Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2654812 A1 2008/01/03

(21) 2 654 812

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2007/06/27
(87) Date publication PCT/PCT Publication Date: 2008/01/03
(85) Entrée phase nationale/National Entry: 2008/12/09
(86) N° demande PCT/PCT Application No.: US 2007/072206
(87) N° publication PCT/PCT Publication No.: 2008/002964
(30) Priorités/Priorities: 2006/06/27 (US60/816,676); 2006/07/07 (US60/819,122); 2006/07/07 (US60/819,119)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01), *A01H 1/00* (2006.01), *C12Q 1/68* (2006.01)

(71) Demandeur/Applicant:
ATHENIX CORPORATION, US

(72) Inventeurs/Inventors:
PETERS, CHERYL L., US;
VANDE BERG, BRIAN, US;
CARR, BRIAN, US;
TOMSO, DANIEL JOHN, US

(74) Agent: SMART & BIGGAR

(54) Titre : GRG33, GRG37, GRG38, GRG39 ET GRG50 : GENES INEDITS DE L'EPSP SYNTHETASE CONFERANT UNE RESISTANCE AUX HERBICIDES

(54) Title: GRG33, GRG35, GRG36, GRG37, GRG38, GRG39, AND GRG50: NOVEL EPSP SYNTHASE GENES CONFERRING HERBICIDE RESISTANCE

(57) Abrégé/Abstract:
Compositions and methods for conferring herbicide resistance 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. 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, the present invention provides for isolated nucleic acid molecules comprising the nucleotide sequence set forth in SEQ ID NO: 1, 3, 4, 6, 1, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22, the herbicide resistance nucleotide sequence deposited in a bacterial host as Accession Nos. NRRL B-30932, B-30933, B-30934, B-30945, B-30946, B-30947, or B-30948, as well as variants and fragments thereof.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 January 2008 (03.01.2008)

PCT

(10) International Publication Number
WO 2008/002964 A3

(51) International Patent Classification:
C12N 15/82 (2006.01) *C12Q 1/68* (2006.01)
A01H 1/00 (2006.01)

(21) International Application Number:
PCT/US2007/072206

(22) International Filing Date: 27 June 2007 (27.06.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/816,676	27 June 2006 (27.06.2006)	US
60/819,119	7 July 2006 (07.07.2006)	US
60/819,122	7 July 2006 (07.07.2006)	US

(71) Applicant *(for all designated States except US)*: **ATHENIX CORPORATION** [US/US]; 108 T.W. Alexander Drive, P.O. Box 110347, Research Triangle Park, NC 27709 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PETERS, Cheryl, L.** [US/US]; 6308 Valley Estates Drive, Raleigh, NC 27612 (US). **VANDE BERG, Brian** [US/US]; 5012 Silhouette Drive, Durham, NC 27713 (US). **CARR, Brian** [US/US]; 3104 Shinleaf Court, Raleigh, NC 27613 (US). **TOMSO, Daniel, John** [US/US]; 114 Vista Del Rio Drive, Bahama, NC 27503 (US).

(74) Agents: **DAVENPORT, Destiny, M.** 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, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, 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, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, 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, MT, 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*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

(88) Date of publication of the international search report:
10 July 2008

(54) Title: GRG33, GRG35, GRG36, GRG37, GRG38, GRG39, AND GRG50: NOVEL EPSP SYNTHASE GENES CONFERRING HERBICIDE RESISTANCE

(57) Abstract: Compositions and methods for conferring herbicide resistance 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. 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, the present invention provides for isolated nucleic acid molecules comprising the nucleotide sequence set forth in SEQ ID NO: 1, 3, 4, 6, 1, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22, the herbicide resistance nucleotide sequence deposited in a bacterial host as Accession Nos. NRRL B-30932, B-30933, B-30934, B-30945, B-30946, B-30947, or B-30948, as well as variants and fragments thereof.

WO 2008/002964 A3

GRG33, GRG35, GRG36, GRG37, GRG38, GRG39 AND GRG50: NOVEL EPSP SYNTHASE GENES CONFERRING HERBICIDE RESISTANCE

FIELD OF THE INVENTION

This invention provides novel genes encoding 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase that provide herbicide resistance. These genes are useful in plant biology, crop breeding, and plant cell culture.

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 to 5-enolpyruvyl-3-phosphoshikimic acid. Inhibition of this enzyme (5-enolpyruvylshikimate-3-phosphate synthase; referred to herein as "EPSP synthase") 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 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 *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).

U.S. Patent 6,040,497 reports mutant maize EPSP synthase enzymes having substitutions of threonine to isoleucine at position 102 and proline to serine at position 106 (the "TIPS" mutation). Such alterations confer glyphosate resistance upon the maize enzyme. A mutated EPSP synthase from *Salmonella typhimurium* strain CT7 confers glyphosate resistance in bacterial cells, and is reported to confer glyphosate

resistance upon plant cells (U.S. Patent Nos. 4,535,060; 4,769,061; and 5,094,945). He *et al.* ((2001) *Biochim et Biophysica Acta* 1568:1-6) have developed EPSP synthases with increased glyphosate tolerance by mutagenesis and recombination between the *E. coli* and *Salmonella typhimurium* EPSP synthase genes, and suggest that mutations at position 42 (T42M) and position 230 (Q230K) are likely responsible for the observed resistance.

Subsequent work (He *et al.* (2003) *Biosci. Biotech. Biochem.* 67:1405-1409) shows that the T42M mutation (threonine to methionine) is sufficient to improve tolerance of both the *E. coli* and *Salmonella typhimurium* enzymes. These enzymes contain amino acid substitutions in their active sites that prevent the binding of glyphosate without affecting binding by PEP or S3P. Mutations that occur in the hinge region between the two globular domains of EPSP synthase have been shown to alter the binding affinity of glyphosate but not PEP (He *et al.*, 2003, *supra*). Therefore, such enzymes have high catalytic activity, even in the presence of glyphosate.

Due to the many advantages herbicide resistance plants provide, methods for identifying herbicide resistance genes with glyphosate resistance activity are desirable.

SUMMARY OF INVENTION

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 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.

In particular, 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. In particular, the present invention provides

for isolated nucleic acid molecules comprising the nucleotide sequence set forth in SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22, the herbicide resistance nucleotide sequence deposited in a bacterial host as Accession Nos. NRRL B-30932, B-30933, B-30934, B-30945, B-30946, B-30947, or B-30948, 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 or a complement of a sequence of the invention are also encompassed.

DESCRIPTION OF FIGURES

Figure 1 shows an alignment of the amino acid sequence of GRG33 (SEQ ID NO:2) and GRG35 (SEQ ID NO:5) with EPSP synthase sequences from *Streptomyces cooelicolor* (SEQ ID NO:24), *Streptomyces avermitilis* (SEQ ID NO:25), *Zea mays* (SEQ ID NO:38), and *E. coli* (SEQ ID NO:37). The alignment shows the most highly conserved amino acid residues highlighted in black and highly conserved amino acid residues highlighted in gray.

Figure 2 shows an alignment of the amino acid sequence of GRG36 (SEQ ID NO:8) with EPSP synthase sequences from *Bacillus halodurans* (SEQ ID NO:26), *Bacillus claussi* (SEQ ID NO:27), *Zea mays* (SEQ ID NO:38), and *E. coli* (SEQ ID NO:37). The alignment shows the most highly conserved amino acid residues highlighted in black and highly conserved amino acid residues highlighted in gray.

Figure 3 shows an alignment of GRG38 (SEQ ID NO:16) and GRG50 (SEQ ID NO:22) with other EPSP synthase enzymes, including GRG8 (SEQ ID NO:29), GRG12 (SEQ ID NO:30), GRG15 (SEQ ID NO:31), GRG5 (SEQ ID NO:32), GRG6 (SEQ ID NO:33), GRG7 (SEQ ID NO:34), GRG9 (SEQ ID NO:35), GRG1 (SEQ ID NO:41), the EPSP synthase described in International Patent Application No. WO2005014820 (SEQ ID NO:36), and EPSP synthase enzymes from *E. coli* (SEQ ID NO:37), *Zea mays* (SEQ ID NO:38), *Agrobacterium tumefaciens* (SEQ ID NO:39), and *Bacillus subtilis* (SEQ ID NO:40). The alignment shows the most highly conserved amino acid residues highlighted in black, and highly conserved amino acid residues highlighted in gray.

DETAILED DESCRIPTION

The present invention is drawn to compositions and methods for regulating herbicide resistance in organisms, particularly in plants or plant cells. The methods involve transforming organisms with a nucleotide sequence encoding a glyphosate resistance gene of the invention. In particular, a nucleotide sequence of the invention is useful for preparing plants that show increased tolerance to the herbicide glyphosate. Thus, transformed bacteria, plants, plant cells, plant tissues and seeds are provided. Compositions include nucleic acids and proteins relating to herbicide tolerance in microorganisms and plants as well as transformed bacteria, plants, plant tissues and seeds. More particularly, nucleotide sequences of the glyphosate resistance genes (*grg33, syngrg33, grg35, syngrg35, grg36, syngrg36, grg37, syngrg37, grg38, syngrg38, grg39, syngrg39, grg50, syngrg50*) 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 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, 3, 4, 6, 7, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, and fragments and variants thereof 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, 5, 8, 11, 14, 16, 19, or 22, and fragments and variants thereof that confer glyphosate resistance or tolerance to a host cell.

Plasmids containing the herbicide resistance nucleotide sequences of the invention were deposited in the permanent collection of the Agricultural Research Service Culture Collection, Northern Regional Research Laboratory (NRRL), 1815 North University Street, Peoria, Illinois 61604, United States of America, on June 9, 2006, and assigned Accession Nos. NRRL B-30932 (for *grg33*), NRRL B-30933 (for *grg35*), and NRRL B-30934 (for *grg36*); and on June 26, 2006, and assigned Accession Nos. NRRL B-30945 (for *grg37*), NRRL B-30946 (for *grg38*), NRRL B-30947 (for *grg39*), and NRRL B-30948 (for *grg50*). This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of 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 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 period of time than cells that do not express the protein. A "glyphosate resistance protein" includes 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 discernable from untreated cells.

Isolated Nucleic Acid Molecules, and Variants and Fragments Thereof

One aspect of the invention pertains to isolated or recombinant nucleic acid molecules comprising nucleotide sequences encoding herbicide resistance proteins and polypeptides or 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, recombinant DNA, or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecules can be single-stranded or double-stranded, but preferably are double-stranded DNA.

Nucleotide sequences encoding the proteins of the present invention include the sequences set forth in SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, the herbicide resistance nucleotide sequence deposited in a bacterial host as Accession Nos. NRRL B-30932, B-30933, B-30934, B-30945, B-30946, B-30947, or B-30948, 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. In some embodiments, the complement hybridizes across the full length of the sequence of the invention. In another embodiment, the

complement hybridizes across at least about 50% of the sequence of the invention, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% of a sequence of the invention. The corresponding amino acid sequences for the herbicide resistance proteins encoded by these nucleotide sequences are set forth in SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22. The invention also encompasses nucleic acid molecules comprising nucleotide sequences encoding partial-length herbicide resistance proteins, and complements thereof.

In some embodiments, the polynucleotides of the present invention encode polypeptides that are Class III EPSP synthase enzymes. For the purposes of the present invention, a "Class III EPSP synthase enzyme" is an herbicide tolerant or herbicide resistant polypeptide containing one or more of the amino acid sequence domains described in U.S. Patent Application No. 11/400,598, which is herein incorporated by reference in its entirety.

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 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, "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 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 intended a portion of a 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 contiguous nucleotides, or up to the number of nucleotides present in a full-length herbicide resistance-encoding nucleotide sequence disclosed herein (for example, 1329 nucleotides for SEQ ID NO:1; 1353 nucleotides for SEQ ID NO:4; 1344 nucleotides for SEQ ID NO:7, etc) depending upon the intended use. 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 protein disclosed herein as SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22. 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, 442 amino acids for SEQ ID NO:2; 450 for SEQ ID NO:5; 447 amino acids for SEQ ID NO:8, etc).

Preferred 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, 4, 6, 7, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23. 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, or 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, 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. BLAST nucleotide searches can be performed with the BLASTN program, score = 100, wordlength = 12, to obtain nucleotide sequences homologous to glyphosate-resistant 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. 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., 9865 Scranton Rd., San Diego, California, USA). When utilizing the ALIGN program for comparing 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) *supra*, will be used to determine sequence identity or similarity using the following parameters: % identity and % similarity for a 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 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 an herbicide resistance protein disclosed herein but that differ conservatively because of the degeneracy of the genetic code, as well as those that are sufficiently identical as discussed above. 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 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 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 protein, without altering the biological activity of the protein. 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, preferably 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. Examples of residues that are conserved and that may be essential for protein activity include, for example, residues that are identical between all proteins contained in the alignment of Figures 1, 2, or 3. Examples of residues that are conserved but that may allow conservative amino acid substitutions and still retain activity include, for example, residues that have only conservative substitutions between all proteins contained in the alignment of Figures 1, 2, or 3.

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 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 J., and Russell, D.W. (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, NY).

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 sequences disclosed herein. Degenerate primers designed on the basis of conserved nucleotides or amino acid residues in the nucleotide sequences or encoded amino acid sequences can additionally be used. The probe typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, at least about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1200, 1300 consecutive nucleotides of an herbicide resistance-encoding nucleotide sequence 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, New York), 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 and are at least about 10 nucleotides in length, or 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 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, New York).

Hybridization of such sequences may be carried out under stringent conditions. By "stringent conditions" or "stringent hybridization conditions" is intended conditions 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, 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, preferably less than 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 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 sulfate) 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. Unless otherwise specified, hybridization conditions are under high stringency.

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: $T_m = 81.5°C + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\%\ form) - 500/L$; 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 ≥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, New York).

Isolated Proteins and Variants and Fragments Thereof

Herbicide resistance proteins are also encompassed within the present invention. By "herbicide resistance protein" is intended a protein having the amino acid sequence set forth in SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22. 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 SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22, 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, 5, 8, 11, 14, 16, 19, or 22. 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%, about 70%, 75%, 80%, 85%, or 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 16, 19, or 22. Variants also include polypeptides encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, 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 *grg* and *syngrg* genes 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 may lead to generation of variants of *grg* and *syngrg* 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, N.Y.; U.S. Patent No. 4,196,265).

Altered or Improved Variants

It is recognized that the DNA sequences of the *grg* or *syngrg* genes of the invention 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 the *grg* or *syngrg* sequences disclosed herein. This protein may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions of one or more amino acids of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, including up to about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 15, about 20, about 25, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 100 or more amino acid substitutions, deletions or insertions.

Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the GRG proteins disclosed herein can be prepared by mutations in the DNA. This may also be accomplished by one of several forms of mutagenesis and/or in directed evolution. In some aspects, the changes encoded in the amino acid sequence will not substantially affect function of the protein. Such variants will possess the desired herbicide resistance activity. However, it is understood that the ability of the GRG proteins disclosed herein to confer herbicide resistance may be improved by one use of such techniques upon the compositions of this invention. For example, one may express the *grg* or *syngrg* sequences disclosed herein 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, one can isolate the DNA of the invention (for example by preparing plasmid DNA, or by amplifying by PCR and cloning the resulting PCR fragment into a vector), culture the *grg* mutations in a non-mutagenic strain, and identify mutated

genes with improved resistance to an herbicide such as glyphosate, for example by growing cells in increasing concentrations of glyphosate and testing for clones that confer ability to tolerate increased 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. This can include insertions, deletions, or alterations introduced by modern molecular methods, such as PCR, including PCR amplifications that alter or extend the protein coding 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 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 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 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 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; Crameri *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; Crameri *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

Transformation of Bacterial or Plant Cells

Provided herein are novel isolated genes that confer resistance to an herbicide. Also provided are amino acid sequences of the GRG proteins of the invention. 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 *grg* or *syngrg* genes are useful as markers to assess transformation of bacterial or plant cells. Methods for detecting the presence of a transgene in a plant, plant organ (e.g., leaves, stems, roots, etc.), seed, plant cell, propagule, embryo or progeny of the same are well known in the art.

By engineering the genes of the invention to be expressed from a promoter known to stimulate transcription in the organism to be tested and properly translated to generate an intact GRG peptide, and placing the cells in an otherwise toxic concentration of herbicide, one can identify cells that have been transformed with the DNA by virtue of their resistance to herbicide. By "promoter" is intended a nucleic acid 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 as "control sequences") are necessary for the expression of a DNA sequence of interest.

Transformation of bacterial cells is accomplished by one of several techniques 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 DNA). In one aspect of the invention, the *grg* or *syngrg* genes disclosed herein are useful as markers to assess transformation of bacterial or plant cells.

Transformation of plant cells can be accomplished in 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 undifferentiated (e.g. callus, suspension culture cells, protoplasts, leaf cells, root cells, phloem cells, pollen). "Transgenic plants" or "transformed plants" or "stably transformed" plants or 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 into the genome of the plant and is capable of being inherited by progeny thereof.

The *grg* genes 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, including recombinant 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, particularly a heterologous promoter) operably-linked to a DNA sequence of the invention, and/or a transcriptional and translational 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 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. It may also be preferable to engineer the plant expression cassette 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.

In one aspect of the invention, synthetic DNA sequences are designed for a given polypeptide, such as the polypeptides of the invention. Expression of the open reading frame of the synthetic DNA sequence in a cell results in production of the polypeptide of the invention. Synthetic DNA sequences can be useful to simply remove unwanted restriction endonuclease sites, to facilitate DNA cloning strategies, to alter or remove any potential codon bias, to alter or improve GC content, to remove or alter alternate reading frames, and/or to alter or remove intron/exon splice recognition sites, polyadenylation sites, Shine-Delgarno sequences, unwanted promoter elements and the like that may be present in a native DNA sequence. It is also possible that synthetic DNA sequences may be utilized to introduce other improvements to a DNA sequence, such as introduction of an intron sequence, creation of a DNA sequence that in expressed as a protein fusion to organelle targeting sequences, such as chloroplast transit peptides, apoplast/vacuolar targeting peptides, or peptide sequences that result in retention of the resulting peptide in the endoplasmic reticulum. Synthetic genes can also be synthesized using host cell-preferred codons for improved expression, or may be synthesized using codons at a host-preferred codon usage frequency. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11; 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.

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 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 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 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 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.

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*-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 is 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 plant 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 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 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 then 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, 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 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.* 12:601-606. 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) *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 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.

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 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" is then probed with, for example, radiolabeled ^{32}P target DNA fragments 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, 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 the genes disclosed herein is then tested by hybridizing the filter to a radioactive probe derived from a polynucleotide of the invention, 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.

Methods for increasing plant yield

Methods for increasing plant yield are provided. The methods comprise introducing into a plant or plant cell a polynucleotide comprising a *grg* sequence disclosed herein. As defined herein, the "yield" of the plant refers to the quality and/or quantity of biomass produced by the plant. By "biomass" is intended any measured plant product. An increase in biomass production is any improvement in the yield of the measured plant product. Increasing plant yield has several commercial applications. For example, increasing plant leaf biomass may increase the yield of leafy vegetables for human or animal consumption. Additionally, increasing leaf biomass can be used to increase production of plant-derived pharmaceutical or industrial products. An increase in yield can comprise any statistically significant

increase including, but not limited to, at least a 1% increase, at least a 3% increase, at least a 5% increase, at least a 10% increase, at least a 20% increase, at least a 30%, at least a 50%, at least a 70%, at least a 100% or a greater increase.

In specific methods, the plant is treated with an effective concentration of an herbicide, where the herbicide application results in enhanced plant yield. By "effective concentration" is intended the concentration which allows the increased yield in the plant. Such effective concentrations for herbicides of interest are generally known in the art. The herbicide may be applied either pre- or post emergence in accordance with usual techniques for herbicide application to fields comprising crops which have been rendered resistant to the herbicide by heterologous expression of a *grg* gene of the invention.

Methods for conferring herbicide resistance in a plant or plant part are also provided. In such methods, a *grg* polynucleotide disclosed herein is introduced into the plant, wherein expression of the polynucleotide results in glyphosate tolerance or resistance. Plants produced via this method can be treated with an effective concentration of an herbicide and display an increased tolerance to the herbicide. An "effective concentration" of an herbicide in this application is an amount sufficient to slow or stop the growth of plants or plant parts that are not naturally resistant or rendered resistant to the herbicide.

In another embodiment, methods for conferring herbicide resistance in a plant or plant part are provided, wherein the plant or plant part is grown under higher or lower than ambient environmental temperatures as described *supra*. Glyphosate tolerant EPSP synthase enzymes having thermal stability at higher or lower temperatures, or have temperature optima at higher or lower temperatures, are useful for conferring glyphosate tolerance in plants that are grown under such conditions.

Methods of controlling weeds in a field

Methods for selectively controlling weeds in a field containing a plant are also provided. In one embodiment, the plant seeds or plants are glyphosate resistant as a result of a *grg* polynucleotide disclosed herein being inserted into the plant seed or plant. In specific methods, the plant is treated with an effective concentration of an herbicide, where the herbicide application results in a selective control of weeds or other untransformed plants. By "effective concentration" is intended the concentration which controls the growth or spread of weeds or other untransformed

plants without significantly affecting the glyphosate-resistant plant or plant seed. Such effective concentrations for herbicides of interest are generally known in the art. The herbicide may be applied either pre- or post emergence in accordance with usual techniques for herbicide application to fields comprising plants or plant seeds which have been rendered resistant to the herbicide.

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

EXPERIMENTAL

Example 1. Isolation of glyphosate resistant EPSP synthases.

Strains capable of growth in presence of glyphosate were isolated by plating samples of soil on HEPES Mineral Salts Medium (HMSM) containing glyphosate as the sole source of phosphorus. Since HMSM contains no aromatic amino acids, a strain must be resistant to glyphosate in order to grow on this media.

Two grams of soil were suspended in approximately 10 ml of water, vortexed for 15 seconds and permitted to settle for 15 minutes. A 10 µl loopful of this suspension was added to 3 ml of HMSM supplemented with 10 mM glyphosate (pH 7.0). HMSM contains (per liter): 10 g glucose, 2 g NH_4SO_4, 9.53 g HEPES, 1.0 ml 0.8 M $MgSO_4$, 1.0 ml 0.1 M $CaCl_2$, 1.0 ml Trace Elements Solution (In 100 ml of 1000x solution: 0.1 g $FeSO_4 \cdot 7H_2O$, 0.5 mg $CuSO_4 \cdot 5H_2O$, 1.0 mg H_3BO_3, 1.0 mg $MnSO_4 \cdot 5H_2O$, 7.0 mg $ZnSO_4 \cdot 7H_2O$, 1.0 mg MoO_3, 4.0 g KCl). The culture was grown in a shaker incubator for four days at 28°C and then 20 µl was used to inoculate 2.5 ml of fresh HMSM containing 10 mM glyphosate as the only phosphorus source. After two days, 20 µl was used to inoculate another fresh 2.5 ml culture. After 5 days, 20 µl was used to inoculate a fresh 2.5 ml culture. After sufficient growth, the culture was plated onto solid media by streaking a 1 µl loop onto the surface of agar plate containing HMSM agar containing 100 mM glyphosate as the sole phosphorus source and stored at 28°C. The culture was then replated for isolation. The strains listed in Table 1 were among the strains selected due to their ability to grow in the presence of high glyphosate concentrations.

Table 1.

Strain Name	Strain ID	EPSP synthase Gene Name
ATX21561	Unknown	*grg33*
ATX21563	Unknown	*grg35*
ATX21567	Unknown	*grg36*

Example 2. Isolation of glyphosate resistant EPSP synthases *grg37* and *grg39*.

Strains capable of growth in presence of glyphosate were isolated by plating samples of soil on various growth media containing glyphosate. Some strains were isolated on mineral salts media supplemented with glyphosate. Other strains were isolated under rich media in the presence of glyphosate and later tested on mineral salts media supplemented with glyphosate. Since the mineral salts media contain no aromatic amino acids, a strain must be resistant to glyphosate in order to grow on this media.

Strains ATX21800 and ATX21804 were isolated by incubation under rich conditions and supplemention with glyphosate. These strains were then tested for their ability to grow in the presence of glyphosate without aromatic amino acids. Strain ATX21804 was isolated from soil (0.01 grams) that was air dried for two days and plated onto nutrient broth agar supplemented with 100 mM glyphosate. A small amount (10 µl) was then used to inoculate an eosin methylene blue agar plate containing 300 mM glyphosate. ATX21800 was isolated by incubating 0.01 grams soil with 3 ml nutrient broth supplemented with 100 mM glyphosate. After initial isolation, each strain was inoculated into Luria Bertani agar plates to confirm single colony type. These strains were then tested on Brunner minimal medium containing 100mM glyphosate and were confirmed to grow in the presence of glyphosate without aromatic amino acids.

The strains listed in Table 2 were among the strains selected due to their ability to grow in the presence of high glyphosate concentrations.

Table 2.

Strain Name	Strain ID	EPSP synthase Gene Name
ATX21800	Unknown	*grg37*
ATX21804	Unknown	*grg39*

Example 3. Isolation of glyphosate resistant EPSP synthases *grg38* and *grg50*.

Strains capable of growth in presence of glyphosate were isolated by plating samples of soil on various growth media containing glyphosate. Some strains were isolated on mineral salts media supplemented with glyphosate. Other strains were isolated under rich media in the presence of glyphosate and later tested on mineral salts media supplemented with glyphosate. Since the mineral salts media contain no aromatic amino acids, a strain must be resistant to glyphosate in order to grow on this media.

Strain ATX20103 was isolated by suspending approximately 2 grams of soil in 10 ml of water, vortexing for 15 seconds and permitting to settle for 15 minutes. A 10 µl loopful of this suspension was added to 3 ml of Tris MSM (TMSM) supplemented with 10 mM glyphosate (pH 7.0). TMSM contains (per liter): 10 g glucose, 2 g NH_4SO_4, 12.12 g Tris, 1.0 ml 0.8 M $MgSO_4$, 1.0 ml 0.1 M $CaCl_2$, 1.0 ml Trace Elements Solution (In 100 ml of 1000x solution: 0.1 g $FeSO_4 \cdot 7H_2O$, 0.5 mg $CuSO_4 \cdot 5H_2O$, 1.0 mg H_3BO_3, 1.0 mg $MnSO_4 \cdot 5H_2O$, 7.0 mg $ZnSO_4 \cdot 7H_2O$, 1.0 mg MoO_3, 4.0 g KCl). The culture was then incubated at 28°C for isolation over repeated rounds of selection and then inoculated onto Luria Bertani agar to confirm single colony type. ATX20103 was then reconfirmed to grow on TMSM in the presence of glyphosate without aromatic acids.

Strain ATX21806 was isolated by incubating under rich conditions and supplementing with glyphosate. This strain was then tested for its ability to growth in the presence of glyphosate without aromatic amino acids. Strain ATX21806 was isolated from soil (0.01 grams) that had been suspended in 10 ml water overnight. A small amount (10 µl) was then used to inoculate an eosin methylene blue agar plate containing 300 mM glyphosate. After each initial isolation, the strain was inoculated into Luria Bertani agar plates to confirm single colony type. The strain was then tested on Brunner minimal medium at 100mM glyphosate and was confirmed to grow in the presence of glyphosate without aromatic amino acids.

The strains listed in Table 3 were among the strains selected due to their ability to grow in the presence of high glyphosate concentrations.

Table 3.

Strain Name	Strain ID	EPSP synthase Gene Name
ATX21806	Unknown	*grg38*
ATX20103	*Rhizobium leguminosarum*	*grg50*

Example 4. Cloning of glyphosate-resistant EPSP Synthases

Genomic DNA was extracted from the strains described in Tables 1, 2, and 3, 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 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 x 10^7 and 3 x 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 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* XL-Blue MRF'(*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.

The *E. coli* XL-Blue MRF'(*aroA*) used for the transfection expresses the F-pilus, and also contains a deletion of the *aroA* gene encoding the endogenous *E. coli* EPSP synthase enzyme. This strain is also referred to as herein as Δ*aroA*. This Δ*aroA* strain is unable to grow on minimal media lacking aromatic amino acids, unless complemented by a functional EPSP synthase. Since glyphosate is a potent inhibitor of typical, glyphosate-sensitive EPSP synthases, such as type I EPSP synthases, transfected clones expressing a non-glyphosate resistant EPSP synthase would be able to grown on M63 plates lacking glyphosate, but would be unable to grow on M63 containing either 20mM or 50mM glyphosate. In order to grow on M63 plates containing 20mM or 50mM glyphosate, the cells must contain a plasmid that

expresses an EPSP synthase that is both (1) capable of complementing the Δ*aroA* mutation of these cells, and (2) resistant to glyphosate. Thus, this screening method allows identification of clones containing glyphosate-resistant EPSP synthases.

Colonies growing on 20mM or 50mM glyphosate 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.

Using this approach, as sometimes modified for each library as known and appreciated in the art, library clones containing EPSP synthase genes were identified for each of the strains listed in Table 4.

Example 5. DNA and protein sequences of EPSP synthases

The DNA sequences of the glyphosate-resistant EPSP synthases was determined for each of the clones described above by methods well known in the art.

grg33. The DNA sequence of *grg33* is provided herein as SEQ ID NO:1. The predicted translation product of *grg33* (GRG33) is provided herein as SEQ ID NO:2. A synthetic sequence encoding GRG33 (*syngrg33*) was also designed and is provided herein as SEQ ID NO:3.

grg35. The DNA sequence of *grg35* is provided herein as SEQ ID NO:4. The predicted translation product of *grg35* (GRG35) is provided herein as SEQ ID NO:5. A synthetic sequence encoding GRG35 (*syngrg35*) was also designed and is provided herein as SEQ ID NO:6.

grg36. The DNA sequence of *grg36* is provided herein as SEQ ID NO:7. The predicted translation product of *grg36* (GRG36) is provided herein as SEQ ID NO:8. A synthetic sequence encoding GRG36 (*syngrg36*) was also designed and is provided herein as SEQ ID NO:9.

grg37. The DNA sequence of *grg37* is provided herein as SEQ ID NO:10. The predicted translation product of *grg37* (GRG37) is provided herein as SEQ ID NO:11. A synthetic sequence encoding GRG37 (*syngrg37*) was also designed and is provided herein as SEQ ID NO:12.

grg38. The DNA sequence of *grg38* is provided herein as SEQ ID NO:15. The predicted translation product of *grg38* (GRG38) is provided herein as SEQ ID NO:16. A synthetic sequence encoding GRG38 (*syngrg38*) was also designed and is provided herein as SEQ ID NO:17.

grg39. The DNA sequence of *grg39* is provided herein as SEQ ID NO:18. The predicted translation product of *grg39* (GRG39) is provided herein as SEQ ID NO:19. A synthetic sequence encoding GRG39 (*syngrg39*) was also designed and is provided herein as SEQ ID NO:20.

grg50. The DNA sequence of *grg50* is provided herein as SEQ ID NO:21. The predicted translation product of *grg50* (GRG50) is provided herein as SEQ ID NO:22. A synthetic sequence encoding GRG50 (*syngrg50*) was also designed and is provided herein as SEQ ID NO:23.

Clones containing each of the *grg33, grg35, grg36, grg37, grg38, grg39,* and *grg50* EPSP synthase genes were deposited at NRRL on June 9, 2006 or June 26, 2007 and assigned deposit numbers as in Table 4.

Table 4. Clones containing glyphosate-resistant EPSP synthases

EPSPS	Strain yielding EPSPS	Original Isolate in pBKCMV	NRRL Number
GRG33	ATX21561	pAX1947	B-30932
GRG35	ATX21563	pAX1948	B-30933
GRG36	ATX21567	pAX1949	B-30934
GRG37	ATX21800	pAX1963	B-30945
GRG38	ATX21806	pAX1964	B-30946
GRG39	ATX21804	pAX1965	B-30947
GRG50	ATX20103	pAX1966	B-30948

Each of the proteins GRG33, GRG35, and GRG36 showed regions of homology to EPSP synthase enzymes in the NCBI database by BLAST search. The EPSPS enzyme with the highest protein sequence identity to each GRG enzyme is listed in Table 5.

Table 5. Homology of GRG33-GRG36 to known EPSP synthases

Protein	Strain with homologous EPSPS enzyme	% Identity
GRG33	GRG35, *S.coelicolor*	88%, 86%
GRG35	GRG33, *S.coelicolor*	88%, 85%
GRG36	*Bacillus halodurans*	53%

The amino acid sequences of GRG33 and GRG35 are 88% identical. A search of public protein databases with the amino acid sequence of GRG33 shows that this protein is 86% identical over 430 amino acids to the EPSP synthase from *Streptomyces coelicolor* (SEQ ID NO:24 GENBANK® Accession No. NP 629359.1), and 82% identical over 430 amino acids to the EPSP synthase from *Streptomyces avermitilis* (SEQ ID NO:25; GENBANK® Accession No. NP824218.1).

The amino acid sequence of GRG35 similarly is 85% identical over 434 amino acids to the EPSP synthase from *Streptomyces coelicolor* (SEQ ID NO:24; GENBANK® Accession No. NP 629359.1), and 81% identical over 441 amino acids to the EPSP synthase from *Streptomyces avermitilis* (SEQ ID NO:25; GENBANK® Accession No. NP 824218.1).

Table 6. Amino acid identity of GRG33 and GRG35 with Streptomyces EPSP synthases

EPSP synthase	Identity with GRG33	Identity with GRG35
GRG33	-	88%
GRG35	88%	-
Streptomyces coelicolor A3(2)	84%	82%
Streptomyces avermitilis MA-4680	80%	78%
E. coli	30%	29%
Maize	30%	29%

A search of public protein databases with the amino acid sequence of GRG36 shows that this protein is related to the EPSP synthase from *Bacillus halodurans*, (64% identical over 441 amino acids, SEQ ID NO:26; GENBANK® Accession No. BAB06432.1), and to a lesser extent to the EPSP synthase from *Bacillus clausii* (55% identical over 439 amino acids; SEQ ID NO:27; GENBANK® Accession No. BAD63759.1)

Table 7. Amino acid identity of GRG36 with EPSP synthases

EPSP synthase	Identity with GRG36
Bacillus halodurans	62%
Bacillus clausii	54%
E. coli	31%
Maize	34%

A search of public protein databases with the amino acid sequence of GRG37 shows that this protein is 81% identical to the EPSP synthase from *Arthrobacter sp. FB24* (SEQ ID NO:28, GENBANK® Accession No. ZP_00413033.1)

The *grg37* open reading frame has two potential start codons. The upstream ATG (predicted amino acid sequence MTASPMGASADNS…….(corresponding to amino acid positions 1 through 13 of SEQ ID NO:10)) contains the best ribosome binding site in correct proximity. However, a second downstream ATG may be used. This ORF yields the predicted amino acid sequence MGASADNS…..(corresponding to amino acid positions 6 through 13 of SEQ ID NO:10)). The upstream ATG appears to have a ribosome binding site ("RBS") that is a better match to the consensus RBS sequence. However, the open reading frame originating from this upstream ATG appears to be translationally coupled to an upstream open reading frame. Translational coupling is one strategy known in the art to be employed by bacteria to ensure good initiation and can substitute for a ribosome binding site. The nucleotide sequence for the downstream start site is provided herein as SEQ ID NO:13, and the encoded amino acid sequence is provided herein as SEQ ID NO:14.

GRG39 shows 96% amino acid identity to the GRG30 EPSP synthase sequence, and is highly homologous to the GRG29 EPSP synthase sequence described in U.S. Patent Application No. 11/760,570 filed June 8, 2007.

GRG38 shows 94% amino acid identity to the GRG12 EPSP synthase described in U.S. Patent Application No. 11/400,598, filed April 7, 2006 (Table 8). GRG38 also contains the domains of the Class III EPSP synthases described in U.S. Patent Application No. 11/400,598.

GRG50 shows 95% amino acid identity to the GRG8 EPSP synthase described in U.S. Patent Application No. 11/315,678 filed December 22, 2005 (Table 8). GRG50 also contains the domains of the Class III EPSP synthases described in U.S. Application 11/400,598, filed April 7, 2006.

Table 8. Comparison with other Class III EPSP synthases

EPSPS	Amino acid identity with GRG38	Amino acid identity with GRG50
GRG38	-	65%
GRG50	65%	-
GRG8	65%	95%
GRG12	87%	62%
GRG6	67%	67%
GRG9	64%	70%
GRG15	64%	71%
GRG5	68%	68%
GRG37	67%	68%
E. coli (non-Class III)	32%	34%
Maize (non-Class III)	32%	31%

Example 6. Cloning of novel glyphosate-resistant EPSP synthases into an *E. coli* expression vector

The EPSP synthase genes contained in the clones of Table 4 were sub-cloned into the *E. coli* expression vector pRSF1b (Invitrogen). Resulting clones were confirmed by DNA sequencing, and used to induce expression of each EPSP synthase in *E. coli*. The expressed His-tagged protein was then purified as known in the art.

Example 7. Glyphosate resistance of GRG33, GRG35, and GRG36 EPSP synthases

The pRSF1b clones were plated onto M63+ plates containing antibiotic and either 0mM or 50mM glyphosate. Growth was scored after two days growth at 37°C. Each of the three EPSP synthases was observed to confer resistance to 50 mM glyphosate in *E. coli* cells (Table 9).

Table 9. Glyphosate screen

EPSPS	Clone in pRSF1B	Growth on 50mM glyphosate
Vector	-	-
GRG33	pAX1951	+++
GRG35	pAX1952	+++
GRG36	pAX1953	+++

Example 8. Glyphosate resistance of GRG37 and GRG39 EPSP synthases

Cells containing the plasmid clones shown in Table 4 were plated onto M63+ plates containing antibiotic and either 0mM or 20mM glyphosate. Growth was scored after two days growth at 37°C. Each of the EPSP synthases was observed to confer resistance to 20 mM glyphosate in *E. coli* cells (Table 10).

Table 10. Glyphosate screen

EPSPS	Plasmid Clone	Growth on 20mM glyphosate
Vector	-	-
GRG37	pAX1963	++
GRG39	pAX1965	++

Example 9. Glyphosate resistance of GRG38 and GRG50 EPSP synthases

Cells containing the plasmid clones shown in Table 4 were plated onto M63+ plates containing antibiotic and either 0mM or 20mM glyphosate. Growth was scored after two days growth at 37°C. Each of the EPSP synthases was observed to confer resistance to 20 mM glyphosate in *E. coli* cells (Table 11).

Table 11. Glyphosate screen

EPSPS	Plasmid Clone	Growth on 20mM glyphosate
Vector	-	-
GRG38	pAX1964	++
GRG50	pAX1966	++

Example 10. Engineering *grg33, grg35, grg36, grg37, grg38, grg39, grg50, syngrg33, syngrg35, syngrg36, syngrg37, syngrg38, syngrg39, and syngrg50* for Plant Transformation

The open reading frame (ORF) for each of the *grg* genes is amplified by PCR from a full-length cDNA template. *Hin*d 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 *grg* PCR product is digested with, for example, *Hin*d III and the fragment containing the intact orF is isolated.

One may generate similar constructs that contain a chloroplast targeting sequence linked to the polynucleotide of the invention by methods known in the art.

A DNA fragment containing the EPSP synthase (and either containing or not containing a chloroplast targeting sequence) is cloned into a plasmid, for example at the *Hin*d III site of pAX200. pAX200 is 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 (or the promoter-leader-gene-terminator fragment) from this intermediate plasmid is subcloned into a plasmid such as pSB11 (Japan Tobacco, Inc.) to form a final plasmid, referred to herein as, for example, pSB11GRG33. pSB11GRG33 is organized such that the DNA fragment containing, for example, the promoter – *grg36* – terminator construct (or the promoter-leader-*grg36* – terminator construct) may be excised by appropriate restriction enzymes and also used for transformation into plants, for example, by aerosol beam injection. The structure of pSB11GRG33 is verified by restriction digest and gel electrophoresis, as well as by sequencing across the various cloning junctions. The same methods can be used to generate a final plasmid for each of the *grg* genes described herein.

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 antibiotic. Plasmid pSB11GRG36 carries spectinomycin resistance but is a narrow host range plasmid and cannot replicate in *Agrobacterium*. Antibiotic resistant colonies arise when

pSB11GRG36 integrates into the broad host range plasmid pSB1 through homologous recombination. The resulting cointegrate product is verified by Southern hybridization. The *Agrobacterium* strain harboring the cointegrate can be used to transform maize, for example, by the PureIntro method (Japan Tobacco).

Example 11. Transformation *grg33, grg35, grg36, grg37, grg38, grg39, grg50, syngrg33, syngrg35, syngrg36, syngrg37, syngrg38, syngrg39, and syngrg50* into Plant Cells

Maize ears are best collected 8-12 days after pollination. Embryos are isolated from the ears, and those embryos 0.8-1.5 mm in size are preferred for use in 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 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. However, it is not necessary *per se* to incubate the embryos overnight.

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

DNA constructs designed to express the GRG proteins of the present invention 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 about 30 min on osmotic media, and placed onto incubation media overnight at 25°C in the dark. To 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 about 5 days, 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

Table 12. 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

The pH of the solution is adjusted to pH 5.8 with 1N KOH/1N KCl, Gelrite (Sigma) is added at a concentration up to 3g/L, and the media is autoclaved. After cooling to 50°C, 2 ml/L of a 5 mg/ml stock solution of silver nitrate (Phytotechnology Labs) is added.

Example 12. Transformation of *grg33, grg35, grg36, grg37, grg38, grg39, grg50, syngrg33, syngrg35, syngrg36, syngrg37, syngrg38, syngrg39, and syngrg50* into Maize Plant Cells by *Agrobacterium*-Mediated Transformation

Ears are best collected 8-12 days after pollination. Embryos are isolated from the ears, and those embryos 0.8-1.5 mm in size are preferred for use in 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 about 5-10 min, and then plated onto co-cultivation media for about 3 days (25°C in the dark). After co-cultivation, explants are transferred to recovery period media for about five days (at

25°C in the dark). 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 as 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.

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. 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 certain changes and modifications may be practiced within the scope of the appended claims.

THAT WHICH IS CLAIMED:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

a) the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, or a complement thereof;

b) a nucleotide sequence having at least 90% sequence identity to the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, or 13, or a complement thereof;

c) a nucleotide sequence having at least 97% sequence identity to the nucleotide sequence of SEQ ID NO:15, 17, 18, 20, 21, or 23, or a complement thereof;

d) the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession Nos. NRRL B-30934, B-30932, B-30933, B-30945, B-30946, B-30947, or B-30948, or a complement thereof;

e) a nucleotide sequence that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:8, 2, 5, 11, 14, 16, 19, or 22;

f) a nucleotide sequence encoding a polypeptide having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO:8, 2, 5, 11, or 14; and,

g) a nucleotide sequence encoding a polypeptide having at least 97% amino acid sequence identity to the amino acid sequence of SEQ ID NO:16, 19, or 22.

2. The isolated nucleic acid molecule of claim 1, wherein said nucleotide 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 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.

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.

10. A transgenic seed comprising the nucleic acid molecule of claim 1.

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

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:8, 2, 5, 11, 14, 16, 19, or 22;

b) a polypeptide encoded by the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23;

c) a polypeptide comprising an amino acid sequence having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:8, 2, 5, 11, or 14, wherein said polypeptide has herbicide resistance activity;

d) a polypeptide comprising an amino acid sequence having at least 97% sequence identity to the amino acid sequence of SEQ ID NO:16, 19, or 22, wherein said polypeptide has herbicide resistance activity;

e) a polypeptide that is encoded by a nucleotide sequence that is at least 90% identical to the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, or 13, wherein said polypeptide has herbicide resistance activity;

f) a polypeptide that is encoded by a nucleotide sequence that is at least 97% identical to the nucleotide sequence of SEQ ID NO:15, 17, 18, 20, 21, or 23, wherein said polypeptide has herbicide resistance activity; and,

g) a polypeptide that is encoded by the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession Nos. NRRL B-30934, B-30932, B-30933, B-30945, B-30946, B-30947, or B-30948.

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

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 NO:8, 2, 5, 11, 14, 16, 19, or 22;

b) a polypeptide encoded by the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23;

c) a polypeptide comprising an amino acid sequence having at least 90% sequence identity to the amino acid sequence of SEQ ID NO:8, 2, 5, 11, or 14, wherein said polypeptide has herbicide resistance activity;

d) a polypeptide comprising an amino acid sequence having at least 97% sequence identity to the amino acid sequence of SEQ ID NO:16, 19, or 22, wherein said polypeptide has herbicide resistance activity;

e) a polypeptide that is encoded by a nucleotide sequence that is at least 90% identical to the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, or 13, wherein said polypeptide has herbicide resistance activity;

f) a polypeptide that is encoded by a nucleotide sequence that is at least 97% identical to the nucleotide sequence of SEQ ID NO:15, 17, 18, 20, 21, or 23, wherein said polypeptide has herbicide resistance activity; and,

g) a polypeptide that is encoded by the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession Nos. NRRL B-30934, B-30932, B-30933, B-30945, B-30946, B-30947, or B-30948.

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, and regenerating a transformed plant, wherein said nucleotide sequence is selected from the group consisting of:

a) the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, or a complement thereof;

b) a nucleotide sequence having at least 90% sequence identity to the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, or 13, or a complement thereof, wherein said polypeptide has herbicide resistance activity;

c) a nucleotide sequence having at least 97% sequence identity to the nucleotide sequence of SEQ ID NO:15, 17, 18, 20, 21, or 23, or a complement thereof, wherein said polypeptide has herbicide resistance activity;

d) the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession Nos. NRRL B-30934, B-30932, B-30933, B-30945, B-30946, B-30947, or B-30948, or a complement thereof;

e) a nucleotide sequence that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:8, 2, 5, 11, 14, 16, 19, or 22;

f) a nucleotide sequence encoding a polypeptide having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO:8, 2, 5, 11, or 14, wherein said polypeptide has herbicide resistance activity; and,

g) a nucleotide sequence encoding a polypeptide having at least 97% amino acid sequence identity to the amino acid sequence of SEQ ID NO:16, 19, or 22, wherein said polypeptide has herbicide resistance activity.

15. The method of claim 14, wherein said plant produces a polypeptide having herbicide resistance activity.

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

17. 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) the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, or a complement thereof;

b) a nucleotide sequence having at least 90% sequence identity to the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, or 13, or a complement thereof, wherein said polypeptide has herbicide resistance activity;

c) a nucleotide sequence having at least 97% sequence identity to the nucleotide sequence of SEQ ID NO:15, 17, 18, 20, 21, or 23, or a complement thereof, wherein said polypeptide has herbicide resistance activity;

d) the herbicide resistance nucleotide sequence of the DNA insert of the plasmid deposited as Accession Nos. NRRL B-30934, B-30932, B-30933, B-30945, B-30946, B-30947, or B-30948, or a complement thereof;

e) a nucleotide sequence that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:8, 2, 5, 11, 14, 16, 19, or 22;

f) a nucleotide sequence encoding a polypeptide having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO:8, 2, 5, 11, or 14, wherein said polypeptide has herbicide resistance activity; and,

g) a nucleotide sequence encoding a polypeptide having at least 97% amino acid sequence identity to the amino acid sequence of SEQ ID NO:16, 19, or 22, wherein said polypeptide has herbicide resistance activity;

wherein said nucleotide sequence is operably linked to a promoter that drives expression of a coding sequence in a plant cell.

18. The plant of claim 17, wherein said plant is a plant cell.

19. A method for generating a polynucleotide variant of a parent *grg33, grg35, grg36, grg7, grg38, grg39, or grg50* polynucleotide, comprising using the nucleotide sequence of SEQ ID NO:7, 1, 3, 4, 6, 9, 10, 12, 13, 15, 17, 18, 20, 21, or 23, or a fragment thereof, in a recombinogenic procedure, and testing a polypeptide encoded by the resulting polynucleotide sequence for an activity of interest.

20. The method of claim 19, wherein the activity of interest is glyphosate resistance activity.

21. The method of claim 20, wherein said polypeptide encoded by the resulting polynucleotide sequence has greater glyphosate resistance activity than a polypeptide encoded by said parent polynucleotide.

22. The method of claim 19, wherein said recombinogenic procedure is DNA shuffling.

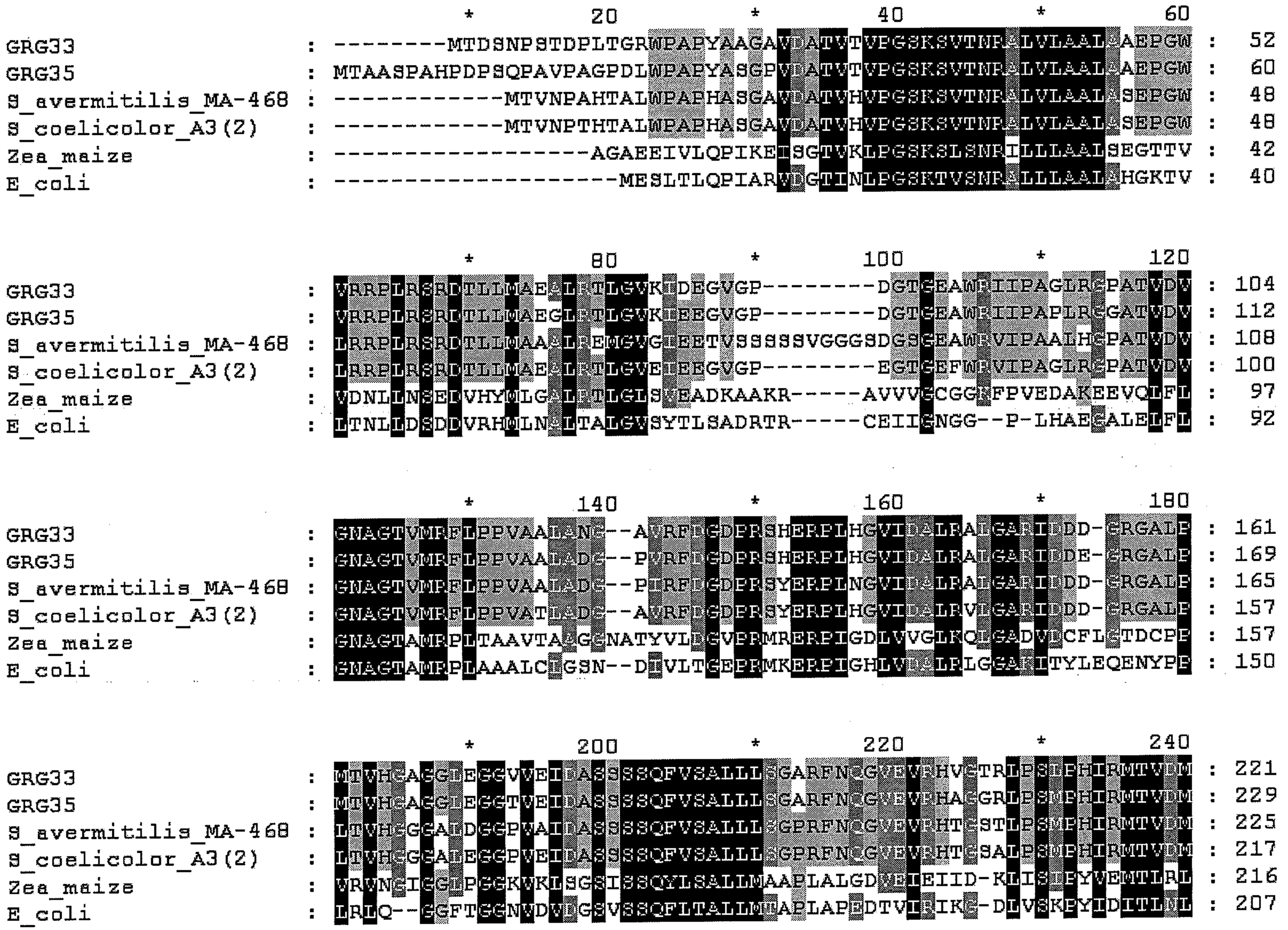

GRG33
GRG35
S_avermitilis_MA-468
S_coelicolor_A3(2)
Zea_maize
E_coli

FIG. 1A

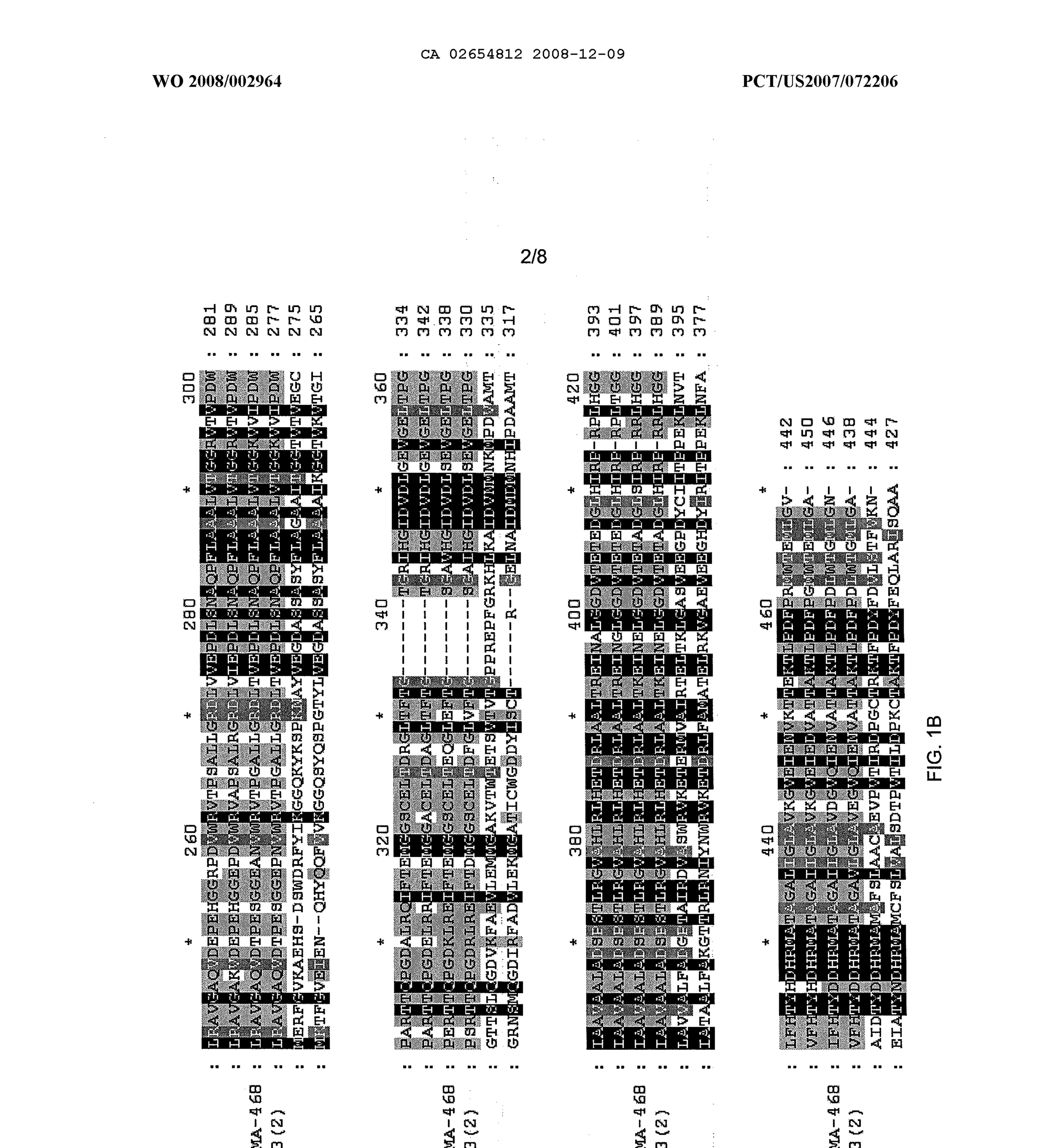


FIG. 1B

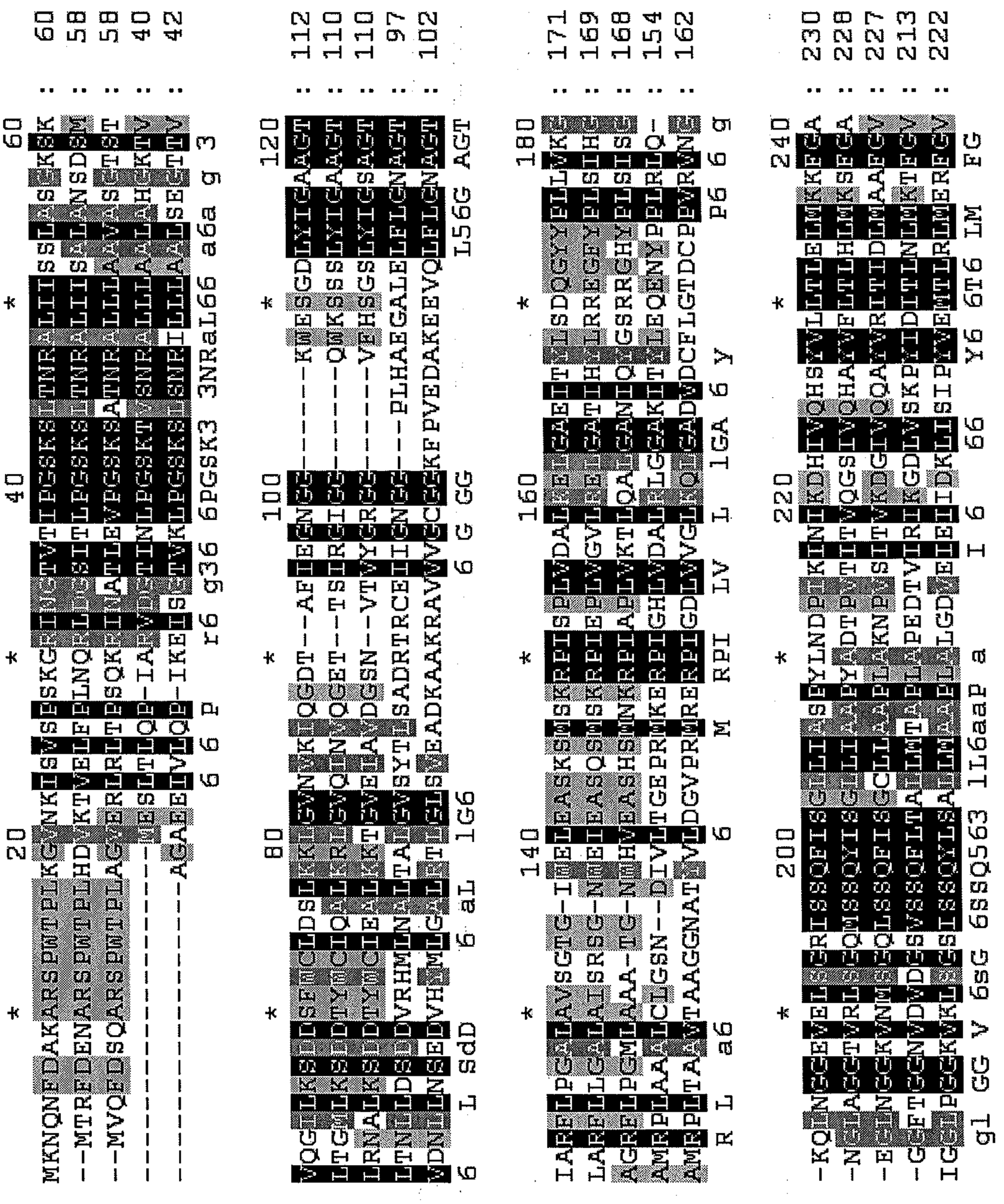


FIG. 2A

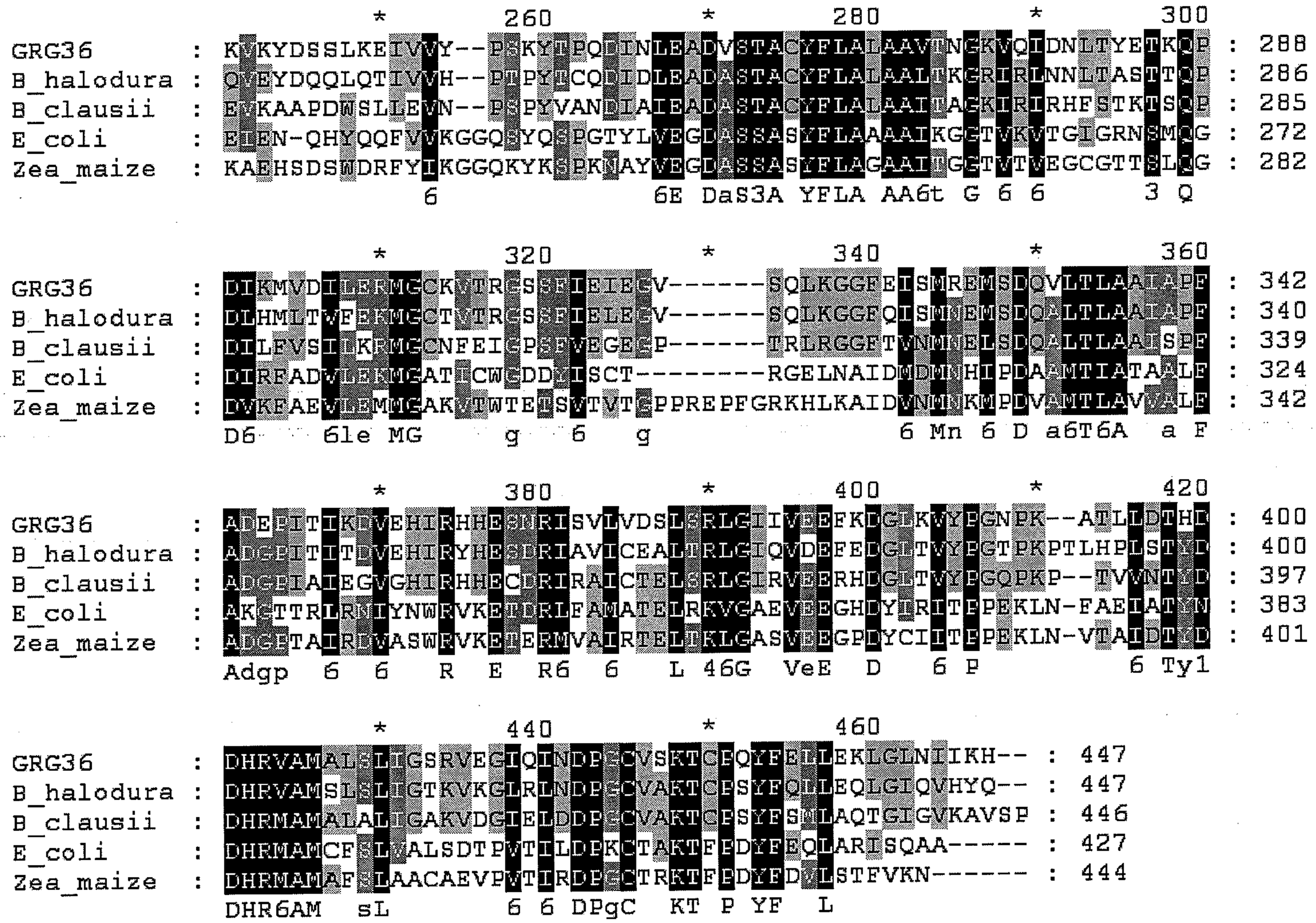

GRG36 : KVKYDSSLKEIVVY--PSKYTPQDINLEADVSTACYFLALAAVTNGKVQIDNLTYETKQP : 288
B_halodura : QVEYDQQLQTIVVH--PTPYTCQDIDLEADASTACYFLALAALTKGRIRLNNLTASTTQP : 286
B_clausii : EVKAAPDWSLLEVN--PSPYVANDIAIEADASTACYFLALAAITAGKIRIRHFSTKTSQP : 285
E_coli : EIEN-QHYQQFVVKGGQSYQSPGTYLVEGDASSASYFLAAAAIKGGTVKVTGIGRNSMQG : 272
Zea_maize : KAEHSDSWDRFYIKGGQKYKSPKNAYVEGDASSASYFLAGAAITGGTVTVEGCGTTSLQG : 282
GRG36 : DIKMVDILERMGCKVTRGSSFIEIEGV------SQLKGGFEISMREMSDQVLTLAAIAPF : 342
B_halodura : DLHMLTVFEFMGCTVTRGSSFIELEGV------SQLKGGFQISMNEMSDQALTLAAIAPF : 340
B_clausii : DILFVSILKFMGCNFEIGPSFVEGEGP------TRLRGGFTVNMNELSDQALTLAAISPF : 339
E_coli : DIRFADVLEKMGATICWGDDYISCT--------RGELNAIDMDMNHIPDAAMTIATAALF : 324
Zea_maize : DVKFAEVLEMMGAKVTWTETSVTVTGPPREPFGRKHLKAIDVNTNKMPDVAMTLAVVALF : 342
GRG36 : ADEPITIKDVEHIRHHESNRISVLVDSLSRLGIIVEEFKDGLKVYPGNPK--ATLLDTHD : 400
B_halodura : ADGPITITDVEHIRYHESDRIAVICEALTRLGIQVDEFEDGLTVYPGTPKPTLHPLSTYD : 400
B_clausii : ADGPIAIEGVGHIRHHECDRIRAICTELSRLGIRVEERHDGLTVYPGQPKP--TVVNTYD : 397
E_coli : AKGTTRLRNIYNWRVKETDRLFAMATELRKVGAEVEEGHDYIRITPPEKLN-FAEIATYN : 383
Zea_maize : ADGPTAIRDVASWRVKETERMVAIRTELTKLGASVEEGPDYCIITPPEKLN-VTAIDTYD : 401
GRG36 : DHRVAMALSLIGSRVEGIQINDPGCVSKTCPQYFELLEKLGLNIIKH-- : 447
B_halodura : DHRVAMSLSLIGTKVKGLRLNDPGCVAKTCPSYFQLLEQLGIQVHYQ-- : 447
B_clausii : DHRMAMALALIGAKVDGIELDDPGCVAKTCPSYFSMLAQTGIGVKAVSP : 446
E_coli : DHRMAMCFSLVALSDTPVTILDPKCTAKTFPDYFEQLARISQAA----- : 427
Zea_maize : DHRMAMAFSLAACAEVPVTIRDPGCTRKTFPDYFDVLSTFVKN------ : 444

FIG. 2B

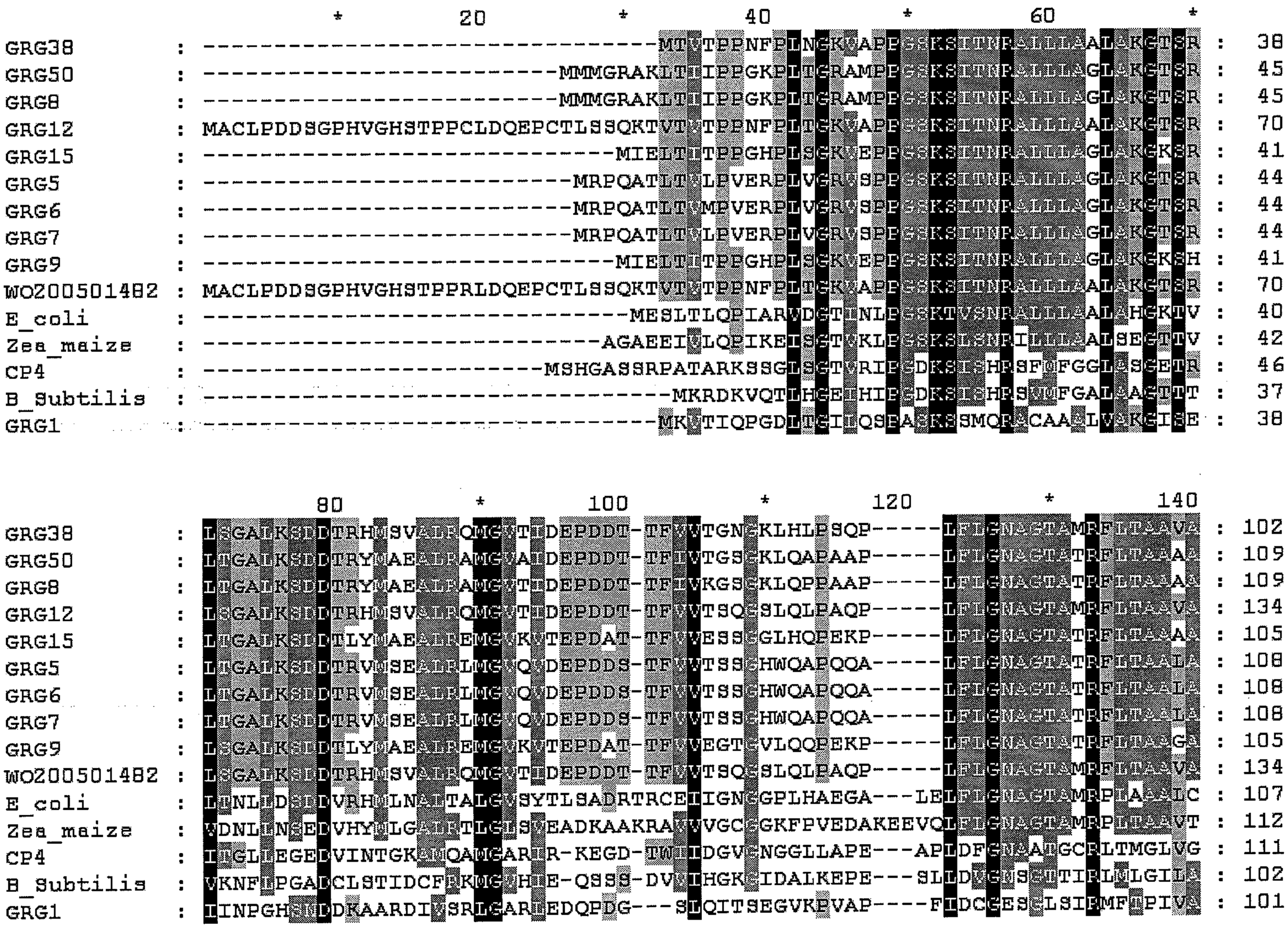

GRG38
GRG50
GRG8
GRG12
GRG15
GRG5
GRG6
GRG7
GRG9
WO200501482
E_coli
Zea_maize
CP4
B_Subtilis
GRG1

FIG. 3A

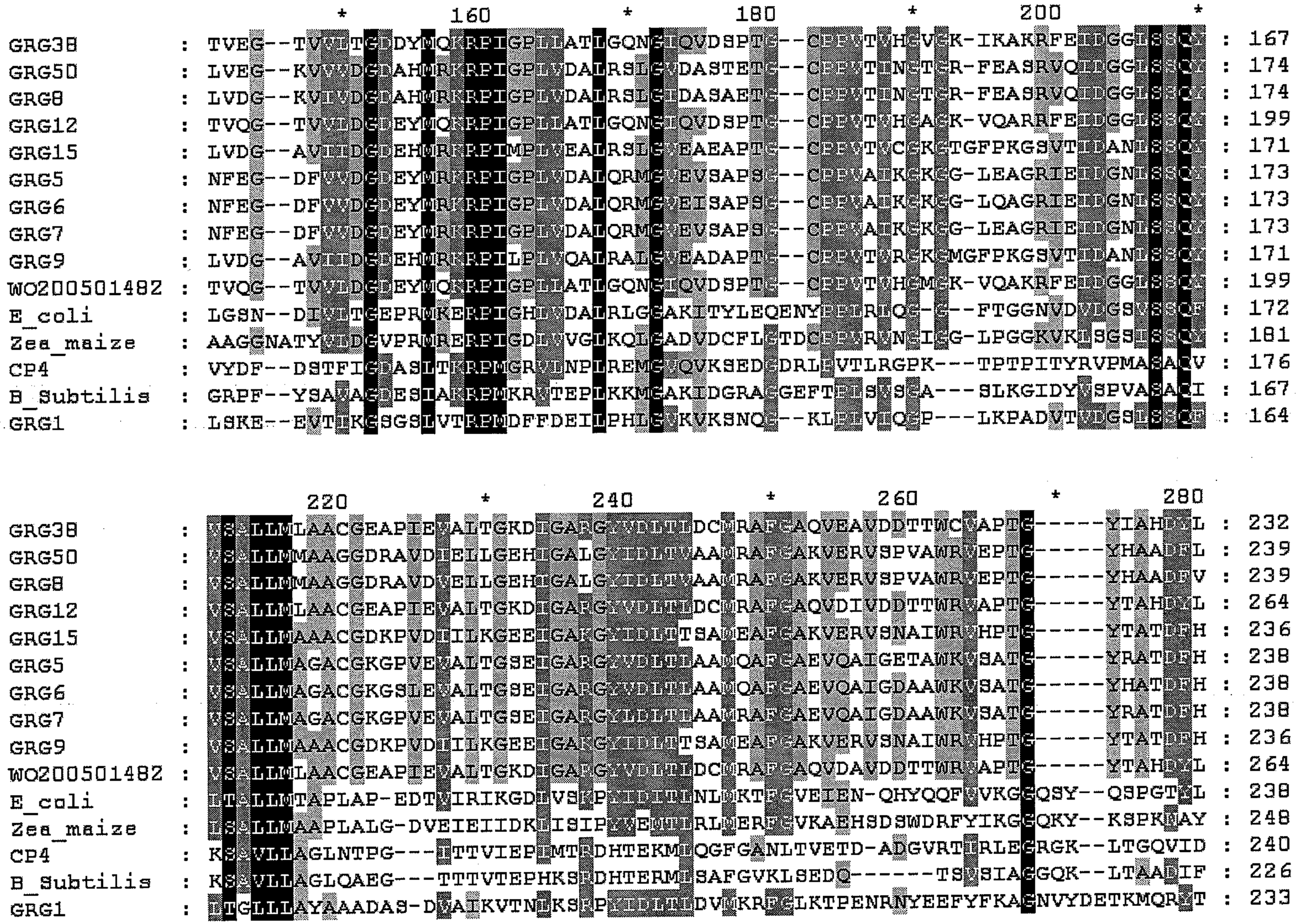

GRG38 : TVEG--TVVITGDDYMQKRPIGPLLATLGQNGIQVDSPTG--CPFVTVHCVGK-IKAKRFEIDGGLSSQY : 167
GRG50 : LVEG--KVVVDGDAHMRKRPIGPLVDALRSLGVDASTETG--CPFVTINCTGR-FEASRVQIDGGLSSQY : 174
GRG8 : LVDG--KVIVDGDAHMRKRPIGPLVDALRSLGIDASAETG--CPFVTINCTGR-FEASRVQIDGGLSSQY : 174
GRG12 : TVQG--TVVIDGDEYMQKRPIGPLLATLGQNGIQVDSPTG--CPFVTVHCAGK-VQARRFEIDGGLSSQY : 199
GRG15 : LVDG--AVIIDGDEHMRKRPIMPLVEALRSLGVEAEAPTG--CPFVTVCCKGTGFPKGSVTIDANLSSQY : 171
GRG5 : NFEG--DFVVDGDEYMRKRPIGPLVDALQRMGVEVSAPSG--CPFVAIKCKGG-LEAGRIEIDGNLSSQY : 173
GRG6 : NFEG--DFVVDGDEYMRKRPIGPLVDALQRMGVEISAPSG--CPFVAIKCKGG-LQAGRIEIDGNLSSQY : 173
GRG7 : NFEG--DFVVDGDEYMRKRPIGPLVDALQRMGVEVSAPSG--CPFVAIKCKGG-LEAGRIEIDGNLSSQY : 173
GRG9 : LVDG--AVIIDGDEHMRKRPILPLVQALRALGVEADAPTG--CPFVTVRCKGMGFPKGSVTIDANLSSQY : 171
WO200501482 : TVQG--TVVIDGDEYMQKRPIGPLLATLGQNGIQVDSPTG--CPFVTVHCMGK-VQAKRFEIDGGLSSQY : 199
E_coli : LGSN--DIVITGEPRMKERPIGHLVDALRLGGAKITYLEQENYPPIRLQG--G--FTGGNVDVIGSVSSQF : 172
Zea_maize : AAGGNATYVIDGVPRMRERPIGDLVVGLKQLGADVDCFLSTDCPPVRVNGIGG-LPGGKVKLSGSISSQY : 181
CP4 : VYDF--DSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPK---TPTPITYRVPMASAQV : 176
B_Subtilis : GRPF--YSAVAGDESIAKRPMKRVTEPLKKMGAKIDGRAGGEFTPLSVSGA---SLKGIDYVSPVASAQI : 167
GRG1 : LSKE--EVTIKGSGSLVTRPMDFFDEILPHLGVKVKSNQG--KLPIVIQGP---LKPADVTVDGSLSSQF : 164
220 240 260 280
GRG38 : VSALLMLAACGEAPIEVALTGKDIGARGYVDLTLDCMRAFGAQVEAVDDTTWCVAPTG-----YIAHDYL : 232
GRG50 : VSALLMMAAGGDRAVDIELLGEHIGALGYIDLTVAAMRAFGAKVERVSPVAWRVEPTG-----YHAADFL : 239
GRG8 : VSALLMMAAGGDRAVDVELLGEHIGALGYIDLTVAAMRAFGAKVERVSPVAWRVEPTG-----YHAADFV : 239
GRG12 : VSALLMLAACGEAPIEVALTGKDIGARGYVDLTLDCMRAFGAQVDIVDDTTWRVAPTG-----YTAHDYL : 264
GRG15 : VSALLMAAACGDKPVDIILKGEEIGARGYIDITTSAMEAFGAKVERVSNAIWRVHPTG-----YTATDFH : 236
GRG5 : VSALLMAGACGKGPVEVALTGSEIGARGYVDLTLAAMQAFGAEVQAIGETAWKVSATG-----YRATDFH : 238
GRG6 : VSALLMAGACGKGSLEVALTGSEIGARGYVDLTLAAMQAFGAEVQAIGDAAWKVSATG-----YHATDFH : 238
GRG7 : VSALLMAGACGKGPVEVALTGSEIGARGYLDLTLAAMRAFGAEVQAIGDAAWKVSATG-----YRATDFH : 238
GRG9 : VSALLMAAACGDKPVDIILKGEEIGARGYIDITTSAMEAFGAKVERVSNAIWRVHPTG-----YTATDFH : 236
WO200501482 : VSALLMLAACGEAPIEVALTGKDIGARGYVDLTLDCMRAFGAQVDAVDDTTWRVAPTG-----YTAHDYL : 264
E_coli : ITALLMTAPLAP-EDTVIRIKGDLVSKPYIDITLNLMKTFGVEIEN-QHYQQFVVKGGQSY--QSPGTYL : 238
Zea_maize : LSALLMAAPLALG-DVEIEIIDKLISIPYVEMTLRLMERFGVKAEHSDSWDRFYIKGGQKY--KSPKNAY : 248
CP4 : KSAVLLAGLNTPG---ITTVIEPIMTRDHTEKMLQGFGANLTVETD-ADGVRTIRLEGRGK--LTGQVID : 240
B_Subtilis : KSAVLLAGLQAEG---TTTVTEPHKSPDHTERMLSAFGVKLSEDQ------TSVSIAGGQK--LTAADIF : 226
GRG1 : LTGLLLAYAAADAS-DVAIKVTNLKSRPYIDLTLDVMKRFGLKTPENRNYEEFYFKAGNVYDETKMQRYT : 233

FIG. 3B

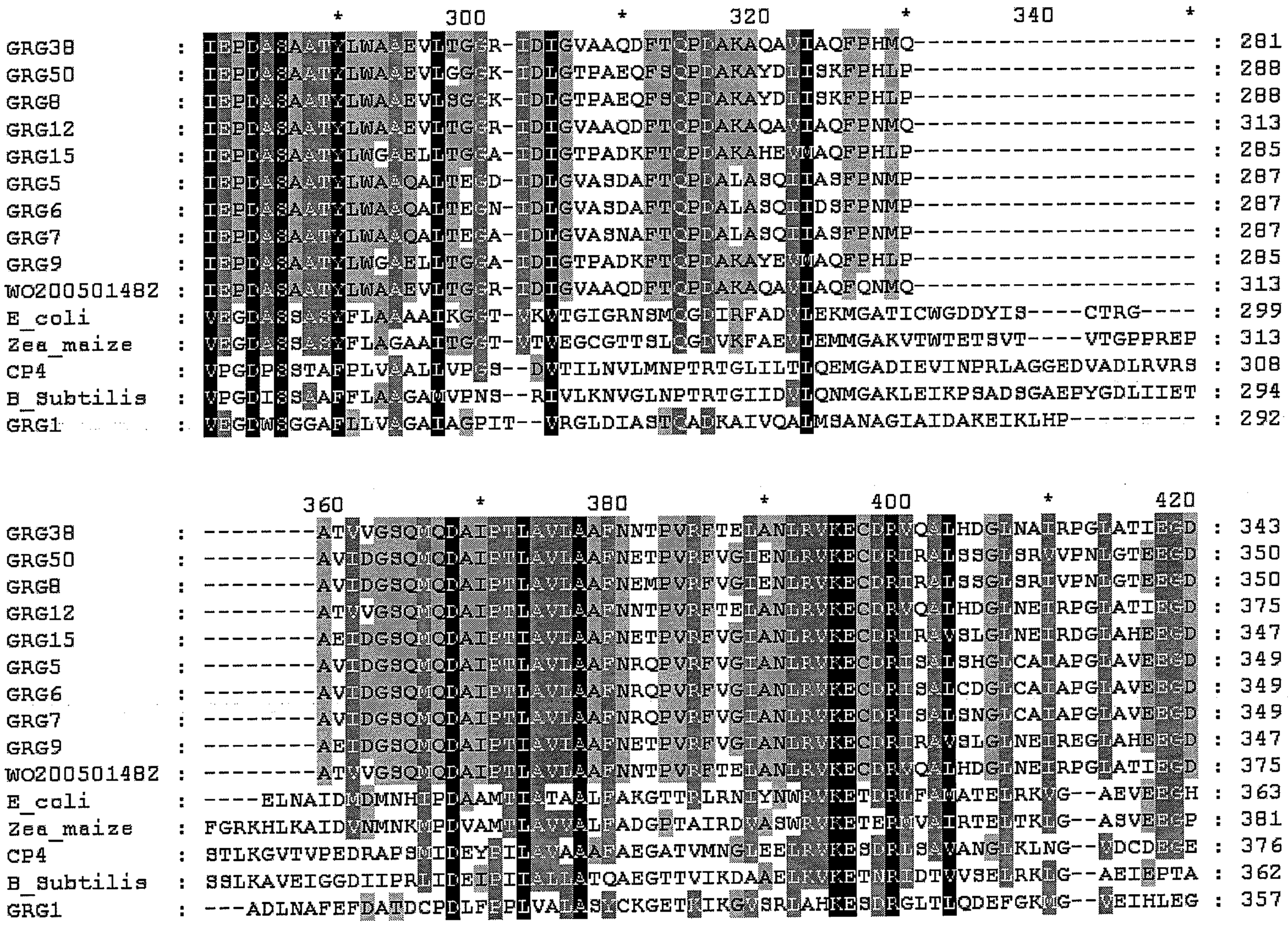

300
320
340
360
380
400
420
GRG38
GRG50
GRG8
GRG12
GRG15
GRG5
GRG6
GRG7
GRG9
WO200501482
E_coli
Zea_maize
CP4
B_Subtilis
GRG1

FIG. 3C

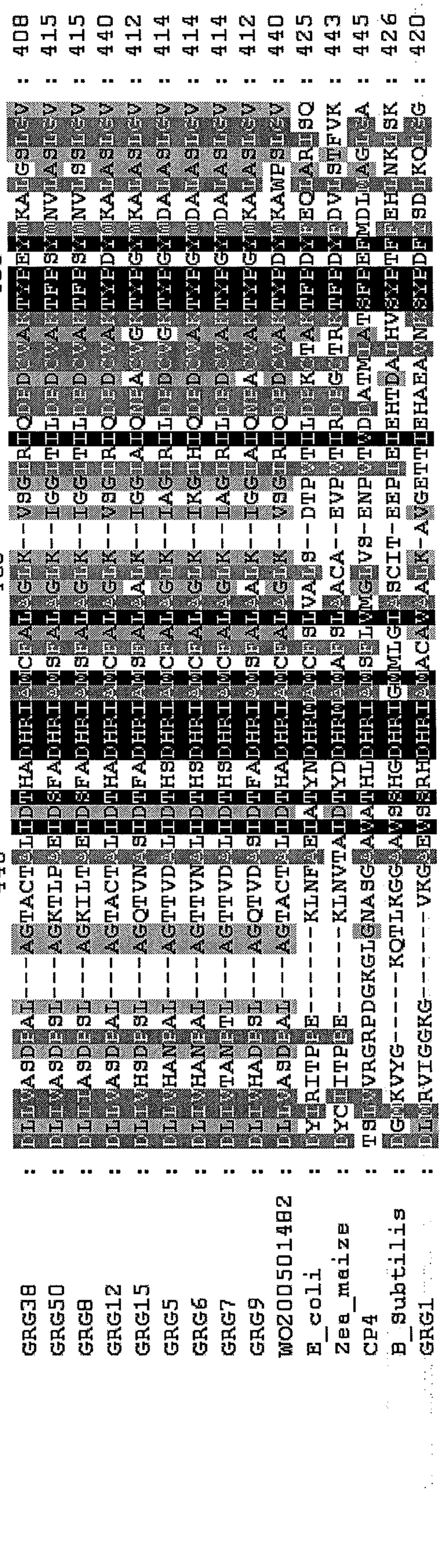

408 415 415 440 412 414 414 414 412 440 425 443 445 426 420
480
460
440
500
GRG38 : 412
GRG50 : 419
GRG8 : 419
GRG12 : 444
GRG15 : 425
GRG5 : 418
GRG6 : 418
GRG7 : 425
GRG9 : 444
WO2005014820 : 427
E_coli : 444
Zea_maize : 455
CP4 : 428
B_Subtilis : 431
GRG1 : 431

FIG. 3D