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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 September 2007 (13.09.2007)

PCT

(10) International Publication Number
WO 2007/103567 A2

(51) International Patent Classification:
C07K 14/43 (2006.01)

(21) International Application Number: PCT/US2007/006090

(22) International Filing Date: 9 March 2007 (09.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/780,946	9 March 2006 (09.03.2006)	US
60/888,634	7 February 2007 (07.02.2007)	US

(71) Applicant *(for all designated States except US)*: **E. I. DUPONT DE NEMOURS & COMPANY** [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **DAM, Thao** [US/US]; 9 Chapman Lane, Bear, Delaware 19701 (US). **GUIDA, JR., Anthony D.** [US/US]; 210 Megan Court, Newark, Delaware 19702 (US). **HAZEL, Christine B.** [US/US]; 252 Water Wheel Drive, Port Deposit, Maryland 21904 (US). **LI, Bailin** [CN/US]; 549 Stenning Drive, Hockessin, Delaware 19707-9220 (US). **WILLIAMS, Mark E.** [US/US]; 37 Country Hills Drive, Newark, Delaware 19711 (US).

(74) Agent: MOXON, Claire L.; Pioneer Hi-Bred International, Inc., 7250 N.W. 62nd Avenue, Johnston, Iowa 50131-0552 (US).

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

[Continued on next page]

(54) Title: POLYNUCLEOTIDE ENCODING A MAIZE HERBICIDE RESISTANCE GENE AND METHODS FOR USE

a

```
Symbol comparison table: blosum62.cmp CompCheck: 1102

                 GapWeight: 8
           GapLengthWeight: 2

 AAK38079_pileup_158748.txt  MSF: 750  Type: P  March 7, 2006 17:35
Check: 2632 ..

XP_469851         SEQ ID NO: 12
XP_469852         SEQ ID NO: 13
aak38080          SEQ ID NO: 5
aak38081          SEQ ID NO: 7
BAD27508          SEQ ID NO: 8
aak38079          SEQ ID NO: 6
BAD27507          SEQ ID NO: 9
BAD27506          SEQ ID NO: 10
rice_nsf_hom      SEQ ID NO: 4
xp_469850         SEQ ID NO: 3
XP_469849         SEQ ID NO: 11
nsf_peptide       SEQ ID NO: 2

//
             1                                                   50
   XP_469851 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   XP_469852 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38080 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38081 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27508 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38079 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27507 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27506 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   xp_469850 maflgwavdi ardsgasssv vltcdgygsa lyfspwdsvp lpataspddg
   XP_469849 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 nsf_peptide ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

             51                                                 100
   XP_469851 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   XP_469852 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38080 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38081 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27508 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38079 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27507 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27506 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   xp_469850 fllprfpdvc vqrsqftnhl apangtgggg srtgvkeeas evlswpptsk
   XP_469849 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 nsf_peptide ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

(57) Abstract: This invention relates to polynucleotide sequences encoding a gene that can confer resistance to at least one herbicide. It further relates to plants and seeds of plants carrying chimeric genes comprising said polynucleotide sequences, which enhance or confer resistance to at least one herbicide, and methods of making said plants and seeds. The invention further presents sequences that can be used as molecular markers that in turn can be used to identify the region of interest in corn lines resulting from new crosses and to quickly and efficiently select the best lines for breeding strategies by avoiding sensitive lines.

WO 2007/103567 A2

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:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

POLYNUCLEOTIDE ENCODING A MAIZE HERBICIDE RESISTANCE GENE AND METHODS FOR USE

Field of the Invention

This invention relates to compositions and methods useful in creating or enhancing herbicide resistance in plants. Additionally, the invention relates to plants that have been genetically transformed with the compositions of the invention.

Background of the invention

In the commercial production of crops, it is desirable to easily and quickly eliminate unwanted plants (*i.e.*, "weeds") from a field of crop plants. An ideal treatment would be one which could be applied to an entire field but which would eliminate only the unwanted plants while leaving the crop plants unharmed. One such treatment system involves the use of crop plants that are tolerant to a herbicide. When the herbicide is sprayed on a field of herbicide-tolerant crop plants, the crop plants continue to thrive while non-herbicide-tolerant weeds are killed or severely damaged.

Crop tolerance to specific herbicides can be conferred by engineering genes into crops which encode appropriate herbicide metabolizing enzymes. In some cases these enzymes, and the nucleic acids that encode them, originate in a plant. In other cases, they are derived from other organisms, such as microbes. *See*, *e.g.*, Padgette *et al.* (1996) "New weed control opportunities: Development of soybeans with a Round UP Ready™ gene" and Vasil (1996) "Phosphinothricin-resistant crops," both in *Herbicide-Resistant Crops*, ed. Duke (CRC Press, Boca Raton, Florida) pp.54-84 and pp. 85-91. Indeed, transgenic plants have been engineered to express a variety of herbicide tolerance genes from a variety of organisms, including a gene encoding a chimeric protein of rat cytochrome P4507A1 and yeast NADPH-cytochrome P450 oxidoreductase (Shiota *et al.* (1994) *Plant Physiol.* 106: 17), among other plant P450 genes (see, for example, Didierjean, L. *et al.* (2002) *Plant Physiol.* 130: 179-189; Morant, M.S. *et al.* (2003) *Opinion in Biotechnology* 14:151-162). Other genes that confer tolerance to herbicides include: acetohydroxy acid synthase ("AHAS"), which has been found

to confer resistance to multiple types of ALS herbicides on plants expressing it and has been introduced into a variety of plants (*see, e.g.*, Hattori *et al.* (1995) *Mol. Gen. Genet.* 246: 419); glutathione reductase and superoxide dismutase (Aono *et al.* (1995) *Plant Cell Physiol.* 36: 1687); and genes for various phosphotransferases (Datta *et al.* (1992) *Plant Mol. Biol.* 20: 619).

While herbicide-tolerant crop plants are presently commercially available, improvements in every aspect of crop production are continuously in demand. Herbicides and crops that are presently commercially available unfortunately have particular characteristics which can limit their usefulness in commercial crop production. Particularly, individual herbicides have different and incomplete spectra of activity against common weed species.

The acetolactate synthase, or ALS (also known as AHAS) family of herbicides control weeds by inhibiting the production of branch chain of amino acids that are essential to plant growth and development. Specifically, they bind to the plant ALS enzyme. Commonly used herbicides in this family include nicosulfuron, rimsulfuron, and chlorsulfuron, among others. Herbicides in this category can be quite crop-specific. Embodiments of the invention relate to plants that are resistant to members of the ALS-inhibiting class of herbicides, which encompasses 5 sub-classes of herbicides including, but not limited to, the sulfonylurea (SU) family of herbicides and the imidazolinone family of herbicides.

The pigment synthesis-inhibiting class of herbicides targets the enzymes that allow plants to synthesize pigments, such as carotenoid pigments or chlorophyll pigments. Loss of pigment results in photo-destruction of chlorophyll and whitening of plant tissues, which is why these herbicides are often called "bleaching" herbicides. An example of a sub-class of the bleaching herbicides is the HPPD-inhibiting class, which inhibits the 4-hydroxyphenylpyruvate dioxygenase (HPPD) enzyme (Lee *et al.* (1997) *Weed Sci.* 45:601-609). Herbicides in this family include, but are not limited to, mesotrione, tembotrione, topramezone and sulcotrione, among others. Corn is generally tolerant to mesotrione due to metabolism of the herbicide (Mitchell *et al.* (2001) *Pest Mgt. Sci.* 57:120-128). The same detoxification system may give tolerance to both mesotrione and some SU herbicides (Green & Williams (2004) *Proceedings Weed Science Society of America* 44:13). Embodiments of the invention relate to plants

that are resistant to members of the pigment synthesis-inhibiting class of herbicides.

The protoporphyrinogen oxidase (PPO)-inhibiting class of herbicides interferes with the synthesis of chlorophyll, resulting in compounds that produce highly active compounds (free-radicals). These reactive compounds disrupt cell membranes which results in the leaf burning associated with post-emergence applications of these products. Herbicides in this family include, but are not limited to, acifluorfen, fomesafen, lactofen, sulfentrazone, carfentrazone, flumiclorac and flumioxazin, among others. Embodiments of the invention relate to plants that are resistant to members of the PPO-inhibiting class of herbicides.

Photosystem II (PSII)-inhibiting herbicides have a mode of action that involves interaction with components in the electron transfer chain of Photosystem II. Photosynthesis requires the transfer of electrons from Photosystem II to Photosystem I. A key step in this electron transfer chain is the reduction of plastoquinone (PQ) by the D_1 protein in the thylakoid membrane. PSII-inhibitor herbicides bind to the D_1 protein, thus inhibiting PQ binding and interrupting the electron transfer process. This results in the plants not being able to fix carbon dioxide and produce the carbohydrates needed for the plant to survive. The block in electron transfer also causes an oxidative stress and the generation of radicals which cause rapid cellular damage. PSII-inhibiting herbicides are represented by several herbicide families, including the symmetrical triazines, triazinones (asymmetrical triazines), substituted ureas,uracils, pyridazinones, phenyl carbamates, nitriles, benzothiadiazoles, phenyl pyridazines,and acid amides. Embodiments of the invention relate to plants that are resistant to members of the PS II-inhibiting class of herbicides.

Synthetic auxin herbicides are a widely used class of herbicides that mimic the natural auxin hormones produced by plants. Auxins regulate many plant processes, including cell growth and differentiation. Auxins are generally present at low concentrations in the plant. Synthetic auxin herbicides mimic natural auxins and cause relatively high concentrations in the cell that result in a rapid growth response. Susceptible plants treated with these herbicides exhibit symptoms that could be called ‘auxin overdose’, and eventually die as a result of increased rates of disorganized and uncontrolled growth. Embodiments of the invention relate to plants that are resistant to members of the synthetic auxin class of herbicides.

Some embodiments of this invention are based on the fine mapping, cloning and characterization of the gene responsible for an important herbicide resistance mechanism in maize.

It has been known since the early 1990s that natural tolerance in maize (*Zea mays* L.) to a subset of sulfonylurea herbicides (nicosulfuron [Dupont Accent® herbicide], rimsulfuron, primisulfuron, and thifensulfuron) is controlled by a single gene (named *nsf* by Kang (1993) *Journal of Heredity* 84(3): 216-217), with resistance dominant and sensitivity recessive (Harms *et al.* (1990) *Theor. Appl. Genet.* 80:353-358; Kang (1993) *supra*; Green & Uhlrich (1993) *Weed Sci.* 41:508-516; Green & Uhlrich (1994) *Pestic. Sci.* 40:187-191). It is also known that tolerant maize plants metabolize nicosulfuron by hydroxylation, with the characteristics of a cytochrome P450 (Fonne-Pfister *et al.* (1990) *Pesticide Biochem. Physiol.* 37:165-173; Brown & Cotterman (1994) *Chem. Plant Prot.* 10:47-81). It has been suggested that the same corn gene responsible for determining tolerance to some sulfonylurea herbicides is also responsible for the tolerance to bentazon (Barrett *et al.* (1997) Role of cytochrome P-450 in herbicide metabolism and selectivity and multiple herbicide metabolizing cytochrome P-450 activities in maize. *In* K. K. Hatzios, ed. Regulation of Enzymatic Systems Detoxifying Xenobiotics in Plants. Dordrecht: Kluwer Academic. pp. 35-50; Green (1998) *Weed Technology* 12:474-477) and HPPD inhibitor herbicides such as mesotrione (Green & Williams (2004) *supra*; Williams *et al.* (2005) *HortScience* 40(6):1801-1805). Recent advances in the development of the maize physical map and integrated markers (Bortiri *et al.* (2006) *Curr Opin Plant Biol.* 9(2):164-71) has allowed a positional cloning approach to be used for identifying the *Nsf1* locus.

The *Nsf1* resistance gene of the embodiments of the present invention encodes a novel gene related to the cytochrome P450 family. While multiple cytochrome P450 genes have been described, they differ widely in their response to different pathogens and exact action. The novel cytochrome P450 gene described in this disclosure has been demonstrated to provide improved tolerance or resistance to numerous herbicides, including nicosulfuron, rimsulfuron, primisulfuron, thifensulfuron and mesotrione.

Summary of the Invention

The present invention is directed to embodiments including an isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide capable of conferring resistance to at least one herbicide, wherein the polypeptide has an amino acid sequence of at least 85, 90 or 95% identity, when compared to SEQ ID NO:1 based on the Needleman-Wunsch alignment algorithm, or a complement of the nucleotide sequence, wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary. The herbicides to which the polynucleotide of the embodiments imparts resistance include members of the ALS-inhibiting class; the pigment synthesis-inhibiting class; the PPO-inhibiting class; the PS II-inhibiting class; and the synthetic auxin class of herbicides. The polynucleotide of the embodiments may impart resistance to one or more herbicides from the same class, or from different classes, including representative members from all 5 classes.

Additional embodiments of the present invention include a vector comprising the polynucleotide of the embodiments and a recombinant DNA construct comprising the polynucleotide of the embodiments, operably linked to at least one regulatory sequence. A plant cell, as well as a plant and a seed each comprising the recombinant DNA construct of an embodiment of the present invention are also encompassed. Also included are plants comprising additional polynucleotides encoding polypeptides responsible for traits of interest, such as polypeptides having glyphosate N-acetyltransferase activity, insecticidal Bt polypeptides, and other polypeptides of interest. Plants comprising these polynucleotides include monocots and dicots, including, but not limited to, maize, wheat, barley, oats, switchgrass, sorghum, rice, soybean, canola, potato, cotton, and sunflower.

The methods embodied by the present invention include 1) a method for transforming a cell, comprising transforming a cell with the polynucleotide of an embodiment of the present invention, 2) a method for producing a plant comprising transforming a plant cell with the recombinant DNA construct of an embodiment of the present invention and regenerating a plant from the transformed plant cell, and 3) methods of conferring or enhancing resistance to at least one herbicide, comprising transforming a plant with the recombinant DNA

construct of an embodiment of the present invention, thereby conferring or enhancing resistance to at least one herbicide, such as a member of the ALS-inhibiting class; the pigment synthesis-inhibiting class; the PPO-inhibiting class; the PS II-inhibiting class; and the synthetic auxin class of herbicides.

5 In addition, an embodiment of the invention is a variant allele of the *Nsf1* sequence in which a specific single amino acid change (see Example 2) renders the gene inoperative, resulting in sensitivity to at least one ALS or HPPD inhibitor herbicide to which most corn is resistant. Accordingly, an additional method embodied by the present invention is a method of using the variant of the *Nsf1*

10 gene as a marker in breeding strategies to avoid incorporating the sensitive allele.

Methods of altering the level of expression of a protein capable of conferring resistance to at least one herbicide in a plant cell comprising (a) transforming a plant cell with the recombinant DNA construct of an embodiment of the present invention and (b) growing the transformed plant cell under conditions

15 that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of a protein capable of conferring resistance to at least one herbicide in the transformed host are also embodied by the present invention. The herbicides for which resistance may be conferred include, for example, members of the ALS-

20 inhibiting class; the pigment synthesis-inhibiting class; the PPO-inhibiting class; the PS II-inhibiting class; and the synthetic auxin class of herbicides.

Herbicides to which a polynucleotide of the embodiments may confer or enhance resistance include, but are not limited to, herbicides selected from the ALS-inhibiting class of herbicides such as nicosulfuron, rimsulfuron, primisulfuron,

25 imazethapyr, chlorsulfuron, chlorimuron ethyl, triasulfuron, flumetsulam and imazaquin. Additionally, such herbicides may be selected from the pigment synthesis-inhibiting class of herbicides, such as isoxaflutole, topramezone, sulcatrione and tembotrione. Such herbicides may also be selected from the PPO-inhibiting class of herbicides, such as acifluorfen, flumioxan and

30 sulfentrazone. Optionally, such herbicides may be selected from the PS II-inhibiting class of herbicides, such as diuron, linuron, bentazon and chlorotoluron. Such herbicides may also be selected from the synthetic auxin class of herbicides, such as dicamba.

Methods of the embodiments include a method of determining the presence of the polynucleotide of the embodiments or the *Nsf1* locus in a plant, comprising at least one of: (a) isolating nucleic acid molecules from the plant and determining if an *Nsf1* gene is present by attempting to amplify sequences homologous to the polynucleotide; or (b) isolating nucleic acid molecules from the plant and performing a Southern hybridization, or (c) isolating proteins from the plant and performing a western blot using antibodies to the NSF1 protein, or (d) isolating proteins from the plant and performing an ELISA assay using antibodies to the NSF1 protein, thereby determining the presence of the polynucleotide of Claim 1 in the plant.

Also encompassed by the embodiments are plants with enhanced tolerance to at least one herbicide, comprising the *Nsf1* gene in a recombinant DNA construct. Such plants further comprise a second herbicide resistance gene providing a certain level of tolerance to a herbicide selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
(b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

such that the presence of the *Nsf1* gene confers upon the plant a higher level of tolerance to the same herbicide than the tolerance level exhibited by a plant comprising the second herbicide resistance gene but not comprising the *Nsf1* gene.

Also encompassed by the embodiments are soybean plants comprising the *Nsf1* gene, wherein such soybean plants also exhibit soybean cyst nematode resistance. Such plants may have been created through transformation or plant breeding techniques, and may have been bred from germplasm such as those selected from the group consisting of, Peking, PI88788, PI89772, PI90763, PI209332, PI404189A, PI437654, PI438489B, PI467312, PI468916, Hartwig, J87-233, and progeny derived from any of the listed sources.

Brief Description of the Drawings

Figure 1(a-e) is a multiple sequence alignment of the polypeptide sequence of the embodiments (SEQ ID NO: 2) comparing it to other known Cytochrome P450 polypeptides (SEQ ID NOs: 3-13). Figure 1d also indicates the position of

the most commonly conserved domain of the cytochrome P450 family (SEQ ID NO: 14). Identical residues in the alignment are indicated in upper case letters.

Figure 2(a-b) is a multiple sequence alignment of the polypeptide sequences of several sensitive and resistant corn lines showing the commonly conserved domain of the cytochrome P450 family (SEQ ID NO: 14) as well as variations among the sequences.

Detailed Description of the Invention

Embodiments of the present invention provide compositions and methods directed to inducing herbicide resistance in plants. The compositions are novel nucleotide and amino acid sequences that confer or enhance resistance to one or more members of one or more classes of herbicides, including the ALS-inhibiting, PPO-inhibiting, pigment synthesis-inhibiting, PS II-inhibiting and synthetic auxin herbicide classes, whose members include, but are not limited to, nicosulfuron, rimsulfuron, primisulfuron, and mesotrione. Specifically, certain embodiments provide polypeptides having the amino acid sequence set forth in SEQ ID NO: 2, and variants and fragments thereof. Isolated nucleic acid molecules, and variants and fragments thereof, comprising nucleotide sequences that encode the amino acid sequence shown in SEQ ID NO: 2 are further provided.

One example of the native nucleotide sequence that encodes the polypeptide of SEQ ID NO: 2 is set forth in SEQ ID NO: 1. Plants, plant cells, seeds, and microorganisms comprising a nucleotide sequence that encodes a polypeptide of the embodiments are also disclosed herein.

The full length polypeptide of the embodiments (SEQ ID NO: 2) shares varying degrees of homology with known polypeptides of the cytochrome P450 family. In particular, the novel polypeptide of the embodiments shares homology with cytochrome P450 proteins isolated from *Oryza sativa*: Accession Nos. XP_469850 (SEQ ID NO: 3), ABC69856 (SEQ ID NO: 4); XP_469849 (SEQ ID NO: 11) and XP_469851 (SEQ ID NO: 12); and XP_469852 (SEQ ID NO: 13) and *Lolium rigidum*: Accession Nos. AAK38080 (SEQ ID NO: 5); AAK38079 (SEQ ID NO: 6); AAK38081 (SEQ ID NO: 7); BAD27508 (SEQ ID NO: 8); BAD27507 (SEQ ID NO: 9) and BAD27506 (SEQ ID NO: 10). Figure 1 provides an alignment of the

amino acid sequence set forth in SEQ ID NO: 2 with the *O. sativa* and *L. rigidum* cytochrome P450 proteins (SEQ ID NOs: 3-13).

Amino acid alignments performed using the GAP program indicate that SEQ ID NO:2 shares the sequence similarities shown in Table 1 with the *O. sativa*
5 and *L. rigidum* cytochrome P450 proteins.

Table 1: Comparison of NSF1 Peptide to other Cytochrome P450 peptides

Other Cytochrome P450 Protein	Percent Identity	Percent Similarity
XP_469850 (SEQ ID NO: 3)	67%	76%
ABC69856 (SEQ ID NO: 4)	67%	76%
AAK38080 (SEQ ID NO: 5)	68%	76%
AAK38079 (SEQ ID NO: 6)	67%	77%
AAK38081 (SEQ ID NO: 7)	67%	76%
BAD27508 (SEQ ID NO: 8)	67%	76%
Other Cytochrome P450 Protein	**Percent Identity**	**Percent Similarity**
BAD27507 (SEQ ID NO: 9)	67%	76%
BAD27506 (SEQ ID NO: 10)	67%	76%
XP_469849 (SEQ ID NO: 11)	66%	75%
XP_469851 (SEQ ID NO: 12)	61%	71%
XP_469852 (SEQ ID NO: 13)	60%	72%

10 The cytochrome P450 family of genes in plants catalyze extremely diverse and often complex regiospecific and/or stereospecific reactions in the biosynthesis or catabolism of plant bioactive molecules. (Morant *et al.* (2003) *Curr. Opin. Biotech.* 14(2): 151-162). P450s are heme proteins that catalyze the activation of molecular oxygen by using electrons from NADPH. In the *Arabidopsis thaliana*
15 genome alone, there are an estimated over 300 cytochromes P450 (Werck-Reichhart *et al.* (2000) *Trends in Plant Science* 5(3): 116-123). Common structural features occur in plant cytochromes P450 and help identify them as such. These features include the F-X-X-G-X-R-X-C-X-G (SEQ ID NO: 14) motif generally found near the C-terminus (see Figure 1d). About 150 residues
20 upstream, another conserved motif generally found follows the A/G-G-X-D/E-T-T/S

(SEQ ID NO: 15) motif and corresponds to the region of the peptide responsible for oxygen-binding and activation.

The nucleic acids and polypeptides of the embodiments find use in methods for conferring or enhancing herbicide resistance to a plant. Accordingly, the compositions and methods disclosed herein are useful in protecting plants from damage caused by herbicides. "Herbicide resistance" is intended to mean that a plant or plant cell has the ability to tolerate a higher concentration of a herbicide than plants or cells which are not resistant, or to tolerate a certain concentration of a herbicide for a longer time than cells or plants which are not resistant. That is, herbicides are prevented from causing plant injury, or the injury caused by the herbicide is minimized or lessened, such as, for example, the reduction of leaf yellowing and associated yield loss. One of skill in the art will appreciate that the compositions and methods disclosed herein can be used with other compositions and methods available in the art for increasing or enhancing plant herbicide resistance. The term "enhance" refers to improve, increase, amplify, multiply, elevate, raise, and the like.

In particular aspects, the embodiments include methods for conferring or enhancing herbicide resistance in a plant comprising introducing into a plant at least one DNA construct, wherein the DNA construct comprises a nucleotide sequence encoding a herbicide resistance polypeptide of the embodiments operably linked to a promoter that drives expression in the plant. The plant expresses the polypeptide, thereby conferring or enhancing herbicide resistance upon the plant, or improving the plant's inherent level of resistance. In particular embodiments, the gene confers or enhances resistance to at least one herbicide of the ALS-inhibiting, pigment synthesis-inhibiting, PPO-inhibiting, PS II-inhibiting or synthetic auxin herbicide classes, whose members include, but are not limited to, the herbicides nicosulfuron, rimsulfuron, primisulfuron, thifensulfuron, bentazon, and mesotrione.

Expression of a polypeptide of the embodiments may be targeted to specific plant tissues, but generally in the case of herbicide resistance, continuous expression is desired throughout the cells of a plant. Therefore, while many promoters could be used in the embodiments of the invention, generally constitutive promoters are utilized. A constitutive promoter is a promoter that

directs expression of a gene throughout the various parts of a plant and continuously throughout plant development.

As used herein, "nucleic acid" includes reference to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, encompasses known analogues (*e.g.*, peptide nucleic acids) having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides.

The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residues is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. Polypeptides of the embodiments can be produced either from a nucleic acid disclosed herein, or by the use of standard molecular biology techniques. For example, a truncated protein of the embodiments can be produced by expression of a recombinant nucleic acid of the embodiments in an appropriate host cell, or alternatively by a combination of *ex vivo* procedures, such as protease digestion and purification.

As used herein, the terms "encoding" or "encoded" when used in the context of a specified nucleic acid mean that the nucleic acid comprises the requisite information to direct translation of the nucleotide sequence into a specified protein. The information by which a protein is encoded is specified by the use of codons. A nucleic acid encoding a protein may comprise non-translated sequences (*e.g.*, introns) within translated regions of the nucleic acid or may lack such intervening non-translated sequences (*e.g.*, as in cDNA).

The embodiments of the invention encompass isolated or substantially purified polynucleotide or protein compositions. An "isolated" or "purified" polynucleotide or protein, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the polynucleotide or protein as found in its naturally occurring environment. Thus, an isolated or purified polynucleotide or protein 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. Optimally, an "isolated" polynucleotide is free of sequences (optimally protein encoding sequences) that naturally flank the polynucleotide (i.e.,

sequences located at the 5' and 3' ends of the polynucleotide) in the genomic DNA of the organism from which the polynucleotide is derived. For example, in various embodiments, the isolated polynucleotide can contain less than about 5 kb, about 4 kb, about 3 kb, about 2 kb, about 1 kb, about 0.5 kb, or about 0.1 kb of nucleotide sequence that naturally flank the polynucleotide in genomic DNA of the cell from which the polynucleotide is derived. A protein that is substantially free of cellular material includes preparations of protein having less than about 30%, about 20%, about 10%, about 5%, or about 1% (by dry weight) of contaminating protein. When the protein of the embodiments, or a biologically active portion thereof, is recombinantly produced, optimally culture medium represents less than about 30%, about 20%, about 10%, about 5%, or about 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

Fragments and variants of the disclosed nucleotide sequences and proteins encoded thereby are also encompassed by the embodiments. By "fragment" is intended a portion of the nucleotide sequence or a portion of the amino acid sequence and hence protein encoded thereby. Fragments of a nucleotide sequence may encode protein fragments that retain the biological activity of the native protein and hence have the ability to confer or enhance resistance to at least one herbicide of the ALS-inhibiting, PPO-inhibiting, pigment synthesis-inhibiting, PS II-inhibiting or synthetic auxin herbicide class. Alternatively, fragments of a nucleotide sequence that are useful as hybridization probes do not necessarily encode fragment proteins retaining biological activity. Thus, fragments of a nucleotide sequence may range from at least about 15 nucleotides, about 50 nucleotides, about 100 nucleotides, and up to the full-length nucleotide sequence encoding the polypeptides of the embodiments.

A fragment of a nucleotide sequence that encodes a biologically active portion of a polypeptide of the embodiments will encode at least about 15, about 25, about 30, about 40, or about 50 contiguous amino acids, or up to the total number of amino acids present in a full-length polypeptide of the embodiments (for example, 521 amino acids for SEQ ID NO: 2). Fragments of a nucleotide sequence that are useful as hybridization probes or PCR primers generally need not encode a biologically active portion of a protein.

As used herein, "full-length sequence" in reference to a specified polynucleotide means having the entire nucleic acid sequence of a native

sequence. By "native sequence" is intended an endogenous sequence, *i.e.*, a non-engineered sequence found in an organism's genome.

Thus, a fragment of a nucleotide sequence of the embodiments may encode a biologically active portion of a polypeptide, or it may be a fragment that can be used as a hybridization probe or PCR primer using methods disclosed below. A biologically active portion of an herbicide resistance polypeptide can be prepared by isolating a portion of one of the nucleotide sequences of the embodiments, expressing the encoded portion of the protein and assessing the ability of the encoded portion of the protein to confer or enhance herbicide resistance in a plant. Nucleic acid molecules that are fragments of a nucleotide sequence of the embodiments comprise at least about 15, about 20, about 50, about 75, about 100, or about 150 nucleotides, or up to the number of nucleotides present in a full-length nucleotide sequence disclosed herein (for example, 1563 nucleotides for SEQ ID NO: 1).

"Variants" is intended to mean substantially similar sequences. For polynucleotides, a variant comprises a deletion and/or addition of one or more nucleotides at one or more internal sites within the native polynucleotide and/or a substitution of one or more nucleotides at one or more sites in the native polynucleotide. As used herein, a "native" polynucleotide or polypeptide comprises a naturally occurring nucleotide sequence or amino acid sequence, respectively. One of skill in the art will recognize that variants of the nucleic acids of the embodiments will be constructed such that the open reading frame is maintained. For polynucleotides, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the amino acid sequence of one of the polypeptides of the embodiments. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant polynucleotides also include synthetically derived polynucleotide, such as those generated, for example, by using site-directed mutagenesis but which still encode a protein of the embodiments. Generally, variants of a particular polynucleotide of the embodiments will have at least about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%,

about 97%, about 98%, about 99% or more sequence identity to that particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

Variants of a particular polynucleotide of the embodiments (i.e., the reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by a variant polynucleotide and the polypeptide encoded by the reference polynucleotide. Thus, for example, isolated polynucleotides that encode a polypeptide with a given percent sequence identity to the polypeptide of SEQ ID NO: 2 are disclosed. Percent sequence identity between any two polypeptides can be calculated using sequence alignment programs and parameters described elsewhere herein. Where any given pair of polynucleotides of the embodiments is evaluated by comparison of the percent sequence identity shared by the two polypeptides they encode, the percent sequence identity between the two encoded polypeptides is at least about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99% or more sequence identity.

"Variant" protein is intended to mean a protein derived from the native protein by deletion or addition of one or more amino acids at one or more internal sites in the native protein and/or substitution of one or more amino acids at one or more sites in the native protein. Variant proteins encompassed by the embodiments are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, the ability to confer or enhance plant herbicide resistance as described herein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Biologically active variants of a native protein of the embodiments will have at least about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99% or more sequence identity to the amino acid sequence for the native protein as determined by sequence alignment programs and parameters described elsewhere herein. A biologically active variant of a protein of the embodiments may differ from that

protein by as few as about 1-15 amino acid residues, as few as about 1-10, such as about 6-10, as few as about 5, as few as 4, 3, 2, or even 1 amino acid residue.

The proteins of the embodiments may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants and fragments of the herbicide resistance proteins can be prepared by mutations in the DNA. Methods for mutagenesis and polynucleotide alterations are well known in the art. See, for example, Kunkel (1985) *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel *et al.* (1987) *Methods in Enzymol.* 154:367-382; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.* (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be optimal.

Thus, the genes and polynucleotides of the embodiments include both the naturally occurring sequences as well as mutant forms. Likewise, the proteins of the embodiments encompass both naturally occurring proteins as well as variations and modified forms thereof. Such variants will continue to possess the desired ability to confer or enhance plant resistance to at least one herbicide of the ALS-inhibiting, PPO-inhibiting, pigment synthesis-inhibiting, PS II-inhibiting or synthetic auxin herbicide classes. Obviously, the mutations that will be made in the DNA encoding the variant must not place the sequence out of reading frame and optimally will not create complementary regions that could produce secondary mRNA structure. See, EP Patent No. 0075444.

The deletions, insertions, and substitutions of the protein sequences encompassed herein are not expected to produce radical changes in the characteristics of the protein. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by screening transgenic plants which have been transformed with the variant protein to ascertain the effect on the herbicide resistance characteristics of the plant.

Variant polynucleotides and proteins also encompass sequences and proteins derived from mutagenic or recombinogenic procedures, including and not limited to procedures such as DNA shuffling. One of skill in the art could envision modifications that would alter the range of herbicides to which the protein
5 responds. With such a procedure, one or more different protein coding sequences can be manipulated to create a new 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*
10 or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the protein gene of the embodiments and other known protein genes to obtain a new gene coding for a protein with an improved property of interest, such as increased ability to confer or enhance plant herbicide resistance. Strategies for such DNA shuffling are known in the art. See,
15 for example, US 2002/0058249; 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.
20 The polynucleotides of the embodiments can be used to isolate corresponding sequences from other organisms, particularly other plants. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences based on their sequence homology to the sequences set forth herein. Sequences isolated based on their sequence identity to the entire
25 sequences set forth herein or to variants and fragments thereof are encompassed by the embodiments. Such sequences include sequences that are orthologs of the disclosed sequences. "Orthologs" is intended to mean genes derived from a common ancestral gene and which are found in different species as a result of speciation. Genes found in different species are considered orthologs when their
30 nucleotide sequences and/or their encoded protein sequences share at least about 60%, about 70%, about 75%, about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or greater sequence identity. Functions of orthologs are often highly conserved among species. Thus, isolated polynucleotides that

encode for a protein that confers or enhances plant herbicide resistance and that hybridize under stringent conditions to the sequences disclosed herein, or to variants or fragments thereof, are encompassed by the embodiments.

In a PCR approach, oligonucleotide primers can be designed for use in PCR reactions to amplify corresponding DNA sequences from cDNA or genomic DNA extracted from any organism of interest. Methods for designing PCR primers and PCR cloning are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York). See also Innis *et al.*, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, New York). Known methods of PCR include, and are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like.

In hybridization techniques, all or part of a known polynucleotide is used as a probe that selectively hybridizes to other corresponding polynucleotides present in a population of cloned genomic DNA fragments or cDNA fragments (i.e., genomic or cDNA libraries) from a chosen organism. The 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. Thus, for example, probes for hybridization can be made by labeling synthetic oligonucleotides based on the polynucleotides of the embodiments. Methods for preparation of probes for hybridization and for construction of cDNA and genomic libraries are generally known in the art and are disclosed in Sambrook *et al.* (1989) *supra*.

For example, an entire polynucleotide disclosed herein, or one or more portions thereof, may be used as a probe capable of specifically hybridizing to corresponding polynucleotides and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique and are optimally at least about 10 nucleotides in length, at least about 15 nucleotides in length, or at least about 20 nucleotides in length. Such probes may be used to amplify corresponding polynucleotides 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,
5 Sambrook *et al.* (1989) *supra*.

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
10 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
15 similarity are detected (heterologous probing). Generally, a probe is less than about 1000 nucleotides in length, optimally 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
20 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 sulphate) at 37 °C, and a wash in
25 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 final wash in 0.1X SSC at 60 to
30 65 °C for at least 30 minutes. 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. The duration of the wash time will be at least a length of time sufficient to reach equilibrium.

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 thermal melting point (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 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 T_m; moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10 °C lower than the T_m; low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20 °C lower than the 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 optimal 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) *supra*.

Various procedures can be used to check for the presence or absence of a particular sequence of DNA, RNA, or a protein. These include, for example, Southern blots, northern blots, western blots, and ELISA analysis. Techniques such as these are well known to those of skill in the art and many references exist

which provide detailed protocols. Such references include Sambrook *et al.* (1989) *supra,* and Crowther, J.R. (2001), *The ELISA Guidebook*, Humana Press, Totowa, NJ, USA.

The following terms are used to describe the sequence relationships between two or more polynucleotides or polypeptides: (a) "reference sequence," (b) "comparison window," (c) "sequence identity," and, (d) "percentage of sequence identity."

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two polynucleotides. Generally, the comparison window is at least about 20 contiguous nucleotides in length, and optionally can be about 30, about 40, about 50, about 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent sequence identity between any two sequences can be accomplished using a mathematical algorithm. Non-limiting examples of such mathematical algorithms are the algorithm of Myers and Miller (1988) *CABIOS* 4:11-17; the local alignment algorithm of Smith *et al.* (1981) *Adv. Appl. Math.* 2:482; the global alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453; the search-for-local alignment method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci.* 85:2444-2448; the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 872264, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877.

Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, and are not limited to: CLUSTAL in the PC/Gene

program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the GCG Wisconsin Genetics Software Package, Version 10 (available from Accelrys Inc., 9685 Scranton Road, San Diego, California, USA). Alignments using these
5 programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins *et al.* (1988) *Gene* 73:237-244 (1988); Higgins *et al.* (1989) *CABIOS* 5:151-153; Corpet *et al.* (1988) *Nucleic Acids Res.* 16:10881-90; Huang *et al.* (1992) *CABIOS* 8:155-65; and Pearson *et al.* (1994) *Meth. Mol. Biol.* 24:307-331. The ALIGN program is based on the algorithm of
10 Myers and Miller (1988) *supra.* A PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used with the ALIGN program when comparing amino acid sequences. The BLAST programs of Altschul *et al* (1990) *J. Mol. Biol.* 215:403 are based on the algorithm of Karlin and Altschul (1990) *supra.* BLAST nucleotide searches can be performed with the BLASTN program,
15 score = 100, wordlength = 12, to obtain nucleotide sequences homologous to a nucleotide sequence encoding a protein of the embodiments. BLAST protein searches can be performed with the BLASTX program, score = 50, wordlength = 3, to obtain amino acid sequences homologous to a protein or polypeptide of the embodiments. To obtain gapped alignments for comparison purposes, Gapped
20 BLAST (in BLAST 2.0) can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389. Alternatively, PSI-BLAST (in BLAST 2.0) 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, PSI-BLAST, the default parameters of the respective programs (e.g.,
25 BLASTN for nucleotide sequences, BLASTX for proteins) can be used. See www.ncbi.nlm.nih.gov. Alignment may also be performed manually by inspection.

Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using GAP Version 10 using the following parameters: % identity and % similarity for a nucleotide sequence using Gap Weight of 50 and
30 Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using Gap Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix; or any equivalent program thereof. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or

amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP Version 10.

GAP uses the algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453, to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP must make a profit of gap creation penalty number of matches for each gap it inserts. If a gap extension penalty greater than zero is chosen, GAP must, in addition, make a profit for each gap inserted of the length of the gap times the gap extension penalty. Default gap creation penalty values and gap extension penalty values in Version 10 of the GCG Wisconsin Genetics Software Package for protein sequences are 8 and 2, respectively. For nucleotide sequences the default gap creation penalty is 50 while the default gap extension penalty is 3. The gap creation and gap extension penalties can be expressed as an integer selected from the group of integers consisting of from 0 to 200. Thus, for example, the gap creation and gap extension penalties can be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 or greater.

GAP presents one member of the family of best alignments. There may be many members of this family, and no other member has a better quality. GAP displays four figures of merit for alignments: Quality, Ratio, Identity, and Similarity. The Quality is the metric maximized in order to align the sequences. Ratio is the quality divided by the number of bases in the shorter segment. Percent Identity is the percent of the symbols that actually match. Percent Similarity is the percent of the symbols that are similar. Symbols that are across from gaps are ignored. A similarity is scored when the scoring matrix value for a pair of symbols is greater than or equal to 0.50, the similarity threshold. The scoring matrix used in Version 10 of the GCG Wisconsin Genetics Software Package is BLOSUM62 (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

(c) As used herein, "sequence identity" or "identity" in the context of two polynucleotides or polypeptide sequences makes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is

used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity." Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

(d) As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

The use of the term "polynucleotide" is not intended to limit the embodiments to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The polynucleotides of the embodiments also encompass all forms of

sequences including, and not limited to, single-stranded forms, double-stranded forms, and the like.

Isolated polynucleotides of the present invention can be incorporated into recombinant DNA constructs capable of introduction into and replication in a host cell. A "vector" may be such a construct that includes a replication system and sequences that are capable of transcription and translation of a polypeptide-encoding sequence in a given host cell. A number of vectors suitable for stable transfection of plant cells or for the establishment of transgenic plants have been described in, e.g., Pouwels *et al.*, *Cloning Vectors: A Laboratory Manual*, 1985, supp. 1987; Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989; and Flevin *et al.*, *Plant Molecular Biology Manual*, Kluwer Academic Publishers, 1990. Typically, plant expression vectors include, for example, one or more cloned plant genes under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable marker. Such plant expression vectors also can contain a promoter regulatory region (e.g., a regulatory region controlling inducible or constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a signal peptide sequence for targeted expression, a transcription termination site, and/or a polyadenylation signal.

The terms "recombinant construct," "expression cassette," "expression construct," "chimeric construct," "construct," "recombinant DNA construct," "DNA construct" and "recombinant DNA fragment" are used interchangeably herein and are nucleic acid fragments. A recombinant construct comprises an artificial combination of nucleic acid fragments, including, and not limited to, regulatory and coding sequences that are not found together in nature. For example, a recombinant DNA construct may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source and arranged in a manner different than that found in nature. Such construct may be used by itself or may be used in conjunction with a vector. If a vector is used then the choice of vector is dependent upon the method that will be used to transform host cells as is well known to those skilled in the art. For example, a plasmid vector can be used. The skilled artisan is well aware of the genetic elements that must be present on the

vector in order to successfully transform, select and propagate host cells comprising any of the isolated nucleic acid fragments of the invention. Screening to obtain lines displaying the desired expression level and pattern of the polynucleotides or of the *Nsf1* locus may be accomplished by amplification, Southern analysis of DNA, Northern analysis of mRNA expression, immunoblotting analysis of protein expression, phenotypic analysis, and the like.

The term "recombinant DNA construct" refers to a DNA construct assembled from nucleic acid fragments obtained from different sources. The types and origins of the nucleic acid fragments may be very diverse.

In some embodiments, DNA constructs comprising a promoter operably linked to a heterologous nucleotide sequence of the embodiments are further provided. The DNA constructs of the embodiments find use in generating transformed plants, plant cells, and microorganisms and in practicing the methods for inducing ALS and HPPD inhibitor herbicide resistance disclosed herein. The DNA construct will include 5' and 3' regulatory sequences operably linked to a polynucleotide of the embodiments. "Operably linked" is intended to mean a functional linkage between two or more elements. "Regulatory sequences" refer to nucleotides located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which may influence the transcription, RNA processing, stability, or translation of the associated coding sequence. Regulatory sequences may include, and are not limited to, promoters, translation leader sequences, introns, and polyadenylation recognition sequences. For example, an operable linkage between a polynucleotide of interest and a regulatory sequence (a promoter, for example) is functional link that allows for expression of the polynucleotide of interest. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, operably linked is intended to mean that the coding regions are in the same reading frame. The coding sequence may additionally contain a sequence used to target the protein to the chloroplast , the vacuole, the endoplasmic reticulum or to the outside of the cell. The cassette may additionally contain at least one additional gene to be cotransformed into the organism. Alternatively, the additional gene(s) can be provided on multiple DNA constructs. Such a DNA construct is provided with a plurality of restriction sites and/or recombination sites for insertion of the polynucleotide that encodes a herbicide

resistance polypeptide to be under the transcriptional regulation of the regulatory regions. The DNA construct may additionally contain selectable marker genes.

The DNA construct will include in the 5'-3' direction of transcription, a transcriptional initiation region (i.e., a promoter), translational initiation region, a polynucleotide of the embodiments, a translational termination region and, optionally, a transcriptional termination region functional in the host organism. The regulatory regions (i.e., promoters, transcriptional regulatory regions, and translational termination regions) and/or the polynucleotide of the embodiments may be native/analogous to the host cell or to each other. Alternatively, the regulatory regions and/or the polynucleotide of the embodiments may be heterologous to the host cell or to each other. As used herein, "heterologous" in reference to a sequence is a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, a promoter operably linked to a heterologous polynucleotide is from a species different from the species from which the polynucleotide was derived, or, if from the same/analogous species, one or both are substantially modified from their original form and/or genomic locus, or the promoter is not the native promoter for the operably linked polynucleotide.

The optionally included termination region may be native with the transcriptional initiation region, may be native with the operably linked polynucleotide of interest, may be native with the plant host, or may be derived from another source (i.e., foreign or heterologous) to the promoter, the polynucleotide of interest, the host, or any combination thereof. 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 Acids Res.* 15:9627-9639. In particular embodiments, the potato protease inhibitor II gene (PinII) terminator is used. See, for example, Keil *et al.* (1986) *Nucl. Acids Res.* 14:5641-5650; and An *et al.* (1989) *Plant Cell* 1:115-122, herein incorporated by reference in their entirety.

A number of promoters can be used in the practice of the embodiments, including the native promoter of the polynucleotide sequence of interest. The promoters can be selected based on the desired outcome. A wide range of plant promoters are discussed in the recent review of Potenza *et al.* (2004) *In Vitro Cell Dev Biol – Plant 40:1-22*, herein incorporated by reference. For example, the nucleic acids can be combined with constitutive, tissue-preferred, pathogen-inducible, or other promoters for expression in plants. Such constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell *et al.* (1985) *Nature* 313:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

Additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

DNA constructs may additionally contain 5' leader sequences. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Gallie *et al.* (1995) *Gene* 165(2):233-238), MDMV leader (Maize Dwarf Mosaic Virus), and human immunoglobulin heavy-chain binding protein (BiP) (Macejak *et al.* (1991) *Nature* 353:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling *et al.* (1987)

Nature 325:622-625); tobacco mosaic virus leader (TMV) (Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.* (1991) *Virology* 81:382-385). See also, Della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968. Other methods
5 known to enhance translation can also be utilized, for example, introns, and the like.

In preparing the DNA construct, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or
10 linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis, primer repair, restriction, annealing, resubstitutions, e.g., transitions and transversions, may be involved.

15 The DNA construct can also comprise a selectable marker gene for the selection of transformed cells. Selectable marker genes are utilized for the selection of transformed cells or tissues. Marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to
20 herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). Additional selectable markers include phenotypic markers such as β-galactosidase and fluorescent proteins such as green fluorescent protein (GFP) (Su *et al.* (2004) *Biotechnol Bioeng 85:*610-9 and Fetter *et al.* (2004) *Plant Cell 16:*215-28), cyan florescent protein (CYP) (Bolte *et*
25 *al.* (2004) *J. Cell Science 117:*943-54 and Kato *et al.* (2002) *Plant Physiol 129:*913-42), and yellow florescent protein (PhiYFP™ from Evrogen, see, Bolte *et al.* (2004) *J. Cell Science 117:*943-54). For additional selectable markers, see generally, Yarranton (1992) *Curr. Opin. Biotech.* 3:506-511; Christopherson *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6314-6318; Yao *et al.* (1992) *Cell* 71:63-72;
30 Reznikoff (1992) *Mol. Microbiol.* 6:2419-2422; Barkley *et al.* (1980) in *The Operon*, pp. 177-220; Hu *et al.* (1987) *Cell* 48:555-566; Brown *et al.* (1987) *Cell* 49:603-612; Figge *et al.* (1988) *Cell* 52:713-722; Deuschle *et al.* (1989) *Proc. Natl. Acad. Aci. USA* 86:5400-5404; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:2549-2553; Deuschle *et al.* (1990) *Science* 248:480-483; Gossen (1993) Ph.D. Thesis,

University of Heidelberg; Reines *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:1917-1921; Labow *et al.* (1990) *Mol. Cell. Biol.* 10:3343-3356; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3952-3956; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:5072-5076; Wyborski *et al.* (1991) *Nucleic Acids Res.* 19:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol.* 10:143-162; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother.* 35:1591-1595; Kleinschnidt *et al.* (1988) *Biochemistry* 27:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Oliva *et al.* (1992) *Antimicrob. Agents Chemother.* 36:913-919; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.* (1988) *Nature* 334:721-724. Such disclosures are herein incorporated by reference.

The above list of selectable marker genes is not meant to be limiting. Any selectable marker gene can be used in the embodiments.

The gene of the embodiments can be expressed as a transgene in order to make plants resistant to at least one herbicide of the ALS-inhibiting, PPO-inhibiting, pigment synthesis-inhibiting, PS II-inhibiting or synthetic auxin herbicide classes. Using the different promoters described elsewhere in this disclosure, this will allow its expression in a modulated form in different circumstances. One can also insert the entire gene, both native promoter and coding sequence, as a transgene. Finally, using the gene of the embodiments as a transgene will allow quick combination with other traits, such as insect or fungal resistance.

In certain embodiments the nucleic acid sequences of the embodiments can be stacked with any combination of polynucleotide sequences of interest, which may be transgenic or non-transgenic, in order to create plants with a desired phenotype. For example, the polynucleotides of the embodiments may be stacked with any other polynucleotides of the embodiments, or with other genes. The combinations generated can also include multiple copies of any one of the polynucleotides of interest. The polynucleotides of the embodiments can also be stacked with any other gene or combination of genes to produce plants with a variety of desired trait combinations including and not limited to traits desirable for animal feed such as high oil genes (*e.g.*, U.S. Patent No. 6,232,529); balanced amino acids (*e.g.* hordothionins (U.S. Patent Nos. 5,990,389; 5,885,801; 5,885,802; and 5,703,409); barley high lysine (Williamson et al. (1987) *Eur. J. Biochem.* 165:99-106; and WO 98/20122); and high methionine proteins

(Pedersen et al. (1986) *J. Biol. Chem.* 261:6279; Kirihara et al. (1988) *Gene* 71:359; and Musumura et al. (1989) *Plant Mol. Biol.* 12: 123)); increased digestibility (*e.g.*, modified storage proteins (U.S. Application Serial No. 10/053,410, filed November 7, 2001); and thioredoxins (U.S. Application Serial No.
5 10/005,429, filed December 3, 2001)), the disclosures of which are herein incorporated by reference. The polynucleotides of the embodiments can also be stacked with traits desirable for insect, disease or herbicide resistance (*e.g.*, *Bacillus thuringiensis* toxin proteins (U.S. Patent Nos. 5,366,892; 5,747,450; 5,737,514; 5723,756; 5,593,881; Geiser et al (1986) *Gene* 48:109); lectins (Van
10 Damme et al. (1994) *Plant Mol. Biol.* 24:825); fumonisin detoxification genes (U.S. Patent No. 5,792,931); avirulence and disease resistance genes (Jones et al. (1994) Science 266:789; Martin et al. (1993) Science 262:1432; Mindrinos et al. (1994) Cell 78:1089); acetolactate synthase (ALS) mutants that lead to herbicide resistance such as the S4 and/or Hra mutations (Lee *et al.*, (1988) *EMBO J.*
15 7(5):1241-1248), resistance to inhibitors of glutamine synthase such as phosphinothricin or basta (e.g., bar gene; De Block *et al.* (1987) *EMBO J.* 6:2513-2518); HPPD genes that confer tolerance to HPPD inhibiting herbicides such as mesotrione or isoxaflutole (Matringe *et al.* (2005) *Pest Management Science* 61:269-276; Dufourmantel *et al.*, (2007) *Plant Biotech. J.* 5:118-133; see also
20 WO1997049816), genes for tolerance to PPO inhibiting herbicides (Li and Nicholl (2005) *Pest Management Science* 61:277-285); synthetic auxin resistance genes (US patent application 2005/014737 and Herman *et al.*, (2005) *J. Biol. Chem.* 280: 24759-24767), and glyphosate resistance (*epsps* genes, *gat* genes such as those disclosed in U.S. Patent Application Publication US2004/0082770, also
25 WO02/36782 and WO03/092360)); and traits desirable for processing or process products such as high oil (*e.g.*, U.S. Patent No. 6,232,529); modified oils (*e.g.*, fatty acid desaturase genes (U.S. Patent No. 5,952,544; WO 94/11516)); modified starches (*e.g.*, ADPG pyrophosphorylases (AGPase), starch synthases (SS), starch branching enzymes (SBE) and starch debranching enzymes (SDBE)); and
30 polymers or bioplastics (*e.g.*, U.S. patent No. 5.602,321; beta-ketothiolase, polyhydroxybutyrate synthase, and acetoacetyl-CoA reductase (Schubert et al. (1988) *J. Bacteriol.* 170:5837-5847) facilitate expression of polyhydroxyalkanoates (PHAs)), the disclosures of which are herein incorporated by reference. One could also combine the polynucleotides of the embodiments with polynucleotides

providing agronomic traits such as male sterility (*e.g.*, see U.S. Patent No. 5.583,210), stalk strength, flowering time, yield improvement, or transformation technology traits such as cell cycle regulation or gene targeting (*e.g.* WO 99/61619; WO 00/17364; WO 99/25821), the disclosures of which are herein incorporated by reference.

These stacked combinations can be created by any method including and not limited to cross breeding plants by any conventional or TopCross® methodology, or genetic transformation. If the traits are stacked by genetically transforming the plants, the polynucleotide sequences of interest can be combined at any time and in any order. For example, a transgenic plant comprising one or more desired traits can be used as the target to introduce further traits by subsequent transformation. The traits can be introduced simultaneously in a co-transformation protocol with the polynucleotides of interest provided by any combination of transformation cassettes. For example, if two sequences will be introduced, the two sequences can be contained in separate transformation cassettes (trans) or contained on the same transformation cassette (cis). Expression of the sequences can be driven by the same promoter or by different promoters. In certain cases, it may be desirable to introduce a transformation cassette that will suppress the expression of the polynucleotide of interest. This may be combined with any combination of other suppression cassettes or overexpression cassettes to generate the desired combination of traits in the plant. It is further recognized that polynucleotide sequences can be stacked at a desired genomic location using a site-specific recombination system. See, for example, WO99/25821, WO99/25854, WO99/25840, WO99/25855, and WO99/25853, all of which are herein incorporated by reference.

Further embodiments include plants obtainable by a method comprising: crossing a plant containing the *Nsf1* gene as a first parent plant, with a different plant that lacks an *Nsf1* gene as a second parent plant, thereby to obtain progeny comprising the *Nsf1* gene of the first parent; and optionally further comprising one or more further breeding steps to obtain progeny of one or more further generations comprising the *Nsf1* gene of the first parent. Such embodied plants can include both inbred and hybrid plants. Seeds of such plants, including those seeds which are homozygous and heterozygous for the *Nsf1* gene, and methods of obtaining plant products resulting from the processing of those seeds are

embodied in the invention. Using such seed in food or feed or the production of a corn product, such as flour, meal and oil is also an embodiment of the invention.

An "ancestral line" or "progenitor" is a parent line used as a source of genes, e.g., for the development of elite lines. "Progeny" are the descendents of the ancestral line, and may be separated from their ancestors by many generations of breeding. An "elite line" or "elite variety" is an agronomically superior line or variety that has resulted from many cycles of breeding and selection for superior agronomic performance. Similarly, "elite germplasm" is an agronomically superior germplasm, typically derived from and/or capable of giving rise to a plant with superior agronomic performance, such as an existing or newly developed elite line of corn or soybeans.

Also embodied in the invention is the use of molecular markers to move the gene or transgene into elite lines using breeding techniques. Molecular markers can be used in a variety of plant breeding applications (eg see Staub *et al.* (1996) *Hortscience* 31: 729-741; Tanksley (1983) *Plant Molecular Biology Reporter.* 1: 3-8). One of the main areas of interest is to increase the efficiency of backcrossing and introgressing genes using marker-assisted selection (MAS). A molecular marker that demonstrates linkage with a locus affecting a desired phenotypic trait provides a useful tool for the selection of the trait in a plant population. This is particularly true where the phenotype is hard to assay, e.g. many disease resistance traits, or, occurs at a late stage in the plants development, e.g. seed characteristics. Since DNA marker assays are less laborious, and take up less physical space, than field phenotyping, much larger populations can be assayed, increasing the chances of finding a recombinant with the target segment from the donor line moved to the recipient line. The closer the linkage, the more useful the marker, as recombination is less likely to occur between the marker and the gene causing the trait, which can result in false positives. Having flanking markers decreases the chances that false positive selection will occur as a double recombination event would be needed. The ideal situation is to have a marker in the gene itself, so that recombination can not occur between the marker and the gene. Such a marker is called a 'perfect marker'.

Optionally, the nucleic acids of the embodiments may be targeted to the chloroplast for expression. In this manner, where the nucleic acid 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 5 *et al.* (1986) *Science* 233:478-481.

Chloroplast targeting sequences are known in the art and include the chloroplast small subunit of ribulose-1,5-bisphosphate carboxylase (Rubisco) (de Castro Silva Filho *et al.* (1996) *Plant Mol. Biol.* 30:769-780; Schnell *et al.* (1991) *J. Biol. Chem.* 266(5):3335-3342); 5-(enolpyruvyl)shikimate-3-phosphate synthase 10 (EPSPS) (Archer *et al.* (1990) *J. Bioenerg. Biomemb.* 22(6):789-810); tryptophan synthase (Zhao *et al.* (1995) *J. Biol. Chem.* 270(11):6081-6087); plastocyanin (Lawrence *et al.* (1997) *J. Biol. Chem.* 272(33):20357-20363); chorismate synthase (Schmidt *et al.* (1993) *J. Biol. Chem.* 268(36):27447-27457); and the light harvesting chlorophyll a/b binding protein (LHBP) (Lamppa *et al.* (1988) *J.* 15 *Biol. Chem.* 263:14996-14999). See also 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.

Methods for transformation of chloroplasts are known in the art. See, for 20 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 25 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 nucleic acids to be targeted to the chloroplast may be optimized for 30 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.

The methods of the embodiments may involve, and are not limited to, introducing a polypeptide or polynucleotide into a plant. "Introducing" is intended to mean presenting to the plant the polynucleotide. In some embodiments, the polynucleotide will be presented in such a manner that the sequence gains access to the interior of a cell of the plant, including its potential insertion into the genome of a plant. The methods of the embodiments do not depend on a particular method for introducing a sequence into a plant, only that the polynucleotide gains access to the interior of at least one cell of the plant. Methods for introducing polynucleotides into plants are known in the art including, and not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

"Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. "Host cell" refers the cell into which transformation of the recombinant DNA construct takes place and may include a yeast cell, a bacterial cell, and a plant cell. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere *et al.*, 1987, *Meth. Enzymol. 143*:277) and particle-accelerated or "gene gun" transformation technology (Klein *et al.*, 1987, *Nature (London) 327*:70-73; U.S. Patent No. 4,945,050), among others.

"Stable transformation" is intended to mean that the nucleotide construct introduced into a plant integrates into the genome of the plant and is capable of being inherited by the progeny thereof. "Transient transformation" or "transient expression" is intended to mean that a polynucleotide is introduced into the plant and does not integrate into the genome of the plant or a polypeptide is introduced into a plant.

Transformation protocols as well as protocols for introducing polypeptides or polynucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., monocot or dicot, targeted for transformation. Suitable methods of introducing polypeptides and polynucleotides into plant cells include microinjection (Crossway *et al.* (1986) *Biotechniques* 4:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, *Agrobacterium*-mediated transformation (U.S. Patent Nos. 5,563,055-and 5,981,840), direct gene

transfer (Paszkowski *et al.* (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, Sanford *et al.*, U.S. Patent Nos. 4,945,050; 5,879,918; 5,886,244; and 5,932,782; Tomes *et al.* (1995) in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-
5 Verlag, Berlin); McCabe *et al.* (1988) *Biotechnology* 6:923-926); and Lec1 transformation (WO 00/28058). Also see Weissinger *et al.* (1988) *Ann. Rev. Genet.* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou *et al.* (1988) *Plant Physiol.* 87:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology* 6:923-926 (soybean); Finer and McMullen
10 (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (soybean); Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324 (soybean); Datta *et al.* (1990) *Biotechnology* 8:736-740 (rice); Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maize); Klein *et al.* (1988) *Biotechnology* 6:559-563 (maize); U.S. Patent Nos. 5,240,855; 5,322,783 and 5,324,646; Klein *et al.* (1988) *Plant Physiol.* 91:440-444 (maize);
15 Fromm *et al.* (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slogteren *et al.* (1984) *Nature (London)* 311:763-764; U.S. Patent No. 5,736,369 (cereals); Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.* (1990)
20 *Plant Cell Reports* 9:415-418 and Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated transformation); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (electroporation); Li *et al.* (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda *et al.* (1996) *Nature Biotechnology* 14:745-750 (maize via *Agrobacterium tumefaciens*); all of
25 which are herein incorporated by reference.

Methods are known in the art for the targeted insertion of a polynucleotide at a specific location in the plant genome. In one embodiment, the insertion of the polynucleotide at a desired genomic location is achieved using a site-specific recombination system. See, for example, WO99/25821, WO99/25854,
30 WO99/25840, WO99/25855, and WO99/25853, all of which are herein incorporated by reference. Briefly, the polynucleotide of the embodiments can be contained in transfer cassette flanked by two non-identical recombination sites. The transfer cassette is introduced into a plant have stably incorporated into its genome a target site which is flanked by two non-identical recombination sites that

correspond to the sites of the transfer cassette. An appropriate recombinase is provided and the transfer cassette is integrated at the target site. The polynucleotide of interest is thereby integrated at a specific chromosomal position in the plant genome.

5 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 progeny having constitutive expression of the desired phenotypic characteristic 10 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 embodiments provides transformed seed (also referred to as "transgenic seed") having a nucleotide construct of the 15 embodiments, for example, a DNA construct of the embodiments, stably incorporated into their genome.

As used herein, the term plant includes plant cells, plant protoplasts, plant cell tissue cultures from which maize plant can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, 20 pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers, and the like. Grain is intended to mean the mature seed produced by commercial growers for purposes other than growing or reproducing the species. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the embodiments, provided that these 25 parts comprise the introduced polynucleotides.

The embodiments of the invention may be used to confer or enhance herbicide resistance in plants, especially soy (*Glycine max*). Other plant species may also be of interest in practicing the embodiments of the invention, including, and not limited to, other dicot and monocot crop plants. The maize gene of the 30 embodiments is commonly found in the majority of commercial corn lines, most of which are naturally tolerant to at least one, and usually several, synthetic auxin, ALS-, PS II- and pigment synthesis-inhibitor herbicides, such as rimsulfuron, nicosulfuron and mesotrione.

It is therefore envisioned that the same tolerance to certain herbicides present in most corn lines can be extended to other crop plants by transgenic means though the use of the endogenous maize *Nsf1* gene and variants thereof. Listings of maize lines with tolerance or sensitivity to selected SU herbicides are widely available, such as those provided by the USDA, ARS, National Genetic Resources Program. *Germplasm Resources Information Network - (GRIN).* [Online Database] National Germplasm Resources Laboratory, Beltsville, Maryland. [retrieved on March 6, 2006]: Retrieved from the internet: <URL: http://www.ars-grin.gov/cgi-bin/npgs/html/dno_eval_acc.pl?89201+153002+21> ; and the "Maize Germplasm Lines" listings available from the Buckler Laboratory website [retrieved on March 6, 2006]: Retrieved from the internet: <URL: http://www.maizegenetics.net/index.php?page=germplasm/lines.html>, and also in reference articles such as Kang (1993) *J. Heredity.* 84(3): 216-217.

Where appropriate, the polynucleotides may be optimized for increased expression in the transformed organism. For example, the polynucleotides can be synthesized using plant-preferred codons for improved expression. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of host-preferred codon usage. Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference.

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 (*Zea mays*), *Brassica* sp. (e.g., *B. napus, B. rapa, B. juncea*), alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor, Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*)), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense, Gossypium hirsutum*), sweet potato (*Ipomoea batatus*), cassava (*Manihot esculenta*), coffee (*Coffea* spp.), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus*

casica), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum* spp.), oats (*Avena* spp.), barley, palm, coconut, castor bean,
5 olive, beans (for example guar, locust bean, fenugreek, soybean, garden beans, mung beans, lima beans, fava beans), peas (such as cowpeas, field peas, lentils, chickpeas, etc.), vegetables, ornamentals, and conifers.

Other plants of interest for the invention include those which have the potential for use as biofuel crops, including, but not limited to, prairie grasses such
10 as switchgrass (*Panicum virgatum*), elephant grass (*Pennisetum purpureum*), Johnson grass (*Sorghum halepense*), *Miscanthus* spp., as well as hybrid poplar and hybrid willow trees.

Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*),
15 peas (*Lathyrus* spp.), and members of the genus *Cucumis* such as cucumber (*C. sativus*), cantaloupe (*C. cantalupensis*), and musk melon (*C. melo*). Ornamentals include azalea (*Rhododendron* spp.), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa* spp.), tulips (*Tulipa* spp.), daffodils (*Narcissus* spp.), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia
20 (*Euphorbia pulcherrima*), and chrysanthemum.

The embodiments provide not only a gene for use in transgenic applications, but sequences and methods that allow the resistance gene to be used as a marker in corn breeding strategies. For example, the gene of the embodiments, or the locus containing it, may be identified in a crop line intended to be used for
25 breeding. Breeders would generally want to avoid using crop lines that are sensitive to herbicides where there is usually natural tolerance. Accordingly, the identification of the sequence of the *Nsf1* gene will help breeders to identify and avoid creating herbicide-sensitive lines.

Nucleic acid based markers can be developed and applied using many
30 different technologies. Such technologies include, and are not limited to, Restriction Fragment Length Polymorphism (RFLP), Simple Sequence Repeat (SSR), Random Amplified Polymorphic DNA (RAPD), Cleaved Amplified Polymorphic Sequences (CAPS) (Rafalski and Tingey, 1993, Trends in Genetics 9:275-280), Amplified Fragment Length Polymorphism (AFLP) (Vos *et al.*, 1995,

Nucleic Acids Res. 23:4407-4414), Single Nucleotide Polymorphism (SNP) (Brookes, 1999, Gene 234:177-186), Sequence Characterized Amplified Region (SCAR) (Paran and Michelmore, 1993, Theor. Appl. Genet. 85:985-993), Sequence Tagged Site (STS) (Onozaki *et al.*, 2004, Euphytica 138:255-262),
5 Single Stranded Conformation Polymorphism (SSCP) (Orita *et al.*, 1989, Proc Natl Acad Sci USA 86:2766–2770), Inter-Simple Sequence Repeat (ISSR) (Blair *et al.*, 1999, Theor. Appl. Genet. 98:780-792), Inter-Retrotransposon Amplified Polymorphism (IRAP), Retrotransposon-Microsatellite Amplified Polymorphism (REMAP) (Kalendar *et al.* (1999) Theor. Appl. Genet. 98:704-711) and the like.
10 As used herein, "locus" shall refer to a genetically defined region of a chromosome carrying a gene or, possibly, two or more genes so closely linked that genetically they behave as a single locus, responsible for a phenotype. A "gene" shall refer to a specific gene within that locus, including its associated regulatory sequences. Thus, the *Nsf1* locus refers to the chromosomal region
15 genetically defined as conferring resistance to at least one herbicide of the ALS-inhibiting, PPO-inhibiting, pigment synthesis-inhibiting, PS II-inhibiting and synthetic auxin herbicide class. One embodiment of the present invention is the isolation of the *Nsf1* gene and the demonstration that it is the gene responsible for the phenotype conferred by the presence of the locus. Genetically defined loci are
20 by their nature not as precisely defined in terms of size as genes, which can be delineated molecularly.

Units, prefixes, and symbols may be denoted in their SI accepted form. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy
25 orientation, respectively. Numeric ranges are inclusive of the numbers defining the range. Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes. The above-defined terms are
30 more fully defined by reference to the specification as a whole.

Examples

The embodiments of the invention are further defined in the following Examples, in which all parts and percentages are by weight and degrees are

Celsius, unless otherwise stated. It should be understood that these Examples, while indicating embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of the embodiments of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications to adapt it to various usages and conditions. Thus, various modifications of the embodiments of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. The disclosure of each reference set forth herein is incorporated by reference in its entirety

Example 1

Identification of the *Nsf1* Gene through Positional Cloning

A BC1 population (expected 50% *Nsf1/nsf1*, 50% *nsf1/nsf1)* was developed using the sensitive inbred W703A as the recurrent parent, and either B73 or Q66 as the resistant line. Plants were misted with a 2.3 mM nicosulfuron, 0.5% v/v Kinetic surfactant solution at approximately the V3 stage. Both resistant and sensitive parents were also grown and sprayed as controls. In order to avoid falsely classifying a plant which may have died due to reasons other than the herbicide application, only resistant progeny were sampled and analyzed. A total of 96 resistant plants were used for the initial mapping. This was sufficient to place *Nsf1* between markers umc1766 and umc2036, and thus on contig 202 of the maizeB73-based physical map ((Retrieved on March 6, 2006) Retrieved from the internet <URL: http://www.gramene.org/Zea_mays/cytoview?contig=ctg202&x=44&y=9>).

Based on BAC-end sequences of a maize Mo17-based contig, flanking CAPS (cleaved amplified polymorphic sequence) markers were identified on BACs of contig 202.

For finer mapping of this interval, a total of 388 resistant plants were used in the next step. Based on sequencing of subcloned fragments of BACs in this interval, two flanking CAPs markers were found on overlapping BACs. Both of these markers had 2/388 recombinants.

Both of these BACs were sequenced and analyzed. Within the 163 kb region of the 2 BACs flanked by two proprietary markers, P1 and P2, there were several putative genes. For the third round of mapping, a total of 2584 resistant plants were used, and markers were developed to separate some of the genes. One marker showed 11/2584 (0.4%) recombinants, helping to eliminate certain genes as being responsible for the resistance. Two other markers each had 2 (0.08%) recombinants, eliminating yet another gene. Finally, a marker between two genes had a single recombinant (0.04%), eliminating one of those two genes. Thus it was determined which gene was the gene of interest. The gene, *Nsf1*, was determined to have homology to some cytochrome P450 genes known in the art.

Example 2

Analysis of the *Nsf1* Gene

Analysis of the Gene 18 (*Nsf1*) sequence in the B73-derived BAC shows an open reading frame of 521 amino acids, and containing the conserved heme-binding motif FXXGXXXCXG (SEQ ID NO: 14) found in all cytochrome P450s (Figures 1d and 2b).

In order to determine if the *Nsf1* allele was consistent across maize lines, three corn lines with unknown sensitivity levels to nicosulfuron were tested to determine their reaction and then evaluate their sequences. Plants were misted with a 2.3 mM nicosulfuron, 0.5% v/v Kinetic surfactant solution at approximately the V3 stage. Both known resistant and sensitive lines were also grown and sprayed as controls. Results of the testing of the three lines showed that lines Q66 and Black Mexican Sweet (BMS) were resistant and line A188 was sensitive.

Of these two other resistant lines, Q66 and BMS, also possess this ORF, although Q66 differs from both B73 and BMS by 3 amino acids (Figure 2a and 2b)

These three variant amino acids are marked with bold type and rectangles in Figures 2a and 2b in the Q66 sequence string to show their positions. Analysis of a sensitive line, GA209, shows an insertion of 392 bp relative to the resistant lines which results in a frameshift and an open reading frame of only 338 amino acids (Figure 2b). A survey of numerous North American sensitive lines showed that many of the sensitive lines contain this same insertion of unknown DNA.

Analysis of the sequence from the F2 sensitive line showed that there is only one nucleotide difference between B73 (SEQ ID NO: 2) and F2 (SEQ ID NO: 22), which changes amino acid 263 from arginine to threonine (Figure 2b). This single change therefore eliminates the resistance phenotype and variant sequences with such a change are expected not to retain biological activity. This change is useful in developing an SNP to assist corn breeders in avoiding the susceptible allele.

Nsf1 is 67% identical to a rice cytochrome P450 which has recently been reported to control sulfonylurea sensitivity in that plant (Accession No: ABC69856, SEQ ID NO: 4).

Genomic sequence from B73 shows a single intron with the expected GT left border and AG right border. The position of the intron is shown in the sequence listing in SEQ ID NO: 16.

The cloning of this gene has a number of potential applications. It could be used as a selectable marker for transformation in a sensitive transformable line such as A188 (Ishida *et al.*, (1996) *Nature Biotechnology* 14:745-750). A transgene designed to suppress the *Nsf1* gene function would function as a dominant negative selectable marker. *Nsf1* could also be used to create transgenic resistance in other plants, such as soybean, which are sensitive to this subclass of sulfonylureas.

Example 3:

Testing of Maize Plants for Sensitivity to Nicosulfuron

Three corn lines with unknown sensitivity levels to nicosulfuron were tested to determine their reaction. Plants were misted with a 2.3 mM nicosulfuron, 0.5% v/v Kinetic surfactant solution at approximately the V3 stage. Both known resistant and sensitive lines were also grown and sprayed as controls. Results of

the testing of the three lines showed that lines Q66 and BMS were resistant and line A188 was sensitive.

Example 4: Preparation of Transgenic Soybean Plants

5 The following stock solutions and media were used for transformation and regeneration of soybean plants:

Stock solutions

10 **Sulfate 100 X Stock:** 37.0 g $MgSO_4.7H_2O$, 1.69 g $MnSO_4.H_2O$, 0.86 g $ZnSO_4.7H_2O$, 0.0025 g $CuSO_4.5H_2O$.

Halides 100 X Stock: 30.0 g $CaCl_2.2H_2O$, 0.083 g KI, 0.0025 g $CoCl_2.6H_2O$,

P, B, Mo 100X Stock: 18.5 g KH_2PO_4, 0.62 g H_3BO_3, 0.025 g $Na_2MoO_4.2H_2O$

Fe EDTA 100X Stock: 3.724 g Na_2EDTA, 2.784 g $FeSO_4.7H_2O$.

15 **2,4-D Stock:** 10 mg/mL.

Vitamin B5 1000X Stock: 10.0 g *myo*-inositol, 0.10 g nicotinic acid, 0.10 g pyridoxine HCl, 1 g thiamine.

Media (per Liter)

SB196: 10 mL of each of the above stock solutions, 1 mL B5 Vitamin stock,

20 0.463 g $(NH_4)_2SO_4$, 2.83 g KNO_3, 1 mL 2,4-D stock, 1 g asparagine, 10 g sucrose, pH 5.7.

SB103: 1 pk. Murashige & Skoog salts mixture, 1 mL B5 Vitamin stock, 750 mg $MgCl_2$ hexahydrate, 60 g maltose, 2 g gelrite, pH 5.7.

SB166: SB103 supplemented with 5 g per liter activated charcoal.

25 **SB71-4:** Gamborg's B5 salts (Gibco-BRL catalog No. 21153-028), 1 mL B5 vitamin stock, 30 g sucrose, 5 g TC agar, pH 5.7.

Soybean embryogenic suspension cultures were maintained in 35 mL liquid medium (SB196) on a rotary shaker (150 rpm) at 28 °C with fluorescent lights

30 providing a 16-hour day/8-hour night cycle. Cultures were subcultured every 2 weeks by inoculating approximately 35 mg of tissue into 35 mL of fresh liquid media.

Soybean embryogenic suspension cultures were transformed by particle gun bombardment (see Klein *et al.* (1987) *Nature 327*:70-73) using a DuPont Biolistic

35 PDS1000/He instrument.

The recombinant DNA plasmid used to express *Nsf1* was on a separate recombinant DNA plasmid from the selectable marker gene. Both recombinant DNA plasmids were co-precipitated onto gold particles as follows. The DNAs in suspension were added to 50 µL of a 20 - 60 mg/mL 0.6 µm gold particle suspension and then combined with 50 µL $CaCl_2$ (2.5 M) and 20 µL spermidine (0.1 M). The mixture was pulse vortexed 5 times, spun in a microfuge for 10 seconds, and the supernatant removed. The DNA-coated particles are then washed once with 150 µL of 100% ethanol, pulse vortexed and spun in a microfuge again, and resuspended in 85 µL of anhydrous ethanol. Five µL of the DNA-coated gold particles are then loaded on each macrocarrier disk.

Approximately 150 to 250 mg of two-week-old suspension culture was placed in an empty 60 mm x 15 mm petri plate and the residual liquid is removed from the tissue using a pipette. The tissue was placed about 3.5 inches away from a retaining screen and each plate of tissue was bombarded once. Membrane rupture pressure was set at 650 psi and the chamber was evacuated to –28 inches of Hg. Eighteen plates were bombarded, and, following bombardment, the tissue from each plate was divided between two flasks, placed back into liquid media, and cultured as described above.

Seven days after bombardment, the liquid medium was exchanged with fresh SB196 medium supplemented with 50 mg/mL hygromycin. The selective medium was refreshed weekly or biweekly. Seven weeks post-bombardment, green, transformed tissue was observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue was removed and inoculated into individual flasks to generate new, clonally-propagated, transformed embryogenic suspension cultures. Thus, each new line was treated as an independent transformation event. These suspensions were then maintained as suspensions of embryos clustered in an immature developmental stage through subculture or were regenerated into whole plants by maturation and germination of individual somatic embryos.

Transformed embryogenic clusters were removed from liquid culture and placed on solid agar medium (SB166) containing no hormones or antibiotics for one week. Embryos were cultured at 26 °C with mixed fluorescent and incandescent lights on a 16-hour day: 8-hour night schedule. After one week, the cultures were then transferred to SB103 medium and maintained in the same

growth conditions for 3 additional weeks. Prior to transfer from liquid culture to solid medium, tissue from selected lines was assayed by PCR for the presence of the chimeric gene. Somatic embryos became suitable for germination after 4 weeks and were then removed from the maturation medium and dried in empty
5 petri dishes for one to five days. The dried embryos were then planted in SB71-4 medium and allowed to germinate under the same light and germination conditions described above. Germinated embryos were transferred to sterile soil and grown to maturity.

10

Example 5

T0 and T1 Transgenic Plant Analysis

T0 Testing

Two different constructs comprising the *Nsf1* gene were created to examine herbicide efficacy of the gene when transformed into soybean. The *Nsf1*
15 constructs were co-bombarded with a 35S:HYG insert to permit event selection using hygromycin.

At the V2 to V6 growth stage, a total of 127 T0 plants were sprayed with 35 g/ha rimsulfuron. All rimsulfuron treatments were applied with 0.2% w/w nonionic surfactant in a spray volume of 287 L/ha. In addition to the T0 plants, replications
20 of three different controls were included – two positive and one negative. Individual plants were evaluated for herbicide response at ten days after treatment, and assigned a visual response score from 1 to 9 (1 = dead plant to 9 = no effect observed). Based upon high tolerance scores to the initial rimsulfuron spray, five T0 events were sprayed with an additional 35 g/ha rimsulfuron. Plants
25 were rated for visual tolerance using a 1 to 9 score at ten days after the second application.

In the T0 generation, 4 of 51 events had improved tolerance compared to the controls at ten days after treatment with 35 g/ha rimsulfuron. Three of 51 T0 events had improved level of tolerance after an additional application of 35 g/ha
30 rimsulfuron. Two of these 51 events were advanced to the T1 generation for more extensive herbicide testing.

T1 Testing

Two events from the T0 generation were advanced to the T1 generation for additional herbicide efficacy testing of the *Nsf1* gene. Replicates of two controls, as well as T1 plants, were grown in greenhouse experiments and sprayed with mesotrione at one of two rates (200 g/ha or 50 g/ha), nicosulfuron (70 g/ha), or rimsulfuron (35 g/ha) at the V3 growth stage. All herbicide treatments were applied with 1% w/w modified seed oil adjuvant in a spray volume of 374 L/ha. Plants were rated for herbicide response at eight days after application using a 1 to 9 score as used in the T0 testing.

An expanded herbicide efficacy test was developed in a second T1 plant experiment for the same two events advanced from the T0 generation. At the V3 growth stage, plants were sprayed with different treatments of herbicides that would typically cause substantial crop injury when applied to commodity soybean at the rates examined. All herbicide treatments were applied in a spray volume of 287 L/ha. Isoxaflutole (140 g/ha), topramezone (140 g/ha), and sulcotrione (140 g/ha) were applied with 1% w/w modified seed oil adjuvant. Diuron treatments (560 g/ha) were applied with 1% w/w petroleum crop oil adjuvant. Acifluorfen (4480 g/ha), sulfentrazone (140 g/ha), flumioxazin (140 g/ha), and dicamba (280 g/ha) were applied with 0.25% w/w nonionic surfactant. Rimsulfuron (35 g/ha) treatments were applied with 0.5% w/w basic blend adjuvant. At eight and fifteen days after treatment, plants were rated visually for crop injury using a 0 to 100 scale (0 = no injury to 100 = dead plant). Since the T1 events were segregating, only the plants with the best overall scores were selected, corresponding to the 75% that would be expected to possess the transgene.

One of the two events had significantly better tolerance compared to the controls at 8 DAT and 15 DAT after application of acifluorfen, dicamba, diuron, flumioxazin, isoxaflutole, mesotrione, rimsulfuron, sulcotrione, sulfentrazone, and topramezone treatments. The second event had significantly better tolerance compared to the controls at 15 DAT after application of acifluorfen, dicamba, isoxaflutole, mesotrione, rimsulfuron, sulcotrione, sulfentrazone, and topramezone treatments. Although the exact expression level of the *Nsf1* gene in the events tested was not determined, transgenic soybean plants comprising the maize *Nsf1* gene displayed better tolerance to a range of different herbicides when compared directly to control plants.

CLAIMS

What is claimed is:

5

1. An isolated polynucleotide comprising:

(a) a nucleotide sequence encoding a polypeptide capable of conferring resistance to at least one herbicide, wherein said herbicide is a member of a class of herbicides selected from the group consisting of:

10 (i) the ALS-inhibiting class;

(ii) the pigment synthesis-inhibiting class;

(iii) the PPO-inhibiting class;

(iv) the PS II-inhibiting class; and

(v) the synthetic auxin class

15 wherein the polypeptide has an amino acid sequence of at least 85% identity, when compared to SEQ ID NO:1 based on the Needleman-Wunsch alignment algorithm, or

(b) a complement of the nucleotide sequence, wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are

20 100% complementary.

2. An isolated polynucleotide of claim 1, wherein the polypeptide is capable of conferring resistance to at least two herbicides, wherein each herbicide is a member of a different class of herbicides selected from the group consisting of:

25 (a) the ALS-inhibiting class;

(b) the pigment synthesis-inhibiting class;

(c) the PPO-inhibiting class;

(d) the PS II-inhibiting class; and

(e) the synthetic auxin class.

30 3. An isolated polynucleotide of claim 1, wherein the polypeptide is capable of conferring resistance to at least three herbicides, wherein each herbicide is a member of a different class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;

35 (b) the pigment synthesis-inhibiting class;

(c) the PPO-inhibiting class;

(d) the PS II-inhibiting class; and

(e) the synthetic auxin class.

4. An isolated polynucleotide of claim 1, wherein the polypeptide is capable of conferring resistance to at least four herbicides, wherein each herbicide is a member of a different class of herbicides selected from the group consisting of:

5 (a) the ALS-inhibiting class;

(b) the pigment synthesis inhibiting class;

(c) the PPO-inhibiting class;

(d) the PS II-inhibiting class; and

(e) the synthetic auxin class.

10 5. An isolated polynucleotide of claim 1, wherein the polypeptide is capable of conferring resistance to at least five herbicides, wherein each herbicide is a member of a different class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;

15 (b) the pigment synthesis-inhibiting class;

(c) the PPO-inhibiting class;

(d) the PS II-inhibiting class; and

(e) the synthetic auxin class.

6. The polynucleotide of Claim 1, wherein the amino acid sequence of

20 the polypeptide and the amino acid sequence of SEQ ID NO: 1 have at least 90% identity based on the Needleman-Wunsch alignment algorithm.

7. The polynucleotide of Claim 1, wherein the amino acid sequence of the polypeptide and the amino acid sequence of SEQ ID NO: 1 have at least 95% identity based on the Needleman-Wunsch alignment algorithm.

25 8. The polynucleotide of Claim 1, wherein the nucleotide sequence comprises SEQ ID NO: 1.

9. A vector comprising the polynucleotide of Claim 1.

10. A recombinant DNA construct comprising the polynucleotide of Claim 1 operably linked to at least one regulatory sequence.

30 11. A method for transforming a cell, comprising transforming a cell with the polynucleotide of Claim 1.

12. A plant cell comprising the recombinant DNA construct of Claim 10.

13. A method for producing a plant comprising transforming a plant cell with the recombinant DNA construct of Claim 10 and regenerating a plant from the

35 transformed plant cell.

14. A plant comprising the recombinant DNA construct of Claim 10.

15. A seed comprising the recombinant DNA construct of Claim 10.

16. The plant of claim 14, wherein said plant is a monocot.

17. The plant of claim 16, wherein said dicot is selected from the group consisting of maize, wheat, barley, oats, switchgrass, sorghum, and rice.

18. The plant of claim 14, wherein said plant is a dicot.

19. The plant of claim 18, wherein said dicot is selected from the group consisting of soybean, canola, potato, cotton, and sunflower.

20. The plant of claim 14, wherein said plant further comprises a second herbicide resistance gene.

21. The plant of claim 14, wherein said plant further comprises a gene encoding a polypeptide with glyphosate N-acetyltransferase activity.

22. The plant of claim 21, wherein said plant further comprises a second herbicide resistance gene encoding a polypeptide conferring tolerance to ALS inhibitors.

23. The plant of claim 14, wherein said plant further comprises a gene encoding an insecticidal polypeptide.

24. A plant with enhanced tolerance to at least one herbicide, comprising the recombinant DNA construct of Claim 10, wherein said plant further comprises a second herbicide resistance gene providing a level of tolerance to a herbicide selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
(b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

wherein the recombinant construct of Claim 10 confers upon the plant a higher level of tolerance to the herbicide than the tolerance level exhibited by a plant comprising the second herbicide resistance gene but not comprising the recombinant construct of Claim 10.

25. A method of conferring or enhancing resistance to at least one herbicide, wherein said herbicide is selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
(b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

comprising transforming a plant with the recombinant DNA construct of Claim 10, thereby conferring or enhancing resistance to the at least one herbicide.

26. A method of conferring or enhancing resistance to at least two herbicides, wherein each herbicide is selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
5 (b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

comprising transforming a plant with the recombinant DNA construct of Claim 10,
10 thereby conferring or enhancing resistance to the at least two herbicides.

27. A method of conferring or enhancing resistance to at least three herbicides, wherein said each herbicide is selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
15 (b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

comprising transforming a plant with the recombinant DNA construct of Claim 10,
20 thereby conferring or enhancing resistance to the at least three herbicides.

28. A method of conferring or enhancing resistance to at least four herbicides, wherein each herbicide is selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
25 (b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

comprising transforming a plant with the recombinant DNA construct of Claim 10,
30 thereby conferring or enhancing resistance to the at least four herbicides.

29. A method of conferring or enhancing resistance to at least five herbicides, wherein each herbicide is selected from a class of herbicides selected from the group consisting of:

(a) the ALS-inhibiting class;
35 (b) the pigment synthesis-inhibiting class;
(c) the PPO-inhibiting class;
(d) the PS II-inhibiting class; and
(e) the synthetic auxin class;

comprising transforming a plant with the recombinant DNA construct of Claim 10, thereby conferring or enhancing resistance to the at least five herbicides.

30. A method of altering the level of expression of a protein capable of conferring resistance to at least one herbicide in a plant cell comprising:

(a) transforming a plant cell with the recombinant DNA construct of Claim 10; and

(b) growing the transformed plant cell under conditions that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of a protein capable of conferring resistance to the at least one herbicide in the transformed host;

wherein the at least one herbicide is selected from a class of herbicides selected from the group consisting of:

(i) the ALS-inhibiting class;
(ii) the pigment synthesis-inhibiting class;
(iii) the PPO-inhibiting class;
(iv) the PS II-inhibiting class; and
(v) the synthetic auxin class.

31. A method of altering the level of expression of a protein capable of conferring resistance to at least two herbicides in a plant cell comprising:

(a) transforming a plant cell with the recombinant DNA construct of Claim 10; and

(b) growing the transformed plant cell under conditions that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of a protein capable of conferring resistance to the at least two herbicides in the transformed host;

wherein each of the at least two herbicides is selected from a class of herbicides selected from the group consisting of:

(i) the ALS-inhibiting class;
(ii) the pigment synthesis-inhibiting class;
(iii) the PPO-inhibiting class;
(iv) the PS II-inhibiting class; and
(v) the synthetic auxin class.

32. A method of altering the level of expression of a protein capable of conferring resistance to at least three herbicides in a plant cell comprising:

(a) transforming a plant cell with the recombinant DNA construct of Claim 10; and

(b) growing the transformed plant cell under conditions that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of a protein capable of conferring resistance to the at least three herbicides in the transformed host;

wherein each of the at least three herbicides is selected from a class of herbicides selected from the group consisting of:

(i) the ALS-inhibiting class;
(ii) the pigment synthesis-inhibiting class;
(iii) the PPO-inhibiting class;
(iv) the PS II-inhibiting class; and
(v) the synthetic auxin class.

33. A method of altering the level of expression of a protein capable of conferring resistance to at least four herbicides in a plant cell comprising:

(a) transforming a plant cell with the recombinant DNA construct of Claim 10; and

(b) growing the transformed plant cell under conditions that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of a protein capable of conferring resistance to the at least four herbicides in the transformed host;

wherein each of the at least four herbicides is selected from a class of herbicides selected from the group consisting of:

(i) the ALS-inhibiting class;
(ii) the pigment synthesis-inhibiting class;
(iii) the PPO-inhibiting class;
(iv) the PS II-inhibiting class; and
(v) the synthetic auxin class.

34. A method of altering the level of expression of a protein capable of conferring resistance to at least five herbicides in a plant cell comprising:

(a) transforming a plant cell with the recombinant DNA construct of Claim 10; and

(b) growing the transformed plant cell under conditions that are suitable for expression of the recombinant DNA construct wherein expression of the recombinant DNA construct results in production of altered levels of a protein capable of conferring resistance to the at least five herbicides in the transformed host;

wherein each of the at least five herbicides is selected from a class of herbicides selected from the group consisting of:

(i) the ALS-inhibiting class;
(ii) the pigment synthesis-inhibiting class;
(iii) the PPO-inhibiting class;
(iv) the PS II-inhibiting class; and
5 (v) the synthetic auxin class.

35. The method of any one of claims 25-34 wherein the herbicide is selected from the ALS-inhibiting class of herbicides and is selected from the group consisting of:
10 (a) nicosulfuron;
(b) rimsulfuron;
(c) primisulfuron;
(d) imazethapyr;
(e) chlorsulfuron;
15 (f) chlorimuron ethyl;
(g) triasulfuron;
(h) flumetsulam; and
(i) imazaquin.

36. The method of any one of claims 25-34 wherein the herbicide is
20 selected from the pigment synthesis-inhibiting class of herbicides and is selected from the group consisting of:
(a) isoxaflutole;
(b) topramezone;
(c) sulcatrione; and
25 (d) tembotrione.

37. The method of any one of claims 25-34 wherein the herbicide is selected from the PPO-inhibiting class of herbicides and is selected from the group consisting of:
(a) acifluofen;
30 (b) flumioxan; and
(c) sulfentrazone.

38. The method of any one of claims 25-34 wherein the herbicide is selected from the PS II-inhibiting class of herbicides and is selected from the group consisting of:
35 (a) diuron;
(b) linuron;
(c) bentazon; and
(d) chlorotoluron.

39. The method of any one of claims 25-34 wherein the herbicide is dicamba.

40. A method of determining the presence of the polynucleotide of Claim 1 in a plant, comprising at least one of:

5 (a) isolating nucleic acid molecules from said plant and determining if an *Nsf1* gene is present by attempting to amplify sequences homologous to the polynucleotide of Claim 1, or

(b) isolating nucleic acid molecules from said plant and performing a Southern or northern hybridization, or

10 (c) isolating proteins from said plant and performing a western blot using antibodies to the NSF1 protein, or

(d) isolating proteins from said plant and performing an ELISA assay using antibodies to the NSF1 protein, thereby determining the presence of the polynucleotide of Claim 1 in said plant.

15 41. A method of determining the presence of the *Nsf1* locus in a plant, comprising at least one of:

(a) isolating nucleic acid molecules from said plant and determining if an *Nsf1* gene is present by attempting to amplify sequences homologous to the polynucleotide of Claim 1, or

20 (b) isolating nucleic acid molecules from said plant and performing a Southern or northern hybridization, or

(c) isolating proteins from said plant and performing a western blot using antibodies to the NSF1 protein, or

(d) isolating proteins from said plant and performing an ELISA assay

25 using antibodies to the NSF1 protein, thereby determining the presence of the *Nsf1* locus in said plant.

42. A soybean plant comprising the polynucleotide of claim 1, wherein said soybean plant also exhibits soybean cyst nematode resistance, wherein said polynucleotide has been incorporated through transformation or plant breeding

30 techniques, and wherein said soybean plant has been bred from germplasm selected from the group consisting of:

(a) Peking;

(b) PI88788;

35 (c) PI89772;

(d) PI90763;

(e) PI209332;

(f) PI404189A;

(g) PI437654;

(h) PI438489B;

(i) PI467312;

(j) PI468916;

5 (k) Hartwig;

(l) J87-233; and

(m) progeny derived from sources (a) through (l).

Figure 1a

```
 Symbol comparison table: blosum62.cmp CompCheck: 1102

                   GapWeight: 8
             GapLengthWeight: 2

 AAK38079_pileup_158748.txt  MSF: 750  Type: P  March 7, 2006 17:35
Check: 2632 ..

XP_469851         SEQ ID NO: 12
XP_469852         SEQ ID NO: 13
aak38080          SEQ ID NO: 5
aak38081          SEQ ID NO: 7
BAD27508          SEQ ID NO: 8
aak38079          SEQ ID NO: 6
BAD27507          SEQ ID NO: 9
BAD27506          SEQ ID NO: 10
rice_nsf_hom      SEQ ID NO: 4
xp_469850         SEQ ID NO: 3
XP_469849         SEQ ID NO: 11
nsf_peptide       SEQ ID NO: 2

//
             1                                                  50
   XP_469851 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   XP_469852 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38080 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38081 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27508 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38079 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27507 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27506 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   xp_469850 maflgwavdi ardsgasssv vltcdgygsa lyfspwdsvp lpataspddg
   XP_469849 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 nsf_peptide ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

             51                                                100
   XP_469851 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   XP_469852 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38080 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38081 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27508 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    aak38079 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27507 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    BAD27506 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   xp_469850 fllprfpdvc vqrsqftnhl apangtgggg srtgvkeeas evlswpptsk
   XP_469849 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 nsf_peptide ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 1b

```
             101                                                150
  XP_469851 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  XP_469852 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   aak38080 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   aak38081 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   BAD27508 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   aak38079 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   BAD27507 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   BAD27506 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  xp_469850 qsvrrlevae hwyrlyktdn qrlspdsqqv svlaeshcdl asgnwkeisi
  XP_469849 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
nsf_peptide ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

             151                                                200
  XP_469851 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  XP_469852 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   aak38080 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   aak38081 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   BAD27508 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   aak38079 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   BAD27507 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   BAD27506 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  xp_469850 hhkkmpsstt tktttpsrda wivsarsdpf hllleaqapl gikadalsqi
  XP_469849 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
nsf_peptide ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

             201                                                250
  XP_469851 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAvfS iAiLFLLvdY frcrrrrgsg
  XP_469852 ~~~~~~~~~~ ~~~~~~~~~m vKAY.IAifS iAvLlLi..h fLfrrr...g
   aak38080 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAiLS cAfLFLvH.Y vLGk..vsdg
   aak38081 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAiLS cAflflvH.Y vLGk..vsdg
   BAD27508 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAiLS cAfLFLvH.Y vLGk..vsdg
   aak38079 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAiLS sAfLFLvH.Y vLGk..vsdg
   BAD27507 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAiLS cAfLFLvH.Y vLgk..vshg
   BAD27506 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAiLS cAfLFLvH.Y vLgk..vshg
rice_nsf_hom ~~~~~~~~~~ ~~~~~~~~~m DnAYiIAiLS vAiLFLLHyY LLgr..gngg
  xp_469850 aavhqshrnt shirelslam DnAYiIAiLS vAiLFLLHyy LLgr..gngg
  XP_469849 ~~~~~~~~~~ ~~~~~~~~~m DKAY.IAvfS iviLFLLvdY Lrrl..rggg
nsf_peptide ~~~~~~~~~~ ~~~~~~~~~M DKAY.IAALS AAALFLLH.Y LLGRRAGGEG

             251                                                300
  XP_469851 snngenkgml qLPPSPPAIP FfGHLHLidk PlHaALsRLA eRHGPVFSlR
  XP_469852 rsng.....m pLPPSPPAIP FfGHLHLidk PfHaALsRLA eRHGPVFSlR
   aak38080 .rrgK.kgav qLPPSPPAvP FLGHLHLVdk PiHatmcRLA ARlGPVFSlR
   aak38081 .rrgK.kgav qLPPSPPAvP FLGHLHLVdk PiHatmcRLA ARlGPVFSlR
   BAD27508 .rrgK.kgav qLPPSPPAIP FiGHLHLVek PiHatmcRLA ARlGPVFSlR
   aak38079 .rrgK.kgav qLPPSPPAvP FLGHLHLVek PiHatmcRLA ARlGPVFSlR
   BAD27507 .rrgK.kgav qLPPSPPAIP FiGHLHLVek PiHatmcRLA ARlGPVFSlR
   BAD27506 .rrgK.kgav qLPPSPPAIP FiGHLHLVek PiHatmcRLA ARlGPVFSlR
rice_nsf_hom aA........ RLPPgPPAvP ilGHLHLVKk PmHatmsRLA eRyGPVFSlR
  xp_469850 aA........ RLPPgPPAvP ilGHLHLVKk PmHatmsRLA eRyGPVFSlR
  XP_469849 tsngK.nkgm RLPPglPAvP iiGHLHLVKk PmHatLsRLA ARHGPVFSlR
nsf_peptide KAKAK.GSRR RLPPSPPAIP FLGHLHLVKA PFHGALARLA ARHGPVFSMR
```

Figure 1c

```
             301                                                        350
   XP_469851 LGsRnAVVVS SPeCARECFT dnDVcFANRP qFPSqmpAtF ygAgfgfanY
   XP_469852 LGsRnAVVVS SPeCARECFT dnDVcFANRP rFPSqmLAtF nGtsLgsanY
    aak38080 LGsRRAVVVS SseCARECFT EHDVtFANRP kFPSqlLvSF nGtaLvtSSY
    aak38081 LGsRRAVVVS SseCARECFT EHDVtFANRP kFPSqlLvSF nGtaLvtSSY
    BAD27508 LGsRRAVVVp SseCARECFT EHDVtFANRP kFPSqlLASF nGtaLvtSSY
    aak38079 LGsRRAVVVS SseCARECFT EHDVtFANRP kFPSqlLvSF nGtaLvtSSY
    BAD27507 LGsRRAVVVS SseCARECFT EHDVtFANRP ssrrklLASF nGtaLvtSSY
    BAD27506 LGsRRAVVVS SseCARECFT EHDVtFANRP kFPSqlLASF nGtaLvtpSY
rice_nsf_hom LGsRRAVVVS SPgCARECFT EHDVtFANRP rFeSqlLvSF nGAaLataSY
   xp_469850 LGsRRAVVVS SPgCARECFT EHDVtFANRP rFeSqlLvSF nGAaLataSY
   XP_469849 LGsRRAVVVS SPgCARECFT EHDVaFANRP rFeSqlLmSF DGtaLamaSY
 nsf_peptide LGTRRAVVVS SPDCARECFT EHDVNFANRP LFPSMRLASF DGAMLSVSSY

             351                                                        400
   XP_469851 GahWRNLRRi AtVhLLSAhR VrgMAgvvsg eiRpMVqRMy RAAAAagvGV
   XP_469852 GPhWRNLRRi AtVhLLSshR VsgMsgiIsg QaRhMVRRMy RAAtAsaaGV
    aak38080 GPhWRNLRRV AtVQLLSAhR VaCMsgvIaA eVRAMaRRlf hAteAspdGa
    aak38081 GPhWRNLRRV AtVQLLSAhR VaCMsgvIaA eVRAMaRRlf hAAeAspdGa
    BAD27508 GPhWRNLRRV AtVQLLSAhR VaCMsgvIaA eVRAMaRRlf hAAeAspdGa
    aak38079 GPhWRNLRRV AtVQLLSAhR VtCMsgvIaA eVRAMaRRlf hAAeAspdGa
    BAD27507 GPhWRNLRRV AtVQLLSAhR VaCMsgvIaA eVRAMaRRlf hAAeAspdGa
    BAD27506 GPhWRNLRRV AtVQLLSAhR VaCMsgvIaA eVRAMaRRlf hAAeAspgGa
rice_nsf_hom GahWRNLRRi vAVQLLSAhR VGlMsglIag eVRAMVRRMy RAAAAspaGa
   xp_469850 GahWRNLRRi vAVQLLSAhR VGlMsglIag eVRAMVRRMy RAAAAspaGa
   XP_469849 GPhWRNLRRV AAVQLLSArR VGlMsglIag eVRAMVRslc R....rpaaa
 nsf_peptide GPYWRNLRRV AAVQLLSAHR VGCMAPAIEA QVRAMVRRMD RAAAAGGGGV

             401                                                        450
   XP_469851 ARVQLKRRLF ELSLSVLMEa IAqTKttRpE adDADtDMSv EAqEFKnvlD
   XP_469852 ARVQLnRRLF ELSLSVLMEa IAqsKTtRrE apDADtDMSm EAqElrhvlD
    aak38080 ARVQLKRRLF ELSLSVLMET IAqTKatRsE .ADADtDMSv EAqEFKevVD
    aak38081 ARVQLKRRLF ELSLSVLMET IAqTKatRsE .ADADtDMSv EAqEFKevVD
    BAD27508 ARVQLKRRLF ELSLSVLMET IAqTKatRsE .ADADtDMSv EAqEFKevVD
    aak38079 ARVQLKRRLF ELSLSVLMET IAqTKatRsE .ADADtDMSl EAqEFKevVD
    BAD27507 tRVQLKRRLF ELSLSVLMET IAqTKatRsE .ADADtDMSv EAqEFKevVD
    BAD27506 ARVQLKRgpF ELSLSVLMET IAqTKatRsE .aDADtDMSv EAqEFKevVD
rice_nsf_hom ARiQLKRRLF EvSLSVLMET IAHTKatRpE .tDpDtDMSv EAqEFKQvVD
   xp_469850 ARiQLKRRLF EvSLSVLMET IAHTKatRpE .tDpDtDMSv EAqEFKQvVD
   XP_469849 ApVQLKRRLF ELSLSVLMET IAqsKatRpE ttDtDtDMSm EAqEyKQvVe
 nsf_peptide ARVQLKRRLF ELSLSVLMET IAHTKTSRAE .ADADSDMST EAHEFKQIVD

             451                                                        500
   XP_469851 ELnPllGaAN lWDYLPaLRv FDVlGVkrKI atlanRRDAF vrRLIDaERq
   XP_469852 ELnPlIGaAN lWDYLPaLRW FDVFGVkrKI vaAVnRRnAF mrRLIDaERq
    aak38080 kLiPhlGaAN mWDYLPVmRW FDVFGVRNKI LhAVsRRDAF LrRLIDaERR
    aak38081 kLiPhlGaAN mWDYLPVmRW FDVFGVRNKI LhAVsRRDAF LrRLIDaERR
    BAD27508 kLiPhlGaAN mWDYLPVmRW FDVFGVRNKI LhAVsRRDAF LrRLIDaERR
    aak38079 kLiPhlGaAN mWDYLPVmRW FDVFGVRsKI LhAVsRRDAF LrRLInaERR
    BAD27507 kLiPhlGaAN mWDYLPVmRW FDVFGVRNKI LhAVsRRDAF LrRLIDaERR
    BAD27506 kpiPhlGaAN mWDYLPVmRW FDVFGVRNKI LhAVsRRDAF LrRLIDaERR
rice_nsf_hom EiiPhIGaAN lWDYLPaLRW FDVFGVRrKI LaAVsRRDAF LrRLIDaERR
   xp_469850 EiiPhIGaAN lWDYLPaLRW FDVFGVRrKI LaAVsRRDAF LrRLIDaERR
   XP_469849 EilerIGTgN lcDYLPaLRW FDVFGVRNrI LaAVsRRDAF LrRLIyaaRw
 nsf_peptide ELVPYIGTAN RWDYLPVLRW FDVFGVRNKI LDAVGRRDAF LGRLIDGERR
```

Figure 1d

```
              501                                                  550
    XP_469851 Rmdng.vDgG DdgEkKSvIs VLLsLQKtEP EVYkDivIvn LCAaLFaAGT
    XP_469852 RmdnndvDgG DdgEkKSMIs VLLTLQKtqP EVYTdTlImt LCApLFGAGT
     aak38080 RL....aDgG sdgdkKSMIA VLLTLQKtEP kVYTDTmITA LCANLFGAGT
     aak38081 RL....aDgG sdgdkKSMIA VLLTLQKtEP kVYTDTmITA LCANLFGAGT
     BAD27508 RL....aDgG sdgdkKSMIA VLLTLQKtEP kVYTDTmITA LCANLFGAGT
     aak38079 RL....aDgG sdgdkKSMIA VLLTLQKtEP kVYTDTmITA LCANLFGAGT
     BAD27507 RL....aDgG sdgdkKSMIA VLLTLQKtEP kVYTDTmITA LCANLFGAGT
     BAD27506 RL....aDgG sdgdkKSMIA VLLTLQKtEP kVYTDTmITA LCANLFGAGT
 rice_nsf_hom RL.....DdG DEgEkKSMIA VLLTLQKtEP EVYTDnmITA LtANLFGAGT
    xp_469850 RL.....DdG DEgEkKSMIA VLLTLQKtEP eVYTDnmITA LtANLFGAGT
    XP_469849 Rm.....Dd. ..gEkKSMIA VLLTLQKtqP EVYTDnmITA LCsNLlGAGT
  nsf_peptide RL.....DAG DESESKSMIA VLLTLQKSEP EVYTDTVITA LCANLFGAGT

              551                                                  600
    XP_469851 ETTamTiEWA MSLLLNHpki LKKAkAEIDA sVGnSRLing DDmPHLsYLQ
    XP_469852 ETTSTTiEWA MSLLLNHpEi LKKAQAEIDm sVGnSRLisv vDVhrLgYLQ
     aak38080 ETTSTTTEWA MSLLLNHpaA LKKAQAEIDA sVGTSRLVsv DDVPsLaYLQ
     aak38081 ETTSTTTEWA MSLLLNHpaA LKKAQAEIDA sVGTSRLVsv DDVPsLaYLQ
     BAD27508 ETTSTTTEWA MSLLLNHpaA LKKAQAEIDA sVGTSRLVsv DDVPsLaYLQ
     aak38079 ETTSTTTEWA MSLLLNHpaA LKKAQAEIDA sVGTSRLVsv DDVPsLaYLQ
     BAD27507 ETTSTTTEWA MSLLLNHpaA LKKAQAEIDA sVGTSRLVsv DDVlsLaYLQ
     BAD27506 ETTSTTTErA MSLLLNHpaA LKKAQAEIDA sVGTSRLVsv DDmPsLaYLQ
 rice_nsf_hom ETTSTTsEWA MSLLLNHpdt LKKAQAEIDA sVGnSRLiTA DDVtrLgYLQ
    xp_469850 ETTSTTsEWA MSLLLNHpdt LKKAQAEIDA sVGnSRLiTA DDVtrLgYLQ
    XP_469849 ETTSTTiEWA MSLLLNHpEt LKKAQAEIDA sVGnSRLiTA DDVPriTYLQ
  nsf_peptide ETTSTTTEWA MSLLLNHREA LKKAQAEIDA AVGTSRLVTA DDVPHLTYLQ

              601                                                  650
    XP_469851 CIinETLRLy PvAPLLiPHE SsADCkVnGY hiPsGTMLLV NViAiqRDPm
    XP_469852 CIinETLRmy PAAPLLLPHE SsADCkVGGY hiPsGaMLLV NVaAiqRDPv
     aak38080 CIVsETLRLy PAAPLLLPHE SsADCkVGGY nVPadTMLiV NayAiHRDPA
     aak38081 CIVnETLRLy PAAPLLLPHE SsADCkVGGY nVPadTMLiV NayAiHRDPA
     BAD27508 CIVnETLRLy PAAPLLLPHE SsADCkVGGY nVPadTMLiV NayAiHRDPA
     aak38079 CIVsETLRLy PAAPLLLPHE SsADCkVGGY nVPadTMLiV NayAiHRDPA
     BAD27507 CIVsETLRLy PAAPLLLPHE SsADCkVGGY nVPadTMLiV NayAiHRDPA
     BAD27506 CIVnETLRLy PAAPLLLPHE SsADCkVGGY nVPadTMLiV NayAiHRDPA
 rice_nsf_hom CIVrETLRLy PAAPmLLPHE SsADCkVGGY niPRGsMLLi NayAiHRDPA
    xp_469850 CIVrETLRLy PAAPmLLPHE SsADCkVGGY niPRGsMLLi NayAiHRDPA
    XP_469849 CIVrETLRLy PAAPmLiPHE SsADCeVGGY sVPRGTMLLV NayAiHRDPA
  nsf_peptide CIVDETLRLH PAAPLLLPHE SAADCTVGGY DVPRGTMLLV NVHAVHRDPA

              651                                                  700
    XP_469851 VWkePneFkP ERFE..nGes EGlfmiPFGM GRRKCPGETm ALqTiGLVLg
    XP_469852 iWkePseFkP ERFE..nGrf EGlfmiPFGM GRRrCPGEmL ALqTiGLVLg
     aak38080 aWEdPleFrP ERFE..dGKA EGlfmiPFGM GRRrCPGETL ALRTiGmVLA
     aak38081 aWEhPlvFrP ERFE..dGKA EGlfmiPFGv GRRrCPGETL ALRTismVLA
     BAD27508 aWEhPleFrP ERFE..dGKA EGlfmiPFGv GRRrCPGETL ALRTismVLA
     aak38079 aWEdPleFkP ERFE..dGKA EGlfmiPFGM GRRrCPGETL ALRTiGmVLA
     BAD27507 aWEhPleFrP ERFE..dGKA EGlfmiPFGM GRRrCPGETL ALRTiGmVLA
     BAD27506 aWEhPleFrP ERFE..dGKA EGlfmiPFGM GRRrCPGETL ALRTiGmVLA
 rice_nsf_hom VWEePekFmP ERFE..dGgc dGnLLMPFGM GRRrCPGETL ALRTVGLVLg
    xp_469850 VWEePekFmP ERFE..dGgc dGnLLMPFGM GRRrCPGETL ALRTVGLVLg
    XP_469849 aWEePeRFVP ERFE..GGgc dGnLsMPFGM GRRrCPGETL ALhTVGLVLg
  nsf_peptide VWEDPDRFVP ERFEGAGGKA EGRLLMPFGM GRRKCPGETL ALRTVGLVLA
```

Figure 1e

```
             701                                                    750
   XP_469851 aLiQCFDWDr VDGAeVDMtq gsGLTnPRAV PLEAMCkPRe AMsdVfReLl
   XP_469852 TmiQCFDWgr VDdAmVDMtq SnGLTslkvi PLEAMCkPRe AMcdVLRkfm
    aak38080 TLvQCFDWep VDGvkVDMte gGGfTiPkAV PLEAvCRPRa vMRdVLqnL~
    aak38081 TLvQCFDWep VDGvnVDMte gGGfTiPkAV PLEAvCRPRa vMRdVLqsi~
    BAD27508 TLvQCFDWep VDGvkVDMte gGGfTiPkAV PLEAvCRPRa vMRdVLqnL~
    aak38079 TLvQCFDWep VDGvkVDMte gGGfTiPkAV PLEAvCRPRv vMRdVLqnL~
    BAD27507 TLvQCFDWep VDGvkVDMte gGGfTiPkAV PLEAvCRPRt vMRdVLqnL~
    BAD27506 TLvQCFDWep VDGvkVDMte gGGfTiPkAV PLEAvCRPRa vMRdVLqnL~
rice_nsf_hom TLiQCFDWer VDGveVDMte gGGLTiPkvV PLEAMCRPRd AMgGVLReLv
   xp_469850 TLiQCFDWer VDGveVDMte gGGLTiPkvV PLEAMCRPRd AMgGVLReLv
   XP_469849 TLiQCFDWer VDGveVDMae gGGLTMPkvV PLEAvCRPRd AMgGVLReL~
 nsf_peptide TLLQCFDWDT VDGAQVDMKA SGGLTMPRAV PLEAMCRPRT AMRGVLKRL~
```

Figure 2a

```
Symbol comparison table: blosum62.cmp CompCheck: 1102

                 GapWeight: 8
           GapLengthWeight: 2

  F2_peptide = SEQ ID NO: 22
 nsf_peptide = SEQ ID NO: 2
 BMS_peptide = SEQ ID NO: 20
 Q66_Peptide = SEQ ID NO: 18
GA209_Peptid = SEQ ID NO: 24
W703A_Peptid = SEQ ID NO: 26


             1                                                  50
  F2_peptide MDKAYIAALS AAALFLLHYL LGRRAGGEGK AKAKGSRRRL PPSPPAIPFL
 nsf_peptide MDKAYIAALS AAALFLLHYL LGRRAGGEGK AKAKGSRRRL PPSPPAIPFL
 BMS_peptide MDKAYIAALS AAALFLLHYL LGRRAGGEGK AKAKGSRRRL PPSPPAIPFL
 Q66_Peptide MDKAYIAALS AAALFLLHYL LGRRAGGEGK AKAKGSRRRL PPSPPAIPFL
GA209_Peptid MDKAYIAALS AAALFLLHYL LGRRAGVEG. .KAKGSRRRL PPSPPAIPFL
W703A_Peptid MDKAYIAALS AAALFLLHYL LGRRAGVEG. .KAKSSRRRL PPSPPAIPFL

             51                                                100
  F2_peptide GHLHLVKAPF HGALARLAAR HGPVFSMRLG TRRAVVVSSP DCARECFTEH
 nsf_peptide GHLHLVKAPF HGALARLAAR HGPVFSMRLG TRRAVVVSSP DCARECFTEH
 BMS_peptide GHLHLVKAPF HGALARLAAR HGPVFSMRLG TRRAVVVSSP DCARECFTEH
 Q66_Peptide GHLHLVKAPF HGALARLAAR HGPVFSMRLG TRRAVVVSSP DCARECFTEH
GA209_Peptid GHLHLVKAPF HGALARLAAR HGPVFSMRLG TRRAVVVSSP DCARECFTEH
W703A_Peptid GHLHLVKAPF HAALARLAAR HGPVFSMRLG TRRAVVVSSP DCARECFTEH

             101                                               150
  F2_peptide DVNFANRPLF PSMRLASFDG AMLSVSSYGP YWRNLRRVAA VQLLSAHRVG
 nsf_peptide DVNFANRPLF PSMRLASFDG AMLSVSSYGP YWRNLRRVAA VQLLSAHRVG
 BMS_peptide DVNFANRPLF PSMRLASFDG AMLSVSSYGP YWRNLRRVAA VQLLSAHRVG
 Q66_Peptide DVNFANRPLF PSMRLASFDG AMLSVSSYGP YWRNLRRVAA VQLLSAHRVG
GA209_Peptid DVNFANRPLF PSMRLASFDG AMLSVSSYGP YWRNLRRVAA VQLLSAHRVG
W703A_Peptid DVNFANRPLF PSMRLASFDG AMLSVSSYGP YWRNLRRVAA VQLLSAHRVA

             151                                               200
  F2_peptide CMAPAIEAQV RAMVRRMDRA AAAGGGGVAR VQLKRRLFEL SLSVLMETIA
 nsf_peptide CMAPAIEAQV RAMVRRMDRA AAAGGGGVAR VQLKRRLFEL SLSVLMETIA
 BMS_peptide CMAPAIEAQV RAMVRRMDRA AAAGGGGVAR VQLKRRLFEL SLSVLMETIA
 Q66_Peptide CMAPAIEAQV RAMVRRMDRA AAAGGGGVAR VQLKRRLFEL SLSVLMETIA
GA209_Peptid CMAPAIEAQV RAMVRRMDRA AAAGGGGVAR VQLKRRLFEL SLSVLMETIA
W703A_Peptid CMVPAIEAQV RAMVRRMDRA AAAGGARRAR PAQAAAVRAL PQRAHGNHRA

             201                                               250
  F2_peptide HTKTSRAEAD ADSDMSTEA. HEFKQIVDEL VPYIGTANRW DYLPVLRWFD
 nsf_peptide HTKTSRAEAD ADSDMSTEA. HEFKQIVDEL VPYIGTANRW DYLPVLRWFD
 BMS_peptide HTKTSRAEAD ADSDMSTEA. HEFKQIVDEL VPYIGTANRW DYLPVLRWFD
 Q66_Peptide HTKTSRAEAD ANSDMSTEA. HEFKQIVNEL VPYIGTANRW DYLPVLRWFD
GA209_Peptid HTKTSRAEAD ANSDMSTEA. HEFKQIVNEL VPYIGTANCW DYLPVLRWFD
W703A_Peptid HQDVPRQLEX VDRGPRVQAX RQRARAVHRR GQPLG..... ..LPAGAAL.
```

Figure 2b

```
              251                                                  300
   F2_peptide VFGVRNKILD AVGTRDAFLG RLIDGERRR. .LDAGDESES KSMIAVLLTL
  nsf_peptide VFGVRNKILD AVGRRDAFLG RLIDGERRR. .LDAGDESES KSMIAVLLTL
  BMS_peptide VFGVRNKILD AVGRR.AFLG RLIDGERRR. .LDAGDESES KSMIAVLLTL
  Q66_Peptide VFGVRNKILD AVGRRDAFLG RLIDGERRR. .LDAGDESES KSMIAVLLTL
GA209_Peptid VFGVRNKILD AVGRRDAFLG RLIDGERRR. .LDAGDESES KSMIAVLLTL
W703A_Peptid VRRVRREEQD P..RRRGQKG RVPEAAHRRG AAEAGRWRRQ RK~~~~~~~~

              301                                                  350
   F2_peptide QKSEPEVYTD TVITALCANL FGAGTETTST TTEWAMSLLL NHREALKKAQ
  nsf_peptide QKSEPEVYTD TVITALCANL FGAGTETTST TTEWAMSLLL NHREALKKAQ
  BMS_peptide QKSEPEVYTD TVITALCANL FGAGTETTST TTEWAMSLLL NHREALKKAQ
  Q66_Peptide QKSEPEVYTD TVITALCANL FGAGTETTST TTEWAMSLLL NHREALKKAQ
GA209_Peptid QKSEPEVYTD TVITALLATL SARARTTVAR DVENARG..V ENRKI~~~~~
W703A_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

              351                                                  400
   F2_peptide AEIDAAVGTS RLVTADDVPH LTYLQCIVDE TLRLHPAAPL LLPHESAADC
  nsf_peptide AEIDAAVGTS RLVTADDVPH LTYLQCIVDE TLRLHPAAPL LLPHESAADC
  BMS_peptide AEIDAAVGTS RLVTADDVPH LTYLQCIVDE TLRLHPAAPL LLPHESAADC
  Q66_Peptide AEIDAAVGAS RLVTADDVPH LTYLQCIVDE TLRLHPAAPL LLPHESAADC
GA209_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
W703A_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

              401                                                  450
   F2_peptide TVGGYDVPRG TMLLVNVHAV HRDPAVWEDP DRFVPERFEG AGGKAEGRLL
  nsf_peptide TVGGYDVPRG TMLLVNVHAV HRDPAVWEDP DRFVPERFEG AGGKAEGRLL
  BMS_peptide TVGGYDVPRG TMLLVNVHAV HRDPAVWEDP DRFVPERFEG AGGKAEGRLL
  Q66_Peptide TVGGYDVPRG TMLLVNVHAV HRDPAVWEDP DRFVPERFEG AGGKAEGRLL
GA209_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
W703A_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

              451                                                  500
   F2_peptide MPFGMGRRKC PGETLALRTV GLVLATLLQC FDWDTVDGAQ VDMKASGGLT
  nsf_peptide MPFGMGRRKC PGETLALRTV GLVLATLLQC FDWDTVDGAQ VDMKASGGLT
  BMS_peptide MPFGMGRRKC PGETLALRTV GLVLATLLQC FDWDTVDGAQ VDMKASGGLT
  Q66_Peptide MPFGMGRRKC PGETLALRTV GLVLATLLQC FDWDTVDGAQ VDMKASGGLT
GA209_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
W703A_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

              501                    524
   F2_peptide MPRAVPLEAM CRPRTAMRGV LKRL
  nsf_peptide MPRAVPLEAM CRPRTAMRGV LKRL
  BMS_peptide MPRAVPLEAM CRPRTAMRGV LKRL
  Q66_Peptide MPRAVPLEAM CRPRTAMRGV LKRL
GA209_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~
W703A_Peptid ~~~~~~~~~~ ~~~~~~~~~~ ~~~~
```