



(86) Date de dépôt PCT/PCT Filing Date: 2006/12/01
(87) Date publication PCT/PCT Publication Date: 2007/06/07
(85) Entrée phase nationale/National Entry: 2008/05/28
(86) N° demande PCT/PCT Application No.: US 2006/045908
(87) N° publication PCT/PCT Publication No.: 2007/064828
(30) Priorités/Priorities: 2005/12/01 (US60/741,166);
2006/06/30 (US60/817,799)

(51) Cl.Int./Int.Cl. *C12N 15/54* (2006.01),
A01H 5/00 (2006.01), *C12N 15/82* (2006.01),
C12N 5/10 (2006.01), *C12N 9/10* (2006.01),
C12Q 1/48 (2006.01)

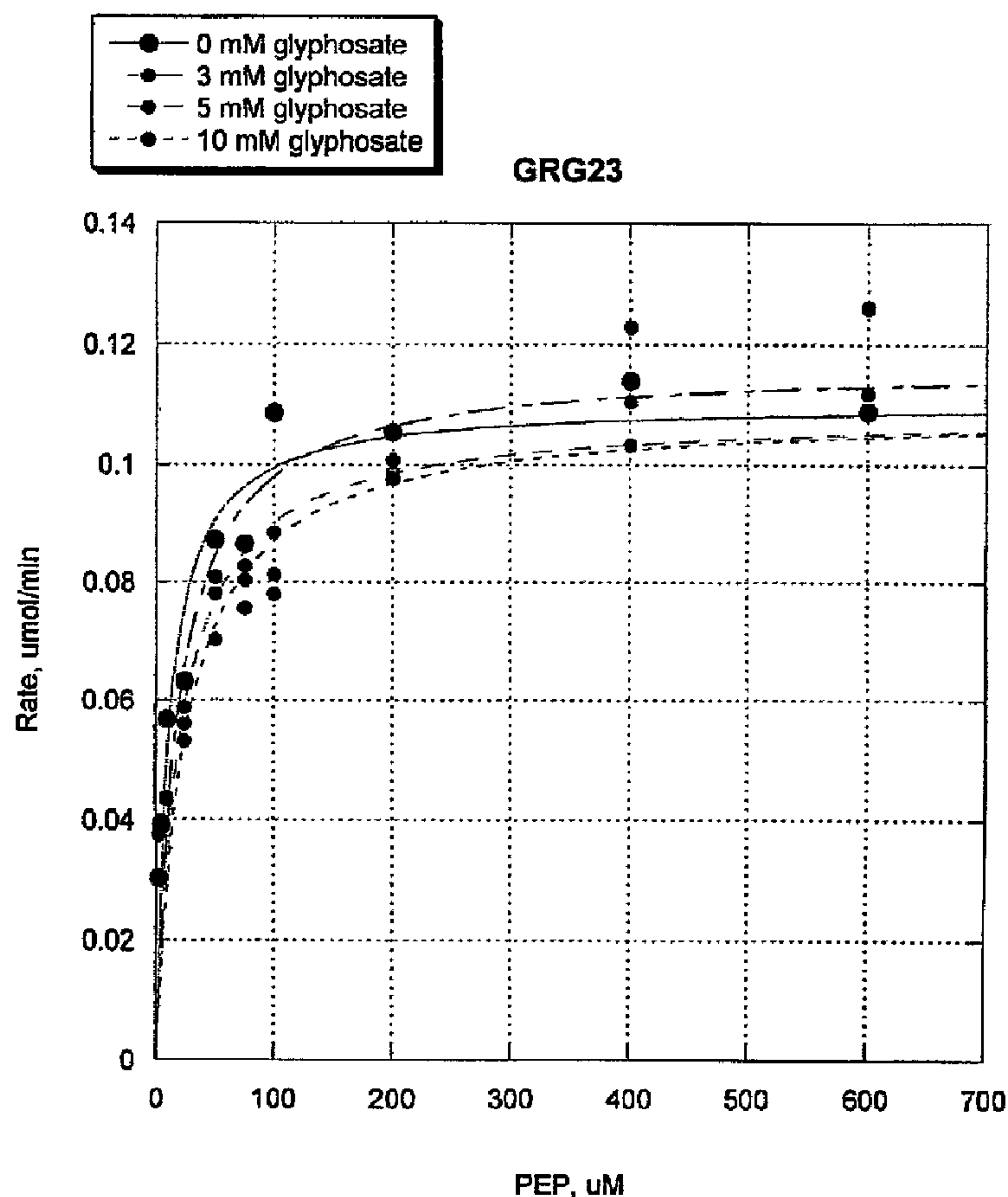
(71) Demandeur/Applicant:
ATHENIX CORPORATION, US

(72) Inventeurs/Inventors:
PETERS, CHERYL L., US;
HINSON, JILL, US;
HAMMER, PHILIP E., US;
VANDE BERG, BRIAN, US;
SCHOUTEN, LAURA COOPER, US;
CARR, BRIAN, US

(74) Agent: SMART & BIGGAR

(54) Titre : GENES GRG23 ET GRG51 CONFERANT LA RESISTANCE AUX HERBICIDES

(54) Title: GRG23 AND GRG51 GENES CONFERRING HERBICIDE RESISTANCE



(57) Abrégé/Abstract:

Compositions and methods for conferring herbicide resistance or tolerance to bacteria, plants, plant cells, tissues and seeds are provided. Compositions include polynucleotides encoding herbicide resistance or tolerance polypeptides, vectors comprising those

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

polynucleotides, and host cells comprising the vectors. The nucleotide sequences of the invention can be used in DNA constructs or expression cassettes for transformation and expression in organisms, including microorganisms and plants. Compositions also comprise transformed bacteria, plants, plant cells, tissues, and seeds.. In particular, isolated polynucleotides encoding glyphosate resistance or tolerance polypeptides are provided. Additionally, amino acid sequences corresponding to the polynucleotides are encompassed. In particular, the present invention provides for isolated polynucleotides comprising nucleotide sequences encoding the amino acid sequence shown in SEQ ID NO:2, 4, or 6, or the nucleotide sequence set forth in SEQ ID NO: 1 , 3, or 5. The present invention additionally provides a method to measure enzyme kinetic activity using fluorogenic substrates.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
7 June 2007 (07.06.2007)

PCT

(10) International Publication Number
WO 2007/064828 A3

(51) International Patent Classification:

C12N 15/54 (2006.01) *A01H 5/00* (2006.01)
C12N 15/82 (2006.01) *C12N 5/10* (2006.01)
C12N 9/10 (2006.01) *C12Q 1/48* (2006.01)

(21) International Application Number:

PCT/US2006/045908

(22) International Filing Date:

1 December 2006 (01.12.2006)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/741,166 1 December 2005 (01.12.2005) US
 60/817,799 30 June 2006 (30.06.2006) US

(71) Applicant (for all designated States except US):
ATHENIX CORPORATION [US/US]; 2202 ELLIS
 ROAD, SUITE B, Durham, NC 27703 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **PETERS, Cheryl,**
 L. [US/US]; 6308 VALLEY ESTATES DRIVE, Raleigh,

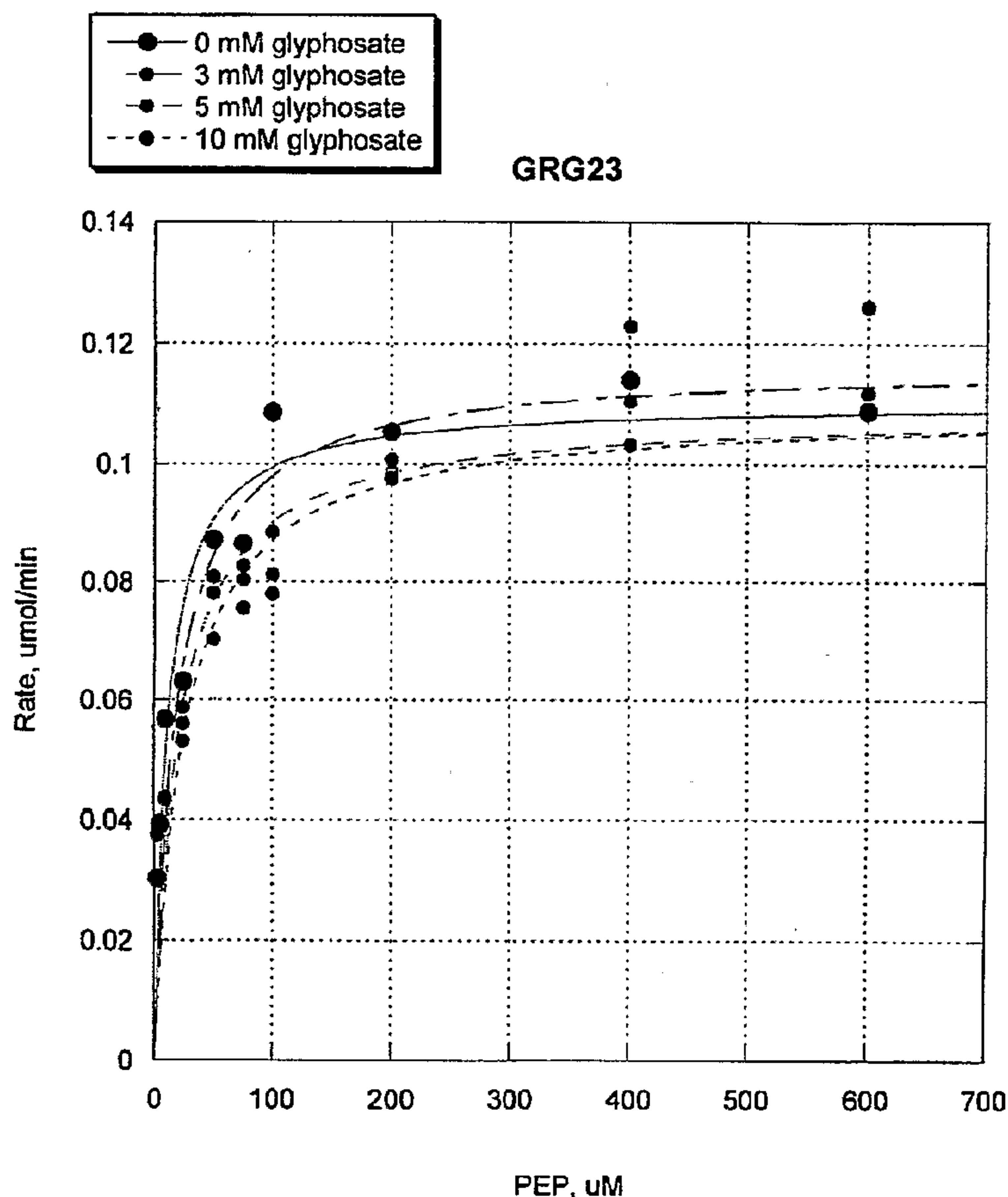
NC 27612 (US). **HINSON, Jill** [US/US]; 3510 Moriah
 Road, Rougemont, North Carolina 27572 (US). **HAM-**
MER, Philip, E. [US/US]; 215 COVINGTON SQUARE
 DRIVE, Cary, NC 27513 (US). **VANDE BERG, Brian**
 [US/US]; 5012 SILHOUETTE DRIVE, Durham, NC
 27713 (US). **SCHOUTEN, Laura, Cooper** [US/US];
 232 PALMER SLOAN ROAD, Sanford, NC 27332 (US).
CARR, Brian [US/US]; 3104 SHINLEAF COURT,
 Raleigh, NC 27613 (US).

(74) Agents: **SPRUILL, W., Murray** et al.; ALSTON
 & BIRD LLP, BANK OF AMERICA PLAZA, 101
 SOUTH TRYON STREET, SUITE 4000, Charlotte, NC
 28280-4000 (US).

(81) Designated States (unless otherwise indicated, for every
 kind of national protection available): AE, AG, AL, AM,
 AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
 CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
 GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS,
 JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS,
 LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY,

[Continued on next page]

(54) Title: GRG23 AND GRG51 GENES CONFERRING HERBICIDE RESISTANCE



(57) Abstract: Compositions and methods for conferring herbicide resistance or tolerance to bacteria, plants, plant cells, tissues and seeds are provided. Compositions include polynucleotides encoding herbicide resistance or tolerance polypeptides, vectors comprising those polynucleotides, and host cells comprising the vectors. The nucleotide sequences of the invention can be used in DNA constructs or expression cassettes for transformation and expression in organisms, including microorganisms and plants. Compositions also comprise transformed bacteria, plants, plant cells, tissues, and seeds. In particular, isolated polynucleotides encoding glyphosate resistance or tolerance polypeptides are provided. Additionally, amino acid sequences corresponding to the polynucleotides are encompassed. In particular, the present invention provides for isolated polynucleotides comprising nucleotide sequences encoding the amino acid sequence shown in SEQ ID NO:2, 4, or 6, or the nucleotide sequence set forth in SEQ ID NO: 1, 3, or 5. The present invention additionally provides a method to measure enzyme kinetic activity using fluorogenic substrates.

WO 2007/064828 A3

WO 2007/064828 A3

MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

(88) Date of publication of the international search report:

2 August 2007

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.

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME __1__ DE __2__

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME __1__ OF __2__

NOTE: For additional volumes please contact the Canadian Patent Office.

GRG23 AND GRG51 GENES CONFERRING HERBICIDE RESISTANCE

FIELD OF THE INVENTION

This invention provides novel genes encoding herbicide resistance, which are useful in plant biology, crop breeding, and plant cell culture.

5 BACKGROUND OF THE INVENTION

N-phosphonomethylglycine, commonly referred to as glyphosate, is an important agronomic chemical. Glyphosate inhibits the enzyme that converts phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid (S3P) to 5-enolpyruvyl-3-phosphoshikimic acid. Inhibition of this enzyme (5-enolpyruvylshikimate-3-phosphate synthase; referred to
10 herein as “EPSPS”) kills plant cells by shutting down the shikimate pathway, thereby inhibiting aromatic acid biosynthesis.

Since glyphosate-class herbicides inhibit aromatic amino acid biosynthesis, they not only kill plant cells, but are also toxic to bacterial cells. Glyphosate inhibits many bacterial EPSP synthases, and thus is toxic to these bacteria. However, certain bacterial
15 EPSP synthases have a high tolerance to glyphosate.

Plant cells resistant to glyphosate toxicity can be produced by transforming plant cells to express glyphosate-resistant bacterial EPSP synthases. Notably, the bacterial gene from *Agrobacterium tumefaciens* strain CP4 has been used to confer herbicide resistance on plant cells following expression in plants. A mutated EPSP synthase from
20 *Salmonella typhimurium* strain CT7 confers glyphosate resistance in bacterial cells, and confers glyphosate resistance on plant cells (U.S. Patent Nos. 4,535,060; 4,769,061; and 5,094,945). However, there is a need for other herbicide resistance genes.

EPSPS kinetic activity can be assayed by measuring the liberation of phosphate. Phosphate liberation is detected using a coupled assay for the fluorescent detection of
25 phosphate based on the generation of *N*-acetyl-3,7-dihydroxyphenoxacine (Amplex® Red), as is known in the art (Vazquez *et al.* (2003) *Analytical Biochemistry* 320: 292–298). The published assay conditions can lead to saturation of the assay in experiments

where phosphate is liberated very quickly. Additional methods are needed for the measurement of EPSPS kinetic activity.

SUMMARY OF INVENTION

5 Compositions and methods for conferring herbicide resistance or tolerance to bacteria, plants, plant cells, tissues and seeds are provided. Compositions include nucleic acid molecules encoding herbicide resistance or tolerance polypeptides, vectors comprising those nucleic acid molecules, and host cells comprising the vectors. Compositions also include antibodies to the herbicide resistance or tolerance
10 polypeptides. As noted, the nucleotide sequences of the invention can be used in DNA constructs or expression cassettes for transformation and expression in organisms, including microorganisms and plants. Compositions also comprise transformed bacteria, plants, plant cells, tissues, and seeds. In addition, methods are provided for producing the polypeptides encoded by the synthetic nucleotides of the invention.

15 Isolated nucleic acid molecules and variants thereof encoding herbicide resistance or tolerance polypeptides are provided. Additionally, amino acid sequences and variants thereof encoded by the polynucleotides that confer herbicide resistance or tolerance are encompassed. The present invention provides for isolated nucleic acid molecules comprising a nucleotide sequence set forth in SEQ ID NO:1, 3, or 5, a nucleotide
20 sequence encoding the amino acid sequence set forth in SEQ ID NO:2, 4, or 6, the herbicide resistance nucleotide sequence deposited in a bacterial host as Accession Nos. NRRL B-30888 or NRRL B-30949, as well as variants and fragments thereof. Nucleotide sequences that are complementary to a nucleotide sequence of the invention, or that hybridize to a sequence of the invention are also encompassed.

25 Methods for measuring enzyme kinetic activity using fluorogenic substrates are also provided.

DESCRIPTION OF FIGURES

Figure 1 shows an alignment of the GRG23 ORF1 amino acid sequence (SEQ ID
30 NO:2) and GRG51 (SEQ ID NO:6) with *Bacillus clausii* (SEQ ID NO:7), *Rubrobacter xylanophilus* (SEQ ID NO:8), *Escherichia coli* (SEQ ID NO:11), *Agrobacterium sp.* strain CP4 (SEQ ID NO:10) and *Zea mays* (SEQ ID NO:9).

Figure 2 shows a scatter plot of GRG23 enzyme activity (y axis) as a function of the concentration of PEP (x axis) at glyphosate concentrations of 0, 3, 5 and 10 mM.

Figure 3 shows a scatter plot of K_m (app) (y axis) as a function of glyphosate concentration (x axis). The $-X$ intercept represents the K_i for glyphosate.

DETAILED DESCRIPTION

5 The present inventions now will be described more fully hereinafter with reference to the accompanying drawings, in which some, but not all embodiments of the inventions are shown. Indeed, these inventions may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these
10 embodiments are provided so that this disclosure will satisfy applicable legal requirements. Like numbers refer to like elements throughout.

Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which these inventions pertain having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the inventions are not to be limited to the specific
15 embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

The present invention is drawn to compositions and methods for regulating
20 herbicide resistance in organisms, particularly in plants or plant cells. The methods involve transforming organisms with nucleotide sequences encoding the glyphosate resistance gene of the invention. The nucleotide sequences of the invention are useful for preparing plants that show increased tolerance to the herbicide glyphosate. Thus, transformed bacteria, plants, plant cells, plant tissues and seeds are provided.

25 Compositions include nucleic acids and proteins relating to herbicide tolerance in microorganisms and plants as well as transformed bacteria, plants, plant tissues and seeds. Nucleotide sequences of the glyphosate resistance gene (*grg23* and *grg51*) and the amino acid sequences of the proteins encoded thereby are disclosed. The sequences find use in the construction of expression vectors for subsequent transformation into plants of
30 interest, as probes for the isolation of other glyphosate resistance genes, as selectable markers, and the like. Thus, by “glyphosate resistance gene of the invention is intended the nucleotide sequence set forth in SEQ ID NO:1 or 3, and variants and fragments thereof (SEQ ID NO:5, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, and 32), that encode a glyphosate resistance or tolerance polypeptide. Likewise, a “glyphosate resistance

polypeptide of the invention” is a polypeptide having the amino acid sequence set forth in

SEQ ID NO:2 or 4, and variants and fragments thereof (SEQ ID NO:6, 15, 17, 19, 21, 23, 25, 27, 29, 31, and 33), that confer glyphosate resistance or tolerance to a host cell.

Plasmids containing the herbicide resistance nucleotide sequences of the invention
 5 were deposited in the permanent collection of the Agricultural Research Service Culture
 Collection, Northern Regional Research Laboratory (NRRL) on November 18, 2005, and
 assigned Accession No. NRRL B-30888 (*grg23*), and on June 26, 2006 and assigned
 Accession No. NRRL B-30949 (*grg51*). This deposit will be maintained under the terms
 of the Budapest Treaty on the International Recognition of the Deposit of
 10 Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as
 a convenience for those of skill in the art and is not an admission that a deposit is required
 under 35 U.S.C. §112.

By “glyphosate” is intended any herbicidal form of N-phosphonomethylglycine
 (including any salt thereof) and other forms that result in the production of the glyphosate
 15 anion *in planta*. An “herbicide resistance protein” or a protein resulting from expression
 of an “herbicide resistance-encoding nucleic acid molecule” includes proteins that confer
 upon a cell the ability to tolerate a higher concentration of an herbicide than cells that do
 not express the protein, or to tolerate a certain concentration of an herbicide for a longer
 time than cells that do not express the protein. A “glyphosate resistance protein” includes
 20 a protein that confers upon a cell the ability to tolerate a higher concentration of
 glyphosate than cells that do not express the protein, or to tolerate a certain concentration
 of glyphosate for a longer period of time than cells that do not express the protein. By
 “tolerate” or “tolerance” is intended either to survive, or to carry out essential cellular
 functions such as protein synthesis and respiration in a manner that is not readily
 25 discernable from untreated cells.

Isolated Nucleic Acid Molecules, and Variants and Fragments Thereof

One aspect of the invention pertains to isolated nucleic acid molecules comprising
 nucleotide sequences encoding herbicide resistance proteins and polypeptides or
 30 biologically active portions thereof, as well as nucleic acid molecules sufficient for use as
 hybridization probes to identify herbicide resistance-encoding nucleic acids. As used
 herein, the term “nucleic acid molecule” is intended to include DNA molecules (e.g.,
 cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or

RNA generated using nucleotide analogs. The nucleic acid molecule can be single-

stranded or double-stranded.

Nucleotide sequences encoding the proteins of the present invention include the sequences set forth in SEQ ID NO:1, 3, and 5 the herbicide resistance nucleotide
5 sequence deposited in a bacterial host as Accession Nos. NRRL B-30888 and NRRL B-30949, and variants, fragments, and complements thereof. By “complement” is intended a nucleotide sequence that is sufficiently complementary to a given nucleotide sequence such that it can hybridize to the given nucleotide sequence to thereby form a stable duplex. The corresponding amino acid sequence for the herbicide resistance protein
10 encoded by these nucleotide sequences is set forth in SEQ ID NO:2, 4, or 6. The invention also encompasses nucleic acid molecules comprising nucleotide sequences encoding partial-length herbicide resistance proteins, and complements thereof.

An “isolated” or “purified” nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when
15 produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Preferably, an “isolated” nucleic acid is free of sequences (preferably protein encoding sequences) that naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For purposes of the invention,
20 “isolated” when used to refer to nucleic acid molecules excludes isolated chromosomes. For example, in various embodiments, the isolated glyphosate resistance-encoding nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequence that naturally flanks the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. An herbicide resistance protein
25 that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of non-herbicide resistance protein (also referred to herein as a “contaminating protein”).

Nucleic acid molecules that are fragments of these herbicide resistance-encoding nucleotide sequences are also encompassed by the present invention. By “fragment” is
30 intended a portion of the nucleotide sequence encoding an herbicide resistance protein. A fragment of a nucleotide sequence may encode a biologically active portion of an herbicide resistance protein, or it may be a fragment that can be used as a hybridization probe or PCR primer using methods disclosed below. Nucleic acid molecules that are fragments of an herbicide resistance nucleotide sequence comprise at least about 15, 20,

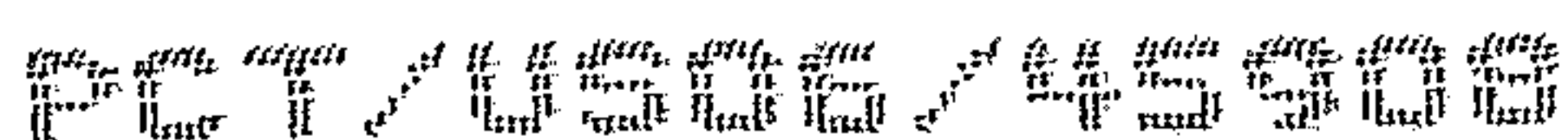
50, 75, 100, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950 contiguous nucleotides, or up to the number of nucleotides present in a full-length herbicide resistance-encoding nucleotide sequence disclosed herein (for example, 1892 nucleotides for SEQ ID NO:1, 1259 nucleotides for SEQ ID NO:3, and 1242 nucleotides for SEQ ID NO:5). By “contiguous” nucleotides is intended nucleotide residues that are immediately adjacent to one another.

Fragments of the nucleotide sequences of the present invention generally will encode protein fragments that retain the biological activity of the full-length glyphosate resistance protein; i.e., herbicide-resistance activity. By “retains herbicide resistance activity” is intended that the fragment will have at least about 30%, at least about 50%, at least about 70%, or at least about 80% of the herbicide resistance activity of the full-length glyphosate resistance proteins disclosed herein as SEQ ID NO:2, 4, or 6. Methods for measuring herbicide resistance activity are well known in the art. See, for example, U.S. Patent Nos. 4,535,060, and 5,188,642, each of which are herein incorporated by reference in their entirety.

A fragment of an herbicide resistance-encoding nucleotide sequence that encodes a biologically active portion of a protein of the invention will encode at least about 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400 contiguous amino acids, or up to the total number of amino acids present in a full-length herbicide resistance protein of the invention (for example, 436 amino acids for SEQ ID NO:2, 413 amino acids for SEQ ID NO:4, and 413 amino acids for SEQ ID NO:6).

Herbicide resistance proteins of the present invention are encoded by a nucleotide sequence sufficiently identical to the nucleotide sequence of SEQ ID NO:1, 3, or 5. The term “sufficiently identical” is intended an amino acid or nucleotide sequence that has at least about 60% or 65% sequence identity, about 70% or 75% sequence identity, about 80% or 85% sequence identity, about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity compared to a reference sequence using one of the alignment programs described herein using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning, and the like.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes. The percent identity



between the two sequences is a function of the number of identical positions shared by the sequences (i.e., percent identity = number of identical positions/total number of positions (e.g., overlapping positions) x 100). In one embodiment, the two sequences are the same length. The percent identity between two sequences can be determined using techniques similar to those described below, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A nonlimiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the BLASTN program, score = 100, wordlength = 12, to obtain nucleotide sequences homologous to GDC-like nucleic acid molecules of the invention. BLAST protein searches can be performed with the BLASTX program, score = 50, wordlength = 3, to obtain amino acid sequences homologous to herbicide resistance protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search that detects distant relationships between molecules. See Altschul *et al.* (1997) *supra*. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) can be used. See www.ncbi.nlm.nih.gov. Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the ClustalW algorithm (Higgins *et al.* (1994) *Nucleic Acids Res.* 22:4673-4680). ClustalW compares sequences and aligns the entirety of the amino acid or DNA sequence, and thus can provide data about the sequence conservation of the entire amino acid sequence. The ClustalW algorithm is used in several commercially available DNA/amino acid analysis software packages, such as the ALIGNX module of the Vector NTI Program Suite (Invitrogen Corporation, Carlsbad, CA). After alignment of amino acid sequences with ClustalW, the percent amino acid identity can be assessed. A non-limiting example of a software program useful for analysis of ClustalW alignments is GeneDoc™. Genedoc™ (Karl Nicholas) allows assessment of amino acid (or DNA) similarity and identity between multiple proteins. Another non-limiting example of a mathematical algorithm

utilized for the comparison of sequences is the algorithm of Myers and Miller (1988)

CABIOS 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0), which is part of the GCG sequence alignment software package (available from Accelrys, Inc., San Diego, CA). When utilizing the ALIGN program for comparing
5 amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

Unless otherwise stated, GAP Version 10, which uses the algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48(3):443-453, will be used to determine sequence identity or similarity using the following parameters: % identity and % similarity for a
10 nucleotide sequence using GAP Weight of 50 and Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity or % similarity for an amino acid sequence using GAP weight of 8 and length weight of 2, and the BLOSUM62 scoring program. Equivalent programs may also be used. By “equivalent program” is intended any sequence comparison program that, for any two sequences in question, generates an
15 alignment having identical nucleotide residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP Version 10.

The invention also encompasses variant nucleic acid molecules. “Variants” of the herbicide resistance-encoding nucleotide sequences include those sequences that encode the herbicide resistance protein disclosed herein but that differ conservatively because of
20 the degeneracy of the genetic code, as well as those that are sufficiently identical as discussed above (for example, SEQ ID NO:5, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, and 32 are variants of SEQ ID NO:1). Naturally occurring allelic variants can be identified with the use of well-known molecular biology techniques, such as polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant nucleotide
25 sequences also include synthetically derived nucleotide sequences that have been generated, for example, by using site-directed mutagenesis but which still encode the herbicide resistance proteins disclosed in the present invention as discussed below. Variant proteins encompassed by the present invention are biologically active, that is they retain the desired biological activity of the native protein, that is, herbicide resistance
30 activity. By “retains herbicide resistance activity” is intended that the variant will have at least about 30%, at least about 50%, at least about 70%, or at least about 80% of the herbicide resistance activity of the native protein. Methods for measuring herbicide resistance activity are well known in the art. See, for example, U.S. Patent Nos.

4,535,060, and 5,188,642, each of which are herein incorporated by reference in their entirety.

The skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of the invention thereby leading to changes in the amino acid sequence of the encoded herbicide resistance proteins, without altering the biological activity of the proteins. Thus, variant isolated nucleic acid molecules can be created by introducing one or more nucleotide substitutions, additions, or deletions into the corresponding nucleotide sequence disclosed herein, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Such variant nucleotide sequences are also encompassed by the present invention.

For example, conservative amino acid substitutions may be made at one or more predicted, nonessential amino acid residues. A “nonessential” amino acid residue is a residue that can be altered from the wild-type sequence of an herbicide resistance protein without altering the biological activity, whereas an “essential” amino acid residue is required for biological activity. A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art.

These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Amino acid substitutions may be made in nonconserved regions that retain function. In general, such substitutions would not be made for conserved amino acid residues, or for amino acid residues residing within a conserved motif, where such residues are essential for protein activity. However, one of skill in the art would understand that functional variants may have minor conserved or nonconserved alterations in the conserved residues.

Lys-22, Arg-124, Asp-313, Arg-344, Arg-386, and Lys-411, are conserved residues of the EPSP synthase from *E. coli* (Schönbrunn *et al.* (2001) *Proc. Natl. Acad. Sci. USA* 98:1376-1380). Conserved residues important for EPSP synthase activity also include Arg-100, Asp-242, and Asp-384 (Selvapandiyan *et al.* (1995) *FEBS Letters*

374:253-256). Arg-27 binds to S3P (Shuttleworth *et al.* (1999) *Biochemistry* 38:296-

302).

Alternatively, variant nucleotide sequences can be made by introducing mutations randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for the ability to confer herbicide resistance activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly, and the activity of the protein can be determined using standard assay techniques.

Using methods such as PCR, hybridization, and the like corresponding herbicide resistance sequences can be identified, such sequences having substantial identity to the sequences of the invention. See, for example, Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) and Innis, *et al.* (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, St. Louis, MO).

In a hybridization method, all or part of the herbicide resistance nucleotide sequence can be used to screen cDNA or genomic libraries. Methods for construction of such cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook and Russell (2001) *supra*. The so-called hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as ^{32}P , or any other detectable marker, such as other radioisotopes, a fluorescent compound, an enzyme, or an enzyme co-factor. Probes for hybridization can be made by labeling synthetic oligonucleotides based on the known herbicide resistance-encoding nucleotide sequence(s) disclosed herein. Degenerate primers designed on the basis of conserved nucleotides or amino acid residues in the nucleotide sequence or encoded amino acid sequence can additionally be used. The probe typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, at least about 25, at least about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, or 1800 consecutive nucleotides of herbicide resistance-encoding nucleotide sequence(s) of the invention or a fragment or variant thereof. Methods for the preparation of probes for hybridization are generally known in the art and are disclosed in Sambrook and Russell (2001) *supra*, and Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), both of which are herein incorporated by reference.

For example, an entire herbicide resistance sequence disclosed herein, or one or

more portions thereof, may be used as a probe capable of specifically hybridizing to
corresponding herbicide resistance sequences and messenger RNAs. To achieve specific
hybridization under a variety of conditions, such probes include sequences that are unique
5 and are at least about 10 nucleotides in length, and at least about 20 nucleotides in length.
Such probes may be used to amplify corresponding herbicide resistance sequences from a
chosen organism by PCR. This technique may be used to isolate additional coding
sequences from a desired organism or as a diagnostic assay to determine the presence of
coding sequences in an organism. Hybridization techniques include hybridization
10 screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook
et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor
Laboratory Press, Cold Spring Harbor, NY).

Hybridization of such sequences may be carried out under stringent conditions.
By “stringent conditions” or “stringent hybridization conditions” is intended conditions
15 under which a probe will hybridize to its target sequence to a detectably greater degree
than to other sequences (e.g., at least 2-fold over background). Stringent conditions are
sequence-dependent and will be different in different circumstances. By controlling the
stringency of the hybridization and/or washing conditions, target sequences that are 100%
complementary to the probe can be identified (homologous probing). Alternatively,
20 stringency conditions can be adjusted to allow some mismatching in sequences so that
lower degrees of similarity are detected (heterologous probing). Generally, a probe is less
than about 1000 nucleotides in length, or less than about 500 nucleotides in length.

Typically, stringent conditions will be those in which the salt concentration is less
than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other
25 salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10
to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50
nucleotides). Stringent conditions may also be achieved with the addition of destabilizing
agents such as formamide. Exemplary low stringency conditions include hybridization
with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl
30 sulphate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium
citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization
in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at
55 to 60°C. Exemplary high stringency conditions include hybridization in 50%
formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C.

Optionally, wash buffers may comprise about 0.1% to about 1% SDS. Duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours.

Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ form}) - 500/L$; where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs.

The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe. T_m is reduced by about 1°C for each 1% of mismatching; thus, T_m , hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. For example, if sequences with $\geq 90\%$ identity are sought, the T_m can be decreased 10°C . Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m , those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a T_m of less than 45°C (aqueous solution) or 32°C (formamide solution), it is preferred to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); and Ausubel *et al.*, eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Herbicide resistance proteins are also encompassed within the present invention.

By “herbicide resistance protein” or “herbicide tolerance protein” is intended a protein having the amino acid sequence set forth in SEQ ID NO:2, 4, or 6. Fragments, biologically active portions, and variants thereof are also provided, and may be used to practice the methods of the present invention.

“Fragments” or “biologically active portions” include polypeptide fragments comprising a portion of an amino acid sequence encoding an herbicide resistance protein as set forth in SEQ ID NO:2, 4, or 6 and that retains herbicide resistance activity. A biologically active portion of an herbicide resistance protein can be a polypeptide that is, for example, 10, 25, 50, 100 or more amino acids in length. Such biologically active portions can be prepared by recombinant techniques and evaluated for herbicide resistance activity. Methods for measuring herbicide resistance activity are well known in the art. See, for example, U.S. Patent Nos. 4,535,060, and 5,188,642, each of which are herein incorporated by reference in their entirety. As used here, a fragment comprises at least 8 contiguous amino acids of SEQ ID NO: 2, 4, or 6. The invention encompasses other fragments, however, such as any fragment in the protein greater than about 10, 20, 30, 50, 100, 150, 200, 250, 300, 350, or 400 amino acids.

By “variants” is intended proteins or polypeptides having an amino acid sequence that is at least about 60%, 65%, at least about 70%, 75%, at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to the amino acid sequence of SEQ ID NO:2, 4, or 6 (for example, SEQ ID NO:6, 15, 17, 19, 21, 23, 25, 27, 29, 31, and 33 are variants of SEQ ID NO:2). Variants also include polypeptides encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule of SEQ ID NO:1, 3, or 5, or a complement thereof, under stringent conditions. Variants include polypeptides that differ in amino acid sequence due to mutagenesis. Variant proteins encompassed by the present invention are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, retaining herbicide resistance activity. Methods for measuring herbicide resistance activity are well known in the art. See, for example, U.S. Patent Nos. 4,535,060, and 5,188,642, each of which are herein incorporated by reference in their entirety.

Bacterial genes, such as the *grg23* or *grg51* gene of this invention, quite often possess multiple methionine initiation codons in proximity to the start of the open reading frame. Often, translation initiation at one or more of these start codons will lead to generation of a functional protein. These start codons can include ATG codons. However,

bacteria such as *Bacillus sp.* also recognize the codon GTG as a start codon, and proteins that initiate translation at GTG codons contain a methionine at the first amino acid.

Furthermore, it is not often determined *a priori* which of these codons are used naturally in the bacterium. Thus, it is understood that use of one of the alternate methionine codons

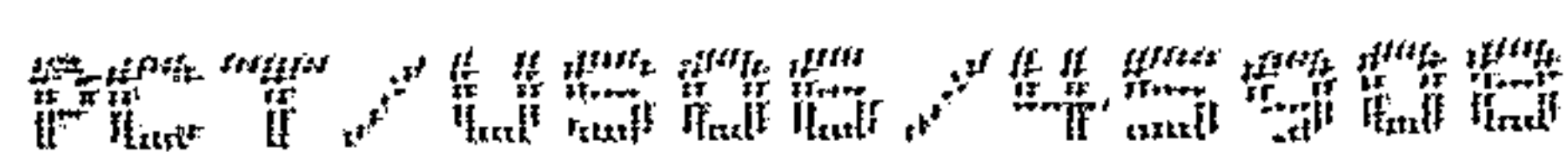
5 may lead to generation of variants of *grg23* or *grg51* that confer herbicide resistance.

These herbicide resistance proteins are encompassed in the present invention and may be used in the methods of the present invention.

Antibodies to the polypeptides of the present invention, or to variants or fragments thereof, are also encompassed. Methods for producing antibodies are well known in the art (see, for example, Harlow and Lane (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY); U.S. Patent No. 4,196,265).

Altered or Improved Variants

It is recognized that the DNA sequence of *grg23* or *grg51* may be altered by various methods, and that these alterations may result in DNA sequences encoding proteins with amino acid sequences different than that encoded by *grg23* or *grg51*. This protein may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the GRG23 or GRG51 protein can be prepared by mutations in the DNA. This may also be accomplished by one of several forms of mutagenesis and/or by directed evolution. In some aspects, the changes encoded in the amino acid sequence will not substantially affect the function of the protein. Such variants will possess the desired herbicide resistance activity. However, it is understood that the ability of GRG23 or GRG51 to confer herbicide resistance may be improved by use of such techniques upon the compositions of the present invention. For example, GRG23 or GRG51 may be expressed in host cells that exhibit high rates of base misincorporation during DNA replication, such as XL-1 Red (Stratagene, La Jolla, CA). After propagation in such strains, *grg23* or *grg51* DNA can be isolated (for example by preparing plasmid DNA, or by amplifying by PCR and cloning the resulting PCR fragment into a vector) and cultured in non-mutagenic strains. Clones containing mutations in *grg23* or *grg51* can be identified by measuring improved resistance to an herbicide such as glyphosate, for example by growing cells in increasing concentrations of glyphosate and testing for clones that confer tolerance to increasing concentrations of glyphosate.



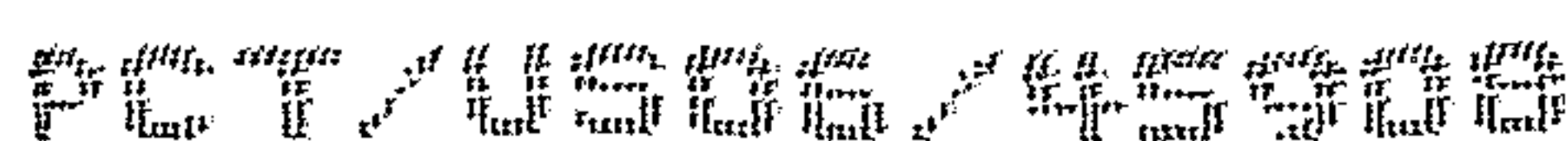
Alternatively, alterations may be made to the protein sequence of many proteins at the amino or carboxy terminus without substantially affecting activity. These alterations can include insertions, deletions, or alterations introduced by modern molecular methods, such as PCR, including PCR amplifications that alter or extend the protein coding
 5 sequence by virtue of inclusion of amino acid encoding sequences in the oligonucleotides utilized in the PCR amplification. Alternatively, the protein sequences added can include entire protein-coding sequences, such as those used commonly in the art to generate protein fusions. Such fusion proteins are often used to (1) increase expression of a protein of interest; (2) introduce a binding domain, enzymatic activity, or epitope to facilitate
 10 either protein purification, protein detection, or other experimental uses known in the art; or, (3) target secretion or translation of a protein to a subcellular organelle, such as the periplasmic space of gram-negative bacteria, or the endoplasmic reticulum of eukaryotic cells, the latter of which often results in glycosylation of the protein.

Variant nucleotide and amino acid sequences of the present invention also
 15 encompass sequences derived from mutagenic and recombinogenic procedures such as DNA shuffling. With such a procedure, one or more different herbicide resistance protein coding regions can be used to create a new herbicide resistance protein possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that
 20 have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the herbicide resistance gene of the invention and other known herbicide resistance genes to obtain a new gene coding for a protein with an improved property of interest, such as an increased glyphosate resistance activity. Strategies for
 25 such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer *et al.* (1997) *Nature Biotech.* 15:436-438; Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

30

Transformation of Bacterial or Plant Cells

Provided herein are novel isolated genes that confer resistance to an herbicide. Also provided are an amino acid sequences of the GRG23 and GRG51 proteins. The protein resulting from translation of this gene allows cells to function in the presence of



concentrations of an herbicide that are otherwise toxic to cells including plant cells and bacterial cells.

In one aspect of the invention, the *grg23* or *grg51* gene is useful as a marker to assess transformation of bacterial or plant cells.

5 By engineering *grg23* or *grg51* to be (1) expressed from a bacterial promoter known to stimulate transcription in the organism to be tested, (2) properly translated to generate an intact GRG23 or GRG51 peptide, and (3) placing the cells in an otherwise toxic concentration of herbicide, cells that have been transformed with DNA by virtue of their resistance to herbicide can be identified. By “promoter” is intended a nucleic acid
10 sequence that functions to direct transcription of a downstream coding sequence. The promoter, together with other transcriptional and translational regulatory nucleic acid sequences, (also termed “control sequences”) are necessary for the expression of a DNA sequence of interest.

Transformation of bacterial cells is accomplished by one of several techniques
15 known in the art, including but not limited to, electroporation or chemical transformation (See, for example, Ausubel (ed.) (1994) *Current Protocols in Molecular Biology* (John Wiley and Sons, Inc., Indianapolis, IN)). Markers conferring resistance to toxic substances are useful in identifying transformed cells (having taken up and expressed the test DNA) from non-transformed cells (those not containing or not expressing the test
20 DNA). In one aspect of the invention, the *grg23* or *grg51* gene is useful as a marker to assess transformation of bacterial or plant cells.

Transformation of plant cells can be accomplished in a similar fashion. By “plant” is intended whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds, plant cells, propagules, embryos and progeny of the same. Plant cells can be differentiated or
25 undifferentiated (e.g. callus, suspension culture cells, protoplasts, leaf cells, root cells, phloem cells, pollen). “Transgenic plants” or “transformed plants” or “stably transformed” plants, cells or tissues refer to plants that have incorporated or integrated exogenous nucleic acid sequences or DNA fragments into the plant cell. By “stable transformation” is intended that the nucleotide construct introduced into a plant integrates
30 into the genome of the plant and is capable of being inherited by progeny thereof.

The *grg23* or *grg51* gene of the invention may be modified to obtain or enhance expression in plant cells. The herbicide resistance sequences of the invention may be provided in expression cassettes for expression in the plant of interest. “Plant expression cassette” includes DNA constructs that are capable of resulting in the expression of a

protein from an open reading frame in a plant cell. The cassette will include in the 5'-3' direction of transcription, a transcriptional initiation region (i.e., promoter) operably-linked to a DNA sequence of the invention, and a translational and transcriptional termination region (i.e., termination region) functional in plants. The cassette may additionally contain at least one additional gene to be cotransformed into the organism, such as a selectable marker gene. Alternatively, the additional gene(s) can be provided on multiple expression cassettes. Such an expression cassette is provided with a plurality of restriction sites for insertion of the herbicide resistance sequence to be under the transcriptional regulation of the regulatory regions.

The promoter may be native or analogous, or foreign or heterologous, to the plant host and/or to the DNA sequence of the invention. Additionally, the promoter may be the natural sequence or alternatively a synthetic sequence. Where the promoter is "native" or "homologous" to the plant host, it is intended that the promoter is found in the native plant into which the promoter is introduced. Where the promoter is "foreign" or "heterologous" to the DNA sequence of the invention, it is intended that the promoter is not the native or naturally occurring promoter for the operably linked DNA sequence of the invention. "Heterologous" generally refers to the nucleic acid sequences that are not endogenous to the cell or part of the native genome in which they are present, and have been added to the cell by infection, transfection, microinjection, electroporation, microprojection, or the like. By "operably linked" is intended a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

Often, such constructs will also contain 5' and 3' untranslated regions. Such constructs may contain a "signal sequence" or "leader sequence" to facilitate co-translational or post-translational transport of the peptide of interest to certain intracellular structures such as the chloroplast (or other plastid), endoplasmic reticulum, or Golgi apparatus, or to be secreted. For example, the gene can be engineered to contain a signal peptide to facilitate transfer of the peptide to the endoplasmic reticulum. By "signal sequence" is intended a sequence that is known or suspected to result in cotranslational or post-translational peptide transport across the cell membrane. In eukaryotes, this transport typically involves secretion into the Golgi apparatus, with some

resulting glycosylation. By "leader sequence" is intended any sequence that when translated, results in an amino acid sequence sufficient to trigger co-translational transport of the peptide chain to a sub-cellular organelle. Thus, this includes leader sequences targeting transport and/or glycosylation by passage into the endoplasmic reticulum, passage to vacuoles, plastids including chloroplasts, mitochondria, and the like. The plant expression cassette can also be engineered to contain an intron, such that mRNA processing of the intron is required for expression.

By "3' untranslated region" is intended a nucleotide sequence located downstream of a coding sequence. Polyadenylation signal sequences and other sequences encoding regulatory signals capable of affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor are 3' untranslated regions. By "5' untranslated region" is intended a nucleotide sequence located upstream of a coding sequence.

Other upstream or downstream untranslated elements include enhancers. Enhancers are nucleotide sequences that act to increase the expression of a promoter region. Enhancers are well known in the art and include, but are not limited to, the SV40 enhancer region and the 35S enhancer element.

The termination region may be native with the transcriptional initiation region, may be native with the herbicide resistance sequence of the present invention, or may be derived from another source. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; and Joshi *et al.* (1987) *Nucleic Acid Res.* 15:9627-9639.

Where appropriate, the gene(s) may be optimized for increased expression in the transformed host cell. That is, the genes can be synthesized using host cell-preferred codons for improved expression, or may be synthesized using codons at a host-preferred codon usage frequency. Generally, the GC content of the gene will be increased. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of host-preferred codon usage. Methods are known in the art for synthesizing host-preferred genes. See, for example, U.S. Patent Nos. 6,320,100; 6,075,185; 5,380,831; and 5,436,391, U.S. Published Application Nos. 20040005600 and 20010003849, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference.

FIG. 1

In one embodiment, the nucleic acids of interest are targeted to the chloroplast for expression. In this manner, where the nucleic acid of interest is not directly inserted into the chloroplast, the expression cassette will additionally contain a nucleic acid encoding a transit peptide to direct the gene product of interest to the chloroplasts. Such transit
5 peptides are known in the art. See, for example, Von Heijne *et al.* (1991) *Plant Mol. Biol. Rep.* 9:104-126; Clark *et al.* (1989) *J. Biol. Chem.* 264:17544-17550; Della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968; Romer *et al.* (1993) *Biochem. Biophys. Res. Commun.* 196:1414-1421; and Shah *et al.* (1986) *Science* 233:478-481.

The nucleic acids of interest to be targeted to the chloroplast may be optimized for
10 expression in the chloroplast to account for differences in codon usage between the plant nucleus and this organelle. In this manner, the nucleic acids of interest may be synthesized using chloroplast-preferred codons. See, for example, U.S. Patent No. 5,380,831, herein incorporated by reference.

Typically this “plant expression cassette” will be inserted into a “plant
15 transformation vector.” By “transformation vector” is intended a DNA molecule that is necessary for efficient transformation of a cell. Such a molecule may consist of one or more expression cassettes, and may be organized into more than one “vector” DNA molecule. For example, binary vectors are plant transformation vectors that utilize two non-contiguous DNA vectors to encode all requisite cis- and trans-acting functions for
20 transformation of plant cells (Hellens and Mullineaux (2000) *Trends in Plant Science* 5:446-451). “Vector” refers to a nucleic acid construct designed for transfer between different host cells. “Expression vector” refers to a vector that has the ability to incorporate, integrate and express heterologous DNA sequences or fragments in a foreign cell.

25 This plant transformation vector may be comprised of one or more DNA vectors needed for achieving plant transformation. For example, it is a common practice in the art to utilize plant transformation vectors that are comprised of more than one contiguous DNA segment. These vectors are often referred to in the art as “binary vectors.” Binary vectors as well as vectors with helper plasmids are most often used for *Agrobacterium*-
30 mediated transformation, where the size and complexity of DNA segments needed to achieve efficient transformation is quite large, and it is advantageous to separate functions onto separate DNA molecules. Binary vectors typically contain a plasmid vector that contains the cis-acting sequences required for T-DNA transfer (such as left border and right border), a selectable marker that is engineered to be capable of expression in a plant

cell, and a “gene of interest” (a gene engineered to be capable of expression in a plant cell for which generation of transgenic plants is desired). Also present on this plasmid vector are sequences required for bacterial replication. The cis-acting sequences are arranged in a fashion to allow efficient transfer into plant cells and expression therein. For example, the selectable marker gene and the gene of interest are located between the left and right borders. Often a second plasmid vector contains the trans-acting factors that mediate T-DNA transfer from *Agrobacterium* to plant cells. This plasmid often contains the virulence functions (Vir genes) that allow infection of plant cells by *Agrobacterium*, and transfer of DNA by cleavage at border sequences and vir-mediated DNA transfer, as in understood in the art (Hellens and Mullineaux (2000) *Trends in Plant Science*, 5:446-451). Several types of *Agrobacterium* strains (e.g. LBA4404, GV3101, EHA101, EHA105, etc.) can be used for plant transformation. The second plasmid vector is not necessary for transforming the plants by other methods such as microprojection, microinjection, electroporation, polyethylene glycol, etc.

Plant Transformation

Methods of the invention involve introducing a nucleotide construct into a plant. By “introducing” is intended to present to the plant the nucleotide construct in such a manner that the construct gains access to the interior of a cell of the plant. The methods of the invention do not require that a particular method for introducing a nucleotide construct to a plant is used, only that the nucleotide construct gains access to the interior of at least one cell of the plant. Methods for introducing nucleotide constructs into plants are known in the art including, but not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

In general, plant transformation methods involve transferring heterologous DNA into target plant cells (e.g. immature or mature embryos, suspension cultures, undifferentiated callus, protoplasts, etc.), followed by applying a maximum threshold level of appropriate selection (depending on the selectable marker gene and in this case “glyphosate”) to recover the transformed plant cells from a group of untransformed cell mass. Explants are typically transferred to a fresh supply of the same medium and cultured routinely. Subsequently, the transformed cells are differentiated into shoots after placing on regeneration medium supplemented with a maximum threshold level of selecting agent (e.g. “glyphosate”). The shoots are then transferred to a selective rooting medium for recovering rooted shoot or plantlet. The transgenic plantlet then grow into

mature plants and produce fertile seeds (e.g. Hiei *et al.* (1994) *The Plant Journal* 6:271-282; Ishida *et al.* (1996) *Nature Biotechnology* 14:745-750). Explants are typically transferred to a fresh supply of the same medium and cultured routinely. A general description of the techniques and methods for generating transgenic plants are found in
 5 Ayres and Park (1994) *Critical Reviews in Plant Science* 13:219-239 and Bommineni and Jauhar (1997) *Maydica* 42:107-120. Since the transformed material contains many cells; both transformed and non-transformed cells are present in any piece of subjected target callus or tissue or group of cells. The ability to kill non-transformed cells and allow transformed cells to proliferate results in transformed plant cultures. Often, the ability to
 10 remove non-transformed cells is a limitation to rapid recovery of transformed plant cells and successful generation of transgenic plants. Molecular and biochemical methods can be used to confirm the presence of the integrated heterologous gene of interest in the genome of transgenic plant.

Generation of transgenic plants may be performed by one of several methods,
 15 including, but not limited to, introduction of heterologous DNA by *Agrobacterium* into plant cells (*Agrobacterium*-mediated transformation), bombardment of plant cells with heterologous foreign DNA adhered to particles, and various other non-particle direct-mediated methods (e.g. Hiei *et al.* (1994) *The Plant Journal* 6:271-282; Ishida *et al.* (1996) *Nature Biotechnology* 14:745-750; Ayres and Park (1994) *Critical Reviews in*
 20 *Plant Science* 13:219-239; Bommineni and Jauhar (1997) *Maydica* 42:107-120) to transfer DNA.

Methods for transformation of chloroplasts are known in the art. See, for example, Svab *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530; Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab and Maliga (1993) *EMBO J.* 12:601-
 25 606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. Additionally, plastid transformation can be accomplished by transactivation of a silent plastid-borne transgene by tissue-preferred expression of a nuclear-encoded and plastid-directed RNA polymerase. Such a system has been reported in McBride *et al.* (1994)
 30 *Proc. Natl. Acad. Sci. USA* 91:7301-7305.

The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the

desired phenotypic characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, the present invention provides

5 transformed seed (also referred to as “transgenic seed”) having a nucleotide construct of the invention, for example, an expression cassette of the invention, stably incorporated into their genome.

Measurement of EPSPS activity

10 In one embodiment of the present invention, the glyphosate-resistant EPSPS enzyme has a K_m for phosphoenolpyruvate (PEP) between about 1 and about 150uM, including about 2 uM, about 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130 or about 140 uM, and a K_i (glyphosate)/ K_m (PEP) between about 500 and about 1000, about 550, about 600, 650, 700, 750, 800, 850, 900, 950, or up to about 1000. As

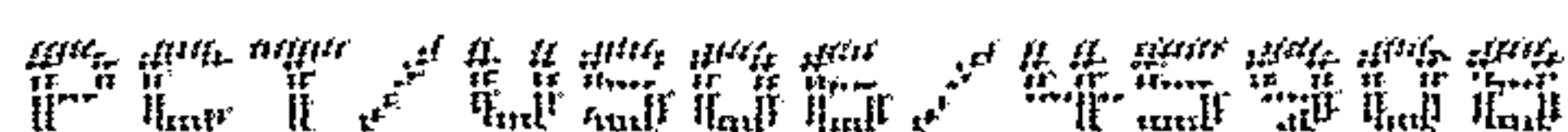
15 used herein, K_m and K_i are measured under conditions in which the enzyme obeys Michaelis-Menten kinetics, around pH 7. One nonlimiting measurement technique uses the enzyme in purified form in potassium chloride and HEPES buffer at pH 7 at room temperature and uses concentrations of glyphosate from 0 to 10mM.

EPSPS kinetic activity can be assayed, for example, by measuring the liberation of

20 phosphate that results during the catalysis of a substrate of EPSPS (for example, PEP and S3P) to its subsequent reaction product (for example, 5-enolpyruvyl-3-phosphoshikimic acid) using a fluorescent assay described by Vazquez *et al.* (2003) *Anal. Biochem.* 320(2):292-298. This assay is based on the oxidation of the non-fluorescent compound *N*-acetyl-3,7-dihydroxyphenoxacine (Amplex® Red, Invitrogen, Carlsbad, CA) to the

25 fluorescent compound resorufin by hydrogen peroxide (Zhou and Panchuk-Voloshina (1997) *Anal. Biochem.* 253:169-174). The reaction relies on the utilization of phosphate by purine nucleoside phosphorylase (PNP), xanthine oxidase (XOD), and horseradish peroxidase (HRP). Phosphate liberation is linked to the level of fluorescence that results from the conversion of Amplex® Red to resorufin. Fluorescence can be measured, for

30 example, using a filter fluorometer, plate reader, spectrofluorometer, spectrophotometer, or the like, using methods well known in the art. The fluorescence generated by the reaction can be detected using a fluorometer set for excitation in the range of about 530 to about 560 nm and an emission of about 590 nm. Absorbance can be detected (for example, using a spectrophotometer or plate reader) at about 565 nm.



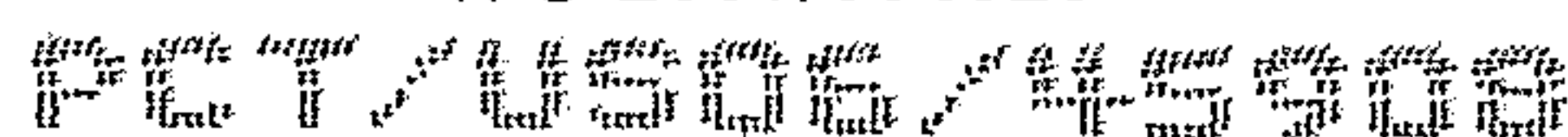
In one embodiment, the present invention encompasses an alteration of the previously reported assay conditions to extend the dynamic range of the assay to accommodate a wider range of substrate concentrations. The alteration comprises a concentration of XOD of at least 1 U/ml, about 1 to about 1.25 U/ml, about 1.25 to about 1.5 U/ml, about 1.5 to about 2 U/ml, or greater than 2 U/ml; a concentration of PNP greater than 0.1 U/ml, about 0.1 to about 0.5 U/ml, about 0.5 to about 1 U/ml, about 1 to about 1.5 U/ml, about 1.5 to about 2 U/ml, or greater than 2 U/ml; and a concentration of Amplex® Red greater than 100 μ M, about 100 to about 200 μ M, about 200 to about 300 μ M, about 300 to about 400 μ M, about 400 to about 500 μ M, about 500 to about 600 μ M, about 700 to about 800 μ M, about 800 to about 900 μ M, about 900 to about 1000 μ M, or greater than about 1000 μ M. This modification can be applied to assays measuring the kinetic activity of any enzyme in which phosphate is liberated during a reaction catalyzed by the enzyme.

15 Plants

The present invention may be used for transformation of any plant species, including, but not limited to, monocots and dicots. Examples of plants of interest include, but are not limited to, corn (maize), sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, and oilseed rape, *Brassica* sp., alfalfa, rye, millet, safflower, peanuts, sweet potato, cassava, coffee, coconut, pineapple, citrus trees, cocoa, tea, banana, avocado, fig, guava, mango, olive, papaya, cashew, macadamia, almond, oats, vegetables, ornamentals, and conifers.

Vegetables include, but are not limited to, tomatoes, lettuce, green beans, lima beans, peas, and members of the genus *Curcumis* such as cucumber, cantaloupe, and musk melon. Ornamentals include, but are not limited to, azalea, hydrangea, hibiscus, roses, tulips, daffodils, petunias, carnation, poinsettia, and chrysanthemum. Preferably, plants of the present invention are crop plants (for example, maize, sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, oilseed rape, etc.).

This invention is particularly suitable for any member of the monocot plant family including, but not limited to, maize, rice, barley, oats, wheat, sorghum, rye, sugarcane, pineapple, yams, onion, banana, coconut, and dates.



Evaluation of Plant Transformation

Following introduction of heterologous foreign DNA into plant cells, the transformation or integration of the heterologous gene in the plant genome is confirmed by various methods such as analysis of nucleic acids, proteins and metabolites associated with the integrated gene.

PCR analysis is a rapid method to screen transformed cells, tissue or shoots for the presence of incorporated gene at the earlier stage before transplanting into the soil (Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)). PCR is carried out using oligonucleotide primers specific to the gene of interest or *Agrobacterium* vector background, etc.

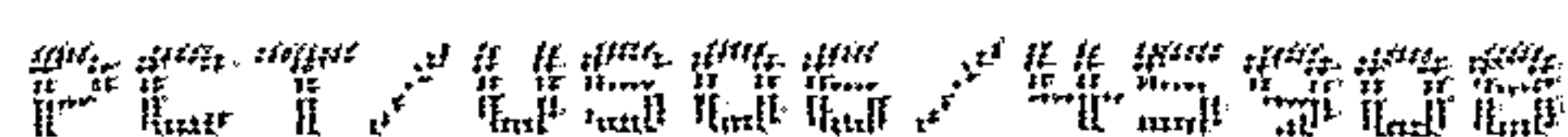
Plant transformation may be confirmed by Southern blot analysis of genomic DNA (Sambrook and Russell (2001) *supra*). In general, total DNA is extracted from the transformant, digested with appropriate restriction enzymes, fractionated in an agarose gel and transferred to a nitrocellulose or nylon membrane. The membrane or "blot" can then be probed with, for example, radiolabeled ^{32}P target DNA fragment to confirm the integration of the introduced gene in the plant genome according to standard techniques (Sambrook and Russell, 2001, *supra*).

In Northern analysis, RNA is isolated from specific tissues of transformant, fractionated in a formaldehyde agarose gel, and blotted onto a nylon filter according to standard procedures that are routinely used in the art (Sambrook and Russell (2001) *supra*). Expression of RNA encoded by *grg23* or *grg51* is then tested by hybridizing the filter to a radioactive probe derived from a GDC by methods known in the art (Sambrook and Russell (2001) *supra*).

Western blot and biochemical assays and the like may be carried out on the transgenic plants to determine the presence of protein encoded by the herbicide resistance gene by standard procedures (Sambrook and Russell (2001) *supra*) using antibodies that bind to one or more epitopes present on the herbicide resistance protein.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL



Example 1: Isolation of ATX21308

ATX 21308 was isolated by plating samples of soil on Enriched Minimal Media 3 (EMM3) containing phosphates and 50 mM glyphosate. Since EMM3 contains no aromatic amino acids, a strain must be resistant to glyphosate in order to grow on this media.

Approximately one gram of soil is suspended in approximately 10 ml of water, and mixed in a vortex mixer for 5 seconds. 100 µl of this suspension is added to 1 ml of EMM3 with phosphate but no glyphosate. EMM3 contains (per liter, pH 7.0): 10 g sucrose, 1 g NH₄Cl, 0.2 g MgSO₄ · 7H₂O, 0.01 g FeSO₄ · 7H₂O, 0.007 g MnSO₄ · H₂O and 10 ml of phosphate solution containing (per liter, pH 7.0) 210 g Na₂HPO₄ and 90 g NaH₂PO₄. The culture is shaken on a tissue culture roller drum at 21°C overnight and then plated onto EMM3 agar containing 50 mM glyphosate. After three days, the isolate is plated onto Luria Bertani (LB) agar to confirm single morphology. After six days, a single colony is streaked onto EMM3 agar containing 50 mM glyphosate. The isolate grew overnight on 50 mM glyphosate plates. One particular strain, designated ATX21308, was selected due to its ability to grow in the presence of high glyphosate concentrations. This strain is tested for its ability to grow in the presence of glyphosate in liquid culture and is capable of growing up to approximately 300mM glyphosate under the conditions tested.

Example 2. Preparation and Screening of Cosmid Libraries

Total DNA was extracted from a culture of ATX21308 using methods commonly known in the art. The DNA was partially digested with restriction enzyme *Sau3A1* and ligated with SuperCos (Stratagene) vector fragment according to the manufacturer's directions. Ligation products were packaged into phage particles using GigaPack III XL packaging extract (Stratagene), transfected into *E. coli* cells, and plated on LB Agar containing 50 µg/ml kanamycin to select for colonies containing cosmids.

Individual colonies were picked into 384-well plates containing LB broth and 50 µg/ml kanamycin, and grown to saturation. Cells from these cultures were diluted 1:10, then pinned onto M63 agar plates containing 50 µg/ml kanamycin, and either 0 mM, 10 mM, 20 mM, or 50 mM glyphosate. [M63 agar medium 100 mM KH₂PO₄, 15 mM (NH₄)₂SO₄, 50 µM CaCl₂, 1 µM FeSO₄, 50 µM MgCl₂, 55 mM glucose, 25 mg/liter L-

proline, 10 mg/liter thiamine HCl, sufficient NaOH to adjust the pH to 7.0, and 15 g/liter agar]. Transformants that grow more rapidly at the higher glyphosate concentrations were isolated and digested with restriction enzyme *EcoR* I to identify cosmids with shared restriction patterns. Several clones which grew in the presence of glyphosate and share similar *EcoR* I restriction patterns were identified. One of these cosmid clones, pAX1924, was selected for further experiments.

Example 3. Identification of *grg23* in cosmid pAX1924

To identify the gene(s) responsible for the glyphosate resistance shown by cosmid pAX1924, DNA from this clone is mutagenized with transposable elements. In this method, clones that have suffered transposon insertions and have lost the ability to confer glyphosate resistance are identified. The location of the transposon insertions identifies the open reading frame responsible for the glyphosate resistance phenotype.

Cosmid pAX1924 is subjected to *in vitro* transposon mutagenesis using an EZ::TN Insertion Kit (Epicentre, Madison, WI) according to the manufacturer's protocol. This process randomly inserts a transposon fragment into the cosmid DNA and thus randomly disrupts the function of genes in the cosmid. This particular transposon contains a gene encoding resistance to trimethoprim, so transposon insertion clones may be selected by the ability to grow in the presence of that antibiotic. The locations of the transposon insertions may be determined by restriction fragment mapping or by sequencing with primers that anneal in the transposon. Transposon insertion clones of pAX1924 are plated on M63 medium containing glyphosate. Multiple transposon-containing clones are identified which have lost the ability to grow in the presence of glyphosate, indicating that the transposon has disrupted the gene responsible for resistance.

The DNA sequence is determined for the region of pAX1924 containing the transposon insertions using sequencing methods well known in the art. Using this sequence information, DNA primers are synthesized and utilized to determine the DNA sequence of pAX1924 in the region encompassing the transposon insertions. Analysis of the resulting DNA sequence shows that this region contains a single gene. This gene is designated herein as *grg23*. Analysis of *grg23* shows that it is capable of yielding two possible proteins in bacterial cells due to the presence of potential alternate translational start sites. The first ORF (ORF1) initiates with a GTG start codon at positions 109-111 of SEQ ID NO:1, and ends at a TAG stop codon at nucleotides 1417-1419 of SEQ ID NO:1.

The second ORF (ORF2) starts at an ATG start codon at nucleotides 178-180 of SEQ ID NO:1, and ends at the TAG stop codon at nucleotides 1417-1419 of SEQ ID NO:1.

Translation of ORF1 yields the amino acid sequence set forth in SEQ ID NO:2.

Translation of ORF2 yields the amino acid sequence set forth in SEQ ID NO:4.

5 Analysis of the DNA region surrounding *grg23* suggests that ORF2 is preceded by a ribosome binding site, whereas there is no obvious ribosome binding site preceding the ORF1 translation start. Furthermore, alignment of both open reading frames with representative EPSPS enzymes shows that few EPSPS enzymes contain this N-terminal extension encoded within ORF1. Thus, the functional ORF encoded by *grg23* in bacteria
10 is ORF2. Therefore, as used herein, GRG23 refers to that which is encoded by ORF2 (nucleotides 178-1419 of SEQ ID NO:1). Nonetheless, it is well known in the art that EPSPS enzymes are quite tolerant of additional amino acids at their N-terminus. Therefore, expression of ORF1 (nucleotides 109-1419 of SEQ ID NO:1) should also yield an EPSPS that confers glyphosate resistance.

15 To test the ability of ORF2 to function as an EPSPS and confer glyphosate resistance upon cells, this open reading frame can be subcloned and expressed in *E. coli*.

Example 4. Subcloning of *grg23* into vectors for expression in *E. coli*.

The gene encoding GRG23 ORF2 (starting with ATG (positions 178-180 of SEQ
20 ID NO:1), expressing a 413 amino acid protein) was subcloned into pUC18 and pRSF1b using the same cloning strategy outlined above. A PCR primer [5' CAGGGATCCGGCATGGAACTGATCGACTAGTG 3'] was synthesized that adds a *Bam*HI site followed by GGC (5'-GGATCCGGC-3') immediately 5' of the start site. A 2nd primer was synthesized [5' ATTGGCGCGCCCTAGCCGGGGAGCGTAAG 3'] that
25 added an *Asc*I site immediately 3' of the stop sequence (5'-GGCGCGCC-3'). The *grg23* coding region is amplified by PCR using PFUULTRA™ DNA polymerase (Stratagene). Following PCR amplification of *grg23* using these primers and restriction digestion with *Bam*HI/*Asc*I, the PCR product was ligated into pUC19 (digested with *Bam*HI and *Asc*I) and pRSF1b (digested with *Bam*HI and *Asc*I), and insert-containing colonies were
30 obtained. The pUC18-*grg23* clone (designated herein as pAX1927) was confirmed by restriction digestion and by DNA sequencing.

Similarly, the expression vector pAX1909 was digested with *Bam*HI and *Asc*I, and the vector containing the fragment was gel-purified by methods well known in the art. pAX1909 is a derivative of PRSF-1b (Novagen, San Diego, CA), modified to contain

a *Bam*HI site directly 3' of the region encoding the histidine rich "His-Tag." Thus, proteins cloned into pAX1909 are in-frame fusions that contain the additional amino acids MAHHHHHHGSG. Vectors such as pAX1909 are typically developed for protein expression and purification, and these methods are well known in the art.

5 The digested PCR product resulting above was ligated into the digested pAX1909 vector, and insert-containing colonies were obtained. The pAX1909-*grg23* clone (designated herein as pAX1926) was confirmed by restriction digestion and by DNA sequencing. The manner of construction of pAX1926 is such that the predicted GRG23 translation product contains an amino-terminal extension comprised of MAHHHHHHH.

10 This N-terminal extension comprises a 'histidine tag' or 'six-His tag' that is useful to facilitate purification of the GRG23 protein, as is well known in the art.

Plasmid pAX1926 containing the *grg23* ORF2 has been deposited at the Agricultural Research Service Culture Collection (NRRL) on November 18, 2005, and assigned Accession No. NRRL B-30888.

15

Example 5. *grg23* Confers Resistance to High Levels of Glyphosate

The pUC18-Grg23 construct (pAX1927) was transformed into *E. coli* strain DH5 α and plated out on LB agar plates supplemented with carbenicillin (0.1 mg/mL). Two colonies were selected, resuspended in sterile water, and streaked onto M63 plates

20 containing either 0 mM, 25 mM, 50 mM or 100 mM glyphosate. Isopropyl-B-D-thiogalactopyranoside (IPTG; 0.1 mM) was also added to the plates. As a control, cells containing pUC18 vector alone were transformed and streaked onto glyphosate plates. Following 2 days of growth, these plates were examined for growth (Table 1).

25 Table 1.

Construct	0 mM glyphosate	25 mM glyphosate	50 mM glyphosate	100 mM glyphosate
pUC18	+	-	-	-
pUC18-Grg23 (pAX1927)	+	++	++	++

This result confirms that expression of *grg23* to produce GRG23-ORF2 confers glyphosate resistance in *E. coli* to at least 100 mM. Additionally, the growth of *E. coli*

containing pAX1927 is stronger in the presence of glyphosate than in the absence of glyphosate.

Example 6. Homology of GRG23 with Other Proteins

5 The deduced amino acid sequence of GRG23 has homology to EPSPS enzymes, indicating that *grg23* encodes an EPSPS.

Examination of the deduced amino acid sequence of GRG-ORF2 (SEQ ID NO:4) reveals that it does not contain the four domains typical of Class II EPSPS enzymes. Thus it is a novel, non-Class II, glyphosate-resistant EPSPS.

10 Searching of publicly available protein databases, such as SWISSPROT, reveal that GRG23 has amino acid similarity to the broad class of EPSPS enzymes. However, no protein in any database has greater than 50% identity to the GRG23 amino acid sequence. A representative alignment of GRG23 with other EPSPS enzymes is shown in Figure 1.

15 Example 7. Purification of GRG23

The pRSF1b-*grg23* construct (pAX1926) was expressed in *E. coli* following induction with IPTG, and purified in a single step using a cobalt chromatography column as known in the art. Following column elution, purified GRG23 was dialyzed against 50 mM HEPES/100 mM KCl, pH 7.0. The protein was greater than 95% pure as assessed by
20 PAGE. The amount of GRG23 was quantified using the method of Bradford, as is well known in the art (Bradford (1976) *Anal. Biochem.* 72:248-254).

Example 8. Kinetic Assays of GRG23 activity.

Samples of purified proteins were assayed for EPSPS activity using a kinetic
25 assay involving incubation of PEP (Sigma, St. Louis, MO) and S3P in a buffer containing potassium chloride and HEPES at pH 7.0. Liberation of phosphate was detected using a coupled assay for the fluorescent detection of phosphate based on the generation of Amplex Red, as is known in the art (Vazquez *et al.* (2003) *Anal. Biochem.* 320: 292-298).

The published assay conditions can lead to saturation of the assay in experiments
30 where phosphate is liberated very quickly. This saturation somewhat limits the dynamic range of the assay, and requires a defined range of enzyme concentrations. It was determined that the kinetic limitation of the fluorescent phosphate assay is apparently due to a combination of factors, including a limitation of inosine and PNP. In the present invention, assay conditions have been developed that yield substantially improved

dynamic range and allow the use of a wider range of enzyme and substrate concentrations. The assay conditions that have been significantly changed include the concentrations of purine nucleoside phosphorylase (PNP), xanthine oxidase (XOD), AMPLEX® Red, and inosine, each of which were increased in concentration in the assay

5 to accommodate higher rates of phosphate turnover. This assay was adapted for use to measure EPSPS activity in a 96 well format with the following improvements:

Table 2. An improved fluorescence assay

	Improved Assay	Vazquez <i>et al.</i> , 2003	Units
XOD	1	.4	IU/ml
PNP	2	.02	IU/ml
Inosine	2.25	1.5	mM
HRP	1	1	IU/ml
Amp Red	1,100	50	uM
Hepes	26.25	--	mM
KCl	26.25	--	mM
pH	7	7.4	
Tris	--	50	mM

10 Enzymatic assays were carried out in 96-well plates in a total volume of 50 uL. Reactions were carried out at room temperature at pH 7.0. All assay components except PEP, EPSPS, and S3P were combined into a Master Mix and aliquoted into a 96-well plate using a multi-channel pipettor. Appropriate PEP concentrations were then added to each well. Fresh dilutions of EPSPS were prepared and added to the appropriate wells.

15 Each assay was initiated by the addition of S3P.

Rate data were plotted and the K_m and K_{cat} kinetic parameters determined by use of the application of the Michaelis-Menten equation using a non-linear curve fit program (KALEIDAGRAPH®, Synergy Software). K_i data were determined by measuring the $K_m(app)$ at multiple glyphosate concentrations, and plotting of $K_m(app)$ as a function of

20 inhibitor concentration.

Patent Application No. 02/000,000

Table 3. Effect of Glyphosate on $K_m(\text{app})$ of GRG23

Glyphosate Concentration (μM)	$K_m(\text{app})$
0	10.95
3000	18.89
5000	20.67
1000	25.23

By plotting the $K_m(\text{app})$ as a function of glyphosate concentration, a linear representation
 5 of the glyphosate resistance of GRG23 can be obtained. The X intercept of the resulting
 line represents the $-K_i$. Plotting this line with the data shown in Table 3 yields the
 following data:

Table 4. Kinetic values for GRG23

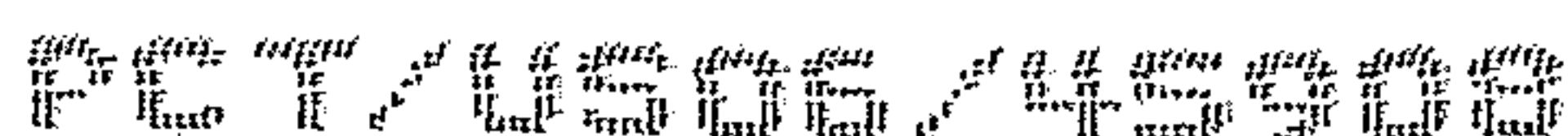
Enzyme	K_m (μM)	K_i (μM)	K_{cat} (sec^{-1})	Ratio of K_i/K_m
GRG23	10.95	9,525	8.2	869

10

GRG23 is highly resistant to glyphosate, with a K_i of over 9mM, and a K_i/K_m ratio over
 800.

15 Example 9. Isolation of ATX21313

For strain ATX 21313, approximately one gram of soil was suspended in 10 ml of
 water, and 100 μl was used to inoculate a 1ml culture of mineral salts medium A
 (MSMA) and no glyphosate. MSMA contains (per 1 liter, pH 7.0) 1 g NH_4Cl , 10 g
 sucrose, 0.2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.01 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.007 g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ supplemented
 20 with phosphates. After an overnight incubation, the culture was plated onto a solid
 medium containing MSMA and 50 mM glyphosate, incubated for a few days, and
 inoculated onto Luria Bertani agar plates to confirm single colony type. Growth in the
 presence of 50 mM glyphosate was reconfirmed by regrowing on MSMA, 50 mM agar
 plates. This isolation method yielded strain ATX21313, which was able to grow well
 25 under these conditions.



Example 10. Cloning of glyphosate-resistant EPSP Synthases

Genomic DNA was extracted from strain ATX21313 and the resulting DNA was partially digested with restriction enzyme *Sau3A* 1 to yield DNA fragments approximately 5 kilobases in size. These DNA molecules were size selected on agarose gels, purified, and ligated into LAMBDA ZAP® vector arms pre-digested with *BamH* I. The ligated arms were then packaged into phage particles, and phage titers were determined as known in the art. The resulting libraries were amplified by methods known in the art to generate a library titer of between 3×10^7 and 3×10^8 PFU/mL. For each independent library, *E. coli* (XL1 Blue MRF') was then co-transfected with phage from an amplified library as well as M13 helper phage into to allow mass excision of the library in the form of an infectious, circular ssDNA as known in the art (Short *et al.* (1988) *Nucleic Acids Research* 16:7583-7600). After centrifugation of the co-infected cells, the phage-containing supernatant was heated to 65-70°C for 15-20 minutes to incapacitate any residual lambda phage particles. Dilutions of the resulting ssDNA plasmid library were transfected into a fresh culture of competent *E. coli* XL1 Blue MRF' cells and also XL-Blue MRF'($\Delta aroA$) cells (XL1 Blue MRF'). The resulting transfected cells were plated onto M63 plates containing kanamycin, 0.1 mM IPTG and either 0 mM, 20 mM or 50 mM glyphosate. This screening method allows identification of clones containing glyphosate-tolerant EPSP synthases, as well as clones carrying tolerance to glyphosate. Colonies growing on 20mM or 50mM glyphosate in the $\Delta aroA$ strain or XL-Blue MRF' were picked and their plasmids analyzed by restriction digest to identify plasmids with shared restriction patterns. Individual plasmids were sequenced by methods known in the art, with preference given to plasmids that conferred resistance to 50 mM glyphosate.

Using this approach, as sometimes modified for each library as known and appreciated in the art, library clones containing EPSP synthase genes were identified.

The sequences of the regions of the resulting clones were determined in the region of the EPSP synthase.

Example 11. DNA and protein sequences of EPSP synthases

The DNA sequence of the glyphosate-resistant EPSP synthase was determined for pAX1967 by methods well known in the art. The DNA sequence of *grg51* is provided herein as SEQ ID NO:5. The predicted translation product of *grg51* (GRG51) is provided

herein as SEQ ID NO:6. GRG51 shows 97% amino acid identity to GRG23 (SEQ ID NO:2).

Plasmid pAX1967 containing *grg51* has been deposited at the Agricultural Research Service Culture Collection (NRRL) on June 26, 2006, and assigned Accession No. NRRL B-30949.

Table 5 summarizes the homology of GRG23 and GRG51 with other EPSP synthase enzymes.

Table 5. Amino acid identity of GRG23-ORF1 and GRG51 to representative EPSPS enzymes

EPSPS	% Identity to GRG23	% Identity to GRG51
GRG23	-	92%
GRG51	92%	-
<i>B. Clausii</i>	36%	35%
<i>R. xylanophilus</i>	39%	38%
<i>E. coli</i>	32%	32%
CP4	20%	21%
<i>Zea maize</i>	32%	32%

Example 12. Cloning of novel Glyphosate-resistant EPSP synthases into an *E. coli* Expression vector

The *grg51* gene contained in pAX1967 was sub-cloned into the *E. coli* expression vector pRSF1b (Invitrogen). Resulting clones were confirmed by DNA sequencing, and used to induce expression of *grg51* in *E. coli*. The expressed His-tagged protein was then purified as known in the art.

Example 13. Glyphosate resistance of EPSP synthases

Cells containing pAX1967 were plated onto M63+ plates containing antibiotic and either 0mM or 20mM glyphosate. Growth was scored after two days growth at 37°C. GRG51 was observed to confer resistance to 20 mM glyphosate in *E. coli* cells (Table 6).

Table 6. Glyphosate screen

EPSPS	Plasmid Clone	Growth on 20mM glyphosate
Vector	-	-
GRG51	pAX1967	++

Example 14. *syngrg23* design and expression

5 A novel gene sequence encoding the GRG23 protein (SEQ ID NO:2; U.S. Patent Application No. 60/741,166 filed December 1, 2005) was designed and synthesized. This sequence is provided as SEQ ID NO:12. This open reading frame, designated “*syngrg23*” herein, was cloned into the expression vector pRSF1b (Invitrogen), by methods known in the art.

10 The *syngrg23* gene encoding GRG23 was cloned into a pUC19 vector to create pAX748. PCR primers that flanked *syngrg23* in this vector were used to amplify *syngrg23* from pAX748 using the MUTAZYME® II system (Stratagene) to introduce random mutations into the *syngrg23* coding region. The template was diluted 1:50 in the error-prone PCR reaction, and amplification was carried out for 30 cycles. The resulting
15 PCR product was digested with the restriction enzymes *Bam*H I and *Sgs* I, gel-purified, and ligated into the vector pRSF1b to create a mutagenized *syngrg23* library.

The mutagenized *syngrg23* libraries were transformed into *E. coli* strain BL21*DE3 star (Invitrogen). Following transformation, individual colonies were plated on 1x M63 medium containing 150 mM glyphosate to select for clones that had retained
20 enzymatic activity and growth tolerance.

Example 15. Screening for glyphosate resistance on plates

Library ligations were transformed into BL21*DE3 competent *E. coli* cells (Invitrogen). The transformations were performed according to the manufacturer’s
25 instructions with the following modifications. After incubation for 1 hour at 37°C in SOC medium, the cells were sedimented by centrifugation (5 minutes, 1000 x g, 4°C). The cells were washed with 1 ml M63+, centrifuged again, and the supernatant decanted. The cells were washed a second time with 1 ml M63+ and resuspended in 200 ul M63+.

For selection of mutant GRG23 enzymes conferring glyphosate resistance in *E. coli*, the cells were plated onto M63+ agar medium plates containing 150 mM glyphosate,
30 0.05 mM IPTG (isopropyl-beta-D-thiogalactopyranoside), and 50 ug/ml kanamycin.

M63+ medium contains 100 mM KH₂PO₄, 15 mM (NH₄)₂SO₄, 50 μM CaCl₂, 1 μM FeSO₄, 50 μM MgCl₂, 55 mM glucose, 25 mg/liter L-proline, 10 mg/liter thiamine HCl, sufficient NaOH to adjust the pH to 7.0, and 15 g/liter agar. The plates were incubated for 36 hours at 37°C.

- 5 Individual colonies were picked and arrayed into 384-well plates. Two 384-well plates were created in this manner. A third plate of 384 clones was picked from colonies that grown on plates lacking glyphosate.

Example 16. Isolation and analysis of glyphosate resistant GRG23 variants

- 10 BL21*DE3 cells transformed with mutagenized *syngrg23* and/ or *grg23* variants were identified by growth on glyphosate plates. Extracts of mutagenized *syngrg23* and *grgr23* variants were prepared and assayed for improved enzymatic activity. Colonies identified on glyphosate plates pinned into 96-well blocks containing LB medium and were grown to an O.D. of about 0.6. IPTG was then added (0.5 mM) and the blocks were
15 incubated overnight at 20°C to induce protein expression. Protein extracts were prepared from the cell pellets using POP culture reagent (Novagen) and Lysonase (Novagen), and the enzymatic activity in the crude lysates was measured after heating the extracts for 30 min at 37°C. Extracts with activity greater than two standard deviations above the mean of a set of extracts containing the appropriate control protein (for example GRG23) were
20 selected for further analysis.

Clones showing increased activity after incubation as crude extracts were grown in 250 mL LB cultures, and protein expression induced with IPTG. Following induction, the mutant GRG23 protein was purified from each culture by affinity chromatography using a cobalt resin (Novagen). The purified proteins were then tested for enzymatic
25 activity following heating for 0, 2, 4, and approximately 16 hours at 37°C.

Example 17. Improved GRG23 variants.

- From a DNA library of mutagenized *syngrg23*, several clones with improved activity were identified. The DNA sequences of the clones corresponding to these
30 extracts was determined. Table 7 shows the amino acid changes identified in six variants of GRG23 that retained glyphosate resistance: *grg23(L3P1.B20)* (SEQ ID NO:26) encoding the amino acid sequence GRG23(L3P1.B20) (SEQ ID NO:27); *grg23(L3P1.B3)* (SEQ ID NO:28) encoding the amino acid sequence GRG23(L3P1.B3)

GRG23(L3P1.F18) (SEQ ID NO:29);

grg23(L3P1.F18) (SEQ ID NO:30) encoding the amino acid sequence GRG23(L3P1.F18) (SEQ ID NO:31); and *grg23(L3P1.O23)* SEQ ID NO:31, encoding the amino acid sequence GRG23(L3P1.O23) SEQ ID NO:32).

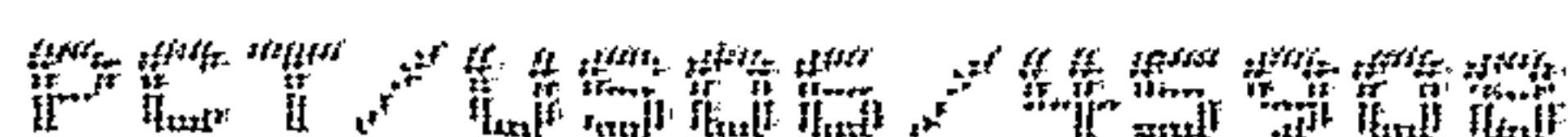


Table 7. Mutations identified in glyphosate-resistant GRG23 variants

Clone	Amino Acid (AA) in GRG23
L3P1B20	V206→I
L3P1B3	D75→H, E217→K
L3P1F18	T274→I
L3P1O23	R5→H

The clones were grown in 250 mL LB cultures, and protein expression induced
 5 and isolated as described above. The purified proteins were then tested for enzymatic
 activity following heating for 0, 2 and 4, and approximately 16 hours at 37°C. One of the
 clones, termed "M5", was found to retain an increased proportion of its enzymatic
 activity after prolonged incubation at 37°C (Table 8). The DNA sequence of this clone
 was determined, and the gene is designated herein as *grg23(ace1)* (SEQ ID NO:14). The
 10 protein expressed from *grg23(ace1)* is designated GRG23(ACE1) (SEQ ID NO:15).

Table 8. Half-life of GRG23(ACE1) vs GRG23 at elevated temperature

Protein	Half-life at 37°C (hours)
GRG23	7
GRG23(ACE1)	15.5

15 GRG23(ACE1) contains 2 amino acid substitutions relative to wild-type GRG23
 protein: **A49→T** and **S276→T**. The pRSF1b vector that contains this gene is designated
 pAX3801. Figure 1 shows the relative stability of GRG23 (ACE1) vs GRG23 at elevated
 temperatures.

20 Example 18. Determination of EPSPS activity of GRG-23 variants

Extracts containing GRG23 variant proteins were assayed for EPSP synthase
 activity as described in U.S. Patent Application No. 60/741,166, filed December 1, 2005,
 herein incorporated by reference in its entirety. Assays were carried out in a final volume
 of 50 ul containing 0.5 mM shikimate-3-phosphate, 200 uM phosphoenolpyruvate (PEP),
 25 1 U/ml xanthine oxidase, 2 U/ml nucleoside phosphorylase, 2.25 mM inosine, 1 U/ml
 horseradish peroxidase, 0-2 mM glyphosate, 50 mM HEPES/KOH pH 7.0, 100 mM KCl,

and AMPLEX® Red (Invitrogen) according to the manufacturer's instructions. Extracts were incubated with all assay components except shikimate-3-phosphate for 5 minutes at room temperature, and assays were started by adding shikimate-3-phosphate. EPSP synthase activity was measured using a Spectramax Gemini XPS fluorescence spectrometer (Molecular Dynamics, excitation: 555 nm; emission: 590 nm).

Following full determination of kinetic parameters were performed on purified protein as previously described (U.S. Patent Application Number 60/ 741,166 filed December 1, 2005), adjusting for the quantity of protein determined by Bradford assay as known in the art. For any one glyphosate concentration, EPSP synthase activity was measured as a function of a broad range of PEP concentrations. The data were fit to the Michaelis-Menten equation using KALEIDAGRAPH® software (Synergy Software) and used to determine the K_m (K_m apparent) of the EPSP synthase at that glyphosate concentration. K_m apparent values were determined at no fewer than 4 glyphosate concentrations, and the K_i of the EPSPS for glyphosate were calculated from the plot of K_m apparent vs. glyphosate concentration, using the equation ($m1 \cdot x / (m2 + x)$; $m1 = 1$; $m2 = 1$) as known in the art.

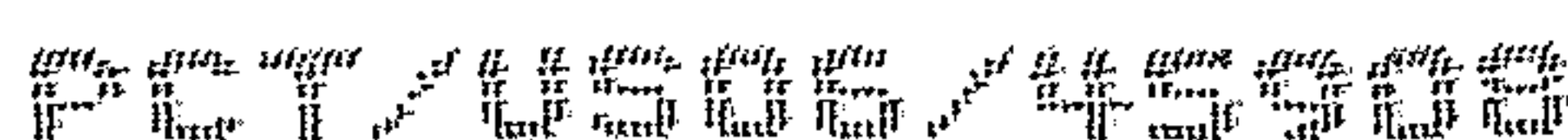
Table 9. Kinetics of GRG23(ACE1) vs GRG23

	K_m (μM)	K_i (μM)	V_{max} (nmol/min/ μg)
GRG23	12.2	13,800	14.77
GRG23(ACE1)	9.7	14,620	13.73

Example 19. Identification of *grg23* (*ace2*).

GRG23(ACE1) contains two amino acid changes relative to GRG23. To determine if additional substitutions at these positions could further improve activity, a DNA library was generated that resulted in clones expressing proteins that were substantially mutated and both positions 49 and 276 of GRG23. Clones conferring glyphosate resistance were selected by growth on glyphosate plates, and grown and assayed for kinetic properties as described.

Surprisingly, one clone, herein designated *grg23*(*ace2*) (SEQ ID NO: 16), encoding the GRG23(ACE2) protein (SEQ ID NO: 17) was identified as having



improved thermostability. The DNA sequence of *grg23(ace2)* shows that GRG23(ACE2) contains a single amino acid change (residue 276 of GRG23 to arginine).

Example 20. Comparison of GRG23 and GRG51, and mutagenesis of differing residues.

Two libraries were generated to assess the permutations of amino acid sequences possible from comparison of the amino acid sequences of GRG23 and GRG51. The first library introduced variation from the GRG51 amino acid sequence into a *grg23(ace2)* coding region. The second library introduced the variation from GRG23(ACE2) amino acid sequence into the *grg51* coding region.

Clones of the resulting libraries were assessed for (1) ability to confer glyphosate resistance upon on a cell, and (2) activity after prolonged incubation at 37°C. A total of ten clones were sequenced and analyzed in more detail. One particular clone, herein designated *grg51.4* (SEQ ID NO:18), encoding the protein GRG51.4 (SEQ ID NO:19), contains several amino acid changes relative to both GRG23(ACE2) and GRG51. The amino acid changes present in GRG51.4 relative to GRG23(ACE2) were subsequently introduced into the *grg23(ace2)* gene, to yield *grg23(ace3)* (SEQ ID NO:20), which encodes the GRG23(ACE3) protein (SEQ ID NO:21). GRG23(ACE3) exhibits superior activity and thermostability relative to GRG23, and GRG23(ACE2).

GRG23(ace1) was mutagenized, and clones were tested to identify clones expressing variants with improved thermostability and/or activity. One clone, *grg23(L5P2.J2)* (SEQ ID NO:22), encoding GRG23(L5P2.J2) (SEQ ID NO:23), was identified by virtue of its improved kinetic properties. GRG23(L5P2.J2) contains three amino acid changes relative to GRG23 (ACE1), as shown in the following Table 10.

Table 10. Amino Acid changes in GRG23(L5P2.J2)

	Amino Acid (AA) in GRG23(L5P2.J2) relative to <u>GRG23(ACE1)</u>
	V101→F
	A213→S
	D284→N

Oligonucleotide mutagenesis was used to create clones that contain each of the amino acid changes identified GRG23(L5P2.J2) into the *grg23(ace3)* coding region. A

clone was identified as encoding a protein having improved kinetic properties over GRG23(ACE3), and designated *grg23(ace4)* (SEQ ID NO:24). The protein encoded by *grg23(ace4)* is designated as GRG23(ACE4) (SEQ ID NO:25) contains a single amino acid change relative to GRG23(ACE3) (Valine 101 to phenylalanine). Based on this result, a separate oligonucleotide mutagenesis was performed to test the kinetics of each possible amino acid substitutions at position 101. None of the amino acid changes resulted in further improvement in kinetic properties compared to GRG23(ACE4).

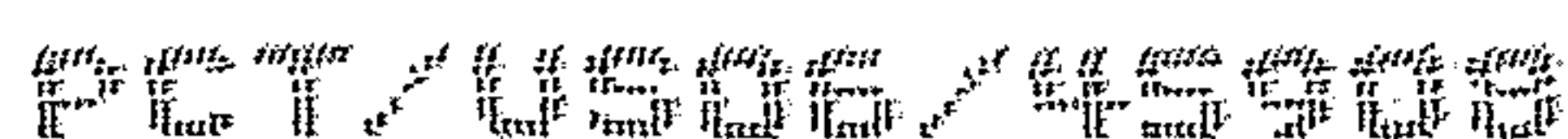
Table 11. Kinetics of improved variants

	Km (uM)	Ki (uM)	Vmax (nmol/min/ug)
GRG23	14	10,800	13
GRG51	15	21,048	13
GRG23(ACE1)	10	14,620	14
GRG23(ACE2)	11	18,104	15
GRG51.4	19	26,610	17
GRG23(ACE3)	15	20,000	17
GRG23(L5P2.J2)	15	2,500	23
GRG23(ACE4)	14	5,010	24

Example 21. Engineering *grg23* or *grg51* for Plant Transformation

The *grg23* or *grg51* open reading frame (ORF) is amplified by PCR from a full-length cDNA template. *Hind* III restriction sites are added to each end of the ORF during PCR. Additionally, the nucleotide sequence ACC is added immediately 5' to the start codon of the gene to increase translational efficiency (Kozak (1987) *Nucleic Acids Research* 15:8125-8148; Joshi (1987) *Nucleic Acids Research* 15:6643-6653). The PCR product is cloned and sequenced using techniques well known in the art to ensure that no mutations are introduced during PCR.

The plasmid containing the *grg23* or *grg51* PCR product is digested with *Hind* III and the fragment containing the intact ORF is isolated. This fragment is cloned into the *Hind* III site of plasmid pAX200, a plant expression vector containing the rice actin promoter (McElroy *et al.* (1991) *Molec. Gen. Genet.* 231:150-160) and the PinII terminator (An *et al.* (1989) *The Plant Cell* 1:115-122). The promoter – gene – terminator fragment from this intermediate plasmid is subcloned into plasmid pSB11 (Japan Tobacco, Inc.) to form a final plasmid, for example, pSB11GRG23. pSB11GRG23 is organized such that the 3.91 kb DNA fragment containing the promoter



– *grg23* – terminator construct may be excised by double digestion with *Kpn* I and *Pme* I and used for transformation into plants by aerosol beam injection. The structure of pSB11GRG23 is verified by restriction digest and gel electrophoresis, and by sequencing across the various cloning junctions.

5 The plasmid is mobilized into *Agrobacterium tumefaciens* strain LBA4404 which also harbors the plasmid pSB1 (Japan Tobacco, Inc.), using triparental mating procedures well known in the art, and plating on media containing spectinomycin. Plasmid pSB11GRG23 carries spectinomycin resistance but is a narrow host range plasmid and cannot replicate in *Agrobacterium*. Spectinomycin resistant colonies arise when
10 pSB11GRG23 integrates into the broad host range plasmid pSB1 through homologous recombination. The cointegrate product of pSB1 and pSB11GRG23 is verified by Southern hybridization. The *Agrobacterium* strain harboring the cointegrate is used to transform maize by the PureIntro method (Japan Tobacco).

15 Example 22. Transformation of *grg23* or *grg51* into Plant Cells

Maize ears are collected 8-12 days after pollination. Embryos are isolated from the ears, and those embryos 0.8-1.5 mm in size are used for transformation. Embryos are plated scutellum side-up on a suitable incubation media, such as DN62A5S media (3.98 g/L N6 Salts; 1 mL/L (of 1000x Stock) N6 Vitamins; 800 mg/L L-Asparagine; 100 mg/L
20 Myo-inositol; 1.4 g/L L-Proline; 100 mg/L Casamino acids; 50 g/L sucrose; 1 mL/L (of 1 mg/mL Stock) 2,4-D). However, media and salts other than DN62A5S are suitable and are known in the art. Embryos are incubated overnight at 25°C in the dark.

The resulting explants are transferred to mesh squares (30-40 per plate), transferred onto osmotic media for 30-45 minutes, then transferred to a beaming plate
25 (see, for example, PCT Publication No. WO/0138514 and U.S. Patent No. 5,240,842).

DNA constructs designed to express GRG23 in plant cells are accelerated into plant tissue using an aerosol beam accelerator, using conditions essentially as described in PCT Publication No. WO/0138514. After beaming, embryos are incubated for 30 min on osmotic media, and placed onto incubation media overnight at 25°C in the dark. To
30 avoid unduly damaging beamed explants, they are incubated for at least 24 hours prior to transfer to recovery media. Embryos are then spread onto recovery period media for 5 days at 25°C in the dark, then transferred to a selection media. Explants are incubated in selection media for up to eight weeks, depending on the nature and characteristics of the particular selection utilized. After the selection period, the resulting callus is transferred

to embryo maturation media until the formation of mature somatic embryos is observed.

The resulting mature somatic embryos are then placed under low light and the process of regeneration is initiated by methods known in the art. The resulting shoots are allowed to root on rooting media, and the resulting plants are transferred to nursery pots and propagated as transgenic plants.

Materials **DN62A5S Media**

Components	per liter	Source
Chu's N6 Basal Salt Mixture (Prod. No. C 416)	3.98 g/L	Phytotechnology Labs
Chu's N6 Vitamin Solution (Prod. No. C 149)	1 mL/L (of 1000x Stock)	Phytotechnology Labs
L-Asparagine	800 mg/L	Phytotechnology Labs
Myo-inositol	100 mg/L	Sigma
L-Proline	1.4 g/L	Phytotechnology Labs
Casamino acids	100 mg/L	Fisher Scientific
Sucrose	50 g/L	Phytotechnology Labs
2,4-D (Prod. No. D-7299)	1 mL/L (of 1 mg/mL Stock)	Sigma

- Adjust the pH of the solution to pH 5.8 with 1N KOH/1N KCl, add Gelrite (Sigma) to 3g/L, and autoclave. After cooling to 50°C, add 2 ml/L of a 5 mg/ml stock solution of Silver Nitrate (Phytotechnology Labs). Recipe yields about 20 plates.

Example 23. Transformation of *grg23* or *grg51* into Maize Plant Cells by

15 *Agrobacterium*-Mediated Transformation

Ears are collected 8-12 days after pollination. Embryos are isolated from the ears, and those embryos 0.8-1.5 mm in size are used for transformation. Embryos are plated scutellum side-up on a suitable incubation media, and incubated overnight at 25°C in the dark. However, it is not necessary *per se* to incubate the embryos overnight. Embryos are

contacted with an *Agrobacterium* strain containing the appropriate vectors for Ti plasmid mediated transfer for 5-10 min, and then plated onto co-cultivation media for 3 days (25°C in the dark). After co-cultivation, explants are transferred to recovery period media for five days (at 25°C in the dark). Explants are incubated in selection media for up to 5 eight weeks, depending on the nature and characteristics of the particular selection utilized. After the selection period, the resulting callus is transferred to embryo maturation media, until the formation of mature somatic embryos is observed. The resulting mature somatic embryos are then placed under low light, and the process of regeneration is initiated as known in the art. The resulting shoots are allowed to root on 10 rooting media, and the resulting plants are transferred to nursery pots and propagated as transgenic plants.

All publications and patent applications mentioned in the specification are indicative of the level of skill of those skilled in the art to which this invention pertains. 15 All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that 20 certain changes and modifications may be practiced within the scope of the appended claims.

DEMANDES OU BREVETS VOLUMINEUX

**LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS
COMPREND PLUS D'UN TOME.**

CECI EST LE TOME __1__ DE __2__

NOTE: Pour les tomes additionels, veuillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

**THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE
THAN ONE VOLUME.**

THIS IS VOLUME __1__ OF __2__

NOTE: For additional volumes please contact the Canadian Patent Office.

P E T R O P O L I S

THAT WHICH IS CLAIMED:

1. An isolated nucleic acid molecule selected from the group consisting of:
- 5 a) a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, or 5, or a complement thereof;
- b) a nucleic acid molecule comprising a nucleotide sequence having at least 80% sequence identity to the nucleotide sequence of SEQ ID NO:1, 3, or 5, or a complement thereof;
- 10 c) the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession No. NRRL B-30888 or B-30949, or a complement thereof;
- d) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 4, or 6; and,
- 15 e) a nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence of SEQ ID NO:2, 4, or 6.
2. The isolated nucleic acid molecule of claim 1, wherein said nucleotide
- 20 sequence is a synthetic sequence that has been designed for expression in a plant.
3. A vector comprising the nucleic acid molecule of claim 1.
4. The vector of claim 3, further comprising a nucleic acid molecule
- 25 encoding a heterologous polypeptide.
5. A host cell that contains the vector of claim 3.
6. The host cell of claim 5 that is a bacterial host cell.
- 30 7. The host cell of claim 5 that is a plant cell.
8. A transgenic plant comprising the host cell of claim 7.

9. The plant of claim 8, wherein said plant is selected from the group

consisting of maize, sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, and oilseed rape.

5 10. A transformed seed of the plant of claim 8.

11. An isolated polypeptide selected from the group consisting of:

- a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 4, or 6;
- 10 b) a polypeptide encoded by the nucleotide sequence of SEQ ID NO:1, 3, or 5;
- c) a polypeptide comprising an amino acid sequence having at least 80% sequence identity to the amino acid sequence of SEQ ID NO:2, 4, or 6, wherein said polypeptide has herbicide resistance activity;
- 15 d) a polypeptide that is encoded by a nucleotide sequence that is at least 80% identical to the nucleotide sequence of SEQ ID NO:1, 3, or 5, wherein said polypeptide has herbicide resistance activity; and,
- e) a polypeptide that is encoded by the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession No. NRRL B-30888 or
- 20 NRRL B-30949.

12. The polypeptide of claim 11 further comprising a heterologous amino acid sequence.

25 13. A method for producing a polypeptide with herbicide resistance activity, comprising culturing the host cell of claim 5 under conditions in which a nucleic acid molecule encoding the polypeptide is expressed, said polypeptide being selected from the group consisting of:

- a) a polypeptide comprising the amino acid sequence of SEQ ID
- 30 NO:2, 4, or 6;
- b) a polypeptide encoded by the nucleic acid sequence of SEQ ID NO:1, 3, or 5;

c) a polypeptide comprising an amino acid sequence having at least 80% sequence identity to a polypeptide with the amino acid sequence of SEQ ID NO:2, 4, or 6, wherein said polypeptide has herbicide resistance activity;

d) a polypeptide encoded by a nucleic acid molecule comprising a nucleotide sequence having at least 80% sequence identity to the nucleic acid sequence of SEQ ID NO:1, 3, or 5, wherein said polypeptide has herbicide resistance activity; and,

e) a polypeptide that is encoded by the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession No. NRRL B-30888 or NRRL B-30949.

10

14. A method for conferring resistance to an herbicide in a plant, said method comprising transforming said plant with a DNA construct, said construct comprising a promoter that drives expression in a plant cell operably linked with a nucleotide sequence at least 80% identical to the nucleotide sequence of SEQ ID NO:1, 3, or 5, and

15 regenerating a transformed plant.

15. The method of claim 14, wherein said herbicide is glyphosate.

16. A plant having stably incorporated into its genome a DNA construct comprising a nucleotide sequence that encodes a protein having herbicide resistance activity, wherein said nucleotide sequence is selected from the group consisting of:

a) a nucleotide sequence of SEQ ID NO:1, 3, or 5;
b) a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence of SEQ ID NO:1, 3, or 5, wherein said nucleotide sequence encodes a polypeptide having herbicide resistance activity;

c) a nucleotide sequence encoding a polypeptide comprising an amino acid sequence of SEQ ID NO:2, 4, or 6;

d) a nucleotide sequence encoding a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence of SEQ ID NO:2, 4, or 6, wherein said polypeptide has herbicide resistance activity; and,

e) the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession No. NRRL B-30888 or NRRL B-30949; wherein said nucleotide sequence is operably linked to a promoter that drives expression of a coding sequence in a plant cell.

47

24. The method of claim 23, wherein said enzyme is EPSPS.

GRG23	---MALERGQHGRSRRLFGASLERITMETDR-----LVIPGSKSITNRALLLAAAAGT	51
GRG51	-----METDR-----LVIPGSKSITNRALLLAAAAGT	28
B_Clausii	---MVQFDSQARSPTPLAGVERLRLTPSQKRINATLEVPGSKSATNRALLLAAVASGT	56
R_xylanophilus	MSGVSGVPGVDFGIEEVRGSPFEEVAPLERPPDATVRLPGSKSITNRALLVAAALAGGT	60
E_coli	-----MESLTLQPIARVDG-TINLPGSKTVSNRALLLAAALAHGK	38
CP4	-----MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGE	44
Zea_maize	-----AGAEIIVLQPIKEISG-TVKLPGSKSLNRILLLAAALSEG	40
	: : ** * : : * : : : : *	
GRG23	SVLVRPLVSADTSAFKTAIQAL--GANVSADGDNWVVEGLGQAPHLAD-----IWCEDA	104
GRG51	SVLVRPLVSADTSAFKTAIQAL--GANVSADGDDWVVEGLGQAPNLDAD-----IWCEDA	81
B_Clausii	STLRNALKSDDTYWCIEALKKT--GVEIAVDGSNVTVYGRGGVFHSGS-----LYIGSA	108
R_xylanophilus	SRIENPLLADDDPFWLNMNALVGLGFGVRVGEAGAVEAGGGGCI PAPSAD-----VFVGNA	115
E_coli	TVLTNLLDSDDVRHMLNALTALGVSYTLSADTRRCEIIGNGGPLHAEGA--LLELFLGNA	95
CP4	TRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAPEAP-----LDFGNA	99
Zea_maize	TVVDNLLNSEDVHYMLGALRTLGLSVEADKAAKRAVVVCGCGKFPVEDAKEEVQLFLGNA	100
	: : * . * * : : * . *	
GRG23	GTVARFLPPFVAAG--QGKFTVDGSEQLRRRPLRPLVDGIRHLGARVSSE-----QLPLTI	158
GRG51	GTVARFLPPFVAAG--QGKFTVDGSEQLRRRPLRPPVVDGIRHLGARVSSE-----QLPLTI	135
B_Clausii	GTAGRFLPGMLAAA--TGNWHVEASHSMNKRPIAPLVKTLQALGANIQYGSRRGHYPLSI	166
R_xylanophilus	GTVARFLPPALALG--SGPYRVVDGTFRMRERPVAELVEALRALGARVECEEREHGLPLVV	173
E_coli	GTAMRPLAAALCLG--SNDIVLTGEPRMKERPIGHLVDALRLGGAKITYLEQENYPPRL	153
CP4	ATGCRLLTMGLVGVI--DFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDG-DRLPVTI	156
Zea_maize	GTAMRPLTAAVTAAGGNATVLDGVPRMRERPIGDLVVGLKQLGADVDCFLGTDCCPPVRV	160
	. * * : : . . : . ** : : : * : *	

FIG. 1A

GRG23	EA-SGLAGGEYEIEAHQSSQFASGLIMAAPYARQGLRVRIPN-PVSQPYLTMTLRMMRDF	216
GRG51	EA-SGLAGGEYEIEAHQSSQFASGLIMAAPYARQGLRVRIPN-PVSQPYLTMTLRMMRDF	193
B_Clausii	SG-EGLNGGKVNMSGQLSSQFISGCLLAAPLAKNPVSIIVKDGIVQQAYVRITIDLMAAF	225
R_xylanophilus	RG-GARGGGEISVSGERSSQFLSGLLISAPCLPGGLTVRPRGALVSRPYVDITVRVMRSF	232
E_coli	Q--GGFTGGNVDVDGSSVSSQFLTALLMTAPLAPEDTVIRIKGDLVSKPYIDITLNLMKTF	211
CP4	RG--PKTPTPTTYRVPMASAQVKSALLAGLNTPGITTVIEPIIMTRDHTKMLQGFGANL	214
Zea_maize	NGIGGLPGGKVKLSGSISSQYLSALLMAAPLALGDVEIEIIDKLISIPYVEMTLRLMERF	220
	:* . . . : *	
GRG23	GLET--STDGATVSVPPGR-YTAR-RYEIEPDASTASYFAAASAVSGRSFEFQGLGTDSI	272
GRG51	GIET--STDGATVSVPPGR-YTAR-RYEIEPDASTASYFAAASAVSGRRFEFQGLGTDSI	249
B_Clausii	GVEVKAAPDWSLLEVNPS-P-YVAN-DIAIEADASTACYFLALAAITAGKIRIRHFSTKTS	283
R_xylanophilus	GASVEEESGAAFRVAPGA-YRAT-AYRVEPDASAASYFLAAAALTAGRVVIPGLGRSSL	290
E_coli	GVEIEN-QHYQQFVVKGGQSYQSPGTYLVEGDASSASYFLAAAIAKGGTVKVTGIGRNSM	270
CP4	TVETDADGVRTIRLEGRGK--LTGQVIDVPGDPSSTAFLPLVAALLVPG-SDVTILNVLMN	271
Zea_maize	GVKAEHSDSWDRFYIKGGQKYKSPKNAYVEGDASSASYFLAGAAITGGTIVTEGCGTTSL	280
	. . . : * . . . : *	
GRG23	QGDTSFFNVLGRLGAEVHWAPNSVTISGP-----ERLNGDIEVDMG---EISDTFMTL	322
GRG51	QGDTSFFNVLGRLGAEVHWASNSVTISGP-----ERLTGDIIEVDMG---EISDTFMTL	299
B_Clausii	QPDILFVSILKRMGCNFEIGPSFVEGEGP-----TRLRGGFTVNMN---ELSDQALT	333
R_xylanophilus	QGDVAFAGILRRMGCRVSLSEDRIELAGP-----PRLRG-VEADMN---AISDTMMTL	339
E_coli	QGDIRFADVLEKMGATICWGDDYISCT-----RGELNAIDMDMN---HIPDAAMTI	318
CP4	PTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPIL	331
Zea_maize	QGDVKFAEVLMMGAKVTWTETSVTVTGPPREP-FGRKHLKAIDVNMN---KMPDVAMTL	336
	: * : * . . . : *	

FIG. 1B

GRG23	AAIAPLADGPITITNIGHARLKESDRISAMETNLRTLGVQTDVGHDMRIYP-----S	375
GRG51	AAIAPLADGPITITNIGHARLKESDRISAMESNLRMLGVQTDVGHDMRIYP-----S	352
B_Clausii	AAISPFADGPIAIEGVGHIRHHECDRIRAICTELSRLGIRVEERHDGLTVYP-----G	386
R_xylanophilus	AAIAPFASSPTLIKNVAHTRLQETDRLAAVAAEELSRLGVRVHETPDSLRIIP-----G	392
E_coli	ATAALFAKGTTRLRNIYNWRVKETDRRLFAMATELRKVGAEVEEGHDYIRITPP-----E	372
CP4	AVAAFAEGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRGRPDGKGLG	391
Zea_maize	AVVALFADGPTAIRDVASWRVKETERMVAIRTELTKLGASVEEGPDYCIITPP-----E	390
	*. : : * . . . : : * : * : * . . . :	
GRG23	TPHGGRVNCHRDHRIAMAFSILGLR-VDGITLDDPQCVCVKTFPGFFDYLGRLFPEKALT	434
GRG51	TPHGGRVNCHRDHRIAMAFSILGLR-VDGITLDDPQCVCVKTFPGFFDYLGRLFPEKALT	411
B_Clausii	QPKPTVNTYDDHRMAMALALIGAK-VDGIELDDPGCVAKTcpsyFSMLAQTGIGVKAVS	445
R_xylanophilus	KVRPAAIRTYGDHRMAMAFSLVGLR-VRGVRILDPGCVTKTLPgyFRLLLEGLRRGG----	447
E_coli	KLNFAEIAATYNDHRMAMCFSLVALS-DTPVTIILDPKCTAKTFPDYFEQLARISQAA----	427
CP4	NASGAAVATHLDHRIAMSFLVMGLVSENpVTVDATMIATSFPEFMDLMAGLGAKIELSD	451
Zea_maize	KLNVTAIDTYDDHRMAMAFSLAACAEVpVTIIRDpGCTRKTFPDYFDVLSTFVKN-----	444
	: : ***:**. : . : * . . . : * : :	
GRG23	PG--	436
GRG51	PG--	413
B_Clausii	P---	446
R_xylanophilus	----	
E_coli	----	
CP4	TKAA	455
Zea_maize	----	

FIG. 1C

4/5

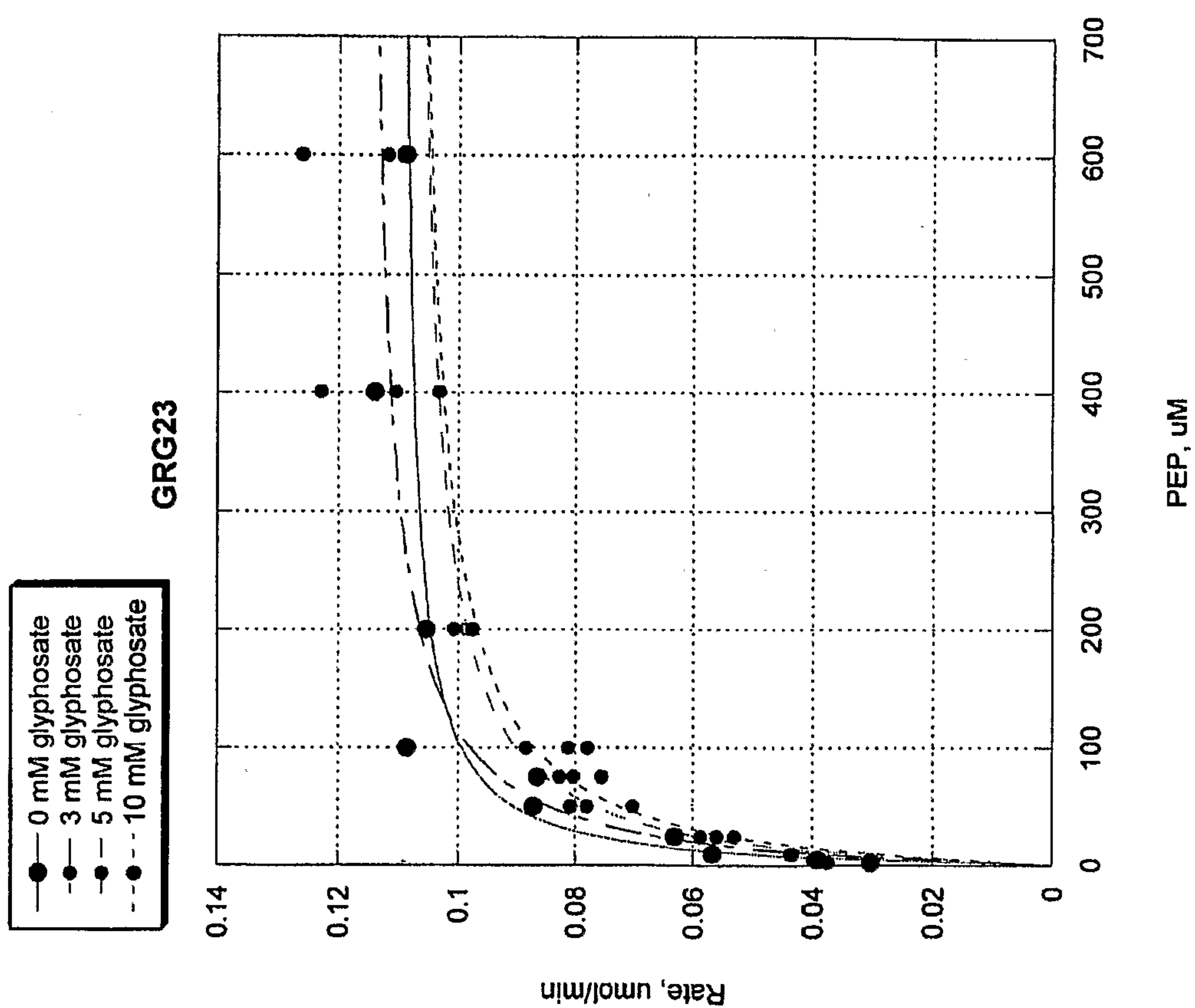


FIG. 2

5/5

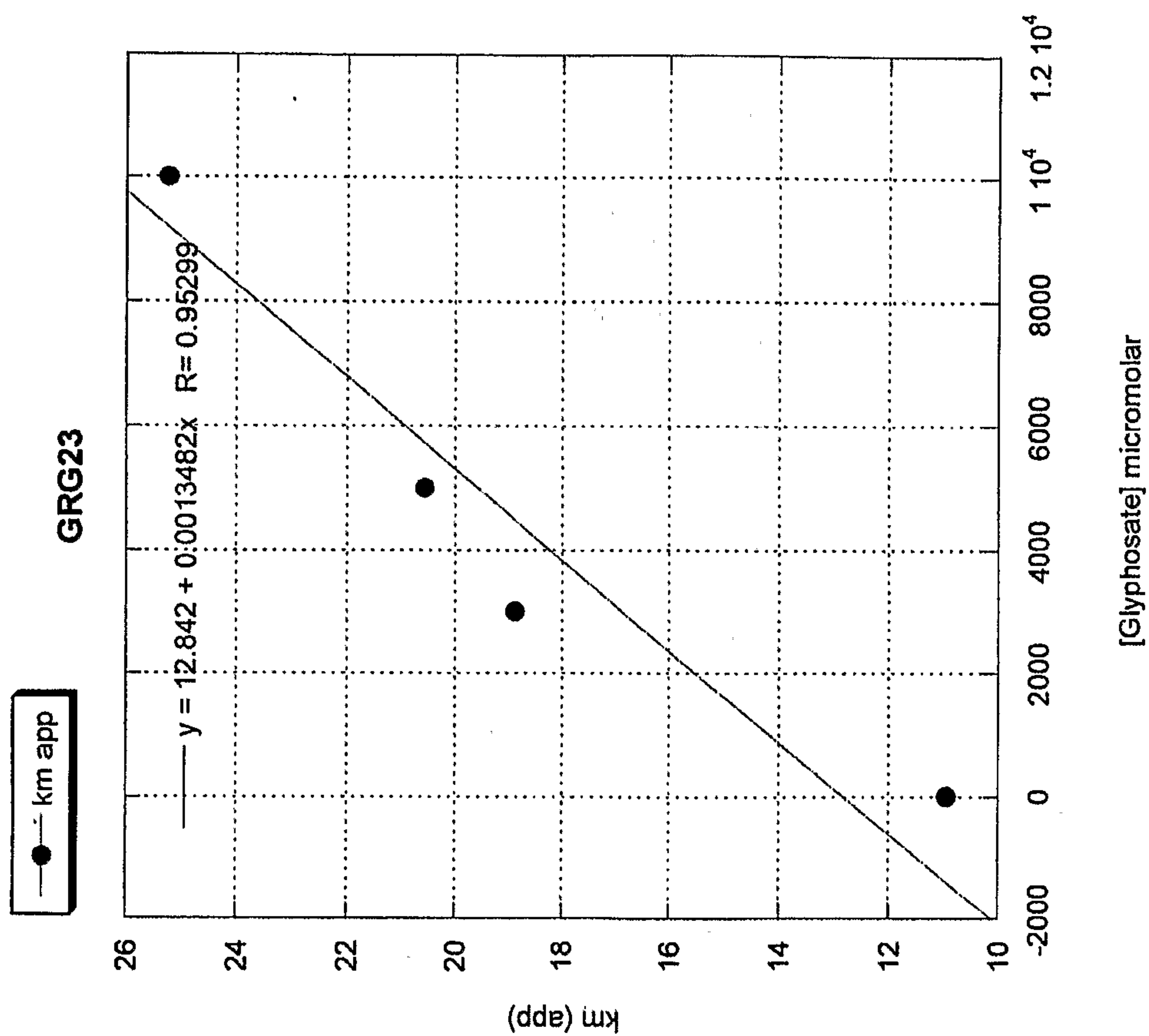
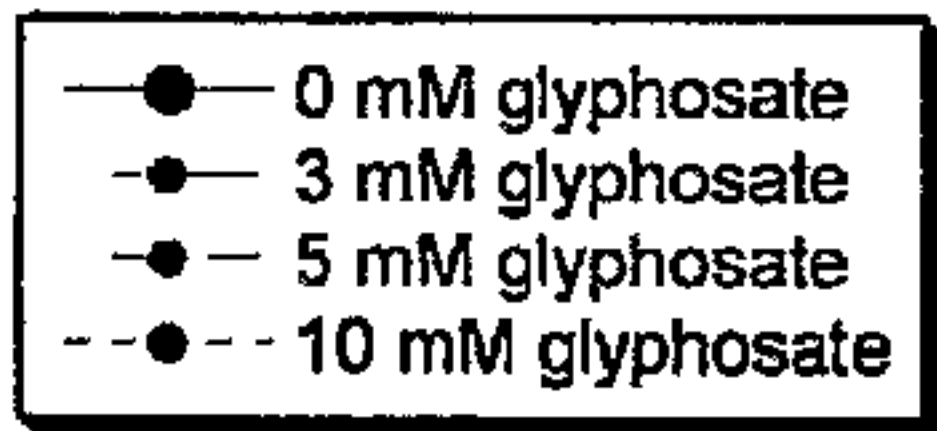


FIG. 3



GRG23

