



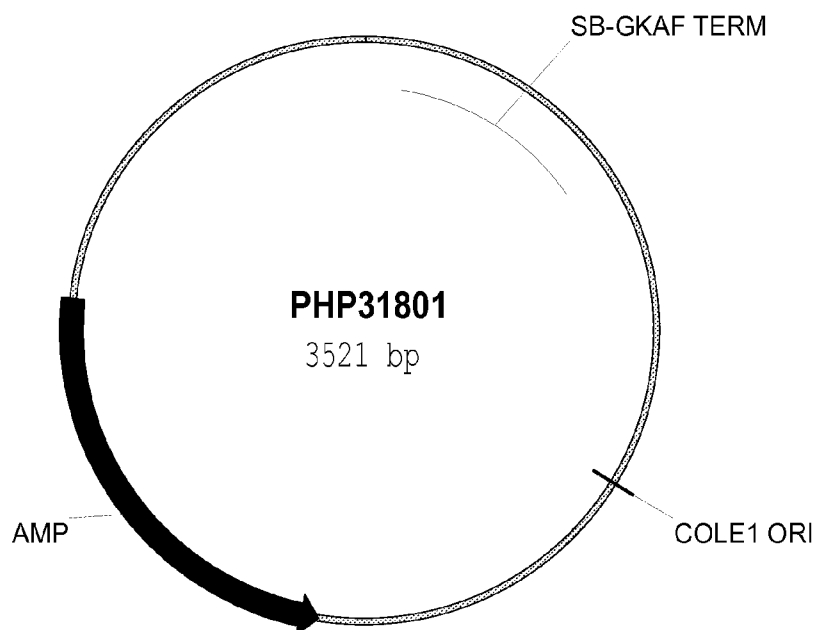
- (51) International Patent Classification:  
C12N 15/82 (2006.01) A01H 5/00 (2006.01)
- (21) International Application Number:  
PCT/US2012/047901
- (22) International Filing Date:  
23 July 2012 (23.07.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
61/514,055 2 August 2011 (02.08.2011) US
- (71) Applicant (for all designated States except US): **PIONEER HI-BRED INTERNATIONAL** [US/US]; 7100 N.W. 62nd Avenue, P.O. Box 1014, Johnston, Iowa 50131 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **ABBITT, Shane E.** [US/US]; 422 N.W. Bayberry Court, Ankeny, Iowa 50023 (US). **JUNG, Rudolf** [DE/US]; 1549 Germania Drive, Des Moines, Iowa 50311 (US).
- (74) Agent: **OROZCO JR, Emil M.**; E. I. du Pont de Nemours and Company, Legal Patent Records Center, 4417 Lancaster Pike, Wilmington, Delaware 19805 (US).

- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: TERMINATOR SEQUENCE FOR GENE EXPRESSION IN PLANTS



(57) Abstract: The present invention discloses polynucleotide sequences that can be used to regulate gene expression in plants. Terminator sequences from *Sorghum bicolor* that are functional in plants are disclosed.

TITLE

## TERMINATOR SEQUENCE FOR GENE EXPRESSION IN PLANTS

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No.

5 61/514,055, filed August 2, 2011, the entire content of which is herein incorporated by reference.

FIELD OF INVENTION

The present invention relates to the field of plant molecular biology and plant genetic engineering. More specifically, it relates to novel plant terminator  
10 sequences and their use to regulate gene expression in plants.

BACKGROUND

Recent advances in plant genetic engineering have opened new doors to engineer plants to have improved characteristics or traits. These transgenic plants characteristically have recombinant DNA constructs in their genome that have  
15 protein coding region operably linked to multiple regulatory regions that allow accurate expression of the transgene. A few examples of regulatory elements that help regulate gene expression in transgenic plants are promoters, introns, terminators, enhancers and silencers.

Plant genetic engineering has advanced to introducing multiple traits into  
20 commercially important plants, also known as gene stacking. This is accomplished by multigene transformation, where multiple genes are transferred to create a transgenic plant that might express a complex phenotype, or multiple phenotypes. But it is important to modulate or control the expression of each transgene optimally. The regulatory elements need to be diverse, to avoid introducing into the same  
25 transgenic plant repetitive sequences, which has been correlated with undesirable negative effects on transgene expression and stability (Peremarti et al (2010) *Plant Mol Biol* 73:363-378; Mette et al (1999) *EMBO J* 18:241-248; Mette et al (2000) *EMBO J* 19:5194-5201; Mourrain et al (2007) *Planta* 225:365-379, US Patent No. 7632982, US Patent No. 7491813, US Patent No. 7674950, PCT Application No.  
30 PCT/US2009/046968). Therefore it is important to discover and characterize novel regulatory elements that can be used to express heterologous nucleic acids in

important crop species. Diverse regulatory regions can be used to control the expression of each transgene optimally.

Regulatory sequences located downstream of coding regions contain signals required for transcription termination and 3' mRNA processing, and are called terminator sequences. The terminator sequences play a key role in mRNA processing, localization, stability and translation (Proudfoot, N. (2004) *Curr. Op. Cell Biol* 16:272-278.; Gilmartin, 2005). The 3' regulatory sequences contained in terminator sequences can affect the level of expression of a gene. Optimal expression of a chimeric gene in plant cells has been found to be dependent on the presence of appropriate 3' sequences (Ingelbrecht, I.L.W. et al (1989) *Plant Cell* 1:671-680). Read through transcription through leaky terminator of a gene can cause unwanted transcription of one transgene from promoter of another one. Also, bidirectional, convergent transcription of transgenes in transgenic plants can occur due to leaky transcription termination of separate convergent genes or from genomic promoters. Convergent, overlapping transcription can decrease transgene expression, or generate antisense RNA (Bieri, S. et al (2002) *Molecular Breeding* 10:107-117). This underlines the importance of discovering novel and efficient transcriptional terminators.

### SUMMARY

The present invention relates to regulatory sequences for modulating gene expression in plants. Specifically, the present invention relates to terminator sequences. Recombinant DNA constructs comprising terminator sequences are provided.

An embodiment of this invention is an isolated polynucleotide sequence comprising: (a) the sequence set forth in SEQ ID NO:1 or SEQ ID NO:18; (b) a sequence with at least 95% sequence identity to SEQ ID NO:1 or SEQ ID NO:18; or (c) a sequence comprising a functional fragment of (a) or (b), wherein the isolated polynucleotide sequence functions as a terminator in a plant cell. Another embodiment of this invention is a recombinant construct comprising an isolated polynucleotide sequence comprising: (a) the sequence set forth in SEQ ID NO:1 or SEQ ID NO:18; (b) a sequence with at least 95% sequence identity to SEQ ID NO:1 or SEQ ID NO:18; or (c) a sequence comprising a functional fragment of (a) or (b),

wherein the isolated polynucleotide sequence functions as a terminator in a plant cell. This recombinant construct may further comprise a promoter and a heterologous polynucleotide, wherein the promoter and the heterologous polynucleotide are operably linked to the isolated polynucleotide sequence.

5 Another embodiment of this invention is a method of expressing a heterologous polynucleotide in a plant, comprising the steps of (a) introducing into a regenerable plant cell the recombinant DNA construct described above; (b) regenerating a transgenic plant from the regenerable plant cell of (a); and (c) obtaining a progeny plant from the transgenic plant of step (b), wherein the  
10 transgenic plant and the progeny plant comprises the recombinant DNA construct and exhibits expression of the heterologous polynucleotide.

In another embodiment, this invention concerns a vector, virus, cell, microorganism, plant, or seed comprising a recombinant DNA construct comprising the terminator sequences described in the present invention.

15 The invention encompasses regenerated, mature and fertile transgenic plants comprising the recombinant DNA constructs described above, transgenic seeds produced therefrom, T1 and subsequent generations. The transgenic plant cells, tissues, plants, and seeds may comprise at least one recombinant DNA construct of interest.

20 In another embodiment, the plant or seed comprising the terminator sequences described in the present invention is a monocotyledonous plant or seed. In another embodiment, the plant or seed comprising the terminator sequences described in the present invention is a maize plant or seed.

In another embodiment, any of the methods of expressing a heterologous  
25 polynucleotide, wherein the plant cell is a monocotyledonous plant cell, e.g., a maize plant cell.

#### **BRIEF DESCRIPTION OF DRAWINGS AND SEQUENCE LISTING**

The invention can be more fully understood from the following detailed description and the accompanying drawings and Sequence Listing which form a  
30 part of this application. The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in Nucleic Acids

Research 13:3021-3030 (1985) and in the Biochemical Journal 219 (No. 2): 345-373 (1984), which are herein incorporated by reference in their entirety. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. § 1.822.

5           FIG.1 shows the map of PHP31801, the vector used for cloning SB-GKAF terminator after amplification.

          FIG.2 shows the map of PHP34074, the vector used for testing the SB-GKAF terminator.

          FIG.3 shows the results of testing SB-GKAF terminator compared to PINII  
10 terminator in transient assays. It shows quantitative analysis of GUS reporter gene expression in BMS cells transformed with PHP34074 (SB-GKAF terminator) and PHP34005 (PINII terminator).

          FIG.4A and FIG.4B show quantitative analysis of GUS reporter gene expression in Gaspe Flint derived maize lines stably transformed with SB-GKAF  
15 (PHP34074) and PINII (PHP34005) terminator constructs. FIG.4A shows GUS reporter gene expression assayed at protein level, and FIG.4B shows GUS reporter gene expression assayed with qRT-PCR.

          FIG.5 shows the results of qRT-PCR assays with stably transformed Gaspe Flint derived maize lines, using two sets of primers downstream of the SB-GKAF  
20 terminator and the PINII terminator.

          FIG.6A – 6C show the alignment between the cloned SB-GKAF terminator (SEQ ID NO:1) and the nucleotides 1863 to 2322 of NCBI GI NO: 671655 (SEQ ID NO:18). The consensus sequence is show at the top, and the residues that match the consensus exactly are boxed.

25           SEQ ID NO:1 is the sequence of the 459 bp SB-GKAF terminator.

          SEQ ID NO:2 and 3 are the sequences of the forward and reverse primers used to amplify SB-GKAF terminator.

          SEQ ID NO:4 is the nucleotide sequence of PHP31801, the vector used for cloning SB-GKAF terminator after PCR amplification.

30           SEQ ID NO:5 is the nucleotide sequence of PHP34074, the vector used for testing SB-GKAF terminator.

SEQ ID NO:6 is the nucleotide sequence of PHP34005, the test vector used as a control with PINII terminator.

SEQ ID NOS:7-9 are the sequences of the forward primer, reverse primer and probe used for assessing GUS expression by qRT-PCR in transgenic maize plants, as described in Table 2.

SEQ ID NOS:10-17 are the sequences of the primers used for quantitating read through transcription through SB-GKAF and PINII terminators, by qRT-PCR in transgenic maize plants, as described in Table 3.

SEQ ID NO:18 corresponds to nucleotides 1863 to 2322 of NCBI GI NO: 671655.

### DETAILED DESCRIPTION

The disclosure of each reference set forth herein is hereby incorporated by reference in its entirety.

As used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a plant" includes a plurality of such plants, reference to "a cell" includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

As used herein:

The terms "monocot" and "monocotyledonous plant" are used interchangeably herein. A monocot of the current invention includes the *Gramineae*.

The terms "dicot" and "dicotyledonous plant" are used interchangeably herein. A dicot of the current invention includes the following families: *Brassicaceae*, *Leguminosae*, and *Solanaceae*.

The terms "full complement" and "full-length complement" are used interchangeably herein, and refer to a complement of a given nucleotide sequence, wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

"Transgenic" refers to any cell, cell line, callus, tissue, plant part or plant, the genome of which has been altered by the presence of a heterologous nucleic acid, such as a recombinant DNA construct, including those initial transgenic events as well as those created by sexual crosses or asexual propagation from the initial

transgenic event. The term "transgenic" as used herein does not encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation.

"Genome" as it applies to plant cells encompasses not only chromosomal DNA found within the nucleus, but organelle DNA found within subcellular components (e.g., mitochondrial, plastid) of the cell.

"Plant" includes reference to whole plants, plant organs, plant tissues, plant propagules, seeds and plant cells and progeny of same. Plant cells include, without limitation, cells from seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores.

"Propagule" includes all products of meiosis and mitosis able to propagate a new plant, including but not limited to, seeds, spores and parts of a plant that serve as a means of vegetative reproduction, such as corms, tubers, offsets, or runners. Propagule also includes grafts where one portion of a plant is grafted to another portion of a different plant (even one of a different species) to create a living organism. Propagule also includes all plants and seeds produced by cloning or by bringing together meiotic products, or allowing meiotic products to come together to form an embryo or fertilized egg (naturally or with human intervention).

"Progeny" comprises any subsequent generation of a plant.

"Transgenic plant" includes reference to a plant which comprises within its genome a heterologous polynucleotide. For example, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant DNA construct.

The commercial development of genetically improved germplasm has also advanced to the stage of introducing multiple traits into crop plants, often referred to as a gene stacking approach. In this approach, multiple genes conferring different characteristics of interest can be introduced into a plant. Gene stacking can be

accomplished by many means including but not limited to co-transformation, retransformation, and crossing lines with different transgenes.

"Transgenic plant" also includes reference to plants which comprise more than one heterologous polynucleotide within their genome. Each heterologous polynucleotide may confer a different trait to the transgenic plant.

"Heterologous" with respect to sequence means 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.

"Polynucleotide", "nucleic acid sequence", "nucleotide sequence", or "nucleic acid fragment" are used interchangeably to refer to a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. Nucleotides (usually found in their 5'-monophosphate form) are referred to by their single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

"Polypeptide", "peptide", "amino acid sequence" 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 residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The terms "polypeptide", "peptide", "amino acid sequence", and "protein" are also inclusive of modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation.

"Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into protein by the cell.

"cDNA" refers to a DNA that is complementary to and synthesized from an mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into the double-stranded form using the Klenow fragment of DNA polymerase I.



“Coding region” refers to the portion of a messenger RNA (or the corresponding portion of another nucleic acid molecule such as a DNA molecule) which encodes a protein or polypeptide. “Non-coding region” refers to all portions of a messenger RNA or other nucleic acid molecule that are not a coding region, including but not limited to, for example, the promoter region, 5’ untranslated region (“UTR”), 3’ UTR, intron and terminator. The terms “coding region” and “coding sequence” are used interchangeably herein. The terms “non-coding region” and “non-coding sequence” are used interchangeably herein.

An “Expressed Sequence Tag” (“EST”) is a DNA sequence derived from a cDNA library and therefore is a sequence which has been transcribed. An EST is typically obtained by a single sequencing pass of a cDNA insert. The sequence of an entire cDNA insert is termed the “Full-Insert Sequence” (“FIS”). A “Contig” sequence is a sequence assembled from two or more sequences that can be selected from, but not limited to, the group consisting of an EST, FIS and PCR sequence. A sequence encoding an entire or functional protein is termed a “Complete Gene Sequence” (“CGS”) and can be derived from an FIS or a contig.

“Mature” protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or pro-peptides present in the primary translation product have been removed.

“Precursor” protein refers to the primary product of translation of mRNA; i.e., with pre- and pro-peptides still present. Pre- and pro-peptides may be and are not limited to intracellular localization signals.

“Isolated” refers to materials, such as nucleic acid molecules and/or proteins, which are substantially free or otherwise removed from components that normally accompany or interact with the materials in a naturally occurring environment. Isolated polynucleotides may be purified from a host cell in which they naturally occur. Conventional nucleic acid purification methods known to skilled artisans may be used to obtain isolated polynucleotides. The term also embraces recombinant polynucleotides and chemically synthesized polynucleotides.

“Recombinant” refers to an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques.

"Recombinant" also includes reference to a cell or vector, that has been modified by the introduction of a heterologous nucleic acid or a cell derived from a cell so modified, but does not encompass the alteration of the cell or vector by naturally occurring events (e.g., spontaneous mutation, natural transformation/transduction/transposition) such as those occurring without deliberate human intervention.

"Recombinant DNA construct" refers to a combination of nucleic acid fragments that are not normally found together in nature. Accordingly, 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, but arranged in a manner different than that normally found in nature. The terms "recombinant DNA construct" and "recombinant construct" are used interchangeably herein.

The terms "entry clone" and "entry vector" are used interchangeably herein.

"Regulatory sequences" or "regulatory elements" are used interchangeably and refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include, but are not limited to, promoters, translation leader sequences, introns, and polyadenylation recognition sequences. The terms "regulatory sequence" and "regulatory element" are used interchangeably herein.

"Promoter" refers to a nucleic acid fragment capable of controlling transcription of another nucleic acid fragment.

"Promoter functional in a plant" is a promoter capable of controlling transcription in plant cells whether or not its origin is from a plant cell.

"Tissue-specific promoter" and "tissue-preferred promoter" are used interchangeably to refer to a promoter that is expressed predominantly but not necessarily exclusively in one tissue or organ, but that may also be expressed in one specific cell.

"Developmentally regulated promoter" refers to a promoter whose activity is determined by developmental events.

Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as “constitutive promoters”.

Inducible promoters selectively express an operably linked DNA sequence in response to the presence of an endogenous or exogenous stimulus, for example by chemical compounds (chemical inducers) or in response to environmental, hormonal, chemical, and/or developmental signals. Examples of inducible or regulated promoters include, but are not limited to, promoters regulated by light, heat, stress, flooding or drought, pathogens, phytohormones, wounding, or chemicals such as ethanol, jasmonate, salicylic acid, or safeners.

“Enhancer sequences” refer to the sequences that can increase gene expression. These sequences can be located upstream, within introns or downstream of the transcribed region. The transcribed region is comprised of the exons and the intervening introns, from the promoter to the transcription termination region. The enhancement of gene expression can be through various mechanisms which include, but are not limited to, increasing transcriptional efficiency, stabilization of mature mRNA and translational enhancement.

An “intron” is an intervening sequence in a gene that is transcribed into RNA and then excised in the process of generating the mature mRNA. The term is also used for the excised RNA sequences. An “exon” is a portion of the sequence of a gene that is transcribed and is found in the mature messenger RNA derived from the gene, and is not necessarily a part of the sequence that encodes the final gene product.

“Operably linked” refers to the association of nucleic acid fragments in a single fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a nucleic acid fragment when it is capable of regulating the transcription of that nucleic acid fragment.

“Expression” refers to the production of a functional product. For example, expression of a nucleic acid fragment may refer to transcription of the nucleic acid fragment (e.g., transcription resulting in mRNA or functional RNA) and/or translation of mRNA into a precursor or mature protein.

“Overexpression” refers to the production of a gene product in transgenic organisms that exceeds levels of production in a null segregating (or non-transgenic) organism from the same experiment.

“Phenotype” means the detectable characteristics of a cell or organism.

5       The term “crossed” or “cross” means the fusion of gametes via pollination to produce progeny (e.g., cells, seeds or plants). The term encompasses both sexual crosses (the pollination of one plant by another) and selfing (self-pollination, e.g., when the pollen and ovule are from the same plant). The term “crossing” refers to the act of fusing gametes via pollination to produce progeny.

10       A “favorable allele” is the allele at a particular locus that confers, or contributes to, a desirable phenotype, e.g., increased cell wall digestibility, or alternatively, is an allele that allows the identification of plants with decreased cell wall digestibility that can be removed from a breeding program or planting (“counterselection”). A favorable allele of a marker is a marker allele that  
15       segregates with the favorable phenotype, or alternatively, segregates with the unfavorable plant phenotype, therefore providing the benefit of identifying plants.

      The term “introduced” means providing a nucleic acid (e.g., expression construct) or protein into a cell. Introduced includes reference to the incorporation of a nucleic acid into a eukaryotic or prokaryotic cell where the nucleic acid may be  
20       incorporated into the genome of the cell, and includes reference to the transient provision of a nucleic acid or protein to the cell. Introduced includes reference to stable or transient transformation methods, as well as sexually crossing. Thus, “introduced” in the context of inserting a nucleic acid fragment (e.g., a recombinant DNA construct/expression construct) into a cell, means “transfection” or  
25       “transformation” or “transduction” and includes reference to the incorporation of a nucleic acid fragment into a eukaryotic or prokaryotic cell where the nucleic acid fragment may be incorporated into the genome of the cell (e.g., chromosome, plasmid, plastid or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (e.g., transfected mRNA).

30       “Suppression DNA construct” is a recombinant DNA construct which when transformed or stably integrated into the genome of the plant, results in “silencing” of a target gene in the plant. The target gene may be endogenous or transgenic to the

plant. "Silencing," as used herein with respect to the target gene, refers generally to the suppression of levels of mRNA or protein/enzyme expressed by the target gene, and/or the level of the enzyme activity or protein functionality. The terms "suppression", "suppressing" and "silencing", used interchangeably herein, include lowering, reducing, declining, decreasing, inhibiting, eliminating or preventing. "Silencing" or "gene silencing" does not specify mechanism and is inclusive, and not limited to, anti-sense, cosuppression, viral-suppression, hairpin suppression, stem-loop suppression, RNAi-based approaches, and small RNA-based approaches.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Sambrook").

"Transcription terminator", "termination sequences", or "terminator" refer to DNA sequences located downstream of a coding sequence, including polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht, I. L., et al., *Plant Cell* 1:671-680 (1989). A polynucleotide sequence with "terminator activity" refers to a polynucleotide sequence that, when operably linked to the 3' end of a second polynucleotide sequence that is to be expressed, is capable of terminating transcription from the second polynucleotide sequence. Transcription termination is the process by which RNA synthesis by RNA polymerase is stopped and both the RNA and the enzyme are released from the DNA template.

Improper termination of an RNA transcript can affect the stability of the RNA, and hence can affect protein expression. Variability of transgene expression is sometimes attributed to variability of termination efficiency (Bieri et al (2002) *Molecular Breeding* 10: 107–117).

The terms "SB-GKAF terminator", "GKAF terminator" and "gamma- kafirin terminator" are used interchangeably herein, and each refers to the sequence encoding the 3' untranslated region (3' UTR) of the *Sorghum Bicolor* gamma-kafirin

gene and about 300 bp of sequence downstream from the 3' UTR. The sequence of the SB-GKAF terminator is given in SEQ ID NO:1. The *Sorghum bicolor* gamma-kafirin gene encodes a gamma-prolamin protein, and the sequence for this gene is given in NCBI GI NO: 671655. Prolamins are the major storage proteins of many cereals. The sorghum gamma-Kafirin, which is the  $\gamma$ -prolamin of sorghum, constitutes about 2-5% of total prolamin in sorghum endosperm, and is composed of a single polypeptide of 27kDa ( de Freitas FA et al (1994) *Mol Gen Genetics* 245(2):177-86).

The present invention encompasses functional fragments and variants of the terminator sequences disclosed herein.

A "functional fragment" of the terminator is defined as any subset of contiguous nucleotides of the terminator sequence disclosed herein, that can perform the same, or substantially similar function as the full length terminator sequence disclosed herein. A "functional fragment" with substantially similar function to the full length terminator disclosed herein refers to a functional fragment that retains the ability to terminate transcription largely at the same level as the full-length terminator sequence. A recombinant construct comprising a heterologous polynucleotide operably linked to a "functional fragment" of the terminator sequence disclosed herein exhibits levels of heterologous polynucleotide expression substantially similar to a corresponding recombinant construct comprising a heterologous polynucleotide operably linked to the full length terminator sequence. A "variant", as used herein, is the sequence of the terminator or the sequence of a functional fragment of a terminator containing changes in which one or more nucleotides of the original sequence is deleted, added, and/or substituted, while substantially maintaining terminator function. One or more base pairs can be inserted, deleted, or substituted internally to a terminator, without affecting its activity. Fragments and variants can be obtained via methods such as site-directed mutagenesis and synthetic construction.

These terminator functional fragments may comprise at least 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425 or 450 contiguous nucleotides of the particular terminator nucleotide sequence disclosed herein. Such fragments may be obtained by use of restriction enzymes to cleave the naturally

occurring terminator nucleotide sequences disclosed herein; by synthesizing a nucleotide sequence from the naturally occurring terminator DNA sequence; or may be obtained through the use of PCR technology. See particularly, Mullis et al., *Methods Enzymol.* 155:335-350 (1987), and Higuchi, R. In PCR Technology: Principles and Applications for DNA Amplifications; Erlich, H. A., Ed.; Stockton Press Inc.: New York, 1989. Again, variants of these terminator fragments, such as those resulting from site-directed mutagenesis, are encompassed by the compositions of the present invention.

The terms "substantially similar" and "corresponding substantially" as used herein refer to nucleic acid fragments, particularly terminator sequences, wherein changes in one or more nucleotide bases do not substantially alter the ability of the terminator to terminate transcription. These terms also refer to modifications, including deletions and variants, of the nucleic acid sequences of the instant invention by way of deletion or insertion of one or more nucleotides that do not substantially alter the functional properties of the resulting terminator relative to the initial, unmodified terminator. It is therefore understood, as those skilled in the art will appreciate, that the invention encompasses more than the specific exemplary sequences.

As will be evident to one of skill in the art, any heterologous polynucleotide of interest can be operably linked to the terminator sequences described in the current invention. Examples of polynucleotides of interest that can be operably linked to the terminator sequences described in this invention include, but are not limited to, polynucleotides comprising regulatory elements such as introns, enhancers, promoters, translation leader sequences, protein coding regions such as disease and insect resistance genes, genes conferring nutritional value, genes conferring yield and heterosis increase, genes that confer male and/or female sterility, antifungal, antibacterial or antiviral genes, and the like. Likewise, the terminator sequences described in the current invention can be used to terminate transcription of any nucleic acid that controls gene expression. Examples of nucleic acids that could be used to control gene expression include, but are not limited to, antisense oligonucleotides, suppression DNA constructs, or nucleic acids encoding transcription factors.

A recombinant DNA construct (including a suppression DNA construct) of the present invention may comprise at least one regulatory sequence. In an embodiment of the present invention, the regulatory sequences disclosed herein can be operably linked to any other regulatory sequence.

5 A number of promoters can be used in recombinant DNA constructs of the present invention. The promoters can be selected based on the desired outcome, and may include constitutive, tissue-specific, inducible, or other promoters for expression in the host organism.

10 The terms “real-time PCR”, “quantitative PCR”, “quantitative real-time PCR”, and “QPCR” are used interchangeably herein, and represent a variation of the standard polymerase chain reaction (PCR) technique used to quantify DNA or RNA in a sample. Using sequence-specific primers and a probe, the relative number or copies of a particular DNA or RNA sequence are determined. The term relative is used since this technique compares relative copy numbers between different genes  
15 with respect to a specific reference gene. The quantification arises by measuring the amount of amplified product at each cycle during the PCR process.

Quantification of amplified product is obtained using fluorescent hydrolysis probes that measure increasing fluorescence for each subsequent PCR cycle. The Ct (cycle threshold) is defined as the number of cycles required for the fluorescent  
20 signal to cross the threshold (i.e., exceeds background level). DNA/RNA from genes with higher copy numbers will appear after fewer PCR cycles; so the lower a Ct value, the more copies are present in the specific sample. To quantify RNA, QPCR or real-time PCR is preceded by the step of reverse transcribing mRNA into cDNA. This is referred to herein as “real-time RT-PCR” or “quantitative RT-PCR” or  
25 “qRT-PCR”.

The Taqman method of PCR product quantification uses a fluorescent reporter probe. This is more accurate since the probe is designed to be sequence-specific and will only bind to the specific PCR product. The probe specificity allows for quantification even in the presence of non-specific DNA amplification. This  
30 allows for multiplexing, which quantitates several genes in the same tube, by using probes with different emission spectra. Breakdown of the probe by the 5' to 3'



exonuclease activity of Taq polymerase removes the quencher and allows the PCR product to be detected.

When plotted on a linear scale, the fluorescent emission increase with PCR cycle number has a sigmoidal shape with an exponential phase and a plateau phase. The plateau phase is determined by the amount of primer in the master mix rather than the nucleotide template. Usually the vertical scale is plotted in a logarithmic fashion, allowing the intersection of the plot with the threshold to be linear and more easily visualized. Theoretically, the amount of DNA doubles every cycle during the exponential phase, but this is affected by the efficiency of the primers used. A positive control using a reference gene, e.g., a "housekeeping" gene that is relatively abundant in all cell types, is also performed to allow for comparisons between samples. The amount of DNA/RNA is determined by comparing the results to a standard curve produced by serial dilutions of a known concentration of DNA/RNA.

The present invention includes a polynucleotide comprising: (i) a nucleic acid sequence of at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity, based on the Clustal V (or Clustal W) method of alignment, when compared to SEQ ID NO:1 or SEQ ID NO:18; or (ii) a nucleic acid sequence of at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, based on the Clustal V (or Clustal W) method of alignment, when compared to a functional fragment of SEQ ID NO:1 or SEQ ID NO:18; or (iii) a full complement of the nucleic acid sequence of (i) or (ii), wherein the polynucleotide acts as a terminator in a plant cell.

Sequence alignments and percent identity calculations may be determined using a variety of comparison methods designed to detect homologous sequences including, but not limited to, the Megalign® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI). Unless stated otherwise, multiple alignment of the sequences provided herein were performed using the Clustal V method of alignment (Higgins and Sharp (1989) *CABIOS*. 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal V method are KTUPLE=1,

GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences, using the Clustal V program, it is possible to obtain "percent identity" and "divergence" values by  
5 viewing the "sequence distances" table on the same program; unless stated otherwise, percent identities and divergences provided and claimed herein were calculated in this manner.

Alternatively, the Clustal W method of alignment may be used. The Clustal W method of alignment (described by Higgins and Sharp, *CABIOS*. 5:151-153  
10 (1989); Higgins, D. G. et al., *Comput. Appl. Biosci.* 8:189-191 (1992)) can be found in the MegAlign™ v6.1 program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, Wis.). Default parameters for multiple alignment correspond to GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergent Sequences=30%, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet  
15 Series, DNA Weight Matrix=IUB. For pairwise alignments the default parameters are Alignment=Slow-Accurate, Gap Penalty=10.0, Gap Length=0.10, Protein Weight Matrix=Gonnet 250 and DNA Weight Matrix=IUB. After alignment of the sequences using the Clustal W program, it is possible to obtain "percent identity" and  
"divergence" values by viewing the "sequence distances" table in the same program.

20 Embodiments of the invention include:

The present invention relates to terminator sequences. Recombinant DNA constructs comprising terminator sequences are provided.

An embodiment of this invention is an isolated polynucleotide sequence comprising (a) the sequence set forth in SEQ ID NO:1 or SEQ ID NO:18; (b) a  
25 sequence with at least 95% sequence identity to SEQ ID NO:1 or SEQ ID NO:18; or (c) a sequence comprising a functional fragment of (a) or (b), wherein the isolated polynucleotide sequence functions as a terminator in a plant cell. In another aspect, this invention concerns a recombinant DNA construct comprising a promoter, at least one heterologous nucleic acid fragment, and any terminator, or combination of  
30 terminator elements, of the present invention, wherein the promoter, at least one heterologous nucleic acid fragment, and terminator(s) are operably linked.

In another embodiment, a functional fragment may comprise at least 450, 425, 400, 375, 350, 325, 300, 275, 250, 225, 200, 175 or 150 contiguous nucleotides of SEQ ID NO:1 or SEQ ID NO:18.

5 Recombinant DNA constructs can be constructed by operably linking the nucleic acid fragment of the invention, the terminator sequences set forth in SEQ ID NO:1, or 18 or a functional fragment of the nucleotide sequence set forth in SEQ ID NO:1, or 18, to a heterologous nucleic acid fragment.

10 Another embodiment is a method for transforming a cell (or microorganism) comprising transforming a cell (or microorganism) with any of the isolated polynucleotides or recombinant DNA constructs of the present invention. The cell (or microorganism) transformed by this method is also included. In particular embodiments, the cell is eukaryotic cell, e.g., a yeast, insect or plant cell, or prokaryotic, e.g., a bacterial cell. The microorganism may be *Agrobacterium*, e.g. *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*.

15 Another embodiment of this invention is a method of expressing a heterologous polynucleotide in a plant, comprising the steps of introducing into a regenerable plant cell the recombinant DNA construct described above and regenerating a transgenic plant from the transformed regenerable plant cell, wherein the transgenic plant comprises the recombinant DNA construct and exhibits  
20 expression of the heterologous polynucleotide.

Another embodiment of this invention is a method of expressing a heterologous polynucleotide in a plant, comprising the steps of introducing into a regenerable plant cell the recombinant DNA construct described above; regenerating a transgenic plant from the regenerable plant cell described above;  
25 and obtaining a progeny plant from the transgenic plant, wherein the transgenic plant and the progeny plant comprises the recombinant DNA construct and exhibits expression of the heterologous polynucleotide.

In another embodiment, any of the methods of expressing a heterologous polynucleotide, wherein the plant cell is a monocotyledonous or dicotyledonous  
30 plant cell, for example, a maize or soybean plant cell. The plant cell may also be from sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, millet, sugar cane or switchgrass.

In another embodiment, this invention concerns a vector, virus, cell, microorganism, plant, or seed comprising a recombinant DNA construct comprising the terminator sequences described in the present invention.

The invention encompasses regenerated, mature and fertile transgenic plants comprising the recombinant DNA constructs described above, transgenic seeds produced therefrom, T1 and subsequent generations. The transgenic plant cells, tissues, plants, and seeds may comprise at least one recombinant DNA construct of interest.

In one embodiment, the plant (or seed derived from the plant) comprising the terminator sequences described in the present invention is a monocotyledonous or dicotyledonous plant, for example, a maize or soybean plant. The plant may also be sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, millet, sugar cane or switchgrass. The plant may be an inbred plant or a hybrid plant.

### EXAMPLES

The present invention is further illustrated in the following Examples, in which 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 this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Furthermore, various modifications 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.

#### EXAMPLE 1

##### Amplification and Cloning of a *Sorghum bicolor*

##### Gamma-Kafirin Terminator Sequence

Primers (SEQ ID NOS:2 and 3) were designed for amplifying the terminator of gamma-Kafirin gene from *Sorghum bicolor* (SB-GKAF) based on the *Sorghum bicolor* genomic sequence database. The primer sequences are given below, the underlined region is not homologous with genomic template:

TMS2039 (forward primer; SEQ ID NO:2):

CAGATCTGATATCGATGGGCCCACTAACTATCTATACTGTAATAATGTTG  
TATAG

TMS2040 (reverse primer; SEQ ID NO:3):

5 CGGACCGGGTGACCAAGCTTAAGCGAACATATGTCCCTC

A 504 bp product comprising the 465 bp SB-GKAF terminator sequence (SEQ ID NO:1) was amplified by PCR using these primers. The product was cloned into pGEMTeasy (Promega) (PHP31801; FIG.1; SEQ ID NO:4) and the sequence was confirmed. The cloned SB-GKAF terminator included 165 bp of the predicted 3' UTR of SB-GKAF along with about 300 bp of downstream sequence. The amplified sequence of SB-GKAF terminator (SEQ ID NO:1) was then cloned into an *Agrobacterium* transformation vector (PHP34074; FIG.2; SEQ ID NO:5), which had the following expression cassettes in divergent orientation:

15 SB-GKAF TERMINATOR: GUSINT: BSV PRO and  
UBI-PRO:UBI INTRON:MOPAT:PINII TERM.

BSV PRO is Banana Streak Virus promoter, which is a strong constitutive promoter. A construct with a potato PINII terminator (Keil et al. (1986) *Nucleic Acids Res.* 14:5641-5650) in place of the SB-GKAF terminator was used as a control (PHP34005; SEQ ID NO:6).

## 20 EXAMPLE 2

### Transient Transformation to Test Efficacy of a SB-GKAF Terminator

The isolated SB-GKAF terminator sequence (SEQ ID NO:1) was tested for its ability to act efficiently as a terminator in a recombinant construct. Its efficacy as a terminator was tested by its ability to stop transcription and by its ability to increase expression of a protein. Since improper termination can lead to improper processing of the 3' end of mRNA, and hence affect RNA stability, terminators have been found to affect protein expression levels. It has been shown that different terminators can cause up to 100-fold variation in the efficiency of transgene expression (Bieri et al, (2002) *Molecular Breeding* 10: 107–117; An et al (1989) *Plant Cell* 1: 115-122; Ingelbrecht et al (1989), *Plant Cell*, 1:671-680; Ali and Taylor (2001) *Plant Mol. Bio.*, 46:251-261). Hence we tested the SB-GKAF sequence (SEQ ID NO:1) for its ability to increase expression of a protein compared to the

well-known PINII terminator. The *Agrobacterium* transformation vectors PHP34074 (SEQ ID NO:5) and PHP34005 (SEQ ID NO:6) described in Example 1 were used for transient transformation of BMS (Black Mexican Sweet) cells. The cells were harvested 5 days after transformation and sent for a quantification of the GUS activity (MUG assay). The SB-GKAF construct (PHP34074; SEQ ID NO:5) had ~35% more expression than that of the PINII construct (PHP34005, SEQ ID NO:6) when the GUS expression was normalized to the MOPAT expression (FIG.3; Table 1). This information was indicative of the ability of the isolated SB-GKAF sequence (SEQ ID NO:1) to act efficiently as a terminator, by allowing protein expression equal to or above that of the PINII terminator.

TABLE 1

Construct	Sequence Tested	Average MUG/PAT*	Standard Deviation
BSV PRO:GUSINT: PINII TERM	PIN II TERM	1.57	0.17
BSV PRO:GUSINT: SB-GKAF TERM	SB-GKAF TERM	2.13	0.41

\* Measured as: nmoles MU/mg total protein/hour/ppm PAT

### EXAMPLE 3

#### Stable Transformation Assays to Test SB-GKAF Terminator Activity

The *Agrobacterium* transformation vectors PHP34074 (SEQ ID NO:5) and PHP34005 (SEQ ID NO:6) described in Example 1, that were used for transient transformation assays as described in Example 2, were also used in Gaspe-Flint derived maize lines for stable transformation to generate transgenic maize plants.

Quantitative Reverse Transcriptase-PCR (qRT-PCR) and GUS assays were done from stably transformed plant tissues to test the ability of isolated SB-GKAF terminator sequence (SEQ ID NO:1) to stop transcription (that is prevent transcription read-through transcription) and to compare GUS expression as compared to that with PINII terminator.

#### GUS Expression Analysis:

The expression of the GUS gene in the transgenic plants was assessed at the protein as well as transcript levels. To assess the expression at the protein level,

MUG assay was performed on seedling leaf material. To assess the expression at the transcript level, qRT-PCR was done using primers shown in Table 2.

**TABLE 2**

<b>Primer/Probe</b>	<b>Type</b>	<b>Sequence</b>	<b>Fluor</b>	<b>qPCR Assay</b>
GUS-1482F	Forward	SEQ ID NO:7	-	Taqman
GUS-1553R	Reverse	SEQ ID NO:8	-	Taqman
GUS-1509P	Probe	SEQ ID NO:9	FAM	Taqman

Plants were grown in the greenhouse and leaves were sampled at the R1 stage of development for expression analysis. Multiple plants were tested for each construct. Each plant was analyzed for expression of the GUS gene. GUS gene with the SB-GKAF terminator had GUS expression in the same range as that of PINII terminator at both the protein (FIG.4A) and transcript (FIG.4B) level.

#### Quantitative Reverse Transcriptase PCR (qRT-PCR) to Determine Read-Through Transcription through the SB-GKAF Terminator:

The qRT-PCR assays were performed with leaf tissue from the stable transformants generated using PHP34074 and PHP34005. Each plant was tested for the presence of read-through transcript that had passed through the PINII terminator and the SB-GKAF terminator (SEQ ID NO:1). To assess presence of products that would indicate that transcription was continuing past the terminator, amplification was targeted downstream of the terminator being tested. Two primer sets were designed downstream of the tested terminators.

- Primer set Term1 ~100 nt from the terminator
- Primer set Term2.1 ~500 nt from the terminator

Multiple plants were tested for each construct. The primers are shown in Table 3.

**TABLE 3**

<b>Primer/Probe</b>	<b>Name</b>	<b>Type</b>	<b>Sequence</b>	<b>Fluor</b>	<b>qPCR Assay</b>
Term2.1 <sup>1</sup>	Term2.1F	fwd	SEQ ID NO:10	-	SYBR
Term2.1 <sup>1</sup>	Term2.1R	rev	SEQ ID NO:11	-	SYBR
Term1 <sup>1</sup>	Term 1F	fwd	SEQ ID NO:12	-	Taqman

Term1 <sup>1</sup>	Term 1R	rev	SEQ ID NO:13	-	Taqman
Term1 <sup>1</sup>	Term_1P	probe	SEQ ID NO:14	FAM	Taqman
Actin <sup>2</sup>	Actin_MGB_F	fwd	SEQ ID NO:15	-	Taqman
Actin <sup>2</sup>	Actin_MGB_R	rev	SEQ ID NO:16	-	Taqman
Actin <sup>2</sup>	Actin_VIC_P	probe	SEQ ID NO:17	VIC	Taqman

<sup>1</sup> Post-Terminator Primer Set

<sup>2</sup> Reference Gene

The test plants were classified into 3 categories depending on the qRT-PCR results:

- 5           1. Plants showing complete termination: where all GUS transcripts are completely terminated before they reached the specific primer set location;
2. Plants showing a high degree of termination: where a large portion of the GUS transcripts are terminated before they reached the specific primer set location, also defined as:

10                   Primer set Term1 -  $\Delta CT > 13$

                  Primer set Term2.1 -  $\Delta CT > 9$ ; and

3. Plants showing poor termination.

As can be see from Fig. 5, the SB-GKAF terminator proved to have fewer “poorly terminating” plants than the PINII terminator (FIG.5). Thus the qRT-PCR score for presence of transcripts that had proceeded through the terminator was lower for the SB-GKAF terminator than that for the PINII terminator.



CLAIMS

We claim:

1. A recombinant construct comprising an isolated polynucleotide sequence comprising:

- 5           (a) a nucleotide sequence comprising the sequence set forth in SEQ ID NO:1 or SEQ ID NO:18;
- (b) a nucleotide sequence comprising a sequence with at least 95% identity to the sequence set forth in SEQ ID NO:1 or SEQ ID NO:18; or
- (c) a nucleotide sequence comprising a functional fragment of either (a)
- 10           or (b);

wherein the isolated polynucleotide sequence functions as a transcriptional terminator in a plant cell.

2. The recombinant construct of Claim 1 wherein the isolated polynucleotide is operably linked to a promoter and a heterologous polynucleotide sequence.
- 15   3. A plant comprising the recombinant construct of Claim 1 or 2.
4. The plant of Claim 3 wherein the plant is a monocot.
5. The plant of Claim 4 wherein the plant is a maize plant.
6. A seed comprising the recombinant construct of Claim 1 or 2.
7. The seed of Claim 6 wherein the seed is from a monocot plant.
- 20   8. The seed of Claim 7 wherein the seed is from a maize plant.
9. A method of expressing a heterologous polynucleotide in a plant, comprising the steps of:
  - (a) introducing into a regenerable plant cell the recombinant construct of Claim 2;
  - 25       (b) regenerating a transgenic plant from the regenerable plant cell of step (a), wherein the transgenic plant comprises the recombinant construct of Claim 2; and
  - (c) obtaining a progeny plant from the transgenic plant of step (b), wherein the progeny plant comprises the recombinant construct of Claim
  - 30       2 and exhibits expression of the heterologous polynucleotide.
10. The method of Claim 9, wherein the plant is a monocot plant.
11. The method of Claim 10, wherein the plant is a maize plant.

FIG. 1

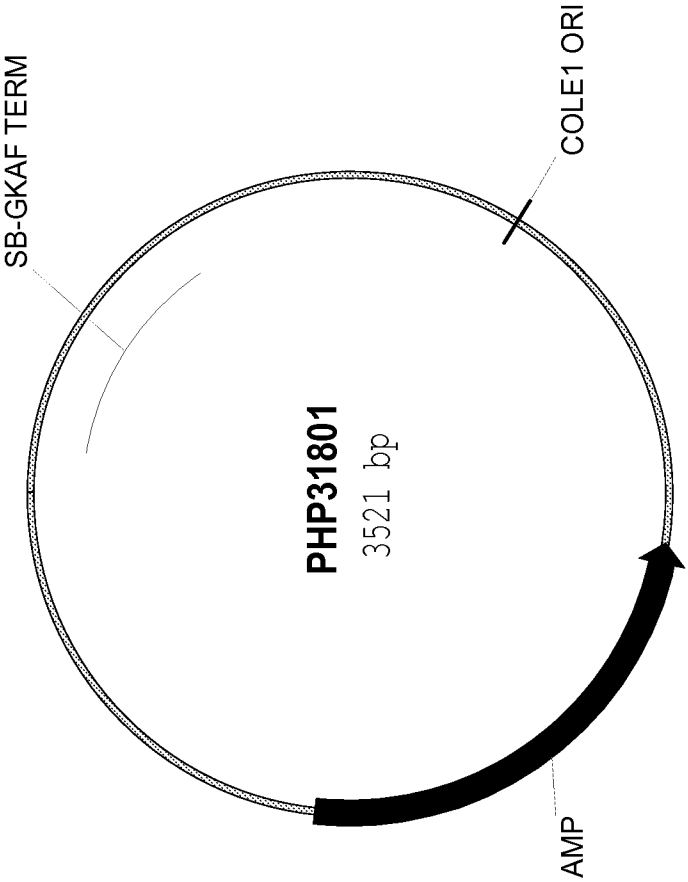


FIG. 2

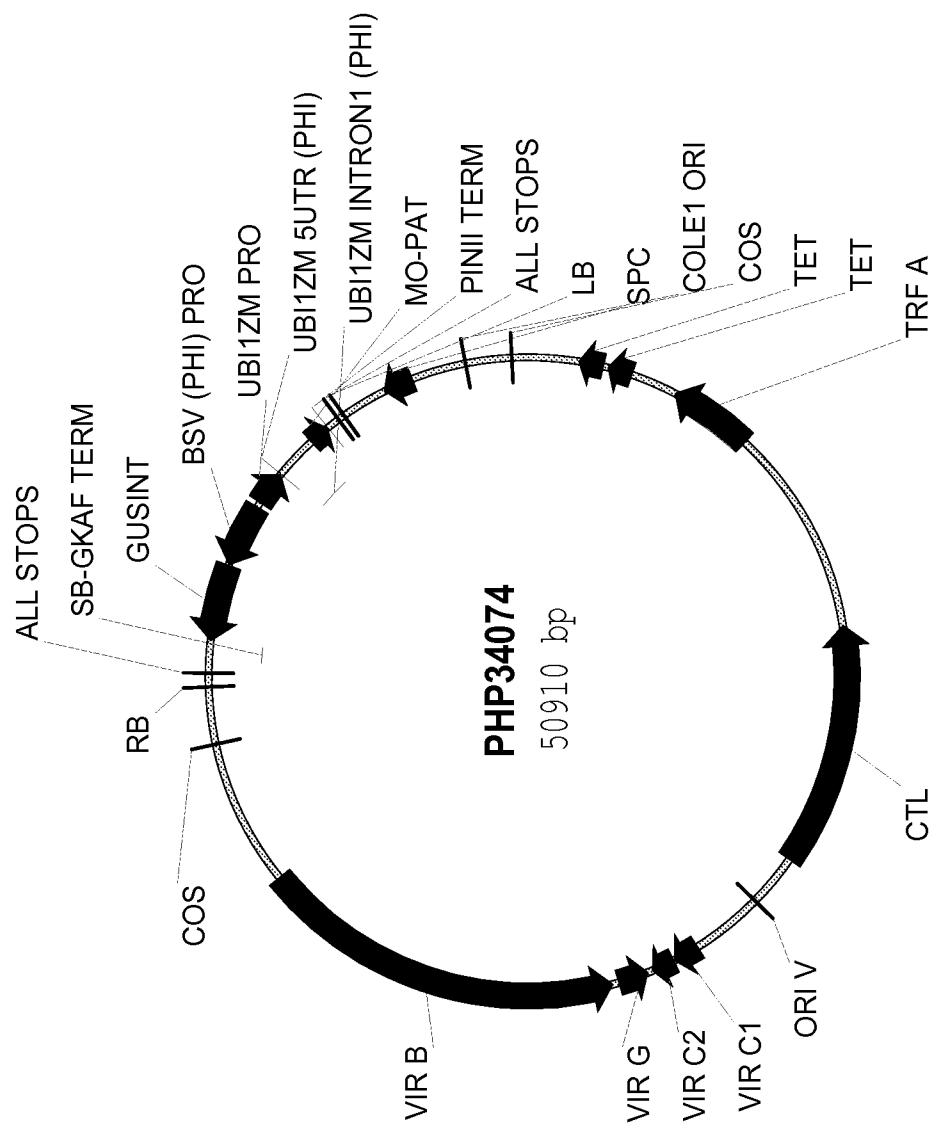


FIG. 3

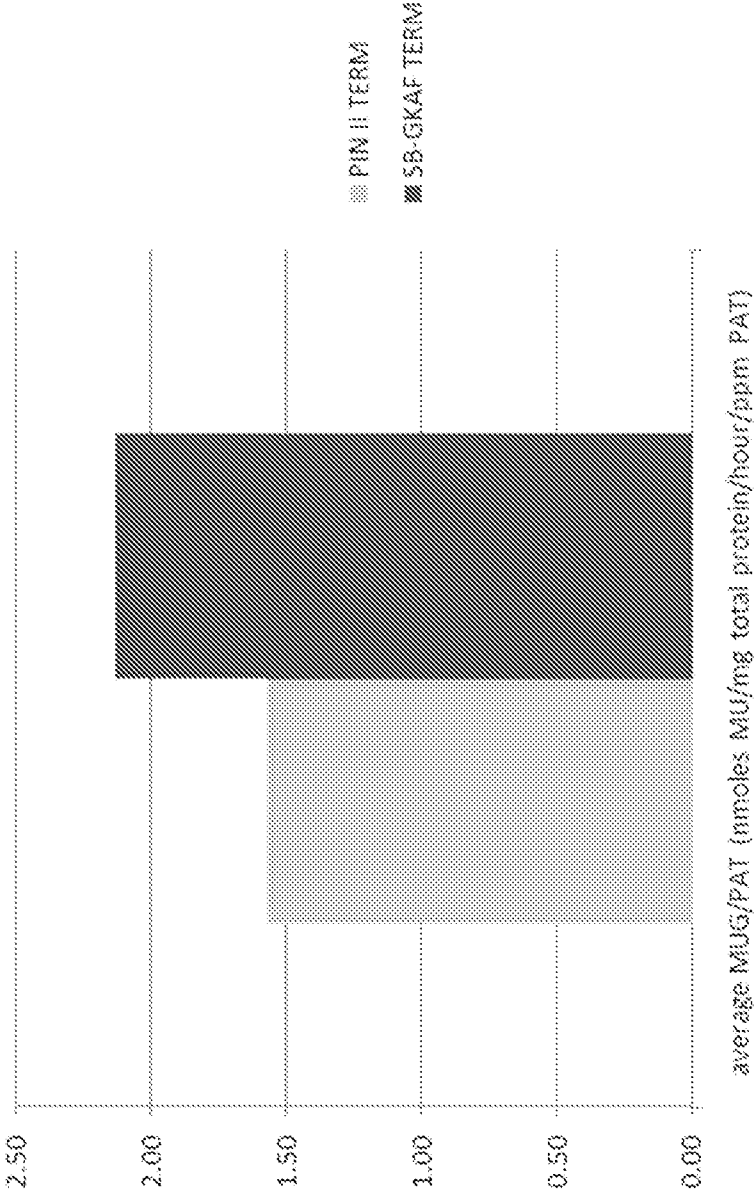


FIG. 4B

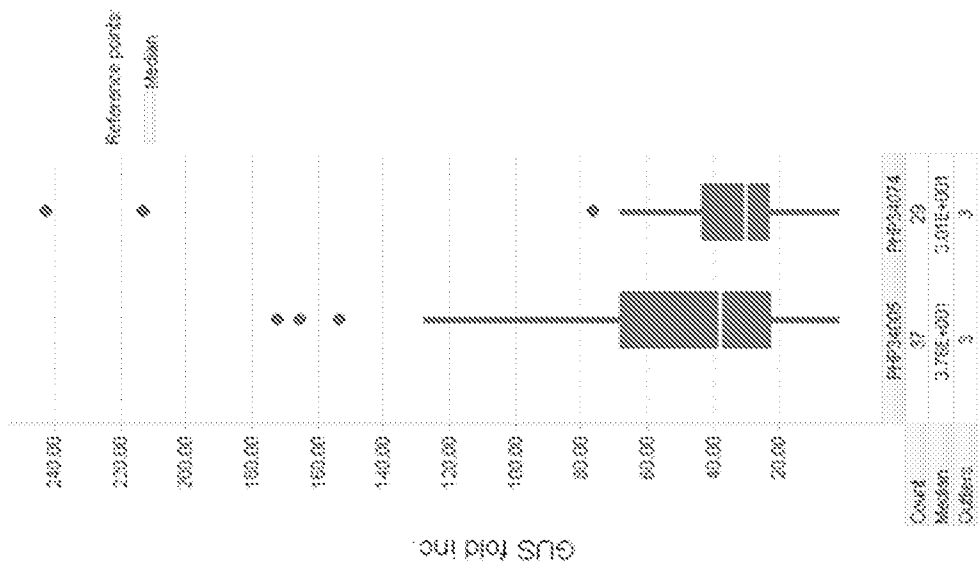


FIG. 4A

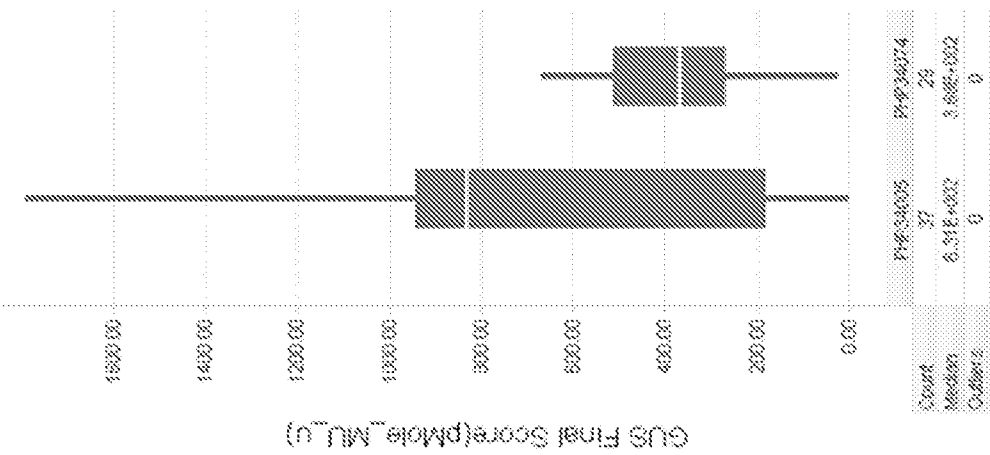


FIG. 5

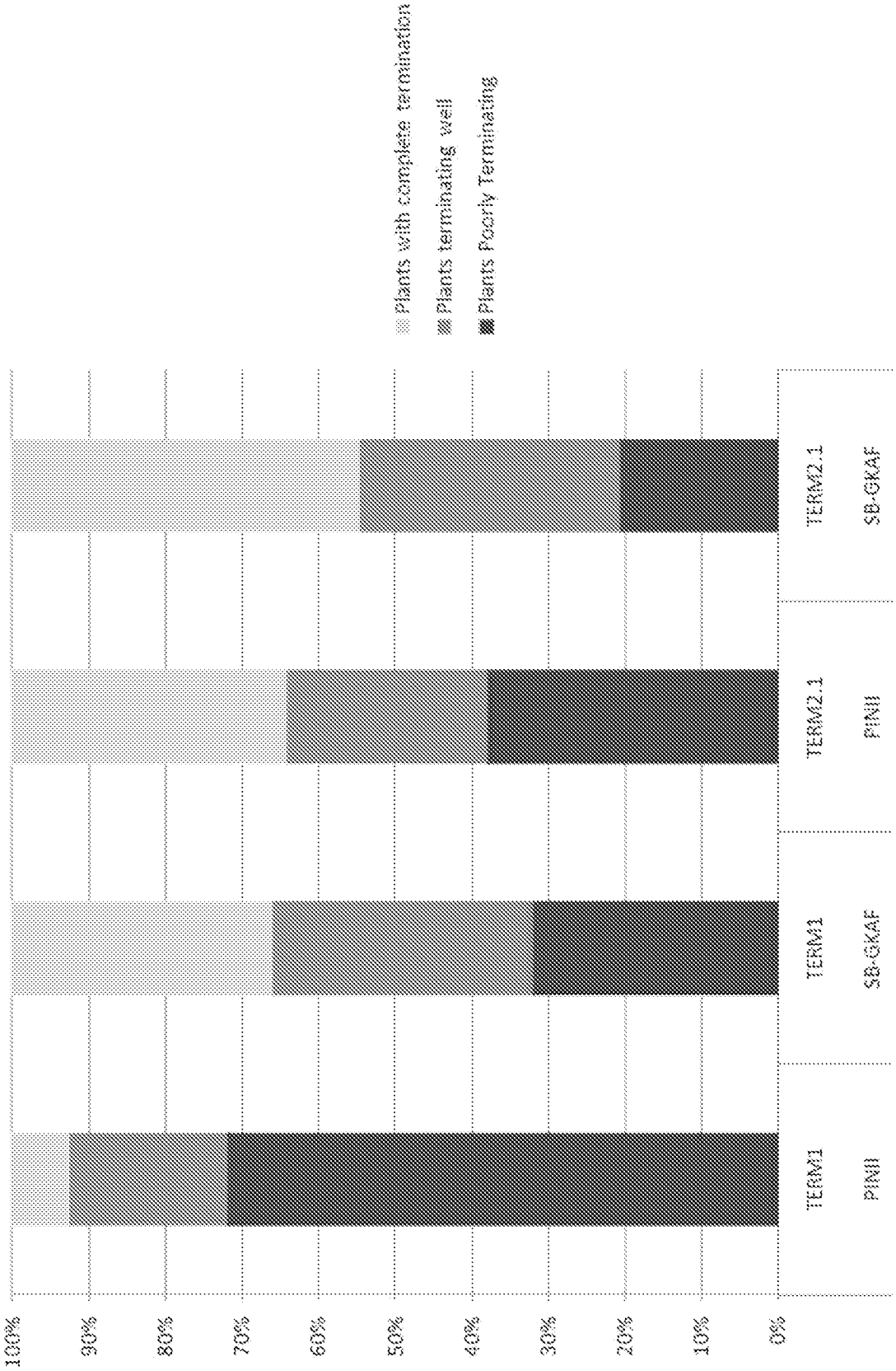


FIG 6A

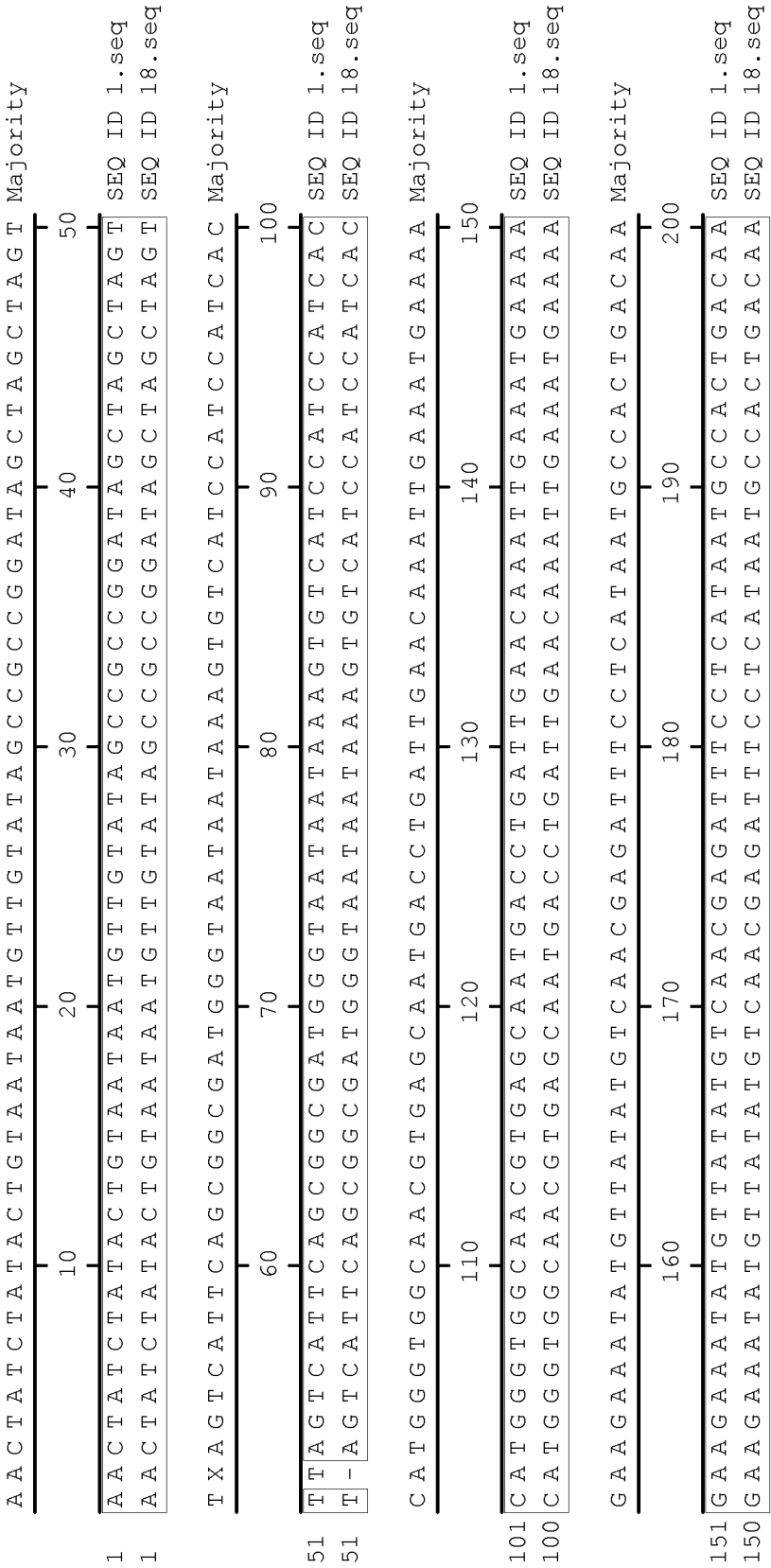
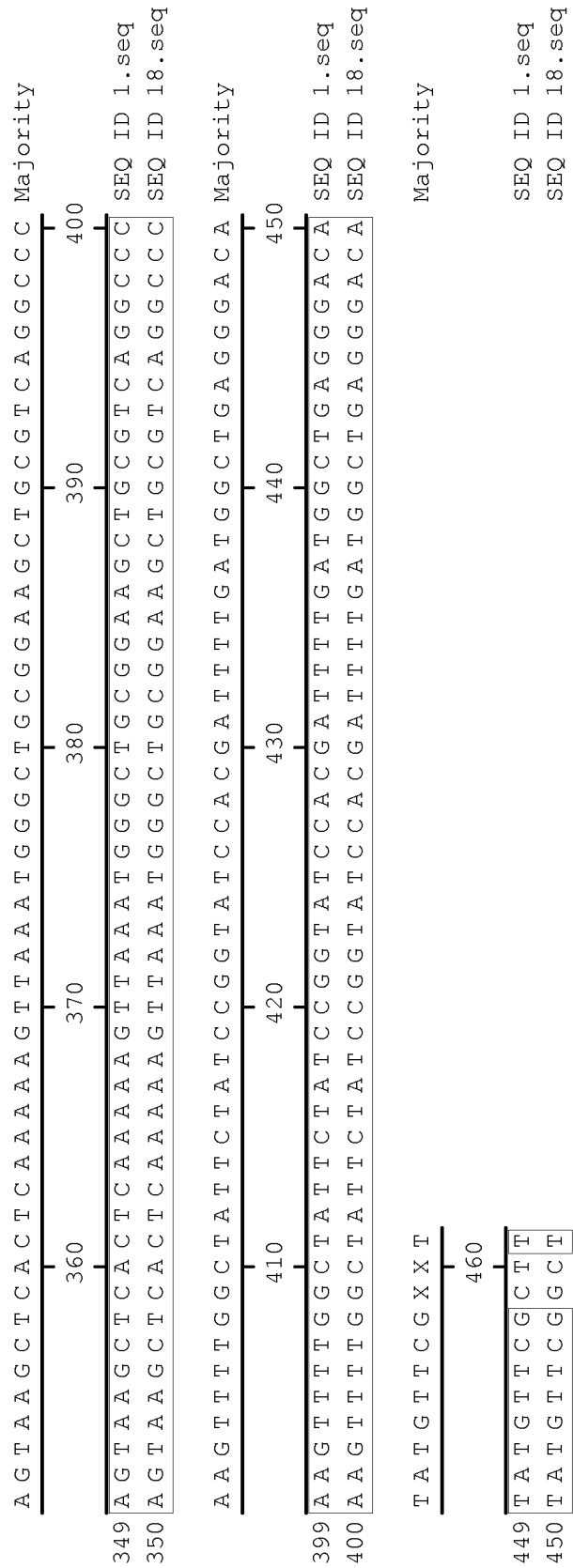


FIG. 6B

	C	G	T	G	T	C	C	A	A	G	A	A	T	G	T	A	T	C	A	G	T	G	A	T	A	T	C	A	C	A	A	T	T	T	T	A	T	Majority									
	210										220										230										240										250						
201	C	G	T	G	T	C	C	A	A	G	A	A	T	G	T	A	T	C	A	G	T	G	A	T	A	T	T	C	A	C	A	A	T	T	T	T	A	T	SEQ ID 1.seq								
200	C	G	T	G	T	C	C	A	A	G	A	A	T	G	T	A	T	C	A	G	T	G	A	T	A	T	T	C	A	C	A	A	T	T	T	T	A	T	SEQ ID 18.seq								
	G	A	C	T	T	A	C	T	C	A	C	A	A	T	T	G	T	T	T	T	T	T	T	T	T	T	A	T	A	C	T	C	X	X	A	C	A	T	T	G	T	Majority					
	260										270										280										290										300						
251	G	A	C	T	T	A	C	T	C	A	C	A	A	T	T	G	T	T	T	T	T	T	T	T	T	T	T	A	C	T	C	-	-	A	C	A	A	T	T	G	T	SEQ ID 1.seq					
250	G	A	C	T	T	A	C	T	C	A	C	A	A	T	T	G	T	T	T	T	T	T	T	T	T	T	T	T	A	C	T	C	G	A	C	A	A	T	T	G	T	SEQ ID 18.seq					
	T	G	T	G	G	T	A	C	C	A	T	A	C	A	A	T	T	C	G	A	T	C	G	A	A	T	A	T	A	T	C	A	G	A	A	A	G	T	T	G	A	C	G	A	Majority		
	310										320										330										340										350						
299	T	G	T	G	G	T	A	C	C	A	T	A	A	C	A	A	T	T	C	G	A	T	C	G	A	A	T	A	T	A	T	C	A	G	A	A	A	G	T	T	G	A	C	G	A	SEQ ID 1.seq	
300	T	G	T	G	G	T	A	C	C	A	T	A	A	C	A	A	T	T	C	G	A	T	C	G	A	T	C	G	A	T	A	T	C	A	G	A	A	A	G	T	T	G	A	C	G	A	SEQ ID 18.seq



FIG. 6C



# INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2012/047901

A. CLASSIFICATION OF SUBJECT MATTER  
INV. C12N15/82 A01H5/00  
ADD.

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

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
C12N A01H

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

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

EP0-Internal, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DE FREITAS F A ET AL: "Structural characterization and promoter activity analysis of the gamma- kafirin gene from sorghum", MOLECULAR AND GENERAL GENETICS, SPRINGER VERLAG, BERLIN, DE, vol. 245, no. 2, 28 October 1994 (1994-10-28), pages 177-186, XP002113596, ISSN: 0026-8925	1,3,4,6, 7
Y	the whole document -----	2,5,8-11
Y	US 2009/089897 A1 (ABBITT SHANE E [US] ET AL) 2 April 2009 (2009-04-02) paragraph [0025] -----	2,5,8-11



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

5 October 2012

Date of mailing of the international search report

12/10/2012

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Maddox, Andrew

## INTERNATIONAL SEARCH REPORT

### Information on patent family members

International application No

PCT/US2012/047901

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009089897	A1	02-04-2009	NONE
-----			