Abstract:
The present invention relates to a Brassica sequence comprising seed-preferential promoter activity. Provided are recombinant genes comprising the seed-preferential promoter operably linked to a heterologous nucleic acid sequence, and cells, plants and seeds comprising the recombinant gene. The promoter can be used to alter gene expression specifically in the seeds and at a specific seed developmental stage and to alter biotic or abiotic stress tolerance, yield or seed quality.
SEED-PREFERENTIAL PROMOTERS AND USES THEREOF

FIELD OF THE INVENTION

[1] The present invention relates to materials and methods for the expression of a gene of interest specifically in seeds of plants. In particular, the invention provides an expression cassette for regulating seed-preferential expression in plants and more specifically expression cassettes for regulating specific seed developmental stage expression.

BACKGROUND OF THE INVENTION

[2] Modification of plants to alter and/or improve phenotypic characteristics (such as productivity or quality) requires the overexpression or down-regulation of endogenous genes or the expression of heterologous genes in plant tissues. Such genetic modification relies on the availability of a means to drive and to control gene expression as required. Indeed, genetic modification relies on the availability and use of suitable promoters which are effective in plants and which regulate gene expression so as to give the desired effect(s) in the transgenic plant.

[3] For numerous applications in plant biotechnology a tissue-specific or a tissue-preferential expression profile is advantageous, since beneficial effects of expression in one tissue may have disadvantages in others.

[4] Seed-preferential promoters are useful for expressing or down-regulating genes preferentially in the seeds to get the desired function or effect, such as improving disease resistance, herbicide resistance, modifying seed or grain composition or quality, such as modifying starch quality or quantity, modifying oil quality or quantity, modifying amino-acid or protein composition, improving tolerance to biotic or abiotic stress, increasing yield, or altering metabolic pathways in the seeds.


[6] There remains an interest to the isolation of novel seed-preferential promoters. For example, in gene stacking approaches, when more than one gene is to be expressed seed-preferentially, the use of...
more than one copy of the same promoter increases the risk on unwanted gene silencing. In addition, it may be desirable to have the availability of different seed-preferential promoters which direct expression at different levels for expression or down-regulation of certain genes of interest. Furthermore, there is a need for promoters which can direct gene expression at specific stages of seed development.

It is thus an objective of the present invention to provide a *Brassica* promoter for seed-preferential and 3 to 5 week-old seed developmental stage-specific expression of genes of interest in plants. This objective is solved by the present invention as herein further explained.

**SUMMARY OF THE INVENTION**

In one aspect, the invention provides an isolated nucleic acid comprising seed-preferential promoter activity selected from the group consisting of (a) a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 1 or a functional fragment thereof; (b) a nucleic acid comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO: 1 or a functional fragment thereof; and (c) a nucleic acid hybridizing under stringent conditions to the nucleotide sequence of SEQ ID NO:1, or a functional fragment thereof.

A further embodiment provides a recombinant gene comprising the nucleic acid according to the invention operably linked to a heterologous nucleic acid sequence encoding an expression product of interest, and optionally a transcription termination and polyadenylation sequence, preferably a transcription termination and polyadenylation region functional in plant cells. In a further embodiment, said expression product of interest is an RNA capable of modulating the expression of a gene or is a protein.

Yet another embodiment provides a host cell, such as an *E. coli* cell, an *Agrobacterium* cell, a yeast cell, or a plant cell, comprising the isolated nucleic acid according to the invention, or the recombinant gene according to the invention.

In a further embodiment, a plant is provided comprising the recombinant gene according to the invention. Yet a further embodiment provides seeds obtainable from the plant according to the invention. In another embodiment, the plants or plant parts according to the invention are seed crop plants or seeds.

Yet another embodiment provides a method of producing a transgenic plant comprising the steps of (a) introducing or providing the recombinant gene according to the invention to a plant cell to create transgenic cells; and (b) regenerating transgenic plants from said transgenic cell.

Further provided is a method of effecting seed-preferential expression of a nucleic acid comprising introducing the recombinant gene according to the invention into the genome of a plant, or
providing the plant according to the invention. Also provided is a method for altering seed properties of
a plant or to produce a commercially relevant product in a plant, said method comprising introducing the
recombinant gene according to the invention into the genome of a plant, or providing the plant according
to the invention. In another embodiment, said plant is a seed crop plant.

[14] Also provided is the use of the isolated nucleic acid according to the invention to regulate
expression of an operably linked nucleic acid in a plant, and the use of the isolated nucleic acid
according to the invention, or the recombinant gene according to the invention to alter seed properties of
a plant or to produce a commercially relevant product in a plant. In a further embodiment, said plant is a
seed crop plant.

[15] Yet another embodiment provides a method of producing food, feed, or an industrial product
comprising (a) obtaining the plant or a part thereof, according to the invention; and (b) preparing the
food, feed or industrial product from the plant or part thereof. In another embodiment, said food or feed
is oil, meal, grain, starch, flour or protein, or said industrial product is biofuel, fiber, industrial
chemicals, a pharmaceutical or a nutraceutical.

BRIEF DESCRIPTION OF THE DRAWINGS

[16] Figure 1: Relative expression levels (tpm; transcript per million) of the Seed_345 transcript in
different tissues. Different tissues: AM33: Apical meristem 33 days after sowing (DAS); BFB42: Big
flower buds 42 DAS; CTYL10: Cotyledons 10 DAS; OF52: Open flowers 52 DAS; Pod2: Pods 14-20
DAS; Pod3: Pods 21-25 DAS; Ro2w: Roots 14 DAS; Seed2: Seeds 14-20 days after flowering (DAF);
Seed3: Seeds 21-25 DAF; Seed4: Seeds 26-30 DAF; Seed5: Seeds 31-35 DAF; Seed6: Seeds 42 DAF;
Seed7: Seeds 49 DAF; SFB42: Small flower buds 42 DAS; St2w: Stem 14 DAS; St5w: Stem 33 DAS;
YL33: Young leaf 33 DAS.

DETAILED DESCRIPTION

[17] The present invention is based on the observation that SEQ ID NO: 1 has 3 to 5 week-old seed-
specific promoter activity.

[18] In one aspect, the invention provides an isolated nucleic acid comprising seed-preferential
promoter activity selected from the group consisting of (a) a nucleic acid comprising a nucleotide
sequence of SEQ ID NO: 1 or a functional fragment thereof; (b) a nucleic acid comprising a nucleotide
sequence having at least 80% sequence identity to SEQ ID NO: 1, or a functional fragment thereof; and
(c) a nucleic acid capable of hybridizing under stringent conditions to the nucleotide sequence of SEQ
ID NO: 1, or a functional fragment thereof.
SEQ ID NO: 1 depicts the region upstream (i.e. located 5' upstream of) from the first ATG start codon of the Seed_345 transcript. Such a promoter region may be at least about 300 bp, at least about 500 bp, at least about 800 bp, at least about 1000 bp, at least about 1500 bp, at least about 2000 bp, at least about 2500 bp, or at least about 3000 bp upstream of the first ATG start codon of the Seed_345 transcript.

The nucleic acid comprising the seed-preferential promoter activity according to the invention may also be comprised in a larger DNA molecule.

"Seed-preferential promoter activity" in the context of this invention means the promoter activity is at least 2 times, or at least 5 times, or at least 10 times, or at least 20 times or even at least 100 times higher in seeds than in other tissues. In other words, in seed-preferential promoter activity, transcription of the nucleic acid operably linked to the promoter of the invention in the seeds is at least 2 times, or at least 5 times, or at least 10 times, or at least 20 times or even at least 100 times higher than in other tissues. In other words, the seed-preferential promoter drives seed-preferential expression of the nucleic acid operably linked to the seed-preferential promoter.

"Seed-preferential promoter activity" encompasses "3 to 5 week-old seed-specific promoter activity".

"3 to 5 week-old seed-specific promoter activity" in the context of this invention means the promoter activity is at least 2 times, or at least 5 times, or at least 10 times, or at least 20 times or even at least 100 times higher in 3 week-old seeds, in 4 week-old seeds and in 5 week-old seeds than in other tissues. In other words, in 3 to 5 week-old seed-specific promoter activity, transcription of the nucleic acid operably linked to the promoter of the invention in the 3 week-old seeds, 4 week-old seeds and 5 week-old seeds is at least 2 times, or at least 5 times, or at least 10 times, or at least 20 times or even at least 100 times higher than in other tissues. In other words, the 3 to 5 week-old seed-specific promoter drives 3 to 5 week-old seed-specific expression of the nucleic acid operably linked to the 3 to 5 week-old seed-specific promoter.

"3 week-old seeds" refers to seeds at the developmental stage reached between 21 and 25 days after flowering.

The phrase "operably linked" refers to the functional spatial arrangement of two or more nucleic acid regions or nucleic acid sequences. For example, a promoter region may be positioned relative to a nucleic acid sequence such that transcription of a nucleic acid sequence is directed by the promoter region. Thus, a promoter region is "operably linked" to the nucleic acid sequence. "Functionally linked" is an equivalent term.
The phrases "DNA", "DNA sequence," "nucleic acid sequence," "nucleic acid molecule" "nucleotide sequence" and "nucleic acid" refer to a physical structure comprising an orderly arrangement of nucleotides. The DNA sequence or nucleotide sequence may be contained within a larger nucleotide molecule, vector, or the like. In addition, the orderly arrangement of nucleic acids in these sequences may be depicted in the form of a sequence listing, figure, table, electronic medium, or the like.

As used herein, "promoter" means a region of DNA sequence that is essential for the initiation of transcription of DNA, resulting in the generation of an RNA molecule that is complementary to the transcribed DNA; this region may also be referred to as a "5' regulatory region." Promoters are usually located upstream of the coding sequence to be transcribed and have regions that act as binding sites for RNA polymerase II and other proteins such as transcription factors (trans-acting protein factors that regulate transcription) to initiate transcription of an operably linked gene. Promoters may themselves contain sub-elements (i.e. promoter motifs) such as cis-elements or enhancer domains that regulate the transcription of operably linked genes. The promoters of this invention may be altered to contain "enhancer DNA" to assist in elevating gene expression. As is known in the art, certain DNA elements can be used to enhance the transcription of DNA. These enhancers often are found 5' to the start of transcription in a promoter that functions in eukaryotic cells, but can often be inserted upstream (5') or downstream (3') to the coding sequence. In some instances, these 5' enhancer DNA elements are introns. Among the introns that are useful as enhancer DNA are the 5' introns from the rice actin 1 gene (see US5641876), the rice actin 2 gene, the maize alcohol dehydrogenase gene, the maize heat shock protein 70 gene (see US5593874), the maize shrunken 1 gene, the light sensitive 1 gene of Solarium tuberosum, the Arabidopsis histon 4 intron and the heat shock protein 70 gene of Petunia hybrida (see US5659122). Thus, as contemplated herein, a promoter or promoter region includes variations of promoters derived by inserting or deleting regulatory regions, subjecting the promoter to random or site-directed mutagenesis, etc. The activity or strength of a promoter may be measured in terms of the amounts of RNA it produces, or the amount of protein accumulation in a cell or tissue, relative to a promoter whose transcriptional activity has been previously assessed.

A promoter as used herein may thus include sequences downstream of the transcription start, such as sequences coding the 5' untranslated region (5' UTR) of the RNA, introns located downstream of the transcription start, or even sequences encoding the protein.

Promoter activity for a functional promoter fragment in seeds may be determined by those skilled in the art, for example using analysis of RNA accumulation produced from the nucleic acid which is operably linked to the promoter as described herein, whereby the nucleic acid which is operably linked to the promoter can be the nucleic acid which is naturally linked to the promoter, i.e. the endogenous gene of which expression is driven by the promoter.
The RNA accumulation, or levels of RNA, such as mRNA, can be measured either at a single time point or at multiple time points and as such the fold increase can be average fold increase or an extrapolated value derived from experimentally measured values. As it is a comparison of levels, any method that measures mRNA levels can be used. In a preferred aspect, the tissue or organs compared are seeds tissues at a specific developmental stage (3 to 5 week-old, 4 week-old, 5 week-old stages) with other tissues of the organism. In another preferred aspect, multiple tissues or organs are compared. A preferred multiple comparison is seed tissue compared with 2, 3, 4, or more tissues or organs selected from the group consisting of apical meristem, flower buds, cotyledons, flowers, pods, roots, seeds at different stages, leaves and stems. As used herein, examples of plant organs are fiber, leaf, root, etc. and example of tissues are leaf primordia, shoot apex, vascular tissue, etc. A multiple comparison may comprise a comparison of 3 week-old seed, 4 week-old seed and 5 week-old seed tissues with apical meristem, flower bud, cotyledon, flower, root, stem, 2 week-old seed, 6 week-old seed and 7 week-old seed tissues.

The seed-preferential expression capacity of the identified or generated fragments of the promoters of the invention can be conveniently tested by determining levels of the transcript of which expression is naturally driven by the promoter of the invention, i.e. endogenous transcript levels, such as, for example, using the methods as described herein in the Examples. Further, the seed-preferential expression capacity of the identified or generated fragments of the promoters of the invention can be conveniently tested by operably linking such DNA molecules to a nucleotide sequence encoding an easy scorable marker, e.g. a beta-glucuronidase gene, introducing such a chimeric gene into a plant and analyzing the expression pattern of the marker in seeds at different developmental stages as compared with the expression pattern of the marker in other parts of the plant or in seeds at other developmental stages. Other candidates for a marker (or a reporter gene) are chloramphenicol acetyl transferase (CAT) and proteins with fluorescent properties, such as green fluorescent protein (GFP) from Aequora victoria.

To define a minimal promoter region, a DNA segment representing the promoter region is removed from the 5' region of the gene of interest and operably linked to the coding sequence of a marker (reporter) gene by recombinant DNA techniques well known to the art. The reporter gene is operably linked downstream of the promoter, so that transcripts initiating at the promoter proceed through the reporter gene. Reporter genes generally encode proteins, which are easily measured, including, but not limited to, chloramphenicol acetyl transferase (CAT), beta-glucuronidase (GUS), green fluorescent protein (GFP), beta-galactosidase (beta-GAL), and luciferase. The expression cassette containing the reporter gene under the control of the promoter can be introduced into an appropriate cell type by transfection techniques well known to the art. To assay for the reporter protein, cell lysates are prepared and appropriate assays, which are well known in the art, for the reporter protein are performed. For example, if CAT were the reporter gene of choice, the lysates from cells transfected with constructs containing CAT under the control of a promoter under study are mixed with isotopically labeled chloramphenicol and acetyl-coenzyme A (acetyl-CoA). The CAT enzyme transfers the acetyl group
from acetyl-CoA to the 2- or 3-position of chloramphenicol. The reaction is monitored by thin-layer chromatography, which separates acetylated chloramphenicol from unreacted material. The reaction products are then visualized by autoradiography. The level of enzyme activity corresponds to the amount of enzyme that was made, which in turn reveals the level of expression and the seed-preferential functionality from the promoter or promoter fragment of interest. This level of expression can also be compared to other promoters to determine the relative strength of the promoter under study. Once activity and functionality is confirmed, additional mutational and/or deletion analyses may be employed to determine the minimal region and/or sequences required to initiate transcription. Thus, sequences can be deleted at the 5' end of the promoter region and/or at the 3' end of the promoter region, and nucleotide substitutions introduced. These constructs are then again introduced in cells and their activity and/or functionality determined.

[32] The activity or strength of a promoter may be measured in terms of the amount of mRNA or protein accumulation it specifically produces, relative to the total amount of mRNA or protein. The promoter preferably expresses an operably linked nucleic acid sequence at a level greater than about 1%, about 2%, more preferably greater than about 5% of the total mRNA. Alternatively, the activity or strength of a promoter may be expressed relative to a well-characterized promoter (for which transcriptional activity was previously assessed).

[33] It will herein further be clear that equivalent seed-preferential promoters can be isolated from other plants. To this end, orthologous promoter fragments may be isolated from other plants using SEQ ID NO: 1 or a functional fragment having at least 300 consecutive nucleotides thereof as a probe and identifying nucleotide sequences from these other plants which hybridize under the herein described hybridization conditions. By way of example, a promoter of the invention may be used to screen a genomic library of a crop or plant of interest to isolate corresponding promoter sequences according to techniques well known in the art. Thus, a promoter sequence of the invention may be used as a probe for hybridization with a genomic library under medium to high stringency conditions. As an alternative equivalent promoters can be isolated using the coding sequences of the genes driven by the promoter of SEQ ID NO: 1 to screen a genomic library (e.g. by hybridization or in silico) of a crop of interest. When sufficient identity between the coding sequences is obtained (for example, higher than 85% identity) then promoter regions can be isolated upstream of the orthologous genes.

[34] Hybridization occurs when the two nucleic acid molecules anneal to one another under appropriate conditions. Nucleic acid hybridization is a technique well known to those of skill in the art of DNA manipulation. The hybridization property of a given pair of nucleic acids is an indication of their similarity or identity. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular
nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA. "Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target nucleic acid sequence. "Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern hybridization are sequence dependent, and are different under different environmental parameters. An example of highly stringent wash conditions is 0.15 M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2 X SSC wash at 65°C for 15 minutes. Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1 X SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4 to 6 X SSC at 40°C for 15 minutes. For short probes (e.g., about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.5 M, more preferably about 0.01 to 1.0 M, Na+ ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C and at least about 60°C for long probes (e.g., >50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2 X (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the proteins that they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. Very stringent conditions are selected to be equal to the T_m for a particular probe. An example of stringent conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or Northern blot is 50% formamide, e.g., hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1 X SSC at 60 to 65°C. 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 1 X to 2 X SSC (20 X 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 0.5 X to 1 X SSC at 55 to 60°C.

The following are examples of sets of hybridization/wash conditions that may be used to clone orthologous nucleotide sequences that are substantially identical to reference nucleotide sequences of the present invention: a reference nucleotide sequence preferably hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M NaP0_4, 1 mM EDTA at 50°C with washing in 2 X SSC, 0.1% SDS at 50°C, more desirably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaP0_4, 1 mM EDTA at 50°C with washing in 1 X SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M NaP0_4, 1 mM EDTA at 50°C with washing in 0.5 X SSC, 0.1% SDS at 50°C.
Suitable to the invention are nucleic acids comprising seed-preferential promoter activity which comprise a nucleotide sequence having at least 40%, at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or at least 98% sequence identity to the herein described promoters and promoter regions or functional fragments thereof and are also referred to as variants. The term "variant" with respect to the transcription regulating nucleotide sequence SEQ ID NO: 1 of the invention is intended to mean substantially similar sequences. 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 herein outlined before. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis of SEQ ID NO: 1. Generally, nucleotide sequence variants of the invention will have at least 40%, 50%, 60%, to 70%, e.g., preferably 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, to 79%, generally at least 80%, e.g., 81% to 84%, at least 85%, e.g., 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, to 98% and 99% nucleotide sequence identity to the native (wild type or endogenous) nucleotide sequence or a functional fragment thereof. Derivatives of the DNA molecules disclosed herein may include, but are not limited to, deletions of sequence, single or multiple point mutations, alterations at a particular restriction enzyme site, addition of functional elements, or other means of molecular modification which may enhance, or otherwise alter promoter expression. Techniques for obtaining such derivatives are well-known in the art (see, for example, J. F. Sambrook, D. W. Russell, and N. Irwin (2000) Molecular Cloning: A Laboratory Manual, 3rd edition Volumes 1, 2, and 3. Cold Spring Harbor Laboratory Press). For example, one of ordinary skill in the art may delimit the functional elements within the promoters disclosed herein and delete any non-essential elements. Functional elements may be modified or combined to increase the utility or expression of the sequences of the invention for any particular application. Those of skill in the art are familiar with the standard resource materials that describe specific conditions and procedures for the construction, manipulation, and isolation of macromolecules (e.g., DNA molecules, plasmids, etc.), as well as the generation of recombinant organisms and the screening and isolation of DNA molecules. As used herein, the term "percent sequence identity" refers to the percentage of identical nucleotides between two segments of a window of optimally aligned DNA. Optimal alignment of sequences for aligning a comparison window are well-known to those skilled in the art and may be conducted by tools such as the local homology algorithm of Smith and Waterman (Waterman, M. S. Introduction to Computational Biology: Maps, sequences and genomes. Chapman & Hall. London (1995), the homology alignment algorithm of Needleman and Wunsch (J. Mol. Biol., 48:443-453 (1970), the search for similarity method of Pearson and Lipman (Proc. Natl. Acad. Sci., 85:2444 (1988), and preferably by computerized implementations of these algorithms such as GAP, BESTFIT, FASTA, and TFASTA available as part of the GCG (Registered Trade Mark), Wisconsin Package (Registered Trade Mark from Accelrys Inc., San Diego, Calif). An "identity fraction" for aligned segments of a test sequence and a reference sequence is the number of identical components that are shared by the two aligned sequences.
divided by the total number of components in the reference sequence segment, i.e., the entire reference sequence or a smaller defined part of the reference sequence. Percent sequence identity is represented as the identity fraction times 100. The comparison of one or more DNA sequences may be to a full-length DNA sequence or a portion thereof, or to a longer DNA sequence.

[36] A nucleic acid comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO: 1 can thus be a nucleic acid comprising a nucleotide sequence having at least 80%, or at least 85%, or at least 90%, or at least 95%, or at least 98%, or 100% sequence identity to SEQ ID NO: 1.

[37] A "functional fragment" of a nucleic acid comprising seed-preferential promoter denotes a nucleic acid comprising a stretch of the nucleic acid sequence of SEQ ID NO: 1 which still exerts the desired function, i.e. which has seed-preferential promoter activity. Assays for determining seed-preferential promoter activity are provided herein. Preferably, the functional fragment of the seed-preferential promoter contains the conserved promoter motifs, such as, for example, conserved promoter motifs as described in DoOP (doop.abc.hu, databases of Orthologous Promoters, Barta E. et al (2005) Nucleic Acids Research Vol. 33, D86-D90). A functional fragment may be a fragment of at least about 300 bp, at least about 500 bp, at least about 800 bp, at least about 1000 bp, at least about 1500 bp, at least about 2000 bp, at least about 2500 bp, or at least about 3000 bp.

[38] "Isolated nucleic acid", used interchangeably with "isolated DNA" as used herein refers to a nucleic acid not occurring in its natural genomic context, irrespective of its length and sequence. Isolated DNA can, for example, refer to DNA which is physically separated from the genomic context, such as a fragment of genomic DNA. Isolated DNA can also be an artificially produced DNA, such as a chemically synthesized DNA, or such as DNA produced via amplification reactions, such as polymerase chain reaction (PCR) well-known in the art. Isolated DNA can further refer to DNA present in a context of DNA in which it does not occur naturally. For example, isolated DNA can refer to a piece of DNA present in a plasmid. Further, the isolated DNA can refer to a piece of DNA present in another chromosomal context than the context in which it occurs naturally, such as for example at another position in the genome than the natural position, in the genome of another species than the species in which it occurs naturally, or in an artificial chromosome.

[39] A further embodiment provides a recombinant gene comprising the nucleic acid according to the invention operably linked to a heterologous nucleic acid sequence encoding an expression product of interest, and optionally a transcription termination and polyadenylation sequence, preferably a transcription termination and polyadenylation region functional in plant cells. In a further embodiment, said expression product of interest an RNA capable of modulating the expression of a gene or is a protein.
The term "expression product" refers to a product of transcription. Said expression product can be the transcribed RNA. It is understood that the RNA which is produced is a biologically active RNA. Said expression product can also be a peptide, a polypeptide, or a protein, when said biologically active RNA is an mRNA and said protein is produced by translation of said mRNA.

Alternatively, the heterologous nucleic acid, operably linked to the promoters of the invention, may also code for an RNA capable of modulating the expression of a gene. Said RNA capable of modulating the expression of a gene can be an RNA which reduces expression of a gene. Said RNA can reduce the expression of a gene for example through the mechanism of RNA-mediated gene silencing.

Said RNA capable of modulating the expression of a gene can be a silencing RNA down-regulating expression of a target gene. As used herein, "silencing RNA" or "silencing RNA molecule" refers to any RNA molecule, which upon introduction into a plant cell, reduces the expression of a target gene. Such silencing RNA may e.g. be so-called "antisense RNA", whereby the RNA molecule comprises a sequence of at least 20 consecutive nucleotides having 95% sequence identity to the complement of the sequence of the target nucleic acid, preferably the coding sequence of the target gene.

However, antisense RNA may also be directed to regulatory sequences of target genes, including the promoter sequences and transcription termination and polyadenylation signals. Silencing RNA further includes so-called "sense RNA" whereby the RNA molecule comprises a sequence of at least 20 consecutive nucleotides having 95% sequence identity to the sequence of the target nucleic acid. Other silencing RNA may be "unpolyadenylated RNA" comprising at least 20 consecutive nucleotides having 95% sequence identity to the complement of the sequence of the target nucleic acid, such as described in WO99/53050 or US6423885 (both documents herein incorporated by reference). Yet another type of silencing RNA is an RNA molecule as described in WO03/076619 (herein incorporated by reference) comprising at least 20 consecutive nucleotides having 95% sequence identity to the sequence of the target nucleic acid or the complement thereof, and further comprising a largely-double stranded region as described in WO03/076619 (including largely double stranded regions comprising a nuclear localization signal from a viroid of the Potato spindle tuber viroid-type or comprising CUG trinucleotide repeats). Silencing RNA may also be double stranded RNA comprising a sense and antisense strand as herein defined, wherein the sense and antisense strand are capable of base-pairing with each other to form a double stranded RNA region (preferably the said at least 20 consecutive nucleotides of the sense and antisense RNA are complementary to each other). The sense and antisense region may also be present within one RNA molecule such that a hairpin RNA (hpRNA) can be formed when the sense and antisense region form a double stranded RNA region. hpRNA is well-known within the art (see e.g. WO99/53050, herein incorporated by reference). The hpRNA may be classified as long hpRNA, having long, sense and antisense regions which can be largely complementary, but need not be entirely complementary (typically larger than about 200 bp, ranging between 200-1000 bp). hpRNA can also be rather small ranging in size from about 30 to about 42 bp, but not much longer than 94 bp (see
WO04/073390, herein incorporated by reference). Silencing RNA may also be artificial micro-RNA molecules as described e.g. in WO2005/052170, WO2005/047505 or US 2005/0144667, or ta-siRNAs as described in WO2006/074400 (all documents incorporated herein by reference). Said RNA capable of modulating the expression of a gene can also be an RNA ribozyme.

[43] Said RNA capable of modulating the expression of a gene can modulate, preferably down-regulate, the expression of other genes (i.e. target genes) comprised within the seeds or even of genes present within a pathogen or pest that feeds upon the seeds of the transgenic plant such as a virus, fungus, insect, bacteria.

[44] The nucleic acid sequence heterologous to the promoters according to the invention may generally be any nucleic acid sequence effecting increased, altered (e.g. in a different organ) or reduced level of transcription of a gene for which such expression modulation is desired. The nucleic acid sequence can for example encode a protein of interest. Exemplary genes for which an increased or reduced level of transcription may be desired in the seeds are e.g. nucleic acids that can provide an agriculturally or industrially important feature in seeds. Suitable heterologous nucleic acid sequences of interest include nucleic acids modulating expression of genes conferring resistance to diseases, stress tolerance genes, genes involved in at different stages of fatty acid biosynthesis, in acyl editing, in storage compound storage or breakdown, genes encoding epoxidas, hydroxylases, cytochrome P450 mono-oxygenases, desaturases, tocopherol biosynthetic enzymes, carotenoid biosynthesis enzymes, amino acid biosynthetic enzymes, steroid pathway enzymes, starch branching enzymes, genes encoding proteins involved in starch synthesis, glycolysis, carbon metabolism, oxidative pentose phosphate cycle, protein synthesis, organelle organization and biogenesis, DNA metabolism, DNA replication, cell cycle, cell organization and biogenesis, chromosome organization and biogenesis, microtubule-based processes, microtubule-based movement, cytoskeleton-dependent intracellular transport, cytoskeleton organization and biogenesis, chromatin assembly or disassembly, DNA-dependent DNA replication, chromosome organization and biogenesis, DNA packaging, establishment and/or maintenance of chromatin architecture, regulation of progression through the cell cycle, regulation of the cell cycle, nucleobase, nucleoside, nucleotide and nucleic acid metabolism, chromatin assembly, macromolecule biosynthesis, intracellular transport, establishment of cellular localization, cellular localization, nucleosome assembly, macromolecule metabolism, or M-phase; genes involved in secondary metabolism or genes involved in seed and/or seed coat architecture.

[45] A "transcription termination and polyadenylation region" as used herein is a sequence that drives the cleavage of the nascent RNA, whereafter a poly(A) tail is added at the resulting RNA 3' end, functional in plant cells. Transcription termination and polyadenylation signals functional in plant cells include, but are not limited to, 3'nos, 3'35S, 3'his and 3'g7.
The term "protein" interchangeably used with the term "polypeptide" as used herein describes a group of molecules consisting of more than 30 amino acids, whereas the term "peptide" describes molecules consisting of up to 30 amino acids. Proteins and peptides may further form dimers, trimers and higher oligomers, i.e. consisting of more than one (poly)peptide molecule. Protein or peptide molecules forming such dimers, trimers etc. may be identical or non-identical. The corresponding higher order structures are, consequently, termed homo- or heterodimers, homo- or heterotrimers etc. The terms "protein" and "peptide" also refer to naturally modified proteins or peptides wherein the modification is effected e.g. by glycosylation, acetylation, phosphorylation and the like. Such modifications are well known in the art.

The term "heterologous" refers to the relationship between two or more nucleic acid or protein sequences that are derived from different sources. For example, a promoter is heterologous with respect to an operably linked DNA region, such as a coding sequence if such a combination is not normally found in nature. In addition, a particular sequence may be "heterologous" with respect to a cell or organism into which it is inserted (i.e. does not naturally occur in that particular cell or organism).

The term "recombinant gene" refers to any gene that contains: a) DNA sequences, including regulatory and coding sequences that are not found together in nature, or b) sequences encoding parts of proteins not naturally adjoined, or c) parts of promoters that are not naturally adjoined. Accordingly, a recombinant gene may comprise regulatory sequences and coding sequences that are derived from different sources, or comprise regulatory sequences, and coding sequences derived from the same source, but arranged in a manner different from that found in nature.

Any of the promoters and heterologous nucleic acid sequences described above may be provided in a recombinant vector. A recombinant vector typically comprises, in a 5' to 3' orientation: a promoter to direct the transcription of a nucleic acid sequence and a nucleic acid sequence. The recombinant vector may further comprise a 3' transcriptional terminator, a 3' polyadenylation signal, other untranslated nucleic acid sequences, transit and targeting nucleic acid sequences, selectable markers, enhancers, and operators, as desired. The wording "5' UTR" refers to the untranslated region of DNA upstream, or 5' of the coding region of a gene and "3' UTR" refers to the untranslated region of DNA downstream, or 3' of the coding region of a gene. Means for preparing recombinant vectors are well known in the art. Methods for making recombinant vectors particularly suited to plant transformation are described in US4971908, US4940835, US4769061 and US4757011. Typical vectors useful for expression of nucleic acids in higher plants are well known in the art and include vectors derived from the tumor-inducing (Ti) plasmid of Agrobacterium tumefaciens. One or more additional promoters may also be provided in the recombinant vector. These promoters may be operably linked, for example, without limitation, to any of the nucleic acid sequences described above. Alternatively, the promoters may be operably linked to other nucleic acid sequences, such as those encoding transit
peptides, selectable marker proteins, or antisense sequences. These additional promoters may be selected on the basis of the cell type into which the vector will be inserted. Also, promoters which function in bacteria, yeast, and plants are all well taught in the art. The additional promoters may also be selected on the basis of their regulatory features. Examples of such features include enhancement of transcriptional activity, inducibility, tissue specificity, and developmental stage-specificity.

[50] The recombinant vector may also contain one or more additional nucleic acid sequences. These additional nucleic acid sequences may generally be any sequences suitable for use in a recombinant vector. Such nucleic acid sequences include, without limitation, any of the nucleic acid sequences, and modified forms thereof, described above. The additional structural nucleic acid sequences may also be operably linked to any of the above described promoters. The one or more structural nucleic acid sequences may each be operably linked to separate promoters. Alternatively, the structural nucleic acid sequences may be operably linked to a single promoter (i.e., a single operon).

[51] Yet another embodiment provides a host cell, such as an E. coli cell, an Agrobacterium cell, a yeast cell, or a plant cell, comprising the isolated nucleic acid according to the invention, or the recombinant gene according to the invention.

[52] Other nucleic acid sequences may also be introduced into the host cell along with the promoter and structural nucleic acid sequence, e.g. also in connection with the vector of the invention. These other sequences may include 3’ transcriptional terminators, 3’ polyadenylation signals, other untranslated nucleic acid sequences, transit or targeting sequences, selectable markers, enhancers, and operators. Preferred nucleic acid sequences of the present invention, including recombinant vectors, structural nucleic acid sequences, promoters, and other regulatory elements, are described above.

[53] In a further embodiment, a plant is provided comprising the recombinant gene according to the invention. In yet a further embodiment, a plant is provided comprising at least two recombinant genes according to the invention, wherein the nucleic acid comprising seed-preferential promoter activity according to the invention and, optionally, the heterologous nucleic acid sequence operably linked thereto, are different in each recombinant gene. Yet a further embodiment provides seeds obtainable from the plant according to the invention. In another embodiment, the plants or seeds according to the invention are seed crop plants or seeds.

[54] The plant cell or plant comprising the recombinant gene according to the invention can be a plant cell or a plant comprising a recombinant gene of which either the promoter, or the heterologous nucleic acid sequence operably linked to said promoter, are heterologous with respect to the plant cell. Such plant cells or plants may be transgenic plant in which the recombinant gene is introduced via transformation. Alternatively, the plant cell of plant may comprise the promoter according to the invention derived from the same species operably linked to a nucleic acid which is also derived from the
same species, i.e. neither the promoter nor the operably linked nucleic acid is heterologous with respect to the plant cell, but the promoter is operably linked to a nucleic acid to which it is not linked in nature. A recombinant gene can be introduced in the plant or plant cell via transformation, such that both the promoter and the operably linked nucleotide are at a position in the genome in which they do not occur naturally. Alternatively, the promoter according to the invention can be integrated in a targeted manner in the genome of the plant or plant cell upstream of an endogenous nucleic acid encoding an expression product of interest, i.e. to modulate the expression pattern of an endogenous gene. The promoter that is integrated in a targeted manner upstream of an endogenous nucleic acid can be integrated in cells of a plant species from which it is originally derived, or in cells of a heterologous plant species. Alternatively, a heterologous nucleic acid can be integrated in a targeted manner in the genome of the plant or plant cell downstream of the promoter according to the invention, such that said heterologous nucleic acid is expressed seed-preferentially. Said heterologous nucleic acid is a nucleic acid which is heterologous with respect to the promoter, i.e. the combination of the promoter with said heterologous nucleic acid is not normally found in nature. Said heterologous nucleic acid may be a nucleic acid which is heterologous to said plant species in which it is inserted, but it may also naturally occur in said plant species at a different location in the plant genome. Said promoter or said heterologous nucleic acid can be integrated in a targeted manner in the plant genome via targeted sequence insertion, using, for example, the methods as described in WO2005/049842.

[55] Plants according to the invention may comprise one or more recombinant genes according to the invention, but may in addition contain a recombinant gene comprising a nucleic acid comprising promoter activity which is preferential or specific to other plant tissues, such as apical meristem, flower buds, cotyledons, flowers, pods, roots, and leaves, or other seed developmental stages, operably linked to a nucleic acid sequence encoding an expression product of interest. The recombinant gene according to the invention and the recombinant gene comprising a nucleic acid comprising another promoter activity may be present at one locus and may be derived from the same transforming DNA molecule.

[56] Yet another embodiment provides a method of producing a transgenic plant comprising the steps of (a) introducing or providing the recombinant gene according to the invention to a plant cell to create transgenic cells; and (b) regenerating transgenic plants from said transgenic cell.

[57] "Introducing" in connection with the present application relates to the placing of genetic information in a plant cell or plant by artificial means. This can be effected by any method known in the art for introducing RNA or DNA into plant cells, protoplasts, calli, roots, tubers, seeds, stems, leaves, seedlings, embryos, pollen and microspores, other plant tissues, or whole plants. "Introducing" also comprises stably integrating into the plant's genome. Introducing the recombinant gene can be performed by transformation.
The term "transformation" herein refers to the introduction (or transfer) of nucleic acid into a recipient host such as a plant or any plant parts or tissues including plant cells, protoplasts, calli, roots, tubers, seeds, stems, leaves, seedlings, embryos and pollen. Plants containing the transformed nucleic acid sequence are referred to as "transgenic plants". Transformed, transgenic and recombinant refer to a host organism such as a plant into which a heterologous nucleic acid molecule (e.g. an expression cassette or a recombinant vector) has been introduced. The nucleic acid can be stably integrated into the genome of the plant.

As used herein, the phrase "transgenic plant" refers to a plant having an introduced nucleic acid stably introduced into a genome of the plant, for example, the nuclear or plastid genomes. In other words, plants containing transformed nucleic acid sequence are referred to as "transgenic plants". Transgenic and recombinant refer to a host organism such as a plant into which a heterologous nucleic acid molecule (e.g. the promoter, the chimeric gene or the vector as described herein) has been introduced. The nucleic acid can be stably integrated into the genome of the plant.

Transformation methods are well known in the art and include Agrobacterium-mediated transformation. Agrobacterium-mediated transformation of cotton has been described e.g. in US patent 5,004,863, in US patent 6,483,013 and WO2000/71733. Plants may also be transformed by particle bombardment: Particles of gold or tungsten are coated with DNA and then shot into young plant cells or plant embryos. This method also allows transformation of plant plastids. Viral transformation (transduction) may be used for transient or stable expression of a gene, depending on the nature of the virus genome. The desired genetic material is packaged into a suitable plant virus and the modified virus is allowed to infect the plant. The progeny of the infected plants is virus free and also free of the inserted gene. Suitable methods for viral transformation are described or further detailed e.g. in WO 90/12107, WO 03/052108 or WO 2005/098004. Further suitable methods well-known in the art are microinjection, electroporation of intact cells, polyethylene glycol-mediated protoplast transformation, electroporation of protoplasts, liposome-mediated transformation, silicon-whiskers mediated transformation etc. Said transgene may be stably integrated into the genome of said plant cell, resulting in a transformed plant cell. The transformed plant cells obtained in this way may then be regenerated into mature fertile transformed plants.

Further provided is a method of effecting seed-preferential expression of a nucleic acid comprising introducing the recombinant gene according to the invention into the genome of a plant, or providing the plant according to the invention. Also provided is a method for altering seed properties of a plant or to produce a commercially relevant product in a plant, comprising introducing the recombinant gene according to the invention into the genome of a plant, or providing the plant according to the invention. In another embodiment, said plant is a seed crop plant.
"Seed properties" as used herein are properties of the seed. Seed properties can, for example, be seed yield, seed storage compound production, seed compound accumulation, seed nutrient accumulation; seed micronutrient accumulation; seed storage compound quality, seed compound composition, seed quality, biotic stress tolerance such as disease tolerance, abiotic stress tolerance, herbicide tolerance, seed dormancy, seed imbition, seed germination, seed vigor. Seed storage compounds can, for example, be, seed oil, seed starch, or seed protein.

Seed properties may be modulated by modulating metabolic pathways, such as starch metabolism, sugar metabolism, inositol phosphate metabolism, glycolysis, amino acid biosynthesis, carbon metabolism, nucleotide metabolism, oxidative pentose phosphate cycle, fatty acid biosynthesis, protein synthesis, or phytate metabolism, and modulating secondary metabolism pathways. Another example is the methyl recycling metabolic activity impacting chromatin remodeling, phospholipid biosynthesis and cell wall lignification. Such metabolic pathways can be modulated by, for example, overexpressing or downregulating a gene involved in one or more of the metabolic pathways using the seed-preferential promoter according to the invention.

Yield as used herein can comprise yield of the plant or plant part which is harvested, such as seed, including seed oil content, seed protein content, seed weight, seed number. Increased yield can be increased yield per plant, and increased yield per surface unit of cultivated land, such as yield per hectare. Yield can be increased by modulating, for example, by increasing seed size or oil content or indirectly by increasing the tolerance to biotic and abiotic stress conditions and decreasing seed abortion.

Quality as used herein can comprise quality of the seed or grain such as beneficial carbohydrate composition or level, beneficial amino acid composition or level, beneficial fatty acid composition or level, nutritional value, seed and fiber content.

Abiotic stress tolerance as used herein can comprise resistance to environmental stress factors such as drought, extreme (high or low) temperatures.

Biotic stress tolerance as used herein can comprise pest resistance, such as resistance or fungal, bacterial, bacterial or viral pathogens or insects.

Also provided is the use of the isolated nucleic acid according to the invention to regulate expression of an operably linked nucleic acid in a plant, and the use of the isolated nucleic acid according to the invention, or the recombinant gene according to the invention to alter seed properties of a plant or to produce a commercially relevant product in a plant. In a further embodiment, said plant is a

A trait as used herein refers to beneficial properties of the plant, such as commercially beneficial properties of a plant.
[69] Also provided is the use of the isolated nucleic acid according to the invention to identify other nucleic acids comprising seed-preferential promoter activity.

[70] Other nucleic acids comprising seed-preferential promoter activity can be identified using methods known in the art. Such nucleotide sequence may be identified and isolated by hybridization under stringent conditions using as probes a nucleic acid comprising the nucleotide sequence of SEQ ID NO: 1 or part thereof. Other nucleic acids comprising seed-preferential promoter activity may also be obtained by DNA amplification using oligonucleotides specific for the sequences according to the invention as primers, such as but not limited to oligonucleotides comprising or consisting of about 20 to about 50 consecutive nucleotides from the nucleotide sequence of SEQ ID NO: 1 or its complement. Other nucleic acids comprising seed-preferential promoter activity can be identified in silico using Basic Local Alignment Search Tool (BLAST) homology search with other nucleotide or amino acid sequences. Functionality of the identified nucleic acids comprising seed-preferential promoter activity can be validated using the methods described herein. Other nucleic acids comprising seed-preferential promoter activity may also be identified by identification of gene sequences orthologous to the gene sequences of the endogenous coding sequences of the genes driven by the promoters of the invention, and isolating the promoter sequences upstream of these orthologous homologous coding sequences.

[71] The promoters according to the invention can further be used to create hybrid promoters, i.e. promoters containing (parts of) one or more of the promoters(s) of the current invention and (parts of) other promoter which can be newly identified or known in the art. Such hybrid promoters may have optimized tissue specificity or expression level.

[72] Yet another embodiment provides a method of producing food, feed, or an industrial product comprising (a) obtaining the plant or a part thereof, according to the invention; and (b) preparing the food, feed or industrial product from the plant or part thereof. In another embodiment, said food or feed is oil, meal, grain, starch, flour or protein, or said industrial product is biofuel, fiber, industrial chemicals, a pharmaceutical or a nutraceutical.

[73] A "seed crop" or "seed crop plant" as used herein is a crop grown for its seeds or material derived from the seeds. Examples of seed crops are rice, maize, wheat, barley, millet, rye, oats, camelina, crambe, Linum, castor bean, calendula, safflower, sunflower, soybean, cotton, or Brassica species, such as Brassica napus, Brassica juncea, Brassica carinata, Brassica rapa, Brassica oleracea, and Brassica nigra.

[74] "Brassicaceae" or "Brassicaceae plant" as used herein refers to plants belonging to the family of Brassicaceae plants, also called Cruciferae or mustard family. Examples of Brassicaceae are, but are not limited to, Brassica species, such as Brassica napus, Brassica oleracea, Brassica rapa, Brassica carinata, Brassica nigra, and Brassica juncea; Raphanus species, such as Raphanus caudatus,
Raphanus raphanistrum, and Raphanus sativus; Matthiola species; Cheiranthus species; Camelina species, such as Camelina sativa; Crambe species, such as Crambe abyssinica and Crambe hispanica; Erura species, such as Erura vesicaria; Sinapis species such as Sinapis alba; Diplotaxis species; Lepidium species; Nasturtium species; Orychophragmus species; Armoracia species, Eutrema species; Lepidium species; and Arabidopsis species.

[75] Said Brassicaceae plant can be a Brassica plant. "Brassica plant" refers to allotetraploid or amphidiploid Brassica napus (AACC, 2n=38), Brassica juncea (AABB, 2n=36), Brassica carinata (BBCC, 2n=34), or to diploid Brassica rapa (syn. B. campestris) (AA, 2n=20), Brassica oleracea (CC, 2n=18) or Brassica nigra (BB, 2n=16).

[76] Crop plants of the Brassica species are, for example, Brassica napus, Brassica juncea, Brassica carinata, Brassica rapa (syn. B. campestris), Brassica oleracea or Brassica nigra.

[77] The plants according to the invention may additionally contain an endogenous or a transgene, which confers herbicide resistance, such as the bar or pat gene, which confer resistance to glufosinate ammonium (Liberty®, Basta® or Ignite®) [EP 0 242 236 and EP 0 242 246 incorporated by reference]; or any modified EPSPS gene, such as the 2mEPSPS gene from maize [EPO 508 909 and EP 0 507 698 incorporated by reference], or glyphosate acetyltransferase, or glyphosate oxidoreductase, which confer resistance to glyphosate (RoundupReady®), or bromoxynitril nitrilase to confer bromoxynitril tolerance, or any modified AHAS gene, which confers tolerance to sulfonylureas, imidazolinones, sulfonylaminocarbonyltriazolinones, triazolopyrimidines or pyrimidyl(oxy/thio)benzoates, such as oilseed rape imidazolinone-tolerant mutants PM1 and PM2, currently marketed as Clearfield® canola. Further, the plants according to the invention may additionally contain an endogenous or a transgene which confers increased oil content or improved oil composition, such as a 12:0 ACP thioesterase increase to obtain high laureate, which confers pollination control, such as such as barnase under control of an anther-specific promoter to obtain male sterility, or barstar under control of an anther-specific promoter to confer restoration of male sterility, or such as the Ogura cytoplasmic male sterility and nuclear restorer of fertility.

[78] The plants or seeds of the plants according to the invention may be further treated with a chemical compound, such as a chemical compound selected from the following lists: Herbicides: Clethodim, Clopyralid, Diclofop, Ethamsulfuron, Fluazifop, Glufosinate, Glyphosate, Metazachlor, Quinmerac, Quizalofop, Tepraloxydim, Trifluralin. Fungicides / PGRs: Azoxy streobin, N-[9-(dichloromethylene)-1,2,3,4-tetrahydro-1,4-methan naphthalen-5-yl]-3-(difluoromethyl)-1-methyl-1H-pyrazole-4-carboxamide (Benzovindiflupyr, Ben zodiflupyr), Bixafen, Boscalid, Car bendazim, Carboxin, Chlor mequat-chloride, Coniothyrium mimitans, Cyproconazole, Cyprodinil, Difenconazole, Dimethomorph, Dimoxystrobin, Ep oxiconazole, Fam oxadone, Flu azinam, Fludioxonil, Fluopicoline, Fluopyram, Fluoxastrobin, Fluquinconazole,
Flusilazole, Fluthianil, Flutriafol, Fluxapyroxad, Iprodione, Isopyrazam, Mefenoxam, Mepiquat-chloride, Metalaxyl, Metconazole, Metominostrobin, Paclorbutrazole, Penflufen, Penthiopyrad, Picoxystrobins, Prochloraz, Prothioconazole, Pyraclostrobin, Sedaxane, Tebuconazole, Tetracozazole, Thiophanate-methyl, Thiram, Triadimenol, Trifloxystrobin, Bacillus firmus, Bacillus firmus strain 1-1582, Bacillus subtilis, Bacillus subtilis strain GB03, Bacillus subtilis strain QST 713, Bacillus pumulis, Bacillus pumulis strain GB34.

Insecticides: Acetamiprid, Aldicarb, Azadirachtin, Carbofuran, Chloranlaniliprole (Rynaxypyr), Clothianidin, Cyaanlaniliprole (Cyazypyr), (beta-)Cyfluthrin, gamma-Cyhalothrin, lambda-Cyhalothrin, Cypermethrin, Deltamethrin, Dimethoate, Dinotefuran, Ethiprole, Fonicamid, Flubendiamide, Fluensulfone, Fluopyram, Fluypyradifurone, tau-Fluralulate, Imicyafos, Imidacloprid, Metaflumizone, Methiocarb, Pymetrozine, Pyrifluquinazon, Spinetoram, Spinosad, Spiroetramate, Sulfoxaflor, Thiacloriprid, Thiamethoxam, 1-(3-chloropyridin-2-yl)-N-[4-cyano-2-methyl-6-(methylcarbamoyl)phenyl]-3-[[5-(trifluoromethyl)-2H-tetrazol-2-yl]methyl]-1H-pyrazole-5-carboxamide, 1-(3-chloropyridin-2-yl)-N-[4-cyano-2-methyl-6-(methylcarbamoyl)phenyl]-3-[[5-(trifluoromethyl)-1H-tetrazol-1-yl]methyl]-1H-pyrazole-5-carboxamide, 1-[(2-fluoro-4-methyl-5-[(2,2,2-trifluoromethyl)sulfinyl]phenyl]-3-(trifluoromethyl)-1H-1,2,4-triazol-5-amine, (1E)-N-[(6-chloropyridin-3-yl)methyl]-N'-cyano-N-(2,2-difluoroethyl)thetanimidamide, Bacillus firmus, Bacillus firmus strain 1-1582, Bacillus subtilis, Bacillus subtilis strain GB03, Bacillus subtilis strain QST 713, Metarhizium anisopliae F52.

Whenever reference to a "plant" or "plants" according to the invention is made, it is understood that also plant parts (cells, tissues or organs, seed pods, seeds, severed parts such as roots, leaves, flowers, pollen, etc.), progeny of the plants which retain the distinguishing characteristics of the parents, such as seed obtained by selling or crossing, e.g. hybrid seed (obtained by crossing two inbred parental lines), hybrid plants and plant parts derived there from are encompassed herein, unless otherwise indicated.

In some embodiments, the plant cells of the invention as well as plant cells generated according to the methods of the invention, may be non-propagating cells.

The obtained plants according to the invention can be used in a conventional breeding scheme to produce more plants with the same characteristics or to introduce the same characteristic in other varieties of the same or related plant species, or in hybrid plants. The obtained plants can further be used for creating propagating material. Plants according to the invention can further be used to produce gametes, seeds (including crushed seeds and seed cakes), seed oil, embryos, either zygotic or somatic, progeny or hybrids of plants obtained by methods of the invention. Seeds obtained from the plants according to the invention are also encompassed by the invention.
"Creating propagating material", as used herein, relates to any means known in the art to produce further plants, plant parts or seeds and includes inter alia vegetative reproduction methods (e.g. air or ground layering, division, (bud) grafting, micropropagation, stolons or runners, storage organs such as bulbs, corms, tubers and rhizomes, striking or cutting, twin-scaling), sexual reproduction (crossing with another plant) and asexual reproduction (e.g. apomixis, somatic hybridization).

As used herein "comprising" is to be interpreted as specifying the presence of the stated features, integers, steps or components as referred to, but does not preclude the presence or addition of one or more features, integers, steps or components, or groups thereof. Thus, e.g., a nucleic acid or protein comprising a sequence of nucleotides or amino acids, may comprise more nucleotides or amino acids than the actually cited ones, i.e., be embedded in a larger nucleic acid or protein. A chimeric gene comprising a nucleic acid which is functionally or structurally defined, may comprise additional DNA regions etc.

The sequence listing contained in the file named „BCS15-2013_ST25.txt”, which is 5 kilobytes (size as measured in Microsoft Windows®), contains the sequence SEQ ID NO: 1, is filed herewith by electronic submission and is incorporated by reference herein.

In the description and examples, reference is made to the following sequences:

SEQUENCES

SEQ ID NO: 1: 3 to 5 week-old seed-specific promoter.

EXAMPLES

Example 1 - RNA isolation from different tissues of *Brassica napus*

The following tissues were isolated from *Brassica napus*:

- a. Apical meristem 33 days after sowing (DAS) (including smallest leaves) (AM33)
- b. Big flower buds (> 5 mm) 42 DAS (BFB42)
- c. Cotyledons (with hypocotyl) 10 DAS (CTYL10)
- d. Open flowers 52 DAS (OF52)
- e. Pods 14-20 DAS (Pod2)
- f. Pods 21-25 DAS (Pod3)
- g. Roots 14 DAS (Pvo2w)
- h. Seeds 14-20 days after flowering (DAF) (Seed2)
- i. Seeds 21-25 DAF (Seed3)
- j. Seeds 26-30 DAF (Seed4)
- k. Seeds 31-35 DAF (Seed5)
- l. Seeds 42 DAF (Seed6)
- m. Seeds 49 DAF (Seed7)
- n. Small flower buds ≤ 5 mm 42 DAS (SFB42)
- o. Stem 14 DAS (St2w)
- p. Stem 33 DAS (St5w)
- q. Young leaf 33 DAS (< 3 cm leaf next to apical meristem) (YL33)

Total RNA from the above tissues was isolated according to standard methods.

Example 2 - Identification of *Brassica napus* 3 to 5 week-old seed-specific transcripts

Transcript activity in the different *B. napus* tissues was measured using transcript profiling (RNA-seq) by Illumina HiSeq paired-end sequencing (2x100 bp sequences).

The counts per transcript (transcript per million; tpm) was determined as described by Li and Dewey, 2011, BMC Bioinformatics 12: 323 (incorporated herein by reference). Sequence reads were mapped to transcripts using an in-house transcriptome dataset.

The tpm values were normalized as described by Anders and Huber, 2010, Genome Biol 11: R106 (incorporated herein by reference).

3 to 5 week-old seed-specific transcripts were selected using the following criteria:

- a. low transcript activity in AM33, BFB42, CTYL0, OF52, Pod2, Pod3, Ro2w, SFB42, St2w, St5w, YL33, Seed2, , Seed6 or Seed7;
- b. transcript should be detected in Seed3, Seed4, and Seed5tissues.
Transcripts were identified having a 3-fold higher expression in each of the 3 week-old, 4 week-old and 5 week-old seed tissues as compared to each of the reference tissues, based on the normalized values.

The one transcript fulfilling the above criteria with the highest expression in the 3, 4 and 5 week-old seed tissue were selected. Figure 1 shows the relative expression levels of the identified transcript in the different tissues.

Example 3 - Identification of the *Brassica napus* seed-preferential promoter

The sequences of the transcript as identified above were blasted against an in-house database of *Brassica napus* sequences. Sequences upstream of the predicted ATG translation start codon of the above-identified transcripts until the next upstream predicted gene, or until 3 kb if the next upstream predicted gene was located more than 3 kb further upstream, were obtained. The sequences obtained in this way, comprising the promoters conferring seed-preferential expression of the transcript Seed_345 is given in SEQ ID NO: 1.
CLAIMS:

1. An isolated nucleic acid comprising seed-preferential promoter activity selected from the group consisting of:
   a. a nucleic acid comprising a nucleotide sequence of SEQ ID NO: 1 or a functional fragment thereof;
   b. a nucleic acid comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO: 1, or a functional fragment thereof; and
   c. a nucleic acid capable of hybridizing under stringent conditions to the nucleotide sequence of SEQ ID NO: 1, or a functional fragment thereof.

2. A recombinant gene comprising the nucleic acid according to claim 1 operably linked to a heterologous nucleic acid sequence encoding an expression product of interest, and optionally a transcription termination and polyadenylation sequence, preferably a transcription termination and polyadenylation region functional in plants.

3. The recombinant gene according to claim 2, wherein the expression product of interest is an RNA molecule capable of modulating the expression of a gene or is a protein.

4. A host cell comprising the isolated nucleic acid according to claim 1, or the recombinant gene according to claim 2 or 3.

5. The host cell of claim 4 which is an E. coli cell, an Agrobacterium cell, yeast cell, or a plant cell.

6. A plant comprising the recombinant gene of claim 2 or 3.

7. Seeds obtainable from the plant according to claim 6.

8. The plant or plant cell or seeds according to any one of claims 5 to 7, which is a seed crop plant, or a cell or plant part.

9. Method of producing a transgenic plant comprising the steps of:
   a. introducing or providing the recombinant gene according to claim 2 or 3 to a plant cell to create transgenic cells; and
   b. regenerating transgenic plants from said transgenic cell.

10. Method of effecting seed-preferential expression of a nucleic acid comprising introducing the recombinant gene according to claim 2 or 3 into the genome of a plant, or providing the plant according to claim 6.
11. Method for altering seed properties of a plant or to produce a commercially relevant product in a plant, said method comprising introducing the recombinant gene according to claim 2 or 3 into the genome of a plant, or providing the plant according to claim 6.

12. Use of the isolated nucleic acid according to claim 1 to regulate expression of an operably linked nucleic acid in a plant.

13. Use of the isolated nucleic acid according to claim 1, or the recombinant gene according to claim 2 or 3 to alter seed properties of a plant or to produce a commercially relevant product in a plant.

14. Use of the isolated nucleic acid according to claim 1 to identify other nucleic acids comprising seed-preferential promoter activity.

15. The method according to any one of claims 9 to 11, or the use according to claims 12 or 13, wherein said plant is a seed crop plant.

16. A method of producing food, feed, or an industrial product comprising
   a) obtaining the plant or a part thereof, of any one of claims 6 to 8; and
   b) preparing the food, feed or industrial product from the plant or part thereof.

17. The method of claim 16 wherein
   a) the food or feed is oil, meal, grain, starch, flour or protein; or
   b) the industrial product is biofuel, fiber, industrial chemicals, a pharmaceutical or a nutraceutical.
**INTERNATIONAL SEARCH REPORT**

**International application No**
PCT/EP2016/073664

**A. CLASSIFICATION OF SUBJECT MATTER**

<table>
<thead>
<tr>
<th>INV.</th>
<th>C12N15/82</th>
<th>C07K14/415</th>
<th>A01H5/Q0</th>
<th>A01H5/10</th>
<th>C12N1/19</th>
</tr>
</thead>
<tbody>
<tr>
<td>C12N1/21</td>
<td>C12N5/04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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)

C07K  C12N

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)

EPO-Internal , Sequence Search , BIOSIS, EMBASE, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>DATABASE EMBL [Onl ine]</td>
<td>1, 4</td>
</tr>
<tr>
<td></td>
<td>26 September 2008 (2008-09-26) &quot;Brassica napus GSS, clone JBnB118A23, end read, primer SACL2 (forward) &quot;, XP002764573, retrieved from EBI access via no. EM_GSS:CU953879 Database accession no. CU953879 sequence</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;/   &quot;</td>
<td></td>
</tr>
</tbody>
</table>

* 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

"Z" document member of the same patent family

Date of the actual completion of the international search

24 November 2016

Date of mailing of the international search report

07/12/2016

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

Oderwald, Harald
Category | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No.
--- | --- | ---
A | WO 92/18634 Al (UNILEVER PLC [GB]; UNILEVER NV [NL]) 29 October 1992 (1992-10-29) the whole document | 1-17
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CA 2708087 A1</td>
<td>25-06-2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 112008003237 T5</td>
<td>17-03-2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 2222856 A2</td>
<td>01-09-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2010263088 A1</td>
<td>14-10-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wo 2009077478 A2</td>
<td>25-06-2009</td>
</tr>
<tr>
<td>wo 9218634 Al</td>
<td>29-10-1992</td>
<td>AT 202144 T</td>
<td>15-06-2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 669478 B2</td>
<td>13-06-1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2106960 A1</td>
<td>10-10-1992</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 69231875 D1</td>
<td>19-07-2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 69231875 T2</td>
<td>04-10-2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 0580649 T3</td>
<td>27-08-2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0580649 A1</td>
<td>02-02-1994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES 2157202 T3</td>
<td>16-08-2001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 3730658 B2</td>
<td>05-01-2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP H06506584 A</td>
<td>28-07-1994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 5767363 A</td>
<td>16-06-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wo 9218634 A1</td>
<td>29-10-1992</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ZA 9202592 A</td>
<td>11-10-1993</td>
</tr>
</tbody>
</table>