB), 0.2 M Tris HCl, 0.05 M EDTA, 2 M Nacl, pH 7.5). 120µl of 5% Sarkosyl is added and the

COR-AAD1-00001813

samples placed in a 65°C water bath for 15 minutes. Extracts are centrifuged for 5 minutes at 6000 rpm after adding 600 µl 24:1 chloroform:isoamyl alcohol. 700 µl of isopropanol is added to the same volume of supernatant and centrifuged for 10 minutes at 13,000 rpm. Then the pellet is washed with 70% ethanol and left to air dry. The pellet was left overnight at 4^0C in 30 µl TE (10 mM Tris HCl, 1 mM EDTA) to resuspend. Samples are kept at -20°C until required.

Total leaf DNA is digested for 6 hours at 37°C with the following restriction enzymes *Sac*I and *Xba*I in 1 x Phor-one-all buffer (20 mM Tris acetate, 20 mM magnesium acetate, 100 mM potassium acetate, Pharmacia) for the extracts from the plants containing the GST-27 gene. DNA is fractionated on a 0.8% agarose gel, denatured by gently shaking in 0.5 M NaOH, 1.5 M NaCl for 30 minutes and the gel is neutralized by shaking in 0.5 M Tris HCl, 1.5 M NaCl for 75 minutes. Then the DNA is transferred onto an Hybond-N (Amersham) nylon membrane by capillary blotting in 20 x SSC (3M Nacl, 0.3M Na_3citrate). DNA is fixed to membranes using a combination of UV strata linking (Stratagene) and baking for 20 minutes at 80°C. Probes are excised from plasmids, used for *Agrobacterium* transformation, containing the GST-27 gene by digestion with *Eco*RI. The probe is labelled with α -^{32}P dNTP (3,000 Ci/mM) using the Prime-a-Gene kit (Promega), random priming protocol described by Feinberg and Vogelstein. Positive controls are prepared by digestion of pIJ21-3A with *Sac*I and *Eco*RI.

Prehybridisations are performed in 5 x SSPE (0.9 M Nacl, 0.05 M sodium phosphate, 0.005 M EDTA, pH 7.7), 0.5% SDS, 1% w/v Marvel (dry milk powder), 200 µg/ml denaturated salmon sperm DNA for 3-4 hours at 65°C. Hybridizations are performed in the same buffer but without the last ingredient. Membranes are washed for 30 minutes at 65°C in 3 x SSC, 0.5% SDS, and twice in 1 x SSC, 0.1% SDS for 20 minutes prior to autoradiography at -70°C.

HPLC assay. To verify the GST-27 expressing plants show GST activity against herbicide substrates an in vitro herbicide assay is performed using HPLC. 1g of leaf tissue is taken from 3-4 month old flowering tobacco plants growing in the glasshouse, and ground in liquid nitrogen and 7 ml of extraction buffer (50mM glycyl glycine, 0.5mM

EDTA, 1 mM DTT, pH 7.5). Extracts are transferred to centrifuge tubes containing 0.1g of PVPP and centrifuged at 16,500 rpm for 30 minutes at 4°. 2.5ml of supernatant is loaded onto Sephadex G-25 (PD10) column (Pharmacia) and eluted with 3.5ml of sodium phosphate buffer (50 mM, pH 7.0) containing 1 mM EDTA and 1 mM DTT. Protein estimation is performed by the Bradford method using BSA as the standard. Extracts are divided into aliquots and kept at -70°C until required. HPLC assays are performed on a Spherisorb 5μ ODS2 column (25 cm * 4.6 mm i.d., manufacturer: Hichrom) using 65:35 acetonitrile:1% aqueous phosphoric acid mobile phase at the rate of 1.5 ml/min. Detection of the compounds is performed on a UV LC-6A Schimadzu detector (wavelength 200 nm).

Reactions are carried out in 0.8 ml HPLC vials at room temperature (20-25°C). 15-94% by volume of plant extract are added to the sodium phosphate buffer (pH 7), 5 mM glutathione or homoglutathione and 2 or 20 ppm of compound (2 ppm for fluorodifen, 20 ppm for acetochlor, alachlor and metolachlor). Controls are also set up in the same proportions but extracts replaced by the sodium phosphate buffer. Reactions are initiated by addition of the herbicide used as substrate. Compound reactivity is monitored for a maximum of 9-19 hours. Specific retention times and peak areas are calculated by the JCL 6000 chromatography data system package (Jones chromatography). HPLC peak area versus time profiles, based on 7-11 time points, are measured for each compound. Half-life and pseudo first-order rate constant data are obtained from exponential fits of corrected peak area versus time data. These data are mastered with the FIT package version 2.01.

Using the methodology described above, the GST activity of the transformed plants is assayed against different herbicide substrates. These herbicides consist of 3 dichloroacetanilides (acetochlor, alachlor, metolachlor) and a diphenyl ether (fluorodifen). These chemicals are known to be conjugated to glutathione, in particular dichloroacetanilides. Extractions are performed in the presence of PVPP and at low temperature to limit denaturation of proteins. Studies on GST stability show that maize GST activity is reduced by 73% in crude extracts when stored at -20°C. Therefore it was decided to divide extracts in aliquots. They were kept at -70°C until required. Each sample was defrosted only once, overnight on ice. The assay is performed within 2 weeks following the extraction.

Concentrations of herbicide in the HPLC vials are set according to their solubility limits. Acetochlor, alachlor and metolachlor were assayed at 20 ppm and fluorodifen at 2 ppm. The assay is run for 9-19 hours according to the reactivity of the herbicide. Metolachlor is assayed for a longer period of time, because its half-life is high under these conditions. Detection of the compounds is performed on a UV detector at 200nm. Specific retention times and peak area are monitored for the herbicide. The GST activity is calculated on the basis of 7-11 time points. Enzymatic conjugation follows an exponential decrease curve. The decrease of the peak area of the assayed herbicide is used for the calculation of the GST activity. The half-life and the first order rate constant are also calculated.

Five tobacco lines are assayed including a wild-type (negative control), 4 GST-27 lines 5, 6, 12 and 17. They are chosen because of their high expression as determined by western analysis. To limit any rapid conjugation before monitoring, the herbicide is added last. The GST-27 line 17 is also assayed for conjugation of acetochlor to homoglutathione. Results are reported in Figure 7 and show GST-27 expressing plants exhibit activity against chloroacetanilide herbicides *in vitro*.

In summary: transgenic tobacco plants express the GST-27 protein and these plants may be distinguished by their relative activities *in vitro* against herbicide substrates.

***In vivo* analysis - Rooting assay**. The GST-27 lines have significant activity *in vitro* against at least 3 chloroacetanilides. Moreover, most of the herbicides of this class are known to inhibit root elongation. Therefore, it is decided to set up a rooting test on acetochlor, alachlor and metolachlor.

A pilot experiment is set up to find out the most effective concentrations. A range of 7 concentrations is chosen: 0, 1, 5, 10, 20, 40 and 100 ppm. Two transformed lines (GST-27 lines 6 and 17) and a wild-type tobacco are tested on alachlor. Lines 6 and 17 are chosen because they represent the lowest and the highest expressing plants, based on western blot analysis. Three explants, consisting of a leaf attached to a piece of shoot, are transferred onto MS medium supplemented with the herbicide. Root growth is observed after 2 weeks (Figure 8). On the general aspect of the plants, an effect of the herbicide is observable on the wild-type from the concentration 1 ppm, the leaves are more yellowish and smaller.

With the increase of the concentrations, these effects are greater and the number of new leaves is reduced. From 10 ppm, the plants do not produce new leaves. In contrast, with respect to the transformed lines, the effect of the herbicide is observable from the concentration 20 ppm for line 2 and 40 ppm for line 6. Between these concentrations, the leaves seem smaller and their number slightly reduced, but they still are green. Secondly, the wild-type produces some roots up to 5 ppm, but their length decreases dramatically between the concentrations 0 and 5 ppm. Regarding the lines 2 and 6, roots are respectively produced up to 10 ppm and 20 ppm, with the decrease of their length for lower concentration. Under these conditions and after 2 weeks, it is noticeable that the concentration limiting the rooting is between 20 and 40 ppm for the "best" line tested at this stage of this experiment.

A subsequent experiment is set up for a wild-type (control), 4 GST-27 (lines 5, 6, 12 and 17). These plants are assayed on acetochlor, alachlor and metolachlor at the following rates: 0, 10, 20, 40 ppm for the acetochlor and metolachlor mentioned herbicide, and 0, 20, 40, 100 ppm for alachlor. These concentrations are chosen because on HPLC the plants show the lowest activity against acetochlor and metolachlor. The same conditions are used: 3 explants per concentration and per line transferred onto MS medium supplemented with herbicide. The observations of the root growth are taken 3 weeks after the beginning of the assay.

As for the pilot experiment the response of the explants in each pot is generally uniform. On acetochlor, the wild-type explants do not show any rooting or any production of new leaves in the presence of herbicide. But the GST-27 lines 6 and 17 produce few roots at 10 and 20 ppm and small leaves as well. The lines 5 and 12 are not as resistant as these 2 lines. On alachlor, the wild-type does not produce any root for the tested rates, but some leaves at 20 ppm. Lines 6 and 17 produce roots up to the concentration of 40 ppm, which roots appear not to be affected by the herbicide. The number of roots seems lower with increasing concentrations of herbicide. For these lines, the rooting concentration limit is between 40 and 100 ppm under these conditions and after 3 weeks. Lines 9 and 16 do not produce any roots but very tiny leaves at 20 and 40 ppm of the herbicide. On metolachlor, the wild-type tobacco produces very few tiny roots at 10 and 20 ppm. Lines 6 and 17 produce short roots, but not as many as are produced on alachlor. For this herbicide, the

rooting concentration limit is between 20 and 40 ppm for the line 6 and more than 40 ppm for line 17.

Treatment of plants with Herbicide. To demonstrate that transgenic plants expressing GST-27 confer resistance to herbicide treatment, pre- and post-emergence herbicide trials are performed in the glasshouse.

Pre-emergence tests are performed by sowing approximately 50 seeds per line for each rate of herbicide in sand (25 % sifted loam, 75 % grift, slow release fertiliser). Four replicates are treated for each chemical rate. Herbicide (0, 300 and 350 g/ha), formulated in 5 % JF 5969 (905.6 g/L cyclohexanone, 33.33 g/L synperonic NPE1800 and 16.7 g/L Tween 85) are applied to seed trays using a tracksprayer. Seeds are left to germinate in the glasshouse and germination is scored after 3 weeks. Results for alachlor show that the transgenic plants are resistant to the pre-emergent application of the herbicide. Similar results are obtained for acetochlor, metolachlor and EPTC (12000g/ha).

Post-emergence tests are carried out by sowing 28 seeds per line and per herbicide rate in compost. After 16 days tobacco plants (1 cm high) are sprayed with alachlor in 5 % formulation JF 5969 using a tracksprayer. Damage is scored 3 weeks following spray treatment using size of the plants, necrosis, apex condition, morphology of leaves relative to unsprayed control. A score of 100 % damage means that the plant is killed by the herbicide and a score of 0% means that the plant resembled an untreated control. Post-emergent results for alachlor demonstrate that the transgenic plants are resistant to this herbicide. Damage to wild type plants and a segregating GST-27 line, is recorded graphically in Figure 9 following metolachlor treatment at 1400 g/ha. Similar studies are performed with acetochlor at 2000 g/ha giving similar results.

EXAMPLE 4

Cloning of glyphosate resistance genes into plant material and the generation of glyphosate resistant plants

A summary of the cassettes and specific plant transformation constructs used in this example is shown in the Figures of European Patent Application No. EP A1 536330.

Dicot Vector 1 Vector 1 is a constitutive control plasmid containing the glyphosate oxidase gene (GOX) fused to the chloroplast transit (CTP) sequence 1 from the Rubisco gene of *Arabidopsis* driven by the enhanced 35S CaMV promoter. The construct contains the omega translational enhancer 5' of the CTP encoding sequence. Vector 1 utilises the NOS terminator. The CTP-GOX construct is synthesised to according to the sequence disclosed in WO92-00377 with the addition of an *Nco* I site at the translation start ATG, and a *Kpn* I at the 3' end. Internal *Sph* I sites and *Nco*I site are deleted during synthesis with no change in the protein sequence. The CTP-GOX sequence is isolated as an *Nco* I *Kpn* I fragment and ligated using standard molecular cloning techniques into *Nco* I *Kpn* I cut pMJB1, a plasmid based on pIBT 211 containing the CaMV 35 promoter with duplicated enhancer linked to the tobacco mosaic virus translational enhancer sequence which replaces the tobacco etch virus 5' non-translated leader, and terminated with the NOS terminator.

A cassette containing the enhanced CaMV35S promoter-Omega enhancer- CTP-GOX-Nos sequence is isolated as a *Hind* III *Eco*RI fragment and ligated into *Hind* III *Eco*RI cut pJRIi, a pBin 19 based plant transformation vector.

Dicot Vector 2. The CP4- EPSPS (which is a class II EPSPS) fused to a chloroplast transit peptide from Petunia is synthesised according to the sequence depicted in WO92-04449 with an *Nco*I site at the translation initiation ATG. An internal *Sph* I site in the EPSPS is silenced with no change in protein sequence. A fragment containing the synthetic CTP-EPSPS sequence is isolated as a *Nco*I *Sac* I fragment and ligated into pMJBI. This sequence is placed under expression control of an enhanced 35S promoter and NOS terminator with an Omega fragment being positioned 5' of the protein encoding regions and isolated as an *Eco*RI *Hind* III fragment which is cloned into pJRIi to give dicot vector 2.

Dicot Vector 3. A control vector with both EPSPS and GOX genes is constructed by cutting dicot vector 2 with *Eco*RI and inserting an *Eco*RI - *Sph* I - *Eco*RI linker. The resultant vector is cut with *Sph* I to liberate a cassette ("B"), which is cloned into an *Sph* I site in dicot vector 1, 5' to the promoter to form pDV3puc (Figure 9). The coding regions, including promoters and terminators derived from vectors (1) and (2) are then excised from pDV3puc as a *Hind* III and *Eco*RI fragment and cloned in to pJRIi .

Plasmid pDV3 in the binary vector pJR1i is introduced into tobacco by *Agrobacterium* mediated transformation using known techniques. 270 Shoots are removed from calli obtained from the thus transformed material, 77 of which rooted. To confirm the presence of the EPSPS and GOX genes in the thus rooted shoots, DNA extracts are prepared from pDV3 plants and analysed by PCR using the following primers:

3' end EPSPS gene

GATCGCTACTAGCTTCCCA (SEQ ID No. 12) EPSPS 2

5' end GOX gene

AATCAAGGTAACCTTGAATCCA (SEQ ID No.13) GOX 1

PCR reactions provide a 1.1 kb band if both genes are present. To confirm the functionality of the glyphosate tolerance genes pDV3 tissue culture explants are transferred to MS media containing 0.01 mM and 0.05mM glyphosate. Plants are scored two weeks following transfer to medium containing glyphosate. Resistant lines, which grow successfully on herbicide-containing media, are analysed by Western using anti-sera raised in rabbits against purified GOX and EPSPS.

Leaf DNA extracts are prepared from each primary transformant and used for PCR reactions to confirm the presence of the vector. Western blot analysis is performed on each PCR positive pDV3 plant to verify the heterologous expression of GOX and EPSPS, using the methods described earlier. High level expressors are self- pollinated and seed screened on kanamycin plates. Single locus plants are kept for homozgote production. Data confirming that plants transformed with the pDV3 construct are resistant to glyphosate is to be found in Example 8.

EXAMPLE 5

Production of plants which are resistant to anilide type herbicides and glyphosate

Heterozygous and homozygous tobacco lines expressing GOX and EPSPS are cross-pollinated onto homozygous tobacco lines expressing GST-27. The seed generated in this cross are sown and leaf material taken for western analysis, using the procedures described earlier. Protein extracts from GST-27 western positive plants are then screened with the GOX/EPSPS antibody to select lines expressing both GST-27, GOX and EPSPS. These lines are then used in pre-emergent herbicide sprays with acetochlor, alachlor, metoloachlor

COR-AAD1-00001820

and EPTC. Subsequently, the plants can be sprayed in a post-emergent manner with formulated glyphosate.

EXAMPLE 6

Production of plants which are resistant to both anilide and glyphosate type herbicides by a process not involving cross-pollination

The vector pDV3puc is cut with EcoRI, phenol chloroform extracted and precipitated. A delta EcoRI- HindIII- EcoRI linker MKOL3 5'AATTACGGAAGCTTCCGT3 ' (SEQ ID No.14) is heated to 70^{0}C and cooled to room temperature allowing it to self-anneal. The annealed linker is then ligated into *Eco*RI cut pDV3puc. Putative recombinants are screened with end labelled oligonucleotide MKOL3. Plasmid DNA is isolated from positively hybridising colonies. Restriction digestion with HindIII release a 5.4 kb fragment containing the 35S CaMV promoter driving expression of Omega-CTP2- EPSPS- NOS and the 35S CaMV promoter driving expression of Omega-CTP1-GOX-NOS. This fragment is cloned in to pGST-27 Bin cut with *Hind*III and dephosphorylated with CIP. Recombinants are selected using an insert probe. The resultant vector pDV6-Bin (Figure 10) is verified by appropriate sequence analysis The resultant plasmid is transformed into Tobacco via *Agrobacterium* using known techniques. 270 Shoots are recovered following transformation, 80 of which are rooted. Leaf DNA extracts are prepared from each primary transformant and are used in PCR reactions to confirm the presence in the leaf of the protein encoding regions of the vector. The primers are as indicated above (SEQ ID Nos. 12 and 13). To confirm the functionality of the trans-genes, primary transformants from pDV6-Bin are assessed on 0, 0.0lmM an 0.05mM glyphosate and 10ppm and 40ppm alachlor in tissue culture medium. A number of transgenic grow successfully on both media under conditions in which the wild type controls fail to. Western blot analysis is performed on each PCR positive plant to verify the heterologous expression of GOX and EPSPS and GST-27, using the methods described earlier. These lines are then used in pre-emergent herbicide sprays with acetochlor, alachlor, metoloachlor and EPTC. Subsequently, the plants can be sprayed in a post-emergent manner with formulated glyphosate.

Table 1 below gives the data for the pre-post herbicide treatments of DV6 plants ie plants expressing both glyphosate resistance genes and GST. The top half of the table shows the rates at which the pre-em herbicides are applied and their continued state in the absence of post-em herbicide application. The lower half of the top table gives the damage incurred after a glyphosate treatment of 800g/ha. The lower table shows the replicate scores for damage inflicted on the plants not subjected to a pre-em treatment as a result of the post-em glyphosate treatment. All replicates of the wild type plants score similarly whereas the transgenic scores reflect the fact that this was a segregating population ie azygous plants not expressing transgenes are able to go through to the post-em spray test.

Table 1 MEAN DATA FOR POST EM HERBICIDE TREATMENT

21 DAT

Post treatment		Pre treatment			% Phytotoxicity	
Chemical	Rate	Chemical	Rate	Pdv6 #2	pDV6 #71	Wild type
None		None	-	0	0	0
		Acetochlor	50	0	0	-
		Metolachlor	300	0	0	-
		Alachlor	400	0	0	-
		Dimethenamid	50	0	0	-
		Cycloate	5000	0	0	-
		EPTC	5000	0	0	-
		Bayer FOE 5043	50	0	0	-
		Tetrazolinone	200	-	0	
Glyphosate	800	None	-	18.75	48.75	86.25
360 g/l		Acetochlor	50	0	0	-
		Metolachlor	300	0	0	-
		Alachlor	400	0	0	-
		Dimethenamid	50	0	0	-
		Cycloate	5000	0	0	-
		EPTC	5000	0	0	-
		Bayer FOE 5043	50	0	0	
		Tetrazolinone	200	-	0	-

Post treatment		Pre treatment				
Chemical	Rate	Chemical	a	b	c	d
Glyphosate	800	None	75	0	0	0
			100	0	0	95
			90	90	90	75

FOE 5043 is an oxyacetamide known as fluthiamide.

EXAMPLE 7

Production of maize which is resistant to glufosinate and anilide type herbicides.

A monocotyledonous (maize, wheat) transformation vector containing GST-27, conferring resistance to pre-emergence herbicides, and phosphinothricin acetyl transferase (PAT), conferring resistance to the post-emergence herbicide glufosinate is generated as follows:

Step 1: Digest pUB1 (a pUC based vector containing the maize ubiquitin promoter and intron) (Figure 11) with *Hind* III. Into the gap produced by the digestion is inserted a *Hind*III-*Age* I-*Hind*III linker (5' AGCTTGTACACCGGTGTACA 3' (SEQ ID No. 15)). The result recombinant vector is designated as pUB2.

Step 2: The GST-27 cDNA is excised from pIJ21-3A using *Kpn* I and *Bam*HI and cloned into *Bam*HI and *Kpn*I cut pUB2 to form pUB3.

Step 3: A *Kpn*I-*Pac* I-*Kpn*I linker (5' CGGACAATTAATTGTCCGGTAC 3' (SEQ ID No. 16)) is self annealed and cloned into *Kpn*I cut pUB3 to form pUB4.

Step 4: The NOS terminator is isolated as a *Sma*I fragment from pIE98 (Figure 12), and blunt end cloned into *Eco*RV cut pUB4 to form pUB5. The orientation of the NOS terminator in pUB5 is confirmed by restriction digestion with *Eco*RI and *Bam*HI. All junctions are sequencesd to confirm the correct insertion of the various construct components.

Step 5: The ubiquitin GST-27 NOS cassette present in pUB5 is removed from it by digestion with *Age* I and *Pac*I and is cloned in the ampicillin minus vector pIGPAT (Figure 13) which contains the PAT gene under the control of the 35S-CaMV promoter. Recombinants are detected by colony hybridisation with an *Eco*RI cDNA insert from pIJ21-3A. Recombinants are detected by colony hybridisation with an EcoR1 cDNA insert from pIJ21-3A. Recombinants are orientated with *Nco* I restriction digestion to form pCAT10 (Figure 14).

Step 6: The 35S-PAT-NOS cassette is removed by digestion with AscI and the AscI ubiquitin-PAT-NOS cassette from pPUN 14 inserted to form pCAT11 (Fig 15). pCAT11 is transformed into wheat and maize using known whiskers and particle bombardment technology. The cells are then transferred into bialophos-containing media to select callus

material which expresses the PAT gene. Calli which grows on media containing this herbicide are then subjected to PCR using the following primers (SEQ ID Nos. 33 and 34 respectively) to conform the presence in the calli of the GST-27 gene.

5'CCAACAAGGTGGCGCAGTTCA3' (SEQ ID No. 33)

5'CATCGCAAGACCGGCAACAG3' (SEQ ID No. 34).

The calli which contain the GST-27 expression cassette are transferred to plant regeneration media and maize plants are recovered. The transformed maize plants are confirmed - by Western blots of total protein extracts from leaves - to constitutively express the GST gene at high levels. Such plants are cross pollinated with an elite maize inbred line and seed is recovered. To confirm enhanced tolerance of the plants to the herbicide acetochlor the said seeds are planted in soil to which has been applied between 2,000 and 8,000 grams per hectare of the herbicide. The seeds are allowed to germinate and grow for 7 days after which time a sample of the resultant seedlings is assessed for damage caused the chemical and compared to the seedlings (if any) which result from non-transgenic seed sown under identical conditions. The "transgenic" seedlings and non-transgenic control seedlings grown in soil treated with the herbicide and a corresponding safener exhibit little, if any damage, whereas non-transgenic seedlings grown in soil which contains herbicide in the absence of safener show very substantial damage. Seedlings which survive the first herbicide treatment are allowed to grow for a further 20 days or so, and then sprayed with a commercial mix of glufosinate at various concentrations ranging from about 0.1 to 1% active ingredient. The seedlings which contain the PAT gene (expression of which is determined by the method described by De Block M.*et al* (The EMBO Journal 6(9): 2513-2518 (1987)) are either completely resistant to glufosinate, or are relatively tolerant of the herbicide - depending upon the concentration applied - when compared with seedlings which do not contain the said gene.

EXAMPLE 8

Production of plants (mono and dicots) which are resistant to both glyphosate and glufosinate

This example demonstrate the production of plants which are resistant to both glufosinate and glyphosate. This multiple herbicide resistance results from the crossing of a

first plant which has been engineered to be resistant to glufosinate with a second plant which has been engineered to be resistant to glyphosate.

Production of a Glufosinate resistance construct pPG6

pPG6 is a Bin 19 based vector derived from pBin19RiPAT, and contains a cassette containing the 35S CaMV promoter driving the GUS gene. Inserted between the promoter and GUS is the second intron of the ST-LS1 gene. This sequence is 189bp, has an A/T content of 80%, typical splice junctions and stop codons in all three reading frames. The presence of the intron prevents expression of GUS in *Agrobacterium* as splicing does not occur in prokaryotes. It also contains a cassette carrying the 35S CaMV promoter driving expression of the PAT gene. Fig 16 shows a map of pPG6.

Glyphosate resistance constructs

Dicot vectors 1-3 are produced as indicated above in Example 4.

Monocot vector 1 Monocot vector 1 is a plasmid containing both CTP1 GOX and CTP2 EPSPS, both driven by the maize polyubiquibitin promoter and enhanced by the maize polyubiquitin intron 1, in a pUC derived plasmid. It also contains a cassette conferring tolerance to phosphinothricin.

Plasmid 1: The vector pUB1 is digested with Kpn1 and a Kpn1-Not1-Kpn1 linker inserted, (sequence 5' CAT TTG CGG CCG CAA ATG GTA C 3 - SEQ ID NO. 17). An EcoR1-Not1-EcoR1 linker (5' AAT TCA TTT GCG GCC GCA AAT G 3' (SEQ ID No.18) is inserted into the EcoR1 site of DV1-pUC. The resulting plasmid is cut with Nco1 and the 5' overhang filled using DNA Polymerase 1 Klenow fragment. The linear vector is then digested with Not1 and a Not1-blunt fragment isolated. This fragment, containing the CTP1-GOX and NOS sequences is ligated into Sma1-Not1 digested modified pUB1. A Hind111-Not1-Hind111 linker (sequence 5' AGC TTG CAG CGG CCG CTG CA 3' (SEQ ID No. 19)) is inserted into the plasmid to give resulting plasmid 1.

Plasmid 2: An EcoR1-Not1-EcoR1 linker (5' AAT TCA TTT GCG GCC GCA AAT G 3' (SEQ ID No. 20)) is inserted into the EcoR1 site of DV2-pUC (another clone is isolated which does not contain the linker mentioned above, thus allowing this cloning strategy).The resulting plasmid is digested with Nco1 and the 5' overhang filled using DNA Polymerase 1 Klenow fragment. The linear vector is then cut with Not1 and the resulting

fragment is cloned into the same vector as described immediately above(pUB1 modified), to generate plasmid 2. The PAT selectable marker cassette, comprising 35S CaMV promoter, Adh1 intron, phosphinothricin acetyl transferase (PAT) gene and nos terminator is excised from pIE108 and cloned into the Hind111 site of plasmid 2 to give cassette 2. Diagnostic restriction analysis is used to confirm that the PAT cassette was in the same orientation as the CTP2 EPSPS cassette.

The cassette carrying the polyubiquitin promoter and intron, CTP1 GOX and nos terminator is excised from plasmid 1 on a Not1 fragment and ligated into Not1 cut cassette 2 to give monocot vector 1, pMV1 (Fig. 17).

Tobacco transformation Plasmids for dicot transformation are transferred to *Agrobacterium tumefaciens* LBA4404 using the freeze thaw method of Holsters et al (1978). *Nicotiana tabaccum var Samsun* is transformed using a leaf disc method described by Bevan et al (1984). Shoots are regenerated on medium containing 100mg/l kanamycin. After rooting and selection plants are transferred to the glass house and grown under 16hr light 8hr dark regime. Transformants of pPG6 are named as 35S-PAT lines.

Maize transformation Maize transformation is performed using the particle bombardment method as described by Klein et al (1988). Selection of the transformed material is on 1mg/l bialophos.

PLANT ANALYSIS

PCR This analysis is performed on all tobacco lines which rooted in tissue culture and maize calli. DNA is extracted by means known to the skilled man. The primary transformants are analysed using the following oligonucleotides:-

pDV1	TMV1 + GOX1,	GOX3 + nos1	
pDV2	TMV1 + EPSPS1,	EPSPS1 + nos1	
pDV3	EPSPS3 + GOX3		
pPG6	35S + BARJAP2R		
pMV1	GOX 4 + GOX5	EPSPS4 + EPSPS5	35S + BARJAP2R

The sequences of the oligonucleotides are:-

TMV1 5' CTCGAGTATTTTTACAACAATTACCAAC (SEQ ID No. 21)
GOX1 5' AATCAAGGTAACCTTGAATCCA (SEQ ID No. 22)
GOX3 5'ACCACCAACGGTGTTCTTGCTGTTGA (SEQ ID No. 23)
nos1 5' GCATTACATGTTAATTATTACATGCTT (SEQ ID No. 24)
EPSPS1 5' GTGATACGAGTTTCACCGCTAGCGAGAC (SEQ ID No. 25)
EPSPS3 5' TACCTTGCGTGGACCAAAGACTCC (SEQ ID No. 26)
EPSPS4 5' ATGGCTTCCGCTCAAGTGAAGTCC (SEQ ID No. 27)
EPSPS5 5' CGAGACCCATAACGAGGAAGCTCA (SEQ ID No. 28)
GOX4 5' ATTGCGTGATTTCGATCCTAACTT (SEQ ID No. 29)
GOX5 5'GAGAGATGTCGATAGAGGTCTTCT (SEQ ID No. 30)
35S 5' GGTGGAGCACGACACACTTGTCTA (SEQ ID No. 31)
BARJAP2R 5' GTCTCAATGTAATGGTTA (SEQ ID No. 32)

PCR +ve plants are selected for further analysis.

Selection on glyphosate A kill curve is constructed for growth of tobacco in tissue culture on glyphosate containing medium. This is done by inserting a stem segment ~6mm long and carrying a leaf node into MS medium containing a range of glyphosate isopropylamine concentrations. Four/five stem segments are used at each concentration. The results are scored after two weeks and are shown in Table 2.

Table 2: Kill curve of glyphosate on wild type tobacco

Glyphosate isopropylamine conc'n(mM)	Growth of explant
0	Good stem growth, 4-5 new leaves, roots~5cm
0.005	No growth in any organ
0.011	"
0.0275	"
0.055	"
0.1	"

Primary transformants of pDV1,2 and 3 are selected by growing on medium containing 0.01 and 0.05 mM glyphosate isopropylamine salt as described above. The results are shown in Table 3.

Table 3:Selection of glyphosate tolerant lines in tissue culture

	pDV1	pDV2	pDV3
tested on herb	50	25	50
tolerant lines	25	18	20

Selection on PAT Regenerating calli are tested on 1mg/l bialophos.

Western analysis Over expression of GOX and EPSPS proteins and antibody generation are performed by means known to the skilled man. Tobacco primary transformants are analysed as follows. ~100mg PVPP is added to the bottom of an Eppendorf tube. Leaf material (four leaf discs obtained by using the tube lid as cutter) are harvested onto ice. 0.5ml extraction buffer (50mM Tris Hcl pH 7.8, 1mM EDTA sodium salt, 3mM DTT) and 2µl 100mM PMSF is added. The samples are ground in a cold room using an electric grinder. Grinding is continued for 10s, unground material pushed back into the tube and grinding continued for another 10-15s until the sample is homogeneous. Tubes are centrifuged for 15' in the cold room, supernatants removed to fresh tubes and frozen at -70C until required. Protein concentrations are determined using the known Bradford method. 25µg protein are fractionated by SDS PAGE and blotted overnight at 40mA onto a Hybond-N membrane. The filter is removed from the blotting apparatus and placed in 100ml 1X Tris-Saline 5% Marvel and shaken at RT for 45'.The filter is washed by shaking at RT in 1X Tris -Saline 0.1% Tween 20- first wash 5', second wash 20'. The primary antibody is used at 1:10000 dilution in 1X Tris-Saline 0.02% Tween 20. The membrane is incubated with the primary antibody at RT 2 hours or over night at 4C. The membrane is washed in 1X T-S 0.1% Tween at RT for 10' then for 1hour. The second antibody (anti rabbit IgG peroxidase conjugate) is used at 1:10000 dilution, incubation with the membrane was for 1 hour at RT. Washing is as described above. Detection is performed using the Amersham ECL detection kit.

A range of protein expression levels are observed in the pDV1 and 2 lines based on the Western results. Expression levels of GOX and EPSPS in the PDV3 showed little variation

COR-AAD1-00001828

in the amount of GOX being expressed but increased variation in the amount of EPSPS. Lines expressing both genes are selected for further analysis.

Maize calli are analysed in the same way, calli expressing GOX and EPSPS are regenerated into whole plants and leaf material analysed again for expression of both genes.

Phosphinothricin acetyl transferase activity assays PAT activity is measured using ^{14}C labelled acetyl CoA. The labelled acetyl group is transferred to the phosphinothricin(PPT) substrate by the PAT in the leaf extracts. Acetylated PPT and ^{14}C migrate at different rates on a TLC plate, and can be visualised by autoradiography. Leaf extraction buffer is prepared using 10X $T_{50}E_2$ buffer (TE)-50ml 1M Tris.HCl pH7.5, 4ml 0.5M EDTA and 46ml dd water. Leupeptin is made up at a rate of 15mg/ml in 1X TE. Stock PMSF is made up in methanol to 30mg/ml. BSA stock solution is made at 30mg/ml in TE, and DTT at 1M. PPT is used as 1mM solution in TE. ^{14}C Acetyl CoA was 58.1mCi/mmol (Amersham). The extraction buffer is made by combining 4315µl dd water, 500µl 10X TE, 50µl leupeprin stock, 25µl PMSF stock, 100µl BSA stock and 10µl DTT stock (final volume 5000µl). Leaf samples are harvested into Eppendorf tubes on ice using the lid as cutter. The samples (three pieces) are ground in 100µl extraction buffer using an electric grinder in cold room. The samples are centrifuged for 10minutes and 50µl removed to a fresh tube on ice. Samples are stored at -70C until use. Bradford analysis is used to quantify the protein present in the extracts. The substrate solution is prepared by mixing 5 volumes of labelled Acetyl CoA with 3 volumes of 1mM PPT solution. To a ~25µg total leaf protein sample (~2µg/µl) is added 2µl substrate solution, the mixture is incubated at 37C for30", then removed to ice to stop the reaction. A sample of 6µl is spotted onto a silica gel TLC plate (Sigma T-6770). Ascending chromatography is performed in a 3:2 mix of isopropanol and 25% ammonia solution, for 3 hour. Plates are wrapped in plastic film and exposed o/n to Kodak XAR-5 film. All 26 primary transformants are assessed for PAT activity using this method of analysis. Table 4 below gives details of the result of this analysis.

COR-AAD1-00001829

Table 4:PAT activity data

PAT activity	pPG6 line
High	1,9,11,12,14,21,22,24,25,27,28,30,32
Medium	5,7,10,15,19,20
Low	6,2,8

Herbicide leaf painting 35S-PAT primary transformants showing a range of PAT activity and control plants are tested by painting of Challenge onto individual leaves. Both surfaces of marked leaves are painted with a 1% and 0.2% solution of the stock solution (150g/l) in water. Scoring is performed after 48 hr and one week and leaf samples are taken for PAT assay.Table 5 shows the results of leaf painting.

Table 5:Leaf paint analysis

Expression	Plant line	0.2%	0.2%	1%	1%
		48hr	1 week	48hr	1 week
High	24,1,14	Undamaged	Undamaged	Undamaged	Undamaged
Low	6,10	Undamaged	Dead	Dead	Dead
Wild type		Dead	Dead	Dead	Dead

Herbicide spray test

Glufosinate (Challenge or Basta). A dose response curve is established for the effect of Challenge on wild type tobacco. Five plants are used in each treatment, the scoring is performed after 14 days. Following construction of the kill curve, selected 35SPAT lines are subjected to spray tests using Challenge, at the same rates of application. Table 6 shows this data, for two lines, #12 and 27. Transgenic plants showed no damage at these rates.

COR-AAD1-00001830

Table 6: Results of spray test on 35S primary transformants

Basta rate	*Wild type*	*35SPAT#12*	*35SPAT#27*
	%damage	%damage	%damage
200g/ha	30	0	0
400g/ha	40	0	0
600g/ha	40	0	0
900g/ha	80	0	0

A kill curve is established for the effect of glyphosate on wild type tobacco. Wild type tobacco growing in tissue culture is sub-cultured by taking stem segments and growing in fresh medium to generate 20 new plants. These are grown in tissue culture for one month before transfer to 3inch pots in John Innes No 3 compost. They are initially covered in fleece to protect them. After uncovering they are allowed to acclimatise for four days before being sprayed. After spraying watering is only into the saucers i.e. no water is allowed to touch the leaves for five days. Scoring was done 8 days and 28 days after treatment. Table 7 shows the mean percentage damage (three reps per treatment) at a range of application concentrations.

Table 7: dose response curve of wild type tobacco treated with glyphosate at the rates indicated.

Trt. No	Compound	Rate g/ha	Adjuvant	Nicotiana wild type
1	Roundup Ultra (USA)	100	'Frigate'	73
2	480g/l glyphosate-	200	(K30512)	90
3	isopropylamine	400		99
4	(a.e. 360g/l)	800		100
5	LDJW010017	1200		100
6		1600		100

COR-AAD1-00001831

Following construction of the glyphosate kill curve, a number of pDV1,2 and 3 lines are spray tested with appropriate rates of glyphosate. Table 8 shows the results for pDV3, the results for pDV1 and pDV2 lines are similar to those of pDV3.

Table 8: dose/response of pDV3 primary transformants treated with glyphosate

	1	2	3	4	5	6	7
Rate g/ha	12.33	37	100	300	1000	3000	9000
Wildtype	2	3	20	80	93	-	-
pDV3#11	-	-	0	0	25	70	80
#14	-	-	7	22	13	33	76
#19	-	-	0	0	0	35	80
#21	-	-	0	5	0	24	78
#31	-	-	12	0	4	26	78
#34	-	-	0	0	6	30	63
#36	-	-	0	0	0	0	0
#37	-	-	0	0	9	24	72
#43	-	-	0	0	0	6	73
#44	-	-	0	40	45	78	85
#45	-	-	0	0	3	9	70
#47	-	-	5	0	0	11	72
#60	-	-	0	0	0	19	63
#64	-	-	0	5	0	12	63

Segregation analysis. Seed from each primary transformant (pPG6, pDV1, pDV2 and pDV3 is sterilised in 10% Domestos for twenty minutes. After several washes in sterile water, 100 seed of each selfed primary transformant is plated onto 0.5XMS (2.3g/l MS salts, 1.5% sucrose, 0.8% Bactoagar, pH 5.9)medium containing 100mg/l kanamycin. Seedlings are scored after three weeks growth at 26°C under16hr light/8 hr dark. Lines segregating in a ratio of 3:1 are assumed to have single transgene insertions. In the case of the pPG6 lines,

COR-AAD1-00001832

#12, 20, 27 segregated in the desired ratio. In the case of the pDV3 lines, # 14,19,21,31,34,43 and 45 segregated in the desired ratio.

Generation of homozygous lines From the segregation tests 10 unbleached seedlings, (heterozygotes or homozygotes) are transferred to fresh medium in tubs and grown on for two-three weeks. After this time they are transferred to JI No 3 compost in 3l pots to flowering. Seeds are retested on Km containing 0.5XMS to identify homozygous lines.

Crossing of tobacco lines Homozygous lines containing pDV1, pDV2 and pDV3 ie plants expressing GOX, EPSPS and GOX/EPSPS genes respectively are cross pollinated onto a homozygous pPG6 line expressing the PAT gene, line # 27. The pollination is also performed using the pPG6 line as the male line.

Analysis of transgenic corn lines Regenerating calli are tested by PCR using the oligos described above ie 35S-AlcR, AlcA-GOX1, and internal oligo's for GOX and EPSPS. Western analysis is also performed on the PCR +ve calli to select those expressing GOX and EPSPS. Those calli are regenerated and the resulting plants are re-analysed by PCR. The plants are then backcrossed and selfed.

Analysis of tobacco progeny

GOX, EPSPS and PAT expression. All progeny are homozygous for both genes. Seed from each crossing and seed from each parent homozygote is sown and leaf material harvested from a number of plants for analysis. Protein extracts are analysed by western blotting and then by PAT activity measurements as described previously. Levels of expression of GOX and EPSPS and PAT activity are found to be similar to each other within a particular cross and to those of the homozygote parent. The plants are scored for appearance, height, vigour of growth etc.

Herbicide treatments. Three broad experiments are designed:-

1.35S-PAT cf pDV1 cf pDV1-PAT

2.35S-PAT cf pDV2 cf pDV2-PAT

COR-AAD1-00001833

3.35S-PAT cf pDV3 cf pDV3-PAT

35S-PAT lines are treated with glufosinate at a range of concentrations and the rates at which particular degrees of damage occurred identified, at different time points. DV lines are treated with glyphosate at a range of concentrations and the damage rate identified. DV-PAT lines are then treated with mixtures of the two herbicides at different ratios and the level of damage assessed. Each of the populations are treated at the 5-6 leaf stage (5 reps per treatment).

Resistance to pathogen attack 35S-PAT, DV1,2 and 3 and DV-PAT lines expressing good levels of each protein and showing good herbicide tolerances are exposed to a number of fungal pathogens and the level of infection scored and compared.

Analysis of maize progeny. The seed resulting from the crossing of the primary transformants is used to generate plants from which to select the best expressing lines. This is done by western analysis of expression levels of GOX, EPSPS and by PAT activity experiments as described above. Similar experiments are performed to determine herbicide tolerance to glyphosate and glufosinate, either applied singly or in various combinations.

EXAMPLE 9

Production of plants tolerant to pre-emergent bleaching herbicides eg fluorochloridone, norflurazon, fluridone, flurtamone and diflufenican and to glyphosate.

Phytoene desaturase (PDS) inhibitors eg flurochloridone and norflurazon are a group of herbicides which block carotenoid biosynthesis and give rise to bleaching symptomology. The PDS gene (*crtI*) is cloned from Erwinia *uredovora,* a non-green phytopathogenic bacterial rot, and over-expressed in transgenic tobacco (and tomato)using a plasmid containing the CaMV 35S promoter and a chloroplast transit peptide (pYPIET4) (Misawa *et al.*, 1993). Homozygous seed of line ET4-208 tobacco plant which over-expresses the *crtI* gene are obtained as are tomato plants containing the same construct.

Herbicide tolerance trials Compounds of formulas (I), (II) and (III) (see below) are tested. Transgenic and wild type tomato seed (*cv* Ailsa Craig) is sown in 3" pots of JIP 3,

COR-AAD1-00001834

three seeds per pot. Each compound is formulated in 4% JF5969 (apart from compound I which is a commercial formulation) and sprayed onto the units in the track sprayer at 200 l/ha. The test is assessed at 13, 20 and 27 DAT (days after treatment). There are clear dose responses from all treatments on the wild type tomato, with the highest rate in all cases giving 87-100% phytotoxicity. The transgenic tomatoes are highly tolerant of all of the PDS inhibitors tested, at least to 1 kg/ha of compounds II and III and up to 9 kg/ha of compound I (see Table 9). Similar results are obtained for transgenic tobacco.

Table 9: Phytotoxity at 27 PAT

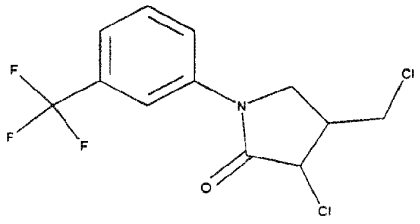
		Tomato	
Chemical	Rate (g/ha)	Wild type	Transgenic
Compound II	37	3.3	
	111	13.3	3.3
	333	20	0
	1000	100	3.3
	3000		0
Compound III	37	0	
	111	8.3	0
	333	56.7	0
	1000	100	10
	3000		3.3
Compound I	333	0	0
	1000	23.3	0
	3000	100	0
	9000	100	0

DV3 # 43B (a glyphosate resistant line comprising the EPSPS and GOX genes - see Example 4) is cross pollinated onto homozygous ET4-208 and *vice versa* in the usual way. Seed is collected and used in herbicide trials similar to those described above. The tobacco seed is sown in rows in small units the day before treatment. Each compound is formulated in 4% JF5969 (apart from Racer which was a commercial formulation) and sprayed onto the units in the track sprayer at 200 l/ha. The test is assessed at 13, 20 and 27 DAT. Seedlings that are tolerant to bleaching herbicides are transferred after the final assessment into fresh

COR-AAD1-00001835

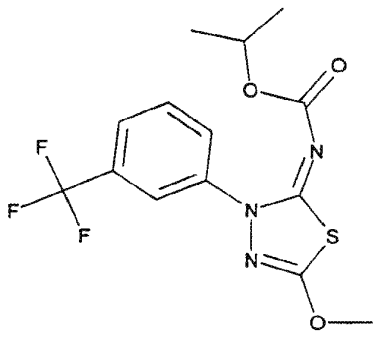
John Innes 111 compost in 3" pots. After two weeks they are subjected to glyphosate herbicide applied at 500 and 800 g/ha. Scoring is performed 14 and 28 DAT. The resultant plants are resistant to both classes of herbicide, and the resistance is inherited in a Mendelian manner.

Compound of Formula I:



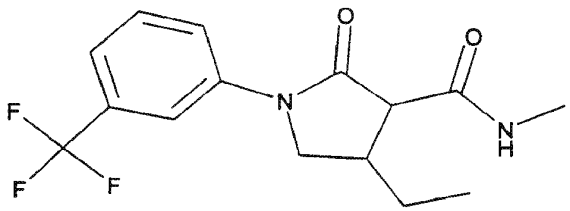
(1)

Compound of Formula II:



(2)

Compound of Formula III:



(3)

EXAMPLE 10

Generation of plants tolerant to triketones, acetanilides and glyphosate

pDV6 # 71G and pDV3 # 19J are cross pollinated onto a homozygous triketone tolerant line and vice versa as described earlier. Seed are collected and used in herbicide trials as described below. The tobacco seeds obtained from the DV6/HPPD cross are sown in rows in small units the day before treatment. Some units are treated with acetochlor (75 g/ha), some with alachlor (300 g/ha) and others with ZA1296 (100 and 300 g/ha). Assessment is at 21 DAT. The scores are given below for the 21DAT assessment and represent phytotoxicity.

		Wild Type	Wild type	DV6/HPPD	DV6/HPPD
Chemical	Rate (g/ha)	Rep a	Rep b	Rep a	Rep b
Acetochlor	75	40	100	0	0
Alachlor	300	80	90	0	0
ZA1296	100	90	95	10	0
	300	100	100	0	0

Seedlings surviving the 300 g/ha treatment of ZA1296 are sprayed with 800 g/ha of glyphosate and demonstrate tolerance to this.

COR-AAD1-00001837

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Zeneca Ltd
(B) STREET: 15 Stanhope Gate
(C) CITY: London
(E) COUNTRY: England
(F) POSTAL CODE (ZIP): W1Y 6LN
(G) TELEPHONE: 0171-304 5000

(ii) TITLE OF INVENTION: Herbicide resistant plants

(iii) NUMBER OF SEQUENCES: 32

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1020 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Synechocystis sp. PCC6803

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..1020

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATG GAA TTC GAC TAT CTT CAT TTA TAC GTT GAC GAT TAT CAG TCA GCT        48
Met Glu Phe Asp Tyr Leu His Leu Tyr Val Asp Asp Tyr Gln Ser Ala
 1               5                  10                  15

CAT CGT TGT TAT CAA CGT CAA TGG GGT TTC ACT TGC GTA AAT AAA ATT        96
His Arg Cys Tyr Gln Arg Gln Trp Gly Phe Thr Cys Val Asn Lys Ile
            20                  25                  30

ATT ACT GAC CAA GGA ATT ACT GGC ATC TAC CAA CAG GGG CAA ATA CTT       144
Ile Thr Asp Gln Gly Ile Thr Gly Ile Tyr Gln Gln Gly Gln Ile Leu
        35                  40                  45

CTG CTA ATT TCG GCA TCG GAA TCT AGT TTG AGT AGA TAT GCC GAC TAT       192
Leu Leu Ile Ser Ala Ser Glu Ser Ser Leu Ser Arg Tyr Ala Asp Tyr
    50                  55                  60
```

```
CTC CAG AAA CAT CCC CCC GGC GTA GGT GAA GTC GCT TGG CAG GTG GCC      240
Leu Gln Lys His Pro Pro Gly Val Gly Glu Val Ala Trp Gln Val Ala
 65                  70                  75                  80

AAT TGG CAA AAA ATT CAG CAT CAA TTA TCA GAA TTA CAG ATA GAA ACC      288
Asn Trp Gln Lys Ile Gln His Gln Leu Ser Glu Leu Gln Ile Glu Thr
                 85                  90                  95

ACA CCA GTT ATT CAT CCT CTG ACT AAA GCA GAA GGA TTA ACT TTT TTG      336
Thr Pro Val Ile His Pro Leu Thr Lys Ala Glu Gly Leu Thr Phe Leu
            100                 105                 110

CTC TGG GGA GAT GTG CAC CAT AGC ATT TAT CCT GTT CGT TCT GAG CTA      384
Leu Trp Gly Asp Val His His Ser Ile Tyr Pro Val Arg Ser Glu Leu
        115                 120                 125

AAT CAG AAT AAA ACA TTG CAT GGT GTT GGT TTA ACG ACC ATC GAC CAT      432
Asn Gln Asn Lys Thr Leu His Gly Val Gly Leu Thr Thr Ile Asp His
    130                 135                 140

GTG GTG CTA AAC ATT GCC GCC GAT CAA TTT ACC CAG GCT TCC CAA TGG      480
Val Val Leu Asn Ile Ala Ala Asp Gln Phe Thr Gln Ala Ser Gln Trp
145                 150                 155                 160

TAT CAA CAG GTG TTT GGC TGG TCG GTG CAG CAG AGT TTT ACT GTC AAT      528
Tyr Gln Gln Val Phe Gly Trp Ser Val Gln Gln Ser Phe Thr Val Asn
                165                 170                 175

ACG CCC CAT TCT GGT CTG TAT AGC GAA GCC CTG GCC AGT GCC AAT GGG      576
Thr Pro His Ser Gly Leu Tyr Ser Glu Ala Leu Ala Ser Ala Asn Gly
            180                 185                 190

AAA GTC CAA TTT AAC CTC AAT TGT CCC ACC AAT AAC AGT TCC CAA ATT      624
Lys Val Gln Phe Asn Leu Asn Cys Pro Thr Asn Asn Ser Ser Gln Ile
        195                 200                 205

CAA ACT TTT TTA GCC AAT AAC CAT GGG GCT GGT ATT CAA CAT GTC GCT      672
Gln Thr Phe Leu Ala Asn Asn His Gly Ala Gly Ile Gln His Val Ala
    210                 215                 220

TTT TCC ACT ACG AGT ATT ACG CGA ACT GTG GCT CAT CTG CGG GAA AGG      720
Phe Ser Thr Thr Ser Ile Thr Arg Thr Val Ala His Leu Arg Glu Arg
225                 230                 235                 240

GGC GTA AAT TTT TTA AAA ATC CCC ACT GGC TAT TAT CAA CAG CAA AGA      768
Gly Val Asn Phe Leu Lys Ile Pro Thr Gly Tyr Tyr Gln Gln Gln Arg
                245                 250                 255

AAC AGT AGC TAT TTT AAT TAT GCA AGT TTG GAT TGG GAT ACC TTA CAG      816
Asn Ser Ser Tyr Phe Asn Tyr Ala Ser Leu Asp Trp Asp Thr Leu Gln
            260                 265                 270

TGC CTA GAA ATT TTG CTG GAT GAT CAA GAT AAT ACG GGG GAG CGA TTA      864
Cys Leu Glu Ile Leu Leu Asp Asp Gln Asp Asn Thr Gly Glu Arg Leu
        275                 280                 285

CTG CTA CAA ATT TTT AGT CAG CCT TGC TAT GGA GTA GGC ACT CTA TTT      912
Leu Leu Gln Ile Phe Ser Gln Pro Cys Tyr Gly Val Gly Thr Leu Phe
    290                 295                 300

TGG GAA ATT ATT GAA CGC CGC CAC CGG GCA AAA GGA TTT GGT CAA GGA      960
Trp Glu Ile Ile Glu Arg Arg His Arg Ala Lys Gly Phe Gly Gln Gly
305                 310                 315                 320
```


COR-AAD1-00001839

```
AAC TTT CAA GCT CTC TAT GAA GCG GTG GAG ACT TTA GAA AAA CAG TTA      1008
Asn Phe Gln Ala Leu Tyr Glu Ala Val Glu Thr Leu Glu Lys Gln Leu
                325                 330                 335

GAA GTG CCA TAA                                                   1020
Glu Val Pro
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 339 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Glu Phe Asp Tyr Leu His Leu Tyr Val Asp Asp Tyr Gln Ser Ala
 1               5                  10                  15

His Arg Cys Tyr Gln Arg Gln Trp Gly Phe Thr Cys Val Asn Lys Ile
            20                  25                  30

Ile Thr Asp Gln Gly Ile Thr Gly Ile Tyr Gln Gln Gly Gln Ile Leu
        35                  40                  45

Leu Leu Ile Ser Ala Ser Glu Ser Ser Leu Ser Arg Tyr Ala Asp Tyr
    50                  55                  60

Leu Gln Lys His Pro Pro Gly Val Gly Glu Val Ala Trp Gln Val Ala
65                  70                  75                  80

Asn Trp Gln Lys Ile Gln His Gln Leu Ser Glu Leu Gln Ile Glu Thr
                85                  90                  95

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

Leu Trp Gly Asp Val His His Ser Ile Tyr Pro Val Arg Ser Glu Leu
        115                 120                 125

Asn Gln Asn Lys Thr Leu His Gly Val Gly Leu Thr Thr Ile Asp His
    130                 135                 140

Val Val Leu Asn Ile Ala Ala Asp Gln Phe Thr Gln Ala Ser Gln Trp
145                 150                 155                 160

Tyr Gln Gln Val Phe Gly Trp Ser Val Gln Gln Ser Phe Thr Val Asn
                165                 170                 175

Thr Pro His Ser Gly Leu Tyr Ser Glu Ala Leu Ala Ser Ala Asn Gly
            180                 185                 190

Lys Val Gln Phe Asn Leu Asn Cys Pro Thr Asn Asn Ser Ser Gln Ile
        195                 200                 205

Gln Thr Phe Leu Ala Asn Asn His Gly Ala Gly Ile Gln His Val Ala
    210                 215                 220

Phe Ser Thr Thr Ser Ile Thr Arg Thr Val Ala His Leu Arg Glu Arg
225                 230                 235                 240
```

Gly Val Asn Phe Leu Lys Ile Pro Thr Gly Tyr Tyr Gln Gln Gln Arg
245 250 255

Asn Ser Ser Tyr Phe Asn Tyr Ala Ser Leu Asp Trp Asp Thr Leu Gln
260 265 270

Cys Leu Glu Ile Leu Leu Asp Asp Gln Asp Asn Thr Gly Glu Arg Leu
275 280 285

Leu Leu Gln Ile Phe Ser Gln Pro Cys Tyr Gly Val Gly Thr Leu Phe
290 295 300

Trp Glu Ile Ile Glu Arg Arg His Arg Ala Lys Gly Phe Gly Gln Gly
305 310 315 320

Asn Phe Gln Ala Leu Tyr Glu Ala Val Glu Thr Leu Glu Lys Gln Leu
325 330 335

Glu Val Pro

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2582 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Pseudomonas fluorescens*

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1217..2290

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATTAGTCGAA GAATATGCCC ATCCTGTCGC CTGTCGAGCA ACTGCTAATG CAACCTCCGT 60

CTGATCGCCT CACCTACCTG AAGCTGGCCG CTGTGACCAT GATTTGGGGT GGCACTTTTG 120

TCGCCGGACG TTACCTGACC AATCAAGTCG ACCCGCTGCT GGCCGCCAGC CTGCGGTTTA 180

TCCTGGCCAG CCTGGCGCTG CTGCTGTTTA TGCTGTGTGC ACGCATCCCG CTGGCGCGGC 240

CACGTCCCCG GCAACTGCTG CATCTGGCGG TGCTGGGGTT TTTCGGGATC TTTTTCTACA 300

ACCTGTGTTT TTTCTACGGC CTGCAGTACA TCAACGCCTC GCGCGCTTCG TTGATCGTGG 360

CGTTGAATCC GGCGGTGATC GGCCTGGCTT CCTGGTGGTT GTTCAAAGAG CGCCTCGGCA 420

CTGCCAGGGT GCTGGGTATC GCGTTGTGCC TGGCCGGCGC TGCGACGGTG ATCGTCAGTC 480

GCAACCCGCA GTTGCTGCAA GGTGCATCGA GTACCTGGCA GGGCGACCTG CTGGTGTTCG 540

GCTGTGTGCT GGGGTGGGGG ATTTACTCGT TGTTTTCCCG CGCATTGAAT CAAAGCCTGG 600

GGCCGTTGCA AACGGTCACC TGGTCAGTGC TGCTGGGCAC CCTGATGCTG ACGGCTGTCA 660

```
CCGCGCTCGC CGGGCGCTTC ACGCTTGCAG GGCTTGGCAG CCTGCACCTG CCGCAGGTTG      720

TGAGCCTGTT GTATTTGGGC GTGCTCGGCT CCGCGCTGGC GTACATCGGC TATTACGATG      780

GCATCCGGCG TATCGGCGCG ACCCGCGCAG GCGTGTTTAT CGCGCTGAAC CCGCTGACGG      840

CGGTGATCTG CGGCGCGCTG CTGCTTGGCG AACAGCTAAC GTTACCCATG GCGCTCGGCG      900

GCGCGGTGAT CCTGTTGGGC ATCTATCTGT GCAACAAACC CCTTGCGCAG CCCAGCGCAA      960

TAGGGATTTG ATGAGAGTGC GGACAAATAC TGTTACGCTG TGTAGAATCG ATTTACGCAT     1020

ACAAGAATAT GGACTTGCGC TCACGCAAGC CTCGGCCGTC AGAGACTGAT GTAATCATGA     1080

AGCTACTCGG CTCCCCCCTG ATCTTTGGTG ACTTCCTCGC GCGCAGCGTG CGGGGTCTCT     1140

CGTGCGCGCC ACCCTGCAAC CTCATCCTTG CCTGTAATTG ACTGCTTGCT ACTTACAAGA     1200

ATGATGAGGT GCCGAA ATG GCC GAC CAA TAC GAA AAC CCA ATG GGC CTG          1249
                  Met Ala Asp Gln Tyr Glu Asn Pro Met Gly Leu
                    1               5                  10

ATG GGC TTT GAA TTT ATT GAA TTC GCA TCG CCG ACT CCG GGC ACC CTG       1297
Met Gly Phe Glu Phe Ile Glu Phe Ala Ser Pro Thr Pro Gly Thr Leu
            15                  20                  25

GAG CCG ATC TTC GAG ATC ATG GGC TTC ACC AAA GTC GCG ACC CAC CGC       1345
Glu Pro Ile Phe Glu Ile Met Gly Phe Thr Lys Val Ala Thr His Arg
        30                  35                  40

TCC AAG AAT GTG CAC CTG TAC CGC CAG GGC GAG ATC AAC CTG ATC CTC       1393
Ser Lys Asn Val His Leu Tyr Arg Gln Gly Glu Ile Asn Leu Ile Leu
    45                  50                  55

AAC AAC CAG CCC GAC AGC CTG GCC TCG TAC TTC GCC GCC GAA CAC GGC       1441
Asn Asn Gln Pro Asp Ser Leu Ala Ser Tyr Phe Ala Ala Glu His Gly
60                  65                  70                  75

CCT TCG GTG TGC GGC ATG GCG TTC CGG GTC AAA GAC TCG CAG CAG GCT       1489
Pro Ser Val Cys Gly Met Ala Phe Arg Val Lys Asp Ser Gln Gln Ala
                80                  85                  90

TAC AAC CGC GCG TTG GAA CTG GGC GCC CAG CCG ATT CAT ATC GAA ACC       1537
Tyr Asn Arg Ala Leu Glu Leu Gly Ala Gln Pro Ile His Ile Glu Thr
            95                  100                 105

GGC CCG ATG GAA CTC AAC CTG CCG GCC ATC AAG GGC ATC GGC GGT GCG       1585
Gly Pro Met Glu Leu Asn Leu Pro Ala Ile Lys Gly Ile Gly Gly Ala
        110                 115                 120

CCG CTG TAC CTG ATC GAC CGC TTC GGT GAA GGC AGC TCG ATA TAT GAC       1633
Pro Leu Tyr Leu Ile Asp Arg Phe Gly Glu Gly Ser Ser Ile Tyr Asp
    125                 130                 135

ATC GAC TTC GTG TAC CTC GAA GGT GTC GAC CGC AAC CCG GTA GGC GCG       1681
Ile Asp Phe Val Tyr Leu Glu Gly Val Asp Arg Asn Pro Val Gly Ala
140                 145                 150                 155

GGC CTC AAG GTC ATC GAC CAC CTG ACC CAC AAC GTG TAT CGC GGC CGC       1729
Gly Leu Lys Val Ile Asp His Leu Thr His Asn Val Tyr Arg Gly Arg
                160                 165                 170

ATG GCC TAC TGG GCC AAC TTC TAC GAG AAA CTG TTC AAC TTC CGT GAA       1777
```

```
Met Ala Tyr Trp Ala Asn Phe Tyr Glu Lys Leu Phe Asn Phe Arg Glu
            175                 180                 185

GCA CGC TAC TTC GAT ATC AAG GGC GAA TAC ACC GGC CTT ACG TCC AAG      1825
Ala Arg Tyr Phe Asp Ile Lys Gly Glu Tyr Thr Gly Leu Thr Ser Lys
        190                 195                 200

GCC ATG AGT GCC CCG GAC GGC ATG ATC CGC ATC CCG CTG AAC GAG GAA      1873
Ala Met Ser Ala Pro Asp Gly Met Ile Arg Ile Pro Leu Asn Glu Glu
    205                 210                 215

TCG TCC AAG GGC GCC GGC CAG ATC GAA GAG TTC CTG ATG CAG TTC AAC      1921
Ser Ser Lys Gly Ala Gly Gln Ile Glu Glu Phe Leu Met Gln Phe Asn
220                 225                 230                 235

GGC GAG GGC ATC CAG CAC GTG GCG TTC CTC ACC GAA GAC CTG GTC AAG      1969
Gly Glu Gly Ile Gln His Val Ala Phe Leu Thr Glu Asp Leu Val Lys
                240                 245                 250

ACC TGG GAT GCG TTG AAG AAG ATC GGC ATG CGC TTC ATG ACC GCG CCG      2017
Thr Trp Asp Ala Leu Lys Lys Ile Gly Met Arg Phe Met Thr Ala Pro
            255                 260                 265

CCG GAC ACC TAC TAC GAA ATG CTC GAA GGC CGC CTG CCA AAC CAC GGC      2065
Pro Asp Thr Tyr Tyr Glu Met Leu Glu Gly Arg Leu Pro Asn His Gly
        270                 275                 280

GAG CCG GTG GAC CAA CTG CAG GCG CGC GGT ATT TTG CTG GAC GGC TCC      2113
Glu Pro Val Asp Gln Leu Gln Ala Arg Gly Ile Leu Leu Asp Gly Ser
    285                 290                 295

TCG ATC GAG GGC GAC AAG CGC CTG CTG CTG CAG ATC TTC TCG GAA ACC      2161
Ser Ile Glu Gly Asp Lys Arg Leu Leu Leu Gln Ile Phe Ser Glu Thr
300                 305                 310                 315

CTG ATG GGC CCG GTG TTC TTC GAA TTC ATC CAG CGC AAA GGC GAC GAT      2209
Leu Met Gly Pro Val Phe Phe Glu Phe Ile Gln Arg Lys Gly Asp Asp
                320                 325                 330

GGG TTT GGC GAG GGC AAC TTC AAG GCG CTG TTC GAG TCG ATC GAG CGC      2257
Gly Phe Gly Glu Gly Asn Phe Lys Ala Leu Phe Glu Ser Ile Glu Arg
            335                 340                 345

GAC CAG GTA CGT CGC GGT GTA CTG ACC ACC GAC TAAGCGTCAG CAACAAAAAA    2310
Asp Gln Val Arg Arg Gly Val Leu Thr Thr Asp
        350                 355

AGCCCGGCGA GAAGGTTTTC AGCCGGGCTT TTTAGTGCCT GCACGTTTTA AGCTTTGCGC    2370

TGACGCACCA AATGTTTGAA GCCTTCATAC ACCAGCACCA TCACGGCCAG CCAGATCGGG    2430

ATATACGTCA GCCATTGCCC GCCCTTGATG CCTTCGCCCA ACAACAAGGC CACAAAGAGC    2490

AGTAATACCG GCTCCACATA GCTGAGCAAC CCGAACAGGC TAAAGGCCAA TAAACGGCTG    2550

GCGATGATGT AGCTCACCAG CGCTGAAGCA CT                                  2582
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 358 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Ala Asp Gln Tyr Glu Asn Pro Met Gly Leu Met Gly Phe Glu Phe
1 5 10 15

Ile Glu Phe Ala Ser Pro Thr Pro Gly Thr Leu Glu Pro Ile Phe Glu
20 25 30

Ile Met Gly Phe Thr Lys Val Ala Thr His Arg Ser Lys Asn Val His
35 40 45

Leu Tyr Arg Gln Gly Glu Ile Asn Leu Ile Leu Asn Asn Gln Pro Asp
50 55 60

Ser Leu Ala Ser Tyr Phe Ala Ala Glu His Gly Pro Ser Val Cys Gly
65 70 75 80

Met Ala Phe Arg Val Lys Asp Ser Gln Gln Ala Tyr Asn Arg Ala Leu
85 90 95

Glu Leu Gly Ala Gln Pro Ile His Ile Glu Thr Gly Pro Met Glu Leu
100 105 110

Asn Leu Pro Ala Ile Lys Gly Ile Gly Gly Ala Pro Leu Tyr Leu Ile
115 120 125

Asp Arg Phe Gly Glu Gly Ser Ser Ile Tyr Asp Ile Asp Phe Val Tyr
130 135 140

Leu Glu Gly Val Asp Arg Asn Pro Val Gly Ala Gly Leu Lys Val Ile
145 150 155 160

Asp His Leu Thr His Asn Val Tyr Arg Gly Arg Met Ala Tyr Trp Ala
165 170 175

Asn Phe Tyr Glu Lys Leu Phe Asn Phe Arg Glu Ala Arg Tyr Phe Asp
180 185 190

Ile Lys Gly Glu Tyr Thr Gly Leu Thr Ser Lys Ala Met Ser Ala Pro
195 200 205

Asp Gly Met Ile Arg Ile Pro Leu Asn Glu Glu Ser Ser Lys Gly Ala
210 215 220

Gly Gln Ile Glu Glu Phe Leu Met Gln Phe Asn Gly Glu Gly Ile Gln
225 230 235 240

His Val Ala Phe Leu Thr Glu Asp Leu Val Lys Thr Trp Asp Ala Leu
245 250 255

Lys Lys Ile Gly Met Arg Phe Met Thr Ala Pro Pro Asp Thr Tyr Tyr
260 265 270

Glu Met Leu Glu Gly Arg Leu Pro Asn His Gly Glu Pro Val Asp Gln
275 280 285

Leu Gln Ala Arg Gly Ile Leu Leu Asp Gly Ser Ser Ile Glu Gly Asp
290 295 300

Lys Arg Leu Leu Leu Gln Ile Phe Ser Glu Thr Leu Met Gly Pro Val
305 310 315 320

Phe Phe Glu Phe Ile Gln Arg Lys Gly Asp Asp Gly Phe Gly Glu Gly

325 330 335

Asn Phe Lys Ala Leu Phe Glu Ser Ile Glu Arg Asp Gln Val Arg Arg
340 345 350

Gly Val Leu Thr Thr Asp
355

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TATGAGAATC CTATGGG 17

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCTTTGAA GTTTCCCTC 17

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTTAGGTACC AGTCTAGACT GACCATGGCC GACCAATACG AAAACC 46

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TAGCGGTACC TGATCACCCG GGTTATTAGT CGGTGGTCAG TAC 43

(2) INFORMATION FOR SEQ ID NO: 9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 516 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Ala Gln Ile Asn Asn Met Ala Gln Gly Ile Gln Thr Leu Asn Pro
1 5 10 15

Asn Ser Asn Phe His Lys Pro Gln Val Pro Lys Ser Ser Ser Phe Leu
20 25 30

Val Phe Gly Ser Lys Lys Leu Lys Asn Ser Ala Asn Ser Met Leu Val
35 40 45

Leu Lys Lys Asp Ser Ile Phe Met Gln Lys Phe Cys Ser Phe Arg Ile
50 55 60

Ser Ala Ser Val Ala Thr Ala Gln Lys Pro Ser Glu Ile Val Leu Gln
65 70 75 80

Pro Ile Lys Glu Ile Ser Gly Thr Val Lys Leu Pro Gly Ser Lys Ser
85 90 95

Leu Ser Asn Arg Ile Leu Leu Leu Ala Ala Leu Ser Glu Gly Thr Thr
100 105 110

Val Val Asp Asn Leu Leu Ser Ser Asp Asp Ile His Tyr Met Leu Gly
115 120 125

Ala Leu Lys Thr Leu Gly Leu His Val Glu Glu Asp Ser Ala Asn Gln
130 135 140

Arg Ala Val Val Glu Gly Cys Gly Gly Leu Phe Pro Val Gly Lys Glu
145 150 155 160

Ser Lys Glu Glu Ile Gln Leu Phe Leu Gly Asn Ala Gly Thr Ala Met
165 170 175

Arg Pro Leu Thr Ala Ala Val Thr Val Ala Gly Gly Asn Ser Arg Tyr
180 185 190

Val Leu Asp Gly Val Pro Arg Met Arg Glu Arg Pro Ile Ser Asp Leu
195 200 205

Val Asp Gly Leu Lys Gln Leu Gly Ala Glu Val Asp Cys Phe Leu Gly
210 215 220

Thr Lys Cys Pro Pro Val Arg Ile Val Ser Lys Gly Gly Leu Pro Gly
225 230 235 240

Gly Lys Val Lys Leu Ser Gly Ser Ile Ser Ser Gln Tyr Leu Thr Ala
245 250 255

Leu Leu Met Ala Ala Pro Leu Ala Leu Gly Asp Val Glu Ile Glu Ile
260 265 270

Ile Asp Lys Leu Ile Ser Val Pro Tyr Val Glu Met Thr Leu Lys Leu
275 280 285

Met Glu Arg Phe Gly Ile Ser Val Glu His Ser Ser Ser Trp Asp Arg
290 295 300

Phe Phe Val Arg Gly Gly Gln Lys Tyr Lys Ser Pro Gly Lys Ala Phe
305 310 315 320

Val Glu Gly Asp Ala Ser Ser Ala Ser Tyr Phe Leu Ala Gly Ala Ala
325 330 335

Val Thr Gly Gly Thr Ile Thr Val Glu Gly Cys Gly Thr Asn Ser Leu
340 345 350

Gln Gly Asp Val Lys Phe Ala Glu Val Leu Glu Lys Met Gly Ala Glu
355 360 365

Val Thr Trp Thr Glu Asn Ser Val Thr Val Lys Gly Pro Pro Arg Ser
370 375 380

Ser Ser Gly Arg Lys His Leu Arg Ala Ile Asp Val Asn Met Asn Lys
385 390 395 400

Met Pro Asp Val Ala Met Thr Leu Ala Val Val Ala Leu Tyr Ala Asp
405 410 415

Gly Pro Thr Ala Ile Arg Asp Val Ala Ser Trp Arg Val Lys Glu Thr
420 425 430

Glu Arg Met Ile Ala Ile Cys Thr Glu Leu Arg Lys Leu Gly Ala Thr

435 440 445

Val Glu Glu Gly Pro Asp Tyr Cys Ile Ile Thr Pro Pro Glu Lys Leu
450 455 460

Asn Val Thr Asp Ile Asp Thr Tyr Asp Asp His Arg Met Ala Met Ala
465 470 475 480

Phe Ser Leu Ala Ala Cys Ala Asp Val Pro Val Thr Ile Asn Asp Pro
485 490 495

Gly Cys Thr Arg Lys Thr Phe Pro Asn Tyr Phe Asp Val Leu Gln Gln
500 505 510

Tyr Ser Lys His
515

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AACAAGGTGG CGCAGTT 17

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CATCGCAAGA CCGGCAACAG 20

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GATCGCTACT AGCTTCCCA 19

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

AATCAAGGTA ACCTTGAATC CA 22

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

AATTACGGAA GCTTCCGT 18

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AGCTTGTACA CCGGTGTACA 20

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CGGACAATTA ATTGTCCGGT AC 22

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CATTTGCGGC CGCAAATGGT AC 22

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

AATTCATTTG CGGCCGCAAA TG 22

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AGCTTGCAGC GGCCGCTGCA 20

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AATTCATTTG CGGCCGCAAA TG 22

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CTCGAGTATT TTTACAACAA TTACCAAC 28

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

AATCAAGGTA ACCTTGAATC CA 22

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ACCACCAACG GTGTTCTTGC TGTTGA 26

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GCATTACATG TTAATTATTA CATGCTT 27

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GTGATACGAG TTTCACCGCT AGCGAGAC 28

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TACCTTGCGT GGACCAAAGA CTCC 24

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

ATGGCTTCCG CTCAAGTGAA GTCC 24

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CGAGACCCAT AACGAGGAAG CTCA 24

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ATTGCGTGAT TTCGATCCTA ACTT 24

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GAGAGATGTC GATAGAGGTC TTCT 24

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31

GGTGGAGCAC GACACACTTG TCTA 24

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GTCTCAATGT AATGGTTA 18

CLAIMS

1. A polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); and (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT).

2. A polynucleotide according to claim 1, wherein each of the regions is under expression control of a plant operable promoter and terminator.

3. A polynucleotide according to either of the preceding claims, wherein the first herbicide is a post emergence herbicide and the second herbicide is a pre-emergence herbicide.

4. A polynucleotide according to any preceding claim, wherein the proteins are selected from the group consisting of glyphosate oxido-reductase, 5-enol-pyruvyl-3-phosphoshikimate synthetase, phosphinothricin acetyl transferase, hydroxyphenyl pyruvate dioxygenase, glutathione S transferase, cytochrome P450, Acetyl-COA carboxylase, Acetolactate synthase, protoporphyrinogen oxidase, dihydropteroate synthase, polyamine transport proteins, superoxide dismutase, bromoxynil nitrilase, phytoene desaturase, the product of the *tfd*A gene obtainable from *Alcaligenes eutrophus,* and known mutagenised or otherwise modified variants of the said proteins.

5. A polynucleotide according to any one of claims 1 to 4, further comprising a region encoding a protein capable of providing the plant with resistance or tolerance to insects, desiccation and/or fungal, bacterial or viral infections.

6. A polynucleotide according to any preceding claim, comprising sequences 5' of and contiguous with the said regions, which sequences encode (i) a peptide which is capable of targeting the translation products of the regions to plastids such as chloroplasts, mitochondria, other organelles or plant cell walls; and/or (ii) non-translated translational enhancing sequences.

7. A polynucleotide according to any preceding claim, which is modified in that mRNA instability motifs and/or fortuitous splice regions are removed, or plant preferred codons are used so that expression of the thus modified polynucleotide in a plant yields substantially similar protein having a substantially similar activity/function to that obtained by expression of the unmodified polynucleotide in the organism in which the protein encoding regions of the unmodified polynucleotide are endogenous, with the *proviso* that if the thus modified polynucleotide comprises plant preferred codons, the degree of identity between the protein encoding regions within the modified polynucleotide and like protein encoding regions endogenously contained within the said plant and encoding substantially the same protein is less than about 70%.

8. A polynucleotide according to any one of claims 3 to 7, wherein the pre-emergence herbicide is selected from the group consisting of a dinitroaniline herbicide, diphenyl ether, sulfonyl urea, phosphosulfonates, oxyacetamides, tetrazolinones and N-carbamoyltetrazolinones, imidazolinone, thiocarbamate, triazine, triazolo-pyrimidines, uracil, a phenylurea, triketone, isoxazole, acetanilide, oxadiazole, triazinone, sulfonanilide, amide, anilide, RP201772, flurochloridone, norflurazon, and triazolinone type herbicide and the post-emergence herbicide is selected from the group consisting of glyphosate and salts thereof, glufosinate, asulam, bentazon, bialaphos, bromacil, sethoxydim or another cyclohexanedione, dicamba, fosamine, flupoxam, phenoxy propionate, quizalofop or another aryloxy-phenoxypropanoate, picloram, fluormetron, atrazine or another triazine, metribuzin, chlorimuron, chlorsulfuron, flumetsulam, halosulfuron, sulfometron, imazaquin, imazethapyr,

isoxaben, imazamox, metosulam, pyrithrobac, rimsulfuron, bensulfuron, nicosulfuron, fomesafen, fluroglycofen, KIH9201, ET751, carfentrazone, ZA1296, sulcotrione, paraquat, diquat, bromoxynil and fenoxaprop.

9. A polynucleotide according to the preceding claim, wherein the pre-emergence herbicide is selected from the group consisting of acetanilides, triketones, PDS inhibitors, thiocarbamates, tetrazolinones, and the post-emergence herbicide is selected from the group consisting of glyphosate, glufosinate, paraquat and bialphos.

10. A vector comprising the polynucleotide of any one of claims 1 to 9.

11. Plants which comprises a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a polynucleotide comprising a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT); and (iv), that when the plant is sugar beet, the herbicide resistance or tolerance conferring genes which it comprises are not solely EPSPS and PAT.

12. Plants according to the preceding claim, wherein the first herbicide is a pre-emergence herbicide and the second herbicide is a post emergence herbicide.

13. Plants including parts, seeds and progeny thereof which are resistant to at least two herbicides and which have been obtained from material which has been transformed

with the polynucleotide according to any one of claims 1 to 9, or the vector according to claim 10.

14. Plants according to the preceding claim, selected from the group consisting of small grain cereals, oil seed crops, fibre plants, fruit, vegetables, plantation crops and trees.

15. Plants according to any one of claims 11 to 14, selected from the group consisting of soybean, cotton, tobacco, sugarbeet, oilseed rape, canola, flax, sunflower, potato, tomato, alfalfa, lettuce, maize, wheat, sorghum, rye, bananas, barley, oat, turf grass, forage grass, sugar cane, pea, field bean, rice, pine, poplar, apple, grape, citrus or nut plants and the progeny, seeds and parts of such plants.

16. A method of selectively controlling weeds in a field comprising weeds and crop plants, wherein the crop plants comprise (i) a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT); and (iv), that when the crop plant is sugar beet, the herbicide resistance or tolerance conferring genes which it comprises are not solely EPSPS and PAT; or (ii) a polynucleotide comprising at least a first region encoding a first protein capable of conferring on a plant, or tissue comprising it, resistance or tolerance to a first herbicide, and a polynucleotide comprising a second region encoding a second protein likewise capable of conferring resistance to a second herbicide, with the *provisos* (i) that the polynucleotide does not encode a fusion protein comprising only a 5-enol-pyruvyl-3-phosphoshikimate synthetase (EPSPS) and a glutathione S transferase (GST); (ii) that the polynucleotide does not

COR-AAD1-00001860

comprise only regions encoding superoxide dismutase (SOD) and glutathione S transferase (GST); (iii) that the polynucleotide does not comprise only regions encoding GST and phosphinothricin acetyl transferase (PAT); and (iv), that when the crop plant is sugar beet, the herbicide resistance or tolerance conferring genes which it comprises are not solely EPSPS and PAT, the method comprising application to the field of at least one of the said herbicides in an amount sufficient to control the weeds without substantially affecting the crop plants.

17. A method according to the preceding claim, wherein the crop plants comprise a gene encoding an EPSPS enzyme and a gene encoding a GST enzyme, the method comprising application to the field of glyphosate and an acetanilide in an amount sufficient to control the weeds without substantially affecting the crop plants.

18. A method according to claim 16, wherein the crop plants comprise a gene encoding an HPPD enzyme and a gene encoding a PAT enzyme, the method comprising application to the field of a triketone and glufosinate in an amount sufficient to control the weeds without substantially affecting the crop plants.

19. A method according to claim 16, wherein the crop plants comprise a gene encoding an PAT enzyme and a gene encoding a GST enzyme, the method comprising application to the field of glufosinate and an acetanilide, thiocarbamate, and/or tetrazolinone in an amount sufficient to control the weeds without substantially affecting the crop plants.

20. A method according to claim 16, wherein the crop plants comprise a gene encoding an EPSPS and/or GOX enzyme and a gene encoding an HPPD enzyme, the method comprising application to the field of glyphosate and a triketone in an amount sufficient to control the weeds without substantially affecting the crop plants.

21. A method according to claim 16, wherein the crop plants comprise a gene encoding a PDS enzyme and a gene encoding an EPSPS and/or GOX enzyme, the method comprising application to the field of a PDS inhibitor and glyphosate in an amount sufficient to control the weeds without substantially affecting the crop plants.

22. A method according to claim 16, wherein the crop plants comprise a gene encoding an EPSPS and/or GOX enzyme and a gene encoding a PAT enzyme, the method comprising application to the field of glyphosate and glufosinate in an amount sufficient to control the weeds without substantially affecting the crop plants, with the proviso that the plants are not sugar beet.

23. A method according to claim 16, wherein the crop plants comprise a gene encoding a PDS enzyme and a gene encoding a PAT enzyme, the method comprising application to the field of a PDS inhibitor and glufosinate in an amount sufficient to control the weeds without substantially affecting the crop plants.

24. A method according to claim 16, wherein the crop plants comprise a gene encoding a PDS enzyme and a gene encoding a GST enzyme, the method comprising application to the field of a PDS inhibitor and an acetanilide herbicide in an amount sufficient to control the weeds without substantially affecting the crop plants.

25. A method according to any one of claims 17 to 24, wherein the crop plants further contain a gene encoding ALS, SOD or BNX, the method comprising application to the field of a sulphonyl urea, paraquat or bromoxynil herbicide in an amount sufficient to control the weeds without substantially affecting the crop plants.

26. A method according to any one of claims 16 to 25, further comprising application to the field of a pesticidally effective amount of one or more of an insecticide, fungicide, bacteriocide, nematicide and anti-viral.

COR-AAD1-00001862

27. A method of producing plants which are substantially tolerant or substantially resistant to two or more herbicides, comprising the steps of:

(i) transforming plant material with the polynucleotide of any one of claims 1 to 9 or the vector of claims 10;

(ii) selecting the thus transformed material; and

(iii) regenerating the thus selected material into morphologically normal fertile whole plants.

28. Use of the polynucleotide of any one of claims 1 to 9, or the vector of claim 10, in the production of plant tissues and/or morphologically normal fertile whole plants (i) which are substantially tolerant or substantially resistant to two or more herbicides.

29. Use of the polynucleotide of any one of claims 1 to 9, or the vector of claim 10, in the production of a herbicidal target for the high throughput *in vitro* screening of potential herbicides.

30. Use according to the preceding claim, wherein the protein encoding regions of the polynucleotide are heterologously expressed in *E. coli* or yeast.

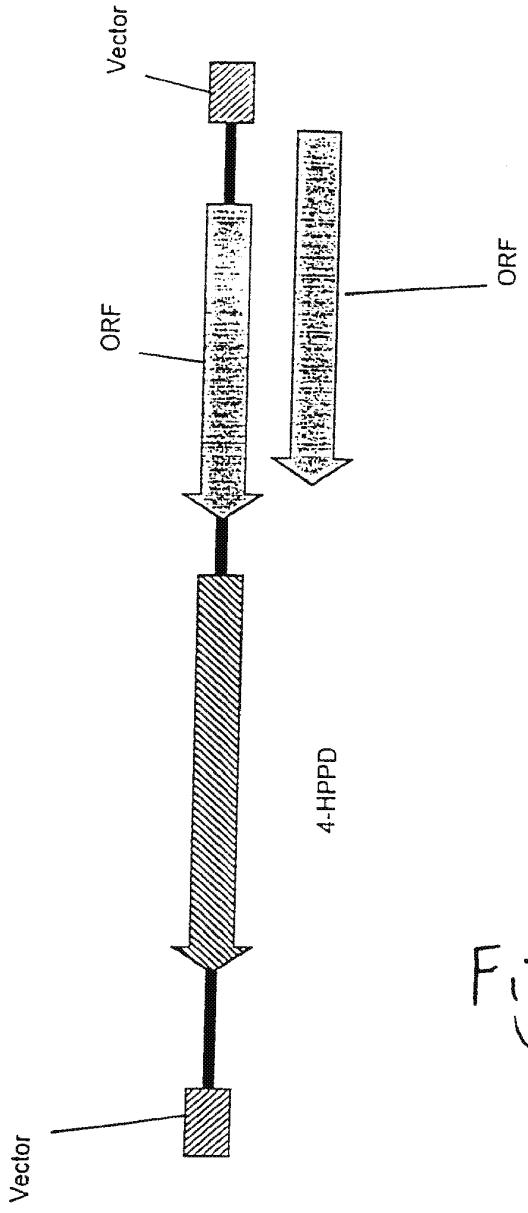
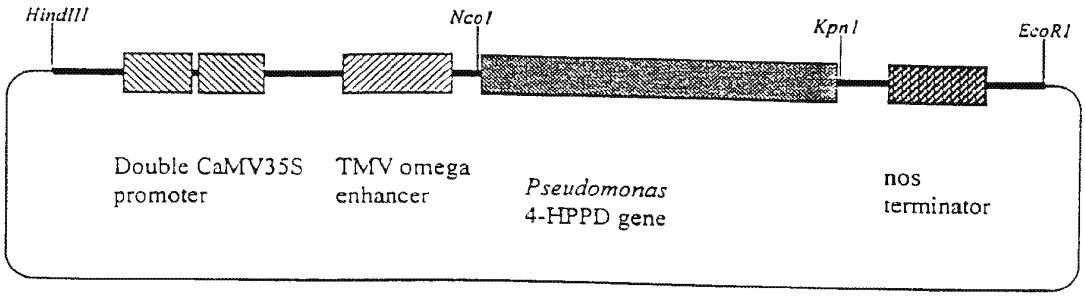
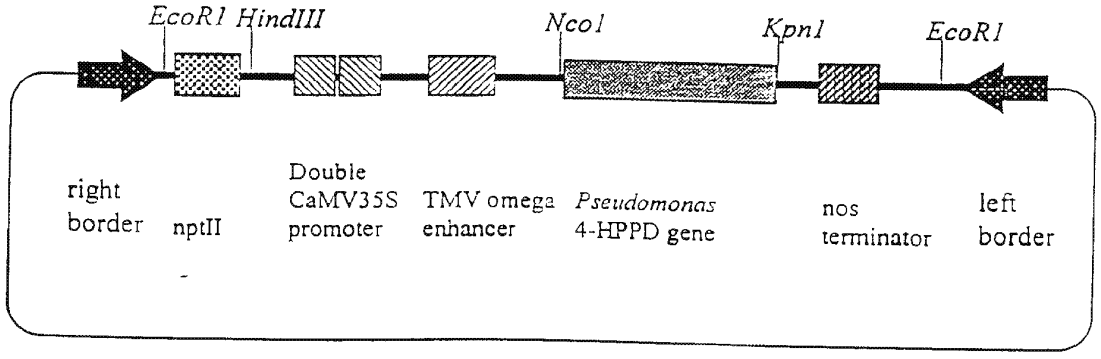


Fig. 1



pUC19

Fig. 2



pBin19

Fig. 3

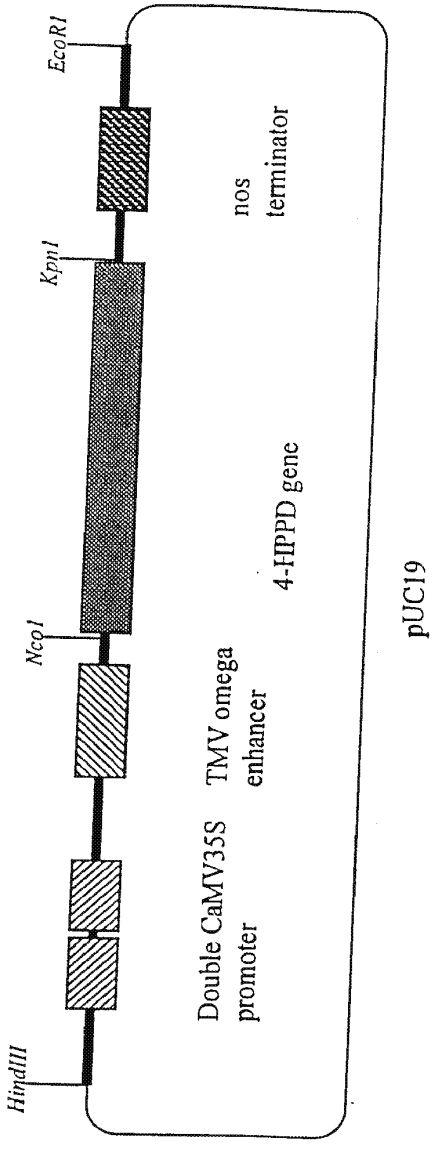

EcoR1
Kpn1
Neo1
HindIII
Double CaMV35S promoter
TMV omega enhancer
4-HPPD gene
nos terminator
pUC19

Fig. 4

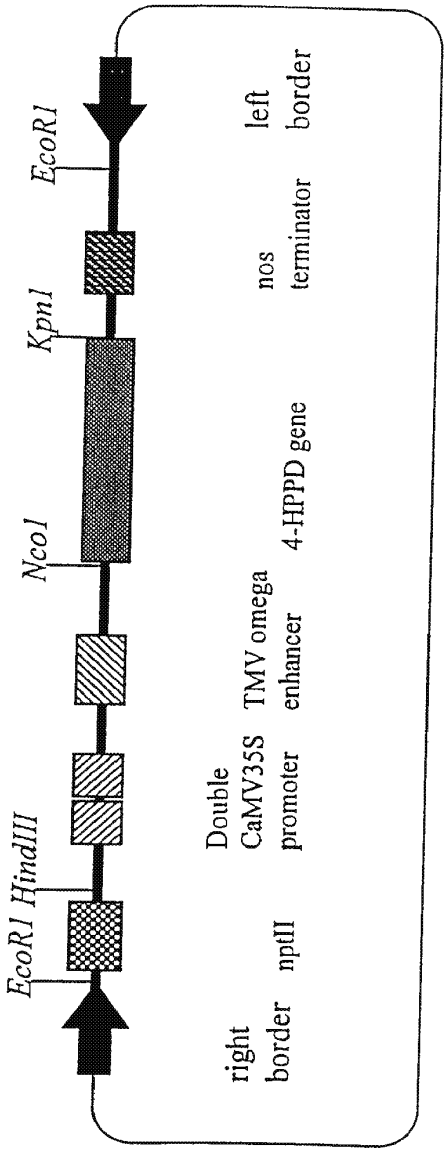

EcoR1
Kpn1
Nco1
EcoR1 HindIII
right border
nptII
Double CaMV35S promoter
TMV omega enhancer
4-HPPD gene
nos terminator
left border
pBin19

Fig. 5

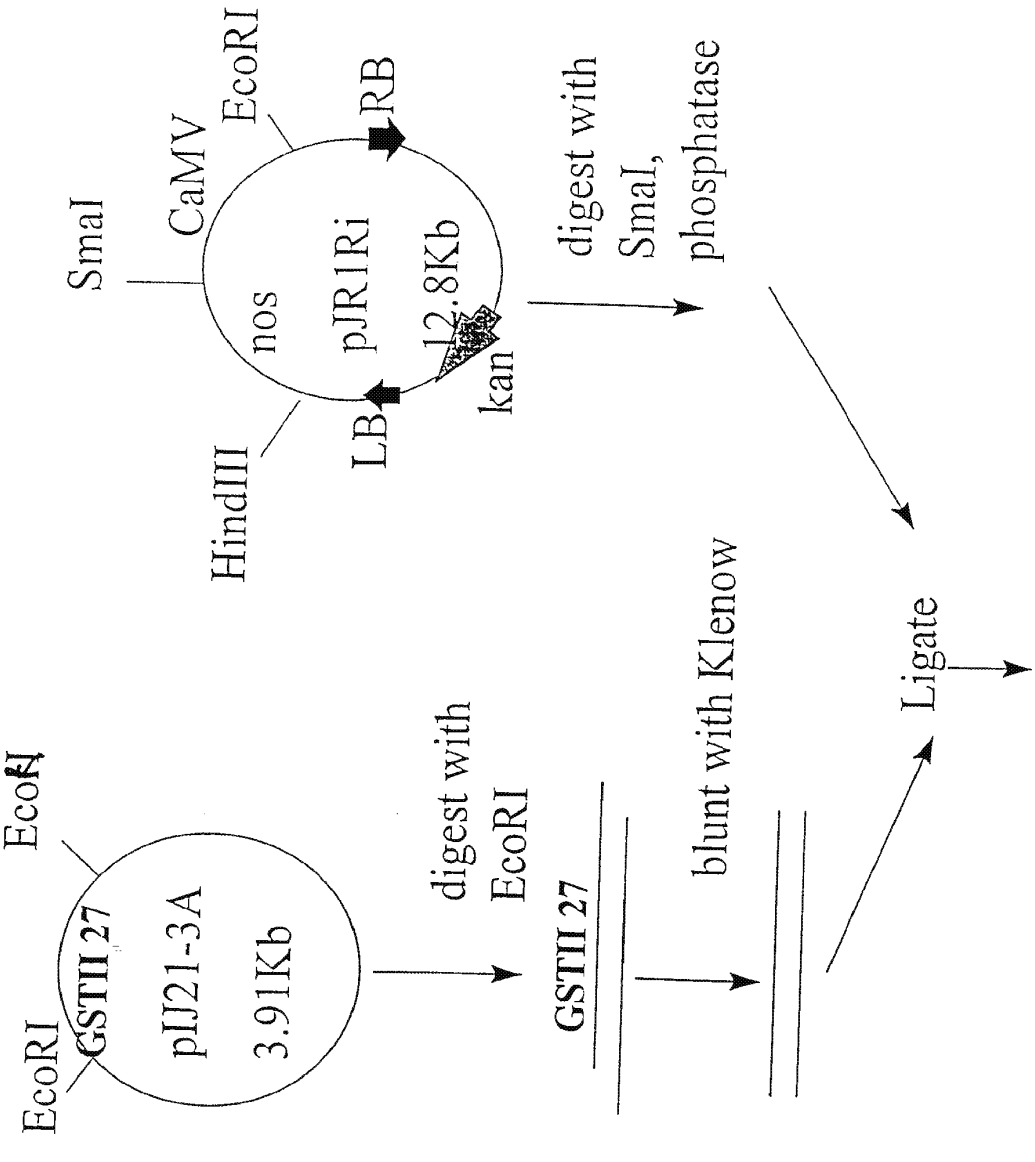

SmaI
CaMV
EcoRI
RB
nos
pJR1Ri
12.8Kb
kan
LB
HindIII
digest with
SmaI,
phosphatase
EcoRI
GSTII 27
pIJ21-3A
3.91Kb
digest with
EcoRI
GSTII 27
blunt with Klenow
Ligate

Fig. 6

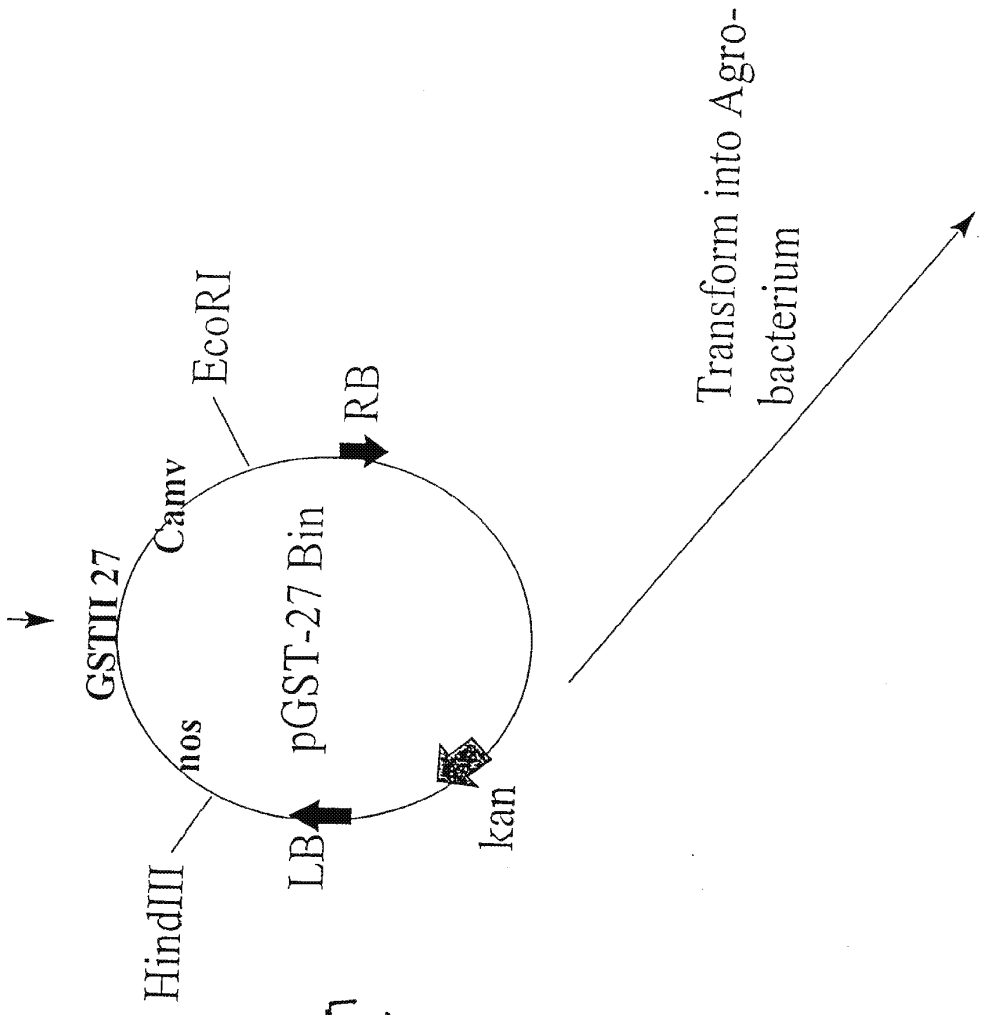
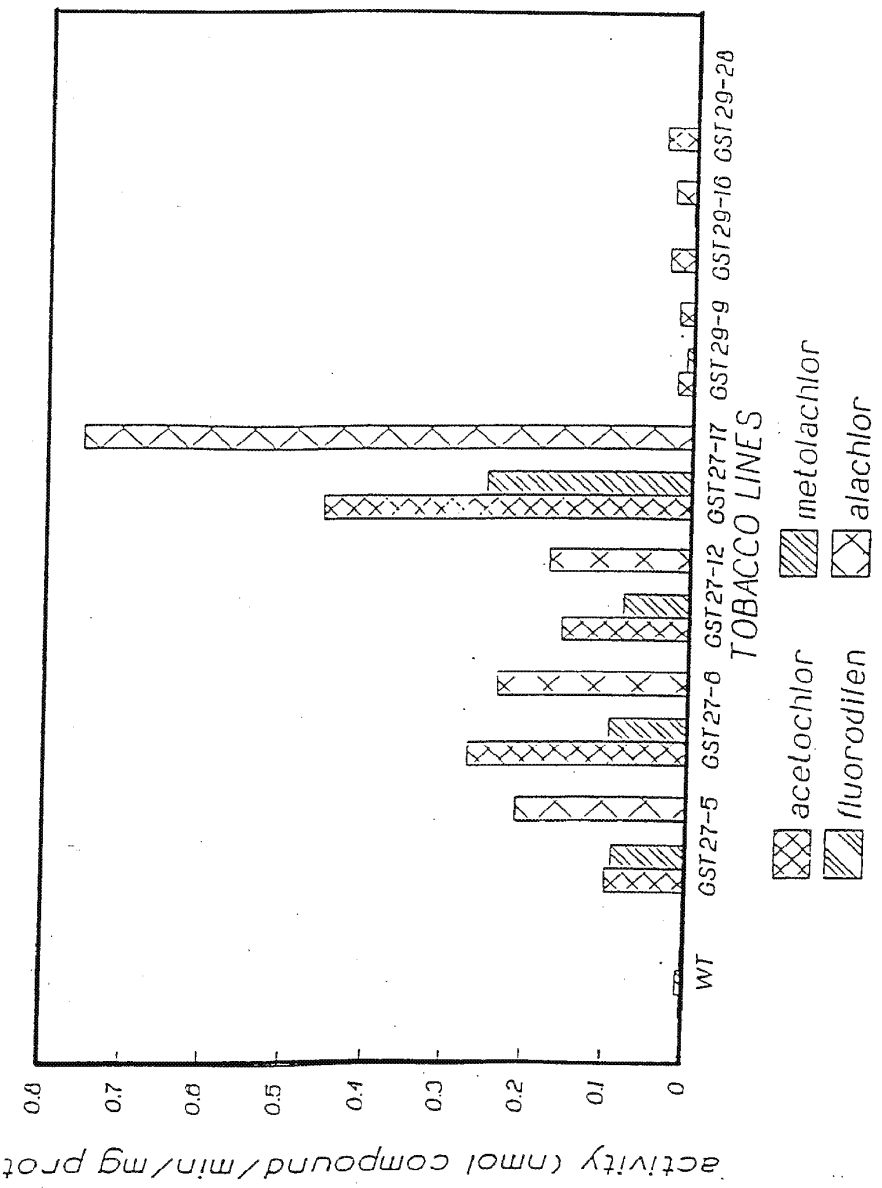

Transform into Agro-bacterium
EcoRI
RB
Camv
GSTII 27
pGST-27 Bin
nos
HindIII
LB
kan

Fig 6 (cont)

Fig. 7


activity (nmol compound/min/mg prot)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
WT
GST27-5
GST27-6
GST27-12
GST27-17
GST29-9
GST29-16
GST29-28
TOBACCO LINES
acetochlor
metolachlor
fluorodifen
alachlor

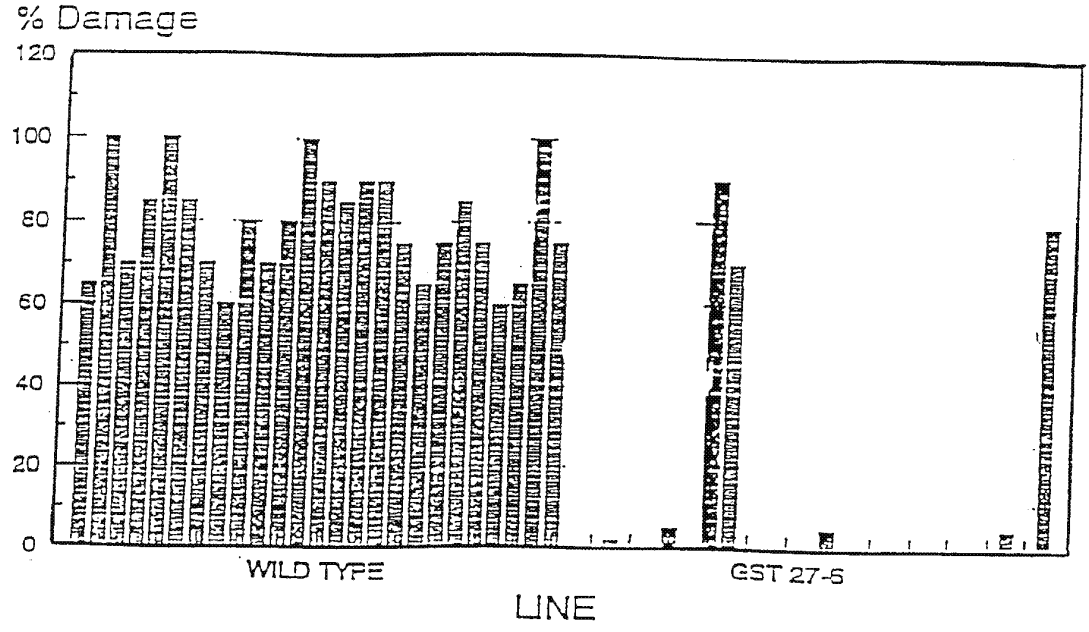
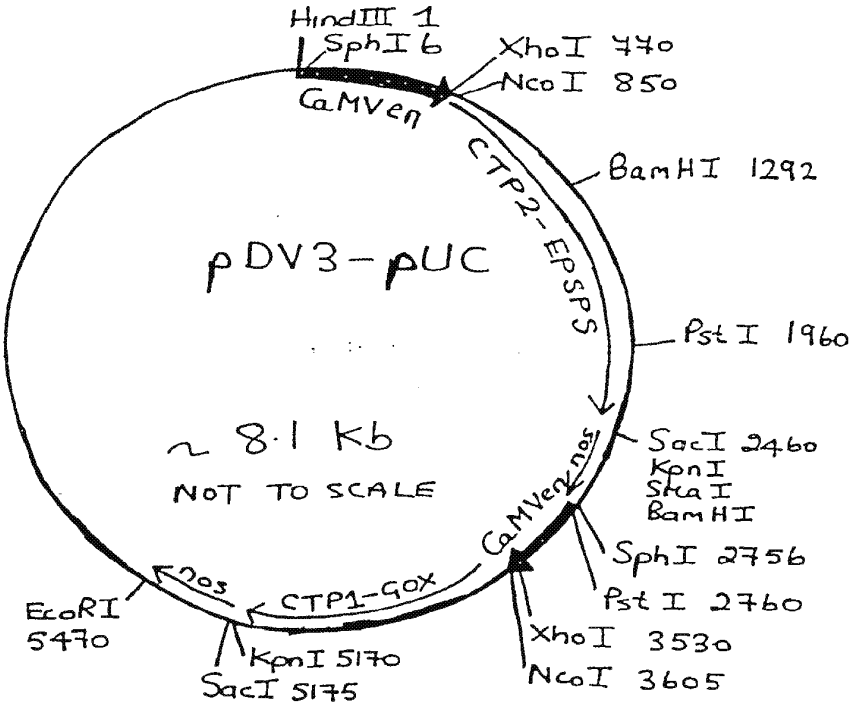


Fig. 8

COR-AAD1-00001871

Fig.9


HindIII 1
SphI 6
XhoI 770
NcoI 850
CaMVen
CTP2-EPSPS
BamHI 1292
pDV3-pUC
PstI 1960
~ 8.1 kb
NOT TO SCALE
SacI 2460
KpnI
SstI
BamHI
nos
CaMVen
SphI 2756
PstI 2760
XhoI 3530
NcoI 3605
nos
CTP1-GOX
EcoRI
5470
KpnI 5170
SacI 5175

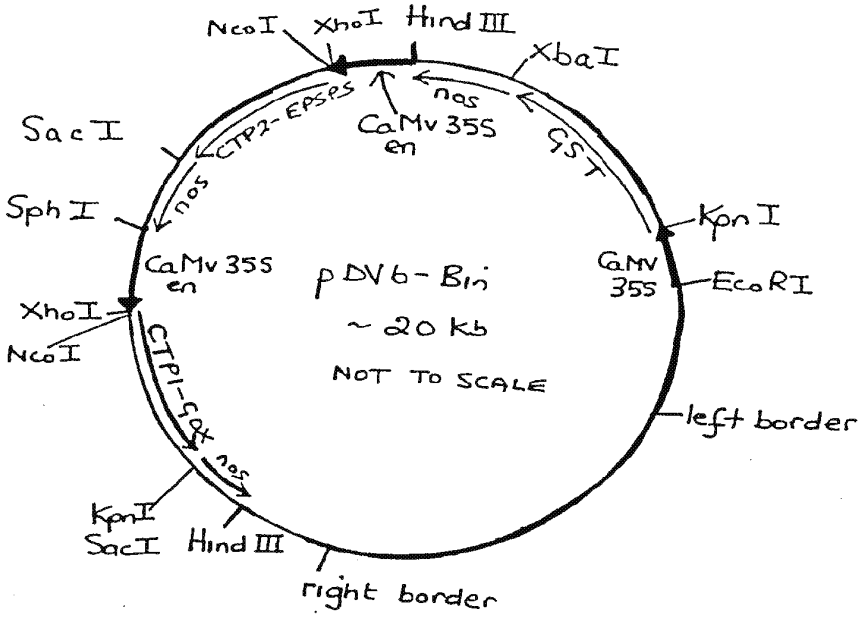

NcoI
XhoI
Hind III
XbaI
SacI
CTP2-EPSPS
nos
CaMv 35S
en
GS
SphI
KpnI
CaMv 35S
en
pDV6-Bin
~20 kb
NOT TO SCALE
GMV
35S
EcoRI
XhoI
NcoI
CTP1-GOX
nos
left border
KpnI
SacI
Hind III
right border

Fig. 10

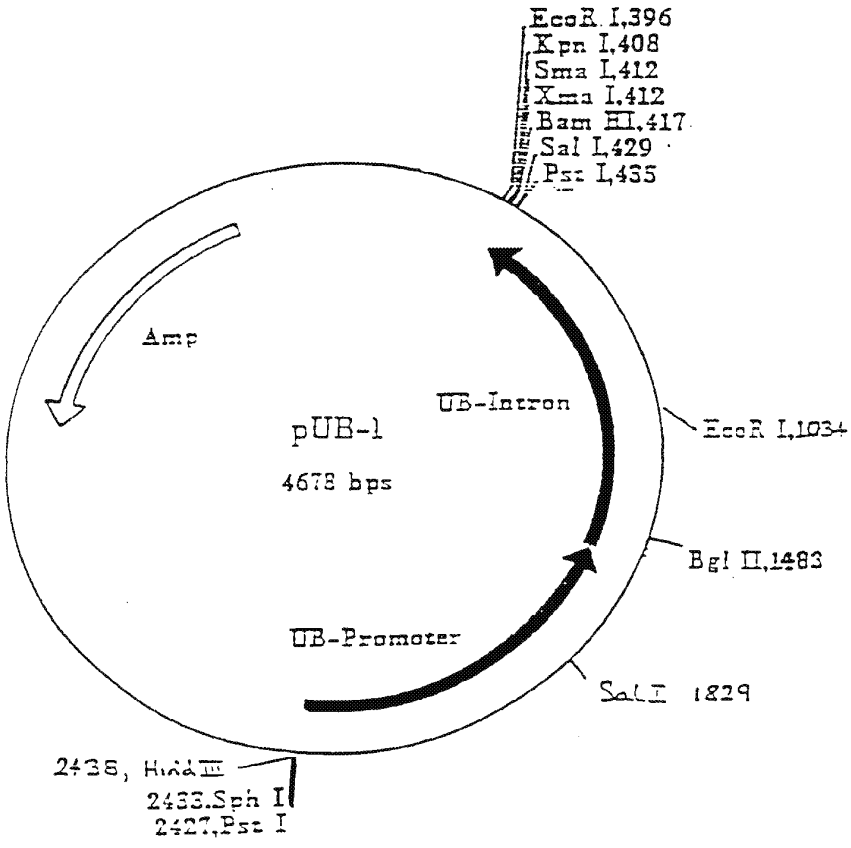

EcoR I,396
Kpn I,408
Sma I,412
Xma I,412
Bam HI,417
Sal I,429
Pst I,435
Amp
UB-Intron
pUB-1
4678 bps
EcoR I,1034
Bgl II,1483
UB-Promoter
Sal I 1829
2438, HindIII
2433,Sph I
2427,Pst I

Fig.11

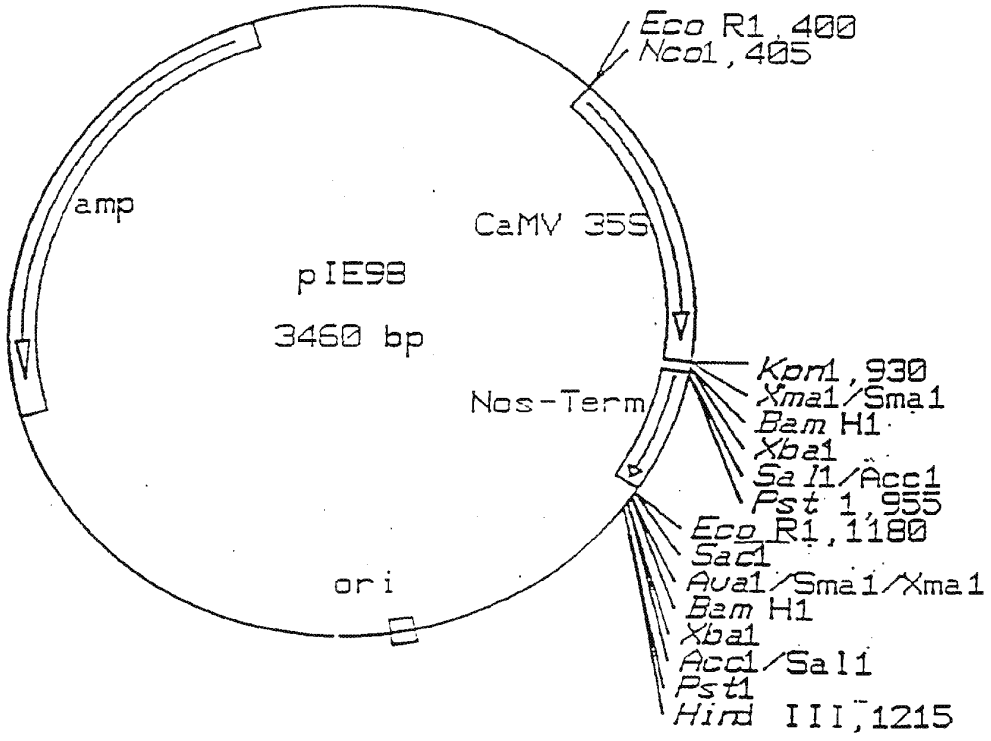

Eco R1, 400
Nco1, 405
amp
CaMV 35S
pIE98
3460 bp
Nos-Term
ori
Kpn1, 930
Xma1/Sma1
Bam H1
Xba1
Sal1/Acc1
Pst 1, 955
Eco R1, 1180
Sac1
Ava1/Sma1/Xma1
Bam H1
Xba1
Acc1/Sal1
Pst1
Hind III, 1215

Fig. 12

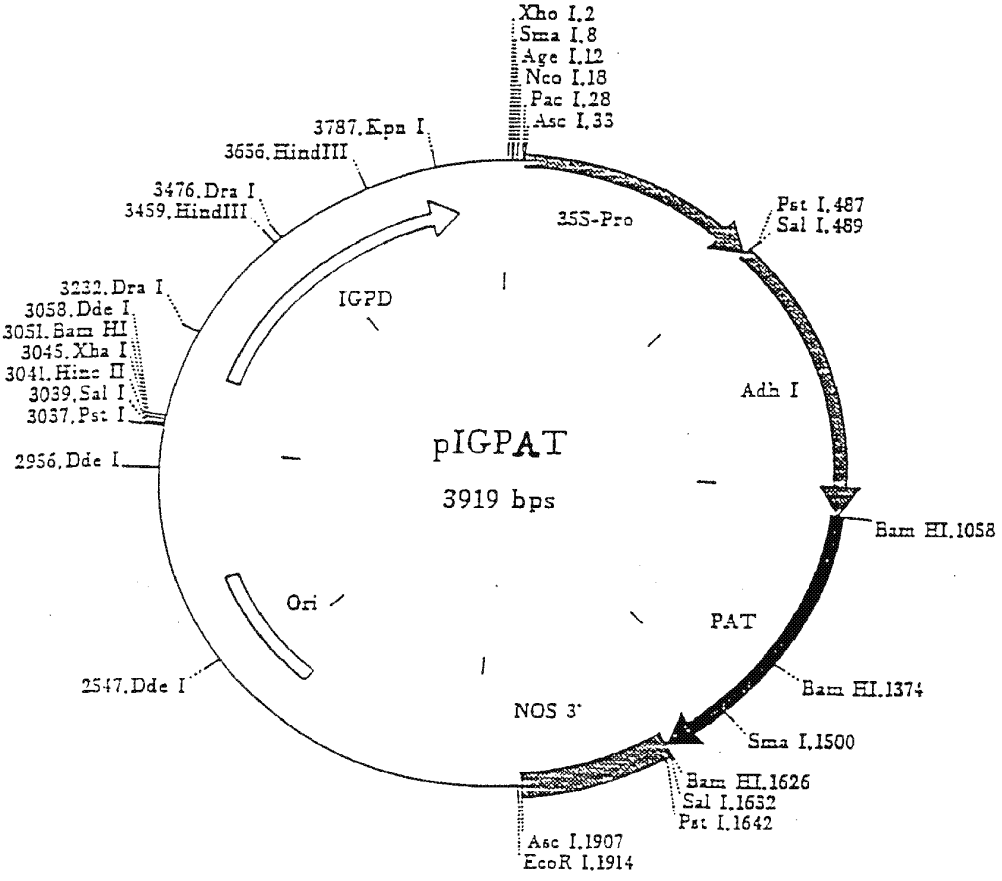

Xho I.2
Sma I.8
Age I.12
Nco I.18
Pac I.28
Asc I.33
3787.Kpn I
3656.HindIII
3476.Dra I
3459.HindIII
3232.Dra I
3058.Dde I
3051.Bam HI
3045.Xba I
3041.Hinc II
3039.Sal I
3037.Pst I
2956.Dde I
2547.Dde I
35S-Pro
Pst I.487
Sal I.489
IGPD
Adh I
pIGPAT
3919 bps
Bam HI.1058
Ori
PAT
Bam HI.1374
NOS 3'
Sma I.1500
Bam HI.1626
Sal I.1632
Pst I.1642
Asc I.1907
EcoR I.1914

Fig. 13

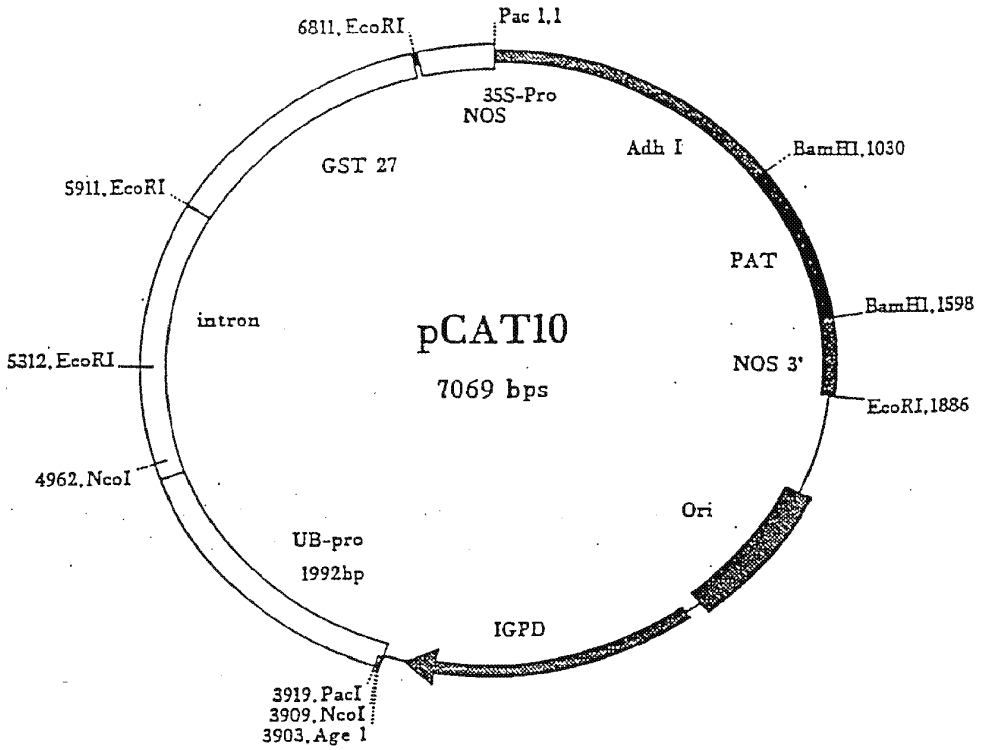

6811. EcoRI
Pac 1.1
35S-Pro
NOS
Adh I
BamHI.1030
GST 27
5911. EcoRI
PAT
BamHI.1598
intron
pCAT10
7069 bps
5312. EcoRI
NOS 3'
EcoRI.1886
4962. NcoI
Ori
UB-pro
1992bp
IGPD
3919. PacI
3909. NcoI
3903. Age 1

Fig. 14

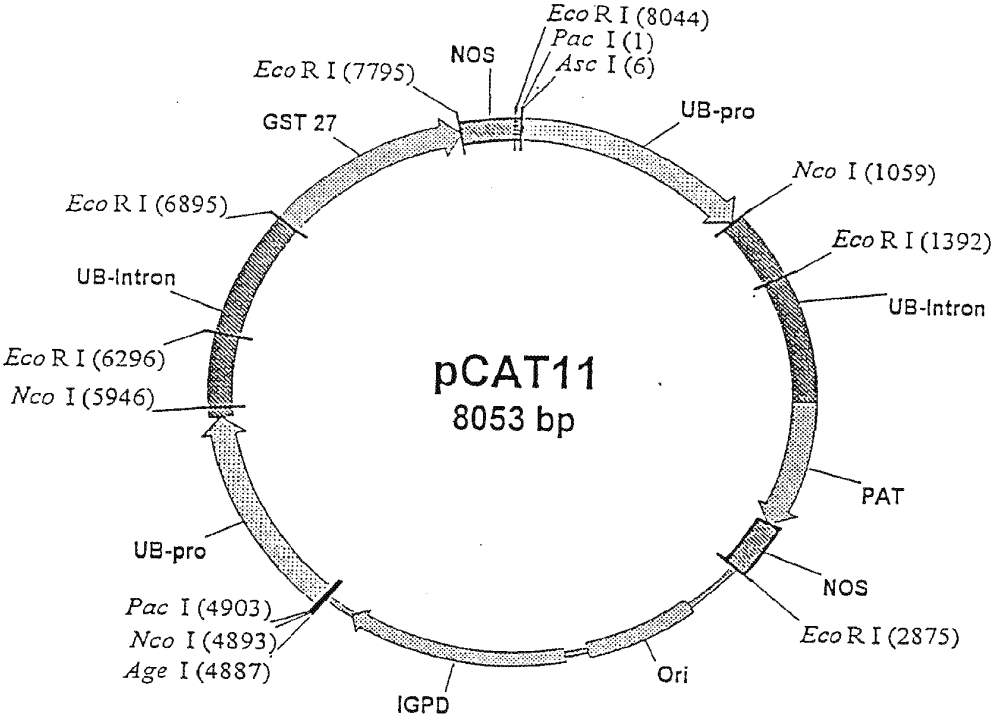

pCAT11
8053 bp
Eco R I (8044)
Pac I (1)
Asc I (6)
NOS
Eco R I (7795)
GST 27
UB-pro
Nco I (1059)
Eco R I (6895)
Eco R I (1392)
UB-Intron
UB-Intron
Eco R I (6296)
Nco I (5946)
PAT
UB-pro
NOS
Pac I (4903)
Nco I (4893)
Age I (4887)
Eco R I (2875)
IGPD
Ori

Fig. 15

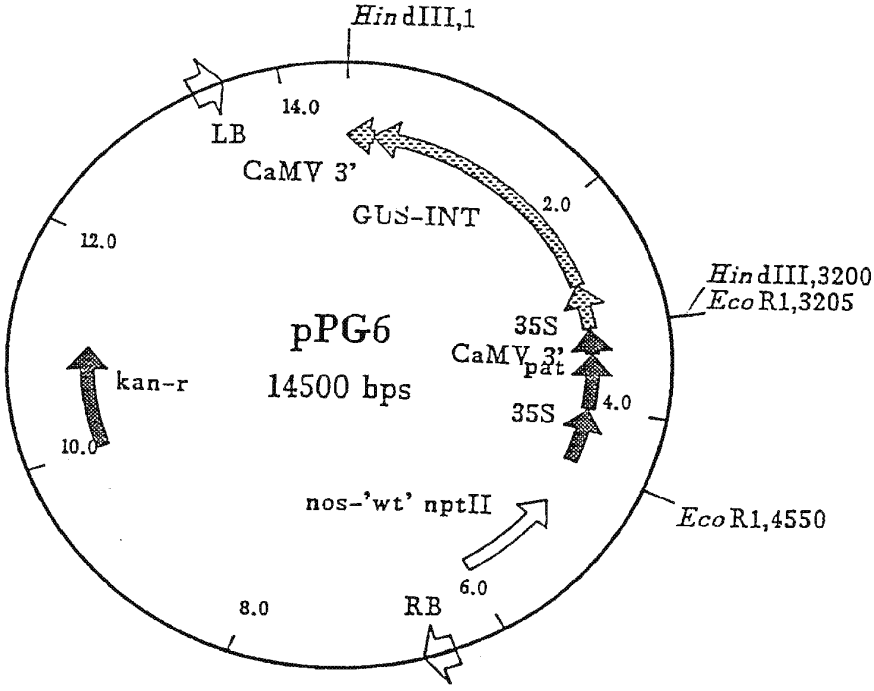

HindIII,1
14.0
LB
CaMV 3'
GUS-INT
2.0
12.0
HindIII,3200
EcoR1,3205
pPG6
14500 bps
35S
CaMV 3'
pat
35S
4.0
kan-r
10.0
nos-'wt' nptII
EcoR1,4550
6.0
8.0
RB

Fig. 16

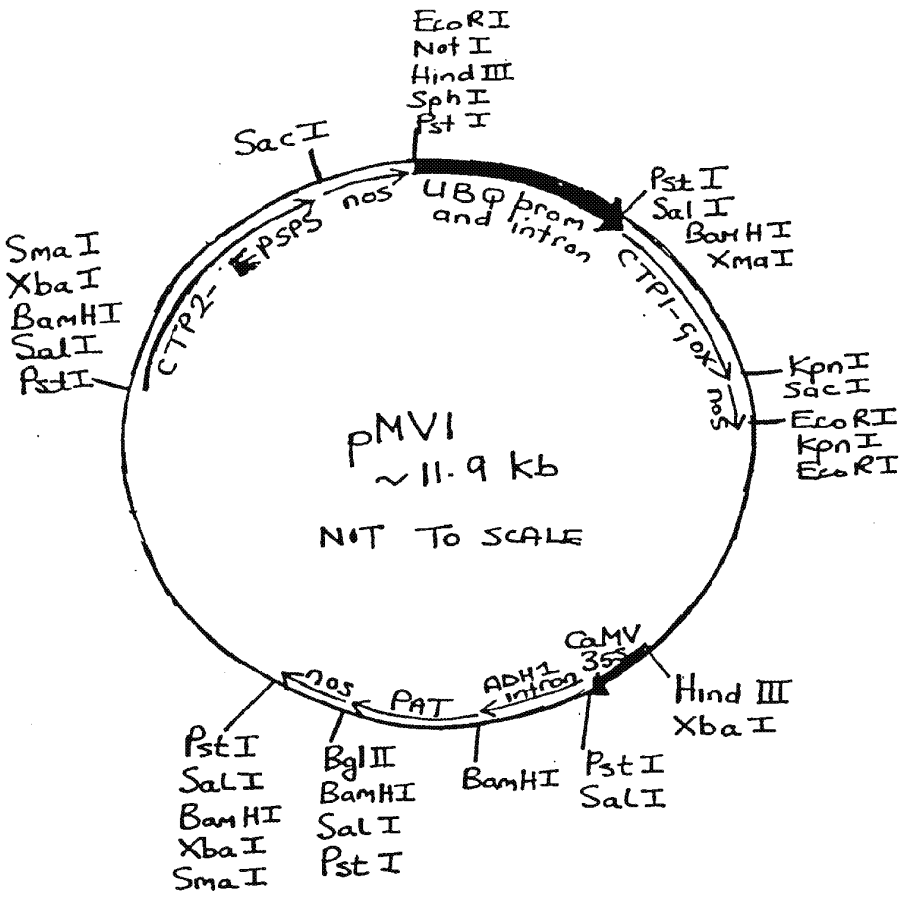

EcoRI
Not I
Hind III
Sph I
Pst I
SacI
Pst I
Sal I
BamHI
XmaI
KpnI
SacI
EcoRI
KpnI
EcoRI
Hind III
XbaI
PstI
SalI
BamHI
XbaI
SmaI
BglII
BamHI
SalI
PstI
BamHI
PstI
SalI
SmaI
XbaI
BamHI
SalI
PstI
UBQ prom and intron
CTP1-gox
nos
CaMV 35S
ADH1 intron
PAT
nos
CTP2-EPSPS
nos
pMV1
~11.9 kb
NOT TO SCALE

Fig. 17

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/00, 15/09, 15/29, 15/82, A01H 1/00, 3/00, 4/00, 5/00

A1

(11) International Publication Number: **WO 98/38294**

(43) International Publication Date: 3 September 1998 (03.09.98)

(21) International Application Number: PCT/US98/03670

(22) International Filing Date: 26 February 1998 (26.02.98)

(30) Priority Data:
08/807,671 27 February 1997 (27.02.97) US

(71) Applicant: UNION CAMP CORPORATION [US/US]; P.O. Box 3301, Princeton, NJ 08543–3301 (US).

(72) Inventor: KAPHAMMER, Bryan, J.; 22 Knight Drive, W. Plainsboro, NJ 08536 (US).

(74) Agent: BLOOM, Allen; Dechert Price & Rhoads, Princeton Pike Corporate Center, P.O. Box 5218, Princeton, NJ 08543–5218 (US).

(81) Designated States: BR, ID, JP, MX, NZ.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: GROWTH–ENHANCED TRANSGENIC PLANTS

(57) Abstract

The present invention relates, in general, to transgenic plant cells and plants, and further relates to, for example, (1) a method of growing a growth–enhanced transgenic perennial plant, comprising cultivating a transgenic plant precursor, wherein cells in the transgenic plant precursor express a 2,4–dichlorophenoxyacetic acid monooxygenase gene, particularly the *tfdA* gene, at a growth–enhancing level; (2) a method of growing a yield–enhanced plantation, comprising cultivating a plot of growth–enhanced transgenic perennial plants; (3) a growth–enhanced transgenic plant comprising a transgenic plant wherein cells in the plant express a 2,4–dichlorophenoxyacetic acid monooxygenase gene, particularly the *tfdA* gene, at a growth–enhancing level; and (4) a yield–enhanced plantation, comprising a plot of growth–enhanced transgenic perennial plants.

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

GROWTH-ENHANCED TRANSGENIC PLANTS

The present invention relates, in general, to transgenic plants and plant precursors, and further relates to, for example, (1) a method of growing a growth-enhanced transgenic perennial plant, comprising cultivating a transgenic plant precursor, wherein cells in the transgenic plant precursor express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing level; (2) a method of growing a yield-enhanced plantation, comprising cultivating a plot of growth-enhanced transgenic perennial plants; (3) a growth-enhanced transgenic perennial plant or plant precursor comprising a transgenic plant or plant precursor wherein cells in the plant or plant precursor express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing level; and (4) a yield-enhanced plantation, comprising a plot of growth-enhanced transgenic perennial plants.

2,4-dichlorophenoxyacetic acid (2,4-D) is a herbicide used to control broadleaf weeds. It can be degraded according to the pathway shown in Figure 1. The first step of the degradation pathway is catalyzed by a monooxygenase which catalyzes the conversion of 2,4-D to 2,4-dichlorophenol (DCP). The *tfdA* gene, isolated from the soil bacterium *Alcaligenes eutrophus*, encodes a 2,4-dichlorophenoxyacetic acid monooxygenase which catalyzes the conversion of 2,4-D to DCP. Transgenic tobacco and cotton plants transformed with the *tfdA* gene have been shown to have increased tolerance to 2,4-D (Streber *et al., Bio/technology* 7: 811-816, 1989 and Bayley *et al., Theor. Appl. Genet.* 83: 645-649, 1992, respectively). In U.S. Application Serial No. 08/358,117, it was taught how the *tfdA* gene can be used directly as a selectable marker to identify transgenic plants.

COR-AAD1-00001883

The present invention relates to the cultivation of growth-enhanced plants and yield-enhanced plantations, particularly perennial plants. It has been unexpectedly discovered that plants transformed with a 2,4-dichlorophenoxyacetic acid monooxygenase such as the *tfdA* gene that encodes enhanced growth as compared to plants not transformed with the gene.

Summary of the Invention

The invention provides a method of obtaining a growth-enhanced transgenic perennial plant, comprising cultivating a transgenic plant precursor, wherein cells in the transgenic plant precursor express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing level.

The methods of the invention further include a method for obtaining a yield-enhanced plantation, comprising cultivating a plot of growth-enhanced transgenic perennial plants, wherein cells in the transgenic plants express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

In preferred embodiments, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

In another preferred embodiment of the invention, the growth enhancement is measurable after more than one growing season.

In yet other preferred embodiments, the plant precursor or plant is a woody perennial plant or plant precursor. Preferably, the plant or precursor is of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome, Broussonetia, Camellia, Camptobera, Carya, Castanea* (Chestnut), *Catalpa, Cinchona, Citrus, Coffea, Corylus* (Hazelnut), *Diospyrus, Eucalyptus, Fagus* (Beech), *Ficus, Fraxinus* (Ash), *Gingko, Gleditsia, Gmelina, Hamamelia, Hedera, Ilex, Juglans, Kalmia, Lexceana, Liquidambar* (Sweetgum), *Liriodendron, Malus, Moghania, Morus,*

COR-AAD1-00001884

Olea, Paulownia, Populus, Prunus, Quercus (Oak), *Rosa, Rhododendron, Robinia, Salix, Santalum, Sapium, Simmondsia, Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus* (Elm), *Vaccinium*, or *Vitus,* or of the order Coniferales. More preferably, the precursor or plant is sweetgum.

The invention also includes a growth-enhanced transgenic perennial plant or plant precursor comprising a transgenic plant or plant precursor, wherein cells in the plant or plant precursor express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

In preferred embodiments, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

In another preferred embodiment, the plant or plant precursor is a woody perennial. Preferably the woody perennial is of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome, Broussonetia, Camellia, Camptobera, Carya, Castanea* (Chestnut), *Catalpa, Cinchona, Citrus, Coffea, Corylus* (Hazelnut), *Diospyrus, Eucalyptus, Fagus* (Beech), *Ficus, Fraxinus* (Ash), *Gingko, Gleditsia, Gmelina, Hamamelia, Hedera, Ilex, Juglans, Kalmia, Lexceana, Liquidambar* (Sweetgum), *Liriodendron, Malus, Moghania, Morus, Olea, Paulownia, Populus, Prunus, Quercus* (Oak), *Rosa, Rhododendron, Robinia, Salix, Santalum, Sapium, Simmondsia, Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus* (Elm), *Vaccinium*, or *Vitus,* or of the order Coniferales. More preferably, the woody perennial is sweetgum.

The invention also includes a yield-enhanced plantation, comprising a plot of growth-enhanced transgenic perennial plants, wherein cells in the transgenic plants express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

In a preferred embodiment, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

In another preferred embodiment, the plant or plant precursor is a woody perennial. Preferably the woody perennial is of the genera *Acacia*,

COR-AAD1-00001885

Acer (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome*, *Broussonetia*, *Camellia*, *Camptobera*, *Carya*, *Castanea* (Chestnut), *Catalpa*, *Cinchona*, *Citrus*, *Coffea*, *Corylus* (Hazelnut), *Diospyrus*, *Eucalyptus*, *Fagus* (Beech), *Ficus*, *Fraxinus* (Ash), *Gingko*, *Gleditsia*, *Gmelina*, *Hamamelia*, *Hedera*, *Ilex*, *Juglans*, *Kalmia*, *Lexceana*, *Liquidambar* (Sweetgum), *Liriodendron*, *Malus*, *Moghania*, *Morus*, *Olea*, *Paulownia*, *Populus*, *Prunus*, *Quercus* (Oak), *Rosa*, *Rhododendron*, *Robinia*, *Salix*, *Santalum*, *Sapium*, *Simmondsia*, *Sycamore*, *Tectona*, *Theobrama*, *Tupidanthus*, *Ulmus* (Elm), *Vaccinium*, or *Vitus*, or of the order Coniferales. More preferably, the woody perennial is sweetgum.

Further objects and advantages of the present invention will be clear from the description that follows.

Brief Description of the Drawings

FIGURE 1. A 2,4-D degradative pathway. Gene designations are shown in parentheses for each enzyme.

FIGURE 2. Construction of pUCW101.

FIGURE 3. Construction of pUCW200.

FIGURE 4. ELISA of sweetgum 2027 selected on 2,4-D or kanamycin.

FIGURE 5. Bar graph of the average heights of five categories of sweetgum clones (ten trees per category).

FIGURE 6. Bar graph of the average diameters at breast heights of five categories of sweetgum clones (ten trees per category).

Definitions

An *analog* of a protein or genetic sequence should be understood as referring to a protein or genetic sequence substantially similar in function to a protein or genetic sequence described herein.

COR-AAD1-00001886

Exogenous DNA should be understood as referring to DNA introduced into a cell or organism, or into an ancestor of such a cell or organism, by a recombinant technology. As will be understood by those of ordinary skill, an exogenous DNA remains in an ancestor if the original introduction to the parental cell or organism created a heritable form of the DNA within the cell or organism.

A *fragment* of a molecule should be understood as referring to any portion of the amino acid or nucleotide genetic sequence.

A *functional derivative* should be understood as referring to a molecule that possesses a biological activity (either functional or structural) that is substantially similar to a biological activity of a protein or nucleic acid sequence.

Growing season should be understood to mean the period of a year commencing with bud break of a plant and continuing until dormancy of the plant, during which period the bulk of a plant's growing occurs. Generally, there is one growing season per calendar year.

A *mutation* should be understood as referring to any detectable change in the genetic material which can be transmitted at least one generation, that is, to daughter cells giving rise to mutant cells or mutant individuals. If the descendants of a mutant cell give rise only to somatic cells in multicellular organisms, a mutant spot or area of cells arises. Mutations in the germ line of sexually reproducing organisms which are transmitted by the gametes to the next generation can result in an individual with the new mutant condition in both its somatic and germ cells. A mutation can be any (or a combination of) detectable, unnatural change affecting the chemical or physical constitution, mutability, replication, phenotypic function, or recombination of one or more deoxyribonucleotides; nucleotides may be added, deleted, substituted for, inverted, or transposed to new positions with and without inversion. Mutations can occur spontaneously or can be induced experimentally by application of mutagens. A mutant variation of a nucleic acid molecule results from a mutation. A mutant polypeptide may result from a mutant nucleic acid

COR-AAD1-00001887

molecule.

Plant should be understood as referring to a multicellular differentiated organism capable of photosynthesis including angiosperms (monocots and dicots) and gymnosperms.

Plant cell should be understood as referring to the structural and physiological unit of plants. The term "plant cell" refers to any cell which is either part of or derived from a plant.

Plant cell progeny should be understood as referring to any cell or tissue derived from plant cells including callus; plant parts such as stems, roots, fruits, leaves or flowers; plants; plant seed; pollen; and plant embryos.

Plant precursor should be understood as referring to any biological matter from which a plant can be generated or grown, including, but not limited to, one or more plant cells, plant cell progeny, propagules, seeds, seedlings, cuttings, plantlets and plants.

Plantation should be understood as referring to a group of plants under cultivation.

Propagules should be understood as referring to any plant material capable of being sexually or asexually propagated, or being propagated *in vivo* or *in vitro*. Such propagules preferably consist of the protoplasts, cells, calli, tissues, embryos or seeds of the regenerated plants.

To *transform* or the *transformation* of a plant or plant precursor should be understood to mean the process of introducing exogenous DNA into a plant or plant precursor to produce a transgenic plant.

Transgenic plant should be understood as referring to a plant having stably incorporated exogenous deoxyribonucleic acid ("DNA") in its genetic material. The term also includes plants where the exogenous DNA which was introduced into a cell or protoplast, or an ancestor of such a cell or protoplast, in various forms, including, for example, naked DNA in circular, linear or supercoiled form, DNA contained in nucleosomes or chromosomes or nuclei or parts thereof, DNA complexed or associated with other molecules, DNA enclosed in liposomes, spheroplasts, cells or protoplasts.

COR-AAD1-00001888

A *variant* of a nucleic acid should be understood as referring to a molecule substantially similar in structure and biological activity to the nucleic acid, or to a fragment thereof. Thus, provided that two molecules possess a common activity and may substitute for each other, they are considered variants as that term is used herein even if the nucleotide sequence is not identical.

A *woody perennial* should be understood to mean a perennial which is a tree, bush or liana.

Detailed Description of the Invention

The present invention relates, among other things, to a method of growing a growth-enhanced transgenic perennial plant, comprising cultivating a transgenic plant precursor, wherein cells in the precursor express a 2,4-dichlorophenoxyacetic acid monooxygenase at a growth-enhancing effective level.

A "growth-enhanced" plant should be understood to mean a plant which exhibits a greater increase in volume, for example, as measured by height and diameter of trunk or stem, over a growing season than a non-growth-enhanced plant. An "effective level" of expression of the 2,4-dichlorophenoxyacetic acid monooxygenase gene in cells of a transgenic plant precursor should be understood to mean a level of expression sufficient to cause the plant to be a growth-enhanced plant. Expression by "cells" in the plant precursor should be understood to mean expression in a sufficient number of cells to cause the plant to be a growth-enhanced plant.

In a preferred embodiment, the growth enhancement of the transgenic plant is measurable after more than one growing season.

Preferably, the transgenic plant precursor is a woody perennial. More preferably, the woody perennial is of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome*, *Broussonetia*, *Camellia*, *Camptobera*, *Carya*, *Castanea* (Chestnut), *Catalpa*,

COR-AAD1-00001889

Cinchona, Citrus, Coffea, Corylus (Hazelnut), *Diospyrus, Eucalyptus, Fagus* (Beech), *Ficus, Fraxinus* (Ash), *Gingko, Gleditsia, Gmelina, Hamamelia, Hedera, Ilex, Juglans, Kalmia, Lexceana, Liquidambar* (Sweetgum), *Liriodendron, Malus, Moghania, Morus, Olea, Paulownia, Populus, Prunus, Quercus* (Oak), *Rosa, Rhododendron, Robinia, Salix, Santalum, Sapium, Simmondsia, Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus* (Elm), *Vaccinium*, or *Vitus*, or of the order Coniferales (collectively referred to here as "Preferred Plants"). Even more preferably, the woody perennial is sweetgum.

In other preferred embodiments, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

The invention further provides a method of growing a yield-enhanced plantation, comprising cultivating a plot of growth-enhanced transgenic perennial plants, wherein cells in the transgenic plants express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

A "yield-enhanced" plantation should be understood to mean a plantation which exhibits more product per acre, for example, more wood, than a plantation comprising similar age and number of non-growth-enhanced plants. Preferably, the plants are woody perennials, more preferably, Preferred Plants, and even more preferably, sweetgum. It should be understood that a plantation can comprise any combination of perennials.

In preferred embodiments, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

The invention also provides a growth-enhanced transgenic perennial plant or plant precursor, wherein cells of the plant or precursor express a growth-enhancing effective level of a 2,4-dichlorophenoxyacetic acid monooxygenase gene. Preferably, the plants are woody perennials, more preferably Preferred Plants, and even more preferably, sweetgum.

In preferred embodiments, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

COR-AAD1-00001890

The invention also provides a yield-enhanced plantation, comprising a plot of growth-enhanced transgenic perennial plants, wherein cells of the plants express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level. Preferably, the plants are woody perennials, more preferably, the plants are Preferred Plants, and even more preferably, the plants are sweetgum.

In preferred embodiments, the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

For all aspects of the invention, the transgenic plants and plant precursors can be cultivated in a greenhouse, or in the field, meaning outside of any protective enclosure such as a greenhouse. In a preferred embodiment, as further described in Example 6, after transformation the plants are grown in a greenhouse until they reach planting size, and are then planted in the field.

The invention relates to the unexpected discovery that plants transformed with an expressible 2,4-dichlorophenoxyacetic acid monooxygenase gene grow faster than corresponding untransformed plants. The transgenic plants which are the subject of the invention are created by genetic engineering methods of transformation appropriate for plants, which generally means that at least a parent of the plants was generated from cells treated by a method intended to insert a 2,4-dichlorophenoxyacetic acid monooxygenase gene into the cells, as discussed further below. Generally, the potentially transformed plant will be monitored or selected for expressing the 2,4-dichlorophenoxyacetic acid monooxygenase gene. This can, for example, be done by: (a) direct selection on 2,4-dichlorophenoxyacetic acid, as described below; (b) monitoring for the presence of 2,4-dichlorophenoxyacetic acid monooxygenase activity extracted from plant tissue, for instance using the HPLC method of detecting 2,4-dichlorophenol (DCP) set forth below; (c) monitoring for the presence of the monooxygenase in plant extract, for instance using an immuno-detection assay; (d) monitoring for the presence of a marker gene that was concurrently introduced into the

COR-AAD1-00001891

plant with the 2,4-dichlorophenoxyacetic acid monooxygenase gene; or (e) selecting for transformants using a selectable marker gene that was concurrently introduced into the plant with the 2,4-dichlorophenoxyacetic acid monooxygenase gene. The gene will generally be present in cells found throughout the plant, though the level of expression in different tissue types will generally vary depending on, for example, the differences in utilization in the various tissues of the promoter driving the expression of the 2,4-dichlorophenoxyacetic acid monooxygenase gene.

All plants which can be transformed are intended to be included within the scope of the invention. Such plants include, for example, species from the genera *Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Hererocallis, Nemesia, Pelargonium, Panicum, Pennisetum, Ranunculus, Sencia, Salpiglossis, Cucumis, Browalia, Glycine, Lolium, Zea, Triticum, Sorghum, Malus, Apium, Datura*, and the like.

Suitable woody perennials include, for example, the Preferred Plants, and the like.

Plant transformation techniques are well known in the art and include but are not limited to, microinjection (Crossway, *Mol. Gen. Genetics* 202: 179-185, 1985), polyethylene glycol transformation (Krens *et al., Nature* 296: 72-74, 1982), high velocity ballistic penetration (Klein *et al., Nature* 327: 70-73, 1987), fusion of protoplasts with other entities, either minicells, cells, lysosomes, or other fusible lipid-surfaced bodies (Fraley *et al., Proc. Natl. Acad. Sci.* USA 79: 1859-1863, 1982), electroporation (Fromm *et al., Proc. Natl. Acad. Sci. USA* 82: 5824, 1985) and techniques set forth in U.S. Patent No. 5,231,019)) and *Agrobacterium tumefaciens* mediated transformation (Hoekema *et al., Nature* 303: 179, 1983, de Framond *et al., Bio/Technology* 1: 262, 1983, Fraley *et al.* WO84/02913, WO84/02919 and WO84/02920,

COR-AAD1-00001892

Zambryski *et al.* EP 116,718, Jordan *et al., Plant Cell Reports* 7: 281-284, 1988, Leple *et al., Plant Cell Reports* 11: 137-141, 1992, Stomp *et al., Plant Physiol.* 92: 1226-1232, 1990, and Knauf *et al., Plasmid* 8: 45-54, 1982). One preferred method of transformation is the leaf disc transformation technique as described by Horsch *et al., Science* 227: 1229-1230, 1985.

The above-described transformation techniques can utilize a 2,4-dichlorophenoxyacetic acid monooxygenase gene, such as the *tfdA* gene or fragment thereof. Included within the scope of the *tfdA* gene is the *tfdA* gene of nucleotide sequence set forth in Streber *et al., J. of Bacteriology* 169:2950-2955, 1987, and also in GenBank Accession No.: M16730, functional derivatives of a *tfdA* gene, as well as variant, analog and mutational derivatives.

The *tfdA* gene that is depicted in Streber and GenBank Accession No.: M16730, can be altered by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence, such as a *tfdA* gene product depicted in GenBank Accession No.: M16730, may be used in the practice of the present invention. These include but are not limited to nucleotide sequences comprising all or portions of the *tfdA* gene depicted in Streber and GenBank Accession No.: M16730, which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change.

Such functional alterations of a given nucleic acid sequence afford an opportunity to promote secretion and/or processing of heterologous proteins encoded by foreign nucleic acid sequences fused thereto. All variations of the nucleotide sequence of the relevant gene and fragments thereof permitted by the genetic code are, therefore, included in this invention.

In addition, the dichlorophenoxyacetic acid monooxygenase gene comprises a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the

COR-AAD1-00001893

nucleic acid formula shown in Streber and GenBank Accession No.: M16730 or a derivative thereof. Any nucleotide or polynucleotide may be used in this regard, provided that its addition, deletion or substitution does not destroy the enzymic activity of the encoded gene. In a more preferred embodiment, the change does not alter the amino acid sequence of Streber and GenBank Accession No.: M16730 which is encoded by the nucleotide sequence. The *tfdA* gene may, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end.

Further, it is possible to delete codons or to substitute one or more codons by codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity of the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally equivalent, as are the two nucleic acid molecules which give rise to their production, even though the differences between the nucleic acid molecules are not related to degeneracy of the genetic code.

The term "2,4-dichlorophenoxyacetic acid monooxygenase gene" includes any nucleotide sequence encoding for a polypeptide which functions as a 2,4-dichlogophenoxyacetic acid monooxygenase. The 2,4-dichlorophenoxyacetic acid monooxygenase gene used for transformation is preferably operably linked to a promoter region functional in plants, a transcription initiation site, and a transcription termination sequence. The particular promoter used in the expression cassette is a noncritical aspect of the invention. Any of a number of promoters which direct transcription in a plant cell is suitable. The promoter can be either constitutive or inducible. Some examples of promoters functional in plants include the nopaline synthase promoter derived from the *Agrobacterium* Ti plasmid (available in vectors pBI101, pBI101.2 and pBI101.3 from Clontech, Palo Alto, CA) and other promoters derived from native Ti plasmids, viral promoters including the 35S and 19S RNA promoters of cauliflower mosaic virus (Odell *et al., Nature* 313: 810-812, 1985, with the 35S promoter available in vector pBI1221 from

COR-AAD1-00001894

Clontech), the p1'2' *mas* bidirectional promoter (Velten et al., *EMBO J.* 3: 2723-2730,1984, where one orientation is referred to as the "*mas* 1' promoter" and the reverse orientation as the "*mas* 2' promoter"), the light-inducible promoter of gene ST-LS1 (Stockhaus et al., *Proc. Natl. Acad. Sci. USA* 84: 7943-7947, 1987), and numerous plant promoters. A promoter that is well suited for a particular use can be determined by standard subcloning methods and the plant transformation and propagation methods taught herein. Efficiency for a particular purpose will often vary on the strength of the promoter and the tissue-by-tissue variations in promoter utilization. Plants transformed using various promoters can be compared for growth enhancement.

General methods for selecting transformed plant cells containing a selectable marker are well known and taught, for example, by Herrera-Estrella, L. and Simpson, J., "Foreign Gene Expression in Plants," in *Plant Molecular Biology, A Practical Approach*, C.H. Shaw, ed., IRL Press, Oxford, England, 1988, pp. 131-160. For example, for use of a 2,4-D gene as a selectable marker, the amount of 2,4-D which inhibits adventitious shoot formation from non-transformed plant cells and allows adventitious shoot formation from transformed plant cells can be determined by 1) plating non--transformed cells on media containing various concentrations of 2,4-D and 2) by determining the lowest concentration of 2,4-D which will inhibit adventitious shoot formation by the plant cells. This lowest concentration can then be used to select transformed plant cells. In general, solubilized 2,4-D should be present in an amount ranging from about 0.001 to about 5 mg/L culture medium. With regard, for example, to the sweetgum plant cells transformed with a *tfdA* gene, 2,4-D should preferably be present in an amount ranging from about 0.01 to about 0.5 mg/L culture medium. A preferred amount of 2,4-D is about 0.01 to about 0.2 mg/L culture medium. The determination of the lowest amount of 2,4-D which will inhibit adventitious shoot formation is taught in U.S. Application Serial No. 08/358,117.

Other foreign marker genes (i.e., exogenously introduced genes)

COR-AAD1-00001895

typically used in selection of transformed cells include selectable markers such as a *neo* gene (Potrykus *et al., Mol. Gen. Genet* 199: 183-188, 1985), which codes for kanamycin resistance; a *bar* gene which codes for bialaphos resistance; a mutant EPSP synthase gene (Hinchee *et al., Bio/technology* 6: 915-922, 1988), which encodes glyphosate resistance; a nitrilase gene which confers resistance to bromoxynil (Stalker *et al., J. Biol. Chem.* 263: 6310-6314, 1988); a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance (EP application number 154,204); a methotrexate resistant DHFR gene (Thillet *et al., J. Biol. Chem.* 263: 12500-12508) and screenable markers which include ß-glucuronidase (GUS) or an R-locus gene, alone or in combination with a C-locus gene (Ludwig *et al., Proc. Natl. Acad. Sci. USA* 86: 7092, 1989; Paz-Ares *et al., EMBO J.* 6: 3553, 1987).

A 2,4-dichlorophenoxyacetic acid monooxygenase gene may be introduced into sweetgum by the transformation techniques outlined above or more preferably as set forth in Chen, Z. and Stomp, A., 1991 "Transformation of *Liquidamber styraciflua* L. (Sweetgum) using *Agrobacterium Tumefaciens,*" In: Proceedings 21st Southern Forest Tree Improvement Conference. June 17-20, 1991 Knoxville, TN. The 2,4-dichlorophenoxyacetic acid monooxygenase gene is preferably contained on a plasmid wherein the gene is operably linked to a promoter region functional in plants, a transcription initiation site, and a transcription termination sequence (examples of which are provided above). In one preferred embodiment, the gene is linked to a foreign marker gene (described above).

After selection, the transformed plant cells can be regenerated into transgenic plants. Plant regeneration techniques are well known in the art and include those set forth in the *Handbook of Plant Cell Culture*, Volumes 1-3, Evans *et al.*, eds., Macmillan Publishing Co., New York, NY, 1983, 1984, 1984, respectively; Predieri and Malavasi, *Plant Cell, Tissue, and Organ Culture* 17: 133-142, 1989; James, D.J., *et al., J. Plant Physiol.* 132: 148-154, 1988; Fasolo, F., *et al., Plant Cell, Tissue, and Organ Culture* 16: 75-87,

COR-AAD1-00001896

1989; Valobra and James, *Plant Cell, Tissue, and Organ Culture* 21: 51-54, 1990; Srivastava, P.S., *et al., Plant Science* 42: 209-214, 1985; Rowland and Ogden, *Hort. Science* 27: 1127-1129, 1992; Park and Son, *Plant Cell, Tissue, and Organ Culture* 15: 95-105, 1988; Noh and Minocha, *Plant Cell Reports* 5: 464-467, 1986; Brand and Lineberger, *Plant Science* 57: 173-179, 1988; Bozhkov, P.V., *et al., Plant Cell Reports* 11: 386-389, 1992; Kvaalen and von Arnold, *Plant Cell, Tissue, and Organ Culture* 27: 49-57, 1991; Tremblay and Tremblay, *Plant Cell, Tissue, and Organ Culture* 27: 95-103, 1991; Gupta and Pullman, U.S. Patent No. 5,036,007; Michler and Bauer, *Plant Science* 77: 111-118, 1991; Wetzstein, H.Y., *et al., Plant Science* 64: 193-201, 1989; McGranahan, G.H., *et al., Bio/Technology* 6: 800-804, 1988; Gingas, V.M., *Hort. Science* 26: 1217-1218, 1991; Chalupa, V., *Plant Cell Reports* 9: 398-401, 1990; Gingas and Lineberger, *Plant Cell, Tissue, and Organ Culture* 17: 191-203, 1989; Bureno, M.A., *et al., Phys. Plant.* 85: 30-34, 1992; and Roberts, D.R., *et al., Can. J. Bot.* 68:1086-1090, 1990.

The present invention is described in further detail in the following non-limiting examples.

Examples

The following protocols and experimental details are referenced in the examples that follow.

COR-AAD1-00001897

a. Strain Sources and Growth Conditions.

The bacterial strains and plasmids used herein are listed in Table 1, and media formulas are shown in Table 2 and 3. *Pseudomonas aeruginosa* PAOlc containing plasmid pRO101 or plasmid pRO1727 were grown on TNA plates containing 50 μg/ml tetracycline (TC^{50}) at 37°C. *P. aeruginosa* PAO1c (pUCW101; Figure 2) was grown on TNA containing 500 μg/ml carbenicillin (Cb^{500}) at 37°C. *P. putida* PPO300 (pUCW200; Figure 3), *Agrobacterium tumefaciens* LBA4404 (pUCW200; Figure 3), *Escherichia coli* HB101 (pBI121, Figure 4), and *E. coli* S17-1 (pUCW200; Figure 3) were grown on TNA containing 50 μg/ml kanamycin (Km50). The *P. putida* and *A. tumefaciens* strains were grown at 30°C, and the *E. coli* strains were grown at 37°C.

Growth of *P. aeruginosa* PAOlc (pUCW101), *P. putida* PP0300 (pUCW200), and *A. tumefaciens* LBA4404 (pUCW200) for analysis of 2,4-D conversion to 2,4-dichlorophenol was done by inoculating 50 ml of Burk's/CAA media containing 1 mM 2,4-D with a loop of culture from an overnight TNA plate containing the appropriate antibiotic. These liquid cultures were shaken at 30°C for 4 hours, and then filter sterilized. The sterile filtrate was analyzed by HPLC as described below.

COR-AAD1-00001898

TABLE 1
BACTERIAL STRAINS AND PLASMIDS

Strain or Plasmid	Relevant Markers[a]	Reference or Source
Pseudomonas aeruginosa PA01c	Prototroph	Holloway[b]
Pseudomonas putida PPO300	Prototroph	ATCC[c]
Escherichia coili S17-1	tbi, pro, recA chromosomally integrated RP4	Simon[d]
Agrobacterium tumefaciens LBA4404	Sm	Clontech[e]
Plasmids pRO1727	Cb, Tc	Cuskey[f]
pRO101	2,4-D$^+$ Tc	Harker[g]
pBI121	Km, Gus$^+$	Clontech[d]
pUCW101	Cb, *tfd*A$^+$	Figure 2
pUCW200	Km, *tfd*A$^+$	Figure 3

a: *Abbreviations:* Sm: streptomycin. Cb: carbenicillin. Tc: tetracycline. Km: kanamycin. Gus: β-glucuronidase.

b: Holloway *et al., Microbiol. Rev.* 43: 73-102, 1979.

c: ATCC 17514, American Type Culture Collection, Rockville, MD.

d: Simon, R. *et al., Bio/Technology* 1: 784-791, 1983.

e: Clontech Laboratories, Inc., Palo Alto, CA.

f: Cuskey *et al., J. Bacteriol.* 169: 2398-2404, 1987.

g: Harker *et al., J. Bacteriol.* 171: 314-320, 1989.

COR-AAD1-00001899

TABLE 2

TNA	Tryptone	5.0 g/L
	Yeast Extract	2.5 g/L
	NaCl	8.5 g/L
	Glucose	1.0 g/L
	Agar	20.0 g/L
Autoclave and temper to 50°C. Add antibiotic if required and pour plates		
LB	Luria Broth Base	15.5 g/L
	Agar	20.0 g/L
Autoclave and temper to 50°C. Add antibiotic if required and pour plates.		
Burk's Salts*		
Stock Solutions:		
a.	$MgSO_4$-$7H_2O$	39.90 g/L
b.	$FeSo_4$-$7H_2O$	0.01 g/L
c.	$NaMoO_4$-$2H_2O$	0.05 g/L
d.	$(NH_4)_2SO_4$	100.00 g/L
e.	1 M Potassium Phosphate buffer, pH 7.1	
Autoclave stock solution and store at room temperature.		
Burk's/CAA		
To 1 L of sterile distilled water containing 0.3% casamino acids add: 5 ml of stock solutions a, b, and c; 10 ml of stock solutions d and e.		
Burk's/succinate plates		
To 1 L of distilled water containing 0.2% succinate and 2% noble agar, which has been autoclaved and tempered to 50°C add: 5 ml of stock solutions a, b, and c; 10 ml of stock solutions d and e. Add appropriate antibiotic if desired and pour plates.		

*Page *et al., J. Bacteriol.* 125: 1080-1087, 1975

COR-AAD1-00001900

TABLE 3

WPM 0.1 mg/L NAA, 2.5 mg/L BA[a] per liter

	100 ml
WPM-macro	10 ml
WPM-macro	10 ml
WPM-Ca	10 ml
Inositol (10 mg/ml)	10 ml
Chelated Iron[b]	1 ml
WPM Vitamin	20 ml
Sucrose	2.5 ml
NAA (0.1 mg/ml)[c]	
BA (0.1 mg/ml)[d]	

Bring volume to one liter with distilled H_2O, pH to 5.8, add 7 g agar, and autoclave.

WPM-macro	g/L	WPM-micro	g/L
NH_4NO_3	4.0	H_3BO_3	0.67
K_2SO_4	9.9	$ZnSO_4$-$7H_2O$	0.86
KH_2PO_4	1.7	$MnSO_4$-H_2O	1.69
$MgSO_4$-$7H_2O$	3.7	Na_2MoO_4-$2H_2O$	0.025
		$CuSO_4$-$5H_2O$	0.025

WPM-Ca	g/100 ml	WPM-Vitamin	g100 ml
$Ca(NO_3)_2$-$4H_2O$	5.56	Thiamine HCl	0.1
$CaCl_2$-H_2O	0.96	Nicotinic acid	0.05
		Pyridoxine HCl	0.05
		Glycine	0.2

a. Lloyd *et al., Comb. Proc. Inter. Plant. Prop. Soc.* 30: 421-427, 1980.

b. Chelated Iron = Na_2 EDTA 3.73 g/L; $FeSO_4$ - $7H_2O$ 2.73 g/L.

c. NAA = Napthaleneacetic acid.

d. BA = Benzylamino purine.

b. Molecular Biology Methods.

Plasmids were isolated by harvesting the bacterial growth of 10 TNA plates containing the appropriate antibiotic by suspending the

growth from each plate in 5 ml of TE buffer (50 mM Tris-HC1, 20 mM EDTA, pH 8.0), pooling the solutions in a 250 ml centrifuge bottle and pelleting the cells by centrifugation at 10,000 x g for five minutes. The pellet was resuspended in 20 ml of lysis buffer (50 mM Tris-HC1, pH 8.0, 20 mM EDTA, 50 mM glucose, 2 mg/ml lysozyme), and incubated at room temperature for five minutes. Freshly prepared (40 ml) alkaline-SDS solution (0.2M NaOH, 1% SDS) was added. The cells were lysed by gentle inversion and incubated in an ice water bath for 10 minutes. Potassium acetate (30 ml of a 5 M solution) was added, the solution was mixed by gentle inversion and incubated in an ice water bath for 10 minutes. This solution was centrifuged for 10 minutes at 10,000 x g, 4°C. The supernatant was decanted to a clean centrifuge bottle and the DNA was precipitated by the addition of two volumes of 95% ethanol and incubation in an ice water bath for 1 hour. The precipitate was collected by centrifugation at 10,000 x g for 30 minutes at 4°C. The resulting pellet was resuspended in 10 ml of ice cold TE buffer by slowly passing the mixture through a pipet. After the pellet was resuspended, 5 ml of 7.5 M ammonium acetate was mixed in by gentle inversion. This solution was incubated in an ice water bath for 20 minutes and then centrifuged for 10 minutes at 10,000 x g and 4°C. The supernatant was decanted to a 50 ml centrifuge tube and 0.313 volumes of 42% polyethylene glycol (MW 6000-8000) was mixed in by gentle inversion. The DNA was allowed to precipitate from 4 hours to overnight at 4°C. The DNA was collected by centrifugation at 10,000 x g for 10 minutes at 4°C. The pellet was resuspended in 8 ml of ice cold TE buffer and then added to 8 g of cesium chloride. After the cesium chloride was in solution, 0.6 ml of a 10 mg/ml solution (in distilled water) of ethidium bromide was added. The solution was centrifuged in an ultracentrifuge at 40,000 rpm for 42 hours at 20°C, using a Ti50 rotor. The plasmid band from this cesium chloride-ethidium bromide gradient was drawn off using a pasteur pipet. The ethidium bromide was removed by several extractions with water saturated n-butanol and then dialyzed for 24 hours, with two

COR-AAD1-00001902

buffer changes, in a TE buffer solution. The purified DNA was stored at -20°C.

Routine analysis of strains for the desired plasmid was done by mini-prep. A loop of culture taken from a TNA antibiotic plate was suspended in 100 μl of lysis buffer by vortexing. After 5 minutes of incubation at room temperature, 200 μl of alkaline-SDS solution was mixed in by gentle inversion and the lysed cells were incubated in an ice water bath for 10 minutes. Potassium acetate (150 μl of a 5M solution) was mixed in by gentle inversion and incubation in the ice water bath was continued for 5 minutes. The lysate was cleared by microfugation at 4°C for 5 minutes, and the supernatant was decanted to a fresh tube. The DNA was precipitated by adding 1 ml of 95% ethanol and incubating the mixture at -70°C for 30 minutes, followed by microfugation for 30 minutes. The pellet was resuspended in 100 μl of ice cold, sterile distilled water, and 50 μl of 7.5 M ammonium acetate was mixed in by gentle inversion of the tube. This mixture was incubated in an ice water bath for 10 minutes, microfuged for 2 minutes, and the supernatant was decanted to a fresh tube. The DNA was precipitated by addition of 300 μl of 95% ethanol, incubation at -70°C for 30 minutes, and microfugation for 30 minutes. The DNA pellet was dried by vacuum desiccation for 10 minutes, and resuspended in 40 μl of TE buffer. Analysis was done by agarose gel electrophoresis as described below.

Restriction endonuclease digestion was done by incubating the purified plasmid DNA in the appropriate Boehringer Mannheim buffer with 1-2 μl of the required Boehringer Mannheim restriction endonuclease, at 37°C for 1 hour. The reaction was inactivated by incubation at 70°C for 10 minutes, followed by incubation in an ice water bath for 10 minutes.

DNA ligation was performed by mixing the two restriction endonuclease digested DNA fragments to be ligated, adding 1/10 volume 7.5 M ammonium acetate, and two volumes of 95% ethanol.

COR-AAD1-00001903

The DNA in this solution was precipitated by incubation at -70°C for 30 minutes and then centrifugation for 30 minutes at 4°C in the microfuge. The pellet was resuspended in 100 μl of ice cold sterile distilled water by vortexing for 15 seconds. The resuspended DNA was reprecipitated by the ammonium acetate-ethanol method described above. After the second precipitation, the DNA pellet was dried by vacuum desiccation for 10 minutes, resuspended in 16 μl of ice cold sterile distilled water and 4 μl of SX Gibco-BRL ligase buffer was added. Gibco-BRL T4 ligase was added to 1 Wiess unit. The ligation mixture was incubated at room temperature for 2 hours, and stopped by addition of 30 μl of ice cold, sterile distilled water.

Analysis of plasmids and DNA fragments was done by agarose gel electrophoresis. The gel is made by adding agarose to a final concentration of 0.7% in TAE buffer (40 mM Tris-acetate, 0.1 mM EDTA). A 15 cm^2 gel had a total volume of 100 ml and a mini-gel had a total volume of 25 ml. The agarose buffer solution was melted in the microwave, tempered to 50°C and then poured into the gel mold and allowed to solidify for 20 minutes. The cast gel was then placed in the gel box, and submerged in TAE buffer. The DNA was loaded into the wells, and electrophoresed for 2.5 hours at 100 volts when 15 cm^2 gels were run, and 45 minutes at 130 volts when mini-gels were run. The DNA was visualized by staining the gel in 300 ml of water containing 40 μl of 10 mg/ml ethidium bromide solution for 20 minutes, and then exposing the gel to UV light at 305 nm. The gel was photographed using a Fisher Brand photodocumentation system and Polaroid 660 film.

Low melting temperature agarose gels were run as described above, except the amount of low melting agarose was 1% and the gels were run at 4°C. The DNA was visualized by ethidium bromide staining, and the desired fragment was cut out of the gel. The excised fragment was eluted from the gel matrix by adding 100 μl of TE buffer and incubating at 70°C for 10 minutes. An equal volume of TE

saturated phenol was added and mixed in by gentle inversion. The phases were separated by microfuging the sample for 3 minutes at 4°C. The top (aqueous) layer was collected, and the phenol layer was extracted twice more with an equal volume of TE buffer. The aqueous phases were pooled and extracted once with a 1:1 mixture of phenol:chloroform and once with chloroform. The DNA was precipitated by adding 1/10 volume of 7.5 M ammonium acetate, 2 volumes of 95% ethanol, incubation at -70°C for 30 minutes, and microfugation for 30 minutes. The pellet was vacuum dried for 10 minutes and then resuspended in 100 µl of TE buffer.

c. Transformation and Conjugation Methods.

Transformation of *P. aeruginosa* PAO1c was as described by Mercer *et al.* (Mercer, A.A. and Loutit, J.S., *J. Bacteriol.* 140: 37-42, 1979). *E. coli* S17-1 was transformed as described by Maniatis *et al.* (Maniatis, T. *et al.,* Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Transfer of plasmid pUCW200 from *E. coli* S17-1 to P. *putida* PP0300 or A. *tumefaciens* LBA4404 by conjugation was done by growing the strains at 30°C in LB media or LB containing 50 µg/ml kanamycin *(E. coli* S17-1) and then mixing equal volumes of each culture, filtering the mixture through a sterile 0.22 µm filter and placing the filter on an LB plate. The plates were incubated overnight at 30°C, followed by washing the filter with 5 ml of sterile distilled water and plating dilutions of the cell suspension on Burk's salts containing 0.2% succinate and 50 µg/ml kanamycin. These plates were incubated at 30°C for 48 hours and the transconjugates were purified by restreaking to identical media.

COR-AAD1-00001905

d. HPLC Analysis for 2,4-D and 2,4-Dichlorophenol.

HPLC analysis was done using a Supelco C8 column, a mobile phase of 70:30 methanol:water at 1 ml/min., and detection at 280 nm of a 20 μl injection. Peaks were identified by comparison of retention times to those of known standards.

Example 1. Construction of Plasmid pUCW101.

To construct pUCW101 (Figure 2), plasmid pRO101, which encodes all the enzymes for the degradation of 2,4-D to chloromaleylacetic acid, was digested with restriction endonucleases *Bam*HI and *Hind*III. The DNA fragment containing the *tfd*A gene was isolated from a low melting temperature agarose gel and ligated into vector plasmid pRO1727 which had been digested with the same restriction endonucleases. The ligated DNA was transformed into *P. aeruginosa* PAO1c and transformants which contained the desired insert were selected by plating for growth on TNA Cb^{500}, followed by replica plating to DNA Tc^{50} and TNA Cb^{500}. Strains with the correct phenotype of Tc sensitivity (due to insertional inactivation) and Cb resistance were characterized further by isolation of the plasmid DNA and digestion with *Bam*HI and *Hind*III. The digested plasmid DNA was analyzed by agarose gel electrophoresis to confirm that the desired fragment had been cloned. This plasmid was designated pUCW101, Figure 2.

Expression of *tfd*A on plasmid pUCW101 in *P. aeruginosa* PAO1c was confirmed by growth of this strain in the presence of 2,4-D and detection of 2,4-dichlorophenol by HPLC.

COR-AAD1-00001906

Example 2. Construction of Plasmid pUCW200.

The *A. tumefaciens* binary vector pBI121 (Clontech Laboratories, Inc., Palo Alto, CA.; shown in Figure 3) was used to transfer and stably express *tfd*A in sweetgum. pBI121 contains the transcriptional and translational start sequences of the cauliflower mosaic virus 35S promotor, and the transcriptional termination and polyadenylation sites and the translational stop codons from the nopaline synthase (NOS) gene. pBI121 also contains left and right T-DNA borders (LB and RB) which are sequences used for transformation of a plant cell (*See* Zambryski *et al., Cell* 56: 193-201, 1989, and Zambryski *et al., Annu. Rev. Plant Physiol. & Plant Mol. Biol.* 43: 465-490, 1992). DNA between these boundaries will be inserted and then replicated with the plant's chromosomal DNA.

Subcloning *tfd*A into plasmid pBI121 creating plasmid pUCW200 was accomplished as diagramed in Figure 3. Plasmid pUCW101 was digested with restriction endonucleases *Xba*I and *Sac*I. The DNA fragment containing the *tfd*A gene was isolated from a low melting temperature agarose gel and ligated into plasmid pBI121 which had been cut with the same enzymes. This mixture was transformed into *E. coli* S17-1, and transformants were selected for growth on LB Km^{50} at 37°C. Transformants were picked, grown overnight on identical media and analyzed for inserts by mini-prep analysis. One strain which seemed to contain the proper insert was further characterized by purifying the plasmid DNA as described above and confirming the insert by digestion with *Xba*I and *Sac*I. This plasmid was designated pUCW200, Figure 3. It contains a *tfd*A gene in the Agrobacterium binary vector pBI121.

Expression of *tfd*A on plasmid pUCW200 was tested by first transferring the plasmid from *E. coli* S17-1 into *P. putida* PPO300 by conjugation, and then growing *P. putida* (pUCW200) in the presence of 2,4-D and detecting 2,4-dichlorophenol by HPLC.

COR-AAD1-00001907

Plasmid pUCW200 was mobilized from *E. coli* S17-1 to *A. tumefaciens* LBA4404 as described above. The presence of plasmid pUCW200 in *A. tumefaciens* LBA4404 was confirmed by mini-prep analysis. Expression of *tfd*A in this strain was confirmed by growing *A. tumefaciens* LBA4404 (pUCW200) in the presence of 2,4-D and demonstrating the accumulation of 2,4-dichlorophenol in the media by HPLC.

Example 3. Transformation of Sweetgum with pUCW200.

Agrobacterium tumefaciens-mediated transformation was used to transform sweetgum. *Agrobacterium tumefaciens* LB4404 has the ability to transfer plasmid pUCW200 into a plant cell. Once inside the plant cell, DNA between the right (RB) and left (LB) borders integrates (randomly) into the plant's chromosome. It is then replicated as if it were a part of the genome of the plant, and thus, when this cell divides and differentiates into a shoot, all the cells of this adventitious shoot contain the gene transformed into the original target cell.

The sweetgum transformation method used is outlined below:

1. The expanding leaves from a known sweetgum clone were surface sterilized by first rinsing them with soapy water and then stirring them in 10% bleach solution (in sterile water) for 10 minutes, followed by three rinses (for 2 minutes each) with sterile distilled water.

2. Each leaf was aseptically cut into small (5 to 10 mm) pieces. Some of the pieces were placed on WPM 0.1/2.5 (Table 2). These pieces acted as the regeneration control.

3. *Agrobacterium tumefaciens* LB4404, containing plasmid pUCW200, was grown overnight at 30°C in 50 ml of LB containing 50 mg/L kanamycin. The next morning this culture was inoculated into 500 ml of the identical media, and the strain was grown for 4 hours at 30°C. The cells were harvested by centrifugation (5 min at 10,000 x g), washed once with LB, and finally, resuspended in 100 ml of fresh LB.

COR-AAD1-00001908

4. The leaf pieces were co-cultivated with *Agrobacterium* for 30 minutes at room temperature. They were then blotted dry on sterile Whatman No. 3 filter paper and placed on WPM 0.1/2.5. The plates were sealed with parafilm and incubated in the growth chamber.

5. After three days, the leaf pieces were transferred to WPM 0.1/2.5 which contained 500 mg/L carbenicillin (Cb^{500}). Carbenicillin, an antibiotic, was used to kill the residual *Agrobacterium.*

6. After two weeks, the non-control leaf pieces were transferred to selective media. Half of the leaf pieces were placed on WPM 0.1/2.5 Cb^{500} with 0.1 mg/L 2 4-D and the other half were placed on the normal WPM 0.1/2.5 Cb^{500}, Kanamycin 75 mg/L (Km^{75}). The results of this selection are shown in Table 5. Plasmid pUCW200 contains the kanamycin resistance gene, NPT-II (Figure 3). Sweetgum is sensitive to kanamycin at 75 mg/L, and will not regenerate in its presence. Therefore, adventitious shoots formed in the presence of kanamycin may contain the resistance gene.

7. Shoots, regenerated under selective pressure, were excised from the leaf piece and transferred to WPM 0.01/2.0 Cb^{500} Km^{75}.

TABLE 5
EFFECT OF 2,4-D SELECTION ON SWEETGUM TRANSFORMATION FREQUENCY

Clone	Selection Media	# shoots/#leaf pieces	Ratio (shoots/leaf piece)
2027 (control)	Cb, control	73/18	4.06
2040 (pUCW200)	Cb, 2,4-D	38/72	0.53
2027 (pUCW200)	Cb, Km	3/85	0.04
2040 (control)	Cb, control	85/30	2.83
2040 (pUCW200)	Cb, 2,4-D	77/66	1.17
2040 (pUCW200)	Cb, Km	16/49	0.33

Abbreviations: Cb = carbenicillin. Km = kanamycin.

COR-AAD1-00001909

Example 4. ELISA Analysis of Transformed Sweetgum Clones.

An analytical method used to confirm the transfer of selected genes into sweetgum clones is an enzyme linked immunosorbant assay (ELISA) for the detection of the NPTII protein encoded by the kanamycin resistance gene on plasmid pUCW200. (NPTT II ELISA Kit, Prime Report 3(2): 3, 1991. Plant tissue (from 100 to 800 mg fresh weight) is placed in 3 ml of extraction buffer (0.25 M Tris-HCl, pH 7.8, 0.1 mM phenylmethylsulfonyl fluoride) and homogenized using a Tekmar Tissuizer, model TR-10 (equipped with a microprobe), for the two pulses of 30 seconds each. An additional 2 ml of extraction buffer is added and the cellular debris is removed by ultracentrifugation at 50,000 RPM for 20 minutes at 4°C. The supernatant is collected and 4 volumes of ice cold acetone are added, followed by incubation at -20°C for four hours, or up to overnight. The precipitated proteins are collected by centrifugation at 4°C, 10,000 RPM for 20 minutes. The pellet is resuspended in 1 ml of extract buffer. This sample is used in the NPT-II ELISA kit purchased from 5 Prime - 3 Prime, Inc., Boulder, CO.

The ELISA method involves using an antibody specific for the kanamycin resistance protein (neomycin phosphotransferase, NPT-II) to detect this protein in the cytoplasmic fraction of putative sweetgum transformants. The presence of this protein is an indication of transformation because the gene encoding it is located on plasmid pUCW200. This plasmid is transferred into the target plant cell by *A. tumefaciens* LBA4404 where it integrates into the plant's genome and expresses its genes. Since the genes on this plasmid are physically linked, the presence of one of the gene products is evidence of the presence of the other genes located on the plasmid. The presence of the NPT-II protein in plant extracts transformed with plasmid pUCW200, the 2,4-D resistance plasmid, which also contains the kanamycin resistance gene, is positive evidence for the presence of the 2,4-D resistance gene.

Example 5. ELISA Results of Sweetgum/pUCW200 Transformation.

The transformation frequencies of pUCW200 isolates selected on 2,4-D were compared to those selected on kanamycin. ELISA for NPT-II was used as a measure of transformation frequency. Four isolates selected on 2,4-D and four isolates selected on kanamycin were assayed. The results are shown in Figure 4. All four of the 2,4-D selected clones (designated 2027 (pUCW200)-TA, -TC, -TD and -TE) were positive for NPT-II, indicating that they are transformed. However, none of the kanamycin selected clones (designated 2027 (pUCW200)-KA, -KB, -KC, and -KD) were positive.

The fact that 2,4-D selection did not yield any false positive isolates demonstrates that it is an outstanding selectable marker. Also, experiments have shown that sweetgum leaf pieces placed on 2,4-D containing media do not regenerate adventitious shoots. Since co-cultivated leaf pieces did regenerate adventitious shoots in the presence of 2,4-D and 100% of those shoots were transformed, these shoots must be using the 2,4-D resistance gene product to convert the 2,4-D in the media to 2,4-dichlorophenol. This indicates that these isolates are expressing the resistance gene.

COR-AAD1-00001911

Example 6. Description of Propagation Method.

Sweetgum were propagated by growing the adventitious shoots on WPM containing 0.01 mg/L napthaleneacetic acid and 2.0 mg/L benzylamino purine. The original shoot formed new shoots on this media. The new shoots were aseptically excised from the original shoot and grown independently. This propagation was repeated until the required number of shoots were generated. These shoots were then elongated by incubating them on WPM containing 0.5 mg/L benzylamino purine. Elongated shoots of greater than 1 cm in length were then incubated on root induction media consisting of 1/3 strength WPM containing 0.1 mg/L IBA (indole-3-butyric acid). When roots began to form, the plantlets were transferred to plug trays containing a soilless mix consisting of equal parts peat, perlite, and vermiculite and incubated in 100% relative humidity until new growth appeared. The plantlets were then transplanted to 10 cubic inch leach tubes and grown in a greenhouse until they reach planting size. The plants are then hardened off to outdoor conditions until dormant, removed from tubes, and planted at the prepared site.

Example 7. A 5 X 10 Plantation of Sweetgum Trees

A sweetgum plantation was established at the Union Camp Nursery, Belleville, Georgia with ten trees each of four clonal propagations (as described in Example 6) of different *tfdA* transformants of the same parent tree line, with the different transformants labelled "TA", "TC", "TD" and "TE", as described in Example 5. A further ten trees were propagated from an adventitious shoot derived from cells of the parent tree line. The trees of each category were grown in a single row. As described in Example 6, for their first growth season, the shoots were initially grown in the greenhouse and transferred to outdoor conditions. In this experiment, the ground was not treated with 2,4-dichlorophenoxyacetic acid. As a

COR-AAD1-00001912

consequence of the tissue culture propagation method, as would be known to one skilled in the art, the trees initially varied in size. After the second growth season, the trees of each category showed the average heights illustrated in Figure 5. The p-values for the average height increases shown by the TA and TD clones compared to the controls were in excess of 0.9. The p-value for the average height increase shown by the TE clones compared to the controls was in excess of 0.95. The TC clones showed no average height increase and were believed to have stopped expressing or to poorly express the *tfdA* gene. Figure 6 shows the average diameters at breast height ("DBH") for the five clones. The p-values for the average HBD increases shown by the TA and TD clones compared to the controls were in excess of 0.9. The p-value for the average HBD increase shown by the TE clones compared to the controls was in excess of 0.95. The TC clones showed no average HBD increase.

* * * * *

While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention and appended claims.

COR-AAD1-00001913

What Is Claimed Is:

1. A method for obtaining a growth-enhanced transgenic perennial plant, the method comprising cultivating a transgenic plant precursor, wherein cells in the precursor or in the plant express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

2. The method of claim 1, wherein the growth enhancement of the transgenic plant is measurable after more than one growing season.

3. The method of claim 1, wherein the transgenic plant precursor is a precursor of a woody perennial.

4. The method of claim 3, wherein the woody perennial is of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome*, *Broussonetia*, *Camellia*, *Camptobera*, *Carya*, *Castanea* (Chestnut), *Catalpa*, *Cinchona*, *Citrus*, *Coffea*, *Corylus* (Hazelnut), *Diospyrus*, *Eucalyptus*, *Fagus* (Beech), *Ficus*, *Fraxinus* (Ash), *Gingko*, *Gleditsia*, *Gmelina*, *Hamamelia*, *Hedera*, *Ilex*, *Juglans*, *Kalmia*, *Lexceana*, *Liquidambar* (Sweetgum), *Liriodendron*, *Malus*, *Moghania*, *Morus*, *Olea*, *Paulownia*, *Populus*, *Prunus*, *Quercus* (Oak), *Rosa*, *Rhododendron*, *Robinia*, *Salix*, *Santalum*, *Sapium*, *Simmondsia*, *Sycamore*, *Tectona*, *Theobrama*, *Tupidanthus*, *Ulmus* (Elm), *Vaccinium*, or *Vitus*, or of the order Coniferales.

5. The method of claim 4, wherein the woody perennial is sweetgum.

6. The method of claim 1, wherein the gene is the *tfdA* gene.

COR-AAD1-00001914

7. The method of claim 6, wherein the growth enhancement of the transgenic plant is measurable after more than one growing season.

8. The method of claim 6, wherein the transgenic plant precursor is a precursor of a woody perennial.

9. The method of claim 8, wherein the woody perennial is of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome, Broussonetia, Camellia, Camptobera, Carya, Castanea* (Chestnut), *Catalpa, Cinchona, Citrus, Coffea, Corylus* (Hazelnut), *Diospyrus, Eucalyptus, Fagus* (Beech), *Ficus, Fraxinus* (Ash), *Gingko, Gleditsia, Gmelina, Hamamelia, Hedera, Ilex, Juglans, Kalmia, Lexceana, Liquidambar* (Sweetgum), *Liriodendron, Malus, Moghania, Morus, Olea, Paulownia, Populus, Prunus, Quercus* (Oak), *Rosa, Rhododendron, Robinia, Salix, Santalum, Sapium, Simmondsia, Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus* (Elm), *Vaccinium*, or *Vitus*, or of the order Coniferales.

10. The method of claim 9, wherein the woody perennial is sweetgum.

11. A method for obtaining a yield-enhanced plantation, the method comprising cultivating a plot of growth-enhanced transgenic perennial plants, wherein cells in the transgenic plants express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

12. The method of claim 11, wherein the transgenic perennial plants are woody perennials.

COR-AAD1-00001915

13. The method of claim 12, wherein the woody perennials are of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome*, *Broussonetia*, *Camellia*, *Camptobera*, *Carya*, *Castanea* (Chestnut), *Catalpa*, *Cinchona*, *Citrus*, *Coffea*, *Corylus* (Hazelnut), *Diospyrus*, *Eucalyptus*, *Fagus* (Beech), *Ficus*, *Fraxinus* (Ash), *Gingko*, *Gleditsia*, *Gmelina*, *Hamamelia*, *Hedera*, *Ilex*, *Juglans*, *Kalmia*, *Lexceana*, *Liquidambar* (Sweetgum), *Liriodendron*, *Malus*, *Moghania*, *Morus*, *Olea*, *Paulownia*, *Populus*, *Prunus*, *Quercus* (Oak), *Rosa*, *Rhododendron*, *Robinia*, *Salix*, *Santalum*, *Sapium*, *Simmondsia*, *Sycamore*, *Tectona*, *Theobrama*, *Tupidanthus*, *Ulmus* (Elm), *Vaccinium*, or *Vitus*, or of the order Coniferales.

14. The method of claim 13, wherein the woody perennials are sweetgum.

15. The method of claim 11, wherein the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

16. The method of claim 15, wherein the transgenic perennial plants are woody perennials.

17. The method of claim 16, wherein the woody perennials are of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome*, *Broussonetia*, *Camellia*, *Camptobera*, *Carya*, *Castanea* (Chestnut), *Catalpa*, *Cinchona*, *Citrus*, *Coffea*, *Corylus* (Hazelnut), *Diospyrus*, *Eucalyptus*, *Fagus* (Beech), *Ficus*, *Fraxinus* (Ash), *Gingko*, *Gleditsia*, *Gmelina*, *Hamamelia*, *Hedera*, *Ilex*, *Juglans*, *Kalmia*, *Lexceana*, *Liquidambar* (Sweetgum), *Liriodendron*, *Malus*, *Moghania*, *Morus*, *Olea*, *Paulownia*, *Populus*, *Prunus*, *Quercus* (Oak), *Rosa*, *Rhododendron*, *Robinia*, *Salix*, *Santalum*, *Sapium*, *Simmondsia*,

COR-AAD1-00001916

Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus (Elm), *Vaccinium*, or *Vitus,* or of the order Coniferales.

18. The method of claim 17, wherein the transgenic perennial plants are sweetgum.

19. A growth-enhanced transgenic perennial plant or a growth-enhanced transgenic perennial plant precursor comprising a transgenic plant or plant precursor, wherein cells in the plant or precursor express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

20. The transgenic plant or precursor of claim 19, wherein the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

21. The transgenic plant of claim 19, wherein the plant or precursor is a woody perennial.

22. The growth-enhanced transgenic plant or precursor of claim 21, wherein the woody perennial is of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome, Broussonetia, Camellia, Camptobera, Carya, Castanea* (Chestnut), *Catalpa, Cinchona, Citrus, Coffea, Corylus* (Hazelnut), *Diospyrus, Eucalyptus, Fagus* (Beech), *Ficus, Fraxinus* (Ash), *Gingko, Gleditsia, Gmelina, Hamamelia, Hedera, Ilex, Juglans, Kalmia, Lexceana, Liquidambar* (Sweetgum), *Liriodendron, Malus, Moghania, Morus, Olea, Paulownia, Populus, Prunus, Quercus* (Oak), *Rosa, Rhododendron, Robinia, Salix, Santalum, Sapium, Simmondsia, Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus* (Elm), *Vaccinium*, or *Vitus,* or of the order Coniferales.

COR-AAD1-00001917

23. The plant or precursor of claim 22, wherein the plant or precursor is sweetgum.

24. A yield-enhanced plantation, comprising a plot of growth-enhanced transgenic perennial plants, wherein cells in the transgenic plants express a 2,4-dichlorophenoxyacetic acid monooxygenase gene at a growth-enhancing effective level.

25. The yield-enhanced plantation of claim 24, wherein the 2,4-dichlorophenoxyacetic acid monooxygenase gene is the *tfdA* gene.

26. The yield-enhanced plantation of claim 24, wherein the plants are woody perennials.

27. The yield-enhanced plantation of claim 26, wherein the woody perennials are of the genera *Acacia*, *Acer* (Maple), *Actinidia*, *Albizzia*, *Alnus*, *Amelanchier*, *Atriplex*, *Betula* (Birch), *Brachycome, Broussonetia, Camellia, Camptobera, Carya, Castanea* (Chestnut), *Catalpa, Cinchona, Citrus, Coffea, Corylus* (Hazelnut), *Diospyrus, Eucalyptus, Fagus* (Beech), *Ficus, Fraxinus* (Ash), *Gingko, Gleditsia, Gmelina, Hamamelia, Hedera, Ilex, Juglans, Kalmia, Lexceana, Liquidambar* (Sweetgum), *Liriodendron, Malus, Moghania, Morus, Olea, Paulownia, Populus, Prunus, Quercus* (Oak), *Rosa, Rhododendron, Robinia, Salix, Santalum, Sapium, Simmondsia, Sycamore, Tectona, Theobrama, Tupidanthus, Ulmus* (Elm), *Vaccinium*, or *Vitus*, or of the order Coniferales.

28. The plantation of claim 27, wherein the plants are sweetgum.

COR-AAD1-00001918

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/03670

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/00, 15/09, 15/29, 15/82; A01H 1/00, 3/00, 4/00, 5/00
US CL : 536/23.6; 435/172.3, 252.2, 320.1; 800/205 DIG 47, DIG 74
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.6; 435/172.3, 252.2, 320.1; 800/205, DIG 47, Dig 74

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, STN, Biosis, Agricola, Wpids

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LYON et al. Expression of a Bacterial Gene in Transgenic Tobacco Plants Confers resistance to the Herbicide 2,4-Dichlorophenoxyacetic acid. Plant Molecular Biology. 1989, Vol. 13, pages 533-540. See entire reference.	1-28
Y	PARSONS et al. Transformation of Poplar by Agrobacterium tumefaciens. Bio/Technology. June 1986, Vol. 4, pages 533-536. See entire reference.	1-28

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"B"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 06 APRIL 1998	Date of mailing of the international search report 24 JUN 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Ousama M-Faiz Zaghmout Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68 | A1

(11) International Publication Number: **WO 98/38336**

(43) International Publication Date: 3 September 1998 (03.09.98)

(21) International Application Number: PCT/US98/03684

(22) International Filing Date: 26 February 1998 (26.02.98)

(30) Priority Data:
60/039,582 28 February 1997 (28.02.97) US

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ELSEMORE, David, Allen [US/US]; 1223 Town Court North, Lawrenceville, NJ 08648 (US). LAROSSA, Robert, Alan [US/US]; 672 Kadar Drive, West Chester, PA 19382 (US). SMULSKI, Dana, Robin [US/US]; 603 Greenbank Road, Wilmington, DE 19808 (US). VAN DYK, Tina, Kangas [-/US]; 119 Cambridge Drive, Wilmington, DE 19803 (US).

(74) Agents: FLOYD, Linda, A. et al.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, GW, HU, ID, IL, IS, JP, KG, KP, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, US, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), 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
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.
Date of receipt by the International Bureau:
26 February 1998 (26.02.1998)

(54) Title: A METHOD FOR IDENTIFYING THE SITE OF ACTION OF XENOBIOTIC CHEMICALS

(57) Abstract

A method is disclosed for the discovery of the site of action of xenobiotic compounds. The method involves subjecting a detector cell comprising a genotoxic–sensitive promoter operably linked to a luminescent reporter gene complex. Exposure of the detector bacteria to compounds that are genotoxic results in an increase in luminescence. Such compounds are grouped and may be analyzed for pharmaceutical activity. Exposure of the detector bacteria to active xenobiotics that are not genotoxic results in inhibition of bioluminescence or stable luminescence. Subjection of the detector bacteria to a screen such as nutritional reversal or genetic titration reveals the site of action of the compound.

COR-AAD1-00001920

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE

A METHOD FOR IDENTIFYING THE SITE OF ACTION OF XENOBIOTIC CHEMICALS

FIELD OF INVENTION

The invention relates to field of molecular biology and to methods for screening compounds for biological activity. More specifically, a method has been developed to rapidly identify the site of action of various xenobiotic compounds in particular, xenobiotic agrochemicals and antimicrobials.

BACKGROUND OF THE INVENTION

Technological advances within the chemical arts have made possible the synthesis of vast arrays of chemical compounds of interest to the agrochemical, pharmaceutical and environmental industries. These methods of synthesis are capable of producing far more compounds than can be reasonably screened to identify their utility. The utility of these compounds is modeled on various structure-function relationships and is often confirmed through screening methods that associate these compounds with desired, known activities. Once a compound or class of compounds has been subjected to a screen and a putative activity has been determined, further analysis is needed to determine the mechanism also known as site of action, i.e., the target, of the compound. The effort to understand the precise site of action of a chemical or environmental agent is driven by a concern for environmental safety and a desire to capture the value associated with this understanding for use in the agrochemical and pharmaceutical industries. A need exits, therefore, for methods of determining the site of action of synthetic xenobiotic compounds that are rapid, accurate and able to accommodate the vast numbers of compounds being produced.

Methods for determining the site of action of various compounds have been developed using the complementary methods of nutritional reversal and genetic titration in microbially based systems. LaRossa et al., [*J Biol. Chem.*, 259 (14), (1984) 8753-8757] teach a pool strategy for nutritional reversal previously used to classify the pathway effected by an auxotrophic mutation, termed auxanography, [Davis, R. W., D. Botstein and J. R. Roth: A Manual For Genetic Engineering: Advanced Bacterial Genetics. X+254p. Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., U.S.A., 1980]. This method precisely defines the site of sulfonylurea herbicide action in *Salmonella typhimurium* LT2. The method uses lawns of bacterial cells grown on solid media and exposed to effective concentrations of herbicide. Sulfonylurea herbicides inhibit *Salmonella* growth by inhibiting the branched chain amino acid biosynthetic pathway. Nutrient supplementation overcome specific amino acid deficiencies imposed by herbicidal action. Results are seen in about 48 hr. The method of LaRossa et al.

COR-AAD1-00001922

is useful for determining a site of action for a specific active xenobiotic compound but, like whole plant assays, is time consuming and requires large amounts of test compound to be screened.

Genetic titration in a microbial population offers an alternative to nutritional reversal. Genetic titration has been used during the past 40 years. Selection of regulatory mutations that increase the titer of pathway enzymes in response to a challenge with an antimetabollite are well known in the *E. coli* and *S. typhimurium* literature. Gene duplications have been selected in both mammalian cell lines (using drugs and transition state analogs) and with antimetabolites in bacteria. These selections are based upon the increased enzyme content of the cell titering out the deleterious effect of the inhibitor. Systematic selection for such titrations using antimetabolites in yeast in conjunction with multicopy recombinant DNA based yeast chromosomal libraries were pioneered by Falco [Falco et al., *Genetics* 109:21-35 (1985)] and Rine [Rine et al., *Proc Natl Acad Sci U S A* 80: 6750-4 (1983)]. More recently, Sternberg and Chattajee PNAS 92:8950-8954 (1995)] have used this recombinant DNA based methodology in *E. coli*. This art teaches the use of the method of genetic titration in yeast to identify various macromolecular targets of a crop protection chemical and a few pharmaceuticals. It also teaches that recombinant DNA mediated genetic titration may identify a bacterial target.

Genetic titration proceeds generally by first preparing DNA libraries of a bacterial genome in multicopy cloning vectors where each portion of the genome is highly amplified. Bacterial hosts are transformed with the library and screened for growth on plates containing effective inhibitory concentrations of the xenobiotic compound to be tested. Colonies are picked and the plasmid isolated and sequenced. The sequence is compared to known sequences to identify genes that could encode a site of action for the compound.

These methods of nutritional reversal and genetic titration are both limited by the need to place cells of bacteria or yeast on agar plates and subsequently score or monitor for cell growth in response to the presence of the compound. This limitation makes the methods slow and cumbersome. Further, neither of these methods offers a way to determine the relative physiological or environmental toxicity of the compounds being screened. One of the significant disadvantages of any current screening method is that many active compounds may be identified in initial screens that must be later discarded due to poor results in toxicological screens. Incorporation of a toxicological pre-screen into the site of action screen would minimize this costly investment of time and resources.

Bioluminescent reporters are known as toxicological detectors in the art. One of the most common are genes encoding the firefly luciferase. Another is a

COR-AAD1-00001923

set of five genes, *luxCDABE*, that has been isolated from the bioluminescent bacteria *Vibrio fischeri*. Both eukaryotic and prokaryotic genes have been used in recombinant systems to serve as detectors.

cDNA encoding firefly luciferase has been expressed in *E. coli* under the control of the *lacZ* promoter [Tatsumi et al., *Biochem. Biophys Acta.*, 1131, (2), pp 161-165, (1992)], and the *luxAB* fusion gene has been expressed in *Bacillus* at levels comparable to those achievable in *E. coli* by placing it under the control of the powerful *Pxyn* promoter (Jacobs et al., *Mol. Gen. Genet.*, 230(1-2), pp 251-256, (1991)].

Bacterial genes have also been isolated and expressed. Rychert et al., [*Environ. Toxic. Water Qual.*, 6 (4), pp 415-422, (1991)] have shown that recombinant *E. coli* harboring the plasmid, pJE202, that contains the promoterless *Vibrio fischeri lux* gene complex, was sensitive to Zn^{2+}, ethidium bromide, sodium pentachoropheate, Cu^{2+} and 2,4-dichloropheoxyacetic acid. Response in this assay was registered by a decrease in baseline light emitted by the transformed *E. coli*.

In addition to the use of the promoterless bacterial gene complex, foreign promoters have been fused to the bacterial bioluminescent structural genes for use in assays specific to various toxicants. For example, recombinant bacteria have been developed by fusing the *lux* gene complex to chemically responsive bacterial promoters and then placing such chimeras in appropriate hosts. These recombinant bacteria are sensor organisms that glow in response to specific stimuli. An example of this type of gene fusion is described by Burlage et al., [*J. Bacteriol,* 172 (9) pp 4749-4757 (1990)] where a DNA fragment from plasmid NAH7 containing a promoter for the naphthalene degradation pathway was fused to the *lux* genes of *Vibrio fischeri* and used to transform a strain of *Pseudomonas*. The resulting transformant displayed an increase in light emission in the presence of naphthalene. The induction of bioluminescence was demonstrated to coincide with naphthalene degradation by the transformed organism.

Another test system specifically responsive to mercury (Hg) is described by H. Molders (EP456667). Here, indicator bacterial strains are provided (by vector-mediated gene transfer) containing a *mer* promoter, specifically inducible by Hg ions, fused to a bacterial luciferase (*lux AB*) genes complex which is responsible for bioluminescence. The test system of Molders relies on the induction of the *mer* promoter by the presence of mercury and the subsequent increase in light emission from the recombinant bacteria.

Applicants have previously disclosed the use of various stress promoters, including those sensitive to genotoxic stress, in commonly owned USSN 08/244,376 and USSN 08/344,428. USSN 08/244,376 teaches the use of detector

COR-AAD1-00001924

organisms containing a stress promoter-bioluminescent gene fusion to detect various environmental stresses including those sensitive to protein damage (heat shock), DNA damage (genotoxic), oxidative damage, cell membrane damage, amino acid starvation, carbon starvation, and nitrogen starvation. USSN 08/344,428 demonstrates the use of similarly transformed detector cells as lyophilized reagents. Finally, in commonly owned USSN 08/735,545 Applicants teach the use of bioluminescent reporters fused to bacterial promoters isolated from restriction of genomic DNA for the identification of new, compound sensitive promoters.

The problem to be overcome therefore is to develop a rapid, facile method for the identification of the site of action of agrochemically or pharmaceutically active xenobiotic compounds that are non-toxic to animals, including humans, or the environment.

SUMMARY OF THE INVENTION

The present invention provides a method for the identification of the site of action of a stressor comprising:

(i) contacting a stressor with a detector bacteria, the detector bacteria comprising a genotoxic-sensitive promoter operably linked to a luminescent reporter gene complex to form a gene fusion that confers a bioluminescent positive phenotype upon the detector bacteria wherein exposure of the genotoxic-sensitive promoter to genotoxic compounds drives heightened expression of the luminescent reporter gene complex producing an increased bioluminescent signal;

(ii) selecting for genotoxic or non-genotoxic stressors capable of inhibiting the growth of the detector bacteria of step (i) by monitoring the growth and light output of the detector bacteria;

(iii) submitting the growth-inhibiting, genotoxic or non-genotoxic stressor selected in step (ii) to a site of action screen;

(iv) identifying the site of action in the detector bacteria; and

(vi) confirming the site of action in a target organism.

The present invention further provides detector bacteria strains comprising a *recA-LuxCDABE* gene fusion as well as non-bioluminescent parent strains possessing a multiplicity of cellular and membrane mutations.

The invention further provides methods for determining whether a compound is genotoxic and comprising the steps of:

(i) culturing a detector cell comprising a promoter regulated by a SOS bacterial regulatory circuit and a *luxCDABE* gene complex wherein the *luxCDABE* gene complex is positioned in the bacterial chromosome downstream of the SOS promoter such that when the SOS promoter is expressed, then the *luxCDABE* gene complex is also expressed;

COR-AAD1-00001925

(ii) contacting the culture with a substance to be tested and;

(iii) determining whether the substance is genotoxic by measuring the amount of luminescence in the culture.

The invention additionally provides methods of identifying a structural gene encoding a stressor target comprising:

(i) contacting a stressor with a detector bacteria, the bacteria comprising a genotoxic-sensitive promoter operably linked to a luminescent reporter gene complex to form a gene fusion that confers a bioluminescent positive phenotype upon the detector bacteria wherein exposure of the genotoxic-sensitive promoter to genotoxic compounds drives heightened expression of the luminescent reporter gene complex producing an increased bioluminescent signal;

(ii) selecting for genotoxic or non-genotoxic stressors capable of inhibiting the growth the detector bacteria of step (i) by monitoring the growth and light output of the detector bacteria;

(iii) submitting the growth-inhibiting, genotoxic or non-genotoxic stressor to a site of action screen wherein the stressor target is identified and;

(vi) isolating the structural gene encoding the stressor target.

In another embodiment, the invention provides methods for identifying compounds having glyphosate-like activity, thienylalanine-like activity and ALS-inhibitory activity.

BRIEF DESCRIPTION OF THE FIGURES, BIOLOGICAL DEPOSITS AND SEQUENCE LISTING

Figure 1 (flow chart-strains 1a and 1b) is a diagram showing the genealogies of the detector cells used in the present invention. Solid arrows indicate construct of non-bioluminescent strains. Broken arrows indicate construction of bioluminescent derivatives. The selection and screens used to isolate bacteria are indicated in capitalized and italicized typeface. Relevant genotypes are italicized. *igm* indicates improved growth on minimal medium.

Figure 2 (method chart) is a flow diagram illustrating the method of the present invention of screening a compound for genotoxicity using the instant detector cells and determining the site of action of non-genotoxic compounds by nutritional reversal of genetic titration.

Figure 3 is an illustration of the method for *in vivo* inhibitor identification of specified targets showing the screening of a compound for a known activity.

Figure 4a is a plot of RLU vs. time of strain DPD1715 containing the *ilvB*, *ilvI*, *ilvH*, *ilvG relA* and *spoT* mutations and the *recA-LuxCDABE* and being tolC+, exposed to the DNA damaging agent mitomycin C.

Figure 4b is a plot of RLU vs. time of strain DPD1730 containing the *recA-LuxCDABE* and tolC+ exposed to the DNA damaging agent mitomycin C.

COR-AAD1-00001926

Figure 5 is a kinetic plot of RLU vs. mitocycin C concentration comparing the sensitivity of tolC- and tolC+ stains to mitomycin C.

Figure 6 is a plot of RLU vs. time for the strain DPD1718 containing the *RecA-LuxCDABE* gene fusion, exposed to varying concentrations of 2,4-dichlorophenoxyacetic acid.

Figure 7 is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of sulfometuron methyl.

Figures 8(a)and (b) show plots of RLU vs. time for the strain DPD1715 containing the *ilvB, ilvI, ilvH, ilvG relA* and *spoT* mutations and the *recA-LuxCDABE* and being tolC+ exposed to varying concentrations of glyphosate.

Figure 9 is a plot of RLU vs. time for the strain DPD1730 containing the *recA-LuxCDABE* gene fusion, and the *ilvB ilvG ilvIH relA spoT tolC*+ and *spoT* mutations exposed to varying concentrations of methyl viologen.

Figure 10(a) is a plot of RLU vs. time for the strain DPD1715 containing the *ilvB, ilvI, ilvH, ilvG relA* and *spoT* mutations and the *recA-LuxCDABE* and being *tolC*+ exposed to varying concentrations of Sulfometuron methyl.

Figure 10(b) is a plot of RLU vs. time for the strain DPD1728 containing the *igm relA* and *spoT* mutations and the *recA-LuxCDABE* and being *tolC*- exposed to varying concentrations of sulfometuron methyl.

Figure 10(c) is a kinetic plot of bioluminescent measure at 80min vs. concentration of sulfometuron methyl comparing the responses of strains DPD1715 and DPD1728.

Figure 11(a) is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of 3-(2-thienyl)-L-alanine in minimal medium.

Figure 11(b) is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of 3-(2-thienyl)-L-alanine in Rich LB medium.

Figure 12(a) is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of the oxime carbamate compound OC in minimal medium.

Figure 12(b) is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of the oxime carbamate compound OC in Rich LB medium.

Figure 13(a) is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of glyphosate in minimal medium.

COR-AAD1-00001927

Figure 13(b) is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, exposed to varying concentrations of glyphosate in Rich LB medium.

Figure 14 is a plot of RLU vs. time for the strain DPD1718 containing the *recA-LuxCDABE* gene fusion, and the *relA* and *ΔilvB* mutations demonstrating reversal of 3-(2-Thienyl)-L-alanine induced inhibition of bioluminescence by cystein.

Figure 15(a) is a plot of RLU vs. time for strain DPD1718 showing light inhibition by sulfometuron methyl.

Figure 15(b) is a plot of RLU vs. time for strain DPD1718 showing restoration of light output after inhibition by sulfometuron methyl as a result of nutritional reversal by the pool comprising the three branched chain amino acids, lysine and histidine.

Figure 16 is a plot of bioluminescent response ratio vs. concentration of thienylalanie contacted with strains transformed with *0245 ygaH, phenA*, and *aroH* genes.

The transformed *E. coli* strains DPD1707, DP1675, and DPD1718 were deposited on 13 February 1997 with the American Type Culture Collection ("ATCC"), international depository located at 12301 Parklawn Drive, Rockville, MD 20852 U.S.A. under the terms of the Budapest Treaty. The strains are designated as ATCC 98328, ATCC 98329, and ATCC 98330 respectively. The designations refer to the accession number of the deposited material.

Applicants have provided a sequence listing in conformity with Rules for the Standard Representation of Nucleotide and Amino Acid Sequences in Patent Applications (Annexes I and II to the Decision of the President of the EPO, published in Supplement No. 2 to OJ EPO, 12/1992) and with 37 C.F.R. 1.821-1.825 and Appendices A and B (Requirements for Application Disclosures Containing Nucleotides and/or Amino Acid Sequences).

SEQ ID NO: 1-10 are primer sequneces.

DETAILED DESCRIPTION OF THE INVENTION

Applicants are the first to recognize that stress promoter-*Lux* gene fusions will be useful in enhancing traditional methods of screening for compound sites of action and the first to develop a detector organism capable of dual functionality in such an assay where the organism is able to denote genotoxic compounds in a first stage and identify site of action in a second assay.

Applicants have developed a method for discovering the site of action of particular xenobiotic compounds by using a bacterial detector cell having a mutation that confers sensitivity to amino acid starvation and containing a gene fusion comprising a genotoxic-sensitive promoter operably linked to a bacterial

COR-AAD1-00001928

bioluminescent reporter gene complex. Typical xenobiotics include chemicals active as herbicides, anticancer agents, antimicrobials and crop protection chemicals. The method has high utility in the agrochemical and pharmaceutical industries for identifying the site of action of compounds and for designing new compounds of similar structure and/or function.

In the method, xenobiotic compounds that are genotoxic interact with the genotoxic-sensitive promoter (driving transcription of the bioluminescent reporter gene complex) and produce an increase in light. Higher concentrations of non-genotoxic xenobiotic compounds may result in the interference of metabolic activity and a decrease in light production.

The present method rapidly determines the site of action of a particular compound via a two part screen involving a uniquely constructed detector cell. The detector cell may posses a *relA* mutation (responsible for diminishing the host cell's response to amino acid starvation) and a *lux* gene fusion comprising a genotoxic-sensitive promoter operably fused to a bacterial *lux* gene complex. In the first stage, detector cells are exposed to specific concentrations of a xenobiotic compound to be screened for and cultures are analyzed for growth inhibition or an increase in bioluminescence. Compounds producing an increase in bioluminescence are discarded as genotoxic. Compounds that result in decreased bioluminescent output relative to the mock-treated control and do not elevate cellular bioluminescence at any tested concentration are subjected to stage two. The second stage uses the detector organism in standard nutritional reversal and genetic titration screens, making use of the bioluminescent gene fusion to identify those nutrients or genes whose supplementation results in prevention of the metabolic interference associated with the xenobiotic compound.

The present method demonstrates for the first time that auxanographic reversal of chemical stressor action can be signaled by restoration of light production of a detector cell strain and that patterns of reversal by defined nutrient pools can define pathways inhibited by the stressor chemical. Advantages of the method include increased numbers of compound screenings per unit of time, increased speed of the biological response, and ease of automation of data collection and processing, while decreasing by a factor of approximately 300 times the amount of compound required for analysis.

The following definitions are used herein and should be referred to for claim interpretation.

The terms "crop protection chemical" and "CPC" refer to compounds having toxic or repellent effect on insects, plant pathogens, or crop-competing plant species known to damage crop plants. CPC's include pesticides (paraquat (methyl viologen), copper sulfate, metidathion), anti-pathogenic compounds such

COR-AAD1-00001929

as fungicides (chlorothalonil, 2-thienylalanine) and profungicides (Oxime Carbamates) or compounds responsible for insect behavior modulation (pheromones, allomones and kairomones), and herbicides referring to compounds having specific or general toxicity to plant species. Typical herbicides include but are not limited to the class of sulfonylurea herbicides and sulfonanilide herbicides (chlorsulfuron, triasulfuron, metsulfuron-methyl), auxin herbicides (e.g., dicamba, 2-methyl-4-chlorophenoxyacetic acid, picloram, quinclorac, quinmerac), pre-emergence herbicides (metribuzin), and post-emergence herbicides (Clethodim Pendimethalin, oryzalin, dithiopyr, oxadiazon, prodiamine, and 2,4-dichlorophenoxyacetic acid).

"Sulfonylurea herbicides" are defined as N-(heterocyclicaminocarbonyl)-arylsulfonamide-containing herbicidal compounds that inhibit the enzyme acetolactate synthase, such as sulfometuron methyl.

The term "sulfometuron methyl" refers to 2-[[[[(4,6-dimethyl-2-pyrimidinyl)amino]carbonyl]amino]sulfonyl]benzoic acid, methyl ester (CAS registry number 74222-97-2), and is abbreviated as "SM".

"Acetolactate synthetase" or "ALS" is a key enzyme responsible for branched chain amino acid biosynthesis.

"Glyphosate" will be abbreviated "GP", has the CAS registry number 1071-83-6 and is a herbicide whose site of action is 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) which catalyzes the conversion of shikimate into anthranilate, a key transformation in plant amino acid synthesis.

The term "thienylalanine-like activity" means any substance, natural or synthetic that has the fungicidal activity of chlorothalonil, 2-thienylalanine.

The term "glypohosate-like activity" means any substance, natural or synthetic that acts the interfere with 5-enolpyruvylshikimate-3-phosphate synthase activity.

The term "ALS-inhibiting activity" means" means any substance, natural or synthetic that inhibits the activity of acetolactate synthetase or expression of the gene encoding acetolactate synthetase.

A "luminescent reporter gene complex" means any reporter gene(s) the products of which result in light production. Examples include but are not limited to the bacterial *lux* genes; the luciferase genes (*luc*), from, for example, the firefly (*Photinus pyralis*) or click beetle (*Pyrophorus plagiophthalamus*); or the gene encoding the luciferase from the sea pansy (*Renilla reniformis*).

"Site of action" refers to the macromolecular target of a particular stressor or xenobiotic compound. Typical sites of action are specific enzymes in a particular biosynthetic pathway.

COR-AAD1-00001930

The terms "plasmid", "vector", and "cassette" refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and are usually in the form of circular double-stranded DNA molecules. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell.

"Integrant" refers to a bacterial strain into whose chromosome has been inserted a foreign gene fragment.

The term "multiple copies" or "multicopy" as it pertains to the presence of expressible genes in an organisms means a number of copies of the gene that exceeds the normal complement of that gene in the cell.

The terms "transformation" and "transfection" refer to the acquisition of new genes in a cell as a result of the incorporation of nucleic acid. The acquired genes may be integrated into chromosomal DNA or introduced as extrachromosomal replicating sequences. The term "transformant" refers to the product of a transformation.

The terms "promoter" and "promoter region" refer to a sequence of DNA, usually upstream of (5' to) the protein coding sequence of a structural gene, which controls the expression of the coding region by providing the recognition for RNA polymerase and/or other factors required for transcription to start at the correct site. Promoter sequences are necessary but not always sufficient to drive the expression of the gene.

"Genotoxic sensitive promoter" refers to a promoter activated by DNA damage. Examples of these promoters include but are not limited to *recA*, *uvrA*, *lexA*, *umuDC*, *uvrA*, *uvrB*, *uvrC*, *sulA*, *recN*, *uvrD*, *ruv*, *alkA*, *ada*, *dinA*, *dinB*, *dinD*, and *dinF* as well as other promoters which are members of the adaptive response regulon group such as those disclosed by Rupp in *E. coli* and *Salmonella*; Cellular and Molecular Biology [Niedhardt et al., Eds., pp 1190-1220, American Society of Microbiology, Washington, D.C. (1996))] and members of the SOS regulon group disclossed by Walker in *E. coli* and *Salmonella typhimurium*; Cellular and Molecular Biology [Niedhardt et al., Eds., pp 1400-1416, American Society of Microbiology, Washington, D.C. (1996))].

A "fragment" constitutes a fraction of the DNA sequence of the particular region.

COR-AAD1-00001931

"Regulation" and "regulate" refer to the modulation of gene expression controlled by DNA sequence elements located primarily, but not exclusively, upstream of (5' to) the transcription start of a gene. Regulation may result in an "all or none" response to a stimulation, or it may result in variations in the level of gene expression.

The term "operably linked" refers to the fusion of two fragments of DNA in a proper orientation and reading frame to be transcribed into functional RNA.

The term "expression" refers to the transcription and translation to gene product from a gene coding for the sequence of the gene product. In the expression of a gene, a DNA chain coding for the sequence of gene product is first transcribed to a complimentary RNA which is often a messenger RNA and, then, the thus transcribed messenger RNA is translated into the above-mentioned gene product if the gene product is a protein.

"Heightened expression" refers to a gene expression greater than that seen in a mock-treated culture. In the case of *lux* gene fusions, heightened expression is indicated by an increase in bioluminescence above background levels that is characterized by a temporal delay that allows for increased transcription and subsequent translation.

The terms "stressor" or "insult" refer to a chemical agent or physical treatment that results in suboptiminal growth of an organism. Stressors may include, but are not limited to, chemicals (such as herbicides, crop protection chemicals, environmental pollutants, heavy metals), physical treatments such as changes in temperature, changes in pH, agents producing oxidative damage or DNA damage (such as from UV exposure), anaerobiosis, biological insults such as the introduction of other life forms (viruses, bacteria, etc.) into the bacterial culture, or changes in nutrient availability. Additionally, stressors may include naturally-occurring compounds such as L-valine, galactose-phosphate, 2-ketobutyrate. A "stressor target" means a specific macomolecular target inhibited by a specific stressor.

The terms "xenobiotic compound" amd "xenobiotic chemical" refer to any stressor chemical which does not typically occur in nature. Typical xenobiotics of interest in the present invention include those useful as herbicides, pesticides, fungicides or any other xenobiotic capable of interfering with a specific metabolic site of action.

The term "bioluminescence" refers to the phenomenon of light emission from a living organism.

"Bioluminescent positive phenotype" refers to a phenotype displaying an increase in light production by a detector cell containing a *lux* gene fusion.

COR-AAD1-00001932

The term "baseline luminescence" means the amount of light produced by a cell having a bioluminescent positive phenotype in an unstressed metabolic state.

The term "*lux* gene complex" refers to the *lux* structural genes which include *luxA*, *luxB*, *luxC*, *luxD* and *luxE* and which are responsible for the phenomenon of bacterial bioluminescence. A *lux* gene complex might include all of the independent *lux* genes, acting in concert, or any subset of the *lux* structural genes so long as *luxA* and *luxB* are part of the complex.

"Gene fusion" is a hybrid DNA fragment comprising a regulatory signal essential for transcription (referred to as a promoter) fused to at least one structural gene sequence coding for a specific polypeptide.

The term "*lux* gene fusion" means the fusion of the *lux* gene complex with a suitable stressor-sensitive promoter. "recA-*LuxCDABE*" refers to the specific fusion of the genotoxic sensitive promoter *recA* fused to the bacterial *Lux* gene complex.

The term "*ilvBN*" refers to the structural genes encoding, respectively, the large and small subunits of the heterotetrameric ALS I-EC 4.1.3.18.

The term "*ilvGM*" refers to the structural genes encoding, respectively, the large and small subunits of the heterotetrameric ALS II-EC 4.1.3.18. *ilvGM* is cryptic in *E. coli* K-12 Δ ilvB-a mutation that deletes part of the *ilvB* gene.

The term "*ilvIH*" refers to the structural genes encoding, respectively, the large and small subunits of the heterotetrameric ALS III-EC 4.1.3.18. *ilvIH* is cryptic in laboratory strains of *Salmonella typhimurium*.

The term "*pheA*" refers to the structural gene encoding the bifuctional polypeptide that displays chorismate mutase (EC 5.4.99.5) and prephenate dehydratase (EC 4.2.1.51) activities.

The term "*relA*" refers to the structural gene encoding the ATP:GTP 3'-pyrophosphotransferase I-EC 2.7.6.5.

The term "*spoT*" refers to the structural gene encoding the ATP:GTP 3'-pyrophosphotransferase II-EC 2.7.6.5.

The term "*aroA*-" refers to the structural gene encoding enolpyruvylshikimate phosphate synthase-EC 2.5.1.19.

The term "*aroH*-" refers to the structural gene encoding DHAP(tryptophan repressible and feed back inhibitable) synthase-EC4.1.2.15.

The term "*tolC*+" refers to the structural gene encoding an outer membrane porin needed for the efflux of many xenobiotics pumped out by a variety of membrane translocases.

The term "*glk*-" refers to the structural gene encoding glucokinase-EC2.7.1.2.

COR-AAD1-00001933

The terms "detector organism", "detector bacteria", and "detector cell" refer to an organism which contains a gene fusion consisting of a genotoxic-sensitive promoter fused to a luminescent reporter gene or gene complex.

The term "non-bioluminescent parent" of the detector cell is a bacterial strain into which a light producing genetic cassette has not been introduced.

"Genetic titration" refers to an alteration of the genetic makeup of a microbe such that the levels of a macromolecular target are elevated to a point whereby they overcome the action of a stressor. Within the context of the present invention genetic titration will involve a process of screening for biochemcal targets of compounds
where a host organism is transfected with a suitable genomic library, transformants are screened for growth in the presence of the compound and portions of the library conveying resistance to the compound are isolated and identified.

"Nutritional reversal" refers to the addition of a nutrient to a culture contacted with a stressor such that the biological output of the culture is restored to the unchallenged level.

"Nutrient" refers to an end product of a biochemical pathway or a compound readily converted to a pathway end product. Typical nutrients are amino acids, vitamins, bases or sugars. Vitamins are readily converted to cofactors which are pathway end products; similarly bases are readily transformed *in vivo* into nucleotide triphosphates.

The term "auxanography" means the diagnostic and systematic administration of nutrient pools to determine the pathway blocked in a microorganism as described in Davis, R. W., D. Botstein And J. R. Roth. A Manual For Genetic Engineering: Advanced Bacterial Genetics. X+254p. Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., U.S.A., 1980.

The term "Relative Light Unit" is abbreviated "RLU" and refers to a measure of light emission as measured by a luminometer, calibrated against an internal standard unique to the luminometer being used.

DETECTOR CELLS AND NONBIOLUMINESCENT STRAINS

Host cells suitable in the present invention include any cell capable of expression of the *lux* gene fusion where prokaryotic cells are preferred and where members of the enteric class of bacteria are most preferred.

Enteric bacteria are members of the family *Enterobacteriaceae* and include such members as *Escherichia*, *Salmonella*, and *Shigella*. They are gram-negative straight rods, 0.3-1.0 X 1.0-6.0 mm, motile by peritrichous flagella (except for *Tatumella*) or nonmotile. They grow in the presence and absence of oxygen and grow well on peptone, meat extract, and (usually) MacConkey's media. Some

COR-AAD1-00001934

grow on D-glucose as the sole source of carbon, whereas others require vitamins and/or mineral(s). They are chemoorganotrophic with respiratory and fermentative metabolism but are not halophilic. Acid and often visible gas is produced during fermentation of D-glucose, other carbohydrates, and polyhydroxyl alcohols. They are oxidase negative and, with the exception of *Shigella dysenteriae* 0 group 1 and *Xenorhabdus nematophilus*, catalase positive. Nitrate is reduced to nitrite (except by some strains of *Erwinia* and *Yersina*). The G + C content of DNA is 38-60 mol% (T_m, Bd). DNAs from species within most genera are at least 20% related to one another and to *Escherichia coli*, the type species of the family. Notable exceptions are species of *Yersina, Proteus, Providenica, Hafnia* and *Edwardsiella*, whose DNAs are 10-20% related to those of species from other genera. Except for *Erwinia chrysanthemi*, all species tested contain the enterobacterial common antigen (Bergy's Manual of Systematic Bacteriology, D. H. Bergy et al., Baltimore: Williams and Wilkins, 1984).

Particularly useful in the present invention are *E. coli* and members of the genus *Salmonella* and those enterics for which the *recA* gene and its upstream promoter region has been identified.

Membrane Mutations

Host cells of the present invention optionally may contain mutations that will facilitate the screening process. The appropriate bacterial strain with which to test the effects of a chemical is one whose growth is affected by that chemical. Hence, the chemical of interest must be able to enter the cell, be retained in the cell, and interact with target molecules of the cellular machinery. Various mutations of *E. coli* are known to affect permeation into and accumulation within the cell. Strains carrying mutant alleles of *rfa* (Ames, B. N., F.D. Lee, and W. E. Durston, *Proc. Nat. Acad. Sci. USA*, 70(3): p. 782-786, (1973)), *envA* (Young, K. and L. L. Silver, *J. Bacteriol.*, 173: p. 3609-3614, (1991)), *imp* (Sampson, B. A., R. Misra, and S. Benson, *Genetics*, 122: p. 491-501, (1989)), *lpp*,(Giam, C. Z., S. Hayashi, and H. C. Wu, *J. Biol. Chem.*, (259): p. 5601-5605, (1984)) or *surA* (Tormo, A., M. Almiron, and R. Kolter, *J. Bacteriol.*, (172): p. 4339-4347, (1990)) have increased sensitivity to a variety of chemicals. Destruction of efflux pumps, with mutations such as *emr* (Ma, D., et al., *Mol. Microbiol.*, 16: p. 45-55, (1995)), or *acrAB* (Paulsen, I.T., et al., *Mol. Micro.*, 19: p. 1167-1175, (1996)), or the channels they use, with mutations such as *tolC* (Schnaitman et al., *J. Bacteriol*, 172 (9), pp 5511-5513, (1990)), also result in increased chemical sensitivity. Moreover, it is contemplated that genetic titration's in conjunction with genomics will permit the development of method for the identification of such sensitizing mutations which can then be integrated in to an appropriate detector cell background. In some instances, the target macromolecule of a chemical may be

COR-AAD1-00001935

intrinsically resistant to the action of that chemical. For example, *E. coli* has two isozymes of the enzyme acetolactate synthase, one of which has a poor binding affinity for the sulfonylureas herbicides. Mutations which destroy the function of *ilvBN*, encoding the resistant isozyme, result in a strain with greatly increased susceptibility to growth inhibition by sulfonylureas that target acetolactate synthase (LaRossa, R. A. and D. R. Smulski, *J. Bacteriol.*, 160: p. 391-394, (1984)). An appropriate host strain of *E. coli* or other bacteria may be constructed to carry a known mutation or combinations of mutations. Furthermore, an appropriately sensitive strain may also be found by screening for growth inhibition following mutagenesis by transposon insertion or chemical or physical treatments.

Mutations Conferring Stress Sensitivity

Detector cells of the present invention optionally may contain mutations that will convey sensitivity to a particular stress to be screened for. The *relA* mutation, for example, prevents the cell from responding to amino acid starvation. Other mutations that will be useful in stress sensitivity include, for example, mutations that result in sensitivity to various anticancer drugs and compounds that cause oxidative stress. It is known that many anticancer drugs interfere with DNA replication while compounds that cause oxidative stress may be useful in controlling fungal pathogens of crops and bacterial pathogens of humans or animals. Mutations in genes conveying sensitivity to stresses are preferred, including but not limited to, mutant genes selected from the group consisting of *cya*, *crp*, *spoT*, *arcAB*, *envZ*, *ompR*, *marR*, *earAB*, *fur*, *oxrG*, *fruR*, *rpoS*, *rpoE*, *creB*, *creC*, *glnG*, *glnL*, *glnB*, *glnD*, *glnF*, *phoB*, *phoP*, *phoQ*, *phoR*, *phoU*, *rpoH*, *lexA*, *recA*, *lrp*, *soxRS*, *oxyR*, *fnr*, *atbR*, *ada*, and *relA* where the *relA* mutation is most preferred. Moreover, genetic titration, may indicate unexpected roles for regulatory genes. Such discovery, coupled with inactivation of the regulatory locus, may optimize the sensor strains.

Within the context of the present invention strains generally contained mutations such that only a single ALS (I or III) sensitive to sulfonylurea herbicides was expressed. Such a sensitive host may either be screened from wild type after standard transposon, chemical (e.g., HNO_2 and NH_2OH), UV, intercalating dye (e.g., acridine dyes) or other mutagenesis protocols have generated the appropriate hypersensitive mutations or can be constructed by combining mutations that together yield the desired sensitivity. For a review of the methods of mutagenesis, see, for example, Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA.

COR-AAD1-00001936

Stress Sensitive Reporter; *Escherichia coli* and *Salmonella*; Cellular and Molecular Biology (Niedhardt et al. Eds., American Society of Microbiology, Washington, D.C. (1996))].

The detector cell of the present invention optionally may also contain a stress-sensitive reporter for the detection of particularly desirable or undesirable characteristics of the stressor compound to be screened. Typically, stress-sensitive reporters are comprised of a stress-sensitive promoter operably linked to a suitable reporter element. The promoter must be chosen so as to be expressible within the specific detector cell desired. Typically in bacterial detector cells, the promoters will be chosen from stress-inducible bacterial promoters. Examples of stress-inducible promoters suitable in the present invention are those responsive to chemicals, environmental pollutants, heavy metals, xenobiotics, changes in temperature, changes in pH as well as agents producing oxidative damage, DNA damage, anaerobiosis, changes in nitrate availability or pathogenesis. Specific examples of suitable bacterial stress-promoters include, but are not limited to, those sensitive to protein damage such as the heat shock genes (*grpE*, *dnaK*, *lon*, *rpoD*, *groESL*, *lysU*, *htpE*, *htpG*, *htpI*, *htpK*, *clpP*, *clpB*, *htpN*, *htpO*, and *htpX*), those sensitive to DNA damage such as those controlled by the SOS regulatory circuit (*recA*, *uvrA*, *lexA*, *umuDC*, *uvrA*, *uvrB*, *uvrC*, *sulA*, *recN*, *uvrD*, *ruv*, di*nA*, *dinB*, *dinD*, and *dinF*), those sensitive to oxidative damage (*katG*, *ahp*, *micF*, *sodA*, *nfo*, *zwf*, and *soi*), those sensitive to membrane damage (*fabA*), those sensitive to amino acid starvation and under the control of the *relA*, or *spoT* regulatory genes(*his*, *ilvBN*, *ilvGMED*, and *thrABC*), those sensitive to carbon starvation and under the control of the *cya*, and *crp* regulatory genes (*lac*, *mal*, and *gal*), those sensitive to phosphate starvation under the control of the *phoB*, *phoM*, *phoR*,and *phoU* regulatory genes (*phoA*, *phoBR*, *phoE*, *phoS*, *aphA*, *himA*, *pepN*, *ugpAB*, *psiD*, *psiE*, *psiF*, *psiK*, *psiG*, *psiI*, *psiJ*, *psiN*, *psiR*, *psiH*, *phiL*, *phiO ara*, *tna*, *dsd*, and *hut*), and those sensitive to nitrogen starvation under the control of the *glnB*, *glnD*, *glnG*, and *glnL* regulatory genes (*glnA*, and *hut*). Preferred within the context of the present invention are those stress-inducible promoters sensitive to genotoxicity or DNA damage under the control of the SOS regulatory circuit where the *recA* gene is most preferred.

Reporter genes suitable for fusion to the stress inducible promoter are structural genes under the control of such a promoter and able to report a detectable signal. Many bacterial reporters such as *lacZ*, *galK*, *xylE*, *luc*, *luxAB*, *luxCDABE*, *phoA*, *uidA* (GUS), *cat*, *npt-II*, *SUC2* and ubiquitin are known in the art (Miller, J. H., A Short Course in Bacterial Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 63-67). Of particular use in the present invention are bioluminescent reporter genes including but not limited

COR-AAD1-00001937

to the bacterial *lux* genes; lucifersae genes (*luc*) from, for example, the firefly, *Photinus pyralis*, or click beetle, *Pyrophorus plagiophthalamus*; or the gene encoding the luciferase from the sea pansy, *Renilla reniformis*. Because its gene products function well in *E. coli* under a wide range of temperatures, most preferred is the promoterless *Photorhabdus luminescens luxCDABE* gene complex obtained from the pCGLS1 plasmid containing the *lux* gene complex. This complex is fully described by Rosson, R. A., in PCT International Application WO 93/03179 (1993).

Bacterial bioluminescence as produced by the *luxCDABE* gene complex is a phenomenon in which the products of 5 structural genes (*luxA, luxB, luxC, luxD and luxE*) work in concert to produce light. The *luxD* product generates a 14 carbon fatty acid from a precursor. The 14 carbon fatty acid is activated in an ATP dependent reaction to an acyl-enzyme conjugate through the action of the *luxE* product which couples bacterial bioluminescence to the cellular energetic state. The acyl-enzyme (*luxE* product) serves as a transfer agent, donating the acyl group to the *luxC* product. The acyl-*Lux*C binary complex is then reduced in a reaction in which NADPH serves as an electron pair and proton donor reducing the acyl conjugate to the C^{14} aldehyde. This reaction couples the reducing power of the cell to bacterial light emission. The light production reaction, catalyzed by luciferase (the product of *luxA* and *luxB*), generates light. The energy for light emission is provided by the conversion of aldehyde to fatty acid and $FMNH_2$ oxidation, providing another link between light production and the cellular energy state.

The usefulness of *luxCDABE* are limited by the inherent thermolability of the protein products of these genes. The temperature requirement of this reporter system has limited overlap with the need to grow bacteria rapidly in defined media. Applicants have solved this problem by using *luxCDABE* that encode protein products capable of functioning in the desired temperature range (28-42°C).

A multiplicity of strains were engineered, each having a specific genotype useful for the specific site of action screen desired. Construction of the strains is reviewed in Figure 1 and the genotypes are summarized in Table 1 in the GENERAL METHODS. All strains engineered to contain the recA-*Lux* -CDABE gene fusion functioned as detector cells. Detector cells may contain only the gene fusion, or optionally may possess other mutations affecting membrane permeability or stressor sensitivity.

Useful strains possess a variety of genotypes including where the expression or suppression of the *ilvB, relA, tolC, igm, spoT* and *ilvG* genes were used to give useful detector cells. Methods of strain construction are well known

COR-AAD1-00001938

in the art and use the basic elements of molecular biology and microbiology fully discussed in Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press (1989); *Escherichia coli* and *Salmonella*; Cellular and Molecular Biology (Niedhardt et al. Eds., American Society of Microbiology, Washington, D.C. (1996))].

Culture Conditions:

Typically, cells are grown at 37 °C in an appropriate media. Preferred growth medium in the present invention are common defined media such as Vogel-Bonner medium (Davis et al., *Advanced Bacterial Genetics*, (1980), Cold Spring Harbor, NY:Cold Spring Harbor Laboratory). Other defined or synthetic growth media may also be used and the appropriate medium for growth of the particular microorganism will be known by one skilled in the art of microbiology or fermentation science. In some instances rich or complete media such as NB (Nutrient broth) are used.

Suitable pH ranges for bacterial growth are between pH 5.0 to pH 9.0, where pH 6.0 to pH 8.0 is preferred as the initial condition.

Growth of the bacterial cells in liquid medium allows a uniform population of cells to be stressed at various growth phases such as early log phase, mid log phase, late log arithmic phase, or stationary phase. Stress is the condition produced in a cell as the result of exposure to a cellular insult or stressor. This cellular stress may be caused by any substance or change in the cellular environment that results in an alteration of normal cellular metabolism in a bacterial cell or population of cells. The addition of chemicals such as herbicides, crop protection chemicals, environmental pollutants, or heavy metals to the growth media can cause such a stress. Additionally, changes in temperature, changes in pH, agents producing oxidative damage or DNA damage (such as from UV exposure), anaerobiosis, or changes in nitrate availability are insults that may cause stress as well.

STRESSOR XENOBIOTICS AND CROP PROTECTION CHEMICALS

The present method is useful for identifying the site of action for a variety of xenobiotic chemicals and particularly those compounds useful as crop protection chemicals. The instant method is useful for determining the site of action of compounds selected from but not limited to sulfonylurea herbicides such as sulfometuron methyl, glyphostate, the profungicides such as the oxime carbamates, oxidants such as methyl viologen (paraquat), hormonal herbicides such as 2,4-dichlorophenoxyacetic acid, and fungicides such as 2-thienylalanine. Any chemical that negatively impacts bacterial metabolism can be anlyszed in this manner. Thus, not only antibacterials, but also antivirals, anticancer agents, anticorrosives , disinfectants and crop protection chemicals may be analyzed with

COR-AAD1-00001939

this invention since the site of action of many such compounds are often enzymes or cofactors shared by bacteria and target organisms. Moreover, products of normal metabolism whose accumulation is detrimental can be analyzed with these methods.

SITE OF ACTION SCREENS

Determination of the site of action of various stressors and xenobiotic compounds by the present method proceeded using the complementary screens of nutritional reversal and genetic titration.

Nutritional Reversal

The above defined media provides those nutrients (inorganic compounds and carbons sources) necessary and sufficient for the growth of wild type *E. coli*. *E. coli* detector cells (containing the recA-*Lux* gene fusion) grow in these media converting the nutrients into organic chemicals required for cell growth and metabolism. If a stressor chemical interferes with one of the biosynthetic processes, the cellular metabolism and the bioluminescence of a culture will be reduced or eliminated.

The present method of nutritional reversal involves first screening compounds for the ability to depress bioluminescence of the detector organism containing the recA-*LUX* gene fusion. Next, panels of cultures are initiated to test which nutrient will reverse the growth inhibition as indicated by a reversal of bioluminescent depression. Stressor chemicals meeting this criteria are defined as nutritionally reversed. The specific pathway affected by a stressor chemical is then determined by auxanographic analysis. Supplementation with biosynthetic intermediates from the identified pathway to determine which intermediates obviate the depression of bioluminescence may be used to define the target within the pathway. Modulation by non-pathway nutrients (Van Dyk and LaRossa, pp. 123-130 and LaRossa et al, pp 109-121 both in Biosynthesis of Branched Chain Amino Acids, Barak et al., eds. Balaban Publishers, Weinheim, FRG, 1990) can also help to define the macromolecular target of the stressor chemical.

Genetic Titration

The genetic titration protocol is predicated upon the fact that chemical stressor action can be overcome by increasing the intracellular concentration of the macromolecular target above that normally found. In this manner the inhibitor action is titrated out and cells grow under conditions that normally inhibit the growth of the wild type. The means of increasing the macromolecular target could be through the selection of genetic duplications (e.g., Anderson and Roth, *Ann Rev Microbiol.*, 31:473-505 (1977); Wahl et al., *J Biol Chem*, 254: 8679-89 (1979), Alt et al., *J Biol Chem*, 253: 1357-70 or through the selection of high level constitutive regulatory mutants (Roth et al., *J. Mol. Biol.* 22:305-323 (1966);

COR-AAD1-00001940

Hartman, Ames, Anton Yanofsky Calvo et al., *J. Bacteriol.* 97:1272-1282 (1969), Calvo et al., *Genetics*, 61:777-787 (1969)). These methods, however, lead to rather laborious analyses to define the overexpressed gene product. A preferred embodiment is that pioneered by Falco [Falco and Dumas, *Genetics* 109:21-35 (1985)] and Rine [Rine et al., *PNAS* USA 80:6750-6754 (1983)] in yeast in which gene amplification is created by (a) construction of genomic libraries in high copy number autonomously replicating plasmids and (b) the introduction of such plasmid libraries into a suitable host strain creating a catalog of mero-multipoids [LaRossa, (1996), in *Escherichia coli* and *Salmonella*; Cellular and Molecular Biology (Niedhardt et al. Eds., pp 1400-1416, American Society of Microbiology, Washington, D.C. (1996))]. In the next stage, (c), the meromultiploids conferring resistance to the chemical stressor are identified by selection for colonies that grow above the compound's MIC (minimal inhibitory concentration). Such methodology has been recently extended to *E. coli* (Chatterjee and Sternberg, *PNAS* USA 92:8950-8954 (1995)).

Prevention Of Inhibitor Action By Partial Genome Amplification.

In addition to the techniques of nutritional reversal and genetic titration it is contemplated that present detector cells may be employed to identify compounds on the basis of the prevention of inhibitor action. For example, the present application exemplifies the incorporation of several genomic fragments suspected of or known to confer growth resistance to thienylalanine into a bioluminescent or other tester strains by transformation selecting for ampicillin resistance. The resultant bioluminescent strains, along with negative controls in which the multicopy plasmid vectors lacking an insert were used for transformation, were exposed to various concentrations of thienylalanine and kinetic plots of bioluminescence versus time were obtained (Example 9) . As demonstrated in Example 9, clone *o245 ygaH* conferred the greatest degree of protection from thienylalanine while other clones confer intermediate levels of resistance. It is evident from this Example that (a) cross resistance can be discovered by simply replacing thienylalanine with other compounds in the test (indeed disk diffusion assays have revealed cross resistance conferred by *o245, ygaH,* to several other amino acid analogs including those identified in table 12) and; (b) other chemicals hitting a specific target can be identified by asking if particular target specifying gene in multiple copies confers inhibitor resistance. In this manner new inhibitors of acetolactate synthase have been revealed).

Membrane Alterations

In another embodiment it is contemplated that the present detector cells may be modified to incorporate disrupted membrane proteins which in turn may be exploited to identify compounds having a specific biological activity. For

COR-AAD1-00001941

example, genetic titration with SM identified *tolC* as a resistance determinant. Similarly titration with GP recognized *yhhTS*, thienylalanine recognized *o245 ygaH*, acivicin recognized *yedA* and mitomycin C recognized *mdfA* as resistance determinants (Examples 7-9) . Each of these genes encodes a predicted membrane protein. It is axiomatic that disruption of these cloned genes and incorporation of the disruptant into the *E. coli* chromosome can be achieved by standard techniques without undo experimentation. The utility of such disruptants can be readily assessed by bioluminescent assays that determine the doses that reduce light output by a factor of 2. These disrupted genes will be useful for a plethora of sensitive bioassays.

Discussion of the Preferred Embodiments

A preferred embodiment of the present method is illustrated in Figure 2. A detector organism (1) having a *recA-luxCDABE* fusion and a *relA* mutation (inhibiting the cell's response to amino acid starvation) is exposed to a battery of compounds (stressors) (2) to be screened for crop protection activity. Compounds producing an increase in light from the detector cell are discarded as genotoxic. Compounds that slow cell growth but do not produce an increase in light are subjected to nutritional reversal screens or genetic titration screens to determine the site of action of the compound. Where nutritional reversal is used, detector cells are grown in a minimal medium supplemented with a multiplicity of different nutrient pools, each pool composed of a different mixture of nutrients (3). The pool supplying the amino acid necessary to reverse metabolic inhibition is detected by the recovery of light production by the bioluminescent gene fusion (4).

The skilled person will appreciate that advances in genomics make a number of related embodiments possible. For example, into the detector organism described in the preceding paragraph an ordered set of overlapping, high copy number plasmids is placed such that each contain different segments of the *E. coli* chromosome. This set of resulting strains may be used in a genetic titration screen of a chemical stressor to identify those chromosomal regions that upon amplification restore bioluminescent output to uninhibited levels. Thus, genetic titration may be performed in alternative selection and screening modes.

In another embodiment, detector cells containing the *recA-LUX* gene fusion were constructed by transformation of suitable hosts with the chimera according to standard methods (Sambrook *supra*). The transformants possessed a variety of mutations including *ilvB*, *relA* and *tolC*. Detector cells were constructed so as to contain one or more of these mutations. These detector cells emitted a baseline luminescence that was altered by the exposure to various xenobiotics.

COR-AAD1-00001942

The sensitivity of the detector cell to genotoxic agents was examined by exposing the cell to mitomycin C. Moderate levels (0.3-20 ug/mL for a *tolC*$^+$ strain, 0.3-1.25 ug/mL for a *tolC* derivative) of mitomycin C resulted in an increase in baseline luminescence. High levels (>2.5 ug/mL in a *tolC* strain) of mitomycin C resulted in a decrease in luminescence (Figures 4 and 5).

Detector cells containing a variety of mutations were treated with 4 different herbicides (SM, MV, GP and 2,4-D) to determine the effect of the compounds on bioluminescent output of the detector cell. In each case kinetic plots indicate a dose-dependent decrease in light emission in response to the herbicides. Data illustrating the effect is seen in Figures 6-9).

Detector cells containing the *tolC* mutation were exposed to varying concentrations of SM and mitomycin C to determine the effect of the mutation on the sensitivity of the assay. As is noted in Tables 2 and 3 and Figures 4 and 10 and 4 the *tolC* mutation not only enhanced the responsiveness of the detector cell to the lipophillic SM but also to mitomycin C. These tests demonstrated the utility of the *tolC* mutation as a component of the detector cell.

The site of action of two CPC's were analyzed using a detector cell comprising the *recA-LuxCDABE* fusion. Using the bioluminescent detector cells, nutritional reversal was applied to each of the compounds to determine which nutrient would reverse the growth inhibiting effects of the compound. As illustrated in Tables 2 and 3, cysteine metabolism was identified as the potential affected site for cells treated with the oxime carbamate compound while phenylalanaine metabolism was the identified site of action for the cells treated with thienylalanine.

Finally, detector cells harboring *ilvB*, *relA* and *tolC* alleles were used in genetic titration assays to determine the site of action of the SM, 2-thienylalanine and GP. The detector cells were transformed with *E. coli* cDNA libraries and transformants were screened for resistance to SM as indicated by changes in bioluminescence of the transformants. Resistant colonies were picked and the plasmids isolated, sequenced and analyzed for genes encoding for possible targets. In this fashion the *ilvBN* and *ilvIH* genes were identified and confirmed as the genes encoding ALS, the target for SM. In similar fashion the *aroA*, the known target for GP encoding EPSPS, was selected. The *aroH* gene was obtained in a selection for 2-thienylalanine resistant clones.

In another alternate embodiment describing *in vivo* inhibitor identification of specified targets, the present detector cells may also be utilized in a method to screen for compounds where the site of action is known. Such an embodiment is illustrated in Figure 3. It is contemplated that a detector organism may be constructed to include not only the stress promoter-*luxCDABE* gene fusion, but

COR-AAD1-00001943

also a plasmid expressing the gene target of the compound to be screened. For example, it is known that the gene target for GP is 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) encoded by the *aroA* gene, and plasmids have been constructed so as to express the EPSPS gene product. Compounds to be screened would be exposed to a standard detector cell (1) and the modified detector cell (2) containing the expressible gene target (3) at suitable concentrations. No addition of compounds (4) will give a baseline luminescence in both detector cells. Addition of compounds that inhibit growth (5) will produce a decrease in luminescence in both detector cells. However, where exposure of the standard detector cell (1) to a glyphostate-like molecule (6) will produce a decrease in light production, exposure of the same compound to the modified detector cell (2) will give a smaller or no decrease in light production. The compound was genetically titrated by the expression of the gene target in the modified detector cell(2).

It will be appreciated by one skilled in the art that the *E. coli aroA* gene may be substituted with plant or other bacterial *aroA* homologues. Moreover, it is contemplated that chemistry for any target of interest can be identified in this manner. The method of *in vivo* inhibitor identification of specified targets here exemplified could be used to identify many useful compounds that include, but are not limited to, new crop protection chemistries, antibacterial chemistries, antifungal chemistries, anticancer chemistries, anti-viral chemistries, chemistries preventing biofilm formation and anticorrosives.

The method of *in vivo* inhibitor identification of specified targets may also be used to detect inhibitors of ALS III, for example. DPD1718 could be transformed with pBR322 selecting for ampicillin or tetracycline resistance yielding strain pBR322/DPD1718. Similarly DPD1718 may be transformed with an *ilvIH* containing plasmid selecting for ampicillin resistance yielding strain pIlvIH/DPD1718 having an ALS III phenotype. These two transformed strains may then be used to screen for specific inhibitors of ALS III since the titer of ALS III will be much greater in strain pIlvIH/DPD1718 than in pBR322/DPD1718. Expected results follow:

Chemical	Relative MLC*	
	pIlvIH/DPD1718	pBR322/DPD1718
Non-ALS toxicant	1	=1
ALS III inhibitor	>1	=1

*Minimal Lumino - Inhibitory Concentration

COR-AAD1-00001944

It is contemplated then that one can score other growth related phenotypes or growth itself to identify novel compounds interfering with EPSP, ALS or other desired cellular functions.

The above embodiment is more fully exemplified in Examples 7-9 which demonstrate how the introduction of genes conveying resistance to Mitomycin C, Acivicin and Thienylalanine into a *recA-LUX* detector cell may be used to screen for compounds having to Mitomycin C, Acivicin and Thienylalanine activity respectively.

In another embodiment the *recA-LUX* detector cell is used as a means for screening compounds for mutagencity. The efficacy of the present SOS regulated bioluminescent gene fusions is seen in the comparison of fusion containing detector cells as indicators of compound mutagenicity as compared with the standardized Ames test (Example 10). Compounds chosen at random which tested positive in the standard Ames test were confirmed as mutagenic by giving a 'lights-on" response in bioluminescent detector cells.

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLES

GENERAL METHODS

Procedures for phosphorylations, ligations and transformations are well known in the art. Techniques suitable for use in the following examples may be found in Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press (1989).

Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following examples may be found in Manual of Methods for General Bacteriology (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994) or Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA. Standard bacterial genetic protocols can be found in Miller, J. H., Experiments in Molecular Genetics, Cold Spring Harbor Laboratory (1972); Miller, J. H., (1992) A Short Course in Bacterial Genetics, Cold Spring Harbor Laboratory Press; Silhavey, T. J., Berman, M. L., and Enquist, L. W.,

COR-AAD1-00001945

Experiments with Gen Fusions (1984), Cold Spring Harbor Laboratory, and Davis, R. W., Botstein, D. and Roth, J. R. (1980), A Manual for Genetic Engineering-Advanced Bacterial Genetics, Cold Spring Harbor Laboratory. All reagents and materials used for the growth and maintenance of bacterial cells were obtained from Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD), or Sigma Chemical Company (St. Louis, MO) unless otherwise specified.

The meaning of abbreviations is as follows: "h" means hour(s), "min" means minute(s), "sec" means second(s), "d" means day(s), "mL" means milliliters, "L" means liters.

Bioluminescence was measured in a luminometer, model ML3000 (Dynatech Laboratories, Chantilly, VA)according to the manufacturer's instructions.

Crop protection chemicals used in the following examples were sulfometuron methyl [obtained from DuPont Agricultural Products, Wilmington, DE], glyphosate [obtained from Sigma], methyl viologen [obtained from Sigma] 2-Thienylalanine [obtained from Aldrich], compound OC, member of the class of oxime carbamates having profungicide activity was prepared by DuPont Agricultural Products. Mitomycin C and acivicin was obtained from Sigma. Stock solutions were prepared as follows:

2-thienylalanine - (10 mg/mL) in water
sulfometuron methyl - (3.6 mg/mL) in 0.01 N NaOH
glyphosate - (5 mg/mL) in water
compound OC - (10 mg/mL) in DMSO (dimethylsulfoxide)
methyl viologen - (2 mg/mL) in water
mitomycin C - (1 mg/mL) in water
acivicin 10 mg/ml in water.

Strains

The transformed *E. coli* strains DPD1707, DP1675, and DPD1718 were deposited on 13 February 1997 with the American Type Culture Collection ("ATCC"), international depository. The strains are designated as ATCC 98328, ATCC 98329, and ATCC 98330 respectively.

Additional strains of detector cells useful in the determination of CPC sites of action are given in Table 1, below.

COR-AAD1-00001946

TABLE 1

Strain	Genotype
CU847	*ilvB2101 ara thi Æ(pro-lac)*, [from E. Umbarger, Purdue University]
DPD1678	as CU847 but Δ*relA251::Kan*
DPD1675	as CU847 but *tolC*::miniTn*10*
DPD1690	as DPD1675 but Δ*relA251::Kan*
DPD1012	as DPD1690 but *igm* (improved growth on minimal medium, an unmapped mutation)
DPD1013	as DPD1690 but *igm* (improved growth on minimal medium, an unmapped mutation)
RK4988	*F' araD139 Æ(lacIPOZYA)U169 strA thi? non gyrA Æ(ilvB-uhpA)2089* [From R. Kadner via Ying-Ying Chang, University of Illinois]
DPD1692	as *Kan* resistant, lactose non-utilizing
DPD1680	as RK4988 but *tolC*::miniTn*10*
DPD1693	as DPD1680 but Δ*relA251::Kan*
FD1062	Hfr *ara-14 ilvI614 ilvH612* λ- *glyA18 relA1 spoT1 ilvB619 bglR20 rbs-5::*Tn*5 ilvG468* (IlvG+) *thi-1*, [from Bob Lawther, University of South Carolina, reference: Babczinski, P. and T. Zelinski, (1991) *Pestic. Sci.* 31: 305-323]
DPD1682	as FD1062 but *tolC*::miniTn*10*
DPD1010	as DPD1682 but *igm* (improved growth on minimal medium, an unmapped mutation)
DPD1011	as DPD1682 but *igm* (improved growth on minimal medium, an unmapped mutation)
DPD1707	as JC7623 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1708	as JC7623 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1715	as FD1062 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1716	as CU847 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1719	as DPD1693 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1728	as DPD1010 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1729	as DPD1011 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1730	as DPD1682 but *lacZ::[recA'ϕluxCDABE cat]*
JC7623	*thr-1 ara-14 leuB6 d(gpt-proA)62 lacY1 tsx-33 supE44 galK2 l- rac- hisG4 (Oc) rfbD1 mgl-51 rpsL31 kdgK51 xyl-5 mtl-1 argE3 (oc) thi-1 qsr' sbcC201 sbcB15 recB21 recC22* [from F. Valle, Insituto de Biotechnologia, UNAM, Cuernavaca, Mexico]
DM800	*F-, metA28, lacY1 or Z4, thi-1, xyl-5 or -7, galK2, tsx-6* [from Brooks Low, Yale University, fully described in *J. Bacteriol.*, 12:886, (1972)]
DM803	as DM800 but *lexA1* [from Brooks Low, Yale University, fully described in *J. Bacteriol.*, 12:886, (1972)]
DPD1714	as DM800 but *lacZ::[recA'ϕluxCDABE cat]*
DPD1709	as DM800 but *lacZ::[recA'ϕluxCDABE cat]*

COR-AAD1-00001947

Construction of libraries

Chromosomal DNA isolated from *E. coli* W3110 [B. Bachmann, in E. coli and *Salmonella typhimurium*; Cellular and Molecular Biology (Niedhardt et al., Eds., pp 1190-1220, American Society of Microbiology, Washington, D.C. (1987))] was partially digested with restriction enzyme *Sau*3A1 and size fractionated on agarose gels. Fractions of two size ranges (average sizes of approximately 2.5 and 4.0 Kbp) were ligated to pBR322 (0.11 pmol) or pUC18 (0.11 pmol) that had previously been digested with restriction enzyme *Bam*HI and treated with calf intestinal alkaline phosphatase. The molar ratio of chromosomal DNA to vector in each of the ligation reactions was approximately 0.2:1. The ligation products were used to transform ultracompetent *E. coli* XL2Blue (Stratagene) to AmpR. Pooled transformants ($>10^5$ for each transformation) were used to isolate plasmid DNA.

EXAMPLE 1

CONSTRUCTION OF DETECTOR CELL STRAINS

Construction of stain DPD1675

Construction of strain DPD1675 containing *tolC* and *ilvB* proceeded by the manipulation of CU847, an *E. coli* strain, possessing the *ilvB* mutation and having the genotype [*ilvB2101 ara thi Δpro lac*] from H.E. Umbarger, Purdue University. A $P1_{vir}$ phage stock was grown on strain DE112 [*tolC*::miniTn10; fully described in Van Dyk et al., *Applied and Environ. Microbiol.* 60:1414-1420 (1994), isogenic with RM443 described in B. Bachmann, in E. coli and Salmonella typhimurium; Cellular and Molecular Biology (Niedhardt et al., Eds., pp 1190-1220, American Society of Microbiology, Washington, D.C. (1987))] by standard methods as described in J. H. Miller, *Experiments in Molecular Genetics*, (1972) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp *201-205*. The phage stock was mixed as described by Miller (*supra*) with strain CU847, tetracycline resistant recombinants were selected and a recombinant that was unable to grow on MaConkey agar (Miller, *supra*) and hence displaying a bile salt sensitive phenotype was designated DPD1675.

Construction of DPD1707

DPD1707 containing the *recA-LuxCDABE* gene fusion was constructed as follows.

Plasmid *precALux3* was isolated from strain DPD2794 (fully described in U.S. 5683868), isogenic with RM443 described in B. Bachmann, in E. coli and Salmonella typhimurium; Cellular and Molecular Biology (Niedhardt et al. Eds., pp 1190-1220, American Society of Microbiology, Washington, D.C. (1987))]. The promoter region (*recA*) was amplified by PCR using primers 1,2 (SEQ ID NOS:1 and 2 respectively).

COR-AAD1-00001948

Primer 1:

5' ACT Ttt aagctt AGA GAA GCC TGT CGG CAC 3' recA-upperH (SEQ ID NO:1)

Primer 2:

5' AGC TTT ggatcc CGC TTT CTG TTT GTT TT 3' recA-lowerB (SEQ ID NO:2)

The resulting product was digested with *BamH1* and *Sal1* and was mixed with a similarly digested pJT205 plasmid (formerly called pCGLS205) containing the *Photorhabdus luminescens luxCDABE* gene complex, fully described by (Rosson, R. A., PCT International Application WO 93/03179 (1993)). After ligation the mixture was transformed into strain DH5 (ATCC) and ampicillin resistant colonies were selected. The colonies were screened for bioluminescence. One such bioluminescent transformant was designated DPD1657. Plasmid pRecALxx1, isolated from strain DPD1657, was digested with *Pst1* and *Ecor1*. This digested plasmid was mixed with similarly digested pBrint.CM. plasmid, [from F. Valle, Insituto de Biotechnologia, UNAM, Cuernavaca, Mexico; Balabas et al., *Gene* 172:65-69 (1996)]. After ligation chloramphenacol resistant transformants were recovered in strain DH5. The resistant colonies were screened for a bioluminescent phenotype. One such transformant was termed DPD1696. The plasmid in this strain was called pDEW14. Plasmid pDEW14 contained a fusion of the *recA* promoter to the *Photorhabdus luminescens luxCDABE* gene complex. This fusion is situated within *lacZ* coding sequences. Plasmid pDEW14 was isolated from DPD1696 and the DNA was introduced into strain JC7623 [B. Bachmann, in *E.coli* and *Salmonella typhimurium*; Cellular and Molecular Biology (Niedhardt et al. Eds., pp 2466, American Society of Microbiology, Washington, D.C. (1987))] by transformation and chloramphenacol resistant colonies were selected. Colonies that were ampicillin sensitive and *lacZ* negative (i.e., could not cleave X-gal) designated as DPD1707. It had the *recALux* fusion integrated into the *lacZ* locus of the *E. coli* chromosome and was bioluminescent.

A phage stock of $P1_{vir}$ was prepared on strain DPD1707 ($P1_{vir}$ and the method are fully described in J. H. Miller, *Experiments in Molecular Genetics*, (1972) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp *201-205*.

Construction of DPD1718 (*lacZ::recA'φ'luxCDABE*) proceeded from the strain DPD1692. A $P1_{vir}$ phage stock grown on strain DPD1707 was mixed as described by Miller (*supra*) with strain DPD1692; chlorampehnicol resistant recombinants were selected. One such recombinant that also displayed bioluminescence was designated DPD1718.

COR-AAD1-00001949

The same transductional methods were used to create strains DPD1715, DPD1730, DPD1728, DPD1729, DPD1716, DPD1708, DPD1719, DPD1709, and DPD1714's reported in Table 1.

EXAMPLE 2

ENHANCED BIOLUMINESCENCE IN THE PRESENCE OF THE DNA DAMAGING AGENT MITOMYCIN C

To test the detector cells for sensitivity to genotoxic substances, strains DPD1715 and DPD1730 were grown to mid-logarithmic phase in LB medium. Strain DPD1715 and DPD1730 differ only in that DPD1730 is *tolC* while DPD1715 is *tolC*+. To the medium was added varying concentrations of Mitomycin C, ranging from 0 to 20 ug/mL. The kinetics of light emission after introduction of the genotoxin was monitored using a microtiter plate format luminometer as described previously (Van Dyk et al., *Applied and Environmental Microbiology* 60, 1414, (1994)) except that the temperature was controlled at 37 °C.

The kinetics are presented in Figures 4 (a) and (b), while a dose-response curve is shown in Figure 5. For DPD1730 [tolC-] (figure 4(b)), high concentrations of mitomycin C, (here above 2.5 ug/mL) will either inhibit or prevent the increase in light production attributable to growth. At low concentrations (below 2.5 ug/mL) the cell does not acknowledge the presence of the chemical. At intermediate concentrations (0.63 and 0.31 ug/mL) the cell responds to the damage by inducing the SOS response as conveniently monitored by the large increase in bioluminescence, above the mock treated control, that develops as a function of time. In contrast, the *tolC*+ strain, DPD1715 (figure 4(a)) was much less sensitive to mitomycin C (see figure 5). Induction was observed throughout the entire range of mitomycin C concentrations. Inhibitory effects were not observed.

As expected, induction is *lexA*-dependent as indicated by kinetic and dose-response curves of the fusion introduced by transduction into an isogenic pair of strains differing in the *lexA* gene that controls the SOS response to DNA damage. The *LexA* repressor encoded by *lexA* must be cleaved if the SOS response is to be activated [Walker, 1996, in *E. coli* and *Salmonella typhimurium*; Cellular and Molecular Biology (Niedhardt et al. Eds., pp 1400-1416, American Society of Microbiology, Washington, D.C. (1996))] The *lacZ::recA'φ'luxCDABE* fusion was introduced by transduction into strains DM800*(lexA*$^+$*)* and DM803 (non inducible *lexA* due to a non-cleavable repressor gene product [Mount et al., *J. Bact.*, 112: 886, (1972)] using a phage stock grown on strain DPD1707. One chlorampehnical transductant from each cross that was bioluminescent was saved and designated as DPD1714 *(lacZ::recA' φ'luxCDABE lexA*$^+$) and DPD1709

COR-AAD1-00001950

(lacZ::recA' ϕ'luxCDABE lexA). Mitomycin C treatment caused enhanced light production from strain DPD1714 but not from strain DPD1709 indicating that the treatment elicited a true SOS response since the inability to cleave the LexA repressor prevented the bioluminescent increase.

EXAMPLE 3

DECREASED BIOLUMINESCENCE AFTER TREATMENT WITH 4 CROP PROTECTION CHEMICALS

Example 3 illustrates the effect of four different crop protection chemicals on the light output of the detector cells.

The detector strains were growth to mid-log arithmic phase as described above treated with sulfometuron methyl, methyl viologen, glyphosate and 2,4-dichlorophenoxyacetic acid (2,4-D) according to the following conditions where all cultures were maintained at 37 °C and tests were run over a time period of 0-90 min.

Strain DPD1718 was exposed to 2,4-D over a concentration range of 78-5000 ug/mL. Modulation in bioluminescence is seen in Figure 6.

Strain DPD1715 was exposed to glyphosate over a concentration range of 2.5-2500 ug/mL. Modulation in bioluminescence is seen in Figures 8(a) and (b).

Strain DPD1718 was exposed to sulfometron methyl over a concentration range of 24-1800 ug/mL. Modulation in bioluminescence is seen in Figure 7.

Strain DPD1730 was exposed to methyl viologen over a concentration range of 0.75-48 ug/mL. Modulation in bioluminescence is seen in Figure 9.

As is evident from the figures illustrating kinetic data, all crop protection chemicals gave a dose-dependent decrease in light emission from the detector strains.

EXAMPLE 4

tolC ENHANCES THE RESPONSE TO SULFOMETURON METHYL AND MITOMYCIN C

Example 4 illustrates the effect of the *tolC* mutation on the detector cell response to sulfometuron methyl and Mitomycin C. Detector cells were growth to mid-logarithmic phase as described above. Detector cells were exposed to a range of crop protection chemical and bioluminescence was measure at 80 min post-exposure. Strain DPD1715 (*tolC*+) and strain DPD1728(*tolC*) were exposed to Sulfometuron methyl over a concentration range of 0.006 to 0.4 ug/mL and responses are compared in Figures 10(a-c). Similarly *tolC*+ and *tolC* strains were exposed to concentrations of 0-20 ug/mL of mitomycin C and responses are compared in Figure 5. From the data it is clear that the *tolC* mutation has the effect of enhancing sensitivity of the detector strain to the CPC.

COR-AAD1-00001951

The bioluminescent assay demonstrating the efficacy of the *tolC*- strains was confirmed using disk diffusion assays [Stephens et al., *PNAS* 72, 4389, (1975); LaRossa R.A. and J.V. Schloss, *J.Biol.Chem.*, 259, 8753 (1984)] (data not shown).

EXAMPLE 5

INDICATIONS OF SITE OF ACTION OF OXIME CARBAMATE OC, GLYPHOSATE AND THIENYLALANINE BY NUTRITIONAL REVERSAL

Example 5 demonstrates the use of a detector cell having to identify the site of action of the oxime carbamate OC, glyphosate and thienylalanine.

Strain DPD1718 was grown in minimal medium or LB medium at 37 °C and was exposed to either 3-(2-thienyl)-L-alanine or the oxime carbamate profungicide compound OC over a variety of concentrations. Concentrations of 3-(2-thienyl)-L-alanine ranged between 0 and 100 ug/mL for growth in both minimal and LB medium.

Concentrations of compound OC ranged from 0 to 100 ug/mL for cells grown in minimal medium and from 0 to 1000 ug/mL for cells grown in LB medium. Concentrations of glyphosate for cells grown in both minimal and LB Rich medium ranged between 0 and 5000 ug/mL.

As seen by the data in Figures 11(a), (b), 12(a), (b) and 13(a), (b), treatment of the detector cell with these compounds caused a decrease in light production in minimal medium, whereas in rich LB medium, light emission was not decreased.

Auxanography was used to determine the nutrient responsible for obviating the light-inhibiting effect of these experimental crop protection chemicals.

Auxanographic protocol:

Strain DPD1718 was grown to mid-logarithmic phase in minimal medium. A total of 11 individual pools are created which represent various bases, vitamins and amino acids (Tables 2, 5). Each pool shares one component with one other pool. Each pool was added singly into a microtiter well in duplicate. Composition of the pools makes metabolic sense as described in Davis, R. W., D. Botstein And J. R. Roth, A Manual For Genetic Engineering: Advanced Bacterial Genetics. X+254p. Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., U.S.A., 1980, see pp. 206-208. Minor changes to the described 11 pools from that of Davis et al. were made to accommodate metabolic differences between *Salmonella typhimurium* and *E. coli*. The pool components follow; their concentrations in the pools are as described by Davis et al., *supra*:

COR-AAD1-00001952

pool-components

1-adenosine, histidine, phenylalanine, glutamine, thymine
2-guanosine, leucine, tyrosine, asparagine, serine
3-cysteine, isoleucine, tryptophan, uracil, glutamate
4-methionine, lysine, threonine, aspartate, DAP
5- isoleucine, valine, proline, arginine, glycine
6-adenosine, guanosine, cysteine, methionine, isoleucine
7-histidine, leucine, isoleucine, lysine, valine
8-phenylalanine, tyrosine, tryptophan, threonine, proline, PABA, DHBA
9-glutamine, asparagine, uracil, aspartate, arginine
10-thymine, serine, glutamate, DAP, glycine
11-pyridoxine, nicotinate, biotin, pantothenate, alanine

The volume added per well was 10 μL of an individual pool plus 40 μL of fresh minimal media which has been supplemented with a predetermined inhibitory concentration of a given compound. Fifty microliters of midlog arithmic cells grown in minimal media were added to the microtiter well. The microtiter plate was then placed into a ML2000 luminometer and the bioluminescence of the sample was monitored over time. The data was compared to the sample of cells not challenged with the compound. The pool that enabled the cell to overcome the inhibitory effects of the compound, as determined by a restoration in bioluminescence, was determined to be the pool that contained the component that reversed the dampening effect of the compound. Finally, individual components of the pool were tested to determine the specific ingredient responsible for restoration of bioluminescence.

Nutritional Reversal of Compound OC Action

Table 2 records the data from the interaction of the compound OC with DPD1718. A lights-off response, (approximately 1000 fold diminution in light output) is observed on minimal but not rich media. Auxanography (Table 3) indicated that the presence of pool 6, containing cysteine, was able to prevent the lights-off response.

COR-AAD1-00001953

TABLE 2

Cycle 10:

Cmpd OC	ADDITION	%LUM
-	-	=100
+	Pool 1	0.051
+	Pool 2	0.13
+	Pool 3	0.051
+	Pool 4	0.042
+	Pool 5	0.022
+	Pool 6	89
+	Pool 7	0.070
+	Pool 8	0.095
+	Pool 9	0.028
+	Pool 10	0.11
+	Pool 11	0.13
+	-	0.10

* Medium: E+B1+glucose+glycine+proline+uracil 37 °C/DPD1718/compound OC at 25 ug/mL

A weak prevention of the light-off response was seen at some time points with pool 3 (data not shown). Thus the individual components of pools 3 and 6 were tested for their ability to prevent the lights off response:

5

TABLE 3

Cmpd. OC	ADDITION	%LUM
Pool 6 components		
-	-	=100
+	Pool 6	90
+	Ade	0.053
+	Gua	0.038
+	Cys	94
+	Met	0.034
+	Ile	0.027
+	-	0.031

COR-AAD1-00001954

Cmpd. OC	ADDITION	%LUM
Pool 3 components		
+	Pool 3	0.048
+	Cys	94
+	Ile	0.027
+	Trp	0.046
+	Glut	0.13
+	Ura	0.027

It is evident that the reversing agent is cysteine, the common component between the 2 pools. Cysteine could exert it effects either by reversing a chemical-induced auxotrophy or by forming an adduct with the oxime carbamate or with its hydrolysis product. To attempt to distinguish between these 2 models, the stereoisomers of cysteine were tested as reversing agents:

TABLE 4

	%LUM		
	Inhibitor:		
Supplement	Cmpd. OC	Analog	None
-	.04	.06	=100
L-cys	45	45	
D-cys	69	75	

It is apparent that D-cysteine can also reverse compound OC effects. Applicants then tested and found that D-cysteine satisfies the auxotrophy of 2 *cysD* strains, a *cysA* mutant, a *cysH* strain and a *cysB* mutant. The data suggests that either the D-cysteine is not optically pure or that *E. coli* has the ability to interconvert the two stereoisomers.

Additional experiments, showed that a free sulfhydryl group, present on a variety of compounds including 2-mercaptoethanol, was responsible for quenching the profungicide inhibition of bioluminescence (data not shown). The finding indicated that compounds with free SH groups obviate compound OC action, suggesting that L-cysteine biosynthesis is not compromised by compound OC action. Thus, nutritional reversal experiments can suggest agents capable of obviating inhibitors action. Such experimentation must be buttressed with further investigation.

COR-AAD1-00001955

Nutritional Reversal of 2-Thienylalanine Action

Nutritional reversal of thienylalanine inhibition was accomplished in a similar fashion as with compound OC. Pools were created as described above. Reversal of thienylalanine inhibition by pools 1 and 8 was observed as shown in Figure 14. The reversal can be attributed to phenylalanine, but not to the other components of the pools.

Nutritional Reversal of Glyphosate Action

Nutritional reversal of glyphosate inhibition was accomplished in a similar fashion as with compound OC and thienylalanine. Reversal of glyphosate inhibition only by pool 8 was observed as shown in Figure 13(a-b) and Table 5.

TABLE 5

Cycle 10:

Glyphosate	ADDITION	%LUM
-	-	=100
+	Pool 1	19
+	Pool 2	19
+	Pool 3	16
+	Pool 4	18
+	Pool 5	29
+	Pool 6	14
+	Pool 7	23
+	Pool 8	105
+	Pool 9	19
+	Pool 10	19
+	Pool 11	19
+	-	18

*Medium: E+B1+glucose+glycine+proline+uracil 37 °C/DPD1718/Glyphosate at 600 ug/mL

Reversal by pool 8 only is indicative of a block in the common aromatic amino acid pathway (Davis et al., *supra*, p. 208). Thus the target pathway of glyphosate is accurately identified by this method.

Nutritional Reversal of Sulfometuron Methyl (SM) Action

The bioluminescence of Strain DPD1718 growing in minimal E medium supplemented with thiamine, glucose and uracil was inhibited by sulfometuron methyl in a dose-dependent manner (see Figure 15(a) in which the concentrations tested were 1800, 900, 450, 225, 113, 61, 31 and 0 ug/mL).

COR-AAD1-00001956

Nutritional reversal of SM inhibition caused by administration of 250 ug/mL of the herbicide is depicted in the bottom panel of Figure 15b. The most effective pool at reversing the inhibition was pool 7 composed of the three branched chain amino acids, lysine and histidine. Reversal by this pool is consistent with the known site of SM action. Incomplete reversal to a lesser degree and with greatly delayed kinetics was seen with pools 1 (adenosine, histidine, phenylalanine, glutamine, thymine), 9 (glutamine, asparagine, uracil, aspartate, arginine) and 11 (pyridoxine, nicotinate, biotin, pantothenate, alanine).

EXAMPLE 6

IDENTIFICATION OF SITE OF ACTION OF SULFOMETURON METHYL AND GLYPHOSATE BY GENETIC TITRATION

Characterization of GP and SM Sensitivity by Analysis of Their Interactions with Selected Mutants

Glyphosate is expected to be inhibitory to *E. coli* K12. In contrast, sulfometuron methyl is not inhibitory towards *E. coli* due to the presence of the refractile ALS I isozyme encoded by *ilvBN*. Both compounds inhibit an *ilvB* mutant of *E coli*. Zone of inhibition assays demonstrate that introduction of a *relA* null allele which prevents the mounting of the stringent response to amino acid starvation into the *ilvB* mutant results in hypersensitivity to GP but does not change the response to sulfometuron methyl.

SM resistant meromultiploids

Strain DPD1675 expresses a single SM-sensitive ALS, ALS III, encoded by *ilvIH* due to the presence of an *ilvB* mutation that prevents production of the wild type ALS I naturally resistant to SM. It also contains a *tolC* mutation that prevents an efflux pump mediated expulsion of SM from the cytoplasm. These two mutations together create a strain quite sensitive to SM. Applicants determined the MIC of strain DPD1675 to be 3 ug/mL in minimal medium containing plates.

The genetic titration of the herbicide sulfometuron methyl was performed in the strain DPD1675 (Table 1). This strain harbors a mutation in the *ilvB* allele and *tolC*, which result in rendering the cell sensitive to the herbicide and more permeable to hydrophobic compounds. The MIC of the herbicide in this strain was determined to be 3 μg/mL following 1 day of growth on minimal plates.

Frozen competent cells of DPD1675 (prepared by the method of Nishimura et al., *Nucleic Acids Research* 18:6169 (1990)) were transformed with 0.5 μL of plasmid DNA from two different *E. coli* libraries, one pBR322-based and one pUC18-based. The transformation mix was washed one time and resuspended in 1 X E [Davis et al., *supra*, pp. 202-203] before being plated to minimal E plates supplemented with thiamine, proline, glycine and glucose at

COR-AAD1-00001957

standard concentrations that can be found in Davis et al., *supra,* pp. 201-210, and Miller, 1972, supra, p. 432. The selection for the desired clones was Ampicillin (100 μg/mL) and Sulfometuron methyl (9 μg/mL) which was also included in the media.

Strain DPD1675 was transformed with 2 libraries containing random fragments of the *E. coli* chromosome ligated into either pBR322 or pUC18. Plasmids from 19 isolates, when reintroduced into strain DPD1675, conferred resistance to SM (Table 6).

TABLE 6

Selection			Plasmid:		
Agent	ug/mL	Host	Vector	Name	Inserted Genes
SM	9	DPD1675	pBR322	pDEW21	*tolC*
				pDEW18	*tolC*
				pDEW22	*tolC*
				pDEW123	*tolC*
				pDEW24	*tolC*
				pDEW25	*tolC*
				pDEW20	*'emrDivbL ilvBN uhpA uhpB uhpC uhpT'*
				pDEW26	*glvA glvB glvG yidL yidK yidJ yidI yudH yidG tidF emrD evbL ilvBN uhpA uhpB uhpC uhpT'*
				pDEW27	*yidG' yidF emrD ivbL ilvBN uhpA uhpB'*
				pDEW28	*'yidF emrDivbL ilvBN uhpA uhpB uhpC'*
				pDEW19	*ilvIH*
SM	9	DPD1675	pUC18	pDEW15	*'emrDivbL ilvBN uhpA uhpB'*
				pDEW34	*'yidF emrDivbL ilvB ilvN*
				pDEW29	*yidG' yidF emrD ivbL ilvBN uhpA uhpB uhpC'*
				pDEW30	*ilvBN'*

COR-AAD1-00001958

				pDEW31	*yidI' yidH yidG yidF emrD ivbL ilvB ilvN*
				pDEW32	*'emrDivbL ilvBN uhpA uhpB uhpC'*
				pDEW17	*ilvIH*
GP	560	DPD1692	pUC18	pDEW110	*'yhhR yhhS yhhT yhhU'*
				pDEW112	*yhhQ yhhR yhhS yhhT yhhU'*
GP	560	DPD1692	pBR322	pDEW114	*'f589 o230 serC aroA*
				pDEW115	*'f307 o552 criR f461 ybeG cspE ybeI ybeH'*
				pDEW116	*'f415 f108 glk o418 o108'*

The response of these transformants to bile acids was scored; bile acid sensitivity indicates a *tolC*$^-$ phenotype while bile acid resistance indicates a *tolC*$^+$ phenotype. The introduction of 6 plasmids into DPD1675 resulted in a bile salt resistant isolate suggesting that these strains are *tolC*$^+$/*tolC* heterozygotes and that *tolC*$^+$ was dominant. This supposition was confirmed by sequencing both strands of each insert The primers used for sequencing the inserts into pBR322 were:

GCC ACT ATC GAC TAC GCG (SEQ ID NO: 3)

CTG TGG CGC CGG TGA TGC (SEQ ID NO: 4)

while those used for sequencing the inserts into pUC18 were:

GTA AAA CGA CGG CCA GT (SEQ ID NO: 5)

AGC GGA TAA CAA TTT CAC ACA GGA (SEQ ID NO: 6).

Informatics methods were used to place the obtained sequences onto the *E. coli* genome (Genbank Accession number U00096). These data (Table 6) led Applicants to conclude that each plasmid conferring both SM and bile salt resistance carried *tolC*$^+$.

Strain DPD1675 containing the other 13 plasmids maintained the bile salt sensitive phenotype of the plasmid free host. The nineteen plasmids were also used to transform strain MF2000. MF2000 harbors mutant *ilvBN* and *ilvIH* alleles and so is devoid of ALS activity. As such, the strain is an isoleucine-valine auxotroph. Introduction of the *tolC*$^+$ containing plasmids does not alter the observed auxotrophy of strain MF2000. In contrast, the transformants of strain

COR-AAD1-00001959

MF2000 containing the other 19 plasmids were isoleucine-valine prototrophs. This indicated that they were expressing either ALS I or ALS III and that the plasmids were likely to harbor *ilvBN* or *ilvIH*. These inserts' ends were also sequenced leading to the identification of the genes harbored on the inserts by comparison with the complete primary sequence of the *E. coli* genome (Genbank Accession number U00096; see Table 6). From the data presented we concluded that there were 9 *ilvBN*$^+$/*ilvBN* heterozygotes (see Table 6) and 2 *ilvIH*$^+$/*ilvIH*$^+$ homozygotes in the original collection of SM-resistant meromultiploids.

Selection and characterization of glyphosate resistant meromultiploids

Strain DPD1692 was used for the selection of glyphosate resistant meromultiploids. Applicants determined the MIC of strain DPD1692 to be 3.0 mM (0.56 mg/mL) glyphosate in minimal media plates.

Competent DPD1692 cells transformed with 0.5 uL (of either library one (pBR22-based; 0.0375 ug DNA) or library three (pUC18-based, 0.15 ug, DNA) yielded 10^4 transformants thus having a frequency of about 3×10^5 and 0.7×10^5 transformants per microgram DNA, respectively when plated on rich media plus ampicillin. Approximately 2×10^3 cells were plated per selective plate which consisted of minimal E media amended with glucose, glycine, proline, uracil and thiamine as well as the selective agents Ampicillin (100 ug/mL) and glyphosate (3.3 mM = 0.56 mg/mL). Glyphosate resistant meromultiploids appeared throughout a 96 h incubation at 37 °C.

Glyphosate selection resulted in the isolation of 7 transformants from the pBR#22 library, and 18 transformants from lthe puC18 library (Table 6). Retransformation of the original host strain DPD1692 with the 25 plasmids conferred glyphosate resistance, insuring that the plasmid conferred the resistance phenotype. Presence of *aroA*, the gene that encodes the known target of glyphosate, was screened by transforming an auxotrophic *aroA* mutant *E. coli* strain AB2829 with the plasmids that conferred glyphosate resistance. Prototrophic transformants indicated that the *aroA* gene was resident on the glyphosate-resistance conferring plasmid. Two of the 25 isolated plasmids restored growth of AB2829 on minimal media. The presence of *aroA* on these two plasmids was confirmed by sequence analysis (see Table 6) using the primers described above. These two plasmids conferred the strongest resistance to glyphosate when the 25 glyphosate resistance plasmids were compared in strain DPD1692 (data not shown).

Of the remaining 23, non-*aroA* containing plasmids, three distinct inserts conferring novel glyphosate resistances have been identified by sequence analysis (See Table 6). Group A had 17 members, group B had 4 members and group C had 2 members. Localization of the inserts relative to the *E. coli* genome map was

COR-AAD1-00001960

accomplished by FASTA and BLAST (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, Wisc.) scans of the databases. The group A inserts, mapping to 77.8 minutes on the *E. coli* chromosome, contain two previously sequenced ORF's, *yhhS*. and *yhhT*. The predicted protein structures for both of these ORF's contain domains predicted to span the membrane and they are thus likely to be permeases. The 4 members of the second group, B, mapping to the 14.2 minute region, contain an uncharacterized region of the *E. coli* chromosome. The third group of two inserts maps to the 52.3 minute region of the *E. coli* chromosome. Sequence analysis has revealed genes related to carbohydrate metabolism are contained in this insert. ORF's identified to date include previously undescribed component IIC and IIB PTS enzymes. The predicted amino acid sequence encoded by these two genes share homology with component IIC and IIB enzymes encoded by a fructose-like PTS operon (Reizer, J. A. et al., *Microbiol.* 141, 961, (1995)). In addition to new PTS enzyme II homologues, this insert contains the single *E. coli* glucokinase gene, *glk*.

EXAMPLE 7

IDENTIFICATION OF MITOMYCIN C - RESISTANCE GENES

Example 7 illustrates the identification and isolation of genes having resistance to Mitomycin C that may be used in the construction of a detector cell containing the *recA-LUX* fusion.

Genetic Titrations of Mitomycin C Action

Mitomycin C (a DNA damaging agent) inhibits colony formation of a $lexA^{ind}$ mutant of *E. coli* ,DM803, with a MIC of 3 ug/ml. This MIC is 5 x lower than that of the $lexA^{+}$ strain DM800. Competent DM 803 cells were transformed with a pBR322 library and pUC18 library. Colonies were selected by plating on LB Amp 150 plates in the presence of 15μg/ml of mitomycin C (3X the MIC) at 37°C. After 2 days, 4 colonies appeared on plates where cells were transformed with the pUC18 based *E. coli* library. These 4 colonies were picked; all plates were incubated for another 5 days without the appearance of further colonies. Plasmids were purified from the four colonies and named plexA3.1, 3.2, 3.3 and 3.4. Reintroduction of each plasmid by retransformation of strain DM803 showed linkage between the ampicillin resistance and mitomycin C resistance indicating that the mitomycin C resistance in each case was a plasmid encoded trait.

Plasmid purification from DM803 host was difficult - poor yield and degradation of samples upon storage were observed. Thus plasmids were transferred to RFM443 for routine purification of plasmid template for sequencing.

COR-AAD1-00001961

Sequences obtained:

(1) The two ends of the LexA3.1 insert were sequened. LexA3.1 forward-primed sequence maps to region 291 out of 400 (minute 18) as defined by the Blattner *E. coli* sequencing project - *glt* region, with putative fimbrial chaperone gene, and *yhcA* while the reverse-primed sequence maps to region 76 out of 400 - this region contains *dacC* (a penicillin binding protein) and *deoR* (deoxyribose operon repressor) as well as several open reading frames encoding proteins of unknown function. Since these are two non-continuous regions of the chromosome, the selected plasmid is a chimera of 2 non adjacent fragments presumably fused during the library construction.

(2) LexA3.2 forward- and reverse-primed sequences map to region 76 out of 400 in *E. coli* database. This region contains *dacC* (a penicillin binding protein) and *deoR* (deoxyribose operon repressor) among the8.6kb found in the insert. TheLex A3.3 reverse-primed sequence: maps to the same region as LexA3.2; while the forward-primed sequence is unavailable at this time.

(3) Similarly the Lex A3.4 reverse-primed sequence also maps to same region as LexA3.2 and appears to contain the same junction as Lex3.3; again forward-primed sequence is unavailable.

Thus it appears that a gene or genes within the region 76 of the chromosome of *E. coli* confers mulitcopy resistance to mitomycin C. It is obvious that division of the segment and retesting will define a novel gene or genes responsible for mitomycin C resistance.

These genes and open reading frames included on pLexA3.2 are: *o127,o371,f210,dacC* = penicillin-binding protein 6 precursor, *deoR*=deoxyribose operon repressor, *f198, o410=mdfA, f94, f262, f402 & o178.*

Further experiments were performed with a strain proficient in the SOS response since the non-inducible allele could compromise a subset of potential resistance mechanisms. Strain RFM443 is $lexA^+$ and thus was used for genetic titration with mitomycin C. The MIC for mitomycin C for this strains was determined to be between 1 and 3 ug/ml on rich, solidified LB medium. Thus LB solidified with agar and supplemented with 6 ug/ml of mitomycin C and 150 ug/ml of ampicillin was used to select mitomycin resistant clones after transformation of strain RFM443 with *E. coli* genomic libraries constructed in either pBR322 or pUC18. Plasmid DNA purified from such resistant isolates was reintroduced into RFM443 to demonstrate if the resistance was a plasmid encoded trait. DNA sequencing of the vector-insert junctions served to define those sequences that conferred resistance to mitomycin C. Such resistances mapped to 3 sites as defined by sequencing of the inserts (see following).

COR-AAD1-00001962

Site A, isolated 7 times, coincides with the site at minute 18 present in the pLexA3 plasmids (above). This set of plasmids, however, (see Table 7 below) demonstrates that one specific gene, *o410* (*mdfA*), a recently described multidrug transporter, in the minute 18 region is capable of conferring resistance to mitomycin C when present in multiple copies. The inhibitory action of another DNA damaging agent, C0360, is not effected by site A clones.

TABLE 7

Site A Clones.

			RESPONSE TO:	
clone (name)	vector	insert (intact orfs included)	MMC 15µg/ul backcross	C0360 200 µg/ml
1+2/A4	pBR322	*f198,**o410**,f94,f262*	+	--
3+4/B1	pUC18	*f198,**o410**,f94,f262*	--	ND
3+4/B7	pUC18	*f198,**o410**,f94*	--	ND
1+2/B13	pBR322	*deoR,f198,**o410***	ND	ND
1+2/B14	pBR322	***o410**,f94,f262*	+	--
1+2/B16	pBR322	***o410**,**	+	--
1+2/B17	pBR322	*deoR,f198,**o410***	+	--

ND-not determined; + indicates growth; - indicates no growth

Site B mediated resistance, due to high dosage of the minute 43 region, is defined by 16 distinct inserts (see Table 8 below). The only intact gene shared in common by these inserts from site B is *sdiA*, which encodes a positive activator of *ftsQAZ* operon whose products are essential for cell division. Both the *sdiA* and *rpoS* gene products act on distinct promoters of *ftsQAZ*. The inhibitory action of the DNA damaging agent C0360 is lessened by the presence of several site B clones.

TABLE 8

Site B Clones

			RESPONSE TO:	
clone (name)	vector	insert (intact ORFs listed)	MMC 15µg/ul backcross	C0360 200 µg/ml
1+2/B3	pBR322	*uvrY,**	+	+
1+2/B4	pBR322	*uvrY,yecF,**sdiA***	--	--
1+2/B6	pBR322	*yecF,**sdiA**,f250*	+	+
1+2/B7	pBR322	*yecF,**sdiA**,f250*	--	+
1+2/B8	pBR322	*uvrY,yecF,**sdiA***	--	--
1+2/B9	pBR322	***sdiA**,**	--	+

COR-AAD1-00001963

1+2/B10	pBR322	***sdiA***,*f250,yecC,f360*	+	+
1+2/B11	pBR322	*yecF,**sdiA**,f250,yecC*	+	+
1+2/B12	pBR322	*yecC,**	+	--
1+2/B15	pBR322	*uvrY,yecF,**sdiA**,f250*	+	+
1+2/B18	pBR322	*uvrY,yecF,**sdiA***	+	+
3+4/B4	pUC18	*uvrY,yecF,**sdiA***	+	ND
3+4/B10	pUC18	*yecF,**sdiA***	+	ND
3+4/B12	pUC18	*yecF,**sdiA**,f250,yecC*	+	ND
3+4/B14	pUC18	*yecF,**sdiA**,f250,yecC*	+	ND
3+4/B15	pUC18	*yecF,**sdiA**,f250,yecC*	+	ND

* only one end of these inserts were sequenced; ND-not determined; + indicates growth; - indicates no growth

10 Site C inserts were identified by sequencing (see Table 9 below). Arising from the minute 44 region, they have but one common intact gene, *sbmC*. Expression of this gene is induced by microcin B17, a small peptide antibiotic that causes double strand DNA breaks, other DNA damaging agents and entry into stationary phase. A limited sampling indicates that site C clones do not confer cross-resstance to the DNA damaging agent C0360.

TABLE 9

The clones in the *sbmC* region, site C.

			RESPONSE TO:	
clone (name)	vector	insert (intact ORFs listed)	MMC 15µg/ul backcross	C0360 200 µg/ml
3+4/B2	pUC18	***sbmC***	+	ND
3+4/B3,	pUC18	*o148,o73,o122,*	+	ND
3+4/B5,		*o124,o64,f131,*	+	ND
3+4/B6,		*yeeA,**sbmC**,yeeC*	+	ND
3+4/B9,			+	ND
3+4/B11,		(note: these clones are	+	ND
3+4/B13,		not identical)	+	ND
3+4/A1			+	--
3+4/B8	pUC18	*o512,o148,o73, o122,o124,o64,f131, yeeA,**sbmC**,yeeC*	+	ND
3+4/A2	pUC18	***sbmC***	+	--

ND-not determined; + indicates growth; - indicates no growth

COR-AAD1-00001964

EXAMPLE 8

IDENTIFICATION OF ACIVICIN RESISTANCE GENES

Example 8 illustrates the identification and isolation of genes having resistance to Acivicin that may be used in the construction of a detector cell containing the *recA-LUX* fusion.

Genetic Titrations of Acivicin Action

Acivicin inhibits colony formation of *E. coli* strain DPD1675 on minimal medium with a MIC of 1 ug/ml.

Competent DPD1675 cells were transformed with a pBR322 library and pUC18 library each containing random fragments of the *E. coli* chromosome. Colonies were selected by plating on E plates supplemented with glucose, thiamine, proline, 100 ug/ml ampicillin and 3 μg/ml of acivicin (3X the MIC) at 37°C. After prolonged incubation, colonies appeared on the plates. Plasmids were purified from the colonies and named. Reintroduction of each plasmid by retransformation of strain DPD1675 showed linkage between the ampicillin resistance and acivicin resistance indicating that the acivicin resistance in each case was a plasmid encoded trait.

Plasmid purification from the resistant clones provided a plasmid template for sequencing.

Sequences obtained

Class 1 clones come from region 287 at about 43 minutes. 9 in number, their inserts vary from about 2800-5800 bp but all contain an intact *yedA*. Perusal of Table 10 indicates that no other gene in this region is intact in all class 5clones that confer resistance.

There is but a single class 2 clone that maps to region 374 at about 69 minutes. It contains a 1831 bp isert that contains the intact *iciA* and *f76* genes. Relevant clones are identified in Table 10:

COR-AAD1-00001965

TABLE 10

CLONE NAME	TOTAL SIZE	GENES
1C2ii	4820bp	'o271 f569 f60 yedI yedA vsr'
1D3ii	4774bp	'o271 f569 f60 yedI yedA vsr'
3C2	2800bp	'yedI yedA vsr dcm'
L1(3B61)	5340bp	'f569 f60 yedI yedA vsr dcm'
L26(3A49)	5197bp	'o271 f569 f60 yedI yedA vsr'
1A7 (E41)	5072bp	'o271 f569 f60 yedI yedA vsr dcm'
L21(3A43)	5340bp	'f569 f60 yedI yedA vsr dcm'
L51(3D39)	5858bp	'o271 f569 f60 yedI yedA vsr dcm'
L56 (3E54)	3838bp	' f60 yedI yedA vsr dcm'
3C12 (E46)	1831bp	'rpiA f76 iciA sbm'

EXAMPLE 9

IN VIVO INHIBITOR IDENTIFICATION OF THIENYLALANINE LIKE COMPOUNDS

Example 9 illustrates the identification and isolation of genes having resistance to Thienylalanine that may be used in the construction of a detector cell containing the *recA-LUX* fusion. The method proceeded by the isolation of Thienylalanine resistant genes and transformation of an appropriate detector cell.

Genetic Titrations of Thienylalanine Action

Thienylalanine inhibits colony formation of *E. coli* strain DPD1675 on minimal medium with a MIC of 75 ug/ml.

Competent DPD1675 cells were transformed with a pBR322 library and pUC18 library each containing random fragments of the *E. coli* chromosome. Colonies were selected by plating on E plates supplemented with glucose, thiamine, proline, 100 ug/ml ampicillin and 150 μg/ml of thienylalanine (2X the MIC) at 37°C. After prolonged incubation, colonies appeared on the plates. Plasmids were purified from the colonies and named. Reintroduction of each plasmid by retransformation of strain DPD1675 showed linkage between the ampicillin resistance and thienylalanine resistance indicating that the thienylalanine resistance in each case was a plasmid encoded trait.

Plasmid purification from the resistant clones provided a plasmid template for sequencing.

Sequences were obtained that yieled the following plasmid insert structures, shown in Table 11:

COR-AAD1-00001966

TABLE 11

strain name	clone	predicted size	E. coli genes on plasmid
DPD1748	1.2.1******	5544 bp	ydiG(A), aroH, ydiE, f478, f237
	1.2.66	14495 bp	ydiG(A), aroH, ydiE, f478, f237, nlpC, btuDEC, himA, pheT
	1.3.41	12831 bp	o312, o429, o97, ydiD, ppsA, ydiG(A), aroH, ydiE, f478, f237
	1.3.51	12831 bp	o312, o429, o97, ydiD, ppsA, ydiG(A), aroH, ydiE, f478, f237
	1.3.66	12832 bp	o312, o429, o97, ydiD, ppsA, ydiG(A), aroH, ydiE, f478, f237
DPD1749	1.2.6	1820 bp	o245, ygaH
	1.2.11	1828 bp	o245, ygaH
	1.2.56	4196 bp	proX,o88, o305, o245, ygaH
DPD1750	1.3.1******	4356 bp	proX,o88, o305, o245, ygaH
	1.3.6	4188 bp	proX,o88, o305, o245, ygaH
	1.3.11	4186 bp	proX,o88, o305, o245, ygaH
	1.3.26	4139 bp	proX,o88, o305, o245, ygaH
	1.3.31	4188 bp	proX,o88, o305, o245, ygaH
	1.3.36	4188 bp	proX,o88, o305, o245, ygaH
	1.3.46	4188 bp	proX,o88, o305, o245, ygaH
	1.3.56	4188 bp	proX,o88, o305, o245, ygaH
	1.3.71	4188 bp	proX,o88, o305, o245, ygaH

Introduction of plasmid pAHH1 (a multicopy plasmid obtained from R. Baurele, Univeristy of Virginia, that contains *aroH* but not adjacent genes) into strains DPD1675 results in a thienylalanine resistant phenotype indicating that *aroH* is the gene responsible for the multicopy mediated resistance.

The clones isolated by thienylalanine resistance were of two classes. One resistant class defined by overlapping regions of 12 independently isolated clones contained the genes *o245 yagaH* which are located at 59.79 minutes of the *E. coli* chromosome. The other resistant class defined by overlapping regions of 5 independently isolated clones contained a region located at 38.22 minutes on the *E. coli* chromosome including *ydiG(A) aroH ydiE f478.*

Subcloning the individual ORFs from DPD1750 (*o245 yagaH*)

The strain DPD1750 contains the largest insert containing *proX o88 o305 0245 ygaH* (pDEW45). PCR amplification utilizing plasmid DNA isolated from DPD1750 (pDEW45) was used as template to subclone either the *o245* or *ygaH*

COR-AAD1-00001967

ORF independently. The primers were designed with defined restriction sites flanking the ORF to allow one to PCR amplify each individual ORF and then directionally clone into pBR322. The forward direction primers were constructed such that they contain a *BamHI* restriction site (5' GGA TCC 3') incorporated into their nucleotide sequence. The reverse primers have an *EcoRI* restriction site (5' GAA TTC 3') incorporated into their nucleotide sequence. The specific primers for the PCR reaction to isolate the *o245* gene were "o245f" and "o245r", respectively. The specific primers for the PCR reaction to isolate the *ygaH* gene were "ygaHf and "ygaHr", respectively. Their specific nucleotide sequences are described in below.

"o245f"	5' GTT AGC GGA TCC CTA ATT TCA GCC TCA GCC3'	[SEQ ID NO: 7]
"o245r"	5'AGC AGT GAA TTC GTG TAT CGT GCA TCA CTT C3'	[SEQ ID NO:8]
"ygaHf"	5'GTT AGC GGA TCC TAA TAG GGG CAT TCT CCG3'	[SEQ ID NO:9]
"ygaHr"	5'AGC AGT GAA TTC CAG CCG ATA TAG TAA CGA CAG3'	[SEQ ID NO:11]

The PCR reaction was run for 40 cycles of 94°C, 1 min.; 50°C, 1 min.; 72°C, 1 min. and the primer conditions were 100 pmole each. The concentrations and sizes of the PCR products were confirmed by electrophoresis on a 2.0% agarose gel. The *o245* PCR reaction yielded a product of the predicted 1236 bp. Likewise, the *ygaH* PCR reaction amplified a product of the predicted size 661 bp.

The PCR products were purified by column filtration (Microcon) and then enzyme digested. Sequential BamHI and EcoRI restriction digestions were performed on the *o245* and *ygaH* PCR products and pBR322 vector which would serve as the host vector during the ligation. A fraction of the samples were run on a 0.7% agarose gel to determine their DNA concentrations. Ligation reactions were performed over night at 4°C into the host vector pBR322 with either the *ygaH* or *0245* digested PCR products as the insert DNA. Alloquots of the ligated pDNA was transformed into the DH5α host selecting ampicillin resistance at 150 μg/ml and screening for tetracycline sensitivity at 20 μg/ml. Plasmid DNA was isolated from the ampicillin resistant, tetracycline sensitive isolates.

Transformations into DPD1718 for Reverse Genetic Titrations

The *o245* ligated plasmid (pDEW40) or the *ygaH* ligated plasmid(pDEW39), were transformed into DPD1718 selecting ampicillin 150 μg/ml resistance and screening for bioluminescence, chloramphenical (20 μg/ml) and kanamycin resistance (25 μg/ml). The host strain DPD1718 contains a chromosomally integrated *recA::lux P. l.* , making the basal level of bioluminescence very high in terms of RLU's (relative light units). Other plasmids were transformed into DPD1718 to serve as positive and negative

COR-AAD1-00001968

controls, namely pBR322, AH1 (*aroH*) (isolated from CB18), and p*pheA*$_1$ isolated by selecting for a pBR322 clone that complements a *pheA* auxotroph.

Experimental Protocol

Strains were grown at 37°C in minimal E medium supplemented with 0.4% thiamine and 0.4%glucose with shaking. The cultures were diluted into fresh minimal media the next morning and grown to early log. Serial two fold dilutions of a thienylalanine stock solution (10 mg/ml dissolved in dH_20) were performed into minimal media in a microtiter plate with highest concentration @ 100 μg/ml. The bioluminescence was monitored for one hour and the kinetic data collected.

The plotted data is shown in figure 16 as ratios (normalized to the unchallenged cells in terms of light production). Only the last time point is shown at all of the concentrations (usually at approximately 60 minutes).

Zone assays were also performed on the cells according to the following protocol. Strains were grown overnight at 37°C in LB medium supplemented with 150 ug/ml of ampicillin. Cultures were collected by centrifugation prior to resuspension in an equal volume of E medium. 0.1 ml portions were plated in 2.5 ml of E medium amended with 0.7% agar to effect an even lawn of cells on E agar plates supplemented with glucose, proline, uracil and 100 ug/ml ampicillin. Filter disks containing the indicated quantities of compounds were placed upon the lawns. Plates were incubated overnight at 37°C before zones of clearing were measured.

Zone assays confirmed the trend in sensitivity to thienylalanine as illustrated in Table 12.

TABLE 12

zone of inhibition (dia. in mm)

compound	ug/disk	DPD1749 (multicopy *ygaH o245*)	DPD3134 (vector alone)
D,L-allyl glycine	200	24t	26c
1,2,4-triazole	200	15t	17c
5-methyl-D,L-trytophan	200	15t, 23vt	14c, 24t
D,L-5-aza-tryptophane hydrate	200	30c, 41t	38c
Beta-2-thiazolyl-D,L-alanine	200	nz	15c, 22t
mimosine	200	11c	15c
azaleucine	200	14t	23c, 43t
thienylalanine	200	nz	60c
sulfometuron methyl	40	40c	42c

COR-AAD1-00001969

glyphosate	200	18c, 25t	27c
rifampicin	1000	35c	40c

nz=nozone

Relevant genotypes of the strains employed are given in Table 13.

TABLE 13

pDEW40	*o245* insert in pBR322-PCR amplified from DPD1750
pDEW 39	*ygaH.3* insert in pBR322- PCR amplified from DPD1750
pDEW45	*proX o88 o305 0245 ygaH* in pBR322 thienylalanine resistant clone
DPD1748	*ydiG(A) aroH ydiE f478 f237* in pBR322 thienylalanine resistant clone
DPD1749	*0245 ygaH* in pBR322 thienylalanine resistant clone
DPD1750	*proX o88 o305 0245 ygaH* in pBR322 thienylalanine resistant clone
DPD1770	pDEW 39/DH5α
DPD1771	pDEW 39/ DH5α -second transformant(*ygaH.5*) of pDEW39
DPD1772	pDEW40/ DH5α
DPD1773	pDEW 39/ DPD1675
DPD1774	pDEW 39/ DPD1675 -a second transformant (*ygaH.5*) of pDEW39
DPD1775	pDEW40/DPD1675
DPD1777	pDEW 39/DPD1718
DPD1778	pDEW 39/ DPD1718-a second transformant (*ygaH.5*) of pDEW39
DPD1779	pDEW40/DPD1718
DPD1780	pDEW45 /DPD1675
DPD1675	*ilvB tolC*
DPD1718	lacZ::Ω($recA_{E.\ coli}$::$lux_{P.\ l.}$)

EXAMPLE 10
IDENTIFICATION OF MUTAGENIC COMPOUNDS vs. STANDARD AMES REVERTANT ASSAY

Example 10 compares the sensitivity and accuracy of a screen for mutagenic compounds using a *recA-LUX* containing detector cell as opposed to a standard revertant based Ames test.

The example evaluated the mutagenic potential of the submitted test substances in *Salmonella typhimurium* strains TA100, TA1535, TA97a, and TA98 and in Escherichia coli strain WP2 uvrA (pKM101). The *Salmonella* strains are unable to synthesize histidine, an essential amino acid, because of mutations in the genes coding for histidine biosynthetic enzymes. Additional mutations in the defective genes can result in individual *Salmonella* bacteria regaining the ability to synthesize histidine [(Maron, D. M. and B. N. Ames, *Mutation Research* 113,

COR-AAD1-00001970

173-215, (1983)]. *E. coli* WP2 uvrA (pKM101) is unable to synthesize tryptophan due to an ochre mutation in a gene required for tryptophan biosynthesis. *E. coli* reversion mutants may arise either from further changes at the ochre site or from suppressor mutations at a locus in tRNA genes. [Brusick et al., *Mutation Research* 76, 169-190, (1980)]. A trace of histidine or tryptophan in the top agar permits several generations of auxotrophic cell division to fix pro-mutagenic lesions. This results in the formation of a microscopic "lawn" of bacteria. However, only those bacteria that revert to a prototrophic state can continue to divide and form macroscopic colonies. These revertants can be scored by their ability to grow on agar plates deficient in histidine (*S. typhimurium* strains) or tryptophan (*E. coli*). By comparing the number of chemically induced revertants to the number of spontaneous revertants, the mutagenicity of the test substance can be assessed.

Test Substance and Negative Control

Dimethyl sulfoxide (DMSO) was the solvent used for all compounds tested and thus was chosen as the test substance solvent and as the negative control. Dilutions of the test substance were performed at room temperature in DMSO. The negative control was assumed to be stable during this study, and no evidence of instability was observed. Any impurities were not expected to have interfered with the study.

Positive Indicators

Positive indicators included the following: 2-aminoanthracene (2AA), 2-nitrofluorene (2NF), sodium azide (NAAZ), ICR 191 Acridine (ICR 191), and methyl methanesulfonate (MMS). Deionized water was the solvent for NAAZ, ICR 191, and MMS. The solvent for other positive indicators was DMSO. The positive indicators were assumed to be stable in this study and no evidence of instability was observed. Any impurities were not expected to have interfered with the study.

Salmonella Tester Strain Characterization

S. typhimurium tester strains were obtained from Dr. Bruce Ames, Berkeley, CA. The phenotypic characteristics and mutational sensitivities of the *S. typhimurium* strains are summarized in Table 14 as follows:

COR-AAD1-00001971

TABLE 14

Strain	His Gene Locus affected	DNA Repair	LPS	R-factor plasmid	Mutational sensitivity
TA100	*his*G46	ΔuvrB	*rfa*	Present	Base pair substitution
TA1535	*his*G46	ΔuvrB	*rfa*	Absent	Base pair substitution
TA97a	*his*D6610	ΔuvrB	*rfa*	Present	Frameshift
TA98	*his*D3052	ΔuvrB	*rfa*	Present	Frameshift

The deletion in uvrB (a gene which codes for DNA excision repair) increases the bacterial sensitivity to mutagens. [Ames, B. N., F. D. Lee, and W. E. Durston, *Proc. Natl. Acad. Sci. USA* 70, 782-786, (1973)]. The *uvrB* trait is confirmed by demonstrating an increased bacterial sensitivity to ultraviolet light. Because this deletion also extends through a gene needed for biotin biosynthesis, the bacteria require exogenous biotin for growth. The *rfa* mutation causes a partial loss in the integrity of the lipopolysaccharide (LPS) cell wall so that permeability to large molecules is increased (Ames et al., 1973). The presence of the R-factor plasmid (plasmid pKM101 which carries ampicillin resistance as a marker gene) results in an enhancement of an error-prone DNA repair system which is endogenous to these bacteria. [McCann et al., *Proc. Natl. Acad. Sci.* USA 72, 979-983, (1975)]

Although the testing guidelines cited require *S. typhimurium* strain TA1537, the more recently developed TA97a strain was used. TA97 was the recommended replacement for TA1537 and has been demonstrated to be more sensitive to frameshift mutagens.(1,5) However, the reconstructed strain, TA97a, is now routinely used in place of TA97 due to its improved growth properties (personal communication with Bruce Ames and associates).

E. coli WP2 uvrA (pKM101) was obtained from the National Collection of Industrial Bacteria, Torrey Research Station, Scotland. Because tryptophan biosynthesis is blocked by an ochre nonsense mutation, revertants arise as a result of base pair substitution. A second class of mutants may arise as a result of nonsense suppressor mutations in genes coding for tRNAs. Frameshift mutagens are not generally expected to be detected by this strain. (Brusick, et al. *supra)*

Salmonella Tester Strain Storage and Culture

All bacterial strains were stored frozen in 8% DMSO in Oxoid nutrient broth at approximately -70°C. Overnight cultures were prepared by inoculating 20 mL of Oxoid nutrient broth with 0.1 mL of thawed bacterial suspension and incubating at 37°C with shaking. Overnight cultures were then stored on ice until

COR-AAD1-00001972

used for mutagenesis assays. Bacterial strain phenotypes were confirmed on previous stored isolates and not concurrently with each trial. Results for the frozen permanent stock of strains has demonstrated the appropriate responses.

Dose Selection

The highest concentration evaluated was a solution or suspension of the test substance at concentrations of 100 µg/plate. Solubility information was confirmed empirically, as needed. Concentrations were calculated with the assumption that addition of the test substance to the solvent did not change the volume of the resulting solution.

Stability And Concentration Verification

Solutions of the test substance were prepared immediately prior to treatment and were presumed to be stable under the conditions of the study. Treatment and control dosing solutions were not analyzed for concentration, uniformity, or stability. Top agar was not assayed for stability or concentration of the test substance or control articles since this assessment was not considered necessary to achieve the objectives of the study.

Bacterial Mutagenicity Assays

This study consisted of one trial without metabolic activation. Three replicates were plated for each tester strain, test concentration, and condition. Positive indicators and negative controls were included in all assays. Treatments with activation were conducted by adding 0.1 mL of negative control or test substance solution, 0.5 mL of S9 mix, and 0.1 mL of an overnight culture containing at least 1 x108 bacteria to 2 mL of top agar [0.6% agar (w/v) and 0.6% NaCl (w/v)] supplemented with 0.05 mM L-histidine and 0.05 mM D-biotin for S. typhimurium strains or 0.05 mM L-tryptophan for the *E. coli* strain. These components were mixed and poured onto a plate containing 25 mL of Davis minimal agar with dextrose (minimal agar plates, purchased from MOLTOX. Revertant colonies were counted after the individually labeled plates were incubated at approximately 37°C for about 48 hours. When necessary, plates were refrigerated prior to counting.

Statistical Analysis

For each tester strain, the average number of revertants and the standard deviation at each concentration were calculated.

Classification Guidelines

A test substance was classified as positive when: (1) the average number of revertants in any strain at any test substance concentration studied was at least two times greater than the average number of revertants in the negative control; and (2) there was a positive dose-response relationship in that same strain.

COR-AAD1-00001973

A test substance was classified as negative when either: (1) there were no test substance concentrations with an average number of revertants which was at least two times greater than the average number of revertants in the negative control; and (2) there was no positive dose-response relationship.

Test substances, AA, BB, CC, DD, EE, FF, and GG, were evaluated for mutagenicity in Salmonella typhimurium strains TA100, TA1535, TA97a, and TA98 and in Escherichia coli strain WP2 uvrA (pKM101) without an exogenous metabolic activation system (S9).

Concentrations of 6.25, 12.5, 25.0, 50.0 and 100.0 (g/plate were evaluated in comparison to negative (solvent) controls. Under the conditions of these studies, evidence of mutagenic activity was detected in five of the seven test articles assessed. Results are summarized in Table 15 below:

TABLE 15

Lowest conc. in ug/ml:

compound	inducing luminescence(a)	preventing luminescence(a)	for mutation(b)	for growth inhibition(b)
AA	0.3	1.3	50	100
BB	-	5.0	-	----
CC	0.1	0. 6	6	100
DD	0.6	1.3	-	6
EE	0.6	-	6	25
FF	1.3	2.5	100	
GG	0.3	1.3	6	13
(1)	(2)	(3)	(4)	(5)

(a) Assay volumn of 100ul

(b) Plate volume of 28ml

AA, BB, CC, EE, FF, and GG displayed evidence of mutagenic activity. Due to what was judged as test-substance related toxicity, there were insufficient acceptable concentrations to assess the mutagenicity of DD.

Five of the six light inducing compounds (col. 2) are also mutagenic (col. 4) while the threshold for inducing light appears equal or lower (col. 2) than that causing mutation (col. 4). Moreover the lights off assay (col. 3) appears superior to the growth inhibition assay (col. 5).

COR-AAD1-00001974

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
(B) STREET: 1007 MARKET STREET
(C) CITY: WILMINGTON
(D) STATE: DELAWARE
(E) COUNTRY: U.S.A.
(F) ZIP: 19898
(G) TELEPHONE: 302-892-8112
(H) TELEFAX: 302-773-0164
(I) TELEX: 6717325

(ii) TITLE OF INVENTION: A METHOD FOR IDENTIFYING THE SITE OF ACTION OF XENOBIOTIC CHEMICALS

(iii) NUMBER OF SEQUENCES: 10

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: DISKETTE, 3.50 INCH
(B) COMPUTER: IBM PC COMPATIBLE
(C) OPERATING SYSTEM: MICROSOFT WINDOWS 95
(D) SOFTWARE: MICROSOFT WORD VERSION 7.0A

(v) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 60/039,582
(B) FILING DATE: FEBRUARY 28, 1997

(vii) ATTORNEY/AGENT INFORMATION:
(A) NAME: FLOYD, LINDA AXAMETHY
(B) REGISTRATION NO.: 33,692
(C) REFERENCE/DOCKET NUMBER: CR-9952

COR-AAD1-00001975

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACTTTTAAGC TTAGAGAAGC CTGTCGGCAC 30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGCTTTGGAT CCCGCTTTCT GTTTGTTTT 29

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCACTATCG ACTACGCG 18

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTGTGGCGCC GGTGATGC 18

COR-AAD1-00001976

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTAAAACGAC GGCCAGT 17

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCGGATAAC AATTTCACAC AGGA 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTTAGCGGAT CCCTAATTTC AGCCTCAGCC 30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCAGTGAAT TCGTGTATCG TGCATCACTT C 31

COR-AAD1-00001977

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTTAGCGGAT CCTAATAGGG GCATTCTCCG 30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGCAGTGAAT TCCAGCCGAT ATAGTAACGA CAG 33

COR-AAD1-00001978

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 7, line s 17 - 22

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution
AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
US

Date of deposit	Accession Number
13 February 1997 (13.02.97)	98328

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

COR-AAD1-00001979

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 7, line s 17- 22

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution
AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
US

Date of deposit	Accession Number
13 February 1997 (13.02.97)	98329

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

COR-AAD1-00001980

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 7, line s 17 - 22

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution
AMERICAN TYPE CULTURE COLLECTION

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
US

Date of deposit	Accession Number
13 February 1997 (13.02.97)	98330

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer	Authorized officer

COR-AAD1-00001981

CLAIMS:

1. A method for the identification of the site of action of a stressor comprising:

(i) contacting a stressor with a detector bacteria, the detector bacteria comprising a genotoxic-sensitive promoter operably linked to a luminescent reporter gene complex to form a gene fusion that confers a bioluminescent positive phenotype upon the detector bacteria wherein exposure of the genotoxic-sensitive promoter to genotoxic compounds drives heightened expression of the luminescent reporter gene complex producing an increased bioluminescent signal;

(ii) selecting for genotoxic or non-genotoxic stressors capable of inhibiting the growth of the detector bacteria of step (i) by monitoring the growth and light output of the detector bacteria;

(iii) submitting the growth-inhibiting, genotoxic or non-genotoxic stressor selected in to a site of action screen;

(iv) identifying the site of action in the detector bacteria; and

(vi) confirming the site of action in a target organism.

2. The method of Claim 1 wherein the stessor of step (ii) is genotoxic and the stressor of step (iii) is non-genotoxic.

3. The method of Claim 1 wherein the luminescent reporter gene complex is selected from the group consisting of thermostable *lux*, *luxCDABE*, thermolabile *lux*, renella *lux*, and *luc*.

4. The method of Claim 1 wherein the detector bacteria is an enteric bacteria and comprises a first mutation conferring the ability to accumulate crop protection chemicals and a second mutation preventing a global cellular response to the active stressor.

5. The method of Claim 4 wherein the second mutation prevents the global cellular response to amino acid starvation.

6. The method of Claim 1 wherein the stressor is selected from the group consisting of crop protection chemicals, antibacterial chemicals, antiviral chemicals, anticancer chemicals, and anticorrosive chemicals.

7. The method of Claim 6 wherein the stressor is selected from the group of crop protection chemicals consisting of herbicides, pesticides, and fungicides.

8. The method of Claim 7 wherein the stressor is selected from the group consisting of sulfonylurea herbicides, profungicides, fungicides, oxidants, and hormonal herbicides.

9. The method of Claim 8 wherein the stressor is selected from the group consisting of Sulfometuron methyl, glyphostate, oxime carbamates, methyl

COR-AAD1-00001982

viologen, 2,4-dichlorophenoxyacetic acid, 2-thienylalanine, acivicin, and mitomycin C.

10. The method of Claim 1 wherein the genotoxic-sensitive promoter is selected from the group of genotoxic-sensitive promoters controlled by the SOS regulatory circuit.

11. The method of Claim 10 wherein the genotoxic-sensitive promoter is selected from the group of genotoxic-sensitive promoters consisting of *recA*, *uvrA*, *lexA*, *umuDC*, *uvrA*, *uvrB*, *uvrC*, *sulA*, *recN*, *uvrD*, *ruv*, din*A*, *dinB*, *dinD*, and *dinF*.

12. The method of Claim 11 wherein the genotoxic-sensitive promoter is *recA*.

13. The method of Claim 1 wherein the site of action screen is a genetic titration screen or a nutritional reversal screen.

14. A detector cell strain comprising a *recA-LuxCDABE* gene fusion selected from the group consisting of DPD1707, DPD1708, DPD1715, DPD1716, DPD1718, DPD1719 DPD1728, DPD1729, and DPD1730.

15. A non-bioluminescent parent strain selected from the group consisting of DPD1678, DPD1675, DPD1690 DPD1012, DPD1013, DPD1692, DPD1680, DPD1693, DPD1682 DPD1010 and DPD1011.

16. A method for the identification of thienylalanine-like compounds comprising:

(i) contacting a compound suspected of having thienylalanine-like activity with a first detector bacteria, and with a second detector bacteria,

the first detector bacteria comprising:

a) an expressible *luxCDABE* gene fusion; and

b) the normal cellular complement of a gene selected from the group consisting of *pheA* and *aroH* wherein the gene is not overexpressed; and

the second detector cell comprising:

a) an expressible *luxCDABE* gene fusion; and

b) multiple copies of an expressible gene selected from the group consisting of *pheA* and *aroH* wherein the expressible gene is overexpressed;

the first and second detector cells being isogenic and wherein the expressible *luxCDABE* gene fusion generates baseline luminescence;

(ii) monitoring the light output of the first and second detector cells contacted with a compound suspected of having thienylalanine-like whereby

a) light output equal to baseline luminescence in both the first and the second detector cells indicates a compound with no biological activity;

COR-AAD1-00001983

b) a decrease in light output as compared with baseline luminescence in both the first and the second detector cells indicates a compound with growth inhibitory activity and;

c) a decrease in light output in the first detector cell as compared with baseline luminescence and an increase or maintenance of light output in the second detector cell indicates a compound having thienylalanine-like activity.

17. A method for the identification of Glyphosate-like compounds comprising:

(i) contacting a compound suspected of having glyphosate-like activity with a first detector bacteria and with a second detector bacteria,

the first detector bacteria comprising:

a) an expressible *luxCDABE* gene fusion; and

b) the normal cellular complement of an *aroA* gene wherein the *aroA* gene is not overexpressed

the second detector cell comprising:

a) an expressible *luxCDABE* gene fusion; and

b) multiple copies of an expressible *aroA* gene selected wherein the *aroA* gene is overexpressed;

the first and the second detector cells being isogenic and wherein the expressible *luxCDABE* gene fusion generates baseline luminescence;

(ii) monitoring the light output of the first and the second detector cells contacted with a compound suspected of having glyphostate-like activity whereby

a) light output equal to baseline luminescence in both the first and the second detector cells indicates a compound with no biological activity;

b) a decrease in light output as compared with baseline luminescence in both the first and the second detector cells indicates a compound with growth inhibitory activity; and

c) a decrease in light output in the first detector cell as compared with baseline luminescence and an increase or maintenance of light output in the second detector cell indicates a compound having glyphosate-like activity.

18. A method for the identification of ALS-inhibiting compounds comprising:

(i) contacting a compound suspected of having ALS-inhibiting activity with a first detector bacteria and with a second detector bacteria,

the first detector bacteria comprising:

COR-AAD1-00001984

a) an expressible *luxCDABE* gene fusion; and

b) the normal cellular complement of a gene selected from the group consisting of *ilvBN* and *ilvIH* wherein the gene is not overexpressed;

the second detector cell comprising:

a) an expressible *luxCDABE* gene fusion; and

b) multiple copies of an expressible gene selected from the group consisting of *ilvBN* and *ilvIH* wherein the expressible gene is overexpressed;

the first and the second detector cells being isogenic and wherein the expressible *luxCDABE* gene fusion generates baseline luminescence;

(ii) monitoring the light output of the first and the second detector cells contacted with a compound suspected of having ALS-inhibiting activity whereby

a) light output equal to baseline luminescence in both the first and the second detector cells indicates a compound with no biological activity;

b) a decrease in light output as compared with baseline luminescence in both the first and the second detector cells indicates a compound with growth inhibitory activity; and

c) a decrease in light output in the first detector cell as compared with baseline luminescence and an increase or maintenance of light output in the second detector cell indicates a compound having ALS-inhibiting activity.

19. A method for identifying a genotoxic substance, comprising the steps of:

(i) culturing a detector cell comprising a promoter regulated by a SOS bacterial regulatory circuit and a *luxCDABE* gene complex wherein the *luxCDABE* gene complex is positioned in the bacterial chromosome downstream of the SOS promoter such that when the SOS promoter is expressed, then the *luxCDABE* gene complex is also expressed;

(ii) contacting the culture of step (i) with a substance to be tested; and

(iii) determining whether the substance is genotoxic by measuring the amount of luminescence in the culture.

20. The method of Claim 19 wherein the promoter regulated by a SOS bacterial regulatory circuit is selected from the group consisting of *recA*, *uvrA*, *lexA*, *umuDC*, *uvrA*, *uvrB*, *uvrC*, *sulA*, *recN*, *uvrD*, *ruv*, d*inA*, *dinB*, *dinD*, and *dinF*.

COR-AAD1-00001985

21. A method of identifying a structural gene encoding a stressor target comprising:

(i) contacting a stressor with a detector bacteria, the detector bacteria comprising a genotoxic-sensitive promoter operably linked to a luminescent reporter gene complex to form a gene fusion that confers a bioluminescent positive phenotype upon the detector bacteria wherein exposure of the genotoxic-sensitive promoter to genotoxic compounds drives heightened expression of the luminescent reporter gene complex producing an increased bioluminescent signal;

(ii) selecting for genotoxic or non-genotoxic stressors capable of inhibiting the growth said detector bacteria of step (i) by monitoring the growth and light output of the detector bacteria;

(iii) submitting the growth-inhibiting, genotoxic or non-genotoxic stressor to a site of action screen wherein the stressor target is identified and;

(vi) isolating the structural gene encoding the stressor target.

22. A method for the identification of the site of action of a stressor comprising:

(i) contacting a stressor with a detector bacteria, the detector bacteria comprising an expressible *luxCDABE* gene fusion wherein the expressible *luxCDABE* gene fusion generates baseline luminescence in the detector bacteria;

(ii) selecting for stressors capable of inhibiting the growth of the detector bacteria of step (i) by monitoring the growth and light output of the detector bacteria;

(iii) submitting the growth-inhibiting stressor selected in step (ii) to a site of action screen;

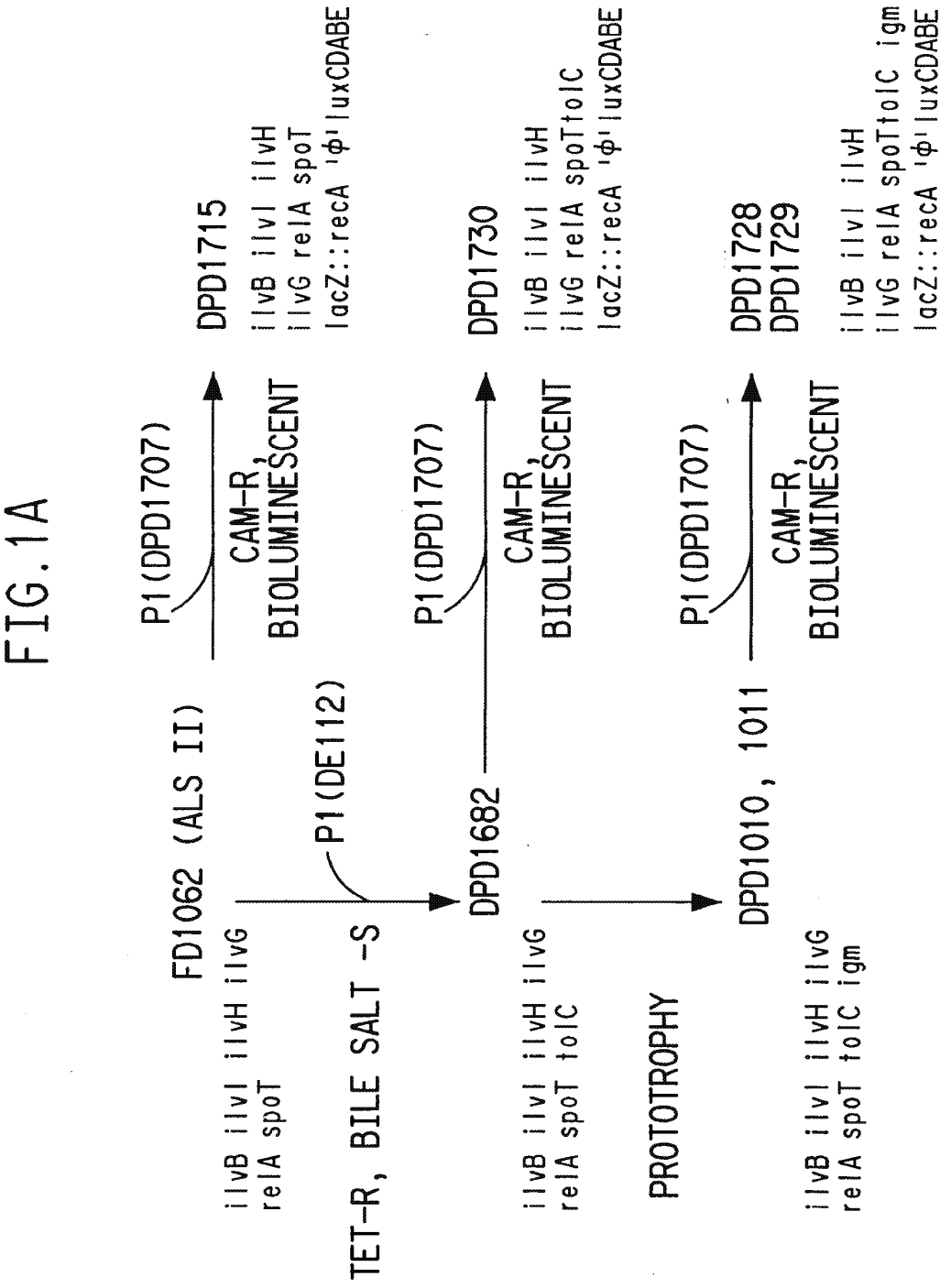
(iv) identifying the site of action in the detector bacteria; and

(vi) confirming the site of action in a target organism.

23. The method of Claim 22 wherein the expressible *luxCDABE* gene fusion further comprises a stressor sensitive promoter selected from the group consisting of *grpE*, *dnaK*, *lon*, *rpoD*, *groESL*, *lysU*, *htpE*, *htpG*, *htpI*, *htpK*, *clpP*, *clpB*, *htpN*, *htpO*, *htpX*, *katG*, *ahp*, *micF*, *sodA*, *nfo*, *zwf*, *soi*, *fabA*, *his*, *ilvBN*, *ilvGMED*, *thrABC*, *lac*, *mal*, and *gal*, *phoA*, *phoBR*, *phoE*, *phoS*, *aphA*, *himA*, *pepN*, *ugpAB*, *psiD*, *psiE*, *psiF*, *psiK*, *psiG*, *psiI*, *psiJ*, *psiN*, *psiR*, *psiH*, *phiL*, *phiO* *ara*, *tna*, *dsd*, *glnA*, and *hut*.

COR-AAD1-00001986

FIG. 1A



COR-AAD1-00001987

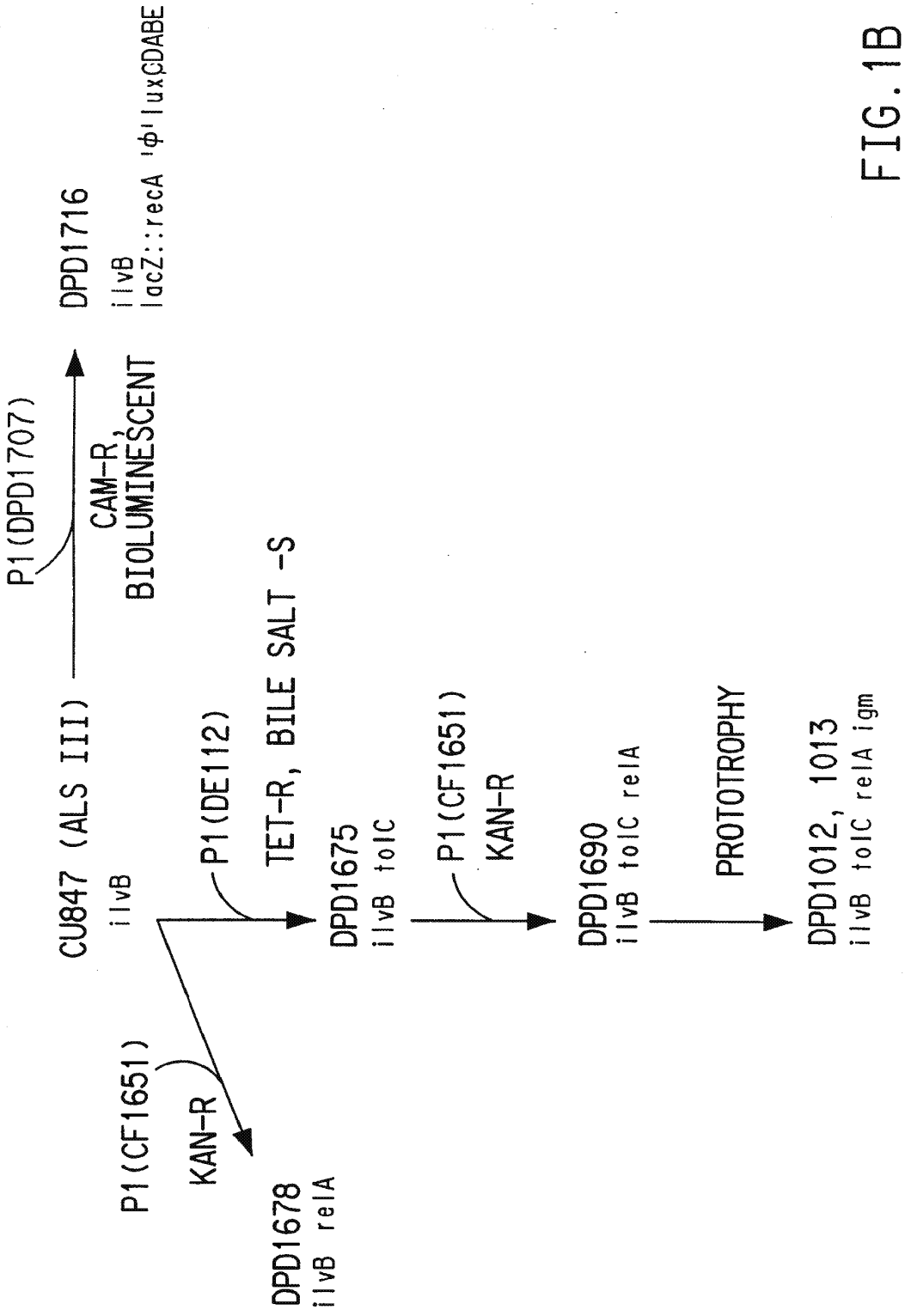


FIG. 1B

COR-AAD1-00001988

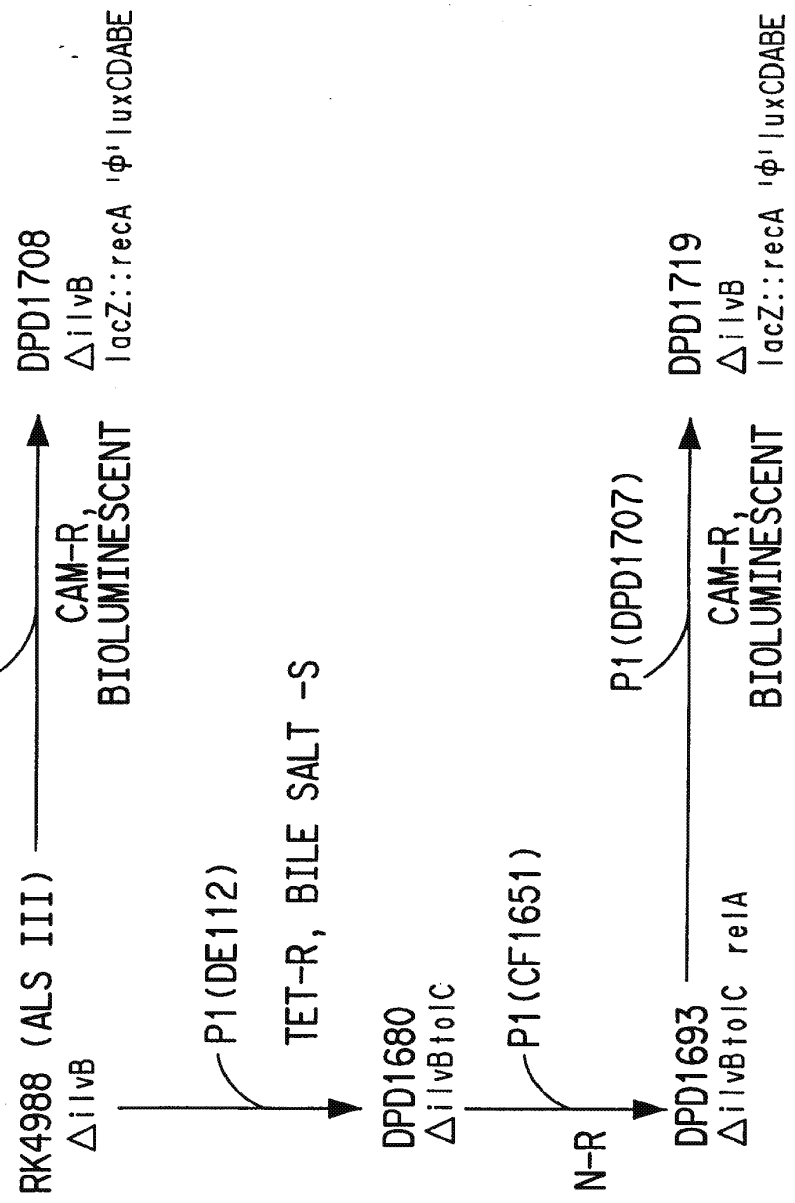
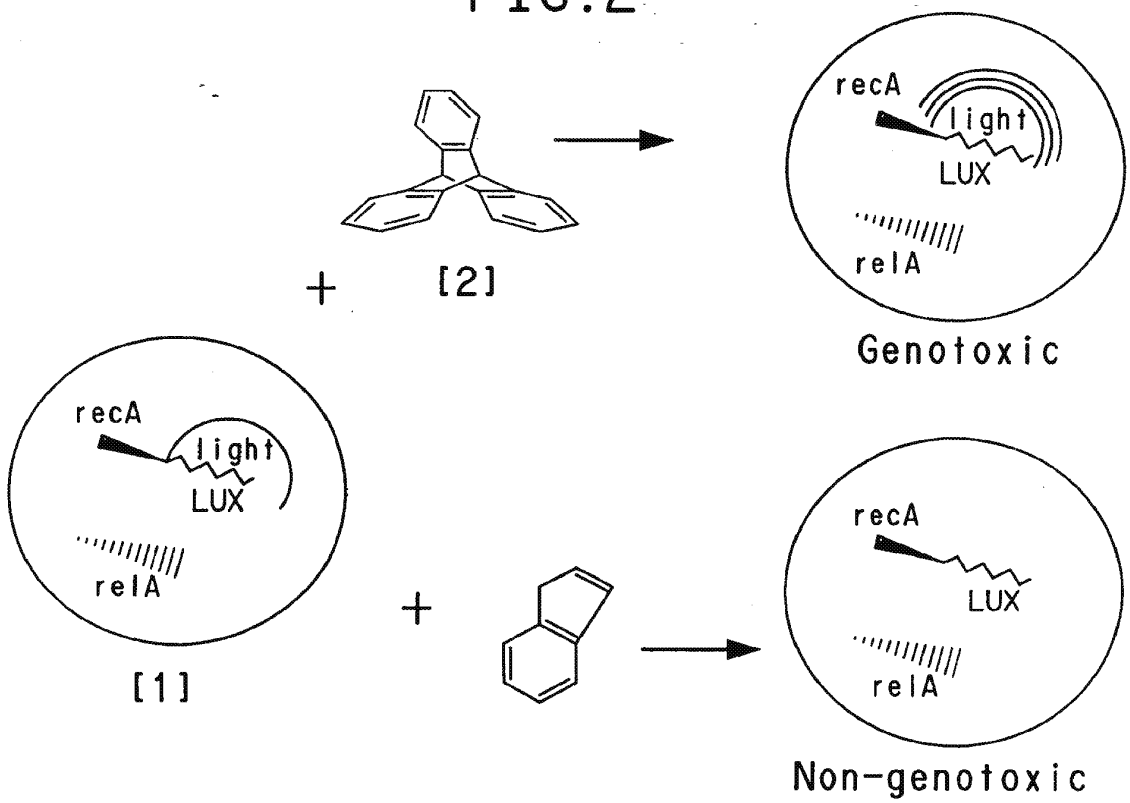


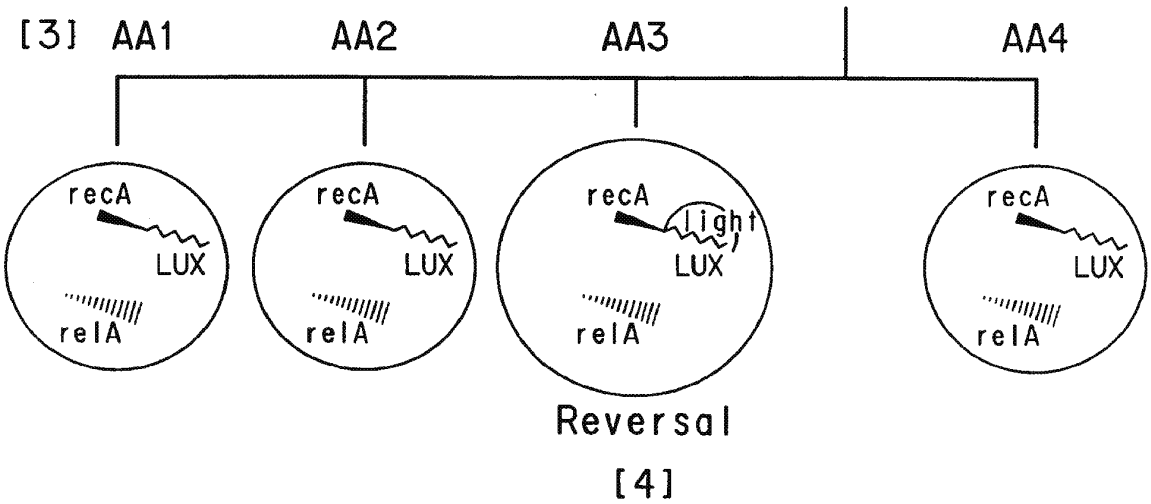
FIG. 1B (CONTINUED)

COR-AAD1-00001989

FIG.2

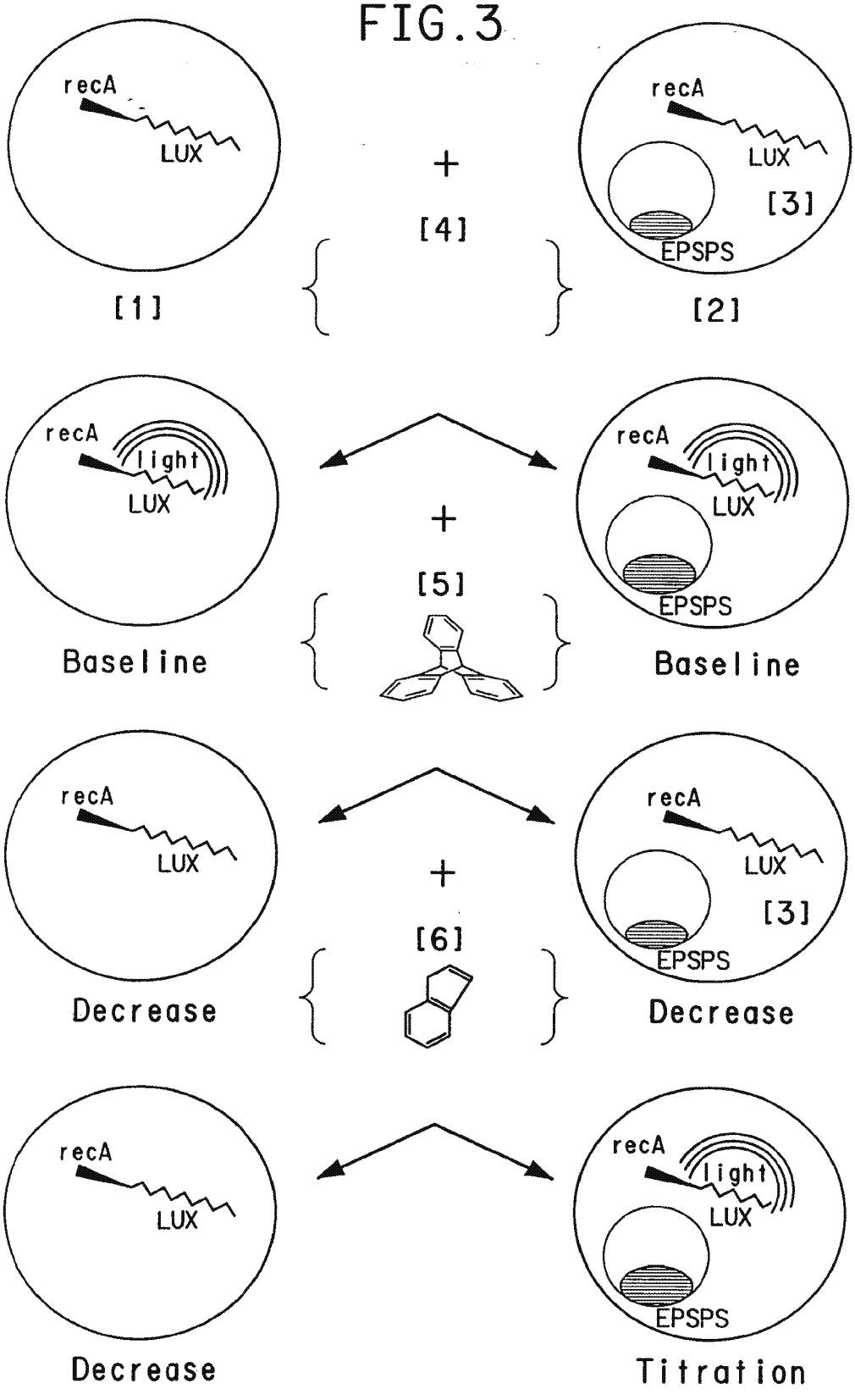


Minimal Medium Supplemented Pools



COR-AAD1-00001990

FIG.3



COR-AAD1-00001991

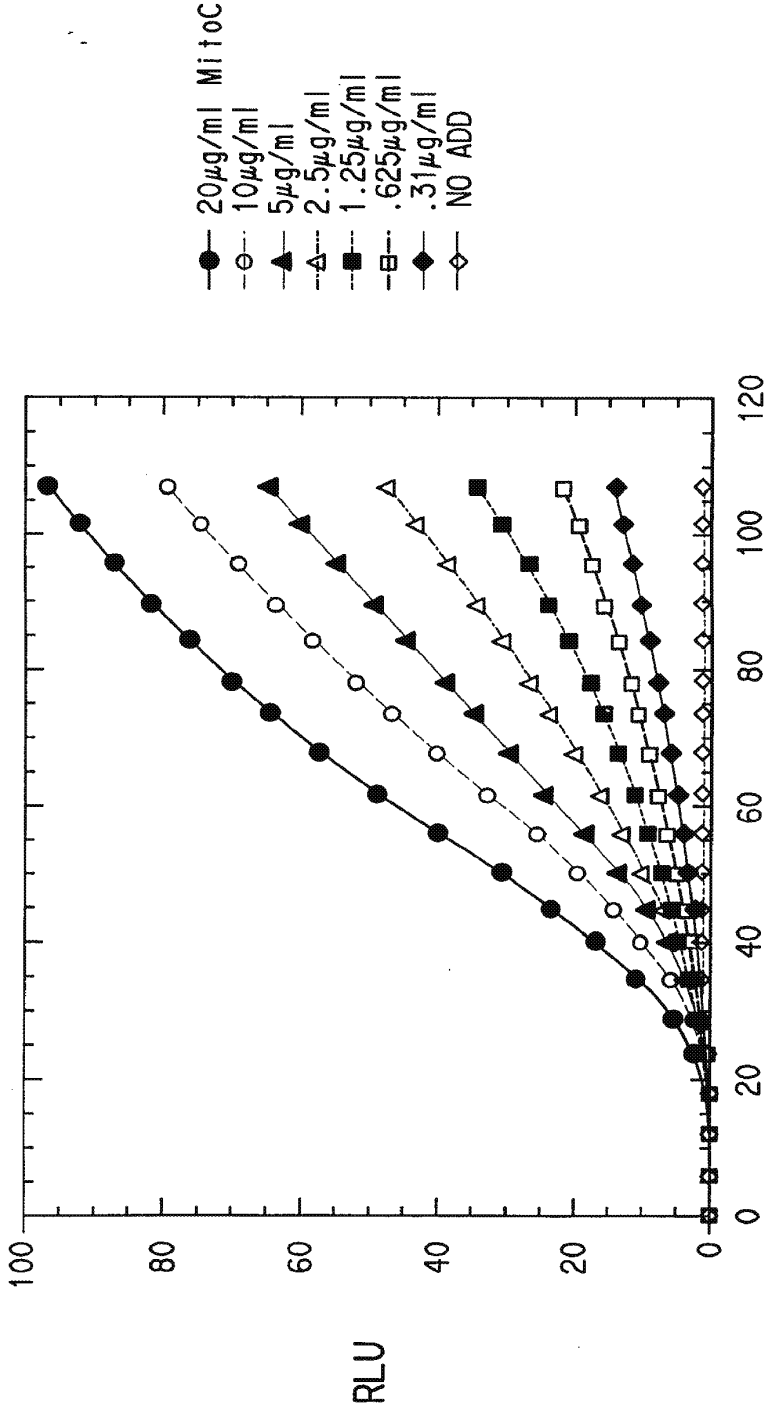

20μg/ml MitoC
10μg/ml
5μg/ml
2.5μg/ml
1.25μg/ml
.625μg/ml
.31μg/ml
NO ADD
RLU
MINUTES
0
20
40
60
80
100
120

FIG. 4A

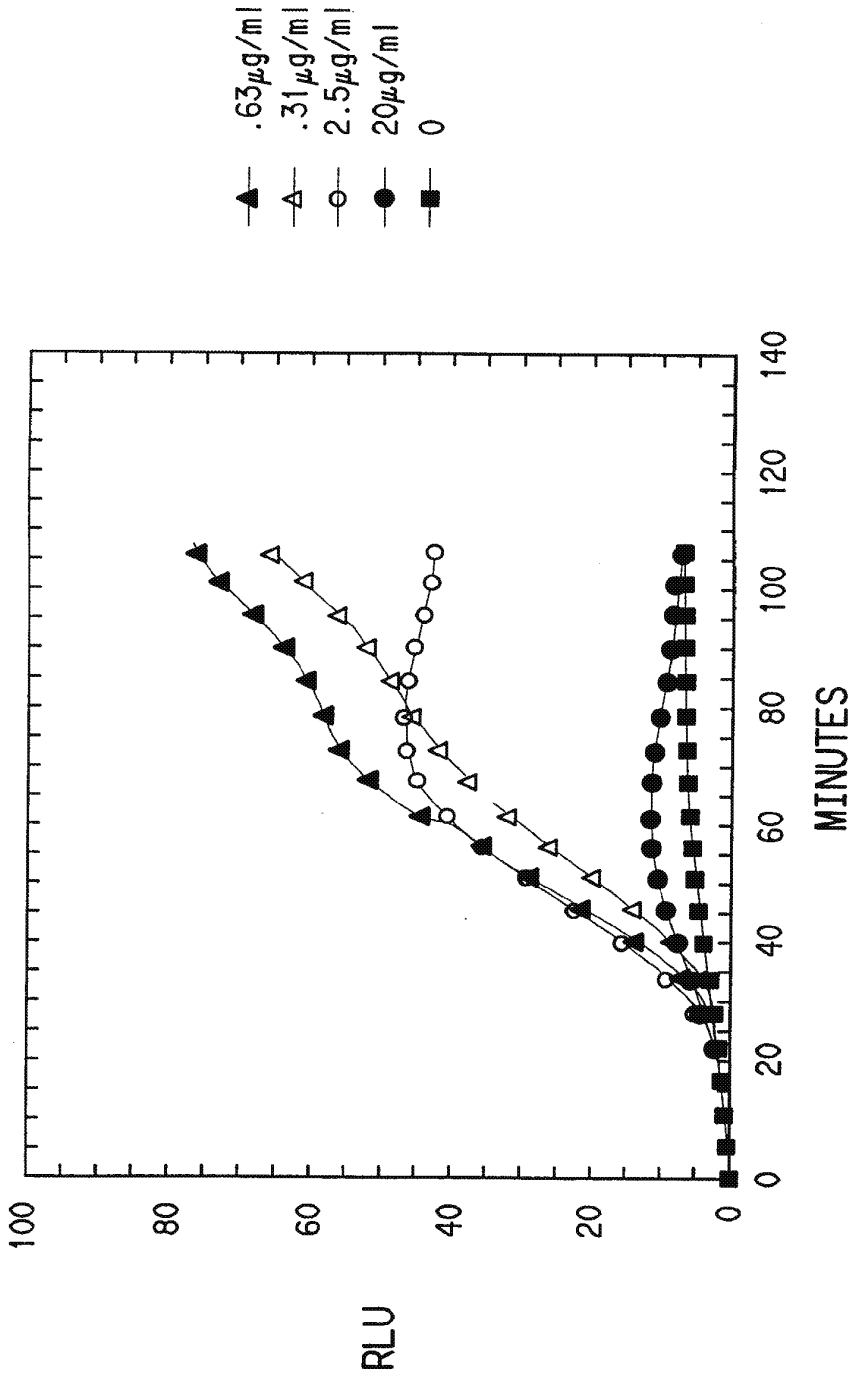


FIG. 4B

COR-AAD1-00001993

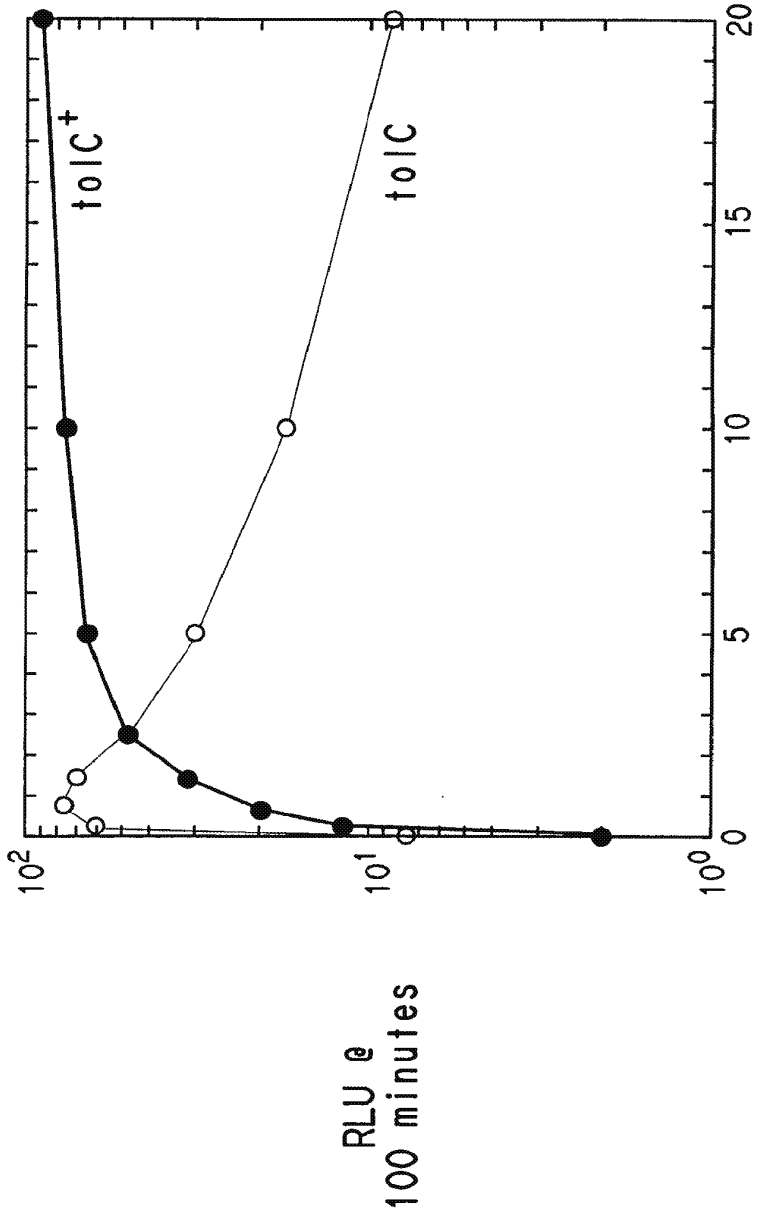


FIG.5

COR-AAD1-00001994

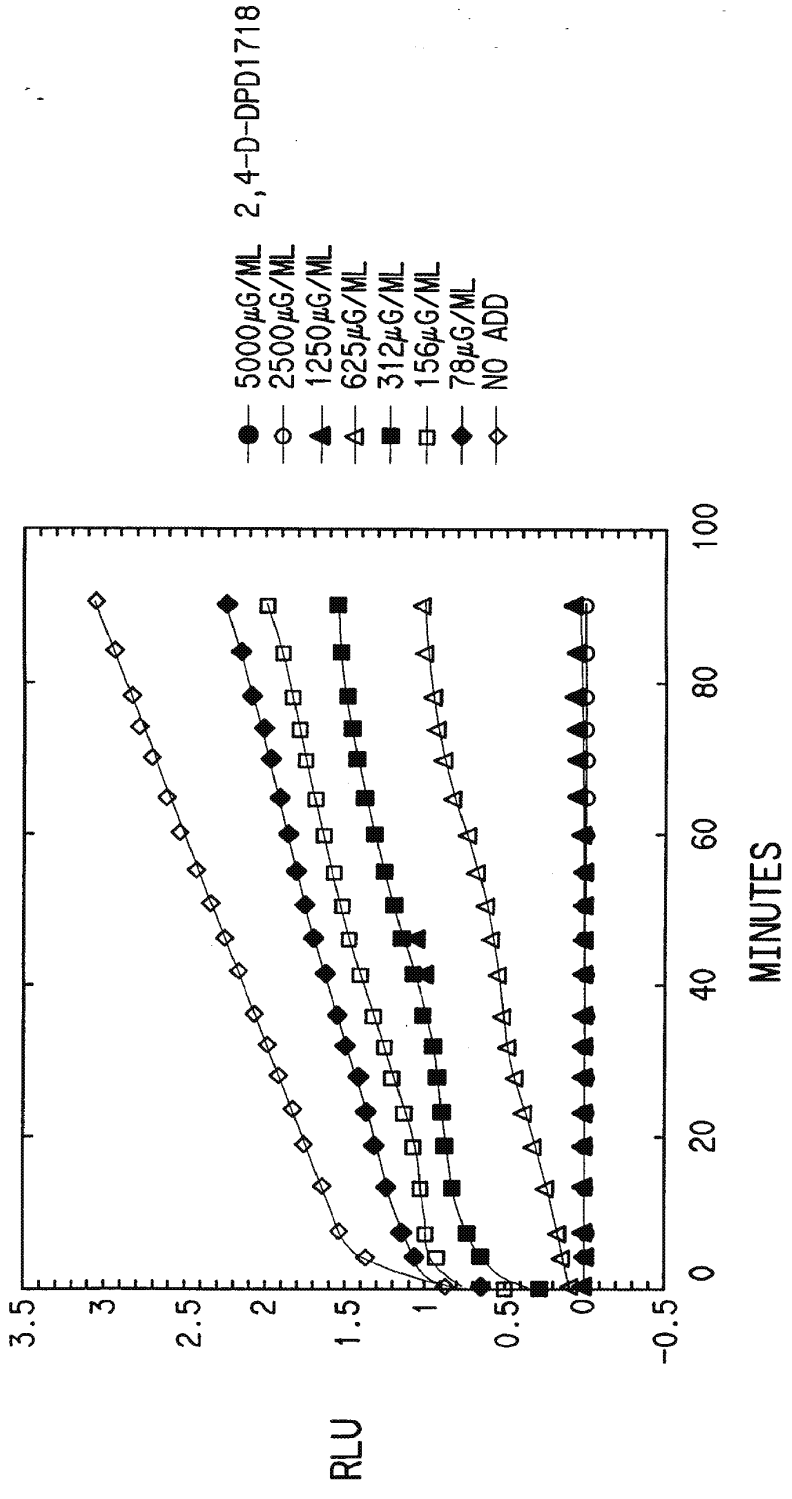

2,4-D-DPD1718
5000µG/ML
2500µG/ML
1250µG/ML
625µG/ML
312µG/ML
156µG/ML
78µG/ML
NO ADD
RLU
MINUTES
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
20
40
60
80
100

FIG.6

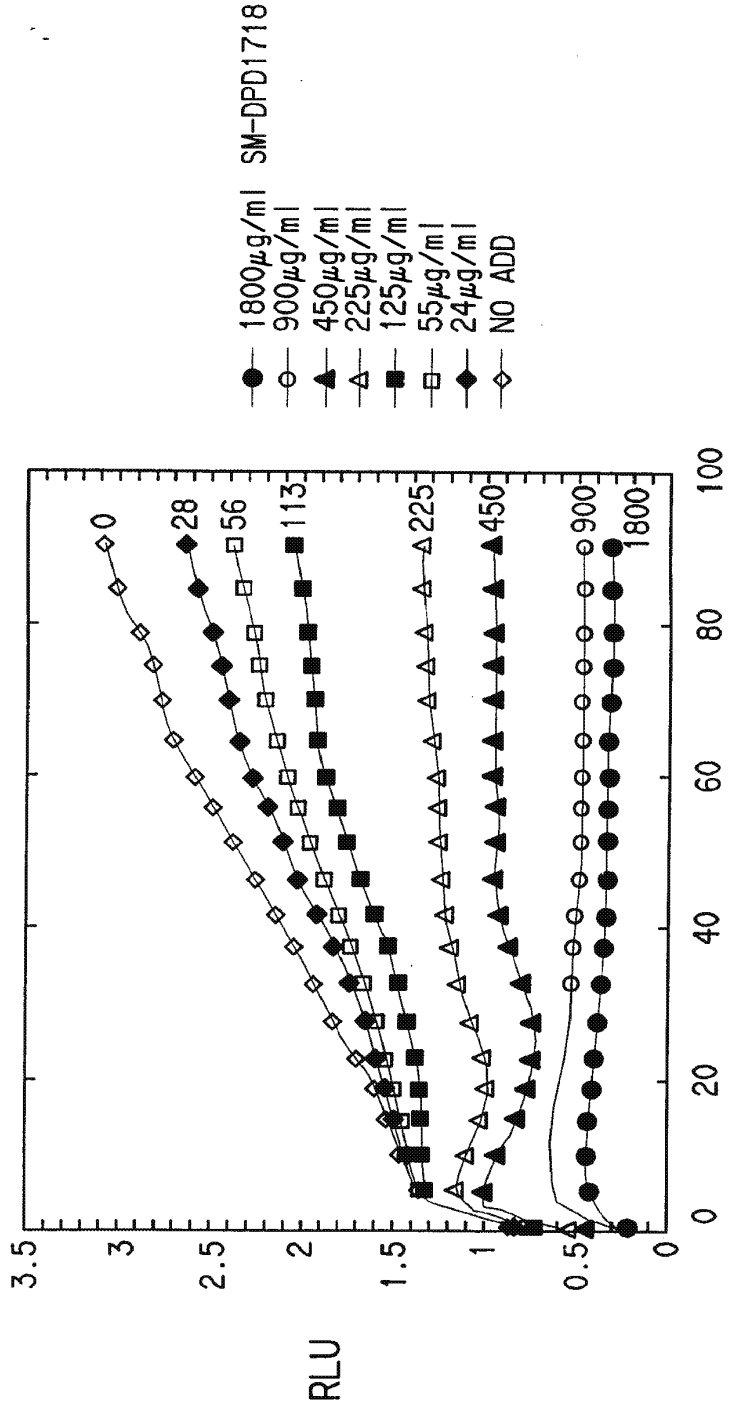

SM-DPD1718
1800µg/ml
900µg/ml
450µg/ml
225µg/ml
125µg/ml
55µg/ml
24µg/ml
NO ADD
0
28
56
113
225
450
900
1800
RLU
MINUTES
3.5
3
2.5
2
1.5
1
0.5
0
0
20
40
60
80
100

FIG. 7

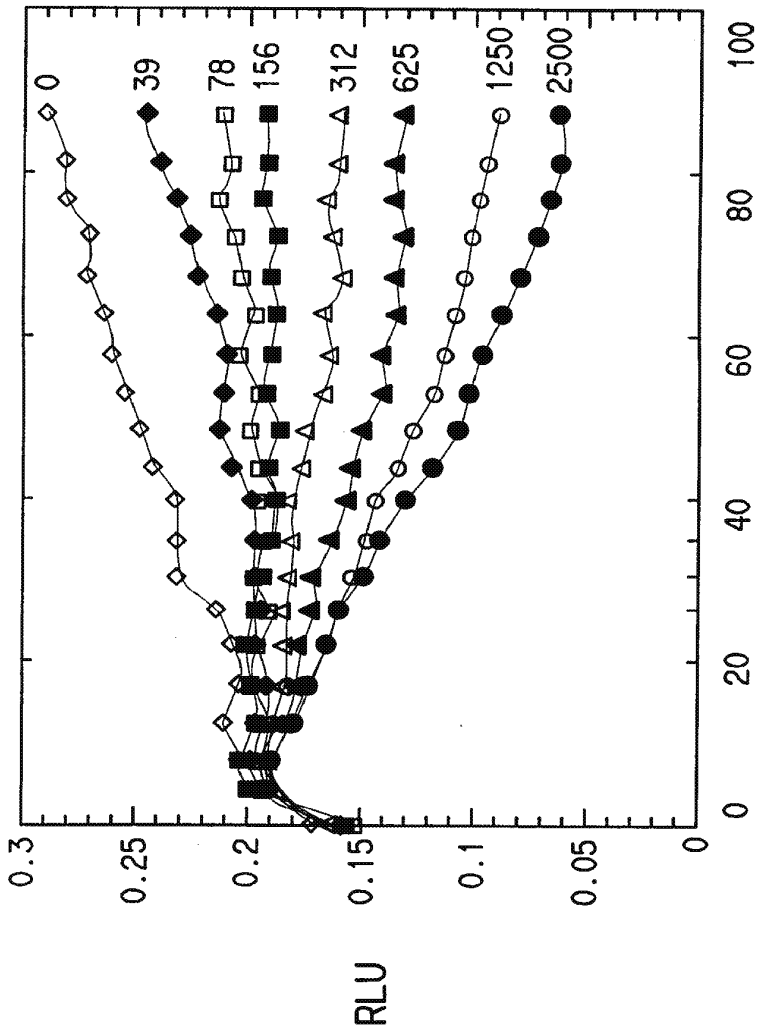


FIG. 8A

COR-AAD1-00001997

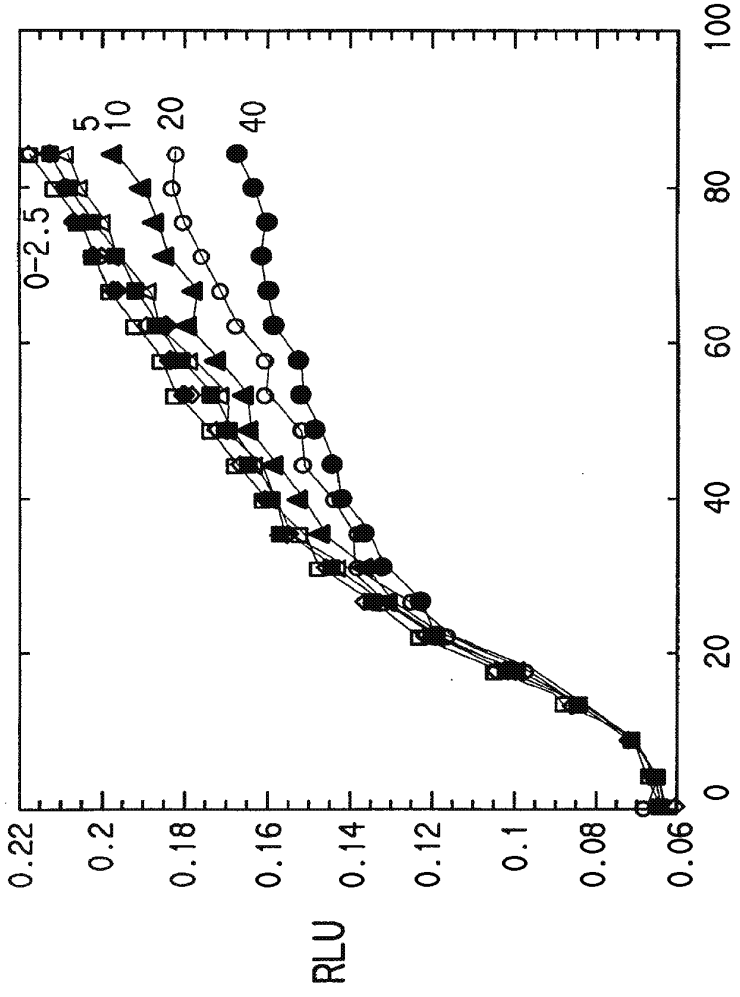


FIG.8B

COR-AAD1-00001998

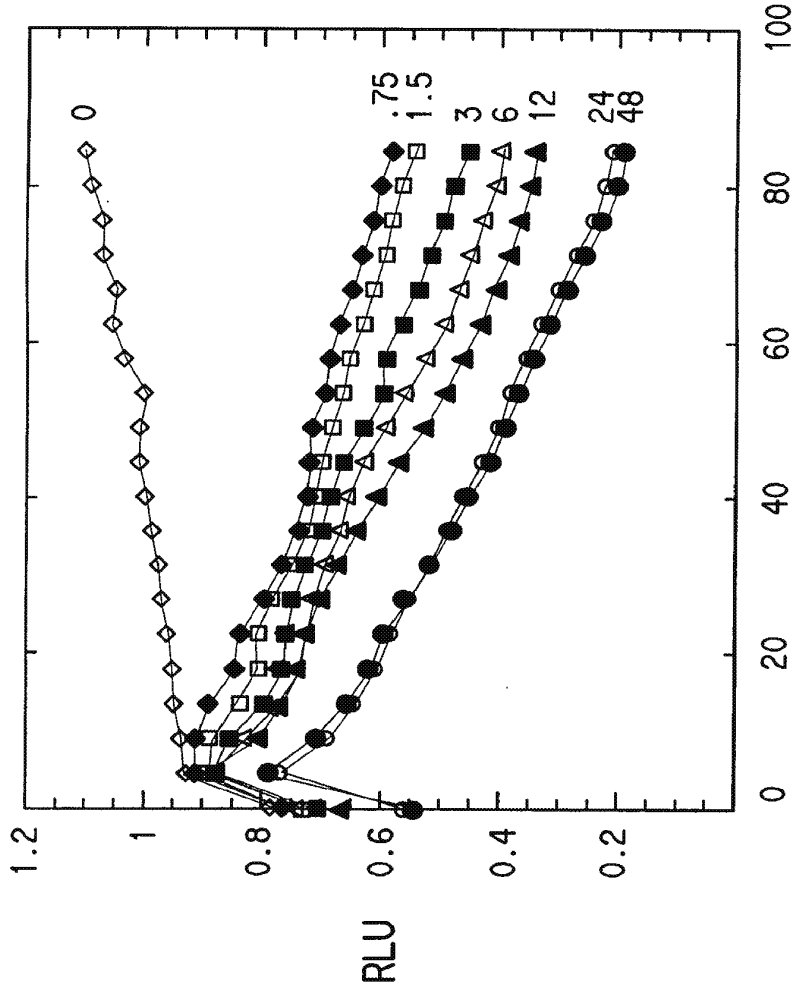


FIG.9

COR-AAD1-00001999

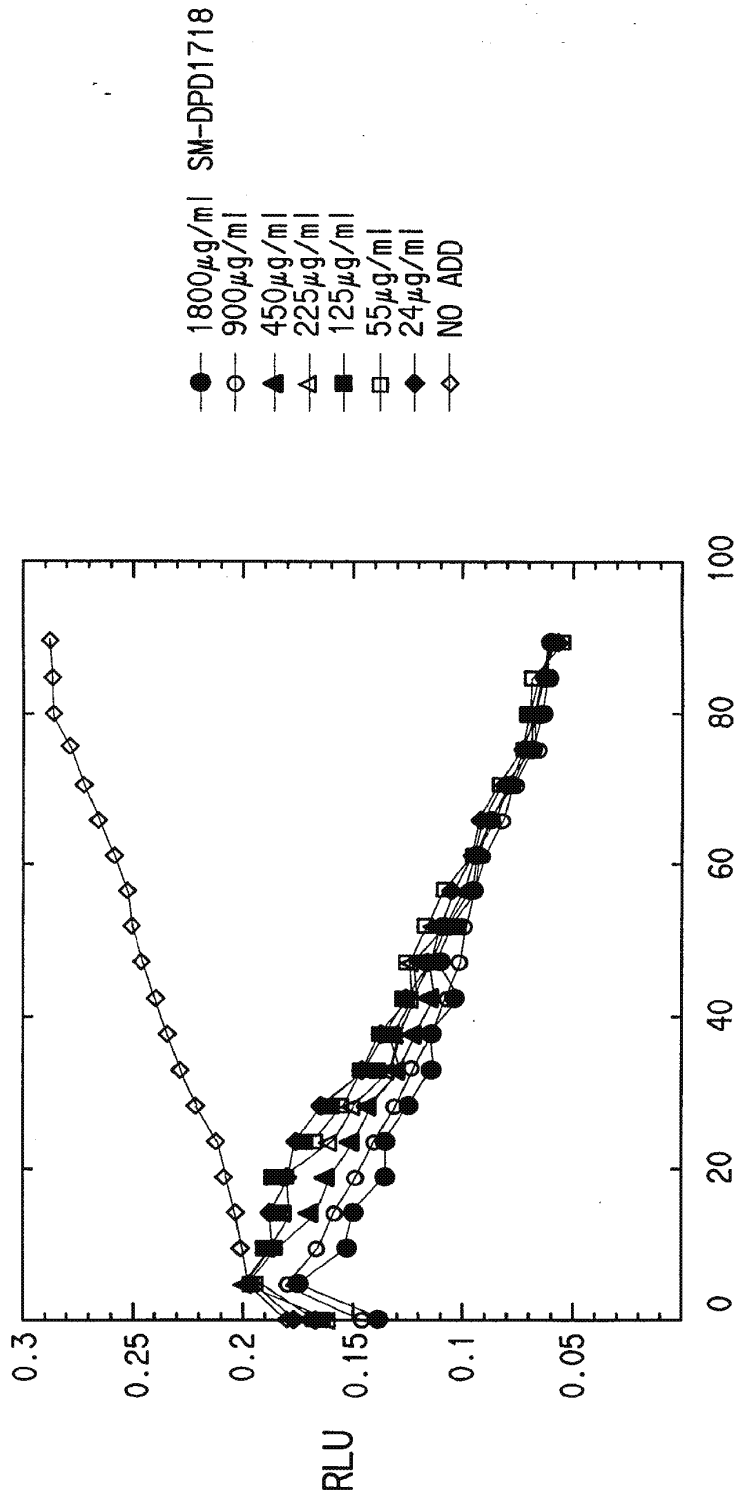

SM-DPD1718
1800µg/ml
900µg/ml
450µg/ml
225µg/ml
125µg/ml
55µg/ml
24µg/ml
NO ADD
RLU
0.3
0.25
0.2
0.15
0.1
0.05
0
20
40
60
80
100
MINUTES

FIG. 10A

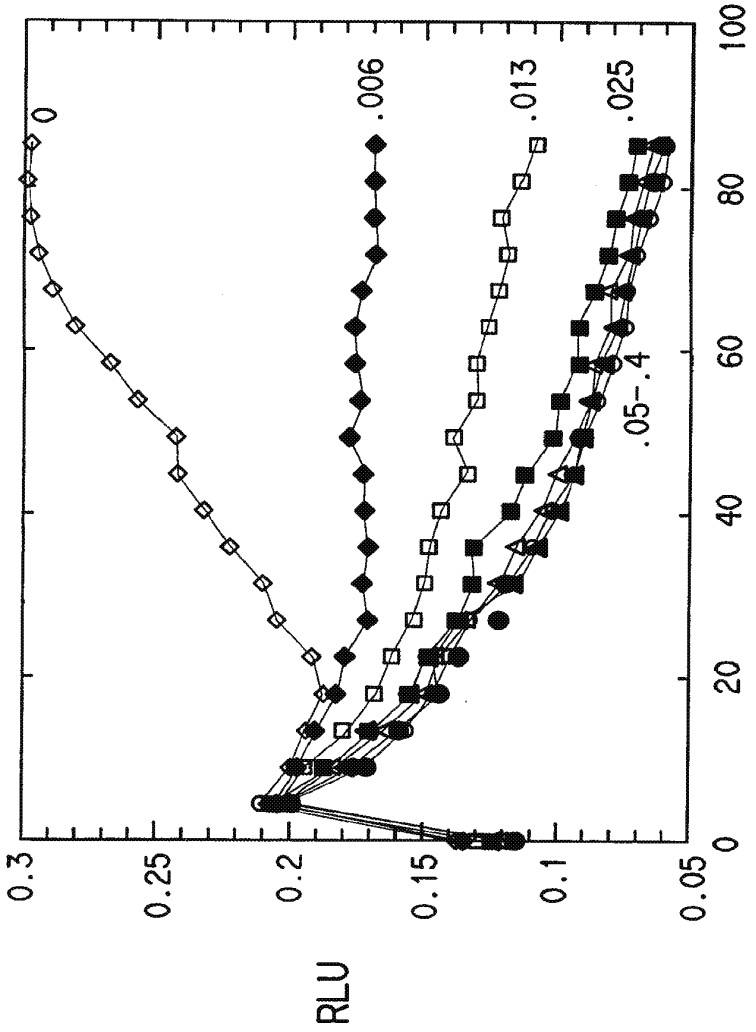


FIG. 10B

COR-AAD1-00002001

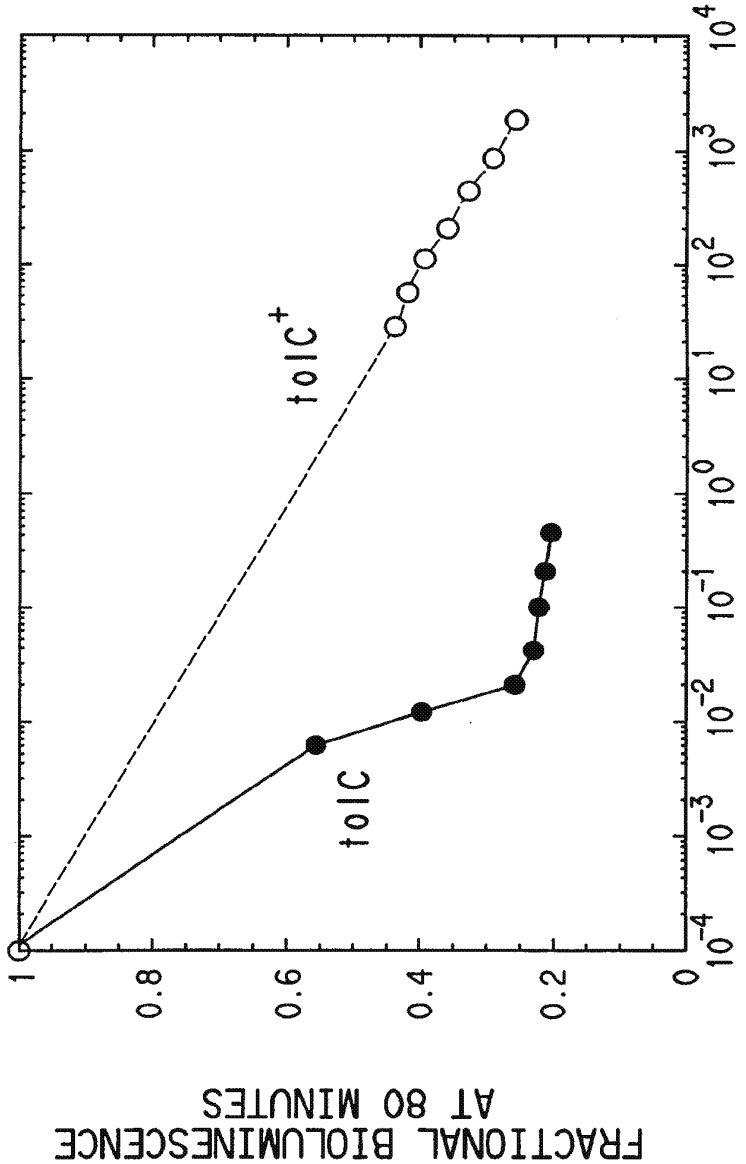


FIG. 10C

COR-AAD1-00002002

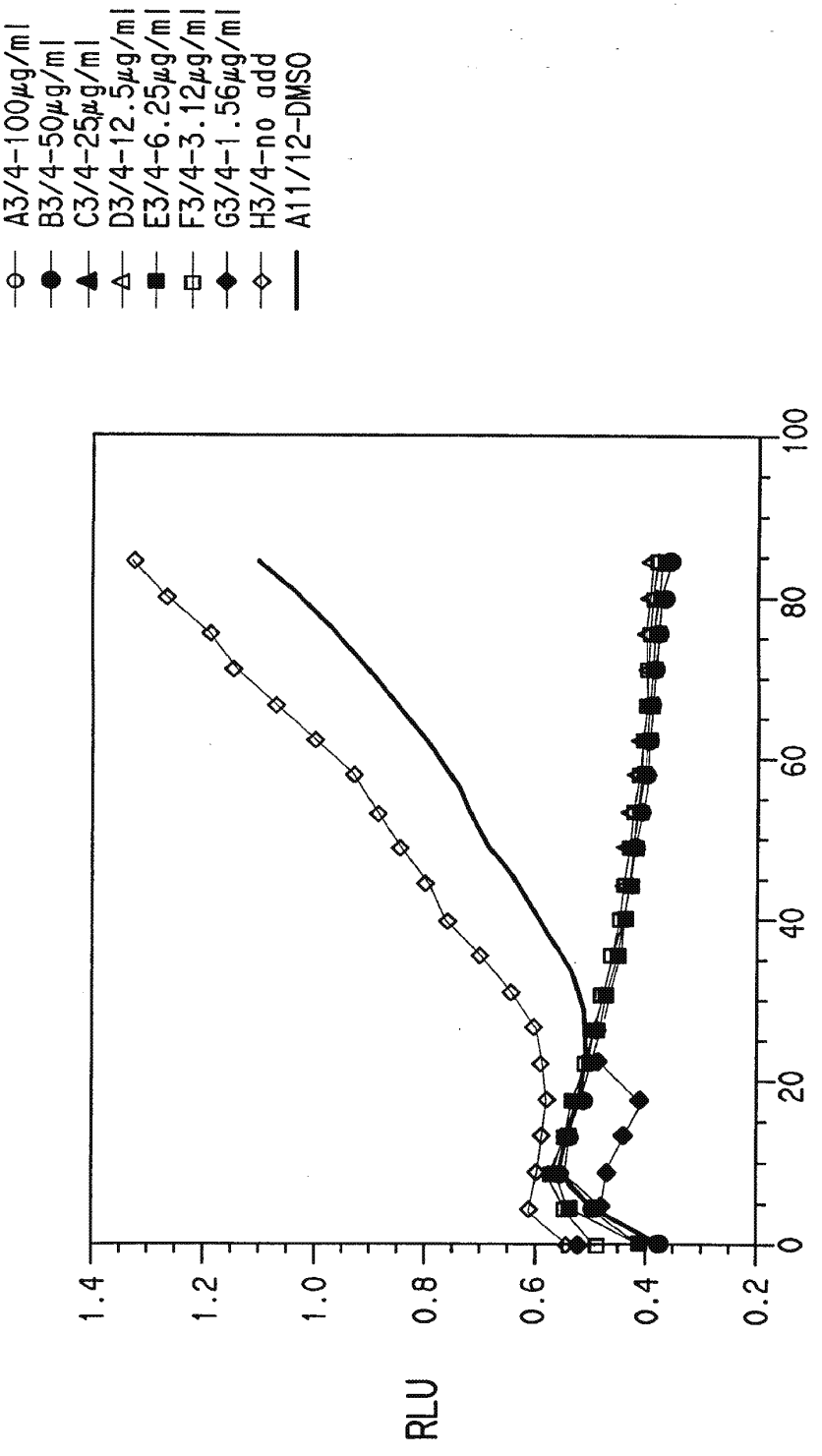

A3/4-100µg/ml
B3/4-50µg/ml
C3/4-25µg/ml
D3/4-12.5µg/ml
E3/4-6.25µg/ml
F3/4-3.12µg/ml
G3/4-1.56µg/ml
H3/4-no add
A11/12-DMSO
RLU
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
20
40
60
80
100
Minutes

FIG. 11A

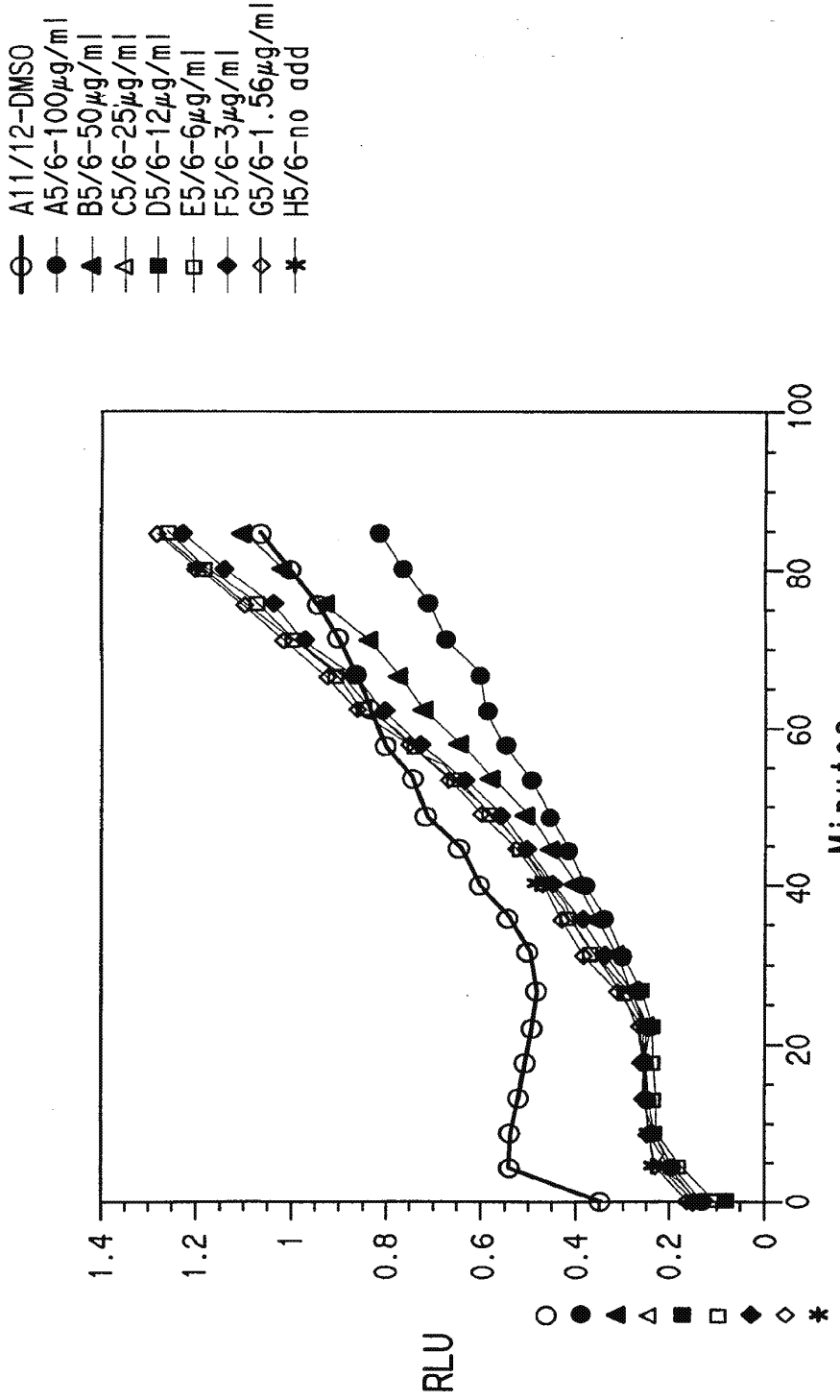

A11/12-DMSO
A5/6-100µg/ml
B5/6-50µg/ml
C5/6-25µg/ml
D5/6-12µg/ml
E5/6-6µg/ml
F5/6-3µg/ml
G5/6-1.56µg/ml
H5/6-no add
RLU
0
0.2
0.4
0.6
0.8
1
1.2
1.4
0
20
40
60
80
100
Minutes

FIG.11B

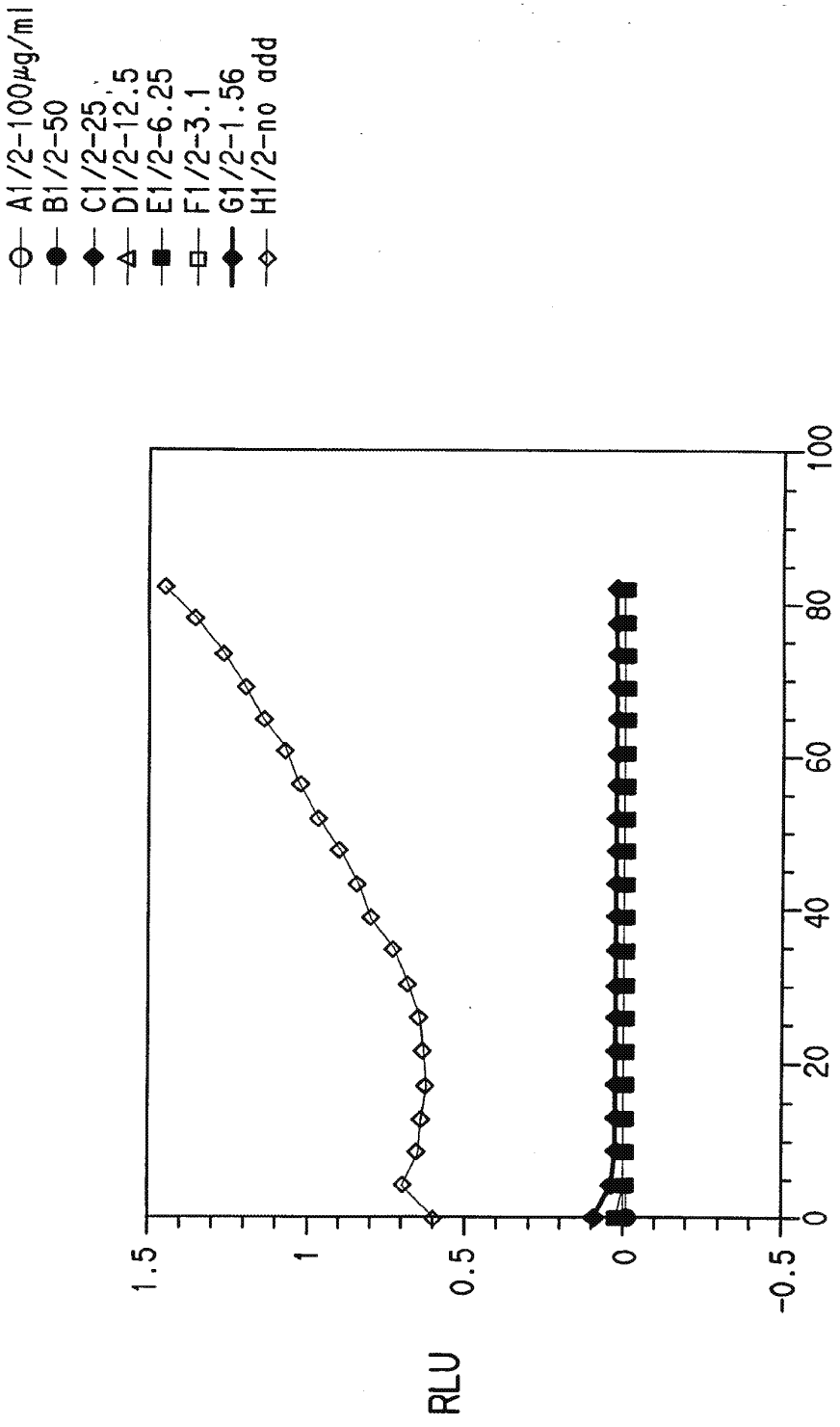


FIG. 12A

COR-AAD1-00002005

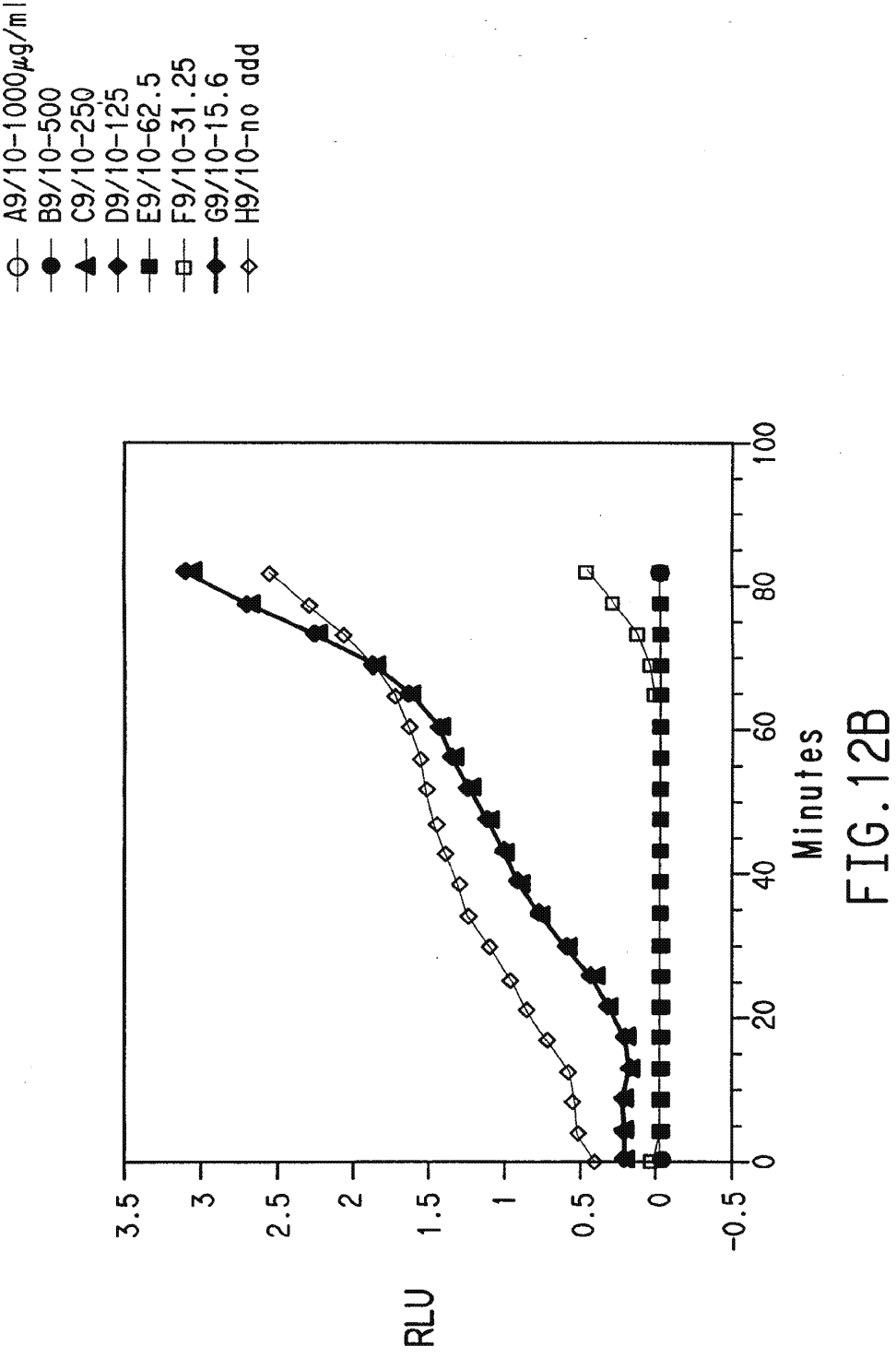

A9/10-1000µg/ml
B9/10-500
C9/10-250
D9/10-125
E9/10-62.5
F9/10-31.25
G9/10-15.6
H9/10-no add
RLU
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
20
40
60
80
100
Minutes

FIG.12B

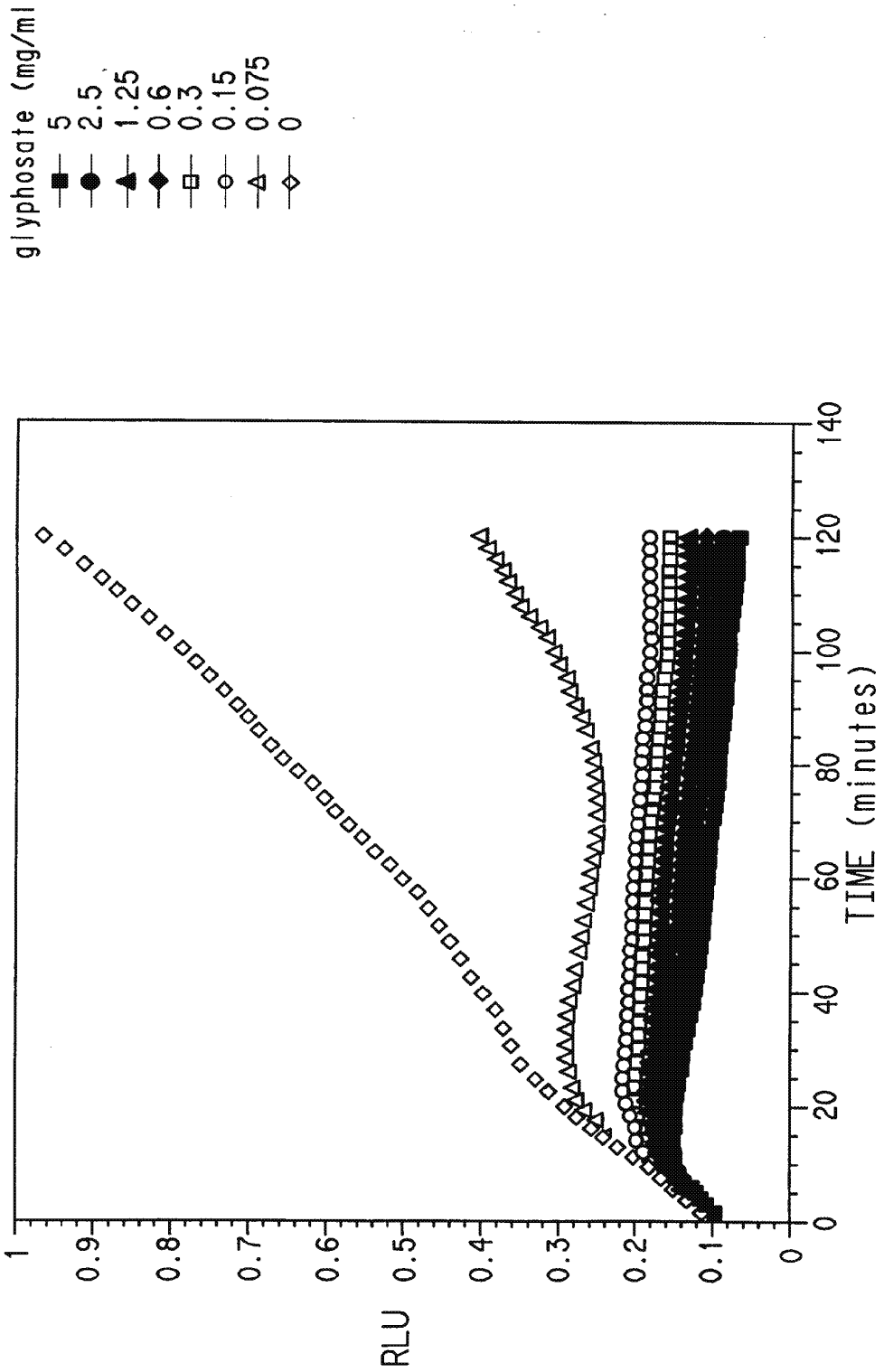


FIG. 13A

COR-AAD1-00002007

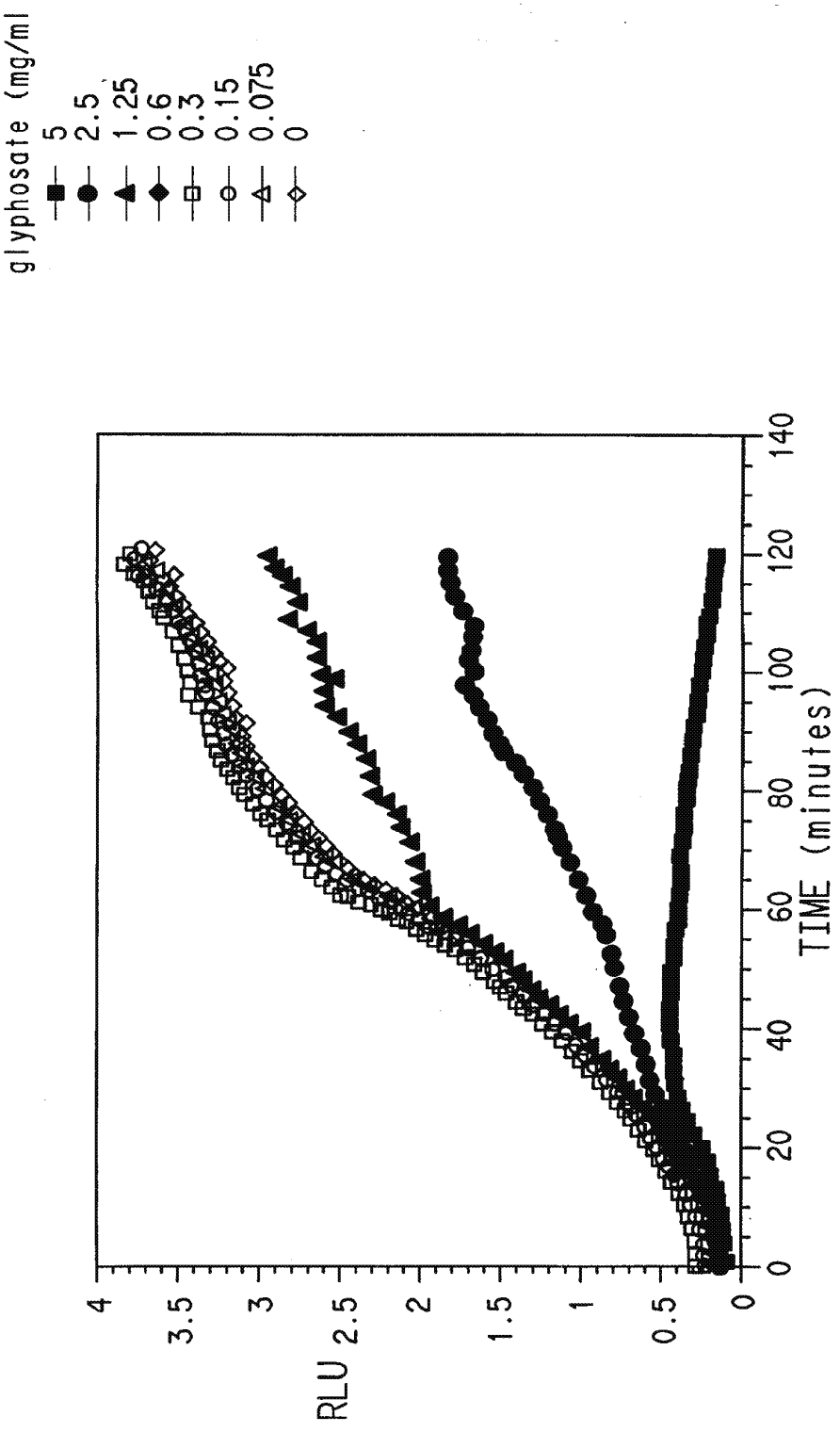

glyphosate (mg/ml
5
2.5
1.25
0.6
0.3
0.15
0.075
0
RLU
TIME (minutes)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
20
40
60
80
100
120
140

FIG. 13B

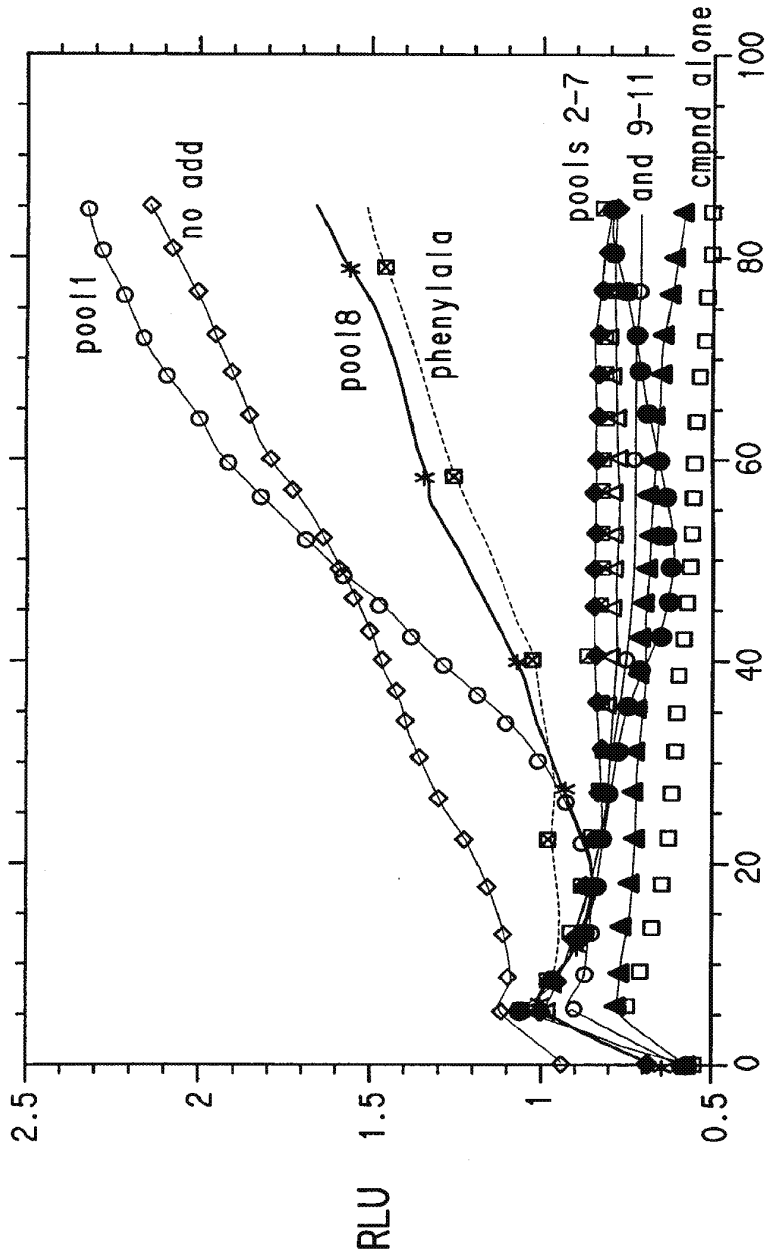

RLU
2.5
2
1.5
1
0.5
0
20
40
60
80
100
Minutes
pool1
no add
pool8
phenylala
pools 2-7
and 9-11
cmpnd alone

FIG.14

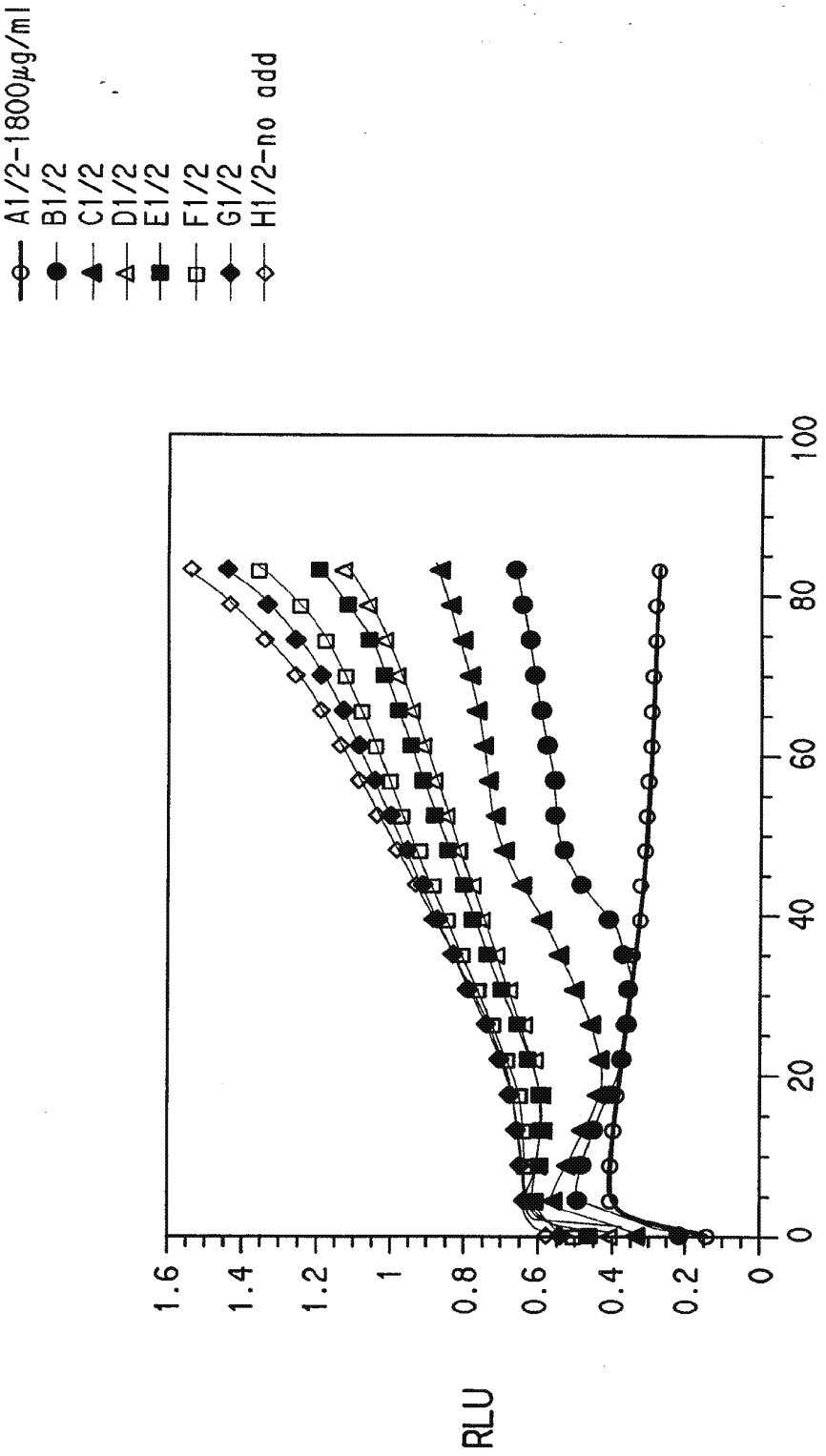


FIG. 15A

COR-AAD1-00002010

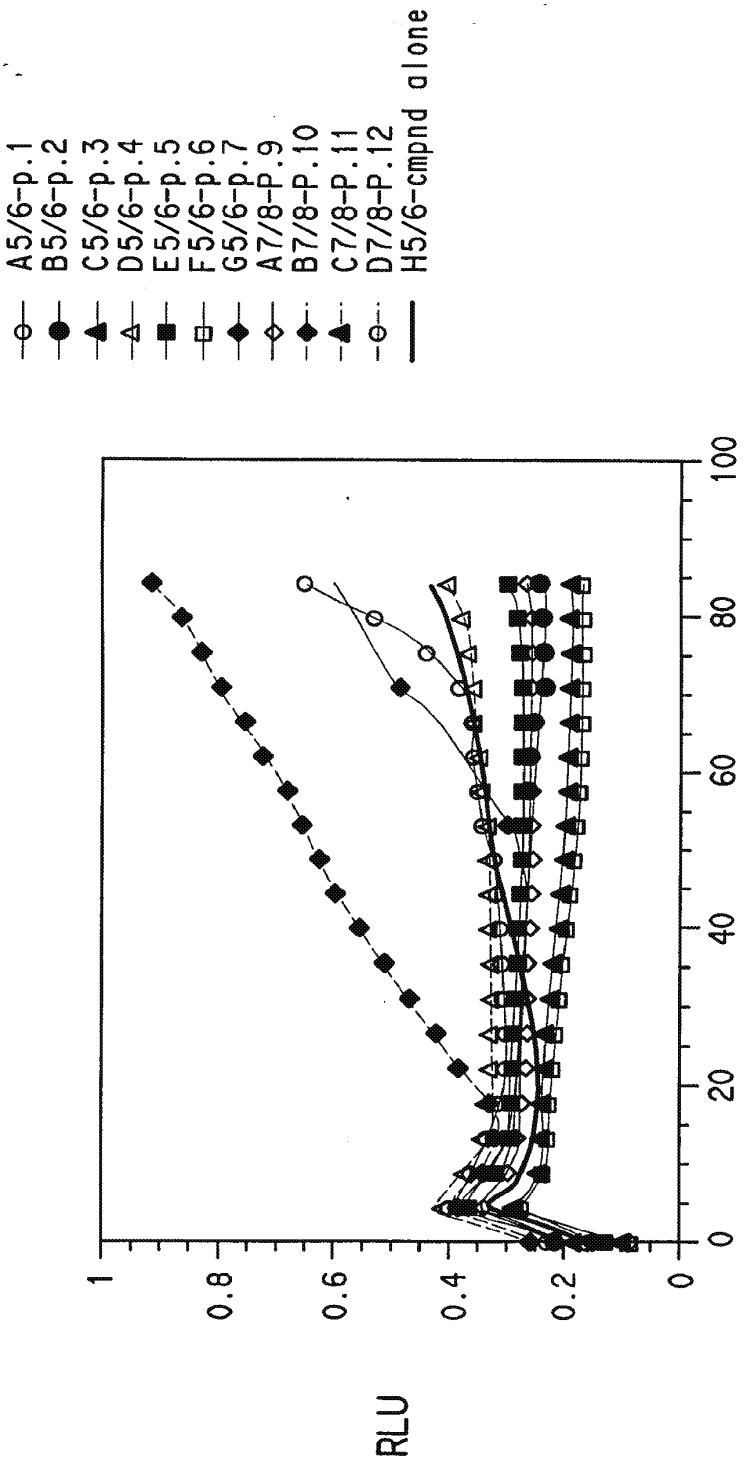

A5/6-p.1
B5/6-p.2
C5/6-p.3
D5/6-p.4
E5/6-p.5
F5/6-p.6
G5/6-p.7
A7/8-P.9
B7/8-P.10
C7/8-P.11
D7/8-P.12
H5/6-cmpnd alone
RLU
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
100
Minutes

FIG. 15B

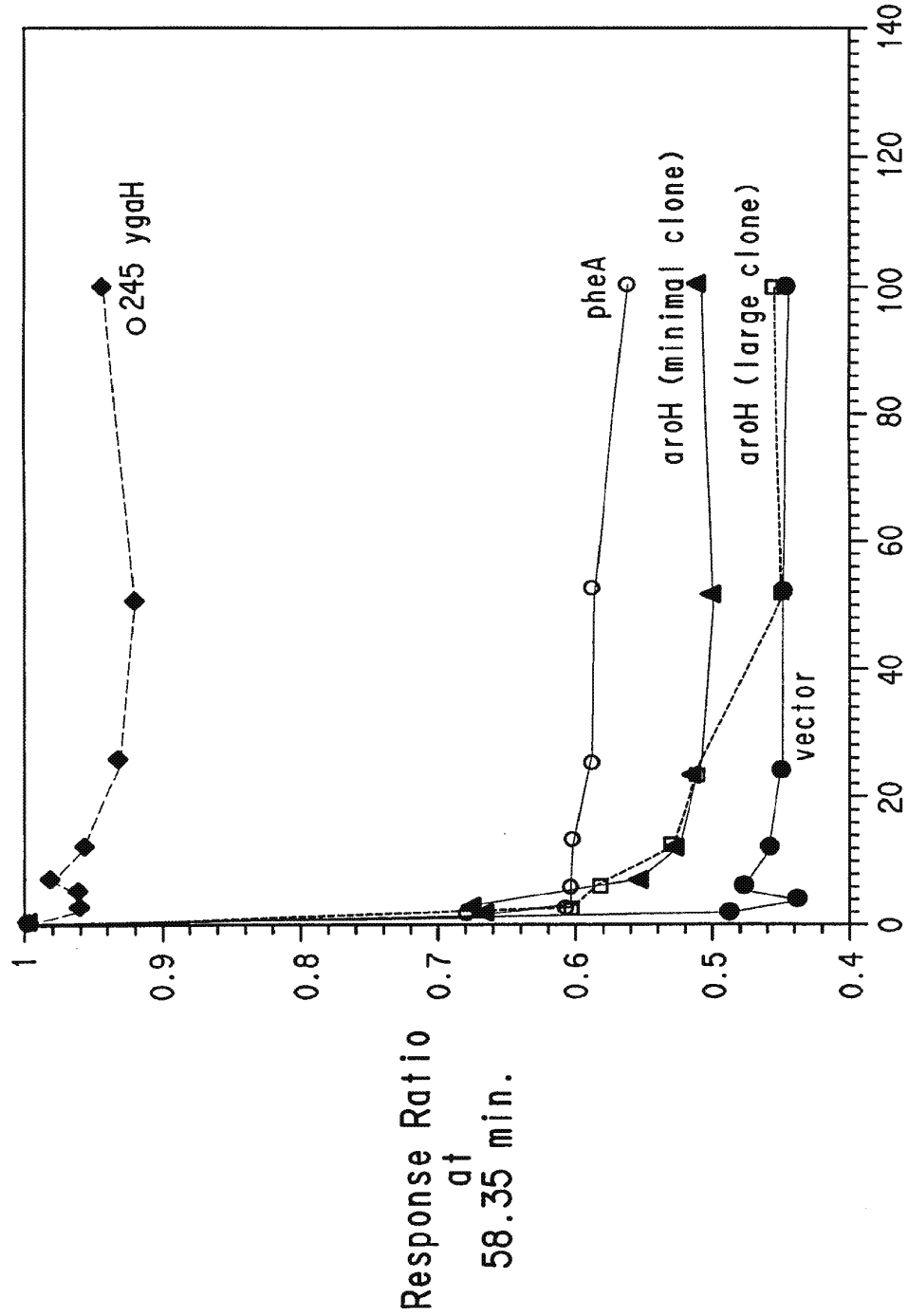


FIG.16

COR-AAD1-00002012

INTERNATIONAL SEARCH REPORT

Int tional Application No
PCT/US 98/03684

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	VAN DYK T K ET AL.: "Rapid and sensitive pollutant detection by induction of heat shock gene-bioluminescence gene fusions" APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 60, no. 5, 1994, pages 1414-1420, XP002069250 see the whole document --- -/--	1-9, 19-23

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publicationdate of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of theinternational search	Date of mailing of the international search report
25 June 1998	09/07/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Knehr, M

Form PCT/ISA/210 (second sheet) (July 1992)

COR-AAD1-00002013

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 98/03684

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BELKIN S ET AL.: "Monitoring subtoxic environmental hazards by stress-responsive luminous bacteria" ENVIRONMENTAL TOXICOLOGY AND WATER QUALITY, vol. 11, no. 3, 1996, pages 179-185, XP002069251 see abstract see page 180, column 1, paragraph 1 - paragraph 2 see page 184, column 1, paragraph 2 - page 185, column 1, paragraph 1; figure 2; table 1	1-4, 6-12, 19-23
X	WO 96 16187 A (DU PONT) 30 May 1996 see the whole document	1-3, 6-12, 19-23
X	EP 0 649 905 A (TOYODA CHUO KENKYUSHO KK) 26 April 1995 see abstract see page 3, line 1 - page 5, line 56; claims 1-20; examples 1,2,5,6	19,20
X	WO 94 13831 A (DYK TINA KANGAS VAN ;LAROSSA ROBERT ALAN (US); MAJARIAN WILLIAM RO) 23 June 1994	1-13, 19-23
Y	see the whole document	14-18
Y	WO 94 17208 A (XENOMETRIX INC ;FARR SPENCER B (US); MARQUE TODD D (US)) 4 August 1994 see abstract see page 1, line 1 - line 12 see page 5, line 7 - line 17 see page 8, line 34 - page 9, line 2; claims 26-50	14-18
A	DE 195 44 374 C (DLR DEUTSCHE FORSCHUNGSANSTALT) 20 February 1997 see the whole document	
A	WO 92 15687 A (VITO) 17 September 1992 see the whole document	
P,X	VOLLMER A C ET AL.: "Detection of DNA damage by use of Escherichia coli carrying recA'::lux, uvrA'::lux, or alkA'::lux reporter plasmids " APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 63, no. 7, 1997, pages 2566-2571, XP002069252	1-12, 19-21
P,Y	see the whole document	13-18, 22,23

-/--

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

COR-AAD1-00002014

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/03684

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 97 41251 A (VITO ;LELIE DANIEL V D (BE); BORREMANS BRIGITTE MARIE FRANC (BE);) 6 November 1997	19,20
P,Y	see the whole document	13-18, 22,23
E	WO 98 21347 A (LAROSSA ROBERT ALAN ;SCOLNIK PABLO ARIEL (US); DU PONT (US)) 22 May 1998	1-3, 6-12, 19-23
	see the whole document	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

COR-AAD1-00002015

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int ional Application No
PCT/US 98/03684

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9616187 A	30-05-1996	CA 2200702 A	30-05-1996
		EP 0793729 A	10-09-1997
		US 5731163 A	24-03-1998
EP 0649905 A	26-04-1995	JP 7227285 A	29-08-1995
		JP 7163359 A	27-06-1995
		US 5702883 A	30-12-1997
WO 9413831 A	23-06-1994	AT 152181 T	15-05-1997
		AU 5730494 A	04-07-1994
		CA 2150232 A	23-06-1994
		DE 69310207 D	28-05-1997
		DE 69310207 T	14-08-1997
		EP 0673439 A	27-09-1995
		ES 2102811 T	01-08-1997
		JP 8504101 T	07-05-1996
		US 5683868 A	04-11-1997
		ZA 9309078 A	05-06-1995
WO 9417208 A	04-08-1994	AT 160178 T	15-11-1997
		AU 6124394 A	15-08-1994
		DE 69406772 D	18-12-1997
		DE 69406772 T	12-03-1998
		EP 0680517 A	08-11-1995
		ES 2111289 T	01-03-1998
		JP 9503121 T	31-03-1997
DE 19544374 C	20-02-1997	DE 19549417 A	05-06-1997
WO 9215687 A	17-09-1992	AT 165394 T	15-05-1998
		DE 69225221 D	28-05-1998
		EP 0573500 A	15-12-1993
WO 9741251 A	06-11-1997	AU 5648296 A	19-11-1997
WO 9821347 A	22-05-1998	NONE	

COR-AAD1-00002016

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/82, 5/14, A01H 5/00

A1

(11) International Publication Number: **WO 98/44139**

(43) International Publication Date: 8 October 1998 (08.10.98)

(21) International Application Number: PCT/US98/06109

(22) International Filing Date: 27 March 1998 (27.03.98)

(30) Priority Data:

60/042,852	28 March 1997 (28.03.97)	US
60/054,422	31 July 1997 (31.07.97)	US

(71) Applicant: MYCOGEN PLANT SCIENCE, INC. [US/US]; 5501 Oberlin Drive, San Diego, CA 92121 (US).

(72) Inventors: HORN, Michael, E.; 850C Indian Trails, Carmel, IN 46032 (US). HALL, Gerald, E., Jr.; 142 Rice Drive, Morrisville, PA 19067 (US).

(74) Agents: LLOYD, Jeff et al.; Saliwanchik, Lloyd & Saliwanchik, Suite A–1, 2421 N.W. 41st Street, Gainesville, FL 32606–6669 (US).

(81) Designated States: AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, GW, HU, ID, IL, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), 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
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMPROVED SAR PLANT TRANSFORMATION PROCESS

CGATTAAAAA TCCCAATTAT ATTTGGTCTA ATTTAGTTTG GTATTGAGTA AAACAAATTC
GAACCAAACC AAAATATAAA TATATAGTTT TTATATATAT GCCTTTAAGA CTTTTTATAG
AATTTTCTTT AAAAAATATC TAGAAATATT TGCGACTCTT CTGGCATGTA ATATTTCGTT
AAATATGAAG TGCTCCATTT TTATTAACTT TAAATAATTG GTTGTACGAT CACTTTCTTA
TCAAGTGTTA CTAAAATGCG TCAATCTCTT TGTTCTTCCA TATTCATATG TCAAAATCTA
TCAAAATTCT TATATATCTT TTTCGAATTT GAAGTGAAAT TTCGATAATT TAAAATTAAA
TAGAACATAT CATTATTTAG GTATCATATT GATTTTTATA CTTAATTACT AAATTTGGTT
AACTTTGAAA GTGTACATCA ACGAAAAATT AGTCAAACGA CTAAAATAAA TAAATATCAT
GTGTTATTAA GAAAATTCTC CTATAAGAAT ATTTTAATAG ATCATATGTT TGTAAAAAAA
ATTAATTTTT ACTAACACAT ATATTTACTT ATCAAAAATT TGACAAAGTA AGATTAAAAT
AATATTCATC TAACAAAAAA AAAACCAGAA AATGCTGAAA ACCCGGCAAA ACCGAACCAA
TCCAAACCGA TATAGTTGGT TTGGTTTGAT TTTGATATAA ACCGAACCAA CTCGGTCCAT
TTGCACCCCT AATCATAATA GCTTTAATAT TTCAAGATAT TATTAAGTTA ACGTTGTCAA
TATCCTGGAA ATTTTGCAAA ATGAATCAAG CCTATATGGC TGTAATATGA ATTTAAAAGC
AGCTCGATGT GGTGGTAATA TGTAATTTAC TTGATTCTAA AAAAATATCC CAAGTATTAA
TAATTTCTGC TAGGAAGAAG GTTAGCTACG ATTTACAGCA AAGCCAGAAT ACAAAGAACC
ATAAAGTGAT TGAAGCTCGA AATATACGAA GGAACAAATA TTTTTAAAAA AATACGCAAT
GACTTGGAAC AAAAGAAAGT GATATATTTT TTGTTCTTAA ACAAGCATCC CCTCTAAAGA
ATGGCAGTTT TCCTTTGCAT GTAACTATTA TGCTCCCTTC GTTACAAAAA TTTTGGACTA
CTATTGGGAA CTTCTTCTGA AAATAGT

(57) Abstract

The subject invention pertains to the use of scaffold attachment regions (SARs), also known as matrix attachment regions (MARs), in expression cassettes to improve the transformation efficiency of such expression cassettes in a transformation process.

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

DESCRIPTION

IMPROVED SAR PLANT TRANSFORMATION PROCESS

Field of the Invention

The present invention relates to the use of scaffold attachment regions (SARs), also known as matrix attachment regions (MARs), in transformation of cells and tissues.

Background of the Invention

SARs/MARs (hereinafter collectively referred to as "SARs") are AT-rich genomic DNA sequences that occur in eukaryotic genomes (see Boulikas [1993] *J. Cell Biochem.* 52:14). SARs are non-coding DNA sequences that flank structural genes and stabilize the transcription process. It is also known that SARs bind to certain components of the proteinaceous scaffold material that encompasses nuclear DNA. SARs have been found to improve the expression of heterologous genes in transformed plants (Allen *et al.* [1993] *Plant Cell* 5:603).

Brief Summary of the Invention

It has been unexpectedly discovered that the use of SARs in plant expression cassettes increases the frequency of recovery of stable transformation events of such expression cassettes in a plant transformation process. In particular, stable transformation events have been observed to increase up to about 7-fold or more while, at the same time, in the case of GUS expression, the transient transformation events have been observed to decrease. This startling observation makes the increase in stable transformation events all the more surprising. This invention is particularly useful in improved electroporation transformation techniques.

It has also been unexpectedly discovered that the use of SARs in plant expression cassettes advantageously decreases the time required to recover a stably transformed plant. The SAR-positive clones invariably appear in culture faster as compared to SAR-negative clones which, in turn, allows for quicker regeneration. Shorter times in culture leads to easier regeneration, less abnormalities in the plants, higher fertility, better seed set, *etc.* Quicker transformation/regeneration processes are highly desirable from a

commercial standpoint due to the fact that shorter time periods allow for dramatically increased throughput of events which result in quicker identification of commercial candidates.

Additionally, the use of SARs in a plant expression cassette employed in electroporation transformation processes results in a high proportion of low copy number transformants which is very desirable from a commercial standpoint. High copy number transformants are associated with high frequencies of gene silencing. Complex Southern blot patterns that are indicative of high copy number insertions have been observed rarely (approximately 10%) in SAR-positive maize transformants from an electroporation process of whole intact maize cells.

Brief Description of the Drawings

Figure 1 shows Southern blots of maize callus samples transformed with SAR-positive plant expression cassettes compared with controls. Treated callus was plated on bialaphos selection and bialaphos-resistant colonies appeared after 7 to 12 weeks. Callus was extracted for DNA and separated on an agarose gel as per Materials and Methods. Lane 1: DNA sizing standards; Lanes 2-12: DNA from selected colonies arising on bialaphos selection; Lanes 13: DNA from PAT+ control callus known to contain the PAT gene; Lane 14: DNA from PAT- control callus known not to contain the PAT gene; Lane 15: plasmid DNA corresponding to the equivalent of 1 copy; Lane 16: plasmid DNA corresponding to the equivalent of 3 copies.

Figure 2 shows Southern analysis of DNA from regenerated T_0 plants from callus (Fig. 1) derived from transfection with pAGM 607 plasmid DNA. Lane 1: DNA sizing standards; Lane 2: plasmid DNA corresponding to the equivalent of 1 copy; Lanes 3-11: not relevant; Lane 12: DNA from PAT+ control leaf tissue known to contain the PAT gene; Lane 13: DNA from PAT- control leaf tissue known not to contain the PAT gene; Lanes 15-27: DNA from leaves of T_0 plants regenerated from Southern positive callus.

Figure 3 shows Southern analysis of DNA from leaf tissue from plants segregating for the PAT gene in the ST2-1 derived T_1 generation. Lane 1: DNA sizing standards; Lane 2: blank; Lanes 3-11: DNA from leaf tissue of plants judged to be sensitive to the herbicide; Lane 22: DNA from PAT- control leaf tissue known not to

contain the PAT gene; Lane 23: Lane 12: DNA from PAT+ control leaf tissue known to contain the PAT gene; Lane 24: plasmid DNA corresponding to the equivalent of 1 copy; Lane 25: plasmid DNA corresponding to the equivalent of 3 copies; Lanes 26-28: not relevant.

Figure 4 shows a nucleotide sequence of a SAR polynucleotide which can be used according to the present invention.

Figure 5 shows a restriction map of plasmid pAGM243.

Figure 6 shows a restriction map of plasmid pAGM285A.

Figure 7 shows a restriction map of plasmid pAGM607.

Figure 8 shows a restriction map of plasmid pAGM608.

Detailed Disclosure of the Invention

The present invention concerns methods and materials for increasing frequency of recovery of stable transformation events in plant transformation processes and for increasing the number of low copy number transformants, as well as reducing or eliminating the occurrence of gene silencing throughout subsequent generations descended from the original transformant. In one embodiment of the present invention, SARs are used in a plant expression cassette to provide a plant transformation process that produces a greater percentage of stable transformants and a greater percentage of low copy number transformants than are obtained with SAR-negative plant expression cassettes.

In one embodiment, a method of the subject invention comprises transforming a cell or tissue with a SAR polynucleotide sequence operably linked to a polynucleotide of interest that encodes a protein, polypeptide, or peptide. SARs useful with the subject invention include, but are not limited to, SARs originating from plants and animals. SAR-encoding polynucleotides useful with the subject invention include, for example, SAR isolated from tobacco (Hall *et al.*, [1991] *Proc. Natl. Acad. Sci. USA* 88:9320). In a preferred embodiment, a SAR polynucleotide useful in practicing the present invention comprises the nucleotide sequence shown in Figure 4, or a functional fragment or mutant thereof. Preferably, the SAR polynucleotide sequence is provided in the form of an "expression cassette" on a suitable vector. Any vector suitable for DNA mediated transformation can be used and such vectors are known in the art. In a highly preferred

embodiment, a SAR polynucleotide sequence is operably linked at either and or both ends of the polynucleotide of interest. Transformation can be accomplished using known methods, including, for example, particle bombardment or biolistics transformation with DNA coated microparticles, *Agrobacterium*-mediated transformation, electroporation, microinjection, magnetophoresis, silicon carbide whiskers, PEG mediated transformation, and protoplast transformation. In a preferred embodiment, plant cells are transformed with the polynucleotides of the invention by electroporation according to standard techniques known in the art. See, for example, Pescitelli, S. M., K. Sukhapinda (1995) "Stable Transformation via Electroporation Into Maize Type II Callus and Regeneration of Fertile Transgenic Plants" *Plant Cell Reports* **14**: 712-716 (hereby incorporated by reference in its entirety). Transformed plant cells can be selected and then cultured under suitable conditions according to routine practice to generate transformed plantlets and plants.

It is well known in the art that the nucleotide sequences of the subject invention can be truncated and/or mutated such that certain of the resulting fragments and/or mutants of the original full-length sequence can retain the desired characteristics of the full-length sequence. A wide variety of restriction enzymes are well known by ordinarily skilled artisans which are suitable for generating fragments from larger nucleic acid molecules. In addition, it is well known that *Bal*31 exonuclease can be conveniently used for time-controlled limited digestion of DNA. See, for example, Maniatis *et al.* (1982) *Molecular Cloning: A Laboratory Manual*, Coldspring Harbor Laboratory, New York, pages 135-139, incorporated herein by reference. See also Wei *et al.* (1983) *J. Biol. Chem.*, **258**:13006-13512. By use of *Bal*31 exonuclease (commonly referred to as "erase-a-base" procedures), the ordinarily skilled artisan can remove nucleotides from either or both ends of the subject acids to generate a wide spectrum of fragments which are functionally equivalent to the subject nucleotide sequences. One of ordinary skill in the art can in this manner, generate hundreds of fragments of controlled, varying length from locations all along the original SAR molecule in one afternoon. The ordinarily skilled artisan is able to routinely test or screen the generated fragments for their characteristics for determining the utility of the fragments as taught herein. It is also well known that mutant sequences of the full length sequence, or fragments thereof, can be easily produced with site directed mutagenesis. See, for example, Larionov, O.A. and

Nikiforov, V.G. (1982) "Directed Mutagenesis" *Genetika* **18(3)**:349-59; Shortle, D., DiMaio, D., and Nathans, D. (1981) "Directed Mutagenesis" *Annu. Rev. Genet.* **15**:265-94; both incorporated herein by reference. The skilled artisan can routinely produce deletion-, insertion-, or substitution-type mutations and test whether the resulting mutants contain the desired characteristics of the full length wild-type sequence, or fragments thereof.

In one embodiment of the invention, a polynucleotide comprising a polynucleotide sequence that has substantially the same sequence as a SAR polynucleotide is operably linked to a polynucleotide that encodes a protein, polypeptide, or peptide. Any desired polynucleotide sequence can be employed to transform cells or tissue according to the present invention. Of particular importance in this regard are genes used, for example, as (i) selective markers (antibiotic and/or herbicide resistance genes), (ii) reporter genes (*e.g.*,GUS), (iii) insecticide resistance genes (*B.t.* delta endotoxins) and (iv) any other genes that improve the value or use of a plant. Such genes can include, but are not limited to, glucuronidase, phosphinothricin, N-acetyltransferase, green fluorescent protein (GFP), luciferase, Pat/bar, and glyphosate resistance genes (NPTII, HPT, biomoxila resistance gene, AHAS, ALS, cyanomide hydrolase, adenine deaminase, 2,4-D monooxygeanse). The encoded protein, polypeptide or peptide can be one that is naturally present in the transformed cell or it can be heterologous to the transformed cell. Optionally, a regulatory sequence, such as, for example, a promoter sequence that can regulate transcription of the polynucleotide sequence, can be included in a polynucleotide of the invention. In a preferred embodiment, the SAR polynucleotide sequence is a polynucleotide component of an expression cassette on a suitable vector. Vectors useful with the SAR polynucleotides of the invention are known in the art and can be prepared and/or selected according to standard techniques.

The methods and materials of the present invention can be used to transform cells or tissue from any organism, and preferably from a eucaryotic organism. In a preferred embodiment, the cells are plant cells. Any plant cell competent to be transformed can be employed in the present invention. It is preferred to employ plant cells that are readily regenerable into whole plants. Suitable plant cells include embryogenic suspension cells, non-embryogenic suspension cells (except in corn where these cells are not regenerable), plant explants, germline cells (pollen, ovules, meristem domes, megaspore cells,

embryos cells, egg cells and embryosacs), microspore cells and callus tissue cells, both compact callus and friable callus. Preferred cells include embryogenic callus, suspension cells (embryogenic suspension) and callused immature zygotic embryos. Particularly preferred plant cells are early embryogenic suspension and young callus (still attached to the zygotic embryo) cells of from about 3-14 and preferably from about 5-10, days old. Wounding the plant cells and/or treating the plant cells prior to being subjected to electroporation is unnecessary and is in fact expressly avoided. When embryogenic suspension tissues are employed it is preferred to gently break up the tissue into small clumps or into fine aggregates as small as possible without damaging the embryogenic suspension cells. This can be done by sieving the tissue through a screen, such as, for example, pushing tissue through a 1,000 micron (u) sieve with a spatula or pestle. The finer the aggregates or suspension of undamaged cells, the more efficient the present process.

Plant tissue useful with the invention includes, but is not limited to, callus, meristematic, leaf, shoot, root, and embryonic tissue. The present invention is applicable to any plant species including angiosperms (dicots, monocots) and gymnosperms. Suitable crops include corn, wheat (especially Type C wheat callus), sorghum, rice, pearl millet, sugar cane, orchardgrass and other Gramineae plants; soybean, peanuts, alfalfa and other members of the Luguminoseae family; cotton, kenaf, and other members of the Malvaceae family; poppy and other members of the Papavaraceae family; cannabis and other members of the Cannabinaceae family; tea and other members of the Theaceae family; rape (canola), vegetables and oilseed crops and other members of the Cruciferae family; sunflower, safflower and other members of the Compositae family; coffee and other members of the Rubiaceae family; cacao, theobroma and other members of the Byttneriaceae family; fruits and vegetables, trees, orchard crops, and turf grass. Preferred crops include cotton, tomato, sugarbeet, potato, peanut, alfalfa, rice, wheat and especially corn (maize).

In a preferred embodiment, the polynucleotide sequences employed in the present invention comprise any sequences which have a 5′ promoter region, a structural gene region and a 3′ nontranslated region (polyadenylation site) which can be expressed in plants. The polynucleotide sequences can be modified in any manner (extra codons, deletion of codons, changed codons, *etc.*) as long as gene expression is not prohibited.

The polynucleotides inserted into the plants according to the present invention can include any desired gene whether eukaryotic or procaryotic in nature. Usually, more than one gene will be inserted into plant cells which are transformed for agronomic purposes. One gene will typically be a selective marker gene (antibiotic resistance gene or an herbicide resistance gene) in order to easily detect transformants from non-transformed cells. Additional genes can also be added to the plant cell genome to impart an additional property, to suppress an existing property (via "antisense" mechanism) or to amplify a known property of the plant cells and the whole plants regenerated therefrom. The genes can be expressed in specific tissues by the use of tissue specific promoters. The genes can be constructed according to techniques well known to one skilled in the art.

Gene constructs may exist as single gene expression cassettes comprised of a promoter, a structural gene coding sequence and a sequence to permit the addition of poly-adenine (poly-A) residues. The promoter is necessary to initiate transcription of the DNA coding for the structural gene into RNA. The promoter may be derived from a variety of sources, as long as it is functional in the cells to be transformed, and may be modified to enhance expression by the addition or deletion of sequences. The DNA may contain intron sequences, either outside or within the coding region for the protein. The removal of these introns and the addition of the poly-A sequence results in the production of a mature messenger RNA (mRNA) which can be translated into the corresponding protein. Gene expression cassettes may be linked in groups of two or more. Polycistronic expression cassettes, in which a single mRNA may code for more than one protein, may also be used. In addition, expression cassettes may be used to produce an "antisense" RNA from the transcription of a strand of DNA which is opposite to the strand of DNA coding for a protein.

Examples of promoters active in plants include maize ubiquitin promoter (Christensen *et al.*, [1992] *Plant Molecular Biology* 18: 675-689). The untranslated leader sequence, including the first intron, of the maize ubiquitin gene may be incorporated, particularly for use in monocot cells. The 35S promoter of Cauliflower Mosaic Virus (Murray *et al.*, [1991] *Plant Molecular Biology* 16: 1035-1050) or the T-DNA Mas2 promoter of the mannopine synthase gene (Leung *et al.*, [1991] *Molecular & General Genetics* 230: 463-474) may also be used. The 35S promoter may contain a deletion with the addition of an upstream enhancer sequence and an intron in the

untranslated leader region (Last *et al.*, [1991] *Theoretical & Applied Genetics* 81: 581-588) and the Mas2 promoter may also contain a deletion to enhance expression (Leung *et al.*, supra).

Examples of structural genes include reporter genes such as that coding for GUS, or -glucuronidase, (Jefferson *et al.*, [1987] *EMBO Journal* 6:3901-3907), a selectable marker gene such as that coding for PAT, or phosphinothricinN-acetyltransferase,which confers resistance to the active ingredient of the commercial herbicide Basta (Droge *et al.*, [1992] *Planta* 187:142-151), or genes which result in expression of a value-added phenotypic trait. Examples of the latter type gene includes those derived from *Bacillus thuringiensis* (B.t.) which confer resistance to insects such as lepidoptera (Adang *et al.*, [1985] *Gene* 36:289-300) or coleoptera (Bradfisch *et al.*, EP 0 500 311 A2 and U.S. Patent No. 5,208,017) and designated B.t.2 and B.t.3 in Figure 1. The B.t. genes may be reconfigured to enhance their expression in plant cells (Adang *et al.*, U.S. Patent No. 5,380,831 issued January 10, 1995).

While herbicide resistance genes serve the purpose of selecting transformants, they also serve the valuable agronomic purpose of allowing herbicide use in the field in otherwise sensitive crops and/or preventing damage to otherwise sensitive crops planted to fields wherein those herbicides were used earlier in the field for weed control (Herbicide carryover). Additional types of genes of value for use in plants include genes isolated from Bacillus thuringiensis that code for delta-endotoxins as well as truncated and/or synthetic derivatives thereof; fungal resistance genes; oil biosynthesis genes; anti-sense genes and genes responsible for nutritional and/or fiber quality.

The DNA and cells are reacted according to the present invention in a suitable buffered medium that is preferably iso-osmotic. See for example, Wong and Neumann's F-medium, Biochemistry and Biophysics Research Communications, Vol. 107, pp. 584-587, 1982; Krens *et al.* [1981] *Nature* 296:72-74. A preferred buffer medium is EPR Buffer (555 mM glucose, 4 mM$CaCl_2$, 10 mM Hepes buffer, pH 7.2). When conducting the present process the fresh plant cells are added to a reaction vessel followed by the addition of a mixture of DNA and buffer medium. The order of addition of the materials is not critical.

The temperature at which the electroporation process takes place is not critical although it is preferred to cool the DNA/plant cell mixture immediately prior to and after

the application of the electric field. Any culturing or regenerating steps are conducted under conditions (including temperature) well known to those of ordinary skill in the art. Heat shock treatments, i.e., 37° C for 10 minutes, of the DNA and plant cell mixture can also be employed.

The electroporation step of the present invention is accomplished by applying an electric field to the DNA/cell mixture according to well-known techniques. Any electric field can be employed. Electric pulses can be from 25-5,000 volts (V) or more depending on the current employed. Preferred methods include rectangular pulse generating systems and capacitor discharge systems. The capacitor discharge system creates pulses of exponentially decaying voltages.

In practicing the present invention, DNA and the plant cells are incubated together at room temperature for at least about 10 minutes and preferably for 20-90 minutes. The DNA/cell mixture is then transferred in aliquots to electroporation cuvettes and optionally cooled on ice prior to applying an electric field to it. The electric field strength can vary depending on a variety of factors, such as, for example, the particular plant species being transformed, the particular type (including age) of cells being employed in the transformation process, the type of electric field being employed including the length of exposure time of the plant cells to the electric field, the concentration and type of DNA, etc. One of ordinary skill in the art can easily determine the optimum process conditions by employing routine titration experiments.

With a capacitor discharge system in the transformation of corn callus cells preferred electroporation conditions include a 250-1500 μF capacitor, 25-500 or more volts and a pulse time of from 50-500 msec. Discharge should be from 25-250 volts. Especially preferred conditions are 850 μF, 150 V and a pulse time of 200 msec. Immediately after application of the electric field the electroporation cuvettes can be optionally placed on ice for about 10 minutes. The cuvettes are then allowed to stand at room temperature for at least about 5 minutes and a small aliquot of cell culture medium is added thereto. Samples are then pipetted from the cuvettes and placed in 2 ml of culture medium in a well of a six-well plate. The treated cells are maintained in culture and regenerated employing standard culturing and plants regenerating techniques.

SAR polynucleotides contemplated within the scope of the present invention encompass known SARs, including functional fragments and allelic variants of a SAR,

as well as any SAR that may be identified in the future so long as the SAR retains substantially the same biological activity as SARs exemplified herein. SARs can be prepared from natural sources or synthesized using standard techniques known in the art, such as an automated DNA synthesizer. The SAR polynucleotides of the subject invention also encompass variant sequences containing mutations in the natural sequences. These mutations can include, for example, nucleotide substitutions, insertions, and deletions as long as the variant SAR sequence retains substantially the same biological activity as the natural SAR sequences of the present invention.

The subject invention also concerns cells and tissue transformed using the methods of the invention. Plants, plantlets, and plant seeds transformed to express heterologous genes according to the methods of the described herein are also contemplated within the scope of the invention.

The disclosure of all references cited throughout the present specification is incorporated herein by reference.

Materials and Methods

PAT-ELISA

Nunc Maxisorb Microtiter plates were coated with 100μl of 1μg/ml protein A purified rabbit anti-PAT IgG in Coating Buffer for two hours while gently shaking at room temperature. After discarding the coating solution, 400μl blocking solution was added and the plates were blocked overnight at 4°C while gently shaking. The plates were then washed using a Skatron SkanWasher plate washer and 100ml of antigen or cell extract was added to the wells and allowed to incubate for two hours at room temperature while gently shaking. Extracts were made by grinding ~100 mgfw of tissue in five volumes Plant Extraction Buffer (PEB) for callus or ten volumes for leaf material and microfuging the crude extract for 5 min. Then the supernatant was transferred and microfuged again for 5 min. Extracts were then diluted 10-fold with grinding buffer prior to adding to the microtiter plate. After washing the plates, 100μl of 1μg/ml protein A purified goat anti-PAT IgG in Ab Buffer was added and the plates incubated for one hour at room temperature with gentle shaking. The plates were washed and 100μl of 1:30,000 dilution of anti-goat antibody conjugated to alkaline phosphatase (Pierce) was added and the

plates incubated for one hour at room temperature with gentle shaking. After washing, 200ml of SIGMA FAST pNPP substrate was added and the plates were allowed to sit without shaking for one to three hours with 595nm (reference 405nm) absorbency readings taken every hour. Quantitation was achieved by comparing the absorbencies with a standard curve made by spiking pure PAT protein into a 10-fold dilution of negative control extract.

Southern Blotting and Hybridization Procedure

Southern Blotting and Pre-hybridization: Southern blotting methodology was performed essentially as described in Murray, M., *et al.*, [1992] *Plant Molecular Biology Reporter*, vol. 10(2). Briefly, genomic DNA (5-10ug) which has been digested with the appropriate restriction enzyme(s) and resuspended in 1X loading buffer is loaded into an agarose/TAE gel (0.85%). The DNA is separated by electrophoresis (75 W/4h) and the gel is then stained (0.1 ug/ml EtBr in 10 mM NaPO) for 30 min. and photographed. The gel is then denatured for 20 min. (150 mM NaPO). The separated DNA is then transferred onto nylon membrane via capillary action overnight. The nylon membrane is then baked for 2 h at 80°C, blocked for 2 h (2% SDS, 0.5% BSA. 1 mM EDTA, 1 mM Orthophenanthroline) and allowed to pre-hybridize for 2 h (100 mM Na phosphate buffer (pH 7.8), 20 mM Na pyrophosphate, 5mM EDTA, 1 mM 1,10 orthophenanthroline, 0.1% SDS, 10% dextran sulfate 500 ug/ml heparin sulfate, 50 ug/ml yeast RNA, 50 ug/ml herring sperm DNA).

Hybridization: DNA template to be used as a probe is labeled with P dCTP using a Prime-It RmT Random Priming Labeling Kit (Stratagene). Labeling efficiency of the probe was measured and approximately 1 X 10^6 CPM/ML is added to the prehybridized membrane. The membrane is hybridized with the probe at 65°C for 12-16 h. The membranes are then washed 3X to remove unbound probe (5mM Na phosphate (pH 7.8), 1.25 mM Na pyrophosphate, 0.25 mM EDTA, 0.1% SDS) and exposed to Kodak scientific imaging film.

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1

Several experiments were conducted to determine the transformation efficiency, both transient and stable, for maize cells transformed with and without SARs in an expression cassette. The transformation process conducted was an electroporation process of whole intact maize cells. The results of this experiment are listed in Table 1 below. As can be seen, the SAR-positive clones (+SARs) had a lower transient transformation efficiency (19.7% vs. 29.49%) but a much higher stable transformation efficiency (3.62% vs. 0.51%).

Table 1

Expt. #	Transient GUS+SARs	Transient GUS−SARs	Stable +SARs	Stable −SARs
ST-1	NA	NA	4/150	--
ST-2	NS	NA	14/135	1/135
ST-3	12.6	29.3	1/25	0/25
ST-4	10.4	32.7	0/20	0/20
ST-5	16.1	37.2	0/25	0/25
ST-6	19.5	28.9	3/90	1/90
ST-7	38.8	28.7	0/50	0/50
ST-8	31.6	36.9	3/155	1/150
ST-9	8.8	12.7	2/95	0.95
TOTALS	137.8	206.4	27/745	5/390
Average	19.7	29.49	3.62%	0.51%

Example 2

ELISA for the detection of phosphinothricin acetyltransferase (PAT) was conducted on several SAR-negative and SAR-positive clones that contained the PAT gene as described in the *Materials and Methods* section herein. Clone T34-1 represents a positive control that was one of the highest PAT-expressing SAR-negative clones. The results of this analysis are shown in Table 2. It can be seen that several SAR-positive clones produce high levels of PAT protein.

Table 2

Callus #	Prot. Adj.	ng PAT	Adj. Value	% of Protein
ST9 neg C	0.017	0.000	0.000	0.0000%
ST9-2 (-SARs)	0.149	-0.010	-0.644	-0.0001%
ST9-3 (-SARs)	0.145	-0.019	-1.314	-0.0001%
ST9-4 (+SARs)	0.099	0.128	13.001	0.0013%
ST9-5 (+SARs)	0.158	0.097	6.140	0.0006%
ST2-2 (+SARs)	0.110	0.025	2.286	0.0002%
ST2-3 (+)	0.169	0.095	5.626	0.0006%
ST2-4 (+)	0.307	0.116	3.772	0.0004%
ST2-5 (+)	0.093	0.039	4.140	0.0004%
ST2-6 (+)	0.160	0.070	4.355	0.0004%
ST2-7 (+)	0.179	0.053	2.934	0.0003%
ST2-8 (+)	0.015	0.018	12.146	0.0012%
ST2-9 (+)	0.165	0.028	1.678	0.0002%
ST2-10 (+)	0.212	0.029	1.381	0.0001%
T34-1 (-)	0.041	0.009	2.210	0.0002%

Example 3

A Southern analysis of several transformed maize callus clones was conducted for the detection of the PAT gene as described in the *Materials and Methods* section herein. The restriction enzyme used was *Bam*HI and the probe was PAT (PCR). Results of Southern blotting experiments, including controls, are shown in Figures 1-3. The results in Fig. 1 show that all the selected calli contained the PAT gene since bands appeared when incubated with a PAT-specific radioactive probe. Figure 2 shows that plants regenerated from selected calli form plants which also contain the PAT gene. Most of the lines show one or a few inserted copies of the PAT gene while one, Lane 24 or ST8-1, showed a complex banding pattern indicative of many PAT gene insertions. Southern analysis of DNA from leaf tissue from plants segregating for the PAT gene in the ST2-1 derived T_1 generation indicated that, with two exceptions, all plants thought to be herbicide sensitive were negative for the presence of the PAT gene (Figure 3). With no exceptions, all plants thought to be herbicide resistant did show the presence of the PAT gene. This data also shows that the inserted gene was capable of passing through meiosis and into the T_1 progeny.

Example 4

Immature zygotic embryos were isolated from 12-day old cobs of the genotype HiII. They were plated on callus initiation medium (CIM), scutellum side up, which contained N_6 salts, 1 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D), and 10μM $AgNO_3$, incubated in the dark for 8 days at 27-29°C and then examined for Type II callus formation. 250 callused embryos were chosen and placed into sterile 1.5ml micro-cuvettes, at five embryos per cuvette. 20μl of AGM285A plasmid DNA (1μg/μl) was introduced, followed by 200μl of EPM buffer (80 mM KCl, 5 mM $CaCl_2$, 10 mM hepes, 0.425 M mannitol, pH 7.2). One cuvette was inoculated with pAGM243 (comprising the gene construct Ubi-*pat*-Nos::Ubi-*gus*-Nos). Another cuvette contained no plasmid DNA. The contents of the cuvettes were agitated and then allowed to stand for 1 hr at room temperature. Electroporation was carried out using the gene ZAPPER 450/2500 (IBI) at 850μF, 150V for one pulse. The cuvettes were then allowed to stand for 20 min at room temperature before removing the embryos and replating on CIM, scutellum side up, and returned to the dark at 27-29°C. Three days later, the embryos treated with pAGM243

were placed into GUS stain and 9 days later were observed for GUS+ events. Those embryos showed 242 GUS+ events or 48.4 GUS+ events per embryo.

Seven days after treatment, all pAGM285A-treated embryos and embryos subjected to electroporation in the absence of plasmid DNA were transferred to callus maintenance medium (CMM) containing N_6 salts, 1 mg/L 2,4-D and 5μM bialaphos. These were transferred again to the same medium type after 22 days. About three weeks later, the tissue was transferred again.

Example 5

Field trials with the maize-SAR+ material showed that 13 out of 19 events gave segregation ratios consistent with heterozygous insertions (3:1 ratio in self crosses; 1:1 ratio in outcross to wild type).

SAR Event	Mendelian Ratios
ST2-02	Yes
ST2-04	No
ST2-05	Yes
ST2-06	Yes
ST2-09	Yes
ST2-10	No
ST8-01	Yes
ST8-02	No
ST9-04	Yes
ST9-05	No
ST2-04*CQ715	Yes
ST2-06*CQ715	Yes
ST2-10*CQ715	No
ST8-01*CQ715	Yes
ST9-04*CQ715	No
ST2-1*CQ806	Yes

ST2-05*CQ806	Yes
ST2-06*CQ806	Yes
ST9-05*CQ806	Yes

Moreover, the only SAR+ line tested showed excellent resistance to the Liberty herbicide in the third generation even at eight times the normal rate of field application (27.3 oz/acre of herbicide Remove). This result shows the extraordinary stability of high Pat gene expression even after several generations.

Normally, transgenic plants exhibit a relatively high frequency of gene silencing of the transgene as a given transgenic line is advanced from one generation to the next. Gene silencing is a major problem in the commercialization of transgenic crops. It occurs at a rate of less than 0.1% up to 3% and sometimes more. SARs have now been shown to prevent this phenomenon when used at the 3′ and 5′ ends of the inserted transgene. Transgenic maize lines containing Ubiquitin promoter-phosphinothricin acetyl transferase-now terminator (Ubi-pat-nos) with the Rb7 SAR regions attached at both ends prevent gene silencing in transgenic corn produced by the methods described in this patent or other known transformation methods over several generations in the field.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Horn, Michael E.

(ii) TITLE OF INVENTION: Improved SAR Transformation Process

(iii) NUMBER OF SEQUENCES: 1

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
(B) STREET: 2421 NW 41st Street, Suite A-1
(C) CITY: Gainesville
(D) STATE: FL
(E) COUNTRY: USA
(F) ZIP: 32606-6669

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/054,422
(B) FILING DATE: 31-JUL-1997

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/042,852
(B) FILING DATE: 28-MAR-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Lloyd, Jeff
(B) REGISTRATION NUMBER: 35,589
(C) REFERENCE/DOCKET NUMBER: MPS-408

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 352-375-8100
(B) TELEFAX: 352-372-5800

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1167 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGATTAAAAA TCCCAATTAT ATTTGGTCTA ATTTAGTTTG GTATTGAGTA AAACAAATTC
60

GAACCAAACC AAAATATAAA TATATAGTTT TTATATATAT GCCTTTAAGA CTTTTTATAG
120

AATTTTCTTT AAAAAATATC TAGAAATATT TGCGACTCTT CTGGCATGTA ATATTTCGTT
180

AAATATGAAG TGCTCCATTT TTATTAACTT TAAATAATTG GTTGTACGAT CACTTTCTTA
240

TCAAGTGTTA CTAAAATGCG TCAATCTCTT TGTTCTTCCA TATTCATATG TCAAAATCTA
300

TCAAAATTCT TATATATCTT TTTCGAATTT GAAGTGAAAT TTCGATAATT TAAAATTAAA
360

TAGAACATAT CATTATTTAG GTATCATATT GATTTTTATA CTTAATTACT AAATTTGGTT
420

AACTTTGAAA GTGTACATCA ACGAAAAATT AGTCAAACGA CTAAAATAAA TAAATATCAT
480

GTGTTATTAA GAAAATTCTC CTATAAGAAT ATTTTAATAG ATCATATGTT TGTAAAAAAA
540

ATTAATTTTT ACTAACACAT ATATTTACTT ATCAAAAATT TGACAAAGTA AGATTAAAAT
600

AATATTCATC TAACAAAAAA AAAACCAGAA AATGCTGAAA ACCCGGCAAA ACCGAACCAA
660

TCCAAACCGA TATAGTTGGT TTGGTTTGAT TTTGATATAA ACCGAACCAA CTCGGTCCAT
720

TTGCACCCCT AATCATAATA GCTTTAATAT TTCAAGATAT TATTAAGTTA ACGTTGTCAA
780

TATCCTGGAA ATTTTGCAAA ATGAATCAAG CCTATATGGC TGTAATATGA ATTTAAAAGC
840

AGCTCGATGT GGTGGTAATA TGTAATTTAC TTGATTCTAA AAAAATATCC CAAGTATTAA
900

TAATTTCTGC TAGGAAGAAG GTTAGCTACG ATTTACAGCA AAGCCAGAAT ACAAAGAACC
960

ATAAAGTGAT TGAAGCTCGA AATATACGAA GGAACAAATA TTTTTAAAAA AATACGCAAT
1020

GACTTGGAAC AAAAGAAAGT GATATATTTT TTGTTCTTAA ACAAGCATCC CCTCTAAAGA
1080

ATGGCAGTTT TCCTTTGCAT GTAACTATTA TGCTCCCTTC GTTACAAAAA TTTTGGACTA
1140

CTATTGGGAA CTTCTTCTGA AAATAGT
1167

Claims

I claim:

1. In a method for transformation of plant cells comprising inserting a polynucleotide of interest into plant cells and maintaining said cells under conditions whereby successfully transformed cells are identified and selected, the improvement comprising providing at least one scaffold attachment region polynucleotide sequence operably linked to said polynucleotide of interest whereby the frequency of recovery of stable transformation events is increased.

2. The improved method according to claim 1, wherein the polynucleotide of interest has at least one scaffold attachment region operably linked to each of its 5′ and 3′ ends.

3. The improved method according to claim 1, wherein said scaffold attachment region polynucleotide comprises the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

4. The improved method according to claim 1, wherein said scaffold attachment region polynucleotide is provided in the form of a DNA expression cassette which can be expressed in a plant.

5. The improved method according to claim 1, wherein said plant transformation process is selected from the group consisting of bombardment with DNA coated microparticles, *Agrobacterium*-mediated transformation, electroporation, microinjection, magnetophoresis, silicon carbide whiskers, PEG mediated transformation, and protoplast transformation.

6. The improved method according to claim 1, wherein said gene of interest is selected from the group consisting of antibiotic resistance genes, herbicide resistance

genes, reporter genes, marker genes, disease resistance genes, oil biosynthesis genes, anti-sense genes, insecticide resistance genes, and nutritional enhancement genes.

7. The improved method according to claim 1, wherein said plant cells are selected from the group consisting of embryogenic suspension cells, non-embryogenic suspension cells, plant explant cells, germ line cells, microspore cells, and callus tissue cells.

8. The improved method according to claim 1, wherein said plant cells are of a species selected from the group consisting of cotton, tomato, peanut, alfalfa, rice, wheat, corn, trees, orchard crops, fruits, vegetables, soybean, canola, turfgrass, cannabis, sunflower, sugarcane, sugarbeet, and potato.

9. The improved method according to claim 4, wherein said expression cassette comprises a 5'- promoter region, a structural gene region comprising said polynucleotide of interest, a 3'- nontranslated region, and a SAR flanking either or both sides of the said structural gene region.

10. The improved method according to claim 6, wherein said reporter gene is selected from the group consisting of glucuronidase, phosphinothricin N-acetyltransferase, green fluorescent protein, and luciferase.

11. The improved method according to claim 6, wherein said marker gene is selected from the group consisting of Pat/bar, glyphosate resistance genes, NPTII, HPT, bromoxil resistance genes, AHAS, ALS, cyamamide hydrotase, adenine deaminase, 2, 4-D monooxygenase.

12. A transformed plant cell produced by the method of claim 1.

13. The transformed cell according to claim 12, said cell comprising a scaffold attachment region polynucleotide having the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

14. A transgenic cell descended form the cell of claim 12.

15. A descendant transgenic cell of claim 14, said descendant cell comprising a scaffold attachment region having the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

16. A transgenic plant regenerated from a cell according to claim 12.

17. The transgenic plant according to claim 16, said plant comprising cells which comprise a scaffold attachment region having the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

18. A method for increasing the number of low copy number transformants in a plant transformation process, said method comprising transforming a plant cell or tissue with at least one scaffold attachment region polynucleotide sequence operably linked to a structural gene of interest.

19. The method according to claim 18, wherein said gene of interest has at least one scaffold attachment region operably linked to each of its 5′ and 3′ ends.

20. The method according to claim 18, wherein said scaffold attachment region polynucleotide comprises the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

21. The method according to claim 18, wherein said scaffold attachment region polynucleotide is provided in the form of a DNA expression cassette comprising a structural gene region which can be expressed in a plant.

22. The method according to claim 18, wherein said plant transformation process is selected from the group consisting of bombardment with DNA coated microparticles, *Agrobacterium* mediated transformation, and electroporation.

23. The method according to claim 18, wherein said gene of interest is selected from the group consisting of antibiotic resistance genes, herbicide resistance genes, reporter genes, fungal resistance genes, oil biosynthesis genes, anti-sense genes, and insecticide resistance genes.

24. The method according to claim 18, wherein said plant cells are selected from the group consisting of embryogenic suspension cells, non-embryogenic suspension cells, plant explant cells, germ line cells, microspore cells, and callus tissue cells.

25. The method according to claim 18, wherein said plant tissue is selected from the group consisting of callus, meristematic tissue, leaf tissue, shoot tissue, root tissue, and embryonic tissue.

26. The method according to claim 18, wherein said plant cell or tissue is of a species selected from the group consisting of cotton, tomato, peanut, alfalfa, rice, wheat, and corn.

27. The method according to claim 21, wherein said expression cassette comprises a five-prime promoter region, a structural gene region and a three-prime nontranslated region, and a SAR flanking either or both ends of said structural gene.

28. The method according to claim 23, wherein said reporter gene is select from the group consisting of glucuronidase and phosphinothricin N-acetyltransferase.

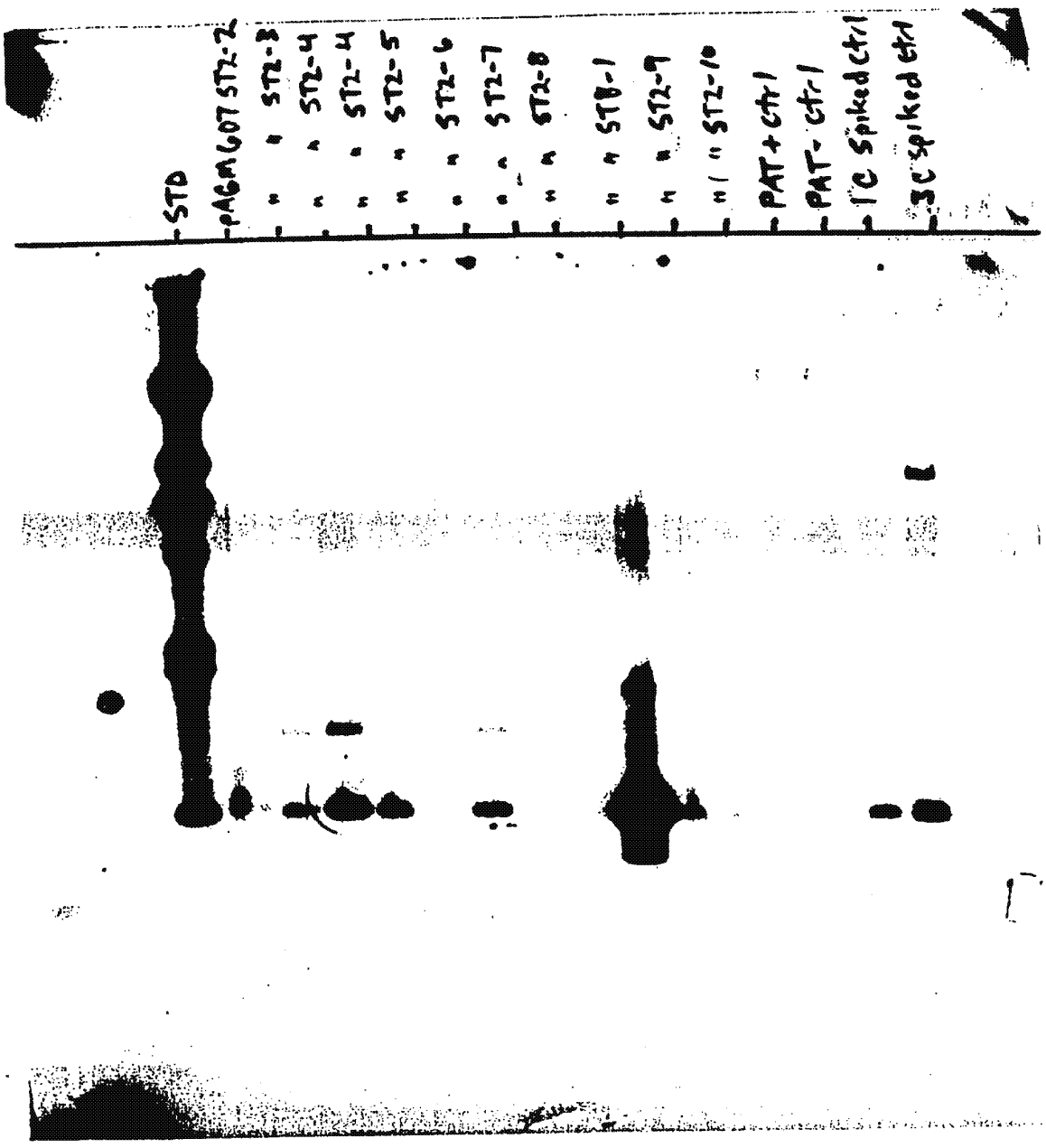

STD
pAGM607 ST2-2
" ST2-3
" ST2-4
" ST2-4
" ST2-5
" ST2-6
" ST2-7
" ST2-8
" ST2-1
" ST2-9
" ST2-10
PAT+ ctrl
PAT- ctrl
1C Spiked ctrl
3C Spiked ctrl

FIGURE 1

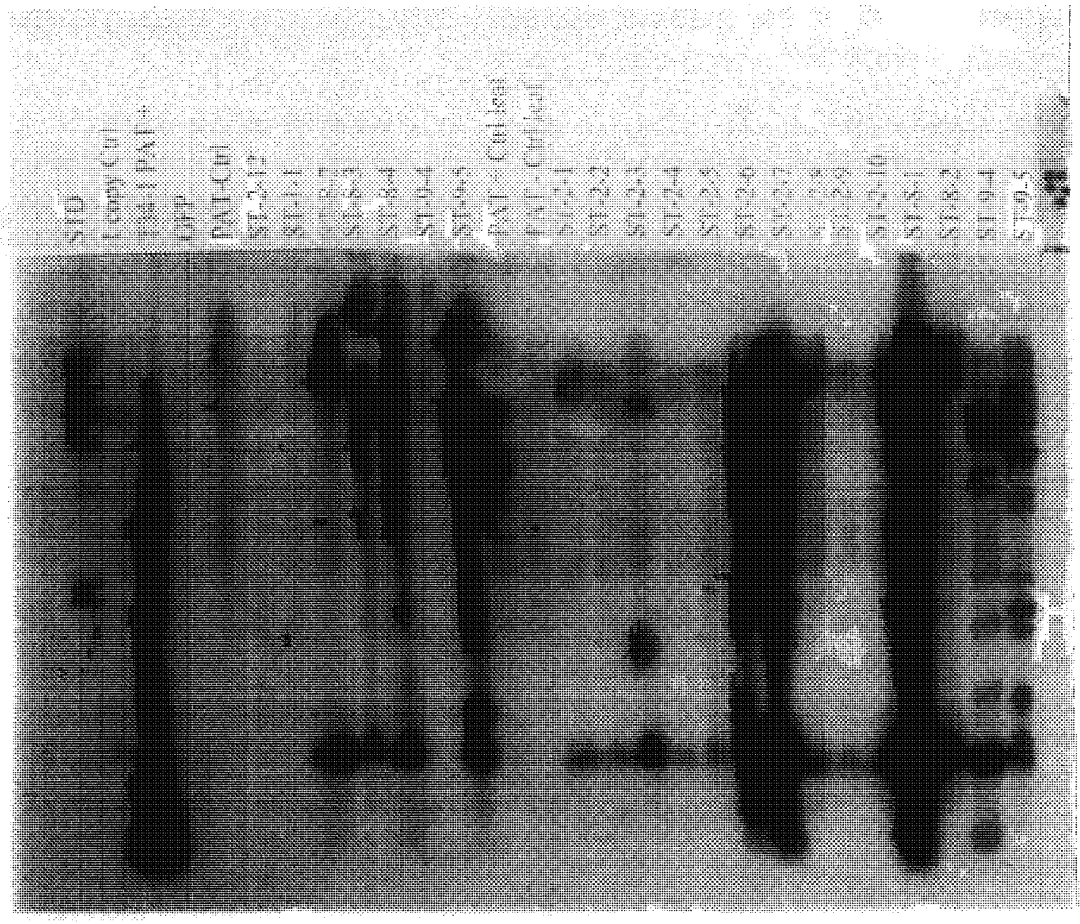


FIG. 2

COR-AAD1-00002043

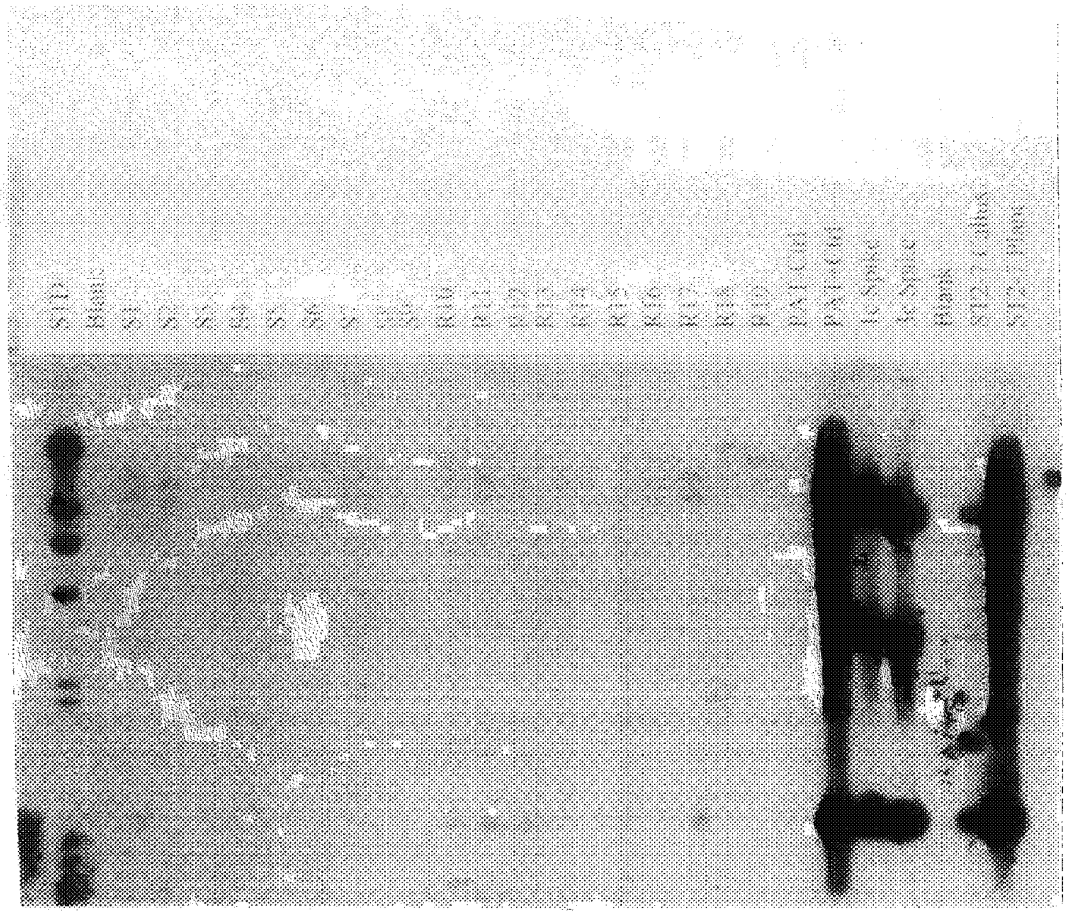


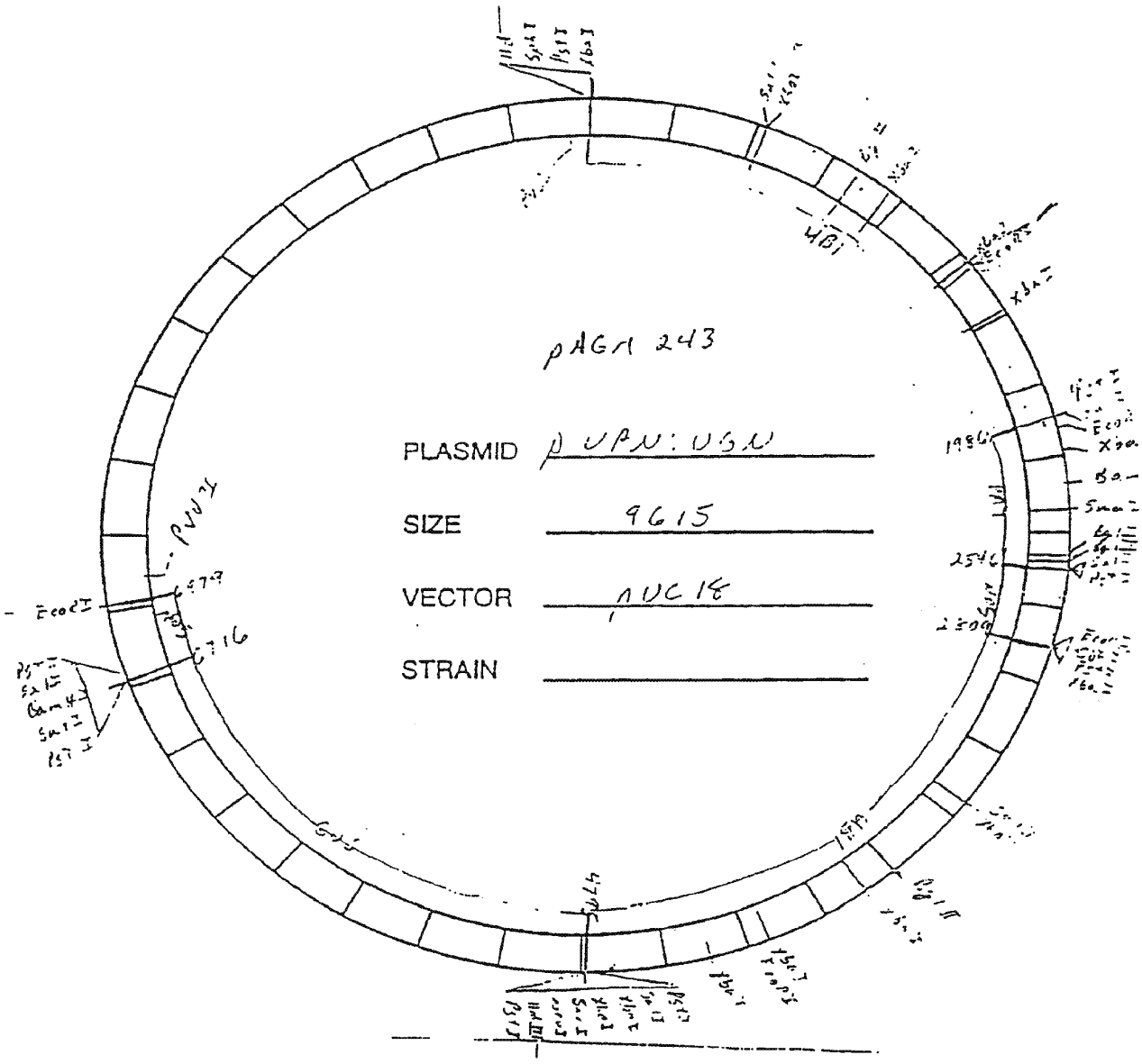
FIG. 3

COR-AAD1-00002044

SUBSTITUTE SHEET (RULE 26)

CGATTAAAAA TCCCAATTAT ATTTGGTCTA ATTTAGTTTG GTATTGAGTA AAACAAATTC
GAACCAAACC AAAATATAAA TATATAGTTT TTATATATAT GCCTTTAAGA CTTTTTATAG
AATTTTCTTT AAAAAATATC TAGAAATATT TGCGACTCTT CTGGCATGTA ATATTTCGTT
AAATATGAAG TGCTCCATTT TTATTAACTT TAAATAATTG GTTGTACGAT CACTTTCTTA
TCAAGTGTTA CTAAAATGCG TCAATCTCTT TGTTCTTCCA TATTCATATG TCAAAATCTA
TCAAAATTCT TATATATCTT TTTCGAATTT GAAGTGAAAT TTCGATAATT TAAAATTAAA
TAGAACATAT CATTATTTAG GTATCATATT GATTTTTATA CTTAATTACT AAATTTGGTT
AACTTTGAAA GTGTACATCA ACGAAAAATT AGTCAAACGA CTAAAATAAA TAAATATCAT
GTGTTATTAA GAAAATTCTC CTATAAGAAT ATTTTAATAG ATCATATGTT TGTAAAAAAA
ATTAATTTTT ACTAACACAT ATATTTACTT ATCAAAAATT TGACAAAGTA AGATTAAAAT
AATATTCATC TAACAAAAAA AAAACCAGAA AATGCTGAAA ACCCGGCAAA ACCGAACCAA
TCCAAACCGA TATAGTTGGT TTGGTTTGAT TTTGATATAA ACCGAACCAA CTCGGTCCAT
TTGCACCCCT AATCATAATA GCTTTAATAT TTCAAGATAT TATTAAGTTA ACGTTGTCAA
TATCCTGGAA ATTTTGCAAA ATGAATCAAG CCTATATGGC TGTAATATGA ATTTAAAAGC
AGCTCGATGT GGTGGTAATA TGTAATTTAC TTGATTCTAA AAAAATATCC CAAGTATTAA
TAATTTCTGC TAGGAAGAAG GTTAGCTACG ATTTACAGCA AAGCCAGAAT ACAAAGAACC
ATAAAGTGAT TGAAGCTCGA AATATACGAA GGAACAAATA TTTTTAAAAA AATACGCAAT
GACTTGGAAC AAAAGAAAGT GATATATTTT TTGTTCTTAA ACAAGCATCC CCTCTAAAGA
ATGGCAGTTT TCCTTTGCAT GTAACTATTA TGCTCCCTTC GTTACAAAAA TTTTGGACTA
CTATTGGGAA CTTCTTCTGA AAATAGT

FIG. 4



pAGM243

CONSTRUCT: Ubi/Pat/Nos::Ubi/Gus/Nos

VECTOR: pUC18

SIZE: 9619bp

ORIGIN: 2809bp SphI/SacI fragment from pAGM106, SacI site blunted, SphI linkers, ligated into SphI site of pAGM100

FIGURE 5

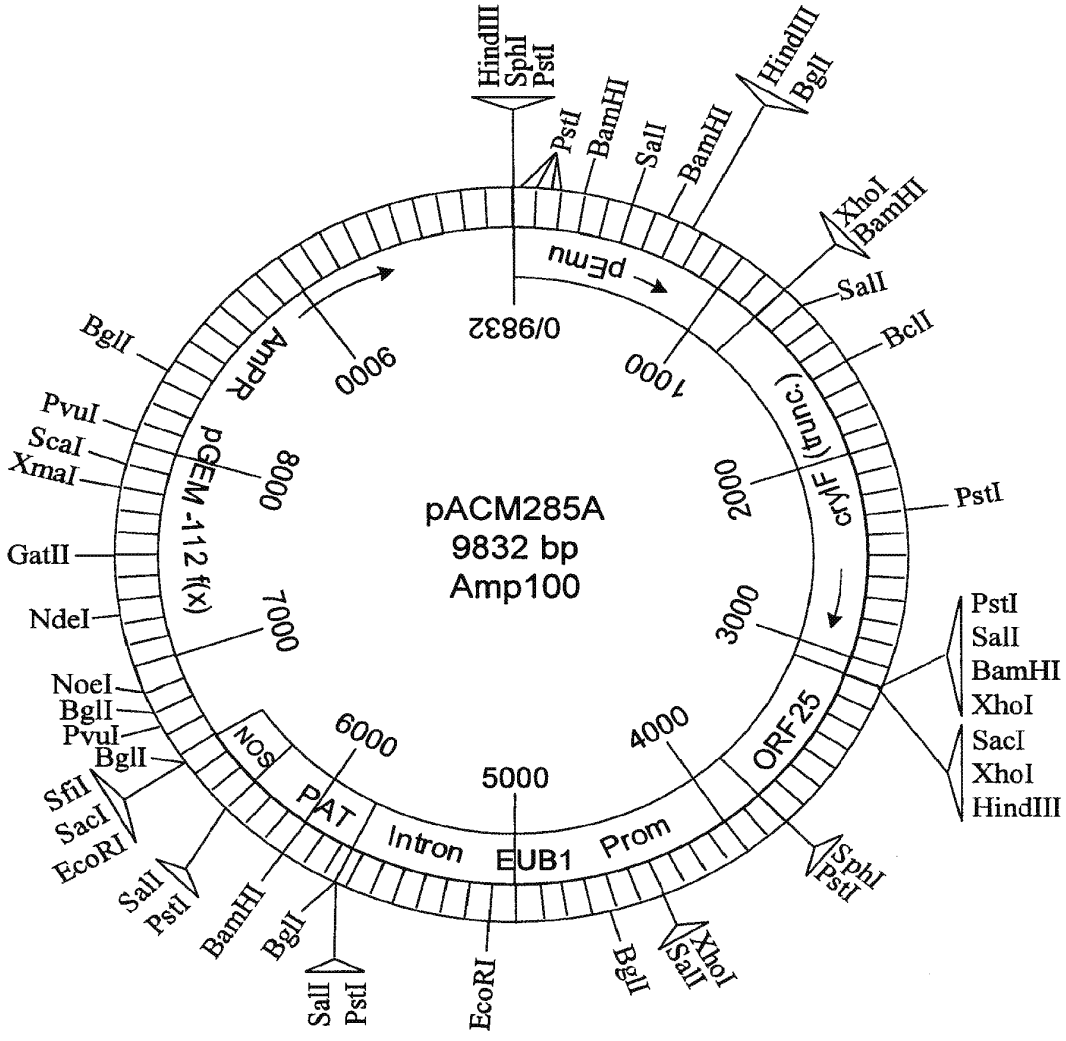


FIG. 6

SUBSTITUTE SHEET (RULE 26)

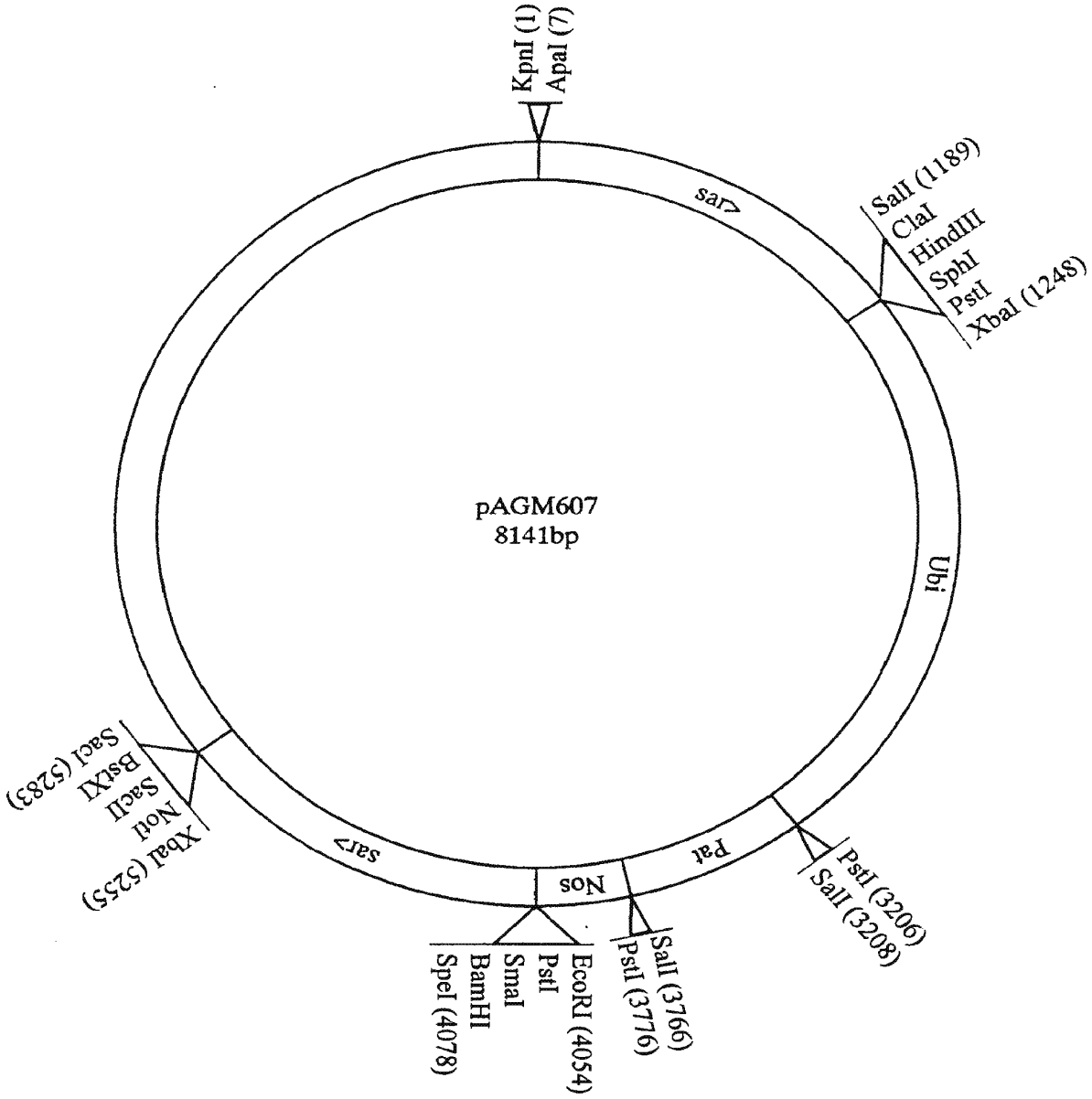


FIG. 7

SUBSTITUTE SHEET (RULE 26)

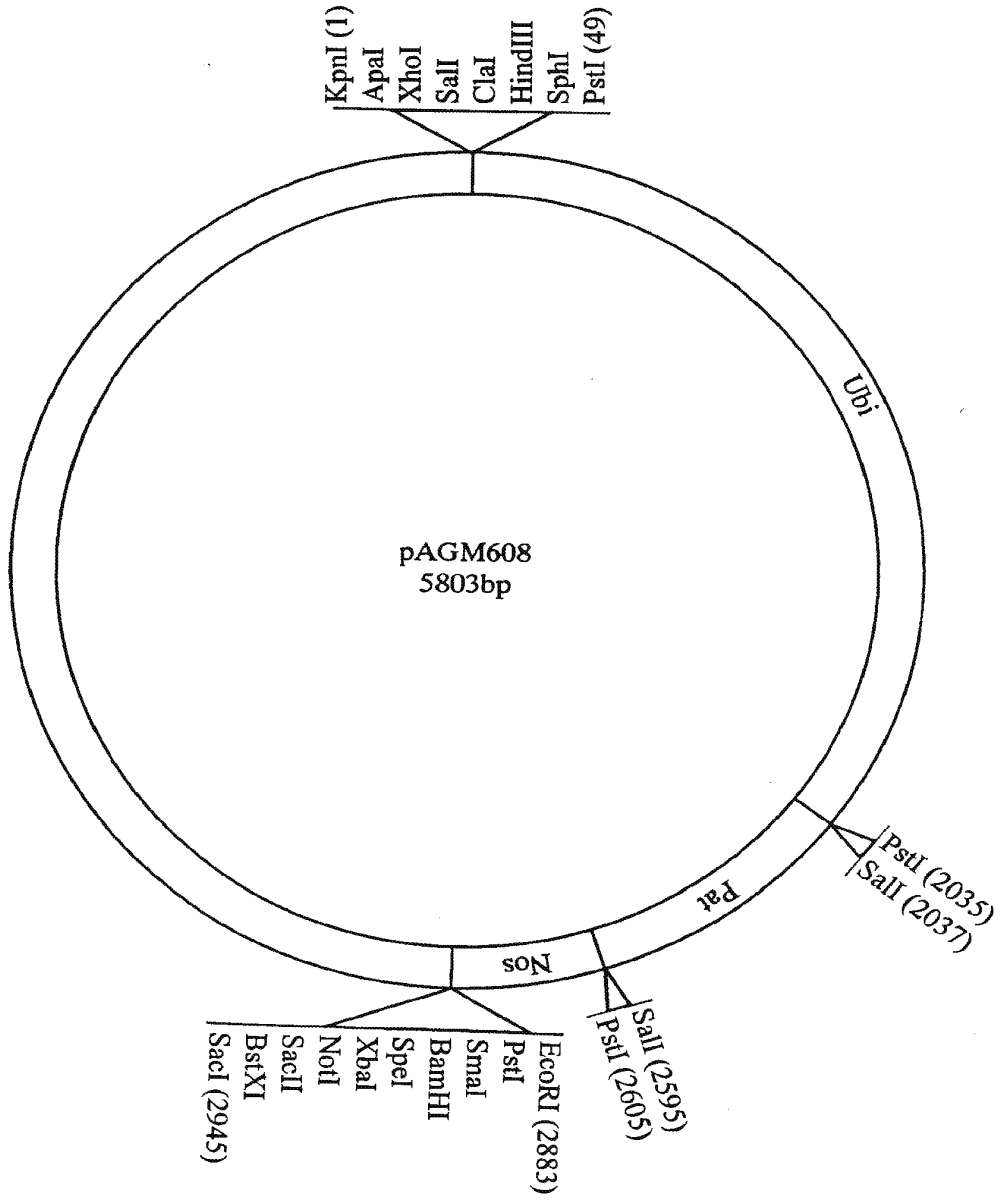


FIG. 8

SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

Inte. ional Application No
PCT/US 98/06109

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N5/14 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 07902 A (UNIV NORTH CAROLINA) 14 April 1994 see the whole document ---	1,2, 4-12,14, 16,18, 19,21-28
X	ALLEN G C ET AL: "High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco." PLANT CELL, (1996 MAY) 8 (5) 899-9133. , XP002072398 see the whole document --- -/--	1,2, 4-12,14, 16,18, 19,21-28

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 July 1998	07/09/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Chakravarty, A

1

Form PCT/ISA/210 (second sheet) (July 1992)

COR-AAD1-00002050

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/US 98/06109

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	BREYNE P. ET AL.: "Characterisation of a plant scaffold attachemt region in a DNA fragment that normalises transgene expression in tobacco" THE PLANT CELL, vol. 4, April 1992, pages 463-471, XP002072399 see the whole document ---	1,2, 4-12,14, 16,18, 19,21-28
X	MAESSEN G.D.F.: "Genomic stability and stability of expression in genetically modified plants" ACTA BOT. NEER., vol. 46, no. 1, March 1997, pages 3-24, XP002072400 see page 12 - page 13 ---	1,2, 4-12,14, 16,18, 19,21-28
P,X	WO 97 27207 A (UNIV NORTH CAROLINA ;THOMPSON WILLIAM F (US); HALL GERALD JR (US);) 31 July 1997 see claims -----	1-28

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

COR-AAD1-00002051

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte ional Application No
PCT/US 98/06109

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9407902 A	14-04-1994	AU 673859 B	28-11-1996
		AU 5165593 A	26-04-1994
		CA 2147006 A	14-04-1994
		EP 0663921 A	26-07-1995
		US 5773689 A	30-06-1998
WO 9727207 A	31-07-1997	US 5773695 A	30-06-1998
		AU 2246997 A	20-08-1997

COR-AAD1-00002052

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, A01H 5/00 — A1

(11) International Publication Number: **WO 99/10513**

(43) International Publication Date: 4 March 1999 (04.03.99)

(21) International Application Number: PCT/IB98/01199

(22) International Filing Date: 5 August 1998 (05.08.98)

(30) Priority Data:

60/055,314	7 August 1997 (07.08.97)	US
60/079,042	23 March 1998 (23.03.98)	US
09/079,640	15 May 1998 (15.05.98)	US

(71) Applicant *(for all designated States except US)*: AUBURN UNIVERSITY [US/US]; 309 Samford Hall, Auburn University, AL 36849–5176 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: DANIELL, Henry [US/US]; 1255 Marina Point 315, Casselberry, FL 32707 (US).

(74) Agents: WEISER, Gerard, J. et al.; Weiser & Associates, P.C., Suite 500, 230 South 15th Street, Philadelphia, PA 19102–3838 (US).

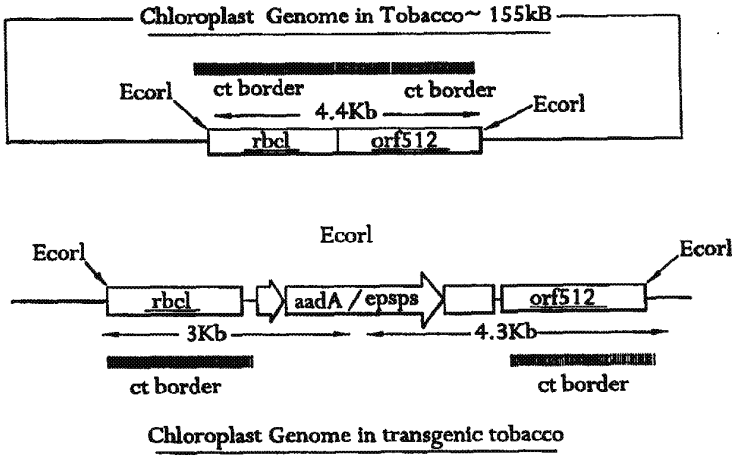
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: UNIVERSAL CHLOROPLAST INTEGRATION AND EXPRESSION VECTORS, TRANSFORMED PLANTS AND PRODUCTS THEREOF



Chloroplast Genome in transgenic tobacco

(57) Abstract

The invention provides universal chloroplast integration and expression vectors which are competent to stably transform and integrate genes of interest into chloroplast genome of multiple species of plants. Transformed plants and their progeny are provided. Monocotyledonous and dicotyledonous plants are transformed which have never been transformed heretofore. Plants transformed with a synthetic gene express valuable biodegradable protein–based polymers (PBPs). Transformed plants produce high value molecules. Resistance is provided to agricultural crops against the major classes of chemical herbicides. Herbicide resistance is used as a lethal selectable marker for chloroplast transformation. The transformed plants are capable of expressing in addition to the targeted trait, a desirable, secondary non–targeted trait. Insect resistance is provided to transformed plants, both against insects that are susceptible to Bt toxins and against insects that have developed resistance to Bt toxins.

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

UNIVERSAL CHLOROPLAST INTEGRATION AND EXPRESSION VECTORS, TRANSFORMED PLANTS AND PRODUCTS THEREOF

This application claims the benefit of pending provisional application Serial No. 60/055,314, filed August 7, 1997, which is incorporated herein by reference in its entirety. This application also claims the benefit of pending provisional application Serial No. 60/079,042, filed March 23, 1998, entitled "Universal Chloroplast Integration and Expression Vector, Transformed Plants and Products Thereof", which is incorporated herein by reference in its entirety. Further, this application is a continuation-in-part of pending patent application Serial No. 08/591,407, filed January 25, 1996, by Henry Daniell, which is a continuation of application Serial No. 08/215,020, filed March 18, 1994, now abandoned, which in turn was a continuation of application Serial No. 07/249,616, filed September 26, 1988, now abandoned.

FIELD OF THE INVENTION

This application pertains to the field of genetic engineering of plant genomes, particularly the genetic engineering of the genome of plant plastids, such as chloroplasts and to the stable transformation of chloroplast genome of any plant species.

RELATED CASES

This application relates in particular to a universal chloroplast expression and integration vector which is competent to transform any plant with one or more genes of interest. The earlier patent application Serial No. 08/591,407, teaches plant cells transformed by means of an expression cassette comprising an exogenous DNA sequence which is stably integrated (covalently linked) to the chloroplast genome of the cell of a target plant. "Stably" integrated DNA sequences are those which are inherited through genome replication by daughter cells or organisms. This stability is exhibited by the ability to establish permanent cell lines,

clones, or transgenic plants comprised of a population containing the exogenous DNA.

Likewise, U.S. patent 5,693,507 (1997) to Daniell and McFadden discloses such stable integration by means of an expression cassette which comprises an exogenous DNA sequence which codes for a desired trait, and the transformed plants.

BACKGROUND OF THE INVENTION

Advantages of Chloroplast Transformation Over Nuclear Transformation. The attractiveness of transformation of the chloroplast genome over transformation of the nuclear genome is attributable to the serious risks resulting from the latter. One common concern is the escape of foreign genes through pollen dispersal from transgenic crop plants to their weedy relatives. It has been demonstrated that transgenic pollen will deliver foreign (transgenic) genes to other sexually-compatible plants (detected by marker gene prevalence in progeny harvested from non-transgenic plants grown in surrounding area). For Example, dispersal of pollen from a central test plot containing transgenic cotton plants to surrounding non-transgenic plants has been observed at varying distances in different directions. (Lewellyn and Fitt, 1996); (Umbeck, P.F., et al., 1991). In addition, the frequencies of marker genes in wild sunflowers averaged about 28 to 38%; in wild strawberries growing within 50 meters of a strawberry field, more than 50% of the wild plants contained marker genes from cultivated strawberries. (King, J., 1996).

The escape of foreign genes through pollen is especially a serious environmental concern, in the case of herbicide resistance genes, because of the high rates of gene flow from crops to wild relatives. The concern is that gene escape from transgenic crops to their weedy relatives will create super weeds. In rice (*Oryza sativa*), gene flow from cultivated varieties to wild relatives has been noted, into *O. perennis* (Barrett, 1983) and red rice (*O. sativa*; Langevin et

al., 1990). In the southern US, red rice has become a major weed because herbicides that kill it also kills cultivated rice. Decreased prices are paid for cultivated rice contaminated with red rice. Some researchers have introduced the *bar* gene conferring resistance to glufosinate (Liberty) into cultivated rice to combat this weed (Oard et al., 1996; Sankula et al., 1996). However, due to sexual compatibility, introduction of a nuclear-expressed gene will allow transmission of that resistance trait into red rice via pollen.

Similarly, transgenic oil seed rape, genetically engineered for herbicide resistance outcrossed with a weedy relative, *Brassica campestris* (field mustard) and conferred herbicide resistance even in the first back-cross generation under field conditions. (Mikkelson, T. R., et al., 1996).

Maternal inheritance of introduced genes prevents gene escape through pollen. Engineering foreign genes through chloroplast genomes (which are maternally inherited for most of the crops) is a solution to this problem. Also, the target enzymes or proteins for most herbicides (e.g. amino acid/fatty acid biosynthetic pathways or photosynthesis) are compartmentalized within the chloroplast. Another important advantage of chloroplast transformation is the higher levels of foreign gene expression due to a very high copy number (5000-10,000) of chloroplast genomes in plant cells. Because the transcriptional and translational machinery of the chloroplast is prokaryotic in nature, herbicide resistant genes of bacterial origin can be expressed at extraordinarily high levels in chloroplasts.

Transformation of the Chloroplast Genome. Early investigations on chloroplast transformation focused on the development of in organello systems using intact chloroplasts capable of efficient and prolonged transcription and translation (Daniell and Rebeiz, 1982; Daniell et al., 1983)

and expression of foreign genes in isolated chloroplasts (Daniell and McFadden, 1987). These experiments were done under the premise that it was possible to introduce isolated intact chloroplasts into protoplasts and regenerate transgenic plants (Carlson, 1973). The discovery of the gene gun as a transformation device opened the possibility of direct plastid transformation in plants (Daniell, 1993). Transient expression of foreign genes in plastids of dicots (Daniell et al., 1990; Ye et al., 1990), monocots (Daniell et al., 1991), prolonged foreign gene expression using autonomously replicating chloroplast expression vectors (Daniell et al., 1990) and stable integration of a selectable marker into the tobacco chloroplast genome (Svab and Maliga, 1993) were accomplished using the gene gun. Tobacco plants resistant to certain insects were obtained by integrating the *cryIAc* gene into the tobacco chloroplast genome (McBride et al., 1995; U.S. Patent 5,451,513, incorporated herein by reference). Stable plastid transformation of higher plants has been accomplished so far only in tobacco.

Prior Studies on the Chloroplast Genome. To date, stable integration of a foreign gene into the chloroplast genome of a higher plant has been reported only in tobacco. This was achieved with a vector which was specific for tobacco and which was derived from the tobacco chloroplast genome, that is, the vector contained a sequence homologous only to the tobacco chloroplast genome and which is not highly conserved in the chloroplast genomes of other plants. Such vector is unsuitable for stably transforming plant species other than tobacco. The only published report of foreign gene expression in a plant species other than tobacco is that of wheat leaves and embryos (Daniell et al., 1991), but stable integration was not accomplished. Stable integration of a foreign gene into the chloroplast genome of a monocotyledonous plant has never been reported. At least in cereals (monocots), previously developed transformation /regeneration protocols may not be

amenable to plastid transformation due to inherent inefficiencies within those systems. Also, sequential/serial selections (repeated selections), deemed important for achieving homoplasmy (Daniell, 1997), may not be feasible using those regeneration systems employed. Recent development of unique corn (Rudraswamy, 1997) and rice (unpublished) transformation/regeneration protocols have the potential to exhibit substantially increased efficiencies and allow more than one round of selection during regeneration.

Maliga et al. in U.S. patent 5,451,513 and Svab et al., 1990 propose a transformation of the plastid genome of tobacco by a non-lethal selection technique which employs plastid DNA encoding a non-lethal selectable phenotype. According to Maliga et al. a non-lethal selection is absolutely essential for obtaining transplastgenic lines.

Unlike the Maliga et al. technique, the method of the invention provides a selection which is lethal to all non-transformed plants, but for tobacco. Only the transformed plants survive and continue to grow. This lethal selection takes place with virtually all antibiotics, including spectinomycin and streptomycin in a medium containing the antibiotic in a concentration of 500 - 1,000 μg/ml. Similar conditions were shown to be non-lethal for tobacco by Maliga et al. Moreover, unlike the technique of Maliga et al., in accordance with the invention, transformation to homoplasmy can be achieved even in the first round of selection.

In European Patent Application No. 0 251 654, Cannon et al. describe transposon-mediated chloroplast transformation of tobacco for instance, using the bacterial transposon Tn5. The vector containing the transposon is targeted at a chromosomal region known to be a "transcriptionally silent" region in order to preserve the transcriptional integrity of the native genes. Such a transcriptionally silent region is identified to be located between two known divergent promoters of chloroplast genes, e.g. the promoters for the genes for the chloroplast large subunit of ribulose bisphosphate carboxylate

(RbcL, "LS RuBisCo") and for β-ATPase (atpB). These promoters transcribe the genes in opposite directions away from the silent region of the chromosome. No transcription terminator is provided in the expression vector of Cannon et al., such terminator regions are known to be absolutely essential for gene expression in plastids. Finally, no stable chloroplast transformation is shown to be accomplished by Cannon et al.

The invention described herein has several distinguishing features over Cannon et al. The invention teaches stable transformation transmittable to the progeny. The integration is not directed into a transcriptionally inactive region of the chloroplast chromosome. The invention integrates a cassette (which contains a transcription terminator as described further hereinafter) into a transcriptionally active region of the chloroplast genome. Promoters controls the expression of one or more genes. Unlike Cannon et al. no transposon is involved in the transformation of the chloroplast in accordance with the invention.

In NATO Asi Series, Daniell et al., 1994 report engineering insect resistance via chloroplast genomes showing the expression of CryIIA protein in plants to control insects. McBride et al, 1995, and U.S. patent 5,545,818 (1996), confirm report of Daniell et al and show the expression of *Bacillus thuringiensis* CRYIAc protein into plant plastids. The vectors reported by McBride are designed to introduce the construct only into the tobacco chloroplast genome.

The Need for a Vector to Transform a Variety of Plants. It is evident from the state of the art that an important need exists for a chloroplast integration and expression vector for transforming, preferably stably, the chloroplast genomes of many different species of plants. Such a "universal vector" would permit the transformation of the chloroplast genome of a selected target plant with a heterologous(foreign) DNA

coding sequence and eliminate the need to construct vectors which each one is specifically suited to transform the chloroplast genome of the particular plant species which is to be transformed.

The problem to construct such a universal vector competent to transform different plants has to the knowledge of the inventor, not yet been solved.

Prior Art Concepts of the Intergenic Spacer Region. While the nucleotide sequence of coding regions of the genome, including the chloroplast genome, are often conserved between species, in contrast the sequences flanking functional genes, i.e. the spacer regions between coding regions typically are not conserved. The accepted dogma for lack of conservation, and thus the low degree of homology between species of spacer regions, is that the spacer regions typically do not perform essential functions. Therefore, there is little, if any, selective pressure to conserve the sequence of spacer regions between species. The sequence of the spacer regions may be altered without undesirable effects.

Stummann et al., 1988, disclose that the gene order of the ribosomal RNA operon of the chloroplast genome is the same between different species of plants, including tobacco, maize, and a liverwort, *Marchantia*, and that the coding sequences of this operon are highly homologous. Stummann also discloses that the interspecies homology of the operon is less than the interspecies homology of the gene coding regions. This is consistent with the lack of conservation of spacer regions; and suggests that the interspecies homology of spacer regions in the ribosomal RNA operon is relatively low.

The invention, contrary to the dogma of lack of conservation of the spacer regions, uses spacer regions that are highly conserved between different plants to construct vectors competent to transform a variety of plants.

OVERVIEW OF THE INVENTION

The invention provides universal chloroplast integration and expression vectors which are competent to stably transform and integrate genes of interest into chloroplast genome of multiple species of plants. Transformed plants and their progeny are provided. Monocotyledonous and dicotyledonous plant are transformed which have never been transformed heretofore. Plants transformed with a synthetic gene express valuable biodegradable protein-based polymers (PBPs). Transformed plants produce high value molecules. Resistance is provided to agricultural crops against the major classes of chemical herbicides. Herbicide resistance is used as a lethal selectable marker for chloroplast transformation. The transformed plants are capable of expressing in addition to the targeted trait, a desirable, secondary non-targeted trait. Insect resistance is provided to transformed plants, both against insects that are susceptible to Bt toxins and against insects that have developed a resistance to Bt toxins.

SUMMARY OF THE INVENTION

The Intergenic Spacer Region. The Invention's Concept.

It has been discovered contrary to the conventional belief, that the chloroplast (ct) genome of plants contains spacer regions with highly conserved nucleotide sequences. The highly conserved nature of the nucleotide sequences of these spacer regions of the chloroplast genome makes such spacer regions, it has been discovered, ideal for the construction of vectors to transform chloroplasts of widely varying species of plant, this without the necessity of constructing individual vectors for different plants or individual crop species, which would require first a determination of the DNA sequence of each of the chloroplast genomes. This finding has numerous useful consequences and important practical applications.

The Several Embodiments of the Invention

The Universal Vector. The invention has several useful embodiments. The invention provides a universal integration and expression vector hereinafter referred to as "UV" and its use for the expression of at least one phenotype in a variety of different plants.

The integration expression universal vector of the invention comprises an expression cassette (further described below) which comprises the necessary genetic elements to transiently or preferably stably transform the plastids e.g. chloroplast genome of a target plant cell with a foreign (heterologous) DNA coding for a molecule of interest, like a phenotype to be expressed by the plant or a non-plant high value molecule, like a biologically active peptide (or polypeptide). The universal vector is constructed with a transcriptionally active region of a chloroplast genome that is highly conserved in a broad range of chloroplast genomes of higher plants. Preferably that region is the spacer 2 region; the intergenic spacer region between the t-RNA^{Ile} and the tRNA^{Ala} region. Such region is often referred to herein as a "spacer" region because in the chloroplast genome it is intergenic between several genes in the rRNA operon which is transcribed by one promoter. When built into the universal vector such region is generally referred to herein as a "border" or preferably as a "flanking sequence" or "flanking sequences". This is because in the universal vector, the operably joined genetic elements for transforming stably the plastid of the target plant are flanked on each side by a sequence i.e. a fragment of the spacer region. The flanking sequences in the vector and the spacer sequences in the chloroplast genome have sufficient homology to each other to undergo homologous recombination. The universal vector is inserted into the spacer of a transcriptionally active region in the chloroplast genome. Generally, the spacer region is positioned in the inverted repeat region of the chloroplast genome. The rest of the construct, i.e. other than the flanking sequences and the expression cassette, is generally

referred to herein as the "vector" which comprises bacterial sequences, like the plasmid cloning vectors pUC, pBR322, pGEM or pBlueScript.

The Expression Vector or Cassette. The universal vector comprises an expression cassette which is flanked on each side by a flanking sequence. A suitable expression cassette for use in the invention is described in U.S. patent 5,693,507 (1997), which is incorporated herein by reference. That cassette comprises, operably joined, a transcriptional initiation region functional in plant chloroplast, at least one heterologous DNA sequence coding for a target molecule of interest, e.g. a gene (or functional fraction thereof) encoding a biologically active compound, and control sequences positioned upstream for the 5' and downstream from the 3' ends and a transcription termination region to provide expression of the coding sequence in the chloroplast genome of a target plant. Preferably, the expression cassette is flanked by plant DNA sequences, like chloroplast DNA sequences, in order to facilitate stable integration of the expression vector into the chloroplast genome. In the construction of the expression cassette, the DNA sequence comprises one or more cloning site(s) for integration of the gene(s) of interest.

The spacer sequences that have been identified in plastids of higher plants are ubiquitously conserved between a great variety of plants. These sequences were found to be ideal to construct the universal vectors of the invention which are, as a result, competent to transform the chloroplast genome of a large variety (or multiplicity) of target plants by homologous recombination. It is thus immaterial from which individual spacer of a particular plant the universal vector is constructed.

As is known, it will be generally advisable to have at least one additional heterologous nucleotide sequence coding for a selectable phenotype, such as a gene providing for antibiotic resistance or a functional portion thereof to

serve as a marker associated with the expression cassette or with the universal integration expression vector. This facilitates identification of the plant cells in which the foreign gene has been stably integrated. Marker genes are known in the literature, for instance β-lactanase, herbicide resistant genes such as the mutant *psb*A gene or EPSPS-*aro*A, the cat gene which encodes chloramphenicol acetotranferase, and the uidA gene encodes β-glucuronidase (gus)and others.

It is recognized that tobacco is unique in being not susceptible to the lethal affect of streptomycin and spectinomycin. Though tobacco leaves lack the pigmentation when exposed to a medium with such an antibiotic, continued growth is observable. However, this property of tobacco is readily circumvented. There are numerous antibiotics available which are lethal for tobacco, like hygromycin. Another approach is to select a gene which expresses a visible marker like a color, fluorescence, etc., like the reporter gene mGFP, that codes for a green fluorescent protein.

Method of Transformation. The invention provides a transformation method which can produce homoplasmy (integration of foreign genes into all of the chloroplast genomes of the plant cell) after a first round of selection without the need for a further selection process. The method for transforming a plant uses the universal vector constructed with flanking sequences from a plant species other than the species of the target plant to be transformed. Alternatively, the vector may contain flanking sequences from the same plant species as the target plant, including from tobacco.

Method to Construct the Universal Vector. The invention further provides a method to construct the universal chloroplast integration and expression vector. To this effect, a spacer portion of the chloroplast genome from any plant is determined to be highly homologous to more than one

species of plants. A nucleotide sequence corresponding to that spacer region is obtained from the identified chloroplast genome (or synthesized) and is incorporated into a suitable vector, such as by subcloning into a plasmid. The spacer region is positioned as flanking sequences to the expression cassette comprising the necessary genetic elements for transformation of the plastid and expression of the foreign gene(s).

Any method of transformation of the chloroplast may be used. Any gene (or functional portion thereof) which may be utilized to transform a plant chloroplast and encode a desired peptide to confer the desired trait to the target plant is suitable for transformation with the universal vector.

Transformed Plants. The invention further provides plants in which the chloroplast genome has been stably, that is, permanently transformed with the universal vector of the invention, including the progeny thereof.

The invention includes monocotyledonous plants like cereals or plant cells, such as maize, rice, barley, oat, wheat and grasses, and their progeny in which the chloroplast genome has been stably transformed with the universal vector derived from the same species or from a different species than the transformed plant. The invention provides dicotyledonous and monocotyledonous plants, stably transformed following a single round of selection, due to homoplasmy achievable with the universal vector comprising a chloroplast origin of replication *(ori)*. The invention also provides stably transformed plants of differing species, including varieties of the same species, genera, families, orders, and divisions of plants.

In accordance with the invention, a plant in which the chloroplast genome has been stably transformed with one or more foreign genes of interest includes mature plants and progeny thereof, like seedlings and embryos. The term "plant"

in this context also includes portions of plants such as explants like cuttings, tissue cultures, cell suspensions, and calli.

Thus, the invention includes the stably transformed multicellular plants, their progeny, the seed, and the transformed plastids, e.g. the chloroplast, etc., and method of regenerating the transformed plants.

In this specification and in the claims, when reference is made to different "species", the term "species" refers not only to "species" but to varieties within a species, genera, families, order, and divisions of the plant kingdom. Thus, a universal vector which can be used to transform plants of different species is understood to be able to transform plants of different varieties within a species, different genera, different families, different orders, and different divisions. The term "plant" (or "plants") is intended to be generic as used herein.

Expression of Non-Plant Products

Biopolymer genes. Another embodiment of the invention using the universal integration and expression vector provides plants transformed with a synthetic biopolymer gene that codes for biodegradable protein-based polymers (PBPs).

These polymers have important properties of practical importance, discussed hereinafter.

Production of High Value Molecules-Biologically Active Molecules. The intriguing discovery that transformation with a synthetic gene which need not have a natural analogue in plant or animal, to produce PBPs, is feasible, has shown the wide applicability of the vector in yet another field of human endeavor: the production of biologically active molecules, like pharmaceuticals in plants, from any gene or functional fraction thereof, synthetic or natural.

A further embodiment of the invention is therefore the use of transformed plants as bioreactors (as factories) for biopharmaceuticals. There are at least two capabilities often needed for the production of proteins of pharmaceutical value, not possible in prokaryotic systems. Plants, unlike bacteria, are able to produce the foreign protein in a biological active conformation. Plants are also often more tolerant to the alteration of their biosynthetic pathways. Thus, the plants can be transformed with a gene non-functional in (or foreign to) plants, that may be synthetic or not, that may normally be functional (competent) in animals (mammals), in oviparous, in pesces or other species.

The invention further provides transformed plants comprising a gene provided by an expression cassette, preferably by a universal vector, which codes for a variety of desired products, especially biologically active molecules like peptides (polypeptides), proteins, insulin, human serum albumin (HSA) and other molecules further described hereinafter. The plants are allowed or caused to grow, and the products are isolated from the transformed crop, like tobacco, maize, etc., and if desirable, harvested first and if necessary, purified.

Herbicide Tolerance. Another important embodiment of the invention provides transgenic herbicide resistant plants in which a foreign transgene conferring resistance to one or more herbicides are integrated, preferably stably, into the chloroplast genome by means of the universal vector. Of particular importance are transformed plants which exhibit glyphosate resistance and thus be resistant to "ROUNDUP™", a herbicide available commercially from Monsanto Company. The universal vector provides an effective means to transform the chloroplast genome of any plant and to confer resistance (or tolerance) to any of the herbicidal chemicals.

A different aspect of the invention provides a method to transform a plant by means of an expression

cassette, preferably by means of the universal vector, to cause it to produce a non-targeted (secondary or other) trait (or phenotype). [See, for example, Penazloza, V., et al. (1995), who report that expression by gromycin β-phosphotransferace gene confers resistance to the herbicide glyphosate.]

In another aspect of the invention, herbicide tolerance is used as a marker gene for chloroplast transformation.

Insect Resistance. A further embodiment of the invention provides insect resistance. With the increased concerns of using chemical pesticides, the use of *Bacillus thuringiensis* (Bt) formulations has been widely advocated. *Bacillus thuringiensis* produces many types of crystalline inclusions which are toxic to insects. The proteins comprising these inclusions have been categorized based on insecticidal host range, and protein homology. The CRYI and CRYII toxins have insecticidal activity against lepidoptera, or lepidoptera and diptera, respectivly. CRYI protoxins are 130-135 kDa in size which are enzymatically cleaved into proteins of 65kDa for insecticidal activity.CRYII protoxin is 65kDa in size with a protein with a molecular mass of 60-62 kDa for insectcidal activity. Many commercially important insects pests (especially in the family Pyralidae) are susceptible to CryIIA toxin, including European corn borer, *Ostrinia nubilalis*, lesser cornstalk borer, *Elasmopalpus lignosellus*, cowpea pod borer, *Maruca testulalis*, tobacco budworm, *Heliothis virescens*, tobacco hornworm, *Manduca sexta* and gypsy moth *Lymantria dispar*, Daniell et al. 1994.

However , Bt formulations have not been as effective as anticipated primarily due to their susceptibility to UV radiation, inadequate coverage, expense and limited host range. Delivery of Bt toxins via Bt-transgenic plants is therefore appealing.

Acceptable insect control has occurred with nuclear-transgenic Bt cotton against the tobacco budworm, *Heliothis virescens*, but these plants do not express enough Bt toxin to control the cotton bollworm, *Helicoverpa zea*. Additionally, these Bt genes may outcross to related plant species via pollen. To circumvent these concerns, chloroplast transformation and expression has been evaluated in accordance with the invention because of the following reasons. 1) Plant cells containing chloroplasts can contain up to 10,000 gene copies per cell, 2) chloroplasts can read intact bacterial DNA, including operons, and 3) chloroplast genomes are maternally inherited, and therefore, foreign gene escape via the pollen is drastically eliminated because chloroplast DNA is generally degraded in pollen.

CRY2A was chosen because CRY2A is relativley non-homologous to CRY1A and therefore, displayes only slight cross resistance against some CRY1A-resistant *H. virescens* populations. Because CRY2A is a "naturally truncated" high expression levels may be achieved without sacrificing the 3' region.

According resistance has been provided to tobacco chloroplast genome transformed with a universal vector of the invention to insects normally susceptible to Bt toxins and also to insects that have developed resistance or less susceptibility to Bt toxins. Insects that had never been killed by any CRY toxin showed 100% mortality on transformed tobacco.

The expression of *cry*2A in a plant could therefore be a valuable tool in controlling multiple insect pests while overcoming the development of Bt resistance. Other insecticidal proteins such as cholesterol oxidase that kill boll weevil by a different mechanism, are expressed to high levels in chloroplasts.

Other embodiments of the invention will become apparent hereinafter.

DESCRIPTION OF THE DRAWINGS

Figure 1 shows a map of the tobacco chloroplast genome. The thick lines in the genome map represent the inverted repeat regions of the chloroplast genome. The arrows labeled "UV" represent the insertion sequences for the preferred embodiment of the universal integration and expression vector (UV); the arrow labelled "TV" represents the insertion sequence for the tobacco vector (TV).

Figure 2A shows the tobacco chloroplast vector (TV) pZS-RD-EPSPS for expression of herbicide resistance.

Figure 2B shows the universal chloroplast expression and integration vector (UV), pSBL-RD-EPSPS for expression of herbicide resistance.

Figure 3A shows the universal chloroplast integration and expression vector, pSBL-CG-EG121 for biopolymer expression.

Figure 3B shows the tobacco integration and expression and vector, pZS-CG-EG121 for biopolymer expression.

Figure 4A-4E shows the sequence homology of spacer regions between tobacco and other crop species .A site for foreign gene insertion is shown by arrows. Upstream of the site for foreign gene insertion the site of an origin of replication (*ori*) is shown.

Figure 4F-4G shows the sequence alignment of the spacer (64 bp) region 16S-23S rDNA from several crop species with the tobacco chloroplast sequence where (+) represents the positive and (-) the negative strands, respectively.

Figures 5A-C shows construction of the vector pSBL-Ct border.

Figures 6A-C shows construction of the vector pSBL-CtV1 selectable marker gene cassette containing a chloroplast 16S rRNA promoter (Prrn), the *aad*A gene and a 3' untranslated region of the chloroplast *psb*A gene.

Figures 7A-7D shows vectors pSBL-CtV2, pSBL-CtV3, pSBL-CtVH, pSBL-CtVHF, respectively.

COR-AAD1-00002071

Figure 8A-B show vectors pSBL-CtVHBt and pSBL-CtVHBtR, respectively.

Figure 9 shows transformed and untransformed tobacco plants growing in the presence of spectinomycin indicating non-lethal selection on the medium (500μg/ml). Note growth of transformed (bleached) and untransformed leaves (green).

Figures 10A-10G show corn plastid transformation and regeneration.

Figure 11 shows corn plastid transformation. Transformed corn plants grow normally (middle shoot) while untransformed plants die on the lethal medium, confirming lethal selection by the antibiotic, (1000 μg/ml spectinomycin).

Figures 12A-12F shows rice plastid transformation and regeneration.

Figures 13A and 13B shows PCR analysis of DNA isolated from the leaves of rice transformants.

Figure 14 shows peanut plastid transformation. Transformed peanut plants grow normally (middle and on left side of plate) while untransformed plants die in the lethal medium (500 μg/ml spectinomycin).

Figure 15 shows soybean plastid transformation. Two transformed plants show shoots, the other plants die on the lethal medium, confirming lethal selection by the antibiotic (500 μg/ml spectinomycin).

Figure 16 shows sweet potato embryo transformation on lethal selection medium (500 μg/ml spectinomycin).

Figure 17 shows grape cells transformation. The transformed culture cells become green while the untransformed cells die in the lethal selection medium (500 μg/ml spectinomycin).

Figure 18 shows protein-based biopolymer (PBP) expression by chloroplast integration and expression vectors in *E. coli*.

Figure 19 shows Southern blot analysis performed with the transformants from PBP transgenic plants using the

COR-AAD1-00002072

tobacco vector (TV). Probes were from chloroplast border sequences (A) or from polymer (EG121) gene sequences (B)

Figure 20 shows Southern blot analysis performed with the transformants from PBP transgenic plants using the universal vector (UV). Probes were from chloroplast border sequences (A) or from the aadA gene (B).

Figure 21A shows foreign gene transcript levels analyzed by northern blotting using total RNA isolated from the control, chloroplast transformants and a nuclear tobacco transgenic plant highly expressing the synthetic biopolymer gene (EG121).

Figure 21B shows an enlargement of lanes 4-7 of Figure 21A.

Figure 22 shows Western blot analysis of purified polymer protein from transgenic plants.

Figure 23A shows the higher growth rate of *E. coli* containing the tobacco vector with the EPSPS gene.

Figure 23B shows the higher growth rate of *E. coli* containing the universal vector with the EPSPS gene.

Figures 24A-24B show the integration of foreign genes into the plastid genome by PCR using *rbcL* and *aadA* primers (A), or 16SRNA and *aadA* primers (B).

Figures 25A-25C show the integration of the aroA gene into the chloroplast by Southern analysis and the high generation of homoplasmy using EPSPS probe (A) or *rcb*L-*orf*512 probe. The site of integration is shown in (C).

Figures 26A and 26B show generation of seeds collected from control and transformed tobacco plants, respectively, in the presence of the selectable markers.

Figures 27A and 27B show transgenic and control tobacco plants sprayed with glyphosate.

Figures 28A and 28B show tobacco susceptibility (control) and resistance (transformed) to insects.

Figure 29 shows (Western blot analysis)total protein isolated from control and transgenic tobacco plants.

COR-AAD1-00002073

Preferred embodiments of the invention are described in greater detail hereinafter.

DETAILED DESCRIPTION OF THE INVENTION

The Universal Integration and Expression Vector. The universal integration and expression vector of the invention is competent for stably transforming chloroplasts of various different target plants. Heterologous DNA coding sequences can be provided in the cassette to code for a phenotype such as herbicide resistance, insect resistance or other traits. The vector further comprises a flanking sequence on each side of the DNA coding sequence which is homologous to a spacer sequence of the chloroplast genome, which spacer sequence is conserved in the chloroplast genomes of different plants. In this manner, stable integration of the heterologous gene into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking border sequences with complimentary spacer sequences in the chloroplast genome. The universal vector is competent to transform any plant species.

The *trnI* and *trnA* spacer region, has been found to be highly conserved in a great variety of plant species, from cyanobacteria to higher plants, such as monocotyledonous and dicotyledonous. The *trnI* and *trnA* genes flank on each side a spacer region, referred to as spacer 2 "or `spa 2' region" (Figure 4F-4G). The regions on either side of the spacer region likewise have almost perfect homology with corresponding regions between species of plants from cyanobacteria to higher plants, except that the higher plants contain two introns in the *trnI* and *trnA* genes. Longer border sequences tend to favor more efficient integration of foreign DNA. Therefore, although not essential, the homology between species of the *trnI* and *trnA* genes, in addition to the homology of the spacer region, contributes to the efficiency of transformation and integration (Figure 4A-4E).

If longer border sequences which include non-homologous portions, are incorporated into the vector, the non-homologous portion of the sequence will be "looped out" and "clipped off", in the recombination process and will not integrate into the target chloroplast genome.

Different universal vectors can be constructed with the spacer region. For instance, shorter or longer flanking sequences can be constituted with part or all of the *trnA* and *trnI* genes adjacent to `spa 2'.

A preferred universal vector comprises the flanking sequences and an expression cassette which comprises the following genetic elements to provide for transcription and translation of the DNA coding sequence organized in the following order (from the 5' to the 3' ends): a 5' part of the flanking sequence, a promoter functional in chloroplast, a DNA sequence with appropriate cloning site(s) for insertion of one or more coding sequence(s) for the desired phenotype or molecule of interest, and for a selectable marker, a terminator of transcription and a 3' part of the flanking sequence. The order of the DNA sequences coding the desired phenotype and the selectable marker can be switched. Additional flanking plant DNA sequences can be provided to promote stable integration. Preferably, the flanking sequence comprises an origin of replication (*ori*).

In a particular illustration, the highly conserved spacer region resides in the inverted repeat of the chloroplast genome. However, the particular location of the spacer in the chloroplast genome is not as important as its high homology with the spacer region of different plants.

Further, as may be seen in Figures 4F-4G, the spacer 2 (or spa 2) sequence which is 64bp long is too short to include the chloroplast genome *ori* which resides upstream, of that spacer. If it is desired to include the *ori*, a longer spacer sequence encompassing the *ori* will be selected which will include the spacer sequence and an additional sequence in the flanking sequences. This will provide a longer

template for homologous recombination into the recipient chloroplast genome and promote homoplasmy.

Another preferred vector is one in which the flanking sequences comprise, each one, in addition to the spacer 2 region, a portion or all of the intergenic spacer region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome (Figure 4A-4E). Further, the flanking sequences may include part or all of the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes, respectively. Optionally, the flanking sequences comprise each one part or all of the 16S and/or 23S rRNA gene sequences.

Illustrative Universal Vectors. A preferred universal vector comprises a DNA sequence which comprises the sequence of the spacer 2 region between the highly conserved *trnI* and the *trnA* genes between the 16S-23S rRNA genes of the chloroplast genome. Preferably, this region comprises part or all of the DNA sequence (Figure 4F-4G) of the *trnI* and the *trnA* genes. That region is excised from a selected plant like tobacco and subcloned into a commonly available plasmid like pUC19, e.g. at the PvuII site. Into the plasmid there is inserted the expression cassette which contains a selectable marker gene, a chloroplast 16SrRNA promoter, a gene encoding an enzyme conferring a resistance to an antibiotic like the *aadA* gene encoding aminoglycoside 3' adenyl transferase conferring resistance to streptomycin/spectinomycin and 3' untranslated region of the chloroplast *psbA* gene.

Specifically, when a universal vector is constructed with plasmid pSBL-Ct-bor (Figure 5), a selectable marker gene cassette containing a chloroplast 16S rRNA promoter, the *aadA* gene encoding aminoglycoside 3'-adenyl transferase conferring resistance for streptomycin /spectinomycin, or a herbicide resistance gene and a 3' untranslated region of the chloroplast *psbA* gene, was inserted into the plasmid. (Figure 6) This selectable marker gene cassette was then inserted into the universal border in two different orientations. In

the vector pSBL-CtV1, the selectable marker gene cassette was inserted into the *trnI* gene (Figure 6A). In the vector pSBL-CtV2 (Figure 7A), the selectable marker gene cassette was inserted between the *trnI* and *trnA* genes in the spacer region, in the direction of the 16S rDNA transcription. In the vector pSBL-CtV2 R (map not shown), the selectable marker gene cassette was inserted between the *trnI* and *trnA* genes in the spacer region, in the direction opposite of the 16S rDNA transcription.

Several genes of interest have been inserted into the pSBL-CtV2 vector, a preferred embodiment of the universal vector. For example, the vector pSBL-CtV3 (Figure 7B) contains the reporter gene mGFP that codes for a green fluorescent protein, isolated from jelly fish. This gene may also be useful for visible selection of transformed plants or in ornamental horticulture, for example, in ornamental crops like Christmas trees or even lawn grass, which may glow with green fluorescence upon illumination with blue light.

The vector pSBL-CtVH (Figure 7C) contains a different selectable marker, hygromycin phosphotransferase (hph gene driven by the chloroplast atpB gene promoter), which confers resistance to the antibiotic hygromycin. This vector may be used to transform plants that are resistant to other antibiotics and is particularly useful for transforming monocots, which are generally resistant to other commonly used antibiotics. This gene may confer additional traits such as herbicide resistance, a non-targeted trait.

Vector pSBL-CtVHF (Figure 7) contains the GFP and hph genes, which can be used for lethal or a combination of lethal/visible selection.

A Chloroplast Vector Specific for Tobacco and a Universal Chloroplast Vector. The tobacco chloroplast vector pZS-RD-EPSPS (Figure 2A) ("TV") and the universal vector pSBL-RD-EPSPS (Figure 2B) ("UV") contain both the Prrn promoter (of the 16S rRNA), the *aadA* gene (for spectinomycin selection), the mutant

aroA gene that codes for the enzyme EPSPS synthase (for glyphosate selection) and the *psbA* 3' region. The flanking sequences in pZS-RD-EPSPS contain *rbc*L and *orf* 512 and in pSBL-RD-EPSPS contain the *trnI* and *trnA* genes facilitating integration into either the Large Single Copy region (Figure 1 at the "TV" arrow) or the inverted repeat regions (Figure 1 at the UV arrows)of the tobacco chloroplast genome, respectively.

Glyphosate is the active ingredient in Monsanto's herbicide ROUNDUP™ and is used as a selectable marker for herbicide selection of transgenic plants.

Construction of Chloroplast Vectors. Standard protocols for vector construction including Klenow filing the dephospphorylation , were used. The tobacco chloroplast expression vector pZS-RD-EPSPS is shown in Figure 2A. The universal chloroplast vector pSBL-RD-EPSPS is shown in Figure 2A. The construction of these vectors is further shown in the examples. Both plasmids were amplified in the XL1 Blue Strain of *E.coli*. Growth curves were recorded in M-9 minimal medium. Both vectors are used for selection on glyphosphate to confirm resistance to ROUNDUP™.

The chloroplast expression vector pZS-RD-EPSPS is specific for tobacco and as noted earlier is not useful to transform other plants(Maier et al., 1995). In contrast, the universal chloroplast expression and integration vector pSBL-RD-EPSPS (Figure 2B) is competent to transform chloroplast genomes of numerous other plant species because of the universality of the vector as described above. The universal vector integrates foreign genes into the 16S-23S spacer region of the chloroplast genome. The universal vector uses *the trnA* and *trnI* genes (chloroplast transfer RNAs) coding for alanine and isoluceine) from the inverted repeat region of the chloroplast genome as flanking sequences for homologous recombination. The chloroplast border sequence used in this

invention also contains the chloroplast origin of replication (*oriA*), as confirmed in several crop species including pea (Nielsen et al., 1993) and tobacco (Lu et al., 1996), which may explain the highly conserved sequence homology in this region. This origin of replication provides increased number of plasmid templates for efficient integration into the recipient chloroplast genome and achieve homoplasmy.

As shown above, in the construction of the universal vector, an expression cassette containing a chloroplast promoter, a selectable maker gene conferring resistance to an antibiotic (or other selected marker), a gene encoding the target molecule, and the other elements (as described herein) are inserted at a convenient restriction site into the DNA fragment containing the spacer region. If desired, the foreign gene encoding the target molecule may be inserted into the expression cassette after insertion of the cassette into the DNA fragment containing the conserved spacer region so that, before insertion, the cassette will include multiple cloning sites for insertion of one or more DNA coding sequences.

The position of the restriction site in the spacer sequence can determine the respective length of the two flanking sequences, which will be fractions (of different or same length) of the spacer region. Thus, the two flanking sequences need not be identical in length as long as each one contains enough of complementarity to the target chloroplast genome to promote homologous recombination.

Because the vector of the invention has such a high degree of homology to the spacer region of the chloroplast genomes of multiple species of plants, it is competent to transform, not only the species of plants from which the border sequence of the vector is derived, but any plant species.

As used in this specification, the term "homologous" means that a DNA sequence from one plant species possesses regions of sequence identity to portions of a DNA sequence from another plant species. That is, if two DNA sequences are

highly homologous, the DNA species may have 100% sequence identity or less than 100% identity. For example, for purposes of chloroplast genome integration, a 400 bp sequence which is only 25% homologous overall, but which contains a 100 bp portion which is 85% to 90% or more homologous, is considered to be highly homologous with the chloroplast genome.

The inclusion of a chloroplast *ori* within the flanking sequences of the universal vector has also been shown to be beneficial. Without being bound by theory, it is believed that the presence of the *ori* in the universal vector promotes homoplasmy following a single round of selection, without the need for a second selection step. This is especially important in the transformation of monocotyledonous plants, such as maize or rice, in which a second selection step is not feasible due to the difficulty of growing these plants in culture from leaf cuttings with the resultant need to grow these plants from embryos. If an *ori* is desired but is lacking, it may be introduced into a flanking sequence or elsewhere. If it is desired to increase the copy number of the introduced universal vector, a chloroplast DNA fragment containing an *ori* will be inserted outside the flanking sequences so that it will function only to amplify the copy number of the universal vector and does not become integrated into the chloroplast genome.

As opposed to being derived from a specific plant, the flanking sequences can be derived from a spacer region synthetically made as shown below.

For transcription and translation of the DNA sequence encoding the polypeptide of interest, the entire promoter region from a gene capable of expression in the chloroplast generally is used. The promoter region may include promoters obtainable from chloroplast genes, such as the psbA gene from spinach or pea, the rbcL and atpB promoter region from maize and rRNA promoters. Competent promoters are also described and other literature sources are identified in patent No. 5,693,507.

The flanking sequences shown by patent No. 5,693,507 and the other publications to promote stable integration are not the flanking sequences of the universal expression and integration vector described herein which are highly conserved from plant species to plant species, whereas the flanking sequences of that patent and the other publications are not.

Identification of Intergenic Spacer Sequences. The invention provides methods to identify appropriate untranscribed intergenic spacer sequences in plants which are appropriate to construct the universal vectors. The method comprises isolating plastid genomic DNA, carrying out hybridization with a radioactive labeled probe of a known spacer, detecting and isolating plastid sequences which exhibit the desired degree of homology with the probe. As an illustration, to determine if a plastid genome of unknown structure and sequence possesses the spacer region, Southern blots utilizing the tobacco spacer region as a probe are carried out. Plastid genomic DNA is isolated and cleaved by an appropriate restriction enzyme according to established procedures. Hybridization with the spacer probe is conducted under both stringent (e.g., 50% formamide at 68°C, wash in 0.1 X SSC at 68°C) and non-stringent condition (e.g., 6 X SSC, at 68°C, wash in 2 X SSC at 50°C) (1 X SSC is 0.15M NaCl, 0.015M sodium citrate) to detect plastid sequences exhibiting approximately 90-100% homology or 60-100% to the tobacco spacer, respectively. The identified plastid sequences are then isolated. If one's requirement of homologous recombination is more permissive, a lower degree of hybridization to the probe, such about 60% can be satisfactory.

Thus, any known or unknown spacer region of sufficient homology for recombination is suitable for the construction of the UV. Likewise, the known sequence of any intergenic highly conserved spacer sequence may be used to identify and isolate plastid sequences which are homologous to a known spacer sequence.

Alternately, the BLAST program as described herein-above can be used to identify highly conserved regions in plastid genomes of which the sequences are known.

Plants that can be Transformed. Plants which may be transformed by the universal vector of the invention include any lower plants, such as cyanobacteria, any higher plants, such as monocotyledonous and dicotyledonous plant species. The plants to be transformed can be solonacious plants or plants that grow underground. A non-exclusive list of examples of higher plants which may be transformed with the universal vector includes cereals such as barley, corn, oat, rice, and wheat, melons such as cucumber, muskmelon, and watermelon; legumes such as bean, cowpea, pea, peanut; oil crops such as canola and soybean; solanaceous plants such as tobacco, tuber crops such as potato and sweet potato, and vegetables like tomato, pepper and radish; fruits such as pear, grape, peach, plum, banana, apple, and strawberry; fiber crops like the Gossypium genus such as cotton, flax and hemp; and other plants such as beet, cotton, coffee, radish, commercial flowering plants, such as carnation and roses; grasses, such as sugar cane or turfgrass; evergreen trees such as fir, spruce, and pine, and deciduous trees, such as maple and oak. Of greatest present interest are the major economically important crops like maize, rice, soybean, wheat and cotton. None of these plants to the inventor's knowledge, other than tobacco, has ever been stably transformed via the chloroplast genome, and none, including tobacco, have been stably transformed by a universal vector, as described herein. A plant from which the DNA sequence is frequently obtained is tobacco because it is the most thoroughly characterized plant, but any other plant chloroplast genome is suitable.

It will be recalled as described above that the vector used to stably transform tobacco was not competent to transform the chloroplast genome of other plants.

Method for Transformation. The expression cassettes may be transformed into a plant cell of interest by any of a number of known methods. These methods include, for example, the following. Transformation by tungsten particle bombardment , polyethylene-glycol- mediated transformation, use of a laser beam, electroporation, microinjection or any other method capable of introducing DNA into a chloroplast. See, for example, Sanford, 1988; Daniell, 1993; Daniell, 1997; U.S. Patent No. 5,693,507; Kin Ying et al., 1996. The use of these techniques permits the application of the invention described herein to a wide variety of both monocotyledonous and dicotyledonous plants.

Expression of Non-Plant Molecules From Transformed Plants

The increased usefulness of the universal expression integration vector of the invention is clearly shown by the competency of the vector to generate transformed plants to express non-plant and valuable molecules.

Biodegradable Protein-Based Polymers. In accordance with another embodiment of the invention, the universal vector has been used to transform tobacco with a synthetic gene expressing protein-based polymers (PBPs). Such polymers, and the genes expressing them, are known in the literature. (Daniell, et al., 1997). Of particular interest, are protein-based polymers (PBP) which have repeating pentamer sequences like GVGVP (Yeh et al., 1987). These PBP polymers show useful inverse-phase temperature transition property. The protein becomes insoluble when the temperature is raised above the transition state. PBPs offer a wide range of materials similar to that of the petroleum-based polymers, such as hydrogels, elastomers, and plastics. They also show remarkable biocompatability, thereby enabling a whole range of medical applications including the prevention of post-surgical adhesions, tissue reconstruction and programmed drug delivery

(Urry et al., 1993). On the non-medical side, potential applications include use in transducers, molecular machines, superabsorbents and biodegradable plastics (Daniell et al., 1997). Such polymers include the polymer $(Val^1\text{-}Pro^2\text{-}Gly^3\text{-}Val^4\text{-}Gly^5)_n$ $(VPGVP)_n$ wherein "n" can vary from 1 to units in the hundreds, like 250 or their analogs. Important commercial possibilities and related aspects are discussed by Daniell, 1995. Useful biodegradable plastics are made from PBPs. The genes experiencing these PBPs are also useful in the invention in that they can be used as carries i.e. the gene of a molecule of interest can be fused to the PBP gene of chloroplast integration and expression

In a prior study, the synthetic polymer gene coding for $(GVGVP)_{121}$ was hyperexpressed in *E. coli* to the extent that polymer inclusion bodies occupied nearly 80-90% of the cell volume (Guda et al., 1995; Daniell et al., 1997). The same gene was also expressed in the nuclear compartment of cultured tobacco cells (Zhang et al., 1995) and leaves of transgenic tobacco plants (Zhang et al., 1996).

In a model system, the intergenic region of the *trnI* and the *trnA* genes in the 16S-23S rRNA spacer region of the tobacco genome (Figure 1) was used to construct an universal vector for integration of the selectable marker gene *aadA* gene and the synthetic polymer gene (EG121). The vector was inserted into the inverted repeat region of the chloroplast genome. Transformed tobacco plants expressed high level of the polymer protein. Chloroplast genome of other plant species, are also transformable with the synthetic gene to express the protein-based polymer, using the universal vector.

Production of High Value Molecules. The studies with the biopolymer have shown that a non-plant product can be expressed by a synthetic gene, thus making it possible by means of the vectors of the invention to express biologically valuable molecules from transformed plants with a great variety of DNA coding sequences. The DNA coding sequence will

be comprised in a universal vector, or, if desired, in an expression cassette, as described above.

Transgenic plants are known to produce valuable biologically active molecules by nuclear transformation but not via chloroplast transformation. See the following literature references, all of which are incorporated by reference. Daniell, 1995; Miele, 1997; Lyons, 1996; Daniell, and Guda, 1997; Arntzen, 1997.

Expression of Biologically Active Molecules. Plants transformed in accordance with the invention with the universal vector or with the expression cassette can be made to express valuable biologically active molecules in chloroplast containing parts of the plants. The plants will then be harvested by known practices. The transformed plants containing these products can thus be administered orally. Arntzen, 1997. Production of pharmaceuticals by transgenic plants has been performed with peptides (proteins) for many pharmaceutical applications, including vaccines, immunomodulators, growth factors, hormones, blood proteins, inhibitors, and enzymes. Typical transgenic plant derived biologicals which have been reported, include vaccines against viral diseases; viral peptide epitopes, like human immunodeficiency virus, non-viral peptide epitopes, bacterial antigenic proteins; bioactive peptides, recombinant toxins, plantibodies (recombinant antibodies), serum proteins, and plant secondary metabolites. All these products can be expressed in chloroplasts of transgenic plants in accordance with the invention.

Typical pharmaceutical peptides or proteins produced in transgenic plants include hepatitis B surface antigen, norwalk virus capsid protein, foot-and-mouth disease virus, human rhinovirus 14, human immunodeficiency virus, *S. mutans* surface protein, *E. coli* enterotoxin, B subunit, malarial circumsporozoite epitopes, mouse ZP3 protein epitope (vaccine); mouse catalytic antibody 6D4, mouse mAB Guy's 13, mAB B1-8, anti-phytochrome Fv protein, anti-substance P

(antibody); human serum albumin (HSA), human protein C (serum protein); α-trichosanthin, ricin (cytotoxin); human epidermal growth factor (growth factor); leu-enkephalin (neuropeptide) and human acid β-glucosidase (hGC) (enzyme). Many of these molecules have been expressed in tobacco, potato tubers, etc.

Of particular interest, in accordance with the invention is the production of insulin and human serum albumin (HSA) with the universal integration and expression vector or with an expression vector. The HSA has already been produced in transgenic (nuclear) potato and tobacco plants. Sijmons et al, 1990. The aforementioned products can be produced via chloroplast transformation in accordance with the invention.

Insulin. The invention provides a method for expressing insulin as a polymer-fusion protein from a transformed plant. The plant may be transformed with an expression cassette or with the universal integration and expression vector which comprises a synthetic biopolymer gene, like $(GVGVP)_{40}$. A suitable plant is tobacco because of the ease of genetic engineering and the need to find alternative uses of this controversial crop.

For expression in bacteria, the method comprises constructing a vector for expression in *E. coli* as a polymer $(GVGVP)_{40}$ gene fusion with a pro-insulin gene, expressing the polymer-insulin fusion proteins in *E. coli*, (grown in a known manner), purifying the recombinant protein utilizing the temperature-transition properties of the protein-based polymer, and cleaving the insulin from the polymer using well known methods.

For plant transformation, an expression cassette or the universal integration and expression vector, which comprises the polymer gene fusion with the pro-insulin gene, is introduced into a target plant, like tobacco. The polymer-insulin fusion protein is expressed from the plant, the polymer fusion protein is extracted from the chloroplast and the cytosol compartments of the plant cells, the insulin is cleaved from the polymer protein with dithiothreitol (DTT) or

the other known methods, and the insulin is collected. Alternatively, insulin polymer fusion product can be expressed in an edible crop or edible parts of the crop.

The technique of fusing a DNA sequence coding for a molecule of biological activity to a synthetic gene expressing a protein-based polymer for expressing, in a suitable bacterial or yeast host or a transformed plant, is a highly promising method of wide applicability.

Recombinant Human Serum Albumin in Plants. In nuclear transgenic tobacco and potato plants, recombinant human serum albumin (rHSA) that is indistinguishable from the authentic human protein has been produced (Sijmons et al., 1990). This showed the expression of a valuable protein in transgenic plants, but also that it was possible to achieve proper processing by fusion of HSA to a plant pro-sequence that resulted in cleavage and secretion of the correct protein. The chloroplast genome of a selected plant like tobacco can be readily transformed with a universal vector as described herein and made to express HSA.

General Applicability. As described herein the universal vector permits, in accordance with the invention, the transformation of plants, make the plant to express a biological molecule which can impart a desired phenotype to the plant and/or produce a desired product which may, but need not have biological activity (or a precursor to a final product). The coding nucleotide sequence can be synthetic, or natural. The produced molecule can be foreign to the plants, non-functional in the plant or functional. The universal vector has broad applications in the domain of plant transformation.

It is contemplated that any biologically active molecule (precursor or derivative thereof) can be produced by transgenic plant transformed with the universal vector of the invention, with suitable adaptations as may be required for a particular case.

Herbicide Tolerance of Chloroplast Transformed Plants

Another embodiment of the invention relates to the use of the universal vector to confer herbicide resistance or tolerance to plants. The advancement of gene transfer technology has made possible the introduction of herbicide resistance genes into plants, thereby making the herbicide selective for a particular crop.

Modifications of the Target Enzyme. The first step towards this approach is the identification and necessary modification of the target enzyme (gene) of the herbicide to confer tolerance, which has been performed quite extensively. Sulfometuron methyl is a sulfonylurea herbicide that blocks the growth of bacteria, yeast, and higher plants by inhibiting the first enzyme of the branched chain amino acids pathway, acetolactate synthase (ALS). Mutant genes for ALS were isolated from *E. coli* and yeast that confer resistance to sulfometuron methyl. Yadav, et al. (1986). Herbicide resistance in tobacco, and several other crops has been achieved by genetic engineering of the ALS gene. Gabard, et al. (1989); Miki, et al. (1990). Yet another approach to engineer herbicide resistance in plants has been the expression of the enzyme phosphinothricin acetyl transferase (PAT) that detoxifies the herbicide phosphinothricin. DeBlock, et al. (1987).

Glyphosate. Glyphosate is a potent, broad spectrum herbicide which is highly effective against annual and perennial grasses and broad leaf weeds. Glyphosate is environmentally safe as it rapidly degrades in soil, has minimal soil mobility, and has very little toxicity to non-plant life forms. Glyphosate works by binding to and inhibiting the enzyme 5-enol-pyruvyl shikimate-3-phosphate (EPSP) synthase. EPSP synthase (EPSPS) catalyzes the formation of EPSP, which is important in the aromatic amino acid biosynthesis pathway, from shikimate-3-phosphate and inorganic phosphate. The non-toxicity of glyphosate to animals is due to the fact that this reaction

occurs only in plants and microorganisms. Unfortunately, because the reaction to form EPSP occurs in all plants, glyphosate does not have selectivity between weeds and desirable plants such as crops and ornamentals.

Two approaches have been used to attempt to develop a glyphosate resistant plant by genetic engineering. One approach is based upon overproduction of wild type EPSP synthase, so that after competitive inhibition of EPSP synthase by glyphosate, the residual EPSP synthase confers glyphosate tolerance. The second approach is based upon the expression of a mutant gene (*aroA*) encoding glyphosate resistant EPSP synthase.

In all of the aforementioned examples, without exception, herbicide resistant genes have been introduced into the nuclear genome.

The Need for Chloroplast Transformation. A serious need exists, to develop a herbicide resistant plant, particularly a plant resistant to the most widely used herbicides, in which the protein conferring herbicide resistance is produced in the chloroplast, and in which the gene conferring herbicide resistance cannot escape by pollen to the environment.

The universal vector of the invention responds to this need by transformation of any target plant to provide tolerance to any selected herbicide like to glyphosate. Important commercial crops like wheat, rice, corn (maize), soybean can be made resistant to a selected herbicide by means of the universal vector.

The invention provides a transgenic herbicide resistant plant in which a foreign transgene conferring resistance to one or more herbicide is integrated into the chloroplast genome by means of the universal vector. The transgenic plant may be a mature plant, an immature plant, such as a seedling, an embryo, a callus, a cultured tissue, or cell suspension, or a portion of a plant such as a cutting or a callus. Herbicides which are suitable for the invention and for which genes conferring resistance may be stably integrated

into the chloroplast genome in accordance with the invention include all known (and also to be developed) herbicides.

Class of Herbicides. Herbicides controllable by the invention have been generally grouped into several chemical classes.

A first class includes the PSII (Photosystem II) herbicides which interfere with the reduction of plastoquinone on the acceptor site of PSII which include such chemicals as triazines, triazinones, urea derivatives, biscarmates, nitriles, nitrophenols, substituted pyridazinones, phenylcarbamates, anilides, cyanoacrylate, DCMU, carboanilides, uracils, specifically for example, dichlorophenyldimethylurea, atrazine, metribuzine, lenacil, phenmedipham, ioxynil and dinoseb. These chemicals bind to a 32-kDa binding protein (Q_B protein, D1, herbicide-binding protein) in the thylakoid membrane of chloroplasts, thereby blocking photosynthetic electron transport. The plastid gene which codes for a precursor of the Q_B protein is named *psbA*, the sequences of which show a very high degree of homology in different plants. The most important PSII herbicide is atrazine, developed by Ciba-Gergy.

Another class are the PSI (Photosystem I) herbicides, a membrane bound protein complex which catalyzes the light-driven oxidation of plastocyanin (PC) and reduction of ferredoxin (Fd). Typical of these chemicals are the bipyridyl herbicides paraquat and diquat.

Another class of herbicides are the aryloxyphenoxypropanoates (APP) or acetyl coenzyme A carboxylase type herbicides, which inhibits acetyl-CoA carboxylase and subsequent fatty acid biosynthesis in plastids. Typical of the APPs are cydohexanedione (CHD), sethoxydin, haloxyfop, quizalofop, phenoxaprop (and the lower alkyl-substituted molecules thereof), dicolofop, sethoxydin, clethodin and tralkoxydim.

Yet another class of herbicides includes the auxin analogs like macropop, chloramben, dicamba, benazolin, 1-

naphthylacetic acid; 3,6-dichloropicolonic acid, picloram, fluoroxypyr, quinclorac, MCPA and 2,4-D.

An additional class are the mitotic herbicides, termed dinitroaniline herbicides, like trifluralin, oryzalin and pendimethalin.

Another chemical class of herbicides to which the invention applies are those acting in the biosynthesis of amino acids, such as tertiary the amino methyl phosphonic acids chlorsulfuron, glufosine and glyphosate.

Another class of herbicides are the acetolactate synthase- inhibiting herbicides (ALS), like the sulfonylureas, imidazolinones, triazolopyrimidines and pyrimidinyl thiobenzoates, such as chlorsulfuron, imazaphyr, flumetsulam (and others listed in chapter 4, Table I of Herbicide Resistance in Plants, 1994, cited below).

Examples of sulfonylurea herbicides are sulfometuron methyl (the active ingredient of Oust™) and chlorsulfuron (the active ingredient of Glean™). Imazapyr, one of the imidazolinones, is the active ingredient of American Cyanamid's herbicide Arsenal™ and imazamethabenz is a mixture of two imidazolinones (Merk Index, 11th Ed. 4825). Mutated forms of ALS located in the structural genes of ALS, *ilvG* and ILV2 may be used to confer herbicide resistance using the universal vector.

In spite of the chemical differences between the imidazolinones and sulfonylureas, these substances inhibit the same enzyme, ALS. It appears that quinones and an imidazolinone herbicide compete with a sulfonylurea herbicide for a common site on ALS. Accordingly, in accordance with the invention, plants can be transformed which will show a resistance to both groups of these and other herbicides.

Another group of chemicals controllable by the invention using the universal vector which has herbicidal activity is typified by L-phosphinothricin, which is a component of the tripeptide "bialaphos". This tripeptide is marketed under the trade name "Herbiace™" by Meiji Seika, Japan, and as "Basta™" by Hoechst AG, Germany. L-

Phosphinothricin is a potent inhibitor of glutamine synthetase, which causes a rapid increase of ammonia concentration plants and leads to the death of the plant cell.

In spite of extensive studies, no satisfactory product has been developed to impart herbicide tolerance to plants by chloroplast transformation. With the glyphosate type herbicide, it has been reported that regenerated nucleus transformed transgenic plants showed tolerance to glyphosate, but also showed impaired growth, and have not resulted in transgenic plants with high level of tolerance to glyphosate, Schultz, et al.,1990.

Chloroplast Transformants. In accordance with the invention, the chloroplast of target plants that are susceptible to herbicides is transformed with a vector of the invention carrying the necessary coding sequence, thereby conferring herbicide resistance. The transformed chloroplast comprises a genome which carries a foreign transgene which confers the herbicide tolerance. The chloroplast may be a mature chloroplast or may be an immature chloroplast, such as an etioplast. Preferably, the gene conferring herbicide resistance codes for EPSP synthase which binds less readily (has reduced affinity) to glyphosate than does wild type the EPSP synthase. If present, the second transgene is generally a gene which confers antibiotic resistance to the plant. Thus, the herbicide resistance can be used by way of lethal selection as a marker for chloroplast transformation by selection of the transformants by exposure to a medium with lethal concentration of the selected herbicide, to which the transformants will survive.

The origin of the second gene can be procaryotic (for example, bacterial) or eucaryotic (for example, plant).

The invention also relates to produce a plant resistant to several herbicides, whether of the same class of herbicide or of different classes of herbicides, and the transformed plant resistant to multiple herbicides.

Some plant species are known to have a developed a natural and non-permanent tolerance to certain types of

herbicides. The teaching of this invention is readily applicable thereto.

The invention includes a method for producing a herbicide resistant plant which comprises transforming the chloroplast of the plant by introducing one or more foreign transgenes which code for a protein conferring herbicide resistance to the genome of the chloroplast of the plant. Preferably, the transgene codes for a mutant form of the enzyme which has decreased affinity for a given herbicide than does the naturally occurring enzyme.

The following Table lists a variety of type of resistance determinants, (chemicals or "molecules") which inhibit or which confer resistance, and typical herbicides related thereto.

Table I

RESISTANCE DETERMINANT	HERBICIDES
glutathione S-transferase	s-triazine simazine chloracetamide metalachlor
Auxin analogs	2, 4-D MCPA mecopop chloramben
EPSP synthase	glyphosate
Q_b (*psbA*) - PS II Type	atrazine terbutyne dichloropheny- dimethylurea metribuzine lenacil phenmedipham loxynil dinoseb
Acetohydroxyacid sythase (ALS)	sulfonylureas chlorosulfuron imazapyr sulfometuron methyl imidazolinones
Glutamine synthase	phosphinothricin

PS I Type	paraquat diquat
Acetyl coenzyme A carboxylase inhibiting enzymes (APP Type)	aryloxyphenoxypropanoate (APR), cyclohexanedione
Mitotic disrupters	trifluralin, oryzalin, pendimethalin, dinitroaniline

For a comprehensive review on the topic of herbicide resistance in plants, see Herbicide Resistance Crops, Agricultural, Environmental, Economic, Regulatory and Technical Aspects, 1996, CRC Press, Inc., Editor, Stephen O. Duke and Herbicide Resistance in Plants, Biology and Biochemistry, 1994, CRC, Press, Edited by Stephen B. Powles and Joseph A.M. Holtum. In the first of these reference books, Chapter 3 on Techniques for Producing Resistant Crops (and the numerous references cited therein) is of particular interest as background to the invention. Both books are incorporated herein by reference in their entirety.

In accordance with the invention it has also been discovered that in the process for expressing a target trait another trait can be expressed by the transformed plant which may be quite desirable, indeed may be more desirable than the initially target trait. In those situations, the plants will be allowed to express and will be selected on the basis of that other trait.

The invention is exemplified in the following non-limiting examples.

Example 1

Universal Chloroplast Integration and Expression Vectors

Tobacco. Exemplary universal chloroplast vectors were constructed by first cutting out the tobacco chloroplast DNA *Bam*HI fragment (130656-140992) containing the 16S and 23S rRNA genes and subcloning it into a commonly available bacterial plasmid pUC19. A map of the tobacco chloroplast genome is shown in Figure 1. A 2.1 kbp *Hin*dIII-*Eco*RI fragment present

within this fragment, containing a universal border sequence comprising *trnI* and *trnA* genes (Figure 5A), including the spacer region between the genes, was subcloned into the pUC19 plasmid at the *Pvu*II site (Figure 5B). The resultant plasmid was designated pSBL-Ct Bor (Figure 5C).

The vector pSBL-RD-EPSPS (Figure 2B) contains a mutant EPSP synthase gene that codes for the enzyme EPSP synthase. Glyphosate, the active ingredient in Mosanto's ROUND UP™, binds to the protein EPSP synthase and blocks the synthesis of essential amino acids, resulting in death of a plant. The EPSP synthase coded for by the mutant gene does not bind glyphosate, and therefore confers herbicide resistance to crop plants.

Other genes, such as those that confer resistance to adverse environmental factors such as salt/drought tolerance (osmotolerance genes such as betaine aldehyde dehydrogenase, (BADH), for the overproduction of glycine betaine) or thermotolerance (genes coding for heat shock proteins) or cold shock tolerance proteins, or to pathogen resistance, such as antimicrobial (lytic peptides, chitinase) or antiviral (coat proteins) can be inserted singly or in non-conflicting combinations into the universal chloroplast vector, or into different cassettes of the same universal chloroplast vector to transform the target plant into one with the desired trait.

Construction of a Universal Chloroplast Integration Vector Containing a Synthetic Spacer 2 Region

A universal chloroplast vector containing only the spacer 2 region of the tobacco chloroplast genome was constructed by first subcloning a synthetic oligonucleotide comprising the spacer 2 region into the bacterial plasmid pUC19. The positive and negative strands of the 64 base pair spacer sequence were synthesized, the sequence of the positive strand was as follows:

5'-GCTGCGCCAGGGAAAAGAATAGAAGAAGCATCTGACTACTTCATGCATGCTCCACTTGGCTCGG-3'

The synthetic fragments were mixed and allowed to anneal, then ligated into pUC19 at the *Pvu*II site. (Figure 5B) Insertion of an appropriate selectable marker gene and a heterologous gene were as described above for pSBL-CtBor. (Figure 5C)

To prepare a longer sequence which includes the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes, the same methodology is followed.

TRANSFORMATION OF DIFFERENT PLANTS

Example 2

Chloroplast Transformation of Tobacco. The following example describes a classic protocol for transformation of tobacco chloroplast for which any vector can be used. Two such vectors are identified below. All new chloroplast vectors were first tested in tobacco as described in Daniell, (1997). Tobacco (*Nicotiana tabacum* var. Petit Havana) plants were grown aseptically by germination of seeds on MSO medium containing MS salts (4.3 g/liter), B5 vitamin mixture (myo-inositol, 100 mg/liter; thiamine-HCl, 10 mg/liter; nicotinic acid, 1 mg/liter; pyridoxine-HCl, 1 mg/liter), sucrose (30 g/liter) and phytagar (6 g/liter) at pH 5.8. Fully expanded green leaves of about two month old plants were selected for bombardment.

Leaves were placed abaxial side up on a Whatman No. 1 filter paper laying on RMOP* medium in standard Petri plates (100x15 mm) for bombardment. Tungsten (1 μm) or Gold (0.6 μm) microprojectiles were coated with plasmid DNA, of interest (e.g. pSBL-RD-EPSPS or pZS-RD-EPSPS) and bombardments were carried out with the biolistic device PDS1000/He (Bio-Rad) as described by Daniell, 1997. Following bombardment, petri plates were sealed with parafilm and incubated at 24°C under 12 h photoperiod. Two days after bombardment, leaves were chopped into small pieces of about 5 mm^2 in size and placed on the lethal selection medium (RMOP containing a selectable marker such as about 500 μg/ml of spectinomycin dihydrochloride) with abaxial side touching the medium in deep

(100x25 mm) petri plates (about 10 pieces per plate). Selected from the shoots that died, the regenerated spectinomycin resistant shoots were chopped into small pieces (about 2mm^2) and subcloned into fresh deep petri plates (about 5 pieces per plate) containing the same lethal selection medium. Resistant shoots from the second culture cycle were transferred to rooting medium (MSO medium supplemented with IBA, 1 μg/liter and an appropriate antibiotic like 500 μg/ml of spectinomycin dihydrochloride,). Rooted plants were transferred to soil and grown at 26°C under continuous lighting conditions for further analysis.

After transfer to the lethal selection medium, the explants gradually became pale and in about 3-8 weeks, green calli and shoots developed from the bombarded side of the 0leaf. Resistant shoots from each callus were considered as a clone.

PCR screening for chloroplast transformants after the first culture cycle showed that 12 out of 20 resistant clones integrate the foreign genes like the *aadA* gene linked to the EG121 gene into the chloroplast genome. These 12 clones were advanced to further steps of regeneration. The entire process of regeneration, starting from bombardment until transfer to soil, takes about 3-5 months.

Figure 9 shows transformed and untransformed tobacco plastids growing in the presence of spectinomycin indicating non-lethal selection on the medium (500μg/ml).

Example 3

Corn Chloroplast Transformation. Surface sterilization and germination of corn seeds. Corn seeds are surface sterilized in a solution containing 20% (v/v) commercial bleach and 0.5% SDS for 15 min under continuous shaking, then serially rinsed in sterile double-distilled water (sddw) four to five times. Liquid MS-based germination medium (modified CSG) containing MS salts (4.3 g/l), sucrose (30 g/l), DM-vitamins (1.0 mg/l thiamine-HCl, 0.5 mg/l nicotinic acid, 0.5 mg/l pyridoxine-HCl and 100 mg/l myo-inositol) and BA (2.0 mg/l) at pH 5.8 is

dispensed per Magenta™ box (45 ml) containing eight layers of cheesecloth, then autoclaved. Seeds are placed in modified CSG (25 seeds of any genotype per box) and cultured for three days (16 h or continuous light; 25C) for germination. Nodal sections are excised aseptically from three day-old seedlings. The nodal section appears as a clear demarcation on the germinating seedling and represents the seventh node (Figure 10A). When excised, the nodal cross-sections are approximately 1.2 -1.5 mm in length (Figure 10B).

Figures 10A-G shows corn plastid transformation and regeneration scheme. A) Three day-old corn seedling; arrows and line depict the seventh node for explant excission; B) Nodal cross sections prior to bombardment; arrows depict margin of one section; C) GUS-positive nodal section (nuclear transformation); histochemical assay conducted three days post-bombardment; D) Multiple shoot induction from one nodal section (control) after eight weeks in culture; E) Control shoot on elongation medium for three weeks; F) Rooted control plantlet; G) Selection of plastid transformed corn in liquid medium containing spectinomycin and streptomycin for eight weeks.

Multiple shoot induction: Nodal section explants are placed on corn shoot induction medium [CSI; MS salts, sucrose and DM-vitamins as above, BA (2.0 mg/l), CPA (0.25 mg/l) and phytagar (8 g/l) at pH 5.8], acropetal end up, and placed under the culture conditions previously mentioned. In all media except modified CSG and RG1 (rice), PGRs and antibiotics are filter-sterilized and added after autoclaving. Tissues are subcultured every two weeks onto fresh CSI medium for multiple shoot formation (Figure 10D). Adventitious shoots are separated from the shoot clumps after eight weeks of culture and elongated on semi-solid MS-based medium containing sucrose, DM-vitamins, glycine (10 mg/l) and asparagine (150 mg/l) at pH 5.8, for three weeks (Figure 10E). The plantlets are rooted (Figure 10F) on the same medium containing IBA (0.5 mg/l). Rooted plantlets can be grown in PGR-free liquid MS in test tubes (150 X 25 mm) containing cheesecloth as the anchor

material to achieve faster growth. Regenerated plantlets are transplanted to potting media, acclimatized then grown to maturity in the greenhouse.

Figure 11 shows corn plastid transformation. Transformed corn plants grow normally (middle shoot) while untransformed plants die on the lethal medium, confirming lethal selection by the antibiotic spectinomycin (1000μg/ml).

Example 4

Rice Chloroplast Transformation. Surface sterilization of rice seeds and preculture. Dehusked seeds from any genotype (indica or japonica types) are surface sterilized first in 70% ethanol for 10 min under continuous shaking then rinsed with ddw about five times. Seeds are then soaked in a 0.2% Benlate (w/v) solution for 20 min, rinsed with sddw five times, then in 50% bleach for 20 min with the sddw rinses repeated. Seeds are pre-cultured in medium RG1 [MS salts, sucrose and DM-vitamins as above, BA (2.0 mg/l) at pH 5.8]. As with corn, liquid RG1 is dispensed to Magenta™ boxes containing cheesecloth prior to autoclaving. Seeds are placed in RG1 (100 seeds of any genotype per box) and pre-cultured overnight (16 h or continuous light; 23C) prior to bombardment the following day.

Bombardment of embryos on intact rice seeds. Pre-cultured seeds are tightly packed vertically, embryo end up (Figure 12A), in the central 2.5 cm area of a petri dish (25 per dish) containing medium RG1.1 (RG1 plus 8.0 g/l phytagar) and bombarded with DNA-coated microprojectiles .

DNA precipitation. The procedure is as described for corn with the following modifications. Ten μl DNA (1.0 μg/μl) and 20 μl isopropanol (2X vol of DNA), 60 μl 2.5 M $CaCl_2$ and 15 μl 0.1M spermidine are used. Each shot delivers 2.0 μg DNA and 720 μg tungsten.

Figures 12A-F shows rice plastid transformation and regeneration scheme. A) Rice seeds, embryo end up, just prior to bombardment; arrows point to embryo margins; B) Multiple shoot induction from one control embryo after seven weeks in culture; C) Selection of plastid transformed rice shoots arising from one initial rice embryo in media containing spectinomycin and streptomycin after eight weeks in selective medium; arrows point to two putative transformants; D) Control rice regenerants; E) Transgenic Priscilla 2.3; F) Transgenic Priscilla 2.4.

Multiple shoot induction and selection of transplastomics post-bombardment. The rice seeds are separated and spread out on the RG1.1 medium (maintaining the polarity) and placed in the dark for two days post-bombardment. The embryo end of the seed is then cut away (embryo plus small amount of endosperm) from the remainder of endosperm which is discarded. Embryos are placed in 50 ml liquid RG2 medium (in 250 ml flask) for multiple shoot induction (Figure 12B). RG2 contains MS salts, sucrose and DM-vitamins as above, and BA (6.0 mg/l) at pH 5.8. RG2 used for selection includes spectinomycin (1000 μg/ml) plus streptomycin sulfate (100 μg/ml). The cultures are placed in the growth chamber (16 h photoperiod; 25°C) and subcultured every two weeks into fresh selection media. Green shoots are selected from the shoot clumps arising from each embryo and placed back in selective media (Figure 12C). Rooting is achieved in medium RG3 [MS salts, sucrose and DM-vitamins as above, IBA (0.5 mg/l) at pH 5.8] plus antibiotics. (Shoots can be either rooted separately or as multiple shoot clusters). Plantlets are transplanted to potting media, acclimatized then repotted to a clay: sand (1:1) mix and grown to maturity in the greenhouse. (Figures 12D,E,F).

Development of plastid transformation and regeneration protocols for corn and rice. As described above unique corn nuclear transformation and regeneration protocols (Figure 10) were developed (Rudraswamy, 1997) and adapted for plastid transformation. (Previous to this work, nodal section explants had not been used for transformation or

regeneration.) Multiple shoots were induced on nodal sections excised from three day-old seedlings of 21 genotypes (none related to A188 or B73) which included hybrid (16 grain, one sweet) and inbred (four) genotypes. After eight weeks in culture, 16 - 32 shoots (avg. 24) were generated per explant. Shoots were rooted and regenerants did not display aberrant phenotypes in greenhouse analyses (limited study of two plants per genotype). DNA could also be delivered into nodal section explants of all genotypes (Figure 10C; transient β-glucuronidase expression). For plastid transformation, nodal section explants were bombarded with pSBL-ctV2, then placed on a multiple shoot induction medium containing spectinomycin and streptomycin. Arising shoots could be excised and re-placed on shoot induction medium for subsequent rounds of selection.

As described above unique rice targets dehusked intact mature seeds, embryo end up, not used in previously reported transformation protocols were coupled with a multiple shoot induction protocol (Figure 12) for mature embryos (excised two days post-bombardment). Multiple shoots were induced on all eight genotypes tested (Litton, Priscilla - two newly released Mississippi cultivars, plus six breeding lines). The noted response should be similar in numerous other cultivars since the initial explant is a mature embryo. Regenerants (non-transformed) are being maintained for collection of F1 seed. After plastid transformation, shoot multiplication occurred in the presence of spectinomycin/streptomycin and, as with corn, shoots could undergo numerous rounds of selection due to shoot proliferation (unknown if axillary or adventitious in origin) from the base of excised shoots. Rooting was also accomplished in selective media.

Figures 13A-B shows PCR analysis of DNA isolated from first generation leaves of rice transformants. PCR analysis was done with DNA isolated from the first generation leaves. The PCR products were not abundant as observed in tobacco chloroplast transgenic plants (Figure 13A, lane 11, 13B, lane 12). This may be because of two reasons. The

protocol used to isolate DNA is not suitable for coarse rice leaves or that the primers designed for tobacco do not anneal as well with rice ct DNA. Nevertheless, for preliminary analysis, tobacco primers were used to test integration of the *aadA* gene into the plant genome from the universal vector. Lack of a product would indicate spontaneous mutants, capable of growing on spectinomycin without the *aadA* gene (Figure 13A, lanes 7-10). A PCR product of 1.57Kb was detected in four lines (Figure 13A, lanes 2-6) transformed with the universal vector. Under the selection conditions used, four mutants were detected out of ten lines transformed with the universal vector. Primers were also designed to specifically identify integration into the plastid genome. For the universal vector, the primer on the native chloroplast genome landed in the 16s rRNA gene, outside the flanking sequence of the chloroplast vector (1.60Kb PCR product). The expected products were observed for the transgenic lines obtained using the universal vector (Figure 13B, lanes 5,6). Unbombarded plants (controls) did not yield any PCR products, as expected (Figure 13B, lane 2; Figure 13A, lane 1). PCR results identified two `Priscilla' rice plants (2.3 and 2.4) which contain transformed plastids (Figures 12 and 13).

Example 5

Peanut Chloroplast Transformation (*Arachis hypogaea*).

Transgenic peanuts having transformed chloroplast genomes were obtained using the universal vector pSBL-CG-CtV2 (Figure 7A). Peanut tissue was grown in culture in accordance with the protocol described by Kanyand et al., 1994. Bombardment conditions were the same as for tobacco chloroplast transformation as described above, except that epicotyl sections were used for bombardment while using rupture discs of variable pressure. Peanut chloroplast transformation has never been previously reported.

Figure 14 shows peanut plastid transformation. Transformed peanut plants grow normally (middle and on left

side of plate) while untransformed plants die in the lethal medium (500 μg/ml).

Example 6

Soybean Chloroplast Transformation. Transgenic soybeans having transformed chloroplast genomes were obtained using the universal vector pSBL-CG-CtV2 (Figure 7A). Bombardment conditions were as for tobacco chloroplast transformation. Soybean chloroplast transformation has never been previously reported.

Figure 15 shows soybean plastic transformation. Two transformed plants show shoots, the other plant die on the lethal medium, confirming lethal selection by the antibiotic spectinomycin (500 μg/ml).

Example 7

Sweet Potato Chloroplast Transformation. Transgenic sweet potato plants having transformed chloroplast genomes were obtained using the universal vector pSBL-CG-CtV2 (Figure 7A). Sweet potato tissue were grown in culture in accordance with the protocol described by Zhang et al.,1996. Bombardment conditions were the same as for tobacco chloroplast transformation as described above, except that calli and primary embryos were bombarded and, after bombardment, were transferred to plates containing 100mg/ml spectinomycin. Sweet potato chloroplast transformation has never been previously reported.

Figure 16 shows sweet potato embryos transformation on the lethal antibiotic spectinomycin selection medium (500 μg/ml). Note bleached calli (right) and green embryos (left).

Example 8

Grape Chloroplast Transformation. Transgenic grape plants having transformed chloroplast genomes are obtained using the same universal vector pSBL-CG-CtV2. Grape tissue are grown in culture according to the protocol of Hebert et al., 1993. All

chloroplast transformation protocols are as for tobacco, except that cells in the exponential phase of growth, about 4 days after subculturing, were used for bombardment. Grape chloroplast transformation has never been previously reported.

Figure 17 shows grape cells transformation. The transformed culture cells become green while the untransformed cells die in the lethal antibiotic spectinomycin selection medium (500 μg/ml).

Example 9

Transformation of Other Plants. Transformation of plants by means of microprojectile bombardment is a favored technique to introduce the universal vector carrying the desired nucleotide sequence coding for the molecule of interest into the target plants. Illustrative transgenic plants obtained through microprojectile bombardment are shown in Table II.

Table II

Transgenic Plants Recovered

Through Microprojectile Bombardment

Plant Species	Explant Utilized for Transformation
Alfalfa	Cali from petiole, stem sections
Arabidopsis	Root sections
Barley	Embryogenic callus, immature embryos
Banana	Embryogenic suspension cells
Bean	Meristems
Citrus	Embryogenic cells
Cotton	Embryogenic suspensions; meristems
Cranberry	Stem sections
Cucumber	Cotyledons
Dendrobium orchid	Protocorms
Eucalyptus	Zygotic embryos
Grape	Embryogenic suspension cells
Maize	Embryogenic suspensions; immature embryos
Oat	Embryogenic
Papaya	Zygotic/somatic embryos; hypocotyls

Pasture grass	Embryogenic calli
Peach	Embryo derived calli
Peanut	Meristems
Poplar	Embryogenic
Rice	Zygotic embryos

The transformation of plants by the use of the gene gun is described in Daniell, 1997. Each crop that was reported to be nuclear transformable via microprojectile bombardment in that Table can have its chloroplast genome transformed using the universal vector as described herein.

Example 10

Expression of Non-Plant Products

The examples that follow, illustrate the expression of biodegradable protein-based biopolymers (PEPs) and analysis of transformants.

Vector pSBL-CG-EG121. The vector pSBL-CG-EG121 (Figure 3A) contains the gene $(GVGVP)_{121mer}$ (designated EG121) which codes for a biodegradable protein-based biopolymer (PBP) that has many medical and non-medical applications.

Construction of Chloroplast Expression Vectors. Standard protocols for vector construction were as outlined by Sambrook et al., 1989. Chloroplast integration and expression vectors pSBL-CtV2 (Figure 7A) and pZS197 were digested, respectively, with *Xba*I (an unique site between the *aad*A gene and the *psbA* 3' region) and *Spe*I (a unique site at 120 bp downstream of the *aad*A geen in the *psbA* 3' regulatory region), Klenow filled and dephosphorylated. The polymer gene EG121 along with the Shine-Dalgarno sequence (GAAGGAG) from the pET11d vector was excised as a *Xba*I-*Bam*HI fragment from the plasmid pET11d-EG121. Sticky ends of the insert fragment were Klenow filled and ligated with vectors pSBL-CtV2 or pZS197 yielding chloroplast expression vectors pSBL-CG-EG121 (Figure 3A) and

pZS-CG-EG121 (Figure 3B), which integrate the *aadA* and EG121 genes at the inverted repeat (IR) or into the spacer region between the *rbcl* and *orf* 512 genes of the tobacco chloroplast genome. Refer to Figure 1, for "UV" and "TV", integration sites, respectively.

Biopolymer Expression in E. coli and Tobacco. Plasmid vector pSBL-CG-EG121 (Figure 3A) was transformed into *E. coli* strain XL-1 Blue and grown in Terrific Broth in the presence of ampicillin (100 μg/ml) at 37°C for 24 h. SDS-PAGE was carried out according to Laemmli, 1970 using a 12% resolving gel and a 5% stacking gel and run for 5 h at a constant current of 30 mAmps. Crude protein extracts from *E. coli* cells were prepared and electrophoresed as described by Guda et al., 1995. After electrophoresis, polypeptides were visualized by negative staining with $CuCl_{2,.}$

Figure 18 shows expression of chloroplast integration and expression vectors in *E. coli* strain HMS174 (DE3). Lane 1, shows the purified polymer protein; lane 2, shows the untransformed *E. coli* control; lane 3, shows the *E. coli* strain XL-1 Blue transformed with universal vector pSBL-CG-EG121 (Figure 3A); lane 4 shows *E. coli* transformed with the tobacco vector and lane 5, shows *E. coli* strain HMS174(DE3) transformed with pET11d-EG121 vector, in which the T7 promoter transcribes the polymer gene. The level of expression by the Prrn promoter in *E. coli* was almost equivalent to that of the highly efficient T7 promoter driving the polymer gene.

Southern Blot Analysis. Total DNA was extracted from leaves of transformed and wild type plants using the CTAB procedure of Rogers and Bendich, 1988.

Figures 19 and 20 show Southern blot anlaysis performed independently with the transformants obtained using the tobacco (Figure 19) and the universal (Figure 20) vectors.

Total DNA was digested with *Eco*RI and HindIII in case of the universal vector (UV) transformants or *Eco*RI and *Eco*RV in case of the tobacco vector (TV) transformants. Presence of an *Eco*RI site at the 3' end of the polymer gene allowed excision of predicted size fragments in the chloroplast transformants only. To confirm foreign gene integration and homoplasmy, individual blots were probed with corresponding border sequences. In the case of the TV transformants after the second or third round of selection, the border sequence hybridized with 4.6 and 1.6 kbp fragments (Figure 19A, lanes 2,3 and 4) and with a 3.1 kbp native fragment in the wild type (Figure 19A, lane 1). On the other hand, in the case of the UV transformants, after the first round of selection, the border sequence hybridized with 4.0 kbp and 1.2 kbp fragments (Figure 20A, lanes 1 and 2) while it hybridized with a native 2.1 kbp fragment in the control (Figure 20A, lane 3). Moreover, TV transformants also showed the native fragment of 3.1 kbp (Figure 19A, lanes 2 and 3) similar to the wild type plant indicating heteroplasmic condition of the transformed chloroplast genomes, even though they have been under several rounds of selection. However, both UV transformants showed homoplasmic condition, even after the first round of selection (Figure 20A, lanes 1,2).

Presence of heteroplasmy even after second selection was reported earlier and it was suggested that selection should be done until attainment of homoplasmy (Svab and Maliga, 1993). This is consistent with the observation that a high degree of heteroplasmy exists after a second selection cycle in the TV transformants (Figure 19A, lanes 2 and 3). However, no heteroplasmic condition was observed in case of the UV transformants which may be because of the copy correction mechanism between the two IR regions and/or the presence of chloroplast origin of replication (*ori*) within the border sequence, which should increase the copy number of the introduced plasmid before integration.

DNA gel blots were also probed with either the *aadA* gene (UV integrated plants) or the EG121 gene (TV integrated plants) to reconfirm integration of foreign genes into the chloroplast genomes. In the TV integrated plants, the polymer gene probe hybridized with a 4.6 kbp fragment, only in the plastid transformant lines (Figure 19B, lanes 2,3 and 4). Also, in the UV integrated plants, *aadA* sequence hybridized with an expected 4.0 kbp fragment (Figure 20B) which also hybridized with the border sequence in plastid transformant lines (Figure 20A, lanes 1 and 2).

Analysis of Transcript Levels in Transgenic Tobacco and Northern Blot. Foreign gene transcript levels were analyzed by northern blotting (Figure 21) using total RNA isolated from the control, chloroplast transformants and a tobacco transgenic plant highly expressing the polymer gene (EG121) via the nuclear genome (Zhang et al., 1996). The polymer gene (EG121) sequence hybridized with a 1.8 kbp fragment in the chloroplast transformants (lanes 1-4, 5-6) and also with larger size fragments in one of the chloroplast transformants (lane 6). In the case of the nuclear transformant, a transcript of about 2.1 kbp was observed (lane 7). This was due to the presence of a poly A tail at the 3' end of the polymer transcript provided by the *nos* terminator. The larger size fragments observed in lane 6 may be the di-, tri- or polycistronic transcripts which are being processed in chloroplasts. This is a common phenomenon in chloroplast gene expression because many plastid genes are organized into polycistronic transcriptional units that give rise to complex sets of overlapping mRNAs. Quantitation of transcript levels revealed that the chloroplast transformants were producing a 11-fold (lane 5) or more than fifty fold (lane 6) of polymer transcripts over that of a highly expressing nuclear transformant (lane 7, highest expressing plant among thirty five TV nuclear transgenic plants examined). This is directly attributed to the presence of higher gene copy numbers in

chloroplasts of transgenic plants. The tobacco vector (TV) integrated plastid transformants showed lower levels of polymer transcript (lanes 1-4) compared to the universal vector integrated transformants (lanes 5,6) because the polymer gene exists as two copies per transformed plastid genome in universal vector transformants as against a single copy in the TV transformants and the heteroplasmic conditions observed in TV transformants.

Western Blot Analysis. Polymer protein was purified from leaves from wild type tobacco, chloroplast transformants and nuclear transgenic plants following the method recently described by Zhang et al., 1995. Purified polymer was analyzed by SDS-PAGE according to Laemmli, 1970 using a 12% resolving gel and a 5% stacking gel and run for 5 h at a constant current of 30 mAmps. Polymer polypeptides of about 60 kDa were visualized by negative staining with 0.3 M $CuCl_2$,. Gels were destained in 0.25 M sodium EDTA and 0.25 M Tris-Cl, pH 9.0 with three changes of buffer at 10 min intervals. Western immunoblotting and staining (Figure 22) was carried out as described by Zhang et al., 1996 using a monoclonal antiserum raised against the polymer AVGVP which cross-reacts well with polymer GVGVP and the "Immuno-Blot Assay Kit" (Bio-Rad). The polymer polypeptides running at about 60 kDa are seen in the plastid transformants of IR integrated plants. Polymer expression from a highly expressing F2 generation nuclear transgenic plant (highest expressing plant among 35 transgenic plants examined) is seen in lane 5 (Figure 22), while no polymer was expressed in the untransformed control as seen in lane 4 (Figure 22). Eleven to fifty fold higher level of polymer transcripts is shown in the chloroplast transformants (Figure 21). In the case of chloroplast native occurring proteins like valine and proline whose biosynthetic pathways are compartmentalized in chloroplasts, higher levels of protein can be expected to be produced.

Example 11

Genetic Engineering for Glyphosate Tolerance via the Nuclear and Chloroplast Genomes

Chloroplast Integration and Expression Vectors with EPSPS in Tobacco. The EPSPS coding sequence has been recently integrated into the tobacco chloroplast genome (Figure 1). The chloroplast vector pZS-RD-EPSPS (Figure 2A) contains the 16S rRNA promoter (Prrn) driving the *aadA and* EPSPS genes with the *psbA* 3' region from the tobacco chloroplast genome. This construct integrates the EPSPS and *aadA* genes into the spacer region between the *rbcL* and *orf*512 genes of the tobacco chloroplast genome. Figure 1, at the "TV" arrow.

Test for Glyphosate Resistance. Gene Expression in *E. coli*. Because of the high similarity in the transcription and translation systems between *E.coli* and chloroplasts (Brixey et al., 1997), chloroplast expression vectors were first tested in *E.coli* for resistance in this case to glyphosate before proceeding with transformation of higher plants. The higher growth rate of *E. coli* containing the tobacco vector compared to the control containing pZS197 (similar to pZS-RD-EPSPS but lacked the EPSPS gene) in the presence of 10 mM and 40 mM glyphosate (Figure 23A) indicates glyphosate tolerance of *E.coli* expressing the EPSPS gene. Another growth curve (Figure 23B), confirms the expression of EPSPS via the universal vector in *E.coli*. Thus, glyphosate tolerance of *E. coli* is due to the expression of the EPSPS gene, present in both the tobacco and universal vectors.

CHARACTERIZATION OF TOBACCO TRANSGENIC PLANTS

Integration of the Gene. Fully expanded green leaves of *Nicotiana tabaccum* var. Petit Havana were bombarded with the tobacco and the universal chloroplast vectors. Two days after bombardment, leaf explants were transferred to selection lethal medium containing spectinomycin (500 μg/ml). Transgenic plants were obtained within 3-5 months after bombardment. Typically, out of 16 bombarded leaves, 10 independently transformed shoots were identified.

PCR analysis was performed with DNA isolated from the first or second generation shoots and also from the mature transgenic plants. Primers were used to confirm integration of the *aadA* gene into the plant genome from the tobacco as well as universal vectors. Lack of a product would indicate spontaneous mutants, capable of growing on spectinomycin without the *aadA* gene. The expected PCR product (887bp) was obtained from six lines (Figures 24A-B, lanes 1-6) transformed with the tobacco vector. A PCR product of 1.57Kb was detected in four lines (Figures 24A-B, lanes 1-4) transformed with the universal vector. Under the selection conditions used, four mutants were detected out of ten lines transformed with the tobacco vector. On the other hand, all the transgenic lines transformed with the universal vector showed integration of the *aadA* gene.

PCR Chloroplast Integration. Primers were also designed to specifically identify integration into the plastid genome. The strategy here was to land one primer on the native chloroplast genome, adjacent to the point of integration of the vector, while landing the other on the *aadA* gene. A primer was designed to land immediately outside the *rbcL* gene in the tobacco vector (2.08Kb PCR product). For the universal vector, the primer on the native chloroplast genome landed in the 16s rRNA gene (1.60Kb PCR product). The expected products were observed for the transgenic lines obtained using the tobacco vector (Figures 24A-B, lanes 2-7) as well as the universal vector (Figures 24A-B, lanes 1-4). Unbombarded plants

(controls) did not yield any PCR products, as expected (Figures 24A, lanes 1 and 9; 24B, lanes 5 and 11). Thus, all transgenic plants examined turned out to be chloroplast and not nuclear transformants perhaps due to the requirement of higher levels of aminoglycoside adenyl transferase (AADA) in transgenic plants under stringent selection conditions. Low levels of AADA present in the codicil of nuclear transgenic plants, should have eliminated nuclear transformants. The results of PCR analysis are conclusive and provide definitive evidence for chloroplast integration of foreign genes using both the tobacco and universal vectors.

Southern Analysis. The integration of the *aroA* gene into the chloroplast was also confirmed by Southern analysis. In addition, the high level of resistance to glyphosate observed (Figure 20A) was confirmed by determination of the copy number of the foreign gene in the transgenic plants. The probe, to determine integration of the foreign gene into the chloroplast genome, comprised 654bp fragment of the EPSPS gene, random primer labeled with P^{32}. The total DNA comprising both organellar and genomic was digested with *Eco*RI. The presence of an *Eco*RI site 200bp upstream from the integration site, in the chloroplast genome, was used to confirm integration of the EPSPS gene into the chloroplast genome in transgenic plants. The probe hybridized to the native EPSPS gene, present in the nuclear genome, is seen as 4.5Kb fragment. In addition, the probe hybridized to the digested chloroplast genomes of the transgenic tobacco plants, generating the 3.5kb and 4.35Kb fragments in Figure 25A, lanes 2,3 and 4. The probe did not hybridize to the digested chloroplast genome of the untransformed control plant (Figure 25A, lane 1) since the foreign gene is not present in the chloroplast genome of tobacco. This clearly establishes the integration, of the EPSPS gene, into the chloroplast genome.

Gene Copy Numbers. The copy number of the integrated gene was determined by establishing homoplasmy for the transgenic

chloroplast genome. Tobacco chloroplasts contain 5000-10,000 copies of their genome per cell. (McBride et al, 1995) If only a fraction of the genomes are actually transformed, the copy number, by default, must be less than 10,000. By establishing that in the transgenics the EPSPS transformed genome is the only one present, one could establish that the copy number is 5000-10,000 per cell. This was shown by digesting the total DNA with *Eco*RI and probing, with the flanking sequences that enable homologous recombination into the chloroplast genome. The probe comprised a 2.9 Kb fragment of the *rbcL-orf* 512 sequences. A chloroplast genome transformed with the EPSPS gene, incorporates an *Eco*RI site between the *rbcL-orf* 512 region of the chloroplast genome, thereby generating an extra fragment when digested with this enzyme (Figure 25C). Southern hybridization analysis revealed a 4.43 Kb fragment in Figure 25B, lane 1 for the untransformed control. In lanes 2,3 and 4, two fragments (4.35 Kb and 3 Kb) were generated due to the incorporation of the EPSPS gene cassette between the *rbcL* and *orf*512 regions (Figure 25C provides a schematic diagram with the dotted lines in gray signifying the point of integration of the foreign DNA). The 4.43 Kb fragment present in the control is absent in the transgenics. This proves that only the transgenic chloroplast genome is present in the cell and there is no native, untransformed, chloroplast genome, without the EPSPS gene present. This establishes the homoplasmic nature of the transformants, simultaneously providing an estimate of 5000~10,000 copies of the foreign EPSPS gene per cell. This would then explain the high levels of tolerance of glyphosate that was observed in the transgenic tobacco plants (Figure 20A).

Progeny. Seeds collected from self-pollinated transgenic plants were germinated in the presence of spectinomycin (500 μg/ml). All seeds germinated, remained green and grew normally (Figure 26B). Uniform spectinomycin resistance

indicated that the *aadA* gene was transmitted to all progeny. Lack of variegation suggested homoplasmy because a heteroplasmic condition would have given rise to variegated progeny on spectinomycin (Svab et al., 1990; Svab and Maliga, 1993). The lack of variation in chlorophyll pigmentation among the progeny also underscores the absence of position effect, an artifact of nuclear transformation. All control seedlings are bleached, and did not grow in the presence of spectinomycin (Figure 26A).

Tolerance of Glyphosate. Eighteen week old control and transgenic plants were sprayed with equal volumes of glyphosate at different concentrations (0.5 to 5 mM). Control tobacco plants were extremely sensitive to glyphosate; they died within seven days even at 0.5 mM glyphosate (Figure 27B). On the other hand, the chloroplast transgenic plants survived concentrations as high as 5 mM glyphosate (Figure 27A). These results are intriguing, considering the fact that the EPSPS gene from petunia used in these chloroplast vectors has a low level of tolerance to glyphosate and also contains the transit peptide for targeting into chloroplasts.

This is the first report of a eukaryotic nuclear gene expression within the prokaryotic chloroplast compartment. It is well known that the codon preference is significantly different between the prokaryotic chloroplast compartment and the eukaryotic nuclear compartment. Ideally, a mutant *aroA* gene (which does not bind glyphosate) from a prokaryotic system should be expressed in the chloroplast compartment. Such genes are now available and exhibit a thousand fold higher level of resistance to glyphosate than the petunia gene used in this work. In light of these observations, it is possible that integration of prokaryotic herbicide resistance genes into the chloroplast genome as performed herein can result in incredibly high levels of resistance to herbicides while still maintaining the efficacy of biological containment, i.e., avoid dissemination by pollen.

Example 12

Tolerance of Corn to Glyphosate. A universal chloroplast vector using corn chloroplast DNA is constructed as follows. First, vector pSBL-Ct-bor (Figure 5C) is constructed as follows: Corn chloroplast DNA subclone containing one of the inverted repeat regions is constructed with bacterial plasmid pUC19. Second, a smaller subclone containing only the rRNA operon is constructed from the first subclone and the fragment present in the second subclone containing the *trnA* and *trnI* genes and spacer regions representing the universal border are subcloned into a pUC19 plasmid at the *Pvu*II site. The resultant plasmid is designated pSBL-Ct-bor. Within plasmid pSBL-Ct-bor, a selectable maker gene cassette containing a chloroplast 16S rRNA promoter, the *aadA* gene (encoding aminoglycoside 3'adenyl transferase conferring resistance for streptomycin/spectinomycin) and a 3' untranslated region of the chloroplast *psbA* gene is inserted to construct vector pSBL-CORN. The selectable maker gene cassette is inserted between the *trnI* and *trnA* genes in the spacer region, in the direction of the 16S rDNA transcription.

The vector pSBL-CORN-aroA, which contains a mutant *aroA* gene from *Salmonella typhimurium* (Stalker et al. 1985; Comai et al. 1983)that encodes the enzyme EPSPS synthase, is constructed by inserting the mutant aroA gene into the pSBL-CORN vector. Transgenic corn plants expressing the mutant aroA gene are resistant to glyphosate treatment like "Roundup™" whereas the untransformed control plants are not.

Example 13

Chloroplast Transformation for Tolerance to Imidizolinones or Sulfonylureas. Plasmid pSBL-CORN is modified by insertion of a DNA fragment containing a mutated form of the acetolactate synthase gene of *Saccharomyces cerevisiae* (Falco and Dumas. 1985; Yadav et al. 1986) to generate plasmid pSBL-CORN-ASL1. This gene encodes an acetolactate synthase that is not

inhibited by imidizolinones or sulfonylureas and confers tolerance to herbicides containing sulfometuron methyl and herbicides containing Imazapyr. Transformed tobacco plants expressing the mutant acetolactate synthase gene are resistant to imidizolinone and sulfonylurea herbicide sprays.

The vector pSBL-CORN-ALS2 is a derivative of pSBL-CORN that contains mutated copies of the tobacco suRA and suRB genes (Chaleff and Ray. 1984). These genes encode the tobacco acetolactate synthase polypeptide. The plasmid pSBL-CORN-ALS2 is constructed by ligating the suRA and suRB genes, isolated from tobacco genomic DNA into the pSBL-CORN vector. The resulting vector confers resistance to imidazolinone and sufonylurea herbicides.

Example 14

Chloroplast Transformation for Tolerance to Photosystem II Inhibitors. Photosystem (PS) II herbicide resistance occurs from mutations within the *psbA* gene, which enclosed the Q_B protein and is highly conserved among many plants. Resistant plants possess mutations that alter amino acids at specific positions within the Q_B proteins, e.g., residues 219, 251, 255, 264, and 275 (Hirschberg and McIntosh. 1993; Galloway and Mets. 1984; Gloden and Haselkorn. 1985; Erickson et al. 1984; Johanningmeier et al. 1987). Genes possessing these mutations can thus be utilized to confer resistance to herbicides that function by inhibiting electron transport carried out by the PS II system.

Examples of these herbicides include dichlorophenyldimethylurea (DCMU), atrazine, metribuzine, lenacil, phenmedipham, loxynil and dinoseb.

The mutant *psbA* gene containing a serine to glycine mutation at residue 264 is isolated from genomic DNA of Chlamydomonas using the appropriate restriction endonucleases. The resulting fragment can be ligated into the universal chloroplast expression vector pSBL-ctV2 and introduced into *E. coli* XL1Blue. Purified plasmid from this *E. coli* strain is

utilized to transform plants. Daniell (1997). Incorporation of the mutant *psbA* genes into the chloroplast genome and selection of the appropriate transformants are carried out as previously described. Transformed plants producing the mutated *psbA* protein containing the serine to glycine substitution are resistant to Atrazine™ whereas control plants are not.

The mutant *psbA* gene containing a valine to isoleucine mutation at residue 219 is isolated from genomic DNA of *Chlamydomonas* using the appropriate restriction endonucleases. A universal vector is constructed as described above. Transgenic plants like corn expressing *psbA* containing the valine to isoleucine mutation at residue 219 are expected to be resistant to DCMU sprays.

Example 15

Tolerance to Auxin Analogs. 2,4-D. The universal chloroplast expression vector *psbL*-ctV2 can be cleaved with *XbaI* and ligated with a DNA fragment containing a gene encoding monooxygenase. The resulting construct can be transformed into chloroplasts to generate transgenic plants that contain multiple copies of the monooxygenase gene. The resulting plants expressing high levels of monooxygenase and are expected to be tolerant to 2,4-D.

Example 16

Chloroplast Transformation for Insect Resistance. Tobacco plants can be transformed with universal vector pSBL-CtVHBt (Figure 8A) which contain the *cry*IIA gene and expresses the CryIIA protoxin, thereby conferring resistance to insects pests like of the family Pyralidoe, such as the tobacco hornworm. Even insects which have developed a resistance or are less susceptible to Bt toxin are killed by the Bt toxin expressed by the gene in the chloroplast vector described herein.

Vector pSBL-CtVHBt is constructed by cleaving pSBL-CtVH with *Sma*I and ligating the product with the *cry*IIA gene encoding the CryIIA protein. The product contained the *cry*IIA gene in the correct transcriptional orientation (Figure 8A)

Integration of the *cry*IIA gene into the tobacco chloroplast genome has been confirmed by PCR analysis. Copy number of *cry*IIA per cell was estimated to be 10,000 by performing Southern blots. The CryIIA protein was estimated to be between 5 and 10% of total cellular protein by performing Western blots. This is the highest level of CryIIA protein ever reported in Bt transgenic plants. Excised leaf bioassays were performed using non-transformed "Petit Havana" and *cry*IIA-transformed tobacco. Five to ten larvae were placed on each leaf and evaluated for mortality and leaf damage after 3-5 days. Using susceptible *H. virescens* (YDK), all larvae died within 3 days on *cry*II-A-transformed tobacco whereas there was no mortality and essentially 100% defoliation on the non-transformed "Petit Havana". Similar results were obtained using CryIAc-resistant (YHD2, 40-50,000 fold resistant) and CryII-A-resistant (CXC, 2000 fold resistant) *H. virescens*. In addition 100% mortality was observed against *Helicoverpa zea* (cotton bollworm), and *Spodoptera exigua* (beet armyworm), neither of which has been previously shown to be killed by any Cry protein.

Figure 28A and 28B show a bioassay of (A) Control untransformed plant and (B) transgenic plant. Insects tested, demonstrated 100% mortality: Tobacco budworm susceptible to CryI, cotton bollworm and beet armyworm resistant to CryI and CryII.

Figure 29 shows total protein isolated by Western blot analysis from Control (lane C) and transgenic plants (lane B). Lanes D-H represent different concentrations of purified CryIIA protein (1-20%). Lane A shows protein standards.

Other controllable insects are described earlier in the description of the invention.

As will be apparent to those skilled in the art, in light of the foregoing description, many modifications, alterations, and substitutions are possible in the practice of the invention without departing from the spirit or scope thereof. It is intended that such modifications, alterations, and substitutions be included in the scope of the claims.

All references cited in this text are expressly incorporated herein by reference.

REFERENCES

Arntzen Ph.D., Charles J. (1997) *Public Health Reports* **112**: 190-197.

Brixey, P.J., Guda, H. and Daniell, H. (1997) *Biotechnol. Lett.* **19**, 395-399.

Carlson, P.S. (1973) *Proc. Natl. Acad. Sci. USA* **70**:598-602.

Chaleff and Ray. (1984) *Science* **223**:1148.

Comai, L., Faciotti, D., Hiatt, W., Thomson, G., Rose, R., and Stalker, D. (1983) *Science* **221**:370.

Daniell, H., Guda, C., McPherson, D.T., Xu, J., Zhang, X. and Urry, D.W. (1997) *Meth. Mol. Biol.*, **63**:359-371.

Daniell, H., and Guda, C. (1997) *Chemistry and Industry*, pages 555-558.

Daniell, H., Krishnan, M. and McFadden, B.A. (1991) *Plant Cell Rep.* **9**: 615-619.

Daniell, H., and McFadden. (1987) *Proc. Nat. Acad. Sci.* (USA) **84**: 6349-6353.

Daniell, H., Vivekananda, J., Neilson, B.,Ye, G.N., Tewari, K.K., and Sanford, J.C. (1990) *Proc. Nat. Acad. Sci.* (USA) **87**: 88-92.

Daniell, H. Porobo Dessai, A., Prakash, C.S. and Moar, W.J. (1994) *NATO Asi Series*. Ed., J.H. Cherry. **H86**: 598-604.

Darkocsik, C., Donovan, W.P. and Jany, C.S. (1990) *Mol. Microbiol.* **4**: 2087-2094.

Daniell, H., (1995) *Inform.* **6**: 1365-1370.

Daniell, H., Ramanujan, P., Krishnan, M., Gnanam, A. and Rebeiz, C.A. (1983) *Biochem. Biophys. Res. Comun* **111**:740-749.

Daniell, H. and Rebeiz, C.A. (1982) *Biochem. Biophys. Res. Comun.* **106**:466-471.

Daniell, H. (1993) *Methods in Enzymology*. **217**:536-556.

Daniell, H. (1997a) *Meth. Mol. Biol.* **62**: 453-488.

DeBlock, M., Botterman, J., Vandewiele, M., Docky, J., Thuen, C., Gossele, V., Movva, N.R., Thomson, C., Van Montagu, M., and Leemans, J. (1987) *EMBO J.* **6**: 2513-2518.

Erickson et al.(1984) *Proc. Nat. Acad. Sci.* (USA) **81**:3617.

Falco and Dumas. (1985) *Genetics* **109**: 21.

Gabard, J.M., Charest, P.J., Iyer, V.N. and Miki, B.L.(1989) *Plant Phys.* **91**:574-580.

Galloway and Mets. (1984) *Plant Physiol.* **74**: 469.

Gloden and Haselkorn. (1985) *Science* **229**: 1104.

Guda, C., Zhang, X., McPherson, D.T., Xu, J., Cherry, J., Urry, D.W. and Daniell, H. (1995) *Biotechnol. Lett.* **17**: 745-750.

Hirschberg and McIntosh. (1993) *Science* **222**: 1346.

Johanningmeier et al. (1987) *FEBS Lett* **211**: 221.

Kanyand et al. (1994) *Plant Cell Reports* **14**: 1-5.

Kin Ying et al. (1996) *The Plant Journal* **10**: 737-743.

King,J. (1996) *Science* **274**: 180-181.

Laemmli, U.K. (1970) *Nature* **227**: 680-685.

Langevin, S.A., Clay, K. and Grace, J.B. (1990) *Evolution* **44**: 1000-1008.

Lewellyn and Fitt (1996) *Molecular Breeding* **2**: 157-166.

Lu, Z., Kunnimalaiyaan, M. and Nielsen, B.L. (1996) *Plant Mol. Biol.* **32**: 693-706.

Lyons, P.C., May, G.D., Mason, H.S., and Arntzen, C.J., (1996) *Pharmaceutical News* **3**: 7-12.

Maier, R.M., Neckerman, K., Igloi, G.L. and Kössel, H. (1995) *J. Mol. Biol.* **251**: 614-628.

May, G.D., Mason, H.S., Lyons, P.C. (1996) *American Chemical Society*, pp. 194-204.

Miele, L. (1997) *Elsevier Trends Journals*, Vol. 15.

Mikkelson, T.R., Anderson, B. and Jörgenson, R.B. (1996) *Nature* **380**: 31.

Miki, B.I., Labbe, H., Hatori, J., Ouellet, T., gabard, J., Sunohara, G., Charest, P.J. and Iyer, V.N. (1990) *Theoretical Applied Genetics* **80**: 449-458.

McBride, K.E., Svab, Z., Schaaf, D.J., Hogan, P.S., Stalker, D.M. and Maliga, P. (1995) *Bio/Technology* **13**: 362-365.

Nielsen, B.L., Lu, Z. and Tewari, K.K. (1993) *Plasmid* **30**: 197-211.

Oard, J.H., Linscombe, S.D., Braveramn, M.P., Jodari, F., Blouin, D.C., Leech, M., Kohli, A., Vain, P. Cooley, J.C. and Christou, P. (1996) *Mol. Breed.* **2**: 359-368.

Penazloza, V., et al. (1995) *Plant Cell Reports* **14**:482-487.

Rudraswamy, V., and Reichert, N.A. (1997) M.S. Thesis. Mississippi State Univ.

Rogers, S.O.,and Bendich, A.J. (1988) in *Plant Molecular Biology Manual,* ed. Gelvin, S.B. and Schilperoot, R.A. (Kulwer Academic Publishers, Dordrecht, Netherlands) pp.A6:1-10.

Sambrook, J., Fritch, E.F. and Maniatis, T. (1989) in *Molecular cloning*. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Sanford, J.C. (1988) *Trends In Biotech.* **6**: 299-302.

Sankula, S., Braverman, M.P., Jordari, F., Linscombe, S.D. and Oard, J.A. (1996) *Weed Technol.* **11**: 70-75.

Schultz, A., Wengenmayer,F. and Goodman, H. (1990) *Critical Review in Plant Sciences* **9**: 1-15.

Shaner and Anderson, P.C., (1985), *Biotechnology in Plant Science,* 287.

Sijmons, P.C., Cekker, B.M.M., Schrammeijer, B., Verwoerd, T.C., van den Elzen, P.J.M., Hoekema, A. (1990) *Biotechnology* **8**: 217-221.

Stalker, et al. (1985) *J. Biol. Chem.* **260**: 4724.

Stummann et al. (1988) *Physiologia Plantarum* **72**:139-146.

Svab, Z. and Maliga, P. (1993) *Proc. Natl. Acad. Sci.* (USA) **90**: 913-917.

Svab, Z., Hajdukiewicz, P. and Maglia, P. (1990) *Proc. Nat. Acad. Sci.* (USA) **87**: 8526-8530.

Umbeck, P.F., et al. (1991) *Econ. Entomology* **84**: 1943-1950.

Urry, D.W., Nicol, A., Gowda, D.C., Hoban, L.D., McKee, A., Williams, T., Olsen, D.B. and Cox, B.A. (1993) in Biotechnological Polymers: Medical, Pharmaceutical and Industrial Applications, ed. Gebelein, C.G. (Technomic Publishing Co., Inc., Atlanta, GA), pp. 82-103.

Widner, W.R. and Whiteley, H.R. (1989) J. *Bacteriol.* **171**: 961-974.

Yadav, et al. (1986) *Proc. Natl. Acad.* (USA) **83**: 4418-4422.

Ye, G.N., Daniell, H. and Sanford, J.C. (1990) *Plant Mol. Biol.* **15**: 809-819.

Yeh, H., Ornstein-Goldstein, N., Indik, Z., Sheppard, P., Anderson, N., Rosenbloom, J., Cicilia, G., Yoon, K. and Rosenbloom, J. (1987) *Collagen Related Res.* **7**: 235-247.

Zhang, X., Guda, C., Datta, R., Dute, R., Urry, D.W. and Daniell, H. (1996) *Plant Cell Reports* **15**: 381-385.

Zhang, X., Guda, C., Datta, R., Dute, R., Urry, D.W., Danell, H.(1995) *Biotechnology Letters* **17**: 1279-1284.

Zhang , X., Urry, D.W. and Daniell, H. (1996) *Plant Cell Rep.* **16**: 174-179.

Current Protocols in Molecular Biology, Asubel et al., eds.,John Wiley and Sons, Inc. (1997) Vol. I, II and III.

Herbicide Resistance Crops, Agricultural, Environmental, Economic, Regulatory and Technical Aspects, Duke, S.O., edt., CRC Press, Inc. (1996).

Herbicide Resistance in Plants, Biology and Biochemistry, Powles, S.B., and Holtum, J.A.M., eds., CRC, Press, Inc. (1994.

CLAIMS

What is claimed is:

1. A universal integration and expression vector competent for stably transforming the chloroplast genome of different plant species which comprises an expression cassette which comprises, operably joined, a heterologous DNA sequence coding for a molecule of interest and control sequences positioned upstream from the 5' and downstream from the 3' ends of the coding sequence to provide expression of the coding sequence in the chloroplast genome of a target plant, and flanking each side of the expression cassette, flanking DNA sequences which are homologous to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plant species, whereby stable integration of the heterologous coding sequence into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome.

2. A universal integration and expression vector competent for stably transforming the chloroplast genome of different plant species which comprises an expression cassette which comprises, operably joined, a heterologous DNA sequence coding for a peptide of interest and control sequences positioned upstream from the 5' and downstream from the 3' ends of the coding sequence to provide expression of the coding sequence in the chloroplast genome of a target plant, and flanking each side of the expression cassette, flanking DNA sequences which are homologous to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plant species, whereby stable integration of the heterologous coding sequence into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome.

3. The vector of claim 2 which comprises a heterologous nucleotide sequence coding for a selectable phenotype.

4. The vector of claim 3 wherein the flanking sequences comprise, each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, whereby double homologous recombination with the conserved spacer 2 region in the target chloroplast genome is facilitated.

5. The vector of claim 4 wherein the flanking sequences comprise each one in addition to the portion of the spacer region, part or all of the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes, respectively.

6. The vector of claim 5 wherein the flanking sequences comprise each one in addition to the portion of the spacer region, part or all of the 16S and/or 23S rRNA gene sequences.

7. The vector of claim 4 wherein the spacer region is located in an inverted repeat of the chloroplast genome.

8. The vector of claim 5 wherein the spacer region is located in an inverted repeat of the chloroplast genome.

9. The vector of claim 6 wherein the spacer region is located in an inverted repeat of the chloroplast genome.

10. The vector of claim 4 which comprises in the spacer region, a chloroplast origin of replication, whereby homoplasmy with the genome is promoted.

11. The vector of claim 5 which comprises in the spacer region, a chloroplast origin of replication, whereby homoplasmy with the genome is promoted.

12. The vector of claim 6 which comprises in the spacer region, a chloroplast origin of replication, whereby homoplasmy with the genome is promoted.

13. The vector of claim 7 which comprises in the spacer region, a chloroplast origin of replication, whereby homoplasmy with the genome is promoted.

14. The vector of claim 8 which comprises in the spacer region, a chloroplast origin of replication, whereby homoplasmy with the genome is promoted.

15. The vector of claim 9 which comprises in the spacer region, a chloroplast origin of replication, whereby homoplasmy with the genome is promoted.

16. The vector of claim 4 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

17. The vector of claim 5 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

18. The vector of claim 6 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

19. The vector of claim 7 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

20. The vector of claim 8 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

21. The vector of claim 9 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

22. The vector of claim 10 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

23. The vector of claim 11 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

24. The vector of claim 12 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

25. The vector of claim 13 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

26. The vector of claim 14 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

27. The vector of claim 15 wherein the DNA of the flanking sequences originate from a plant species other than that of the target plant.

28. The vector of claim 13 wherein the DNA of the flanking sequences originate from a plant other than the target plant from the same plant species as the target plant species.

29. The vector of claim 14 wherein the DNA of the flanking sequences originate from a plant other than the target plant from the same plant species as the target plant species.

30. The vector of claim 15 wherein the DNA of the flanking sequences originate from a plant other than the target plant from the same plant species as the target plant species.

31. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 4, or the progeny thereof.

32. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 5, or the progeny thereof.

33. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 6, or the progeny thereof.

34. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 7, or the progeny thereof.

35. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 8, or the progeny thereof.

36. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 9, or the progeny thereof.

37. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 10, or the progeny thereof.

38. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 11, or the progeny thereof.

39. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 12, or the progeny thereof.

40. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 25, or the progeny thereof.

41. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 26, or the progeny thereof.

42. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 27, or the progeny thereof.

43. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 28, or the progeny thereof.

44. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 29, or the progeny thereof.

45. A stably transformed plant which comprises chloroplast stably transformed with the vector of claim 30, or the progeny thereof.

46. The stably transformed plant of claim 31 which is a solanaceous plant.

47. The stably transformed plant of claim 32 which is a solanaceous plant.

48. The stably transformed plant of claim 33 which is a solanaceous plant.

49. The stably transformed plant of claim 40 which is a solanaceous plant.

50. The stably transformed plant of claim 41 which is a solanaceous plant.

51. The stably transformed plant of claim 42 which is a solanaceous plant.

52. The stably transformed plant of claim 31, which is monocotyledonous.

53. The stably transformed plant of claim 32, which is monocotyledonous.

54. The stably transformed plant of claim 33, which is monocotyledonous.

55. The stably transformed plant of claim 31 which is dicotyledonous.

56. The stably transformed plant of claim 32 which is dicotyledonous.

57. The stably transformed plant of claim 33 which is dicotyledonous.

58. The stably transformed plant of claim 34 which is monocotyledonous.

59. The stably transformed plant of claim 35 which is monocotyledonous.

60. The stably transformed plant of claim 36 which is monocotyledonous.

61. The stably transformed plant of claim 34 which is dicotyledonous.

62. The stably transformed plant of claim 35 which is dicotyledonous.

63. The stably transformed plant of claim 36 which is dicotyledonous.

64. The stably transformed plant of claim 37 which is monocotyledonous.

65. The stably transformed plant of claim 38 which is monocotyledonous.

66. The stably transformed plant of claim 39 which is monocotyledonous.

67. The stably transformed plant of claim 37 which is dicotyledonous.

68. The stably transformed plant of claim 38 which is dicotyledonous.

69. The stably transformed plant of claim 39 which is dicotyledonous.

70. The stably transformed plant of claim 40 which is monocotyledonous.

71. The stably transformed plant of claim 41 which is monocotyledonous.

72. The stably transformed plant of claim 42 which is monocotyledonous.

73. The stably transformed plant of claim 64 which is maize, rice, grass, rye, barley, oat or wheat.

74. The stably transformed plant of claim 65 which is maize, rice, grass, rye, barley, oat or wheat.

75. The stably transformed plant of claim 66 which is maize, rice, grass, rye, barley, oat or wheat.

76. The stably transformed plant of claim 67 which is soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

77. The stably transformed plant of claim 68 which is soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

78. The stably transformed plant of claim 69 which is soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

79. The stably transformed plant of claim 70 which is maize, rice, grass, rye, barley, oat or wheat.

80. The stably transformed plant of claim 71 which is maize, rice, grass, rye, barley, oat or wheat.

81. The stably transformed plant of claim 72 which is maize, rice, grass, rye, barley, oat or wheat.

82. The stably transformed plant of claim 67 which is soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

83. The stably transformed plant of claim 68 which is soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

84. The stably transformed plant of claim 69 which is soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

85. A process for stably transforming a target plant species which comprises introducing an integration and expression universal vector into the chloroplast genome of the target plant and allowing the transformed plant to grow, the vector being competent to stably transform the chloroplast of different plant species and comprising an expression cassette which comprises, operably joined, a heterologous DNA sequence coding for a molecule of interest and control sequences positioned upstream from the 5' and downstream from the 3' ends of the coding sequence to provide expression of the coding sequence in the chloroplast of the target plant, a heterologous sequence coding for a selectable phenotype, and flanking each side of the expression cassette, flanking DNA sequences which comprise each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, which are homologous to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plants species, whereby stable integration of the heterologous coding sequence into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome.

86. A process for stably transforming a target plant species which comprises introducing an integration and expression universal vector into the chloroplast genome of the target plant species and allowing the transformed plant to grow, the vector being competent to stably transform the chloroplast of different plants species and comprising an expression cassette which comprises, operably joined, a heterologous DNA sequence coding for a peptide of interest, and control sequences positioned upstream from the 5' and downstream from the 3' ends of the coding sequence to provide expression of the coding sequence in the chloroplast of the target plant, a heterologous nucleotide sequence coding for a selectable phenotype, and flanking each side of the expression cassette, flanking DNA sequences which comprise each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, which are homologous to a spacer sequence of the target chloroplast genome, which sequence is

conserved in the chloroplast genome of different plant species, whereby stable integration of the heterologous coding sequence into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome.

87. The process of claim 86 wherein the transformed plant is heteroplasmic.

88. The process of claim 86 wherein the transformed plant is a homoplasmic plant.

89. The process of claim 88 wherein the transformed plant is a first generation plant.

90. The process of claim 86 wherein the target plant is a solanaceous plant.

91. The process of claim 86 wherein the transformed plant is monocotyledonous.

92. The process of claim 86 wherein the transformed plant is dicotyledonous.

93. The process of claim 91 wherein the transformed plant is one of the following monocotyledonous plants: maize, rice, grass, rye, barley, oat or wheat.

94. The process of claim 92 wherein the transformed plant is one of the following dicotyledonous plants: soybean, peanut, grape, sweet potato, pea, canola, tobacco, tomato or cotton.

95. The process of claim 86 which comprises isolating the peptide of interest.

96. The process of claim 86 wherein the peptide of interest is a polypeptide.

97. The process of claim 96 wherein the polypeptide is a synthetic protein-base polymer (PBP).

98. The process of claim 97 wherein the PBP has repeating pentamer sequences $(GVGVP)_n$ wherein "n" is an integer of 1 to 250, "G" is glycine, "V" is valine and "P" is proline.

99. The process of claim 98 wherein "n" is 121.

100. The process of claim 96 wherein the expressed polypeptide of interest is insulin.

101. The process of claim 100 which comprises isolating the insulin.

102. The process of claim 100 wherein the insulin is in the form of pro-insulin.

103. The process of claim 102 wherein the pro-insulin is fused to a PBP.

104. The process of claim 100 wherein the transformed plant is tobacco.

105. The process of claim 96 wherein the polypeptide of interest is human serum albumin (HSA).

106. The process of claim 105 wherein the transformed plant is tobacco.

107. The vector of claim 4 wherein the peptide of interest is a biologically active molecule.

108. The vector of claim 107 wherein the peptide is a polypeptide which is a synthetic protein-based polymer (PBP).

109. The vector of claim 108 wherein the PBP has repeating pentamer sequences $(GVGVP)_n$ wherein "n" is an integer of 1 to 250, "G" is glycine, "V" is valine and "P" is proline.

110. The vector of claim 109 wherein the polypeptide is $(GVGVP)_n$ wherein the integer "n" is 121.

111. The vector of claim 108 wherein the PBP is fused to a biologically active molecule.

112. The vector of claim 111 wherein the biologically active molecule is insulin.

113. The vector of claim 112 wherein the insulin is in the form of pro-insulin.

114. The polypeptide obtained by the process of claim 96.

115. The PBP obtained by the process of claim 97.

116. The pro-insulin obtained by the process of claim 102.

117. The HSA obtained by the process of claim 105.

118. A stably transformed plant which comprises a chloroplast genome stably transformed with the vector of claim 107 which comprises a biologically active molecule.

119. The harvested plant of claim 118.

120. The biologically active molecule isolated from the plant of claim 118.

121. The biologically active molecule of claim 120 which is insulin.

122. The stably transformed plant of claim 118 which is tobacco.

123. A stably transformed herbicide resistant target plant species or the progeny thereof, which comprises chloroplast stably transformed with a universal integration and expression vector competent for stable transformation of the chloroplast of different plants species which comprises an expression cassette which comprises a heterologous protein of interest expressed by a heterologous DNA sequence

in the chloroplast genome of the target plant species, a selected phenotype other than tolerance to said herbicide, and flanking each side of the expression cassette, flanking DNA sequences which comprise each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, which are homologous to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plant species, whereby stable integration of the heterologous protein into the chloroplast genome of the target plant was facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome.

124. The herbicide resistant target plant of claim 123 in which the protein of interest is a mutant form of an enzyme which has decreased affinity for the herbicide than does the naturally occurring enzyme.

125. The herbicide resistant target plant of claim 123 wherein the herbicide is glyphosate.

126. The herbicide resistant target plant of claim 125 wherein the enzyme is EPSP synthase.

127. The herbicide resistant target plant of claim 124 wherein the herbicide is selected from at least one of the following types: PSI, PSII, APP, auxin analog, mitotic, tertiary amino methyl phosphoric acids type and ALS inhibiting types.

128. The herbicide resistant target plant of claim 127 wherein the herbicide is the PSI type selected from paraquat and diquat.

129. The herbicide resistant target plant of claim 127 wherein the herbicide is selected from atrazine, dinoseb, lenacil and metribuzine.

130. The herbicide resistant target plant of claim 127 wherein the herbicide is of the APP type selected from cyclohexanedione, haloxyfop, clethodim and phenoxaprop, and the lower alkyl-substituted compound thereof.

131. The herbicide resistant target plant of claim 127 wherein the herbicide is an auxin analog selected from MCPA and 2,4-D.

132. The herbicide resistant target plant of claim 127 wherein the herbicide is a mitotic type herbicide, which is dinitroaniline.

133. The herbicide resistant target plant of claim 124 wherein the herbicide is a tertiary amino methyl phosphoric acid type, which is glyphosate.

134. The herbicide resistant target plant of claim 127 wherein the herbicide is an ALS inhibiting type selected from sulfonylureas and imidazolines.

135. The herbicide resistant target plant of claim 127 wherein the herbicide is selected from bromoxynil, methyl sulfuron, chlorsulfuron, phosphinothricin and imazapyr.

136. The herbicide resistant target plant of claim 123 wherein the expressed phenotype is antibiotic resistance.

137. The herbicide resistant plant of claim 123 wherein the selectable phenotype is encoded by the hygromycin gene that confers herbicide resistance.

138. The herbicide resistant target plant of claim 127 which is maize, rice, grass, rye, barley, oat, wheat, soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

139. The herbicide resistant target plant of claim 138 which is a homoplasmic plant.

140. A process for conferring herbicide resistance to a target plant species which comprises introducing into the plant a universal integration and expression vector competent for stably transformed with a universal integration and expression vector competent for stably transforming the chloroplast of different plants

species which comprises an expression cassette which comprises, operably joined, a heterologous DNA sequence coding for a protein of interest which confers resistance to a herbicide and control sequences positioned upstream from the 5' and downstream from the 3' ends of the coding sequence to provide expression of the coding sequence in the chloroplast of the target plant, a heterologous nucleotide sequence coding for a selectable phenotype other than for tolerance to said herbicide, and flanking each side of the expression cassette, flanking sequences which comprise each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, which are homologous to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plant species, whereby stable integration of the heterologous coding sequence into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome and growing the transformed plant.

141. The process of claim 140 wherein the DNA sequence codes for a mutant form of an enzyme which has decreased affinity for the herbicide than does the naturally occurring enzyme.

142. The process of claim 141 wherein the enzyme is EPSP synthase and the herbicide is glyphosate.

143. The process of claim 142 wherein DNA sequence is the EPSP synthase gene which is a mutant EPSP synthase gene.

144. The process of claim 140 which comprise selecting the viable, transformed target plants from a medium which is lethal to non-transformed plants.

145. The process of claim 144 wherein the viable transformed target plants are homoplasmic plants.

146. The process of claim 144 wherein the viable transformed target plants are heteroplasmic plants.

147. The process of claim 140 wherein the herbicide resistant target plant is tobacco, the nucleotide sequence codes for a selectable phenotype which is lethal to tobacco or for a visual trait that permits to select the transformed tobacco plants from the non-transformed plants.

148. The process of claim 147 wherein the lethal selectable phenotype is hygromycin to which tobacco is not naturally resistant.

149. The process of claim 147 wherein the visual trait is the expression of a color.

150. The process of claim 140 for stably transforming the phenotype of a target plant species, whereby the grown transformed plant expresses the selectable phenotype and another trait in addition to the expression of that phenotype.

151. The process of claim 150 wherein the selected phenotype is conferred by expression of the hygromycin β-phosphotransferase gene and the additional trait that is conferred is resistance to the herbicide glyphosate.

152. A process of determining chloroplast transformation and expression of a target trait on the basis of a target plant species acquisition of resistance to a selected herbicide due to the transformation of the plant species which comprises introducing into the plant a universal integration and expression vector competent for stably transforming the chloroplast of different plant species which comprises an expression cassette which comprises, operably joined, a heterologous DNA sequence coding for the desired target trait and control sequences positioned upstream from the 5' and downstream from the 3' ends of the coding sequence to provide expression of the coding sequence in the chloroplast of a target plant, a heterologous nucleotide sequence coding for a selectable phenotype and flanking each side of the expression cassette, flanking DNA sequences which comprise each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, which are homologous

to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plant species, whereby stable integration of the heterologous coding sequence into the chloroplast genome of the target plant is facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome, exposing the plants into which the vector has been introduced to a lethal concentration of the herbicide and selecting the plants which do not die from exposure thereto, thereby having selected the transformed plants which express the desired target trait.

153. The process of claim 152 wherein the selected herbicide is selected from at least one of the following types: PSI, PSII, APP, auxin analog, mitotic, tertiary amino methyl phosphoric acids and ALS inhibiting types.

154. A stably transformed insect resistant target plant species or the progeny thereof, which comprises chloroplast genome stably transformed with a universal integration and expression vector competent for stable transformation of the chloroplast of different plants species which comprises an expression cassette which comprises a heterologous DNA sequence which expresses a target protein which confers resistance to a target insect in the chloroplast genome of the target plant species, a selected phenotype other than tolerance to the target insect, and flanking each side of the expression cassette, flanking DNA sequences which comprise each one a portion of the intergenic spacer 2 region between the $tRNA^{Ile}$ and the $tRNA^{Ala}$ genes of the chloroplast genome, which are homologous to a spacer sequence of the target chloroplast genome, which sequence is conserved in the chloroplast genome of different plants species, whereby stable integration of the heterologous protein into the chloroplast genome of the target plant was facilitated through homologous recombination of the flanking sequences with the homologous sequences in the target chloroplast genome, wherein the heterologous protein is the CryIIA protein toxin expressed by the *cry*IIA gene, which protein confers resistance to insects.

155. The stably transformed insect-resistant target plant species of claim 154 wherein the insect is susceptible to the Bt toxin.

156. The stably transformed insect-resistant target plant species of claim 154 wherein the insect's lack of susceptibility to the Bt toxin is overcome.

157. The stably transformed insect-resistant target plant species of claim 155 wherein the target plant is tobacco and the insect is tobacco budworm.

158. The stably transformed insect-resistant target plant species of claim 156 wherein the target plant is tobacco, cotton or beet and the insect is a Bt toxin-resistant type cottonboll worm or beet armyworm.

159. The stably transformed insect-resistant target plant of claim 154 wherein the insect is tobacco budworm, cottonboll worm or beet armyworm.

160. A vector selected from the following: pSBL-RD-EPSPS, pSBL-CG-EG121, pSBL-Ct-Border, pSBL-Ct-V1, pSBL-Ct-V2, pSBL-CtV3, pSBL-Ct-VH, pSBL-Ct-VHF, and pSBL-Ct-VHBt.

161. The vector of claim 160 which is a universal integration and expression vector which is pSBL-RD-EPSPS, pSBL-CG-EG121, pSBL-CtV1, pSBL-CtV2, pSBL-CtV3, pSBL-CtVH, pSBL-Ct-VHF, and pSBL-CtVHBt.

162. An isolated DNA sequence which comprises the intergenic DNA sequence between the 16S and 23S rDNA genes of plant chloroplast genome, which intergenic sequence is highly conserved in a multiplicity of different species of plants.

163. The isolated DNA sequence of claim 162 which comprises the spacer 2 region which is the intergenic DNA sequence between the trnI and the trnA genes of plant chloroplast genome.

164. The isolated intergenic DNA sequence of claim 163 which comprises the trnI and the trnA genes.

165. The isolated DNA sequence of claim 163 which comprises one of the inverted repeats of the chloroplast genome of the plant.

166. The isolated DNA sequence of claim 165 which comprises an origin of replication.

167. The isolated DNA sequence of claim 163 which comprises the rRNA operon in the spacer 2 region.

168. The universal integration and expression vector of claim 4 wherein the spacer 2 region comprises the rRNA operon.

169. The universal integration and expression vector of claim 4 wherein the flanking sequences are synthetic.

170. The herbicide resistant target plant species of claim 85 wherein the DNA sequence which codes for the protein of interest is of prokaryotic origin.

171. The universal integration and expression vector of claim 2 which does not include a transposon.

172. The stably transformed target plant species of claim 41 which does not include a transposon.

173. The process for stably transforming a target plant species of claim 86 wherein the universal vector does not include a transposon.

174. The universal integration and expression vector of claim 4 which comprises a promoter functional in chloroplast.

175. The Universal expression and integration vector of claim 4 which comprises a nucleotide sequence encoding a selectable phenotype which allows for

identification and selection of viable transformed plants from non-viable, non-transformed plants.

176. The process of claim 86 which comprises selecting from a medium which is lethal to non-transformed plants, the viable transformed target plant.

177. The stably transformed herbicide tolerant target plant or progeny thereof of claim 123 which allows identification and selection on the basis of viable transformed plants from non-viable, non-transformed plant.

178. The process of claim 140 which comprises selecting from a medium which is lethal to non- transformed plants, the viable transformed target plant.

179. An expression cassette competent for stably transforming chloroplast genome of a target plant which comprises, operably joined, a heterologous DNA sequence coding for a molecule of interest, control sequences to provide expression of the coding sequence in the chloroplast genome of the target plant, a promoter functional in chloroplast, and flanking each side of the cassette, plant DNA sequences to facilitate stable integration of the DNA with the target chloroplast genome by homologous recombination, whereby the DNA is stably integrated therein and is inherited through organelle replication in daughter cells.

180. The expression cassette of claim 175 wherein the heterologous DNA sequence is a synthetic sequence coding for a protein-based polymer (PBP).

181. The expression cassette of claim 176 wherein the PBP has repeating pentamer sequences $(GVGVP)_n$ wherein "n" is an integer of 1 to 250, "G" is glycine, "V" is valine and "P" is proline.

182. The expression cassette of claim 177 wherein the synthetic coding sequence is the synthetic biopolymer gene EG121 and the expressed PBP is the polymer protein $(GVGVP)_{121}$.

183. A stably transformed transcription/ translation active chloroplast genome of a target plant, which is competent for stable integration of a heterologous DNA sequence, which comprises an expression cassette which comprises a heterologous molecule of interest encoded by a heterologous DNA sequence and expressed by control sequences in the chloroplast genome of the target plant, and plant DNA flanking each side of the expression cassette which facilitated stable integration of the DNA into the target chloroplast genome by homologous recombination, which DNA is inherited through organelle replication in daughter cells.

184. The stably transformed chloroplast of claim 179 wherein the molecule of interest is a protein-based polymer (PBP).

185. The stably transformed chloroplast of claim 180 wherein the PBP has repeating pentamer sequences $(GVGVP)_n$ wherein n is an integer of 1 to 250, "G" is glycine, "V" is valine and "P" is proline.

186. The stably transformed chloroplast of claim 181 wherein the synthetic coding sequence is the synthetic biopolymer gene EG121 and the expressed PBP is the polymer protein $(GVGVP)_{121}$.

187. The synthetic biopolymer $(GVGPP)_n$ which is expressed from the stably transformed chloroplast of claim 182, wherein "n", "G", "V" and "P" are defined therein.

188. The synthetic biopolymer of claim 183 which is $(GVGVP)_{121}$.

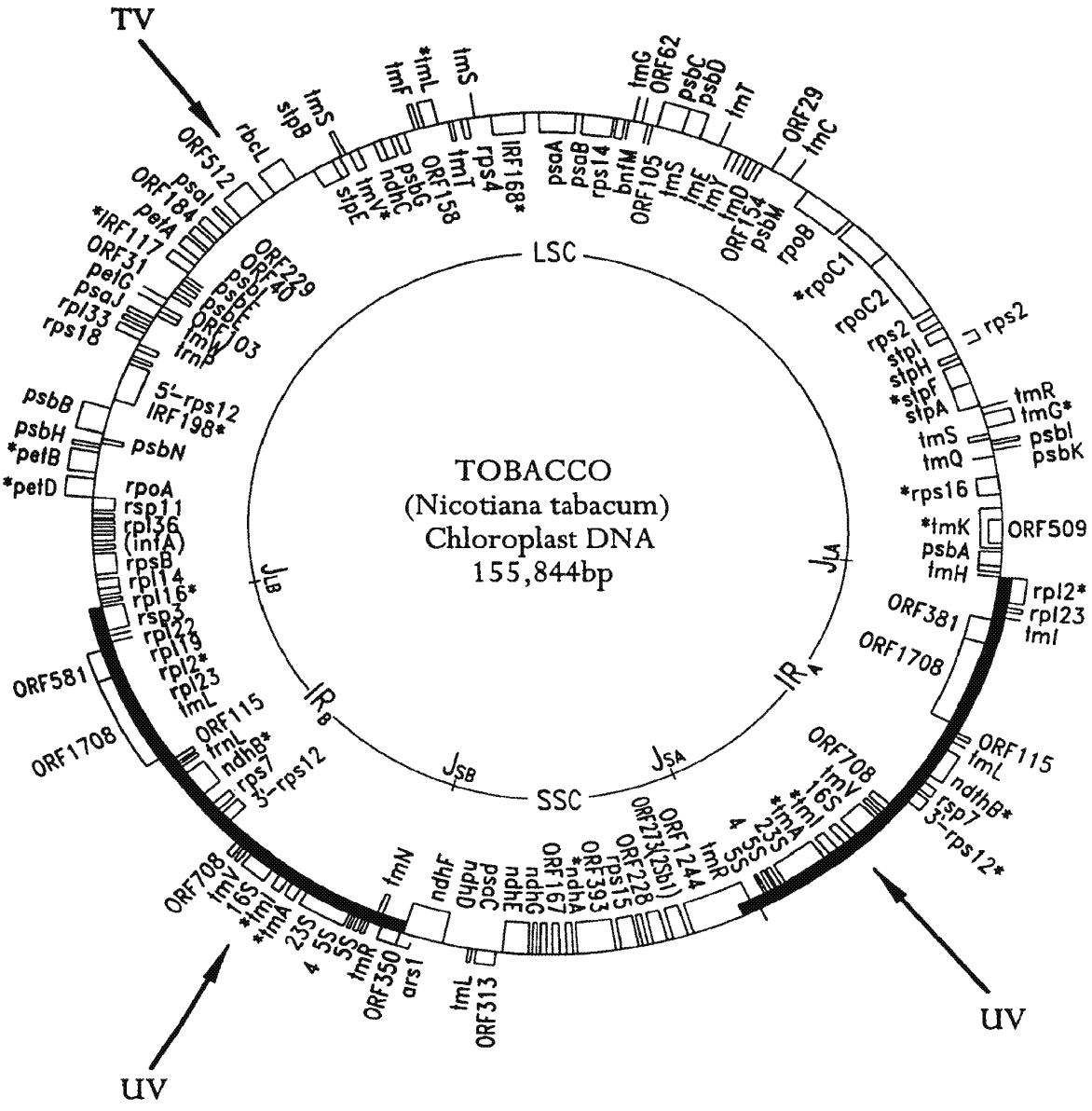

TV
UV
UV
LSC
SSC
IR A
IR B
JLA
JLB
JSA
JSB
TOBACCO
(Nicotiana tabacum)
Chloroplast DNA
155,844bp
rbcL
ORF512
psaI
ORF184
petA
*IRF117
ORF31
petG
psaJ
rpl33
rps18
psbB
psbH
*petB
*petD
rpoA
rsp11
rpl36
(infA)
rpsB
rpl14
rpl16*
rsp3
rpl22
rpl19
rpl2*
rpl23
tmL
ORF581
ORF1708
ORF115
trnL
ndhB*
rps7
5'-rps12
ORF708
tmV
16S
*tmI
*tmA
23S
4.5S
5S
tmR
ORF350
ars1
tmL
ORF313
tmN
ndhF
ndhD
psaC
ndhE
ndhG
ORF167
*ndhA
ORF393
rps15
ORF228
ORF273(2b)
ORF1244
tmR
4.5S
23S
5S
*tmI
*tmA
ORF708
tmV
16S
ORF115
tmL
ndhB*
rsp7
3'-rps12*
ORF1708
ORF381
rpl2*
rpl23
tmI
*rps16
*tmK
psbA
tmH
ORF509
tmR
tmG*
tmS
tmQ
psbI
psbK
rps2
stpI
stpH
*stpF
stpA
rpoC2
*rpoC1
rpoB
psbM
tmE
tmY
tmD
ORF154
ORF105
psbF
psbE
psbL
ORF40
ORF229
ORF103
tmW
tmP
5'-rps12
IRF198*
psbN
tmS
tmF
*tmL
tmT
rps4
IRF168*
psaA
psaB
rps14
bnfM
tmG
ORF62
psbC
psbD
tmT
ORF29
tmC
tmS
stpB
tmV*
stpE
ndhC
psbG
ORF158

FIG. 1

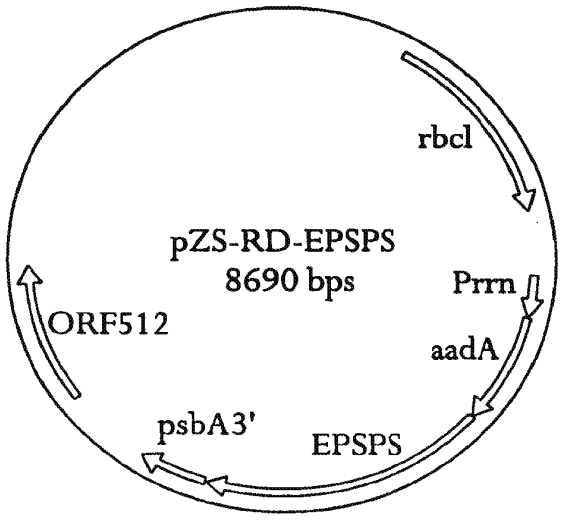


FIG. 2A

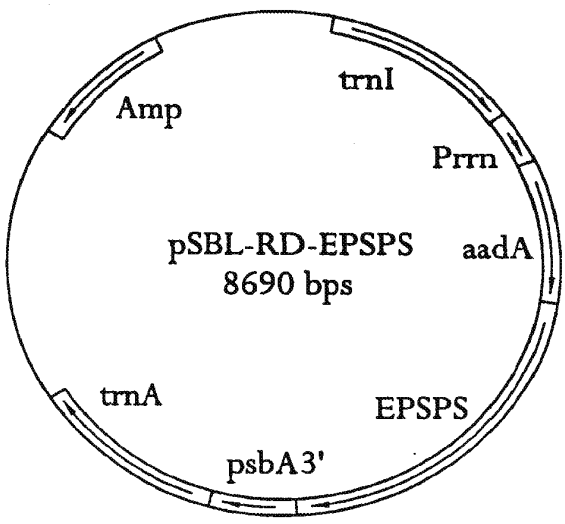


FIG. 2B

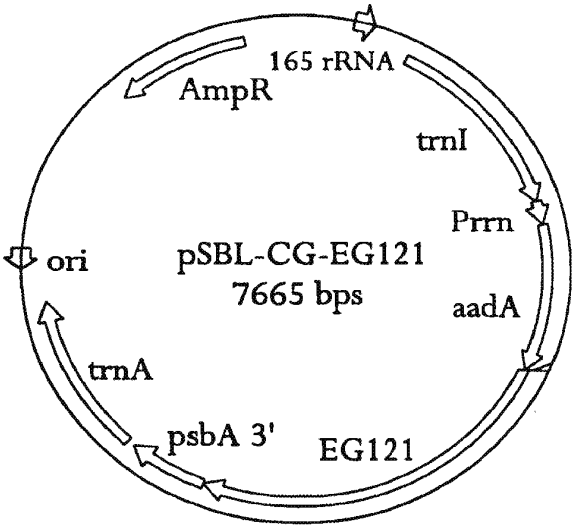


FIG. 3A

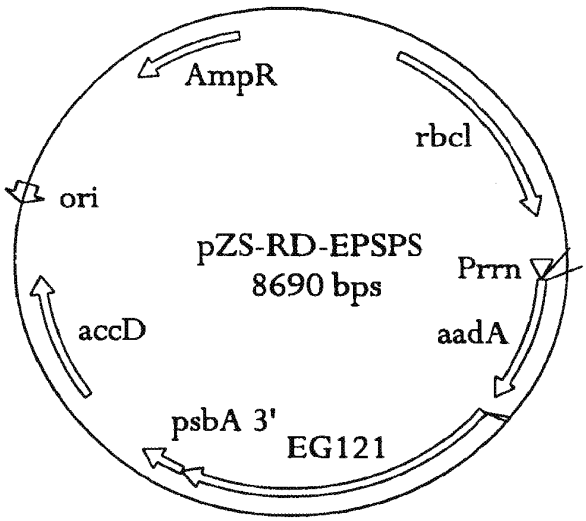


FIG. 3B

Sequence alignment of 16S-23S rDNA spacer region from several crop species

```
  ************************ *********** *   ****** ***** ******* *** *********** *****  **** * *********
S GTACACACCGCCCGTCACACTATGGGAGCTGGCCATGCC-GAAGTCGTTACC-TTAACCGCAAG-AGGGGGATGCCGAAGGCAGGGCTAGTGACTGGAGT
1 T GTACACACCGCCCGTCACACTATGGGAGCTGGCCATGCCCGAAGTCGTTACC-TTAACCGCAAGGAGGGGGATGCCGAAGGCAGGGCTAGTGACTGGAGT
M GTACACACCGCCCGTCACACTATAGGAGCTGGCCAGGTTTGAAGTCATTACCCTTAACCGTAAGGAGGGGGATGCCTAAGGCTAGGCTTGCGACTGGAGT

  ************************************************************************** *      ****** *** ****************
S GAAGTCGTAACAAGGTAGCCGTACTGGAAGGTGCGGCTGGATCACCTCCTTTTCAGGGAGAGCTAATGCTTGT------TGGGTAGTTTAGTTTGACACTGCTTCA
101 T GAAGTCGTAACAAGGTAGCCGTACTGGAAGGTGCGGCTGGATCACCTCCTTTTCAGGGAGAGCTAATGCTTGT------TGGGTATTTTGGTTTGACACTGCTTCA
M GAAGTCGTAACAAGGTAGCCGTACTGGAAGGTGCGGCTGGATCACCTCCTTTTCAGGGAGAGCTAAGTCTTATGCTTATTGGGTATTTTGGTTTGACACTGCTTCA

  * ***     **********  ** ** *******  *** ***** ****************** ***    * ************** *  ************
S CACCC-----AAAAAGAAGCGAGTTATGTCTGAGTCAAATTTGGAGATGGAAGTCTTCTTTCGTTTCTCGATGGTGAAGTAAGACTAAACTCATGAGCTTA
201 T CACCCCCAAAAAAAAGAAGGGAGCTACGTCTGAGTTAAACTTGGAGATGGAAGTCTTCTTTCCTTTCTCGACGGTGAAGTAAGACCA-GCTCATGAGCTTA
M CGCCC-----AAAAAGAAGGCAGCTACGTCTGAGCTAAACTTGGATATGGAAGTCTTCTTTCGTTT----AGGGTGAAGTAAGACCAAGCTCATGAGCTTA

  ********************* **          ******* * *******     *******                   *******  *  ********** ** ***
S TTATCCTAGGTCGGAACAAGTT----------GATAGGAGCTACTTTTTTCA-CCCCCAT--------27bp-------ATGGGGGTGAAAAAAGGAAAGAGAGGGATGGG
301 T TTATCCTAGGTCGGAACAAGTT----------GATAGGACCCCCTTTTTTACGTCCCCATGTTCCCCCCGTGTGGCGACATGGGGC-GAAAAAAGGAAAGAGAGGGATGGG
M TTATCCTAGGTCGGAACAAATTAGTTGATAGTGATAGGATCCCCTTTTTGACGTCCCCATGT-CCCCCCGTGTGGCGGCATGGGGATGTCAAAAGGAAAGGGATGGA----

  **** *** ********* ******** **    ******** *** ********************************************** **********
S GTTTCTCTTGCTTTTGGCATAGCGGGCCCCGGC-GGGAGGCCCGCACGACGGGCTATTAGCTCAGTGGTAGAGCGCGCCCCTGATAATTGCGTCGTTGTG
401 T GTTTCTCTCGCTTTTGGCATAGCGGGCCCCCAGTGGGAGGCTCGCACGACGGGCTATTAGCTCAGTGGTAGAGCGCGCCCCTGATAATTGCGTCGTTGTG
M GTTTTTCTCGCTTTTGGCGTAGCGGCCTCCCTTTGGGAGGC-CGCGCGACGGGCTATTAGCTCAGTGGTAGAGCGCGCCCCTGATAATT-CGTCGTTGTG
```

FIG. 4A

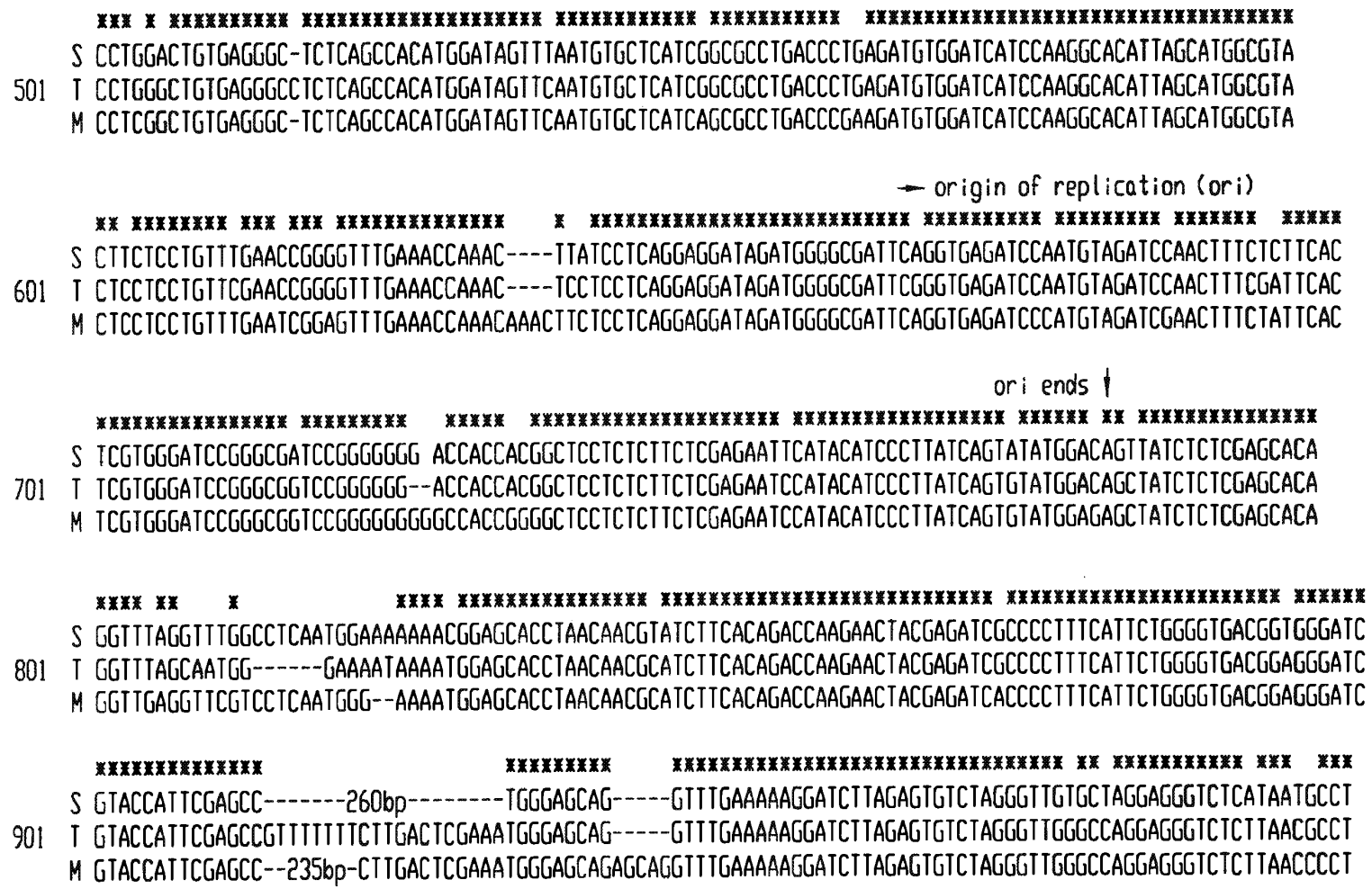

origin of replication (ori)
ori ends
S CCTGGACTGTGAGGGC-TCTCAGCCACATGGATAGTTTAATGTGCTCATCGGCGCCTGACCCTGAGATGTGGATCATCCAAGGCACATTAGCATGGCGTA
501 T CCTGGGCTGTGAGGGCCTCTCAGCCACATGGATAGTTCAATGTGCTCATCGGCGCCTGACCCTGAGATGTGGATCATCCAAGGCACATTAGCATGGCGTA
M CCTCGGCTGTGAGGGC-TCTCAGCCACATGGATAGTTCAATGTGCTCATCAGCGCCTGACCCGAAGATGTGGATCATCCAAGGCACATTAGCATGGCGTA
S CTTCTCCTGTTTGAACCGGGGTTTGAAACCAAAC----TTATCCTCAGGAGGATAGATGGGGCGATTCAGGTGAGATCCAATGTAGATCCAACTTTCTCTTCAC
601 T CTCCTCCTGTTCGAACCGGGGTTTGAAACCAAAC----TCCTCCTCAGGAGGATAGATGGGGCGATTCGGGTGAGATCCAATGTAGATCCAACTTTCGATTCAC
M CTCCTCCTGTTTGAATCGGAGTTTGAAACCAAACAAACTTCTCCTCAGGAGGATAGATGGGGCGATTCAGGTGAGATCCCATGTAGATCGAACTTTCTATTCAC
S TCGTGGGATCCGGGCGATCCGGGGGGG ACCACCACGGCTCCTCTCTTCTCGAGAATTCATACATCCCTTATCAGTATATGGACAGTTATCTCTCGAGCACA
701 T TCGTGGGATCCGGGCGGTCCGGGGGG--ACCACCACGGCTCCTCTCTTCTCGAGAATCCATACATCCCTTATCAGTGTATGGACAGCTATCTCTCGAGCACA
M TCGTGGGATCCGGGCGGTCCGGGGGGGGGCCACCGGGGCTCCTCTCTTCTCGAGAATCCATACATCCCTTATCAGTGTATGGAGAGCTATCTCTCGAGCACA
S GGTTTAGGTTTGGCCTCAATGGAAAAAAACGGAGCACCTAACAACGTATCTTCACAGACCAAGAACTACGAGATCGCCCCTTTCATTCTGGGGTGACGGTGGGATC
801 T GGTTTAGCAATGG------GAAAATAAAATGGAGCACCTAACAACGCATCTTCACAGACCAAGAACTACGAGATCGCCCCTTTCATTCTGGGGTGACGGAGGGATC
M GGTTGAGGTTCGTCCTCAATGGG--AAAATGGAGCACCTAACAACGCATCTTCACAGACCAAGAACTACGAGATCACCCCTTTCATTCTGGGGTGACGGAGGGATC
S GTACCATTCGAGCC-------260bp--------TGGGAGCAG-----GTTTGAAAAAGGATCTTAGAGTGTCTAGGGTTGTGCTAGGAGGGTCTCATAATGCCT
901 T GTACCATTCGAGCCGTTTTTTTCTTGACTCGAAATGGGAGCAG-----GTTTGAAAAAGGATCTTAGAGTGTCTAGGGTTGGGCCAGGAGGGTCTCTTAACGCCT
M GTACCATTCGAGCC--235bp-CTTGACTCGAAATGGGAGCAGAGCAGGTTTGAAAAAGGATCTTAGAGTGTCTAGGGTTGGGCCAGGAGGGTCTCTTAACCCCT

FIG. 4B

```
  ** ******* * ***************** *** ********  ***** *  ** * ** **** *********** ******************  ******
S TCCTTTTTCTTCTCATCGGAGTTATTTCCCAAAGACTTGCCATGGTAAAGAAGA-AGGG-GGAACAAGCACACTTGGAGAGCGCAGTACAACGGATAGTTG
1001 T TCTTTTTTCTTCTCATCGGAGTTATTTCACAAAGACTTGCCAGGGTAAGGAAGA-AGGGGGGAACAAGCACACTTGGAGAGCGCAGTACAACGGAGAGTTG
M TCTTTTTTCTGCCCATCGGAGTTATTTCCCAAGGACTTGCCGTGGTAAGGGGGAGAAGGGGGAAGAAGCACACTTGAAGAGCGCAGTACAACGGGGAGTTG

  ********************** ****************** ***************************** ***** ************************
S TATGCTGCGTTCGGGAAGGATGAATCGCTCCCGAAAAGGAATCTATTGATTCTCTCCCAATTGGTTGGACTGTAGGTGCGATGATTTACTTCACGGGCGA
1101 T TATGCTGCGTTCGGGAAGGATGAATCGCTCCCGAAAAGGAATCTATTGATTCTCTCCCAATTGGTTGGACCGTAGGTGCGATGATTTACTTCACGGGCGA
M TATGCTGCGTTCGGGAAGGATGGATCGCTCCCGAAAAGGAGTCTATTGATTCTCTCCCAATTGGTTGGATCGTAGGGGCGATGATTTACTTCACGGGCGA

                                    site of foreign gene insertion
  ********************* ***************** ** **************** *   *  ***** ********** ******* *********** ***
S GGTCTCTGGTTCAAGTCCAAGATGGCCCAGCTGCGTCAAGGAAAAGAATAGAAAACTGACTTGACTCCTTCATGCATGCTCCACTCGGCTCGGGGGG-ATA
1201 T GGTCTCTGGTTCAAGTCCAGGATGGCCCAGCTGCGCCAGGGAAAAGAATAGAAGAAGCATCTGACTACTTCATGCATGCTCCACTTGGCTCGGGGGG-ATA
M GGTCTCTGGTTCAAGTCCAGGATGGCCCAGCTGCG-CAGGGAAAAGAATAGAAGAAGCATCTGACTCTTTCATGCATACTCCACTTGGCTCGGGGGGGATA

  ************************************************************ ************ ******************************
S TAGCTCAGTTGGTAGAGCTCCGCTCTTGCAATTGGGTCGTTGCGATTACGGGTTGGATGTCTAATTGTCTAGGCGGTAATGATAGTATCTTGTACCTGAA
1301 T TAGCTCAGTTGGTAGAGCTCCGCTCTTGCAATTGGGTCGTTGCGATTACGGGTTGGATGTCTAATTGTCCAGGCGGTAATGATAGTATCTTGTACCTGAA
M TAGCTCAGTTGGTAGAGCTCCGCTCTTGCAATTGGGTCGTTGCGATTACGGGTTGGCTGTCTAATTGTCCAGGCGGTAATGATAGTATCTTGTACCTGAA

  ****************************** ******** ***** **************************** ** ********** ** **********
S CCGGTGGCTCACTTTTTCTAAGTAATGGGAAAGAGGACCGAAACATGCCACTGAAAGACTCTACTGAGACAAA--GACGGGCTGTCAAGAACGTAGAGGAGG
1401 T CCGGTGGCTCACTTTTTCTAAGTAATGGGGAAGAGGACCGAAACGTGCCACTGAAAGACTCTACTGAGACAAA--GATGGGCTGTCAAGAACGTAGAGGAGG
M CCGGTGGCTCACTTTTTCTAAGTAATGGGGAAGAGGACTGAAACATGCCACTGAAAGACTCTACTGAGACAAAAAGATGGGCTGTCAAAAAGGTAGAGGAGG
```

FIG. 4C

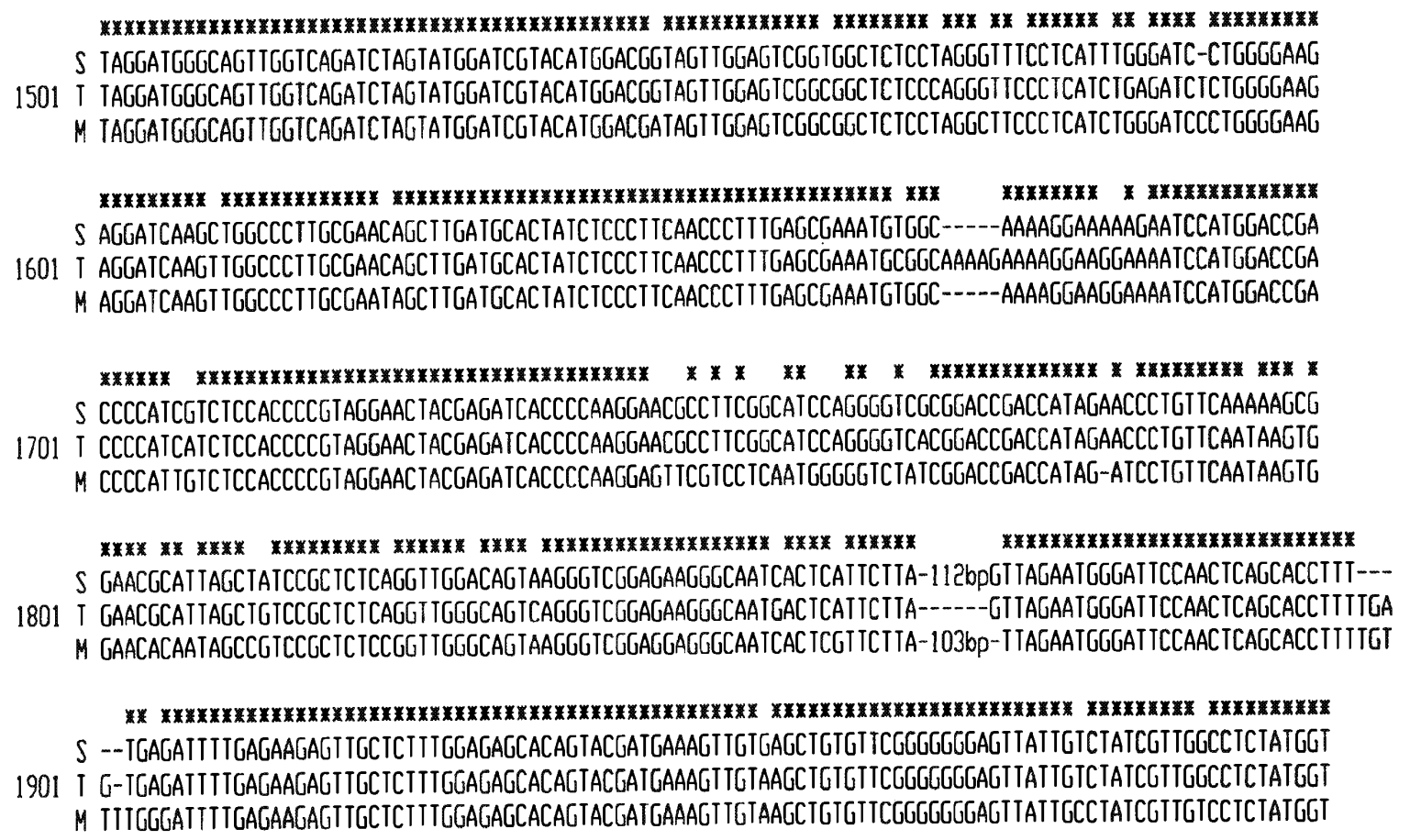

S TAGGATGGGCAGTTGGTCAGATCTAGTATGGATCGTACATGGACGGTAGTTGGAGTCGGTGGCTCTCCTAGGGTTTCCTCATTTGGGATC-CTGGGGAAG
1501 T TAGGATGGGCAGTTGGTCAGATCTAGTATGGATCGTACATGGACGGTAGTTGGAGTCGGCGGCTCTCCCAGGGTTCCCTCATCTGAGATCTCTGGGGAAG
M TAGGATGGGCAGTTGGTCAGATCTAGTATGGATCGTACATGGACGATAGTTGGAGTCGGCGGCTCTCCTAGGCTTCCCTCATCTGGGATCCCTGGGGAAG
S AGGATCAAGCTGGCCCTTGCGAACAGCTTGATGCACTATCTCCCTTCAACCCTTTGAGCGAAATGTGGC-----AAAAGGAAAAAGAATCCATGGACCGA
1601 T AGGATCAAGTTGGCCCTTGCGAACAGCTTGATGCACTATCTCCCTTCAACCCTTTGAGCGAAATGCGGCAAAAGAAAAGGAAGGAAAATCCATGGACCGA
M AGGATCAAGTTGGCCCTTGCGAATAGCTTGATGCACTATCTCCCTTCAACCCTTTGAGCGAAATGTGGC-----AAAAGGAAGGAAAATCCATGGACCGA
S CCCCATCGTCTCCACCCCGTAGGAACTACGAGATCACCCCAAGGAACGCCTTCGGCATCCAGGGGTCGCGGACCGACCATAGAACCCTGTTCAAAAAGCG
1701 T CCCCATCATCTCCACCCCGTAGGAACTACGAGATCACCCCAAGGAACGCCTTCGGCATCCAGGGGTCACGGACCGACCATAGAACCCTGTTCAATAAGTG
M CCCCATTGTCTCCACCCCGTAGGAACTACGAGATCACCCCAAGGAGTTCGTCCTCAATGGGGGTCTATCGGACCGACCATAG-ATCCTGTTCAATAAGTG
S GAACGCATTAGCTATCCGCTCTCAGGTTGGACAGTAAGGGTCGGAGAAGGGCAATCACTCATTCTTA-112bpGTTAGAATGGGATTCCAACTCAGCACCTTT---
1801 T GAACGCATTAGCTGTCCGCTCTCAGGTTGGGCAGTCAGGGTCGGAGAAGGGCAATGACTCATTCTTA------GTTAGAATGGGATTCCAACTCAGCACCTTTTGA
M GAACACAATAGCCGTCCGCTCTCCGGTTGGGCAGTAAGGGTCGGAGGAGGGCAATCACTCGTTCTTA-103bp-TTAGAATGGGATTCCAACTCAGCACCTTTTGT
S --TGAGATTTTGAGAAGAGTTGCTCTTTGGAGAGCACAGTACGATGAAAGTTGTGAGCTGTGTTCGGGGGGGAGTTATTGTCTATCGTTGGCCTCTATGGT
1901 T G-TGAGATTTTGAGAAGAGTTGCTCTTTGGAGAGCACAGTACGATGAAAGTTGTAAGCTGTGTTCGGGGGGGAGTTATTGTCTATCGTTGGCCTCTATGGT
M TTTGGGATTTTGAGAAGAGTTGCTCTTTGGAGAGCACAGTACGATGAAAGTTGTAAGCTGTGTTCGGGGGGGAGTTATTGCCTATCGTTGTCCTCTATGGT

FIG. 4D

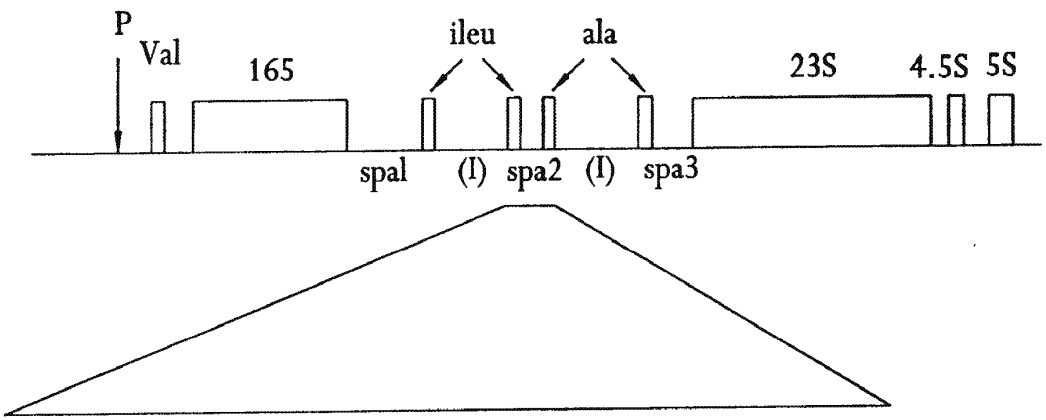
```
      **** * *******   ************************* ************************************ * ***********  *****
    S AGAATCAGTCGGGG---CCTGAGAGGCGGTGGTTTACCCTGTGGCGGATGTCAGCGGTTCGAGTCCGCTTATCTCCAACTCGTGAACTTAGTCGATACAAA
2001 T AGAATCAGTCGGGG-GACCTGAGAGGCGGTGGTTTACCCTGCGGCGGATGTCAGCGGTTCGAGTCCGCTTATCTCCAACTCGTGAACTTAGCCGATACAAA
    M AGAACCCGTCGGGGAGGCCTGAGAGGCGGTGGTTTACCCTGTGGCGGATGTCAGCGGTTCGAGTCCGCTTATCTCCAGCCCGTGAACTTAGCGGATAC---

      ***
    S GCTA
2101 T GCTT
    M ----

* indicates homology
- indicates gaps in the sequence compared to each other sequences
Nucleotide number corresponds to tobacco sequences only
S-soybean, T-tobacco, M-maize
```

FIG. 4E

GENE SEQUENCE OF rRHA CODING REGIONS IN PLASTID DNA FROM HIGHER PLANTS



SEQUENCE ALIGNMENT OF SPACER-2 (64 bp) REGION FROM SEVERAL CROP SPECIES WITH TOBACCO

```
EPIFAGUS  (90   %)     GCTGCGCTA-GGAAAAAAATATAAAAAGCATCTGATTACTTCATGCATGCT
TOBACCO (+)---      GCTGCGCCAGGGAAAAGAATAGAAGAAGCATCTGACTACTTCATGCATGCTCCA-CTTGGCTCGG
                    **** ** * * **** *  ** *      *****  ********* ***** ** ******
HELIANTHUS(96   %)     CGTGCGCCAGGGAAAAGAATAGAAGAAGCGTCTGACTCCTTCATGCATGCTCCA-CTTGGCTCGG
OENOTHERA (96   %)     GCTGCGCAAAGGAAAAGAATAGAAGAAGCATCTGACTCCTTCATGCATGCTCCA-CTTGGCTCGG
ALNUS     (95   %)     GCTGCGCCAAGTAAAAGAATAGAAGAAGCATCTGACTCCTTCATGCATGCTCCA-CTTGGCTCGG
RICE      (95   %)     GCTGCGCCAGGGAAAAGAATAGAAGAAGCATCTGACTCTTTCATGCATACTCCA-CTTGGCTCGG
MAIZE     (94   %)     GCTGCG-CAGGGAAAAGAATAGAAGAAGCATCTGACTCTTTCATGCATACTCCA-CTTGGCTCGG
SOYBEAN   (84   %)     GCTGCGTCAAGGAAAAGAATAGAAAACTGACTTGACTCCTTCATGCATGCTCCA-CTCGGCTCGG
```

FIG. 4F

```
PEA       (89   %)     GCTGCGCCAAGGAAAAGACTAAAAGACGGATTTGACTCCTTCATGCATGCTCCAACTTGGCTCGG
SPINACH   (89   %)     ACTGCGCCAAGAATAAGAATCGAAGAAGCGTCTGACTCCTTCATGCATGCTCCA-CTTGGCTCGG
                                 / \
                               CGCCAGGGAA

TOBACCO   (-)---       CCGAGCCAAGTGGAGCATGCATGAAGTAGTCAGATGCTTCTTCTATTCTTTTCCCTGGCGCAGC
CUSCUTA   (96   %)     CCGAGCCAAGTGGAGCATGCATGAAGTAGTCAGATACTTCTTCGATTCTTTTCCCTGGCGCAGC
```

FIG. 4G

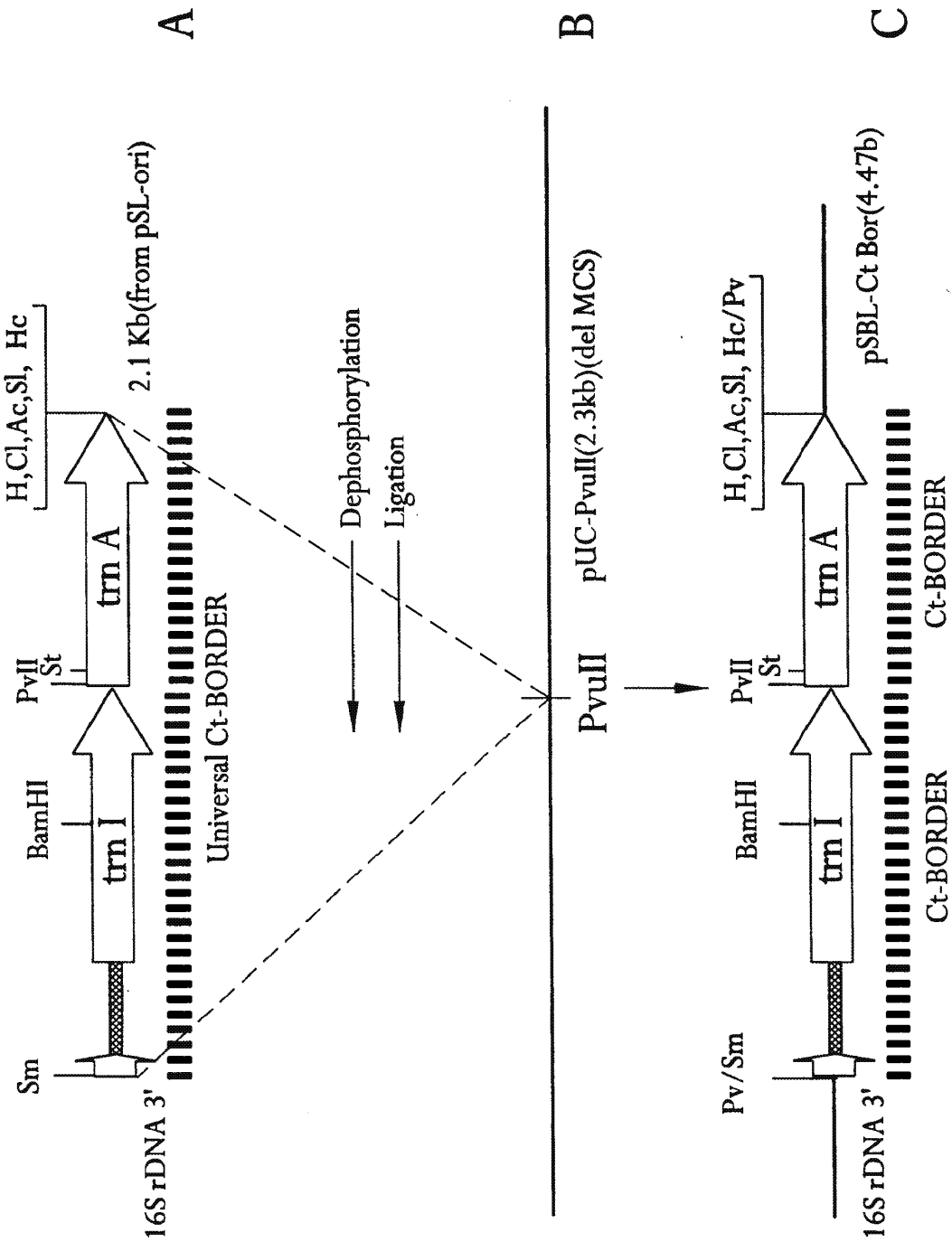


FIG. 5

COR-AAD1-00002155

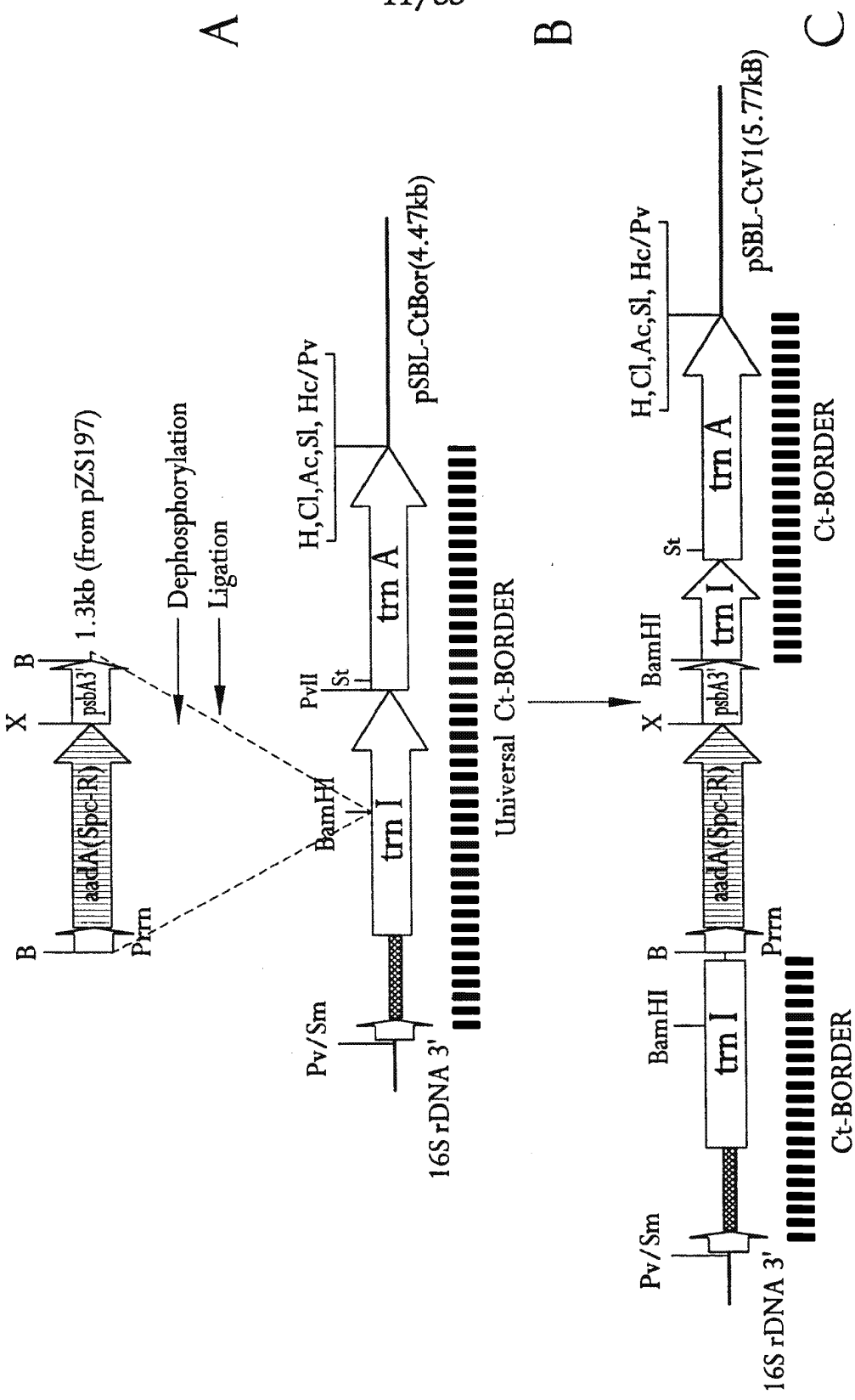


FIG. 6

COR-AAD1-00002156

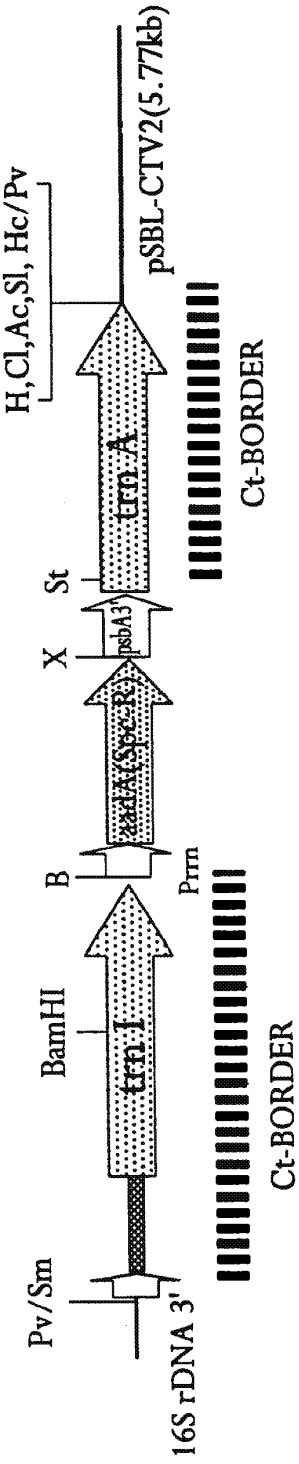

H,Cl,Ac,Sl, Hc/Pv
pSBL-CTV2(5.77kb)
trn A
St
X
psbA3'
aadA(Spc-R)
B
Prrn
BamHI
trn I
Pv/Sm
16S rDNA 3'
Ct-BORDER
Ct-BORDER

FIG. 7A

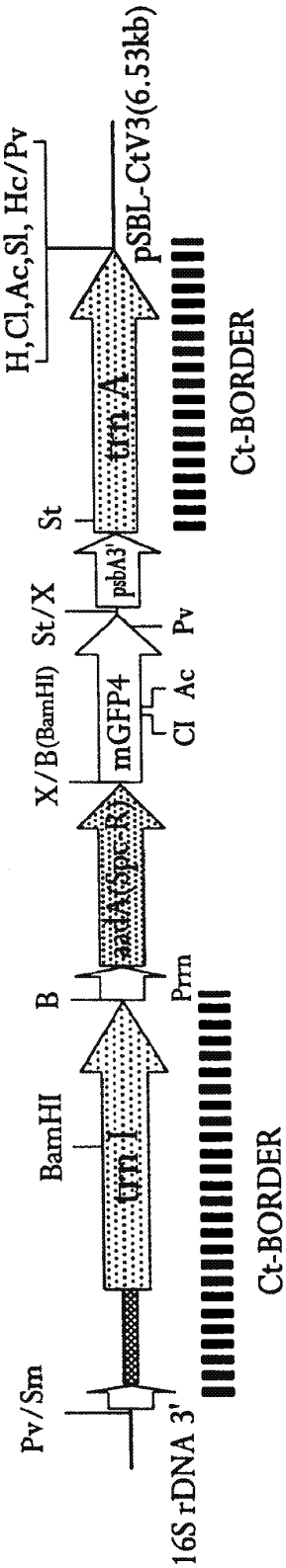

H,Cl,Ac,Sl, Hc/Pv
pSBL-CtV3(6.53kb)
trn A
St
psbA3'
St/X
Pv
Cl Ac
mGFP4
X/B(BamHI)
aadA(Spc-R)
B
Prrn
BamHI
trn I
Pv/Sm
16S rDNA 3'
Ct-BORDER
Ct-BORDER

FIG. 7B

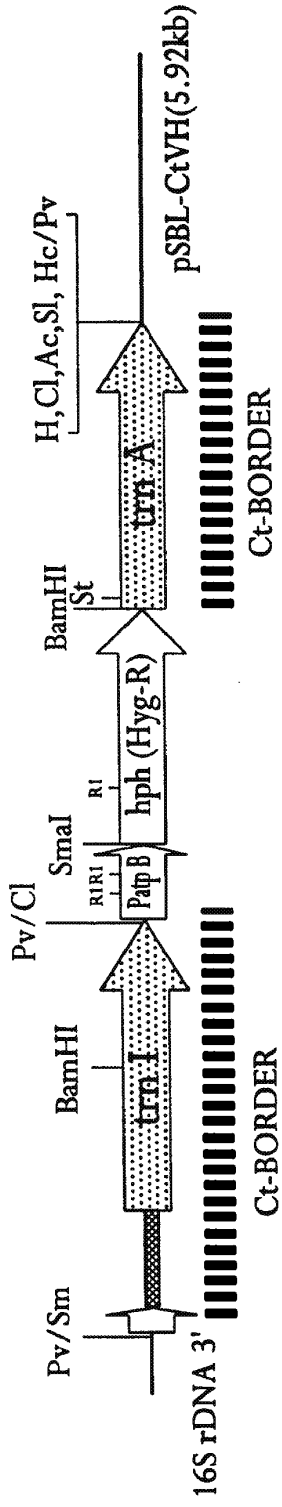

Pv/Sm
BamHI
trn I
Pv/Cl
SmaI
R1
R1R1
Patp B
hph (Hyg-R)
BamHI
St
trn A
H,Cl,Ac,Sl, Hc/Pv
pSBL-CtVH(5.92kb)
16S rDNA 3'
Ct-BORDER
Ct-BORDER

FIG. 7C

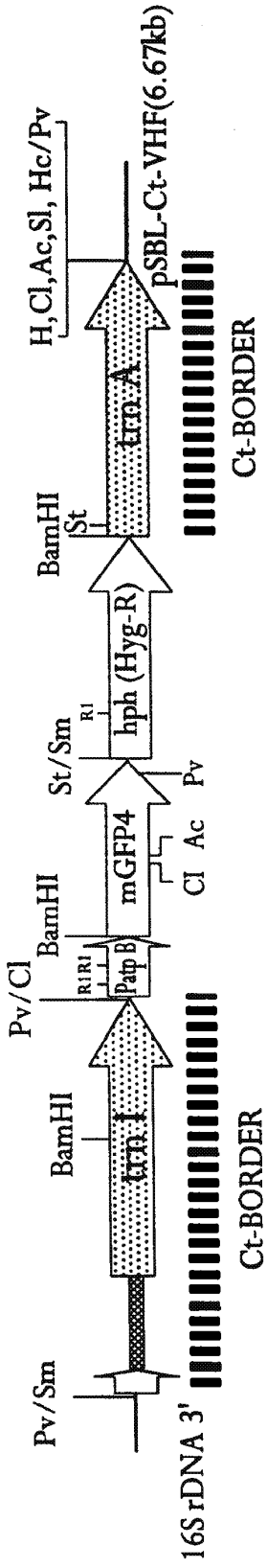

Pv/Sm
BamHI
trn I
Pv/Cl
BamHI
R1R1
Patp B
mGFP4
Cl Ac
St/Sm
Pv
R1
hph (Hyg-R)
BamHI
St
trn A
H,Cl,Ac,Sl, Hc/Pv
pSBL-Ct-VHF(6.67kb)
16S rDNA 3'
Ct-BORDER
Ct-BORDER

FIG. 7D

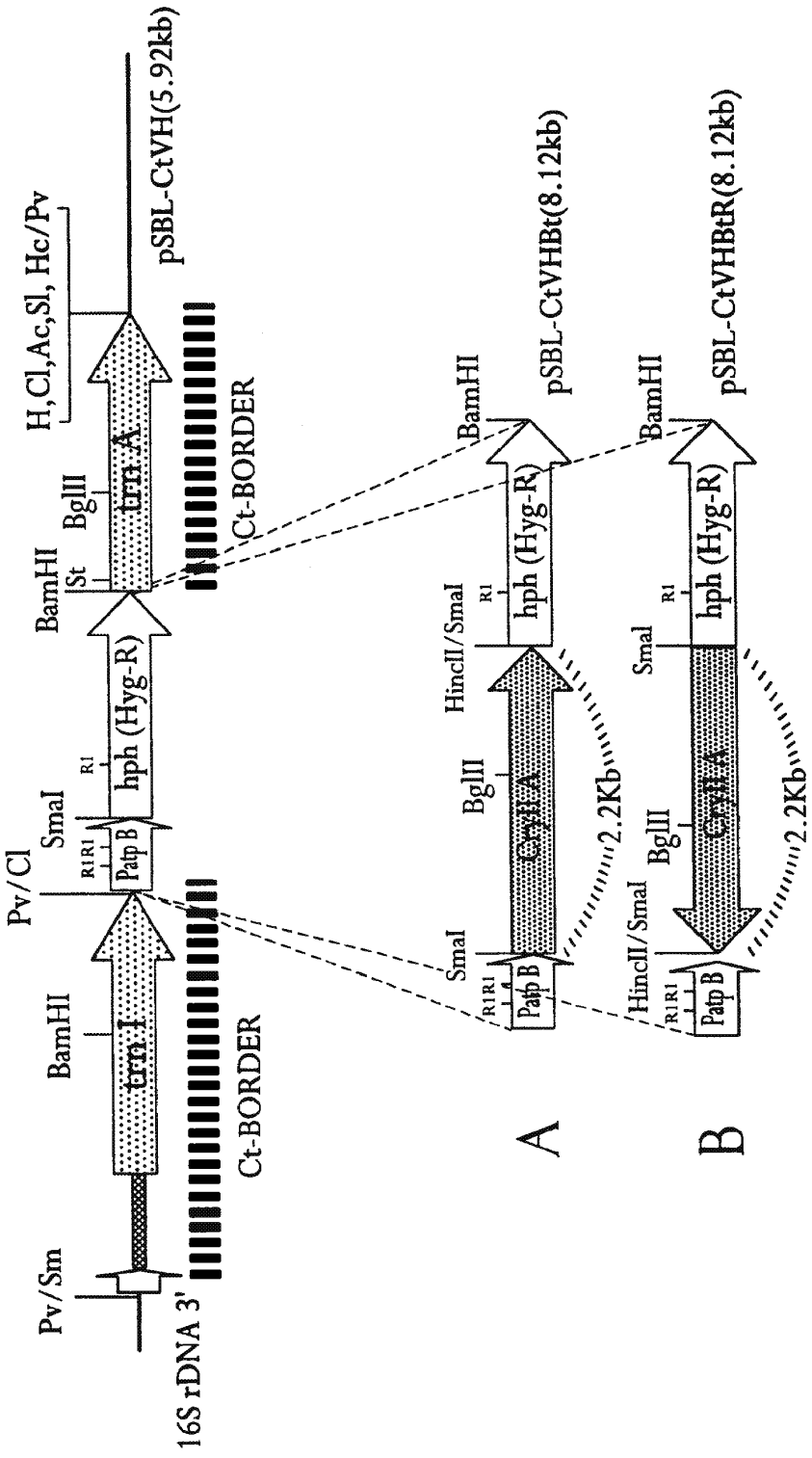

pSBL-CtVH(5.92kb)
H,Cl,Ac,Sl, Hc/Pv
BamHI
BglII
trn A
St
Ct-BORDER
hph (Hyg-R)
R1
SmaI
Pv/Cl
R1R1
Patp B
BamHI
trn I
Ct-BORDER
Pv/Sm
16S rDNA 3'
pSBL-CtVHBt(8.12kb)
BamHI
hph (Hyg-R)
R1
HincII/SmaI
BglII
Cry1 A
2.2Kb
SmaI
R1R1
Patp B
A
pSBL-CtVHBtR(8.12kb)
BamHI
hph (Hyg-R)
R1
SmaI
BglII
Cry1 A
2.2Kb
HincII/SmaI
R1R1
Patp B
B

FIG. 8

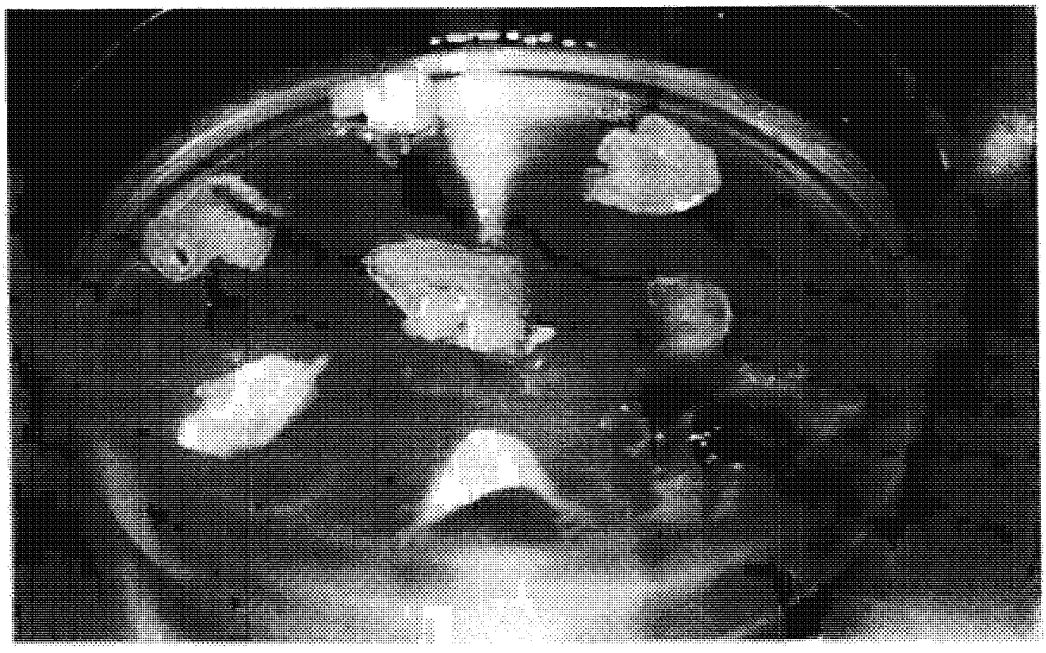


FIG. 9

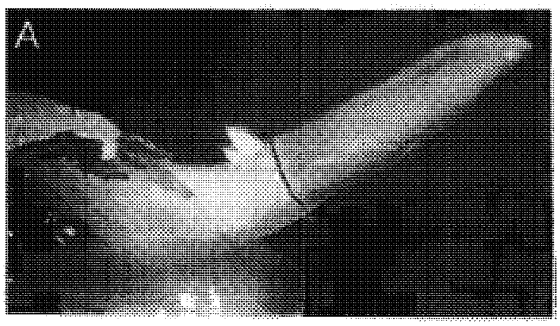


FIG. 10A

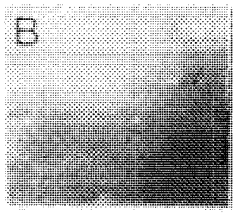


FIG. 10B

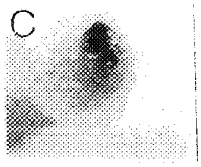


FIG. 10C

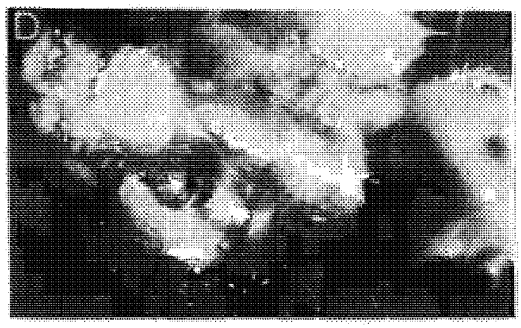


FIG. 10D



FIG. 10E

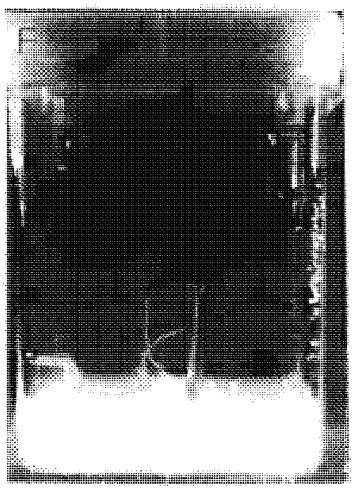


FIG. 10F

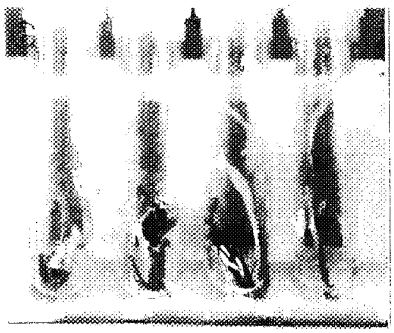


FIG. 10G

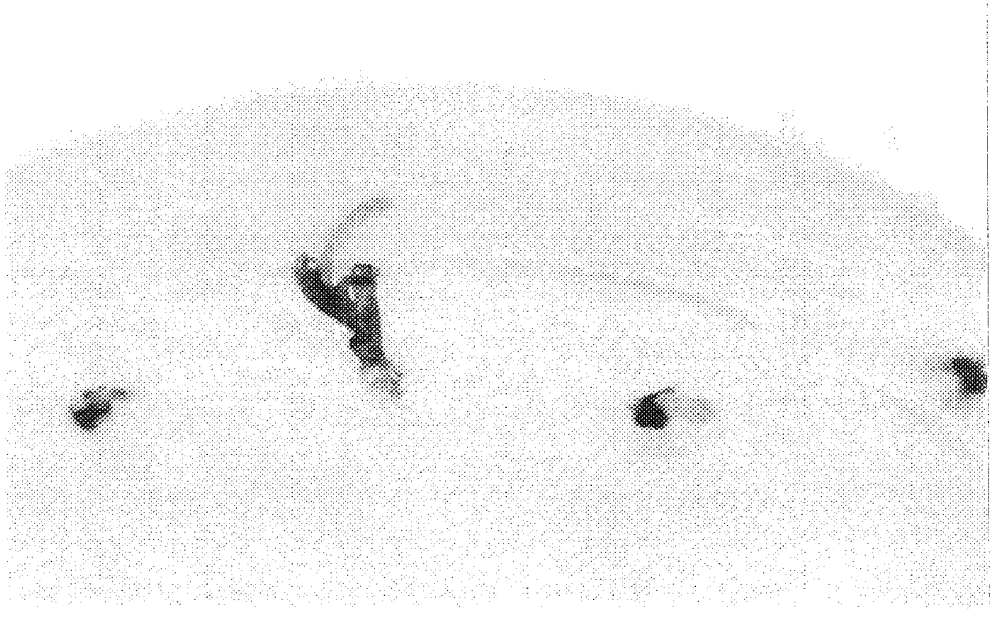


FIG. 11

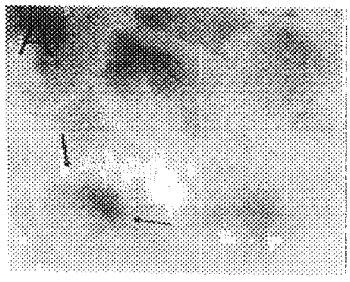


FIG. 12A

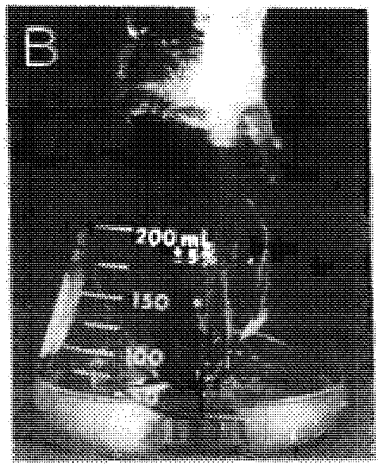


FIG. 12B

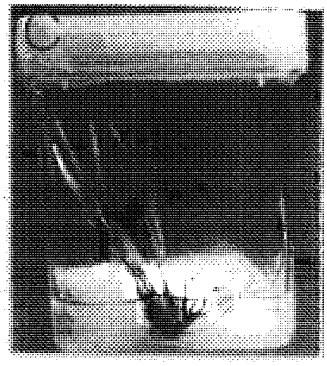


FIG. 12C



FIG. 12D

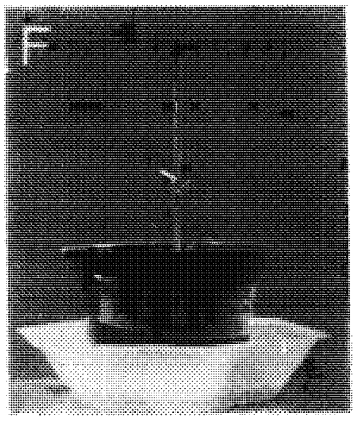


FIG. 12E

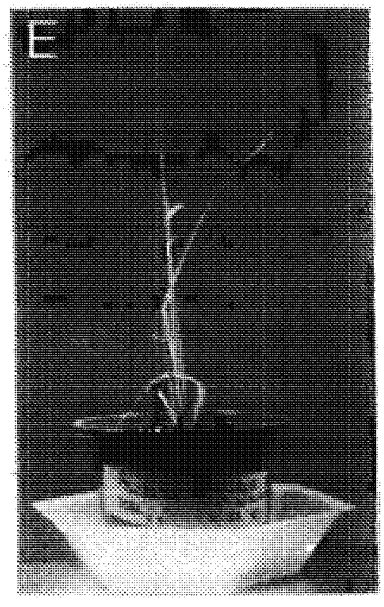


FIG. 12F

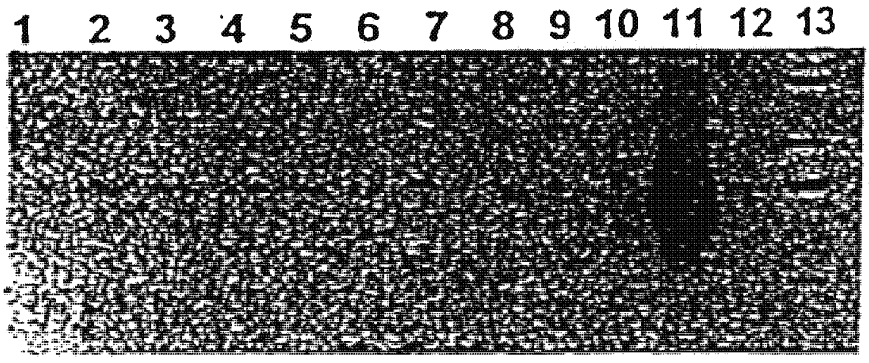


FIG. 13A

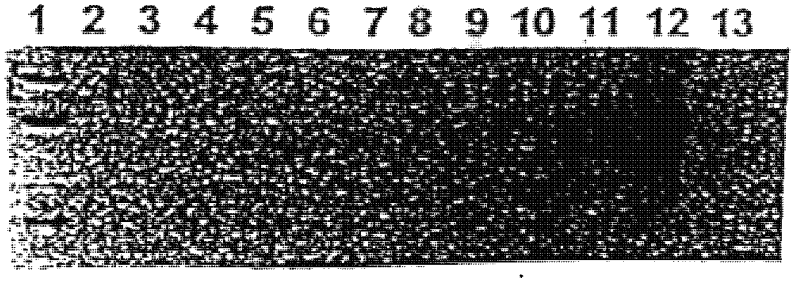


FIG. 13B

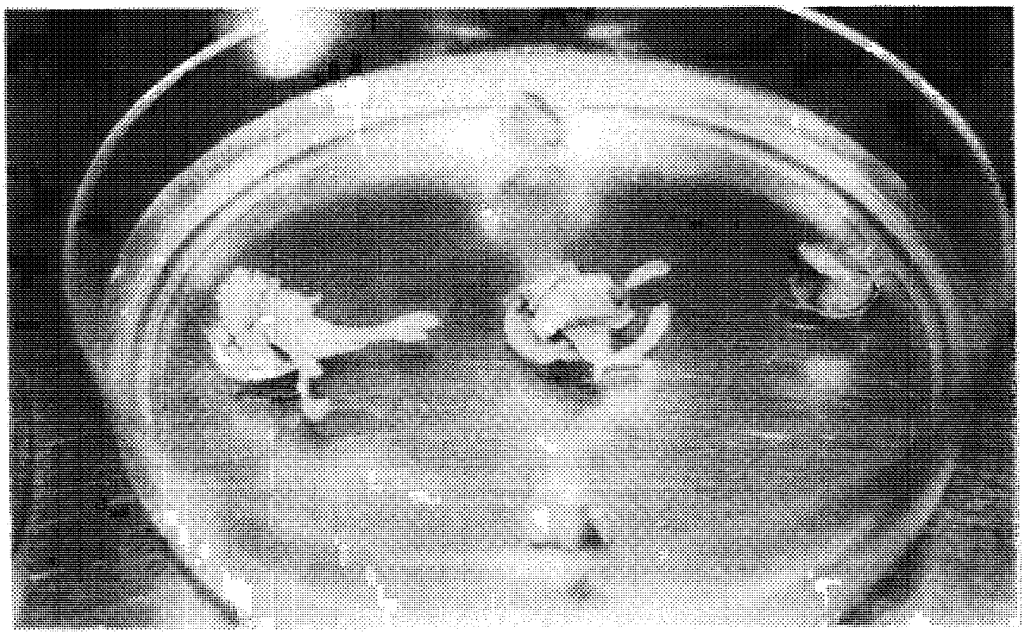


FIG. 14

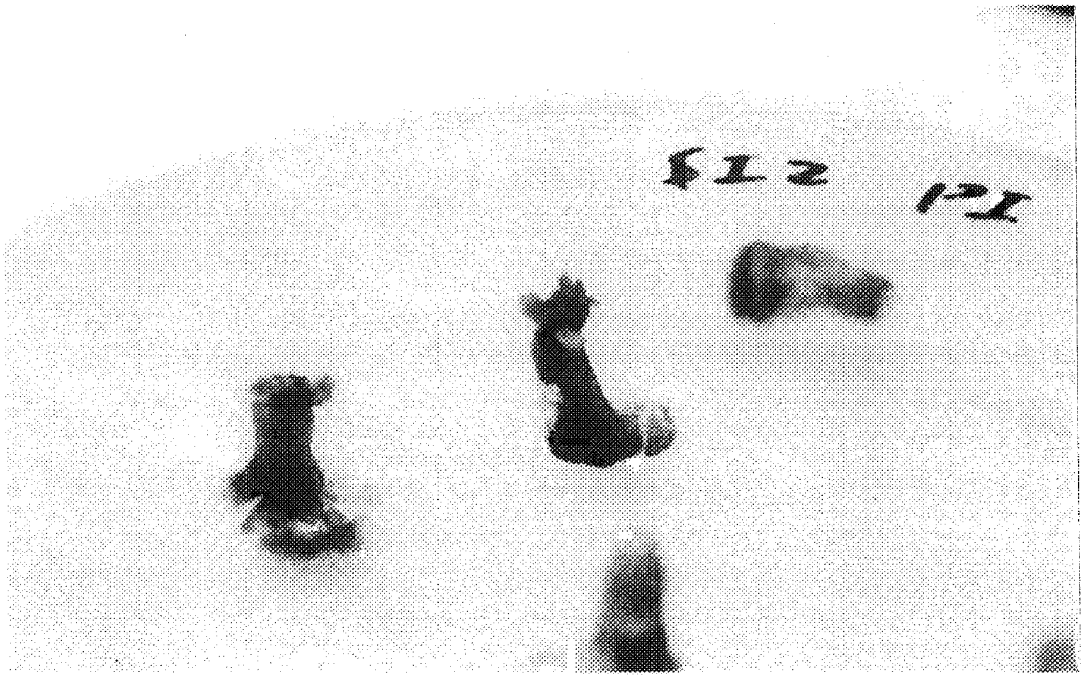


FIG. 15

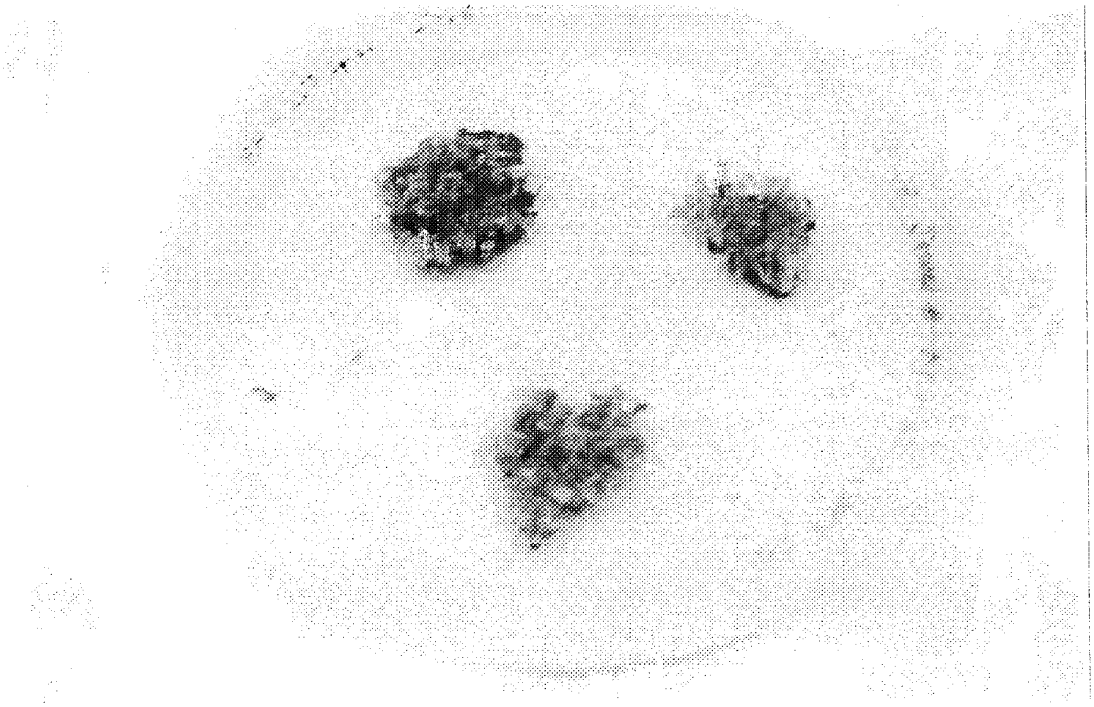


FIG. 16

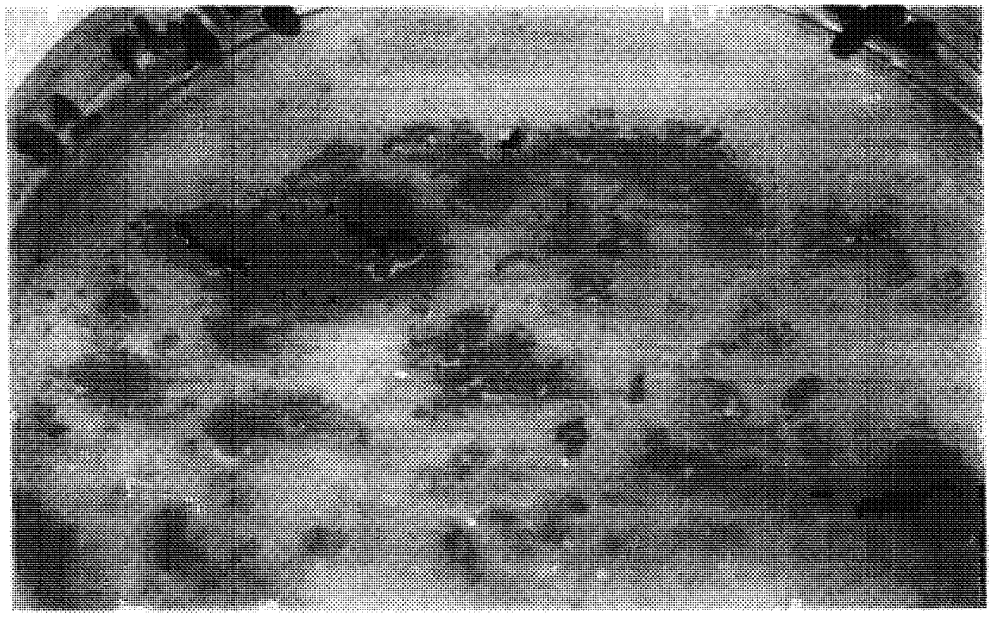


FIG. 17

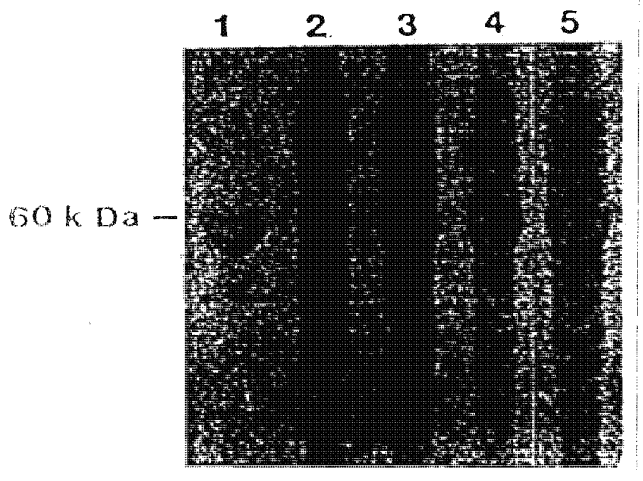


FIG. 18

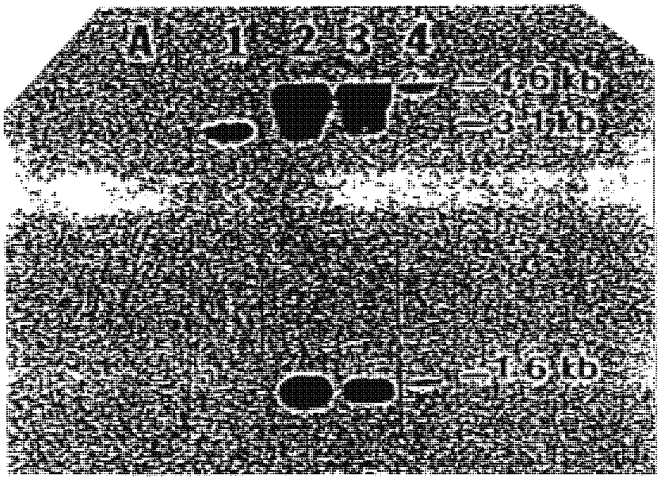


FIG. 19A

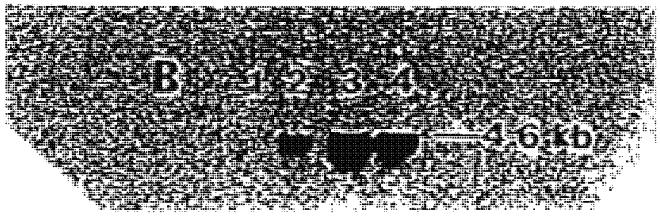


FIG. 19B

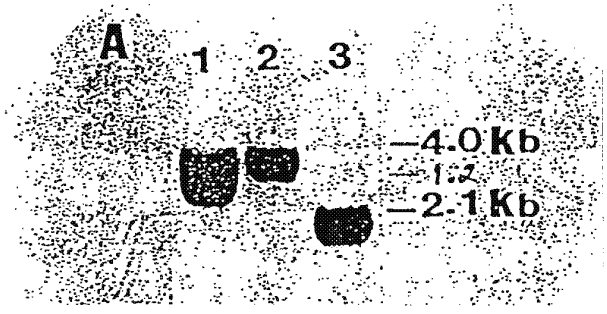


FIG. 20A

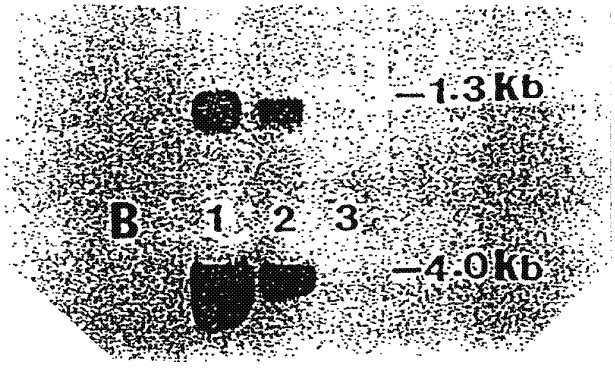


FIG. 20B

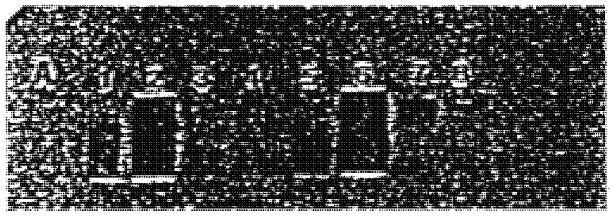


FIG. 21A

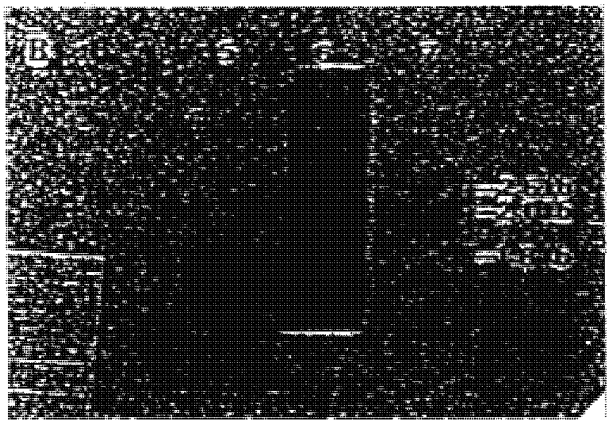


FIG. 21B

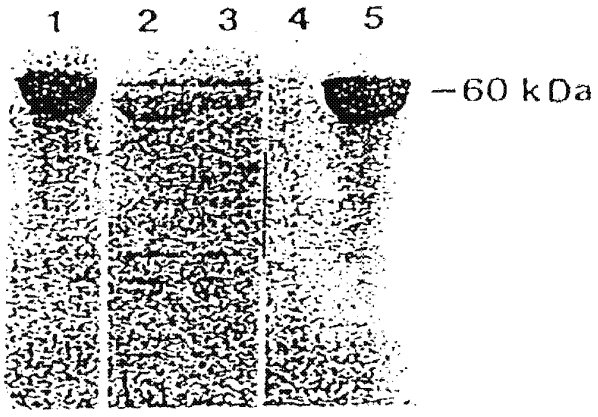


FIG. 22

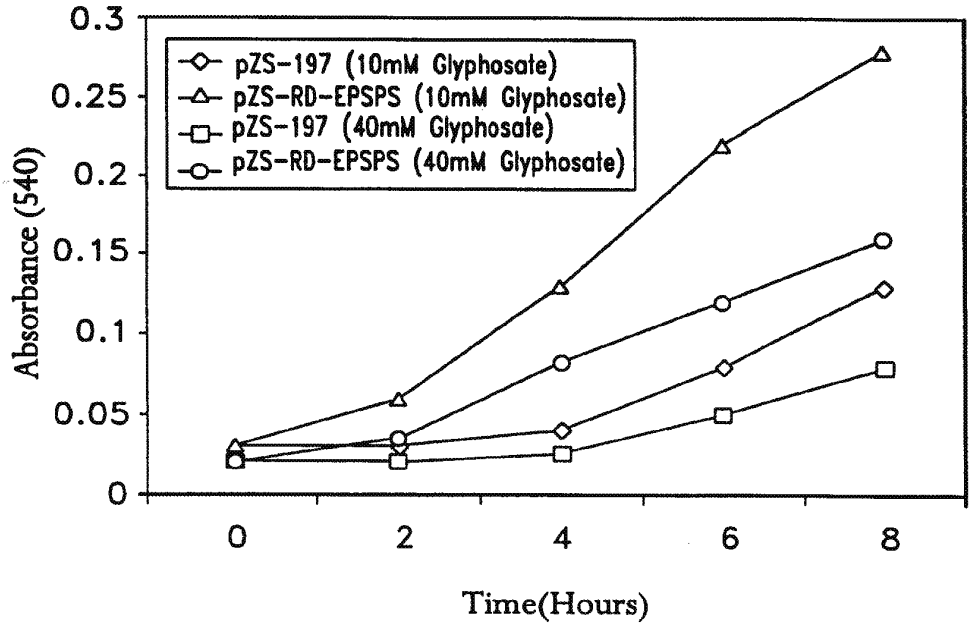


FIG. 23A

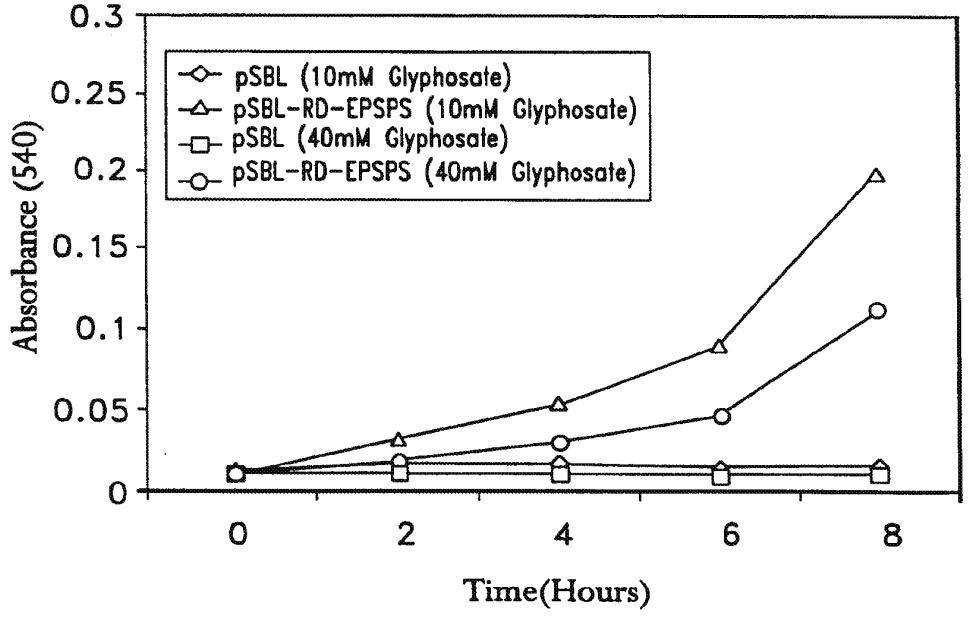


FIG. 23B

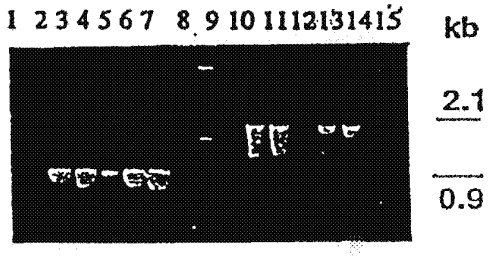


FIG. 24A

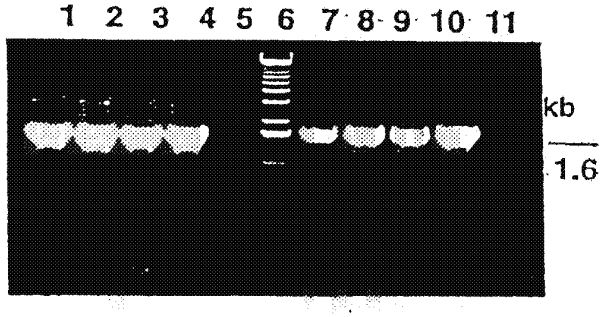


FIG. 24B

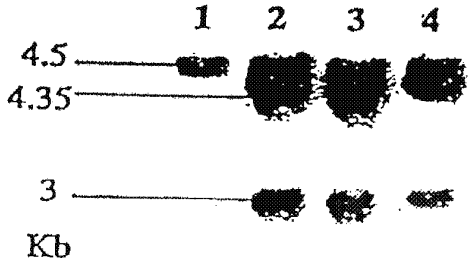


FIG. 25A

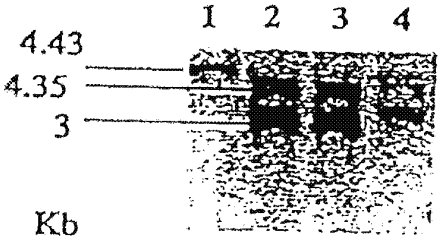


FIG. 25B

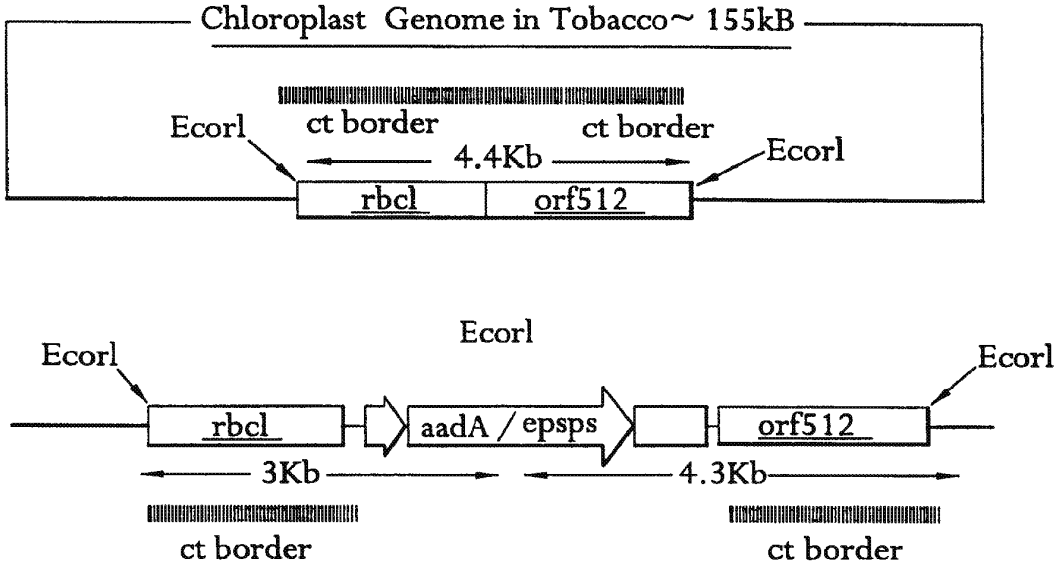

Chloroplast Genome in Tobacco~ 155kB
EcorI
ct border
ct border
4.4Kb
EcorI
rbcl
orf512
EcorI
EcorI
EcorI
rbcl
aadA / epsps
orf512
3Kb
4.3Kb
ct border
ct border
Chloroplast Genome in transgenic tobacco

FIG. 25C

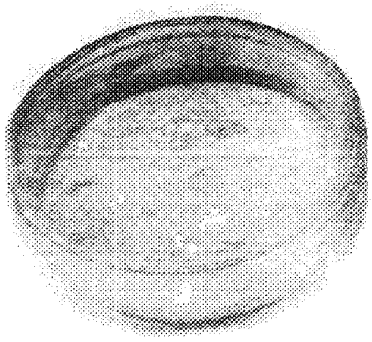


FIG. 26A

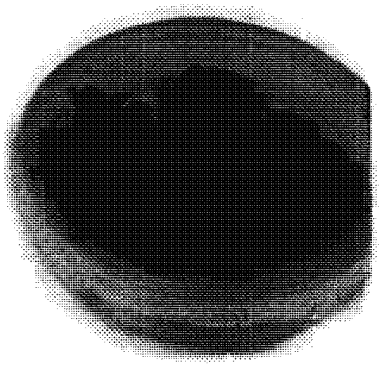


FIG. 26B



FIG. 27A

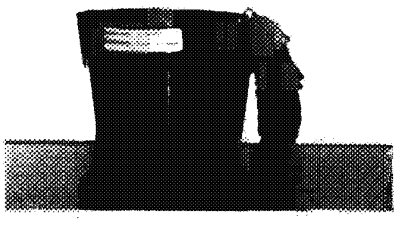


FIG. 27B

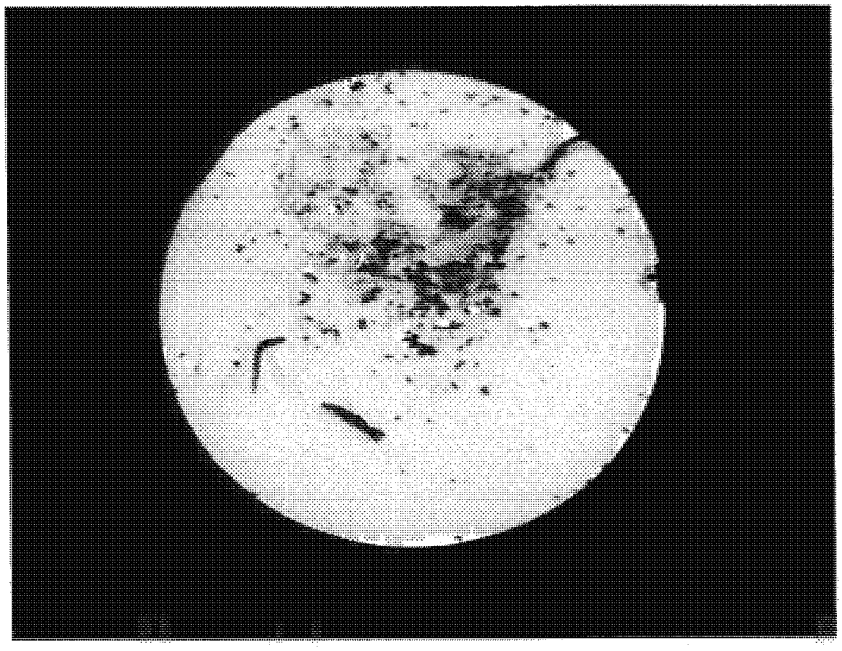


FIG. 28A

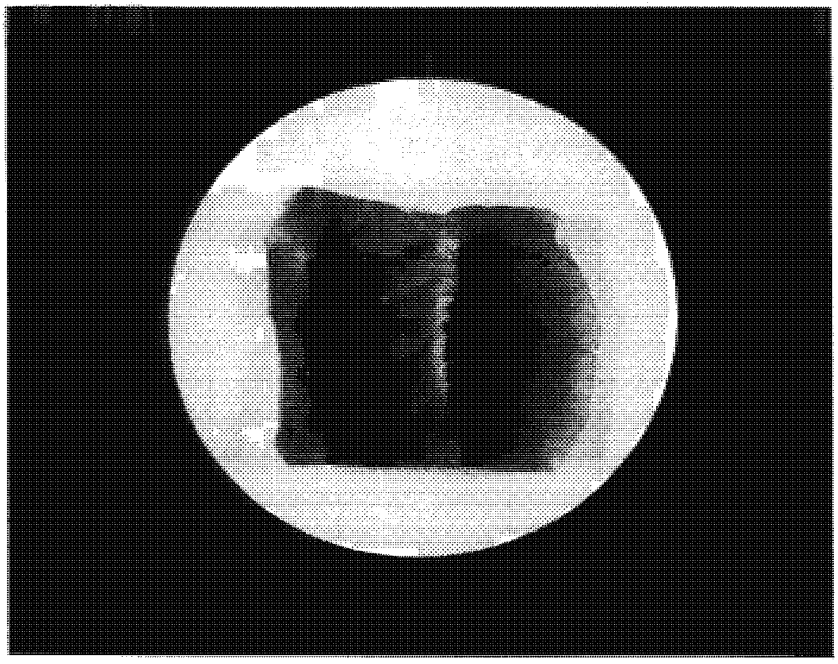


FIG. 28B

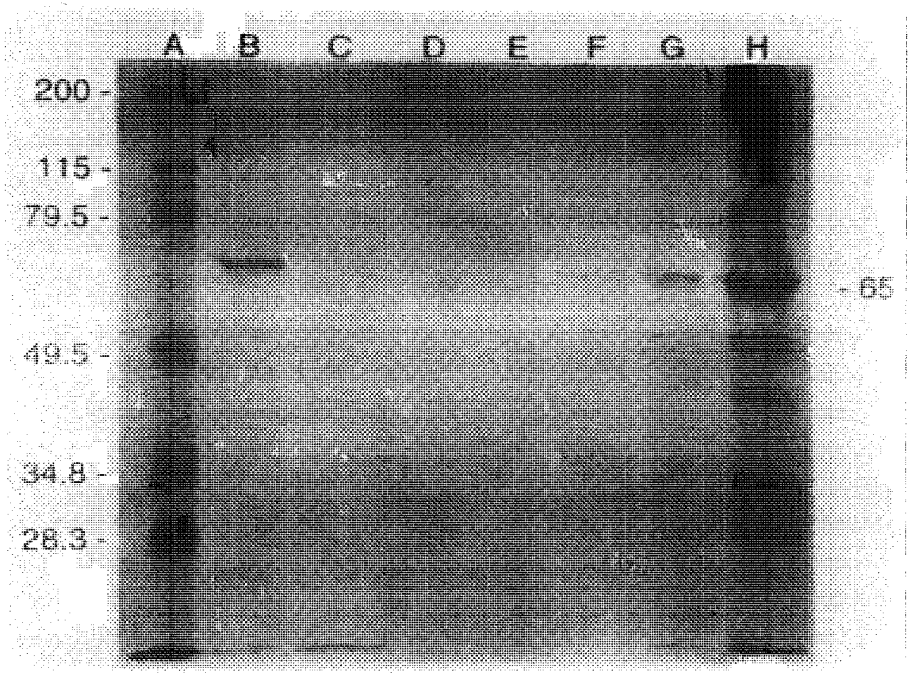


FIG. 29

INTERNATIONAL SEARCH REPORT

Interne al Application No
PCT/IB 98/01199

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCBRIDE, K.E., ET AL. : "amplification of a chimeric Bacillus gene in chloroplasts leads to an extraordinary level of an insecticial protein in tobacco" BIO/TECHNOLOGY, vol. 13, 13 April 1995, pages 362-365, XP002090404 see the whole document ---	1-3,14, 171,179, 183
X	ZOUBENKO, O.V., ET AL. : "efficient targeting of foreign genes into the tobacco plastid genome" NUCLEIC ACID RESEARCH, vol. 22, no. 19, 1994, pages 3819-3824, XP002090405 see the whole document ---	1-3,114, 171,179, 183

-/--

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 January 1999	01/02/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

1

Form PCT/ISA/210 (second sheet) (July 1992)

COR-AAD1-00002181

INTERNATIONAL SEARCH REPORT

Internat ·l Application No
PCT/1B 98/01199

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DATTA, R. , ET AL.: "transformation of he tobacco chloroplast genome with the aroA gene to confer glyphosate tolerance" PLANT PHYSIOLOGY, SUPPLEMENT, vol. 111, no. 2, June 1996, page 168 XP002090406 see the whole document	1-3,114, 171,179, 183
A	STAUB, J.M. AND MALIGA, P.: "long egions of homologous DNA are incorporated into the tobacco plastid genome by transformation" THE PLANT CELL, vol. 4, 1992, pages 39-45, XP002090407 see figure 1	1-188
A	LU, Z., ET AL. : "characterization of replication origins flanking the 23S rRNA gene in tobacco chloroplast DNA" PLANT MOLECULAR BIOLOGY, vol. 32, 1996, pages 693-706, XP002090408 page 704; Fig. 4; Fig. 1B	1-188
P,X	DANIELL,H., ET AL. : "containment of herbicide resistance through genetic engineering of the chloroplast genome" NATURE BIOTECHNOLOGY, vol. 16, April 1998, pages 345-348, XP002090409 see the whole document	1-45, 55-57, 61-63, 67-69, 76-78, 82-86, 88,89, 92,94, 96,107, 114, 118-120, 122 123-126, 133, 140-147, 152, 160-173, 175-179, 183

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

COR-AAD1-00002182

INTERNATIONAL SEARCH REPORT

Int. ational application No.
PCT/IB 98/01199

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
Claims 28-30 were read as referring to sequences that originate from the same plant species as the target plant.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

COR-AAD1-00002183

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/29, A01H 5/00, C07K 14/415

A1

(11) International Publication Number: **WO 99/63092**

(43) International Publication Date: 9 December 1999 (09.12.99)

(21) International Application Number: PCT/US99/12277

(22) International Filing Date: 3 June 1999 (03.06.99)

(30) Priority Data:
60/087,789 3 June 1998 (03.06.98) US

(71) Applicant: WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH [US/US]; Nine Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors: LUSCHNIG, Christian; Ortliebgasse 5–9/118, A–1170 Vienna (AT). GAXIOLA, Roberto, A.; Apartment 3, 228 Commonwealth Avenue, Boston, MA 02116 (US). GRISAFI, Paula; 22 King Philip Road, Sudbury, MA 01776 (US). FINK, Gerald, R.; 40 Ashton Road, Chestnut Hill, MA 02167 (US).

(74) Agents: HOGLE, Doreen, M.; Hamilton, Brook, Smith & Reynolds, P.C., Two Militia Drive, Lexington, MA 02421 (US) et al.

(81) Designated States: CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ROOT–SPECIFIC PROTEIN INVOLVED IN AUXIN TRANSPORT

(57) Abstract

A root–specific plant gene which encodes an auxin–transport carrier protein that is required for gravitropism; an auxin–transport efflux carrier protein; genetically engineered plants whose genomes comprise heterologous DNA encoding a root–specific auxin efflux carrier protein, or heterologous DNA encoding a portion of an auxin efflux carrier protein sufficient to encode a functional carrier protein and to confer a phenotype characterized by a decreased sensitivity to a herbicide which is an auxin derivative, an auxin analogue or a formulation comprising an auxin transport inhibitor in combination with a second herbicide; and methods useful for identifying molecular targets involved in gravitropic signal transduction, for evaluating the effects of agents on auxin transport and for elucidating the role of gene expression and the molecular mechanism of polar auxin transport.

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

ROOT SPECIFIC PROTEIN INVOLVED IN AUXIN TRANSPORT

BACKGROUND OF THE INVENTION

The plant hormone auxin (indole-3-acetic acid (IAA), is involved in plant growth (cell division and cell expansion), morphogenesis (e.g., the differentiation of vascular tissue and lateral or adventitious root formation) and physiological responses to the environment, such as phototropism and gravitropism. Plant tropisms, growth towards or away from a stimulus such as light or gravity, have been ascribed to asymmetric plant growth in which one side of a plant organ elongates to a greater extent than the other, resulting in a curvature toward or away from the stimulus (Darwin, C., *et al.*, (1880) *Power of movements in plants*. John Murray, London; Poff *et al.*, (1994) The physiology of tropisms, In *Arabidopsis,* 639-664, eds. Meyerowitz, E. M., and Somerville, C.R., Cold Spring Harbor Laboratory Press). Root gravitropism (growth in a direction defined by gravity) can be demonstrated by manipulating plants so that they lie horizontal to the surface of the earth (gravistimulation) (Okada, K. and Y. Shimura (1992) *Aust. J. Plant Physiol. 19*: 439-448). Within a short time, the roots curve downward exhibiting a positive gravitropic growth response. Transport studies suggest that IAA is redistributed in response to gravity so that it accumulates along the lower side of the root tip (Young, L.M., *et al.*, (1990) *Plant Physiol. 92*: 792-796). Removal of the root tip abolishes gravitropism; and it is well established that polar auxin transport can be specifically inhibited by synthetic compounds, known as auxin transport inhibitors (Galwelier, L. et al., (1998) Science 282: 2226-2230). Thus, redistribution of IAA in the root tip may be critical to gravitropism (Blancaflor, E.B., *et al.*, (1998) *Plant Physiol. 116*: 213-222). These observations are consistent with earlier views (the Cholodny-Went hypothesis, see Estelle, M., (1996) *Curr Biol. 6*: 1589-91) which suggested that when roots are oriented horizontally (e.g., gravistimulated), IAA accumulates along the lower side of the

elongating zone, resulting in inhibition of cell elongation in those cells while those on top elongate, a process that eventuates in the downward bending of the root.

It is well established that specialized auxin transport systems, comprising auxin influx and efflux activities, exist in several species of plants (Estelle, M., (1996) *Curr Biol. 6*: 1589-91). IAA is thought to be polarly transported, from its point of synthesis in the plant shoot, down to the root (acropetal transport) tip via the vascular system, and then transported up from the root tip to the elongation zone (basipetal transport) where it probably localizes in the epidermis. This polarized cell to cell transport can be explained by the chemiosmotic hypothesis (Goldsmith, M.H.M., (1977) *Ann. Rev. Plant Physiol. 28*: 439-478; Lomax, T.L., *et al.*, (1995) In *Plant Hormones, Physiology, Biochemistry and Molecular Biology,* P.J. Davies, Ed. (Martinus, Nijhoff, Kluwer, Dordrecht, Boston, London, pp 509-530). This model posits that uncharged IAA in the acidic extracellular space enters a cell either by passive diffusion or facilitated transport. Upon entry into the relatively more basic cytosol, IAA dissociates to form IAA^-. The transit of IAA^- (auxin anion) out of a cell and into an adjoining cell is thought to depend on, and be regulated by, an efflux carrier protein (Lomax, T.L., *et al.*, (1995) In Plant Hormones, Kluwer Academic Publishers. Dordrecht, Boston, London; Jacobs, M. and S.F. Gilbert, (1983) *Science. 220*: 1297-1300). Thus, gravitropism could result from the differential activity of an IAA efflux carrier in response to gravity.

Support for the chemiosmotic hypothesis comes from the effects of auxin transport inhibitors such as 2,3,5-triiodobenzoic acid (TIBA) and N-1-naphtylphtalamic acid (NPA) (Thomson, K.-S., *et al.*, (1973) *Planta (Berl). 109*: 337-352; Katekar, G.F. and A.E. Geissler, (1980) *Plant Physiol. 66*: 1190-1195), which interfere with auxin efflux. (Sussman, M.R. and M.H.M. Goldsmith, (1981) *Planta. 152*: 13-18) Plants grown in the presence of TIBA or NPA are agravitropic (Mulkey, T.J. and M.L. Evans, (1982) *J. Plant Growth Regul. 1*: 259-265; Lee, J.S., *et al.*, (1984) *Planta. 160*: 536-543). Moreover, mutants with altered response to these auxin transport inhibitors have phenotypes consistent with the hypothesis that transport of auxin is critical for the gravitropic response. Mutants resistant to IAA have phenotypes that also support the involvement of IAA in gravitropism. Several auxin resistant mutants are agravitropic

(Estelle, M. and H.J. Klee, (1994) *Arabidopsis* Cold Spring Harbor Laboratory Press). Despite the connection between auxin transport inhibitors and gravitropism, the targets of these inhibitors and the molecules involved in directed auxin transport have not yet been identified.

SUMMARY OF THE INVENTION

Described herein is the isolation and characterization of a plant gene which encodes an auxin-transport-efflux carrier protein that is required for gravitropism. The disclosed protein and gene are targets for regulation of auxin transport in response to the hormones ethylene and auxin (including auxin analogues and auxin derivatives) and the inhibition of auxin transport mediated by synthetic transport inhibitors. Also described are uses of the gene and the encoded protein; mutant forms of the gene and the encoded protein; modified EIR1 nucleic acid molecules; assays which are useful for identifying and characterizing mutant forms of the gene (variant nucleic acid molecules; mutants or alleles) and the encoded protein; methods of altering auxin homeostasis in plant roots; methods of producing genetically engineered plants, such as crop plants and flowering plants, which show greater resistance to herbicides which are auxin derivatives or auxin analogues or formulations comprising an auxin transport inhibitor in combination with a second herbicide than is shown by the corresponding wild type plants; genetically engineered plants with greater resistance to herbicides; seeds, leaves and other plant tissues or parts obtained from such plants; and seeds from which plants with increased herbicide resistance can be produced.

A specific embodiment of the present invention relates to a plant gene, referred to as *EIR*1 (for: Ethylene Insensitive Root) , and its encoded auxin transport (e.g., efflux) carrier protein, EIR1, which is required for gravitropism; assays useful for assessing *EIR1* activity and determining structure/function relationships characteristic of mutagenized alleles of *EIR1*; inhibitors and enhancers of EIR1 protein identified by the assays; methods of increasing transport of (efflux) auxin in plant roots by introducing *EIR1* DNA into the root of a plant, directly or indirectly (e.g., by producing plants from seeds containing exogenous *EIR1* DNA); methods of producing plants which exhibit greater resistance to herbicides (relative to the susceptibility exhibited by

the corresponding wild type plant) which are auxin derivatives or auxin analogues or formulations comprising an auxin transport inhibitor in combination with a second herbicide relative to the susceptibility that is exhibited by the corresponding wild type plants; genetically engineered (e.g., transgenic) plants in which the roots contain and express heterologous DNA, or a portion or fragment thereof which encodes a protein which is involved in auxin transport (e.g. EIR1 DNA, REH1 DNA) wherein the DNA is expressed in the roots as a functional root–specific auxin transport protein; and the transgenic plant exhibits greater resistance (or tolerance) to herbicides which are auxin derivatives or auxin analogues or compositions comprising an auxin transport inhibitor, than the sensitivity exhibited by the corresponding wild type plant; plant tissues or parts obtained from such plant tissues; and seeds from which plants with increased herbicide resistance can be produced. Also the subject of this invention are mutant *eir*1 genes (mutant alleles), the encoded mutant protein and *eir*1 mutant plants, in which the roots are agravitropic and have a reduced sensitivity to ethylene (relative to wild type plants).

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A - 1C are dose response curves of normalized root growth from wild type plants (black circle, Col-O; black square, Ws;) and *eir1* mutants (open circle, *eir1-1*; open square, *eir1-3*) in the presence of 1-aminocyclopropane-1-carboxylic-acid (ACC; the immediate biosynthetic precursor of ethylene), 2,3,5 triiodobenzoic acid (TIBA ;an inhibitor of auxin transport), and napthaleneacetic acid (NAA; an auxin analogue). Root elongation determined at 12 days after germination (DAG) was normalized to root growth on unsupplemented medium (100%). Standard deviations are shown as bars; molarities used are indicated.

Figure 2 is a schematic representation of an *Eco*RI fragment isolated from phage

λ5-3. The bars indicate the 9 exons of *EIR1*. Those segments presumed to be translated are black. Two mutations are indicated beneath the line: Insertion of Ac in *eir1-3* after amino acid 133 and base substitution of the intron/exon junction in *eir1-1*. The grey bar above the line indicates the genomic fragment amplified by inverse PCR

as described herein.. Abbreviations for restriction sites are as follows: RI: *Eco*RI; H: *Hin*DIII; Ba: *Bam*HI; X: *Xba*I; B: *Bcl*I.

Figure 3 shows the alignment of the deduced amino acid sequences of *EIR1* (SEQ ID NO.: 1), the rice homologue *REH1* (SEQ ID NO.: 2) and the two putative *Arabidopsis* homologues *AEH1* (SEQ ID NO.: 3) and *AEH2* (SEQ ID NO.:4). For *EIR1* and *REH1*, ORFs of the cDNAs were deduced. The protein sequences of *AEH1* and *AEH2* were deduced from the genomic sequences by identifying canonical splice donor and acceptor sites. Identical residues are boxed and dashes indicate gaps in the sequence. Black lines correspond to the 10 potential transmembrane domains shared by all 4 proteins. Potential, conserved N-glycosylation sites are typed in bold letters. An arrow indicates the cleavage site of a potential N-terminal signal peptide found for EIR1, REH1 and AEH1.

Figure 4 shows the alignment of the conserved – (top) EIRI, SEQ ID NO.:5; REHI, SEQ ID NO.:6 and C-terminal (bottom) EIRI, SEQ ID NO.: 25; REH1, SEQ ID NO: 26) protein transmembrane domains of EIR1 and REH1 with a number of selected bacterial transporters (mdcF, SEQ ID NO.: 7; livM, SEQ ID NO.: 8; arsB, SEQ ID NO: 9; and sbmA, SEQ ID NO.: 10). Identical residues are boxed. The bold letters represent positions where exchanges are conservative (L, I, V, M; A, S, T; F, W , Y; N, Q; D, E; and K, R) and shared by EIR1 and at least two other sequences. Dashes indicate gaps in the alignment.

Figure 5 is the nucleotide sequence of *EIR1* genomic DNA (SEQ ID NO.: 11), including the promoter.

Figure 6 is the nucleotide sequence of *EIR1* cDNA (SEQ ID NO.: 12) (GenBank accession number AF056026).

Figure 7 is the amino acid sequence of EIR1 protein (SEQ ID NO.: 1).

Figure 8 is the nucleotide sequence of the rice homologue (*REH1)* cDNA (SEQ ID NO.: 13) (GenBank accession number AF056027).

Figure 9 is the amino acid sequence of the rice homologue (REH1) protein (SEQ ID NO. 2).

Figures 10 A and 10 B are graphs demonstrating auxin transport activity in yeast strains *gef*1 and *gef1 EIR1*. The graphs summarize the amount (in percent) of ^{14}C-IAA

remaining in yeast cell samples taken at different points from cells maintained under various assay conditions. The amount of total radioactivity incorporated by the cells was determined in a sample of cells prior to their introduction into the assay. Bars indicate standard deviations derived from 3 parallel samples. Each experiment was performed at least four times. Figure 10 A shows a lack of auxin transport in *gef1* cells assayed in the presence of an external carbon source (2% glucose) in the efflux buffer. Figure 10 B shows the results of an assay performed in the absence of an external carbon source; auxin transport under these conditions depends exclusively on the pre-established membrane potential. Fig. 10 B (*gef1*; *gef1 EIR1*) demonstrates that the expression of *EIR1* in *gef1* yeast results in the retention of about 10 to about 20 percent less ^{14}C-IAA within the cells. The gef1 + CCCP and gef1 EIR1 + CCCP 1 data (Fig. 10B) demonstrate that the inclusion of the protonophore CCCP in the efflux buffer eliminates auxin transport activity.

Figures 11 A and 11 B are graphs comparing the growth of yeast cells expressing either wild type EIR1 or one of three Ser97 negative alleles of EIR1. The conserved amino acid Ser97 of EIR1 was replaced with another amino acid residue, thereby producing three mutants: EIR1-S97G; EIR1-S97A and EIR1-S97E. Figure 11 A shows the growth curve of *gef1* transformed with either EIR1 or one of the Ser97 mutants in Synthetic Complete medium (SC) . Figure 11 B shows the growth curve of either EIR1 or one of the Ser97 mutants in SC supplemented with 200μM 5-fluoro-indole.

DETAILED DESCRIPTION OF THE INVENTION

Described herein is the isolation and characterization of *EIR1*, a plant gene whose function is required for gravitropism. Genetic and physiological analyses of the *EIR1* gene and *eir1* mutants (*eir1-1* , *wav6-52* and *eir1-3)* support a role for *EIR1* involvement in root-specific auxin transport (efflux). Furthermore, the data provided herein indicate that EIR1 protein, which functions as a root-specific auxin efflux carrier, is a target for the regulation of auxin transport. These findings provide molecular evidence for the critical role of auxin transport in gravitropism and provide important targets and reagents useful for elucidating the role of gene expression in gravitropic signal transduction and the molecular mechanism of polar auxin transport.

The present invention relates to an isolated root-specific protein involved in auxin transport, isolated nucleic acid (e.g., DNA, RNA), for example, DNA encoding the protein, mutants of the DNA and altered forms of the encoded root-specific protein, and uses for the proteins and encoding DNA. In a particular embodiment, root-specific DNA designated *EIR1* and modified EIR1 nucleic acids have been isolated and characterized. The EIR1 protein is required for gravitropism and is involved in root-specific auxin transport. In addition, data presented herein supports the role of the EIR1 protein, which functions as an efflux carrier, as a target for regulation of auxin transport by ethylene and synthetic transport inhibitors. Genomic *EIR1* DNA and *EIR1* cDNA nucleotide sequences and the encoded EIR1 protein (amino acid) sequence are presented, as are the nucleotide sequence and amino acid sequences of a rice homologue, designated *REH1* (for: Rice *EIR1* Homologue) and REH1, respectively.

As used herein, the term "DNA encoding a root-specific protein involved in auxin transport" encompasses such DNA from any and all plant types (e.g., mustard plants, corn, rice, wheat and other grains or grasses, other crop plants, flowering plants). Isolated DNA which is the subject of the invention encodes a protein which is involved in root-specific transport, such as EIR1-protein encoding DNA. For example, DNA encoding a protein involved in root-specific auxin transport includes: (a) the sequences presented herein (SEQ ID NOS.: 11-13) and portions of any of those sequences, provided that they encode a functional root-specific auxin transport carrier protein; (b) DNA which, due to degeneracy of the genetic code, encodes EIR1 protein of the present invention (e.g., EIR1 protein having the amino acid sequence of SEQ ID NOS.: 1, 2, 5 or 6); (c) DNA which hybridizes under high stringency conditions to the complement of any DNA of (a) or (b) and; (d) DNA which is from Arabidopsis or from a plant species other than Arabidopsis which is sufficiently similar in sequence to DNA of (a), (b) or (c) to encode a root-specific protein involved in auxin transport (e.g., as demonstrated by the assay described herein). Homologous DNA can be identified by substantial nucleic acid sequence homology to an EIR1 nucleic acid. For example, homologous DNA can be identified based upon overall nucleic acid sequence homology with the EIR1 DNA sequence disclosed herein, allowing for the degeneracy of the genetic code and codon bias in different species of plants, and on the requirement that homologous sequences encode a functional root-specific auxin transport (efflux)

carrier protein. For example, the overall homology of the nucleotide sequence is preferably greater than about 40%, preferably greater than 60% , still more preferably greater than about 80% and most preferably greater than 90% homologous. Thus, the invention also comprises the use of the disclosed nucleic acid sequences, or portions thereof, as probes and primers for the identification and isolation of homologous sequences from other species of plants.

DNA of the present invention also includes coding or noncoding DNA which is the complement of any of the DNA of (a) - (d) and portions (or fragments) thereof which are of sufficient length (e.g., at least four to six nucleotides) to hybridize to complementary DNA and remain hybridized (e.g., in order that hybridization can be detected, such as for diagnostic or assay purposes). Such fragments also include those which hybridize to characteristic portions of the DNA of the present invention (e.g., to a characteristic portion of DNA of SEQ ID NOS.: 11, 12 or 13). The complement of DNA encoding a root-specific protein of the present invention is also a subject of this invention. For example, DNA complementary to all or a portion of EIR1 protein encoding DNA, such as DNA of SEQ ID NO.11 or SEQ ID NO. 12, is the subject of this invention. Such complementary DNA is useful as probes and primers, for example, in hybridization and amplification (e.g., PCR) reactions.

As used herein, the term "modified EIR1 nucleic acid" refers to a variant EIR1 nucleic acid molecule which includes addition, substitution, insertion or deletion of one or more nucleotide(s), thereby producing a modified nucleotide sequence. As used herein, the term "nucleic acid" encompasses DNA (genomic and cDNA), RNA and analogues (e.g., comprising base analogues such as inosine) thereof. The "modified EIR1 nucleic acid" can embody either a naturally occurring allelic variant or a synthetically produced sequence. For example, the disclosed naturally occurring (e.g., wild type) nucleic acid isolated from *Arabidopsis thaliana* can be used as a precursor nucleic acid molecule which can be modified by standard techniques that are well-known to those of skill in the art to produce a synthetic variant. For example, site-directed mutagenesis or cassette-mutagenesis can be used to substitute one or more nucleotides.

Promoters and other regulatory sequences (e.g., cis acting elements and/or transcriptional enhancers) of DNA encoding a root-specific auxin transport protein are

also the subject of this invention, as are their use in vectors and expression systems designed to direct the tissue-preferential transcription of foreign (e.g., heterologous) genes operably linked thereto, in the roots of plants.

The isolate nucleic acid which is the subject of the invention can be obtained from a plant as it occurs in nature, or can be produced by synthetic (e.g., chemical) methods or recombinant methods. Also included herein are mutant genes, such as the mutant gene designated *eir1*-3 which is present in an agravitropic mutation. This mutated gene and the agravitropic mutation are useful to study the pathway of which *EIR1* is a component.

The isolated root-specific proteins involved in auxin transport and allelic variants thereof which are the subject of the invention include the encoded protein products of the DNA sequences disclosed herein and functional portions and fragments thereof. In particular embodiments the invention comprises proteins having the amino acid sequence comprising SEQ ID NOS.: 1 and 2.

Genetically engineered plants (e.g., transgenic, transformed plants expressing heterologous DNA episomally, transiently, or stably integrated into plant nuclear DNA), plant tissues and seeds characterized by an increased resistance (or tolerance) to the effects of herbicides which are auxin derivatives, auxin analogues, or an herbicidal formulation comprising at least one auxin transport inhibitor applied in combination with at least one additional herbicide, relative to the corresponding wild type plants are also the subject of this invention. More specifically, the invention relates to plants, plant tissues , and seeds which are resistant to growth inhibition by an herbicide (which is an auxin derivative or an auxin analogue), or an herbicidal composition (which includes an auxin, analogue derivative, auxin analogue or auxin transport inhibitor), at concentrations which normally inhibit the growth of those plants, plant tissues or seeds. In one embodiment, the present invention relates to a method of producing a transgenic plant characterized by altered auxin homeostasis. The method comprises introducing DNA encoding a root-specific auxin transport carrier protein into a plant cell under conditions in which the DNA is expressed, thereby producing a transformed plant cell; and producing a transgenic plant from the resulting transformed cell. Transgenic (genetically engineered) plants can be produced using DNA described herein and methods known to those of skill in the art. For example, DNA encoding a

root-specific auxin transport protein can be introduced into plants or plant tissues (e.g., roots) or seeds by transformation (e.g., transfection or transduction) using Ti plasmids, Ri plasmids, plant virus vectors, direct DNA transformation, microinjection, protoplast fusion, electroporation or bombardment (e.g., microprojectile bombardment) with nucleic acid-coated particles.

As used herein, the term "herbicide" refers to compounds which combat or control undesired plant growth. The term "auxin transport inhibitor" refers to compounds which act by inhibiting the transmembrane movement (e.g., transport) of auxin which accumulates in cells as a result of polar auxin transport and affects plant growth. Thus, as used herein, auxin transport inhibitors are themselves herbicides. The observation that auxin transport inhibitors are usually highly active herbicides is consistent with this usage. As used herein, the terms "resistance" and "tolerance" refer to the sensitivity of a plant to the toxic effects of an herbicide, such that a genetically engineered plant, whose genome comprises a nucleotide sequence encoding a root-specific heterologous auxin transport carrier protein is resistant to an herbicide. Genetically engineered plants (transgenic plants) of the present invention include, but are not limited to, vascular plants, including gymnosperms and agronomically important plant crops, such as rice, wheat, barley, rye, corn, soybeans, canola, sunflower, sorghum, sugarcane, fruits (oranges, grapefruit, lemons, limes, apples, pears, melons, plums, cherries, peaches, apricots, strawberries, grapes, raspberries, pineapples, bananas), vegetables (potatoes, carrots, sweet potatoes, beans, peas, lettuce, cabbage, cauliflower, broccoli, turnip, radishes, spinach, onions, garlic, peppers, pumpkins) and angiosperms or flowering plants, both monocots and dicots.

In one embodiment, plants with greater resistance are genetically engineered plants whose root cells comprise heterologous DNA which encodes a protein involved in auxin transport (e.g., *EIR1* DNA, *REH1* DNA) which is expressed as a functional root-specific auxin transport (e.g., efflux) protein. The corresponding wild type plant differs from the genetically engineered plant in that the wild type plant has not been altered to comprise the heterologous DNA present in the genetically engineered plant. In one embodiment, the heterologous DNA which encodes an auxin specific efflux carrier protein is constitutively expressed in a tissue-specific (e.g., root tissue) fashion and the expression trait and resulting phenotype is stably transmitted (sexually and

somatically) to progeny cells. In a second embodiment, the invention comprises transgenic plants, the cells of which comprise heterologous DNA stably integrated into the plant nuclear DNA. In an alternative embodiment, the expression of the heterologous DNA encoding an auxin specific efflux carrier is inducible.

In a second embodiment transgenic plants characterized by an altered auxin homeostasis exhibit a distinctive phenotype, attributed to increased auxin efflux, such as an increased number of lateral or adventitious roots. Such plants may also be further characterized by an increased auxin transport rate relative to the auxin transport rate of a corresponding wild type plant.

As used herein, the term "heterologous DNA" means DNA isolated from a source other than the plant, or plant cell, in which it is expressed (e.g., from a source other than the cell into which it is introduced or in which it is present as a result of having been introduced into a precursor cell, such as seeds or plant tissue from which a plant develops or seeds or plant tissue obtained from a genetically engineered plant). The heterologous DNA can be from the same plant type (e.g., *Arabidopsis* DNA introduced into *Arabidopsis*) or from a different plant type (e.g., *Arabidopsis* DNA introduced into corn, wheat, rice or other plant type, rice DNA introduced into corn, wheat or other plant type). Heterologous DNA can be used, for example, to avoid or reduce the silencing or inactivation to which the endogenous gene or its encoded protein (e.g., post-translational modification) can be subjected. As a result of the presence and expression of heterologous DNA encoding a root-specific auxin transport protein in roots of a genetically engineered plant, auxin transport (efflux) is enhanced and the plant exhibits enhanced resistance to auxin derivatives auxin analogues or formulations comprising an auxin transport inhibitor in combination with a second herbicide. For example, plants (e.g., crop plants, flowering plants, gymnosperms) which are genetically engineered to include or are produced from seeds, plant tissues, or plant parts which include *EIR1* or *REH1* DNA can be produced to provide genetically engineered plants with enhanced herbicide resistance. Plant part is meant to include any portion of a plant from which a regenerated plant can be produced.

Plants which show increased auxin transport and/or enhanced root tissue growth and/or differentiation (compared to the corresponding wild type plants) resulting from altered auxin homeostasis are also the subject of this invention. More specifically, the

invention also comprises genetically engineered plants comprising a heterologous DNA sequence encoding a root-specific protein involved in auxin transport, wherein the genetically engineered plant exhibits a distinctive phenotype, relative to the phenotype of an isogenic plant which does not comprise a heterologous DNA encoding a protein involved in root-specific auxin transport, attributed to the effects of altered auxin homeostasis. For example, transgenic plants characterized by a phenotype comprising an increased number of lateral or adventitious roots.

Also described are alleles of EIR1, in which the conserved residue Ser97 of EIR1 is replaced with another amino acid. Three alleles, EIR1-S97G, EIR1-S97A and EIR1-S97E, were created and characterized, as described in Example 10. These alleles were expressed in diploid yeast strains, defective for the *gef1* gene, under the control of the ADH-promoter. The strains were tested in a filter assay carried out with either 5-fluoro-indole or 5-fluoro-indole acetic acid. The strains exhibited a hypersensitivity to these compounds.

Also described herein is an assay for assessing agents (compounds and molecules) for their effects on auxin transport. As described in the examples, an assay is available in which auxin transport is assessed in yeast by measuring transport of detectably labeled (e.g., radiolabeled) auxin. This assay is useful to determine whether an agent inhibits or enhances the activity of EIR1 protein and, as a result, inhibits or enhances auxin transport. The auxin transport assay can be used for example to characterize EIR1 alleles identified by their ability to confer an altered growth phenotype. For example, one would expect to find an increased auxin transport rate associated with an allele which confers significantly increased resistance of *gef1* yeast cells to fluoroindolics. The yeast cell-based overexpression model disclosed herein provides a functional assay useful for assessing structure/function relationships in isolated DNA molecules and mutated EIR1 sequences encoding auxin transport proteins and their variants. In adition the yeast cell-based overexpression model can be used to identify an allele (mutant) of EIR1 which confers altered auxin-mediated responses in a plant. Briefly the overexpression assay comprises: introducing a mutated EIR1 nucleic acid into yeast cells, thereby producing transformed yeast cells; contacting the transformed yeast cells with a fluorinated indolic compound under assay conditions which favor the diffusion of the compound into the yeast cells; determining the growth phenotype of the cells;

and comparing the growth phenotype of the transformed cells to the growth phenotype of wild type cells, wherein detection of an altered growth phenotype in the transformed cells relative to the growth phenotype of wild type cells is indicative of a nucleic acid which is an allele that results in altered auxin-mediated responses in a plant. The altered growth phenotype observed in the overexpression assay can be either an increased tolerance or an increased sensitivity to concentrations of the fluorinated indolic compounds, relative to the sensitivity of wild type cells. Diploid yeast cells which are defective for the GEF1 gene, and therefore have an altered ion hemostasis are particularly useful for the establishment of an overexpression assay. The overexpression assay is useful, for example, to identify mutant nucleotide sequences, produced by random mutagenesis of wild-type DNA sequences encoding auxin transport proteins which exhibit altered growth phenotypes (either enhanced or decreased sensitivity) to fluorinated indolic compounds. Yeast strains exhibiting altered growth phenotypes (tolerance or increased sensitivity) comprise mutated DNA sequences which upon introduction into a transgenic plant will alter auxin homestasis and auxin-mediated responses such as growth, morphogenesis (lateral or adventitious root formation) and tropisms (gravitropism). The present invention also comprises transgenic plants comprising mutant *EIR1* alleles identified in the yeast cell-based overexpression assay.

The sequences (nucleotide and amino acid) and topology of EIR1, its homology to several bacterial carrier proteins and its function establish that EIR1 functions as a root-specific auxin transport (efflux) carrier protein involved in gravitropism

The present invention is illustrated by the following examples, which are not intended to be limiting in any way. Further, all references referred to herein are expressly incorporated by reference.

EXAMPLES

METHODS AND MATERIALS

The following methods and materials were used in the work described herein, particularly in the examples:

Plant Strains and Growth Conditions

Plants were grown aseptically on unsupplemented PNA (Plant Nutrient Agar) without sucrose. (Haughn, G.W. and C. Somerville, (1986) *Mol. Gen. Genet. 204*: 430-434). Growth responses were tested by adding various supplements to the medium as indicated. Plates were wrapped in gas-permeable surgical tape and kept under continuos illumination. For gravitropic response experiments, plates were kept in a vertical position. For the "root waving assay" plates were kept at an angle of about 30 degrees. Root elongation was assayed at 10-12 days after germination (DAG). Formation of lateral roots was compared by counting lateral roots on both wild type and mutant plants grown under conditions described herein.

Seed stocks for *eir1-1* and *eto3-1* were obtained from the Arabidopsis Biological Resource Center at OSU, Columbus, OH). *ctr1-1* was a kind gift from J. Hua at Caltech, Pasadena, CA. *agr1-52* was obtained from K. Okada, National Institute for Basic Biology, Okazaki, Japan. *PIG4*::GUS was a kind gift from J. Normanly, University of Massachusetts, Amherst, MA. Transposon line B222 was obtained from DNA Plant Technology Corporation, Oakland, CA.

Inverse PCR Cloning and Structural Analysis of *EIR1*

Genomic DNA was prepared according to a protocol from Quiagen. After grinding the frozen tissue, the resulting powder was incubated at 74°C for 20 minutes in lysis buffer (100 mM Tris/HCl pH 9.5, 1.4 M NaCl, 0.02 M EDTA, 2% CTAB, 1% PEG 8000). After extraction with an equal amount of chloroform, DNA was precipitated with isopropanol. After resuspension in 1 M NaCl and treatment with RNase A, the DNA was loaded onto equilibrated Quiagen columns and purified according to the manufacturer's instructions. DNA extracted from the Ac line *B222* and *eir1-3* was digested with *EcoR*I and *Bcl*I. The ends of the DNA were made blunt with Klenow fragment. This DNA was religated and used for inverse PCR performed with oligonucleotides CCTCGGGTTCGAAATCG (SEQ ID NO.: 14) and GGGGAAGAACTAATGAAGTGTG (SEQ ID NO.: 15). After 40 cycles of amplification at 60°C annealing temperature, the products were separated on 1% agarose gels. A fragment specific for *eir1-3* DNA was cloned into pGEMT (Promega) to give pGsac1 and used for Southern hybridization on *eir1-3* and wild type DNA.

Phage genomic and cDNA libraries of *A. thaliana* (Kieber, J.J., *et al.*, (1993) *Cell 72*: 427-441) were probed with pGsac1 using standard techniques. (Ausubel, F.M., *et al.*, (1987) *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc). Genomic clone λ5-3, which hybridized to pGsac1, was subcloned into pBluescriptII (Stratagene) to give pB5-3. The sequence of an *Eco*RI fragment approximately 8kb in length was determined on an ABI Automated DNA sequencer. For sequence analysis of *eir1-1*, the coding region of this allele and its corresponding wild type (Col-O) were amplified with PCR. The point mutation in *eir1-1* was confirmed by subsequent PCR amplification of sequences covering the mutation.

Two full-length cDNA clones subcloned into pBSII (pBc5-2 and pBc6-1) were completely sequenced. The rice EST (D25054) homologous to *EIR1* was obtained from MAFF DNA Bank at the National Institute of Agrobiological Resources (NIAR), Japan.

Sequence comparisons with Database entries were performed using Gapped BLAST and PSI-BLAST algorithms. (Altschul, S.F., *et al.*, (1997) *Nucleic Acids Res. 25*: 3389-40.2) Multiple alignments and structural predictions were performed using the algorithms at BCM Search Launcher.

Complementation of *eir1-1* in Transgenic Plants

An *Eco*RI fragment of the genomic clone pB5-3, which carries the entire coding region and more than 2kb of upstream sequences of the *EIR1* gene was subcloned into pBIBhyg (Becker, D., (1990) *Nucl. Acids Res. 18*: 203). The resulting T-DNA vector pBRL was transformed into *Agrobacterium tumefaciens* strain GV3101 via electroporation, and used for subsequent vacuum infiltration of *eir1-1* plants. (Bechtold, N., *et al.*, (1993) *C.R. Acad. Sci. Paris, Sciences de la vie/Life sciences 316*: 1194-1199) Correct integration of the full-length transgene was confirmed on Southern blots.

RNA Template-Specific Polymerase Chain Reaction (RS-PCR)

For expression analysis, total RNA from tissue of sterile grown plants was isolated. (Niyogi, K.K. and G.R. Fink, (1992) *Plant Cell. 4*: 721-33) Vegetative tissue isolated

from plants 15 DAG was used. Flower-specific RNA was isolated at approximately 20 DAG and silique-specific RNA at about 25 DAG. polyA$^+$ RNA was isolated with the polyATract kit from Promega. About 50 ng of polyA$^+$ RNA of each tissue was used for RNA Template-Specific PCR (RS-PCR).

RS-PCR with slight modifications was performed as described by (Shuldiner, A.R., *et al.*, (1993) *In: Methods in Molecular Biology: PCR Protocols: Current Methods and Applications* Human Press Inc. Totowa, NJ). Oligonucleotides GAACATCGATGACCAAGCTTAGGTATCGATAGCCCCACGGAACTCAAA (SEQ ID NO.: 16) (underlined bases are complementary to nucleotides 454 to 470 of the *EIR1* coding region) and CTTATACGGATATCCTGGCAATTCGGACTTGTTAG CTTTAGGGTTAA (SEQ ID NO.: 17 (underlined bases are complementary to nucleotides 335 to 351 of *ACT2* coding region) were added to polyA$^+$ RNA to a final concentration of 2 μM in a volume of 10 μl. The tubes were placed at 65°C for 10 minutes and allowed to cool down to 37°C. First strand cDNA synthesis was performed using Gibco BRL AMV Reverse Transcriptase. Primer pairs GAACATCGATGACC AAGCTTAGGTATCGATA (SEQ ID NO.: 18) and GGCAAAGACATGTACGATGT TTTAGCGG (SEQ ID NO.: 19) (bases 10 to 37 of *EIR1* coding region) or CTTATACGGATATCCTGGCAATTCGGACTT (SEQ ID NO.: 20) and GTCTGTGACAATGGAACTGGAATG (SEQ ID NO.: 21) (bases 31 to 54 of *ACT2* coding region) were used in a standard PCR for 30 cycles with 40 seconds at 94°C, 40 seconds at 60°C and 1 minute at 72°C. For *EIR1*, 1/100 of this reaction was used for reamplification under the same conditions. ^{32}P-end labeled oligonucleotide GTGAAAAGAGCGTTAT CATCCATTCTAG (SEQ ID NO.: 22) (complementary to bases 292 to 319 of *EIR1* coding region) allowed verification of the identity of the *EIR1*-specific band on a Southern blot.

Preparation and Microscopic Analysis of Roots

Whole plants were incubated twice in methanol:glacial acetic acid (3:1) and rinsed several times in PBT (130 mM NaCl, 10 mM sodium phosphate pH 7.0, 0.1% Tween 20). Roots were then mounted onto microscope slides into clearing solution (stock solution: 8 g chloralhydrate in 2.5 ml 20% glycerol). After 10 minutes, roots were viewed under a Zeiss microscope using Nomarski Optics. Dark field photographs of

live plants were made using a Wild M5-A microscope. These images were used for determination of root cell length. GUS stainings were performed as described. (Lehman, A., *et al.*, (1996) *Cell. 85*: 183-94) Images were recorded on Kodak Ektachrome 160T film and processed using Adobe Photoshop.

Complementation Analysis and Construction of Double Mutants

For complementation analysis of the three putative *eir1* alleles, *eir1-3/eir1-3* plants (*eir1-3* still contains the Ac-donor T-DNA-construct conferring hygromycin resistance) were crossed into plants homozygous for either *eir1-1* or *wav6-52*.

Heterozygous F1 plants (*eir1-3/wav6-52 and eir1-3/eir1-1*) identified as resistant to hygromycin were defective in root gravitropism, giving evidence for the allelism of the three mutants analyzed. F2 plants derived from each of the F1 heterozygotes were all Eir1 $^{-}$ whereas the hygromycin resistance marker segregated as a single, dominant locus. Double mutants (e.g. *ein2-1/ein2-1 eir1-1/eir1-1*) were derived from crosses of homozygous single mutant lines and scored for segregation in the F2 generation of the initial crosses. Double mutant candidates were backcrossed into their two parental single mutant lines and their genotype verified by complementation with parental testers. For *eir1-1 alf1-1* double mutants, we used *eir1-1/eir1-1* plants for pollination of *alf1-1/ALF1* heterozygotes. F2 seeds were scored for segregation of Eir1 $^{-}$ and Alf1 $^{-}$ phenotypes. The double mutant was verified by segregation of the aerial Alf1 $^{-}$ phenotype in Eir1 $^{-}$ F3 plants derived from the initial cross.

Auxin Transport Assay

Yeast strains transformed with pAD-EI and pAD4M (Luschnig et al., 1998) were grown to an O.D. 600 of 0.8 to 0.9. Cells were pelleted and an aliquot corresponding to 15ml starting culture was washed in 10mM Na Citrate buffer pH 4.5. The pellet was resuspended in 1ml of 10mM Na Citrate (pH 4.5) supplemented with 1mM IAA (final concentration) and 2.5 micro Ci ^{14}C-IAA (Sigma).

The cells were allowed to incorporate the tracer for 10 or 20 minutes. The cells were subsequently washed on MF-filters (Millipore) on a multifiltration unit, and

resuspended in Synthetic Complete (SC)-medium adjusted to pH 4.0 with HCl. Aliquots of the suspension were dropped onto MF-filters and washed twice with SC-medium
(pH 4.0). The dry filters were transferred into Scintilation Cocktail and radioactivity was determined in a Scintilation Counter. Each experiment was performed for at least 4 times. The radioactivity remaining in the cells is expressed as percentage of total radioactivity present in the washed pellet prior to the efflux assay. Each time point was determined by 3 parallel samples. For the experiments performed in the presence of glucose, 2% (w/v) glucose was added to the efflux buffer. Similarly, for assays performed in the presence of the protonophor CCCP 0.5mM (final concentration) of CCCP were added to the efflux buffer prior to the efflux experiment.

Yeast Manipulations and Constructs

All experiments were carried out in W303 (a/α *ura3-1 can1-100 leu2-3, 112 trp1-1 his3-11, 15*). Plasmid pRG52 was used for disruption of *GEF1*, For analysis of *CLC-0* the vector PRS1024 was used (for more details, see Gaxiola, R.A., *et al.*, (1998) *Proc. Natl. Acad. Sci. USA 95*: 4046-4050).

Yeast strains were grown over night at 30 °C in Synthetic Complete medium (SC) and approximately $2x10^6$ cells were plated onto SC plates. Solutions of inhibitors used in the filter growth assays were spotted onto Schleicher & Schuell Filter Paper #740. After they dried, the filters were transferred onto the yeast plates, which then were incubated at 25%C in the dark for two to five days. After that, yeast growth was monitored and documented.

For expression of *EIR1* in *S cerevisiae* the insert of pBc5-2 was cloned into pAD4M (described in Ballester *et al.*, (1989) *Cell*, *59*: 681-686) to give pAD-E1. A frameshift mutation in *EIR1* was obtained by filling in the internal *Hin*dIII site resulting in a nonsense mutation after codon 178 (plasmid pADE1-H). For construction of the HA-tagged version of *EIR1*, we used primers
GGGTCTAGAGTACTCTACTACGTTCTTTTGGGGCTTT
ACCCATACGATGGTCCTGAC (SEQ ID NO.: 23) and
GGGTCTAGAGTCGACGCA CTGAGCAGCGTAAT (SEQ ID NO.: 24) for PCR

amplification of a fragment encoding 3 copies of the HA-epitope. The PCR product was ligated into pAD-E1 resulting in pAD-E1HA coding for a protein with the 3xHA-tag fused to the authentic C-terminus of EIR1. Immunostaining of the tagged protein in haploid and diploid cells was performed as described by Gaxiola, R.A., *et al.*, (1998) *Proc. Natl. Acad. Sci. USA 95*: 4046-4050. Cells were viewed by using charge-coupled device microscopy and sectioned by using SCANALYTICS (Billerica, MA).

EXAMPLE 1

Isolation and phenotypic characterization of *eir1-3*

An agravitropic mutant (e.g., a plant whose roots do not respond to gravistimulation) was isolated from the Ac transposon pool B222-24 (Keller, J., *et al.*, (1992) *Genetics, 131*: 449-59). This agravitropism segregated as if it resulted from a mutation in a single gene. A comparison of DNA isolated from the mutant transposon-tagged line B222-24 with the untransposed parental line B222 on Southern blots revealed that the mutant contained an additional copy of the transposon. This extra Ac element cosegregated with the mutant phenotype, suggesting that the mutation, designated *eir1-3* was caused by the insertion of the transposon element.

This agravitropic mutation, *eir1-3*, is allelic to two previously described mutations, *wav6-52* (allelic with *agr1*), which was isolated as an agravitrophic mutant (Bell, C.J. and P.E. Maher, (1990) *Mol. Gen. Genet. 220*:289-293) and *eir1-1* , which was isolated as an ethylene insensitive mutant (Roman, G., *et al.*, (1995) *Genetics. 139*: 1393-1409). The new mutation, *eir1-3*, fails to complement *wav6-52*, and *eir1-1* showing that all three are alleles of *EIR1*. All three mutants have similar phenotypes with the severity of the mutant phenotype in the order *eir1-3* = *eir1-1* > *wav6-52*.

eir1 mutant roots do not respond to gravity when germinated and grown on agar plates oriented vertically. Instead, *eir1* roots grow in random directions, whereas *EIR1* roots grow downward. If the seedlings are reoriented so that the roots are now parallel to the surface of the earth, after 24 hours, the roots of wild type reorient downward (roughly 90%), whereas roots of *eir1* fail to reorient their growth.

These severe defects in gravitropism appear to be restricted to the root, as the hypocotyl in all three *eir1* mutant strains tested, still reorients when germinated in the dark. In another assay, seedlings were kept on 2% agar plates that were tilted vertically at an angle of less than 90°. Under these conditions, *EIR1* roots do not penetrate the agar but grow on the surface in a wavy pattern, that is caused by reversible turns of the root tip. (Okada, K. and Y. Shimura (1990) *Science 250*: 274-276) By contrast, *eir1* roots exhibit a roughly linear growth pattern interrupted by random turns. When wild type seeds are germinated on plates whose surface is parallel to the surface of the earth, they enter the agar and form a characteristic array of almost concentric curls. (Garbers, C., *et al.*, (1996) *Arabidopsis. EMBO J. 15*: 2215-2124) However, *eir1* mutant roots failed to curl on the bottom of the plate and grew out in irregular patterns.

Root growth of *eir1* mutant plants is less sensitive to ethylene than that of the wild type, suggesting an involvement of ethylene in the regulation of root tropic responses. *eir1* roots have a phenotype that is similar to *EIR1* roots grown in the presence of NPA and TIBA, inhibitors of auxin transport that block cell elongation (Sussman, M.R. and M.H.M. Goldsmith, (1981) *Planta 152*: 13-18). Moreover, *eir1* root elongation was much more resistant than *EIR1* to NPA and TIBA (Figures 1A - 1C). By contrast, these auxin transport inhibitors inhibit lateral root formation to the same extent in both wild type and *eir1* mutants. Also, *eir1* root growth is more resistant than wild type to 1-aminocyclopropane-1-carboxylic acid (ACC), the immediate biosynthetic precursor of the growth regulator ethylene (Figure 1A). However, the root growth inhibition of *eir1* mutants is no different from *EIR1* with respect to other growth regulators (abscissic acid, gibberellic acid, kinetin), the auxin-analogue NAA (-napthaleneacetic acid) (Figure 1C), and 2,4-D (2,4-dichloro-phenoxyacetic acid).

The *eir1* mutants have longer roots than wild type plants (Table 1), which could be due to an increased rate of cell division and/or to greater elongation of individual root cells. Direct measurement showed that *eir1-3* root cells were longer than wild type cells (Table 1). However, it is possible that increased cell division contributes to the increased length as well.

Table 1 Root Growth and Cell Elongation

Strain	average root length[a]	average cell length[b]
Ws	79 ± 7	102.9 ± 12.4
eir1-3	97 ± 11	135.9 ± 15.3

[a] Length of primary roots was determined at approximately 12 DAG;

[b] Elongation of 35-40 young trichoblasts was determined on images. Root lengths are indicated in mm, cell length in µm.

EXAMPLE 2

Cloning of *EIR1*

The *eir1-3* allele was cloned using an inverse Polymerase Chain Reaction (PCR) approach. A 600 bp fragment amplified from *eir1-3* DNA hybridized to the additional band caused by the Ac transposon element insert in *eir1-3*. This subcloned fragment was used to screen an *A. thaliana* genomic phage library. Three genomic clones of the putative *EIR1* gene (λ5-3, λ6-1 and λ6-3) had the same restriction pattern. The subcloned insert of λ5-3 was used for screening cDNA libraries. Eight hybridizing phage clones were isolated from approximately $5x10^5$ plaques screened. These clones all show similar restriction patterns. Two inserts of approximately 2.2 kb were completely sequenced. The largest cDNA clone contained a continuous Open Reading Frame (ORF) starting 29 bp downstream of its

5' end. Comparison of the cDNA with the genomic clone revealed that the ORF is split into 9 exons coding for a predicted protein of 69.3 kDa.

The Ac insertion in *eir1-3* is located after codon 113 in exon 2 (Figure 2). The insertion is flanked by a perfect 8 bp direct repeat and probably results in a null allele of the affected gene. Results showed that *eir1-1* (as compared with the progenitor Columbia wild type) contains a transition mutation at the intron 5/exon 6 border that replaces the absolutely conserved G at splice position -1. (Brown, J.W.S., (1996) *Plant J. 10*: 771-780) The *eir1-1* mutation presumably results in a truncated EIR1 protein that would lack a conserved portion of the molecule (Figure 2).

To determine whether the cloned segment was the *EIR1* gene, *eir1-1* was transformed with the putative *EIR1* ORF and more than 2kb of upstream sequences. All five independent hygromycin-resistant transformants of *eir1-1* tested had a root growth phenotype typical of wild type. Therefore, the defects of the *eir1-1* mutant were complemented by the genomic fragment. No other large ORFs were present on the genomic fragment used in the transformation. Therefore, the open reading frame has been designated as the coding region of *EIR1*.

Isolation of *eir1-3* a new transposon-tagged allele of *EIR1*, permitted the cloning and characterization of both the mutant and wild type genes. Sequence analysis shows that *eir1-3* is an Ac insertion in the second of nine exons and *eir1-1* is a base substitution at a conserved splice site junction. Both of these mutations are likely to be null alleles because they should result in completely defective proteins. Expression of *EIR1* appears to be restricted to the root, which is consistent with the finding that all of the *eir1* mutant phenotypes, the most striking of which is gravitropism, affect the root and not other parts of the plant.

The amino acid sequence of *EIR1* is consistent with a role for this protein in transport of IAA. EIR1 is predicted to be an integral membrane protein. The presence of potential N-glycosylation sites and a potential N-terminal signal peptide indicates localization in the plasma membrane. EIR1 also has similarities to several membrane proteins involved in translocation of a variety of different substances across the plasma membrane. The transporters related to EIR1 are diverse in their substrate specificity and translocate amino acids, heavy metals, antibiotics, and dicarboxylic acids.

Perhaps the most compelling evidence that EIR1 plays a role in transport is that expression of *EIR1* in *S. cerevisiae* confers increased resistance to fluorinated analogues of indolic compounds. The resistance phenotypes are strongest in the *gef1* mutant, which has increased sensitivity to various compounds probably as a result of altered ion homeostasis (Gaxiola, R.A., *et al.*, (1998) *Proc. Natl. Acad. Sci. USA 95*: 4046-4050). Resistance to these indoles is completely dependent upon a functional *EIR1* gene product as neither *ClC-O* nor a mutated version of *EIR1* were capable of restoring yeast growth in the presence of fluorinated indolic compounds.

The EIR1 protein could prevent the inhibition of yeast by these compounds either by preventing their uptake or facilitating their efflux from the cytosol. The preferential localization of EIR1 in the plasma membrane of yeast is consistent with either of these mechanisms.

EXAMPLE 3

EIR1, a Highly Conserved Plant Gene Family with Similarities to Bacterial Transporters

Several lines of evidence suggest that *EIR1* belongs to a highly conserved gene family. Arabidopsis has several genes with considerable homology to *EIR*. In addition to several *Arabidopsis* ESTs (Genbank accession numbers: T04468, T43636, R84151, and Z38079), similar ORFs were found in database entries of the *Arabidopsis* Genome Initiative. Two close relatives dubbed *AEH1* and *AEH2* (for *Arabidopsis EIR1* Homologue) were located on clones T26J12 and MKQ4 on chromosome 1 and 5 respectively. These relatives probably account for the extra restriction fragments that hybridize to the EIR1 probe under conditions of high stringency. A related rice EST (accession number: D25054), which is derived from root-specific cDNA, was identified and sequenced (Figure 3). No other closely related sequences could be found outside the plant kingdom, suggesting that *EIR1* and its homologues represent a family of genes unique to higher plants.

Alignment of the deduced amino acid sequences of *EIR1*, *AEH1*, *AEH2*, and *REH1* (Rice *EIR1* Homologue) revealed that the regions of identity are restricted to the

N- and C-termini (Figure 3). Hydropathy plots and topology predictions identified 10 potential transmembrane domains shared by the 4 members of the gene family. The predicted EIR1 gene product comprises 647 amino acids and includes ten potential transmembrane domains flanking a central region enriched for hydrophilic amino acids. The predicted REH1 gene product comprises 595 amino acids and similarly comprises ten potential transmembrane domains flanking a central region which is predominantly hydrophilic. The transmembrane domains are located in the highly conserved portions of the proteins -- 5 at the N-terminus and 5 at the C-terminus (Figure 4).

The internal segments of the protein, though less conserved in sequence than the putative membrane spanning domains, exhibits a number of similarities. The central hydrophilic segments have a remarkably high content of serine and proline. EIR1 possesses a number of potential N-glycosylation sites, two of which are also found in REH1 and AEH1 (Figure 3). EIR1 has no ER-retention signal but does have a potential N-terminal signal peptide (von Heijne, G., (1986) *Nucleic Acids Res.14*: 4683-4690), which likely allows the protein to transit the secretory pathway to the plasma membrane. The open reading frame (ORF) of the EIR1 cDNA (SEQ ID NO.: 12) comprises nucleotides 19-1962 (Figure 6); the ORF of the REH1 cDNA (SEQ ID NO.: 13) comprises nucleotides 158-1945 (Figure 8).

The two hydrophobic portions of EIR1 show restricted similarity to a number of bacterial membrane proteins (Figure 4). The *mdcF* (U95087) protein is a potential malonate transporter from *Klebsiella pneumoniae* (Hoenke, S. *et al.*, (1997) *Eur. J. Biochem. 246*: 530-538), whereas *livM* (P22729) is involved in high affinity uptake of leucine into *Escherichia coli* (Adams, M.D., *et al.*, (1990) *J. Biol. Chem. 265*: 11436-11443).

Particularly noteworthy is the similarity of EIR1 to the class of efflux carriers that remove toxic compounds from the interior of the cell. For example, *E.coli arsB* (P37310) represents a part of the arsenic efflux system. (Diorio, C., *et al.*, (1995) *J. Bacteriol. 177*: 2050-2056). *sbmA* (X54153), another integral membrane protein of *E. coli*, has been shown to be necessary for uptake of the antibiotic Microcin 25 (Salomon, R.A. and R.N. Farias, (1995) *J. Bacteriol. 177*: 3323-3325). Portions of EIR1 show 35-40% similarity to these proteins. The finding that the N- and the C-

terminus of EIR1 exhibit similarities to the corresponding parts of bacterial transporters indicates that EIR1 is a membrane protein with a related function.

EXAMPLE 4

EIR1 Affects the Root-specific Response to Endogenous Ethylene

The reduced sensitivity of *eir1* roots to inhibition by ethylene suggested that *EIR1* might be a gene involved in regulation of ethylene responses specific to the root. In order to test this hypothesis, the response of the entire *eir1* mutant plant to endogenous ethylene was examined by constructing double mutants of *eir1* with *eto3* and *ctr1*. *eto3* causes overproduction of ethylene, giving rise to the typical triple response (the hypocotyl of plants germinated in the dark remains short, undergoes radial swelling and apical hook formation is exaggerated). Mutations in the Raf-like protein kinase *CTR1* phenocopy the ethylene-grown phenotype without elevating endogenous ethylene concentrations, suggesting that *CTR1* acts as a negative regulator of ethylene signal transduction (Kieber, J.J., *et al.*, (1993) *Cell 72*: 427-441).

The double mutants *eir1-3/eto3-1* and *eir1-3/ctr1-1* were germinated both in the dark and under constant illumination. Dark germinated plants still undergo the triple response, indicating that the *eir1* mutation has no influence on germination and early development of the aerial parts of the seedling . However, the inhibition of root elongation caused by *eto3* and *ctr1* mutations is considerably reduced in the double mutants.

These results suggest that reduced ethylene sensitivity of the *eir1* mutant is completely restricted to the root. Moreover, the phenotype is not caused by a block in biosynthesis or transport of ethylene because *eir1-3* bypasses the root phenotypes of *ctr1-1*, a mutation thought to be constitutive for the transduction of the ethylene signal.

Gravitropism, the curvature of the root in response to gravity, results from greater elongation of the upper side of the root than the lower side. Differential root elongation has been postulated to arise as the consequence of a gravity-induced auxin gradient with more auxin on the lower than the upper side (Kaufman, P.B., *et al.*,

(1995) In Plant Hormones Kluwer Academic Publishers, Dordrecht, Boston, London). The factors responsible for creating the auxin gradient are not known.

The simplest model to explain the phenotypes of the *eir1* mutant is that *EIR1* is required for efflux of auxin from the cells of the root tip into the elongation zone. If the root is oriented so that there is an increase in the auxin concentration on one side of the root tip, then EIR1 would pump auxin into the adjacent elongation zone with the concomitant inhibition of cell elongation. In *eir1* mutants the increased auxin in the lower portion of the root tip would fail to be transported into the elongation zone, and there would be no differential elongation. The predicted phenotypes of such a defect agree with those observed for an *eir1* mutation. The root should be agravitropic, and longer overall than an *EIR1* root. Furthermore, as described herein, increased levels of internal auxin should fail to inhibit the root or to induce root specific auxin inducible transcripts. The insensitivity of the *eir1* root to ethylene can be reconciled with the model if ethylene inhibits root growth by increasing the internal auxin concentrations (Suttle, J.C., (1991) *Plant Physiol. 96*: 875-880).

This model is also consistent with the response of *eir1* mutants to externally added auxins. If the *eir1* block were not in efflux, but rather in uptake of auxin, as has been proposed for *aux1* mutants (Bennett, M.J., *et al.*, (1996) *Science 273*: 948-50), then like the *aux1* mutants, the *eir1* mutants should be resistant to external auxin. However, the *eir1* mutants respond normally to external auxin. Root elongation is inhibited as in wild type, and induction of the *AtIAA2*-reporter construct appears to be unaffected.

EXAMPLE 5

EIR1 Expression is Localized to the Root

RNA-specific-PCR (RS-PCR) was used to analyze *EIR1* expression in the plant. Primers located on the 5' end of the *EIR1*-cDNA were used to amplify transcripts from reverse transcribed poly-A^+ RNA derived from roots, leaves, stems, flowers, and siliques. Primers for first strand cDNA synthesis were chimeric, having a 5' extension with no complementary sequences in the *Arabidopsis* genome. This sequence extension was used for subsequent PCR to avoid contamination. Genomic

DNA from ecotype Col-O served as a negative control. Results revealed a specific RS-PCR product in the root, but not in any other tissues. The root-specificity of *EIR1*-expression correlates well with the root-specific alterations detected in *eir1* mutants, suggesting that these defects are likely to be a consequence of the absence of *EIR1* function in the roots.

EXAMPLE 6

EIR1 Function is Required for Auxin Homeostasis in Root Cells

The involvement of EIR1 in root-specific auxin distribution was tested by analysis of the expression pattern of an auxin inducible gene, *AtIAA2*. The expression of *AtIAA2* has been shown to be strongly induced within a few minutes after exposure to auxin (Abel, S., *et al.*, (1996) *BioEssays. 18*: 647-654) The *AtIAA2* expression pattern was visualized using a reporter construct, *PIG4::GUS*, a transgene expressing β-glucuronidase under control of the AtIAA2-promoter. *AttIAA2* expression is strongest in the root meristem in wild type and *eir1-3*. When wild type is gravistimulated, expression of *AtIAA2* extends into the elongation and differentiation zone. Moreover, the expression is asymmetric with the lower portion of the elongation zone showing more intense staining than the upper. This asymmetric staining suggests that the lower portion of the elongation zone has elevated auxin levels as compared with the upper level. By contrast, reporter expression in *eir1-3* does not respond to the gravistimulus and remains restricted to the root tip.

The *eir1* root is known to be less sensitive to ethylene and to have an increased resistance to synthetic auxin transport inhibitors. These phenotypes could be explained if ethylene, like auxin transport inhibitors, interferes with tissue distribution of auxin. The effect of exogenous auxin on *PIG4::GUS* was assessed. Expression of *AtIAA2* has been shown to be strongly induced within a few minutes after exposure to auxin (Abel, S., *et al.*, (1995). *J Mol Biol. 251*: 533-49). In plants grown on regular medium, GUS staining is found in the root meristem and in the stele proximal to the root meristem. Addition of NAA (an auxin analogue) to the medium induces reporter gene expression in both the root meristem and elongation

zone of the root tip in wild type and the *eir1* mutant. Therefore, *eir1* mutants retain their ability to respond to exogenous auxin.

Plants (wild type and mutant) with the reporter responded quite differently to growth in ACC (the immediate biosynthetic precursor of ethylene) (1 μM ACC for 24 hours). In wild type, the entire elongation and differentiation zone shows considerable GUS staining upon ACC treatment. Furthermore, expression of GUS in the cell division zone appeared to be enhanced. In striking contrast, *eir1-3* mutant plant roots grown in ACC shows virtually no response in these tissues. Expression is restricted to the root tip at an intensity similar to that of plants grown in the absence of ACC.

The results with the auxin transport inhibitor TIBA are similar to those obtained with exogenous ACC. The reporter construct is induced in wild type but the mutant has a very reduced response. As auxin is the only known endogenous inducer of *AtIAA2* (Abel, S., *et al.*, (1996) *BioEssays. 18*: 647-654), ectopic expression of *AtIAA2* in wild-type roots treated with auxin transport inhibitors should be a consequence of elevated auxin concentrations in those cells that express the reporter. Unaltered *AtIAA2* expression in TIBA- and ACC-treated *eir1-3* roots suggests that auxin concentrations in cells of the root elongation zone remain unaffected when treated with these compounds.

The expression pattern of the auxin-inducible *AtIAA2::GUS* fusion in *eir1-3* is consistent with a block in auxin transport in the roots of this mutant. In wild type and *eir1-3* plants this reporter is expressed in root tips and at a low level in the younger parts of the vascular tissue. Wild type plants in the presence of ethylene, show increased expression of the reporter in the elongation zone, suggesting that these cells have an increased level of IAA.

The expression of the auxin-inducible reporter upon gravistimulation supports and extends these results. In wild type the auxin reporter is expressed asymmetrically, with more intense GUS-staining localized to the lower side of the elongation zone. This distribution is consistent with a model that proposes an inhibitory role for auxin in the regulation of root cell elongation and differential inhibition as the basis for gravitropism. Consistent with this interpretation, the

agravitropic *eir1-3* mutant grown under the same conditions fails to show differential staining or induction of the reporter in the elongation zone.

The failure of cells in the elongation zone of *eir1* roots to respond to IAA could be a consequence either of a failure to synthesize or to redistribute this growth regulator in response to ethylene. The effect of the *eir1* mutation on the root phenotype of the *alf1* mutant supports the redistribution hypothesis. The *alf1* mutation results in an approximately ten-fold increase in the endogenous concentration of IAA (Boerjan, W., *et al.*, (1995) *Plant Cell. 7*: 1405-1419). The high auxin level enhances the formation of lateral and adventitious roots but, also inhibits root elongation. Primary root growth in the *eir1 alf1* double mutant is not inhibited, showing that *eir1* suppresses the inhibitory effect of IAA on root elongation caused by the *alf1* mutation. However, *eir1* does not block the hyper-induction of lateral roots caused by *alf1*, showing that there are high levels of auxin in the root of the *eir1 alf1* double mutant.

These data are consistent with a model in which *EIR1* functions in auxin homeostasis in the root and auxin distribution in the root elongation zone. Two directions of auxin transport have been suggested for roots (Estelle, M., (1996) *Curr Biol. 6*: 1589-91): acropetal transport in the central cylinder from the base to the tip of the root and basipetal transport from the root tip to the elongation zone. If the inhibition of root growth in the *alf1-1* mutant results from the inhibition of cell expansion by excess auxin in the cells of the elongation zone, then the suppression of *alf1* by *eir1* is a consequence of *eir1's* defect in basipetal auxin transport into the elongation zone.

The root phenotype of *eto3* and *ctr1*, like that of the *alf1* mutant, is also suppressed by *eir1*. In both mutants the entire plant exhibits a strong ethylene response. *eto3* causes ethylene overproduction, whereas *ctr1* is probably a negative regulator of the ethylene response because *ctr*1 strains act as if they were in the presence of high ethylene although they do not have elevated ethylene concentrations (Kieber, J.J., *et al.*, (1993) *Cell. 72*: 427-441). The *eir1* mutant partially suppresses the *ctr1* phenotypes suggesting that EIR1 acts either downstream of ETO3 and CTR1 or in a pathway parallel to that in which ETO3 and CTR1 function (Roman, G., *et al.*, (1995) *Genetics. 139*: 1393-1409).

The decreased sensitivity of the *eir1* root to the inhibitory effects of ethylene as well as to the synthetic auxin transport inhibitors TIBA and NPA suggests a connection between auxin and ethylene. This behavior is similar to that of the *HOOKLESS1* (*HLS1*), mutants of *Arabidopsis* (Lehman, A., *et al.*, (1996) *Cell. 85*: 183-94). *HLS1* is thought to control bending in the apical tip of the hypocotyl because *hls1* mutants fail to form the apical hook during germination. Expression of the *HLS1* gene and enhanced hook formation are induced by treatment of plants with ethylene, which causes differential cell elongation. Remarkably, wild type seedlings grown in the presence of NPA have the same effect on apical hook formation and tissue distribution of auxin-induced genes as does the *hls1* mutant. Thus, auxin transport inhibitors phenocopy the *hls1* mutant, which is defective in the response of the apical hook to ethylene. These observations led to the speculation (Lehman, A., *et al.*, (1996) *Cell. 85*: 183-94) that an ethylene response gene may control differential cell growth by regulating auxin activity or distribution.

The growth characteristics of the *eir1* mutants also suggest a connection between auxin and ethylene. The *eir1* mutant root, like the apical hook of the *hls1* mutant is less sensitive to both exogenous and endogenous ethylene. Growth of wild type in the presence of auxin transport inhibitors blocks apical hook formation and the negative gravitropic response of the root. Moreover, like *hls1* the *eir1* roots are resistant to auxin transport inhibitors. In fact, this cross-resistance to both ethylene and auxin transport inhibitors is characteristic of mutants defective for auxin and ethylene responses (Fujita, H. and K. Syono, (1996) *Plant Cell Physiol. 37*: 1094-1101). This phenomenon probably represents an underlying mechanistic connection between the ethylene response and the auxin response, which is not yet understood.

EXAMPLE 7

eir1 Blocks the Inhibition of Root Growth Caused by High Endogenous Levels of Auxin

If *EIR1* is responsible for the redistribution of endogenous auxin, then the *eir1* mutation should block the defects in strains producing high levels of auxin. The effect of endogenous auxin was examined in *eir1-1 alf1-1* double mutants. The *alf1* mutation results in an enormously increased concentration of internal auxin, which leads to severe morphological alterations, which include the development of numerous short adventitious and lateral roots (Celenza, J.L., *et al.*, (1995) *Genes Dev. 9*: 2131-2142; Boerjan, W. *et al.*, (1995) *Plant Cell 7*: 1405-1419). The short root phenotype is caused by inhibition of cell elongation. The *eir1-1* mutation completely suppresses the short root phenotype caused by *alf1-1*, and retains the agravitropic phenotype, whereas the aerial portion of the *eir1 alf1* double mutant resembles *alf1*. These results suggest that in the eir1/alf1 double mutant elevated auxin levels do not reach the root elongation zone and that EIR1 is a tissue specific auxin transporter which is active in the root but not in the vascular tissue. Furthermore, the increased adventitious and lateral root formation, typical of *alf1* is not blocked by *eir1-3* suggesting that *eir1* represents a root tip-specific suppressor of the elevated auxin concentrations present in *alf1*.

EXAMPLE 8

Auxin transport in *Saccharomyces cerevisiae* Expressing *EIR1*

When auxin (IAA) is maintained under relatively acidic assay conditions (e.g., pH 4.0) it is protonated and thus capable of entering cells via diffusion across the plasma membrane. Once inside the cell the higher cytoplasmic pH acts as an ion-trap. IAAH dissociates and efflux of IAA-depends on anion transporters. We found that only in the absence of external carbon sources (compare Figures 10 A and B) there is a significant difference in the transport kinetics. Under these condiditions the ATP-requiring transporters of yeast are down as no new ATP is synthesized in the absence of an exogenous carbon source. However, yeast can maintain its intracellular (higher) pH for at least 30 minutes. This pH gradient is sufficient for EIR1-mediated ^{14}C-IAA transport as shown by the *gef1* and *gef1 EIR1* data (Figure

10B). Data resulting from the same experiment performed in the presence of the presence of the plasma-membrane specific protonophore CCCP demonstrates that under these under these conditions all differences in axuin transport activity between the EIR1-expressing and the control strain are gone (*gef1*+CCCP; *gef1 EIR1*+CCCP (Figure 10B)). Adding CCCP causes uptake of protons from the more acidic extracellular space into the cells. As a result the intracellular pH drops which gives rise to a protonation of IAA^-. IAAH in turn can diffuse across the plasma membrane following a concentration gradient.

EXAMPLE 9

EIR1 in *Saccharomyces cerevisiae* Confers Increased Resistance to Flouroindolic Compounds

The growth of yeast strains that overexpress a plasmid borne *Arabidopsis EIR1* gene under the control of the *ADH1* promoter was analyzed. Wild type yeast strains are only slightly sensitive to fluorinated indolic compounds such as 5-DL-fluoro-tryptophan or 5-fluoro-indole, toxic analogues of potential precursors of IAA (Bartel, B., (1997) *Plant Mol. Biol. 48*: 51-66). However, strains, which carry the Δ*gef1* deletion (a mutant which alters ion homeostasis in yeast (Gaxiola, R.A., *et al.*, (1998) *Proc. Natl. Acad. Sci. USA 95*: 4046-4050)), are much more sensitive to 5-fluoro-indole, 5-fluoro-DL-tryptophan and 5-fluoro-indoleacetic acid. Remarkably, *gef1* strains that contain the *EIR1* gene were much more resistant to these indolic compounds than isogenic *gef1* strains with only a vector. The increased resistance conferred by EIR1 can also be observed in wild type, but the effect is more subtle because of the greater intrinsic resistance of strains with a functional *GEF1* gene.

Expression of the *EIR1* gene is required for this resistance because yeast strains containing a mutant form of the *EIR1* gene (a frameshift in the *EIR1* OF, plasmid pADE1-H) fail to show the increased resistance to fluoro-indoles. Moreover, this resistance is specific to these indolic compounds because strains carrying the *EIR1* gene are no more resistant than controls to fluconazole, another inhibitor of yeast growth. In addition, the increased resistance is not simply the consequence of expression of a foreign transporter in yeast. Expression of the

Torpedo marmorata chloride channel (*C1C-0*), which suppresses many of the *gef1* defects, failed to confer increased resistance to indolic compounds.

In order to localize the EIR1 protein in yeast, a functional, hemagglutinin (HA) epitope-tagged version of *EIR1* was introduced into *S. cerevisiae*. Examination of immunodecorated yeast cells using charge-coupled microscopy localized the most intense staining of EIR1 to the plasma membrane. This membrane localization is consistent with a role for EIR1 in excluding compounds from the cell and, thereby, preventing the toxicity of the indolic compounds.

EXAMPLE 10

Creation and Characterization of *EIR1* Alleles

Site-specific mutagenesis was performed in order to replace the conserved residue Ser97 of EIR1 with other amino acids. Three alleles were made: EIR1-S97G, EIR1-S97A and EIR1-S97E. Table 2 shows a comparison of the nucleotide and the deduced amino acid sequence of EIR1 and the three negative alleles proximal to Serine 97. The affected amino acid residue is typed in bold letters, alterations in the nucleotide sequence are indicated as lower case letters. Mutations were introduced by site-directed mutagenesis. No other alterations in the nucleotide sequences could be detected.

Table 2 EIR1 Ser97 mutants

EIR1	AGAGGAAGCCTA R G **S** L
EIR1-S97G	AGAGGAgGCCTA R G **G** L
EIR1-S97A	AGAGGAgcCCTA R G **A** L
EIR1-S97E	AGAGGAgagCTA R G **E** L

Expression of these alleles under control of the ADH-promoter (pADEI/S97G, pADEI/S97A and pADEI/S97E) was performed in diploid yeast strains, defective for

the *GEF1* gene (as described in Luschnig *et al.*, (1998) *Genes and Dev., 12(14):* 2175-2187). When testing these strains in a filter assay performed with either 5-fluoro-indole or 5-fluoro-indole acetic acid, these strains exhibit a hypersensitivity towards these toxic compounds. Moreover, these phenotypes appear to be conditional, as there are no growth differences detectable between strains expressing either EIR1 or one of the mutant alleles grown in regular medium. However, addition of 5-fluoro-indole (final concentration: 200 μM) to the liquid cultures, results in a reduced growth rate of yeast strains expressing the negative alleles.

A comparison of an HA-tagged version of EIR1 (pADEI-HA, Luschnig *et al.*, (1998) *Genes and Dev., 12(14):* 2175-2187) with an HA-tagged version of EIR1-S97G (pADEI/S97G-HA; a derivative of pADEI-HA in which an An *AgeI-PmlI* DNA fragment of pADEI-HA was replaced with the same fragment from pADEI/S97G carrying the Serine to Glycine substitution) revealed that EIR1-S97G no longer localizes to the plasma membrane, but is enriched in intracellular vesicle-like structures. A possible consequence of protein retention within the cell would be an increased concentration of the toxic, indolic compounds which, in turn, would explain the hypersensitivity of yeast strains, expressing the negative alleles. Increased intracellular concentrations of these compounds could be mediated by either binding of Flouroindolic to the mutant EIR1-protein or by increased uptake of the toxins into the vesicle-like structures.

Results showed that Serine 97 is of critical importance for correct targeting of EIR1 in yeast. Expression of three different alleles EIR1-S97G, -S97E, as well as -S97A, results in a reversion of (loss of) the 5 fluoro-indole resistance phenotype observed upon expression of the wild type EIR1 protein. *gef1* strains transformed with the different alleles under control of the ADH promoter were plated and tested for their growth in a filter assay. There was a dramatic increase in the zone of inhibition of the Flouroindolic for all of the gef1 transformants expressing one of the negative alleles. The growth delay caused by a replacement of Serine 97 does not interfere with yeast growth in the absence of Flouroindolic. The Growth curves of *gef1* strains transformed with either *EIR1* or one of the Ser97 mutants in Synthetic Complete medium (SC) (Figure 11A) or SC supplemented with 200 μM 5-fluoro-indole (Figure 11B) indicates that although growth in unsupplemented medium is not

affected by the mutations; growth in the presence of 5-fluoro-indole is severely reduced in all three mutant strains.

Immunodetection of hemagglutin (HA) epitope-tagged versions of EIR1-HA and EIR1/S97G-HA, performed according to the method of Example 9, revealed that EIR1/S97G-HA does not localize to the plasma membrane, as does EIR1-HA, rather it is enriched in intracellular vesicle-like structures.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

CLAIMS

What is claimed is:

1. Isolated DNA encoding a protein involved in root-specific auxin transport, selected from the group consisting of:
   (a) DNA of SEQ ID NOS.: 11, 12 or 13 or portions thereof which encode a functional root-specific auxin transport protein;
   (b) DNA which, due to the degeneracy of the genetic code, encodes a protein having an amino acid sequence of SEQ ID NOS.:1 or 2;
   (c) DNA which hybridizes to DNA of (a) or (b) under high stringency conditions;
   (d) DNA which is sufficiently similar in sequence to DNA of (a), (b) or (c) to encode a root-specific protein involved in auxin transport;
   (e) DNA which encodes the amino acid sequence of SEQ ID NO.:2; and
   (f) isolated genomic DNA comprising DNA which encodes the amino acid sequence of SEQ ID NO.:2.

2. Isolated DNA selected from the group consisting of: SEQ ID NO.: 5, SEQ ID NO.:6, SEQ ID NO.:25 and SEQ ID NO.:26.

3. A genetically engineered plant comprising heterologous DNA encoding a root- specific protein involved in auxin transport, wherein the genetically altered plant is more resistant to an herbicidal composition which comprises at least one chemical compound which is auxin, an auxin derivative, an auxin analogue, or an auxin transport inhibitor.

4. The genetically engineered plant of Claim 3, wherein the heterologous DNA is selected from the group consisting of:
   (a) DNA of SEQ ID NOS.: 11, 12 or 13 or portions thereof which encode a functional root-specific auxin transport protein;
   (b) DNA which, due to the degeneracy of the genetic code, encodes a protein having an amino acid sequence of SEQ ID NOS.: 1 or 2;

(c) DNA which hybridizes to DNA of (a) or (b) under high stringency conditions;

(d) DNA which, is sufficiently similar in sequence to DNA of (a), (b) or (c) to encode a root-specific protein involved in auxin transport;

(e) DNA which encodes the amino acid sequence of SEQ ID NO.:2; and

(f) isolated genomic DNA comprising DNA which encodes the amino acid sequence of SEQ ID NO.:2.

5. The genetically engineered plant of Claim 4 which is a crop plant or a flowering plant.

6. A method of enhancing transport of an auxin derivative or an auxin analogue in plant roots comprising introducing into a plant part, including a seed, a gene which encodes a root-specific auxin transport protein and growing the plant part or seed under conditions appropriate for production of a plant, wherein the roots of the resulting plant contain the gene and express the encoded protein in sufficient quantity to enhance transport of auxin.

7. A genetically engineered plant comprising a heterologous DNA encoding a root-specific protein involved in auxin transport, wherein the genetically engineered plant exhibits altered auxin homeostasis relative to the auxin homeostasis of a wild type plant.

8. The genetically engineered plant according to Claim 7 wherein the heterologous DNA comprises isolated DNA selected from the group consisting of:

(a) DNA of SEQ ID NOS.: 11, 12 or 13 or portions thereof which encode a
functional root-specific auxin transport protein;

(b) DNA which, due to the degeneracy of the genetic code, encodes a protein having an amino acid sequence of SEQ ID NOS.:1 or 2;

(c) DNA which hybridizes to DNA of (a) or (b) under high stringency conditions;

(d) DNA which, is sufficiently similar in sequence to DNA of (a), (b) or (c) to encode a root-specific protein involved in auxin transport;

(e) DNA which encodes the amino acid sequence of SEQ ID NO.:2; and

(f) isolated genomic DNA comprising DNA which encodes the amino acid sequence of SEQ ID NO.:2.

9. The genetically engineered plant according to Claim 8, wherein the altered auxin homeostasis results in an increased number of lateral or adventitious roots.

10. A genetically engineered plant comprising isolated DNA according to Claim 1 wherein the plant is characterized by an increased auxin transport rate relative to the auxin transport rate of a corresponding wild type plant.

11. A method of identifying an allele of EIR1 which confers altered auxin-mediated responses in a plant, comprising the steps of:

(a) introducing a mutated EIR1 nucleic acid into yeast cells under conditions in which the DNA is expressed,thereby producing transformed yeast cells;

(b) contacting the transformed yeast cells of (a) with a fluorinated indolic compound under assay conditions which favor diffusion of the compound into the transformed yeast cells;

(c) determining the growth phenotype of the cells of (b); and

(d) comparing the growth phenotype of the transformed yeast cells to the growth phenotype of wild type cells

wherein detection of an altered growth phenotype in the transformed cells relative to the growth phenotype in wild-type cells indicates that the mutant EIR1 nucleic acid is an allele of EIR1 which confers altered auxin-mediated responses in a plant.

12. The method of Claim 11 in which the yeast cell is a diploid yeast strain defective for the GEF1 gene.

13. The method of Claim 12 wherein the fluorinated indolic compound is selected from the group consisting of: 5-DL-fluoro-tryptophan, 5-fluoro-indole and 5-fluoro-indolacetic acid.

14. The method of Claim 12 wherein the altered growth phenotype associated with expression of the mutated EIR1 nucleotide sequence comprises tolerance to concentrations of a fluorinated indolic compound which are toxic to wild type cells.

15. The method of Claim 12 wherein the altered growth phenotype associated with expression of the mutated EIR1 nucleotide sequence comprises increased
sensitivity to concentrations of a fluorinated indolic compound which are not toxic to wild type cells.

16. A transgenic plant comprising an allele of EIR1 DNA identified by the method of Claim 12.

17. Isolated or recombinantly produced root-specific protein involved in auxin transport and alleleic variants thereof.

18. The protein of Claim 17 comprising an amino acid sequence selected from the group consistig of : SEQ ID NO.: 1 and SEQ ID NO.: 2.

19. An expression vector comprising DNA selected from the group consisting of: SEQ ID NO.: 11, SEQ ID NO.: 12 and SEQ ID NO.: 13.

20. A method of producing a transgenic plant characterized by altered auxin homeostasis comprising the steps of:

(a) introducing DNA encoding a root-specific auxin transport carrier protein into a plant cell under conditions in which the DNA is expressed, thereby producing a transformed plant cell; and

(b) producing a transgenic plant from the transformed plant cell.

21. The method of Claim 20, wherein the DNA encoding a root-specific auxin transport carrier protein is selected from the group consisting of:

(a) DNA of SEQ ID NO.: 11, 12 or 13 or portions thereof which encode a functional root-specific auxin transport protein;

(b) DNA which, due to the degeneracy of the genetic code, encodes a protein having an amino acid sequence of SEQ ID NOS.:1 or 2;

(c) DNA which hybridizes to DNA of (a) or (b) under high stringency conditions;

(d) DNA which, is sufficiently similar in sequence to DNA of (a), (b) or (c) to encode a root-specific protein involved in auxin transport;

(e) DNA which encodes the amino acid sequence of SEQ ID NO.:2; and

(f) isolated genomic DNA comprising DNA which encodes the amino acid sequence of SEQ ID NO.:2.

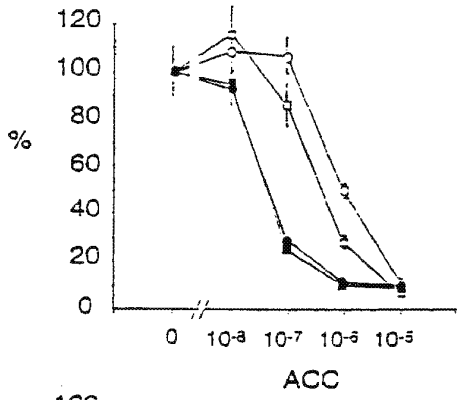


FIG. 1A

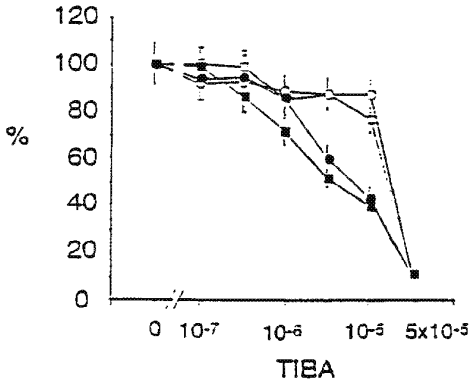


FIG. 1B

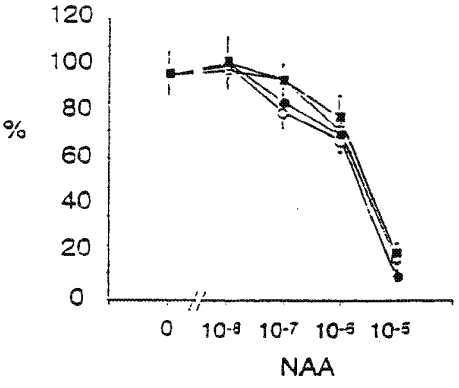


FIG. 1C

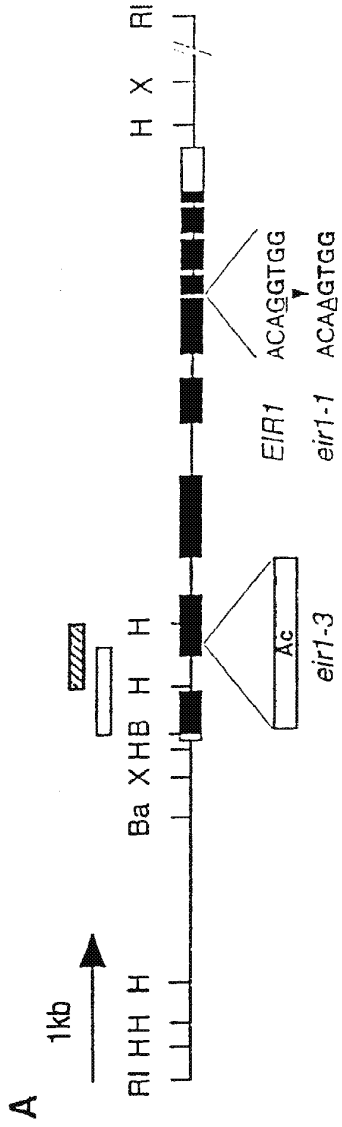

A
1kb
RI H H H
Ba X H B H H
Ac
eir1-3
EIR1 ACAGGTGG
eir1-1 ACAAGTGG
H X RI

FIG. 2

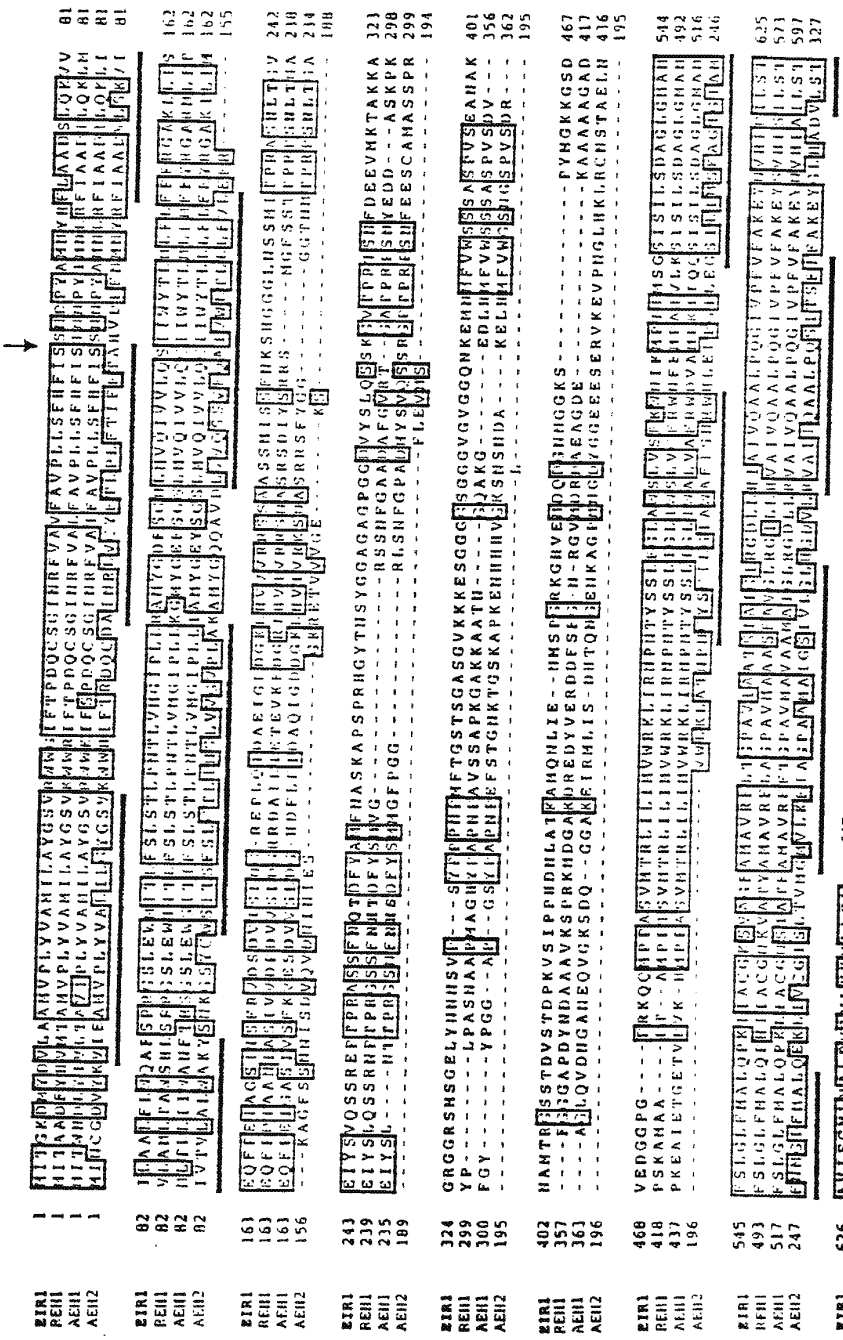


FIG. 3



FIG. 4

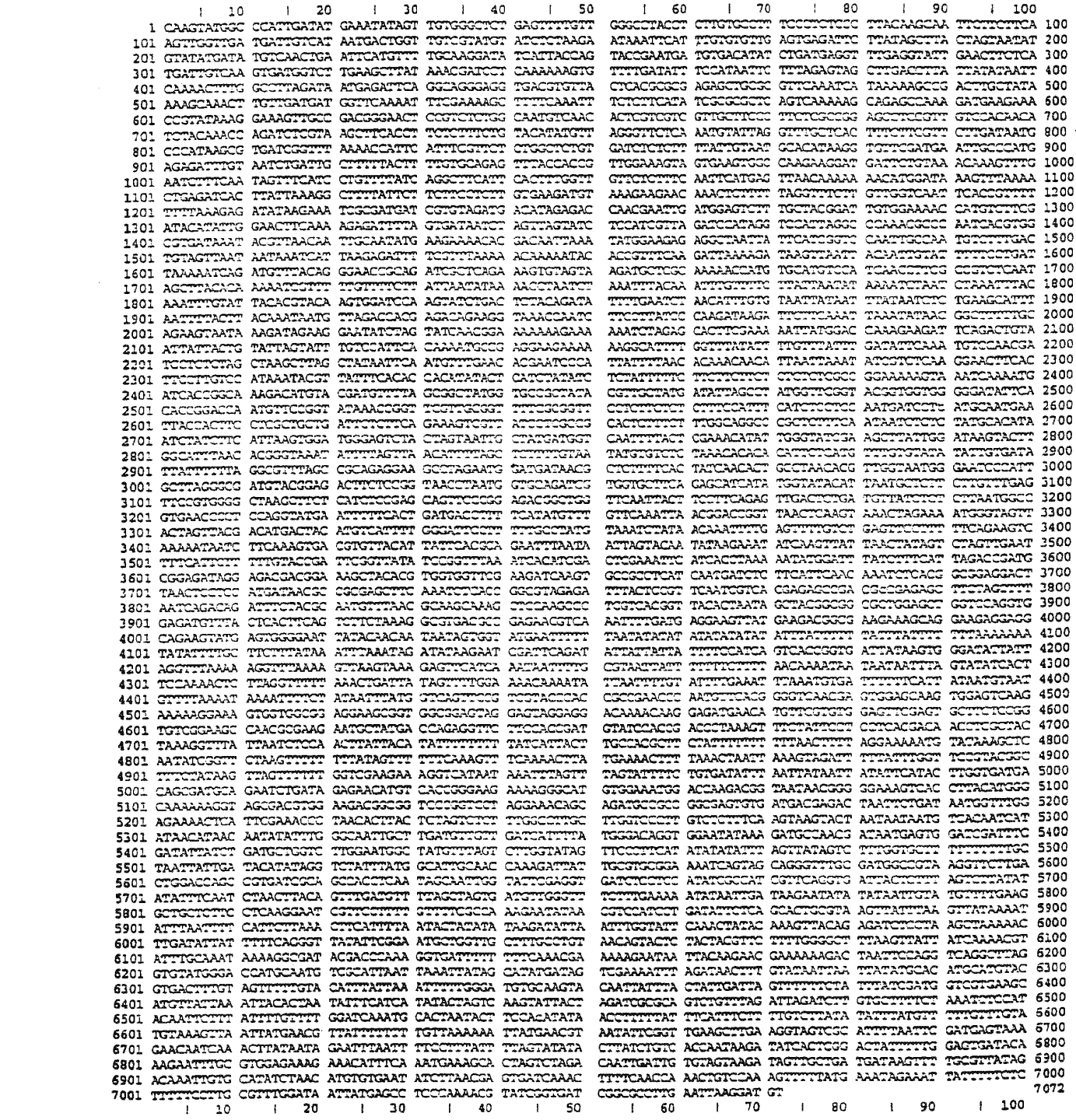


FIG. 5

COR-AAD1-00002230

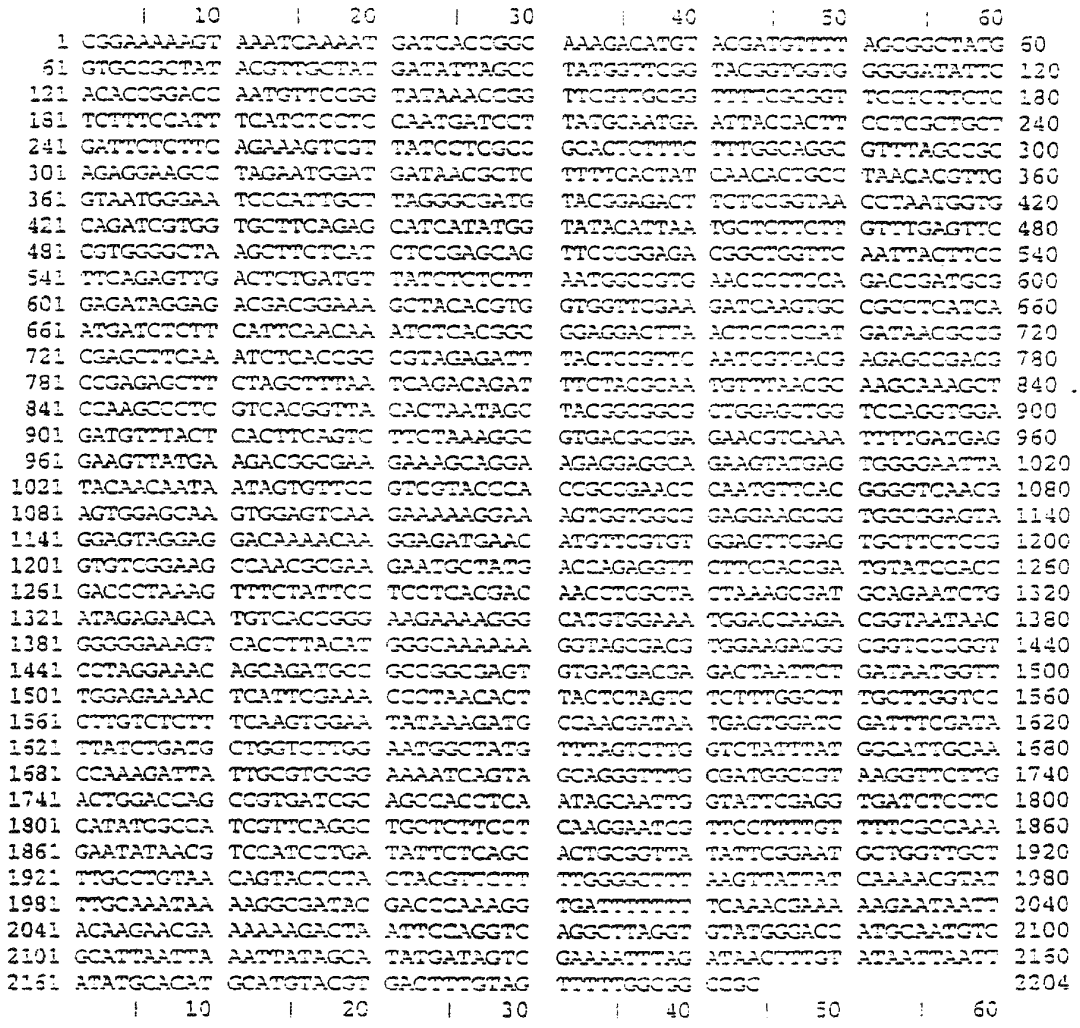


FIG. 6

```
        10         20         30         40         50         60
MITGKDMYDV LAAMVPLYVA MILAYGSVRW WGIFTPDQCS GINRFVAVFA VPLLSFHFIS 60
SNDPYAMNYH FLAADSLQKV VILAALFLWQ AFSRRGSLEW MITLFSLSTL PNTLVMGIPL 120
LRAMYGDFSG NLMVQIVVLQ SIIWYTLMLF LFEFRGAKLL ISEQFPETAG SITSFRVDSD 180
VISLNGREPL QTDAEIGDDG KLHVVVRRSS AASSMISSFN KSHGGGLNSS MITPRASNLT 240
GVEIYSVQSS REPTPRASSF NQTDFYAMFN ASKAPSPRHG YTNSYGGAGA GPGGDVYSLQ 300
        310        320        330        340        350        360
SSKGVTPRTS NFDEEVMKTA KKAGRGGRSM SGELYNNNSV PSYPPPNPMF TGSTSGASGV 360
KKKESGGGGS GGGVGVGGQN KEMNMFVWSS SASPVSEANA KNAMTRGSST DVSTDPKVSI 420
PPHDNLATKA MQNLIENMSP GRKGHVEMDQ DGNNGGKSPY MGKKGSDVED GGPGPRKQQM 480
PPASVMTRLI LIMVWRKLIR NPNTYSSLFG LAWSLVSFKW NIKMPTIMSG SISILSDAGL 540
GMAMFSLGLF MALQPKIIAC GKSVAGFAMA VRFLTGPAVI AATSIAIGIR GDLLHIAIVQ 600
        610        620        630        640        650        660
AALPQGIVPF VFAKEYNVHP DILSTAVIFG MLVALPVTVL YYVLLGL. 648
```

FIG. 7

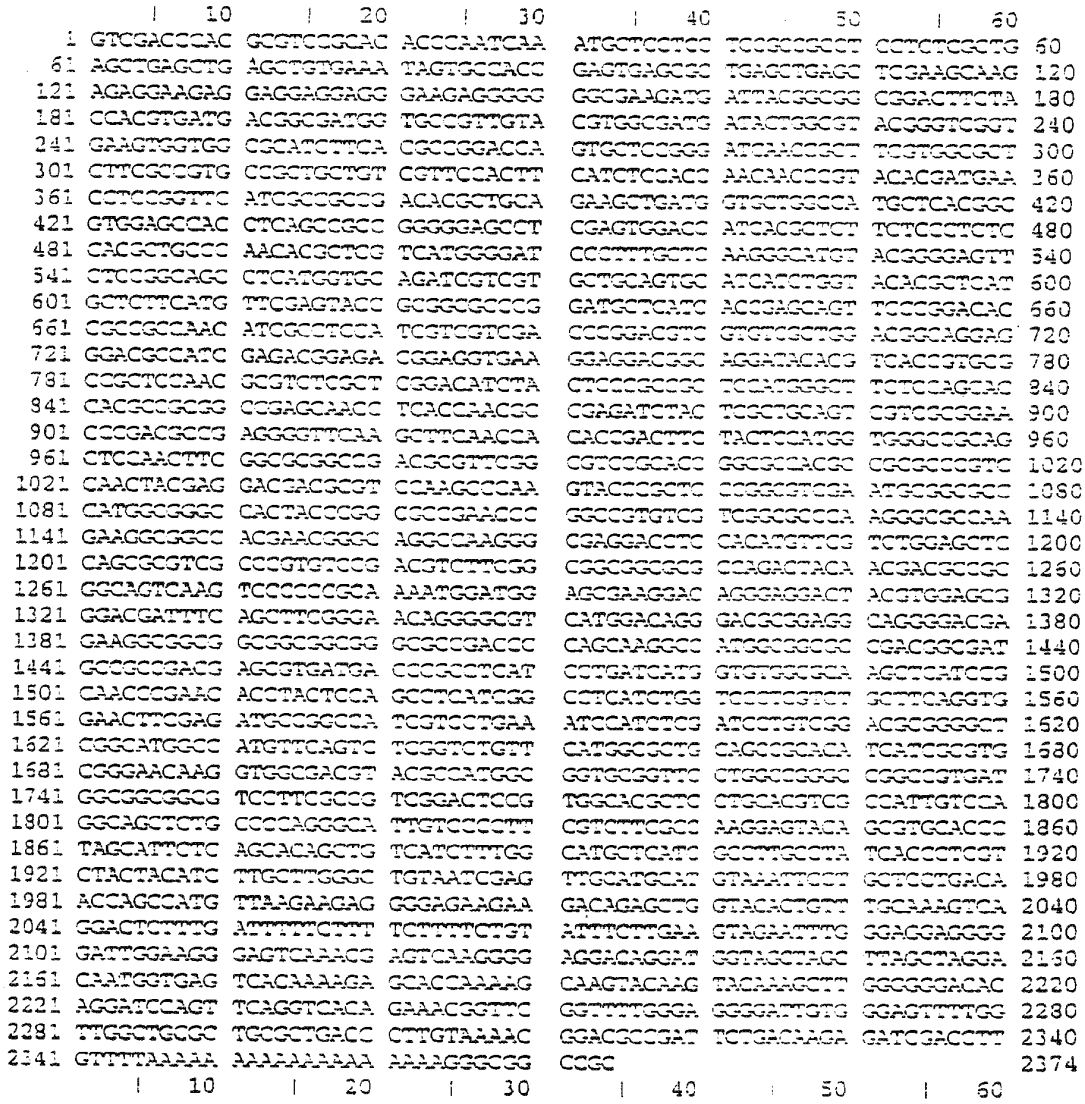

| 10 | 20 | 30 | 40 | 50 | 60
1 GTCGACCCAC GCGTCCGCAC ACCCAATCAA ATGCTCCTCC TCCGCCGCCT CCTCTCGCTG 60
61 AGCTGAGCTG AGCTGTGAAA TAGTGCCACC GAGTGAGCGC TGAGCTGAGC TCGAAGCAAG 120
121 AGAGGAAGAG GAGGAGGAGG GAAGAGGGGG GGCGAAGATG ATTACGGCGG CGGACTTCTA 180
181 CCACGTGATG ACGGCGATGG TGCCGTTGTA CGTGGCGATG ATACTGGCGT ACGGGTCGGT 240
241 GAAGTGGTGC CGCATCTTCA CGCCGGACCA GTGCTCCGGG ATCAACCGCT TCGTGGCGCT 300
301 CTTCGCCGTG CCGCTGCTGT CGTTCCACTT CATCTCCACC AACAACCGCT ACACGATGAA 360
361 CCTCCGGTTC ATCGCCGCCG ACACGCTGCA GAAGCTGATG GTGCTGGCCA TGCTCACGGC 420
421 GTGGAGCCAC CTCAGCCGCC GGGGGAGCCT CGAGTGGACC ATCACGCTCT TCTCCCTCTC 480
481 CACGCTGCCC AACACGCTCG TCATGGGGAT CCGTTTGCTC AAGGGCATGT ACGGGGAGTT 540
541 CTCCGGCAGC CTCATGGTGC AGATCGTCGT GCTGCAGTGC ATCATCTGGT ACACGCTCAT 600
601 GCTCTTCATG TTCGAGTACC GCGGCGCCCG GATGCTCATC ACCGAGCAGT TCCCGGACAC 660
661 CGCCGCCAAC ATCGCCTCCA TCGTCGTCGA CCCGGACGTC GTGTCGCTGG ACGGCAGGAG 720
721 GGACGCCATC GAGACGGAGA CGGAGGTGAA GGAGGACGGC AGGATACACG TCACCGTGCG 780
781 CCGCTCCAAC GCGTCTCGCT CGGACATCTA CTCCCGCCGC TCCATGGGCT TCTCCAGCAC 840
841 CACGCCGCGG CCGAGCAACC TCACCAACGC CGAGATCTAC TCGCTGCAGT CGTCGCGGAA 900
901 CCCGACGCCG AGGGGTTCAA GCTTCAACCA CACCGACTTC TACTCCATGG TGGGCCGCAG 960
961 CTCCAACTTC GGCGCGGCCG ACGCGTTCGG CGTCCGCACC GGCGCCACGC CGCGCCCGTC 1020
1021 CAACTACGAG GACGACGCGT CCAAGCCCAA GTACCCGCTC CCGGCGTCGA ATGCGGCGCC 1080
1081 CATGGCGGGC CACTACCCGG CGCCGAACCC GGCCGTGTCG TCGGCGCCCA AGGGCGCCAA 1140
1141 GAAGGCGGCC ACGAACGGGC AGGCCAAGGG CGAGGACCTC CACATGTTCG TCTGGAGCTC 1200
1201 CAGCGCGTCG CCCGTGTCCG ACGTCTTCGG CGGCGGCGCG CCAGACTACA ACGACGCCGC 1260
1261 GGCAGTCAAG TCCCCCCGCA AAATGGATGG AGCGAAGGAC AGGGAGGACT ACGTGGAGCG 1320
1321 GGACGATTTC AGCTTCGGGA ACAGGGGCGT CATGGACAGG GACGCGGAGG CAGGGGACGA 1380
1381 GAAGCCGGCG GCGGCGGCGG GCGCCGACCC CAGCAAGGCC ATGCCGGCGC CGACGGCGAT 1440
1441 GCCGCCGACG AGCGTGATGA CCCGCCTCAT CCTGATCATG GTGTGGCGCA AGCTCATCCG 1500
1501 CAACCCGAAC ACCTACTCCA GCCTCATCGG CGTCATCTGG TCCTCGTCTT GCTTCAGGTG 1560
1561 GAACTTCGAG ATGCCGGCCA TCGTCCTGAA ATCCATCTCG ATCCTGTCGG ACGCGGGGCT 1620
1621 CGGCATGGCC ATGTTCAGTC TCGGTCTGTT CATGGCGCTG CAGCCGCACA TCATCGCGTG 1680
1681 CGGGAACAAG GTGGCGACGT ACGCCATGGC GGTGCGGTTC CTGGCCGGGC CGGCCGTGAT 1740
1741 GGCGGCGGCG TCCTTCGCCG TCGGACTCCG TGGCACGCTC CTGCACGTCG CCATTGTCCA 1800
1801 GGCAGCTCTG CCCCAGGGCA TTGTCCCCTT CGTCTTCGCC AAGGAGTACA GCGTGCACCC 1860
1861 TAGCATTCTC AGCACAGCTG TCATCTTTGG CATGCTCATC GCCTTGCCTA TCACCTCGT 1920
1921 CTACTACATC TTGCTTGGGC TGTAATCGAG TTGCATGCAT GTAAATTCGT GCTCCTGACA 1980
1981 ACCAGCCATG TTAAGAAGAG GGGAGAAGAA GACAGAGCTG GTACACTGTT TGCAAAGTCA 2040
2041 GGACTCTTTG ATTTTTCTTT TCTTTTCTGT ATTTCTTGAA GTAGAATTTG GGAGGAGGGG 2100
2101 GATTGGAAGG GAGTCAAACG AGTCAAGGGG AGGACAGGAT GGTAGCTAGC TTAGCTAGGA 2160
2161 CAATGGTGAG TCACAAAAGA GCACCAAAAG CAAGTACAAG TACAAAGCTT GGGGGGACAC 2220
2221 AGGATCCAGT TCAGGTCACA GAAACGGTTC GGTTTTGGGA GGGGATTGTG GGAGTTTTGG 2280
2281 TTGGCTGCGC TGCGCTGACC CTTGTAAAAC GGACGCCGAT TCTGACAAGA GATCGACCTT 2340
2341 GTTTTAAAAA AAAAAAAAAA AAAAGGGCGG CCGC 2374
| 10 | 20 | 30 | 40 | 50 | 60

FIG. 8

```
          10         20         30         40
          |          |          |          |
MITAADFYHV MTAMVPLYVA MILAYGSVKW WRIFTPDQCS 40
GINRFVALFA VPLLSFHFIS TNNPYTMNLR FIAADTLQKL 80
MVLAMLTAWS HLSRRGSLEW TITLFSLSTL PNTLVMGIPL 120
LKGMYGEFSG SLMVQIVVLQ CIIWYTLMLF MFEYRGARML 160
ITEQFPDTAA NIASIVVDPD VVSLDGRRDA IETETEVKED 200
          210        220        230        240
          |          |          |          |
GRIHVTVRRS NASRSDIYSR RSMGFSSTTP RPSNLTNAEI 240
YSLQSSRNPT PRGSSFNHTD FYSMVGRSSN FGAADAFGVR 280
TGATPRPSNY EDDASKPKYP LPASNAAPMA GHYPAPNPAV 320
SSAPKGAKKA ATNGQAKGED LHMFVWSSSA SPVSDVFGGG 360
APDYNDAAAV KSPRKMDGAK DREDYVERDD FSFGNRGVMD 400
          410        420        430        440
          |          |          |          |
RDAEAGDEKA AAAAGADPSK AMAAPTAMPP TSVMTRLILI 440
MVWRKLIRNP NTYSSLIGLI WSLVCFRWNF EMPAIVLKSI 480
SILSDAGLGM AMFSLGLFMA LQPHIIACGN KVATYAMAVR 520
FLAGPAVMAA ASFAVGLRGT LLHVAIVQAA LPQGIVPFVF 560
AKEYSVHPSI LSTAVIFGML IALPITLVYY ILLGL. 596
```

FIG. 9

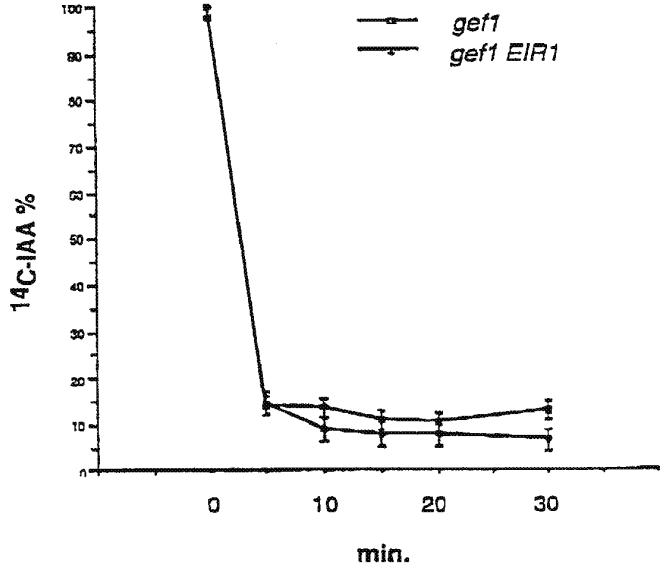


FIG. 10A

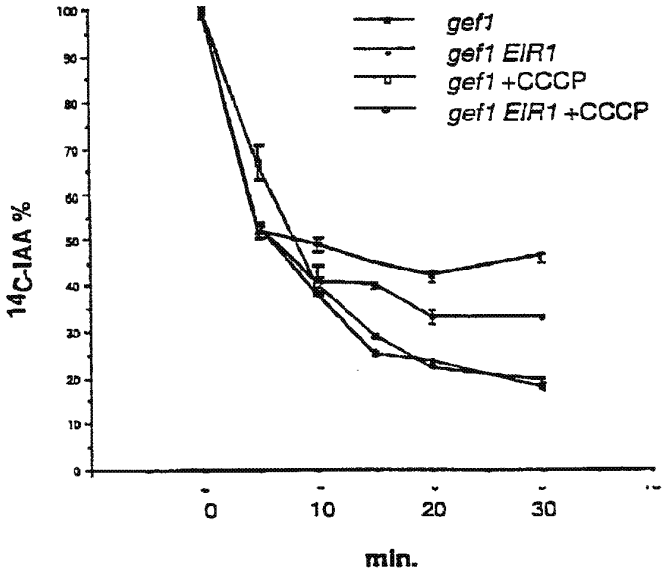


FIG. 10B

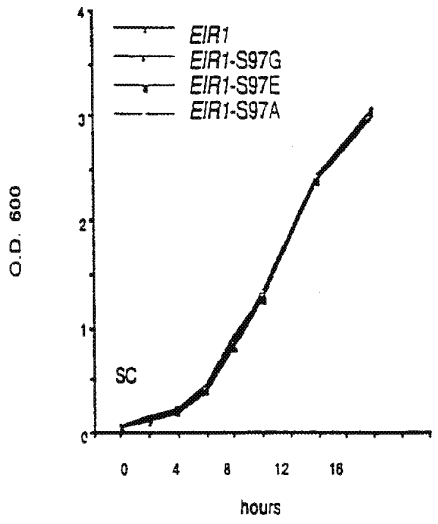


FIG. 11A

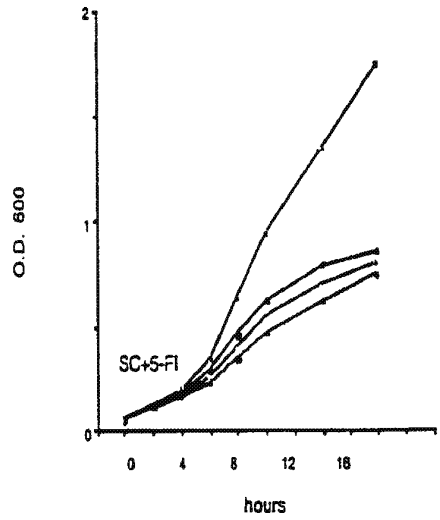


FIG. 11B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/12277

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/29 A01H5/00 C07K14/415

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 49810 A (MAX PLANCK GESELLSCHAFT ;GAELWEILER LEO (DE); PALME KLAUS (DE); WI) 31 December 1997 (1997-12-31) the whole document ---	1-11,17, 20,21
P,X	LUSCHNIG C ET AL.: "EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in Arabidopsis thaliana" GENES AND DEVELOPMENT, vol. 12, no. 14, 15 July 1998 (1998-07-15), pages 2175-2187, XP002116368 the whole document ---	1-21
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 September 1999	15/10/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl. Fax: (+31–70) 340–3016	Bilang, J

2

Form PCT/ISA/210 (second sheet) (July 1992)

COR-AAD1-00002237

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/12277

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication,where appropnate, of the relevant passages	Relevant to claim No.
P,X	MÜLLER A ET AL.: "AtPIN2 defines a locus of Arabidopsis for root gravitropism control" EMBO JOURNAL, vol. 17, no. 23, 1 December 1998 (1998-12-01), pages 6903-6911, XP002116369 the whole document	1-11,17, 20,21
P,X	UTSUNO K ET AL.: "AGR, an Agravitropic locus of Arabidopsis thaliana, Encodes a Novel Membrane-Protein Family Member" PLANT AND CELL PHYSIOLOGY, vol. 39, no. 10, October 1998 (1998-10), pages 1111-1118, XP002116370 the whole document	1-11,17, 20,21
A	ROMAN G ET AL.: "Genetic Analysis of Ethylene Signal Transduction in Arabidopsis thaliana: Five Novel Mutant Loci Integrated into a Stress REsponse Pathway" GENETICS, vol. 139, no. 3, March 1995 (1995-03), pages 1393-1409, XP002116371 page 1397, column 1, paragraph 2 - column 2, paragraph 1	

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/12277

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9749810 A	31-12-1997	EP 0814161 A	29-12-1997
		AU 3539597 A	14-01-1998
		CA 2259125 A	31-12-1997

COR-AAD1-00002239

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/82, A01H 1/00, 4/00, 5/00 — A1

(11) International Publication Number: **WO 00/06757**

(43) International Publication Date: 10 February 2000 (10.02.00)

(21) International Application Number: PCT/US99/07598

(22) International Filing Date: 6 April 1999 (06.04.99)

(30) Priority Data:
09/127,080 31 July 1998 (31.07.98) US

(71) Applicant: MYCOGEN PLANT SCIENCE, INC. [US/US]; 5501 Oberlin Drive, San Diego, CA 92121 (US).

(72) Inventors: HORN, Michael, E.; 1500 Research Parkway #220, College Station, TX 77845 (US). HALL, Gerald, E., Jr.; 142 Rice Drive, Morrisville, PA 19067 (US).

(74) Agents: LLOYD, Jeff et al.; Saliwanchik, Lloyd & Saliwanchik, Room A–1, 2421 N.W. 41st Street, Gainesville, FL 32606 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: IMPROVED PLANT TRANSFORMATION PROCESS BY SCAFFOLD ATTACHMENT REGIONS (SAR)

(57) Abstract

The subject invention pertains to the use of scaffold attachment regions (SARs), also known as matrix attachment regions (MARs), in expression cassettes to improve the transformation efficiency of such expression cassettes in a transformation process.

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			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

DESCRIPTION

IMPROVED PLANT TRANSFORMATION PROCESS BY SCAFFOLD ATTACHMENT REGIONS (SAR)

Cross-Reference to a Related Application

This application is related to and claims the benefit of U. S. Serial No. 09/127,080 filed July 31, 1998.

Field of the Invention

The present invention relates to the use of scaffold attachment regions (SARs), also known as matrix attachment regions (MARs), in transformation of cells and tissues.

Background of the Invention

SARs/MARs (hereinafter collectively referred to as "SARs") are AT-rich genomic DNA sequences that occur in eukaryotic genomes (see Boulikas [1993] *J. Cell Biochem.* 52:14). SARs are non-coding DNA sequences that flank structural genes and stabilize the transcription process. It is also known that SARs bind to certain components of the proteinaceous scaffold material that encompasses nuclear DNA. SARs have been found to improve the expression of heterologous genes in transformed plants (Allen *et al.* [1993] *Plant Cell* 5:603).

Brief Summary of the Invention

It has been unexpectedly discovered that the use of SARs in plant expression cassettes increases the frequency of recovery of stable transformation events of such expression cassettes in a plant transformation process. In particular, stable transformation events have been observed to increase up to about 7-fold or more while, at the same time, in the case of GUS expression, the transient transformation events have been observed to decrease. This startling observation makes the increase in stable transformation events all the more surprising. This invention is particularly useful in improved electroporation transformation techniques.

It has also been unexpectedly discovered that the use of SARs in plant expression cassettes advantageously decreases the time required to recover a stably transformed plant. The SAR-positive clones invariably appear in culture faster as compared to SAR-negative clones which, in turn, allows for quicker regeneration. Shorter times in culture leads to easier regeneration, less abnormalities in the plants, higher fertility, better seed set, *etc.* Quicker transformation/regeneration processes are highly desirable from a commercial standpoint due

to the fact that shorter time periods allow for dramatically increased throughput of events which result in quicker identification of commercial candidates.

Additionally, the use of SARs in a plant expression cassette employed in electroporation transformation processes results in a high proportion of low copy number transformants which is very desirable from a commercial standpoint. High copy number transformants are associated with high frequencies of gene silencing. Complex Southern blot patterns that are indicative of high copy number insertions have been observed rarely (approximately 10%) in SAR-positive maize transformants from an electroporation process of whole intact maize cells.

Brief Description of the Drawings

Figure 1 shows Southern blots of maize callus samples transformed with SAR-positive plant expression cassettes compared with controls. Treated callus was plated on bialaphos selection and bialaphos-resistant colonies appeared after 7 to 12 weeks. Callus was extracted for DNA and separated on an agarose gel as per Materials and Methods. Lane 1: DNA sizing standards; Lanes 2-12: DNA from selected colonies arising on bialaphos selection; Lanes 13: DNA from PAT+ control callus known to contain the PAT gene; Lane 14: DNA from PAT- control callus known not to contain the PAT gene; Lane 15: plasmid DNA corresponding to the equivalent of 1 copy; Lane 16: plasmid DNA corresponding to the equivalent of 3 copies.

Figure 2 shows Southern analysis of DNA from regenerated T_0 plants from callus (Fig. 1) derived from transfection with pAGM 607 plasmid DNA. Lane 1: DNA sizing standards; Lane 2: plasmid DNA corresponding to the equivalent of 1 copy; Lanes 3-11: not relevant; Lane 12: DNA from PAT+ control leaf tissue known to contain the PAT gene; Lane 13: DNA from PAT- control leaf tissue known not to contain the PAT gene; Lanes 15-27: DNA from leaves of T_0 plants regenerated from Southern positive callus.

Figure 3 shows Southern analysis of DNA from leaf tissue from plants segregating for the PAT gene in the ST2-1 derived T_1 generation. Lane 1: DNA sizing standards; Lane 2: blank; Lanes 3-11: DNA from leaf tissue of plants judged to be sensitive to the herbicide; Lane 22: DNA from PAT- control leaf tissue known not to contain the PAT gene; Lane 23: Lane 12: DNA from PAT+ control leaf tissue known to contain the PAT gene; Lane 24: plasmid DNA corresponding to the equivalent of 1 copy; Lane 25: plasmid DNA corresponding to the equivalent of 3 copies; Lanes 26-28: not relevant.

Figure 4 shows a nucleotide sequence of a SAR polynucleotide which can be used according to the present invention.

Figure 5 shows a restriction map of plasmid pAGM243.

Figure 6 shows a restriction map of plasmid pAGM285A.

Figure 7 shows a restriction map of plasmid pAGM607.

Figure 8 shows a restriction map of plasmid pAGM608.

Detailed Disclosure of the Invention

The present invention concerns methods and materials for increasing frequency of recovery of stable transformation events in plant transformation processes and for increasing the number of low copy number transformants, as well as reducing or eliminating the occurrence of gene silencing throughout subsequent generations descended from the original transformant. In one embodiment of the present invention, SARs are used in a plant expression cassette to provide a plant transformation process that produces a greater percentage of stable transformants and a greater percentage of low copy number transformants than are obtained with SAR-negative plant expression cassettes.

In one embodiment, a method of the subject invention comprises transforming a cell or tissue with a SAR polynucleotide sequence operably linked to a polynucleotide of interest that encodes a protein, polypeptide, or peptide. SARs useful with the subject invention include, but are not limited to, SARs originating from plants and animals. SAR-encoding polynucleotides useful with the subject invention include, for example, SAR isolated from tobacco (Hall *et al.*, [1991] *Proc. Natl. Acad. Sci. USA* 88:9320). In a preferred embodiment, a SAR polynucleotide useful in practicing the present invention comprises the nucleotide sequence shown in Figure 4, or a functional fragment or mutant thereof. Preferably, the SAR polynucleotide sequence is provided in the form of an "expression cassette" on a suitable vector. Any vector suitable for DNA mediated transformation can be used and such vectors are known in the art. In a highly preferred embodiment, a SAR polynucleotide sequence is operably linked at either and or both ends of the polynucleotide of interest. Transformation can be accomplished using known methods, including, for example, particle bombardment or biolistics transformation with DNA coated microparticles, *Agrobacterium*-mediated transformation, electroporation, microinjection, magnetophoresis, silicon carbide whiskers, PEG mediated transformation, and protoplast transformation. In a preferred embodiment, plant cells are transformed with the polynucleotides of the invention by electroporation according to standard techniques known in the art. See, for example, Pescitelli, S. M., K. Sukhapinda (1995) "Stable Transformation via Electroporation Into Maize Type II Callus and Regeneration of Fertile Transgenic Plants" *Plant Cell Reports* **14**: 712-716 (hereby incorporated by reference in its entirety). Transformed plant cells can be

selected and then cultured under suitable conditions according to routine practice to generate transformed plantlets and plants.

It is well known in the art that the nucleotide sequences of the subject invention can be truncated and/or mutated such that certain of the resulting fragments and/or mutants of the original full-length sequence can retain the desired characteristics of the full-length sequence. A wide variety of restriction enzymes are well known by ordinarily skilled artisans which are suitable for generating fragments from larger nucleic acid molecules. In addition, it is well known that *Bal*31 exonuclease can be conveniently used for time-controlled limited digestion of DNA. See, for example, Maniatis *et al.*(1982) *Molecular Cloning: A Laboratory Manual*, Coldspring Harbor Laboratory, New York, pages 135-139, incorporated herein by reference. See also Wei *et al.* (1983) *J. Biol. Chem.*, **258**:13006-13512. By use of *Bal*31 exonuclease (commonly referred to as "erase-a-base" procedures), the ordinarily skilled artisan can remove nucleotides from either or both ends of the subject acids to generate a wide spectrum of fragments which are functionally equivalent to the subject nucleotide sequences. One of ordinary skill in the art can, in this manner, generate hundreds of fragments of controlled, varying length from locations all along the original SAR molecule in one afternoon. The ordinarily skilled artisan is able to routinely test or screen the generated fragments for their characteristics for determining the utility of the fragments as taught herein. It is also well known that mutant sequences of the full length sequence, or fragments thereof, can be easily produced with site directed mutagenesis. See, for example, Larionov, O.A. and Nikiforov, V.G. (1982) "Directed Mutagenesis" *Genetika* **18(3)**:349-59; Shortle, D., DiMaio, D., and Nathans, D. (1981) "Directed Mutagenesis" *Annu. Rev. Genet.* **15**:265-94; both incorporated herein by reference. The skilled artisan can routinely produce deletion-, insertion-, or substitution-type mutations and test whether the resulting mutants contain the desired characteristics of the full length wild-type sequence, or fragments thereof.

In one embodiment of the invention, a polynucleotide comprising a polynucleotide sequence that has substantially the same sequence as a SAR polynucleotide is operably linked to a polynucleotide that encodes a protein, polypeptide, or peptide. Any desired polynucleotide sequence can be employed to transform cells or tissue according to the present invention. Of particular importance in this regard are genes used, for example, as (i) selective markers (antibiotic and/or herbicide resistance genes), (ii) reporter genes (*e.g.*,GUS), (iii) insecticide resistance genes (*B.t.* delta endotoxins) and (iv) any other genes that improve the value or use of a plant. Such genes can include, but are not limited to, glucuronidase, phosphinothricin N-acetyltransferase, green fluorescent protein (GFP), luciferase, Pat/bar, and glyphosate resistance

genes (NPTII, HPT, biomoxila resistance gene, AHAS, ALS, cyanomide hydrolase, adenine deaminase, 2,4-D monooxygeanse). The encoded protein, polypeptide or peptide can be one that is naturally present in the transformed cell or it can be heterologous to the transformed cell. Optionally, a regulatory sequence, such as, for example, a promoter sequence that can regulate transcription of the polynucleotide sequence, can be included in a polynucleotide of the invention. In a preferred embodiment, the SAR polynucleotide sequence is a polynucleotide component of an expression cassette on a suitable vector. Vectors useful with the SAR polynucleotides of the invention are known in the art and can be prepared and/or selected according to standard techniques.

The methods and materials of the present invention can be used to transform cells or tissue from any organism, and preferably from a eucaryotic organism. In a preferred embodiment, the cells are plant cells. Any plant cell competent to be transformed can be employed in the present invention. It is preferred to employ plant cells that are readily regenerable into whole plants. Suitable plant cells include embryogenic suspension cells, non-embryogenic suspension cells (except in corn where these cells are not regenerable), plant explants, germline cells (pollen, ovules, meristem domes, megaspore cells, embryos cells, egg cells and embryosacs), microspore cells and callus tissue cells, both compact callus and friable callus. Preferred cells include embryogenic callus, suspension cells (embryogenic suspension) and callused immature zygotic embryos. Particularly preferred plant cells are early embryogenic suspension and young callus (still attached to the zygotic embryo) cells of from about 3-14 and preferably from about 5-10, days old. Wounding the plant cells and/or treating the plant cells prior to being subjected to electroporation is unnecessary and is in fact expressly avoided. When embryogenic suspension tissues are employed it is preferred to gently break up the tissue into small clumps or into fine aggregates as small as possible without damaging the embryogenic suspension cells. This can be done by sieving the tissue through a screen, such as, for example, pushing tissue through a 1,000 micron (u) sieve with a spatula or pestle. The finer the aggregates or suspension of undamaged cells, the more efficient the present process.

Plant tissue useful with the invention includes, but is not limited to, callus, meristematic, leaf, shoot, root, and embryonic tissue. The present invention is applicable to any plant species including angiosperms (dicots, monocots) and gymnosperms. Suitable crops include corn, wheat (especially Type C wheat callus), sorghum, rice, pearl millet, sugar cane, orchardgrass and other Gramineae plants; soybean, peanuts, alfalfa and other members of the Luguminoseae family; cotton, kenaf, and other members of the Malvaceae family; poppy and other members of the Papavaraceae family; cannabis and other members of the Cannabinaceae family; tea and other

members of the Theaceae family; rape (canola), vegetables and oilseed crops and other members of the Cruciferae family; sunflower, safflower and other members of the Compositae family; coffee and other members of the Rubiaceae family; cacao, theobroma and other members of the Byttneriaceae family; fruits and vegetables, trees, orchard crops, and turf grass. Preferred crops include cotton, tomato, sugarbeet, potato, peanut, alfalfa, rice, wheat and especially corn (maize).

In a preferred embodiment, the polynucleotide sequences employed in the present invention comprise any sequences which have a 5′ promoter region, a structural gene region and a 3′ nontranslated region (polyadenylation site) which can be expressed in plants. The polynucleotide sequences can be modified in any manner (extra codons, deletion of codons, changed codons, *etc.*) as long as gene expression is not prohibited. The polynucleotides inserted into the plants according to the present invention can include any desired gene whether eukaryotic or procaryotic in nature. Usually, more than one gene will be inserted into plant cells which are transformed for agronomic purposes. One gene will typically be a selective marker gene (antibiotic resistance gene or an herbicide resistance gene) in order to easily detect transformants from non-transformed cells. Additional genes can also be added to the plant cell genome to impart an additional property, to suppress an existing property (via "antisense" mechanism) or to amplify a known property of the plant cells and the whole plants regenerated therefrom. The genes can be expressed in specific tissues by the use of tissue specific promoters. The genes can be constructed according to techniques well known to one skilled in the art.

Gene constructs may exist as single gene expression cassettes comprised of a promoter, a structural gene coding sequence and a sequence to permit the addition of poly-adenine (poly-A) residues. The promoter is necessary to initiate transcription of the DNA coding for the structural gene into RNA. The promoter may be derived from a variety of sources, as long as it is functional in the cells to be transformed, and may be modified to enhance expression by the addition or deletion of sequences. The DNA may contain intron sequences, either outside or within the coding region for the protein. The removal of these introns and the addition of the poly-A sequence results in the production of a mature messenger RNA (mRNA) which can be translated into the corresponding protein. Gene expression cassettes may be linked in groups of two or more. Polycistronic expression cassettes, in which a single mRNA may code for more than one protein, may also be used. In addition, expression cassettes may be used to produce an "antisense" RNA from the transcription of a strand of DNA which is opposite to the strand of DNA coding for a protein.

Examples of promoters active in plants include maize ubiquitin promoter (Christensen *et al.*, [1992] *Plant Molecular Biology* 18: 675-689). The untranslated leader sequence, including the first intron, of the maize ubiquitin gene may be incorporated, particularly for use in monocot cells. The 35S promoter of Cauliflower Mosaic Virus (Murray *et al.*, [1991] *Plant Molecular Biology* 16: 1035-1050) or the T-DNA Mas2 promoter of the mannopine synthase gene (Leung *et al.*, [1991] *Molecular & General Genetics* 230: 463-474) may also be used. The 35S promoter may contain a deletion with the addition of an upstream enhancer sequence and an intron in the untranslated leader region (Last *et al.*, [1991] *Theoretical & Applied Genetics* 81: 581-588) and the Mas2 promoter may also contain a deletion to enhance expression (Leung *et al.*, supra).

Examples of structural genes include reporter genes such as that coding for GUS, or -glucuronidase, (Jefferson *et al.*, [1987] *EMBO Journal* 6:3901-3907), a selectable marker gene such as that coding for PAT, or phosphinothricin N-acetyltransferase, which confers resistance to the active ingredient of the commercial herbicide Basta (Droge *et al.*, [1992] *Planta* 187:142-151), or genes which result in expression of a value-added phenotypic trait. Examples of the latter type gene includes those derived from *Bacillus thuringiensis* (B.t.) which confer resistance to insects such as lepidoptera (Adang *et al.*, [1985] *Gene* 36:289-300) or coleoptera (Bradfisch *et al.*, EP 0 500 311 A2 and U.S. Patent No. 5,208,017) and designated B.t.2 and B.t.3 in Figure 1. The B.t. genes may be reconfigured to enhance their expression in plant cells (Adang *et al.*, U.S. Patent No. 5,380,831 issued January 10, 1995).

While herbicide resistance genes serve the purpose of selecting transformants, they also serve the valuable agronomic purpose of allowing herbicide use in the field in otherwise sensitive crops and/or preventing damage to otherwise sensitive crops planted to fields wherein those herbicides were used earlier in the field for weed control (Herbicide carryover). Additional types of genes of value for use in plants include genes isolated from Bacillus thuringiensis that code for delta-endotoxins as well as truncated and/or synthetic derivatives thereof; fungal resistance genes; oil biosynthesis genes; anti-sense genes and genes responsible for nutritional and/or fiber quality.

The DNA and cells are reacted according to the present invention in a suitable buffered medium that is preferably iso-osmotic. See for example, Wong and Neumann's F-medium, Biochemistry and Biophysics Research Communications, Vol. 107, pp. 584-587, 1982; Krens *et al.* [1981] *Nature* 296:72-74. A preferred buffer medium is EPR Buffer (555 mM glucose, 4 mM$CaCl_2$, 10 mM Hepes buffer, pH 7.2). When conducting the present process the fresh plant

cells are added to a reaction vessel followed by the addition of a mixture of DNA and buffer medium. The order of addition of the materials is not critical.

The temperature at which the electroporation process takes place is not critical although it is preferred to cool the DNA/plant cell mixture immediately prior to and after the application of the electric field. Any culturing or regenerating steps are conducted under conditions (including temperature) well known to those of ordinary skill in the art. Heat shock treatments, i.e., 37° C for 10 minutes, of the DNA and plant cell mixture can also be employed.

The electroporation step of the present invention is accomplished by applying an electric field to the DNA/cell mixture according to well-known techniques. Any electric field can be employed. Electric pulses can be from 25-5,000 volts (V) or more depending on the current employed. Preferred methods include rectangular pulse generating systems and capacitor discharge systems. The capacitor discharge system creates pulses of exponentially decaying voltages.

In practicing the present invention, DNA and the plant cells are incubated together at room temperature for at least about 10 minutes and preferably for 20-90 minutes. The DNA/cell mixture is then transferred in aliquots to electroporation cuvettes and optionally cooled on ice prior to applying an electric field to it. The electric field strength can vary depending on a variety of factors, such as, for example, the particular plant species being transformed, the particular type (including age) of cells being employed in the transformation process, the type of electric field being employed including the length of exposure time of the plant cells to the electric field, the concentration and type of DNA, etc. One of ordinary skill in the art can easily determine the optimum process conditions by employing routine titration experiments.

With a capacitor discharge system in the transformation of corn callus cells preferred electroporation conditions include a 250-1500 μF capacitor, 25-500 or more volts and a pulse time of from 50-500 msec. Discharge should be from 25-250 volts. Especially preferred conditions are 850 μF, 150 V and a pulse time of 200 msec. Immediately after application of the electric field the electroporation cuvettes can be optionally placed on ice for about 10 minutes. The cuvettes are then allowed to stand at room temperature for at least about 5 minutes and a small aliquot of cell culture medium is added thereto. Samples are then pipetted from the cuvettes and placed in 2 ml of culture medium in a well of a six-well plate. The treated cells are maintained in culture and regenerated employing standard culturing and plants regenerating techniques.

SAR polynucleotides contemplated within the scope of the present invention encompass known SARs, including functional fragments and allelic variants of a SAR, as well as any SAR

that may be identified in the future so long as the SAR retains substantially the same biological activity as SARs exemplified herein. SARs can be prepared from natural sources or synthesized using standard techniques known in the art, such as an automated DNA synthesizer. The SAR polynucleotides of the subject invention also encompass variant sequences containing mutations in the natural sequences. These mutations can include, for example, nucleotide substitutions, insertions, and deletions as long as the variant SAR sequence retains substantially the same biological activity as the natural SAR sequences of the present invention.

The subject invention also concerns cells and tissue transformed using the methods of the invention. Plants, plantlets, and plant seeds transformed to express heterologous genes according to the methods of the described herein are also contemplated within the scope of the invention.

The disclosure of all references cited throughout the present specification is incorporated herein by reference.

Materials and Methods

PAT-ELISA

Nunc Maxisorb Microtiter plates were coated with 100μl of 1μg/ml protein A purified rabbit anti-PAT IgG in Coating Buffer for two hours while gently shaking at room temperature. After discarding the coating solution, 400μl blocking solution was added and the plates were blocked overnight at 4°C while gently shaking. The plates were then washed using a Skatron SkanWasher plate washer and 100ml of antigen or cell extract was added to the wells and allowed to incubate for two hours at room temperature while gently shaking. Extracts were made by grinding ~100 mgfw of tissue in five volumes Plant Extraction Buffer (PEB) for callus or ten volumes for leaf material and microfuging the crude extract for 5 min. Then the supernatant was transferred and microfuged again for 5 min. Extracts were then diluted 10-fold with grinding buffer prior to adding to the microtiter plate. After washing the plates, 100μl of 1μg/ml protein A purified goat anti-PAT IgG in Ab Buffer was added and the plates incubated for one hour at room temperature with gentle shaking. The plates were washed and 100μl of 1:30,000 dilution of anti-goat antibody conjugated to alkaline phosphatase (Pierce) was added and the plates incubated for one hour at room temperature with gentle shaking. After washing, 200ml of SIGMA FAST pNPP substrate was added and the plates were allowed to sit without shaking for one to three hours with 595nm (reference 405nm) absorbency readings taken every

hour. Quantitation was achieved by comparing the absorbencies with a standard curve made by spiking pure PAT protein into a 10-fold dilution of negative control extract.

Southern Blotting and Hybridization Procedure

Southern Blotting and Pre-hybridization: Southern blotting methodology was performed essentially as described in Murray, M., *et al.*,[1992] *Plant Molecular Biology Reporter,* vol. 10(2). Briefly, genomic DNA (5-10ug) which has been digested with the appropriate restriction enzyme(s) and resuspended in 1X loading buffer is loaded into an agarose/TAE gel (0.85%). The DNA is separated by electrophoresis(75 W/4h) and the gel is then stained (0.1 ug/ml EtBr in 10 mM NaPO) for 30 min. and photographed. The gel is then denatured for 20 min. (150 mM NaPO). The separated DNA is then transferred onto nylon membrane via capillary action overnight. The nylon membrane is then baked for 2 h at 80°C, blocked for 2 h (2% SDS, 0.5% BSA. 1 mM EDTA, 1 mM Orthophenanthroline)and allowed to pre-hybridize for 2 h (100 mM Na phosphate buffer (pH 7.8), 20 mM Na pyrophosphate, 5mM EDTA, 1 mM 1,10 orthophenanthroline,0.1% SDS, 10% dextran sulfate 500 ug/ml heparin sulfate, 50 ug/ml yeast RNA, 50 ug/ml herring sperm DNA).

Hybridization: DNA template to be used as a probe is labeled with P dCTP using a Prime-It RmT Random Priming Labeling Kit (Stratagene). Labeling efficiency of the probe was measured and approximately 1 X 10^6 CPM/ML is added to the prehybridized membrane. The membrane is hybridized with the probe at 65°C for 12-16 h. The membranes are then washed 3X to remove unbound probe (5mM Na phosphate (pH 7.8), 1.25 mM Na pyrophosphate, 0.25 mM EDTA, 0.1% SDS) and exposed to Kodak scientific imaging film.

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1

Several experiments were conducted to determine the transformation efficiency, both transient and stable, for maize cells transformed with and without SARs in an expression cassette. The transformation process conducted was an electroporation process of whole intact maize cells. The results of this experiment are listed in Table 1 below. As can be seen, the

SAR-positive clones (+SARs) had a lower transient transformation efficiency (19.7% vs. 29.49%) but a much higher stable transformation efficiency (3.62% vs. 0.51%).

Table 1

Expt. #	Transient GUS+SARs	Transient GUS-SARs	Stable +SARs	Stable -SARs
ST-1	NA	NA	4/150	--
ST-2	NS	NA	14/135	1/135
ST-3	12.6	29.3	1/25	0/25
ST-4	10.4	32.7	0/20	0/20
ST-5	16.1	37.2	0/25	0/25
ST-6	19.5	28.9	3/90	1/90
ST-7	38.8	28.7	0/50	0/50
ST-8	31.6	36.9	3/155	1/150
ST-9	8.8	12.7	2/95	0.95
TOTALS	137.8	206.4	27/745	5/390
Average	19.7	29.49	3.62%	0.51%

Example 2

ELISA for the detection of phosphinothricin acetyltransferase (PAT) was conducted on several SAR-negative and SAR-positive clones that contained the PAT gene as described in the *Materials and Methods* section herein. Clone T34-1 represents a positive control that was one of the highest PAT-expressing SAR-negative clones. The results of this analysis are shown in Table 2. It can be seen that several SAR-positive clones produce high levels of PAT protein.

Table 2

Callus #	Prot. Adj.	ng PAT	Adj. Value	% of Protein
ST9 neg C	0.017	0.000	0.000	0.0000%
ST9-2 (-SARs)	0.149	-0.010	-0.644	-0.0001%
ST9-3 (-SARs)	0.145	-0.019	-1.314	-0.0001%
ST9-4 (+SARs)	0.099	0.128	13.001	0.0013%
ST9-5 (+SARs)	0.158	0.097	6.140	0.0006%
ST2-2 (+SARs)	0.110	0.025	2.286	0.0002%
ST2-3 (+)	0.169	0.095	5.626	0.0006%
ST2-4 (+)	0.307	0.116	3.772	0.0004%
ST2-5 (+)	0.093	0.039	4.140	0.0004%
ST2-6 (+)	0.160	0.070	4.355	0.0004%
ST2-7 (+)	0.179	0.053	2.934	0.0003%
ST2-8 (+)	0.015	0.018	12.146	0.0012%
ST2-9 (+)	0.165	0.028	1.678	0.0002%
ST2-10 (+)	0.212	0.029	1.381	0.0001%
T34-1 (-)	0.041	0.009	2.210	0.0002%

Example 3

A Southern analysis of several transformed maize callus clones was conducted for the detection of the PAT gene as described in the *Materials and Methods* section herein. The restriction enzyme used was *Bam*HI and the probe was PAT (PCR). Results of Southern blotting experiments, including controls, are shown in Figures 1-3. The results in Fig. 1 show that all the selected calli contained the PAT gene since bands appeared when incubated with a PAT-specific radioactive probe. Figure 2 shows that plants regenerated from selected calliform plants which also contain the PAT gene. Most of the lines show one or a few inserted copies of the PAT gene while one, Lane 24 or ST8-1, showed a complex banding pattern indicative of many PAT gene insertions. Southern analysis of DNA from leaf tissue from plants segregating for the PAT gene in the ST2-1 derived T_1 generation indicated that, with two exceptions, all plants thought to be herbicide sensitive were negative for the presence of the PAT gene (Figure 3). With no exceptions, all plants thought to be herbicide resistant did show the presence of the PAT gene.

This data also shows that the inserted gene was capable of passing through meiosis and into the T_1 progeny.

Example 4

Immature zygotic embryos were isolated from 12-day old cobs of the genotype HiII. They were plated on callus initiation medium (CIM), scutellum side up, which contained N_6 salts, 1 mg/L 2,4-dichlorophenoxyaceticacid (2,4-D), and 10μM $AgNO_3$, incubated in the dark for 8 days at 27-29°C and then examined for Type II callus formation. 250 callused embryos were chosen and placed into sterile 1.5ml micro-cuvettes, at five embryos per cuvette. 20μl of AGM285A plasmid DNA (1μg/μl) was introduced, followed by 200μl of EPM buffer (80 mM KCl, 5 mM $CaCl_2$, 10 mM hepes, 0.425 M mannitol, pH 7.2). One cuvette was inoculated with pAGM243 (comprising the gene construct Ubi-*pat*-Nos::Ubi-*gus*-Nos). Another cuvette contained no plasmid DNA. The contents of the cuvettes were agitated and then allowed to stand for 1 hr at room temperature. Electroporation was carried out using the gene ZAPPER 450/2500 (IBI) at 850μF, 150V for one pulse. The cuvettes were then allowed to stand for 20 min at room temperature before removing the embryos and replating on CIM, scutellum side up, and returned to the dark at 27-29°C. Three days later, the embryos treated with pAGM243 were placed into GUS stain and 9 days later were observed for GUS+ events. Those embryos showed 242 GUS+ events or 48.4 GUS+ events per embryo.

Seven days after treatment, all pAGM285A-treated embryos and embryos subjected to electroporation in the absence of plasmid DNA were transferred to callus maintenance medium (CMM) containing N_6 salts, 1 mg/L 2,4-D and 5μM bialaphos. These were transferred again to the same medium type after 22 days. About three weeks later, the tissue was transferred again.

Example 5

Maize embryos were transformed using the pAGM 607 and pAGM608 constructs. Hill (B X A) or Hill S_1 immature zygotic embroyos of maize were isolated 9-12 days after pollination using standard techniques and plated on callus induction medium (CIM). They were then incubated in the dark at 28°C. Seven or eight days later, the embryos showing pre-embryogenic callus formation were selected and placed into sterile spectrophotometer cuvetttes (MFGR) at five callused embroyos per cuvette. These cuvettes had been prepackaged and sterilized by gassing. Then 20μl of plasmid DNA was added to the cuvette followed by 200μl of EPM buffer. The cuvettes were flicked lightly to distribute the embryos in the cuvette for better coverage by

the DNA/buffer solution. The cuvettes were then allowed to incubate for one hour at room temperature.

We utilized the IBI gene ZAPPER 450/2500 at the electroporation step. Electroporation was conducted at 150V and 850μF. Following electroporation, the cuvettes were allowed to sit at room temperature for 20 min. The callused embryos were then removed and placed on CIM medium, callus side up. It is important that the callus remain above the surface of the medium.

The callused embryos were cultured for one week on CIM medium (N6 sals, 2.88g/l proline, 100mg/l casamino acids, 20 g/l sucrose, modified B5 vitamins, 1 mg/l 2, 4-D, 1mg/L $AgNO_3$, pH 5.8) and then transferred to callus maintenance medium (CMM; CIM medium with no $AgNO_3$) containing 5μM to 10μM Bialaphos. Non-DNA-treated controls were always included and these were split into two groups, one on CMM with Bialaphos and one on CMM without Bialaphos. All callused embryos were transferrred to the same medium every three weeks or so.

Bialaphos-resistant colonies began to appear approximately 10 weeks or so after treatment. These were isolated and put on CMM with 5μM for one more round of selection. Any colonies which grew and had a healthy appearance, i.e. good cream color and embryogenic morphology, were then given a clone number and allowed to proliferate for the purposes of analysis and plant regeneration.

Most of the clones produced from pAGM 607 treatments appeared on selection much earlier than clones produced from pAGM 608 treatment. Clones were observed in as little as 5-6 weeks (Table 2) after treatment whereas normally we see clones between 9 and 12 weeks after treatment (SAR−). Moreover, the majority of SAR+ clones continued to grow rapidly and maintain a desirable phenotype even under high (10μM Bialaphos) selection pressure.

Table 2. Cumulative data concerning SAR+ and SAR- transgenic maize clones

Clone #	+/- SARs	Date Init.	Date Iso.	d to iso	Southern	Pat-ELISA	Plants?	Southern
ST2-1	+	6/28/96	9/10/96	73	?	ND	Yes	Yes
ST2-2		6/28/96	9/10/96	73	Simple	0.0002%	Yes	Yes
ST2-3		6/28/96	9/10/96	73	Simple	0.0006%	Yes	Yes
ST2-4		6/28/96	9/10/96	73	Simple	0.0004%	Yes	Yes
ST2-5		6/28/96	9/10/96	73	Simple	0.0004%	Yes	Yes
ST2-6		6/28/96	9/10/96	73	Simple	0.0004%	Yes	Yes
ST2-7		6/28/96	9/10/96	73	Simple	0.0007%	Yes	Yes
ST2-8		6/28/96	9/30/96	93	Simple	0.0012%	Yes	Yes

ST2-9	6/28/96	11/5/96	129	Simple	0.0002%	Yes	Yes
ST2-10	6/28/96	11/5/96	129	Simple	0.0001%	Yes	Yes
ST2-12	6/28/96	12/18/96	172	ND	0.0507%	Yes	Yes
ST2-13	6/28/96	11/5/96	129	ND	ND	not tested	Yes
ST2-14	6/28/96	12/18/96	172	ND	ND	not tested	Yes
ST8-1	8/13/96		46	Complex	0.0871%	Yes	Yes
ST8-2	8/13/96		46	Simple	0.0194%	Yes	Yes
ST9-4	8/16/96		69	Medium	0.0013%	Yes	Yes
ST9-5	8/16/96		69	Medium	0.0006%	Yes	Yes
Average			92.06				
Std. Dev.			2.83				
ST2-15	6/28/96	12/18/96	172	ND	ND	No	
ST8-4	8/13/96	11/13/96	92	Simple	ND	No	
Average			132				
Std. Dev.			40				

ND=Not done

All but two of the SAR+ clones were highly regenerable, more so than any previous Transfection-derived clones without SARs. Plant recovery was very rapid and several dozen plants were obtained from each clone with little effort. The two pAGM 607 clones that were not readily regenerable, appeared much later than the other 607 clones, did not show a particularly good phenotype and were relatively slow growing. We postulate that the SARs on these inserts had been partially degraded prior to incorporation into the cell's chromosome and thus were not able to exert their positive effects.

The plants derived from pAGM 607 were quite normal in appearance, were very vigorous but flowered and set seed early. The S_1 progeny however, were normal in every way, segregated for Liberty resistance with expected segregation ratios (data not shown) and showed very high Liberty resistance.

Example 6

Field trials with the maize-SAR+ material showed that 13 out of 19 events gave segregation ratios consistent with heterozygous insertions (3:1 ratio in self crosses; 1:1 ratio in outcross to wild type).

SAR Event	Mendelian Ratios
ST2-02	Yes
ST2-04	No
ST2-05	Yes
ST2-06	Yes
ST2-09	Yes
ST2-10	No
ST8-01	Yes
ST8-02	No
ST9-04	Yes
ST9-05	No
ST2-04*CQ715	Yes
ST2-06*CQ715	Yes
ST2-10*CQ715	No
ST8-01*CQ715	Yes
ST9-04*CQ715	No
ST2-1*CQ806	Yes
ST2-05*CQ806	Yes
ST2-06*CQ806	Yes
ST9-05*CQ806	Yes

Moreover, the only SAR+ line tested showed excellent resistance to the Liberty herbicide in the third generation even at eight times the normal rate of field application (27.3 oz/acre of herbicide Remove). This result shows the extraordinary stability of high Pat gene expression even after several generations.

Normally, transgenic plants exhibit a relatively high frequency of gene silencing of the transgene as a given transgenic line is advanced from one generation to the next. Gene silencing is a major problem in the commercialization of transgenic crops. It occurs at a rate of less than 0.1% up to 3% and sometimes more. SARs have now been shown to prevent this phenomenon when used at the 3′ and 5′ ends of the inserted transgene. Transgenic maize lines containing Ubiquitin promoter-phosphinothricin acetyl transferase-now terminator (Ubi-pat-nos) with the Rb7 SAR regions attached at both ends prevent gene silencing in transgenic corn produced by the methods described in this patent or other known transformation methods over several generations in the field.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application.

Claims

1. In a method for transformation of plant cells comprising inserting a polynucleotide of interest into plant cells and maintaining said cells under conditions whereby successfully transformed cells are identified and selected, the improvement comprising providing at least one scaffold attachment region polynucleotide sequence operably linked to said polynucleotide of interest whereby the frequency of recovery of stable transformation events is increased.

2. The improved method according to claim 1, wherein the polynucleotide of interest has at least one scaffold attachment region operably linked to each of its 5′ and 3′ ends.

3. The improved method according to claim 1, wherein said scaffold attachment region polynucleotide comprises the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

4. The improved method according to claim 1, wherein said scaffold attachment region polynucleotide is provided in the form of a DNA expression cassette which can be expressed in a plant.

5. The improved method according to claim 1, wherein said plant transformation process is selected from the group consisting of bombardment with DNA coated microparticles, *Agrobacterium*-mediated transformation, electroporation, microinjection, magnetophoresis, silicon carbide whiskers, PEG mediated transformation, and protoplast transformation.

6. The improved method according to claim 1, wherein said gene of interest is selected from the group consisting of antibiotic resistance genes, herbicide resistance genes, reporter genes, marker genes, disease resistance genes, oil biosynthesis genes, anti-sense genes, insecticide resistance genes, and nutritional enhancement genes.

7. The improved method according to claim 1, wherein said plant cells are selected from the group consisting of embryogenic suspension cells, non-embryogenic suspension cells, plant explant cells, germ line cells, microspore cells, and callus tissue cells.

8. The improved method according to claim 1, wherein said plant cells are of a species selected from the group consisting of cotton, tomato, peanut, alfalfa, rice, wheat, corn, trees, orchard crops, fruits, vegetables, soybean, canola, turfgrass, cannabis, sunflower, sugarcane, sugarbeet, and potato.

9. The improved method according to claim 4, wherein said expression cassette comprises a 5'- promoter region, a structural gene region comprising said polynucleotide of interest, a 3'- nontranslated region, and a SAR flanking either or both sides of the said structural gene region.

10. The improved method according to claim 6, wherein said reporter gene is selected from the group consisting of glucuronidase, phosphinothricin N-acetyltransferase, green fluorescent protein, and luciferase.

11. The improved method according to claim 6, wherein said marker gene is selected from the group consisting of Pat/bar, glyphosate resistance genes, NPTII, HPT, bromoxil resistance genes, AHAS, ALS, cyamamide hydrotase, adenine deaminase, 2, 4-D monooxygenase.

12. A transformed plant cell produced by the method of claim 1.

13. The transformed cell according to claim 12, said cell comprising a scaffold attachment region polynucleotide having the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

14. A transgenic cell descended form the cell of claim 12.

15. A descendant transgenic cell of claim 14, said descendant cell comprising a scaffold attachment region having the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

16. A transgenic plant regenerated from a cell according to claim 12.

17. The transgenic plant according to claim 16, said plant comprising cells which comprise a scaffold attachment region having the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

18. A method for increasing the number of low copy number transformants in a plant transformation process, said method comprising transforming a plant cell or tissue with at least one scaffold attachment region polynucleotide sequence operably linked to a structural gene of interest.

19. The method according to claim 18, wherein said gene of interest has at least one scaffold attachment region operably linked to each of its 5′ and 3′ ends.

20. The method according to claim 18, wherein said scaffold attachment region polynucleotide comprises the nucleotide sequence of SEQ ID NO. 1, or a functional fragment or mutant thereof.

21. The method according to claim 18, wherein said scaffold attachment region polynucleotide is provided in the form of a DNA expression cassette comprising a structural gene region which can be expressed in a plant.

22. The method according to claim 18, wherein said plant transformation process is selected from the group consisting of bombardment with DNA coated microparticles, *Agrobacterium* mediated transformation, and electroporation.

23. The method according to claim 18, wherein said gene of interest is selected from the group consisting of antibiotic resistance genes, herbicide resistance genes, reporter genes, fungal resistance genes, oil biosynthesis genes, anti-sense genes, and insecticide resistance genes.

24. The method according to claim 18, wherein said plant cells are selected from the group consisting of embryogenic suspension cells, non-embryogenic suspension cells, plant explant cells, germ line cells, microspore cells, and callus tissue cells.

25. The method according to claim 18, wherein said plant tissue is selected from the group consisting of callus, meristematic tissue, leaf tissue, shoot tissue, root tissue, and embryonic tissue.

26. The method according to claim 18, wherein said plant cell or tissue is of a species selected from the group consisting of cotton, tomato, peanut, alfalfa, rice, wheat, and corn.

27. The method according to claim 21, wherein said expression cassette comprises a five-prime promoter region, a structural gene region and a three-prime nontranslated region, and a SAR flanking either or both ends of said structural gene.

28. The method according to claim 23, wherein said reporter gene is select from the group consisting of glucuronidase and phosphinothricin N-acetyltransferase.

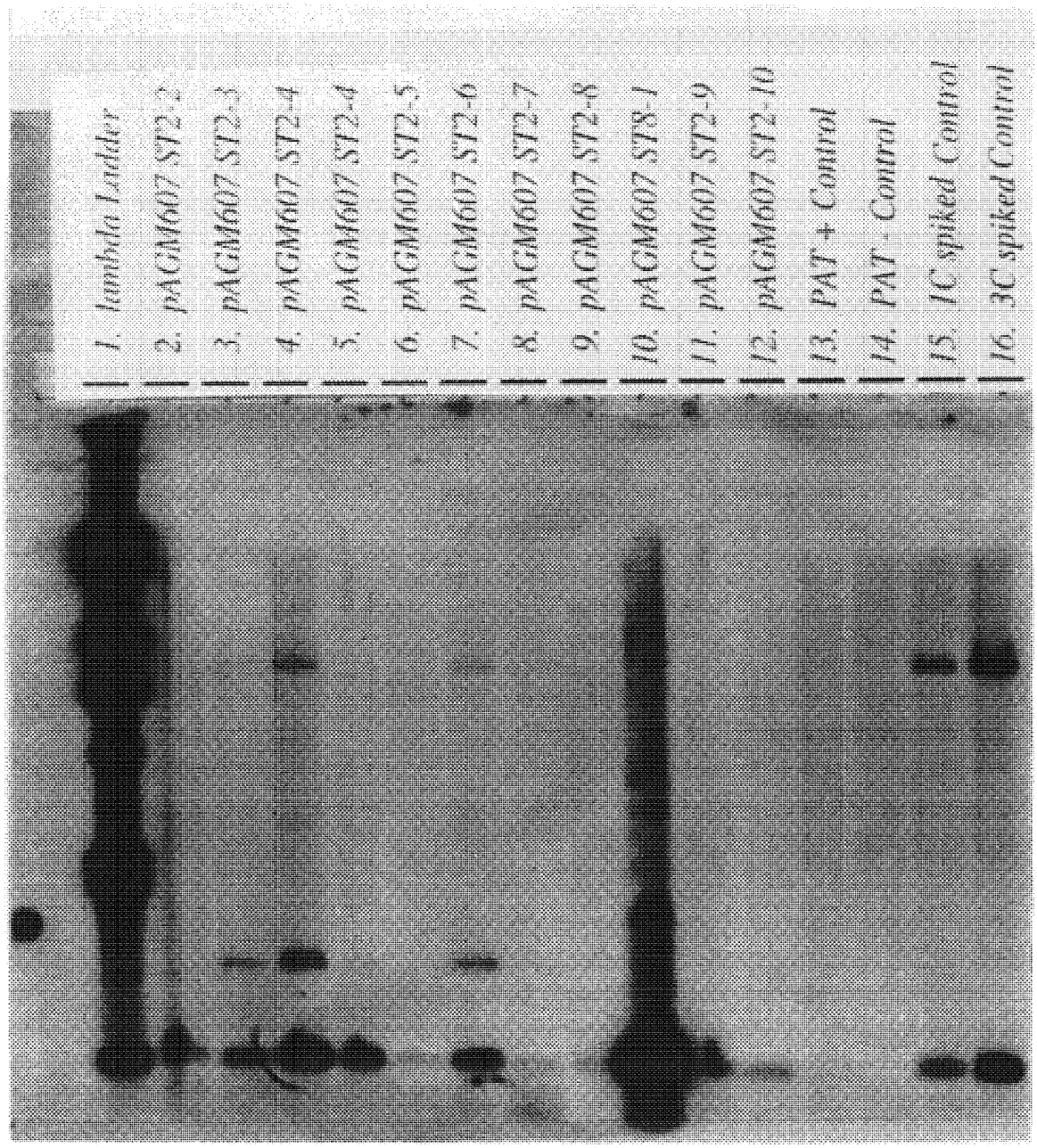

1. lambda Ladder
2. pAGM607 ST2-2
3. pAGM607 ST2-3
4. pAGM607 ST2-4
5. pAGM607 ST2-1
6. pAGM607 ST2-5
7. pAGM607 ST2-6
8. pAGM607 ST2-7
9. pAGM607 ST2-8
10. pAGM607 STS-1
11. pAGM607 ST2-9
12. pAGM607 ST2-10
13. PAT + Control
14. PAT - Control
15. 1C spiked Control
16. 3C spiked Control

FIG. 1

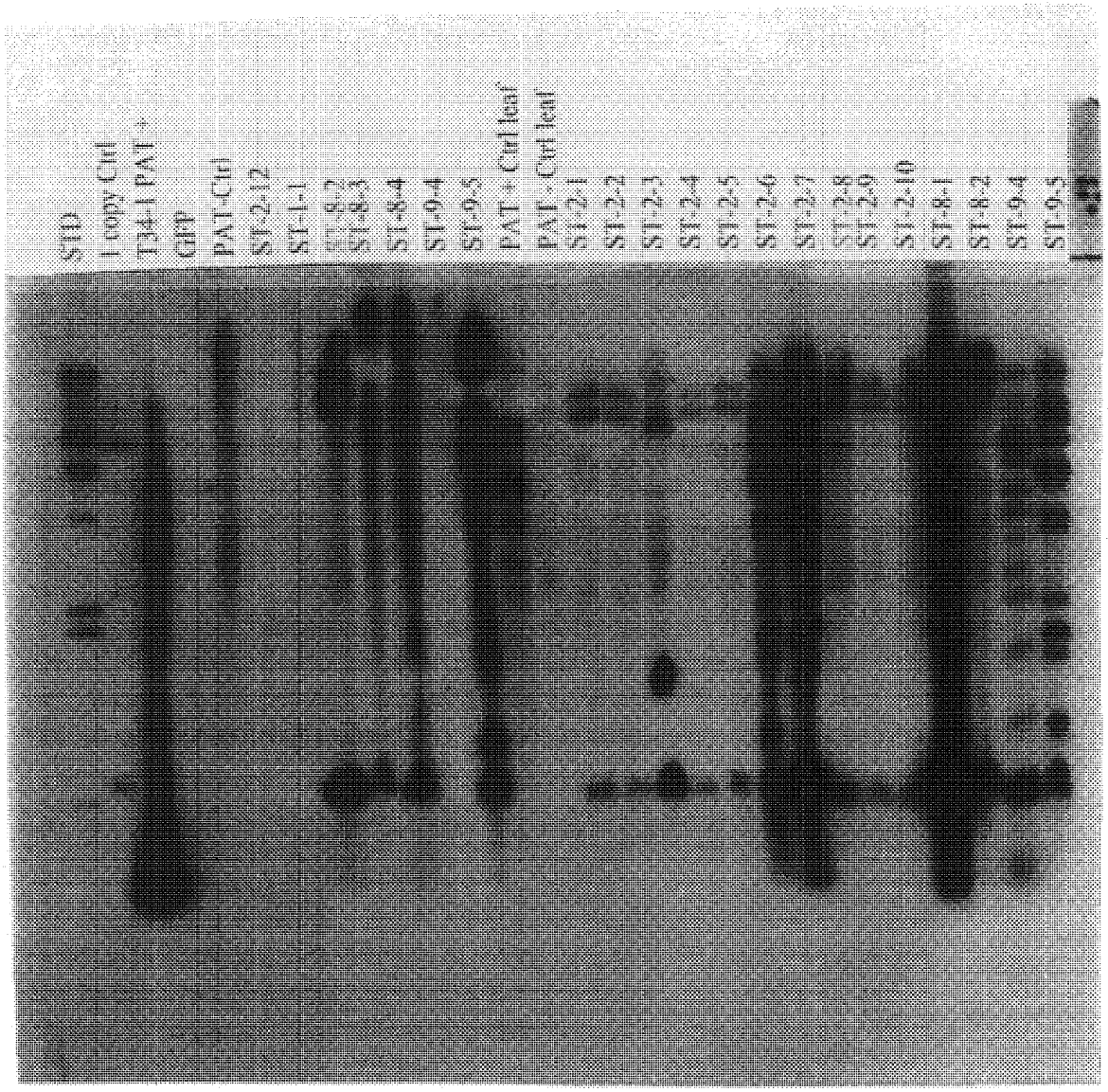


FIG. 2

COR-AAD1-00002264

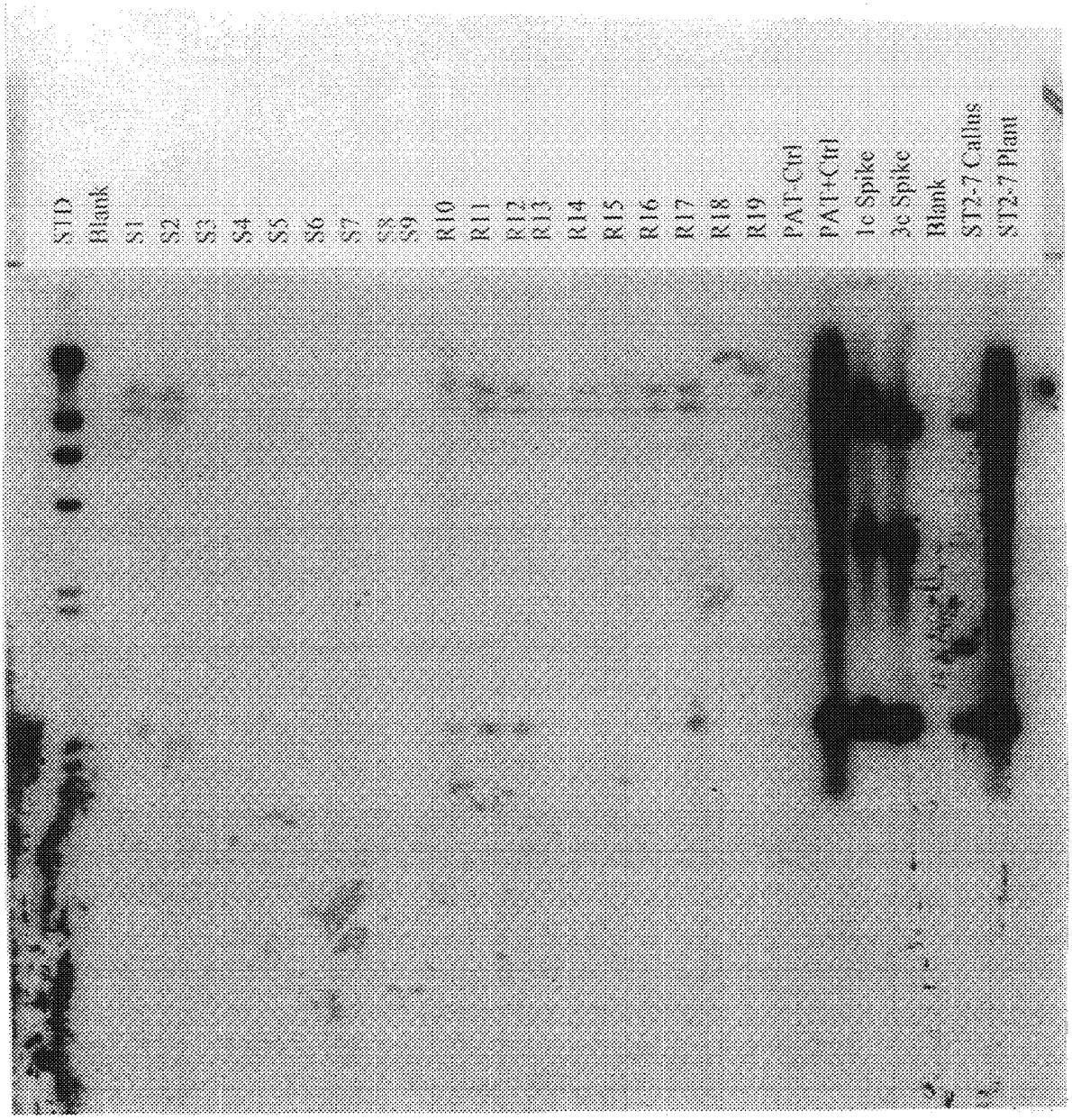

STD
Blank
S1
S2
S3
S4
S5
S6
S7
S8
S9
R10
R11
R12
R13
R14
R15
R16
R17
R18
R19
PAT-Ctrl
PAT+Ctrl
1c Spike
3c Spike
Blank
ST2-7 Callus
ST2-7 Plant

FIG. 3

```
CGATTAAAAA TCCCAATTAT ATTTGGTCTA ATTTAGTTTG GTATTGAGTA AAACAAATTC
GAACCAAACC AAAATATAAA TATATAGTTT TTATATATAT GCCTTTAAGA CTTTTTATAG
AATTTTCTTT AAAAAATATC TAGAAATATT TGCGACTCTT CTGGCATGTA ATATTTCGTT
AAATATGAAG TGCTCCATTT TTATTAACTT TAAATAATTG GTTGTACGAT CACTTTCTTA
TCAAGTGTTA CTAAAATGCG TCAATCTCTT TGTTCTTCCA TATTCATATG TCAAAATCTA
TCAAAATTCT TATATATCTT TTTCGAATTT GAAGTGAAAT TTCGATAATT TAAAATTAAA
TAGAACATAT CATTATTTAG GTATCATATT GATTTTTATA CTTAATTACT AAATTTGGTT
AACTTTGAAA GTGTACATCA ACGAAAAATT AGTCAAACGA CTAAAATAAA TAAATATCAT
GTGTTATTAA GAAAATTCTC CTATAAGAAT ATTTTAATAG ATCATATGTT TGTAAAAAAA
ATTAATTTTT ACTAACACAT ATATTTACTT ATCAAAAATT TGACAAAGTA AGATTAAAAT
AATATTCATC TAACAAAAAA AAAACCAGAA AATGCTGAAA ACCCGGCAAA ACCGAACCAA
TCCAAACCGA TATAGTTGGT TTGGTTTGAT TTTGATATAA ACCGAACCAA CTCGGTCCAT
TTGCACCCCT AATCATAATA GCTTTAATAT TTCAAGATAT TATTAAGTTA ACGTTGTCAA
TATCCTGGAA ATTTTGCAAA ATGAATCAAG CCTATATGGC TGTAATATGA ATTTAAAAGC
AGCTCGATGT GGTGGTAATA TGTAATTTAC TTGATTCTAA AAAAATATCC CAAGTATTAA
TAATTTCTGC TAGGAAGAAG GTTAGCTACG ATTTACAGCA AAGCCAGAAT ACAAAGAACC
ATAAAGTGAT TGAAGCTCGA AATATACGAA GGAACAAATA TTTTTAAAAA AATACGCAAT
GACTTGGAAC AAAAGAAAGT GATATATTTT TTGTTCTTAA ACAAGCATCC CCTCTAAAGA
ATGGCAGTTT TCCTTTGCAT GTAACTATTA TGCTCCCTTC GTTACAAAAA TTTTGGACTA
CTATTGGGAA CTTCTTCTGA AAATAGT
```

FIG. 4

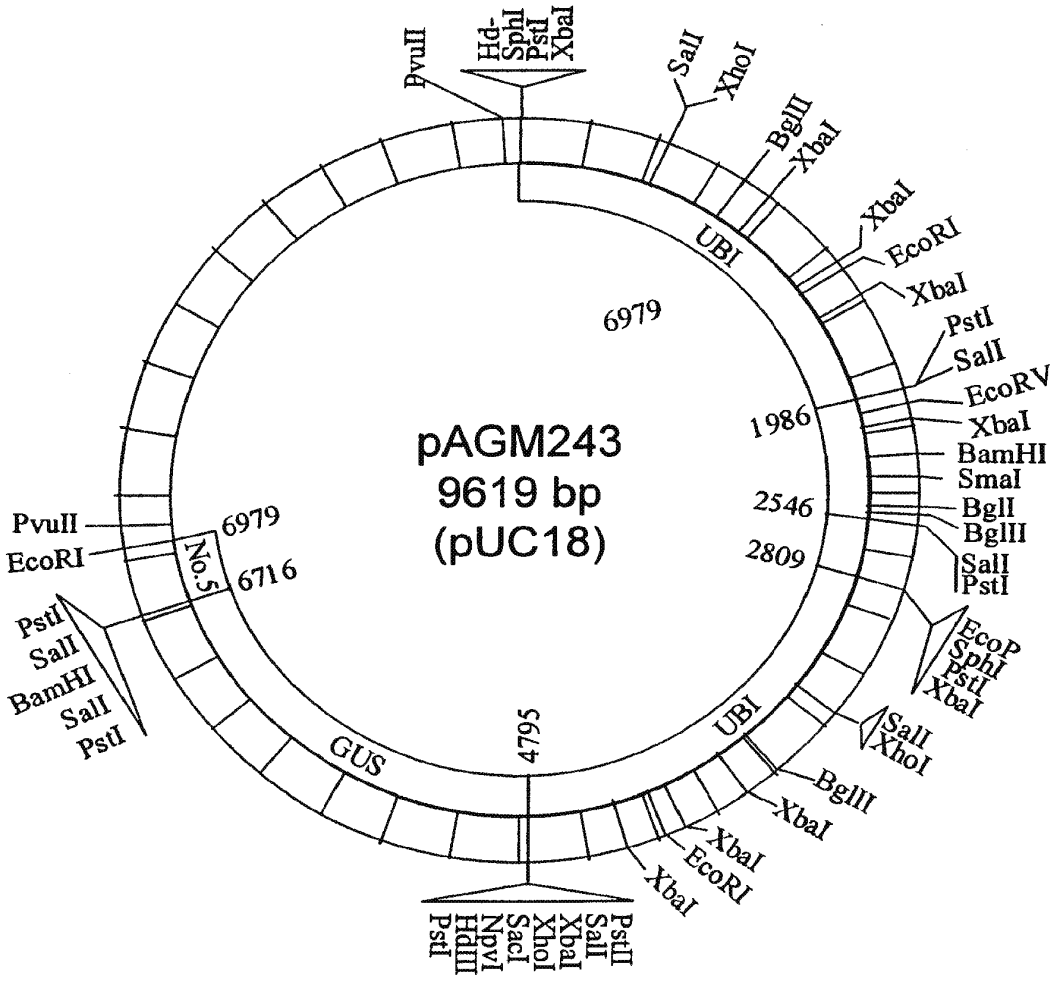

pAGM243
9619 bp
(pUC18)
6979
1986
2546
2809
4795
6716
UBI
UBI
GUS
No.5
PvuII
Hd-
SphI
PstI
XbaI
SalI
XhoI
BglII
XbaI
XbaI
EcoRI
XbaI
PstI
SalI
EcoRV
XbaI
BamHI
SmaI
BglI
BglII
SalI
PstI
EcoP
SphI
PstI
XbaI
SalI
XhoI
BglII
XbaI
XbaI
EcoRI
XbaI
PstII
SalI
XbaI
XhoI
SacI
NpvI
HdIII
PstI
PvuII
EcoRI
6979
PstI
SalI
BamHI
SalI
PstI

FIG. 5

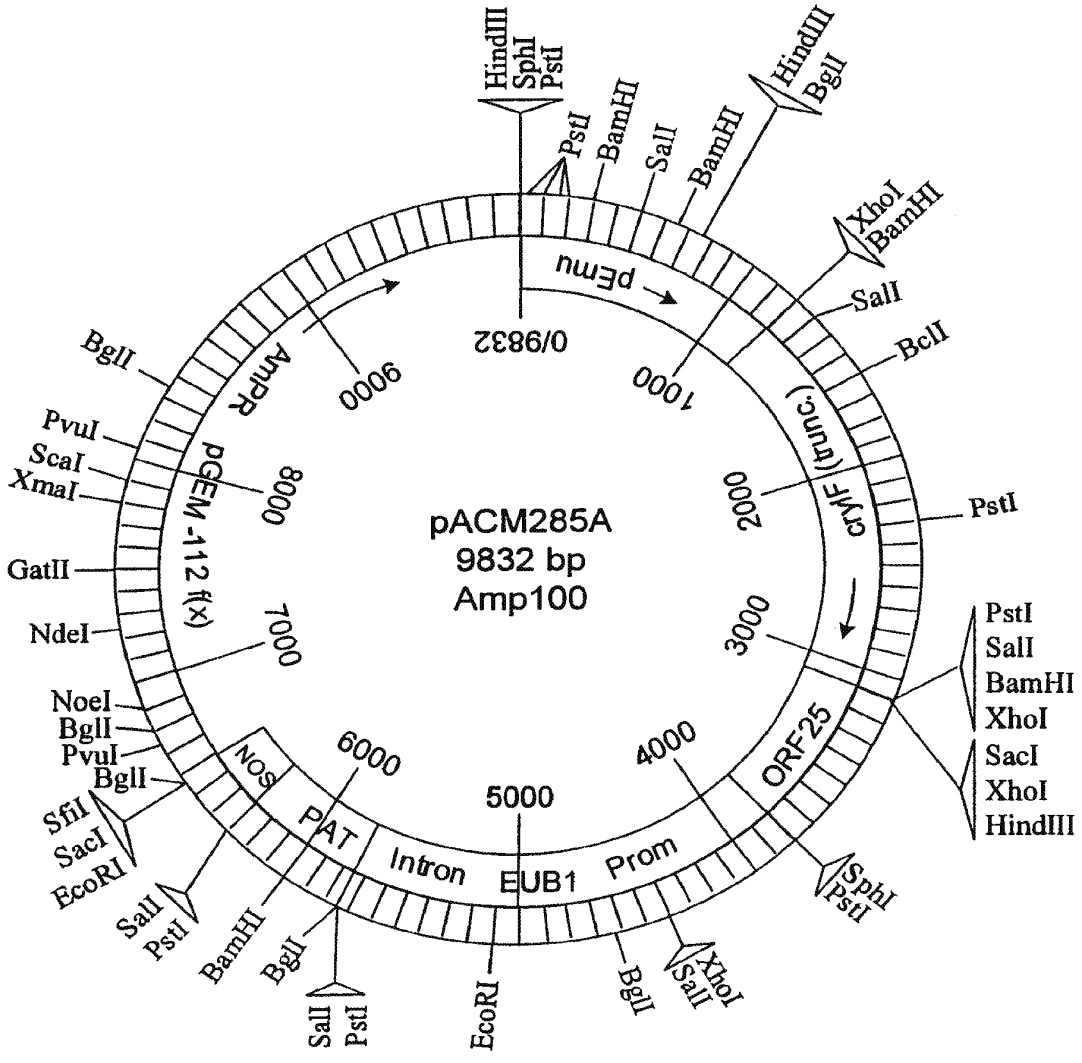

pACM285A
9832 bp
Amp100
0/9832
1000
2000
3000
4000
5000
6000
7000
8000
9000
pEmu
ORF25
Prom
EUB1
Intron
PAT
NOS
pGEM-11Zf(+)
AmpR
orfF (trunc.)
HindIII
SphI
PstI
PstI
BamHI
SalI
BamHI
HindIII
BglI
XhoI
BamHI
SalI
BclI
PstI
PstI
SalI
BamHI
XhoI
SacI
XhoI
HindIII
SphI
PstI
XhoI
SalI
BglI
EcoRI
SalI
PstI
BglI
BamHI
BglI
SalI
PstI
SfiI
SacI
EcoRI
BglI
PvuI
NoeI
BglI
NdeI
GatII
XmaI
ScaI
PvuI
BglI

FIG. 6

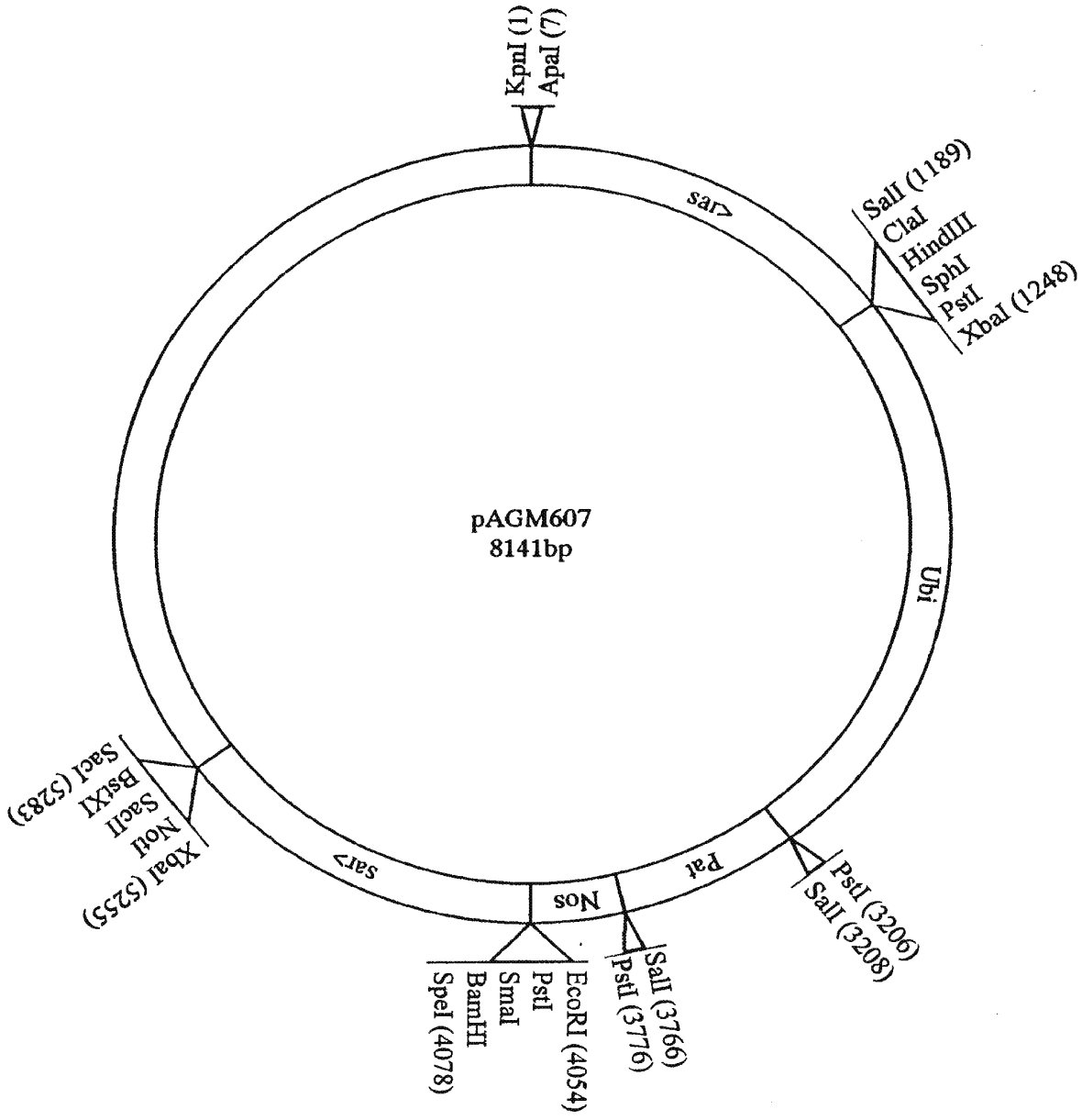

pAGM607
8141bp
KpnI (1)
ApaI (7)
SalI (1189)
ClaI
HindIII
SphI
PstI
XbaI (1248)
Ubi
PstI (3206)
SalI (3208)
Pat
SalI (3766)
PstI (3776)
Nos
EcoRI (4054)
PstI
SmaI
BamHI
SpeI (4078)
sar7
XbaI (5255)
NotI
SacII
BstXI
SacI (5283)
sar7

FIG. 7

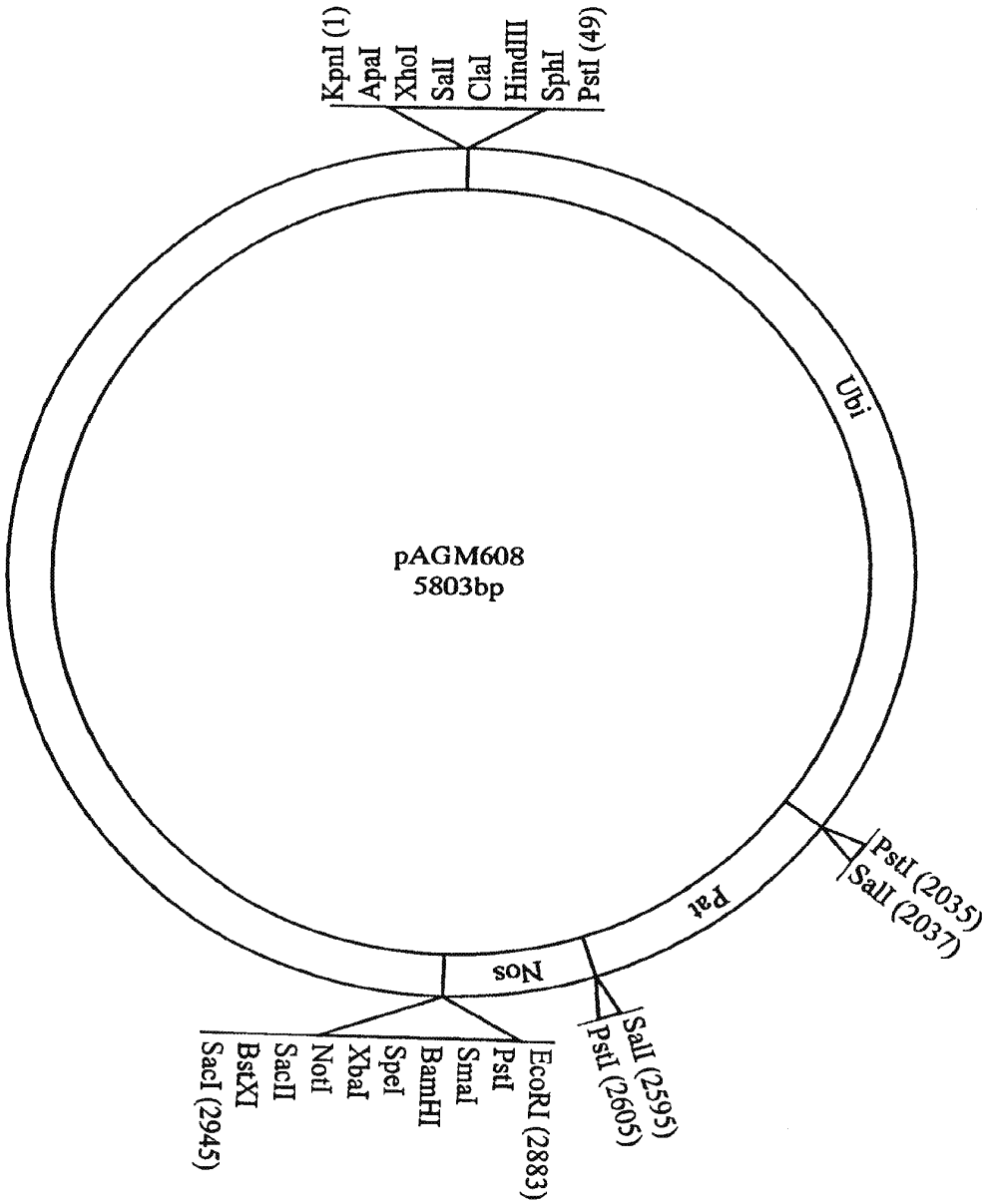

KpnI (1)
ApaI
XhoI
SalI
ClaI
HindIII
SphI
PstI (49)
Ubi
pAGM608
5803bp
PstI (2035)
SalI (2037)
Pat
Nos
SalI (2595)
PstI (2605)
EcoRI (2883)
PstI
SmaI
BamHI
SpeI
XbaI
NotI
SacII
BstXI
SacI (2945)

FIG. 8

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Mycogen Plant Science, Inc.

(ii) TITLE OF INVENTION: Improved SAR Transformation Process

(iii) NUMBER OF SEQUENCES: 1

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
(B) STREET: 2421 NW 41st Street, Suite A-1
(C) CITY: Gainesville
(D) STATE: FL
(E) COUNTRY: USA
(F) ZIP: 32606-6669

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 09/127,080
(B) FILING DATE: 31-JUL-1998

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Lloyd, Jeff
(B) REGISTRATION NUMBER: 35,589
(C) REFERENCE/DOCKET NUMBER: MPS-408

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 352-375-8100
(B) TELEFAX: 352-372-5800

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1167 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGATTAAAAA TCCCAATTAT ATTTGGTCTA ATTTAGTTTG GTATTGAGTA AAACAAATTC 60

GAACCAAACC AAAATATAAA TATATAGTTT TTATATATAT GCCTTTAAGA CTTTTTATAG 120

AATTTTCTTT AAAAAATATC TAGAAATATT TGCGACTCTT CTGGCATGTA ATATTTCGTT 180

AAATATGAAG TGCTCCATTT TTATTAACTT TAAATAATTG GTTGTACGAT CACTTTCTTA 240

TCAAGTGTTA CTAAAATGCG TCAATCTCTT TGTTCTTCCA TATTCATATG TCAAAATCTA 300

TCAAAATTCT TATATATCTT TTTCGAATTT GAAGTGAAAT TTCGATAATT TAAAATTAAA 360

TAGAACATAT CATTATTTAG GTATCATATT GATTTTTATA CTTAATTACT AAATTTGGTT 420

AACTTTGAAA GTGTACATCA ACGAAAAATT AGTCAAACGA CTAAAATAAA TAAATATCAT 480

GTGTTATTAA GAAAATTCTC CTATAAGAAT ATTTTAATAG ATCATATGTT TGTAAAAAAA 540

ATTAATTTTT ACTAACACAT ATATTTACTT ATCAAAAATT TGACAAAGTA AGATTAAAAT 600

AATATTCATC TAACAAAAAA AAAACCAGAA AATGCTGAAA ACCCGGCAAA ACCGAACCAA 660

TCCAAACCGA TATAGTTGGT TTGGTTTGAT TTTGATATAA ACCGAACCAA CTCGGTCCAT 720

TTGCACCCCT AATCATAATA GCTTTAATAT TTCAAGATAT TATTAAGTTA ACGTTGTCAA 780

TATCCTGGAA ATTTTGCAAA ATGAATCAAG CCTATATGGC TGTAATATGA ATTTAAAAGC 840

AGCTCGATGT GGTGGTAATA TGTAATTTAC TTGATTCTAA AAAAATATCC CAAGTATTAA 900

TAATTTCTGC TAGGAAGAAG GTTAGCTACG ATTTACAGCA AAGCCAGAAT ACAAAGAACC 960

ATAAAGTGAT TGAAGCTCGA AATATACGAA GGAACAAATA TTTTTAAAAA AATACGCAAT 1020

GACTTGGAAC AAAAGAAAGT GATATATTTT TTGTTCTTAA ACAAGCATCC CCTCTAAAGA 1080

ATGGCAGTTT TCCTTTGCAT GTAACTATTA TGCTCCCTTC GTTACAAAAA TTTTGGACTA 1140

CTATTGGGAA CTTCTTCTGA AAATAGT 1167

INTERNATIONAL SEARCH REPORT

Inter ional Application No
PCT/US 99/07598

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/82 A01H1/00 A01H4/00 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 95 06128 A (DEKALB GENETICS CORP) 2 March 1995 (1995-03-02) abstract examples figures 1,2 claims 1,15-20,26,27 ---	1,2, 4-12,14, 16,18, 19,21-28
X,P	WO 98 55608 A (UNIV NORTH CAROLINA ;THOMPSON WILLIAM F (US)) 10 December 1998 (1998-12-10) the whole document --- -/--	1-27

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 October 1999	05/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Panzica, G

1

Form PCT/ISA/210 (second sheet) (July 1992)

COR-AAD1-00002273

INTERNATIONAL SEARCH REPORT

Int ional Application No
PCT/US 99/07598

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 16650 A (DU PONT ;KREBBERS ENNO (US); ODELL JOAN TELLEFSEN (US)) 23 April 1998 (1998-04-23) the whole document	1,2, 4-12,14, 16,18, 19,21, 23-27
X	WO 97 04122 A (UNIV WASHINGTON ;LEE JAMES M (US); MAGNUSON NANCY S (US); AN GYNHE) 6 February 1997 (1997-02-06) the whole document	1,2, 4-12,14, 16,18, 19,21, 23-27
X	WO 94 07902 A (UNIV NORTH CAROLINA) 14 April 1994 (1994-04-14) the whole document	1-27
X	WO 97 27207 A (UNIV NORTH CAROLINA ;SPIKER STEVEN (US); ALLEN GEORGE C (US); HALL) 31 July 1997 (1997-07-31) the whole document	1-28
X,P	WO 98 44139 A (MYCOGEN PLANT SCIENCE INC) 8 October 1998 (1998-10-08) the whole document	1-27
X	ALLEN G C ET AL: "SCAFFOLD ATTACHMENT REGIONS INCREASE REPORTER GENE EXPRESSION IN STABLY TRANSFORMED PLANT CELLS" PLANT CELL, vol. 5, no. 6, 1 June 1993 (1993-06-01), pages 603-613, XP002023140 ISSN: 1040-4651 the whole document	1,2, 4-12,14, 16,18, 19,21-27
X	HALL G ET AL: "NUCLEAR SCAFFOLDS AND SCAFFOLD-ATTACHMENT REGIONS IN HIGHER PLANTS" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 88, 1 October 1991 (1991-10-01), pages 9320-9324, XP002002962 ISSN: 0027-8424 cited in the application the whole document	1,2, 4-12,14, 16,18, 19,21, 23-27

-/--

1

COR-AAD1-00002274

INTERNATIONAL SEARCH REPORT

Inte tional Application No
PCI/US 99/07598

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	ALLEN G C ET AL: "HIGH-LEVEL TRANSGENE EXPRESSION IN PLANT CELLS:EFFECTS OF A STRONG SCAFFOLD ATTACHMENT REGION FROM TOBACCO" PLANT CELL, vol. 8, 1 May 1996 (1996-05-01), pages 899-913, XP002910490 ISSN: 1040-4651 the whole document ---	1,2, 4-12,14, 16,18, 19,21, 23-27
X	BREYNE P ET AL: "Characterisation of a plant scaffold attachemt region in a DNA fragment that normalises transgene expression in tobacco" PLANT CELL, no. 4, 1 April 1992 (1992-04-01), page 463 471 XP002072399 ISSN: 1040-4651 the whole document ---	1,2, 4-12,14, 16,18, 19,21, 23-27
X	MAESSEN G D F: "Genomic stability and stability of expression in genetically modified plants" ACTA BOTANICA NEERLANDICA, vol. 1, no. 46, 1 March 1997 (1997-03-01), page 3 24 XP002072400 ISSN: 0044-5983 abstract page 8, paragraph 3 -page 13, paragraph 1 ---	1,2, 4-12,14, 16,18, 19,21, 23-27
X	BROUWER C ET AL: "MARS AND DAM METHYLASE IN MAIZE" JOURNAL OF CELLULAR BIOCHEMISTRY, vol. 21B, 2 April 1995 (1995-04-02), page 129 XP002054632 ISSN: 0730-2312 abstract ---	1,2, 4-12,14, 16,18, 19, 21-23,27
X	GEEST VAN DER A H M ET AL: "THE BETA-PHASEOLIN GENE IS FLANKED BY MATRIX ATTACHMENT REGIONS" PLANT JOURNAL, vol. 6, no. 3, 1 January 1994 (1994-01-01), pages 413-423, XP002054633 ISSN: 0960-7412 abstract column 6, paragraph 3 column 11, paragraph 2 -column 16, paragraph 3 --- -/--	1,2, 4-12,16, 18,19, 21,23-27

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

COR-AAD1-00002275

INTERNATIONAL SEARCH REPORT

Intr tional Application No
PCI/US 99/07598

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	LIU J -W ET AL: "THE INFLUENCE OF TWO PLANT NUCLEAR MATRIX ATTACHMENT REGIONS (MARS) ON GENE EXPRESSION IN TRANSGENIC PLANTS" MOLECULAR AND GENERAL GENETICS, vol. 39, no. 1, 1 January 1998 (1998-01-01), pages 115-123, XP002910494 ISSN: 0026-8925 the whole document	1,2, 4-12,14, 16,18, 19,21, 23-27
X	CLAPHAM D ET AL: "ENHANCEMENT OF SHORT- AND MEDIUM-TERM EXPRESSION OF TRANSGENES IN EMBRYOGENIC SUSPENSIONS OF PICEA ABIES (L.) KARST" JOURNAL OF EXPERIMENTAL BOTANY, vol. 46, no. 287, 1 June 1995 (1995-06-01), pages 655-662, XP002910489 ISSN: 0022-0957 abstract column 3, paragraph 2	1,2, 4-12,16, 18,19, 21,23-27
X	HAN K -H ET AL: "MATRIX ATTACHMENT REGIONS (MARS) ENHANCE TRANSFORMATION FREQUENCY AND TRANSGENE EXPRESSION IN POPLAR" TRANSGENIC RESEARCH, vol. 6, 1 January 1997 (1997-01-01), pages 415-420, XP002910491 ISSN: 0962-8819 the whole document	1,2, 4-12,14, 16,18, 19,21, 23-27
A	VAIN P ET AL: "Foreign gene delivery into monocotyledonous species" BIOTECHNOLOGY ADVANCES, vol. 13, no. 4, 1 January 1995 (1995-01-01), page 653-671 XP004044845 ISSN: 0734-9750	5-8

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

COR-AAD1-00002276

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/07598

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9506128 A	02-03-1995	AU 5640498 A	04-06-1998
		AU 684105 B	04-12-1997
		AU 7716994 A	21-03-1995
		BR 9407355 A	19-08-1997
		CA 2170260 A	02-03-1995
		EP 0721509 A	17-07-1996
		HU 74392 A	30-12-1996
		US 5780709 A	14-07-1998
		ZA 9406488 A	30-11-1995
		ZA 9604217 A	26-08-1996
WO 9855608 A	10-12-1998	AU 7714098 A	21-12-1998
WO 9816650 A	23-04-1998	AU 4893397 A	11-05-1998
		EP 0931155 A	28-07-1999
WO 9704122 A	06-02-1997	AU 6504796 A	18-02-1997
WO 9407902 A	14-04-1994	AU 673859 B	28-11-1996
		AU 5165593 A	26-04-1994
		CA 2147006 A	14-04-1994
		EP 0663921 A	26-07-1995
		US 5773689 A	30-06-1998
WO 9727207 A	31-07-1997	US 5773695 A	30-06-1998
		AU 2246997 A	20-08-1997
		BR 9707208 A	06-04-1999
		CA 2244204 A	31-07-1997
		EP 0904276 A	31-03-1999
WO 9844139 A	08-10-1998	AU 6871698 A	22-10-1998

COR-AAD1-00002277

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 May 2001 (31.05.2001)

PCT

(10) International Publication Number
WO 01/38513 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/US00/32298

(22) International Filing Date:
22 November 2000 (22.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/167,452 23 November 1999 (23.11.1999) US
60/202,233 5 May 2000 (05.05.2000) US

(71) Applicant *(for all designated States except US)*: **MAXYGEN, INC.** [US/US]; 515 Galveston Drive, Redwood City, CA 94063 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CASTLE, Linda, A.** [US/US]; 784 Hans Avenue, Mountain View, CA 94040 (US). **LASSNER, Michael** [US/US]; 721 Falcon Avenue, Davis, CA 95616 (US). **MCBRIDE, Kevin** [US/US]; 1309 Marina Circle, Davis, CA 95616 (US). **ENGLISH, James** [US/US]; 1249 Capuchino Avenue, Burlingame, CA 94010 (US).

(74) Agents: **QUINE, Jonathan, Alan**; The Law Offices of Jonathan Alan Quine, P.O. Box 458, Alameda, CA 94501 et al. (US).

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

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

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

WO 01/38513 A2

(54) Title: SHUFFLING OF AGROBACTERIUM AND VIRAL GENES, PLASMIDS AND GENOMES FOR IMPROVED PLANT TRANSFORMATION

(57) Abstract: Methods for evolving plant vectors with improved characteristics by recursive recombination are provided. Plant vectors that are RNA or DNA polynucleotides, conjugated-DNA polynucleotides, and plasmids are provided, as are vectors that are agrobacterium strains and plant viruses. Agrobacterium vectors that have evolved such desired properties as broad host range, increased transformation efficiency, insert precision, targeted insertion, and targeting of T-DNA sequences to the chloroplast are provided. Agrobacterium strains, which are amenable to transforming a broad range of host species using simple transformation techniques such as vacuum infiltration or direct infection in planta, are provided. Plant virus vectors are provided that have evolved desired properties, including: rapid systemic spread, reduction of symptoms, and increased protein expression. Use of the evolved vectors to produce transgenic plants is provided. Methods and vectors for producing proteins in transgenic plants and for conferring pathogen-derived resistance are provided.

SHUFFLING OF AGROBACTERIUM AND VIRAL GENES, PLASMIDS AND GENOMES FOR IMPROVED PLANT TRANSFORMATION

COPYRIGHT NOTIFICATION PURSUANT TO 37 C.F.R. § 1.71(E)

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to and benefit of United States application number 60/167,452, filed 11/23/99, and United States application number 60/202,233, filed May 5, 2000, which are incorporated in their entirety for all purposes.

FIELD OF THE INVENTION

The present invention relates to vectors for use in the production of transgenic plants, and to methods of producing novel plant transformation vectors.

BACKGROUND OF THE INVENTION

Genetic modification of plant species utilizing the techniques of molecular biology has become widespread in recent years. A number of methods exist for the introduction of exogenous gene sequences into the genetic material of plants. The choice of technique is largely dependent on the application: whether stable heritable changes are sought or whether transient expression is desired, and on the plant species in question. Each technique offers a unique set of advantages and disadvantages.

For example, for plant species which can be regenerated from cultured protoplasts, electroporation, the PEG method and microinjection are all suitable for introducing exogenous DNA into the protoplast nucleus. Electroporation techniques involve mixing plant protoplasts with DNA polynucleotides and subjecting the mixture to an electric pulse. The electric pulse causes the formation of holes in the protoplast cell membranes through which DNA can pass. It is a highly reproducible technique and can be adapted to both dicot and monocot species. However, it can be used only with species for which protoplast culture and regeneration protocols exist. Because it relies on the

COR-AAD1-00002279

regeneration of protoplasts, production of transgenic plants is a lengthy process, often requiring several months.

The polyethylene glycol (PEG) method, while similar to electroporation in that DNA and protoplasts are mixed in vitro, relies on changes in cell membrane permeability induced by the addition of PEG to the medium. This method is simple and economical, as it requires no specialized equipment for its success. However, like electroporation, it requires that protoplast culture and regeneration techniques exist for the plant species in question.

Microinjection techniques, in which DNA or RNA polynucleotides are injected directly into the nucleus of protoplasts, are also applicable to a wide variety of species. A drawback of this technique is that it is time consuming and success is highly dependent on the skill of the individual experimentalist. Thus it is not widely applicable to the large scale generation of modified crops or other plants of commercial interest.

So-called "biolistic," or particle mediated bombardment techniques, shoot DNA-coated metal particles into plant cells or tissues at high speed. Thus, biolistic techniques are at least theoretically applicable to any tissue or species, regardless of whether culture and regeneration techniques are available. Transformation efficiency is variable and depends on selection of successful transformants.

A wide variety of dicotyledons can be readily transformed using agrobacterium mediated transformation. Agrobacterium mediated transformation relies on the ability of *A. tumefaciens* or *A. rhizogenes* to transfer DNA molecules called T-DNA to a host plant cell. *A. tumefaciens* and *A. rhizogenes* are the causative agents of the plant neoplastic diseases crown gall and hairy root disease, respectively. Agrobacteria, which reside normally in the soil, detect soluble molecules secreted by wounded plant tissues through a specialized signal detection/transformation system. In the presence of these chemical signals, agrobacteria attach to the cell walls of wound exposed plant tissues. The agrobacteria then excise and transfer a portion of specialized DNA, designated T-DNA and delimited by "T-DNA borders," to the host plant cell nucleus where it is integrated into the chromosomal DNA. This DNA transfer system can be manipulated to transfer exogenous DNA situated between T-DNA borders to a host plant cell of choice.

However, monocots, especially those of commercially valuable species belonging to the family *graminae*, have proved less amenable to transformation by agrobacteria. Agrobacterial vectors which extend host range and increase transformation

COR-AAD1-00002280

efficiency are of significant interest in adapting agrobacterium mediated transformation to presently recalcitrant plant species. Similarly, agrobacterium strains which are capable of introducing T-DNAs into a plants other than Arabidopsis by simple methods, such as vacuum infiltration, which do not require culture in vitro, would significantly reduce the material and labor costs associated with generating genetically modified plants.

Viral vectors derived from a large selection of plant viruses have proven adaptable for use in the transformation of dicot and monocot species. The majority of plant viruses have genomes consisting of RNA polynucleotides, although a number of DNA viruses, including cauliflower mosaic virus, have been widely used. With the advent of simple in vitro transcription protocols, transformation by RNA viruses has become widespread. Typically, viral replication and expression occurs in the plant cytoplasm. Therefore virally mediated techniques are most useful where transient, rather than stable, genetic changes are the goal. For example, high levels of transient expression, suitable for the production and recovery of exogenous proteins can be achieved using plant viral vectors.

Although expression of exogenous sequences via viral vectors is of significant benefit, it is not without risk. Viral vectors currently employed are capable of spreading not only within a host plant but also between plants. This makes unintentional infection of neighboring plants a potential hazard, restricting the use of many vectors to non-food crop plants such as tobacco. Thus, the development of plant viral expression vectors which safely infect valuable food and other crop species is of great interest.

The present invention provides methods for developing novel and improved plant vectors for use in the production of transgenic plants. These methods are applicable to polynucleotide vectors suitable for introduction by electroporation, PEG, microinjection, particle mediated bombardment; plasmids and agrobacterium strains for use in agrobacterium mediated transformation and novel plant virus vectors. Vectors can be developed using the methods of the invention which address the issues desicribed above, as well as other disadvantages.

SUMMARY OF THE INVENTION

The present invention utilizes recursive recombination, e.g., "DNA shuffling," techniques to develop new plant vectors with improved properties. The vectors of the invention include polynucleotides such as naked RNA and DNA, conjugated DNA molecules, and plasmids, as well as plant viruses and agrobacterium

COR-AAD1-00002281

strains. The methods of the invention provide for the recombination, in silico, in vitro or in vivo, of DNA fragments corresponding to plant vector components and their homologues. One or more cycle of recursive recombination is followed by screening or selection to identify recombinant vector components, or vectors, which have evolved towards a desired property. Multiple reiterations of the recombination (e.g., via shuffling) process can be employed to achieve the property desired. In some embodiments, diversity among substrate fragments is increased using error prone PCR, mutagenic amplification and/or site directed mutagenesis. Preferred embodiments include vectors which are agrobacterium, and agrobacterium derived plasmids. Other preferred embodiments are plant viruses.

Methods of the invention use various screening or selection protocols to identify recombinant vectors or vector components with desired properties. Embodiments of the invention employ a variety of screening techniques, including: PCR, LCR, hybridization, proteomics and detection of reporters, (such as green fluorescent protein (GFP), β-glucuronidase (GUS), luciferase, or proteins incuded by the maize *Lc* gene). Screening techniques of the invention also include selection protocols including herbicide resistance and antibiotic resistance. In some cases, the selection protocols rely on detection of negatively selectable markers, such as *dhl1, codA, tms2* and *NIA2*, and their homologues.

The methods of the invention provide for the generation of Agrobacterium vectors which have evolved new and improved properties. Such properties include, but are not restricted to: insert precision, targeted insertion, broad host range, increased transformation efficiency targeting of T-DNA to the chloroplast, and improved amenability to simplified transformation techniques. These methods pertain to Agrobacterium strains, and to agrobacterium-derived plasmids, including the plasmids of binary vector systems. Embodiments of the invention provide for binary vector systems that supply necessary virulence functions on a plasmid, or alternatively, integrated into an Agrobacterium chromosome. In some embodiments, one or more agrobacterium T-DNA borders are included in the fragments to be shuffled. In preferred embodiments, the T-DNA borders are contiguous with PCR primer binding sites. In some embodiments, the fragments include agrobacterium virulence genes, e.g., *virA, virB, virC, virD, virE, virG* and *ChvE*. The fragments optionally comprise an entire agrobacterium genome.

COR-AAD1-00002282

Alternative embodiments of the invention provide for Agrobacterium vectors which include *A. tumefaciens* or *A. rhizogenes*.

Preferred embodiments of the invention relate to the development of vectors which insert T-DNA sequences precisely into the host plant chromosome, or insert T-DNA sequences into a predetermined site in the host plant chromosome. In one preferred embodiment, regions of sequence similarity with a desired insertion site flank the T-DNA border or borders. In another preferred embodiment, a promoter or enhancer is included adjacent to the region of sequence similarity. The invention also provides for screening the recombinants of such embodiments in host plants or plant cells having detectable markers which are activated by insertion of the evolved vector at a predetermined site. An especially preferred embodiment relates to screening such recombinants in transgenic plants.

Alternative embodiments of the invention are directed towards evolution of agrobacterium vectors with broad host ranges, including both monocot and dicot plant species. In a preferred embodiment, virulence gene sequences are recombined, e.g., shuffled, to evolve agrobacterium vectors with a broad host range. *VirA* fragments are optionally recombined and the resulting recombinant sequences assessed for their ability to activate a reporter linked to a *vir* promoter. Embodiments include methods of screening for improved host range in intact plants and in plant cell cultures.

Methods for evolving agrobacterium vectors which target T-DNA sequences to the chloroplast are also a feature of the invention. In preferred embodiments, *virD2* and/or *virE2* genes are the substrates for nucleic acid diversification (e.g., DNA shuffling) procedures, optionally including one or more of a nuclear localization signal, and a chloroplast localization signal. In an especially preferred embodiment, screening for chloroplast targeting is performed by detecting a marker localized to the chloroplast.

The invention provides methods of producing plant virus vectors with novel, desired properties. The methods of the invention are equally applicable to both RNA and DNA viruses. Properties such as rapid systemic spread; reduction of viral symptoms; increased viral and exogenous gene expression; and production of proteins restricted to transgenic host plants are among those provided for by the methods of the invention.

A method for evolving a plant virus vector with rapid systemic spread is an embodiment of the invention. An embodiment provides for inoculating viruses

COR-AAD1-00002283

incorporating the recombinant library generated by DNA shuffling onto leaves of an intact plant followed by recovery and analysis of viruses from leaves distal to the site of inoculation. Viral movement proteins or complete viral genomes are optionally provided as substrates for nucleic acid diversification, e.g., by DNA shuffling.

Other embodiments of the invention relate to methods of evolving viral vectors which result in a reduction in symptoms following infection. Preferred targets include viral coat proteins, viral movement proteins, viral replicases and viral recombinases, as well as entire viral genomes.

The invention includes methods for developing virus vectors which give rise to an increase in expression of either viral or exogenous DNA sequences. Optionally viral or plant promoters are provided as substrates for diversification, e.g., by DNA shuffling. For example, the promoters can be, e.g., a cauliflower mosaic virus 35S promoter. Alternatively, fragments constituting a viral genome are used as the substrate for generating recombinant libraries. A marker gene is optionally cloned between recombinant, e.g., shuffled, 5' and 3' regions of a viral genome and the resulting recombinant virus is inoculated onto the leaf of a host plant.

Methods for evolving vectors which confer pathogen-derived resistance to plant pathogens are also features of the invention, as are plant virus vectors which confer pathogen-derived resistance to plant pathogens.

A feature of the present invention is the construction of novel viral vectors for protein production in transgenic plants. Screening of such viral vectors is accomplished by expressing a recombinant virus which lacks a coat protein in a transgenic plant which supplies the missing coat protein in trans. Another aspect of the invention is a virus vector which has a polylinker in place of a coat protein. Preferred embodiments include vectors which have an exogenous polynucleotide sequence cloned into the polylinker and expressed under regulatory control of the coat protein promoter. Especially preferred embodiments include vectors which express polynucleotides that encode proteins which are biofuels, industrial enzymes, nutritional enzymes, nutraceuticals, pharmaceuticals, plantibodies, antigens, biosynthetic enzymes, or the like. A preferred embodiment is a virus which is a tobamovirus. Tobamoviruses which infect *Brassicas*, including *Arabidopsis* as well as tobamoviruses which infect legumes and other plants, including monocots are embodiments of the invention. Another feature of the invention is a method of producing proteins in transgenic plants using the recombinant plant virus lacking a coat protein of the invention.

COR-AAD1-00002284

In some embodiments, the methods for evolving improved plant vectors involve diversification and screening or selection of RNA viral vectors in vivo. A plurality of RNA viral vectors containing genes of interest are introduced into a cell and the cells are grown under conditions permitting replication and recombination of the viral sequences. Optionally, the viral vectors are recovered, and the recombination is performed recursively. After recombination of the viral RNA molecules, a viral vector comprising a gene with a desired property is identified. The viral vectors are introduced into cells by inoculating the cell with infectious RNA transcripts. Alternatively, a plurality of cDNA molecules corresponding to viral transcripts are used to introduce the genes of interest into the cell. In the latter case, the plurality of cDNA molecules can be introduced by a variety of techniques including, electroporation, microinjection, biolistics, agrobacterium mediated transfromation or agroinfection. In preferred embodiments the RNA viral vectors are plant virus vectors, and the cells are plant cells. Such vectors include, tabamoviruses, potyviruses, tobraviruses, and potexviruses. In some embodiments, the plant cells are isolated cells grown in culture. In other embodiments, the plant cells are plant protoplasts, plant tissues, plant organs or intact plants.

In an embodiment, two viral vectors having complementary mutations in proteins involved in systemic infection are used to introduce nucleic acids comprising genes of interest. Upon recombination, infectivity is restored, thereby facilitating selection of recombinant genes of interest. Exemplary proteins involved in systemic infection include viral coat proteins and viral movement proteins.

Another aspect of the invention relates to libraries of recombinant sequences generated by nucleic acid diversification procedures, such as DNA shuffling. Such libraries may be vectors, vector components, vector genomes, or the like. For example, libraries composed of agrobacterial or viral sequences are included. Agrobacterium strains and virus vectors incorporating recombinant, e.g., shuffled library sequences are also a feature of the invention. Similarly, libraries derived from transgenic plants transformed by the vectors of the invention are a feature of the invention.

Vectors produced by the methods of the invention are an aspect of the invention. Preferred embodiments include vectors which are agrobacterium plasmids or agrobacterium strains. Other embodiments of the invention are evolved plant viruses. Preferred embodiments refer to evolved plant viruses which include an exogenous DNA sequence. A preferred embodiment has a library polynucleotide inserted to give rise to a

COR-AAD1-00002285

protein that is expressed as a fusion with a coat protein. Optionally, the polypeptide encoded by the inserted polynucleotide is expressed on the outside surface of a virus particle. In alternative embodiments, the exogenous polynucleotide sequence is a viral or plant promoter linked to a reporter gene, such as GFP, luciferase, β-glucuronidase, or the maize *Lc* regulatory gene. Other embodiments include vectors which have incorporated exogenous polynucleotide sequences which are selectable markers. Such selectable markers can be genes which confer resistance to an antibiotic or to a herbicide. Alternatively, the selectable markers can be negatively selectable markers, including homologues of the *dhl1, codA, tms2* and *NIA2* genes. Other embodiments of the invention include evolved plant viruses which express exogenous DNA sequences under the control of a cauliflower mosaic virus 35S promoter. Alternative embodiments relate to evolved plant virus vectors which are positioned adjacent to agrobacterium borders and introduced into a plant host by Agroinfection. The invention also provides for introducing evolved viral vectors into plant protoplasts by other methods including electroporation, microinjection, biolistics and direct mechanical inoculation.

Another aspect of the present invention includes plants, plant cells and plant explants transformed by the vectors of the invention to produce transgenic plants. Additionally, libraries of sequences derived from transgenic plants produced using the vectors of the invention are a feature of the invention.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 schematically illustrates a replicating Geminivirus vector system.

Figure 2 schematically illustrates recombination of Geminivirus genome to alter size selectivity of the movement proteins BR1 and BL1.

Figure 3 schematically illustrates production of a recombinant Geminivirus genome library in a common region (CR) cassette.

Figure 4 schematically illustrates control (CV) and test (TV) reporter vectors for relaxing size selectivity of Geminivirus movement proteins.

Figures 5A and 5B schematically illustrate recombination of Geminivirus genomes including the common region (CR), leftward elements, including the coat protein promoter (P_{CP}), and Replicase (AC1) gene. The vecors illustrated in figures 5A and 5B differ by the position of a restriction enzyme recognition site (RS1).

COR-AAD1-00002286

Figure 6 schematically illustrates production of a recombinant Geminivirus genome library in alternative common region cassettes lacking and possessing a visual reporter.

DETAILED DISCUSSION OF THE INVENTION

The present invention relates to the production of plant vectors with improved characteristics. Recursive recombination, e.g., DNA shuffling, is employed to develop plant vectors with novel and improved properties. Some embodiments of the invention are agrobacterium vectors, while other embodiments include evolved plant viruses. Preferred embodiments include agrobacterium vectors which extend the host range, and increase transformation efficiency, as well as which improve insert precision and simplify transformation techniques. Other preferred embodiments are viral vectors which have evolved properties such as rapid systemic spread, altered host range, and reduction of viral symptoms after infection. Vectors which facilitate the production of commercially valuable products in transgenic plants, while simultaneously reducing the risk of contamination of neighboring crops are provided. Vectors which confer pathogen-derived resistance to plant pathogens are provided. Another aspect of the invention relates to the production of transgenic plants using the evolved vectors and to the transgenic plants produced thereby. In one aspect, a transgenic plant expressing a viral coat protein is used to produce a vector encoded polypeptide.

DEFINITIONS

Unless defined otherwise, all scientific and technical terms are understood to have the same meaning as commonly used in the art to which they pertain. For the purpose of the present invention, the following terms are defined below.

The term "shuffling" is used herein to indicate recombination between non-identical sequences. For example, "DNA shuffling" involves recombination of deoxyribonucleic acid (DNA) segments. In some embodiments shuffling may include crossover via homologous recombination or via non-homologous recombination, such as via cre/lox and/or flp/frt systems. Shuffling can be carried out by employing a variety of different formats, including for example, *in vitro* and *in vivo* shuffling formats, in silico shuffling formats, shuffling formats that utilize either double-stranded or single-stranded templates, primer based shuffling formats, nucleic acid fragmentation-based shuffling formats, and oligonucleotide-mediated shuffling formats, all of which are based on recombination events between non-identical sequences and are described in more detail

COR-AAD1-00002287

or referenced herein below, as well as other similar recombination-based formats. In one class of embodiments, nucleic acid shuffling involves the recursive recombination of nucleic acid sequences.

"Screening" is, in general, a two-step process in which one first determines which cells, organisms or molecules, do and do not express a detectable marker, or phenotype (or a selected level of marker or phenotype), and then physically separates the cells, organisms or molecules, having the desired property. Selection is a form of screening in which identification and physical separation are achieved simultaneously by expression of a selectable marker, which under some circumstances, allows cells expressing the marker to survive while other cells die (or vice versa). Screening reporters include luciferase, β-glucuronidase, green fluorescent protein (GFP), and the maize *Lc* (anthocyanin regulatory) gene. Selectable markers include antibiotic and herbicide resistance genes. A special class of selectable markers are negatively selectable markers. Cells or organisms expressing a negatively selectable marker die under appropriate selection conditions while organisms lacking or having a non-functional form of the marker survive. Examples of negatively selectable markers useful in the context of plant genetic engineering include a number of genes involved in herbicide metabolism, including: *dlh1, codA, tms2* and *NIA2.*

The term "gene" is used broadly to refer to any segment of DNA associated with a biological function. Thus, genes include coding sequences and/or the regulatory sequences required for their expression. Genes also include nonexpressed polynucleotides that, for example, form recognition sequences for other proteins. Non-expressed regulatory sequences include "promoters" and "enhancers", to which regulatory proteins such as transcription factors bind, resulting in transcription of adjacent or nearby sequences.

A "wild-type" gene, or DNA or RNA sequence, is a gene, or sequence, which occurs in an organism in nature.

An "exogenous" gene or "transgene" is a gene foreign (or heterologous) to the cell, or homologous to the cell, but in a position within the host cell nucleic acid in which the element is not ordinarily found. Exogenous genes can be expressed to yield exogenous polypeptides. A "transgenic" organism is one which has a transgene introduced into its genome. Such an organism may be either an animal or a plant. The site at which a transgene is located in the genome is referred to as its "site of insertion" or

COR-AAD1-00002288

alternatively, its "insertion site." The integration of a transgene or other vector-borne sequence into a predetermined location within the host genome is referred to as "targeted insertion," while "insert precision" refers to the insertion of a single, defined copy of a T-DNA without rearrangement of the T-DNA ends.

A "vector" is a means by which an exogenous DNA is introduced into a "host cell. A vector can be a nucleic acid polynucleotide, most typically a plasmid or virus, but also including a naked RNA polynucleotide, a naked DNA polynucleotide, a poly-lysine –conjugated DNA, a peptide-conjugated DNA, a liposome-conjugated DNA, or the like, or it can be an organism which comprises one or more of the above polynucleotide constructs such as an agrobacterium or a bacterium. Once inside the host cell, replication and in some cases expression, are established under control of vector origin and regulatory sequences. "Transformation" refers to the process by which a vector is introduced into a host cell. Transformation (or transduction, or transfection), can be achieved by any one of a number of means including electroporation, microinjection, biolistics (or particle bombardment-mediated delivery), or agrobacterium mediated transformation. A "parental" cell, or organism, is an untransformed member of the host species giving rise to a transgenic cell, or organism.

A "host species" is the recipient of a transforming vector. A vector which is capable of transforming a wide variety of host species is said to have a "broad host range" while vectors which transform only one or a few host species is said to have a "narrow host range."

"Agrobacterium" are soil-borne plant-pathogenic bacterium, the two predominant species of which are *A. tumefaciens* and *A. rhizogenes*. Agrobacterium carry a tumor causing plasmid designated Ti, modified versions of which serve as common plant vectors. During infection, the "T region", a portion of the Ti plasmid from which "T-DNA" is derived is transferred, along with any intervening sequences, to the host plant cell. Imperfect 25 base pair repeat sequences designated the "right T-DNA border" and the "left T-DNA border" define the limits of the T-DNA, and are utilized in the transfer of the T-DNA from the agrobacterium to the plant cell.

Also located on the Ti plasmid are the "virulence" or "*vir*" genes of agrobacterium. The products of the *vir* genes contribute various functions required for transfer of the T-DNA, and include *virA, virB, virC, virD, virE* and the chromosomal gene, *ChvE*. The *vir* genes are regulated by a signal transduction system in which the product of the *virA* gene, the VirA protein, acts as sensor to detect secreted molecules

COR-AAD1-00002289

from the plant. Binding of a secreted signal molecule and activation of VirA results in the induction of other *vir* genes via transcriptional activation. The regulatory region responsible for activation of a *vir* gene is referred to as a "*vir* promoter."

The term "Agroinfection" refers to the introduction of plant infectious agents, for example, viruses, into plants by Agrobacterium. More generally "agrobacterium mediated transformation" refers to the transfer of any DNA sequences present between T-DNA ends into a plant cell by an agrobacterium.

"Binary vector system" generally refers to a two plasmid system in which Vir function is supplied on a "helper plasmid," frequently a modified Ti plasmid, to mediate transfer of a T-DNA located on a separate plasmid. Alternatively, it is possible to supply Vir function by sequences integrated into an Agrobacterium chromosome.

"PCR," or the "polymerase chain reaction," is an in-vitro method for the amplification of specific DNA sequences. PCR utilizes multiple cycles of 1) denaturation of the target DNA sequence, 2) annealing of DNA sequences serving to prime DNA polymerization to the target DNA, and 3) polymerase mediated extension of the annealed sequence to produce multiple copies of a target DNA sequence. Single stranded oligonucleotides, typically in the range of 15-30 bases, which hybridize to portions of the target DNA, are referred to as "PCR primers." "Error prone PCR," "mutagenic amplification" and "site directed mutagenesis" are methods for introducing mutations, or alterations, in the nucleotide composition of target DNA molecules. An "inverse polymerase chain reaction" is a method of amplifying a target of an unknown sequence, for example a transgene insertion site, in which PCR primers corresponding to sequences flanking the target are provided. "LCR" or "ligase chain reaction" refers to a variation of PCR, in which T4 ligase mediates joining of fragments serving as substrates for amplification.

The term "sequence similarity" means that two nucleic acid sequences are identical or share regions of sequence identity.

Viral vectors are introduced into cultured plant tissues or cells or directly onto intact plants at a "site of inoculation." "Systemic spread" refers to the replication and dissemination of the virus throughout the plant in a process which depends on the activity of viral "movement proteins." Viral "coat proteins" are used to package the replicated viral RNA or DNA genome of many plant viruses for transport within and between plant hosts.

COR-AAD1-00002290

In the context of a viral vector, as elsewhere in this application, 5' refers to the left-hand or "upstream" direction of a nucleic acid molecule and 3' refers to the right-hand or "downstream" direction of a nucleic acid molecule. Hence a "5' region" of a viral nucleic acid lies upstream of a "3' region," and the direction of transcription reads from 5' to 3'.

The use of nucleic acid sequences derived from a plant pathogen, most typically a virus, to confer resistance to plant pathogens in transgenic plants is referred to as "pathogen-derived resistance."

INTRODUCTION

The present invention provides a method for evolving various plant vectors, for example, agrobacterium and viruses, to acquire desired properties. Prior to the invention, construction of vectors was limited, on one hand, to the introduction of specific rational changes by site directed mutagenesis techniques, and on the other hand, to random, and often deleterious, mutagenesis techniques. The invention utilizes recursive recombination, e.g., DNA shuffling, technologies, in a process of directed evolution, to select and optimize mutations leading to desired characteristics in plant transformation vectors. These vectors are used to produce transgenic plants which are useful for a wide variety of purposes, including the production of commercially important products such as proteins, lipids, and biosynthetic polymers. Additionally, the invention relates to the production and use of evolved plant viruses which confer pathogen-derived resistance to plant pathogens.

General texts which describe molecular biological techniques useful herein, including the use of vectors, promoters and many other relevant topics related to, e.g., the cloning and expression of agrobacterium *vir* genes, viral coat proteins, replicases and movement proteins, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook et al., Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook") and Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999) ("Ausubel")). Similarly, examples of techniques sufficient to direct persons of skill through in vitro amplification methods, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Qβ-replicase amplification and

COR-AAD1-00002291

other RNA polymerase mediated techniques (e.g., NASBA), e.g., for the production of the homologous nucleic acids of the invention are found in Berger, Sambrook, and Ausubel, as well as Mullis et al., (1987) U.S. Patent No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (October 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3, 81-94; (Kwoh et al. (1989) Proc. Natl. Acad. Sci. USA 86, 1173; Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87, 1874; Lomell et al. (1989) J. Clin. Chem 35, 1826; Landegren et al., (1988) Science 241, 1077-1080; Van Brunt (1990) Biotechnology 8, 291-294; Wu and Wallace, (1989) Gene 4, 560; Barringer et al. (1990) Gene 89, 117, and Sooknanan and Malek (1995) Biotechnology 13: 563-564. Improved methods of cloning in vitro amplified nucleic acids are described in Wallace et al., U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684-685 and the references therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See*, Ausubel, Sambrook and Berger, *all supra.*

The present invention also relates to host cells and organisms which are transformed with vectors of the invention, and the production of polypeptides of the invention, e.g.viral coat proteins and other proteins, and polypeptides encoded by exogenous DNAs, by recombinant techniques. Host cells are genetically engineered (i.e., transformed, transduced or transfected) with the vectors of this invention, which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, an agrobacterium, a virus, a naked polynucleotide, or a conjugated polynucleotide. The vectors are introduced into plant tissues, cultured plant cells or plant protoplasts by standard methods including electroporation (From et al., *Proc. Natl. Acad. Sci. USA* 82, 5824 (1985), infection by viral vectors such as cauliflower mosaic virus (CaMV) (Hohn et al., *Molecular Biology of Plant Tumors*, (Academic Press, New York, 1982) pp. 549-560; Howell, US 4,407,956), high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein et al., *Nature* 327, 70-73 (1987)), use of pollen as vector (WO 85/01856), or use of *Agrobacterium tumefaciens* or *A. rhizogenes* carrying a T-DNA plasmid in which DNA fragments are cloned. The T-DNA plasmid is transmitted to plant cells upon infection by *Agrobacterium tumefaciens*, and a portion is stably integrated into

COR-AAD1-00002292

the plant genome (Horsch et al., *Science* 233, 496-498 (1984); Fraley et al., *Proc. Natl. Acad. Sci. USA* 80, 4803 (1983)).

The engineered host cells can be cultured in conventional nutrient media modified as appropriate for such activities as, for example, activating promoters or selecting transformants. These cells can optionally be cultured into transgenic plants. Plant regeneration from cultured protoplasts is described in Evans et al., "Protoplast Isolation and Culture," *Handbook of Plant Cell Cultures* 1, 124-176 (MacMillan Publishing Co., New York, 1983); Davey, "Recent Developments in the Culture and Regeneration of Plant Protoplasts," *Protoplasts*, (1983) pp. 12-29, (Birkhauser, Basal 1983); Dale, "Protoplast Culture and Plant Regeneration of Cereals and Other Recalcitrant Crops," *Protoplasts* (1983) pp. 31-41, (Birkhauser, Basel 1983); Binding, "Regeneration of Plants," *Plant Protoplasts*, pp. 21-73, (CRC Press, Boca Raton, 1985).

The present invention also relates to the production of transgenic organisms, which may be bacteria, yeast, fungi, or plants. A thorough discussion of techniques relevant to bacteria, unicellular eukaryotes and cell culture may be found in references enumerated above and are briefly outlined as follows. Several well-known methods of introducing target nucleic acids into bacterial cells are available, any of which may be used in the present invention. These include: fusion of the recipient cells with bacterial protoplasts containing the DNA, electroporation, projectile bombardment, and infection with viral vectors (discussed further, below), etc. Bacterial cells can be used to amplify the number of plasmids containing DNA constructs of this invention. The bacteria are grown to log phase and the plasmids within the bacteria can be isolated by a variety of methods known in the art (*see*, for instance, Sambrook). In addition, a plethora of kits are commercially available for the purification of plasmids from bacteria. For their proper use, follow the manufacturer's instructions (see, for example, EasyPrep™, FlexiPrep™, both from Pharmacia Biotech; StrataClean™, from Stratagene; and, QIAprep™ from Qiagen). The isolated and purified plasmids are then further manipulated to produce other plasmids, used to transfect plant cells or incorporated into *Agrobacterium tumefaciens* related vectors to infect plants. Typical vectors contain transcription and translation terminators, transcription and translation initiation sequences, and promoters useful for regulation of the expression of the particular target nucleic acid. The vectors optionally comprise generic expression cassettes containing at least one independent terminator sequence, sequences permitting replication of the cassette in eukaryotes, or prokaryotes, or both, (e.g., shuttle vectors) and selection

COR-AAD1-00002293

markers for both prokaryotic and eukaryotic systems. Vectors are suitable for replication and integration in prokaryotes, eukaryotes, or preferably both. *See,* Giliman & Smith, *Gene* 8:81 (1979); Roberts, *et al.*, *Nature*, 328:731 (1987); Schneider, B., *et al.*, *Protein Expr. Purif.* 6435:10 (1995); Ausubel, Sambrook, Berger (*all supra*). A catalogue of Bacteria and Bacteriophages useful for cloning is provided, e.g., by the ATCC, e.g., *The ATCC Catalogue of Bacteria and Bacteriophage* (1992) Gherna *et al.* (eds) published by the ATCC. Additional basic procedures for sequencing, cloning and other aspects of molecular biology and underlying theoretical considerations are also found in Watson *et al.* (1992) *Recombinant DNA Second Edition Scientific American Books, NY.*

TRANSFORMING NUCLEIC ACIDS INTO PLANTS.

Preferred embodiments pertain to the production of transgenic plants using the evolved vectors of the invention. Techniques for transforming plant cells with nucleic acids are generally available and can be adapted to the invention by the use of evolved plasmids, viruses, and components thereof, and by the use of agrobacterium strains comprising evolved vectors. In addition to Berger, Ausubel and Sambrook, useful general references for plant cell cloning, culture and regeneration include Jones (ed) (1995) Plant Gene Transfer and Expression Protocols-- Methods in Molecular Biology, Volume 49 Humana Press Towata, NJ ("Jones"); Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, NY ("Payne"); and Gamborg and Phillips (eds) (1995) Plant Cell, Tissue and Organ Culture; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) ("Gamborg"). A variety of cell culture media are described in Atlas and Parks (eds) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL ("Atlas"). Additional information for plant cell culture is found in available commercial literature such as the Life Science Research Cell Culture Catalogue (1998) from Sigma- Aldrich, Inc (St Louis, MO) (Sigma-LSRCCC) and, e.g., the Plant Culture Catalogue and supplement (1997) also from Sigma-Aldrich, Inc (St Louis, MO) (Sigma-PCCS). Additional details regarding plant cell culture are found in Croy, (ed.) (1993) Plant Molecular Biology Bios Scientific Publishers, Oxford, U.K.

The nucleic acid constructs of the invention, e.g., plasmids, viruses, DNA and RNA polynucleotides, are introduced into plant cells, either in culture or in the organs of a plant by a variety of conventional techniques. To use the evolved, e.g., shuffled, sequences of the invention, recombinant DNA or RNA vectors suitable for transformation

COR-AAD1-00002294

of plant cells are isolated and/or prepared. To introduce an exogenous DNA, which can itself be a recombinant, e.g., shuffled, DNA, the exogenous DNA sequence can be incorporated into an evolved vector of the invention and transformed into the plant as indicated above. Where the sequence is expressed, the sequence is optionally combined with transcriptional and translational initiation regulatory sequences which direct the transcription or translation of the sequence from the exogenous DNA in the intended tissues of the transformed plant.

The DNA constructs of the invention, for example plasmids, or naked or variously conjugated-DNA polynucleotides, (e.g., polylysine-conjugated DNA, peptide-conjugated DNA, liposome-conjugated DNA, etc.) can be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant cells using ballistic methods, such as DNA particle bombardment.

Microinjection techniques for injecting e.g., cells, embryos, and protoplasts, are known in the art and well described in the scientific and patent literature. For example, a number of methods are described in Jones (ed) (1995) Plant Gene Transfer and Expression Protocols-- Methods in Molecular Biology, Volume 49 Humana Press Towata NJ, as well as in the other references noted herein and available in the literature.

For example, the introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski, et al., EMBO J. 3:2717 (1984). Electroporation techniques are described in Fromm, et al., Proc. Nat'l. Acad. Sci. USA 82:5824 (1985). Ballistic transformation techniques are described in Klein, et al., Nature 327:70-73 (1987). Additional details are found in Jones (1995) *supra*.

AGROBACTERIUM MEDIATED TRANSFORMATION

In preferred embodiments, DNA constructs comprising suitable T-DNA border sequences are introduced into an Agrobacterium host vector. The virulence functions of the Agrobacterium host directs the insertion of the construct into the plant genomic DNA. For example, any one or more of the following are optionally combined with T-DNA flanking sequences: a coding sequence, a non-coding sequence, a structural gene, a disabled gene, a promoter, an enhancer and a marker and inserted into the plant cell DNA when the cell is infected by the bacteria. Agrobacterium-mediated transformation techniques, including disarming and use of binary vectors, are also well

COR-AAD1-00002295

described in the scientific literature. *See*, for example Horsch, et al., "A simple and general method for transferring genes into plants." Science 233:496-498 (1984), and Fraley, et al., "Expression of bacterial genes in plant cells." Proc. Nat'l. Acad. Sci. USA 80:4803 (1984) and recently reviewed in Hansen and Chilton, "Lessons in gene transfer to plants by a gifted microbe." Current Topics in Microbiology 240:22-51 (1998) and Das (1998) "DNA transfer from *Agrobacterium* to plant cells in crown gall tumor disease." Subcellular Biochemistry 29: Plant Microbe Interactions:343-363. These techniques are adapted in the present invention with the use of evolved vectors and vector components, and the use of such modified vectors in the production of transgenic plants.

Agrobacteria are gram-negative, soil-borne plant pathogens that cause neoplastic growth in susceptible plants. The most prevalent pathogenic strains, *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* cause crown gall and hairy root disease, respectively. In recent years, the art of plant transgenesis has taken advantage of genetically modified agrobacterium strains to transfer exogenous DNA to host plants by means of agrobacterium mediated transformation, a process which utilizes the naturally occurring DNA transfer system of these pathogenic bacteria. The present invention specifically provides for the development of Agrobacterium plasmids and Agrobacterium strains with improved properties, for use in the production of transgenic plants.

Genes located on the Ti, or "tumor inducing," plasmid of *A. tumefaciens*, are involved in DNA transfer and tumor induction. The transferred DNA or T-DNA is derived from the T-region of the Ti plasmid. This region is delimited by the presence of two imperfect 25 base pair repeats designated T-DNA borders. DNA sequences, including, for example, the exogenous sequences selected for targeting to a plant genome of the invention, can be cloned between the T-DNA borders and subsequently transferred to a host plant cell. Alternatively, the exogenous sequence can be cloned adjacent to a single right T-DNA border and transferred to the host plant cell.

The virulence, or *vir*, gene products activate and facilitate transfer of T-DNA to a host plant cell. VirA is a transmembrane protein which acts as a sensor, detecting the presence of molecules, such as acetosyringone (3',5'-dimethoxy-4'-hydroxyacetophenone), which are secreted by wounded or metabolically active target cells. When activated by such molecules, VirA exhibits kinase activity that activates the VirG protein. In turn, VirG induces transcription of the *virB, virC, VirD* and *virE* operons. The VirB proteins are thought to form a conjugation-like pore in the bacterial

COR-AAD1-00002296

surface through which the T-DNA passes upon transfer. VirC, VirD and VirE proteins are involved in the production and transfer of the T-DNA strand. In the present invention, these virulence genes, either singly or in combination are shuffled to produce virulence proteins with novel functional activities. These evolved virulence proteins are encoded by components of plasmids and the agrobacterium strains which harbor them and confer various desirable phenotypic modifications, such as increased host range, and improved T-DNA transfer efficiency under standard or simplified transformation protocols.

As the T-DNA is released from the plasmid, the VirD2 protein binds the right T-DNA border. For effective transformation, VirE2 function is transported to the host plant cell by the Agrobacterium. While it is clear that VirE2 is involved in integration of the T-DNA into the host cell chromosome, it does not appear to be essential for transfer of the T-DNA from the Agrobacterium to the host cell, and may be supplied by a second Agrobacterium or by the host cell itself (Ream "Import of *Agrobacterium tumefaciens* virulence proteins and transferred DNA into plant cell nuclei." Subcellular Biochemistry 29: Plant-Microbe Interactions:365-384 (1998), and references therein). Multiple molecules of VirE2, a single stranded DNA binding protein, then bind to the T-DNA strand. Both VirD2 and VirE2 possess nuclear localization signals (NLS) thought to facilitate targeting of the T-DNA to the host plant cell nucleus, where the T-DNA is integrated into the plant cell nucleus. A preferred embodiment provides for the modification or replacement of VirD2 and VirE2 NLS with sequences selected from variant, e.g., shuffled, chloroplast DNA sequences which target the T-DNA to the chloroplast.

BINARY VECTOR SYSTEMS

Because the *vir* gene products are soluble factors, their function is optionally supplied in trans, giving rise to the use of binary vector systems in the production of transgenic plants. In a binary vector system, *vir* genes are supplied on a helper plasmid, often a disarmed Ti plasmid, or, alternatively, integrated into a plant chromosome. A disarmed Ti plasmid which lacks the genes that mediate oncogenesis is much reduced in size from the native 200 kilobase pairs, thereby facilitating its manipulation in vitro. An exogenous DNA sequence, for example, a plant or bacterial structural gene, or a plant or viral promoter or enhancer, is cloned between T-DNA border repeats on a second plasmid, which typically also has a marker, e.g. an antibiotic

COR-AAD1-00002297

resistance gene, to facilitate selection after introduction into Agrobacterium. Subsequently, the exogenous DNA sequence is transferred as part of the T-DNA strand to a host plant (or plant cell or explant) where it integrates into a random site in the host plant chromosome. Preferred embodiment of the invention include evolved binary vectors and the components thereof. Individual components or combinations thereof are recombined, e.g., shuffled, in silico, in vitro or in vivo to produce novel polynucleotides with new characteristics. For example, various *vir* genes of a helper plasmid can be recombined (or recursively recombined) separately, together in various combinations, together as an intact helper plasmid in vitro, or together as an intact helper plasmid in the context of whole genome shuffling of bacterial genomes, and the resultant recombinant vector components and vectors screened for desirable properties.

The use of agrobacterium mediated transfer has proven a valuable technique in the production of genetically modified plant species. In addition to their utility in the transformation of plant species, Agrobacterium are readily manipulated in vitro by well established techniques of molecular biology. Such techniques are well known to those skilled in the art, and are referenced in e.g., Ausubel, Sambrook, and Berger, *supra*; Croy (ed) (1993) Plant Molecular Biology, Bios Scientific Publishers, Oxford, U.K., and Jones (ed) (1995) Plant Gene Transfer and Expression Protocols--Methods in Molecular Biology, Volume 49 Humana Press, Towata NJ. These are useful in the context of the present invention for the manipulation and culture of Agrobacterium cells, transformation techniques, and techniques useful for the analysis of plant cells subject to agrobacterium mediated transformation.

While dicotyledonous plants have proven most amenable to manipulation by agrobacterium mediated transformation, reports of transformation of important monocotyledonous crop plants have been forthcoming. In addition, Agrobacterium strains which are capable of transforming fungal species have also been described. Preferred embodiments of the invention include agrobacterium strains which are capable of efficient transformation of a wide variety of fungal and plant species, (so-called broad host range) including both dicots and monocots. For example, by selecting recombinant *vir* genes which respond to signals from a broad range of host species, or which more effectively transduce the host plant derived signal to the activation of *vir* genes involved in transfer of the T-DNA, Agrobacterium strains which are capable of transforming an extended range of host plants can be developed.

COR-AAD1-00002298

PLANT VIRUSES

Embodiments of the present invention also comprise vectors which are plant viruses. Plant viruses evolved to have new and desirable transformation and expression properties are preferred embodiments. Viruses are typically useful as vectors for expressing exogenous DNA sequences in a transient manner in plant hosts. In contrast to agrobacterium mediated transformation which results in the stable integration of DNA sequences in the plant genome, viral vectors are generally replicated and expressed without the need for chromosomal integration. Nonethless, in certain embodiments, stable transformation of one or more viral components can prove beneficial, e.g., to transactivate expression of an exogenous or endogenous sequence comprising appropriate cis-acting sequences.

Plant virus vectors offer a number of advantages, specifically: DNA copies of viral genomes can be readily manipulated in *E.coli*, and transcribed in vitro, where necessary, to produce infectious RNA copies; naked DNA, RNA, or virus particles can be easily introduced into mechanically wounded leaves of intact plants; high copy numbers of viral genomes per cell results in high expression levels of introduced genes; common laboratory plant species as well as monocot and dicot crop species are readily infected by various virus strains; infection of whole plants permits repeated tissue sampling of single library clones; recovery and purification of recombinant virus particles is simple and rapid; and because replication occurs without chromosomal insertion, expression is not subject to position effects. These many advantages are exploited by the present invention for the production of improved viral vectors.

Over six-hundred-fifty plant viruses have been identified, and are amenable directly and/or indirectly as substrates for the directed evolution processes of the invention. Plant viruses cause a range of diseases, most commonly mottled damage to leaves, so-called mosaics. Other symptoms include necrosis, deformation, outgrowths, and generalized yellowing or reddening of leaves. Plant viruses are known which infect every major food-crop, as well as most species of horticultural interest. The host range varies between viruses, with some viruses infecting a broad host range (e.g., alfalfa mosaic virus infects more than 400 species in 50 plant families) while others have a narrow host range, sometimes limited to a single species (e.g., barley yellow mosaic virus). Host range is among the many traits for which it is possible to select appropriate vectors according to the methods provided by the present invention.

COR-AAD1-00002299

Approximately 75% of the known plant viruses have genomes which are single-stranded (ss) messenger sense (+) RNA polynucleotides. Major taxonomic classifications of ss-RNA(+) plant viruses include the bromovirus, capillovirus, carlavirus, carmovirus, closterovirus, comovirus, cucumovirus, fabavirus, furovirus, hordeivirus, ilarvirus, luteovirus, potexvirus, potyvirus, tobamovirus, tobravirus, tombusvirus, and many others. Other plant viruses exist which have single-stranded antisense (-) RNA (e.g., rhabdoviridae), double-stranded (ds) RNA (e.g., cryptovirus, reoviridae), or ss or ds DNA genomes (e.g., geminivirus and caulimovirus, respectively).

Preferred embodiments of the invention include evolved vectors which are either RNA and DNA viruses. Examples of such embodiments include viruses selected from among: an alfamovirus, a bromovirus, a capillovirus, a carlavirus, a carmovirus, a caulimovirus, a closterovirus, a comovirus, a cryptovirus, a cucumovirus, a dianthovirus, a fabavirus, a fijivirus, a furovirus, a geminivirus, a hordeivirus, a ilarvirus, a luteovirus, a machlovirus, a maize chlorotic dwarf virus, a marafivirus, a necrovirus, a nepovirus, a parsnip yellow fleck virus, a pea enation mosaic virus, a potexvirus, a potyvirus, a reovirus, a rhabdovirus, a sobemovirus, a tenuivirus, a tobamovirus, a tobravirus, a tomato spotted wilt virus, a tombusvirus, and a tymovirus.

Plant viruses can be engineered as vectors to accomplish a variety of functions. Examples of both DNA and RNA viruses have been used as vectors for gene replacement, gene insertion, epitope presentation and complementation, (see, e.g., Scholthof, Scholthof and Jackson, (1996) "Plant virus gene vectors for transient expression of foreign proteins in plants," Annu.Rev.of Phytopathol. 34:299-323.) To evolve plant vectors which are improved for one or more properties related to the aforementioned uses, as well as for traits which are useful in the high yield production of proteins, or other products of interest, in planta, or for pathogen-derived resistance, one or more virus encoded proteins, or an entire viral genome is selected as the substrate for the diversification procedures, e.g., DNA or RNA shuffling, of the present invention.

Typically, plant viruses encode multiple proteins required for initial infection, replication and systemic spread, e.g. coat proteins, helper factors, replicases, and movement proteins. Any of these, as well as other coding and non-coding sequences, such as promoters and enhancers, can readily serve as substrates in the methods of the invention. The nucleotide sequences encoding many of these proteins are matters of public knowledge, and accessible through any of a number of databases, e.g. (Genbank: www.ncbi.nlm.nih.gov/genbank/ or EMBL: www.ebi.ac.uk.embl/).

COR-AAD1-00002300

For example, due to the development of symptoms after viral infection by the host plant, it is sometimes difficult to recover expressed products in quantity. Reduction in the host response or alterations in the hypersensitivity reaction are of significant value in effort to increase yield from transformed plants. Sequences encoding viral proteins, in particular, coat proteins can be recombined, e.g., shuffled, separately, in combination or in the context of whole genome shuffling of plant viruses to produce the vectors of the invention which have diminished viral induced symptoms following infection.

Methods for the transformation of plants and plant cells using sequences derived from plant viruses include the direct transformation techniques described above relating to DNA molecules, *see* e.g., Jones, ed. (1995) Plant Gene Transfer and Expression Protocols, Humana Press, Totowa, NJ, for a recent compilation. In addition viral sequences can be cloned adjacent T-DNA border sequences and introduced via Agrobacterium mediated transformation, or Agroinfection (*see*, e.g., Grimsley et al. (1986) Proc Natl Acad Sci USA 83:3282).

Viral particles comprising the plant virus vectors of the invention can also be introduced by mechanical inoculation using techniques well known in the art, (*see* e.g., Cunningham and Porter, eds. (1997) Methods in Biotechnology, Vol.3. Recombinant Proteins from Plants: Production and Isolation of Clinically Useful Compounds, for detailed protocols). Briefly, for experimental purposes, young plant leaves are dusted with silicon carbide (carborundum), then inoculated with a solution of viral transcript, or encapsidated virus and gently rubbed. Large scale adaptations for infecting crop plants are also well known in the art, and typically involve mechanical maceration of leaves using a mower or other mechanical implement, followed by localized spraying of viral suspensions, or spraying leaves with a buffered virus/carborundum suspension at high pressure. Any of these above mentioned techniques can be adapted to the evolved vectors of the invention, and are useful for alternative applications depending on the choice of plant virus, and host species, as well as the scale of the specific transformation application.

COR-AAD1-00002301

REGENERATION OF TRANSGENIC PLANTS

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans, et al., Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, pp. 124-176, Macmillian Publishing Company, New York, (1983); and Binding, Regeneration of Plants, Plant Protoplasts, pp. 21-73, CRC Press, Boca Raton, (1985). Regeneration can also be obtained from plant callus, explants, somatic embryos (Dandekar, et al., J. Tissue Cult. Meth. 12:145 (1989); McGranahan, et al., Plant Cell Rep. 8:512 (1990)), organs, or parts thereof. Such regeneration techniques are described generally in Klee, et al., Ann. Rev. of Plant Phys. 38:467-486 (1987). Additional details are found in Payne (1992) and Jones (1995), both *supra*. These methods are adapted to the invention to produce transgenic plants using evolved vectors, including agrobacteria and viruses.

Preferred plants for the transformation and expression of the novel recombinases of this invention include agronomically and horticulturally important species. Such species include, but are not restricted to members of the families: Graminae (including corn, rye, triticale, barley, millet, rice, wheat, oats, etc.); Leguminosae (including pea, beans, lentil, peanut, yam bean, cowpeas, velvet beans, soybean, clover, alfalfa, lupine, vetch, lotus, sweet clover, wisteria, and sweetpea); Compositae (the largest family of vascular plants, including at least 1,000 genera, including important commercial crops such as sunflower) and Rosaciae (including raspberry, apricot, almond, peach, rose, etc.), as well as nut plants (including, walnut, pecan, hazelnut, etc.), and forest trees (including *Pinus, Quercus, Pseutotsuga, Sequoia, Populus,*etc.)

Additionally, preferred targets for modification the evolved vectors of the invention, as well as those specified above, plants from the genera: *Agrostis, Allium, Antirrhinum, Apium, Arachis, Asparagus, Atropa, Avena* (e.g., oats), *Bambusa, Brassica, Bromus, Browaalia, Camellia, Cannabis, Capsicum, Cicer, Chenopodium, Chichorium, Citrus, Coffea, Coix, Cucumis, Curcubita, Cynodon, Dactylis, Datura, Daucus, Digitalis, Dioscorea, Elaeis, Eleusine, Festuca, Fragaria, Geranium, Glycine, Helianthus, Heterocallis, Hevea, Hordeum* (e.g., barley), *Hyoscyamus, Ipomoea, Lactuca, Lens,*

COR-AAD1-00002302

Lilium, Linum, Lolium, Lotus, Lycopersicon, Majorana, Malus, Mangifera, Manihot, Medicago, Nemesia, Nicotiana, Onobrychis, Oryza (e.g., rice), *Panicum, Pelargonium, Pennisetum* (e.g., millet), *Petunia, Pisum, Phaseolus, Phleum, Poa, Prunus, Ranunculus, Raphanus, Ribes, Ricinus, Rubus, Saccharum, Salpiglossis, Secale* (e.g., rye), *Senecio, Setaria, Sinapis, Solanum, Sorghum, Stenotaphrum, Theobroma, Trifolium, Trigonella, Triticum* (e.g., wheat), *Vicia, Vigna, Vitis, Zea* (e.g., corn), and the *Olyreae*, the *Pharoideae* and many others. As noted, plants in the family *Graminae* are a particularly preferred target plants for the methods of the invention.

Common crop plants which are targets of the present invention include corn, rice, triticale, rye, cotton, soybean, sorghum, wheat, oats, barley, millet, sunflower, canola, peas, beans, lentils, peanuts, yam beans, cowpeas, velvet beans, clover, alfalfa, lupine, vetch, lotus, sweet clover, wisteria, sweetpea and nut plants (e.g., walnut, pecan, etc).

In construction of recombinant expression cassettes of the invention, which include, for example, helper plasmids comprising virulence functions, and plasmids or viruses comprising exogenous DNA sequences such as structural genes, a plant promoter fragment is optionally employed which directs expression of a nucleic acid in any or all tissues of a regenerated plant. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill. Alternatively, the plant promoter may direct expression of the polynucleotide of the invention in a specific tissue (tissue-specific promoters) or may be otherwise under more precise environmental control (inducible promoters). Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only in certain tissues, such as fruit, seeds, or flowers.

Any of a number of promoters which direct transcription in plant cells can be suitable. The promoter can be either constitutive or inducible. In addition to the promoters noted above, promoters of bacterial origin which operate in plants include the octopine synthase promoter, the nopaline synthase promoter and other promoters derived from native Ti plasmids. *See,* Herrara-Estrella et al. (1983), Nature, 303:209-213. Viral promoters include the 35S and 19S RNA promoters of cauliflower mosaic virus. *See,* Odell et al. (1985) Nature, 313:810-812. Other plant promoters include the ribulose-1,3-bisphosphate carboxylase small subunit promoter and the phaseolin

COR-AAD1-00002303

promoter. The promoter sequence from the E8 gene and other genes may also be used. The isolation and sequence of the E8 promoter is described in detail in Deikman and Fischer, (1988) EMBO J. 7:3315- 3327. Many other promoters are in current use and can be coupled to an exogenous DNA sequence to direct expression of the nucleic acid.

If expression of a polypeptide, including various viral, bacterial and exogenous gene products, such as viral coat proteins, biosynthetic enzymes, and markers of the present invention, is desired, a polyadenylation region at the 3'-end of the coding region is typically included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from, e.g., T-DNA.

The vector comprising the sequences (e.g., promoters or coding regions) from genes encoding expression products and transgenes of the invention will typically include a nucleic acid subsequence, a marker gene which confers a selectable, or alternatively, a screenable, phenotype on plant cells. For example, the marker may encode biocide tolerance, particularly antibiotic tolerance, such as tolerance to kanamycin, G418, bleomycin, hygromycin, or herbicide tolerance, such as tolerance to chlorosluforon, or phosphinothricin (the active ingredient in the herbicides bialaphos or Basta). See, e.g., Padgette et al. (1996) "New weed control opportunities: Development of soybeans with a Round UP Ready™ gene" In: Herbicide-Resistant Crops (Duke, ed.), pp 53-84, CRC Lewis Publishers, Boca Raton ("Padgette"). For example, crop selectivity to specific herbicides can be conferred by engineering genes into crops which encode appropriate herbicide metabolizing enzymes from other organisms, such as microbes. See, Vasil (1996) "Phosphinothricin-resistant crops" In: Herbicide-Resistant Crops (Duke, ed.), pp 85-91, CRC Lewis Publishers, Boca Raton) ("Vasil").

The invention described herein furthers the current technology by providing for improved plant transformation vectors for the introduction of exogenous DNA sequences such as the expression cassettes described above. One of skill will recognize that after the exogenous DNA sequence is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

DEVELOPMENT OF IMPROVED VECTORS

The methodology of the present invention is directed towards the development of vectors for plant transformation which have improved properties. A

COR-AAD1-00002304

significant advantage of the methods of the invention is that they can be applied to the evolution of vectors which meet a wide variety of general, as well as specialized needs. The directed evolution processes described herein are limited only by the availability of screening and/or selection protocols available or developed to analyze the trait under consideration. No prior knowledge of the genes, or gene products contributing to the phenotype is essential for the successful evolution of novel plant vectors with improved characteristics, such as expanded host range, or increased transformation efficiency.

Agrobacterium vectors have traditionally proven most useful in the transformation of dicotyledonous plant species. A limited number of reports have recently emerged demonstrating agrobacterium strains which can transform monocots and other recalcitrant species. It is of significant interest to develop agrobacterium strains which are capable of transforming a broad range of monocot species, in particular food crop species, for example barley, corn, oats, rye, and wheat. Embodiments of the invention relate to evolved agrobacterium vectors with a broad host range. Another limitation of the presently available agrobacterium vectors is that they insert T-DNAs randomly in the host genome. This is further complicated by the fact that duplications, deletions and insertions frequently occur in the DNA flanking the insertion site. This results in significant variability in expression between independent transformation events, even of the same exogenous DNA sequences. This insertion site variability is not always correctable by regulatory sequences incorporated into the exogenous DNA, leading to significant unpredictability in the expression of exogenous DNAs in transgenic plants. It would be of enormous technical and commercial value to be able to select the site of transgene insertion, as well as to limit disruptions in flanking DNA during the T-DNA insertion process. Certain embodiments of the invention relate to the production of agrobacterium vectors which result in targeted insertion events and to those which have evolved the property of insert precision, that is, they insert exogenous DNA sequences into the host genome with high efficiency and result in a simple insert pattern.

Other embodiments of the invention address the need to target exogenous sequences to the chloroplast rather than the nucleus of host plant. Targeting of chloroplast derived sequences to the nucleus frequently results in unstable integration events. Such unstable integration events present numerous difficulties in the cultivation and propagation of transgenic plants, especially for commercial purposes. This problem is particularly amenable to the methods of the present invention because chloroplast targeting sequences are poorly defined and lack consistent homology consensus

COR-AAD1-00002305

sequences. By evolving, e.g., by shuffling, various chloroplast DNA (cpDNA) derived sequences in concert with e.g., nuclear targeting sequences of *vir* genes, vectors can be evolved which target the chloroplast with increased efficiency and stability.

In addition to the traits discussed above, the methods of the invention are suitable for the development of agrobacterium strains with increased transformation efficiency as well as of strains which are amenable to simplified transformation techniques. Typically, dicotyledonous plants of such species as the experimentally popular *Arabidopsis thaliana* can be readily transformed by numerous agrobacterium strains through simple methods, such as vacuum infiltration of intact plants, well known in the art. However, valuable crop species belonging to monocot families, e.g., graminae, require more labor intensive and less efficient protocols, requiring plant cell culture under antibiotic conditions for successful transformation. Significant material and labor costs could be saved through the simplification of transformation protocols used for the production of transgenic crop species. The present invention addresses this need through the evolution of agrobacterium strains suitable for such protocols.

Aspects of the present invention also relate to the development of improved virus vectors for the expression and recovery of non-native proteins and other products in plants. The invention provides for evolved plant viruses which exhibit, e.g., faster systemic movement, greater expression levels, increased insert size, fewer disease symptoms, reduced toxicity, higher replication levels, extended host range, and pathogen-derived resistance.

For example, high throughput screening and virus re-isolation would be faster if the time from initial infection to systemic spread was reduced. The rate of systemic spread depends largely on movement proteins and their interaction with the viral nucleic acid and the plant plasmodesmata. For example, shuffling of single virus movement proteins or family shuffling of related viral movement proteins can be used to improve the speed of systemic spread. Screening for recombinant viruses or virus components which result in rapid systemic spread can be performed by inoculating recombined, e.g., shuffled, viruses or viruses comprising recombinant, e.g., shuffled, libraries onto leaves of intact plants. At set time intervals, distal leaves can be harvested and the viruses recovered. Additional cycles of shuffling with the fastest viruses leads to a significant improvement in the rate of systemic spread. Similarly, host range can be altered by changes in viral protein-plasmodesmata interactions.

COR-AAD1-00002306

Similarly, high throughput screening is facilitated by consistent high-level expression of the viral vector, and any incorporated transgene. To insure high expression of sequences derived from the introduced vector, it is desirable to improve the efficiency with which the replicase protein acts on the cis-acting origin of replication (e.g., the common region in geminiviruses). The AC2 trans-activating factor can be diversified by any of the diversification, e.g., mutation, recombination, recursive recombination, or shuffling procedures described herein, and variants which result in an increased efficiency of replication of viral derived sequences.

Preferred embodiments of the invention also include plant virus vectors which confer increased expression of viral, or alternatively, exogenous proteins. In addition to strategies for recombining (e.g., shuffling) whole genomes, targeted recombination (e.g., by shuffling) of at least two different types of sequences is appropriate. Firstly, by increasing copy number of viral genomes, and/or transcripts, expression levels can be increased. For this approach, populations of DNA fragments corresponding to viral replicases are provided as substrates for recombination reactions. Alternatively, viral and/or plant promoters can be recombined, e.g., shuffled, together to produce promoter sequences which result in expression levels which are higher, or otherwise more favorable, (e.g., tissue specific, or inducible) than wild-type viral or plant promoters. Variants, e.g., shuffled variants, of viruses or viral components can be screened for quantitative differences in protein expression using reporter genes such as green fluorescent protein (GFP), β-glucuronidase (GUS), luciferase, or the maize anthocyanin regulatory gene, *Lc*.

Among the hundreds of known plant viruses, a large proportion are unsuitable as vectors due to the small insert size which can be stably incoporated and transmitted by the vector. For example, the geminiviruses are constrained to a genome size of approximately 2.7 Kb. Exogenous sequences, e.g., transgenes, that result in a larger genome are frequently truncated during systemic spread of the virus. This limitation is due to a lack of movement of the DNA from cell to cell and not from a failure in packaging or replication of the larger viral genome. Recently, it has been determined that both the nucleus to cyptoplasm ssDNA shuttle protein (BR1) and plasmodesmata movement proteins (BL1) have size but not sequence selectivity. By screening or selecting for variants in these proteins with relaxed size selectivity, a vector that faithfully replicates and moves systemically with large inserts can be produced.An

COR-AAD1-00002307

especially preferred embodiment provides for viral vectors which express exogenous DNA sequences, only in a specific, transgenic host plant. Prior to the present invention, plant virus vectors useful for the expression of exogenous proteins in plants have been fully infectious to all plant species susceptible to the virus of the vector. Thus, the recombinant viruses are capable of escaping into the environment. The present invention provides for a viral vector that replicates and is encapsidated only in a corresponding transgenic host. A viral vector is constructed that lacks a coat protein necessary for its encapsidation, and thereby, its infectivity. The subgenomic promoter corresponding to the coat protein drives expression of an exogenous DNA sequence cloned, via a polylinker, adjacent to the promoter. The virus is then used to infect a transgenic host plant which expresses the coat protein lacking in the virus. In the transgenic host, virus particles containing the viral vector are packaged, and are capable of both short and long distance movement. However, the virus vector is not competent to form viral particles in a wild-type host plant and is not capable of escape into the environment as are autonomous viruses. One useful application is in the extension of host range of a plant virus vector. A given plant virus of the invention can be evolved by various recombination (e.g., recursive recombination or shuffling) procedures described herein, and recombinant viruses selected which have an expanded host range, without the risk of evolving new plant pathogens for those hosts. Such tailor-made viruses also have valuable applications in the screening of gene libraries inserted into such vectors in plants. Preferred embodiments relate to gene libraries which are recursively recombined, e.g., shuffled, gene libraries. The evolved vectors of this invention further provide environmentally safe viral vector systems for expression of proteins in plants, such as crop plants, in a field. The risk of escape into the environment is significantly reduced when compared to pre-existing viral expression systems. This reduces the risk associated with viral vectors that infect major crops. Currently the only commercial viral vectors for protein production are used on tobacco. The vector-transgenic plant system of the invention allows evolution of viral vectors that can infect model system plants such as *Arabidopsis thaliana* as well as the evolution of viral vectors for major crops such as corn and soybean without the risk of creating new plant pathogens for those crops.

Alternatively, viral vectors that are capable of transactivating expression of a gene of interest can be produced. For example, the Geminivirus AC1 or "replicase" protein is capable of transactivating replication of a latent mini-replicon integrated into the genome of a plant. To limit spread of an infectious virus expressing the protein, two

COR-AAD1-00002308

transgenic plant lines can be produced. The first transgenic plant expresses the replicase while the second transgenic plant possesses the integrated mini-replicon. Progeny inheriting both the replicase and the mini-replicon can express the transgene of interest at very high levels. However, the AC1 replicase is highly toxic to cell growth and differentiation as it signals the cell to enter a continuous s-phase of the cell cycle. The AC1 gene is diversified by any one or more of the recombination and/or mutagenesis methods described herein, thereby generating diverse libraries of recombinant AC1 replicase variants which are selected or screened for variants that are permissive to plant transformation while still retaining partial or full replication competence.

Another embodiment relates to the production of plant viruses that confer pathogen-derived resistance to plant viruses and other plant pathogens. Pathogen-derived resistance to plant viruses can be achieved by overexpression of coat proteins, (and occasionally other viral proteins) in transgenic plants. However, resistance is not always complete and protection is often limited to a single virus. Recombination, e.g., shuffling of viral coat proteins from a single or multiple viral strains can be used to select viral coat proteins which have, for example, increased affinity for viral RNA or DNA, or have altered assembly or encapsidation properties. Screening for recombinant coat protein genes which confer resistance to a broad range of viral pathogens can be conducted by expressing the shuffled viruses or shuffled coat proteins of the invention in transgenic plants which are then subjected to infection by wild type viruses.

DIVERSITY GENERATION

The invention provides for the evolution of plant vectors of various types that have acquired new and advantageous properties by a variety of diversification and screening or selection procedures. A variety of diversity generating protocols are available and described in the art. The procedures can be used separately, and/or in combination to produce one or more variants of a nucleic acid or set of nucleic acids, as well variants of encoded proteins. Individually and collectively, these procedures provide robust, widely applicable ways of generating diversified nucleic acids and sets of nucleic acids (including, e.g., nucleic acid libraries) useful, e.g., for the engineering or rapid evolution of nucleic acids, proteins, pathways, cells and/or organisms with new and/or improved characteristics.

While distinctions and classifications are made in the course of the ensuing discussion for clarity, it will be appreciated that the techniques are often not mutually

COR-AAD1-00002309

exclusive. Indeed, the various methods can be used singly or in combination, in parallel or in series, to access diverse sequence variants.

The result of any of the diversity generating procedures described herein can be the generation of one or more nucleic acids, which can be selected or screened for nucleic acids that encode proteins with or which confer desirable properties. Following diversification by one or more of the methods herein, or otherwise available to one of skill, any nucleic acids that are produced can be selected for a desired activity or property, e.g., improved or altered host range, increased efficiency orprecision of insertion, targeted insertion, increased or decreased infectivity, decreased pathogenicity, etc. This can include identifying any activity that can be detected, for example, in an automated or automatable format, e.g., by any of the assays in the art as described above and in the examples below. A variety of related (or even unrelated) properties can be evaluated, in serial or in parallel, at the discretion of the practitioner.

Descriptions of a variety of diversity generating procedures for generating modified nucleic acid sequences contributing to improved plant transformation vectors are found the following publications and the references cited therein: Stemmer, et al. (1999) "Molecular breeding of viruses for targeting and other clinical properties" Tumor Targeting 4:1-4; Ness et al. (1999) "DNA Shuffling of subgenomic sequences of subtilisin" Nature Biotechnology 17:893-896; Chang et al. (1999) "Evolution of a cytokine using DNA family shuffling" Nature Biotechnology 17:793-797; Minshull and Stemmer (1999) "Protein evolution by molecular breeding" Current Opinion in Chemical Biology 3:284-290; Christians et al. (1999) "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling" Nature Biotechnology 17:259-264; Crameri et al. (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" Nature 391:288-291; Crameri et al. (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," Nature Biotechnology 15:436-438; Zhang et al. (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening" Proc. Natl. Acad. Sci. USA 94:4504-4509; Patten et al. (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines" Current Opinion in Biotechnology 8:724-733; Crameri et al. (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling" Nature Medicine 2:100-103; Crameri et al. (1996) "Improved green fluorescent protein by molecular evolution using DNA shuffling" Nature Biotechnology 14:315-319; Gates et al. (1996) "Affinity selective isolation of ligands from peptide libraries through display

COR-AAD1-00002310

on a lac repressor 'headpiece dimer'" Journal of Molecular Biology 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" In: The Encyclopedia of Molecular Biology. VCH Publishers, New York. pp.447-457; Crameri and Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes" BioTechniques 18:194-195; Stemmer et al., (1995) "Single-step assembly of a gene and entire plasmid form large numbers of oligodeoxy-ribonucleotides" Gene, 164:49-53; Stemmer (1995) "The Evolution of Molecular Computation" Science 270: 1510; Stemmer (1995) "Searching Sequence Space" Bio/Technology 13:549-553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" Nature 370:389-391; and Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution." Proc. Natl. Acad. Sci. USA 91:10747-10751.

Mutational methods of generating diversity include, for example, site-directed mutagenesis (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" Anal Biochem. 254(2): 157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" Methods Mol. Biol. 57:369-374; Smith (1985) "In vitro mutagenesis" Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" Science 229:1193-1201; Carter (1986) "Site-directed mutagenesis" Biochem. J. 237:1-7; and Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)); mutagenesis using uracil containing templates (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Methods in Enzymol. 154, 367-382; and Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" Science 242:240-245); oligonucleotide-directed mutagenesis (Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Zoller & Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" Nucleic Acids Res. 10:6487-6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" Methods in Enzymol. 100:468-500; and Zoller & Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" Methods in Enzymol. 154:329-350);

COR-AAD1-00002311

phosphorothioate-modified DNA mutagenesis (Taylor et al. (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" Nucl. Acids Res. 13: 8749-8764; Taylor et al. (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" Nucl. Acids Res. 13: 8765-8787 (1985); Nakamaye & Eckstein (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" Nucl. Acids Res. 14: 9679-9698; Sayers et al. (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" Nucl. Acids Res. 16:791-802; and Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" Nucl. Acids Res. 16: 803-814); mutagenesis using gapped duplex DNA (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" Nucl. Acids Res. 12: 9441-9456; Kramer & Fritz (1987) Methods in Enzymol. "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154:350-367; Kramer et al. (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" Nucl. Acids Res. 16: 7207; and Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro" Nucl. Acids Res. 16: 6987-6999).

Additional suitable methods include point mismatch repair (Kramer et al. (1984) "Point Mismatch Repair" Cell 38:879-887), mutagenesis using repair-deficient host strains (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" Nucl. Acids Res. 13: 4431-4443; and Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" Methods in Enzymol. 154: 382-403), deletion mutagenesis (Eghtedarzadeh & Henikoff (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), restriction-selection and restriction-selection and restriction-purification (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317: 415-423), mutagenesis by total gene synthesis (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" Nucl. Acids Res. 14: 6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" Gene 34:315-

COR-AAD1-00002312

323; and Grundström et al. (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" Nucl. Acids Res. 13: 3305-3316), double-strand break repair (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" Current Opinion in Biotechnology 4:450-455. "Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA, 83:7177-7181). Additional details on many of the above methods can be found in Methods in Enzymology Volume 154, which also describes useful controls for trouble-shooting problems with various mutagenesis methods.

Additional details regarding various diversity generating methods can be found in the following U.S. patents, PCT publications, and EPO publications: U.S. Pat. No. 5,605,793 to Stemmer (February 25, 1997), "Methods for In Vitro Recombination;" U.S. Pat. No. 5,811,238 to Stemmer et al. (September 22, 1998) "Methods for Generating polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" U.S. Pat. No. 5,830,721 to Stemmer et al. (November 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly;" U.S. Pat. No. 5,834,252 to Stemmer, et al. (November 10, 1998) "End-Complementary Polymerase Reaction;" U.S. Pat. No. 5,837,458 to Minshull, et al. (November 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer and Crameri, "Mutagenesis by Random Fragmentation and Reassembly;" WO 96/33207 by Stemmer and Lipschutz "End Complementary Polymerase Chain Reaction;" WO 97/20078 by Stemmer and Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" WO 97/35966 by Minshull and Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering;" WO 99/41402 by Punnonen et al. "Targeting of Genetic Vaccine Vectors;" WO 99/41383 by Punnonen et al. "Antigen Library Immunization;" WO 99/41369 by Punnonen et al. "Genetic Vaccine Vector Engineering;" WO 99/41368 by Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" EP 752008 by Stemmer and Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly;" EP 0932670 by Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" WO 99/23107 by Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling;" WO 99/21979 by Apt et al., "Human Papillomavirus Vectors;" WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" WO 98/27230 by Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering;" WO 98/27230 by Stemmer et al.,

COR-AAD1-00002313

"Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection," WO 00/00632, "Methods for Generating Highly Diverse Libraries," WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences," WO 98/42832 by Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers," WO 99/29902 by Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences," WO 98/41653 by Vind, "An in Vitro Method for Construction of a DNA Library," WO 98/41622 by Borchert et al., "Method for Constructing a Library Using DNA Shuffling," and WO 98/42727 by Pati and Zarling, "Sequence Alterations using Homologous Recombination."

Certain U.S. applications provide additional details regarding various diversity generating methods, including "SHUFFLING OF CODON ALTERED GENES" by Patten et al. filed September 28, 1999, (USSN 09/407,800); "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION", by del Cardayre et al. filed July 15, 1998 (USSN 09/166,188), and July 15, 1999 (USSN 09/354,922); "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by Crameri et al., filed September 28, 1999 (USSN 09/408,392), and "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by Crameri et al., filed January 18, 2000 (PCT/US00/01203); "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" by Welch et al., filed September 28, 1999 (USSN 09/408,393); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed January 18, 2000, (PCT/US00/01202) and, e.g., "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed July 18, 2000 (USSN 09/618,579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" by Selifonov and Stemmer, filed January 18, 2000 (PCT/US00/01138); and "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" by Affholter, filed Sept. 6, 2000 (USSN 09/656,549).

In brief, several different general classes of sequence modification methods, such as mutation, recombination, etc. are applicable to the present invention and set forth, e.g., in the references above. The following exemplify some of the different types of preferred formats for diversity generation in the context of the present invention,

COR-AAD1-00002314

including, e.g., certain recombination based diversity generation formats, including a variety of DNA and RNA shuffling procedures.

Nucleic acids can be recombined in vitro by any of a variety of techniques discussed in the references above, including e.g., DNAse digestion of nucleic acids to be recombined followed by ligation and/or PCR reassembly of the nucleic acids. For example, sexual PCR mutagenesis can be used in which random (or pseudo random, or even non-random) fragmentation of the DNA molecule is followed by recombination, based on sequence similarity, between DNA molecules with different but related DNA sequences, in vitro, followed by fixation of the crossover by extension in a polymerase chain reaction. This process and many process variants is described in several of the references above, e.g., in Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751. Thus, any of the vectors, e.g., agrobacterium plasmids or viral vectors, or their components, described herein can be recombined in vitro according to the methods of the invention.

Similarly, nucleic acids can be recursively recombined in vivo, e.g., by allowing recombination to occur between nucleic acids in cells. Many such in vivo recombination formats are set forth in the references noted above. Such formats optionally provide direct recombination between nucleic acids of interest, or provide recombination between vectors, viruses, plasmids, etc., comprising the nucleic acids of interest, as well as other formats. Details regarding such procedures are found in the references noted above. Thus, nucleic acids comprising plant vectors, or components of such vectors can be recombined in vivo in cells to produce plant transformation vectors with improved properties as described herein. Additional details regarding one particularly favorable embodiment of in vivo recombination based on recombination of RNA viral vectors and components thereof are provided hereinbelow.

Whole genome recombination methods can also be used in which whole genomes of cells or other organisms are recombined, optionally including spiking of the genomic recombination mixtures with desired library components (e.g., genes corresponding to the pathways of the present invention). These methods have many applications, including those in which the identity of a target gene is not known. Details on such methods are found, e.g., in WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" and in, e.g., PCT/US99/15972 by del Cardayre et al., also entitled "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination."

COR-AAD1-00002315

Synthetic recombination methods can also be used, in which oligonucleotides corresponding to targets of interest are synthesized and reassembled in PCR or ligation reactions which include oligonucleotides which correspond to more than one parental nucleic acid, thereby generating new recombined nucleic acids. Oligonucleotides can be made by standard nucleotide addition methods, or can be made, e.g., by tri-nucleotide synthetic approaches. Details regarding such approaches are found in the references noted above, including, e.g., "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by Crameri et al., filed September 28, 1999 (USSN 09/408,392), and "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" by Crameri et al., filed January 18, 2000 (PCT/US00/01203); "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" by Welch et al., filed September 28, 1999 (USSN 09/408,393); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al. , filed January 18, 2000, (PCT/US00/01202); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" by Selifonov and Stemmer (PCT/US00/01138), filed January 18, 2000; and, e.g., "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed July 18, 2000 (USSN 09/618,579).

In silico methods of recombination can be effected in which genetic algorithms are used in a computer to recombine sequence strings which correspond to homologous (or even non-homologous) nucleic acids. The resulting recombined sequence strings are optionally converted into nucleic acids by synthesis of nucleic acids which correspond to the recombined sequences, e.g., in concert with oligonucleotide synthesis/ gene reassembly techniques. This approach can generate random, partially random or designed variants. Many details regarding in silico recombination, including the use of genetic algorithms, genetic operators and the like in computer systems, combined with generation of corresponding nucleic acids (and/or proteins), as well as combinations of designed nucleic acids and/or proteins (e.g., based on cross-over site selection) as well as designed, pseudo-random or random recombination methods are described in "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al. , filed January 18, 2000, (PCT/US00/01202)

COR-AAD1-00002316

"METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" by Selifonov and Stemmer (PCT/US00/01138), filed January 18, 2000; and, e.g., "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed July 18, 2000 (USSN 09/618,579). Extensive details regarding in silico recombination methods are found in these applications. This methodology is generally applicable to the present invention in providing for recombination of plant transformation vectors and their components in silico and/ or the generation of corresponding nucleic acids or proteins.

In addition to the nucleic acid recombination methods described above, the present invention specifically provides a format for in vivo RNA recombination or "RNA shuffling" that is favorably employed in the generation of, e.g., plant virus vectors with improved properties. Nucleic acids encoding, e.g., coat proteins, movements proteins, replicases, or other components of viral vectors, or subsequences thereof, are inserted into RNA viral vectors. In the context of developing plant viral vectors and related sequences, or other sequences that have a site of action in plants, plant viruses are the vector of choice. However, it will be understood that this method is equally applicable to any type of RNA virus, and can be adapted by appropriate selection of virus and cell type to perform in vivo recombination of RNA substrates. Selection of an appropriate viral vector is within the discretion of the practitioner and can largely be determined by the cell type wherein expression is desired and/or by the mode of action or site of action of the gene of interest.

In some instances it will be desirable to insert cDNA or other DNA sequences of interest into a DNA transcription vector capable of giving rise to infectious viral RNA transcripts. The methods for so doing are well established in the art, and referenced below. For example, cDNAs, oligonucleotides, genomic fragments, or other sequences encoding, e.g., coat proteins, or subportions of coat proteins, or inactive or active gene homologs that are coat protein gene related, (or movement proteins, or replicases, or the like) can be cloned into reverse transcribed, double stranded viral cDNA molecules, which are optionally components of autonomously replicating vectors such as plasmids, episomes, T DNAs, transposons, and the like.

In either case, a population of viral vectors, each comprising a variant of the gene of interest, is introduced into plant cells or tissues such that a single plant cell or tissue receives multiple different variants of the gene of interest. If infectious transcripts

COR-AAD1-00002317

are used, following inoculation, RNA transcripts are cytoplasmically replicated under the control of viral replication sequences located, typically, within the 5' terminal region of the transcript. Alternatively, after introduction, e.g., by electroporation, microinjection, or agrobacterium mediated transformation, the cDNA vector gives rise to RNA transcripts, which are then replicated in the cytoplasm of the cell by the viral RNA polymerase.

Both homologous and non-homologous recombination occur in RNA viruses, and both processes are believed to be mediated by template switching of the viral RNA-dependent RNA polymerase during replication. Specific mutations have been identified within viral RNA polymerases that affect the frequency of homologous or non-homologous RNA recombination. Accordingly, the RNA polymerase can be selected to bias the recombination process to acheive the desired outcome with respect to diversity generation. Alternatively, RNA shuffling as described herein, or other nucleic acid shuflling methods can be used to derive viral RNA polymerases with enhanced homologous and/or non-homologous RNA recombination activity.

In an embodiment, viral vectors containing complementary mutations in proteins required for systemic spread of the virus are used to select variants of the vector, or of a gene of interest inserted into the vector. For example, to evolve an exogenous gene of interest, a viral vector is constructed including, in the direction of transcription: a RNA-dependent RNA polymerase (RdRp, e.g., from Potato Virus X); essential movement protein encoding sequences under regulatory control of a first subgenomic promoter; a variant of a gene of interest under regulatory control of a second subgenomic promoter; and coat protein under regulatory control of a third subgenomic promoter.

Multiple members of a population of vectors having alternative mutations in, e.g., a movement protein, a coat protein, or other viral component of interest (or in an exogenous gene of interest inserted into the vector) are introduced into, e.g., a basal leaf of an intact plant. In a preferred embodiment, the mutations are complementary such that only variants that have undergone recombination between the complementary mutations, e.g., in the viral component, or gene, of interest, will be capable of systemic infection and movement throughout the plant. Thus, sampling of distal leaves, e.g., those higher on the plant, provides a simple means of screening and selecting recombined viral vectors. In addition, this technology provides the benefit that recombination and expression are acheived in vivo in a single step.

COR-AAD1-00002318

Many methods of accessing natural diversity, e.g., by hybridization of diverse nucleic acids or nucleic acid fragments to single-stranded templates, followed by polymerization and/or ligation to regenerate full-length sequences, optionally followed by degradation of the templates and recovery of the resulting modified nucleic acids can be similarly used. In one method employing a single-stranded template, the fragment population derived from the genomic library(ies) is annealed with partial, or, often approximately full length ssDNA or RNA corresponding to the opposite strand. Assembly of complex chimeric genes from this population is then mediated by nuclease-base removal of non-hybridizing fragment ends, polymerization to fill gaps between such fragments and subsequent single stranded ligation. The parental polynucleotide strand can be removed by digestion (e.g., if RNA or uracil-containing), magnetic separation under denaturing conditions (if labeled in a manner conducive to such separation) and other available separation/purification methods. Alternatively, the parental strand is optionally co-purified with the chimeric strands and removed during subsequent screening and processing steps. Additional details regarding this genral approach are found, e.g., in "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" by Affholter, USSN 09/656,549, filed Sept. 6, 2000.

In another approach, single-stranded molecules are converted to double-stranded DNA (dsDNA) and the dsDNA molecules are bound to a solid support by ligand-mediated binding. After separation of unbound DNA, the selected DNA molecules are released from the support and introduced into a suitable host cell to generate a library enriched sequences which hybridize to the probe. A library produced in this manner provides a desirable substrate for further diversification using any of the procedures described herein.

Any of the preceding general recombination formats can be practiced in a reiterative fashion (e.g., one or more cycles of mutation/recombination or other diversity generation methods, optionally followed by one or more selection methods) to generate a more diverse set of recombinant nucleic acids.

Mutagenesis employing polynucleotide chain termination methods have also been proposed (*see*, e.g., U.S. Patent No. 5,965,408, "Method of DNA reassembly by interrupting synthesis" to Short, and the references above), and can be applied to the present invention. In this approach, double stranded DNAs corresponding to one or more genes sharing regions of sequence similarity are combined and denatured, in the presence

COR-AAD1-00002319

or absence of primers specific for the gene. The single stranded polynucleotides are then annealed and incubated in the presence of a polymerase and a chain terminating reagent (e.g., ultraviolet, gamma or X-ray irradiation; ethidium bromide or other intercalators; DNA binding proteins, such as single strand binding proteins, transcription activating factors, or histones; polycyclic aromatic hydrocarbons; trivalent chromium or a trivalent chromium salt; or abbreviated polymerization mediated by rapid thermocycling; and the like), resulting in the production of partial duplex molecules. The partial duplex molecules, e.g., containing partially extended chains, are then denatured and reannealed in subsequent rounds of replication or partial replication resulting in polynucleotides which share varying degrees of sequence similarity and which are diversified with respect to the starting population of DNA molecules. Optionally, the products, or partial pools of the products, can be amplified at one or more stages in the process. Polynucleotides produced by a chain termination method, such as described above, are suitable substrates for any other described recombination format.

Diversity also can be generated in nucleic acids or populations of nucleic acids using a recombinational procedure termed "incremental truncation for the creation of hybrid enzymes" ("ITCHY") described in Ostermeier et al. (1999) "A combinatorial approach to hybrid enzymes independent of DNA homology" Nature Biotech 17:1205. This approach can be used to generate an initial a library of variants which can optionally serve as a substrate for one or more in vitro or in vivo recombination methods. See, also, Ostermeier et al. (1999) "Combinatorial Protein Engineering by Incremental Truncation," Proc. Natl. Acad. Sci. USA, 96: 3562-67; Ostermeier et al. (1999), "Incremental Truncation as a Strategy in the Engineering of Novel Biocatalysts," Biological and Medicinal Chemistry, 7: 2139-44.

Mutational methods which result in the alteration of individual nucleotides or groups of contiguous or non-contiguous nucleotides can be favorably employed to introduce nucleotide diversity. For example, any of a variety of in vivo (e.g., exposure to chemical mutagens, passage through mutator cells lines) and/or in vitro (e.g., mutagenic PCR, site-specific mutagenesis) procedures can be used to diversify substrates corresponding to the vectors of the invention. Many mutagenesis methods are found in the above-cited references; additional details regarding mutagenesis methods can be found in following, which can also be applied to the present invention.

For example, error-prone PCR can be used to generate nucleic acid variants. Using this technique, PCR is performed under conditions where the copying

COR-AAD1-00002320

fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. Examples of such techniques are found in the references above and, e.g., in Leung et al. (1989) Technique 1:11-15 and Caldwell et al. (1992) PCR Methods Applic. 2:28-33. Similarly, assembly PCR can be used, in a process which involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions can occur in parallel in the same reaction mixture, with the products of one reaction priming the products of another reaction.

Oligonucleotide directed mutagenesis can be used to introduce site-specific mutations in a nucleic acid sequence of interest. Examples of such techniques are found in the references above and, e.g., in Reidhaar-Olson et al. (1988) Science, 241:53-57. Similarly, cassette mutagenesis can be used in a process that replaces a small region of a double stranded DNA molecule with a synthetic oligonucleotide cassette that differs from the native sequence. The oligonucleotide can contain, e.g., completely and/or partially randomized native sequence(s).

Recursive ensemble mutagenesis is a process in which an algorithm for protein mutagenesis is used to produce diverse populations of phenotypically related mutants, members of which differ in amino acid sequence. This method uses a feedback mechanism to monitor successive rounds of combinatorial cassette mutagenesis. Examples of this approach are found in Arkin & Youvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815.

Exponential ensemble mutagenesis can be used for generating combinatorial libraries with a high percentage of unique and functional mutants. Small groups of residues in a sequence of interest are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. Examples of such procedures are found in Delegrave & Youvan (1993) Biotechnology Research 11:1548-1552.

In vivo mutagenesis can be used to generate random mutations in any cloned DNA of interest by propagating the DNA, e.g., in a strain of *E. coli* that carries mutations in one or more of the DNA repair pathways. These "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in one of these strains will eventually generate random mutations within the DNA. Such procedures are described in the references noted above.

COR-AAD1-00002321

Other procedures for introducing diversity into a genome, e.g. a bacterial, fungal, animal or plant genome can be used in conjunction with the above described and/or referenced methods. For example, in addition to the methods above, techniques have been proposed which produce nucleic acid multimers suitable for transformation into a variety of species (*see*, e.g., Schellenberger U.S. Patent No. 5,756,316 and the references above). Transformation of a suitable host with such multimers, consisting of genes that are divergent with respect to one another, (e.g., derived from natural diversity or through application of site directed mutagenesis, error prone PCR, passage through mutagenic bacterial strains, and the like), provides a source of nucleic acid diversity for DNA diversification, e.g., by an in vivo recombination process as indicated above.

Alternatively, a multiplicity of monomeric polynucleotides sharing regions of partial sequence similarity can be transformed into a host species and recombined in vivo by the host cell. Subsequent rounds of cell division can be used to generate libraries, members of which, include a single, homogenous population, or pool of monomeric polynucleotides. Alternatively, the monomeric nucleic acid can be recovered by standard techniques, e.g., PCR and/or cloning, and recombined in any of the recombination formats, including recursive recombination formats, described above.

Methods for generating multispecies expression libraries have been described (in addition to the reference noted above, *see*, e.g., Peterson et al. (1998) U.S. Pat. No. 5,783,431 "METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC PATHWAYS," and Thompson, et al. (1998) U.S. Pat. No. 5,824,485 METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC PATHWAYS) and their use to identify protein activities of interest has been proposed (In addition to the references noted above, *see*, Short (1999) U.S. Pat. No. 5,958,672 "PROTEIN ACTIVITY SCREENING OF CLONES HAVING DNA FROM UNCULTIVATED MICROORGANISMS"). Multispecies expression libraries include, in general, libraries comprising cDNA or genomic sequences from a plurality of species or strains, operably linked to appropriate regulatory sequences, in an expression cassette. The cDNA and/or genomic sequences are optionally randomly ligated to further enhance diversity. The vector can be a shuttle vector suitable for transformation and expression in more than one species of host organism, e.g., bacterial species, eukaryotic cells. In some cases, the library is biased by preselecting sequences which encode a protein of interest, or which hybridize to a nucleic acid of interest. Any such libraries can be provided as substrates for any of the methods herein described.

COR-AAD1-00002322

The above descibed procedures have been largely directed to increasing nucleic acid and/ or encoded protein diversity. However, in many cases, not all of the diversity is useful, e.g., functional, and contributes merely to increasing the background of variants that must be screened or selected to identify the few favorable variants. In some applications, it is desirable to preselect or prescreen libraries (e.g., an amplified library, a genomic library, a cDNA library, a normalized library, etc.) or other substrate nucleic acids prior to diversification, e.g., by recombination-based mutagenesis procedures, or to otherwise bias the substrates towards nucleic acids that encode functional products. For example, in the case of antibody engineering, it is possible to bias the diversity generating process toward antibodies with functional antigen binding sites by taking advantage of in vivo recombination events prior to manipulation by any of the described methods. For example, recombined CDRs derived from B cell cDNA libraries can be amplified and assembled into framework regions (e.g., Jirholt et al. (1998) "Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework" Gene 215: 471) prior to diversifying according to any of the methods described herein.

Libraries can be biased towards nucleic acids which encode proteins with desirable enzyme activities. For example, after identifying a clone from a library which exhibits a specified activity, the clone can be mutagenized using any known method for introducing DNA alterations. A library comprising the mutagenized homologues is then screened for a desired activity, which can be the same as or different from the initially specified activity. An example of such a procedure is proposed in Short (1999) U.S. Patent No. 5,939,250 for "PRODUCTION OF ENZYMES HAVING DESIRED ACTIVITIES BY MUTAGENESIS." Desired activities can be identified by any method known in the art. For example, WO 99/10539 proposes that gene libraries can be screened by combining extracts from the gene library with components obtained from metabolically rich cells and identifying combinations which exhibit the desired activity. It has also been proposed (e.g., WO 98/58085) that clones with desired activities can be identified by inserting bioactive substrates into samples of the library, and detecting bioactive fluorescence corresponding to the product of a desired activity using a fluorescent analyzer, e.g., a flow cytometry device, a CCD, a fluorometer, or a spectrophotometer.

Libraries can also be biased towards nucleic acids which have specified characteristics, e.g., hybridization to a selected nucleic acid probe. For example,

COR-AAD1-00002323

application WO 99/10539 proposes that polynucleotides encoding a desired activity (e.g., an enzymatic activity, for example: a lipase, an esterase, a protease, a glycosidase, a glycosyl transferase, a phosphatase, a kinase, an oxygenase, a peroxidase, a hydrolase, a hydratase, a nitrilase, a transaminase, an amidase or an acylase) can be identified from among genomic DNA sequences in the following manner. Single stranded DNA molecules from a population of genomic DNA are hybridized to a ligand-conjugated probe. The genomic DNA can be derived from either a cultivated or uncultivated microorganism, or from an environmental sample. Alternatively, the genomic DNA can be derived from a multicellular organism, or a tissue derived therefrom. Second strand synthesis can be conducted directly from the hybridization probe used in the capture, with or without prior release from the capture medium or by a wide variety of other strategies known in the art. Alternatively, the isolated single-stranded genomic DNA population can be fragmented without further cloning and used directly in, e.g., a recombination-based approach, that employs a single-stranded template, as described above.

"Non-Stochastic" methods of generating nucleic acids and polypeptides are alleged in Short "Non-Stochastic Generation of Genetic Vaccines and Enzymes" WO 00/46344. These methods, including proposed non-stochastic polynucleotide reassembly and site-saturation mutagenesis methods be applied to the present invention as well. Random or semi-random mutagenesis using doped or degenerate oligonucleotides is also described in, e.g., Arkin and Youvan (1992) "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis" *Biotechnology* 10:297-300; Reidhaar-Olson et al. (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes" *Methods Enzymol.* 208:564-86; Lim and Sauer (1991) "The role of internal packing interactions in determining the structure and stability of a protein" *J. Mol. Biol.* 219:359-76; Breyer and Sauer (1989) "Mutational analysis of the fine specificity of binding of monoclonal antibody 51F to lambda repressor" *J. Biol. Chem.* 264:13355-60); and "Walk-Through Mutagenesis" (Crea, R; US Patents 5,830,650 and 5,798,208, and EP Patent 0527809 B1.

It will readily be appreciated that any of the above described techniques suitable for enriching a library prior to diversification can also be used to screen the products, or libraries of products, produced by the diversity generating methods.

Kits for mutagenesis, library construction and other diversity generation methods are also commercially available. For example, kits are available from, e.g., Stratagene (e.g., QuickChange™ site-directed mutagenesis kit; and Chameleon™ double-

COR-AAD1-00002324

stranded, site-directed mutagenesis kit), Bio/Can Scientific, Bio-Rad (e.g., using the Kunkel method described above), Boehringer Mannheim Corp., Clonetech Laboratories, DNA Technologies, Epicentre Technologies (e.g., 5 prime 3 prime kit); Genpak Inc, Lemargo Inc, Life Technologies (Gibco BRL), New England Biolabs, Pharmacia Biotech, Promega Corp., Quantum Biotechnologies, Amersham International plc (e.g., using the Eckstein method above), and Anglian Biotechnology Ltd (e.g., using the Carter/Winter method above).

The above references provide many mutational formats, including recombination, recursive recombination, recursive mutation and combinations or recombination with other forms of mutagenesis, as well as many modifications of these formats. Regardless of the diversity generation format that is used, the nucleic acids of the invention can be recombined (with each other, or with related (or even unrelated) sequences) to produce a diverse set of recombinant nucleic acids, including, e.g., sets of homologous nucleic acids, as well as corresponding polypeptides.

Following recombination and/or mutagenesis, any nucleic acids which are produced can be selected for a desired activity. In the context of the present invention, this can include testing for any property desired in a plant vector, as given in the illustrative examples above, by any of the assays in the art.

A recombinant nucleic acid produced, e.g., by recursively recombining one or more polynucleotide of the invention with one or more additional nucleic acid also forms a part of the invention. The one or more additional nucleic acid can include another polynucleotide of the invention; optionally, alternatively, or in addition, the one or more additional nucleic acid can include, e.g., a nucleic acid encoding a naturally-occurring vector or vector component, or a subsequence thereof, or any homologous sequence or subsequence thereof, (e.g., as found in Genbank or other available literature, or newly identified), or, e.g., any other homologous or non-homologous nucleic acid (certain recombination formats noted above, notably those performed synthetically or in silico, do not require homology for recombination).

The recombining steps may be performed in vivo, in vitro, or in silico as described in more detail in the references above. Also included in the invention is a cell containing any resulting recombinant nucleic acid, nucleic acid libraries produced by recursive recombination of the nucleic acids set forth herein, and populations of cells, vectors, viruses, plasmids or the like comprising the library or comprising any recombinant nucleic acid resulting from recombination (or recursive recombination) of a

COR-AAD1-00002325

nucleic acid as set forth herein with another such nucleic acid, or an additional nucleic acid. Corresponding sequence strings in a database present in a computer system or computer readable medium are a feature of the invention.

For the purposes of the present invention, the above methodologies either singly or in combination are used to evolve novel plant vectors. Such vectors can be any polynucleotide, RNA or DNA, or conjugated DNA, (e.g. poly-lysine-conjugated DNA, peptide-conjugated DNA, liposome-conjugated DNA) suitable for introduction into a host plant or host plant cell by any methodology known in the art. Such vectors can also be plasmids, derived from or incorporated into various agrobacterium strains, or indeed the Agrobacteria themselves. Similarly, the vectors can be plant viruses comprising either RNA or DNA genomes. Any of the foregoing are readily amenable to the aforementioned DNA shuffling techniques.

The methods of the invention are adapted to each application through the choice of substrates and the methods of screening or selection. For example, to evolve an Agrobacterium vector with an expanded host range, sequences, actual or virtual, corresponding to *vir* genes are provided as the substrates for recombination in vitro or in vivo, or in silico, respectively. The *vir* genes can originate as components of plasmids derived directly or indirectly from one or more agrobacterium strains, or from partial or complete Agrobacterium genomes. Additional diversity can be introduced by any one of: random mutagenesis in the generation of synthetic oligonucleotides, by error-prone PCR, mutagenic amplification, or site directed mutagenesis. After one or more cycles of recombination, e.g., shuffling, screening is performed based on the property desired. In the foregoing example, recombinant virulence genes (or plasmids, or genomic sequences) are screened or selected based on their ability to induce other virulence gene promoters, or respond to signals from a plant host.

Vectors of the invention with other desirable characteristics are developed by the choice of substrate sequences, and/or screening techniques which will be apparent to one of skill in the art. In vitro techniques based on the characteristics of polynucleotides, such as PCR, LCR, nucleic acid hybridization analysis, or on the characteristics of proteins, e.g. western hybridization, proteomics, are the method of choice in some instances, e.g., evaluation of T-DNA insertion sites, as will be readily apparent to one knowledgeable in the art.

Alternatively, screening in bacteria (e.g., *E.coli*) or agrobacteria, is appropriate, for example: assays for *vir* gene induction used in screening for variant host

COR-AAD1-00002326

range. References contained herein and previously discussed, (e.g. Sambrook, Berger, and Ausubel) provide numerous techniques relevant for the introduction of the nucleic acids of the invention, e.g., evolved viral or agrobacterial vectors or their components, into bacterial and agrobacterial vectors and strains. Bacteria, including *E.coli* and Agrobacteria can be evaluated for expression of recombinant genes as well as for numerous other characteristics particular to the individual circumstances.

In other cases, the screening is best performed in plant cells or intact plants. For example, evaluation of the rate of systemic spread of a virus after infection can be performed by assessing viral production in leaves distal to the site of inoculation on an intact plant. These examples are provided by way of illustration, and variants and adaptations appropriate to the specific characteristic under evaluation will be apparent to one of skill in the art.

EXAMPLES

AGROBACTERIUM VECTORS WITH EXTENDED HOST RANGE

The methods of the present invention provide for the evolution of agrobacterium vectors with a broad host range. Agrobacterium vectors currently in widespread use transform many dicot species with efficiency. However, monocot species, including most important food crop plants have been less readily amenable to agrobacterium mediated transformation. At present, transformation of many monocot species requires labor intensive plant cell culture and/or the addition of exogenous compounds to stimulate expression of *vir* region genes necessary for T-DNA transfer, making the procedure too inefficient for routine use. It would be extremely useful to have defined agrobacterium strains that easily and efficiently transform both dicot and monocot crop species. For example, Agrobacterium strains which can infect, in planta, leaves or cut stems of both dicot and monocot species without the addition of exogenous phytohormones, or in vitro culture, is of particular interest, and would drastically simplify the production of transgenic food crop species.

The capacity to transform, or to transfer a T-DNA, is largely dependent on the ability of a given agrobacterium strain to detect and respond to phytochemical signals emitted by the host plant. To extend the host range of an agrobacterium strain involves tuning the receptor/signaling pathway to respond to molecules given off by the plant species to be targeted. One approach to evolving agrobacterium strains which transform

COR-AAD1-00002327

monocot species is to adapt the VirA and ChvE sensor molecules to recognize the phenolic and saccharide signals of the plant. The *virA* gene is a plasmid-borne sequence belonging to the virulence region of the Ti plasmid in wild-type *A. tumefaciens*, while the *chvE* locus is chromosomal. Fragments representing either of the *virA* gene or the *chvE* gene can be recombined by any of the recursive recombination methodologies described above to generate novel *virA* or *chvE* sequences. Alternatively, both genes can be recombined, or recursively recombined, e.g., shuffled, simultaneously, as isolated or cloned sequences, as PCR fragments, on plasmids, or in the context of whole genome shuffling of agrobacterium strains.

Screening for *virA* and *chvE* sequences which extend host range is performed in a two step process. Firstly, recombinant VirA and/or ChvE molecules are screened for the ability to induce other genes of the virulence region. For example, a recombinant library comprising shuffled *virA* and/or *chvE* sequences is transformed, e.g., by electroporation, into an agrobacterium or other bacterial strain carrying a suitably responsive reporter gene, such as a β-galactosidase (*lacZ*) structural gene regulated by a *vir* or *pinF* promoter. The agrobacterium are then incubated in the presence of plant extracts derived from the host species of interest and β -galactosidase activity is measured, (*see*, e.g., Grimsley et. al. (1989) "DNA transfer from *Agrobacterium* to *Zea mays* or *Brassica* by agroinfection is dependent on bacterial virulence functions." Mol.Gen.Genet. 217:309-316.) Secondly, agrobacterium strains which demonstrate high levels of β -galactosidase activity are evaluated for their ability to transform a broad range of host species. For example, an agrobacterium strain carrying a recombinant *virA* library is incubated in culture with representative plant cells, for example leaf, root, meristem, callus, protoplast or embryo, whatever is most easily regenerated for each plant species. Culture and inoculation can be performed in any high throughput format, e.g. 96 well plates. Cultures are then subjected to selection, for example for antibiotic resistance, e.g., kanamycin, imparted by integration of a T-DNA comprising the selectable marker. Those agrobacterium strains which transform a broad range of host plant species are then recovered, and if so desired, used as the starting material for additional rounds of shuffling.

In a similar manner, the *virG* locus, responsible for the "supervir" phenotype, (Jin et. al. (1987) "Genes responsible for the supervirulence phenotype of

COR-AAD1-00002328

Agrobacterium tumefaciens A281." J. Bacteriol. 169:4417-4425.) can be recombined. Screening and selection is performed as indicated above.

AGROBACTERIUM EXHIBITING INSERT PRECISION

For the purposes of this discussion, insert precision is defined as the insertion of a T-DNA into a host plant chromosome in a simple, defined manner without disruption of flanking sequences. It is closely related to the concept of targeted insertion of a T-DNA at a predetermined locus in the host plant chromosome. The present invention includes methods for evolving agrobacterium strains which integrate T-DNAs in a precise and/or targeted manner, as well as the agrobacterium strains which exhibit these properties.

T-DNA borders are not always cut precisely in the process of their release from the agrobacterium. Furthermore, insertion into the plant chromosome is frequently flanked by deletions, duplications and insertions. The T-DNA can be concatamerized, truncated, or inserted at multiple and unlinked locations in the host genome. Agrobacterium strains influence the process of insertion as evidenced by the observation that different strains consistently result in more or less complicated insertion patterns. Typically, nopaline strains are more efficient at transformation but result in complex insertion patterns, while octopine strains are less efficient at transformation but result in simpler insertion patterns. This is of concern, because there is an inverse correlation between the number of inserts and the expression of the inserted genes. Combining the properties of high transformation efficiency and simple, single inserts would make high throughput screening more consistent and reproducible.

Insertion of a T-DNA into a plant chromosome appears to be at least partially dependent on short regions of sequence similarity between the T-DNA border and the plant chromosome, so-called "microhomology" regions. To evolve agrobacterium that insert T-DNAs simply and precisely into the host plant genome, and/or insert T-DNAs into a predetermined locus of the plant chromosome, recombination, e.g., shuffling, of at least two potential types of targets is warranted. On the one hand, the T-DNA border sequences themselves, (and in concert with adjacent regions of DNA), and on the other hand the *virD* and *virE* genes which encode proteins responsible for the excision and insertion of the T-DNA strand are selected as the substrates for recursive recombination as previously described. Sequences from both octopine and nopaline strains can be recombined and screened to identify recombinant

COR-AAD1-00002329

sequences which result in targeted insertions and/or insert precision. Characterization of the insertion sites resulting from transformation by agrobacterium carrying the library of recombined sequences can be performed by southern analysis to determine copy number, and by inverse PCR to define the limits of microhomology and the nature of the flanking sequences to determine the complexity of the insertion event. Those agrobacterium which give rise to precisely integrated T-DNAs at high efficiency can be isolated, and optionally subjected to additional rounds of DNA recombination.

Agrobacterium which integrate T-DNAs to a predetermined site in the host plant genome can be developed by including regions of sequence similarity adjacent to the T-DNA borders in the fragments provided as substrates for recombination or recursive recombination (e.g., shuffling) reactions. Selection protocols based on the activation of a marker at a predetermined site in the plant genome are particularly suited to screening for targeted integration events. For example, to identify agrobacterium strains, e.g., comprising a shuffled *virE* library, which mediate targeted integration events, a binary vector system is employed. Shuffled *virE* sequences are cloned into a helper plasmid and transformed into agrobacterium carrying a plasmid comprising a T-DNA. The T-DNA is constructed to include regions of sequence similarity with a predetermined locus in the host plant genome adjacent to the T-DNA borders, or alternatively, adjacent to a single right T-DNA border. The T-DNA also includes a promoter sequence between or adjacent to the regions of sequence similarity. The agrobacterium comprising the shuffled library is then used to transform transgenic plants or plant cells which have a selectable marker, for example a gene encoding kanamycin resistance, at the aforementioned predetermined site in their genome. The resulting transformants are subjected to selection in kanamycin: only those transformants which have activated the kanamycin resistance gene by targeted integration of the promoter T-DNA will survive the selection procedure.

TARGETING OF T-DNAS TO THE CHLOROPLAST

The development of agrobacterium vectors which target T-DNAs to the chloroplast is particularly suited to the methods of the invention. The sequences responsible for guiding proteins to the chloroplast, while sharing certain features, are poorly defined, and appear to lack any clear consensus sequence, placing addition of chloroplast targeting sequences beyond simple cloning operations. Most polypeptides targeted to the chloroplast have N-terminal sequences that are rich in serine and

COR-AAD1-00002330

threonine; are deficient in aspartate, glutamate, and tyrosine; and generally have a central domain rich in positively charged amino acids.

The agrobacterium VirD2 and VirE2 proteins play a role in transporting and integrating T-DNAs into the host cell nucleus. Both VirD2 and VirE2 contain nuclear localization signals (NLS) that appear to be involved in targeting the T-DNA to the plant cell nucleus. For example, the NLS of VirE2 is comprised of two amino acid sub-sequences rich in basic amino acids with the sequence NS1: KLRPEDRYVQTERYGRR; NS2: KRRYGGETEIKLKSK. While it is not certain that the VirE2 NLS are essential for transport of the T-DNA to the nucleus, they probably increase the efficiency of targeting.

Recombination, e.g., shuffling, of *virD* and *virE* sequences with the 5' regions, corresponding to the N-terminal transit sequence, of genes encoding chloroplast proteins gives rise to hybrid VirD2 and VirE2 proteins capable of localizing to the chloroplast. Selection is based e.g., on detecting a T-DNA encoded marker, such as green fluorescent protein (GFP) in the chloroplast. Assessment of GFP in the chloroplast can be performed either in situ by fluorescence microscopy or by isolation of chloroplasts and fluoroscopic detection of GFP.

INCREASED INSERT CAPACITY OF VIRUS VECTORS

Another aspect of the invention relates to the evolution of improved viral vectors. For example, plant viruses known to have double stranded circular DNA forms during their life cycle such as Geminiviruses (GV) or Caulimoviruses (CaMV) have many attractive attributes as plant transformation vectors, including high level and chromosome position-independent gene expression. However, due to strict genome size limitations, both CaMV and GV vectors have achieved only limited applicability as vectors.

Geminiviruses are related to a family of viruses that includes the *E. coli* phages phiX174 and M13, which replicate via a rolling circle mechanism. They are the only class of plant viruses that replicate in the nucleus depending exclusively on the host encoded DNA-dependent DNA polymerases. Geminiviruses typically possess bipartite genomes, e.g., which split the genes between two DNA molecules designated A and B: subgroup III (e.g., Tomato Golden Mosaic Virus (TGMV), African Cassava Mosaic Virus (ACMV), Bean Dwarf Mosaic Virus (BDMV), Bean Golden Mosaic Virus (BGMV), etc.). In some cases, the genome is monopartite with all coding capacity residing on one DNA molecule: subgroups I and II (e.g., Maize Streak Virus (MSV), Wheat Dwarf Virus

COR-AAD1-00002331

(WDV), Digitaria Streak Virus (DSV) and Tomato Leaf Curl Virus (TLCV) and tomato yellow Leaf Curl Virus (TYLCV), respectively). All Geminiviruses possess an "intergenic region" (IR) or "common region" (CR) that contains the origin of viral (+)-strand replication. In the genome of a bipartite GV, the CRs of genomes A and B are identical at the DNA sequence level and are highly conserved between closely related viruses.

The viral genomes(s) encode a Replicase protein (AC1 or AL1), a coat protein (designated CP or AV, or AR), and proteins involved in cell-to-cell movement as well as systemic spread. Typically, the Rep and coat protein are encoded by genome A and the movement functions by genome B of a bipartite genome. An exemplary experimental geminivirus system is illustrated in Figure 1. Two plasmids, designated MAXY-A and MAXY-B are constructed encoding, the replicase and a visual marker, e.g., GFP, and the BR1 and BL1 movement proteins, respectively.

Cell to cell movement of the bipartite GVs is governed by the BR1 and BL1 proteins. The *BR1* gene encodes a movement protein that interacts with and increases the pore size of the plasmodesmata to allow passage of unit viral genomes between cells. BR1 is now known to encode a ssDNA binding protein responsible for transport of ssDNA GV genomes across the nuclear envelope membrane. BR1 (or its equivalent in monopartite GVs lacking a BR1 gene) is involved in systemic movement in an appropriate host plant (Timmermans et al. (1994) Ann Rev Plant Physiol Plant Mol Biol 45:79). BL1 localizes to the cell periphery and interacts with BR1 to shuttle the GV genome from cell to cell. The AC2 gene also plays a role in systemic movement, encoding a transcriptional activator specific for the BR1 and CP promoters. Widespread use of GVs as vectors has been hampered by the observation that genomes larger that wild type become rearranged, e.g., deleted, indicating that movement is size restricted.

To increase the insert size that can be stably incorporated and transmitted between cells in an infected plant, the *BL1* an *BR1* genes can be diversified, e.g., by any of the recombination, recursive recombination or shuffling procedures described herein to produce a library of movement protein variants. The components can be diversified individually, simultaneously on a single cassette or vector, or as components of intact viruses. For example, figure 2 schematically illustrates recombination of movement proteins, e.g., from 3 "parental" sequences (**1**, **2**, and **3**) to generate variant movement proteins (**4**). To identify variants that successfully replicate and spread a larger genome, a vector is constructed including the *BL1* and/or *BR1* variants and incorporating a visual

COR-AAD1-00002332

marker, e.g., β-glucuronidase (GUS) or Green Fluorescent Protein (GFP)-GUS fusion that exceeds the size range of the wild-type GV of interest. For example, a library of recombinant movement proteins as exemplified by **4** is cut with restriction enzymes 1 and 2 (RS1 and RS2) and cloned into a "common region" (CR) cassette, e.g., in a T-DNA vector suitable for agroinfection as illustrated in Figure 3. A suitable plant explant, such as leaf disks, are infected, e.g., by agro-infiltration, and arrayed for evaluation. Alternatively, transformation is performed, e.g, in bean hypocotyledons by biolistic transformation. The arrayed explants are then co-cultivated with a replication competent/movement incompetent reporter, e.g., encoding a GUS-GFP or GFP-35S fusion protein as illustrated in figure 4, and evaluated for ability to spread a construct with a larger than wild-type insert. Expression of the GFP fusion protein is evaluated distal to the site of inoculation, e.g., using a handheld long-wave uv monitor.

RAPID SYSTEMIC SPREAD OF EVOLVED PLANT VIRUS VECTORS

One major drawback in assessing novel viral characteristics is the length of time required for the engineered virus to replicate and spread internally after inoculation. If the rate of systemic spread following initial infection were increased, high throughput screening and virus reisolation could be made significantly faster. Systemic spread depends mostly on movement proteins and their interactions with viral nucleic acids and plant plasmodesmata. Embodiments of the invention use recursive recombination, e.g., shuffling, of single virus movement proteins and/or of families of related viral movement proteins to increase the rate of systemic spread. In cases where the origin of viral assembly lies within the movement protein-coding region, oligonucleotides with the wild-type origin can be added to the recombination reaction to insure that recombinants with improved properties will be able to form intact viral particles. The resulting libraries of recombinant viral movement proteins are cloned into a viral vector and inoculated onto leaves of intact plants. At set time intervals, distal leaves are harvested and virus recovered. Those library members which yield high titers at earlier time points than a wild-type reference virus are subjected to additional cycles of shuffling. Similarly, host range can be expanded by manipulating the viral protein-plasmodesmata interaction, and assessing the ability of recombinant viruses to infect host plant species of interest.

COR-AAD1-00002333

REDUCTION OF VIRAL SYMPTOMS FOLLOWING INFECTION

In some host-virus reactions, viral symptoms are severe, making it difficult to assess infected leaves for biochemical properties or to purify protein from infected leaves. In most cases, viral coat proteins are the most important determinants of the host response, including the hypersensitivity reaction. Viral replicases and movement proteins may also play significant roles. An embodiment of the present invention relates to the evolution of plant viruses that result in diminished viral symptoms. Genes encoding coat protein, replicases, and movement proteins are all suitable substrates for the development of plant viruses which evoke reduced symptoms. Such sequences can be recombined, e.g., shuffled, individually, in combinations of families of related genes, simultaneously including members of multiple gene families, or together in the context of whole genome shuffling.

The critical determinant in evolving plant virus vectors which do not elicit significant detrimental symptoms or hypersensitive responses is the manner in which the recombinant library is screened. To evolve plant viruses which provoke diminished plant responses, screening is based on an observed reduction in symptoms. Typically, a hypersensitive response (HR) results in necrotic lesions near the infection site. Lesions are characterized by an accumulation of cytotoxic phytoalexins, alterations in the cell walls, the accumulation of pathogenesis-related (PR) proteins and rapid cell death, and as such can be evaluated both visually and biochemically. Investigations into individual host-pathogen interactions have led to the cloning of individual genes involved in the HR, (*see*, for example Karrer et. al., (1998) "Cloning of tobacco genes that elicit the hypersensitive response." Plant Molecular Biology 36: 681-190.) However, such investigations are time consuming and laborious. In contrast, recursive recombination of viral genes, and genomes rapidly results in the rapid isolation of viruses which fail to elicit the HR response, as well as those which exhibit reduced symptoms in general.

Alternatively, toxicity due to a viral component can prevent high level expression of a virus or virus-derived gene of interest. For example, high levels of the AC1 (AL1) replicase of Geminiviruses induces efficient replication and, thus, expression of a gene of interest incorporated by a Geminivirus (e.g., African cassava mosaic virus, ACMV) vector. However, the AC1 replicase is highly toxic to the cell, inducing a continuous S-phase of the cell cycle.

Replication of a geminivirus genome requires three basic components: a Rep protein (Rep) encoded by AC1 (AL1), a cognate CR, and host DNA synthesis

COR-AAD1-00002334

machinery. Although Rep is the only viral encoded protein required for DNA replication, another small protein encoded by the viral AC3 gene greatly enhances replication (Sunter et al. (1990) Virology 179:69). Amplification of double stranded viral DNA as well as rolling circle production of viral (+) ssDNA is initiated by Rep protein. This protein binds to a specific recognition sequence in the CR, creates a ssDNA nick then covalently links a conserved tyrosine residue within itself to a specific base within the nearby conserved loop structure on the (+) DNA strand. This results in the initiation of (+) strand DNA replication by the host cell DNA synthesis machinery. In addition to this function, the Rep proteins have a putative retinoblastoma susceptibility protein (Rb) binding domain as has been demonstrated for the related tumorigenic SV40 T-antigen. In mammailian cells, Rb inhibits the activity of a family of transcription factors that induce expression of genes required for DNA synthesis. The T-antigen of SV40, and other related proteins, send quiescent cells into S-phase thru binding and deactivating Rb and related proteins. In addition, the Rep proteins may also directly activate transcription via interaction with plant proliferating cell nuclear antigen (PCNA). Rep proteins of the geminiviurses WDV and TGMV have been shown to bind to Rb proteins. Furthermore, the WDV Rep protein contains the conserved Rb binding motif LXCXE found in SV40 T-antigen (and related proteins). In addition, over-expression of the maize homologue of Rb inhibits WDV replication in plant cells. The tumorigenic capacity of SV40 T-antigen can be completely uncoupled from its Rep function provided the host DNA replication machinery is active (Cooper et al. (1997) Proc Natl Acad Sci USA 94:6450). The mutant T-antigen is more active in promoting viral replication since binding to Rb as well as p53 compete for its function as a "replicase" protein.

Thus, to decrease toxicity and permit increased expression of Rep proteins in transgenic plants, variants of the AC1 gene that exhibit replication competence, but lack the activities, e.g., Rb, PCNA binding, that result in induction of S-phase are selected following diversification of the AC1 gene by any one or more of the recombination or recursive recombination, e.g., shuffling, procedures described herein. For example, Figures 5A and 5B illustrate two alternative constructs for recombining parental nucleic acids encoding GV Rep proteins (**1**, **2**, and **3**) to generate a library of nucleic acids variants (**4**) encoding Rep proteins. Figures 5A and 5B differ by the placement of restriction enzyme recognition sites useful for facilitating manipulation of the substrate and recombinant sequences. Figure 6 illustrates alternative embodiments of vectors favorable for evolution of Rep proteins with reduced toxicity. Preceding or following

COR-AAD1-00002335

recombination, the AC1 (optionally including the AC2 and AC3 genes) are inserted into common region (CR) cassettes with or without a reporter, (e.g., GFP). Such variants can be selected by phenotypic evaluation of transformed plant cells or explants as described above, or by assessing DNA content of the transformed cells by any of the methods known in the art. If necessary to insure high-level replication of a such a GV vector, induction of the DNA synthesis machinery can be regulated by alternative methods such as spraying the plant with cytokinins, etc.

PLANT VIRUS VECTORS WITH ENHANCED EXPRESSION LEVELS

High expression levels of vector derived sequences are valuable in a number of commercially relevant applications. Embodiments of the invention provide for evolved viruses which express high levels of a viral protein, or alternatively engender high levels of expression of an exogenous protein inserted into the vector. Viruses offer several advantages over stable transformation procedures for expressing proteins in plants. Proteins encoded by chimeric viruses can be produced in 1-2 weeks rather than the 3-12 month time frame required for production using stable transformed plants. The expression levels of proteins encoded by viruses can exceed the levels obtained by stable transgenes. In addition viruses, are also amenable to high throughput systems that are useful for screening libraries of diverse genes such as those produced by DNA shuffling and other directed evolution technologies.

Because of the organization and limited size of viral genomes, wild-type viral or closely related promoters are typically used to drive expression. Promoters which lead to high levels of constitutive expression or to other expression patterns (e.g. tissue specific, or inducible) of interest can be evolved using the methods of the invention. Promoter sequences from various plant viruses as well as plant promoter sequences are recombined to generate a library of recombinant promoters. The recombinant promoters are cloned into a viral vector adjacent to a reporter gene, such as GFP, β -glucuronidase, or the maize anthocyanin regulatory gene, *Lc*, such that expression of the reporter is dependent on the recombinant promoter. Quantitative monitoring of the reporter gene is used to evaluate promoter strength, selectivity, and inducibility, (and any other desired property). Promoters identified in the screening can then be used to drive expression of exogenous genes other than the reporter gene, or cloned to replace wild-type, or other sub-optimal promoters of other vectors.

COR-AAD1-00002336

In one preferred embodiment, whole genome shuffling is used to recombine a virus, or a group of closely related viruses. The entire genome (less than 10 kb) is treated to induce fragmentation (for example, by DNAse treatment or by sonication) and then recombined as described above. A marker gene is then cloned between the 5' and 3' shuffled regions. The recombinant viruses can be transcribed in vitro and the transcripts used to inoculate leaves of a host plant via manual inoculation procedures as previously described. Assessment of the marker gene, by either qualitative or quantitative methods allows concurrent screening of promoter strength, replication level, and systemic spread. Concurrent screening of multiple parameters offers the unique benefit of simultaneously optimizing the variables which act in concert to determine protein production. The viral expression system is particularly suited for promoter shuffling and evaluation because there is no affect on expression due to chromosomal position effects. Thus, comparisons between transformants will provide a more accurate reflection of variation in promoter activity.

The invention also provides for plant virus vectors for the production of proteins in planta. Such vectors are useful for the expression and screening of recombinant, including, e.g., shuffled, libraries as well as for large scale production of shuffled and other proteins in plants. Some viruses such as the potato virus X (PVX) and tobacco mosaic virus (TMV) can accept large genes and still package virions effectively. Other viruses can be employed by replacing genes non-essential for replication with exogenous sequences, including library sequences.

The present invention offers a number of advantages for the screening and evaluation of recombinant, e.g., shuffled, libraries. For example, the evolved plant viruses of the invention can be introduced by electroporation, or other equivalent methods, into plant protoplasts where they can replicate autonomously and accumulate to high levels. Chromosomal insertion is unnecessary and evolved protein function can be assayed within hours of introduction.

The choice of vector can be made to optimize production of protein. For example the cauliflower mosaic virus (CaMV) 35S promoter is a well characterized constitutive promoter, which in some cases can yield as much as 2% of total soluble protein recovered from a plant cell. The CaMV 35S promoter or the evolved promoters of the invention can be used to generate high levels of protein expression. Production of some proteins (for example, sparingly soluble proteins) is simplified by cloning the exogenous gene, or gene library (including a shuffled library) as a fusion with the viral

COR-AAD1-00002337

coat protein. Fusion proteins can then be simply and effectively recovered from virus particle preparations. Such coat protein fusions are particularly useful in the production of phage display vectors. Antigenic epitopes are fused to the coat protein such that they are displayed on the outside of the virus particle. This application is especially useful for vaccine production.

PROTEIN PRODUCTION IN TRANSGENIC PLANTS

A potential drawback of using viral vectors to produce proteins in plants is the possibility that symptomatic infection can result, and that this infection could spread from the engineered plants to nearby crop plants. A preferred embodiment of the invention provides for viral vectors which are capable of propagation only within a transgenic host, reducing the risk of escape into the environment. This technology allows the use of DNA shuffling to evolve wider host ranges for virus vectors without the risk of evolving new plant pathogens for those hosts.

The present invention supplies a viral coat protein in trans by expression of the coat protein in a transgenic plant. The corresponding viral promoter drives expression of an exogenous sequence to be expressed for protein production. For purposes of illustration, development of a tobacco mosaic virus (TMV) vector which is operable in transgenic cruciferae is described. However, the methods can be applied equally to viruses which infect other plant species, and to transgenic plants of such species which serve as hosts.

At least 4 strains of TMV are known to infect cruciferous crops and are able to replicate in Arabidopsis. The sequences of these strains are available in Genbank, and accordingly, polynucleotides corresponding to their sequences can be isolated and cloned by techniques well known in the art. Similarly, DNA sequences of strains which infect soybean are available in Genbank and can be used to develop viral vectors which infect soybeans and other legumes. Several strains that infect odontoglossum and ribgrass are known, and can be used to evolve vectors for monocot crops such as corn, rice and wheat.

Several publications suggest that coat proteins and movement proteins are the major determinants of host range, (*see*, e.g., Spitsin et al. (1999) Proc. Nat'l. Acad. Sci. USA 96:2549-2553; Hilf and Dawson (1993) Virology 193:106-114.) Cloning sites are introduced into an infectious TMV clone, allowing the introduction of heterologous movement proteins and coat proteins into the vector backbone. This provides the basis of

COR-AAD1-00002338

evolving tobamovirus vectors for any host. Libraries of oligonucleotides of crucifer infecting tobamovirus movement and coat proteins are constructed and cloned into the TMV backbone. Arabidopsis plants are infected with the library and variants capable of replication and long distance movement are harvested from the distal leaves. Several passages of infection and recovery can be used to select for virus fitness, and several independent isolates developed and sequenced.

After developing a virus capable of forming high titers on Arabidopsis, the coat protein is excised and cloned into a vector for the production of transgenic plants. The recombinant, e.g., coat protein gene, under control of a constitutive promoter, is then introduced (e.g., by agrobacterium mediated transformation using standard vacuum infiltration techniques) into Arabidopsis plants (or plant cells). Transgenic plants are then propagated by techniques well established in the art, and described above.

The site formerly occupied by the excised coat protein, and retaining the coat protein promoter then provides the cloning site for exogenous sequences to be expressed in the transgenic plants. For example, sequences corresponding to a visual marker, such as green fluorescent protein (GFP), can be cloned adjacent to the coat protein promoter and the virus vector used to infect the transgenic Arabidopsis. Visual assessment under appropriate light conditions enables selection of recombinant GFP with improved or increased fluorescent properties. Similarly, the evolved vector can be used for screening and production of protein pharmaceuticals, nutraceuticals, biosynthetic enzymes, and biodegradable polymers and polymer components among others.

Further cycles of recombination and selection are then used to evolve vectors which infect crop plants of interest. The present invention provides for the production of protein and other biosynthetic products of interest in plant species which have previously posed significant risks of infection. The virus vector of the invention is unable to spread systemically and incapable of infecting wild-type plants in experimental or field settings. Only when used to infect transgenic plants can the virus vector replicate, be encapsidated and spread systemically, as is required for efficient protein production. In addition to the crucifer vector described above, preferred embodiments relate to virus vectors and to methods for their development, which are suitable for protein production in legumes, such as soybeans, and graminae, such as corn. Proteins to be expressed include such valuable products as biofuels; biodegradable polymers (including plastics) and their components; industrial enzymes; enzymes for improved animal and human nutrition; nutraceuticals; protein pharmaceuticals; antigens; plantibodies; and biosynthetic enzymes

COR-AAD1-00002339

(e.g., of lipid, or carbohydrate synthetic pathways). The invention also provides for transgenic plants used for the expression of such evolved viruses.

PATHOGEN-DERIVED RESISTANCE TO PLANT PATHOGENS

Plant viruses can cause considerable damage to commercially valuable crops. The present invention provides for evolved plant viruses and viral genes that confer pathogen-derived resistance to plant viruses, as well as for a method of evolving viruses or viral genes to confer pathogen-derived resistance to plant viruses. Pathogen-derived resistance refers to the observation that overexpression of a pathogen derived component, in many cases a viral coat protein, can protect a transgenic plant which expresses it from infection by that virus. However, resistance is not always complete and protection is most frequently limited to a single virus from which the coat protein originated.

It is likely that resistance occurs because the high concentration of coat proteins prevents the origin of replication from being exposed long enough for the virus to replicate in the plant cell cytoplasm. However, proof of this mechanism is not required for the effective application of the methods or for the deployment of the evolved viruses of the invention. Recursive recomibnation, e.g., DNA or RNA shuffling, is applied to a single viral coat protein, or to families of several coat proteins, and assayed for increased affinity for RNA or DNA, or for processivity in assembly. For example, recombinant viral proteins can be recovered and pooled for evaluation by binding to RNA or DNA-conjugated columns, on matrices such as agarose or sepharose. Washing conditions are determined empirically for each input coat protein to select only recombinant coat proteins which bind the viral nucleic acid with greater affinity. Specificity for the viral origin of replication can be assessed by using synthetic oligonucleotides corresponding to the viral origin as the column supported substrate. By evaluating binding to oligonucleotides corresponding to the origin of replication of various viruses, the spectrum of viruses to which resistance is provided can be expanded. Details pertaining to employment of the above techniques, are available in the previously cited references (e.g., Ausubel, Berger, and Sambrook) and are known to those of skill in the art. Similarly, other well known techniques for the binding of proteins to nucleic acids can be used to evaluate evolved coat proteins, and are available to persons skilled in the art.

After pre-screening based on molecular characteristics, the ability to confer resistance in planta is assayed. Recombinant, e.g., shuffled, coat protein encoding

COR-AAD1-00002340

sequences with favorable binding (or other characteristics) identified in the previous screening step, are introduced into transgenic plants by any of the techniques routinely used to generate stable transformants, e.g., microinjection, electroporation, biolistics, Agrobacterium mediated transformation. Stable integrants are selected and regenerated to constitute whole plants which express the recombinant coat protein. Such transgenic plants are subjected to infection by wild-type virus. For example, to evolve viruses which protect important crucifer crops, (e.g., cabbage, cauliflower, broccoli, mustards and rapes, including forage and oilseed rape), the model plant *Arabidopsis thaliana* is used. Many viruses that infect Arabidopsis also infect other crucifers, and their sequences and corresponding polynucleotides are readily available, *see* e.g., Aguilar et al. (1996) "Nucleotide sequence of chinese rape mosaic virus (oilseed rape mosaic virus), a crucifer tobamovirus infectious on *Arabidopsis thaliana*. Plant Mol. Bio. 30:191-197. After one or more rounds of diversification and screening in Arabidopsis has yielded recombinant viral sequences that confer resistance, candidate clones are tested in specific crop species to confirm protection.

While the foregoing invention has been described in some detail for puposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, vectors with various desirable properties, not explicitly described, which are useful for specific applications can readily be developed by the methods of the present invention. All publications, patents, patent applications, or other documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, or other document were individually indicated to be incorporated by reference for all purposes.

COR-AAD1-00002341

WHAT IS CLAIMED IS:

1. A method of evolving a plant vector, the method comprising:

(a) providing a population of nucleic acid fragments comprising at least one component of a plant vector;

(b) recursively recombining the population of nucleic acid fragments to generate a library of recombinant polynucleotides;

(c) optionally repeating the recombination process of steps (a) and (b) one or more times;

(d) screening the library of recombinant polynucleotides to identify at least one recombinant polynucleotide, wherein the recombinant polynucleotide is a component of an evolved vector which has acquired or evolved a desired property;

(e) repeating steps (a) through (d) until the evolved plant vector has acquired the desired property.

2. The method of claim 1, wherein the nucleic acid fragments provided in step (a) are produced by one or more of: error prone PCR, mutagenic amplification, and site directed mutagenesis.

3. The method of claim 1, comprising providing a population of DNA fragments.

4. The method of claim 1, comprising providing a population of RNA fragments.

5. A library of recombinant polynucleotides produced by the method of claim 1 or 2.

6. The library of claim 5, wherein the library is introduced into an Agrobacterium.

7. The library of claim 5, wherein the library is introduced into a plant virus.

8. An evolved vector produced by the method of claim 1 or 2.

COR-AAD1-00002342

9. A transgenic plant or a transgenic plant cell or a transgenic plant explant made by transforming a parental plant cell with the evolved vector of claim 1 or 2.

10. The transgenic plant cell or transgenic plant explant of claim 9, wherein the plant cell or plant explant is a stem, a leaf, a root, a meristem, a callus, or an embryo.

11. A library of recombinant polynucleotides, wherein the recombinant polynucleotides comprise at least one sequence derived from a genome of the transgenic plant, or the transgenic plant cell, or the transgenic plant explant of claim 9.

12. The method of claim 1, wherein the vector comprises one or more of: a plasmid, a virus, a bacterium, an agrobacterium, a RNA polynucleotide, a naked DNA polynucleotide, a poly-lysine-conjugated DNA, a peptide-conjugated DNA and a liposome-conjugated DNA.

13. The method of claim 12, wherein the vector is an agrobacterium vector, the population of nucleic acid fragments provided in step (a) comprising at least one agrobacterium gene, wherein the recombinant polynucleotide identified in step (d) is a component of an evolved agrobacterium strain which has acquired or evolved a desired property, step (e) comprising repeating steps (a) through (d) until the evolved agrobacterium strain has acquired the desired property.

14. The method of claim 13, wherein the evolved agrobacterium vector comprises a binary vector system having a helper plasmid, which helper plasmid comprises *vir* function.

15. The method of claim 13, wherein a DNA sequence comprising *vir* function is integrated into a chromosome of an agrobacterium.

16. The method of claim 13, wherein the evolved agrobacterium strain is *A. tumefaciens*.

17. The method of claim 13, wherein the evolved agrobacterium strain is *A. rhizogenes*.

COR-AAD1-00002343

18. The method of claim 13, wherein the nucleic acid fragments provided in step (a) include at least one agrobacterium right T-DNA border.

19. The method of claim 18, wherein the nucleic acid fragments provided in step (a) include at least one agrobacterium right T-DNA border and at least one agrobacterium left T-DNA border.

20. The method of claim 19, wherein the at least one right agrobacterium T-DNA border and the at least one left agrobacterium T-DNA border provided in step (a), each further comprise a PCR primer binding site.

21. The method of claim 13, wherein the fragments provided in step (a) comprise at least one of: a *virA* gene, a *virB* gene, a *virC* gene, a *virD* gene, a *virE* gene, a *virG* gene or a *chvE* gene.

22. The method of claim 13, wherein the nucleic acid fragments provided in step (a) comprise at least one agrobacterium genome.

23. The method of claim 13, wherein the desired property comprises one or more of: insert precision, targeted insertion, improved host range, transformation efficiency, in planta transformation of leaves, in planta transformation of cut stems, in planta transformation in the absence of exogenous phytohormones, transformation without in vitro culture, and chloroplast targeting.

24. The method of claim 23, wherein the desired property is insert precision.

25. The method of claim 23, wherein the nucleic acid fragments provided in step (a) include at least one agrobacterium right T-DNA border.

26. The method of claim 25, wherein the screening of step (d) is performed by one or more of polymerase chain reaction, hybridization, ligase chain reaction, and proteomics.

27. The method of claim 26, wherein the screening of step (d) is performed by an inverse polymerase chain reaction of at least one junction between an integrated T-DNA and its site of insertion.

COR-AAD1-00002344

28. The method of claim 23, wherein the desired property is targeted insertion.

29. The method of claim 28, wherein the nucleic acid fragments provided in step (a) include at least one Agrobacterium right T-DNA border.

30. The method of claim 29, wherein the screening of step (d) is performed by one or more of polymerase chain reaction, hybridization, ligase chain reaction and proteomics.

31. The method of claim 30, wherein the screening of step (d) is performed by an inverse polymerase chain reaction of at least one junction between an integrated T-DNA and its site of insertion.

32. The method of claim 31, wherein the nucleic acid fragments provided in step (a) further comprise a region of sequence similarity with a desired insertion site, which region of sequence similarity is attached adjacent and contiguous to the at least one agrobacterium T-DNA border.

33. The method of claim 32, wherein a promoter or an enhancer is included in the nucleic acid fragments provided in step (a), which promoter or enhancer is located adjacent to the region of sequence similarity.

34. The method of claim 33, wherein the screening of step (d) further comprises: transforming the library of recombinant polynucleotides into a host plant or a host plant cell, which host plant or host plant cell has a detectable marker at a desired insertion site, whereby insertion of a library member at the desired insertion site results in activation of the detectable marker.

35. The method of claim 34, wherein the host plant or host plant cell in step (d) is a transgenic plant or cultured transgenic plant explant having a marker transgene at the desired insertion site.

36. The method of claim 35, wherein the marker transgene of step (d) is green fluorescent protein (GFP), β-glucuronidase, luciferase or a maize *Lc* (anthocyanin regulatory) gene.

COR-AAD1-00002345

37. The method of claim 35, wherein the marker transgene of step (d) is a selectable marker.

38. The method of claim 37, wherein the marker transgene of step (d) confers resistance to an antibiotic or a herbicide.

39. The method of claim 37, wherein the marker transgene of step (d) is a negatively selectable marker.

40. The method of claim 39, wherein the marker transgene is a homologue of a *dhl1* gene, a *codA* gene, a *tms2* gene, or a *NIA2* gene.

41. The method of claim 40, wherein the marker transgene is a non-functional homologue of a *dhl1* gene, a *codA* gene, a *tms2* gene, or a *NIA2* gene.

42. The method of claim 23, wherein the desired property is the ability to transform a broad range of host plant species.

43. The method of claim 42, wherein the nucleic acid fragments provided in step (a) comprise at least one gene selected from among: *virA, virB, virC, virD, virE, virG* and *chvE.*

44. The method of claim 43, wherein the nucleic acid fragments provided in step (a) comprise at least one *virA* gene, and wherein the screening of step (d) is performed by detecting expression of a marker operably linked to a *vir* promoter.

45. The method of claim 44, wherein the *vir* promoter is bound by a VirG protein, which VirG protein is activated by a recombinant VirA protein.

46. The method of claim 42, wherein the screening of step (d) is performed in an intact plant or in a plant culture of one or more of a stem, a leaf, a root, a protoplast, a callus, a meristem, and an embryo.

47. The method of claim 42, wherein the host plant species comprises a dicot or a monocot.

48. The method of claim 23, wherein the desired property is the ability of the Agrobacterium to target a T-DNA to the chloroplast.

COR-AAD1-00002346

49. The method of claim 48, wherein the nucleic acid fragments provided in step (a) comprise at least one of a *virD2* DNA sequence and a *virE2* DNA sequence.

50. The method of claim 48, wherein the nucleic acid fragments provided in step (a) comprise at least one of a chloroplast localization signal and a nuclear localization signal.

51. The method of claim 48, wherein the screening of step (d) is performed by screening for localization to a chloroplast.

52. A library of recombinant polynucleotides produced by the method of claim 13.

53. The library of claim 52, wherein the library is introduced into an Agrobacterium.

54. An evolved agrobacterium strain produced by the method of claim 13.

55. A transgenic plant or a transgenic plant cell or a transgenic plant explant made by transforming a parental plant cell with the evolved agrobacterium strain of claim 54.

56. The transgenic plant cell or transgenic plant explant of claim 55, wherein the plant cell or plant explant is a stem, a leaf, a root, a meristem, a callus, or an embryo.

57. A library of recombinant polynucleotides, wherein the recombinant polynucleotides comprise at least one sequence derived from a genome of the transgenic plant, or the transgenic plant cell, or the transgenic plant explant of claim 55.

58. The method of claim 12, wherein the vector comprises a virus, the population of nucleic acid fragments provided in (a) comprising at least one component of a plant virus, wherein the recombinant polynucleotide identified in the screening of step (d) is a component of an evolved plant virus, which evolved plant virus has acquired or evolved a desired property, step (e) comprising repeating steps (a) through (d) until the evolved plant virus has acquired the desired property.

COR-AAD1-00002347

59. The method of claim 58, wherein the plant virus is selected from the group: a RNA virus, a DNA virus, an alfamovirus, a bromovirus, a capillovirus, a carlavirus, a carmovirus, a caulimovirus, a closterovirus, a comovirus, a cryptovirus, a cucumovirus, a dianthovirus, a fabavirus, a fijivirus, a furovirus, a geminivirus, a hordeivirus, a ilarvirus, a luteovirus, a machlovirus, a maize chlorotic dwarf virus, a marafivirus, a necrovirus, a nepovirus, a parsnip yellow fleck virus, a pea enation mosaic virus, a potexvirus, a potyvirus, a reovirus, a rhabdovirus, a sobemovirus, a tenuivirus, a tobamovirus, a tobravirus, a tomato spotted wilt virus, a tombusvirus, and a tymovirus

60. The method of claim 58, wherein the desired property is systemic spread of an evolved virus incorporating an insert, which insert is larger than an insert spread by a wild-type virus.

61. The method of claim 60, wherein the nucleic acid fragments provided in step (a) encode at least one movement protein.

62. The method of claim 61, wherein the nucleic acid fragments provided in step (a) comprise at least one viral genome.

63. The method of 58, wherein the desired property is rapid systemic spread of an evolved plant virus following infection.

64. The method of claim 60, wherein the library of recombinant polynucleotides is cloned into a virus vector, which recombinant virus vector is inoculated onto at least one leaf of an intact plant followed by harvest and recovery of a plurality of virus particles from leaves distal to the at least one leaf whereon lies the site of inoculation

65. The method of claim 60, wherein the nucleic acid fragments provided in step (a) encode at least one movement protein.

66. The method of claim 60, wherein the nucleic acid fragments provided in step (a) comprise at least one viral genome.

67. The method of claim 58, wherein the desired property is reduction of viral symptoms after infection.

COR-AAD1-00002348

68. The method of claim 67, wherein the nucleic acid fragments provided in step (a) comprise at least one viral protein.

69. The method of claim 68, wherein the nucleic acid fragments provided in step (a) encode at least one of a viral coat protein, a viral replicase, a viral recombinase and a movement protein.

70. The method of claim 67, wherein the nucleic acid fragments provided in step (a) comprise at least one viral genome.

71. The method of claim 58, wherein the desired property is reduced toxicity following introduction of the virus.

72. The method of claim 71, wherein the nucleic acid fragments provided in (a) encode at least one viral protein.

73. The method of claim 72, wherein the nucleic acid fragments encode at least one of a viral coat protein, a viral replicase a viral recombinase and a viral movement protein.

74. The method of claim 71, wherein the nucleic acid fragments provided in (a) comprise at least one viral genome.

75. The method of claim 58, wherein the desired property is a higher level of viral gene expression as compared to the level of expression generated by a wild-type virus.

76. The method of claim 75, wherein the nucleic acid fragments provided in step (a) comprise at least one viral or plant promoter.

77. The method of claim 76, wherein the nucleic acid fragments provided in step (a) comprise a cauliflower mosaic virus 35S promoter.

78. The method of claim 77, wherein the nucleic acid fragments provided in step (a) comprise at least one viral genome.

COR-AAD1-00002349

79. The method of claim 58, wherein the desired property is a higher constitutive level of exogenous gene expression than is expressed under control of a wild-type promoter provided in step (a).

80. The method of claim 79, wherein the nucleic acid fragments provided in step (a) comprise at least one viral or plant promoter.

81. The method of claim 80, wherein the nucleic acid fragments provided in step (a) comprise a cauliflower mosaic virus 35S promoter.

82. The method of claim 79, wherein the nucleic acid fragments provided in step (a) comprise at least one viral genome.

83. The method of claim 82, wherein the screening of step (d) comprises the steps of cloning a marker gene between a 5' shuffled region and a 3' shuffled region of the at least one viral genome to generate a recombinant virus, and inoculating the recombinant virus onto a leaf of a host plant.

84. The method of claim 58, wherein the screening step of (d) is performed in a transgenic plant constitutively expressing a viral coat protein, and wherein the evolved vector comprises a polycloning site, which polycloning site replaces a viral coat protein structural gene and is contiguous with a corresponding viral coat protein promoter, such that an exogenous DNA inserted into the polycloning site is expressed under control of the viral coat protein promoter.

85. A library of recombinant polynucleotides produced by the method of claim 58.

86. The library of claim 85, wherein the library is introduced into a plant virus.

87. An evolved plant virus produced by the method of claim 58.

88. The evolved plant virus of claim 87, wherein a polynucleotide is inserted into the evolved plant virus, which polynucleotide comprises a product of a gene library.

COR-AAD1-00002350

89. The evolved plant virus of claim 88, wherein the gene library is produced by recursive recombination.

90. The evolved plant virus of claim 89, wherein the gene library is produced by DNA or RNA shuffling.

91. The evolved plant virus of claim 89, wherein the polynucleotide is inserted into a viral coat protein gene resulting in translation of a fusion protein, which fusion protein comprises a polypeptide encoded by the polynucleotide fused to a viral coat protein.

92. The evolved plant virus of claim 91, wherein the polypeptide encoded by the polynucleotide is displayed on an outside surface of a virus particle.

93. The evolved plant virus of claim 89, wherein the polynucleotide inserted into the evolved plant virus is a plant or viral promoter, which promoter is operably linked to a reporter gene.

94. The evolved plant virus of claim 93, wherein the reporter gene is green fluorescent protein (GFP), luciferase, β-glucuronidase or a maize *Lc* (anthocyanin regulatory) gene.

95. The evolved plant virus of claim 93, wherein the reporter gene of step (d) is a selectable marker.

96. The evolved plant virus 95, wherein the reporter gene of step (d) confers resistance to an antibiotic or a herbicide.

97. The evolved plant virus of claim 95, wherein the reporter gene of step (d) is a negatively selectable marker.

98. The evolved plant virus of claim 97, wherein the reporter gene is a homologue of a *dhl1* gene, a *codA* gene, a *tms2* gene, or a *NIA2* gene.

99. The evolved plant virus of claim 98, wherein the reporter gene is a non-functional homologue of a *dhl1* gene, a *codA* gene, a *tms2* gene, or a *NIA2* gene.

COR-AAD1-00002351

100. The evolved plant virus of claim 89, wherein the polynucleotide is inserted contiguous to a cauliflower mosaic virus 35S promoter resulting in expression of a protein encoded by the polynucleotide under the control of the cauliflower mosaic virus 35S promoter.

101. The evolved plant virus of claim 89, wherein a viral genome of the evolved plant virus is positioned adjacent to at least a right agrobacterium T-DNA border.

102. A method of using the evolved plant virus of claim 101, comprising introducing the evolved plant virus into a host plant cell by Agroinfection.

103. A method of using the evolved plant virus of claim 89. Wherein the vector is introduced into plant protoplasts by one or more of electroporation, microinjection, biolistics, or direct mechanical inoculation.

104. A transgenic plant or a transgenic plant cell or a transgenic plant explant made by transforming a host plant or host plant cell or host plant explant with the evolved plant virus of claim 87.

105. The transgenic plant cell or transgenic plant explant of claim 104, wherein the plant cell or plant explant is a stem, a leaf, a root, a meristem, a callus, or an embryo.

106. A library of recombinant polynucleotides, wherein the polynucleotides comprise at least one sequence derived from a genome of the transgenic plant, or the transgenic plant cell, or the transgenic plant explant of claim 104.

107. A recombinant plant virus, which recombinant plant virus comprises a polycloning site, which polycloning site replaces a viral coat protein structural gene and is contiguous with a corresponding viral coat protein promoter, such that an exogenous DNA inserted into the polycloning site is expressed under control of the viral coat protein promoter.

108. The recombinant plant virus of claim 107, wherein the viral vector is a tobamovirus.

COR-AAD1-00002352

109. The recombinant plant virus of claim 108, wherein the desired property is the ability to infect a transgenic *Brassica.*

110. The recombinant plant virus of claim 109, wherein the *Brassica* is *Arabidopsis thaliana.*

111. The recombinant plant virus of claim 107, wherein the desired property is the ability to infect a transgenic plant, which transgenic plant is selected from the family *Leguminosae.*

112. The recombinant plant virus of claim 107, wherein the desired property is the ability to infect a transgenic plant, which transgenic plant is a monocot.

113. The recombinant plant virus of claim 112, wherein the transgenic plant is a crop plant selected from among the genera: *Eleusine, Lollium, Bambusa, Dactylis, Sorghum, Pennisetum, Zea, Oryza, Triticum, Secale, Avena, Hordeum, Saccharum, Coix.*

114. The recombinant plant virus of claim 107, further comprising an exogenous DNA, which exogenous DNA is inserted into the polycloning site.

115. The recombinant plant virus of claim 114 wherein the exogenous DNA encodes a polypeptide.

116. The recombinant plant virus of claim 115, wherein the polypeptide is a protein, which protein is a biofuel, an industrial enzyme, a nutritional enzyme, a nutraceutical, a protein pharmaceutical, a plantibody, an antigen, or a biosynthetic enzyme.

117. The recombinant plant virus of claim 116, wherein the protein is a biosynthetic enzyme, which biosynthetic enzyme comprises a component of a biochemical pathway producing at least one of an oil, a starch, a biodegradable plastic, a component of a biodegradable plastic, a biopolymer, and a component of a biopolymer.

118. A method of producing a polypeptide or a biosynthetic product, the method comprising:

COR-AAD1-00002353

transforming a transgenic plant, or plant cell or plant explant, which transgenic plant or plant cell or plant explant expresses a viral coat protein, with the recombinant plant virus of claim 107.

119. A method of evolving a plant virus or viral component to confer pathogen-derived resistance to plant viruses, the method comprising:

(a) providing a population of plant virus nucleic acid fragments;

(b) recursively recombining the population of nucleic acid fragments to generate a library of recombinant polynucleotides;

(c) optionally repeating the recombination process of steps (a) and (b) one or more times;

(d) screening the library to identify at least one recombinant polynucleotide, wherein the at least one recombinant polynucleotide comprises one or more evolved viral components, which one or more evolved viral components have acquired or evolved the property of conferring resistance to infection by at least a second plant virus when the one or more evolved viral components are stably introduced and expressed in a host plant, wherein the at least second plant virus is the same or different from the evolved plant virus or the one or more evolved viral components;

(e) repeating steps (a) through (d) until the one or more evolved viral components have acquired the ability to confer pathogen-derived resistance to plant viruses.

120. An evolved plant virus or viral component produced by the method of claim 119.

121. A method for producing a gene with a desired property, the method comprising:

(a) introducing a plurality of RNA viral vectors comprising one or more gene of interest into at least one cell;

(b) growing the cell under conditions permitting cytoplasmic recombination between the plurality of RNA viral vectors, thereby producing a library of recombinant RNA viral vectors;

(c) optionally recovering at least one recombinant viral vector and repeating steps (a) and (b);

(d) identifying at least one RNA viral vector comprising a gene with a desired property.

COR-AAD1-00002354

122. The method of claim 121, comprising introducing the plurality of RNA viral vectors by inoculating at least one cell with infectious viral transcripts.

123. The method of claim 121, comprising introducing the plurality of RNA viral vectors by introducing a plurality of cDNA molecules corresponding to viral transcripts.

124. The method of claim 123, wherein viral transcripts comprising the plurality of cDNA molecules are produced in the cytoplasm of the at least one cell.

125. The method of claim 123, wherein the plurality of cDNA molecules are introduced by electroporation, microinjection, biolistics, agrobacterium mediated transformation or agroinfection.

126. The method of claim 121, wherein the RNA viral vector comprises a plant viral vector.

127. The method of claim 126, wherein the RNA viral vector is selected from among a tobamovirus, a potyvirus, a tobravirus and a potexvirus.

128. The method of claim 126, wherein the RNA viral vector comprises a Tobacco Mosaic Virus (TMV), a TMV homolog or an engineered viral vector derived from a TMV or TMV homolog.

129. The method of claim 121, wherein the gene of interest comprises a protein coding sequence.

130. The method of claim 121, wherein the at least one cell comprises a plant cell.

131. The method of claim 130, wherein the plant cell comprises an isolated plant cell, a protoplast, a plant explant, a plant tissue or an intact plant.

132. The method of claim 130, comprising growing the plant cell in suspension culture.

133. The method of claim 130, comprising growing at least one intact plant comprising the plant cell.

COR-AAD1-00002355

134. The method of claim 121, wherein the cytoplasmic recombination is mediated by template switching of an RNA polymerase expressed by the at least one cell.

135. The method of claim 134, wherein the RNA polymerase is a plant viral RNA polymerase.

136. The method of claim 134, wherein the RNA polymerase is a mutant or engineered viral RNA polymerase that enhances the frequency of homologous or non-homologous RNA recombination relative to a wild-type plant viral RNA polymerase.

137. The method of claim 136, wherein the mutant or engineered viral RNA polymerase is produced by a directed evolution process.

138. The method of claim 137, wherein the directed evolution process comprises DNA or RNA shuffling.

139. The method of claim 121, comprising recovering at least one recombinant viral vector by isolating RNA from the at least one cell.

140. The method of claim 121, comprising identifying the at least one RNA viral vector comprising a gene with a desired property by selection or screening.

141. The method of claim 121, comprising introducing at least a first RNA viral vector incapable of systemic infection in a plant and a second RNA viral vector incapable of systemic infection in a plant, which first and second viral vectors have complementary mutations in genes essential for systemic infection, and identifying at least one recombinant RNA viral vector by selecting or screening for RNA viral vectors capable of systemic infection.

142. The method of claim 141, wherein the genes having complementary mutations comprise one or more of a gene encoding a viral movement protein or a gene encoding a viral coat protein.

143. The method of claim 141, wherein selecting or screening is performed by sampling a plant cell or tissue remote from the site of introduction.

COR-AAD1-00002356

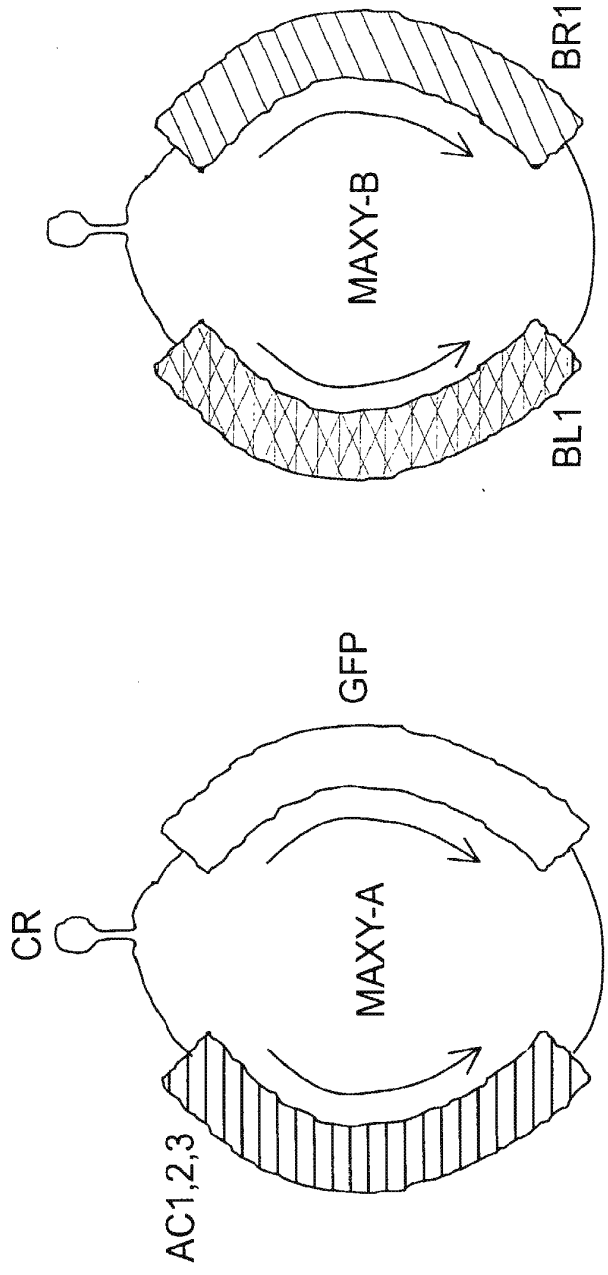


Fig. 1

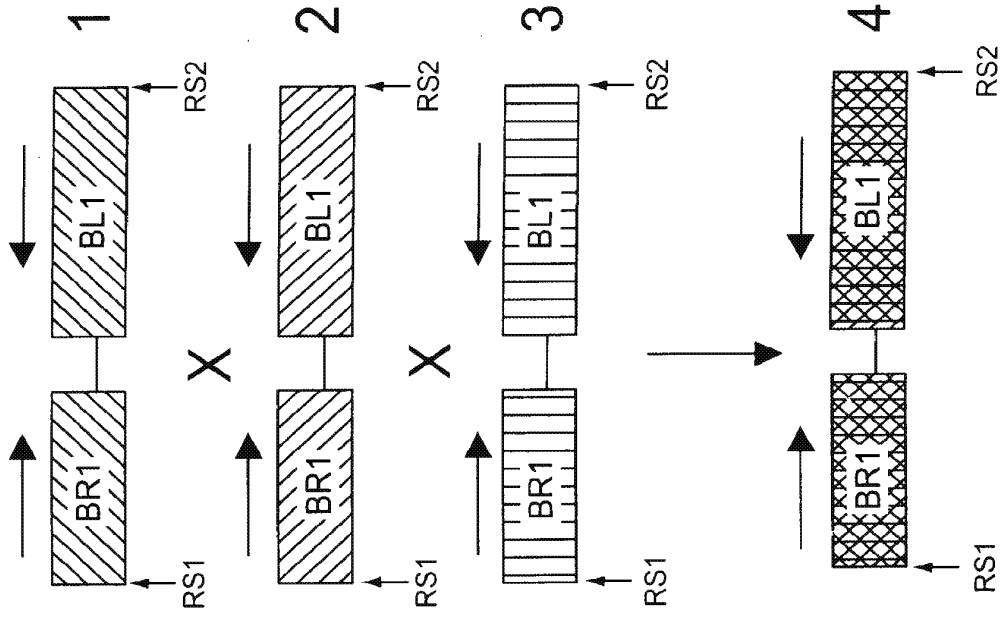


Fig. 2

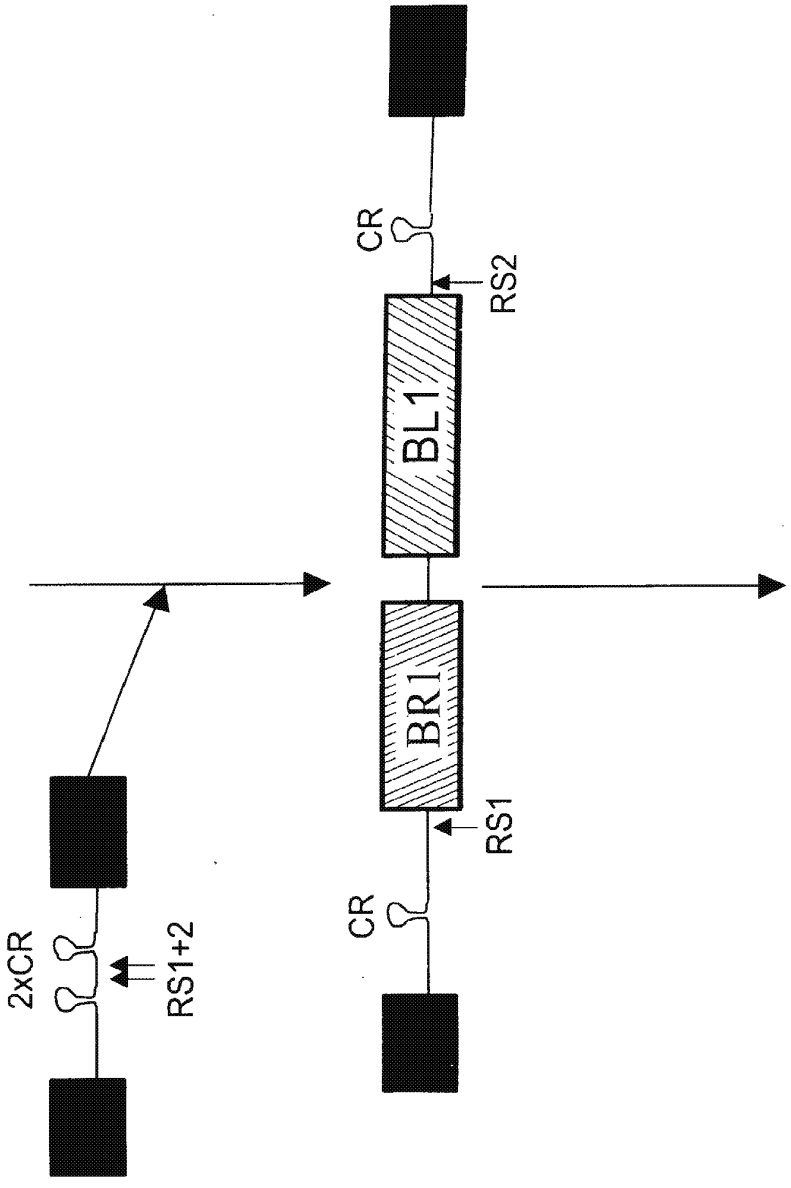


Fig. 3

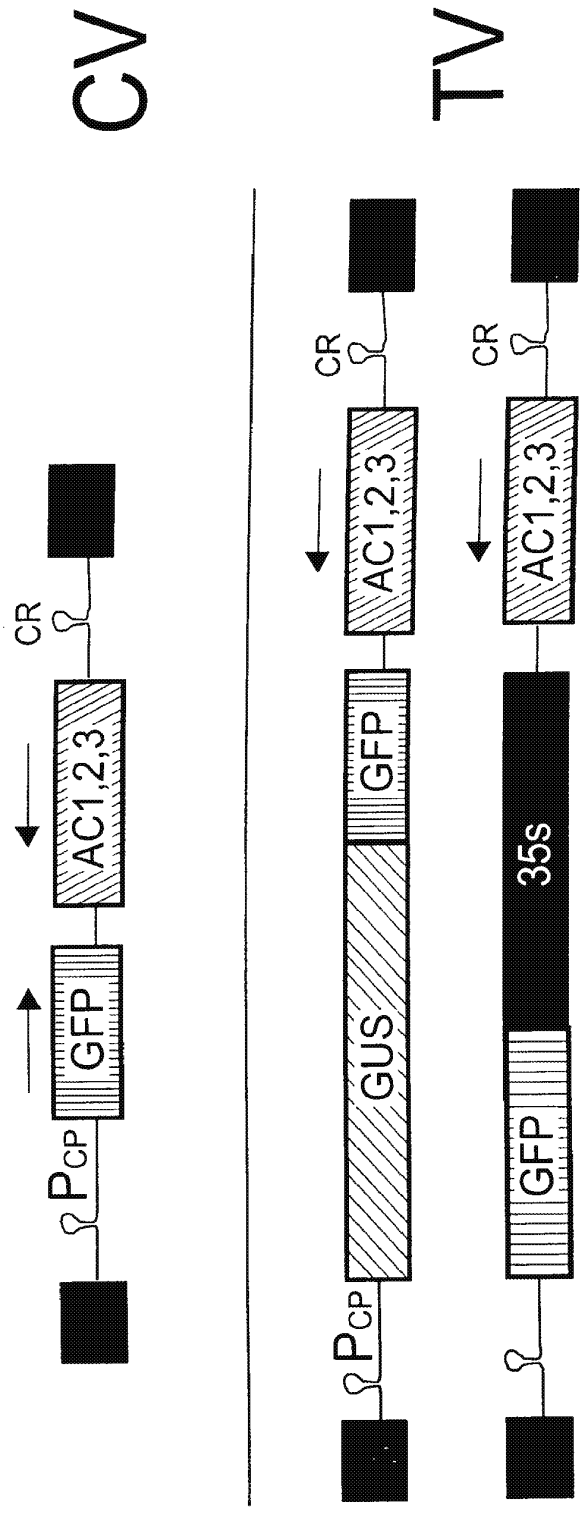


Fig. 4

Fig. 5A

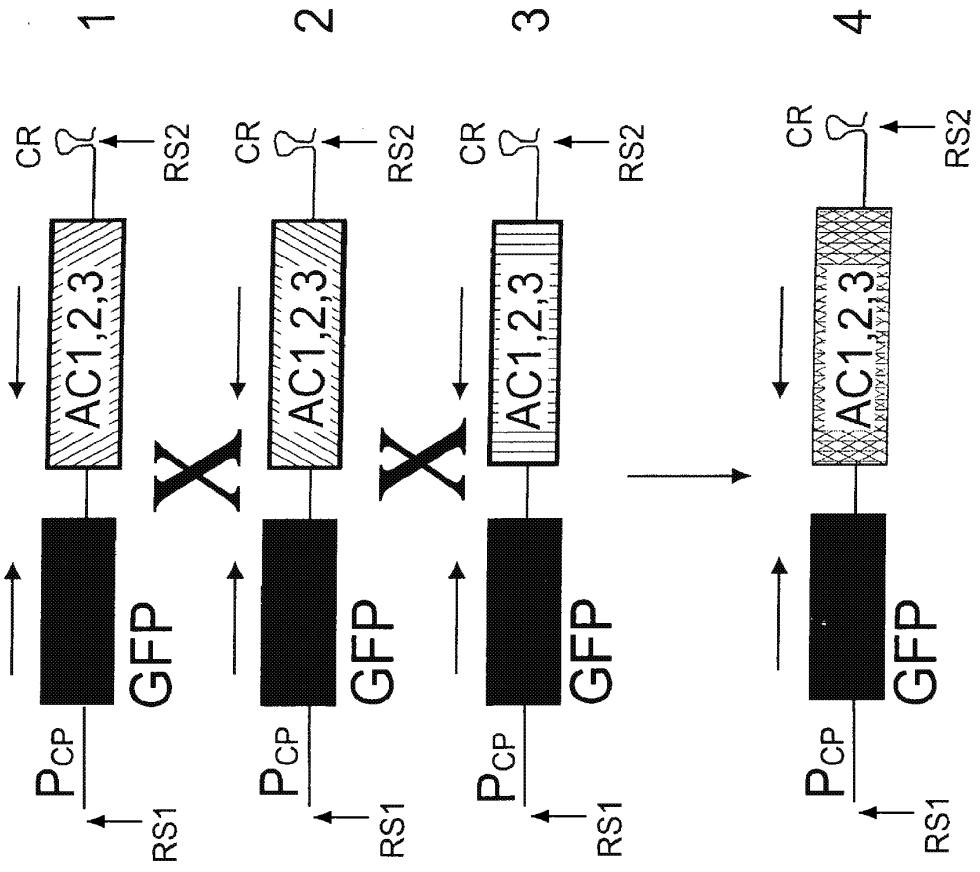


Fig. 5B

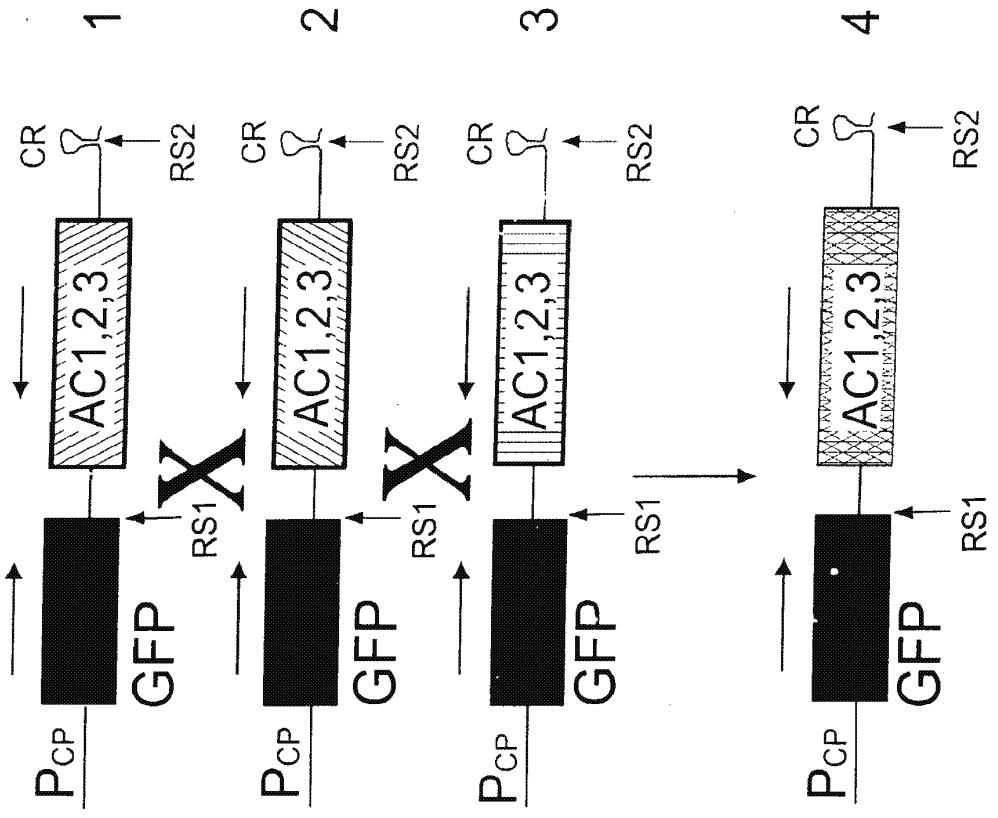
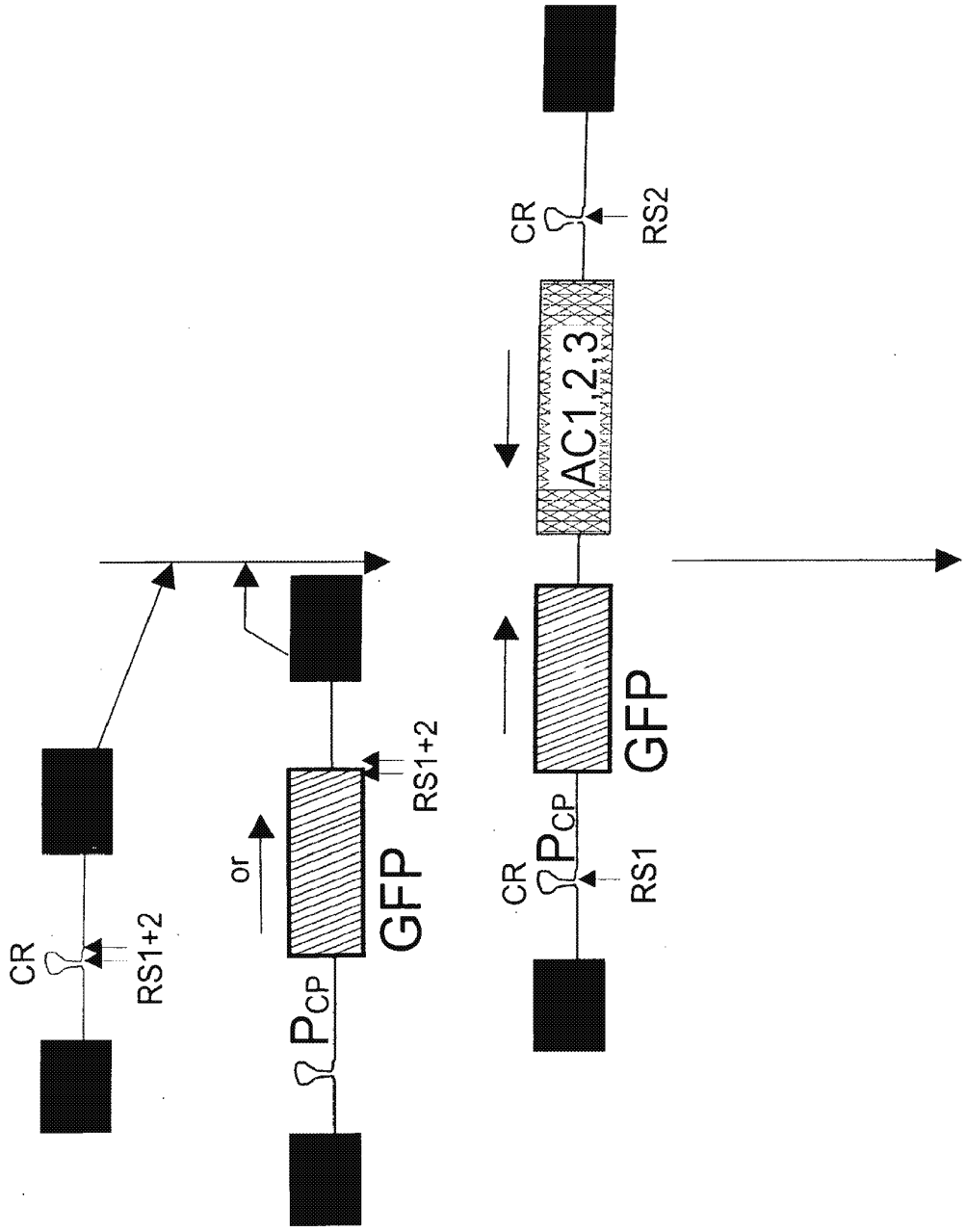


Fig. 6



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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 July 2003 (17.07.2003)

PCT

(10) International Publication Number
WO 03/056904 A2

(51) International Patent Classification[7]: A01H

(21) International Application Number: PCT/US03/00432

(22) International Filing Date: 7 January 2003 (07.01.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/346,541 8 January 2002 (08.01.2002) US

(71) Applicant and
(72) Inventor: **RAAB, Michael, R.** [US/US]; 1 Fifth Street, Cambridge, MA 02141 (US).

(74) Agents: **VOLPE, Anthony, S.** et al.; Volpe and Koenig, P.C., Suite 400, One Penn Center, 1617 John F. Kennedy Boulevard, Philadelphia, PA 19103 (US).

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

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

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

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

(54) Title: TRANSGENIC PLANTS EXPRESSING CIVPS OR INTEIN MODIFIED PROTEINS AND RELATED METHOD

CO_2H -	Native Protein Sequence		Intein Protein Sequence	$-NH_2$
CO_2H -	Intein Protein Sequence		Native Protein Sequence	$-NH_2$
CO_2H -	Carboxy Terminal Native Protein Seq.	Intein Protein Sequence	Amino Terminal Native Protein Seq.	$-NH_2$

(57) Abstract: Transgenic plants that express CIVPS or intein modified proteins, compositions of matter comprising them, products of diverse applications made from the transgenic plants, methods to construct the transgenic plants containing CIVPS or intein modified genes, methods to express CIVPS or intein modified proteins in plants, and methods of using the transgenic plants.

WO 03/056904 A2

[0001] TRANSGENIC PLANTS EXPRESSING CIVPS OR INTEIN MODIFIED PROTEINS AND RELATED METHOD

[0002] BACKGROUND

[0003] FIELD OF THE INVENTION

[0004] The present invention relates to transgenic plants expressing CIVPS or intein modified proteins, methods for the production of the transgenic plants, methods for the expression of CIVPS or intein modified proteins in plants, and various uses of and products containing the transgenic plants expressing CIVPS or intein modified proteins.

[0005] DESCRIPTION OF THE BACKGROUND

[0006] Since fossil fuels are non-renewable resources, adequate supplies of energy and organic feedstocks need to be secured for the future. A transition to sustainable resources requires new technologies for the construction of improved feedstocks, the design of efficient processes to convert the feedstocks into valuable products, and/or the design of products that efficiently utilize an altered substrate spectrum. This transformation will create benefits such as decreased pollution from energy production and use, decreased pollution from chemical manufacturing processes, increased sustainability through the utilization of renewable natural resources and organic waste products as substrates, decreased dependence on foreign country's raw materials, and an increase in local economies and markets involved in the production of new substrates.

[0007] Plant biomass is one sustainable resource that can help meet future feedstock requirements. The use of plants as substrates for energy and organic feedstock takes advantage of existing large-scale agricultural production, uses energy from the sun to incorporate carbon dioxide into plants via photosynthesis, and has fewer environmentally hazardous by-products. By using photosynthesis, plants make the carbon dioxide removed from the air available for the production of energy, chemicals, and agricultural products. Finding ways to effectively redistribute this carbon in forms that are readily and economically employable remains a challenge.

COR-AAD1-00002365

[0008] The production of chemical feedstocks and fuels from plant biomass is still in its infancy. Starch-based raw materials, for example, may be applied to the production of commodity chemical products. Poor substrate and strain availability hampering bioconversion, along with real or perceived safety issues related to containment, and a lack of economic viability, have made progress in this area particularly slow. Non-cellulosic biomass, such as cornstarch, compares favorably with fossil resources on a mass basis, but is too costly. Cellulosic biomass, such as short-rotation poplar, pine, switchgrass, corn stover, sugar cane bagasse, waste paper sludge, and municipal solid waste, in contrast, is cost competitive in terms of both mass and energy. Cellulosic biomass, because of its complex structure, is nevertheless difficult to process. Currently, cellulosic biomass requires pretreatment with strong acids, bases, and/or other chemicals for use as a substrate for fuel, e.g. ethanol, or for chemical production, e.g. paper products. This pretreatment efficiently exposes monomeric subunits, primarily hexoses, pentoses, and phenolic compounds, which are then cleaved and used as substrates, but is expensive. One alternative to the use of more hazardous chemicals is the use of enzymes, although it is not cost effective.

[0009] Recombinant DNA technology has been applied to alter microorganisms to perform substrate bioconversion at reduced costs, thus expanding the use of microorganisms, and increasing the number of products that are produced. For example, plant cells that express lignocellulosic degrading enzymes have been constructed, although they rarely differentiate and regenerate into complete plants due to decomposition of structural components. In cases where they differentiate into complete plants, e.g. with lignin and cellulose substrates, the enzyme titres are low and the plants require further processing. Attempts to combine pretreatment of substrate biomass with fermentation have encountered difficulties as well, in part because of mass transfer limitations and interference with the fermenting organism.

[0010] CIVPS or inteins are in-frame, self-cleaving peptides that generally occur as part of a larger precursor protein molecule. CIVPS or inteins differ from

COR-AAD1-00002366

other proteases or zymogens in several fundamental ways. Unlike proteases that cleave themselves or other proteins into multiple, unligated polypeptides, CIVPS or inteins have the ability to both cleave and ligate in either cis or trans conformations. Thus as opposed to terminal cleavage that would result from the reaction of a protease on a protein, CIVPS or inteins have the ability to cleave at multiple sites, and ligate the resulting protein fragments. This cleavage is induced under specific conditions and can be engineered using molecular biology techniques. CIVPS or inteins have been described in the literature in *Sacchromyces cerevisiae* (Kane et. al., Science 250:651; Hirata et al., J. Bio. Chem. 265:6726 (1990)), *Mycobacterium tuberculosis* (Davis et al., J. Bact. 173:5653 (1991), Davis et al., Cell 71:1 (1992)), *Thermococcus litoralis* (Perler, et al., PNAS 89:5577 (1992)), and in other organisms.

[0011] Accordingly, there is a need for providing novel methods for producing energy from more easily renewable sources, such as by modifying plants in a manner such that they may be used as energy and chemical feedstocks.

[0012] SUMMARY

[0013] The present invention provides for genetically recombinant plants, their parts, plantlets, seeds, seedlings, and their progeny (collectively referred to as "plants"), which may contain single or multiple exogenous gene sequences, each being interrupted by, or fused to single or multiple CIVPS or intein sequences, or a combination of a Controllable InterVening Protein Sequence (CIVPS) or intein sequence, and optionally regulatory sequences suitable for gene expression and transformation of a plant. The modified gene sequences may be expressed constitutively or transiently, throughout the entire plant or in specific tissues, or any combination thereof encompassing both single and multiple CIVPS or intein modified gene sequences. In different embodiments of the invention, any modified gene sequence, or set of modified gene sequences, may be expressed in any or all tissues constitutively or at specific times.

COR-AAD1-00002367

[0014] The invention also relates to methods of producing transgenic plants comprising CIVPS or intein modified genes, e. g. by first constructing a piece of DNA comprising the parent CIVPS or intein modified gene, and transforming the plant with the construct.

[0015] The invention also relates to methods of producing an CIVPS or intein modified protein(s) in transgenic plants, e. g. by transforming the plant, or plant cells, with a single or multiple modified gene sequence(s), and expressing the CIVPS or intein modified protein(s). [In one preferred embodiment the gene sequences may be expressed at any time. In another embodiment, prior to the protein(s) being spliced it preferably is(are) provided with a substantially different activity(ies) and/or structural property(ies). The spliced protein product(s) have its(their) activity(ies) unveiled, unless inhibited by an exogeneously added or endogeneously produced molecule(s) analogous to the non-CIVPS or intein modified protein parent sequence. The CIVPS or intein modified gene products may be expressed in large quantities and recovered from the plant material. Alternatively, the plant or plant material may itself be used as a source of CIVPS or intein modified gene products.

[0016] The invention also provides for the use of CIVPS or intein modified gene products expressed in plants, the use of transgenic plants expressing CIVPS or intein modified genes in animal feed, or the use of transgenic plants expressing CIVPS or intein modified genes in batch, semi-batch, and continuous industrial processes for the production of fuels, chemical products, animal food or food additives, pharmaceuticals, paper, paper products, and for vaccine delivery and the remediation of waste materials.

[0017] Other objects, advantages and features of the present invention will become apparent to those skilled in the art from the following brief description of the drawings and discussion.

COR-AAD1-00002368

[0018] BRIEF DESCRIPTION OF THE DRAWING(S)

[0019] Figure 1 illustrates the construction of an CIVPS or intein modified protein coding DNA sequence by constructing an CIVPS or intein modified protein DNA coding sequence constructed by fusion of an CIVPS or intein coding sequence to the coding sequence of a protein of a purported activity, at either the 3' end of the gene, the 5' end of the gene, or internally, within the protein gene. Other variants are possible by combining any of the three resulting CIVPS or intein modified protein coding sequences shown in Figure 1.

[0020] Figure 2 illustrates one configuration of the resulting CIVPS or intein modified proteins, or components thereof. This figure demonstrates the case of a single CIVPS or intein modified protein. Multiple native protein sequences, however, may be combined with single or multiple CIVPS or inteins as well.

[0021] Figure 3 illustrates the cleavage of an CIVPS or intein modified protein, or components thereof, which may be attained *in vitro* or *in vivo* when subjected to an appropriate cleavage stimulus(i). Illustrated here schematically is an example of the cleavage process for a single CIVPS or intein modified protein. Other variants may be constructed as combinations of the CIVPS or intein modified proteins shown in this figure.

[0022] DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT(S)

[0023] This invention arose from a desire by the inventor to provide novel methods for generating energy from renewable resources, e. g. plant materials or biomass, and to do this in a cost effective manner. He posited that one way to effectively attain this goal was by modifying plant biomass through the use of CIVPS or intein modified proteins, where the CIVPS or intein is attached to a desired protein. Within the text of this patent the terms CIVPS and intein are intended to refer to similar products, and will be used interchangeably. From the knowledge that intein modified proteins may be expressed in cells at high titer, yet with substantially decreased activity, he concluded that, if cloned into a

COR-AAD1-00002369

plant, this decrease in activity would allow the thus formed transgenic plant cells, plant fragments, or plant tissues, to develop into intein modified protein producing complete plants. Moreover, he thought that such transgenic plants could be provided as several different embodiments, such as those where the recombinant plants are made to express the modified proteins either 1) constitutively or transiently, 2) through chemical induction or biological induction by the plant's growth cycle, 3) throughout the entire plant or specifically in distinct plant tissues, and/or 4) with or without subcellular localization, among others. As envisioned by the inventor, in one embodiment of this invention, the expressed intein modified protein(s) is(are) comprised of a parent protein sequence(s), whose activity(ies) may be known, inferred through sequence or structure homology and/or produced by mutagenesis or by de novo synthesis; each parent sequence(s) being interrupted by, or fused to, an intein sequence(s). Once inserted, the intein portion(s) of the modified protein(s) inactivate(s), in vivo, the activity or structural utility of the parent proetin. The parent protein's original activity may be, however, substantially recovered, if and when desired, by induction of intein splicing. For example, in one application, following plant harvest and during substrate pretreatment, each CIVPS may be induced to splice itself from its parent protein sequence, which parent protein now has recovered its original activity. Methods for intein splicing with, or without, recombining of the protein to a functioning activity are known to one skilled in the art, and need not be repeated here. These methods include the use of light, temperature, change in pH, and/or the addition of chemical reagents.

[0024] More specifically, this invention is directed to a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, comprising an expression construct(s) that encode(s) at least one modified protein comprising a target protein(s) or protein segment(s), which is(are) fused, either internally or terminally, to a controllable intervening protein sequence(s) (CIVPS) or intein sequence(s) or segment(s) thereof, or to an amino terminus(i) or a carboxyl terminus(i) thereof. In one

COR-AAD1-00002370

embodiment, each expression construct of the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent comprises, operatively linked to one another, a first nucleic acid segment(s) encoding a target protein(s), and a second nucleic acid segment(s) encoding a CIVPS or intein sequence(s), and optionally a selectable marker(s) or reporter gene(s) and/or a promoter(s). It is understood that in a more specific embodiment the sequences may be fused, either directly or via a linker(s), and more preferably in reading frame. The modified protein(s) may be expressed by the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent either constitutively, or inductively. In the latter case, the expression and/or splicing of the at least one modified protein(s) may be triggered or induced by a stimulus(i). Examples of suitable stimuli comprise a pH change, change in osmolality, or temperature, the addition of a fertilizer, pesticide, or chemical, or a change in light, and/or sound. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may express the modified protein(s) either at a pre-determined point of the plant life cycle, in one or more specific tissues or parts thereof, and/or in at least one specific sub-cellular compartment(s). Alternatively or in conjunction with the latter the modified protein(s) may be expressed and secreted extracellularly. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent specific tissue(s) may be seeds, roots, fruits, stems, tubers and/or leaves, and the specific subcellular compartments may be a cellular cytosol, mitochondrion, plastid, endoplasmic reticulum, inclusion body, vacuole and/or nucleus. Other variations, however, are also included within the confines of this invention.

[0025] The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may also carry a selectable marker that confers it resistance to a chemical. Examples of these are bromoxynil, 2,2-dichloropropionic acid, G418, glyphosphate, haloxyfop, hygromycin, imidazoline, kanamycin, methotrexate, neomycin, phosphinothricin,

COR-AAD1-00002371

sethoxydim, 2,2-dichloropropionic acid, glyphosphate, hygromycin, trichothecne, sulfonylurea, s-triazine, and/or triazolopyrimidine. Others, however, may also be employed. The promoter may be included to precede a CIVPS or intein-modified protein polynucleotide. In some cases, the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may be tolerant or resistant to normally extremely toxic levels of a selected chemical(s). In another embodiment, the plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent is fertile, and has at least one heritable modified protein encoding polynucleotide sequence(s). However, it may just as well not be fertile. Further, as indicated above, also part of this invention are inbred and hybrid genetically recombinant plants, or plant parts, plantlets, tissues, cells, sub-cellular fractions, seeds, seedlings, protoplasts, progeny and descendents, which may or may not be produced by the method of this invention. Of particular interest are plant parts, plant seeds, plant seedlings and plant protoplasts, which have substantial commercial importance. Also of commercial and other interest are plants, plant tissues, plant cells, and sub-cellular fractions. The spliced protein may have the ability of changing the content or activity of one or more plant component(s). In one example, the content may be altered, e.g. reduced, of a plant component such as glucose, fructose, glycerol, glycine-betaine, sucrose, lactose, maltose, galactose, amino acids, lipids, vitamins and/or starch, and the like. In another, the plant component whose activity is altered, e.g. reduced, may be ne or more of cellulose, hemicellulose, lignin, starch, pectin and/or lipids, among others. In one aspect, the CIVPS or intein sequence and the target protein or protein segment form at least one splice junction with the target protein. In a desirable embodiment, the amino acid residue at the carboxyl terminus(i) of the splice junction(s) is(are) provided with a hydroxyl or a sulfhydryl side chain(s). In another particularly useful embodiment, the splice junction(s) is placed downstream of the CIVPS or intein sequence(s) or segment(s) thereof, and may comprise(s) an amino acid residue(s) lacking, for example, hydroxyl or sulfhydryl side chains at the amino

COR-AAD1-00002372

terminus(i) of the target protein or protein segment(s). In another important variation, the splice junction(s) is(are) placed upstream of the CIVPS or intein sequence(s) or segment(s) thereof, and may comprise an amino acid residue(s) having hydroxyl or sulfhydryl side chains at the amino terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof. Another important possibility is that where the splice junction(s) is(are) placed upstream of the CIVPS or intein sequence(s) or segment(s) thereof, and it may comprise(s) a cysteine. Still another important variation is that wherein the splice junction(s) is(are) placed downstream of the CIVPS or intein sequence(s) or segment(s) thereof, and may be provided with His-Asn at the carboxyl terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof, and/or with an amino acid residue(s) having hydroxyl or sulfhydryl side chains at the amino terminus(i) of the adjoining region(s) of the target protein(s). In yet another interesting variant, the splice junction(s) is placed downstream of the CIVPS or intein sequence(s) or protein segment(s) thereof, and may be provided with an Asp at the carboxyl terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof, and/or with an amino acid residue(s) having hydroxyl or sulfhydryl side chains at the amino terminus(i) of the adjoining region(s) of the target protein(s) or protein segment(s). Further modifications are those where the Asp at the carboxyl terminus(i) is replaced by an amino acid(s) lacking carboxyl or amino side chains, and where the CIVPS or intein sequence(s) or its segment(s) comprise(s) an externally controllable CIVPS or intein sequence(s) or segment(s) thereof, which may be from, among other species, a Saccharomyces fungi, and more specifically a Saccharomyces cerevisiae fungi. Other constructs suitable for insertion in the products of the invention are those where the CIVPS or intein sequence(s) or segment(s) thereof is(are) inserted immediately before Ser, Thr or Cys of the target protein(s) or protein segment(s), and where the CIVPS or intein amino or carboxy terminus(s) comprise(s) Ser, Thr or Cys, among others. As described in more detail below, the protein may be expressed in a microorganism, such as a bacterium, as is known in the art. Examples of microorganisms that may be employed are Bacillus

COR-AAD1-00002373

thuringiensis, or Phytolacca insularis. One preferred target protein is Bacillus thuringensis endotoxin, which results in a modified Bacillus thuringiensis endotoxin being expressed. Another embodiment includes the expression of the modified protein by a virus. Although any virus could be employed, examples are potato virus Y, geminivirus, aspermy virus 2b, and cucumber mosaic virus, among others.

[0026] The recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may be produced by a method comprising

[0027] providing an expression construct that encode(s) at least one modified protein comprising a target protein, or protein segment(s), which is(are) fused, either internally or terminally, to a CIVPS or intein sequence(s) or segment(s) thereof, or to an amino terminus(i) or a carboxyl terminus(i) thereof;

[0028] transforming a plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, with the expression construct; and

[0029] regenerating a genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, from the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, that encode(s) at least one modified protein sequence(s).

[0030] It is highly preferred that the transformation be a stable transformation. However, transformations that have some temporary stability are also desirable. The regeneration step may be conducted by breeding of a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling protoplast, progeny or descendent; crossing of a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent and a non-genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent; and/or back-crossing of two genetically recombinant plant, or plant

COR-AAD1-00002374

part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent. The expression construct employed in this method may comprise one or more of promoter, selectable marker, resistance marker, heritable marker, poly-adenylation sequence, repressor, enhancer, localization sequence, and/or signaling sequence. These are intended for use in the application of recombinant technologies as is known in the art, and exemplified elsewhere and below in the examples. In an important aspect of the method, the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is(are) transformed with the expression construct by either viral transformation, bombardment with DNA-coated microprojectiles, liposomal gene transformation, bacterial gene transfer, electroporation, or chemical gene transformation, or more than one of these. As indicated above, the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, may be transformed by means of a bacterium, e. g. Agrobacterium tumefaciens, although other microorganisms may also be employed. In the present method, the transformation may be conducted by chemical gene transformation, and it may be done with the aid of, e.g. calcium phosphate, and/or polyethylene glycol, or other chemicals known in the art as being suitable for this purpose. The selection may be attained with the aid of a selectable marker, or a resistance marker, or of the expression of at least one nucleic acid encoding an CIVPS or intein modified protein. In the method of the invention, the genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may be regenerated from a transformed embryogenic tissue(s); plant protoplasts; cells derived from immature embryos; or from transformed seeds, among other sources.

[0031] Another method is also provided in this patent, which method is suitable for producing a modified protein(s) or protein segment(s) from a recombinant transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent expressing the

COR-AAD1-00002375

protein(s) or protein segment(s), that comprises conducting the method described above, and further harvesting the modified protein(s) or protein segment(s) from the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent. The method may further comprise purifying the modified protein, which may be done by one of many techniques known in the art. As described here, this method may produce a modified protein(s) or protein segment(s) that comprises a CIVPS or intein modified protein(s) or protein segment(s).

[0032] Still a further method is provided here for producing a modified protein comprising a target protein(s) or protein segment(s) fused, either internally or terminally, to a CIVPS or intein sequence(s0 or segment(s) thereof, or to its amino terminus(i) or carboxyl terminus(i), which method comprises

[0033] obtaining an expression construct encoding a target protein having an in-frame fused CIVPS or intein sequence(s) or segment(s) thereof, or its amino terminus(i) or carboxyl terminus(i);

[0034] transforming a host plant cell(s) with the expression construct; and

[0035] culturing the transformed plant host cell under conditions effective for expressing the modified protein.

[0036] In one preferred aspect, in the expression construct the at least one first nucleic acid segment(s) encoding the CIVPS or intein sequence(s) or segment(s) thereof is(are) fused to the 5'-end of the second nucleic acid segment(s) encoding the target protein(s) or protein segment(s). Alternatively, in the expression construct the first nucleic acid segment(s) encoding the CIVPS or intein sequence(s) or segment(s) thereof may be fused to the 3'-end of the second nucleic acid segment(s) encoding the target protein(s) or protein segment(s). It is particularly suitable to practice the present method to employ a Saccharomyces CIVPS or intein sequence(s) or segment(s) thereof, which is known to effect, either in cis or in trans, excision, cleavage, ligation, excision-ligation, cleavage-ligation, and/or cyclization. When the CIVPS or intein or its(their) segment(s)are employed to induce protein splicing, this event may be induced or triggered by a

COR-AAD1-00002376

change of temperature, light or pH, the addition/removal of a chemical reagent that facilitates/inhibits splicing or cleavage, amino acid dephosphorylation or deglycosylation, or by contact with, or removal of, a peptide or peptidomimetic activating or blocking of splicing or of cleavage. Another manner of inducing protein splicing is either in vitro or in vivo contact with, or removal of, a peptide or peptidomimetic agent that may either activate or block splicing or cleavage. Interesting variations that produce superior results are those where the amino or carboxy terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) Ser, Thr or Cys, or where the carboxyl terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) Asp preceding Ser, Thr or Cys of the target protein(s) or protein segment(s). However, other modifications are also possible, as is known in the art. See, for example, US patent No. 5,834,247 that discloses for the prokaryotic and eukaryotic realms some methodology incorporated in this invention to the production of hybrid plants of useful characteristics. In the present method, the expression construct may further comprise a promoter, a selectable marker, a resistance marker, a heritable marker, a poly-adenylation sequence, a repressor, an enhancer, a localization sequence, or a signaling sequence. Moreover, the method presented here may also comprise the transformation of the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent with the expression construct being implemented by viral transformation, bombardment with DNA-coated microprojectiles, liposomal gene transfer, bacterial gene transfer, electroporation, and/or chemical gene transformation, and/or other methods known in the art, or that will be subsequently developed. As described above, in the method described here, the bacterium used to transfer the expression construct may be an Agrobacterium tumefaciens bacterium; the chemical used for transformation may be calcium phosphate, or polyethylene glycol; the transformed plant cells, plant parts, plants, etc. may be selected through their expression of a selectable marker, or resistance marker; the selection of the transformed plant, or plant part, plantlet,

COR-AAD1-00002377

tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may be conducted through their expression of the modified protein gene sequence; and the regeneration of the genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent may be attained from transformed embryogenic tissue; from cells derived from immature embryos; or from transformed seeds, among others.

[0037] Also disclosed in this patent is a method for producing seed that express a modified protein(s), this method comprising

[0038] obtaining the genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of the invention;

[0039] culturing or cultivating the genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent; and

[0040] obtaining from the cultivated plant seed that expresses a modified protein(s).

[0041] Still another method provided by this patent is one for using a plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent expressing a modified protein for producing a compound, the method comprising

[0042] harvesting a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent in accordance with the teachings of this patent;

[0043] mechanically processing the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent;

[0044] combining the mechanically processed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, with a non-genetically recombinant plant in a proportion greater than or equal to zero recombinant:non-recombinant; and

COR-AAD1-00002378

[0045] chemically processing the plant or specific portions of the plant under conditions effective for obtaining the compound.

[0046] This method may be practiced by mechanical processing of the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent by extrusion, grinding, shredding, mulching, chipping, dicing, compressing, exploding, and/or tearing. Other processing techniques, however, are also suitable. The chemical processing of the combined components may be attained by various techniques or a combination thereof. Some of them are pre-treatment with steam, dilute or concentrated acid, ammonia explosion, sterilization, soaking in water, mixing with a solvent, a change of pH, temperature or osmolality, exposure to or changes in light, inorganic and/or enzyme catalysis, saccharification, bleaching, scouring, fermentation, distillation, chromatography, adsorption, and/or addition of a chemical(s). Others, of course, are also employed successfully. Various steps are of use when practiced as follows: the pre-treatment may include steaming the combined products for sterilization purposes; the chemical processing may be attained by pre-treatment with at least one of sulfuric acid, hydrochloric acid, phosphoric acid, or carbonic acid, or by soaking in water at a temperature greater than or equal to about 20 °C, and/or by mixing the combined products with at least one of water, or an organic or inorganic solvent(s). As already explained, an external stimulus(i) may be applied to induce splicing of the modified protein(s) or protein segment(s). Examples of external stimuli are a change of pH, osmolality, or temperature, exposure to sound, light, or addition of a chemical(s). In some cases the spliced protein(s) or protein segment(s) may exhibit altered activity(ies) with respect to the modified protein(s) or protein segment(s), such as altered catabolic or anabolic activity(ies) with respect to the original target protein(s). Examples of spliced protein(s) or protein segment(s)are those capable of degrading starch, dextrin, pectin, lipids, protein, chitin, lignin, cellulose, or hemicellulose, or modifying lignin, or having saccharification activity. Thus, the spliced protein may be capable of producing glucose, fructose, xylose, phenol,

COR-AAD1-00002379

glycerol, mannose, lactic acid, acetic acid, ethylene, propylene, toluene, ethyl benzene, styrene, xylene, ethylene glycol, butadiene, formaldehyde, isopropanol, acetone, butanediol, methanol, ethanol, propanol, butanol, propanediol, vitamins, methane, ethane, propane, butane, pentane, hexane, heptane, octane, benzene, and/or biopolymers, among other compounds. In one specific embodiment of the pre-treatment, saccharification, and fermentation may be conducted in one step, and the fermentation may be attained by employing a prokaryotic or eukaryotic microorganism capable of producing lactic acid, acetic acid, ethylene, propylene, toluene, ethyl benzene, styrene, xylene, ethylene glycol, butadiene, formaldehyde, isopropanol, acetone, butanediol, methanol, ethanol, propanol, butanol, octanol, propanediol, vitamins, methane, ethane, propane, butane, pentane, hexane, heptane, octane, benzene, and/or biopolymers, among other compounds.

[0047] This invention also encompasses the production of animal feedstock that comprises a nutritious amount of the recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of the invention. When the feedstock provided by the inventor is ingested by an animal, the modified protein(s) or protein segment(s) is(are) spliced by an internal stimulus(i) from the animal. Examples of internal stimuli are the animal's saliva, bile, chymotrypsin, trypsin, bicarbonate, hydrochloric acid, or stomach pH or temperature, among others. The feedstock of the invention may comprise spliced protein(s) such as phytases, endocellulases, exocellulases, amylases, glucanases, hemi-cellulases, pectinases, proteases, xylanases, or lipases, or a growth hormone. Other proteins, however, could also be employed as desired.

[0048] Yet another aspect of this invention provides for the use of the feedstock described above in the manufacture of an immune response enhancing composition, wherein the spliced protein(s) or protein segment(s) comprise(s) at least one recombinant immunogen(s). The immunogen may included one or more viral or bacterial immunogens, and it may fe formulated in various suitable forms. Preferred are an oral formulation, a trans-mucosal formulation, a

COR-AAD1-00002380

gastrointestinal (G.I.) tract absorbed formulation. However, this composition of matter may be formulated in any systemic or topical form suitable for administration to an animal, including its addition to animal feed.

[0049] The animal feedstock of the invention may be produced by first conducting the steps indicated above to obtain a genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, and then processing the genetically modified plant, or a portion of the resulting product under conditions effective to obtain an animal digestible feedstock.

[0050] The product of this invention may also be employed for promoting animal growth, for example by producing feedstock that comprises a growth promoting product, and allowing an animal access to the modified feedstock. The product of this invention may also be employed for enhancing an animals immune response. This may be done by administering to an animal in need of the treatment, an immune enhancing amount of the composition of the invention.

[0051] A further aspect of this invention involves a method for producing a target protein(s) or protein segment(s), the method comprising

[0052] producing a first modified protein(s) or protein segment(s), wherein the amino terminus of a CIVPS or intein sequence(s) or segment(s) thereof is(are) fused to the carboxyl terminus(i) of a target protein(s) or protein segment(s) by the method described above;

[0053] producing a second modified protein(s) comprising a segment(s) of the CIVPS or intein sequence(s); and

[0054] contacting first and second modified proteins under conditions effective for trans cleavage of the CIVPS or intein sequence(s) or segment(s) thereof by the second modified protein(s).

[0055] Yet another variation of the above method for producing a target protein(s), comprises

[0056] producing a first modified protein(s), wherein the carboxyl terminus of a CIVPS or intein sequence(s) or protein segment(s) thereof is(are) fused to the

COR-AAD1-00002381

amino terminus(i) of the target protein(s) or protein segment(s) by the already described method;

[0057] similarly producing a second modified protein(s) or protein segment(s) comprising a segment(s) of the CIVPS or intein sequence(s); and

[0058] contacting first and second modified proteins under conditions effective for trans cleaving the CIVPS or intein sequence(s) or segment(s) thereof from the first modified protein(s) or protein segment(s). The cleavage may be induced in this procedure by a change in temperature, light, or pH, addition/removal of chemical that facilitates/inhibits splicing or blocking of cleavage, amino acid dephosphorylation or deglycosylation, and/or contact/removal of peptide or peptidomimetic that activates/blocks splicing/cleavage, among others.

[0059] Thus, the invention is directed towards transgenic plants, which term is intended in this patent to be synonymous with genetically recombinant plants, their seeds and progeny plants, or any plant portion, tissue or cell, containing a gene(s) for an CIVPS modified protein(s). The invention is further directed towards methods for the production of the transgenic plants that produce CIVPS modified proteins, methods for the production of intein modified proteins in plants, and uses of the plants as substrates for fuels, chemicals, animal food or food additives, paper, and pharmaceutical production. The invention allows for the production of transgenic plants that can be used as a source of structural or catalytic components, or can have their intein modified proteins purified and used separately as structural or catalytic proteins. Transgenic plants are multi-cellular plants that express single or multiple exogenous genes and their associated protein (or ribonucleic acid) activities. Within the context of this invention, gene or enzyme classes may be specifically referred to, however this is not a limiting aspect of the invention. When specific classes are stated, this is understood to identify any gene or enzyme within the specific classification. CIVPS or inteins are protein sequences internal or adjacent to a parent protein sequence, that may spontaneously cleave themselves at either, or both, the

COR-AAD1-00002382

carboxyl or amino terminal ends and are capable of selectively ligating the resulting extein protein fragments when appropriate, under specific conditions. See, for example, Perler, et al., Nucl. Acids Res., 22:1125-1127 (1994); Wallace, C. J., Protein Sci., 2:697-705 (1993); Xu, et al., Cell, 75: 1371-1377 (1993); Pietrokovski, S.,Protein Sci., 2:697-705 (1994). Thus, CIVPSs may be said to be in-frame, self-cleaving peptides that generally occur as part of a larger precursor protein molecule. CIVPS or inteins differ from other proteases or zymogens in several fundamental ways. Unlike proteases that cleave themselves or other proteins into multiple, unligated polypeptides, inteins have the ability to both cleave and ligate in either cis or trans conformations. Thus as opposed to terminal cleavage that would result from the reaction of a protease on a protein, inteins have the ability to cleave at multiple sites, and ligate the resulting protein fragments. This cleavage is induced under specific conditions and may be brought about implementing techniques that are known in molecular biology. Inteins from various sources, their sequences, characteristics and functions have been described fully in the literature. See, for example, Kane et. al., Science 250:651 (1990); Hirata et al., J. Bio. Chem. 265:6726 (1990) (Sacchromyces cerevisiae); Davis et al., J. Bact. 173:5653 (1991), Davis et al., Cell 71:1 (1992) (Mycobacterium tuberculosis); Perler, et al., PNAS 89:5577 (1992) (Thermococcus litoralis). As shown in Figure 1, the combination of an CIVPS with a protein of purported activity or structural role yields an intein modified protein, whose purported activity or structural role may be substantially altered. Transgenic plants that express CIVPS modified proteins (from their associated intein modified genes) are an improvement upon previous transgenic plants, because the parent intein modified protein can have two substantially different states that are controllably mediated by intein cleavage. This cleavage may or may not be associated with recombination of the purported protein sequence. The invention may be formed from any plant species, combined with any combination of single or multiple proteins and CIVPS. Plant species may include, but are not limited to: poplar, birch, cedar, pine, hardwoods, softwoods, soybeans,

COR-AAD1-00002383

switchgrass, corn, tobacco, alfalfa, sugar cane, cauliflowers, artichokes, bananas, apples, cherries, cranberries, cucumbers, lettuce, grapes, lemons, melons, nuts, tangerines, rice, oranges, peaches, pears, blueberries, strawberries, tomatoes, carrots, cabbages, potatoes, endive, leeks, spinach, weeds, arrowroot, beets, carrots, cassava, turnips, yams, radishes, sweet potatoes, wheat, barley, soya, beans, rapeseed, millet, sunflower, oats, peas, tubers, bamboo, seaweed, algae, or any other plant species. Proteins may include any known, putative, modified, or de novo created proteins. Although the selection of the native protein is not restricted, preferred proteins include lignocellulosic degrading proteins (cellulases, lignases, amylases), starch degrading enzymes (amylases, glucanases), enzymes in the biosynthetic pathways required for fuel or chemical production, bacterial or viral antigens, enzymes in the biosynthetic pathways for vitamins or other food additives (phytases, cellulases, amylases, glucanases, hemi-cellulase, pectinase, protease, xylanase, lipase, growth hormone), proteins that impart pest or insect resistance, and therapeutic proteins implicated in disease pathogenesis. The choice of CIVPS or intein used to modify the protein, the fusion of which is expressed in the desired plant, is also not limited. Any single or multiple CIVPS or intein may be used in any configuration with respect to the desired protein or proteins. The CIVPS or inteins should have the capability to be spliced at one or both ends in response to some stimuli, and may or may not permit ligation of the proteins to which single or multiple CIVPS or inteins are fused.

[0060] Transgenic plants expressing CIVPS or intein modified proteins, and the production of CIVPS or intein modified proteins in transgenic plants can be accomplished by combining methods (Ausbel, et al.) known in the art. Generally, these methods include construction of a DNA containing the CIVPS or intein modified protein of interest and the necessary regulatory elements required for its expression, amplification and selection of the constructed DNA, transformation of the desired plant species, regeneration and selection of the appropriately transformed plant species, and if necessary, purification of the

COR-AAD1-00002384

CIVPS or intein modified protein in its native form or the cleaved form. Both the production of transgenic plants expressing CIVPS or intein modified proteins, and the production of CIVPS or intein modified proteins in transgenic plants form part of this invention. For the production of the transgenic plants, or CIVPS or intein modified proteins in transgenic plants, the CIVPS or intein modified protein DNA sequence must be constructed. This is easily accomplished by cloning the gene sequence of the desired activity and the desired intein sequence into *E. coli* or any other suitable host (e.g., yeast may be beneficial in some cases, or expression in mammalian or plant cells with or without the use of viral or non-viral vectors). Once the gene and intein coding sequences have been cloned, they must be joined in the desired configuration. The chosen intein sequence should be able to perform the desired functions such as splicing in response to an imposed stimuli (for example, light, pH change, temperature, pressure, or changes in the local chemical composition surrounding the intein modified protein), and if necessary permitting ligation of the fused protein. Joining of the CIVPS or intein's DNA sequence and the protein's DNA sequence is easily accomplished by methods known in the art, resulting in CIVPS or intein modified protein DNA coding sequences, or combinations thereof, as shown in Figure 1. As already indicated, an CIVPS or intein modified protein is one which fuses the CIVPS or intein to one of either the carboxy terminal, amino terminal, or internal portions of the native protein or proteins. Although many alternative methods exist, one way of creating the fusion between the CIVPS or intein and desired protein coding sequences would be to purify the DNA encoding the desired protein sequence, use a restriction enzyme to cut the protein coding sequence at the desired point of intein insertion, and then ligate the intein coding sequence into the restricted site. The polynucleotide, or either of the nucleic acid segments may be cloned directly to appropriate regulatory and/or selection sequences, or via a vector them. Examples of regulatory segments are promoters to control the temporal expression of the CIVPS or intein-modified protein, origins of replication, and/or signaling sequences to control the spatial

COR-AAD1-00002385

distribution of CIVPS or intein-modified proteins in vivo in specific plant tissues and/or specific subcellular compartments, and/examples of selection elements are herbicidal or antibacterial genes, fluorescent makers, dye markers, and other suitable selective markers. The resulting polynucleotide or vector comprising the CIVPS or intein modified protein(s) encoding polynucleotide(s), and optionally any desired regulatory, and selection elements, then may be amplified to obtain larger amounts of product, which may be used for subsequent transformation of a desired plant species. Modification of any and all of these steps is possible to facilitate specific orientation and fusion between any desired CIVPS or intein(s) and protein(s)outine , and it is conducted employing methods that are known in the art. Alteration of either the coding sequences and/or the CIVPS or intein coding sequence and the ligation of either or both of these sequences may also be easily accomplished by techniques known in the art, such as site-directed mutagenesis, random mutagenesis, polymerase chain reaction (PCR), error-prone PCR, and/or any other suitable method that would be considered routine by an artisan. These techniques facilitate the placement of a number of joining sequences, and any desirable and suitable combination may be used. Likewise, any combination or orientation of regulatory and selective elements may also be implemented in accordance with this invention. Gene regulatory elements, such as promoters (Guilley et al., Higgins, T.J.V., Coruzzi et al., Tingey et al., Ryan et al., Rocha-Sosa et al., Wenzler et al., Bird et al.), enhancers (Brederode, et al.), RNA splicing sites, ribosomal binding sites, glycosylation sites, protein splicing sites, subcellular signalling sequences (Smeekens et al., van den Broeck et al., Schreier et al., Tague et al.), secretory signal sequences (Von Heijne, G., Sijmons, et al.), or others may be advantageous in controlling either the temporal or spatial distribution of the CIVPS or intein modified protein concentration and activity *in vivo* in the transformed plant. Use of these elements may be desired to facilitate the production and processing of intein modified proteins in transgenic plants. The expression of the intein-modified protein(s) may be conducted either in a constitutive or induced manner. In ordeer to attain either of

COR-AAD1-00002386

these modes, any of the methods that are either described in this patent or known in the art, or later made available, may be implemented. The induction of protein expression may be attained with the aid of a foreign stimulus(i). Exemples of these are the exposure to a pesticide(s), to light, a temperature change(s), and/or sound(s). Other foreign stimuli, however, may also be employed. In addition, the recombinant plant may also express any one or more of the selectable marker gene or reporter gene(s) mentioned above.

[0061] Once the CIVPS or intein modified protein DNA sequence has been constructed, combined with the desired regulatory and selection DNA sequences, successfully cloned and selected, then transformation of the desired plant species and generation of full plants is required. Methods for transformation of a desired plant species, and the generation of full plants can be accomplished by techniques known in the art (Draper, et al., Potrykus, et al.). Transformation techniques include, but are not limited to: *Agrobacterium tumefaciens* mediated gene transfer, *Agrobacterium rhizogenes* mediated gene transfer, direct gene transfer to plant protoplasts, Ti plasmid mediated gene transfer (with or without a helper plasmid), biolistic or particle bombardment plant tranformation (Gordon-Kamm et al.), microinjection and fiber-mediated transformation, and tissue electroploration (Shimamoto et al.). Gene transfer may occur in whole plants, plant explants (such as, but not limited to root explants), any plant portion (such as, but not limited to plant leaf segments, seeds, or seed segments), plant protoplasts or apoplasts, or single or multiple plant cells. Each different method has been substantially described in detail by the prior art. Methods of selection of properly transformed plants are known in the art. Selection methods may be facilitated by including a selectable marker in the transformed DNA containing the CIVPS or intein modified protein (such as a resistance gene, gene coding the production of a colored compound, gene coding the production of a fluorescent compound, or any other suitable method). Additionally, DNA from transformed plants may be isolated and sequenced to confirm the presence of the desired CIVPS or intein modified protein coding sequence. Other techniques are also

COR-AAD1-00002387

suitable for confirmation of the selection process, such as polymerase chain reaction, restriction digest analysis and southern analysis. Any method of selection that allows identification of the desired transgenic plant may be used. Once the plant is transformed with the CIVPS or intein modified protein and desired regulatory and selection sequences, whole plants can be regenerated by methods know to the art (Horsch et al.). Most methods consist of culturing the transformed plant cells, explants, tissues, parts, or whole plants in the proper medium and under appropriate conditions of light and temperature. The method used to regenerate the plant should not limit the invention and any effective method may be used. The resulting transgenic plant should produce CIVPS or intein-modified proteins that are substantially described as, or a combination of, those shown schematically in Figure 2. Once the whole, transgenic plant has been selected, it can be monitored for CIVPS or intein modified protein expression. This is not required for the production of transgenic plants expressing CIVPS or intein modified proteins, but is prudent to confirm that the desired transgenic plant expressing the desired CIVPS or intein modified protein has been obtained and expression is properly controlled by the desired control elements used. Monitoring of CIVPS or intein modified protein expression is necessary for the purification of the CIVPS or intein modified proteins in the cleaved or uncleaved state, as described schematically in Figure 3 for either whole intein modified proteins, or components of intein modified proteins that are composed of combinations of elements shown in Figure 3. Protein expression of the intein modified protein can be monitored by western analysis, 2-dimensional gel electrophoresis (and staining), or mass spectrometry, conducted on plant extracts or protein fractions purified from the transgenic plant. In addition, either some of the purified proteins, or the transgenic plant itself, should be exposed to the intein cleavage stimulus. After exposure, both the CIVPS or intein modified protein and the resulting protein that appears as a consequence of CIVPS or intein cleavage can both be analyzed by western analysis, and other assays, to verify the presence of the appropriate proteins, and the difference in

COR-AAD1-00002388

activity between the intein modified protein and the resulting cleaved protein. The activity assays must be designed so as to monitor the desired protein activity and should be specific to that activity and not vulnerable to competing interferences. A control can be used as a standard to compare the native activity with both the intein modified activity and the activity following intein cleavage. Methods and processes using transgenic plants expressing CIVPS or intein modified proteins include the use of the plants as substrates for fuel production (including, but not limited to: burnable biomass, ethanol, methanol, propanol, propane, methane, or octane production), the use of the plants as substrates for commodity chemical production (including, but not limited to: lactic acid, ethanol, glucose or other hexoses, pentoses, propane diols, ethene, ethane, ethylene, phenolic compounds, amino acids, paper pulp, pesticides, insecticides, other alcohols, other ethers, other esters), the use of the plants as substrates for food production and or food additive production (including but not limited to: amino acids, sugars, vitamins, fiber, or cattle feed), the use of the plants for vaccine delivery, the use of the plants for paper production, and the use of the plants for remediation of waste materials. Any batch, semi-batch, or continuous process in which transgenic plants that express intein modified proteins are used as substrates for one of the purposes described above is claimed. These processes may include, but are not limited in scope to, processes in which the transgenic plants expressing intein modified proteins are harvested, exerted to the intein cleavage stimuli, mixed with other substrates in a transgenic plant to substrate ratio greater than or equal to zero, and then converted either chemically, enzymatically, or biologically to one of the products detailed above.

[0062] The examples provided below illustrate the process of the invention, as well as the manufacture of transgenic plants expressing CIVPS or intein modified cellulase enzymes, and the thus produced plants. In these plants the cellulase enzymes are expressed as dictated by the regulatory elements controlling the CIVPS or intein modified genes. The cellulase activity is substantially reduced *in vivo* by interruption of the native cellulase enzyme by

COR-AAD1-00002389

the fused intein. This allows the plant to grow, uninhibited or with little inhibition by cellulase activity. The plants may be harvested and exerted to the intein cleavage stimuli, such as exposure to a certain wavelength of light, mixed with a sulfurous or pH altering chemical, mixed with a salt, mixed with any other chemical, or exerted to a change in temperature. In this case, the CIVPSs or intein is be cleaved and the cellulase activity recovered, which then catalyzes the cleavage of cellulose and/or lignin. At this point the cleaved protein plant mash may be mixed in any proportion, preferably greater than or equal to zero, with other plant substrates, chemical substrates, municipal waste, manufacturing by-products, enzymes, and/or prokaryotic or eukaryotic cells, among others, to aid in the conversion of the plant substrate to the desired product, e.g. a fuel, commodity chemical, food for human or animal consumption, food additive, paper pulp, or vaccine antigen, among others. It should also be noted that the use of the present invention is not limited to manufacturing processes or mechanical processes. Non-limiting examples of applications of this invention are in the delivery of vaccines, hormones, or therapeutic proteins, in which case the intein modified protein may comprise a combination of therapeutic protein(s) and/or protein antigen(s), potentially protective protein sequences, and CIVPSs or intein(s) that may be expressed by the transgenic plant, e.g. a banana plant. The delivery process may occur, for example, by ingestion of the plant product by a human or non-human animal. The plant is then masticated in the mouth and exposed to a stimulus(i) in vivo in the stomach, which in turn triggers or induces cleavage by the CIVPS or intein. In the case of humans the stimulus may be the reduced pH of the stomach, which induces the cleavage of the CIVPS or intein from the antigen or therapeutic protein, and provides for appropriate ligation, if necessary. The therapeutic protein or antigen would then flow into the duodenum, or small intestine, where the pH would be neutralized and protein products are now ready to be absorbed into the blood stream.

[0063] Background For Exemplarly Information Provided Below

COR-AAD1-00002390

[0064] Many different variations in the protocol presented in Example 1 below are suitable for practicing the present invention, as an artisan would know. In general, a DNA sequence encoding a CIVPS or intein modified protein is constructed and packaged into an appropriate vector, plant material, whether it is single cells grown in suspension, protoplasts, plant segments or parts, whole plants, or other forms suitably described here are transformed with the vector, and complete plants, seeds, or other plant forms described here are regenerated. Example 1 shows one embodiment of the inventive method, variations of which are possible that may be used to generate a transgenic tree, e.g. a poplar species expressing an intein modified cellulase. The choice of desired protein, however, depends upon the application the transgenic plant species is intended for. In this regard native proteins, *de novo* synthetic proteins, or evolved proteins, e.g., by gene shuffling, error prone PCR, or any other analogous method, may be used. Cellulases catalyze a cleavage reaction in breaking down cellulose, a chemical component of the plant. While other plants have been constructed expressing cellulases the enzymes typically have to be transiently expressed, or sequestered in parts of the cell so as not to disrupt plant tissue differentiation and development. See, for example, Ziegler et al. (2000); Dai et al. (a), (2000); Dai et al. (b) (2000); Montalvo-Rodriguez et al. (2000). Hence, in the case where the cellulase activity is not controlled by localization or transient expression, whole plants are often very difficult to regenerate, or the cellulase activity is often too low to be useful. By using an intein modified cellulase, the whole plant can be regenerated while the less activate intein modified cellulase is produced throughout the plant and at high titer. See, Aspergen et al., *Molecular Breeding* 1:91-99 (1995). The enzyme can be subsequently activated by the self-splicing ability of the intein to yield a cellulase of increased activity relative to the intein modified cellulase. It is noteworthy that any native protein will meet the requirement for this invention, and selection of the protein is dependent upon the plant's intended purpose. In this case, a poplar species that could be induced to

COR-AAD1-00002391

de-polymerize its own cellulose would be beneficial for ethanol production from biomass, or as a substrate for fermentation of other chemicals.

[0065] Contruction of CIVPS or Intein Modified Proteins

[0066] Various recombinant DNA techniques may be used in combination to construct the vector carrying the DNA encoding the modified protein. One of the easiest and most direct utilizes the polymerase chain reaction (PCR) to assemble the nucleic acid sequence encoding the intein-modified protein with appropriate complementary ends that facilitates ligation into the desired vector. The PCR method is illustrated here. Other methods may be used to accomplish this same goal, and some rely on specific restriction and ligation of the desired protein and intein encoding sequences, but may still include PCR steps. PCR Kits for conducting the reaction are readily available (Epicentre, Madison, WI). The only requirements on the primers is that one should match the 5' end of the sense strain to be amplified, and the other should match the 5' end of the corresponding antisense strain; relative sequence uniqueness is beneficial.

[0067] Clean-up and Purification from a Gel

[0068] The purification of DNA from a gel may be accomplished using electroelution, phenol extraction, agarase digestion, glass bead extraction, or from a number of commercially available kits. The commercially available QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and associated method is one example.

[0069] Selection of Intein According to Intended Use

[0070] Two features are of importance in this step: the property the CIVPC or intein possesses to induce splicing that will facilitate optimization of the transgenic plant for its intended purpose, and where to place the intein within the nucleic acid sequence encoding the target protein. Any coding sequence for a self-splicing protein, i.e. an intein, may be used in this invention. A compilation of some known inteins is given on the following website http://blocks.fhcrc.org/~pietro/inteins/. Other inteins remain to be discovered and new inteins may be created through sequence analysis, recombinant DNA

COR-AAD1-00002392

methods, and mutation of known sequences. This intein of Example 1 is advantageous for the intended transgenic poplar species because upon splicing it yields predominantly ligated, native protein (>75%), and is temperature sensitive so that intein splicing is inhibited at temperatures less than 30°C, and is not substantial until 50 °C, at which temperature the half-life of the uncleaved protein is less than 2 hours.

[0071] Construction of Intein Modified Protein

[0072] In order to ensure proper intein splicing, the intein is inserted in Example 1 in frame next to a serine, cysteine, or threonine residue of the native target protein. This leaves the native target protein's serine, cysteine, or threonine on the carboxylic acid side, of this intein's histidine – asparagine, conserved residues at the terminal 536 and 537 intein amino acid positions, respectively. Other terminal residues may be used, depending upon the desired stimulus and mechanism for intein splicing. If desired, the codons at the extein-intein junction may be altered to facilitate these requirements. Care is advised when altering the junction codons, so that the intein modified protein may cleave as desired, and allow the resulting products to perform the appropriate activity. The intein position within the native target protein such is to substantially change the activity of the resulting intein modified protein. In most circumstances virtually any interruption within or near the active site of the molecule meets this criterion. The combination of the amplified intein sequence and the amplified native protein sequence is easily accomplished if a serine residue resides close to a unique restriction site of the native protein's coding sequence. Conversely, the intein coding sequence is readily incorporated at any desired position in the native protein sequence by using several polymerase chain reactions. A preferred PCR method is set forth here. Preferrably 50 oligonucleotide primers are used. Shorter primers may be used, however it is beneficial, although not necessary, to use primers of the same length. The sense primer of the C-extein may hybridize to both the C-extein and intein sequence at the junction to facilitate the fusion of the amplified sequences in subsequent PCR

COR-AAD1-00002393

amplifications. For intein amplification, both primers preferably overlap with their respective desired adjacent extein sequences to facilitate fusion of the intein sequence and extein sequences in subsequent PCR amplifications. The polymerase chain reaction is preferably carried out using the standard protocol outlined above, but may have some optimization. Typical optimization parameters are the amount of template and primer DNA added to the mixture (generally the primer DNA is added in great excess relative to the template DNA), the temperatures and times for the reaction cycles, the number of cycles, and the $MgCl_2$ concentration. The length and composition of the primers used may also be varied to yield an effective intein modified protein, so long as the constraints on placement are observed. Kits are commercially available which include all necessary reagents: *Taq* DNA polymerase, $MgCl_2$, 25 mM dNTP mixture (containing equamolar amounts of dATP, dCTP, dGTP, and dTTP), reaction buffer, and water.

[0073] At this point the next round of PCR is started to fuse the extein and intein sequences. In this case the intein fragment is preferably mixed with an equimolar portion of the C-extein cellulase fragment. Combination of these fragments represents both the template and primers (overlapping regions) to be used. Addition of reaction buffer, 25 mM dNTPs, $MgCl_2$, and *Taq* DNA polymerase is still required, as are the changing temperature cycles. This reaction is preferably augmented by addition of the following sense and anti-sense primers, respectively, along with *E. coli* DNA ligase (New England Biolabs, Beverly, MA), however this addition is not necessary and depending upon the exact reaction conditions employed may not lead to an increase in the yield.

5'-ACAGAATGGGGAACGAGCGATGCTAGCATTTTACCGGAAGAATGGGTTC-3' [SEQ ID NO: 1]

5'-CGTGTCTGCTCCGTTTACCGCTTTTTTTAATTGGACGAATTTGTGCGTGA-3' [SEQ ID NO: 2]

COR-AAD1-00002394

[0074] Once completed, the PCR products are preferably again run on an agarose gel, and the appropriate band, 2665 nucleotides long, purified from the gel and analyzed according to the methods described above. A small amount of the purified reaction product is preferably used for quantitation by measuring the absorbance at 260 nm and 280 nm wavelengths on a UV spectrophotometer. To complete assembly a PCR reaction of the intein modified cellulase coding sequence is carried out combining equimolar amounts of the fused C-extein and intein fragments just constructed, with the N-extein fragment purified previously. The PCR reaction is preferably conducted using the same temperature cycles as in the previous reaction after addition of reaction buffer, 25 mM dNTPs, $MgCl_2$, and *Taq* DNA polymerase. This reaction is preferably augmented by addition of the following sense and anti-sense primers, and *E. coli* DNA ligase (New England Biolabs, Beverly, MA); however this addition is not necessary and depending upon the exact reaction conditions employed may not lead to an increase in the yield.

5'-AGCATTCAGACCTCCCATTTCATACGAAAAGAGGAAATAGATAGATTTTC-3' [SEQ ID NO: 3]

5'-CGTGTCTGCTCCGTTTACCGCTTTTTTTAATTGGACGAATTTGTGCGTGA-3' [SEQ ID NO: 4]

[0075] Vector Construction

[0076] Other elements may be included in the expression cassette prepared in Example 1, e.g. extracellular secretion signaling sequences, intracellular localization signaling sequences, other inducible promoters, etc. As the vector is now contained within the recombinant strain *A. tumefaciens*, the gene transfer to the poplar plant relies on the bacteria's specialized delivery system. Other gene transfer methods are available, and selection of a suitable transformation method depends upon the source of the plant material. For example, protoplasts or individual plant cells may be transformed directly with the recombinant pTiBo542 plasmid using electroploration, calcium chloride, or Biolistic particle bombardment (Bio-Rad, Hercules, CA). Conversely plant callus, plant segments,

COR-AAD1-00002395

or in some cases, whole plants may be used as starting material, when appropriate. For efficient gene transfer to occur, the time of incubation and cell density of the culture is preferably optimized.

[0077] Advantages and Uses for Transgenic Poplar of Example 1

[0078] The resulting transgenic poplar species may be grown and passaged indefinitely while producing the intein modified cellulase in high titer. The cellulase may be subsequently activated by harvesting the plant, mechanically chipping or grinding it to increase the exposed surface area, and then incubating the resulting mash in a vat or tank at an elevated temperature (preferably 30°C to 50°C) and/or lower pH (6.5 or below). Exposure to the elevated temperature, and lower pH, if used, will induce the intein splicing and yield the native cellulase at a substantially increased activity. Under these conditions the cellulase may now catalyze the cleavage reaction of cellulose to economically produce substrates that may be subsequently fermented into ethanol or other chemical entities. In addition, this plant may be used as a source of either the intein modified cellulase, or the recovered native cellulase, post splicing. In either case, the protein is preferably purified from the plant using methods well known in the prior art, such as precipitation, membrane filtration, chromatography including affinity chromatography, extraction, etc.

[0079] The use of transgenic plants producing intein modified proteins has two advantages over previously reported transgenic plants. Because the intein modified protein has substantially less activity than the native protein, it may be expressed at higher titer and localized anywhere in the plant species. Previous reports of transgenic plants expressing cellulase enzymes have taught elimination of the secretion signals to contain the cellulase enzymes in the cytosol of cells. This is not necessary with the use of intein modified proteins and is a substantial improvement as the modified protein may be placed in close proximity with its substrate, but not catalyze the reaction until desired. In addition, these plants have a higher degree of environmental safety. Because the genes transferred encode proteins of substantially less activity under

COR-AAD1-00002396

physiological conditions, horizontal gene transfer between species is less likely to impart any selective advantage. For this reason it is unlikely that either the transgenic plants would outperform native plants in the wild, or that gene transfer would yield a selective advantage favoring a transformed population.

[0080] Example 2 demonstrates the broad applications of this invention. Example 2 shows a variation of the method of Example 1 to generate a transgenic Douglas-fir species expressing an intein modified lignin peroxidase. The choice of a specific target protein depends upon the application intended for the transgenic plant species. For this example, a lignin peroxidase gene that facilitates the catalytic breakdown of lignin, a chemical component of wood was selected. By using an intein modified lignin peroxidase, the whole plant may be regenerated while the inactivated intein modified lignin peroxidase is produced throughout the plant, at high titer if desirable. The enzyme may be subsequently activated by the self-splicing ability of the intein to yield the native lignin peroxidase at increased activity than the intein modified lignin peroxidase. This allows improved control of the lignin peroxidase activity that is not currently available. Such a transgenic plant species is valuable for the production of pulp, animal feeds, substrates for other processes, improvements on mechanical pulping, biobleaching of pulp, improvement from decreased pulp processing wastes, and the production of biopolymers with unique properties.

[0081] Construction of Gene & Intein Modified Protein

[0082] As indicated above, any native protein is suitable as the target protein, and its selection is dependent upon the plant's intended purpose. For this example, a Douglas-fir species that may modify its own lignin is beneficial as a substrate for different pulping processes. The protein encoding nucleic acid of interest may be isolated from *Phanerochaete chrysosporium* (GenBank Accession # M37701). One primer preferably matches the 5' end of the sense strain to be amplified, and the other the 5' end of the complementing DNA strand at the end of the gene,It is beneficial to have relative sequence uniqueness.

[0083] Purification of PCR Fragments from Gel

COR-AAD1-00002397

[0084] The purification of the nucleic acid from the gel is accomplished using electroelution, phenol extraction, agarase digestion, glass bead extraction, or from a number of commercially available kits. Preferably the commercially available QIAquick Gel Extraction Kit, available from Qiagen (Valencia, CA) is used.

[0085] Intein Selection

[0086] The choice of intein is very dependent upon both the intended purpose of the plant and the intein modified protein. Many different inteins exist and may be used. For this example an intein with the same properties as in Example 1 is beneficial for the intended use of a transgenic Douglas-fir species. Hence, a variant of the Psp pol intein (GenBank Accession # PSU00707) from Pyrrococcus spp. is preferably used. The advantage of this intein is that upon splicing it yields predominantly ligated, native protein (>75%), and is temperature sensitive so that intein splicing is inhibited at temperatures less than 30°C, and is not substantial until 50°C, where the half-life of the uncleaved protein is less than 2 hours. This intein induces splicing *in vitro* by a pH shift, thus adding increased flexibility to subsequent processing of the transgenic plant.

[0087] Vector Transformation

[0088] With the vector contained within the recombinant strain of *A. tumefaciens*, gene transfer to Douglas-fir relies on the bacteria's specialized delivery system. Other gene transfer methods are available, and selection of a suitable transformation method depends upon the source of the plant material and ease with which the method can be applied. Some modification and optimization of the transformation parameters is usually necessary.

[0089] Uses of the Recombinant Trees

[0090] The tree of Example 2 may be used as a source material for the purification of lignin peroxidase or intein-modified lignin peroxidase. Alternatively, it may be used also by itself as a substrate for producing wood pulp in any number of applications, e.g. paper production, animal feed, composite materials, etc. Both Example 1 and Example 2 have illustrated the use of trees,

COR-AAD1-00002398

certainly other plants are useful options and depend upon the intended use of the invention. In many areas these types of trees do not grow well and grasses, vines, seaweed, or other plant species do, and may be used equally well. In addition, many fruits and vegetables may benefit from intein modified protein technology, such as for example to induce ripening, pesticide resistance, or any number of other applications. Hence the choice of host plant is not limiting. The use of plants as sources for recombinant proteins is facilitated by use of the CIVPS or intein technology of this invention. Plants are made to express any number of fusion proteins where the fusion point is comprised of an intein that does not facilitate recombination of fused protein exteins, but instead links the desired protein to a binding protein for purification via affinity chromatography. In this case the desired protein may or may not have full activity *in vivo*. Once expressed in the plant, the fusion protein is eluted onto an affinity column where the binding portion of the fusion protein binds the column. The column is then treated to induce intein splicing and the desired protein is washed away and recovered. Another variation of the invention that is of medical interest is a fusion protein comprising a therapeutic protein or vaccine fused by inteins to protect protein groups. Such a therapeutic protein can be expressed in a plant and eaten by a human and non-human animal, e. g. in the case of animal vaccination or hormonal treatment. Intein splicing then occurs inside the patient, or animal, relying on the change of pH within the stomach, or a thiol gradient induced by ingestion of a third chemical. Splicing removes the protective protein groups, yielding the native therapeutic protein or vaccine, which is then absorbed in the gut.

[0091] Either of the transgenic trees expressing intein modified proteins from Examples 1 and 2 may be effectively used in an industrial scale process as is shown in Example 3. The pulping itself may be enhanced by a modification similar to that used in Example 2 for the Douglas-fir species.

[0092] Tree Processing

COR-AAD1-00002399

[0093] Typical pretreatment processes for the degradation of lignocellulosic substrates include concentrated acid pretreatment (usually Sulfuric Acid), dilute acid pretreatment, ammonia explosion pretreatment, and hot water pretreatment. Other pretreatment processes are possible, and design of the transgenic tree expressing an intein modified protein should be optimized to take full advantage of a pretreatment process when necessary. Intein splicing may occur in a vessel via any known method, such as, but not limited to: pH shift, temperature change, light exposure, acoustic stimulation, or any exogeneous chemical addition.

[0094] Intended Uses and Process Variations

[0095] Preferred variations of the process of Example 3 include combining the pretreatment, splicing, digestion, and fermentation steps. This preferred processing consolidation may occur between any of the steps, however a preferred manifestation incorporates all steps simultaneously in a single unit operation. This preferred combination may realize cost savings through a decrease in capital expenditure and depreciation, decrease in the cost of substrates, and a process dependent decrease in the cost of energy and chemicals input to the process. In addition, as opposed to competing chemical processes making the same products, environmental benefits may be realized through decreased emissions and hazardous waste generation. In Example 3 the choice of product is dependent upon the organism used in the fermentation for the desired bioconversion. Any organism that may adequately utilize the degraded cellulose as a substrate may be used to effectively produce a desired product. For this reason the spectrum of end products that may be made is very large. Applications that will benefit from substrates with preferred processing traits facilitated by the intein modified proteins carrying plant of this invention include, but are not limited to, fuel production, chemicals production, textiles production, biopolymer production, food production, and saccharification. Although Example 3 is mostly focused on the fermentation of the degraded transgenic plants, intein modified plants may also be used as substrates for traditional chemical processes. For example, the

COR-AAD1-00002400

plants of Example 2 may be preferentially used in paper pulping. In such a process, benefits are derived from a decrease in the harsh chemicals used to bleach the wood. This will likely result in a decrease in the costs of chemical input, hazardous material generation and containment, and potentially some consolidation in processing. Another use is that of pectinase for cotton scouring, or cellulases for other textile production processes. In these instances, the end products are derived from more traditional chemical processes, although benefits accrue through the use of intein modified protein plant substrates, as opposed to the normally harsh chemical processing environment generally employed.

[0096] Animal feed is commonly supplemented with a variety of enzymes used to increase the nutritional value of the feed, as well as decrease the environmental burden experienced in proximities where animal manure accumulation is substantial. Nutritional value is increased through the putative enzyme action on plant polymers, which assist the animal in digesting the feed and thereby utilizing more of the beneficial feed components. The environmental burden may be decreased by limiting the amounts of added minerals, such as inorganic phosphate, which may be obtained from the plants themselves in the presence of the active enzyme. The benefits associated with using intein modified proteins, as opposed to unmodified proteins, result in multi-protein expression, at high levels, which do not interfere with plant regeneration yet impart a desired enzyme activity upon splicing within the animals stomach. This decreases the cost of feed by delivering the enzymes within the meal itself, as opposed to their being produced exogenously and added to the meal. In addition, the added benefit of using genes that code for nearly inactive proteins *in vivo* in plants, provides a technology platform that is less likely to be associated with environmental risks associated with horizontal gene transfer to native plant species. This advantageous environmental affect, whether real or perceived, holds for all intein modified protein plant products. Example 4 illustrates the construction an intein modified phytase in rapeseed, for use as animal feed.

[0097] Uses and Variations

COR-AAD1-00002401

[0098] Phytase is an enzyme that assists in the evolution of inorganic phosphorous from myoinositol phosphates contained inherently in animal food. An economic impact is brought about through a decrease in the amount of phosphate supplementation required for the production of animal feed, and a decrease in the phosphate content of the animal's manure, which contributes to the contamination of local waters. Although Example 4 below illustrates the construction and use of an intein modified phytase expressed in rapeseed for animal feed, a number of other valuable native proteins may be used as well. For example, phytase may be substituted with, or used in addition to any number of cellulases, amylases, glucanases, hemi-cellulases, pectinases, proteases, xylanases, lipases, growth hormones, or immunogenic antigens, among others. Each of these other proteins has a potential economic value in the use of animal feed supplementation.

[0099] Example 5 illustrates one of the preferred embodiments of the invention. A transgenic corn is constructed and used as a substrate for ethanol processing. In this case the intein modified gene sequence of Example 1 is again used for demonstration purposes only. In a preferred embodiment, however, several intein modified proteins may be expressed simultaneously to optimize the desired plant degradation processing trait for use in the fermentation process. The target enzymes may be selected from enzymes commonly known as cellulases (E.C. 3.2.1.4), exocellobiohydrolases (E.C. 3.2.1.91), glucosidases (E.C. 3.2.1.21), and may be expressed optimally with other enzymes selected from the Enzyme Classification heading 3.2.1.x, or any other classification group necessary. In addition to the simultaneous expression of multiple intein modified proteins, the preferred composition of matter embodiment is a fertile plant capable of reproduction and stable gene inheritance.

[00100] Transformation Information

[00101] The macroprojectiles are used to accelerate the microprojectiles, which enter the plant cells.

COR-AAD1-00002402

[00102] Having now generally described this invention, the same will be better understood by reference to certain specific examples, which are included herein for purposes of illustration only, and are not intended to be limiting of the invention or any embodiment thereof, unless and where it is so specified.

[00103] EXAMPLES

[00104] Example 1: Production of Transgenic Poplar Expressing an Intein-Modified Cellulase

[00105] For this example a cellulase enzyme is used. A vector is first assembled containing the DNA coding sequence for the intein modified protein. In order to construct such a vector, an intein modified protein DNA sequence is first prepared, and then packaged into the desired vector. The desired protein for this plant is a cellulase (GenBank Accession # AY039744) isolated from Bacillus sp. NBL420. The gene corresponding to this protein is amplified using PCR from a genomic DNA template isolated from the Bacillus sp. NBL420. The PCR reaction is performed by mixing the template DNA, two primers complimentary to the 3'ends of the template DNA to be amplified, Taq DNA polyermase, reaction buffer (500 mM KCl, 100 mM Tris-Cl pH 9.0, 0.1% Triton X-100), and $MgCl_2$ in a thin-walled 250 L PCR tube. Once mixed, each reaction tube is placed in a thermocycler, and the thermocylcer is set for 35 cycles comprised of three segments: 94 °C for 30 seconds, 60 °C for 60 seconds, 72 °C for 120 seconds. Following amplification, the resulting PCR product is analyzed by electrophoresis on a 1% agarose gel, along with molecular weight standards (Invitrogen, Carlsbad, CA), with the aid of 1X TAE (or TBE) running buffer, and stained with ethidium bromide (0.5 g/ mL). Care should be taken to ensure that the appropriately sized band, of approximately 3200 base pairs (bp), has been obtained. This band is then cut out from the gel with a scalpel, and purified (separated from the gel material) using a commerically available gel purification kit (Qiagen). Once the fragment has been purified from the gel, the band is analyzed using restriction digestion or sequencing as described by Ausbel et. al., Current Protocols in Molecular Biology, Wiley, New York (1998). After gene

COR-AAD1-00002403

amplification, the gene is modified by insertion of the intein sequence segment. In this case, a variant of the Psp pol intein (GenBank Accession # PSU00707) from Pyrrococcus spp. is used This variant, described in the literature , contains a mutation at the tyrosine 534 residue which converts that tyrosine to methionine. See, Xu, M, Perler, F, (1996), The mechanism of protein splicing and its modulation by mutation, The EMBO Journal 15:5146-5153. This intein may be cleaved in vitro by a pH shift. The coding sequence of this intein is then amplified by PCR using genomic DNA from Pyrrococcus spp, as a template. The PCR reaction is conducted using a standard protocol (e.g., 30 cycles comprised of a 94 °C for 30 seconds, 50 °C for 60 seconds, 72 °C for 120 seconds) and the following primers.

[00106] 5'-ATTATGTGCATAGAGGAATCCAAAG-3' [SEQ ID NO: 5]

[00107] 5'-AGCATTTTACCGGAAGAATGGGTTC-3' [SEQ ID NO: 6]

[00108] Once the amplification is complete, the PCR product is transferred to and elecrophoresed on a 1% agarose gel in 1X TAE or TBE buffer. The resulting band is then purified and analyzed as described above for the native cellulase coding sequence. At this point the two PCR fragments shown in Figure 1, one encoding the cellulase protein and one encoding the intein polypeptide sequence, are joined. Here, the intein is inserted in frame into the native protein, such that a serine residue of the native protein becomes the terminal C-extein amino acid at the junction point between the native intein and C-extein of the native protein. This intein modified protein segment is produced using PCR by first amplifying the C-extein coding sequence of the cellulase gene. Primers that overlap both the C-extein, and the intein end containing the histidine and asparagine codons immediately adjacent to the C-extein are used to amplify the C-extein sequence:

5'-CTTTGGATTCCTCTATGCACATAATTCCGGAAACGGCGGTGTCTACCTCG-3' [SEQ ID NO: 7]

5'-CGTGTCTGCTCCGTTTACCGCTTTTTTTAATTGGACGAATTTGTGCGTGA-3' [SEQ ID NO: 8]

COR-AAD1-00002404

[00109] The resulting sequence is 604 nucleotides long. The intein is then amplified using a sense primer that contains both the intein end containing the terminal serine codon, and the N-extein end of the cellulase gene, along with an antisense primer that contains specific nucleotides of the intein and C-extein. For this PCR reaction the following primers are used to obtain a sequence 1661 nucleotides long:

5'-ACAGAATGGGGAACGAGCGATGCTAGCATTTTACCGGAAGAATGGGTTC-3' [SEQ ID NO: 9]

5'-CGAGGTAGACACCGCCGTTTCCGGAATTATGTGCATAGAGGAATCCAAAG-3' [SEQ ID NO: 10]

[00110] The N-extein is then amplified using PCR and one primer that contains specific nucleotides of the sense N-extein strand, and another primer that contains specific nucleotides of the N-extein and adjacent intein sequence. The N-extein portion of the cellulase gene is amplified with the following primers resulting in a sequence 1541 nucleotides long.

5'- AGCATTCAGACCTCCCATTTCATACGAAAAGAGGAAATAGATAGATTTTC-3' [SEQ ID NO: 11]

5'-GAACCCATTCTTCCGGTAAAATGCTAGCATCGCTCGTTCCCCATTCTGTG-3' [SEQ ID NO: 12]

[00111] Once these three reactions are complete, each PCR fragment is cleaned to remove residual primers, and the C-extein, intein, and N-extein PCR fragments are joined by conducting two more polymerase chain reactions. The intein and one of the cellulase extein regions, either the C-extein or the N-extein, are amplified in a single reaction by mixing equimolar portions of the two PCR fragments generated above and performing PCR as described earlier. This reaction requires no extra external primers and results in the first intein-extein fusion. This reaction mixture is cleaned, and then equimolar portions of the cleaned fusion product are mixed with the remaining extein portion, and PCR is conducted once again without adding additional primers. No exogeneous primer is required in either of the last two PCR reactions, and annealing occurs at the

COR-AAD1-00002405

intein-extein junctions. The annealed regions are extended by Taq polymerase resulting in the final fusion products. This sequence of reactions results in the coding sequence of the intein modified protein with the intein inserted at the exact position desired. The product of the final reaction is cleaned again, and amplified using PCR one last time with primers specific to the cellulase extein termini with specific ends to facilitate ligation into the cloning vector. Once this reaction is complete, the PCR products are run on an agarose gel, and the appropriate band, 3806 nucleotides long, is purified from the gel and analyzed according to the methods described above. The resulting intein modified protein coding sequence (nucleic acid segment) contains a ribosome binding site, a start codon at the beginning of the N-extein, the complete sequence of the intein modified cellulase with the intein inserted in frame in the proper orientation, and a stop codon at the end of the C-extein coding sequence. The intein modified cellulase coding sequence is then cloned into pTiBo542, replacing the tms and tmr genes in the T DNA, using the methods described in Ausbel, et. al., Current Protocols in Molecular Biology (1998). See, Parsons TJ, Sinkar, VP, Stettler, RF, Nester, EW, Gordon, MP, "Transformation of Poplar by Agrobacterium tumefaciens," Biotechnology 4:533-536, 1986. Here the expression cassette includes a "MAC" promoter, a mannopine synthetase terminator, and a kanamycin resistance marker. This vector is transformed into A. tumefaciens A281 using any suitable method known in the art (e.g., electroploration, or the calcium chloride method). Various transformation methods are also described by Ausbel, et. al. (1998), above.

[00112] To transform the desired Populas trichocarpa x deltoides, H11 plant species, with the recombinant A. tumefaciens, a variation of the leaf disk method is employed. The recombinant A. tumefaciens is cultured in selective medium containing 50% MG medium (10 g/ L mannitol, 2.32 g/ L sodium glutamate, 0.5 g/ L KH_2PO_4, 0.2 g/ L NaCl, 0.2 g/ L $MgSO_4$-$7H_2O$, 0.002 g/ L biotin, pH 7.0), 50% luria broth (20 g/ L tryptone, 10 g/ L yeast extract, and 10 g/ L NaCl), and appropriate antibiotic, at 30 °C in an incubator-shaker. For plant

COR-AAD1-00002406

transformation, small greenwood stem sections, approximately 7 mm in length and 2 - 3 mm in diameter, are sterilized with a 20% bleach, 0.1% Tween 20, and 30 mg/ L Benomyl systemic fungicide (Chas. H. Lilly Co., Portland, OR) solution. After washing with sterile water, the stem sections are aseptically transferred to a culture of A. tumefaciens at a cell concentration of approximately 5 x 10^8 cells per mL, and the sections allowed to incubate for 16 hours. After exposure to the recombinant A. tumefaciens culture, the plant stems are transferred to solid Murashige-Skoog medium supplemented with zeatin riboside and kanamycin in a vertical position. See, Murashige T, Skoog F, "A revised medium for rapid growth and bioassays with tobacco tissue cultures," Physiol. Plant, 15:473-497, 1962. Once roots have begun to grow, shoots will develop. The regenerating plants are transferred to fresh plates every two to three weeks, and a normal light cycle is maintained during plant growth and at elevated humidity in the incubator. Once roots form, the explants are transferred to a solid medium lacking zeatin riboside, but containing kanamycin for another two to three weeks, after which period the plants are transferred to boxes containing soil for four to five days prior to replanting in soil and full growth in a greenhouse or controlled plot of soil. Initial plants are screened by several methods to ensure the intein modified cellulase DNA sequence has been transferred to the genome and protein expression is active. Genetic screening is conducted by southern analysis on genomic DNA isolated from the transgenic plant using the intein modified cellulase coding sequence as a probe, as described by Ausbel, et. al. (1998), above. PCR is conducted using probes specific to the intein modified cellulase coding sequence and the transgenic plant's genomic DNA as a template, as described above. Appearance of the appropriately sized band on an ethidium bromide stained gel verifies the presence of the intein modified cellulase coding sequence. Direct sequencing of the plant's genomic DNA may also be performed. Protein production is monitored by western analysis using antibodies specific to both the intein modified cellulase and the native cellulase. In addition, enzymatic assays

COR-AAD1-00002407

for cellulase activity are known in the art and may be used to quantify the activity of the unspliced intein modified cellulase and the spliced cellulase.

[00113] Example 2: Production of Transgenic Douglase Fir

[00114] Expressing Intein-Modified Lignin Peroxidase

[00115] This example uses the same method for constructing the vector containing the intein modified lignin peroxidase coding sequence as used in example one. The primary differences are in the A. tumefaciens plasmid employed, the native protein sequence that is modified, and the primers selected to amplify the new intein modified lignin peroxidase coding sequence.

[00116] The lignin peroxidase gene (GenBank Accession # M37701) is amplified by PCR using genomic DNA from P. chrysosporium as a template. The primers

5'-ATGGCCTTCAAGCAGCTCGTCGCAG-3' [SEQ ID NO: 13]

5'-TTAAGCACCCGGCGGCGGGGGGCTG-3' [SEQ ID NO: 14]

are are used in the PCR reaction as described in example one. Following amplification, the resulting PCR product is analyzed using gel electrophoresis on an agarose gel, along with molecular weight standards as described in example one. After staining the gel with ethidium bromide, the 1567 base pair (bp) band is cut from the gel with a scalpel, and purified from the gel as described above. After purifying the fragment from the gel, the fragment is analyzed using restriction digestion or sequencing for direct verification as described by Ausbel, et al., 1998.

[00117] After the gene is amplified, it is modified by insertion of the intein sequence into the gene sequence. For this example, the same intein is used as in example one. The coding sequence of this intein is amplified in the same manner as described in example one. The resulting intein DNA sequence is purified by gel electrophoresis and analyzed as described previously.

[00118] The two PCR fragments, one encoding the lignin peroxidase and one encoding the intein polypeptide sequence, are joined. To ensure proper intein splicing, the intein is inserted in frame next to a serine residue of the lignin

COR-AAD1-00002408

peroxidase such that this serine is on the carboxylic acid side, of this intein's histidine - asparagine conserved residues at the terminal 536 and 537 intein amino acid positions, respectively. The intein is inserted into the native protein, such that the serine residue of the native protein becomes the terminal C-extein amino acid at the junction point between the native intein and C-extein of the native protein. The intein is positioned within the native protein such that its presence substantially reduces the activity of the resulting intein modified protein. In most circumstances virtually any serine residue within or near the active site of the molecule will meet this criterion, however some optimization may be necessary.

[00119] The intein modified protein sequence is produced using PCR the same as described in example one, with the only difference being the choice of primers. The C-extein portion of the lignin peroxidase gene is amplified using the cleaned gene product from the PCR reaction above, and the following primers resulting in a 445 nucleotide sequence:

5'-CTTTGGATTCCTCTATGCACATAATTCTCGCCCGCGACTCCCGCACCGCT-3' [SEQ ID NO: 15]

5'-TAAGCACCCGGCGGCGGGGGGCTGGAAGAGGAATATGTCAGCTGGGGGC-3' [SEQ ID NO: 16]

[00120] The N-extein portion of the lignin peroxidase gene is amplified using the same template, by PCR using the following primers.

5'- ATGGCCTTCAAGCAGCTCGTCGCAGCGATTTCCCTCGCACTCTCGCTCAC-3' [SEQ ID NO: 17]

5'-GAACCCATTCTTCCGGTAAAATGCTGTGTGGTCGGTCTGGATGCGGATCT-3' [SEQ ID NO: 18]

[00121] The resulting sequence is 1196 nucleotides long. The intein coding sequence to be placed into the lignin peroxidase gene is amplified using PCR as described in example one. In this reaction use a Pyrrococcus spp genomic DNA template and the following primers:

COR-AAD1-00002409

5'-AGATCCGCATCCAGACCGACCACACAGCATTTTACCGGAAGAATGGGTTC-3' [SEQ ID NO: 19]

5'-GCGGTGCGGGAGTCGCGGGCGAGAATTATGTGCATAGAGGAATCCAAAG-3' [SEQ ID NO: 20]

[00122] The resulting sequence is 1661 nucleotides long. Once these reactions are complete, the reaction products are electrophoresed on an agarose gel, purified from the gel, and analyzed as described above. The extein and intein portions are joined as described in example one. In this case the intein fragment is mixed with an equamolar portion of the C-extein lignin peroxidase fragment. Combination of these fragments represents both the template and primers required for the PCR reaction. PCR is performed using the same reaction conditions as in example one. Once complete, the PCR products are electrophoresed on a 1% agarose gel, and the appropriate band, 2106 nucleotides long, is purified from the gel. The purified band is analyzed as described in example one. A small amount of the purified reaction product is then quantified by measuring the absorbance at 260 nm and 280 nm on a UV spectrophotometer.

[00123] The intein modified protein DNA coding sequence is completed with one more PCR reaction. Equamolar amounts of the fused C-extein and intein fragment just constructed are combined with the N-extein fragment purified previously. The PCR reaction is conducted using the same conditions in the previous reactions. The reaction products are electrophoresed on a 1% agarose gel, the appropriate band, 3302 nucleotides long, is purified from the gel, and analyzed according to the methods described in example one. The final intein modified protein coding sequence has the complete intein sequence in frame, in the proper orientation, within the lignin peroxidase coding sequence.

[00124] The intein modified lignin peroxidase coding sequence is cloned into a plant expression cassette. In this case, the pTiA6 plasmid is used with kanamycin resistance and lacking the octupine synthetase genes, but containing the octupine transcription control sequences. The intein modified lignin peroxidase is ligated into a restricted pTiA6 under the octupine transcription

COR-AAD1-00002410

control sequences (promoter and 3' polyadenylation site). A. tumefaciens K12X562 is transformed using the resulting ligated vector, and any suitable method known in the art (e.g., electroploration, or the calcium chloride method). Transformation methods are described by Ausbel, et. al. (1998).

[00125] Douglas-fir is transformed, with the recombinant A. tumefaciens, and nodal segments or seeds sampled from these trees. The shoot multiplication and elongation is conducted as previously described (Gupta PK, Durzan, DJ, "Shoot multiplication from mature trees of Douglas-fir, and sugar pine,"Plant Cell Reports, 9:177-179, 1985) in culture on DCR basal medium plates. A culture of the recombinant A. tumefaciens is grown according to the method described in example one. For plant transformation, the regenerated shoots from culture, approximately 50 mm in length, or seeds are surface sterilized by treatment with a 10% bleach and 0.1% Tween 20. Once sterilized, the shoots or seeds are aseptically rinsed with sterile, distilled, and deionized water. The seeds or the shoots are transformed by first wounding the epidermal tissue with a sterile needle or by cutting the surface with a sterile scalpel. The wounded shoots or seeds are soaked in a culture of the recombinant A. tumefaciens at a cell concentration of approximately 5 x 10^8 cells per mL. After a 12 hour exposure to the recombinant A. tumefaciens culture, the shoots and seeds are cultured in DCR basal medium with 2.2% sucrose and 0.8% Bacto (Difco) agar. The culture conditions include a 16 hour light cycle at 25 °C, followed by and 8 hour dark cycle at 20°C in a green house or growth chamber. The regenerating plants are transferred to fresh plates every two to three weeks. Once roots form, the explants are transferred to boxes containing soil for four to five days prior to replanting in soil and full growth in a greenhouse or controlled plot of soil. The first year of growth is conducted within a green house under controlled temperature conditions, not exceeding 30 °C.

[00126] The plants are screened using methods similar to those of example one, except specific to the lignin peroxidase protein or intein modified lignin peroxidase protein in the case of western analysis.

COR-AAD1-00002411

[00127] The resulting transgenic Douglas-fir species is grown indefinitely while producing the intein modified lignin peroxidase in high titer. The lignin peroxidase is subsequently activated using the same methods described in example one because the same intein was employed for modification in this example.

[00128] Example 3: Fermentation Substrate Preparation Process

[00129] Using Plants Expressing Intein Modified Protein

[00130] In the case of example one, the transgenic poplar species can be used as substrate for ethanol production via fermentation. For this process the transgenic tree is harvested using a suitable tool, such as a chain saw or ax. The tree is subsequently pulped using a mechanical pulper. The pulp is then placed it in a tank. After any necessary pretreatment has been conducted, intein splicing is induced by raising the temperature of the tank and reducing the pH to a value of 4. Depending on the pretreatment used, intein splicing may be stimulated by the pretreatment and thereby occur in parallel with that process operation. Once spliced the native enzyme activity begins digesting the cellulose of the pulp, increasing the concentration of monosaccharides.

[00131] Following the induction of splicing, the contents of the saccharification vessel are mixed in any proportion with native poplar pulp or other substrates, to facilitate cellulose degradation of those substrates. The proportion of the mixing depends upon the cellulase activity of the transgenic poplar which is a function of the amount of intein modified cellulase expressed in the plant, the efficiency of splicing, the efficiency of recombination, and the activity of the recombined, native cellulase on the substrate. Each one of those parameters has a broad spectrum of possible values and can be optimized to facilitate process economics.

[00132] The resulting glucose is then filter sterilized from the degraded cellulose through a 0.22 (or less) m filter, or heat sterilized in batch or continuous mode through a heat exchanger. The sterilized glucose is fed to a fermentation process, where it can be used as a substrate for ethanol production

COR-AAD1-00002412

as described in the literature. See, H.K. Sreenath and T. W. Jeffries, "Production of ethanol from wood hydrolysate by yeasts," Bioresource Technology, 72(3): 253-260, 2000; Lisbeth Olsson and Barbel Hahn-Hagerdal, "Fermentation of lignocellulosic hydrolysates for ethanol production," Enzyme and Microbial Technology, 18(5):312-331, 1996; Kutluo O. Ulgen, et. al., "Bioconversion of starch into ethanol by a recombinant Saccharomyces cerevisiae strain YPG-AB", Process Biochemistry, 37(10):1157-1168, 2002; M. Mete Altintas, et al, "Improvement of ethanol production from starch by recombinant yeast through manipulation of environmental factors," Enzyme and Microbial Technology, 31(5):640-647, 2002; Farooq Latif, et al., "Production of ethanol and xylitol from corn cobs by yeasts," Bioresource Technology, 77(1):57-63, 2001.

[00133] The fermentation process is conducted in batch, fed-batch, or continuous modes.

[00134] Example 4: Plants Expressing an Intein Modified Protein used for Animal Feed

[00135] A transgenic rapeseed is constructed following essentially the same methods described in Examples 1 and 2 above, with the following modifications. In constructing the CIVPS or intein modified gene sequence, the same intein coding sequence can be used, however in this case it is fused within the phytase expressed by Aspergillus ficuum. In this example the selected intein modified protein relies upon the acidity of the animal's stomach to induce protein splicing. The selected phytase was chosen because of its high level of activity at low pH (van Ooijen et al. (2000), United States Patent Publication No. 6,022,846). The C-extein portion of the phytase is amplified using the following primers under the same conditions as previously described.

5'-CTTTGGATTCCTCTATGCACATAATTTCCTTCGACACCATCTCCACCAGCA-3' [SEQ ID NO: 21]

5'- CTAAGCAAAACACTCCGCCCAATCACCGCCAGATCTGGCAAAGCTCAACC-3' [SEQ ID NO: 22]

COR-AAD1-00002413

[00136] The resulting sequence is 627 nucleotides long. The intein sequence is amplified using primers under the same conditions as previously described.
5'- AGTGACCTACCTCATGGACATGTGCAGCATTTTACCGGAAGAATGGGTTC- 3' [SEQ ID NO: 23]
5'-GCTGGTGGAGATGGTGTCGAAGGAATTATGTGCATAGAGGAATCCAAAG- 3' [SEQ ID NO: 24]

[00137] Finally the N-extein is amplified using primers resulting in a PCR fragment 928 nucleotides long.
5'-ATGGGTGTCTCTGCCGTTCTACTTCCTTTGTACCTCCTGTCCGGAGTATG- 3' [SEQ ID NO: 25]
5'- GAACCCATTCTTCCGGTAAAATGCTGCACATGTCCATGAGGTAGGTCACT- 3' [SEQ ID NO: 26]

[00138] The resulting DNA fragments are cleaned and analyzed, then combined using PCR and the associated methods described in Examples 1 and 2. This procedure results in the intein modified phytase coding sequence. The final composite intein modified phytase sequence is then amplified, cleaned and analyzed as described in Examples 1 and 2. The intein modified phytase DNA coding sequence is cloned into the same expression cassette, and used to transform A. tumefaciens as described in Example 2. Rapeseed stem segments are transformed using the resulting recombinant A. tumefaciens. Transformation occurs substantially as described in Examples 1 and 2 with the following modifications. The rapeseed stem segments are surface sterilized from five to six week-old plants using a 20% bleach solution for 25 minutes at room temperature. Following sterilization, the stem segments are aseptically rinsed with sterile, distilled, and deionized water. The segments are preconditioned by incubation for 24 hours on Murashige-Skoog medium supplemented with 1 mg/ L of BAP. Once the 24 hours has transpired, the stem segments are incubated for 48 hours with the newly transformed strain of A. tumefaciens containing the intein modified phytase. Following this incubation step, regenerate transgenic plants and select them using the kanamycin resistance marker, following

COR-AAD1-00002414

substantially the same procedures described in Examples 1 and 2. Confirmation of incorporation of the intein modified phytase can also be conducted as described in Examples 1 and 2.

[00139] The resulting transgenic rapeseed is grown in an approved area according to local legislation. The rapeseed is harvested when it's mature and used to supplement animal feed. Conversely, the rapeseed can be grown on grazing land for the animals since intein splicing should occur spontaneously in the animal's stomach, allowing for activation of the phytase activity.

[00140] Example 5: Production of Transgenic Maize Expressing an Intein Modified

[00141] Cellulase and Utilization in the Production of Ethanol

[00142] This example illustrates one way in which the invention may be practiced. Here, transgenic corn is constructed and used as a substrate for ethanol processing, or as a substrate in other fermentations. In this example the intein modified gene sequence of Example 1 is again used for demonstration. The growth of Zea mays friable, embryogenic type II callus cultures is initiated from immature embryos, approximately 1.6 mm to 1.8 mm in length, from greenhouse grown A188 (University of Minnesota, Crop Improvement Association) x B73 (Iowa State University) plants. After harvest, fragments are surface sterilized using 50% bleach, for 25 minutes at room temperature, and then washed with sterile, distilled, deionized water. New cultures are aspetically initiated from the harvested fragments and maintained under no more than 10 E m^{-2} s^{-1} light, at 24 °C, on modified N6 medium (Chu, et al., (1975), "Establishment of an Efficient Medium for Anther Culture of Rice through Comparative Experiments on Nitrogen Sources," Sci. Sin., 18:659-668) at pH 5.8, with 2 mg/ L glycine, 2.9 g/ L L-proline, 1 mg/ L 2,4-dichlorophenoxyacetic acid (2,4-D), 100 mg/ L casein hydrolysate, 20 g/ L sucrose, and solidified with 2 g/ L Gelgro (ICN Biochemicals).

[00143] After approximately two weeks of incubation, the cultures are manually evaluated for proper morphology. This entails visual observation for friable consistency in the presence of well-characterized somatic embryos.

COR-AAD1-00002415

Proliferations demonstrating proper morphology are transferred to fresh modified N6 medium (described above). Tissues resulting with the proper morphology are routinely subcultured every two to three weeks, until the microprojectile bombardment is prepared. The desired intein modified gene sequence and expression vector can be constructed as described in Example 1. In this example, the preferred expression vector also has the following alterations. Replace the kanamycin resistance marker with a hygromycin resistance marker using methods known in the art (for example, PCR of the hygromycin resistance marker from a suitable template, DNA endonuclease restriction of the vector, followed by purification, and ligation of the hygromycin resistance marker) as described by Ausbel, et. al., 1998. Once constructed, the vector is precipitated in a 1:1 molar ratio onto either tungsten (average diameter 1.2 m, GTE Sylvania), or gold, particles. As with other steps in this procedure, the precipitation parameters may require some minor optimization. The precipitation is performed by combining 1.25 mg of the tungsten particles, and 25 g of the vector DNA in solution with 1.1 M $CaCl_2$ and 8.7 mM spermidine at a total volume of 575 L. The precipitate is vortexed for 10 minutes at 0 °C. Once vortexed, the mixture is centrifuged at 500xg for five minutes. After centrifugation, the supernatant, approximately 550 L, is removed and the remaining 25 L of precipitate is dispensed in 1 L aliquots onto macroprojectiles (Biolistics, Inc, Ithaca, NY) for bombardmentas described by Klein et al. (1987), except for the changes noted above. All manipulations are performed aseptically and on ice.

[00144] Once the biolistic projectiles are ready, the desired plant tissues are prepared for the bombardment procedure. Any number of callus clumps are aseptically arranged, each weighing 50 mg (wet weight), in an x-pattern near the center of a sterile 60 x 15 mm petri dish (Falcon 1007). Several dishes should be prepared for each bombardment procedure. These dishes are each paced in turn, 5 cm below the stopping plate of the microprojectile instrument. The dishes are centered below the device, with the lids removed, and a 3 x 3 mm mesh screen

COR-AAD1-00002416

covering the top of the plate. The mesh screen helps contain bombarded tissue within the dish during the procedure. The tissue bombardment is performed with the microprojectile instrument as described by the manufacturer's instructions; commercial microprojectile instruments are available through Bio-Rad (Hercules, CA). Following bombardment, the callus are transferred to fresh modified N6 medium plates and cultured under the same conditions used above.

[00145] The selection of transformed cells for subsequent regeneration is began fter two days of culture. The callus plates subjected to the bombardment procedure are aseptically transferred to fresh, sterile, modified N6 medium plates formulated to a final concentration of 10mg/ L hygromycin B (Calbiochem). After two weeks of exposure, all callus are aseptically transferred from the selective plates to fresh, sterile, modified N6 medium plates formulated to a final concentration of 50 mg/ L hygromycin B. This transfer is conducted so that only five 30 mg pieces of callus are contained on a single plate, resulting in an expansion of the number of plates used. Following three weeks on the 50 mg/ L hygromycin B plates, all callus are aseptically transferred to fresh, sterile, modified N6 medium plates formulated to a final concentration of 60 mg/ L hygromycin B. After two weeks of incubation, the callus are inspected for proliferating clumps. Selected proliferating clumps are transferred to a modified Murashige-Skoog medium supplemented with 0.5 mg/ L thiamine-HCl, 0.75 mg/ L 2,4-D, 50 g/ L sucrose, 150 mg/ L asparagines, and 2.0 g/ L Gelgro.

[00146] At this point it is prudent to ensure transformation of the selected plants. The presence of the intein modified cellulase is verified using the methods described in Examples 1 and 2. In this case, either or both of the intein modified cellulase coding sequence, and the hygromycin resistance marker can be used as the subject of the transformation validation, using methods known in the art, as described by Ausbel, et al., 1998. After two weeks on the modified Murashige-Skoog medium, the plates are exposed to a light cycle incubation regimen composed of 14 hours of light, followed by 10 hours of dark, at 24 °C. Plantlets that form are aseptically transferred to 1 L, wide mouthed Erlenmeyer

COR-AAD1-00002417

flasks containing 100 mL of the modified Murashige-Skoog medium. The resulting plants are transferred to vermiculite for one to two weeks prior to plantation in soil and growth to maturity. The mature plants are analyzed substantially as described in Example 1 to ensure stable transformation of the intein modified protein sequence, and preferentially, expression of the intein modified cellulase.

[00147] The resulting mature plants may be cross-pollinated using standard techniques. This can be done either between transformed plants, or between a single transformed plant and an untransformed plant. The progeny resulting from the breeding are screened for containment of the intein modified cellulase, as well as the hygromycin resistance marker. Note, at this point the hygromycin resistance marker used in the selection is no longer an essential element for the use and application of the constructed transgenic corn plants. So long as the intein modified cellulase sequence is contained, retention of the hygromycin resistance marker is not an essential component of the transgenic corn. Seed can be harvested from the fertile transgenic plants and used for plant expansion. The resulting transgenic plants can be grown for use in processes similar to those described in Example 3. The process using a transgenic corn species expressing multiple intein modified proteins, would have the economic advantages of utilizing both the starch and cellulosic portions of the corn plant, consolidating the pretreatment, saccharification, and fermentation steps, and decreased energy and raw material input costs. Effective use of this process for the production of ethanol would be enabled by the inclusion of the intein modified proteins in the transgenic plant.

[00148] The enclosed examples do not in any way limit the scope of this patent, as they solely provided to help illustrate applications of the invention disclosed in this patent. Other variations are possible as an artisan would know, and are included within the four corners of this invention.

[00149] BIBLIOGRAPHY

[00150] Dale, Biotechnol. Prog. 15:775-776 (1999).

COR-AAD1-00002418

[00151] Committee on Biobased Industrial Products, Biobased Industrial Products: Priorities for Research and Commercialization, National Academy Press, Washington D C (1999).

[00152] Cameron, et al., Biotechnol. Prog. 14:116-125 (1998).

[00153] Park, et al., Biotechnol. Prog. 14:699-704 (1998).

[00154] Taylor, et al., Biotechnol. Prog. 16:541-547 (2000).

[00155] Poirier, Nature Biotechnology 17:960-961 (1999).

[00156] Lynd, Biotechnol. Prog. 15:777-793 (1999).

[00157] Ingram, Biotechnol. Prog. 15:855-866 (1999).

[00158] Aspergren, et al., Molecular Breeding 1:91-99 (1995).

[00159] Mansfield, et al., Biotechnol. Prog. 15:804-816 (1999).

[00160] Evans, et al., Protein Science 7:2256-2264 (1998).

[00161] Perler, et al., Nucl. Acids Res. 22:1125-1127 (1994).

[00162] Xu, et al., EMBO 15:5146-5153 (1996).

[00163] Wood, et al., Biotechnol. Prog. 16:1055-1063 (2000).

[00164] Clarke, P.N.A.S. (USA) 91:11084-11088 (1994).

[00165] Derbyshire, et al., P.N.A.S. (USA) 95:1356-1357 (1998).

[00166] Perler, et al., Nucleic Acids Res. 22:1125-1127 (1994).

[00167] Wallace, C. J., Protein Sci. 2:697-705, (1993).

[00168] Xu, et al., Cell 75:1371-1377 (1993).

[00169] Pietrokovski, S., Protein Sci. 3735:2340-2350 (1994).

[00170] Ausbel, et al., "Current Protocols in Molecular Biology," Wiley, New York (1998).

[00171] Draper, et al., "Plant Genetic Transformation and Gene Expression: A Laboratory Manual," Blackwell Scientific Publications, Boston (1998).

[00172] Potrykus, et al., "Gene Transfer to Plants," Springer, New York (1995).

[00173] Guilley, et al., Cell, 30:763 (1982).

[00174] Higgins, T. J. V., Annu. Rev. Plant Physiol. 35:191(1984).

[00175] Coruzzi, et al., EMBO J. 3:1671 (1984).

COR-AAD1-00002419

[00176] Tingey, et al., EMBO J. 6:3565 (1987).

[00177] Ryan, et al., Nuc. Acids Res. 17:3584 (1989).

[00178] Rocha-Sosa, et al., EMBO J. 8:23 (1989).

[00179] Wenzler, et al., Plant Mol. Biol. 12:41 (1989).

[00180] Bird et al., Plant Mol. Biol. 11:651 (1988).

[00181] Brederode, et al., Nucl. Acids Res. 8:2213 (1980).

[00182] Smeekens, et al., T.I.B.S. 15:73 (1990).

[00183] van den Broeck et al., Nature 313:358 (1985).

[00184] Schreier, et al., EMBO J. 4:25 (1985).

[00185] Tague, et al., Plant Phys. 86:506 (1988).

[00186] Von Heijne, G., J. Mol. Biol. 189:239 (1986).

[00187] Sijmons, et al., Bio/Technol. 8:217 (1990).

[00188] Gordon-Kamm, et al., The Plant Cell 2:603 (1990).

[00189] Shimamoto, et al., Nature 338:274 (1989).

[00190] Horsch, et al., Science 2:1229 (1985).

[00191] Ziegler, et al., Molecular Breeding 6:37-46 (2000).

[00192] Dai, et al., (a), Molecular Breeding 6:277-285 (2000).

[00193] Dai, et al., (b), Molecular Breeding 9:43-54 (2000).

[00194] Montvalvo-Rodriguez, et al., Biotech. and Bioeng. 2:151-159 (2000).

[00195] U.S. Patent No. 6,022,846.

[00196] Chu, et al., Sci. Sin. 18:659-668 (1975).

[00197] Klein, et al., Nature 327:70 – 73 (1987).

[00198] The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein.

* * *

COR-AAD1-00002420

CLAIMS

What is claimed is:

1. A recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, comprising an expression construct(s) that encode(s) at least one modified protein comprising a target protein(s) or protein segment(s), which is(are) fused, either internally or terminally, to a controllable intervening protein sequence(s) (CIVPS) or intein sequence(s) or segment(s) thereof, or to an amino terminus(i) or a carboxyl terminus(i) thereof.

2. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein each expression construct comprises, operatively linked to one another, a first nucleic acid segment(s) encoding a target protein(s), and a second nucleic acid segment(s) encoding a CIVPS or intein sequence(s), and optionally a selectable marker(s) or reporter gene(s) and/or a promoter(s).

3. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the modified protein(s) is(are) expressed constitutively.

4. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, which inductively expresses at least one modified protein(s); and optionally wherein a stimulus(i) triggers splicing of the modified protein.

5. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 4, wherein the stimulus(i) comprise(s) a change in pH, osmolality, temperature, a fertilizer(s), pesticide(s), chemical(s), light, and/or sound.

COR-AAD1-00002421

6. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein at least one modified protein(s) is(are)
expressed at a pre-determined point of the plant life cycle;
in at least one specific tissue or part thereof; and/or
in at least one specific sub-cellular compartment(s); and/or
expressed and secreted extracellularly.

7. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 6, wherein the plant's at least one
specific tissue(s) comprise(s) seeds, roots, fruits, stems, tubers and/or leaves; and
specific subcellular compartment(s) comprise(s) a cytosol, mitochondrion, plastid, endoplasmic reticulum, inclusion body, vacuole and/or nucleus.

8. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 2, wherein
the at least one selectable marker confers resistance to a chemical comprising bromoxynil, 2,2-dichloropropionic acid, G418, glyphosphate, haloxyfop, hygromycin, imidazoline, kanamycin, methotrexate, neomycin, phosphinothricin, sethoxydim, 2,2-dichloropropionic acid, glyphosphate, hygromycin, trichothecne, sulfonylurea, s-triazine and/or triazolopyrimidine; and/or
the promoter precedes the at least one CIVPS or intein-modified protein polynucleotide.

9. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 8, which is tolerant or resistant to normally toxic levels of at least one chemical(s).

10. The plant, or plant part, plantlet, tissue, cell, subcellular fraction,

COR-AAD1-00002422

seed, seedling, protoplast, progeny or descendent of claim 1, which is fertile, and has at least one heritable modified protein encoding polynucleotide sequence(s).

11. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the plant is not fertile.

12. An inbred genetically recombinant plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1.

13. A hybrid genetically recombinant plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1.

14. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 4, wherein the spliced protein(s) changes(s) the content or activity of at least one pre-determined plant component(s).

15. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the plant component(s), whose content(s) is(are) reduced comprise(s) glucose, fructose, glycerol, glycine-betaine, sucrose, lactose, maltose, galactose, amino acids, lipids, vitamins and/or starch.

16. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the plant component(s), whose activity(ies) is(are) reduced comprise(s) cellulose, hemicellulose, lignin, starch, pectin and/or lipids.

COR-AAD1-00002423

17. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the CIVPS or intein sequence(s) and the target protein(s) or protein segment(s) form at least one splice junction(s).

18. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the amino acid residue at the carboxyl terminus of the splice junction(s) comprise(s) an amino acid residue provided with a hydroxyl or a sulfhydryl side chain(s).

19. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the splice junction(s) downstream of the CIVPS or intein sequence(s) or segment(s) thereof comprises an amino acid residue lacking a hydroxyl or sulfhydryl side chain(s) at the amino terminus of the target protein or protein segment(s).

20. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the splice junction(s) upstream of the CIVPS or intein sequence(s) or segment(s) thereof comprises an amino acid residue having a hydroxyl or a sulfhydryl side chain at the amino terminus of the CIVPS or intein sequence(s) or protein segment(s).

21. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the splice junction(s) upstream of the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) a cysteine.

22. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the splice junction(s) downstream of the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) His-Asn at the carboxyl terminus(i) of the CIVPS or intein

COR-AAD1-00002424

sequence(s) or segment(s) thereof, and an amino acid residue having a hydroxyl or sulfhydryl side chain at the amino terminus(i) of the adjoining region(s) of the target protein(s).

23. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 17, wherein the splice junction(s) downstream of the CIVPS or intein sequence(s) or protein segment(s) thereof comprise(s) Asp at the carboxyl terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof, and an amino acid residue having a hydroxyl or sulfhydryl side chain at the amino terminus(i) of the adjoining region(s) of the target protein(s) or protein segment(s).

24. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 23, wherein the Asp at the carboxyl terminus(i) Asp is replaced by an amino acid lacking a carboxyl or an amino side chain(s).

25. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) an CIVPS or intein(s).

26. The plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the CIVPS or intein sequence(s) or segment(s) thereof comprises an externally controllable Saccharomyces CIVPS or intein sequence(s).

27. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the CIVP or intein sequence(s) comprise(s) a externally controllable Saccharomyces cerevisiae CIVPS or intein sequence(s).

COR-AAD1-00002425

28. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the CIVPS or intein sequence(s) is(are) inserted immediately before Ser, Thr or Cys of the target protein(s) or protein segment(s).

29. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the CIVPS or intein amino or carboxy terminus(s) comprise(s) Ser, Thr or Cys.

30. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the protein is expressed by a microorganism.

31. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 30, wherein the microorganism is selected from Bacillus thuringiensis, or Phytolacca insularis.

32. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the expressed protein comprises a modified Bacillus thuringiensis endotoxin.

33. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, wherein the protein is expressed by a virus.

34. The plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 33, wherein the virus is selected from potato virus Y, geminivirus, aspermy virus 2b, or cucumber mosaic virus.

35. A method for producing the recombinant plant, or plant part,

COR-AAD1-00002426

plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1, comprising

providing an expression construct that encode(s) at least one modified protein comprising a target protein, or protein segment(s), which is(are) fused, either internally or terminally, to a CIVPS or intein sequence(s) or segment(s) thereof, or to an amino terminus(i) or a carboxyl terminus(i) thereof;

transforming a plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, with the expression construct; and

regenerating a genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, from the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, that encode(s) at least one modified protein sequence(s).

36. The method of claim 35, wherein the transformation comprises a stable transformation.

37. The method of claim 35, wherein the regeneration step is conducted by at least one of

breeding of a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling protoplast, progeny or descendent;

crossing of a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent and a non-genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent; or

back-crossing of two genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent.

38. The method of claim 35, wherein the expression construct comprises

COR-AAD1-00002427

at least one of promoter, selectable marker, resistance marker, heritable marker, poly-adenylation sequence, repressor, enhancer, localization sequence, or signaling sequence.

39. The method of claim 35, wherein the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is transformed with the expression construct by at least one of viral transformation, bombardment with DNA-coated microprojectiles, liposomal gene transformation, bacterial gene transfer, electroporation, or chemical gene transformation.

40. The method of claim 35, wherein the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, is transformed by means of a bacterium comprising Agrobacterium tumefaciens.

41. The method of claim 35, wherein the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, is transformed by chemical gene ttransformation with the aid of at least one of calcium phosphate, or polyethylene glycol.

42. The method of claim 35, wherein the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is selected with the aid of at least one of a selectable marker, or resistance marker.

43. The method of claim 35, wherein the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is selected with the aid of the expression of at least one nucleic acid encoding an CIVPS or intein modified protein.

COR-AAD1-00002428

44. The method of claim 35, wherein the genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is regenerated from

one of transformed embryogenic tissue(s);

plant protoplasts;

cells derived from immature embryos; or

transformed seeds.

45. A method for producing a modified protein(s) or protein segment(s) from a recombinant transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent expressing the protein(s) or protein segment(s), comprising

conducting the method of claim 35; and

harvesting the modified protein(s) or protein segment(s) from the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent.

46. The method of claim 45, further comprising purifying the modified protein.

47. The method of claim 45, wherein the modified protein(s) or protein segment(s) is(are) CIVPS or intein modified protein(s) or protein segment(s).

48. A method for producing a modified protein comprising a target protein(s) or protein segment(s) fused, either internally or terminally, to a CIVPS or intein sequence(s) or segment(s) thereof, or to its amino terminus(i) or carboxyl terminus(i), comprising

obtaining an expression construct encoding a target protein having an in-frame fused CIVPS or onteon sequence(s) or segment(s) thereof, or its amino terminus(i) or carboxyl terminus(i);

transforming a host plant cell(s) with the expression construct; and

COR-AAD1-00002429

culturing the transformed plant host cell under conditions effective for expressing the modified protein.

49. The method of claim 48, wherein in the expression construct, the at least one first nucleic acid segment(s) encoding the CIVPS or intein sequence(s0 or segment(s) thereof is(are) fused to the 5'-end of the at least one second nucleic acid segment(s) encoding the target protein(s) or protein segment(s).

50. The method of claim 48, wherein in the expression construct, the at least one first nucleic acid segment(s) encoding the CIVPS or intein sequence(s) or segment(s) thereof is(are) fused to the 3'-end of the at least one second nucleic acid segment(s) encoding the target protein(s) or protein segment(s).

51. The method of claim 48, wherein the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) a Saccharomyces protein(s) or segment(s) thereof.

52. The method of claim 48, wherein the CIVPS or intein sequence(s) or segment(s) thereof is(are) capable of effecting, either in cis or in trans, at least one of excision, cleavage, ligation, excision-ligation, cleavage-ligation, or cyclization.

53. The method of claim 48, wherein protein splicing is induced by at least one of change of temperature, light or pH, addition/removal of chemical reagent that facilitates/inhibits splicing or cleavage, amino acid dephosphorylation or deglycosylation, or contact with/removal of peptide or peptidomimetic activating/blocking splicing or cleavage.

54. The method of claim 53, wherein protein splicing is induced by at least one of in vitro or in vivo contact/removal of peptide or peptidomimetic that activates/blocks splicing or cleavage.

COR-AAD1-00002430

55. The method of claim 48, wherein the amino or carboxy terminus(i) of the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) Ser, Thr or Cys.

56. The method of claim 48, wherein the carboxyl terminus of the CIVPS or intein sequence(s) or segment(s) thereof comprise(s) Asp preceding Ser, Thr or Cys of the target protein(s) or protein segment(s).

57. The method of claim 48, wherein the expression construct further comprises at least one of promoter, selectable marker, resistance marker, heritable marker, poly-adenylation sequence, repressor, enhancer, localization sequence, or signaling sequence; and optionally further comprising
transforming the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent with the expression construct by at least one of viral transformation, bombardment with DNA-coated microprojectiles, liposomal gene transfer, bacterial gene transfer, electroploration, or chemical gene transformation.

58. The method of claim 57, wherein the bacterium used to transfer the expression construct is Agrobacterium tumefaciens.

59. The method of claim 57, wherein the chemical used for transformation is calcium phosphate, or polyethylene glycol; and
the transformed plant cells, plant parts, or plants are selected through their expression of a selectable marker, or resistance marker.

60. The method of claim 48, wherein
the transformed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is(are) selected through their expression of the modified protein gene sequence; and

COR-AAD1-00002431

the genetically recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is(are) regenerated from transformed embryogenic tissue.

61. The method of claim 48, wherein the genetically recombinant plants are regenerated from cells derived from immature embryos.

62. The method of claim 48, wherein the genetically recombinant plants are regenerated from transformed seeds.

63. A method for producing seed expressing a modified protein(s), comprising

obtaining the genetically recombinant plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent comprising the method of claim 1;

culturing or cultivating the genetically recombinant plant, or plant part, plantlet, tissue, cell, subcellular fraction, seed, seedling, protoplast, progeny or descendent; and

obtaining therefrom seed that expresses a modified protein(s).

64. A method for using a modified protein(s) expressing plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent for producing a compound, comprising

harvesting a recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1;

mechanically processing the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent;

combining the mechanically processed plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, with a non-genetically recombinant plant in a proportion greater than or equal to zero recombinant:non-recombinant; and

COR-AAD1-00002432

chemically processing the plant or specific portions of the plant under conditions effective for obtaining the compound.

65. The method of claim 64, wherein the mechanical processing of the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent, comprises at least one of extrusion, grinding, shredding, mulching, chipping, dicing, compressing, exploding, or tearing.

66. The method of claim 64, wherein the chemical processing of the combine comprises at least one of pre-treatment with steam, dilute or concentrated acid, or ammonia explosion, sterilization, soaking in water, mixing with a solvent, a change of pH, temperature and/or osmolality, light exposure, inorganic and/or enzyme catalysis, saccharification, bleaching, scouring, fermentation, distillation, chromatography, adsorption, or addition of a chemical(s).

67. The method of claim 66, wherein the pre-treatment comprises steaming the combine to sterilize it.

68. The method of claim 66, wherein the chemical processing comprises one of

pre-treatment with at least one of sulfuric acid, hydrochloric acid, phosphoric acid, or carbonic acid;

soaking in water at a temperature greater than or equal to about 20 °C; and/or

mixing the combine with at least one of water, or an organic or inorganic solvent(s).

69. The method of claim 67, wherein an external stimulus(i) is applied to induce splicing of the modified protein(s) or protein segment(s).

COR-AAD1-00002433

70. The method of claim 69, wherein the external stimulus(i) comprises at least one of change of pH, osmolality, or temperature, exposure to sound, light, or addition of a chemical(s).

71. The method of claim 69, wherein the spliced protein(s) or protein segment(s) has(have) altered activity(ies) with respect to the modified protein(s) or protein segment(s).

72. The method of claim 71, wherein the spliced protein(s) or protein segment(s) has(have) at least one of catabolic or anabolic activity(ies).

73. The method of claim 72, wherein the spliced protein(s) or protein segment(s) is (are) capable of degrading starch, dextrin, pectin, lipids, protein, chitin, lignin, cellulose, or hemicellulose, or modifying lignin, or has(have) saccharification activity.

74. The method of claim 72, which is capable of producing at least one of glucose, fructose, xylose, phenol, glycerol, mannose, lactic acid, acetic acid, ethylene, propylene, toluene, ethyl benzene, styrene, xylene, ethylene glycol, butadiene, formaldehyde, isopropanol, acetone, butanediol, methanol, ethanol, propanol, butanol, propanediol, vitamins, methane, ethane, propane, butane, pentane, hexane, heptane, octane, benzene, or biopolymers.

75. The method of claim 67, wherein the pre-treatment, saccharification, and fermentation are conducted in one step.

76. The method of claim 67, wherein the saccharification and fermentation are conducted in one step.

77. The method of claim 67, wherein the fermentation is conducted with a prokaryotic or eukaryotic microorganism.

COR-AAD1-00002434

78. The method of claim 74, wherein the microorganism produces at least one of lactic acid, acetic acid, ethylene, propylene, toluene, ethyl benzene, styrene, xylene, ethylene glycol, butadiene, formaldehyde, isopropanol, acetone, butanediol, methanol, ethanol, propanol, butanol, octanol, propanediol, vitamins, methane, ethane, propane, butane, pentane, hexane, heptane, octane, benzene, or biopolymers.

79. An animal feedstock, comprising a nutritious amount of the recombinant plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent of claim 1.

80. The feedstock of claim 79 wherein, when the plant, or plant part, plantlet, tissue, cell, sub-cellular fraction, seed, seedling, protoplast, progeny or descendent is ingested the modified protein(s) or protein segment(s) is(are) spliced by an internal stimulus(i) from the animal.

81. The feedstock of claim 79, wherein the animal's internal stimulus(i) comprise(s) at least one of saliva, bile, chymotrypsin, trypsin, bicarbonate, hydrochloric acid, or stomach pH or temperature.

82. The feedstock of claim 79, wherein the spliced protein(s) comprise(s) at least one of a phytase, endocellulase, exocellulase, amylase, glucanase, hemi-cellulase, pectinase, protease, xylanase, or lipase enzyme, or a growth hormone.

83. An immune response enhancing composition, comprising the feedstock of claim 79, wherein the spliced protein(s) or protein segment(s) comprise(s) at least one recombinant immunogen(s).

84. The compostion of claim 83, wherein the recombinant immunogen comprises a viral or bacterial immunogen.

COR-AAD1-00002435

85. The composition of claim 83, in the form of an oral formulation.

86. The composition of claim 83, in the form of a trans-mucosally absorbed formulation.

87. The composition of claim 83, in the form of a gastrointestinal (G.I.) tract absorbed formulation.

88. The composition of claim 83, wherein the spliced protein(s) or protein segment(s) is(are) a prokaryotic protein(s) or protein segment(s).

89. The composition of claim 83, wherein the spliced protein(s) or protein segment(s) is(are) an eukaryotic protein(s) or protein segment(s).

90. A method of producing an animal feedstock, comprising

conducting the method of claim 35; and

processing the genetically modified plant, or a portion of the resulting genetically recombinant plant, under conditions effective to obtain an animal digestible feedstock.

91. A method of promoting animal growth, comprising allowing an animal access to the feedstock of claim 79.

92. A method of enhancing an animals immune response, comprising administering to an animal in need of immune enhancement, an immune enhancing amount of the composition of claim 83.

93. A method of producing a target protein(s) or protein segment(s), comprising

COR-AAD1-00002436

producing a first modified protein(s) or protein segment(s), wherein the amino terminus of a CIVPS or intein sequence(s) or segment(s) thereof is(are) fused to the carboxyl terminus(i) of a target protein(s) or protein segment(s) by the method of claim 45;

producing a second modified protein(s) comprising a segment(s) of the CIVPS or intein sequence(s); and

contacting first and second modified proteins under conditions effective for trans cleavage of the CIVPS or intein sequence(s) or segment(s) thereof by the second modified protein(s).

94. A method of producing a target protein(s), comprising

producing a first modified protein(s), wherein the carboxyl terminus of a CIVPS or intein sequence(s) or protein segment(s) thereof is(are) fused to the amino terminus(i) of the target protein(s) or protein segment(s) by the method of claim 45;

similarly producing a second modified protein(s) or protein segment(s) comprising a segment(s) of the CIVPS or intein sequence(s); and

contacting first and second modified proteins under conditions effective for trans cleaving the CIVPS or intein sequence(s) or segment(s) thereof from the first modified protein(s) or protein segment(s).

95. The method of claim 92, wherein the cleavage is induced by a change in at least one of temperature, light, or pH, addition/removal of chemical that facilitates/inhibits splicing or blocking of cleavage, amino acid dephosphorylation or deglycosylation, or contact/removal of peptide or peptidomimetic that activates/blocks splicing/cleavage.

COR-AAD1-00002437

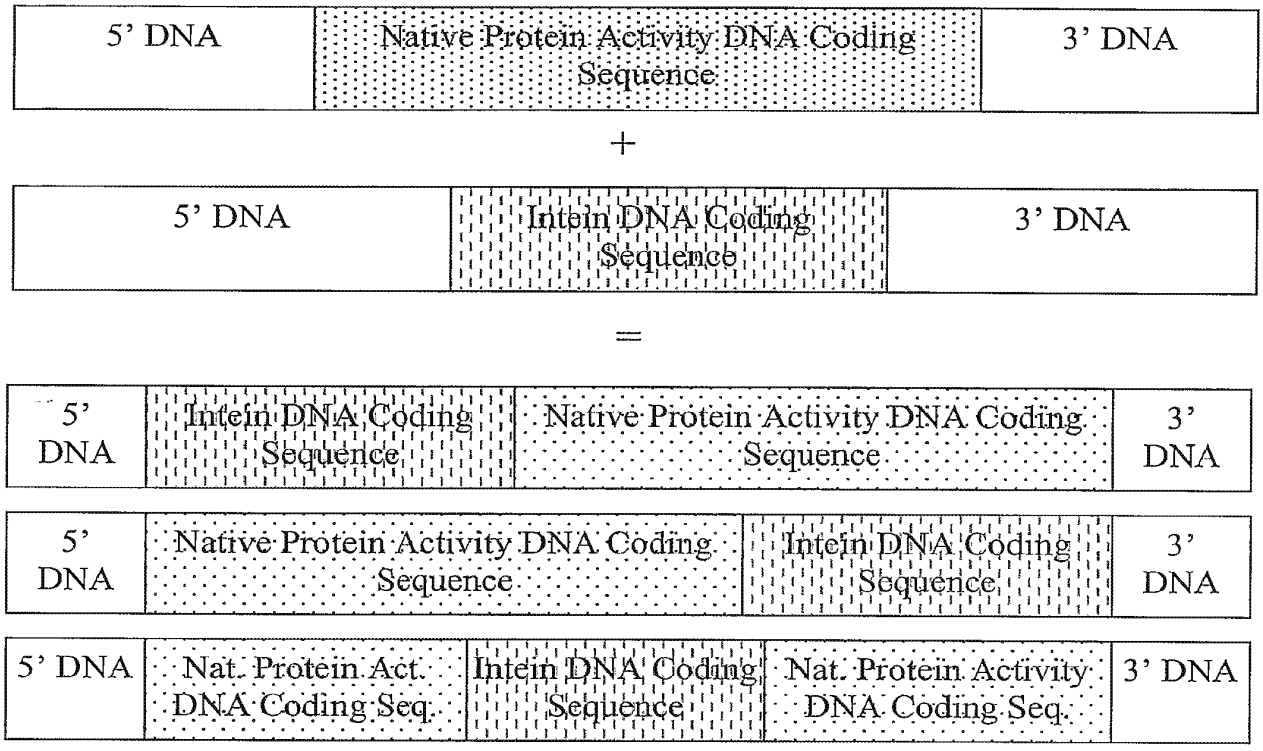


Figure 1

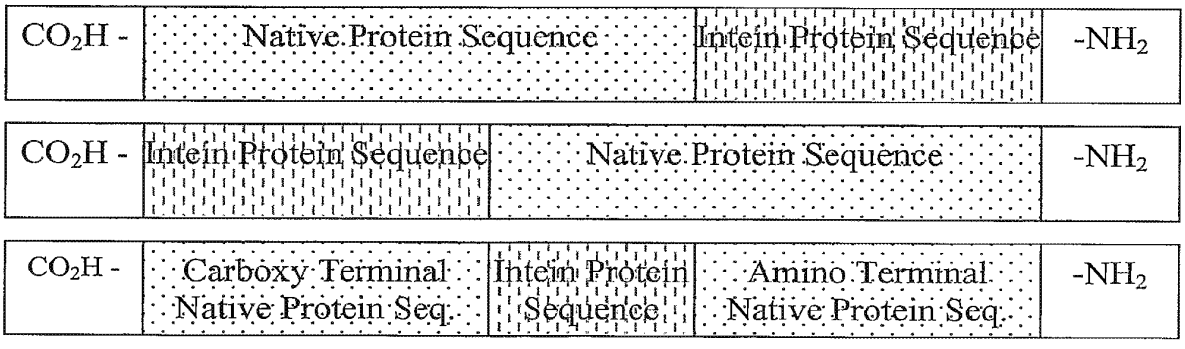


Figure 2

COR-AAD1-00002438

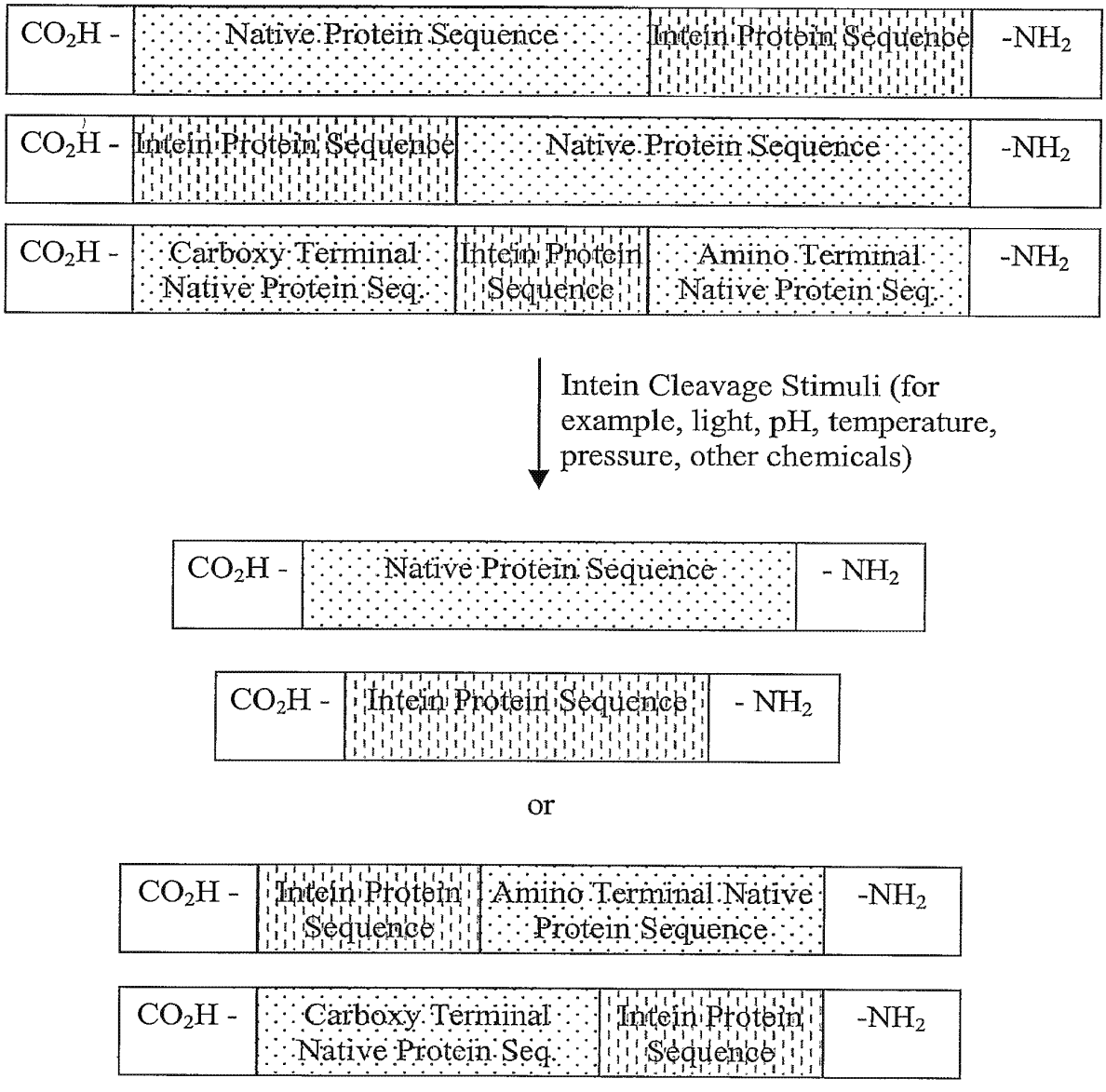


Figure 3

COR-AAD1-00002439

Theor Appl Genet (1992) 83:645–649

© Springer-Verlag 1992

Engineering 2,4-D resistance into cotton

C. Bayley [1], N. Trolinder [2], C. Ray [1], M. Morgan [1], J. E. Quisenberry [2], and D. W. Ow [1,*]

[1] Plant Gene Expression Center, USDA/ARS, and Department of Plant Pathology, University of California at Berkeley, 800 Buchanan St., Albany, CA 94710, USA
[2] Cropping Systems Research Laboratory, USDA/ARS, Route 3, Box 213, Lubbock, TX 79401, USA

Received June 11, 1991; Accepted August 8, 1991
Communicated by P. Maliga

Summary. To reduce damage by drift-levels of the herbicide 2,4-dichlorophenoxyacetic acid, we have engineered the 2,4-D resistance trait into cotton (*Gossypium hirsutum* L.). The 2,4-D monooxygenase gene *tfd*A from *Alcaligenes eutrophus* plasmid pJP5 was isolated, modified and expressed in transgenic tobacco and cotton plants. Analyses of the transgenic progeny showed stable transmission of the chimeric *tfd*A gene and production of active 2,4-D monooxygenase. Cotton plants obtained were tolerant to 3 times the field level of 2,4-D used for wheat, corn, sorghum and pasture crops.

Key words: Genetic modification – Gene transfer – *tfd*A – *Gossypium hirsutum* – Herbicide

Introduction

The herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) and several related phenoxy compounds have been used extensively for more than 40 years to control broadleaf weeds. Since monocotyledonous plants are naturally tolerant to 2,4-D, it is applied to control weeds which grow among cereals, sugarcane, turf grass, and to manage forest undergrowth. Because 2,4-D is highly volatile and tends to drift in the prevailing winds, nearby crops that are susceptible are often damaged. One of the crops that is most susceptible to damage by this herbicide is cotton, and annual loss to 2,4-D damage in cotton is extensive. One approach to reducing these annual losses is to incorporate resistance to 2,4-D into cotton germplasm. Numerous strains of soil bacteria have been described which contain plasmid-borne genes that confer the ability to catabolize 2,4-D (Amy et al. 1985; Don and Pemberton 1981). The gene for the first step of 2,4-D degradation has been isolated from pJP4, a plasmid found in *A. eutrophus* (Streber et al. 1987; Lyon et al. 1989). This gene, *tfd*A, was shown to encode 2,4-D monooxygenase, which degrades 2,4-D into 2,4-dichlorophenol and glyoxylate. The transfer of modified versions of this gene into tobacco conferred resistance to the herbicide (Lyon et al. 1989; Streber and Willmitzer 1989).

In this communication, we report the engineering of 2,4-D resistance into cotton. A homologue *tfd*A gene was isolated from pJP5, another *A. eutrophus* 2,4-D degradative plasmid, and initially examined in tobacco. We show that both tobacco and cotton plants obtained by gene transfer exhibited 2,4-D degradative activity and were tolerant to high levels of the herbicide. Cultivation of 2,4-D tolerant varieties of cotton would reduce annual losses caused by drift-levels of the herbicide.

Materials and methods

Recombinant DNA

Standard techniques of recombinant DNA were used throughout (Maniatis et al. 1982). *A. eutrophus* DSM 4059 and *E. coli* DSM 4063 harboring plasmid pJP5 were obtained from the German collection of microorganisms (DSM, Mascheroder Weg 1 b, W-3300 Braunschweig, FRG). DNA oligonucleotides corresponding to the 5′- and 3′-coding regions, 5′-AGCGTCGTCGCAAATCCC-3′ and 5′-ATCGTCCAGGGTGGTCGC-3′, respectively, were used as synthetic probes to detect the *tdf*A gene by Southern hybridization. A partial genomic clone bank was constructed from DSM 4063 by inserting approximately 2 kb gel-purified XbaI and SalI doubly-digested DNA fragments into pUC19. Colony hybridization was used to detect *tfd*A-containing plasmids. The same primers were used for limited DNA sequence analysis of the 5′ and 3′ ends of the coding region to verify *tfd*A identity.

A modified *tfd*A gene was constructed for expression in plant cells. The approximately 860-bp 5′-FokI to NcoI-3′ frag-

* To whom correspondence should be addressed

COR-AAD1-00002440

ment, obtained from partial digestion with FokI and containing the *tfd*A coding region except for the first 11 bp, was fused to a synthetic adapter

5'-CTAGACACAAAGTGGTTAACAAGGAAGGAAAAATGAGCGTCGT-3'

3'-TGTGTTTCACCAATTGTTCCTTCCTTTTTACTCGCAGCAGCGT-5'

and a SalI linker (5'-GGTCGACC-3') at the 5' and 3' ends, respectively (Fig. 1). The 5' end adapter regenerated the original *tfd*A coding region except for an ATG in place of GTG as the putative start codon. Preceding the coding region is a leader sequence with a prokaryotic ribosome binding site. This modified XbaI to SalI *tfd*A fragment was then inserted between the XbaI and SalI sites of pUC19, allowing excision of the modified *tfd*A gene as a BamHI/SphI fragment, which was then placed within the corresponding sites in pDO432 (Ow et al. 1986). This placed the modified *tfd*A gene between a 1585-bp 35S (CaMV 35S promoter) fragment and a 754-bp *nos*3' (nopaline synthase gene polyA) fragment to yield pRO17 (Fig. 1). To transfer the *tfd*A gene into an *Agrobacterium* gene transfer vector, pRO17 was linearized by HindIII and ligated to HindIII-cleaved pBIN19 (Bevan 1984) to form a co-integrate plasmid, with transcription of the *tfd*A gene towards the left T-DNA border (Fig. 1). The cointegrate pBIN19 : : pRO17 was transferred into the "disarmed" *A. tumefaciens* strain GV3111 harboring the helper Ti plasmid pTiB6S3S3.

Gene transfer into tobacco and cotton

Agrobacterium-mediated gene transfer into *Nicotiana tabacum* (cv 'Wi38') has been described (Horsch et al. 1985). Kanamycin (100 µg/ml) was used to select transformed cells, and cefotaxime (500 µg/ml) was used to remove *Agrobacterium*. The same *Agrobacterium* vector was used for gene transfer into cotton (cv 'Coker 312'). Transformation of cotton was similar to that reported previously (Umbeck et al. 1987) with the following modifications. Hypocotyl sections were precultured for 3 days on MS salts (Gibco) containing B5 vitamins, 30 g/l glucose, 0.1 mg/ml 2,4-D and 0.5 mg/l kinetin. Kanamycin and cefotaxime were used at 50 µg/ml and 500 µg/ml, respectively. Plant regeneration was performed as described (Trolinder and Goodin 1987) except that embryos were partially dehydrated over a 2-week period by placing them on a 2% agar medium containing 15 ml of S medium (Stewart and Hsu 1977) and 30 g/l glucose. S medium containing 0.5% agar, 1.5 g/l Gelrite and 0.5% sucrose was used for embryo germination.

Assay for 2,4-D degradative activity

Bacterial cultures were assayed for $^{14}CO_2$ production from ^{14}C-labeled 2,4-D as described previously (Amy et al. 1985). The same assay was used on plant material, but with two modifications: the use of plant protoplasts, and the addition of excess unlabelled glyoxylate as a competitive substrate for glyoxylate metabolism. Leaf mesophyll protoplasts were prepared as described (Takebe et al. 1971). A 10-ml aliquot of protoplasts at a density of 1×10^5/ml was incubated in culture medium with 1 ml of a 20 mg/ml glyoxylate solution and 2 µl of ^{14}C-labeled 2,4-D (1.8 µmoles, 0.1 µCi). After 16 h of incubation in the dark, the amount of acid-hydrolyzed $^{14}CO_2$ released from the culture was measured as in the bacterial assay.

Assay of explants for 2,4-D tolerance

Tobacco leaf slices were placed on shoot-induction medium in the absence of 2,4-D or in the presence of 2 µg/ml (10^{-5} *M*) to 6 µg/ml of 2,4-D. The ability to form shoots was scored after 2 weeks. Regenerated cotton plants were screened for resistance to 2,4-D by placing axenic shoot tips into rooting medium containing 0.05 µg/ml 2,4-D or by treatment of the shoot tip with 0.01 µg/ml 2,4-D. Root and new leaf morphologies were observed after 1 week.

Spray-test of greenhouse-grown plants

Tobacco plants were grown in the greenhouse to a height of 6 inches. Each plant was sprayed with approximately 10 ml of either 0.2%, 0.5% or 1.0% 2,4-D solution (pH 5.6) or with water as a control. The area covered by each spraying was approximately 500 cm^2, equivalent to 2,4-D dosages of 4, 10 and 20 kg/ha. Plant growth was observed 3 weeks after spraying. For cotton, primary transformants that were resistant to the root or meristem 2,4-D treatment were transferred to the greenhouse, grown to a height of approximately 8 inches and then sprayed with a range of 2,4-D concentrations (0, 10, 300, 1,000 and 3,000 mg/l) using a spray chamber designed to simulate field spray conditions. The spray pattern covered 3.7 ft^2 and at 3,000 ppm delivered the equivalent of 1.5 kg/ha of 2,4-D. Testing for herbicide resistance on the progeny from a transformed plant was conducted at the two-leaf stage using a 1.5 kg/ha commercial preparation of 2,4-D, and the plants were monitored for damage over a period of several weeks.

Molecular analysis of plant DNA

Plant genomic DNA was extracted from young leaves and PCR analysis (Saiki et al. 1988) was performed using two 20-bp primers: 5'-CACAATCCCACTATCCTTCG-3' (35S promoter sequence) and 5'-GACGGCATCGTCCAGGGTGG-3' (*tfd*A gene sequence). Approximately 1 µg of template DNA, 0.25 µg of each primer and 0.2 m*M* dNTP were used along with Taq polymerase and the buffer solution as recommended by the manufacturer (Promega). The PCR profile used was 94 °C, 1 min, 55 °C, 2 min, 72 °C, 2 min for a total of 30 cycles on a Perkin-Elmer Cetus DNA thermal cycler.

Results and discussion

Isolation of the 2,4-D monooxygenase (tfdA) gene from A. eutrophus plasmid pJP5

Genomic DNA from *A. eutrophus* and *E. coli* strains harboring pJP5 were screened for homology using synthetic oligonucleotides corresponding to the sequence of the pJP4-derived *tfd*A gene (Streber et al. 1987). A XbaI to SalI fragment of approximately 2 kg was detected from both strains by Southern hybridization. Gel-purified XbaI/SalI-cleaved DNA from *E. coli* (pJP5) was inserted into pUC19 to obtain a partial genomic library, and colony hybridization was used to identify plasmids containing DNA homologous to the pJP4-derived *tfd*A sequence. Homology to *tfd*A of pJP4 was confirmed by sequencing the 5' and 3' ends of the putative coding region.

As in the case of pJP4, GTG appeared to be the translation start codon of the *tfd*A gene in pJP5. For translation in plant cells, part of the 5' coding region was removed and replaced with a synthetic oligonucleotide containing an ATG start codon (Fig. 1). The modified gene was inserted between the cauliflower mosaic virus 35S promoter (35S) and the polyadenylation signal of the

COR-AAD1-00002441

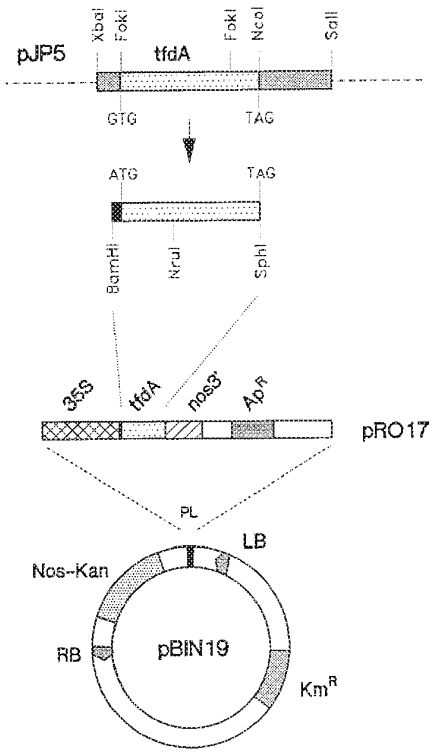


Fig. 1. Schematic representation of the *tfd*A gene from pJP5 leading to constructs pRO17 and pBIN19 : : pRO17. Only restriction sites relevant to the text are shown. *35S* cauliflower mosaic virus 35S promoter, *nos3'* polyadenylation signal-containing fragment from nopaline synthase gene, Ap^R bacterial ampicillin resistance gene, Km^R bacterial kanamycin resistance gene, *Nos-Kan* nopaline synthase promoter-kanamycin chimeric gene which confers resistance to kanamycin to plant cells, *RB* T-DNA right border sequence, *LB* T-DNA left border sequence, *PL* polylinker cloning site

nopaline synthase gene (*nos3'*) to produce pRO17 (Fig. 1). Since the CaMV 35S promoter has activity in bacteria (Assaad and Signer 1990), pRO17 was tested for *tfd*A expression in *E. coli* prior to introducing it into plants. Activity of 2,4-D monooxygenase in *E. coli* was determined by measuring the conversion of ^{14}C-labeled 2,4-D into $^{14}CO_2$ (Amy et al. 1985). As with *A. eutrophus* harboring pJP5, *E. coli* carrying pRO17 was efficient in releasing $^{14}CO_2$, showing that the modified *tfd*A gene in pRO17 encodes functional 2,4-D monooxygenase (Table 1).

Transgenic tobacco plants

Since tobacco can be transformed readily, it was used to evaluate the effectiveness of this chimeric construct. pRO17 was transferred into *Agrobacterium tumefaciens* as a cointegrate plasmid with pBIN19, a vector used for plant transformation (Bevan 1984) (Fig. 1). Kanamycin-

Table 1. 2,4-D monooxygenase activity in (A) bacteria, (B) tobacco and (C) cotton. For tobacco, leaf mesophyll protoplasts from primary transformants were used. For cotton, protoplasts from cotyledons and true leaves were isolated from two seedlings derived from self-pollination of plants 2, 3 and 14 and previously determined by PCR analysis to harbor the *tfd*A gene. Activity is defined as counts per minute (cpm) of $^{14}CO_2$ released after incubation in medium containing ^{14}C-labeled 2,4-D

	$^{14}CO_2$ (10^3 cpm)	Relative activity
(A) Bacteria		
E. coli	2.7	1
E. coli (pRO17)	198	73
A. eutrophus (pJP5)	111	41
(B) Tobacco		
Nontransformed control	0.15	1
Transformant 3	1.6	11
Transformant 4	10.4	69
Transformant 5	5.3	35
Transformant 7	4.3	29
Transformant 12	10.4	69
Transformant 13	12.7	85
(C) Cotton		
Nontransformed control	0.6	1
Transformant 2.1	22.1	37
Transformant 2.6	22.7	38
Transformant 3.1	9.9	17
Transformant 3.2	14.3	24
Transformant 14.1	21.8	36
Transformant 14.4	21.1	35

resistant plants were obtained following *Agrobacterium* infection of tobacco leaves. Leaves from putative transformed plants were placed on shoot-regeneration medium containing 2 mg/l, 4 mg/l and 6 mg/l 2,4-D. Seven transformants were able to produce shoots at 6 mg/l 2,4-D and two others at 4 mg/l 2,4-D. In contrast, nontransformed tobacco leaves were unable to produce shoots at 2 mg/l 2,4-D. Vegetatively derived clones of these primary transformants were transferred to the greenhouse and sprayed with varying doses of 2,4-D. Whereas non-transformed plants were sensitive at 4 kg/ha, one line was resistant to 10 kg/ha and seven lines were unaffected at levels up to 20 kg/ha (Fig. 2).

As evidence that the mechanism of 2,4-D resistance in the transgenic plants was due to degradation of the herbicide, $^{14}CO_2$ released from acid hydrolysis of glyoxylate, an enzymatic product of 2,4-D monooxygenase action on ^{14}C-labeled 2,4-D, was measured. Since plant cells can metabolize glyoxylate, excess unlabeled glyoxylate was supplemented to the incubation solution to reduce metabolism of the ^{14}C-labeled product. In a sampling of six lines of primary tobacco transformants, $^{14}CO_2$ production was 11- to 85-fold higher than that

COR-AAD1-00002442

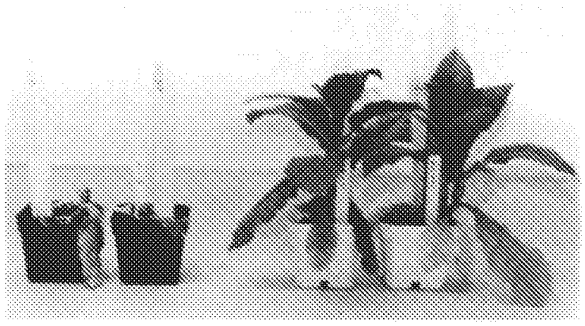


Fig. 2. Comparison of transgenic and wild-type tobacco plants sprayed with 2,4-D. Two wild-type plants (*left*) and two clonal replicas (*right*) of a representative *tfd*A transgenic line (no. 4) are shown 3 weeks after treatment with 2,4-D equivalent to 20 kg/ha

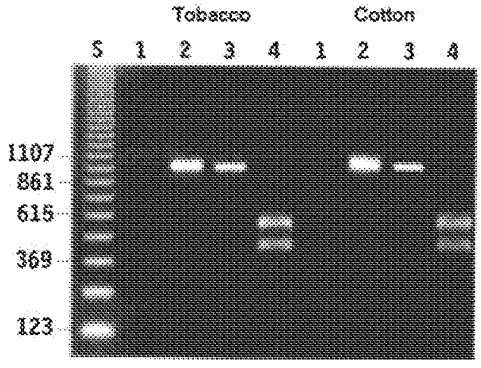


Fig. 3. PCR analysis of tobacco and cotton plants. PCR reactions examined for the presence of the predicted 940-bp *tfd*A-specific sequence in a non-transformed and a representative transformant are shown in *lanes 1* and *2*, respectively. *Lanes 3* and *4* show, respectively, no treatment and treatment with endonuclease NruI of the 940-bp PCR products purified from reactions on the same transformed tobacco or cotton DNA. NruI cleavage of the 940-bp product yielded two new fragments of the expected sizes of about 500 bp and 400 bp. Shown in lane *S* are size markers (BRL 123-bp DNA ladder), and relevant sizes (in bp) are as indicated

from non-transformed material (Table 1), indicating that resistance to 2,4-D was due to herbicide degradation.

In addition to enzyme activity, polymerase chain reaction (PCR) analysis was used to detect the presence of the 35S-*tfd*A gene fusion. Primers that anneal to the 35S promoter and the carboxyl end of the *tfd*A gene were used to amplify a fragment with an expected size of 940 bp. This 940-bp PCR fragment was indeed cleaved by endonuclease NruI to yield two fragments of approximately the expected sizes of 513 bp and 427 bp (Fig. 3). Seeds of two transgenic lines were examined in a germination assay. Whereas all 112 non-transformed seeds germinated only at 10^{-7} *M* 2,4-D, seeds derived from primary transformants numbers 3 and 4 germinated in 10^{-3} *M* 2,4-D with frequencies of 77% (62/80) and 74% (73/99), respectively. These 3 to 1 segregation ratios suggested that both primary transformants harbor a single 2,4-D resistance locus.

Transgenic cotton plants

The central focus of this work was to engineer 2,4-D tolerance into cotton. The same chimeric *tfd*A gene in pBIN19 : : pRO17 was therefore introduced into cotton by *Agrobacterium*-mediated transformation. Shoots of putative transformed plants were placed on rooting medium containing 0.05 mg/l 2,4-D and/or meristems were treated with 0.01 mg/l 2,4-D. Approximately 300 transformants from 30 cell lines rooted normally and/or were unaffected by the meristem treatment. Vegetatively derived clones of the primary transformants were transferred to the greenhouse and sprayed with varying levels of 2,4-D. Transformants were unaffected up to 3 times the recommended spray rate of 2,4-D, whereas non-transformed plants were severely affected at 300 times less the recommended spray rate of the herbicide. Higher levels of 2,4-D were not tested.

Progeny derived from self-fertilization of three transformed lines (numbers 2, 3, 14) were analyzed further. PCR analysis confirmed the presence of the *tdf*A gene in representative progeny seedlings (Fig. 3). Two *tfd*A transgenic seedlings from each line, as shown by PCR analysis, were assayed for 2,4-D monooxygenase. Activity 17- to 38-fold above that of the non-transformed control was found from the *tfd*A-containing plants (Table 1). In a germination assay, progeny from wild-type plants were unable to form roots on medium containing 0.1 mg/l 2,4-D. In contrast, nearly three-quarters of the germinating progeny seedlings from each of the three transformed lines were unaffected by the herbicide. The 3 : 1 segregation pattern is indicative of a single 2,4-D resistance locus. Direct spraying of the herbicide was conducted on the progeny of primary transformant no. 2. Plants germinated in the absence of 2,4-D were sprayed at the two-leaf stage with a commercial preparation of the herbicide set at a dosage equivalent of 1.5 kg/ha (Fig. 4). This is approximately 3 times the manufacturer's highest recommended rate for wheat, corn, sorghum and pastures (0.2, 0.2, 0.2 and 0.6 kg/ha, respectively). Whereas all 18 wild-type plants were killed within 5 days, 83 out of 107 progeny plants derived from transformant no. 2 were unaffected by the herbicide treatment.

Concluding remarks

Our overall goal was to reduce annual losses of cotton due to drift-levels of 2,4-D. Towards this aim, we isolated and modified a bacterial gene encoding 2,4-D monooxygenase and tested its effectiveness in tobacco. The result-

COR-AAD1-00002443



Fig. 4. Comparison of wild-type and transgenic cotton plants 3 weeks after treatment with a commercial preparation of 2,4-D: wild-type cotton sprayed with 0.15 kg/ha 2,4-D (*leftmost*), and representative progeny plants from transgenic line no. 2 treated with 0.15 kg/ha (*second from left*), 0.5 kg/ha (*second from right*) and 1.5 kg/ha of the herbicide (*rightmost*)

ing transgenic tobacco plants produced functional 2,4-D monooxygenase and were resistant to the herbicide. In previous reports, Streber and Willmitzer (1989) and Lyon et al. (1989) also obtained 2,4-D-resistant tobacco plants using a *tfd*A gene isolated from a different *A. eutrophus* plasmid. Transgenic tobacco plants were found to be resistant to 2,4-D at 10 kg/ha and 6 kg/ha, respectively, which are in the same range of resistance as the tobacco plants described here. In the case of cotton, tolerance to at least 1.5 kg/ha was found, which is 3 times the highest recommended level used for wheat, corn, sorghum and pasture. Moreover, this level of resistance was achieved using commercial preparations of the herbicide that contain emulsifiers and surfactants. The degree of tolerance achieved should be more than adequate to protect cotton from drift-levels of the herbicide.

Acknowledgements. We thank Greg Cartwright for technical assistance. This work was funded by the Agricultural Research Service, U.S. Department of Agriculture.

References

Amy PS, Schulke JW, Frazier LM, Seidler RJ (1985) Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-Dichlorophenoxyacetic acid. Appl Environ Microbiol 49:1237–1245

Assaad FF, Signer ER (1990) Cauliflower mosaic virus P35S promoter activity in *Escherichia coli*. Mol Gen Genet 223:517–520

Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation. Nucleic Acids Res 12:8711–8721

Don RH, Pemberton JM (1981) Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J Bacteriol 145:681–686

Horsch RB, Fry JE, Hoffman NL, Eichholtz D, Rogers SG, Fraley RT (1985) A simple and general method for transferring genes into plants. Science 227:1229–1231

Lyon BR, Llewellyn DJ, Huppatz JL, Dennis ES, Peacock WJ (1989) Expression of a bacterial gene in transgenic tobacco plants confers resistance to the herbicide 2,4-dichlorophenoxyacetic acid. Plant Mol Biol 13:533–540

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Ow DW, Wood KV, DeLuca M, deWet JR, Helinski DR, Howell SH (1986) Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. Science 234:856–859

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer directed amplication of DNA with a thermostable DNA polymerase. Science 239:487–491

Stewart J McD, Hsu CL (1977) In ovulo embryo culture and seedling development of cotton (*Gossypium hirsutum* L.). Planta 137:113–117

Streber WR, Willmitzer L (1989) Transgenic tobacco plants expressing a bacterial detoxifying enzyme are resistant to 2,4-D. Biotechnology 7:811–816

Streber WR, Timmis KN, Zenk MH (1987) Analysis, cloning, and high-level expression of 2,4-dichlorophenoxyacetate monooxygenase gene *tfd*A of *Alcaligenes eutrophus* JMP134. J Bacteriol 169:2950–2955

Takebe I, Labib G, Melchers G (1971) Regeneration of whole plants from isolated mesophyll protoplasts of tobacco. Naturwissenschaften 88:318–320

Trolinder N, Goodin JR (1987) Somatic embryogenesis and plant regeneration in cotton (*Gossypium hirsutum*). Plant Cell Reports 6:231–234

Umbeck P, Johnson G, Barton K, Swain W (1987) Genetically transformed cotton (*Gossypium hirsutum* L.) plants. Bio/Technology 5:263–266

COR-AAD1-00002444

ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS
Vol. 300, No. 2, February 1, pp. 738–746, 1993

Purification of 3,5-Dichlorocatechol 1,2-Dioxygenase, a Nonheme Iron Dioxygenase and a Key Enzyme in the Biodegradation of a Herbicide, 2,4-Dichlorophenoxyacetic acid (2,4-D), from *Pseudomonas cepacia* CSV90[1]

Manzoor A. Bhat,[2] Tetsuo Ishida, Kihachiro Horiike, Chelakara S. Vaidyanathan,* and Mitsuhiro Nozaki[3]

*Department of Biochemistry, Shiga University of Medical Science, Seta, Ohtsu, Shiga 520-21, Japan; and *Department of Biochemistry, Indian Institute of Science, Bangalore 560 012, India*

Received September 9, 1992

An enzyme which cleaves the benzene ring of 3,5-dichlorocatechol has been purified to homogeneity from *Pseudomonas cepacia* CSV90, grown with 2,4-dichlorophenoxyacetic acid (2,4-D) as the sole carbon source. The enzyme was a nonheme ferric dioxygenase and catalyzed the intradiol cleavage of all the examined catechol derivatives, 3,5-dichlorocatechol having the highest specificity constant of 7.3 $\mu M^{-1} s^{-1}$ in an air-saturated buffer. No extradiol-cleaving activity was observed. Thus, the enzyme was designated as 3,5-dichlorocatechol 1,2-dioxygenase. The molecular weight of the native enzyme was ascertained to be 56,000 by light scattering method, while the M_r value of the enzyme denatured with 6 M guanidine–HCl or sodium dodecyl sulfate was 29,000 or 31,600, respectively, suggesting that the enzyme was a homodimer. The iron content was estimated to be 0.89 mol per mole of enzyme. The enzyme was deep red and exhibited a broad absorption spectrum with a maximum at around 425 nm, which was bleached by sodium dithionite, and shifted to 515 nm upon anaerobic 3,5-dichlorocatechol binding. The catalytic constant and the K_m values for 3,5-dichlorocatechol and oxygen were 34.7 s^{-1} and 4.4 and 652 μM, respectively, at pH 8 and 25°C. Some heavy metal ions, chelating agents and sulfhydryl reagents inhibited the activity. The NH_2-terminal sequence was determined up to 44 amino acid residues and compared with those of the other catechol dioxygenases previously reported.

© 1993 Academic Press, Inc.

2,4-Dichlorophenoxyacetic acid (2,4-D)[4] and its related compounds have been used extensively as hormone herbicides, resulting in the pollution of soil and natural water (1). Once their accumulation touches enormity, these compounds are very hazardous to life (2). Thus, the metabolisms of 2,4-D by various organisms have been studied by many workers (e.g., 3). A metabolic pathway of 2,4-D in some bacteria such as *Pseudomonas* sp. (4), *Arthrobacter* sp. (5), *Flavobacterium* sp. (6), and *Alcaligenes eutrophus* JMP134 (7) has been proposed mainly by the identification of metabolites from the spent media and/or the enzyme activities in the crude extracts (Fig. 1). First, the ether bond in 2,4-D is cleaved to produce 2,4-dichlorophenol by a 2,4-D monooxygenase, but details of this reaction are unclear. The gene for the monooxygenase was cloned from plasmid pJP4 of *Alcaligenes eutrophus*

[1] This research was supported in part by a Grant-in-Aid for Scientific Research 02454565 from the Ministry of Education, Science and Culture of Japan.

[2] Senior Research Fellow of the Indian Institute of Science, Bangalore, and of Council of Scientific and Industrial Research, New Delhi, India. Recipient of a scholarship for Research Students from the Ministry of Education, Science and Culture of Japan. On leave from the Department of Biochemistry, Indian Institute of Science, Bangalore, India.

[3] To whom all correspondence and reprint requests should be addressed.

[4] Abbreviations used: 2,4-D, 2,4-dichlorophenoxyacetic acid; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; LS, the peak height of light scattering intensity in arbitrary unit; RI, the peak height of differential refractometry signal in arbitrary unit; SD, standard deviation.

0003-9861/93 $5.00
Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.

COR-AAD1-00002445

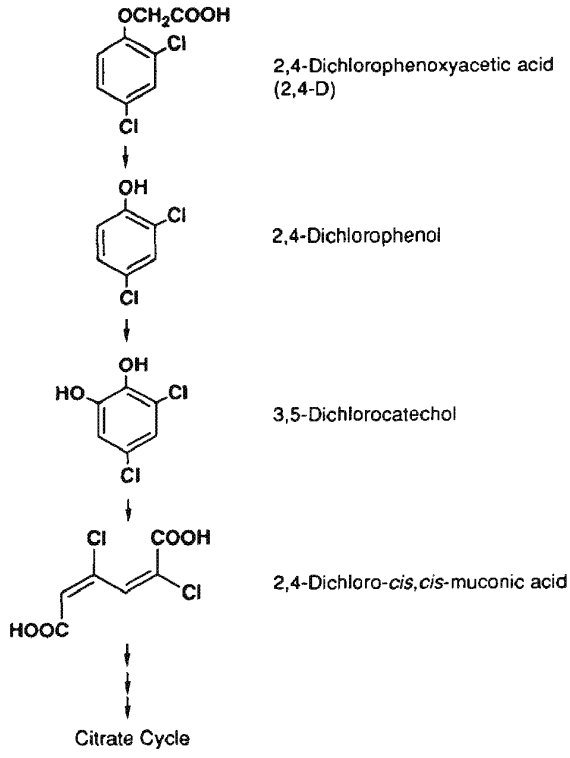


FIG. 1. A proposed metabolic pathway of 2,4-D in some soil bacteria.

JMP134 (8). The hydroxylation of 2,4-dichlorophenol to 3,5-dichlorocatechol, the second step, is catalyzed by 2,4-dichlorophenol 6-monooxygenase, an NADPH-dependent flavomonooxygenase. We already purified and characterized this enzyme (9). 3,5-Dichlorocatechol appears to be cleaved by a dioxygenase. However, no detailed information on the enzyme has been available.

In this paper, we describe, for the first time, the purification and characterization of the dioxygenase from the organism involved in the metabolism of a herbicide, 2,4-D. The purified enzyme, designated as 3,5-dichlorocatechol 1,2-dioxygenase, was distinct from the other chlorocatechol dioxygenases reported previously (10, 11). We discuss the comparison of the properties of halogenated catechol dioxygenases.

EXPERIMENTAL PROCEDURES

Materials. 2,4-D and a standard solution of iron were purchased from Nacalai Tesque. 3,5-Dichlorosalicylaldehyde was obtained from Sigma. 3-Chloro-, 4-chloro-, 3-methyl-, 4-methyl-, and 3-methoxy-catechol were from Tokyo Kasei Kogyo. The standard proteins for light scattering measurement, high-speed gel filtration chromatography, and SDS–PAGE were obtained from either Oriental Yeast, Sigma, or Boehringer–Mannheim. Crystalline catechol 2,3-dioxygenase (EC 1.13.11.2) was prepared from *Escherichia coli* W3110 carrying the dioxygenase gene from TOL plasmid, according to the previously published method (12), except that Q-Sepharose was used instead of DEAE-cellulose. All other chemicals used were of analytical grade.

Biological material. The bacterium isolated from an industrial effluent contaminated soil as reported previously (9), now designated as *Pseudomonas cepacia* strain CSV90,[5] was used. The bacterium was cultured by the method previously described (9): Briefly, the strain was grown in basal salts medium, pH 7.2, containing 0.2% 2,4-D as the sole source of carbon, on large scale using a 50-liter jar fermentor incubated at 30°C and air supplied at the rate of 8 liter/min. The cells were harvested after 24 h of growth, and either kept frozen at −20°C or processed directly for enzyme purification. No loss of activity was observed during storage for at least 2 months.

Synthesis of 3,5-dichlorocatechol. 3,5-Dichlorocatechol was synthesized by a modified procedure of Dakin oxidation (13): 3,5-Dichlorosalicylaldehyde (5.0 g) was taken in a 100-ml round-bottomed flask, 30 ml of 1 M NaOH was added, and the mixture was stirred. After 3,5-dichlorosalicylaldehyde had completely dissolved, 6 ml of 30% H_2O_2 was added and stirring was continued. The color of the reaction mixture changed from yellow to reddish pink with increase in temperature. The reaction mixture was allowed to cool down, acidified by diluted H_2SO_4, and then neutralized by addition of a saturated $NaHCO_3$ solution (100 ml). The products were then extracted from reaction mixture with diethyl ether three times using 150 ml each time. The ether extracts were combined, washed with a saturated NaCl solution (100 ml), dried over anhydrous Na_2SO_4 (100 g), and decanted followed by evaporation to dryness under reduced pressure. The material left behind was crystallized from toluene.

We confirmed the synthesized compound as 3,5-dichlorocatechol by analyzing it with ^1H and ^{13}C NMR spectroscopies and mass spectrometry. The melting point of the compound was determined to be 90.8 ± 0.5°C (uncorrected). The elemental analysis of the compound also gave satisfactory results.

Enzyme assay and kinetics. The activity of 3,5-dichlorocatechol 1,2-dioxygenase was assayed by measuring the increase in absorbance at 268 nm due to the formation of 2,4-dichloromuconic acid. The reaction mixture (3.0 ml) contained 333 μM 3,5-dichlorocatechol in 0.05 M Tris–HCl buffer, pH 8.0, and the reaction was initiated by the addition of the enzyme at 25°C. The difference of molar absorption coefficients of 2,4-dichloromuconic acid and 3,5-dichlorocatechol at 268 nm was experimentally obtained to be 13,400 M^{-1} cm^{-1}. One unit is defined as the amount of the enzyme that produces 1 μmol of 2,4-dichloromuconic acid per minute under the standard assay conditions.

The stoichiometry of the reaction was determined by measuring O_2 uptake with a Yellow Springs Model 53 oxygen electrode with limited amounts of the substrate. The electrode was calibrated by measuring the O_2 consumption due to the oxygenation of the known amounts of catechol catalyzed by catechol 2,3-dioxygenase (14). The O_2 concentration in the assay buffer was determined to be 260 μM.

Different concentrations of O_2 required for the determination of kinetic parameters were obtained by bubbling the reaction mixture with known concentrations of an O_2–N_2 gas mixture for 5 min prior to the addition of the enzyme. The O_2 concentrations were checked with an oxygen electrode and by absorbance of the product formed by catechol 2,3-dioxygenase (14). Apparent K_m and V_{max} values for various catechol derivatives were determined in the air-saturated buffer.

Absorption spectrum. Visible and ultraviolet spectra of the enzyme, enzyme–substrate complex, various substrates, and the products were recorded on Shimadzu recording spectrophotometers, uv-300 or uv-2200.

Protein concentration. The protein concentration was determined by the method of Lowry *et al.* (15) using bovine serum albumin as the standard protein. The purified enzyme concentration was also deter-

[5] We here designate this isolate as *Pseudomonas cepacia* CSV90, since reported strains of *P. cepacia* were 2,4-D nondegrading.

COR-AAD1-00002446

mined by amino acid analysis. The values obtained differed by less than 15%.[6]

Metal analysis. The enzyme preparations were dialyzed overnight at 4°C against 500 ml of 0.05 M Tris–HCl buffer (pH 8.0). The samples were prepared by adding known amounts of iron (0–5 μg) and treated with concentrated H_2SO_4 and H_2O_2 according to the method of Yonetani (16). The iron contents were determined at least with triplicate samples on a Shimadzu atomic absorption spectrophotometer (Model AA660) by using the method of standard addition. Equal volumes of the outer solution treated entirely analogous to the sample solution were used as the blank and the standard.

Molecular weight determination. Low-angle laser light scattering measurement combined with gel chromatography (17) was carried out at 21°C. A TSK-GEL G3000SW column (7.5 × 675 mm), an LS-8000 low-angle laser light scattering photometer, a uv-8000 uv–vis detector, and an RI-8000 differential refractometer, all from Tosoh, were used. The eluent used was 0.1 M sodium phosphate containing 0.2 M NaCl (pH 7.1), and the flow rate was 0.5 ml/min. The value of the specific refractive increment of the enzyme solution was assumed to be the same as the standard proteins.

The molecular weight of the enzyme was also estimated by high-speed gel filtration method (18, 19) with a TSK-GEL G3000SW column (7.5 × 675 mm). The eluent used was 0.1 M sodium phosphate containing 0.2 M NaCl (pH 7.1). Chromatographic data were analyzed by the methods described previously (18–23).

The subunit molecular weight was estimated according to the method of Ui (22, 24), using TSK-GEL G3000SW column (7.5 × 600 mm) under denaturing conditions using 6 M guanidine–HCl in 0.01 M sodium phosphate, pH 6.5.

Polyacrylamide gel electrophoresis. SDS–PAGE was performed according to the method of Laemmli (25). Proteins were stained with 0.25% Coomassie brilliant blue R-250 in 45% methanol–9% acetic acid.

Amino acid analysis and NH_2-terminal amino acid sequence. The amino acid analysis was carried out with a Hitachi L-8500 amino acid analyzer. The enzyme samples were hydrolyzed in the vapor phase of 6 M HCl containing 0.01% phenol at 110°C for 24, 48, and 72 h in evacuated reaction vials from Waters Pico Tag work station. For the determination of tryptophan and cysteine, the spectrophotometric methods of Edelhoch (26) and Ellman (27) were used, respectively.

The amino acid sequence of the NH_2-terminal portion of the protein was determined using a gas phase protein sequencer, Applied Biosystems Model 470A, equipped with an on-line HPLC, Model 120A (28).

RESULTS

Purification of 3,5-dichlorocatechol 1,2-dioxygenase. All purification steps were carried out at about 4°C.

Step 1: Crude extract. The frozen cells, 50 g wet wt, were suspended in 0.05 M Tris–HCl buffer (pH 8.0) in a final volume of 500 ml and disrupted in 250-ml batches by sonication. Cell debris was removed by centrifugation at 20,000*g* for 15 min. The supernatant was again centrifuged at 160,000*g* for 1 h to remove particulate materials. The supernatant was collected and designated crude extract.

Step 2: DEAE-Sephacel chromatography. The crude extract (400 ml) was directly loaded on a DEAE-Sephacel column (Pharmacia, 2.5 × 25 cm) equilibrated with 0.05 M Tris–HCl buffer, pH 8.0, and then washed with 200 ml of the same buffer. The enzyme was eluted with a linear concentration gradient established between 500 ml each of the buffer and the one containing 0.5 M KCl. The elution profiles of proteins and the enzyme activity are given in Fig. 2A. The active fractions were pooled together.

Step 3: Phenyl-Sepharose chromatography. The pooled fraction obtained from the previous step was brought to 33% ammonium sulfate saturation and stirred for 30 min. The resulting precipitate was removed by centrifugation at 20,000*g* for 15 min. The supernatant was loaded on a phenyl-Sepharose column (Pharmacia, 2 × 16 cm) equilibrated with 0.05 M Tris–HCl buffer (pH 8.0) containing ammonium sulfate (33% saturation). After washing the column with 150 ml of the same buffer, the enzyme was eluted with a negative gradient of ammonium sulfate from 33 to 0% saturation and a positive gradient of ethylene glycol from 0 to 70%, each in 250 ml of 0.05 M Tris–HCl buffer, pH 8.0. The elution profiles are presented in Fig. 2B. The active fractions were collected, dialyzed overnight against 0.05 M Tris–HCl buffer (pH 8.0), and concentrated to ~3 ml with Amicon Centriflo membrane cones (Type CF25).

Step 4: Gel filtration chromatography. Preparative gel filtration was carried out on a TSK-GEL G3000SW column equilibrated with 0.05 M Tris–HCl, pH 7.5, contain-

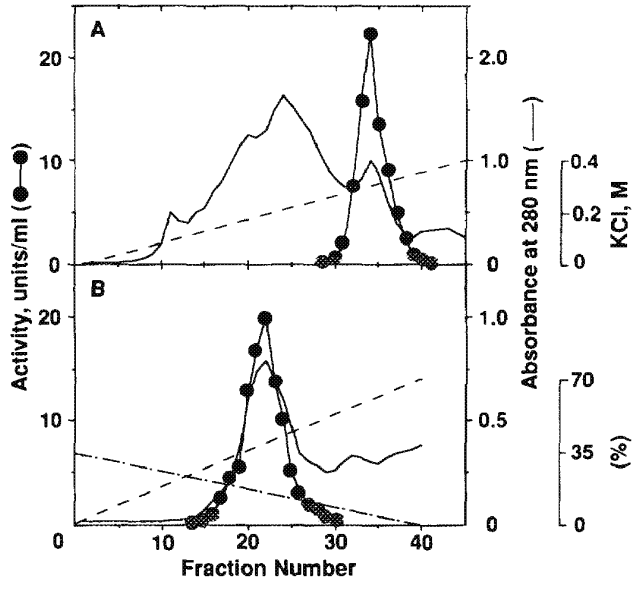


FIG. 2. Chromatograms of 3,5-dichlorocatechol 1,2-dioxygenase on a DEAE-Sephacel column (A) and on a phenyl-Sepharose column (B). Each chromatography was performed as described in the text. The volume of each fraction was 16 ml for A and 12.5 ml for B. In (B) the gradients of ammonium sulfate (% saturation) and ethylene glycol (%) are shown by (- · -) and (- - -), respectively.

[6] The correction factor of protein determination by the Lowry method (15) was calculated to be 1.15 from the average of the amino acid analyses of Asp, Glu, Gly, Ala, Val, Ile, Leu, Lys, His, Arg, and Pro (the result in Table III), and M_r = 56,000 (Table II).

COR-AAD1-00002447

TABLE I

Purification of 3,5-Dichlorocatechol 1,2-Dioxygenase

Purification steps	Total volume (ml)	Total activity (units)	Total protein (mg)	Specific activity (units/mg)	Yield (%)
Crude extract	400	1390	2720	0.51	100
DEAE-Sephacel	190	1310	332	3.95	94.2
Phenyl-Sepharose	190	1140	162	7.04	82.0
Gel filtration	4	800	72	11.1	57.6

ing 0.3 M KCl, and the same buffer was used as the eluent. The active fractions were concentrated with Amicon Centriflo membrane cones (Type CF25) and simultaneously equilibrated with 0.05 M Tris–HCl buffer, pH 8.0.

The purification procedure is summarized in Table I. The overall purification was 22-fold with a yield of 58%.

Purity, molecular weight, and subunit composition. The purified enzyme eluted as a single peak on TSK-GEL G3000SW gel chromatography without or with 6 M guanidine–HCl, and the M_r values were estimated to be 66,000 (data not shown) and 29,000 (Fig. 3B and Table II), respectively. On SDS–PAGE, the enzyme gave a single band; the estimated M_r value was 31,600 (Fig. 3C). On the other hand, the M_r value of the native enzyme was obtained to be 56,000 by light scattering method (Fig. 3A). These results suggested that the purified enzyme was homogeneous and consists of two identical subunits, i.e., a homodimer α_2 (Table II).

Amino acid composition and NH_2-terminal amino acid sequence. The amino acid composition of the enzyme is shown in Table III. During the automated Edman degradation of the enzyme up to 44 residues, a single amino acid was released at each step (Fig. 4), suggesting that the enzyme is composed of identical subunits. This also supported that the enzyme is a homodimer.

Iron content. After overnight dialysis against 0.05 M Tris–HCl buffer (pH 8.0), the enzyme preparation showed essentially no loss of the activity, and the iron content was determined to be 0.89 mol of iron per mole of enzyme, based on the enzyme concentration determined by Lowry method (15) and M_r = 56,000 (Table II). Overnight dialysis against the same buffer containing an iron chelator, *o*-phenanthroline (0.1 mM), did not change the iron content in the preparation.

Absorption spectrum. The enzyme showed a broad absorption spectrum with a maximum at about 425 nm

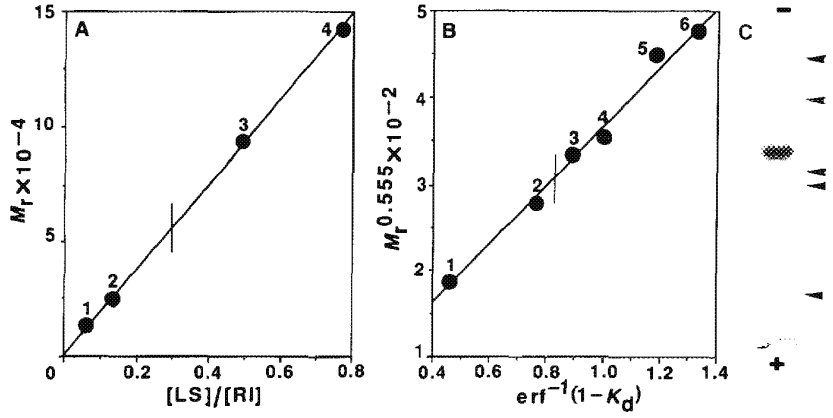


FIG. 3. Molecular weight determination of 3,5-dichlorocatechol 1,2-dioxygenase. Experiments were performed as described under Experimental Procedures. (A) The light scattering method coupled with high-speed gel chromatography on TSK-GEL G3000SW. The value of M_r is plotted against the ratio of the peak height of light scattering (LS) and that of differential refractometry (RI). Each value is an average of at least three determinations. The vertical bar shows the value obtained for the enzyme. The standard proteins were: 1, horse heart cytochrome c; 2, yeast adenylate kinase; 3, yeast enolase; and 4, pig heart lactate dehydrogenase. (B) Gel filtration on TSK-GEL G3000SW column in 6 M guanidine–HCl. The value of $M_r^{0.555}$ is plotted against $erf^{-1}(1 - K_d)$, where K_d is the distribution coefficient and $erf^{-1}(1 - K_d)$ is the inverse error function of $1 - K_d$ (20, 21). The vertical bar shows the value obtained for the enzyme. The protein standards used were: 1, horse heart cytochrome c; 2, bovine pancreas chymotrypsin; 3, catechol 2,3-dioxygenase from *P. putida* mt-2; 4, rabbit muscle aldolase; 5, bovine liver catalase; 6, bovine serum albumin. (C) SDS–PAGE of the dioxygenase (5.5 μg) on a 12.5% gel. The marker proteins used were bovine serum albumin, ovalbumin, bovine erythrocyte carbonic anhydrase, bovine pancreas trypsinogen, and egg white lysozyme from top to the bottom. The positions of each marker protein are shown by the arrowheads.

COR-AAD1-00002448

TABLE II

Molecular and Kinetic Properties of 3,5-Dichlorocatechol 1,2-Dioxygenase

Molecular and kinetic properties	Value
Molecular weight of native enzyme	
HPLC–LS–RI method	56,000
Gel filtration	66,000
Molecular weight of subunit	
Gel filtration in 6 M guanidine–HCl	29,000
SDS–PAGE	31,600
Subunit composition	α_2
Iron content (mol/mol enzyme)	0.89
Kinetic properties at 25°C and pH 8.0	
Specific activity (μmol/min/mg)	11.1
k_{cat} (s^{-1})	34.7
K_m for 3,5-dichlorocatechol (μM)	4.40
K_m for oxygen (μM)	652



FIG. 4. NH_2-terminal sequence of 3,5-dichlorocatechol 1,2-dioxygenase. Sequencing was carried out as described under Experimental Procedures.

($\epsilon \approx 4700\ M^{-1}\ cm^{-1}$ based on iron concentration) in the visible region (Fig. 5A). The spectral properties were similar to those of nonheme ferric dioxygenases (11, 29, 30). When the enzyme was treated with sodium dithionite, the visible absorption almost disappeared, and reappeared again by aeration, suggesting that the iron in the enzyme is in the ferric state.

The visible spectrum was not altered by exhaustive evacuation nor did it change by the addition of oxygen. However, when the organic substrate, 3,5-dichlorocatechol, was added to the enzyme under anaerobic conditions, an immediate change of color from deep red to purple was observed: The spectrum was characterized with the large red shift of the broad 425-nm band of the enzyme to ~515 nm, an increase in the absorption in the long wavelength region, and a decrease in the absorption at around 390 nm (Fig. 5A). This spectral change was completely reversed after the consumption of 3,5-dichlorocatechol by the introduction of atmospheric oxygen. Since the mixture of the catechol and inorganic iron did not show such spectral change under the same conditions, and neither the organic substrate nor the product had absorption in the visible region (Fig. 5B), the above results suggested a complex formation of 3,5-dichlorocatechol with the ferric ion at the active site.

Identification of the product and the reaction stoichiometry. 3,5-Dichlorocatechol, the product of 2,4-dichlorophenol 6-monooxygenase reaction (9), served as the substrate for the present enzyme. The absorption spectrum of 3,5-dichlorocatechol revealed two peaks at 247 and 298 nm, and a trough at 274 nm at pH 8.0 (Fig. 5B). No absorption was observed in the visible region. The molar absorption coefficient at 298 nm was calculated to be $2.61 \times 10^3\ M^{-1}\ cm^{-1}$.

Upon addition of a catalytic amount of the purified enzyme to a 3,5-dichlorocatechol solution in 0.05 M Tris–HCl buffer, pH 8.0, the 247- and 298-nm absorption bands started disappearing with concomitant appearance of a new absorption band at around 268 nm. Two isosbestic points were observed at 230 and 297 nm, indicating a single product formation. The final absorption spectrum

TABLE III

Amino Acid Composition per Subunit of 3,5-Dichlorocatechol 1,2-Dioxygenase[a]

Amino acid	No.	Amino acid	No.	Amino acid	No.	Amino acid	No.
Asp	28.8	Ala	19.9	Tyr	8.8	Pro	16.1
Thr	22.6[b]	Val	28.6[c]	Phe	8.1	Cys	2.2[d]
Ser	10.0[b]	Met	5.1[b]	Lys	10.7	Trp	3.8[e]
Glu	24.3	Ile	11.8	His	6.7		
Gly	19.6	Leu	14.1	Arg	13.0	Total	256[f]

[a] Unless otherwise indicated, the results are presented as averages of two separate analyses after hydrolysis for 24 and 48 h, and residues were calculated based on subunit molecular weight of 28,000.

[b] Values extrapolated to zero-time hydrolysis.

[c] Values at 72 h hydrolysis.

[d] Determined by the spectrophotometric method of Ellman (27).

[e] Determined by the spectrophotometric method of Edelhoch (26).

[f] Summation of the neartest integer of each residue.

COR-AAD1-00002449

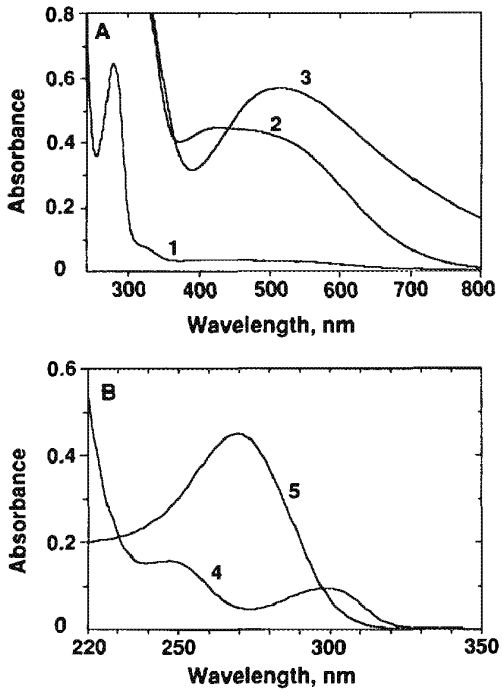


FIG. 5. Absorption spectra of 3,5-dichlorocatechol 1,2-dioxygenase (A), and 3,5-dichlorocatechol and the reaction product (B) in 0.05 M Tris-HCl buffer, pH 8.0, at 25°C. (curves 1 and 2) The enzyme concentrations were 0.44 and 5.9 mg/ml, respectively. (curve 3) The enzyme (5.9 mg/ml) was mixed with 3,5-dichlorocatechol (667 μM) after exhaustive evacuation in a Thunberg cuvette. (curve 4) 24.9 μM 3,5-dichlorocatechol; (curve 5) the product formed by complete conversion of the substrate (see text).

revealed a peak at 268 nm and no absorption in the visible region (Fig. 5B). The molar absorption coefficient of the product was calculated to be 1.48×10^4 M^{-1} cm^{-1} at 268 nm at pH 8.0 (Table IV). This type of spectrum is characteristic of muconic acids formed by the intradiol cleavage of the catechol rings (31). Thus, we concluded that the product of this reaction was 2,4-dichloromuconic acid when the substrate was 3,5-dichlorocatechol.[7]

We analyzed the stoichiometry of this reaction using various substrate concentrations (40.5–119 μM) and monitored the amount of oxygen consumed. From nine separate analyses, the mole ratio of the amount of substrate used to O_2 uptake was found to be 1.09 ± 0.04 (SD). When the substrate to product conversion was monitored at 268 nm, the amount of dichlorocatechol added to the reaction mixture was equivalent to that of the muconic acid formed in terms of the mole to mole ratio within experimental errors, indicating that only one product was formed during the reaction.

These data suggested that the enzyme was an intradiol-cleaving dioxygenase catalyzing the reaction; 3,5-dichlorocatechol + O_2 → 2,4-dichloromuconic acid.

Effect of pH on activity and stability of the enzyme. The optimal pH for 3,5-dichlorocatechol was around 8.0. Below pH 6.5, the activity decreased drastically and at pH 6.0 almost no activity was observed when assayed polarographically (data not shown). The enzyme was stable at pH 8.0 for several weeks and thereafter progressive loss of activity was observed.

Kinetic constants and substrate specificity. Initial velocity studies of the enzyme were carried out as functions of 3,5-dichlorocatechol and oxygen concentrations (Fig. 6). The apparent V_{max} values almost linearly increased with the increase of oxygen concentration, suggesting a high K_m value for oxygen. The experimentally available highest $[O_2]$ was ~500 μM because 3,5-dichlorocatechol

[7] We tried to isolate and analyze the product by ^1H and ^{13}C NMR spectroscopies and mass spectrometry. However, these attempts were unsuccessful because of chemical lability of the product as well as of the substrate. The authentic sample of the product was also unavailable.

TABLE IV

Substrate Specificity of 3,5-Dichlorocatechol 1,2-Dioxygenase[a]

Substrates	K_m^a (μM)	k_{cat}^a (s^{-1})	k_{cat}^a/K_m^a (μM^{-1} s^{-1}) (%)	Products	
				λ_{max} (nm)	ϵ (M^{-1} cm^{-1})
3,5-Dichlorocatechol	1.5	11	7.3 (100)	268	14,800
4-Chlorocatechol	1.8	6.7	3.7 (51)	257	14,000
3-Chlorocatechol	2.6	7.8	3.0 (41)	267	17,900
4-Methylcatechol	9.0	15	1.7 (23)	254	13,800
3-Methylcatechol	37	19	0.51 (7.0)	259	18,500
Catechol	16	4.9	0.31 (4.2)	255	16,400
3-Methoxycatechol	14	2.2	0.16 (2.1)	279	18,900

[a] The activities were assayed spectrophotometrically at 25°C in air-saturated buffer of 0.05 M Tris-HCl, pH 8.0. The values of apparent kinetic constants (K_m^a and k_{cat}^a) were obtained from [S]/v versus [S] plots, assuming the molecular weight of 56,000 for the enzyme.

COR-AAD1-00002450