

(12) PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. AU 199666184 B2
(10) Patent No. 718521

(54) Title
Isolated DNA sequence for use as a regulator region in a chimeric gene useful for transforming plants

(51)⁷ International Patent Classification(s)
C12N 015/82 C12N 005/10
A01H 005/00 C12N 015/29
C12N 001/21

(21) Application No: **199666184**

(22) Application Date: **1996.07.17**

(87) WIPO No: **WO97/04114**

(30) Priority Data

(31) Number	(32) Date	(33) Country
95/08980	1995.07.19	FR

(43) Publication Date : **1997.02.18**

(43) Publication Journal Date : **1997.04.17**

(44) Accepted Journal Date : **2000.04.13**

(71) Applicant(s)
Rhone-Poulenc Agro

(72) Inventor(s)
Richard Derose; Nicole Chaubet; Claude Gigot

(74) Agent/Attorney
DAVIES COLLISON CAVE, 1 Little Collins Street, MELBOURNE VIC 3000

(56) Related Art
AU 11443/92

OPI DATE 18/02/97 APPLN. ID 66184/96
AOJP DATE 17/04/97 PCT NUMBER PCT/FR96/01109



DEN

AU9666184

(PCT)

(51) Classification internationale des brevets ⁶ :
C12N 15/82, 15/29, 1/21, 5/10, A01H 5/00

A2

(11) Numéro de publication internationale: **WO 97/04114**

(43) Date de publication internationale: 6 février 1997 (06.02.97)

(21) Numéro de la demande internationale: PCT/FR96/01109

(22) Date de dépôt international: 17 juillet 1996 (17.07.96)

(30) Données relatives à la priorité:
95/08980 19 juillet 1995 (19.07.95) FR

(71) Déposant (pour tous les Etats désignés sauf US): RHONE
POULENC AGROCHIMIE [FR/FR]; 14-20, rue Pierre-
Baizet, F-69009 Lyon (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (US seulement): DEROSE, Richard
[FR/FR]; 216, rue de Saint-Cyr, F-69009 Lyon (FR).
CHAUBET, Nicole [FR/FR]; Esc. C, 13 A, boulevard
Wilson, F-67800 Strasbourg (FR). GIGOT, Claude [FR/FR];
Esc. C, 13 A, boulevard Wilson, F-67800 Strasbourg (FR).

(74) Représentant commun: RHONE POULENC AGROCHIMIE;
Boîte postale 9163, F-69263 Lyon Cédex 09 (FR).

(81) Etats désignés: AL, AU, BB, BG, BR, CA, CN, CU, CZ, EE,
GE, HU, IL, IS, JP, KP, KR, LK, LR, LT, LV, MG, MK,
MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US,
UZ, VN, brevet ARIPO (KE, LS, MW, SD, SZ, UG), brevet
eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet
européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG,
CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Publiée

*Sans rapport de recherche internationale, sera republiée dès
réception de ce rapport.*



(54) Title: ISOLATED DNA SEQUENCE FOR USE AS A REGULATOR REGION IN A CHIMERIC GENE USEFUL FOR TRANSFORMING PLANTS

(54) Titre: SEQUENCE ADN ISOLEE POUVANT SERVIR DE ZONE DE REGULATION DANS UN GENE CHIMERE UTILISABLE POUR LA TRANSFORMATION DES PLANTES

(57) Abstract

A chimeric gene for transforming plants is disclosed. The gene includes, in the transcription direction, at least one promoter region, one transgene and one regulator region. Said regulator region consists of at least one intron 1 in the non-coding 5' region of a plant histone gene enabling expression of the protein in rapid growth regions. The production of transgenic plants is also disclosed.

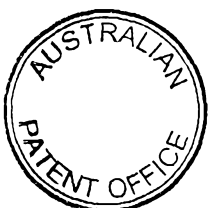
(57) Abrégé

Séquence ADN isolée pouvant servir de zone de régulation dans un gène chimère utilisable pour la transformation des plantes, 1) gène chimère utilisable pour la transformation des plantes. 2) Il comprend au moins, dans le sens de la transcription, une zone promotrice, un transgène et une zone de régulation, caractérisé en ce que la zone de régulation est constituée d'au moins un intron (1) dans la région 5' non-codant d'un gène d'histone végétale permettant l'expression de la protéine dans les zones de croissance rapide. 3) Production de plantes transgéniques.

**Isolated DNA sequence capable of serving as regulatory
zone in a chimeric gene which can be used for the
transformation of plants.**

The present invention relates to the use of a
5 regulatory zone isolated from transcribed plant genes,
of new chimeric genes containing them and to their use
for the transformation of plants.

Numerous phenotypic characters associated
with the expression of one or more gene elements can be
10 integrated into the genome of plants and thus confer on
these transgenic plants advantageous agronomic
properties. In a nonexhaustive manner, there may be
mentioned: the resistances to pathogenic agents for
crops, the resistance to phytotoxic plant-protection
15 products, the production of substances of dietary or
pharmacological interest. In addition to the isolation
and characterization of the gene elements encoding
these various characters, an appropriate expression
should be ensured. This appropriate expression may be
20 situated both at the qualitative and quantitative
levels. At the qualitative level, for example the
spatial level: preferential expression in a specific
tissue, or temporal level: inducible expression; at the
quantitative level, by the accumulated quantity of the
25 product of expression of the gene introduced. This
appropriate expression depends, for a large part, on
the presence of regulatory gene elements associated



with the transgenes, in particular as regards the quantitative and qualitative elements. Among the key elements ensuring this appropriate regulation, the use of single or combined homologous or heterologous promoter zones has been widely described in the scientific literature. The use of a regulatory zone downstream of the transgene was used for the sole purpose of putting a boundary which makes it possible to stop the process of transcription of the transgene, without presupposition as to their role as regards the quality or the quantity of the expression of the transgene.

The present invention relates to the use of an intron 1 as regulatory zone isolated from plant genes, of new chimeric genes containing them and to their use for the transformation of plants. It relates to an isolated DNA sequence capable of serving as regulatory zone in a chimeric gene which can be used for the transformation of plants and allowing the expression of the product of translation of the chimeric gene in particular in the regions of the plant undergoing rapid growth, which comprises, in the direction of transcription of the chimeric gene, at least one intron such as the first intron (intron 1) of the noncoding 5' region of a plant histone gene. It relates more particularly to the simultaneous use of intron 1 as regulatory zone and of promoters isolated from the same plant gene. It allows the appropriate



expression, both quantitative and qualitative, of the transgenes under the control of these elements for gene regulation. This appropriate expression, obtained by the use of the present invention, may relate to

5 characters such as: the resistance to pathogenic agents for crops, the resistance to phytotoxic plant-protection products, the production of substances of dietary or pharmacological interest. In particular, it makes it possible to confer on the transgenic plants an

10 enhanced tolerance to herbicides by a qualitative and quantitative preferential expression of the product of expression of the chimeric genes in the regions of the plant undergoing rapid growth. This specific appropriate expression of the gene for herbicide

15 resistance is obtained by the simultaneous use of the promoter regulatory elements and of at least one intron 1 as regulatory zone of the histone gene of the "H3.3-like" type. Such a pattern of expression can be obtained for all the characters which are of interest,

20 as described above, with the regulatory elements used to confer an enhanced herbicide tolerance. The present invention also relates to the plant cells transformed with the aid of these genes and the transformed plants regenerated from these cells as well as the plants

25 derived from crossings using these transformed plants.

Among the plant-protection products used for the protection of crops, the systemic products are characterized in that they are transported in the plant



after application and, for some of them, accumulate in the parts undergoing rapid growth, especially the caulinary and root apices, causing, in the case of herbicides, deterioration, up to the destruction, of the sensitive plants. For some of the herbicides exhibiting this type of behaviour, the primary mode of action is known and results from inactivation of characterized enzymes involved in the biosynthesis pathways of compounds required for good development of the target plants. The target enzymes of these products may be located in various subcellular compartments and observation of the mode of action of known products most often shows a location in the plastid compartment.

Tolerance of plants sensitive to a product belonging to this group of herbicides, and whose primary target is known, may be obtained by stable introduction, into their genome, of a gene encoding the target enzyme, of any phylogenetic origin, mutated or otherwise with respect to the characteristics of inhibition, by the herbicide, of the product of expression of this gene. Another approach consists in introducing, in a stable manner, into the genome of sensitive plants a gene of any phylogenetic origin encoding an enzyme capable of metabolizing the herbicide into a compound which is inactive and nontoxic for the development of the plant. In the latter case, it is not necessary to have characterized the target of the herbicide.



Given the mode of distribution and accumulation of products of this type in the treated plants, it is advantageous to be able to express the product of translation of these genes so as to allow their preferential expression and their accumulation in the regions of the plant undergoing rapid growth where these products accumulate. Furthermore, and in the case where the target of these products is located in a cellular compartment other than the cytoplasm, it is advantageous to be able to express the product of translation of these genes in the form of a precursor containing a polypeptide sequence allowing directing of the protein conferring the tolerance into the appropriate compartment, and in particular in the plastid compartment.

By way of example illustrating this approach, there may be mentioned glyphosate, sulfosate or fosametine which are broad-spectrum systemic herbicides of the phosphonomethylglycine family. They act essentially as competitive inhibitors, in relation to PEP (phosphoenolpyruvate), of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS, EC 2.5.1.19). After their application to the plant, they are transported into the plant where they accumulate in the parts undergoing rapid growth, especially the caulinary and root apices, causing the deterioration, up to the destruction, of the sensitive plants.

EPSPS, the principal target of these



products, is an enzyme of the pathway of biosynthesis of aromatic amino acids which is located in the plastid compartment. This enzyme is encoded by one or more nuclear genes and is synthesized in the form of a cytoplasmic precursor and then imported into the plastids where it accumulates in its mature form.

The tolerance of plants to glyphosate and to products of the family is obtained by the stable introduction, into their genome, of an EPSPS gene of plant or bacterial origin, mutated or otherwise with respect to the characteristics of inhibition, by glyphosate, of the product of this gene. Given the mode of action of glyphosate, it is advantageous to be able to express the product of translation of this gene so as to allow its high accumulation in the plastids and, furthermore, in the regions of the plant undergoing rapid growth where the products accumulate.

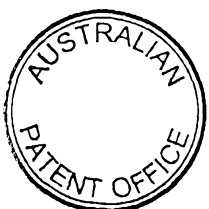
It is known, for example, from American patent 4,535,060 to confer on a plant a tolerance to a herbicide of the above type, in particular N-phosphonomethylglycine or glyphosate, by introduction, into the genome of the plants, of a gene encoding an EPSPS carrying at least one mutation making this enzyme more resistant to its competitive inhibitor (glyphosate), after location of the enzyme in the plastid compartment. These techniques require, however, to be improved for greater reliability in the use of these plants during a treatment with these products



under agronomic conditions.

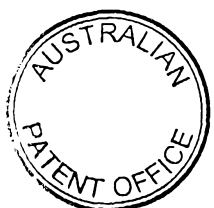
In the present description, "plant" is understood to mean any differentiated multicellular organism capable of photosynthesis and "plant cell" any cell derived from a plant and capable of constituting undifferentiated tissues such as calli, or differentiated tissues such as embryos or plant portions or plants or seeds. "Intron 1 of Arabidopsis as regulatory zone" is understood to mean an isolated DNA sequence of variable length, situated upstream of the coding part or corresponding to the structural part of a transcribed gene. Gene for tolerance to a herbicide is understood to mean any gene, of any phylogenetic origin, encoding either the target enzyme for the herbicide, having or otherwise one or more mutations with respect to the characteristics of inhibition by the herbicide, or an enzyme capable of metabolizing the herbicide into a compound which is inactive and nontoxic for the plant. Zones of the plants undergoing rapid growth are understood to mean the regions which are the seat of substantial cell multiplications, in particular the apical regions.

The present invention relates to the production of transformed plants having especially an enhanced tolerance to herbicides accumulating in the zones of the treated plants undergoing rapid growth, by regeneration of cells transformed with the aid of new chimeric genes comprising a gene for tolerance to these



products. The subject of the invention is also the production of transformed plants having an enhanced tolerance to herbicides of the phosphonomethylglycine family by regeneration of cells transformed with the aid of new chimeric genes comprising a gene for tolerance to these herbicides. The invention also relates to these new chimeric genes, as well as to transformed plants which are more tolerant because of a better tolerance in the parts of these plants undergoing rapid growth, as well as to the plants derived from crossings using these transformed plants. Its subject is also new intron 1 of plant histone and its use as regulatory zone for the construction of the above chimeric genes.

More particularly, the subject of the invention is a chimeric gene for conferring on plants especially an enhanced tolerance to a herbicide having EPSPS as target, comprising, in the direction of transcription, a promoter zone, a transit peptide zone, a sequence encoding an enzyme for tolerance to the products of the phosphonomethylglycine family and a regulatory zone, characterized in that the regulatory zone consists of a fragment of an intron 1 of a plant histone gene in any orientation relative to its initial orientation in the gene from which it is derived, allowing the preferential expression and the accumulation of the protein for tolerance to the herbicide in the zones for accumulation of the said herbicide.



The histone gene, from which intron 1 according to the invention is derived, comes from a monocotyledonous plant such as for example wheat, maize or rice, or preferably from a dicotyledonous plant such as for example lucerne, sunflower, soya bean, rapeseed or preferably *Arabidopsis thaliana*. Preferably, a histone gene of the "H3.3-like" type is used.

The transit peptide zone comprises, in the direction of transcription, at least one transit peptide of a plant gene encoding an enzyme of plastid location, a portion of the sequence of the mature N-terminal part of a plant gene encoding an enzyme of plastid location, and then a second transit peptide of a plant gene encoding an enzyme of plastid location, preferably a gene for the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) according to European patent application PCT 508 909. The role of this characteristic zone is to allow the release, into the plastid compartment, of a mature polypeptide with a maximum efficiency, preferably in a native form.

The coding sequence which can be used in the chimeric gene according to the invention comes from a herbicide tolerance gene of any phylogenetic origin. This sequence may be especially that of the mutated EPSPS having a degree of tolerance to glyphosate.

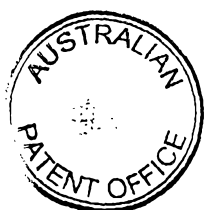
The promoter zone according to European patent application PCT 507 698 may be of any origin, in



a single or duplicated or combined form of a gene naturally expressed in plants, that is to say, for example of bacterial origin such as that of the nopaline synthase gene, or of viral origin such as that of the 35S transcript of the cauliflower mosaic virus, or preferably of plant origin such as that of the small subunit of the ribulose-1,5-bisphosphate carboxylase/oxygenase or preferably such as that of a plant histone gene and preferably from *Arabidopsis thaliana*. A histone gene of the "H4" type is preferably used.

The chimeric gene according to the invention may comprise, in addition to the above essential parts, an untranslated intermediate zone (linker) between the promoter zone and the coding zone as well as between the coding zone and intron 1 and which may be of any phylogenetic origin.

The following examples show by way of limitation, but with no limitation being implied, several aspects of the invention: isolation of the introns according to the invention and their use for the genetic transformation of plants as well as the improved qualities of expression of the heterologous genes of plants transformed with the aid of these introns.



EXAMPLE 1:

**1. Production of an EPSPS fragment from
*Arabidopsis thaliana***

a) two 20-mer oligonucleotides of respective
5 sequences:

5'-GCTCTGCTCATGTCTGCTCC-3'

5'-GCCCCGCCCTTGACAAAGAAA-3'

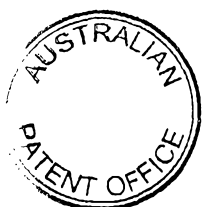
were synthesized from the sequence of an
EPSPS gene from *Arabidopsis thaliana* (Klee H.J. et al.,
10 (1987) Mol. Gen. Genet., 210, 437-442). These two
oligonucleotides are in positions 1523 to 1543 and 1737
to 1717, respectively, of the published sequence and in
convergent orientation.

b) The total DNA from *Arabidopsis thaliana*
15 (var. *columbia*) was obtained from Clontech (catalogue
reference: 6970-1)

c) 50 nanograms (ng) of DNA are mixed with
300 ng of each of the oligonucleotides and subjected to
35 amplification cycles with a Perkin-Elmer 9600
20 apparatus under the standard medium conditions for
amplification recommended by the supplier. The
resulting 204 bp fragment constitutes the EPSPS
fragment from *Arabidopsis thaliana*.

**2. Construction of a library of a cDNA from a
25 BMS maize cell line.**

a) 5 g of filtered cells are ground in liquid
nitrogen and the total nucleic acids extracted
according to the method described by Shure et al. with



the following modifications:

- the pH of the lysis buffer is adjusted to
pH = 9.0;
- after precipitation with isopropanol, the
5 pellet is taken up in water and after
dissolution, adjusted to 2.5M LiCl. After
incubation for 12 h at 0°C, the pellet from
the 15 min centrifugation at 30,000 g at 4°C
is resolubilized. The LiCl precipitation
10 stage is then repeated. The resolubilized
pellet constitutes the RNA fraction of the
total nucleic acids.

b) the RNA-poly A+ fraction of the RNA
fraction is obtained by chromatography on an oligo-dT
15 cellulose column as described in "Current Protocols in
Molecular Biology".

c) Synthesis of double-stranded cDNA with an
EcoRI synthetic end: it is carried out by following the
procedure of the supplier of the various reagents
20 necessary for this synthesis in the form of a kit: the
"copy kit" from the company Invitrogen.

Two single-stranded and partially
complementary oligonucleotides of respective sequences:

5'-AATTCCTCGGG-3'

25 5'-CCCGGG-3' (the latter being
phosphorylated)

are ligated to double-stranded cDNAs with blunt ends.

This ligation of the adaptors results in the



creation of SmaI sites attached to the double-stranded cDNAs and of EcoRI sites in cohesive form at each end of the double-stranded cDNAs.

d) Creation of the library:

5 The cDNAs having at their ends the cohesive artificial EcoRI sites are ligated to the λ gt10 bacteriophage cDNA cut with EcoRI and dephosphorylated according to the procedure of the supplier New England Biolabs.

10 An aliquot from the ligation reaction was encapsidated *in vitro* with encapsidation extracts: Gigapack Gold according to the supplier's instructions, this library was titrated using the bacterium *E.coli* C600hfl. The library thus obtained is amplified and
15 stored according to the instructions of the same supplier and constitutes the cDNA library from BMS maize cell suspension.

**3. Screening of the cDNA library from BMS
maize cell suspension with the EPSP probe from**

20 ***Arabidopsis thaliana*:**

 The procedure followed is that of "Current Protocols in Molecular Biology" Volumes 1 and 2, Ausubel F.M. et al., published by Greene Publishing Associates and Wiley-Interscience (1989) (CPMB).
25 Briefly, about 10^6 recombinant phages are plated on an LB plate at a mean density of 100 phages/cm². The lysis plaques are replicated in duplicate on a Hybond N membrane from Amersham.



The DNA was fixed onto the filters by a 1600 kJ UV treatment (Stratalinker from Stratagene). The filters were prehybridized in: 6×SSC/0.1 % SDS/0.25 [lacuna] skimmed milk for 2 h at 65°C. The EPSPS probe
5 from *Arabidopsis thaliana* was labelled with ³²P-dCTP by random priming according to the instructions of the supplier (Kit Ready to Go from Pharmacia). The specific activity obtained is of the order of 10⁸ cpm per µg of fragment. After denaturation for 5 min at 100°C, the
10 probe is added to the prehybridization medium and the hybridization is continued for 14 hours at 55°C. The filters are fluorographed for 48 h at -80°C with a Kodak XAR5 film and intensifying screens Hyperscreen RPN from Amersham. The alignment of the positive spots
15 on the filter with the plates from which they are derived make it possible to collect, from the plate, the zones corresponding to the phages exhibiting a positive hybridization response with the EPSPS probe from *Arabidopsis thaliana*. This step of plating,
20 transfer, hybridization and recovery is repeated until all the spots of the plate of phages successively purified prove 100 % positive in hybridization. A lysis plaque per independent phage is then collected in the diluent λ medium (Tris-Cl pH=7.5; 10 mM MgSO₄; 0.1M
25 NaCl; 0.1 % gelatine), these phages in solution constituting the positive EPSPS clones from the maize cell suspension.



4. Preparation and analysis of the DNA of the EPSP clones from the BMS maize cell suspension.

About 5×10^8 phages are added to 20 ml of C600hfl bacteria at OD 2 (600 nm/ml) and incubated for 15 minutes at 37°C. This suspension is then diluted in 200 ml of growth medium for the bacteria in a 1 l Erlenmeyer flask and shaken in a rotary shaker at 250 rpm. Lysis is observed by clarification of the medium, corresponding to lysis of the turbid bacteria and occurs after about 4 h of shaking. This supernatant is then treated as described in "Current Protocols in Molecular Biology". The DNA obtained corresponds to the EPSPS clones from the BMS maize cell suspension.

One to two μg of this DNA are cut with EcoRI and separated on a 0.8 % LGTA/TBE agarose gel (ref. CPMB). A final verification consists in ensuring that the purified DNA indeed exhibits a hybridization signal with the EPSPS probe from *Arabidopsis thaliana*. After electrophoresis, the DNA fragments are transferred onto Hybond N membrane from Amersham according to the Southern procedure described in "Current Protocols in Molecular Biology". The filter is hybridized with the EPSPS probe from *Arabidopsis thaliana* according to the conditions described in paragraph 3 above. The clone exhibiting a hybridization signal with the EPSPS probe from *Arabidopsis thaliana* and containing the longest EcoRI fragment has a gel-estimated size of about 1.7 kbp.



5. Production of the pRPA-ML-711 clone:

Ten μg of DNA from the phage clone containing the 1.7 kbp insert are digested with EcoRI and separated on a 0.8 % LGTA/TBE agarose gel (ref. CPMB).

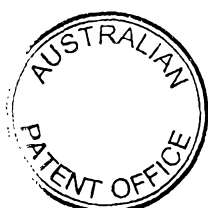
5 The gel fragment containing the 1.7 kbp insert is excised from the gel by BET staining and the fragment is treated with β -agarase according to the procedure of the supplier New England Biolabs. The DNA purified from the 1.7 kbp fragment is ligated at 12°C for 14 h with

10 DNA from the plasmid pUC 19 (New England Biolabs) cut with EcoRI according to the ligation procedure described in "Current Protocols in Molecular Biology". Two μl of the above ligation mixture are used for the transformation of one aliquot of electrocompetent

15 *E.coli* DH10B; the transformation occurs by electroporation using the following conditions: the mixture of competent bacteria and ligation medium is introduced into an electroporation cuvette 0.2 cm thick (Biorad) previously cooled to 0°C. The physical

20 electroporation conditions using an electroporator of Biorad trade mark are 2500 volts, 25 μFarad and 200 Ω . Under these conditions, the mean condenser discharge time is of the order of 4.2 milliseconds. The bacteria are then taken up in 1 ml of SOC medium (ref. CPMB) and

25 shaken for 1 hour at 200 rpm on a rotary shaker in 15 ml Corning tubes. After plating on LB/agar medium supplemented with 100 $\mu\text{g/ml}$ of carbenicillin, the mini-preparations of the bacteria clones having grown



overnight at 37°C is carried out according to the procedure described in "Current Protocols in Molecular Biology". After digestion of the DNA with EcoRI and separation by electrophoresis on a 0.8 % LGTA/TBE agarose gel (ref. CPMB), the clones having a 1.7 kbp insert are conserved. A final verification consists in ensuring that the purified DNA indeed exhibits a hybridization signal with the EPSPS probe from *Arabidopsis thaliana*. After electrophoresis, the DNA fragments are transferred onto a Hybond N membrane from Amersham according to the Southern procedure described in "Current Protocols in Molecular Biology". The filter is hybridized with the EPSPS probe from *Arabidopsis thaliana* according to the conditions described in paragraph 3 above. The plasmid clone having a 1.7 kbp insert and hybridizing with the EPSPS probe from *Arabidopsis thaliana* was prepared on a larger scale and the DNA resulting from the lysis of the bacteria purified on a CsCl gradient as described in "Current Protocols in Molecular Biology". The purified DNA was partially sequenced with a Pharmacia kit, following the supplier's instructions and using, as primers, the direct and reverse M13 universal primers ordered from the same supplier. The partial sequence produced covers about 0.5 kbp. The derived amino acid sequence in the region of the mature protein (about 50 amino acid residues) exhibits 100 % identity with the corresponding amino sequence of the mature maize EPSPS



described in American patent USP 4,971,908. This clone, corresponding to a 1.7 kbp EcoRI fragment of the DNA for the EPSP from the BMS maize cell suspension, was called **pRPA-ML-711**. The complete sequence of this clone
5 was obtained on both strands by using the Pharmacia kit procedure and by synthesizing oligonucleotides which are complementary and of opposite direction every 250 bp approximately. The complete sequence of this 1713 bp clone obtained is presented by SEQ ID No. 1.

10 **6. Production of the clone pRPA-ML-715:**

Analysis of the sequence of the clone pRPA-ML-711 and in particular comparison of the derived amino acid sequence with that from maize shows a sequence extension of 92 bp upstream of the GCG codon
15 encoding the NH₂-terminal alanine of the mature part of the maize EPSPS (American patent UPS 4,971,908). Likewise, a 288 bp extension downstream of the AAT codon encoding the COOH-terminal asparagine of the mature part of the maize EPSPS (American patent USP
20 4,971,908) is observed. These two parts might correspond, for the NH₂-terminal extension, to a portion of the sequence of a transit peptide before plastid location and, for the COOH-terminal extension, to the untranslated 3' region of the cDNA.

25 In order to obtain a cDNA encoding the mature part of the cDNA for the maize EPSPS, as described in USP 4,971,908, the following operations were carried out:



**a) Elimination of the untranslated 3' region:
construction of pRPA-ML-712:**

The clone pRPA-ML-711 was cut with the restriction enzyme AseI and the resulting ends of this cut made blunt by treating with the Klenow fragment of DNA polymerase I according to the procedure described in CPMB. A cut with the restriction enzyme SacII was then performed. The DNA resulting from these operations was separated by electrophoresis on a 1 % LGTA/TBE agarose gel (ref. CPMB).

The gel fragment containing the insert "AseI-blunt ends/SacII" of 0.4 kbp was excised from the gel and purified according to the procedure described in paragraph 5 above. The DNA of the clone pRPA-ML-711 was cut with the restriction enzyme HindIII situated in the polylinker of the cloning vector pUC19 and the ends resulting from this cut were made blunt by treating with the Klenow fragment of DNA polymerase I. A cut with the restriction enzyme SacII was then performed. The DNA resulting from these manipulations was separated by electrophoresis on a 0.7 % LGTA/TBE agarose gel (ref. CPMB).

The gel fragment containing the insert HindIII-blunt ends/SacII of about 3.7 kbp was excised from the gel and purified according to the procedure described in paragraph 5 above.

The two inserts were ligated, and 2 μ l of the ligation mixture served to transform *E.coli* DH10B as



described above in paragraph 5.

The plasmid DNA content of the various clones was analysed according to the procedure described for pRPA-ML-711. One of the plasmid clones retained
5 contains an EcoRI-HindIII insert of about 1.45 kbp. The sequence of the terminal ends of this clone shows that the 5' end of the insert corresponds exactly to the corresponding end of pRPA-ML-711 and that the 3' terminal end has the following sequence:

10 "5'...AATTAAGCTCTAGAGTCGACCTGCAGGCATGCAAGCTT-3'".

The sequence underlined corresponds to the codon for the COOH-terminal amino acid asparagine, the next codon corresponding to the stop codon for translation. The nucleotides downstream correspond to
15 sequence components of the polylinker of pUC19. This clone, comprising the sequence of pRPAML-711 up to the site for termination of translation of the mature maize EPSPS and followed by sequences of the polylinker of pUC19 up to the HindIII site, was called **pRPA-ML-712**.

20 **b) Modification of the 5' end of pRPA-ML-712:
construction of pRPA-ML-715**

The clone pRPA-ML-712 was cut with the restriction enzymes PstI and HindIII. The DNA resulting from these manipulations was separated by
25 electrophoresis on a 0.8 % LGTA/TBE agarose gel (ref. CPMB). The gel fragment containing the PstI/EcoRI insert of 1.3 kbp was excised from the gel and purified according to the procedure described in paragraph 5



above. This insert was ligated in the presence of an equimolar quantity of each of the two partially complementary oligonucleotides of sequence:

Oligo 1: 5'-GAGCCGAGCTCCATGGCCGGCGCCGAGGAGATCGTGCTGCA-3'

5 Oligo 2: 5'-GCACGATCTCCTCGGCGCCGGCCATGGAGCTCGGCTC-3'

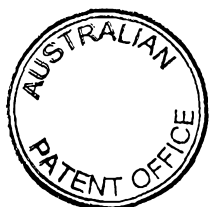
as well as in the presence of DNA from the plasmid pUC19 digested with the restriction enzymes BamHI and HindIII.

Two μ l of the ligation mixture served to
10 transform *E.coli* DH10B as described above in paragraph 5. After analysis of the plasmid DNA content of various clones according to the procedure described above in paragraph 5, one of the clones having an insert of about 1.3 kbp was conserved for subsequent analyses.
15 The sequence of the terminal 5' end of the clone retained shows that the DNA sequence in this region is the following: sequence of the polylinker of pUC19 of the EcoRI to BamHI sites, followed by the sequence of the oligonucleotides used during the cloning, followed
20 by the rest of the sequence present in pRPAML-712. **This clone was called pRPA-ML-713.** This clone has a methionine codon ATG included in an NcoI site upstream of the N-terminal alanine codon of the mature EPSPSynthase. Furthermore, the alanine and glycine
25 codons of the N-terminal end were conserved, but modified on the third variable base: initial GCGGGT gives modified GCGGGC.

The clone pRPA-ML-713 was cut with the



restriction enzyme HindIII and the ends of this cut made blunt by treating with the Klenow fragment of DNA polymerase I. A cut with the restriction enzyme SacI was then performed. The DNA resulting from these manipulations was separated by electrophoresis on a 0.8 % LGTA/TBE agarose gel (ref. CPMB). The gel fragment containing the insert "HindIII-blunt ends/SacI" of 1.3 kbp was excised from the gel and purified according to the procedure described in paragraph 5 above. This insert was ligated in the presence of DNA from the plasmid pUC19 digested with the restriction enzyme XbaI and the ends of this cut made blunt by treating with the Klenow fragment of DNA polymerase I. A cut with the restriction enzyme SacI was then performed. Two μ l of the ligation mixture served to transform *E.coli* DH10B as described above in paragraph 5. After analysis of the plasmid DNA content of various clones according to the procedure described above in paragraph 5, one of the clones having an insert of about 1.3 kbp was conserved for subsequent analyses. The sequence of the terminal ends of the clone retained shows that the DNA sequence is the following: sequence of the polylinker of pUC19 of the EcoRI to SacI sites, followed by the sequence of the oligonucleotides used during the cloning, from which the 4 bp GATCC of oligonucleotide 1 described above have been deleted, followed by the rest of the sequence present in pRPA-ML-712 up to the HindIII site and



sequence of the polylinker of pUC19 from XbaI to HindIII. This clone was called pRPA-ML-715.

7) Production of a cDNA encoding a mature maize EPSPS

5 All the mutagenesis steps were carried out with the U.S.E. mutagenesis kit from Pharmacia, following the instructions of the supplier. The principle of this mutagenesis system is as follows: the plasmid DNA is heat-denatured and recombined in the
10 presence of a molar excess, on the one hand, of the mutagenesis oligonucleotide and, on the other hand, of an oligonucleotide which makes it possible to eliminate a unique restriction enzyme site present in the polylinker. After the reassociation step, the synthesis
15 of the complementary strand is performed by the action of T4 DNA polymerase in the presence of T4 DNA ligase and protein of gene 32 in an appropriate buffer provided. The synthesis product is incubated in the presence of the restriction enzyme, whose site is
20 supposed to have disappeared by mutagenesis. The *E.coli* strain exhibiting, in particular, the mutS mutation is used as host for the transformation of this DNA. After growth in liquid medium, the total plasmid DNA is prepared and incubated in the presence of the
25 restriction enzyme used above. After these treatments, the *E.coli* DH10B strain is used as host for the transformation. The plasmid DNA of the isolated clones is prepared and the presence of the mutation introduced



is checked by sequencing.

A) - Site or sequence modifications with no effect a priori on the resistance character of maize EPSPS to the products which are competitive inhibitors of the activity of EPSP synthase: elimination of an internal NcoI site from pRPA-ML-715.

The sequence of pRPA-ML-715 is arbitrarily numbered by placing the first base of the N-terminal alanine codon GCC in position 1. This sequence has an NcoI site in position 1217. The site-modifying oligonucleotide has the sequence:

5'-CCACAGGATGGCGATGGCCTTCTCC-3'.

After sequencing according to the references given above, the sequence read after mutagenesis corresponds to that of the oligonucleotide used. The NcoI site was indeed eliminated and translation into amino acids in this region conserves the initial sequence present in pRPA-ML-715.

This clone was called pRPA-ML-716.

The 1340 bp sequence of this clone is represented SEQ ID No. 2 and SEQ ID No. 3.

B) Sequence modifications allowing an increase in the resistance character of maize EPSPS to products which are competitive inhibitors of the activity of EPSP synthase.

The following oligonucleotides were used:



a) Thr 102 $\Rightarrow$ Ile mutation.

5'-GAATGCTGGAATCGCAATGCGGCCATTGACAGC-3'

b) Pro 106 $\Rightarrow$ Ser mutation.

5'-GAATGCTGGAAGTCAATGCGGTCCTTGACAGC-3'

5 c) Gly 101 $\Rightarrow$ Ala and Thr 102 $\Rightarrow$ Ile mutations.

5'-CTTGGGGAATGCTGCCATCGCAATGCGGCCATTG-3'

d) Thr 102 $\Rightarrow$ Ile and Pro 106 $\Rightarrow$ Ser mutations.

5'-GGGGAATGCTGGAATCGCAATGCGGTCCTTGACAGC-3'

After sequencing, the sequence read after
10 mutagenesis on the three mutated fragments is identical
to the sequence of the parental DNA pRPA-ML-716 with
the exception of the mutagenesis region which
corresponds to that of the mutagenesis oligonucleotides
used. These clones were called: pRPA-ML-717 for the Thr
15 102 $\Rightarrow$ Ile mutation, pRPA-ML-718 for the Pro 106 $\Rightarrow$ Ser
mutation, pRPA-ML-719 for the Gly 101 $\Rightarrow$ Ala and Thr 102
 $\Rightarrow$ Ile mutations and pRPA-ML-720 for the Thr 102 $\Rightarrow$ Ile
and Pro 106 $\Rightarrow$ Ser mutations.

The 1340 bp sequence of pRPA-ML-720 is
20 represented SEQ ID No. 4 and SEQ ID No. 5.

The NcoI-HindIII insert of 1395 bp will be
called in the rest of the descriptions "the double
mutant of maize EPSPS".



EXAMPLE 2: Construction of chimeric genes

The construction of chimeric genes according to the invention is carried out using the following elements:

5 1). The genomic clone (cosmid clone c22) from *Arabidopsis thaliana*, containing two genes of the "H3.3-like" type was isolated as described in Chaubet et al. (J. Mol. Biol. 1992. 225 569-574).

2). Intron No. 1:

10 A DNA fragment of 418 base pairs is purified from digestion of the cosmid clone c22 with the restriction enzyme DdeI followed by treatment with a Klenow fragment of DNA polymerase from *E.coli*, according to the manufacturer's instructions for
15 creating a blunt-ended DNA fragment and then cut with MseI. The purified DNA fragment is ligated to a synthetic oligonucleotide adaptor having the following sequence:

Adaptor 1: 5' TAATTTGTTGAACAGATCCC 3'

20 TAAACAACTTGTCTAGGG

The ligation product is cloned into pGEM7Zf(+) (Stratagene catalogue No. P2251) which was digested with SmaI. This clone, called "intron No. 1", is checked by sequencing (SEQ ID 6).

25 **3). Intron No. 2:**

A DNA fragment of 494 base pairs is purified from the digestion of the cosmid clone c22 with the restriction enzymes AluI and CfoI. The purified DNA



fragment is ligated to a synthetic oligonucleotide adaptor having the following sequence:

Adaptor 2: 5' CAGATCCCGGGATCTGCG 3'
GCGTCTAGGGCCCTAGACGC

5 The ligation product is cloned into pGEM7Zf(+) (Stratagene catalogue No. P2251) which was digested with SmaI. This clone, called "intron No. 2", is checked by sequencing (SEQ ID 7).

4). pRA-1

10 The construction of this plasmid is described in French patent 9,308,029. This plasmid is a derivative of pBI 101.1 (Clontech catalogue No. 6017-1) which contains the histone promoter from *Arabidopsis* H4A748 regulating the synthesis of the *E.coli* β -
15 glucoronidase gene and of the nopaline synthase ("NOS") polyadenylation site. Thus, a chimeric gene is obtained having the structure:

"H4A748 promoter-GUS gene-NOS"

5). pCG-1

20 This plasmid contains the above intron No. 1 placed between the H4A748 promoter and the GUS coding region of pRA-1. This plasmid is obtained by digestion of cosmid clone c22 with BamHI and SmaI. The intron
No. 1 of 418 base pairs is directly ligated into pRA-1
25 which was digested with BamHI and SmaI.

Thus, a chimeric gene is obtained having the structure:

"H4A748 promoter-intron No. 1-GUS gene-NOS"



6) . pCG-13

This plasmid contains the above intron No. 2 placed between the H4A748 promoter and the GUS coding region of pRA-1. This plasmid is obtained by digestion
5 of cosmid clone c22 with BamHI and SmaI. The intron No. 2 of 494 base pairs is directly ligated into pRA-1 which was digested with BamHI and SmaI.

Thus, a chimeric gene is obtained having the structure:

10 "H4A748 promoter-intron No. 2-GUS gene-NOS"

7) . pCG-15

This plasmid contains only intron No. 1 before the above GUS coding sequence placed between the H4A748 promoter and the GUS coding region of pCG-1.
15 This plasmid is obtained by digestion of pCG-1 with BamHI and HindIII followed by treatment with a Klenow fragment of DNA polymerase from *E.coli*, according to the manufacturer's instructions for creating a blunt-ended DNA fragment.

20 This vector is then religated to give a chimeric gene having the structure:

"intron No. 1-GUS-NOS"

8) . pCG-18

This plasmid contains only the above intron
25 No. 2 in front of the GUS coding sequence of pCG-13. This plasmid is obtained by partial digestion of pCG-13 with BamHI and SphI, followed by treatment with a fragment of T4 phage DNA polymerase, according to the



manufacturer's instructions in order to create a blunt-ended DNA fragment.

This vector is then religated and checked by enzymatic digestion in order to give a chimeric gene
5 having the structure:

"intron No. 2-GUS-NOS"

9) . pRPA-RD-124

Addition of a "nos" polyadenylation signal to pRPA-ML-720 with creation of a cloning cassette
10 containing the maize double mutant EPSPS gene (Thr 102 → Ile and Pro 106 → Ser). pRPA-ML-720 is digested with HindIII and treated with the Klenow fragment of DNA polymerase from *E.coli* in order to produce a blunt end. A second digestion is carried out with NcoI and the
15 EPSPS fragment is purified. The EPSPS gene is then ligated with purified pRPA-RD-12 (a cloning cassette containing the nopaline synthase polyadenylation signal) to give pRPA-RD-124. To obtain the purified useful vector pRPA-RD-12, it was necessary for the
20 latter to be previously digested with SalI, treated with Klenow DNA polymerase, and then digested a second time with NcoI.

10) . pRPA-RD-125

Addition of an optimized transit peptide
25 (OTP) from pRPA-RD-124 with creation of a cloning cassette containing the EPSPS gene targeted on the plasmids. pRPA-RD-7 (European Patent Application EP 652 286) is digested with SphI, treated with T4 DNA



polymerase and then digested with SpeI and the OTP fragment is purified. This OTP fragment is cloned into pRPA-RD-124 which was previously digested with NcoI, treated with Klenow DNA polymerase in order to remove
5 the 3' protruding part, and then digested with SpeI. This clone is then sequenced in order to ensure the correct translational fusion between the OTP and the EPSPS gene. pRPA-RD-125 is then obtained.

11). pRPA-RD-196

10 In this plasmid, the "intron No. 1 + β -glucoronidase gene from *E.coli*" portion of pCG-1 is replaced by a chimeric gene of 2 kilobases containing an optimized transit peptide, a double mutant EPSPS gene (Ile₁₀₂+Ser₁₀₆) and a nopaline synthase
15 polyadenylation site ("NOS") isolated from pRPA-RD-125. To obtain pRPA-RD-196, the digestion of pCG-1 is performed with EcoRI and BamHI, followed by treatment with a Klenow fragment of DNA polymerase from *E.coli*, according to the manufacturer's instructions in order
20 to create a blunt-ended DNA fragment. The 2-kilobase DNA fragment containing an optimized transit peptide of a double mutant EPSPS gene (Ile₁₀₂+Ser₁₀₆) and a nopaline synthase polyadenylation site ("NOS") is obtained from pRPA-RD-125 by digestion with NcoI and
25 NotI, followed by treatment with DNA polymerase from *E.coli*, according to the manufacturer's instructions in order to create a blunt-ended DNA fragment. This blunt-ended fragment is then ligated into pCG-1 prepared



above.

A chimeric gene is thus obtained having the structure:

"H4A748 promoter-OTP-maize EPSPS gene-NOS"

5 12). pRPA-RD-197

In this plasmid, the " β -glucoronidase gene from *E.coli*" portion of pCG-1 is replaced by a chimeric gene of 2 kilobases containing an optimized transit peptide, a double mutant EPSPS gene (Ile₁₀₂+Ser₁₀₆) and
10 a nopaline synthase polyadenylation site ("NOS") isolated from pRPA-RD-125. To obtain pRPA-RD-197, the digestion of pCG-1 is performed with EcoRI, followed by treatment with a Klenow fragment of DNA polymerase from *E.coli*, according to the manufacturer's instructions in
15 order to create a blunt-ended DNA fragment, then cut with SmaI. The 2-kilobase DNA fragment containing an optimized transit peptide, a double mutant EPSPS gene (Ile₁₀₂+Ser₁₀₆) and a nopaline synthase polyadenylation site ("NOS") is obtained from pRPA-RD-125 by digestion
20 with NcoI and NotI, followed by a treatment with DNA polymerase from *E.coli*, according to the manufacturer's instructions in order to create a blunt-ended DNA fragment. This blunt-ended fragment is then ligated into pCG-1 prepared above.

25 A chimeric gene is thus obtained having the structure:

"H4A748 promoter-intron No. 1-maize EPSPS gene-NOS"



13). pRPA-RD-198

In this plasmid, the " β -glucoronidase gene from *E.coli*" portion of pCG-13 is replaced by a chimeric gene of 2 kilobases containing an optimized transit peptide, a double mutant EPSPS gene (Ile₁₀₂+Ser₁₀₆) and a nopaline synthase polyadenylation site ("NOS") isolated from pRPA-RD-125. To obtain pRPA-RD-198, the digestion of pCG-13 is performed with EcoRI, followed by treatment with a Klenow fragment of DNA polymerase from *E.coli*, according to the manufacturer's instructions in order to create a blunt-ended DNA fragment, then cut with SmaI. The 2-kilobase DNA fragment containing an optimized transit peptide, a double mutant EPSPS gene (Ile₁₀₂+Ser₁₀₆) and a nopaline synthase polyadenylation site ("NOS") is obtained from pRPA-RD-125 by digestion with NcoI and NotI, followed by a treatment with DNA polymerase from *E.coli*, according to the manufacturer's instructions in order to create a blunt-ended DNA fragment. This blunt-ended fragment is then ligated into pCG-13 prepared above.

A chimeric gene is thus obtained having the structure:

"H4A748 promoter-intron No. 2-OTP-maize EPSPS gene-NOS"

EXAMPLE 3: Expression of the activity of a reporter gene

1) Transformation and regeneration

The vector is introduced into the nononcogenic strain of *Agrobacterium tumefaciens* LBA

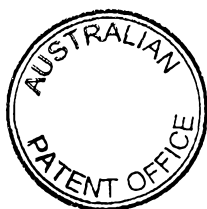


4404 available from a catalogue (Clontech #6027-1) by triparental crossing using the "helper" plasmid pRK 2013 in *Escherichia coli* HB101 according to the procedure described by Bevan M. (1984) Nucl. Acids Res., 12, 8711-8721.

The transformation technique using root explants of *Arabidopsis thaliana* L.-ecotype C24 was carried out according to the procedure described by Valvekens D. et al. (1988) Proc. Natl. Acad. Sci USA, 85, 5536-5540. Briefly, 3 steps are necessary: induction of the formation of calli on Gamborg B5 medium supplemented with 2,4-D and kinetin; formation of buds on Gamborg B5 medium supplemented with 2iP and IAA; rooting and formation of seeds on hormone-free MS.

2) Measurement of the GUS activity in plants
a - histochemical observations

Visualization of the GUS activity by histochemical spots (Jefferson R.A. et al. (1987) EMBO J., 6, 3901-3907) on 10-day transgenic plants shows an increase in the intensity of the histochemical pattern which is tissue-specific for the plasmids containing the intron sequences (pCG-1 and pCG-13) compared with those without these introns (pRA-1). In particular, the pattern of spots for pCG-1 and pCG-13 is identical, showing an increase in intensity of the spots for the vascular and meristematic tissues, leaves and roots compared with that of the construct pRA-1. The constructs containing only the sequences of intron



No. 1 (pCG-15 and pCG-18) show an extremely clear histochemical spot only in the apical meristem region.

b - fluorometric measurements

The GUS activity measured by fluorometry on
5 extracts of floral and leaf buds of the rosette
(Jefferson R.A. et al. (1987) EMBO J., 6, 3901-3907)
from 12 plants, shows that the activity of the H4A748
promoter is increased under the influence of intron
Nos. 1 and 2. Compared with the construct pRA-1, the
10 GUS activity of pCG-1 and pCG-13 are at least six times
greater in the floral buds, twenty times greater in the
leaves of the rosette and twenty-six times greater in
the roots.

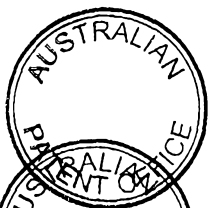
These measurements clearly show that introns
15 Nos. 1 and 2 of Arabidopsis histone genes of the "H3.3-
like" type used as regulatory zone induces an increase
in the activity of expression of the chimeric gene.

EXAMPLE 4: Tolerance of transgenic plants to
a herbicide

20 1) Transformation and regeneration

The vector is introduced into the
nononcogenic strain of *Agrobacterium tumefaciens* LBA
4404 available from a catalogue (Clontech #6027-1) by
triparental crossing using the "helper" plasmid pRK
25 2013 in *Escherichia coli* HB101 according to the
procedure described by Bevan M. (1984) Nucl. Acids
Res., 12, 8711-8721.

The transformation technique using foliar

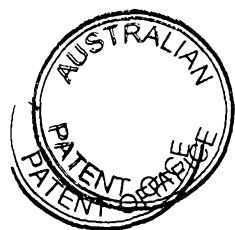


explants of tobacco is based on the procedure described by Horsh R. et al. (1985) Science, **227**, 1229-1231. The regeneration of the PBD6 tobacco (origin SEITA-France) from foliar explants is carried out on a Murashige and Skoog (MS) basal medium comprising 30 g/l of sucrose as well as 200 µg/ml of kanamycin in three successive steps: the first comprises the induction of shoots on an MS medium supplemented with 30 g of sucrose containing 0.05 mg of naphthylacetic acid (NAA) and 2 mg/l of benzylaminopurine (BAP) for 15 days. The shoots formed during this step are then developed by culturing on an MS medium supplemented with 30 g/l of sucrose but not containing any hormone, for 10 days. The developed shoots are then removed and they are cultured on an MS rooting medium diluted one half, with half the content of salts, vitamins and sugars and not containing any hormone. After about 15 days, the rooted shoots are planted in the soil.

2) Measurement of the tolerance to glyphosate:

Twenty transformed plants were regenerated and transferred to a greenhouse for each of the constructs pRPA-RD-196, pRPA-RD-197 and pRPA-RD-198. These plants were treated in a greenhouse at the 5-leaf stage with an aqueous suspension of herbicide, sold under the trademark RoundUp, corresponding to 0.8 kg of active substance glyphosate per hectare.

The results correspond to the observation of



- 36 -

phytotoxicity values noted 3 weeks after treatment. Under these conditions, it is observed that the plants transformed with the constructs have on average an acceptable tolerance (pRPA-RD-196) or even a good tolerance (pRPA-RD-197 and pRPA-RD-198) whereas the untransformed control plants are completely destroyed.

These results show clearly the improvement offered by the use of a chimeric gene according to the invention for the same gene encoding tolerance to glyphosate.

The transformed plants according to the invention may be used as parents for producing lines and hybrids having the phenotypic character corresponding to the expression of the chimeric gene introduced.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.



Sequence listing:

SEQ ID No. 1.

AATCAATTTT ACACAGGAAA CAGCTATGAC CATGATTACG AATTCGGGGC CGGGCGCGTG	60
ATCCGGCGGC GGCAGCGGCG GCGGCGGTGC AGGCGGGTGC CGAGGAGATC GTGCTGCAGC	120
CCATCAAGGA GATCTCCGGC AECGTCAAGC TGCCGGGGTC CAAGTCGCTT TCCAACCGGA	180
TCCTCCTACT CGCCGCCCTG TCCGAGGGGA CAACAGTGGT TGATAACCTG CTGAACAGTG	240
AGGATGTCCA CTACATGCTC GGGGCCCTGA GGACTCTTGG TCTCTCTGTC GAAGCGGACA	300
AAGCTGCCAA AAGAGCTGTA GTTGTGGCT GTGGTGGAAA GTTCCAGTT GAGGATGCTA	360
AAGAGGAAGT GCAGCTCTTC TTGGGGAATG CTGGAAGTGC AATGCGGCCA TTGACAGCAG	420
CTGTACTGTC TGCTGGTGGG AATGCAACTT ACGTGCTTGA TGGAGTACCA AGAATGAGGG	480
AGAGACCCAT TGGCGACTTG GTTGTGGAT TGAAGCAGCT TGCTGCAGAT GTTGATTGTT	540
TCCTTGGCAC TGAAGCCCA CCTGTCTGTG TCAATGGAAT CGGAGGGCTA CCTGGTGGCA	600
AGTCTAGCT GTCTGGCTCC ATCAGCAGTC AGTACTTGAG TGCCTTGCTG ATGGCTGCTC	660
CTTGGCTCTT TGGGATGTG GAGATTGAAA TCATTGATAA ATTAATCTCC ATTCCGTACG	720
TCGAAATGAC ATTGAGATTG ATGGAGCGTT TTGGTGTGAA AGCAGAGCAT TCTGATAGCT	780
GGGACAGATT CTACATTAAG GGAGTCAAA AATACAAGTC CCCTAAAAAT GCCTATGTTG	840
AAGGTGATGC CTCAAGCGCA AGCTATTTCT TGGCTGGTGC TGCAATTACT GGAGGGACTG	900
TCACTGTGGA AGGTGTGGC ACCACCGTT TGCAGGGTGA TGTGAAGTTT GCTGAGGTAC	960
TGGAGATGAT GGGAGCGAAG GTTACATGGA CCGAGACTAG CGTAACTGTT ACTGGCCAC	1020
CGCGGGAGCC ATTTGGGAGG AAACACCTCA AGGCGATTGA TGTCAACATG AACAAGATGC	1080
CTGATGTGCG CATGACTCTT GCTGTGGTTG CCCTCTTTGC CGATGGCCCG ACAGCCATCA	1140
GAGACGTGGC TTCCTGGAGA GTAAAGGAGA CCGAGAGGAT GGTTCGGATC CGGACGGAGC	1200
TAACCAAGCT GGGAGCATCT GTTGAGGAAG GCGCGGACTA CTGCATCATC ACGCCGCCGG	1260
AGAAGCTGAA CGTGACGGCG ATCGACAGT ACGACGACCA CAGGATGGCC ATGGCCTTCT	1320
CCCTTGCCGC CTGTCCGAG GTCCTCGTCA CCACTCGGGA CCTGCGTGC ACCCGGAAGA	1380
CCTTCCCCGA CTAATTCGAT GTGCTGAGCA CTTCGTCAA GAATTAATAA AGCGTCGGAT	1440
ACTACCACGC AGCTTGATTG AAGTGATAGG CTGTGCTGA GGAAATACAT TTCTTTTGT	1500
CTGTTTTTCT CTTCACGGG ATTAAGTTTT GAGTCTGTAA CGTTAATTGT TTGTAGCAAG	1560
TTTCTATTTT GGATCTTAAG TTTGTGCACT GTAAGCCAAA TTTCATTTCA AGAGTGGTTC	1620
GTGGAATAA TAAGATAAT AAATTACGTT TCAAGTAAAA AAAAAAAAAA AAAAAAAAAA	1680
AAAAAAAAA AAAAAAAAAA AACCCGGGAA TTC	1713



SEQ ID No. 2.

CCATG GCC GGC GCC GAG GAG ATC GTG CTG CAG CCC ATC AAG GAG ATC Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile 1 5 10	47
TCC GGC ACC GTC AAG CTG CCG GGG TCC AAG TCG CTT TCC AAC CGG ATC Ser Gly Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile 15 20 25 30	95
CTC CTA CTC GCC GCC CTG TCC GAG GGG ACA ACA GTG GTT GAT AAC CTG Leu Leu Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu 35 40 45	143
CTG AAC AGT GAG GAT GTC CAC TAC ATG CTC GGG GCC TTG AGG ACT CTT Leu Asn Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu 50 55 60	191
GGT CTC TCT GTC GAA GCG GAC AAA GCT GCC AAA AGA GCT GTA GTT GTT Gly Leu Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val 65 70 75	239
GGC TGT GGT GGA AAG TTC CCA GTT GAG GAT GCT AAA GAG GAA GTG CAG Gly Cys Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln 80 85 90	287
CTC TTC TTG GGG AAT GCT GGA ACT GCA ATG CGG CCA TTG ACA GCA GCT Leu Phe Leu Gly Asn Ala Gly Thr Ala Met Arg Pro Leu Thr Ala Ala 95 100 105 110	335
GTT ACT GCT GCT GGT GGA AAT GCA ACT TAC GTG CTT GAT GGA GTA CCA Val Thr Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro 115 120 125	383
AGA ATG AGG GAG AGA CCC ATT GGC GAC TTG GTT GTC GGA TTG ARG CAG Arg Met Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln 130 135 140	431
CTT GGT GCA GAT GTT GAT TGT TTC CTT GGC ACT GAC TGC CCA CCT GTT Leu Gly Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val 145 150 155	479
CGT GTC AAT GGA ATC GGA GGG CTA CCT GGT GGC AAG GTC AAG CTG TCT Arg Val Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser 160 165 170	527
GGC TCC ATC AGC AGT CAG TAC TTG AGT GCC TTG CTG ATG GCT GCT CCT Gly Ser Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro 175 180 185 190	575
TTG GCT CTT GGG GAT GTG GAG ATT GAA ATC ATT GAT AAA TTA ATC TCC Leu Ala Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser 195 200 205	623
ATT CCG TAC GTC GAA ATG ACA TTG AGA TTG ATG GAG CGT TTT CGT GTG Ile Pro Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val 210 215 220	671
AAA GCA GAG CAT TCT GAT AGC TGG GAC AGA TTC TAC ATT AAG GGA GGT Lys Ala Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly 225 230 235	719



SEQ ID No. 2 (continuation).

CAA AAA TAC AAG TCC CCT AAA AAT GCC TAT GTT GAA GGT GAT GCC TCA Gln Lys Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser 240 245 250	767
AGC GCA AGC TAT TTC TTG GCT GGT GCT GCA ATT ACT GGA GGG ACT GTG Ser Ala Ser Tyr Phe Leu Ala Gly Ala Ala Ile Thr Gly Gly Thr Val 255 260 265 270	815
ACT GTG GAA GGT TGT GGC ACC ACC AGT TTG CAG GGT GAT GTG AAG TTT Thr Val Glu Gly Cys Gly Thr Thr Ser Leu Gln Gly Asp Val Lys Phe 275 280 285	863
GCT GAG GTA CTG GAG ATG ATG GGA GCG AAG GTT ACA TGG ACC GAG ACT Ala Glu Val Leu Glu Met Met Gly Ala Lys Val Thr Trp Thr Glu Thr 290 295 300	911
AGC GTA ACT GTT ACT GGC CCA CCG CCG GAG CCA TTT GGG AGG AAA CAC Ser Val Thr Val Thr Gly Pro Arg Glu Pro Phe Gly Arg Lys His 305 310 315	959
CTC AAG GCG ATT GAT GTC AAC ATG AAC AAG ATG CCT GAT GTC GCC ATG Leu Lys Ala Ile Asp Val Asn Met Asn Lys Met Pro Asp Val Ala Met 320 325 330	1007
ACT CTT GCT GTG GTT GCC CTC TTT GCC GAT GGC CCG ACA GCC ATC AGA Thr Leu Ala Val Val Ala Leu Phe Ala Asp Gly Pro Thr Ala Ile Arg 335 340 345 350	1055
GAC GTG GCT TCC TGG AGA GTA AAG GAG ACC GAG AGG ATG GTT GCG ATC Asp Val Ala Ser Trp Arg Val Lys Glu Thr Glu Arg Met Val Ala Ile 355 360 365	1103
CGG ACG GAG CTA ACC AAG CTG GGA GCA TCT GTT GAG GAA GGG CCG GAC Arg Thr Glu Leu Thr Lys Leu Gly Ala Ser Val Glu Glu Gly Pro Asp 370 375 380	1151
TAC TGC ATC ATC ACG CCG CCG GAG AAG CTG AAC GTG ACG GCG ATC GAC Tyr Cys Ile Ile Thr Pro Pro Glu Lys Leu Asn Val Thr Ala Ile Asp 385 390 395	1199
ACG TAC GAC GAC CAC AGG ATG GCC ATG GGC TTC TCC CTT GCC GCC TGT Thr Tyr Asp Asp His Arg Met Ala Met Ala Phe Ser Leu Ala Ala Cys 400 405 410	1247
GCC GAG GTC CCC GTC ACC ATC CCG GAC CCT GGG TGC ACC CCG AAG ACC Ala Glu Val Pro Val Thr Ile Arg Asp Pro Gly Cys Thr Arg Lys Thr 415 420 425 430	1295
TTC CCC GAC TAC TTC GAT GTG CTG AGC ACT TTC GTC AAG AAT Phe Pro Asp Tyr Phe Asp Val Leu Ser Thr Phe Val Lys Asn 435 440	1337
TAA	1340



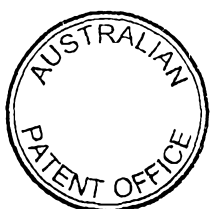
SEQ ID NO. 3.

Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile Ser Gly
 1 5 10 15
 Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile Leu Leu
 20 25 30
 Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu Leu Asn
 35 40 45
 Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu Gly Leu
 50 55 60
 Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val Gly Cys
 65 70 75 80
 Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln Leu Phe
 85 90 95
 Leu Gly Asn Ala Gly Thr Ala Met Arg Pro Leu Thr Ala Ala Val Thr
 100 105 110
 Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro Arg Met
 115 120 125
 Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln Leu Gly
 130 135 140
 Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val Arg Val
 145 150 155 160
 Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser Gly Ser
 165 170 175
 Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro Leu Ala
 180 185 190
 Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser Ile Pro
 195 200 205
 Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val Lys Ala
 210 215 220
 Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly Gln Lys
 225 230 235 240
 Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser Ser Ala
 245 250 255
 Ser Tyr Phe Leu Ala Gly Ala Ala Ile Thr Gly Gly Thr Val Thr Val
 260 265 270
 Glu Gly Cys Gly Thr Thr Ser Leu Gln Gly Asp Val Lys Phe Ala Glu
 275 280 285
 Val Leu Glu Met Met Gly Ala Lys Val Thr Trp Thr Glu Thr Ser Val
 290 295 300
 Thr Val Thr Gly Pro Pro Arg Glu Pro Phe Gly Arg Lys His Leu Lys
 305 310 315 320
 Ala Ile Asp Val Asn Met Asn Lys Met Pro Asp Val Ala Met Thr Leu
 325 330 335
 Ala Val Val Ala Leu Phe Ala Asp Gly Pro Thr Ala Ile Arg Asp Val
 340 345 350
 Ala Ser Trp Arg Val Lys Glu Thr Glu Arg Met Val Ala Ile Arg Thr
 355 360 365
 Glu Leu Thr Lys Leu Gly Ala Ser Val Glu Glu Gly Pro Asp Tyr Cys
 370 375 380
 Ile Ile Thr Pro Pro Glu Lys Leu Asn Val Thr Ala Ile Asp Thr Tyr
 385 390 395 400
 Asp Asp His Arg Met Ala Met Ala Phe Ser Leu Ala Ala Cys Ala Glu
 405 410 415
 Val Pro Val Thr Ile Arg Asp Pro Gly Cys Thr Arg Lys Thr Phe Pro
 420 425 430
 Asp Tyr Phe Asp Val Leu Ser Thr Phe Val Lys Asn
 435 440



SEQ ID No. 4.

CCATG GCC GGC GCC GAG GAG ATC GTG CTG CAG CCC ATC AAG GAG ATC Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile 1 5 10	47
TCC GGC ACC GTC AAG CTG CCG GGG TCC AAG TCG CTT TCC AAC CGG ATC Ser Gly Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile 15 20 25 30	95
CTC CTA CTC GCC GCC CTG TCC GAG GGG ACA ACA GTG GTT GAT AAC CTG Leu Leu Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu 35 40 45	143
CTG AAC AGT GAG GAT GTC CAC TAC ATG CTC GGG GCC TTG AGC ACT CTT Leu Asn Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu 50 55 60	191
GGT CTC TCT GTC GAA GCG GAC AAA GCT GCC AAA AGA GCT GTA GTT GTT Gly Leu Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val 65 70 75	239
GGC TGT GGT GGA AAG TTC CCA GTT GAG GAT GCT AAA GAG GAA GTG CAG Gly Cys Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln 80 85 90	287
CTC TTC TTG GGG AAT GCT GGA ATC SCA ATG CGG TCC TTG ACA GCA GCT Leu Phe Leu Gly Asn Ala Gly Ile Ala Met Arg Ser Leu Thr Ala Ala 95 100 105 110	335
GTT ACT GCT GCT GGT GGA AAT GCA ACT TAC GTG CTT GAT GGA GTA CCA Val Thr Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro 115 120 125	383
AGA ATG AGG GAG AGA CCC ATT GGC GAC TTG GTT GTC GGA TTG AAG CAG Arg Met Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln 130 135 140	431
CTT GGT GCA GAT GTT GAT TGT TTC CTT GGC ACT GAC TGC CCA CCT GTT Leu Gly Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val 145 150 155	479
CGT GTC AAT GGA ATC GGA GGG CTA CCT GGT GGC AAG GTC AAG CTG TCT Arg Val Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser 160 165 170	527
GGC TCC ATC AGC AGT CAG TAC TTG AGT GCC TTG CTG ATG GCT GCT CCT Gly Ser Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro 175 180 185 190	575
TTG GCT CTT GGG GAT GTG GAG ATT GAA ATC ATT GAT AAA TTA ATC TCC Leu Ala Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser 195 200 205	623
ATT CCG TAC GTC GAA ATG ACA TTG AGA TTG ATG GAG CGT TTT GGT GTG Ile Pro Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val 210 215 220	671
AAA GCA GAG CAT TCT GAT AGC TGG GAC AGA TTC TAC ATT AAG GGA GGT Lys Ala Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly 225 230 235	719



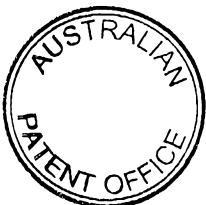
SEQ ID No. 4 (continuation).

CAA AAA TAC AAG TCC CCT AAA AAT GCC TAT GTT GAA GGT GAT GCC TCA Gln Lys Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser 240 245 250	767
AGC GCA AGC TAT TTC TTG GCT GGT GCT GCA ATT ACT GGA GGG ACT GTG Ser Ala Ser Tyr Phe Leu Ala Gly Ala Ala Ile Thr Gly Gly Thr Val 255 260 265 270	815
ACT GTG GAA GGT TGT GGC ACC ACC AGT TTG CAG GCT GAT GTG AAG TTT Thr Val Glu Gly Cys Gly Thr Thr Ser Leu Gln Gly Asp Val Lys Phe 275 280 285	863
GCT GAG GTA CTG GAG ATG ATG GGA GCG AAG GTT ACA TGG ACC GAG ACT Ala Glu Val Leu Glu Met Met Gly Ala Lys Val Thr Trp Thr Glu Thr 290 295 300	911
AGC GTA ACT GTT ACT GGC CCA CCG CGG GAG CCA TTT GGG AGG AAA CAC Ser Val Thr Val Thr Gly Pro Pro Arg Glu Pro Phe Gly Arg Lys His 305 310 315	959
CTC AAG GCG ATT GAT GTC AAC ATG AAC AAG ATG CCT GAT GTC GCC ATG Leu Lys Ala Ile Asp Val Asn Met Asn Lys Met Pro Asp Val Ala Met 320 325 330	1007
ACT CTT GCT GTG GTT GCC CTC TTT GCC GAT GGC CCG ACA GCC ATC AGA Thr Leu Ala Val Val Ala Leu Phe Ala Asp Gly Pro Thr Ala Ile Arg 335 340 345 350	1055
GAC GTG GCT TCC TGG AGA GTA AAG GAG ACC GAG AGG ATG GTT CCG ATC Asp Val Ala Ser Trp Arg Val Lys Glu Thr Thr Arg Met Val Ala Ile 355 360 365	1103
CGG ACG GAG CTA ACC AAG CTG GGA GCA TCT GTT GAG GAA GGG CCG GAC Arg Thr Glu Leu Thr Lys Leu Gly Ala Ser Val Glu Glu Gly Pro Asp 370 375 380	1151
TAC TGC ATC ATC ACG CCG CCG GAG AAG CTG AAC GTG ACG GCG ATC GAC Tyr Cys Ile Ile Thr Pro Pro Glu Lys Leu Asn Val Thr Ala Ile Asp 385 390 395	1195
ACG TAC GAC GAC CAC AGG ATG GCG ATG GCC TTC TCC CTT GCC GCC TGT Thr Tyr Asp Asp His Arg Met Ala Met Ala Phe Ser Leu Ala Ala Cys 400 405 410	1247
GCC GAG GTC CCC GTC ACC ATC CCG GAC CCT GGG TGC ACC CGG AAG ACC Ala Glu Val Pro Val Thr Ile Arg Asp Pro Gly Cys Thr Arg Lys Thr 415 420 425 430	1295
TTC CCC GAC TAC TTC GAT GTG CTG AGC ACT TTC GTC AAG AAT Phe Pro Asp Tyr Phe Asp Val Leu Ser Thr Phe Val Lys Asn 435 440	1337
TAA	1340



SEQ ID No. 5.

Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile Ser Gly
 1 5 10 15
 Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile Leu Leu
 20 25 30
 Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu Leu Asn
 35 40 45
 Ser Glu Asn Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu Gly Leu
 50 55 60
 Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val Gly Cys
 65 70 75 80
 Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln Leu Phe
 85 90 95
 Leu Gly Asn Ala Gly Ile Ala Met Arg Ser Leu Thr Ala Ala Val Thr
 100 105 110
 Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro Arg Met
 115 120 125
 Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln Leu Gly
 130 135 140
 Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val Arg Val
 145 150 155 160
 Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser Gly Ser
 165 170 175
 Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro Leu Ala
 180 185 190
 Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser Ile Pro
 195 200 205
 Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val Lys Ala
 210 215 220
 Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly Gln Lys
 225 230 235 240
 Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser Ser Ala
 245 250 255
 Ser Tyr Phe Leu Ala Gly Ala Ala Ile Thr Gly Gly Thr Val Thr Val
 260 265 270
 Glu Gly Cys Gly Thr Thr Ser Leu Gln Gly Asp Val Lys Phe Ala Glu
 275 280 285
 Val Leu Glu Met Met Gly Ala Lys Val Thr Trp Thr Glu Thr Ser Val
 290 295 300
 Thr Val Thr Gly Pro Pro Arg Glu Pro Phe Gly Arg Lys His Leu Lys
 305 310 315 320
 Ala Ile Asp Val Asn Met Asn Lys Met Pro Asp Val Ala Met Thr Leu
 325 330 335
 Ala Val Val Ala Leu Phe Ala Asp Gly Pro Thr Ala Ile Arg Asp Val
 340 345 350
 Ala Ser Trp Arg Val Lys Glu Thr Glu Arg Met Val Ala Ile Arg Thr
 355 360 365
 Glu Leu Thr Lys Leu Gly Ala Ser Val Glu Glu Gly Pro Asp Tyr Cys
 370 375 380
 Ile Ile Thr Pro Pro Glu Lys Leu Asn Val Thr Ala Ile Asp Thr Tyr
 385 390 395 400
 Asp Asp His Arg Met Ala Met Ala Phe Ser Leu Ala Ala Cys Ala Glu
 405 410 415
 Val Pro Val Thr Ile Arg Asp Pro Gly Cys Thr Arg Lys Thr Phe Pro
 420 425 430
 Asp Tyr Phe Asp Val Leu Ser Thr Phe Val Lys Asn
 435 440



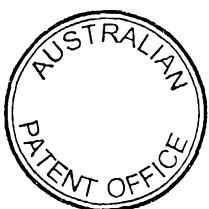
SEQ ID No. 6.

TGAGGTACGA TTCTTCGATC CTCCTTGATT TTCCTGGAAA TATTTTTTCG GTGATCGTGA	60
AACTACTGGA ATCGCTCGAT AGGTGGTACG AAATTAGGCG AGATTAGTTT CTATCTTGG	120
CCATTATCTT GTTCTTCGC CGAATGATCT TCCGTATAAA GATTTTAGGT TAGAGATGAA	180
TCGTATAGCT AGATTTCATC ACCAGTAGT TTCCTTGCT AGAATCTCTG AAATTCCTGA	240
TAGTTTTCAC ATGTGTAAAT AGATTGTCT TATTCGGCGA TTGTGATTA GGGTTTGTAT	300
TTCTTGATT ATGCGATTGC AATTAGGGAT TTCTTTTGGT TTTGTGTGA TCTTACGATA	360
CATTCCTGCA ATTGAATAGG TATGGATCTA AATCTTGTTA ATTTGTTGAA CAGATCCC	418



SEQ ID No. 7.

CTCAGGCGAA GAACAGGTAT GATTGTGTTG TAATTACATC AGGGGTTTAG GTCTTTCCAT	60
TACTTTTAA TGTGTTTCT GTTACTGTCT CCGCGATCTG ATTTTACGAC AATAGAGTTT	120
CGGGTTTGT CCCATTCCAG TTTGAAAATA AAGGTCCGTC TTTTAAGTTT GCTGGATCGA	180
TAAACCTGTG AAGATTGAGT CTAGTCGATT TATTGGATGA TCCATTCTTC ATCGTTTFTT	240
TCTTGCTTCG AAGTTCTGTA TAACCAGATT TGTCTGTGTG CGATTGTCAT TACCTAGCCG	300
TGTATCGAGA ACTAGGGTTT TCGAGTCAAT TTTGCCCCCTT TTGGTTATAT CTGGTTCGAT	360
AACGATTCAT CTGGATTAGG GTTTTAAGTG GTGACGTTTA GTATTCCAAT TTCTTCAAAA	420
TTTAGTTATG GATAATGAAA ATCCCAATT GACTGTTCAA TTTCTTGTTA AATGCGCAGA	480
TCCCGGGATC TCGG	494



- 46 -

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An isolated DNA molecule comprising a nucleotide sequence which is capable of serving as a regulatory zone in a chimeric gene for the transformation of a plant, wherein said nucleotide sequence comprises the nucleotide sequence of at least one intron of the 5'-non-coding region of a plant histone H3.3 gene, and wherein said nucleotide sequence is capable of facilitating the expression of a translation product of said chimeric gene in a region of the plant which undergoes rapid growth.
2. The isolated DNA molecule of claim 1, wherein the intron is the first intron (intron 1) of a plant histone H3.3 gene.
3. The isolated DNA molecule of claim 1 or 2, wherein the plant from which the histone H3.3 gene is derived is a dicotyledonous plant.
4. The isolated DNA molecule of claim 3 wherein the dicotyledonous plant is *Arabidopsis thaliana*.
5. The isolated DNA molecule of claim 1 or 2, wherein the plant from which the histone H3.3 gene is derived is a monocotyledonous plant.
6. The isolated DNA molecule of claim 5, wherein the monocotyledonous plant is *Zea mays*.
7. The isolated DNA molecule according to any one of claims 1 to 6, wherein the intron is in the same orientation within the chimeric gene as occurs in the plant histone H3.3 gene from which it was derived.
8. The isolated DNA molecule according to any one of claims 1 to 6, wherein the intron is in the opposite orientation within the chimeric gene to the orientation of said intron in the plant histone H3.3 gene from which it was derived.



- 47 -

9. The isolated DNA molecule according to any one of claims 1 to 8, wherein the number of introns of said DNA molecule is two.

10. The isolated DNA molecule according to claim 9, wherein each of the introns comprises a different nucleotide sequence.

11. The isolated DNA molecule according to claim 9, wherein both of the introns comprises the same nucleotide sequence.

12. A chimeric gene for the transformation of a plant comprising the isolated DNA molecule according to any one of claims 1 to 11 operably connected to a transgene so as to facilitate expression of the transgene in a region of said plant which undergoes rapid growth.

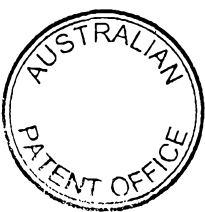
13. The chimeric gene of claim 12 further comprising a first promoter sequence operably connected to the transgene so as to regulate expression of the transgene in a region of said plant which undergoes rapid growth.

14. The chimeric gene according to claim 13 wherein the promoter is derived from a plant histone gene.

15. The chimeric gene according to claim 14 wherein the promoter is derived from the same plant histone H3.3 gene as the intron of the DNA molecule according to any one of claims 1 to 11.

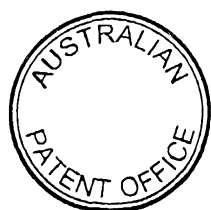
16. The chimeric gene according to claim 14 or 15, wherein the promoter is derived from a duplicated plant histone gene.

17. The chimeric gene according to any one of claims 13 to 16 further comprising a second promoter sequence which is operable in a plant, in combination with the first promoter sequence.



- 48 -

18. The chimeric gene according to any one of claims 12 to 17, wherein the transgene is a herbicide tolerance gene which is expressed to enhance herbicide tolerance on a plant.
19. The chimeric gene of claim 18 wherein the herbicide tolerance gene is fused with a DNA molecule comprising a nucleotide sequence which encodes a transit peptide to facilitate the accumulation of the translation product of said herbicide tolerance gene in a sub-cellular compartment.
20. The chimeric gene of claim 19 wherein the sub-cellular compartment is a plastid.
21. The chimeric gene of claim 20 comprising a first nucleotide sequence encoding at least one transit peptide of an enzyme of plastid location and optionally a downstream second nucleotide sequence encoding the mature N-terminal amino acid sequence of a plastid-localised polypeptide and optionally a third nucleotide sequence downstream of said second nucleotide sequence which encodes a second transit peptide of a plastid-localised enzyme.
22. The chimeric gene according to any one of claims 18 to 21, wherein the herbicide tolerance gene encodes an enzyme which enhances tolerance of a plant to a herbicide which inhibits the activity of a 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme in a plant.
23. The chimeric gene according to claim 22, wherein the herbicide is glyphosate.
24. A vector for the transformation of a plant wherein said vector comprises the chimeric gene according to any one of claims 12 to 23.
25. A bacterial cell or bacterial strain comprising the chimeric gene according to any one of claims 12 to 23 or the vector of claim 24.
26. The bacterial cell or bacterial strain of claim 25 consisting of an *Agrobacterium sp.* cell or *Agrobacterium sp.* strain.



- 49 -

27. The bacterial cell or bacterial strain of claim 26, wherein the *Agrobacterium sp.* is an *A. tumefaciens* cell or strain.

28. A transformed plant cell comprising the chimeric gene according to any one of claims 12 to 23 or the vector of claim 24.

29. A plant or plant part regenerated from the plant cell of claim 28, wherein said plant or plant part comprises the chimeric gene according to any one of claims 12 to 23 or the vector of claim 24.

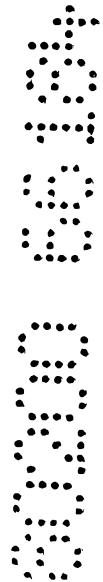
30. A propagule or progeny of the plant or plant part according to claim 29, wherein said propagule or progeny comprises the chimeric gene according to any one of claims 12 to 23 or the vector of claim 24.

31. A method of producing a chimeric gene which is capable of being expressed in a plant cell, wherein said method comprises operably connecting the isolated DNA molecule according to any one of claims 1 to 11 to a first nucleotide sequence encoding a transit peptide and a second nucleotide sequence comprising a transgene in an arrangement so as to facilitate expression of the transgene in a region of said plant which undergoes rapid growth and the targeting of the translation product of said transgene to a sub-cellular compartment of said cell.

32. The method according to claim 31 wherein the first nucleotide sequence is positioned between the DNA molecule and the second nucleotide sequence.

33. The isolated DNA molecule according to any one of claims 1 to 11 substantially as hereinbefore described with reference to the Figures and/or Examples.

34. The chimeric gene according to any one of claims 12 to 23 substantially as hereinbefore described with reference to the Figures and/or Examples.



- 50 -

35. The vector of claim 24 substantially as hereinbefore described with reference to the Figures and/or Examples.

36. The bacterial cell or strain according to any one of claims 25 to 27 substantially as hereinbefore described with reference to the Figures and/or Examples.

37. The plant cell of claim 28 or the plant or plant part of claim 29 or the propagule or progeny of claim 30 substantially as hereinbefore described with reference to the Figures and/or Examples.

38. The method of claim 31 or 32 substantially as hereinbefore described with reference to the Figures and/or Examples.

DATED this EIGHTH day of FEBRUARY, 2000

Rhône-Poulenc Agro

by *Davies Collison Cave*

Patent Attorneys for the Applicant

