



(86) Date de dépôt PCT/PCT Filing Date: 2009/06/30
(87) Date publication PCT/PCT Publication Date: 2010/01/07
(45) Date de délivrance/Issue Date: 2018/02/13
(85) Entrée phase nationale/National Entry: 2010/12/24
(86) N° demande PCT/PCT Application No.: EP 2009/058138
(87) N° publication PCT/PCT Publication No.: 2010/000708
(30) Priorité/Priority: 2008/07/01 (EP08159440.0)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01),
C07K 14/415 (2006.01)
(72) Inventeurs/Inventors:
BAUER, JOERG, DE;
SENGER, TORALF, DE
(73) Propriétaire/Owner:
BASF PLANT SCIENCE GMBH, DE
(74) Agent: ROBIC

(54) Titre : PROMOTEURS DE BRASSICA NAPUS POUR L'EXPRESSION DE GENES SPECIFIQUE DE GRAINES
(54) Title: PROMOTERS FROM BRASSICA NAPUS FOR SEED SPECIFIC GENE EXPRESSION

(57) **Abrégé/Abstract:**

The present invention is concerned with means and methods for allowing tissue specific and, in particular, seed specific expression of genes. The present invention, accordingly, relates to a polynucleotide comprising an expression control sequence which allows seed specific expression of a nucleic acid of interest being operatively linked thereto. Moreover, the present invention contemplates vectors, host cells, non-human transgenic organisms comprising the aforementioned polynucleotide as well as methods and uses of such a polynucleotide.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
7 January 2010 (07.01.2010)

PCT

(10) International Publication Number
WO 2010/000708 A3

(51) International Patent Classification:

C12N 15/82 (2006.01) A01H 5/00 (2006.01)
C07K 14/415 (2006.01)

(21) International Application Number:

PCT/EP2009/058138

(22) International Filing Date:

30 June 2009 (30.06.2009)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

08159440.0 1 July 2008 (01.07.2008) EP

(71) Applicant (for all designated States except US): **BASF PLANT SCIENCE GMBH** [DE/DE]; 67056 Ludwigshafen (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BAUER, Jörg** [DE/DE]; Maxburgstrasse 6, 67117 Limburgerhof (DE). **SENGER, Toralf** [DE/DE]; Grindelbachstrasse 112b, 69469 Weinheim (DE).(74) Agent: **HERZOG FIESSER & PARTNER**; Bad Kreuznacher Straße 27-29, 68309 Mannheim (DE).

(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, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

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

22 April 2010

(54) Title: PROMOTERS FROM BRASSICA NAPUS FOR SEED SPECIFIC GENE EXPRESSION

(57) Abstract: The present invention is concerned with means and methods for allowing tissue specific and, in particular, seed specific expression of genes. The present invention, accordingly, relates to a polynucleotide comprising an expression control sequence which allows seed specific expression of a nucleic acid of interest being operatively linked thereto. Moreover, the present invention contemplates vectors, host cells, non-human transgenic organisms comprising the aforementioned polynucleotide as well as methods and uses of such a polynucleotide.



WO 2010/000708 A3

Promoters from Brassica napus for seed specific gene expression

The present invention is concerned with means and methods for allowing tissue specific and, in particular, seed specific expression of genes. The present invention, accordingly, relates to a polynucleotide comprising an expression control sequence which allows seed specific expression of a nucleic acid of interest being operatively linked thereto. Moreover, the present invention contemplates vectors, host cells, non-human transgenic organisms comprising the aforementioned polynucleotide as well as methods and uses of such a polynucleotide.

10

In the field of "green" (agricultural) biotechnology, plants are genetically manipulated in order to confer beneficial traits. These beneficial traits may be yield increase, tolerance increase, reduced dependency on fertilizers, herbicidal, pesticidal- or fungicidal- resistance, or the capability of producing chemical specialties such as nutrients, drugs, oils for food and petrochemistry etc..

In many cases, it is required to express a heterologous gene in the genetically modified plants at a rather specific location in order to obtain a plant exhibiting the desired beneficial trait. One major location for gene expression is the plant seed. In the seeds, many important synthesis pathways, e.g., in fatty acid synthesis, take place. Accordingly, expression of heterologous genes in seeds allow for the manipulation of fatty acid synthesis pathways and, thus, for the provision of various fatty acid derivatives and lipid-based compounds.

25

However, for many heterologous genes, a seed specific expression will be required. Promoters which allow for a seed specific expression are known in the art. Such promoters include the oilseed rape napin promoter (US 5,608,152), the Vicia faba USP promoter (Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67), the Arabidopsis oleosin promoter (WO 98/45461), the Phaseolus vulgaris phaseolin promoter (US 5,504,200), the Brassica Bce4 promoter (WO 91/13980) or the legumine B4 promoter (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2):233-9), and promoters which bring about the seed-specific expression in monocotyledonous plants such as maize, barley, wheat, rye, rice and the like. Suitable noteworthy promoters are the barley lpt2 or lpt1 gene promoter (WO 95/15389 and WO 95/23230) or the promoters from the barley hordein gene, the rice glutelin gene, the rice oryzin gene, the rice prolamine gene, the wheat gliadine gene, the wheat glutelin gene, the maize zeine gene, the oat glutelin gene, the sorghum kasirin gene or the rye secalin gene, which are described in WO 99/16890.

40

However, there is a clear need for further expression control sequences such as promoters and terminators which allow for a reliable and efficient control of expression of foreign nucleic acids in seeds.

5

The technical problem underlying this invention can be seen as the provision of means and methods complying with the aforementioned needs. The technical problem is solved by the embodiments characterized in the claims and herein below.

10

Accordingly, the present invention relates to a polynucleotide comprising an expression control sequence which allows seed specific expression of a nucleic acid of interest being operatively linked thereto, said expression control sequence being selected from the group consisting of:

- 15 (a) an expression control sequence having a nucleic acid sequence as shown in any one of SEQ ID NOs: 1, 6, 9, 14, 16, 22, 25, 70, 77, 85, 95, 103, 111, 119, 124 or 131;
- (b) an expression control sequence having a nucleic acid sequence which is at least 80% identical to a nucleic acid sequence shown in any one of SEQ ID
20 NOs: 1, 6, 9, 14, 16, 22, 25, 70, 77, 85, 95, 103, 111, 119, 124 or 131;
- (c) an expression control sequence having a nucleic acid sequence which hybridizes under stringent conditions to a a nucleic acid sequence as shown in any one of SEQ ID NOs: 1, 6, 9, 14, 16, 22, 25, 70, 77, 85, 95, 103, 111, 119, 124 or 131;
- 25 (d) an expression control sequence having a nucleic acid sequence which hybridizes to a nucleic acid sequences located upstream of an open reading frame sequence shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125;
- (e) an expression control sequence having a nucleic acid sequence which hybridizes to a nucleic acid sequences located upstream of an open reading
30 frame sequence being at least 80% identical to an open reading frame sequence as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125;
- (f) an expression control sequence obtainable by 5' genome walking or by
35 thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) on genomic DNA from the first exon of an open reading frame sequence as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125; and
- (g) an expression control sequence obtainable by 5' genome walking or TAIL
40 PCR on genomic DNA from the first exon of an open reading frame se-

3

quence being at least 80% identical to an open reading frame as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125.

5 In accordance to a particular embodiment, the invention relates to an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or SEQ ID NO: 25; wherein said expression control sequence governs seed-specific expression of said
10 operatively linked nucleic acid of interest.

In accordance to another embodiment, the invention relates to a vector comprising the expression cassette of the present invention.

15 In accordance to another embodiment, the invention relates to a non-human host cell comprising the expression cassette of the present invention, or the vector of the present invention.

20 In accordance to another embodiment, the invention relates to a plant cell comprising the expression cassette of the invention, or the vector of the invention.

In accordance to another embodiment, the invention relates to a seed cell comprising the expression cassette of the invention, or the vector of the invention.

25 In accordance to another embodiment, the invention relates to a method of expressing a nucleic acid of interest in a host cell comprising:

- 30 (a) introducing an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or 25 or a vector comprising said expression control sequence, into a host cell; and
(b) expressing said nucleic acid of interest in said host cell.

35 In accordance to another embodiment, the invention relates to a method of expressing a nucleic acid of interest in a plant cell or seed cell comprising:

- (a) introducing an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said

3a

nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or 25 or a vector comprising said expression control sequence, into said plant cell or seed cell; and

- 5 (b) expressing said nucleic acid of interest in said plant cell or seed cell.

In accordance to another embodiment, the invention relates to a method of expressing a nucleic acid of interest in a non-human organism comprising:

- 10 (a) introducing an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or 25 or a vector comprising said expression control sequence, into a non-human organism; and
- 15 (b) expressing said nucleic acid of interest in said non-human transgenic organism.

In accordance to another embodiment, the invention relates to a method of expressing a nucleic acid of interest in a plant cell or seed cell comprising:

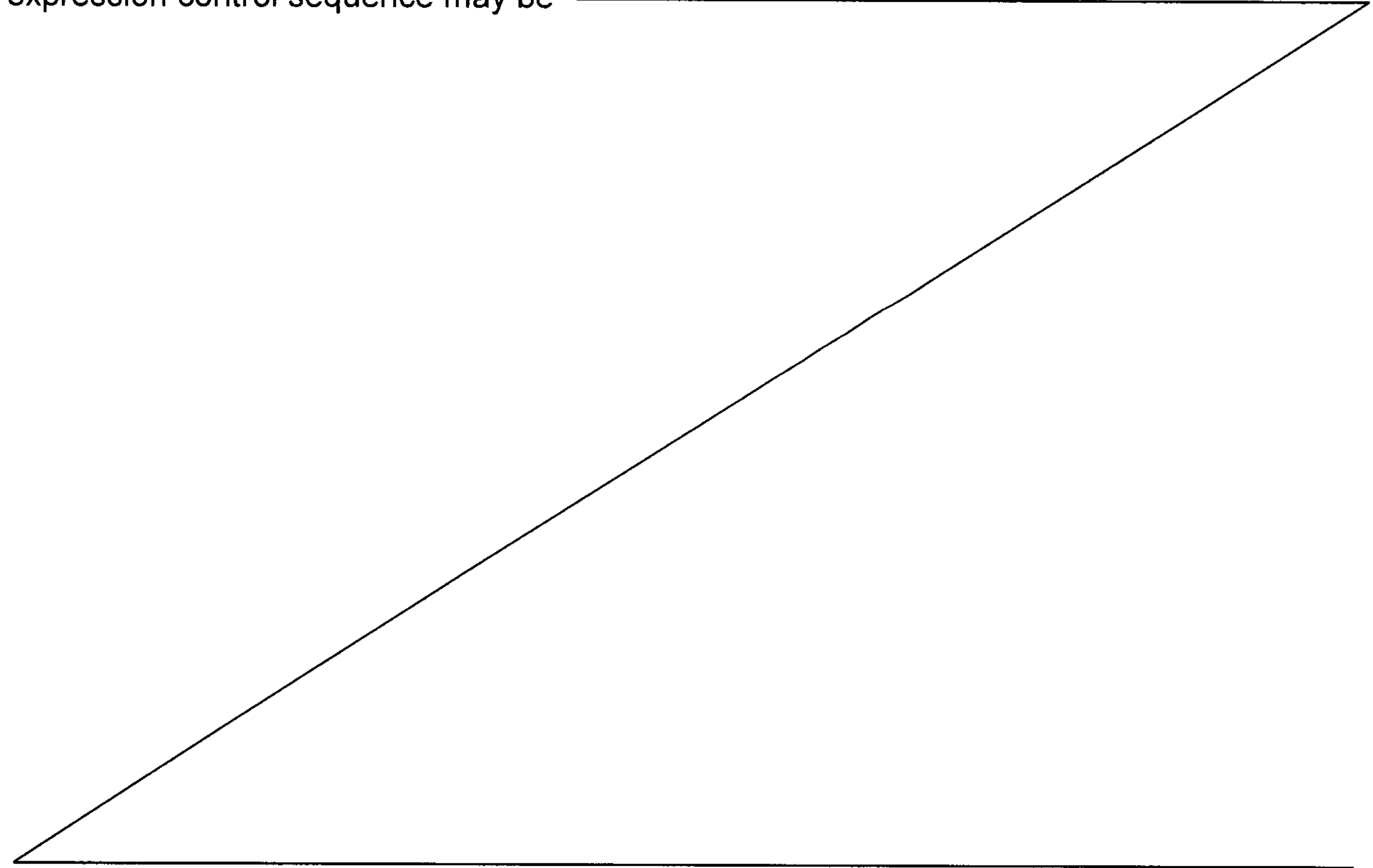
- 20 (a) introducing an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or 25 or a vector comprising said expression control sequence, into said plant cell or seed cell; and
- 25 (b) expressing said nucleic acid of interest in said plant cell or seed cell.

The term "polynucleotide" as used herein refers to a linear or circular nucleic acid molecule. It encompasses DNA as well as RNA molecules. The polynucleotide of the present invention is characterized in that it shall comprise an expression control sequence as defined elsewhere in this specification. In addition to the expression control sequence, the polynucleotide of the present invention, preferably, further comprises at least one nucleic acid of interest being operatively linked to the expression control sequence and/or a termination sequence or transcription. Thus, the polynucleotide of the present invention, preferably, comprises an expression cassette for the expression of at least one nucleic acid of interest. Alternatively, the polynucleotide may comprise in addition to the said expression control sequence a multiple cloning site and/or a termination sequence for transcription. In such a case, the multiple cloning site is, preferably, arranged in a manner as to allow for operative linkage of a nucleic acid to be introduced in the multiple cloning site with the

3b

expression control sequence. In addition to the aforementioned components, the polynucleotide of the present invention, preferably, could comprise components required for homologous recombination, i.e. flanking genomic sequences from a target locus. However, also preferably, the polynucleotide of the present invention can essentially consist of the said expression control sequence.

The term "expression control sequence" as used herein refers to a nucleic acid which is capable of governing the expression of another nucleic acid operatively linked thereto, e.g. a nucleic acid of interest referred to elsewhere in this specification in detail. An expression control sequence as referred to in accordance with the present invention, preferably, comprises sequence motifs which are recognized and bound by polypeptides, i.e. transcription factors. The said transcription factors shall upon binding recruit RNA polymerases, preferably, RNA polymerase I, II or III, more preferably, RNA polymerase II or III, and most preferably, RNA polymerase II. Thereby the expression of a nucleic acid operatively linked to the expression control sequence will be initiated. It is to be understood that dependent on the type of nucleic acid to be expressed, i.e. the nucleic acid of interest, expression as meant herein may comprise transcription of RNA polynucleotides from the nucleic acid sequence (as suitable for, e.g., anti-sense approaches or RNAi approaches) or may comprises transcription of RNA polynucleotides followed by translation of the said RNA polynucleotides into polypeptides (as suitable for, e.g., gene expression and recombinant polypeptide production approaches). In order to govern expression of a nucleic acid, the expression control sequence may be



located immediately adjacent to the nucleic acid to be expressed, i.e. physically linked to the said nucleic acid at its 5' end. Alternatively, it may be located in physical proximity. In the latter case, however, the sequence must be located so as to allow functional interaction with the nucleic acid to be expressed. An expression control sequence referred to herein, preferably, comprises between 200 and 5,000 nucleotides in length. More preferably, it comprises between 500 and 2,500 nucleotides and, more preferably, at least 1,000 nucleotides. As mentioned before, an expression control sequence, preferably, comprises a plurality of sequence motifs which are required for transcription factor binding or for conferring a certain structure to the polynucleotide comprising the expression control sequence. Sequence motifs are also sometimes referred to as cis-regulatory elements and, as meant herein, include promoter elements as well as enhancer elements. Preferred expression control sequences to be included into a polynucleotide of the present invention have a nucleic acid sequence as shown in any one of SEQ ID NOs: 1, 6, 9, 14, 16, 22, 25, 70, 77, 85, 95, 103, 111, 119, 124 and 131.

Further preferably, an expression control sequence comprised by a polynucleotide of the present invention has a nucleic acid sequence which hybridizes to a nucleic acid sequences located upstream of an open reading frame sequence shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125, i.e. is a variant expression control sequence. It will be understood that expression control sequences may slightly differ in its sequences due to allelic variations. Accordingly, the present invention also contemplates an expression control sequence which can be derived from an open reading frame as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125. Said expression control sequences are capable of hybridizing, preferably under stringent conditions, to the upstream sequences of the open reading frames shown in any one of SEQ ID NOs. 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125, i.e. the expression control sequences shown in any one of SEQ ID NOs.: 1, 6, 9, 14, 16, 22, 25, 70, 77, 85, 95, 103, 111, 119, 124 and 131. Stringent hybridization conditions as meant herein are, preferably, hybridization conditions in $6 \times$ sodium chloride/sodium citrate (= SSC) at approximately 45°C, followed by one or more wash steps in $0.2 \times$ SSC, 0.1% SDS at 53 to 65°C, preferably at 55°C, 56°C, 57°C, 58°C, 59°C, 60°C, 61°C, 62°C, 63°C, 64°C or 65°C. The skilled worker knows that these hybridization conditions differ depending on the type of nucleic acid and, for example when organic solvents are present, with regard to the temperature and concentration of the buffer. For example, under "standard hybridization conditions" the temperature differs depending on the type of nucleic acid between 42°C and 58°C in aqueous buffer with a concentration of 0.1 to $5 \times$ SSC (pH 7.2). If organic solvent is present in the abovementioned buffer, for example 50% formamide, the temperature

under standard conditions is approximately 42°C. The hybridization conditions for DNA:DNA hybrids are preferably for example 0.1 × SSC and 20°C to 45°C, preferably between 30°C and 45°C. The hybridization conditions for DNA:RNA hybrids are preferably, for example, 0.1 × SSC and 30°C to 55°C, preferably between 45°C and 55°C.

5 The abovementioned hybridization temperatures are determined for example for a nucleic acid with approximately 100 bp (= base pairs) in length and a G + C content of 50% in the absence of formamide. Such hybridizing expression control sequences are, more preferably, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94% at least 95%, at least 96%, at least 97%, at least 98%, or at

10 least 99% identical to the expression control sequences as shown in any one of SEQ ID NOs.: 1, 6, 9, 14, 16, 22, 25, 70, 77, 85, 95, 103, 111, 119, 124 and 131. The percent identity values are, preferably, calculated over the entire nucleic acid sequence region. A series of programs based on a variety of algorithms is available to the skilled worker for comparing different sequences. In this context, the algorithms of Needleman

15 and Wunsch or Smith and Waterman give particularly reliable results. To carry out the sequence alignments, the program PileUp (J. Mol. Evolution., 25, 351-360, 1987, Higgins 1989, CABIOS, 5: 151-153) or the programs Gap and BestFit (Needleman 1970 J. Mol. Biol. 48; 443-453 and Smith 1981, Adv. Appl. Math. 2; 482-489), which are part of the GCG software packet (Genetics Computer Group, 575 Science Drive, Madison,

20 Wisconsin, USA 53711 version 1991), are to be used. The sequence identity values recited above in percent (%) are to be determined, preferably, using the program GAP over the entire sequence region with the following settings: Gap Weight: 50, Length Weight: 3, Average Match: 10.000 and Average Mismatch: 0.000, which, unless otherwise specified, shall always be used as standard settings for sequence alignments.

25

Moreover, expression control sequences which allow for seed specific expression can not only be found upstream of the aforementioned open reading frames having a nucleic acid sequence as shown in any one of SEQ ID NOs. 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125. Rather, expression control sequences which allow for seed

30 specific expression can also be found upstream of orthologous, paralogous or homologous genes (i.e. open reading frames). Thus, also preferably, an variant expression control sequence comprised by a polynucleotide of the present invention has a nucleic acid sequence which hybridizes to a nucleic acid sequences located upstream of an open reading frame sequence being at least 70%, more preferably, at least 80%, at

35 least 90%, at least 91%, at least 92%, at least 93%, at least 94% at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to a sequence as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125. The said variant open reading shall encode a polypeptide having the biological activity of the corresponding polypeptide being encoded by the open reading frame shown in any

40 one of SEQ ID NOs.: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125. In this

context it should be mentioned that the open reading frame shown in SEQ ID NO: 5 encodes a polypeptide showing similarity to gibberlin responsive proteins, the open reading frame shown in SEQ ID NO: 13 encodes a polypeptide belonging to the pectinesterase family, the open reading frame shown in SEQ ID NO: 20 encodes “sinapoyl choline transferase 1” (SCT1), and the open reading frames shown in SEQ ID NO: 21 encodes “sinapoyl choline transferase 2” (SCT2). These biological activities can be determined by those skilled in the art without further ado. The open reading frames shown in SEQ ID NO: 27 and 28 encode polypeptides showing homology to seed proteins with yet unknown functions.

10

Table 1: Protein function of genes identified to be seed specifically expressed

SEQ ID	Protein function
5	putative gibberelin responsive protein
13	putative pectinesterase
20	Sinapoyl choline transferase
21	Sinapoyl choline transferase
27	Seed protein
28	Seed protein
72	Seed protein
80	CRU4 subunit of Cruciferin (seed storage protein)
88	Myrosinase
98	Seed protein
106	serine proteinase inhibitor
114	Transcription factor involved in embryonic development
121	Glutathione S-transferase
133	Seed protein

Also preferably, a variant expression control sequence comprised by a polynucleotide of the present invention is (i) obtainable by 5' genome walking or TAIL PCR from an open reading frame sequence as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125 or (ii) obtainable by 5' genome walking or TAIL PCR from a open reading frame sequence being at least 80% identical to an open reading frame as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 and 125. Variant expression control sequences are obtainable without

20

further ado by the genome walking technology or by thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) which can be carried out as described in the accompanying Examples by using, e.g., commercially available kits.

- 5 Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 1, preferably, comprise at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140 or all of the sequence motifs recited in Table 1. Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 6, preferably, com-
 10 prise at least 80, at least 90, at least 100, at least 110 or all of the sequence motifs recited in Table 2. Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 9, preferably, comprise at least 40, at least 50, at least 60 or all of the sequence motifs recited in Table 3. Variant expression control sequences referred to in this specification for the expression control
 15 sequence shown in SEQ ID NO: 14, preferably, comprise at least 50, at least 60, at least 70, at least 80, at least 90 or all of the sequence motifs recited in Table 4. Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 16, preferably, comprise at least 50, at least 60, at least 70, at least 80, at least 90 or all of the sequence motifs recited in Table 5. Variant
 20 expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 22, preferably, comprise at least 80, at least 90, at least 100, at least 110, at least 120, at least 130 or all of the sequence motifs recited in Table 6. Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 25, preferably, comprise at least
 25 80, at least 100, at least 120, at least 130, at least 140, at least 150 or all of the sequence motifs recited in Table 7.

Variant expression control sequences referred to in this specification also, preferably, comprise at least the cis-regulatory elements referred to in Table 8, below. Even more
 30 preferably, the variant regulatory expression control sequences comprise said elements with the same frequency and distribution as referred to in Table 9 for the individual regulatory sequences.

The term "seed specific" as used herein means that a nucleic acid of interest being
 35 operatively linked to the expression control sequence referred to herein will be predominantly expressed in seeds when present in a plant. A predominant expression as meant herein is characterized by a statistically significantly higher amount of detectable transcription in the seeds with respect to other plant tissues. A statistically significant higher amount of transcription is, preferably, an amount being at least two-fold, three-
 40 fold, four-fold, five-fold, ten-fold, hundred-fold, five hundred-fold or thousand-fold the

amount found in at least one of the other tissues with detectable transcription. Alternatively, it is an expression in seeds whereby the amount of transcription in non-seed tissues is less than 1%, 2%, 3%, 4% or most preferably 5% of the overall (whole plant) amount of expression. The amount of transcription directly correlates to the amount of transcripts (i.e. RNA) or polypeptides encoded by the transcripts present in a cell or tissue. Suitable techniques for measuring transcription either based on RNA or polypeptides are well known in the art. Seed specific alternatively and, preferably in addition to the above, means that the expression is restricted or almost restricted to seeds, i.e. there is essentially no detectable transcription in other tissues. Almost restricted as meant herein means that unspecific expression is detectable in less than ten, less than five, less than four, less than three, less than two or one other tissue(s). Seed specific expression as used herein includes expression in seed cells or their precursors, such as cells of the endosperm and of the developing embryo.

15 An expression control sequences can be tested for seed specific expression by determining the expression pattern of a nucleic acid of interest, e.g., a nucleic acid encoding a reporter protein, such as GFP, in a transgenic plant. Transgenic plants can be generated by techniques well known to the person skilled in the art and as discussed elsewhere in this specification. The aforementioned amounts or expression pattern are, preferably, determined by Northern Blot or in situ hybridization techniques as described in WO 02/102970 in *Brassica napus* plants, more preferably, at 20, 25, 30, 35 or 40 days after flowering. Preferred expression pattern for the expression control sequences according to the present invention are shown in the Figure or described in the accompanying Examples, below.

25 The term "nucleic acid of interest" refers to a nucleic acid which shall be expressed under the control of the expression control sequence referred to herein. Preferably, a nucleic acid of interest encodes a polypeptide the presence of which is desired in a cell or non-human organism as referred to herein and, in particular, in a plant seed. Such a polypeptide may be an enzyme which is required for the synthesis of seed storage compounds or may be a seed storage protein. It is to be understood that if the nucleic acid of interest encodes a polypeptide, transcription of the nucleic acid in RNA and translation of the transcribed RNA into the polypeptide may be required. A nucleic acid of interest, also preferably, includes biologically active RNA molecules and, more preferably, antisense RNAs, ribozymes, micro RNAs or siRNAs. Said biologically active RNA molecules can be used to modify the amount of a target polypeptide present in a cell or non-human organism. For example, an undesired enzymatic activity in a seed can be reduced due to the seed specific expression of an antisense RNAs, ribozymes, micro RNAs or siRNAs. The underlying biological principles of action of the aforementioned biologically active RNA molecules are well known in the art. Moreover, the per-

son skilled in the art is well aware of how to obtain nucleic acids which encode such biologically active RNA molecules. It is to be understood that the biologically active RNA molecules may be directly obtained by transcription of the nucleic acid of interest, i.e. without translation into a polypeptide. It is to be understood that the expression control sequence may also govern the expression of more than one nucleic acid of interest, i.e. at least one, at least two, at least three, at least four, at least five etc. nucleic acids of interest.

The term "operatively linked" as used herein means that the expression control sequence of the present invention and a nucleic acid of interest, are linked so that the expression can be governed by the said expression control sequence, i.e. the expression control sequence shall be functionally linked to said nucleic acid sequence to be expressed. Accordingly, the expression control sequence and, the nucleic acid sequence to be expressed may be physically linked to each other, e.g., by inserting the expression control sequence at the 5' end of the nucleic acid sequence to be expressed. Alternatively, the expression control sequence and the nucleic acid to be expressed may be merely in physical proximity so that the expression control sequence is capable of governing the expression of at least one nucleic acid sequence of interest. The expression control sequence and the nucleic acid to be expressed are, preferably, separated by not more than 500 bp, 300 bp, 100 bp, 80 bp, 60 bp, 40 bp, 20 bp, 10 bp or 5 bp.

Advantageously, it has been found in the studies underlying the present invention that seed specific expression of a nucleic acid of interest can be achieved by expressing said nucleic acid of interest under the control of an expression control sequence from *Brassica napus* or a variant expression control sequence as specified above. The expression control sequences provided by the present invention allow for a reliable and highly specific expression of nucleic acids of interest. Thanks to the present invention, it is possible to (i) specifically manipulate biochemical processes in seeds, e.g., by expressing heterologous enzymes or biologically active RNAs, or (ii) to produce heterologous proteins in seeds. In principle, the present invention contemplates the use of the polynucleotide, the vector, the host cell or the non-human transgenic organism for the expression of a nucleic acid of interest. Preferably, the envisaged expression is seed specific. More preferably, the nucleic acid of interest to be used in the various embodiments of the present invention encodes a seed storage protein or is involved in the modulation of seed storage compounds.

As used herein, seed storage compounds include fatty acids and triacylglycerides which have a multiplicity of applications in the food industry, in animal nutrition, in cos-

metics and the pharmacological sector. Depending on whether they are free saturated or unsaturated fatty acids or else triacylglycerides with an elevated content of saturated or unsaturated fatty acids, they are suitable for various different applications. More preferably, the polynucleotide of the present invention comprising the expression control sequence referred to above is applied for the manufacture of **polyunsaturated fatty acids** (PUFAs). For the manufacture of PUFAs in seeds, the activity of enzymes involved in their synthesis, in particular, elongases and desaturases, needs to be modulated. This will be achieved by seed specific expression of the nucleic acids of interest encoding the aforementioned enzymes or by seed specific expression of antisense, ribozyme, RNAi molecules which downregulate the activity of the enzymes by interfering with their protein synthesis. PUFAs are seed storage compounds which can be isolated by a subsequently applied purification process using the aforementioned seeds.

Particularly preferred PUFAs in accordance with the present invention are polyunsaturated long-chain ω -3-fatty acids such as eicosapentaenoic acid (= EPA, C20:5 $\Delta^{5,8,11,14,17}$), ω -3 eicosetraenic acid (= ETA, C20:4 $\Delta^{8,11,14,17}$), arachidonic acid (= ARA C20:4 $\Delta^{5,8,11,14}$) or docosahexaenoic acid (= DHA, C22:6 $\Delta^{4,7,10,13,16,19}$). They are important components of human nutrition owing to their various roles in health aspects, including the development of the child brain, the functionality of the eyes, the synthesis of hormones and other signal substances, and the prevention of cardiovascular disorders, cancer and diabetes (Poulos, A Lipids 30:1-14, 14 $\Delta^{8,11,14,17}$ 995; Horrocks, LA and Yeo YK Pharmacol Res 40:211-225, 1999). There is, therefore, a need for the production of polyunsaturated long-chain fatty acids.

25

Particular preferred enzymes involved in the synthesis of PUFAs are disclosed in WO 91/13972 (Δ 9-desaturase), WO 93/11245 (Δ 15-desaturase), WO 94/11516 (Δ 12-desaturase), EP A 0 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP A 0 794 250, Stukey et al., J. Biol. Chem., 265, 1990: 20144-20149, Wada et al., Nature 347, 1990: 200-203 or Huang et al., Lipids 34, 1999: 649-659. Δ 6-Desaturases are described in WO 93/06712, US 5,614,393, US 5,614,393, WO 96/21022, WO 00/21557 and WO 99/27111, and also the application for the production in transgenic organisms is described in WO 98/46763, WO 98/46764 and WO 98/46765. Here, the expression of various desaturases is also described and claimed in WO 99/64616 or WO 98/46776, as is the formation of polyunsaturated fatty acids. As regards the expression efficacy of desaturases and its effect on the formation of polyunsaturated fatty acids, it must be noted that the expression of a single desaturase as

35

described to date has only resulted in low contents of unsaturated fatty acids/lipids such as, for example, γ -linolenic acid and stearidonic acid. Furthermore, mixtures of ω -3- and ω -6-fatty acids are usually obtained.

5

Furthermore, the present invention relates to a polynucleotide comprising an expression termination sequence which allows for termination of transcription of a nucleic acid of interest being operatively linked thereto, said expression termination sequence being selected from the group consisting of:

- 10 (a) a expression termination sequence having a nucleic acid sequence as shown in any one of SEQ ID NOs: 2, 7, 10, 15, 17, 23, 71, 78, 86, 96, 104, 112, or 125;
- (b) a expression termination sequence having a nucleic acid sequence which is at least 80% identical to a nucleic acid sequence as shown in any one of
15 SEQ ID NOs: which hybridizes under stringent conditions to a nucleic acid sequence as shown in any one of SEQ ID NOs: 2, 7, 10, 15, 17, 23, 71, 78, 86, 96, 104, 112, or 125;
- (c) a expression termination sequence having a nucleic acid sequence which hybridizes under stringent conditions to a nucleic acid sequence as shown
20 in any one of SEQ ID NOs: 2, 7, 10, 15, 17, 23, 71, 78, 86, 96, 104, 112, or 125;
- (d) a expression termination sequence having a nucleic acid sequence which hybridizes to a nucleic acid sequences located downstream of an open reading frame sequence shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125;
25
- (e) a expression termination sequence having a nucleic acid sequence which hybridizes to a nucleic acid sequences located downstream of an open reading frame sequence being at least 80% identical to an open reading frame sequence as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125;
30
- (f) a expression termination sequence obtainable by 3' genome walking or TAIL PCR on genomic DNA from the last exon of an open reading frame sequence as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125; and
- 35 (g) a expression termination sequence obtainable by 3' genome walking or TAIL PCR on genomic DNA from the last exon of an open reading frame sequence being at least 80% identical to an open reading frame as shown in any one of SEQ ID NOs: 5, 13, 20, 21, 27, 28, 71, 78, 86, 96, 104, 112 or 125.

40

The term "expression termination sequence" as used herein refers to a nucleic acid which is capable of governing the termination of the process of RNA transcription of a nucleic acid operatively linked thereto, e.g. a nucleic acid of interest referred to elsewhere in this specification in detail. A termination sequence as referred to in accordance with the present invention, preferably, contains a polyadenylation signal and furthermore mediates dissociation of RNA polymerases, preferably, RNA polymerase I, II or III, more preferably, RNA polymerase II or III, and most preferably, RNA polymerase II from the transcribed DNA. Thereby the elongation of a RNA transcript, transcribed from a nucleic acid operatively linked to the termination sequence will be terminated and the RNA will be released. In order to govern termination of transcription of a nucleic acid, the expression control sequence may be located immediately adjacent to the nucleic acid whos expression is to be terminated, i.e. physically linked to the said nucleic acid at its 3' end. Alternatively, it may be located in physical proximity. In the latter case, however, the sequence must be located so as to allow functional interaction with the nucleic acid whos transcription is to be terminated. A termination sequence referred to herein, preferably, comprises between 50 and 2,000 nucleotides in length. More preferably, it comprises between 100 and 800 nucleotides and, more preferably, at least 100 nucleotides. Preferred expression termination sequences are those comprised by the polynucleotide referred to above.

20

Furthermore, the definitions and explanations of the terms made above apply mutatis muandis except as specified herein below.

For termination sequences, the term "operatively linked" means that the termination sequence of the present invention and a nucleic acid of interest, are linked so that the termination of transcription of the mRNA can be governed by said termination sequence, i.e. the termination sequence shall be functionally linked to said nucleic acid sequence whos transcription is to be terminated. Accordingly, the expression control sequence, the nucleic acid sequence to be expressed and the termination sequence may be physically linked to each other, e.g., by inserting the expression control sequence at the 5' end of the nucleic acid sequence to be expressed and/or inserting the termination sequence at the 3' end of the nucleic acid sequence whos transcription is to be terminated. Alternatively, the expression control sequence and the nucleic acid to be expressed may be merely in physical proximity so that the expression control sequence is capable of governing the expression of at least one nucleic acid sequence of interest. The termination sequence and the nucleic acid whos transcription is to be terminated are, preferably, separated by not more than 50 bp, 40 bp, 20 bp, 10 bp or 5 bp.

35

Advantageously, the polynucleotide of the present invention comprising a expression termination sequence can be also applied for efficient expression control in plants and, in particular, plant seeds. Specifically, the expression termination sequence allows for accurate termination of transcription of the DNA into RNA after the nucleic acid sequence of interest has been transcribed. Thus, the transcription of undesired nucleic acid sequences is avoided.

The present invention also relates to a vector comprising the polynucleotide of the present invention.

10

The term "vector", preferably, encompasses phage, plasmid, viral or retroviral vectors as well as artificial chromosomes, such as bacterial or yeast artificial chromosomes. Moreover, the term also relates to targeting constructs which allow for random or site-directed integration of the targeting construct into genomic DNA. Such target constructs, preferably, comprise DNA of sufficient length for either homologous or heterologous recombination as described in detail below. The vector encompassing the polynucleotides of the present invention, preferably, further comprises selectable markers for propagation and/or selection in a host. The vector may be incorporated into a host cell by various techniques well known in the art. If introduced into a host cell, the vector may reside in the cytoplasm or may be incorporated into the genome. In the latter case, it is to be understood that the vector may further comprise nucleic acid sequences which allow for homologous recombination or heterologous insertion. Vectors can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. The terms "transformation" and "transfection", conjugation and transduction, as used in the present context, are intended to comprise a multiplicity of prior-art processes for introducing foreign nucleic acid (for example DNA) into a host cell, including calcium phosphate, rubidium chloride or calcium chloride coprecipitation, DEAE-dextran-mediated transfection, lipofection, natural competence, carbon-based clusters, chemically mediated transfer, electroporation or particle bombardment (e.g., "gene-gun"). Suitable methods for the transformation or transfection of host cells, including plant cells, can be found in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) and other laboratory manuals, such as Methods in Molecular Biology, 1995, Vol. 44, Agrobacterium protocols, Ed.: Gartland and Davey, Humana Press, Totowa, New Jersey. Alternatively, a plasmid vector may be introduced by heat shock or electroporation techniques. Should the vector be a virus, it may be packaged in vitro using an appropriate packaging cell line prior to application to host cells. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host/cells.

15

20

25

30

35

40

Preferably, the vector referred to herein is suitable as a cloning vector, i.e. replicable in microbial systems. Such vectors ensure efficient cloning in bacteria and, preferably, yeasts or fungi and make possible the stable transformation of plants. Those which must be mentioned are, in particular, various binary and co-integrated vector systems which are suitable for the T-DNA-mediated transformation. Such vector systems are, as a rule, characterized in that they contain at least the vir genes, which are required for the Agrobacterium-mediated transformation, and the sequences which delimit the T-DNA (T-DNA border). These vector systems, preferably, also comprise further cis-regulatory regions such as promoters and terminators and/or selection markers with which suitable transformed host cells or organisms can be identified. While co-integrated vector systems have vir genes and T-DNA sequences arranged on the same vector, binary systems are based on at least two vectors, one of which bears vir genes, but no T-DNA, while a second one bears T-DNA, but no vir gene. As a consequence, the last-mentioned vectors are relatively small, easy to manipulate and can be replicated both in *E. coli* and in *Agrobacterium*. These binary vectors include vectors from the pBIB-HYG, pPZP, pBecks, pGreen series. Preferably used in accordance with the invention are Bin19, pBI101, pBinAR, pGPTV, pSUN and pCAMBIA. An overview of binary vectors and their use can be found in Hellens et al, Trends in Plant Science (2000) 5, 446–451. Furthermore, by using appropriate cloning vectors, the polynucleotide of the invention can be introduced into host cells or organisms such as plants or animals and, thus, be used in the transformation of plants, such as those which are published, and cited, in: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), chapter 6/7, pp. 71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, vol. 1, Engineering and Utilization, Ed.: Kung and R. Wu, Academic Press, 1993, 15-38; B. Jené et al., Techniques for Gene Transfer, in: Transgenic Plants, vol. 1, Engineering and Utilization, Ed.: Kung and R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225.

More preferably, the vector of the present invention is an expression vector. In such an expression vector, the polynucleotide comprises an expression cassette as specified above allowing for expression in eukaryotic cells or isolated fractions thereof. An expression vector may, in addition to the polynucleotide of the invention, also comprise further regulatory elements including transcriptional as well as translational enhancers. Preferably, the expression vector is also a gene transfer or targeting vector. Expression vectors derived from viruses such as retroviruses, vaccinia virus, adeno-associated virus, herpes viruses, or bovine papilloma virus, may be used for delivery of the polynucleotides or vector of the invention into targeted cell population. Methods which are well known to those skilled in the art can be used to construct recombinant viral

vectors; see, for example, the techniques described in Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. and Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1994).

5

Suitable expression vector backbones are, preferably, derived from expression vectors known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3 (Invitrogen) or pSPORT1 (GIBCO BRL). Further examples of typical fusion expression vectors are pGEX (Pharmacia Biotech Inc; Smith, D.B., and Johnson, K.S. (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ), where glutathione S-transferase (GST), maltose E-binding protein and protein A, respectively, are fused with the nucleic acid of interest encoding a protein to be expressed. The target gene expression of the pTrc vector is based on the transcription from a hybrid trp-lac fusion promoter by host RNA polymerase. The target gene expression from the pET 11d vector is based on the transcription of a T7-gn10-lac fusion promoter, which is mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is provided by the host strains BL21 (DE3) or HMS174 (DE3) from a resident λ -prophage which harbors a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter. Examples of vectors for expression in the yeast *S. cerevisiae* comprise pYepSecl (Baldari et al. (1987) Embo J. 6:229-234), pMFa (Kurjan and Herskowitz (1982) Cell 30:933-943), pJRY88 (Schultz et al. (1987) Gene 54:113-123) and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and processes for the construction of vectors which are suitable for use in other fungi, such as the filamentous fungi, comprise those which are described in detail in: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of fungi, J.F. Peberdy et al., Ed., pp. 1-28, Cambridge University Press: Cambridge, or in: More Gene Manipulations in Fungi (J.W. Bennett & L.L. Lasure, Ed., pp. 396-428: Academic Press: San Diego). Further suitable yeast vectors are, for example, pAG-1, YEp6, YEp13 or pEMBLYe23. As an alternative, the polynucleotides of the present invention can be also expressed in insect cells using baculovirus expression vectors. Baculovirus vectors which are available for the expression of proteins in cultured insect cells (for example Sf9 cells) comprise the pAc series (Smith et al. (1983) Mol. Cell Biol. 3:2156-2165) and the pVL series (Lucklow and Summers (1989) Virology 170:31-39).

35

The polynucleotides of the present invention can be used for expression of a nucleic acid of interest in single-cell plant cells (such as algae), see Falciatore et al., 1999, Marine Biotechnology 1 (3):239-251 and the references cited therein, and plant cells from higher plants (for example Spermatophytes, such as arable crops) by using plant expression vectors. Examples of plant expression vectors comprise those which are

40

described in detail in: Becker, D., Kemper, E., Schell, J., and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20:1195-1197; and Bevan, M.W. (1984) "Binary Agrobacterium vectors for plant transformation", *Nucl. Acids Res.* 12:8711-8721; Vectors for Gene Transfer in Higher Plants; in: *Transgenic Plants, Vol. 1, Engineering and Utilization*, Ed.: Kung and R. Wu, Academic Press, 1993, p. 15-38. A plant expression cassette, preferably, comprises regulatory sequences which are capable of controlling the gene expression in plant cells and which are functionally linked so that each sequence can fulfill its function, such as transcriptional termination, for example polyadenylation signals. Preferred polyadenylation signals are those which are derived from *Agrobacterium tumefaciens* T-DNA, such as the gene 3 of the Ti plasmid pTiACH5, which is known as octopine synthase (Gielen et al., *EMBO J.* 3 (1984) 835 et seq.) or functional equivalents of these, but all other terminators which are functionally active in plants are also suitable. Since plant gene expression is very often not limited to transcriptional levels, a plant expression cassette preferably comprises other functionally linked sequences such as translation enhancers, for example the overdrive sequence, which comprises the 5'-untranslated tobacco mosaic virus leader sequence, which increases the protein/RNA ratio (Gallie et al., 1987, *Nucl. Acids Research* 15:8693-8711). Other preferred sequences for the use in functional linkage in plant gene expression cassettes are targeting sequences which are required for targeting the gene product into its relevant cell compartment (for a review, see Kermode, *Crit. Rev. Plant Sci.* 15, 4 (1996) 285-423 and references cited therein), for example into the vacuole, the nucleus, all types of plastids, such as amyloplasts, chloroplasts, chromoplasts, the extracellular space, the mitochondria, the endoplasmic reticulum, oil bodies, peroxisomes and other compartments of plant cells.

The abovementioned vectors are only a small overview of vectors to be used in accordance with the present invention. Further vectors are known to the skilled worker and are described, for example, in: *Cloning Vectors* (Ed., Pouwels, P.H., et al., Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018). For further suitable expression systems for prokaryotic and eukaryotic cells see the chapters 16 and 17 of Sambrook, J., Fritsch, E.F., and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

35

The present invention also contemplates a host cell comprising the polynucleotide or the vector of the present invention.

Host cells are primary cells or cell lines derived from multicellular organisms such as plants or animals. Furthermore, host cells encompass prokaryotic or eukaryotic single cell organisms (also referred to as micro-organisms). Primary cells or cell lines to be used as host cells in accordance with the present invention may be derived from the multicellular organisms referred to below. Host cells which can be exploited are furthermore mentioned in: Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Specific expression strains which can be used, for example those with a lower protease activity, are described in: Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128. These include plant cells and certain tissues, organs and parts of plants in all their phenotypic forms such as anthers, fibers, root hairs, stalks, embryos, calli, cotyledons, petioles, harvested material, plant tissue, reproductive tissue and cell cultures which are derived from the actual transgenic plant and/or can be used for bringing about the transgenic plant. Preferably, the host cells may be obtained from plants. More preferably, oil crops are envisaged which comprise large amounts of lipid compounds, such as oilseed rape, evening primrose, hemp, thistle, peanut, canola, linseed, soybean, safflower, sunflower, borage, or plants such as maize, wheat, rye, oats, triticale, rice, barley, cotton, cassava, pepper, Tagetes, Solanaceae plants such as potato, tobacco, eggplant and tomato, Vicia species, pea, alfalfa, bushy plants (coffee, cacao, tea), Salix species, trees (oil palm, coconut) and perennial grasses and fodder crops. Especially preferred plants according to the invention are oil crops such as soybean, peanut, oilseed rape, canola, linseed, hemp, evening primrose, sunflower, safflower, trees (oil palm, coconut). Suitable methods for obtaining host cells from the multicellular organisms referred to below as well as conditions for culturing these cells are well known in the art.

The micro-organisms are, preferably, bacteria or fungi including yeasts. Preferred fungi to be used in accordance with the present invention are selected from the group of the families Chaetomiaceae, Choanephoraceae, Cryptococcaceae, Cunninghamellaceae, Dematiaceae, Moniliaceae, Mortierellaceae, Mucoraceae, Pythiaceae, Saccharomycetaceae, Saprolegniaceae, Schizosacharomycetaceae, Sodariaceae or Tuberculariaceae. Further preferred micro-organisms are selected from the group: Choanephoraceae such as the genera Blakeslea, Choanephora, for example the genera and species *Blakeslea trispora*, *Choanephora cucurbitarum*, *Choanephora infundibulifera* var. *cucurbitarum*, Mortierellaceae, such as the genus Mortierella, for example the genera and species *Mortierella isabellina*, *Mortierella polycephala*, *Mortierella ramaniana*, *Mortierella vinacea*, *Mortierella zonata*, Pythiaceae such as the genera Phytium, Phytophthora for example the genera and species *Pythium debaryanum*, *Pythium intermedium*, *Pythium irregulare*, *Pythium megalacanthum*, *Pythium paroecandrum*, *Pythium sylvaticum*, *Pythium ultimum*, *Phytophthora cactorum*, *Phytophthora cinnamomi*,

Phytophthora citricola, *Phytophthora citrophthora*, *Phytophthora cryptogea*, *Phytophthora drechsleri*, *Phytophthora erythroseptica*, *Phytophthora lateralis*, *Phytophthora megasperma*, *Phytophthora nicotianae*, *Phytophthora nicotianae* var. *parasitica*, *Phytophthora palmivora*, *Phytophthora parasitica*, *Phytophthora syringae*, Saccharomycetaceae such as the genera *Hansenula*, *Pichia*, *Saccharomyces*, *Saccharomycodes*, *Yarrowia* for example the genera and species *Hansenula anomala*, *Hansenula californica*, *Hansenula canadensis*, *Hansenula capsulata*, *Hansenula ciferrii*, *Hansenula glucozyma*, *Hansenula henricii*, *Hansenula holstii*, *Hansenula minuta*, *Hansenula nonfermentans*, *Hansenula philodendri*, *Hansenula polymorpha*, *Hansenula saturnus*, *Hansenula subpelliculosa*, *Hansenula wickerhamii*, *Hansenula wingei*, *Pichia alcoholophila*, *Pichia angusta*, *Pichia anomala*, *Pichia bisporea*, *Pichia burtonii*, *Pichia canadensis*, *Pichia capsulata*, *Pichia carsonii*, *Pichia cellobiosa*, *Pichia ciferrii*, *Pichia farinosa*, *Pichia fermentans*, *Pichia finlandica*, *Pichia glucozyma*, *Pichia guilliermondii*, *Pichia haplophila*, *Pichia henricii*, *Pichia holstii*, *Pichia jadinii*, *Pichia lindnerii*, *Pichia membranaefaciens*, *Pichia methanolica*, *Pichia minuta* var. *minuta*, *Pichia minuta* var. *nonfermentans*, *Pichia norvegensis*, *Pichia ohmeri*, *Pichia pastoris*, *Pichia philodendri*, *Pichia pini*, *Pichia polymorpha*, *Pichia quercuum*, *Pichia rhodanensis*, *Pichia sargentensis*, *Pichia stipitis*, *Pichia strasburgensis*, *Pichia subpelliculosa*, *Pichia toletana*, *Pichia trehalophila*, *Pichia vini*, *Pichia xylosa*, *Saccharomyces aceti*, *Saccharomyces baillii*, *Saccharomyces bayanus*, *Saccharomyces bisporus*, *Saccharomyces capensis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces cerevisiae* var. *ellipsoideus*, *Saccharomyces chevalieri*, *Saccharomyces delbrueckii*, *Saccharomyces diastaticus*, *Saccharomyces drosophilarum*, *Saccharomyces elegans*, *Saccharomyces ellipsoideus*, *Saccharomyces fermentati*, *Saccharomyces florentinus*, *Saccharomyces fragilis*, *Saccharomyces heterogenicus*, *Saccharomyces hienipiensis*, *Saccharomyces inusitatus*, *Saccharomyces italicus*, *Saccharomyces kluyveri*, *Saccharomyces krusei*, *Saccharomyces lactis*, *Saccharomyces marxianus*, *Saccharomyces microellipsoides*, *Saccharomyces montanus*, *Saccharomyces norbensis*, *Saccharomyces oleaceus*, *Saccharomyces paradoxus*, *Saccharomyces pastorianus*, *Saccharomyces pretoriensis*, *Saccharomyces rosei*, *Saccharomyces rouxii*, *Saccharomyces uvarum*, *Saccharomycodes ludwigii*, *Yarrowia lipolytica*, Schizosaccharomycetaceae such as the genera *Schizosaccharomyces* e.g. the species *Schizosaccharomyces japonicus* var. *japonicus*, *Schizosaccharomyces japonicus* var. *versatilis*, *Schizosaccharomyces malidevorans*, *Schizosaccharomyces octosporus*, *Schizosaccharomyces pombe* var. *malidevorans*, *Schizosaccharomyces pombe* var. *pombe*, Thraustochytriaceae such as the genera *Althornia*, *Aplanochytrium*, *Japonochytrium*, *Schizochytrium*, *Thraustochytrium* e.g. the species *Schizochytrium aggregatum*, *Schizochytrium limacinum*, *Schizochytrium mangrovei*, *Schizochytrium minutum*, *Schizochytrium octosporum*, *Thraustochytrium aggregatum*, *Thraustochytrium amoeboides*, *Thraustochytrium antacticum*, *Thraustochytrium arudimentale*, *Thraustochy-*

trium aureum, *Thraustochytrium benthicola*, *Thraustochytrium globosum*, *Thraustochytrium indicum*, *Thraustochytrium kerguelense*, *Thraustochytrium kinnei*, *Thraustochytrium motivum*, *Thraustochytrium multirudimentale*, *Thraustochytrium pachydermum*, *Thraustochytrium proliferum*, *Thraustochytrium roseum*, *Thraustochytrium rossii*,
5 *Thraustochytrium striatum* or *Thraustochytrium visurgense*. Further preferred micro-organisms are bacteria selected from the group of the families Bacillaceae, Enterobacteriaceae or Rhizobiaceae. Examples of such micro-organisms may be selected from the group: Bacillaceae such as the genera *Bacillus* for example the genera and species
10 *Bacillus acidocaldarius*, *Bacillus acidoterrestris*, *Bacillus alcalophilus*, *Bacillus amylo-liquefaciens*, *Bacillus amylolyticus*, *Bacillus brevis*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus sphaericus* subsp. *fusiformis*, *Bacillus galactophilus*, *Bacillus globisporus*, *Bacillus globisporus* subsp. *marinus*, *Bacillus halophilus*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus polymyxa*, *Bacillus psychrosaccharolyticus*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus subtilis* subsp. *spizizenii*, *Bacillus subtilis* subsp. *subtilis* or *Bacillus thuringiensis*;
15 Enterobacteriaceae such as the genera *Citrobacter*, *Edwardsiella*, *Enterobacter*, *Erwinia*, *Escherichia*, *Klebsiella*, *Salmonella* or *Serratia* for example the genera and species *Citrobacter amalonaticus*, *Citrobacter diversus*, *Citrobacter freundii*, *Citrobacter genomospecies*, *Citrobacter gillenii*, *Citrobacter intermedium*, *Citrobacter koseri*, *Citrobacter murlinae*, *Citrobacter sp.*, *Edwardsiella hoshinae*, *Edwardsiella ictaluri*, *Edwardsiella tarda*, *Erwinia alni*, *Erwinia amylovora*, *Erwinia ananatis*, *Erwinia aphidicola*, *Erwinia billingiae*, *Erwinia cacticida*, *Erwinia cancerogena*, *Erwinia carnegieana*, *Erwinia carotovora* subsp. *atroseptica*, *Erwinia carotovora* subsp. *betavasculorum*, *Erwinia carotovora* subsp. *odorifera*, *Erwinia carotovora* subsp. *wasabiae*, *Erwinia chrysanthemi*, *Erwinia cypripedii*, *Erwinia dissolvens*, *Erwinia herbicola*, *Erwinia mallotivora*, *Erwinia milletiae*,
25 *Erwinia nigrifluens*, *Erwinia nimipressuralis*, *Erwinia persicina*, *Erwinia psidii*, *Erwinia pyrifoliae*, *Erwinia quercina*, *Erwinia rhapontici*, *Erwinia rubrifaciens*, *Erwinia salicis*, *Erwinia stewartii*, *Erwinia tracheiphila*, *Erwinia uredovora*, *Escherichia adecarboxylata*, *Escherichia anindolica*, *Escherichia aurescens*, *Escherichia blattae*, *Escherichia coli*,
30 *Escherichia coli* var. *communior*, *Escherichia coli-mutabile*, *Escherichia fergusonii*, *Escherichia hermannii*, *Escherichia sp.*, *Escherichia vulneris*, *Klebsiella aerogenes*, *Klebsiella edwardsii* subsp. *atlantae*, *Klebsiella ornithinolytica*, *Klebsiella oxytoca*, *Klebsiella planticola*, *Klebsiella pneumoniae*, *Klebsiella pneumoniae* subsp. *pneumoniae*, *Klebsiella sp.*, *Klebsiella terrigena*, *Klebsiella trevisanii*, *Salmonella abony*, *Salmonella arizonae*, *Salmonella bongori*, *Salmonella choleraesuis* subsp. *arizonae*, *Salmonella choleraesuis* subsp. *bongori*, *Salmonella choleraesuis* subsp. *choleraesuis*, *Salmonella choleraesuis* subsp. *diarizonae*, *Salmonella choleraesuis* subsp. *houtenae*, *Salmonella choleraesuis* subsp. *indica*, *Salmonella choleraesuis* subsp. *salamae*, *Salmonella daressalaam*, *Salmonella enterica* subsp. *houtenae*, *Salmonella enterica* subsp. *salamae*,
40 *Salmonella enteritidis*, *Salmonella gallinarum*, *Salmonella heidelberg*, *Salmo-*

nella panama, Salmonella senftenberg, Salmonella typhimurium, Serratia entomophila, Serratia ficaria, Serratia fonticola, Serratia grimesii, Serratia liquefaciens, Serratia marcescens, Serratia marcescens subsp. marcescens, Serratia marinorubra, Serratia odorifera, Serratia plymouthis, Serratia plymuthica, Serratia proteamaculans, Serratia proteamaculans subsp. quinovora, Serratia quinivorans or Serratia rubidaea; Rhizobiaceae such as the genera *Agrobacterium, Carbophilus, Chelatobacter, Ensifer, Rhizobium, Sinorhizobium* for example the genera and species *Agrobacterium atlanticum, Agrobacterium ferrugineum, Agrobacterium gelatinovorum, Agrobacterium larrymoorei, Agrobacterium meteori, Agrobacterium radiobacter, Agrobacterium rhizogenes, Agrobacterium rubi, Agrobacterium stellulatum, Agrobacterium tumefaciens, Agrobacterium vitis, Carbophilus carboxidus, Chelatobacter heintzii, Ensifer adhaerens, Ensifer arboris, Ensifer fredii, Ensifer kostiensis, Ensifer kummerowiae, Ensifer medicae, Ensifer meliloti, Ensifer saheli, Ensifer teranga, Ensifer xinjiangensis, Rhizobium ciceri, Rhizobium etli, Rhizobium fredii, Rhizobium galegae, Rhizobium gallicum, Rhizobium giardinii, Rhizobium hainanense, Rhizobium huakuii, Rhizobium huautlense, Rhizobium indigoferae, Rhizobium japonicum, Rhizobium leguminosarum, Rhizobium loessense, Rhizobium loti, Rhizobium lupini, Rhizobium mediterraneum, Rhizobium meliloti, Rhizobium mongolense, Rhizobium phaseoli, Rhizobium radiobacter, Rhizobium rhizogenes, Rhizobium rubi, Rhizobium sullae, Rhizobium tianshanense, Rhizobium trifolii, Rhizobium tropici, Rhizobium undicola, Rhizobium vitis, Sinorhizobium adhaerens, Sinorhizobium arboris, Sinorhizobium fredii, Sinorhizobium kostiense, Sinorhizobium kummerowiae, Sinorhizobium medicae, Sinorhizobium meliloti, Sinorhizobium morelense, Sinorhizobium saheli or Sinorhizobium xinjiangense.*

How to culture the aforementioned micro-organisms is well known to the person skilled in the art.

The present invention also relates to a non-human transgenic organism, preferably a plant or seed thereof, comprising the polynucleotide or the vector of the present invention.

The term "non-human transgenic organism", preferably, relates to a plant, a plant seed, a non-human animal or a multicellular micro-organism. The polynucleotide or vector may be present in the cytoplasm of the organism or may be incorporated into the genome either heterologous or by homologous recombination. Host cells, in particular those obtained from plants or animals, may be introduced into a developing embryo in order to obtain mosaic or chimeric organisms, i.e. non-human transgenic organisms comprising the host cells of the present invention. Suitable transgenic organisms are, preferably, all organisms which are suitable for the expression of recombinant genes.

Preferred plants to be used for making non-human transgenic organisms according to the present invention are all dicotyledonous or monocotyledonous plants, algae or mosses. Advantageous plants are selected from the group of the plant families

5 Adelotheceaceae, Anacardiaceae, Asteraceae, Apiaceae, Betulaceae, Boraginaceae, Brassicaceae, Bromeliaceae, Caricaceae, Cannabaceae, Convolvulaceae, Chenopodiaceae, Crypthecodiniaceae, Cucurbitaceae, Ditrichaceae, Elaeagnaceae, Ericaceae, Euphorbiaceae, Fabaceae, Geraniaceae, Gramineae, Juglandaceae, Lauraceae, Leguminosae, Linaceae, Prasinophyceae or vegetable plants or ornamentals such as

10 Tagetes. Examples which may be mentioned are the following plants selected from the group consisting of: Adelotheceaceae such as the genera *Physcomitrella*, such as the genus and species *Physcomitrella patens*, Anacardiaceae such as the genera *Pistacia*, *Mangifera*, *Anacardium*, for example the genus and species *Pistacia vera* [pistachio], *Mangifera indica* [mango] or *Anacardium occidentale* [cashew], Asteraceae, such as the

15 genera *Calendula*, *Carthamus*, *Centaurea*, *Cichorium*, *Cynara*, *Helianthus*, *Lactuca*, *Locusta*, *Tagetes*, *Valeriana*, for example the genus and species *Calendula officinalis* [common marigold], *Carthamus tinctorius* [safflower], *Centaurea cyanus* [cornflower], *Cichorium intybus* [chicory], *Cynara scolymus* [artichoke], *Helianthus annuus* [sunflower], *Lactuca sativa*, *Lactuca crispera*, *Lactuca esculenta*, *Lactuca scariola* L. ssp.

20 *sativa*, *Lactuca scariola* L. var. *integrata*, *Lactuca scariola* L. var. *integrifolia*, *Lactuca sativa* subsp. *romana*, *Locusta communis*, *Valeriana locusta* [salad vegetables], *Tagetes lucida*, *Tagetes erecta* or *Tagetes tenuifolia* [african or french marigold], Apiaceae, such as the genus *Daucus*, for example the genus and species *Daucus carota* [carrot], Betulaceae, such as the genus *Corylus*, for example the genera and species

25 *Corylus avellana* or *Corylus colurna* [hazelnut], Boraginaceae, such as the genus *Borago*, for example the genus and species *Borago officinalis* [borage], Brassicaceae, such as the genera *Brassica*, *Melanosinapis*, *Sinapis*, *Arabidopsis*, for example the genera and species *Brassica napus*, *Brassica rapa* ssp. [oilseed rape], *Sinapis arvensis*, *Brassica juncea*, *Brassica juncea* var. *juncea*, *Brassica juncea* var. *crispifolia*, *Brassica juncea* var. *foliosa*, *Brassica nigra*, *Brassica sinapioides*, *Melanosinapis communis* [mustard], *Brassica oleracea* [fodder beet] or *Arabidopsis thaliana*, Bromeliaceae, such as the genera *Anana*, *Bromelia* (pineapple), for example the genera and species

30 *Anana comosus*, *Ananas ananas* or *Bromelia comosa* [pineapple], Caricaceae, such as the genus *Carica*, such as the genus and species *Carica papaya* [pawpaw], Cannabaceae, such as the genus *Cannabis*, such as the genus and species *Cannabis sativa* [hemp], Convolvulaceae, such as the genera *Ipomoea*, *Convolvulus*, for example the genera and species *Ipomoea batatas*, *Ipomoea pandurata*, *Convolvulus batatas*, *Convolvulus tiliaceus*, *Ipomoea fastigiata*, *Ipomoea tiliacea*, *Ipomoea triloba* or *Convolvulus panduratus* [sweet potato, batate], Chenopodiaceae, such as the genus *Beta*, such as

35 the genera and species *Beta vulgaris*, *Beta vulgaris* var. *altissima*, *Beta vulgaris*

40

var. Vulgaris, *Beta maritima*, *Beta vulgaris var. perennis*, *Beta vulgaris var. conditiva* or *Beta vulgaris var. esculenta* [sugarbeet], Crypthecodiniaceae, such as the genus *Crypthecodinium*, for example the genus and species *Cryptecodinium cohnii*, Cucurbitaceae, such as the genus *Cucurbita*, for example the genera and species *Cucurbita maxima*, *Cucurbita mixta*, *Cucurbita pepo* or *Cucurbita moschata* [pumpkin/squash], Cymbellaceae such as the genera *Amphora*, *Cymbella*, *Okedenia*, *Phaeodactylum*, *Reimeria*, for example the genus and species *Phaeodactylum tricornutum*, Ditrichaceae such as the genera *Ditrichaceae*, *Astomiopsis*, *Ceratodon*, *Chrysoblastella*, *Ditrichum*, *Distichium*, *Eccremidium*, *Lophidion*, *Philibertiella*, *Pleuridium*, *Saelania*, *Trichodon*, *Skottsbergia*, for example the genera and species *Ceratodon antarcticus*, *Ceratodon columbiae*, *Ceratodon heterophyllus*, *Ceratodon purpureus*, *Ceratodon purpureus*, *Ceratodon purpureus ssp. convolutus*, *Ceratodon purpureus spp. stenocarpus*, *Ceratodon purpureus var. rotundifolius*, *Ceratodon ratodon*, *Ceratodon stenocarpus*, *Chrysoblastella chilensis*, *Ditrichum ambiguum*, *Ditrichum brevisetum*, *Ditrichum crispatisimum*, *Ditrichum difficile*, *Ditrichum falcifolium*, *Ditrichum flexicaule*, *Ditrichum giganteum*, *Ditrichum heteromallum*, *Ditrichum lineare*, *Ditrichum lineare*, *Ditrichum montanum*, *Ditrichum montanum*, *Ditrichum pallidum*, *Ditrichum punctulatum*, *Ditrichum pusillum*, *Ditrichum pusillum var. tortile*, *Ditrichum rhynchostegium*, *Ditrichum schimperi*, *Ditrichum tortile*, *Distichium capillaceum*, *Distichium hagenii*, *Distichium inclinatum*, *Distichium macounii*, *Eccremidium floridanum*, *Eccremidium whiteleggei*, *Lophidion strictus*, *Pleuridium acuminatum*, *Pleuridium alternifolium*, *Pleuridium holdridgei*, *Pleuridium mexicanum*, *Pleuridium ravenelii*, *Pleuridium subulatum*, *Saelania glaucescens*, *Trichodon borealis*, *Trichodon cylindricus* or *Trichodon cylindricus var. oblongus*, Elaeagnaceae such as the genus *Elaeagnus*, for example the genus and species *Olea europaea* [olive], Ericaceae such as the genus *Kalmia*, for example the genera and species *Kalmia latifolia*, *Kalmia angustifolia*, *Kalmia microphylla*, *Kalmia polifolia*, *Kalmia occidentalis*, *Cistus chamaerhodendros* or *Kalmia lucida* [mountain laurel], Euphorbiaceae such as the genera *Manihot*, *Janipha*, *Jatropha*, *Ricinus*, for example the genera and species *Manihot utilissima*, *Janipha manihot*, *Jatropha manihot*, *Manihot aipil*, *Manihot dulcis*, *Manihot manihot*, *Manihot melanobasis*, *Manihot esculenta* [manihot] or *Ricinus communis* [castor-oil plant], Fabaceae such as the genera *Pisum*, *Albizia*, *Cathormion*, *Feuillea*, *Inga*, *Pithecolobium*, *Acacia*, *Mimosa*, *Medicago*, *Glycine*, *Dolichos*, *Phaseolus*, *Soja*, for example the genera and species *Pisum sativum*, *Pisum arvense*, *Pisum humile* [pea], *Albizia berteriana*, *Albizia julibrissin*, *Albizia lebbeck*, *Acacia berteriana*, *Acacia littoralis*, *Albizia berteriana*, *Albizia berteriana*, *Cathormion berteriana*, *Feuillea berteriana*, *Inga fragrans*, *Pithecellobium berterianum*, *Pithecellobium fragrans*, *Pithecolobium berterianum*, *Pseudalbizia berteriana*, *Acacia julibrissin*, *Acacia nemu*, *Albizia nemu*, *Feuillea julibrissin*, *Mimosa julibrissin*, *Mimosa speciosa*, *Sericanrda julibrissin*, *Acacia lebbeck*, *Acacia macrophylla*, *Albizia lebbeck*, *Feuillea lebbeck*, *Mimosa lebbeck*, *Mimosa speciosa* [silk tree], *Medicago sativa*, *Medicago*

falcata, *Medicago varia* [alfalfa], *Glycine max* *Dolichos soja*, *Glycine gracilis*, *Glycine hispida*, *Phaseolus max*, *Soja hispida* or *Soja max* [soybean], Funariaceae such as the genera *Aphanorrhagma*, *Entosthodon*, *Funaria*, *Physcomitrella*, *Physcomitrium*, for example the genera and species *Aphanorrhagma serratum*, *Entosthodon attenuatus*,
5 *Entosthodon bolanderi*, *Entosthodon bonplandii*, *Entosthodon californicus*, *Entosthodon drummondii*, *Entosthodon jamesonii*, *Entosthodon leibergii*, *Entosthodon neoscoticus*, *Entosthodon rubrissetus*, *Entosthodon spathulifolius*, *Entosthodon tucsoni*, *Funaria americana*, *Funaria bolanderi*, *Funaria calcarea*, *Funaria californica*, *Funaria calvescens*, *Funaria convoluta*, *Funaria flavicans*, *Funaria groutiana*, *Funaria hygrometrica*,
10 *Funaria hygrometrica* var. *arctica*, *Funaria hygrometrica* var. *calvescens*, *Funaria hygrometrica* var. *convoluta*, *Funaria hygrometrica* var. *muralis*, *Funaria hygrometrica* var. *utahensis*, *Funaria microstoma*, *Funaria microstoma* var. *obtusifolia*, *Funaria muhlenbergii*, *Funaria orcuttii*, *Funaria plano-convexa*, *Funaria polaris*, *Funaria ravenelii*, *Funaria rubrisseta*, *Funaria serrata*, *Funaria sonora*, *Funaria sublimbatus*, *Funaria tucsoni*, *Physcomitrella californica*, *Physcomitrella patens*, *Physcomitrella readeri*, *Physcomitrium australe*, *Physcomitrium californicum*, *Physcomitrium collenchymatum*, *Physcomitrium coloradense*, *Physcomitrium cupuliferum*, *Physcomitrium drummondii*, *Physcomitrium eurystomum*, *Physcomitrium flexifolium*, *Physcomitrium hookeri*, *Physcomitrium hookeri* var. *serratum*, *Physcomitrium immersum*, *Physcomitrium kellermanii*,
20 *Physcomitrium megalocarpum*, *Physcomitrium pyriforme*, *Physcomitrium pyriforme* var. *serratum*, *Physcomitrium rufipes*, *Physcomitrium sandbergii*, *Physcomitrium subsphaericum*, *Physcomitrium washingtoniense*, Geraniaceae, such as the genera *Pelargonium*, *Cocos*, *Oleum*, for example the genera and species *Cocos nucifera*, *Pelargonium grossularioides* or *Oleum cocois* [coconut], Gramineae, such as the genus
25 *Saccharum*, for example the genus and species *Saccharum officinarum*, Juglandaceae, such as the genera *Juglans*, *Wallia*, for example the genera and species *Juglans regia*, *Juglans ailanthifolia*, *Juglans sieboldiana*, *Juglans cinerea*, *Wallia cinerea*, *Juglans bixbyi*, *Juglans californica*, *Juglans hindsii*, *Juglans intermedia*, *Juglans jamaicensis*, *Juglans major*, *Juglans microcarpa*, *Juglans nigra* or *Wallia nigra* [walnut], Lauraceae, such as the genera *Persea*, *Laurus*, for example the genera and species *Laurus nobilis* [bay], *Persea americana*, *Persea gratissima* or *Persea persea* [avocado],
Leguminosae, such as the genus *Arachis*, for example the genus and species *Arachis hypogaea* [peanut], Linaceae, such as the genera *Linum*, *Adenolinum*, for example the genera and species *Linum usitatissimum*, *Linum humile*, *Linum austriacum*, *Linum bi-*
35 *enne*, *Linum angustifolium*, *Linum catharticum*, *Linum flavum*, *Linum grandiflorum*, *Adenolinum grandiflorum*, *Linum lewisii*, *Linum narbonense*, *Linum perenne*, *Linum perenne* var. *lewisii*, *Linum pratense* or *Linum trigynum* [linseed], Lythraeae, such as the genus *Punica*, for example the genus and species *Punica granatum* [pomegranate],
Malvaceae, such as the genus *Gossypium*, for example the genera and species *Gossypium hirsutum*, *Gossypium arboreum*, *Gossypium barbadense*, *Gossypium her-*
40

baceum or *Gossypium thurberi* [cotton], Marchantiaceae, such as the genus *Marchantia*, for example the genera and species *Marchantia berteroana*, *Marchantia foliacea*, *Marchantia macropora*, Musaceae, such as the genus *Musa*, for example the genera and species *Musa nana*, *Musa acuminata*, *Musa paradisiaca*, *Musa* spp. [banana],
5 Onagraceae, such as the genera *Camissonia*, *Oenothera*, for example the genera and species *Oenothera biennis* or *Camissonia brevipes* [evening primrose], Palmae, such as the genus *Elacis*, for example the genus and species *Elaeis guineensis* [oil palm], Papaveraceae, such as the genus *Papaver*, for example the genera and species *Papaver orientale*, *Papaver rhoeas*, *Papaver dubium* [poppy], Pedaliaceae, such as the
10 genus *Sesamum*, for example the genus and species *Sesamum indicum* [sesame], Piperaceae, such as the genera *Piper*, *Artanthe*, *Peperomia*, *Steffensia*, for example the genera and species *Piper aduncum*, *Piper amalago*, *Piper angustifolium*, *Piper auritum*, *Piper betel*, *Piper cubeba*, *Piper longum*, *Piper nigrum*, *Piper retrofractum*, *Artanthe adunca*, *Artanthe elongata*, *Peperomia elongata*, *Piper elongatum*, *Steffensia elon-*
15 *gata* [cayenne pepper], Poaceae, such as the genera *Hordeum*, *Secale*, *Avena*, *Sorghum*, *Andropogon*, *Holcus*, *Panicum*, *Oryza*, *Zea* (maize), *Triticum*, for example the genera and species *Hordeum vulgare*, *Hordeum jubatum*, *Hordeum murinum*, *Hordeum secalinum*, *Hordeum distichon*, *Hordeum aegiceras*, *Hordeum hexastichon*, *Hordeum hexastichum*, *Hordeum irregulare*, *Hordeum sativum*, *Hordeum secalinum* [bar-
20 ley], *Secale cereale* [rye], *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena fatua* var. *sativa*, *Avena hybrida* [oats], *Sorghum bicolor*, *Sorghum halepense*, *Sorghum saccharatum*, *Sorghum vulgare*, *Andropogon drummondii*, *Holcus bicolor*, *Holcus sorghum*, *Sorghum aethiopicum*, *Sorghum arundinaceum*, *Sorghum caffrorum*, *Sorghum cernuum*, *Sorghum dochna*, *Sorghum drummondii*, *Sorghum durra*, *Sorghum*
25 *guineense*, *Sorghum lanceolatum*, *Sorghum nervosum*, *Sorghum saccharatum*, *Sorghum subglabrescens*, *Sorghum verticilliflorum*, *Sorghum vulgare*, *Holcus halepensis*, *Sorghum miliaceum*, *Panicum militaceum* [millet], *Oryza sativa*, *Oryza latifolia* [rice], *Zea mays* [maize], *Triticum aestivum*, *Triticum durum*, *Triticum turgidum*, *Triticum hybernium*, *Triticum macha*, *Triticum sativum* or *Triticum vulgare* [wheat], Porphyridi-
30 aceae, such as the genera *Chrootheca*, *Flintiella*, *Petrovanella*, *Porphyridium*, *Rhodella*, *Rhodosorus*, *Vanhoeffenia*, for example the genus and species *Porphyridium cruentum*, Proteaceae, such as the genus *Macadamia*, for example the genus and species *Macadamia intergrifolia* [macadamia], Prasinophyceae such as the genera *Nephroselmis*, *Prasinococcus*, *Scherffelia*, *Tetraselmis*, *Mantoniella*, *Ostreococcus*, for
35 example the genera and species *Nephroselmis olivacea*, *Prasinococcus capsulatus*, *Scherffelia dubia*, *Tetraselmis chui*, *Tetraselmis suecica*, *Mantoniella squamata*, *Ostreococcus tauri*, Rubiaceae such as the genus *Cofea*, for example the genera and species *Cofea* spp., *Coffea arabica*, *Coffea canephora* or *Coffea liberica* [coffee], Scrophulariaceae such as the genus *Verbascum*, for example the genera and species *Verbas-*
40 *cum blattaria*, *Verbascum chaixii*, *Verbascum densiflorum*, *Verbascum lagurus*, *Ver-*

Verbascum longifolium, *Verbascum lychnitis*, *Verbascum nigrum*, *Verbascum olympicum*,
Verbascum phlomoides, *Verbascum phoenicum*, *Verbascum pulverulentum* or *Verbas-*
cum thapsus [mullein], Solanaceae such as the genera *Capsicum*, *Nicotiana*, *Solanum*,
Lycopersicon, for example the genera and species *Capsicum annuum*, *Capsicum an-*
5 *nuum* var. *glabriusculum*, *Capsicum frutescens* [pepper], *Capsicum annuum* [paprika],
Nicotiana tabacum, *Nicotiana alata*, *Nicotiana attenuata*, *Nicotiana glauca*, *Nicotiana*
langsdorffii, *Nicotiana obtusifolia*, *Nicotiana quadrivalvis*, *Nicotiana repanda*, *Nicotiana*
rustica, *Nicotiana sylvestris* [tobacco], *Solanum tuberosum* [potato], *Solanum melon-*
gena [eggplant], *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon*
10 *pyriforme*, *Solanum integrifolium* or *Solanum lycopersicum* [tomato], Sterculiaceae,
such as the genus *Theobroma*, for example the genus and species *Theobroma cacao*
[cacao] or Theaceae, such as the genus *Camellia*, for example the genus and species
Camellia sinensis [tea]. In particular preferred plants to be used as transgenic plants in
accordance with the present invention are oil fruit crops which comprise large amounts
15 of lipid compounds, such as peanut, oilseed rape, canola, sunflower, safflower, poppy,
mustard, hemp, castor-oil plant, olive, sesame, *Calendula*, *Punica*, evening primrose,
mullein, thistle, wild roses, hazelnut, almond, macadamia, avocado, bay, pump-
kin/squash, linseed, soybean, pistachios, borage, trees (oil palm, coconut, walnut) or
crops such as maize, wheat, rye, oats, triticale, rice, barley, cotton, cassava, pepper,
20 *Tagetes*, Solanaceae plants such as potato, tobacco, eggplant and tomato, *Vicia* spe-
cies, pea, alfalfa or bushy plants (coffee, cacao, tea), *Salix* species, and perennial
grasses and fodder crops. Preferred plants according to the invention are oil crop
plants such as peanut, oilseed rape, canola, sunflower, safflower, poppy, mustard,
hemp, castor-oil plant, olive, *Calendula*, *Punica*, evening primrose, pumpkin/squash,
25 linseed, soybean, borage, trees (oil palm, coconut). Especially preferred are plants
which are high in C18:2- and/or C18:3-fatty acids, such as sunflower, safflower, to-
bacco, mullein, sesame, cotton, pumpkin/squash, poppy, evening primrose, walnut,
linseed, hemp, thistle or safflower. Very especially preferred plants are plants such as
safflower, sunflower, poppy, evening primrose, walnut, linseed, or hemp.
30
Preferred mosses are *Physcomitrella* or *Ceratodon*. Preferred algae are *Isochrysis*,
Mantoniella, *Ostreococcus* or *Cryptothecodinium*, and algae/diatoms such as *Phaeodac-*
tylum or *Thraustochytrium*. More preferably, said algae or mosses are selected from
the group consisting of: *Shewanella*, *Physcomitrella*, *Thraustochytrium*, *Fusarium*, *Phy-*
35 *tophthora*, *Ceratodon*, *Isochrysis*, *Aleurita*, *Muscarioides*, *Mortierella*, *Phaeodactylum*,
Cryptothecodinium, specifically from the genera and species *Thallasiosira pseudonona*,
Euglena gracilis, *Physcomitrella patens*, *Phytophthora infestans*, *Fusarium gramineum*,
Cryptocodinium cohnii, *Ceratodon purpureus*, *Isochrysis galbana*, *Aleurita farinosa*,
Thraustochytrium sp., *Muscarioides viallii*, *Mortierella alpina*, *Phaeodactylum tricornu-*
40 *tum* or *Caenorhabditis elegans* or especially advantageously *Phytophthora infestans*,

Thallasiosira pseudonona and Cryptocodinium cohnii.

Transgenic plants may be obtained by transformation techniques as published, and cited, in: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), chapter 6/7, pp.71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, vol. 1, Engineering and Utilization, Ed.: Kung and R. Wu, Academic Press, 1993, 15-38; B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, vol. 1, Engineering and Utilization, Ed.: Kung and R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225. Preferably, transgenic plants can be obtained by T-DNA-mediated transformation. Such vector systems are, as a rule, characterized in that they contain at least the vir genes, which are required for the Agrobacterium-mediated transformation, and the sequences which delimit the T-DNA (T-DNA border). Suitable vectors are described elsewhere in the specification in detail.

Preferably, a multicellular micro-organism as used herein refers to protists or diatoms. More preferably, it is selected from the group of the families Dinophyceae, Turaniellidae or Oxytrichidae, such as the genera and species: *Cryptocodinium cohnii*, *Phaeodactylum tricornutum*, *Stylonychia mytilus*, *Stylonychia pustulata*, *Stylonychia putrina*, *Stylonychia notophora*, *Stylonychia sp.*, Colpidium campylum or Colpidium sp.

The present invention also relates to a method for expressing a nucleic acid of interest in a host cell comprising

- (a) introducing the polynucleotide or the vector of the present invention into the host cell, whereby the nucleic acid sequence of interest will be operatively linked to the expression control sequence; and
- (b) expressing the said nucleic acid sequence in said host cell.

The polynucleotide or vector of the present invention can be introduced into the host cell by suitable transfection or transformation techniques as specified elsewhere in this description. The nucleic acid of interest will be expressed in the host cell under suitable conditions. To this end, the host cell will be cultivated under conditions which, in principle, allow for transcription of nucleic acids. Moreover, the host cell, preferably, comprises the exogenously supplied or endogenously present transcription machinery required for expressing a nucleic acid of interest by the expression control sequence. More preferably, the host cell is a plant cell and, most preferably, a seed cell or precursor thereof.

Moreover, the present invention encompasses a method for expressing a nucleic acid of interest in a non-human organism comprising

- 5 (a) introducing the polynucleotide or the vector of the present invention into the non human organism, whereby the nucleic acid sequence of interest will be operatively linked to the expression control sequence ; and
- (b) expressing the said nucleic acid sequence in said non-human transgenic organism.

10 The polynucleotide or vector of the present invention can be introduced into the non-human transgenic organism by suitable techniques as specified elsewhere in this description. The non-human transgenic organism, preferably, comprises the exogenously supplied or endogenously present transcription machinery required for expressing a nucleic acid of interest by the expression control sequence. More preferably, the non-human transgenic organism is a plant or seed thereof. It is to be understood that the
15 nucleic acid of interest will be expressed, preferably, seed specific in the said non-human transgenic organism.

Further, the present invention relates to a method for the manufacture of stearidonic acid (SDA) in a plant seed comprising the steps of:

- 20 a) growing a transgenic plant expressing a polynucleotide encoding a delta 6 desaturase under the control of the polynucleotide of the present invention; and
- b) obtaining said SDA from the harvested seeds of the said plant.

25 The SDA may be manufactured in the form of a triglyceride ester, a phospholipids or as Acyl-CoA bound or free fatty acid comprised by seed oil or seed fatty acid preparations obtainable from the harvested seeds of the grown transgenic plants by standard techniques such as an oil mill or chromatographic extraction and/or purification techniques.

30 Moreover, how to grow transgenic plants and how to harvest their seeds is well known in the art. How to make transgenic plants expressing a gene of interest such as a delta 6 desaturase under the control of the polynucleotide of the present invention is set forth elsewhere herein.

35 Surprisingly, the polynucleotides of the present invention were found to influence the ratio of the omega-3 fatty acid stearidonic acid (SDA) to the omega-6 fatty acid gamma linolenic acid (GLA). Accordingly, the seeds of the aforementioned transgenic plants expressing a delta 6 desaturase under the control of the promoter comprised by the polynucleotide of the invention are particularly suitable as a source for SDA or SDA
40 enriched fatty acid preparations such as oil.

Thus, the present invention also pertains to the use of the polynucleotide of the present invention driving expression of a delta 6 desaturase in a transgenic plant for increasing (preferably to a statistically significant extent) the amount of SDA at the expense of
5 GLA in plant seeds of said plants.

Moreover, the present invention pertains to seed oil having the said altered SDA to GLA ratio (i.e. an increased SDA amount at the expense of GLA) obtainable by an oil mill from the harvested seeds of a transgenic plant as set forth above. It will be under-
10 stood that such an oil will be characterized in addition to the altered SDA to GLA ratio by the presence of remaining DNA contaminations including the polynucleotide of the present invention and/or the delta 6 desaturase encoding polynucleotide.

Nucleic acids encoding suitable delta 6 desaturases are well known in the art and are
15 d6-Desaturases d6Des(Cp) from *Ceratodon purpureus* (WO2000075341), d6Des(Ol) from *Ostreococcus lucimarinus* (WO2008040787), d6Des(Ot) from *Ostreococcus tauri* (WO2006069710), d6Des(Pf) from *Primula farinosa* (WO2003072784), d6Des(Pir)_BO from *Pythium irregulare* (WO2002026946), d6Des(Pir) from *Pythium irregulare* (WO2002026946), d6Des(Plu) from *Primula luteola* (WO2003072784), d6Des(Pp) from
20 *Physcomitrella patens* (WO200102591), d6Des(Pt) from *Phaeodactylum tricornutum* (WO2002057465), d6Des(Pv) from *Primula vialii* (WO2003072784) and d6Des(Tp) from *Thalassiosira pseudonana* (WO2006069710) and, in particular, those mentioned in the accompanying Examples.

25 Moreover, a transgenic plant expressing a polynucleotide encoding a delta 6 desaturase under the control of the polynucleotide of the present invention may also comprise further desaturases or elongases of the omega-3 pathway which are required for the synthesis of end products such as eicosapentaenoic acid (EPA) or docosahexaenoic
30 acid (DHA).

Thus, a method is provided for increasing stearidonic acid in a plant seed comprising the steps of growing a transgenic plant expressing a polynucleotide encoding a delta 6 desaturase under the control of the polynucleotide of the present invention.

35 By increasing the omega-3 pathway substrate SDA at the expense of the omega-6 pathway substrate GLA, further fatty acid products of the omega-3 pathway can be produced more efficiently in the aforementioned transgenic plants. Preferably, the desaturases and/or elongases required for the production of a desired fatty acid can also
40 be expressed under the control of a polynucleotide of the present invention. Most pref-

erably, however, it is envisaged that the delta 6 desaturase and the further desaturases and/or elongases are expressed under the control of polynucleotides of the present invention comprising different expression control sequences with respect to each other.

5

In the following tables 1 to 9, the cis-regulatory elements found in the expression control sequences of the present invention are shown.

10

Table 1: cis-regulatory elements of SEQ ID NO: 1

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
SEQ_1	P\$AHBP/ATHB9.01	HD-ZIP class III protein ATHB9	0.77	7	17	(-)	1.000	0.772	ttgATGAtttc
SEQ_1	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	0.82	57	77	(+)	1.000	0.897	aaaaaCCATatcttgaaaacc
SEQ_1	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes Arabidopsis Telo-box interacting protein re- lated to the conserved animal protein Pur- alpha	0.81	57	75	(+)	0.750	0.827	aaAAACcatacttgaaaa
SEQ_1	P\$TELO/ATPURA.01	Trihelix DNA-binding factor GT-3a	0.85	57	71	(+)	0.750	0.851	aaaaACCAtatcttg
SEQ_1	P\$GTBX/GT3A.01	Plant specific floral meristem identity gene LEAFY (LFY)	0.83	86	102	(-)	1.000	0.847	accattGTTActccct
SEQ_1	P\$LFYB/LFY.01	SEF3, Soybean embryo factor 3	0.93	91	103	(-)	0.914	0.936	tACCAttgttact
SEQ_1	P\$SEF3/SEF3.01	Circadian clock associated 1	0.87	147	161	(+)	1.000	0.875	acccaACCCaaagag
SEQ_1	P\$CCAF/CCA1.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photo- synthesis-related genes	0.85	177	191	(+)	1.000	0.943	tcagctaaAATCtaa
SEQ_1	P\$LREM/ATCTA.01	RY and Sph motifs conserved in seed-specific promoters	0.85	184	194	(+)	1.000	0.921	aaATCTaagga
SEQ_1	P\$LEGB/RV.01	CCAAT-box in plant promoters	0.87	202	228	(-)	1.000	0.936	ggctactcCATGcaatattggatgctc
SEQ_1	P\$CAAT/CAAT.01	Nodulin consensus sequence 1	0.97	206	214	(+)	1.000	0.983	atCCAAtat
SEQ_1	P\$NCS1/NCS1.01	E2F class I sites	0.85	230	240	(+)	0.878	0.865	aAAATgatgcg
SEQ_1	P\$E2FF/E2F.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photo- synthesis-related genes	0.82	234	248	(-)	0.757	0.833	ttgtTTCTcgcatca
SEQ_1	P\$LREM/ATCTA.01	Protodermal factor 2	0.85	274	284	(+)	1.000	0.918	cgATCTacaat
SEQ_1	P\$L1BX/PDF2.01		0.85	278	294	(+)	1.000	0.899	ctacaaTAAAtaccaga

SEQ_1	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	287	301	(+)	1.000	0.881	ataccagaAATCtca
SEQ_1	P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.85	330	346	(+)	0.843	0.903	catgagGTGAgctcttt
SEQ_1	P\$OPAQ/GCN4.01	GCN4, conserved in cereal seed storage protein gene promoters, similar to yeast GCN4 and vertebrate AP-1	0.81	332	348	(+)	1.000	0.830	tgaggTGAGtcttttt
SEQ_1	P\$SALT/ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	0.95	360	374	(-)	1.000	0.977	ggtatgcGGTGtttc
SEQ_1	P\$NACF/TANAC69.01	Wheat NACdomain DNA binding factor	0.68	367	389	(+)	0.812	0.736	cgcataccagaAACGtaaagaaa
SEQ_1	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0.75	375	391	(+)	1.000	0.815	agaaacgtAAAGaaaat
SEQ_1	P\$HEAT/HSE.01	Heat shock element	0.81	375	389	(+)	1.000	0.918	agaaacgtaaAGAAa
SEQ_1	P\$MADS/MADS.01	Binding sites for AP1, AP3-PI and AG dimers	0.75	411	431	(-)	1.000	0.791	tttccCATattttttacatt
SEQ_1	P\$TBPF/TATA.01	Plant TATA box	0.88	464	478	(+)	1.000	0.941	aaaaTATAaaaaaaa
SEQ_1	P\$GAPB/GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	0.88	474	488	(+)	0.807	0.895	aaaaATTAAaagaaa
SEQ_1	P\$PSRE/GAAA.01	GAAA motif involved in pollen specific transcriptional activation	0.83	480	496	(+)	1.000	0.838	taaaaGAAAaatttgac
SEQ_1	P\$WBXF/WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	487	503	(+)	1.000	0.920	aaattTTGAcgctgaaa
SEQ_1	P\$MYBL/MYBPH3.01	Myb-like protein of Petunia hybrida	0.80	498	514	(+)	0.750	0.801	ctgaaaccGTAAatctt
SEQ_1	P\$TELO/ATPURA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.85	499	513	(+)	0.750	0.863	tgaaACCGtaaatact
SEQ_1	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	501	515	(+)	1.000	0.889	aaaccgtaAATCtta
SEQ_1	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	517	533	(+)	0.750	0.851	aatcaaCAAAtgcataa
SEQ_1	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	525	541	(+)	1.000	0.954	aatgaTAAAtgcaaag
SEQ_1	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage	0.75	530	546	(+)	1.000	0.773	ataaatgcAAAGttatt

SEQ_1	P\$AHBP/BLR.01	age protein gene promoters Transcriptional repressor BELLRINGER Plant TATA box GA-regulated myb gene from barley DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TAC-TATT) sequences of sporamin and beta-amylase genes Maize INDETERMINATE1 zinc finger protein Sequence motif from the promoters of different sugar-responsive genes Sunflower homeodomain leucine-zipper protein Hahb-4 SBF-1 Maize INDETERMINATE1 zinc finger protein GT1-Box binding factors with a trihelix DNA-binding domain DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TAC-TATT) sequences of sporamin and beta-amylase genes Arabidopsis thaliana meristem layer 1 CA-rich element PBF (MPBF) OCS-like elements OCS-like elements Sunflower homeodomain leucine-zipper protein Hahb-4 GT1-Box binding factors with a trihelix DNA-binding domain	0.90	538	548	(+)	0.826	0.936	aaaGTTAttga
	SEQ_1	P\$TBPF/TATA.01	0.88	557	571	(-)	1.000	0.964	ctaaTATAaaatat
	SEQ_1	P\$MYBL/GAMYB.01	0.91	583	599	(-)	1.000	0.934	tagtatttGTTAgcagt
SEQ_1	P\$SPF1/SP8BF.01		0.87	593	605	(+)	1.000	0.902	aaTACTatacaga
	SEQ_1	P\$IDDF/ID1.01	0.92	600	612	(-)	1.000	0.955	tgttTTGTctgta
	SEQ_1	P\$SUCB/SUCROSE.01	0.81	605	623	(+)	0.750	0.836	acAAAAAcacattattaa
SEQ_1	P\$AHBP/HAHB4.01		0.87	611	621	(+)	1.000	0.909	cacattATTaA
	SEQ_1	P\$GTBX/SBF1.01	0.87	611	627	(+)	1.000	0.886	cacattaTTAAaaaaac
	SEQ_1	P\$IDDF/ID1.01	0.92	627	639	(-)	1.000	0.965	tattTTGTctttg
SEQ_1	P\$GTBX/GT1.01		0.85	636	652	(-)	1.000	0.858	tgatatGTTAatatatt
SEQ_1	P\$SPF1/SP8BF.01		0.87	649	661	(+)	0.814	0.901	atCACTattacta
	SEQ_1	P\$L1BX/ATML1.02	0.76	669	685	(+)	0.890	0.762	acaCAATaaaaacacca
	SEQ_1	P\$CARM/CARICH.01	0.78	672	690	(+)	1.000	0.815	caataaaaAACAccaataa
SEQ_1	P\$DOFF/PBF.01		0.97	698	714	(+)	1.000	0.990	aacaaataAAGtgatc
SEQ_1	P\$OCSE/OC SL.01		0.69	701	721	(+)	0.807	0.729	aaataaaaagtgatcACA Taat
SEQ_1	P\$OCSE/OC SL.01		0.69	704	724	(-)	0.769	0.713	gtaattatgtgatcACTTtta
SEQ_1	P\$AHBP/HAHB4.01		0.87	714	724	(+)	1.000	0.902	cacataATTAc
SEQ_1	P\$GTBX/GT1.01		0.85	714	730	(-)	0.968	0.867	gataatGTAAttatgtg

SEQ_1	P\$IBOX/GATA.01	Class I GATA factors	0.93	719	735	(-)	1.000	0.942	ttttgGATAatgtaatt
SEQ_1	P\$MYBS/MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	0.90	722	738	(+)	1.000	0.962	tacattATCCaaaaaat
SEQ_1	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	733	751	(+)	1.000	0.908	aaAAATcatacttttaaca
SEQ_1	P\$GTBX/SBF1.01	SBF-1	0.87	740	756	(-)	1.000	0.967	attgttgTTAAaagtat
SEQ_1	P\$MYBS/TAMYB80.01	MYB protein from wheat	0.83	759	775	(+)	1.000	0.857	aacaATATccgcgcgga
SEQ_1	P\$CGCG/OSCBT.01	Oryza sativa CaM-binding transcription factor	0.78	768	784	(-)	1.000	0.781	gtcCGCGcttcgcgcgg
SEQ_1	P\$NCS3/NCS3.01	Nodulin consensus sequence 3	0.89	782	792	(+)	1.000	0.913	gaCACCCcccct
SEQ_1	P\$OPAQ/O2_GCNA4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	802	818	(-)	1.000	0.829	catacaACATgactaca
SEQ_1	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	842	862	(-)	0.838	0.843	gcATGGgaaatcaggtccatc
SEQ_1	P\$EINL/TEIL.01	TEIL (tobacco EIN3-like)	0.92	843	851	(+)	0.863	0.966	aTGGAcctg
SEQ_1	P\$LEGB/LEGB.01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.59	847	873	(-)	0.750	0.592	acgactaCTATgcatgggaaatcaggt
SEQ_1	P\$LEGB/RV.01	RY and Sph motifs conserved in seed-specific promoters	0.87	850	876	(+)	1.000	0.944	tgatttccCATGcatagtagtgcgtcat
SEQ_1	P\$MADS/AGL15.01	AGL15, Arabidopsis MADS-domain protein	0.79	851	871	(-)	1.000	0.850	gacTACTatgcatgggaaatc
SEQ_1	P\$MADS/AGL3.01	AGAMOUS-like 15	0.83	852	872	(+)	1.000	0.864	atttcCCATgcatagtagtcg
SEQ_1	P\$SPF1/SP8BF.01	AGL3, MADS Box protein							
SEQ_1	P\$GBOX/TGA1.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TAC-TATT) sequences of sporamin and beta-amylase genes	0.87	858	870	(-)	1.000	0.922	acTACTatgcatg
SEQ_1		Arabidopsis leucine zipper protein TGA1	0.90	861	881	(-)	1.000	0.900	ccgagaTGACgactactatgc

SEQ_1	P\$TELO/ATPURA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.85	873	887	(-)	1.000	0.860	ataaACCCGagatga
SEQ_1	P\$OPAQ/O2_GCN4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	883	899	(-)	0.829	0.887	gatataACTTgaataaa
SEQ_1	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	912	928	(-)	1.000	0.993	ttcagattAAAGaacgt
SEQ_1	P\$MADS/AGL15.01	AGL15, Arabidopsis MADS-domain protein	0.79	912	932	(+)	0.925	0.793	acgTTCtTtaatctgaaccct
SEQ_1	P\$AHBP/WUS.01	AGAMOUS-like 15 Homeodomain protein WUSCHEL	0.94	915	925	(+)	1.000	0.963	ttcttTAATct
SEQ_1	P\$TELO/ATPURA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.85	924	938	(+)	1.000	0.876	ctgaACCCtatacacc
SEQ_1	P\$STKM/STK.01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	0.85	937	951	(-)	1.000	0.877	tccTAAAtaaatcgg
SEQ_1	P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.85	968	984	(-)	0.968	0.858	aaagtgGTAAttttgt
SEQ_1	P\$DOFF/PBF.01	PBF (MPBF)	0.97	976	992	(-)	1.000	0.988	gagacagaAAAGtggtga
SEQ_1	P\$GTBX/SBF1.01	SBF-1	0.87	999	1015	(+)	1.000	0.907	ctcgtttTTAAtttggt
SEQ_1	P\$MIIG/MYBC1.01	Maize C1 myb-domain protein	0.92	1009	1023	(+)	1.000	0.942	atttgGTAGtttcag
SEQ_1	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	0.87	1048	1060	(-)	0.814	0.872	aaCACTatgaaaa
SEQ_1	P\$OPAQ/O2_GCN4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	1056	1072	(+)	1.000	0.891	gtgttaACATgtttaag

SEQ_1	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	1095	1105	(-)	1.000	0.858	atATCTatgtt
SEQ_1	P\$MYBL/NTMYBAS1.01	Anther-specific myb gene from tobacco	0.96	1125	1141	(+)	1.000	0.967	tagtggtgGTTAacaaa
SEQ_1	P\$GTBX/SBF1.01	SBF-1	0.87	1127	1143	(+)	1.000	0.894	gtggtggTTAAcaaaag
SEQ_1	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	1130	1146	(-)	1.000	0.918	ggccttttGTTAaccac
SEQ_1	P\$GBOX/BZIP911.01	bZIP transcription factor from Antirrhinum majus	0.77	1138	1158	(-)	0.750	0.781	ataagtTGAAtgaccttttg
SEQ_1	P\$OPAQ/O2.01	Opaque-2 regulatory protein	0.87	1141	1157	(+)	0.852	0.882	aaggccattTCAAActta
SEQ_1	P\$NCS1/NCS1.01	Nodulin consensus sequence 1	0.85	1163	1173	(+)	1.000	0.909	tAAAAGataga
SEQ_1	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	1176	1194	(+)	0.750	0.815	gcAAAAGcattgttgataaa
SEQ_1	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	1184	1200	(+)	1.000	0.994	ttgttgatAAAAGcctct
SEQ_1	P\$GAGA/BPC.01	Basic pentacycstine proteins	1.00	1184	1208	(-)	1.000	1.000	ataaagAGAGaggcttatcaacaa
SEQ_1	P\$IBOX/GATA.01	Class I GATA factors	0.93	1184	1200	(+)	1.000	0.938	ttgttGATAaagcctct
SEQ_1	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	1199	1215	(+)	0.777	0.841	ctctcttTATAaaaga
SEQ_1	P\$TBPF/TATA.01	Plant TATA box	0.88	1199	1213	(-)	1.000	0.958	tttaTATAaagagag
SEQ_1	P\$TBPF/TATA.02	Plant TATA box	0.90	1201	1215	(-)	1.000	0.917	tcttTATAtaaagag
SEQ_1	P\$TBPF/TATA.02	Plant TATA box	0.90	1202	1216	(+)	1.000	0.917	tcttTATAtaaagag
SEQ_1	P\$TBPF/TATA.01	Plant TATA box	0.88	1204	1218	(+)	1.000	0.934	tttaTATAaagaggg
SEQ_1	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	1215	1231	(-)	1.000	0.995	aaggcgctAAAGccccct
SEQ_1	P\$WBXF/WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	1236	1252	(+)	1.000	0.975	atgctTTGActttacct
SEQ_1	P\$TELO/RPBX.01	Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression	0.84	1260	1274	(-)	1.000	0.842	cgaaaCCCTtcactt
SEQ_1	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	1317	1333	(+)	0.750	0.830	caagaaTCAAtgtaagc
SEQ_1	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	1332	1346	(-)	1.000	0.879	caactatAATCtgc

SEQ_1	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	1338	1354	(+)	0.817	0.929	tatagtTTGTtagtttt
SEQ_1	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	1342	1358	(+)	1.000	0.818	gtttgtTAGTttttcag
SEQ_1	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	1382	1398	(-)	1.000	0.818	aggcgtTAGTtcagaaa
SEQ_1	P\$MYBL/NTMYBAS1.01	Anther-specific myb gene from tobacco	0.96	1386	1402	(-)	1.000	0.967	tcaaaggcGTTAgttca
SEQ_1	P\$MADS/SQUA.01	MADS-box protein SQUAMOSA	0.90	1400	1420	(-)	1.000	0.902	attgaccATTTtttcttca
SEQ_1	P\$WBXF/ERE.01	Elicitor response element	0.89	1408	1424	(-)	1.000	0.973	gtcaatTGACCcattttt
SEQ_1	P\$GTBX/GT3A.01	Trihelix DNA-binding factor GT-3a	0.83	1425	1441	(-)	1.000	0.897	ttagtgGTTAcaatggc
SEQ_1	P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.85	1432	1448	(-)	1.000	0.882	catgtgGTTAgtggtta
SEQ_1	P\$MIIG/PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	0.81	1433	1447	(-)	0.936	0.820	atGTGGttagtggtt
SEQ_1	P\$LEGB/RV.01	RY and Sph motifs conserved in seed-specific promoters	0.87	1445	1471	(-)	1.000	0.944	tgtgtttgCATGcatagccagtcgatg
SEQ_1	P\$LEGB/RV.01	RY and Sph motifs conserved in seed-specific promoters	0.87	1452	1478	(+)	1.000	0.909	tggctatgCATGcaaacacaatgagat
SEQ_1	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TAC-TATT) sequences of sporamin and beta-amylase genes	0.87	1488	1500	(-)	1.000	0.871	ttTACTcttaggc
SEQ_1	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	1493	1509	(-)	0.778	0.838	caaagtTGGTttactct
SEQ_1	P\$MADS/AGL1.01	AGL1, Arabidopsis MADS-domain protein AGAMOUS-like 1	0.84	1495	1515	(-)	0.975	0.840	ggaTTCcaaagttggttact
SEQ_1	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	0.82	1496	1516	(+)	0.968	0.845	gtaaaCCAAccttggatccc
SEQ_1	P\$STKM/STK.01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	0.85	1513	1527	(+)	0.833	0.873	tccCAAaaaattata
SEQ_1	P\$ERSE/ERSE_I.01	ERSE I (ER stress-response element I)-like motif	0.79	1515	1533	(+)	0.750	0.803	ccaaaaaattatagcCATG
SEQ_1	P\$GBOX/EMBP1.01	bZIP transcription factor implicated in ABA	0.84	1522	1542	(-)	0.750	0.854	agaacgacACATggctataat

SEQ_1	P\$LEGB/LEGB.01	induced gene expression Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.59	1522	1548	(+)	1.000	0.609	attatagCCATggtgctgttcttgatga
SEQ_1	P\$ABRE/ABF1.01	ABA (abscisic acid) inducible transcriptional activator	0.79	1525	1541	(-)	1.000	0.853	gaacgACACatggctat
SEQ_1	P\$OCSE/OCSE.01	OCS-like elements	0.69	1549	1569	(-)	0.807	0.693	aaattttattggaaACGAatt
SEQ_1	P\$GTBX/SBF1.01	SBF-1	0.87	1562	1578	(-)	0.826	0.872	ttgtttCTAAatttta
SEQ_1	P\$MIIG/PALBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TAC- TATT) sequences of sporamin and beta- amylase genes	0.81	1570	1584	(-)	0.936	0.819	ttGTGGttgtttct
SEQ_1	P\$SPF1/SP8BF.01	OCS-like elements	0.87	1586	1598	(-)	1.000	0.881	atTACTttgtatt
SEQ_1	P\$OCSE/OCSE.01	Transcriptional repressor BELLRINGER	0.69	1589	1609	(-)	0.769	0.716	taagttaaaaaaattACTTtgt
SEQ_1	P\$AHBP/BLR.01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and su- crose-inducible gene expression	0.90	1591	1601	(-)	1.000	0.976	aaaATTActtt
SEQ_1	P\$STKM/STK.01	Binding sites for AP1, AP3-PI and AG dimers	0.85	1593	1607	(-)	1.000	0.901	agtTAAaaaattact
SEQ_1	P\$MADS/MADS.01	Protodermal factor 2	0.75	1599	1619	(-)	1.000	0.777	ttccCCATttaagttaaaaa
SEQ_1	P\$L1BX/PDF2.01	I-Box in rbcS genes and other light regulated genes	0.85	1602	1618	(+)	1.000	0.864	ttaactTAAAtggtggaa
SEQ_1	P\$IBOX/IBOX.01	Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression	0.81	1640	1656	(+)	1.000	0.824	aaagaGATAgggcttaa
SEQ_1	P\$TELO/RPBX.01	Dof2 - single zinc finger transcription factor	0.84	1642	1656	(-)	1.000	0.886	ttaagCCCTatctct
SEQ_1	P\$DOFF/DOF2.01	Wheat NACdomain DNA binding factor	0.98	1647	1663	(+)	1.000	0.994	tagggcttAAGGcagca
SEQ_1	P\$NACF/TANAC69.01	GA-regulated myb gene from barley	0.68	1673	1695	(-)	0.812	0.695	aacgaaaaacgaAACGtatcacag
SEQ_1	P\$MYBL/GAMYB.01	Class I GATA factors	0.91	1689	1705	(+)	1.000	0.963	tttcgtttGTTAtcaca
SEQ_1	P\$IBOX/GATA.01		0.93	1691	1707	(-)	1.000	0.931	attgtGATAaacaacga

SEQ_1	P\$GTBX/S1F.01	S1F, site 1 binding factor of spinach rps1 promoter	0.79	1707	1723	(+)	1.000	0.851	tttcATGGactatatac
SEQ_1	P\$TBPF/TATA.02	Plant TATA box	0.90	1713	1727	(+)	1.000	0.901	ggacTATAtacattt
SEQ_1	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	1718	1734	(-)	1.000	0.875	ctaagcTAAAtgtatat
SEQ_1	P\$PSRE/GAAA.01	GAAA motif involved in pollen specific transcriptional activation	0.83	1739	1755	(+)	1.000	0.847	caaaaGAAAccatctac
SEQ_1	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	1748	1758	(+)	1.000	0.854	ccATCTacttg
SEQ_1	P\$AHBP/ATHB1.01	Arabidopsis thaliana homeo box protein 1	0.90	1768	1778	(-)	1.000	0.990	ggaATTAttgt
SEQ_1	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	1768	1778	(+)	0.829	0.940	acaATAAAttcc

Table 2: cis- regulatory elements of SEQ ID NO: 6

5

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
SEQ_6	P\$SEF3/SEF3.01	SEF3, Soybean embryo factor 3	0.87	5	19	(+)	1.000	0.921	ctataACCCaacc
SEQ_6	P\$SEF3/SEF3.01	SEF3, Soybean embryo factor 3	0.87	10	24	(+)	1.000	0.891	accacACCCaaa
SEQ_6	P\$GTBX/SBF1.01	SBF-1	0.87	22	38	(-)	1000	0.880	aggatacTTAAacttgt
SEQ_6	P\$MYBL/ATMYB77.01	R2R3-type myb-like transcription factor (l-type binding site)	0.87	35	51	(-)	1000	0.892	ttttgtCGGTtcagga
SEQ_6	P\$DREB/CRT_DRE.01	C-repeat/dehydration response element	0.89	38	52	(+)	1.000	0.902	tgaaaCCGAcaaaag
SEQ_6	P\$DOFF/PBF.01	PBF (MPBF)	0.97	41	57	(+)	1000	0.986	aaccgacaAAGagaat
SEQ_6	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	41	53	(-)	1.000	0.987	tcTTTGTcggtt
SEQ_6	P\$EPFF/ZPT22.01	Member of the EPF family of zinc finger transcription factors	0.75	50	72	(-)	1.000	0.770	agtcaaaaaagaCAGTattcttt

SEQ_6	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT)	0.87	55	67	(+)	1.000	0.881	aaTACTgtctttt
SEQ_6	P\$NCS2/NCS2.01	sequences of sporamin and beta-amylase genes Nodulin consensus sequence 2 WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box Heat shock element Transcriptional repressor BELLRINGER Prolamin box, conserved in cereal seed storage protein gene promoters SBF-1 Basic pentacysteine proteins GAAA motif involved in pollen specific transcriptional activation	0.79	57	71	(+)	0.750	0.795	tactgtCTTTtgac
SEQ_6	P\$WBXF/WRKY.01		0.92	62	78	(+)	1.000	0.965	tccttTTGActtctcg
SEQ_6	P\$HEAT/HSE.01		0.81	72	86	(-)	0.826	0.819	agattattcaGGAAa
SEQ_6	P\$AHBP/BLR.01		0.90	77	87	(-)	1000	0.901	aagATTAttca
SEQ_6	P\$DOFF/PBOX.01		0.75	80	96	(-)	1.000	0.769	tatttttAAAGattat
SEQ_6	P\$GTBX/SBF1.01		0.87	80	96	(+)	1.000	0.892	ataatctTTAAaaaata
SEQ_6	P\$GAGA/BPC.01		1.00	88	112	(-)	1.000	1000	ttccagAGAGaactgatattttta
SEQ_6	P\$PSRE/GAAA.01		0.83	105	121	(+)	1000	0.878	ctctgGAAAtagtaaag
SEQ_6	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT)	0.87	108	120	(-)	1.000	0.943	ttTACTatttcca
SEQ_6	P\$OCSE/OCSL.01	sequences of sporamin and beta-amylase genes OCS-like elements Octamer motif found in plant histone H3 and H4 genes Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes M-phase-specific activators (NtmybA1, NtmybA2, NtmybB) Sunflower homeodomain leucine-zipper protein Hahb-4 Arabidopsis thaliana signal-responsive gene1, Ca2+/- calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene) OCS-like elements	0.69	116	136	(-)	0.769	0.700	ttgcatcagttcttACTTtac
SEQ_6	P\$HOCT/HOCT.01		0.76	145	161	(+)	0.750	0.762	atcaccgATCTacgaag
SEQ_6	P\$LREM/ATCTA.01		0.85	150	160	(+)	1.000	0.864	cgATCTacgaa
SEQ_6	P\$MSAE/MSA.01		0.80	174	188	(-)	1.000	0.834	aaaaaAACGggtgga
SEQ_6	P\$AHBP/HAHB4.01		0.87	187	197	(+)	1.000	0.934	ttgataATTAt
SEQ_6	P\$CGCG/ATSR1.01		0.84	194	210	(-)	1.000	0.912	aagCGCGtacgatataa
SEQ_6	P\$OCSE/OCSL.01		0.69	196	216	(-)	0.807	0.709	gttttaagcgcgTACGAtat

SEQ_6	P\$GTBX/SBF1.01	SBF-1	0.87	202	218	(+)	1.000	0.875	tacgcgcTTAAaacct
SEQ_6	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	217	233	(-)	0.778	0.781	tttagtTCGTaaaaaag
SEQ_6	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0.75	245	261	(-)	1.000	0.831	ttatctgaAAAGtaaaa
SEQ_6	P\$IBOX/GATA.01	Class I GATA factors	0.93	252	268	(+)	1.000	0.933	tttcaGATAatgttgca
SEQ_6	P\$AGP1/AGP1.01	AG-motif binding protein 1	0.91	284	294	(-)	1.000	0.988	acaGATCtatac
SEQ_6	P\$AGP1/AGP1.01	AG-motif binding protein 1	0.91	285	295	(+)	1.000	0.920	ataGATCtgtt
SEQ_6	P\$GARP/ARR10.01	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	0.97	287	295	(+)	1.000	0.970	AGATctgtt
SEQ_6	P\$CAAT/CAAT.01	CCAAT-box in plant promoters	0.97	298	306	(+)	1.000	0.973	ttCCAAAtga
SEQ_6	P\$LFYB/LFY.01	Plant specific floral meristem identity gene LEAFY (LFY)	0.93	298	310	(+)	1.000	0.930	tTCCAatgagaat
SEQ_6	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	300	314	(+)	1.000	0.864	ccaatgagAATCtgt
SEQ_6	P\$GARP/ARR10.01	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	0.97	304	312	(-)	1.000	0.971	AGATtctca
SEQ_6	P\$RAV5/RAV1-5.01	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	0.96	308	318	(-)	1.000	0.960	aacAACAgatt
SEQ_6	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	339	351	(-)	1.000	0.934	tattTTGTcagat
SEQ_6	P\$CAAT/CAAT.02	CCAAT-box in plant promoters	1.00	400	408	(-)	1.000	1.000	gtcCAATta
SEQ_6	P\$WBXF/WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	405	421	(-)	1.000	0.957	cacatTTGActatgtcc
SEQ_6	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	407	425	(-)	1.000	0.953	ccaacACATttgactatgt
SEQ_6	P\$CARM/CARICH.01	CA-rich element	0.78	412	430	(-)	1.000	0.785	aaagtccAACAcatttgac
SEQ_6	P\$GTBX/SBF1.01	SBF-1	0.87	424	440	(-)	1.000	0.875	tcggaaaTTAAaagtc
SEQ_6	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	439	457	(+)	1.000	0.913	gaAAATcattaaaaacaat
SEQ_6	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	440	450	(-)	0.829	0.902	ttaATGAtttt
SEQ_6	P\$AHBP/HAHB4.01	Sunflower homeodomain leucine-zipper protein	0.87	440	450	(+)	1.000	0.967	aaaatcATTaA

42

SEQ_6	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes
SEQ_6	P\$DOFF/DOF2.01	Dof2 - single zinc finger transcription factor
SEQ_6	P\$L1BX/ATML1.02	Arabidopsis thaliana meristem layer 1
SSEQ_6	P\$GAPB/GAP.01	Cis-element in the GAPDH promoters conferring light inducibility
SEQ_6	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes
SEQ_6	P\$CCAF/CCA1.01	Circadian clock associated 1
SEQ_6	P\$GTBX/SBF1.01	SBF-1
SSEQ_6	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida
SEQ_6	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGA-MOUS-like 2
SEQ_6	P\$MADS/AG.01	Agamous, required for normal flower development, similarity to SRF (human) and MCM (yeast) proteins
SEQ_6	P\$LFYB/LFY.01	Plant specific floral meristem identity gene LEAFY (LFY)
SEQ_6	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters
SEQ_6	P\$SALT/ALFIN1.01	Zinc-finger protein in alfalfa roots, regulates salt tolerance
SEQ_6	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences
SEQ_6	P\$WBXF/ERE.01	Elicitor response element
SEQ_6	P\$MYBL/ATMYB77.01	R2R3-type myb-like transcription factor (l-type binding site)
SEQ_6	P\$MSAE/MSA.01	M-phase-specific activators (NtmybA1, NtmybA2,

SEQ_6	P\$AHBP/HAHB4.01	NtmybB) Sunflower homeodomain leucine-zipper protein
SEQ_6	P\$GTBX/GT1.01	Hahb-4 GT1-Box binding factors with a trihelix DNA-binding domain
SEQ_6	P\$LEGB/R.Y.01	R.Y and Sph motifs conserved in seed-specific promoters
SEQ_6	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)
SEQ_6	P\$LEGB/R.Y.01	R.Y and Sph motifs conserved in seed-specific promoters
SEQ_6	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)
SEQ_6	P\$CCAF/CCA1.01	Circadian clock associated 1
SEQ_6	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes
SEQ_6	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes

Table 3: cis- regulatory elements of SEQ ID NO: 9

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
SEQ_9	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	17	35	(+)	0.750	0.859	caATTcaggtagcttaag
SEQ_9	P\$MIIG/MYBC1.01	Maize C1 myb-domain protein	0.92	21	35	(+)	1.000	0.928	ttcagGTAGcttaag

SEQ_9	P\$MADS/AGL3.01	AGL3, MADS Box protein	0.83	30	50	(-)	0.973	0.858	agccaCCAAtttagagcttaag
SEQ_9	P\$CAAT/CAAT.01	CCAAT-box in plant promoters	0.97	39	47	(-)	1.000	0.981	caCCAAtta
SEQ_9	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	44	60	(-)	1.000	0.995	tattacctAAAGccacc
SEQ_9	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	69	85	(+)	0.817	0.794	ctcagtTTGTaaatgta
SEQ_9	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	72	88	(+)	1.000	0.925	agtttgTAAAtgtagtt
SEQ_9	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	78	94	(+)	1.000	0.794	taaatgTAGTtaaaact
SEQ_9	P\$GTBX/SBF1.01	SBF-1	0.87	80	96	(+)	1.000	0.929	aatgtagTTAAaacttt
SEQ_9	P\$NCS1/NCS1.01	Nodulin consensus sequence 1	0.85	88	98	(-)	1.000	0.850	cAAAAGtttta
SEQ_9	P\$SALT/ALFIN1.01	Zinc-finger protein in alfalfa roots, regulates salt tolerance	0.93	93	107	(+)	1.000	0.948	cttttGTGGtgtaaa
SEQ_9	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0.75	97	113	(+)	0.776	0.793	tgtgggtAAATcatgt
SEQ_9	P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.85	97	113	(+)	0.968	0.857	tgtgggtTAAAtcatgt
SEQ_9	P\$CAAT/CAAT.01	CCAAT-box in plant promoters	0.97	121	129	(-)	1.000	0.979	aaCCAAtcg
SEQ_9	P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.85	121	137	(+)	1.000	0.866	cgattgGTTAataaaaa
SEQ_9	P\$SEF4/SEF4.01	Soybean embryo factor 4	0.98	129	139	(-)	1.000	0.985	acTTTTtatta
SEQ_9	P\$OPAQ/GCN4.01	GCN4, conserved in cereal seed storage protein gene promoters, similar to yeast GCN4 and vertebrate AP-1	0.81	141	157	(+)	1.000	0.813	gttgaTGAGttaaaaa
SEQ_9	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	158	174	(+)	1.000	0.805	caaaatTAGTtgcagtt
SEQ_9	P\$AHBP/BLR.01	Transcriptional repressor BELLRINGER	0.90	159	169	(+)	1.000	0.981	aaaATTAgttg
SEQ_9	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	165	181	(-)	0.750	0.844	taattTAACtgcaact
SEQ_9	P\$GTBX/SBF1.01	SBF-1	0.87	166	182	(+)	1.000	0.954	gttcagTTAAaattac
SEQ_9	P\$NACF/TANAC69.01	Wheat NACdomain DNA binding factor	0.68	169	191	(+)	1.000	0.707	gcagttaaaatTACGaatac
SEQ_9	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGA-MOUS-like 2	0.82	176	196	(-)	1.000	0.820	ggagcCCATgattcgtaattt

SEQ_9	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	196	206	(+)	1.000	0.897	ctATCTatatt	46
SEQ_9	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	196	212	(+)	0.777	0.905	ctatctaTATTttacat	
SEQ_9	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0.75	200	216	(-)	0.761	0.834	tgtaatgtAAAAatag	
SEQ_9	P\$DOFF/DOF2.01	Dof2 - single zinc finger transcription factor	0.98	211	227	(+)	1.000	0.988	attacaatAAAGctttt	
SEQ_9	P\$DOFF/DOF2.01	Dof2 - single zinc finger transcription factor	0.98	242	258	(+)	1.000	0.995	attacaatAAAGctata	
SEQ_9	P\$TBPF/TATA.02	Plant TATA box	0.90	251	265	(+)	1.000	0.913	aagcTATAtatcact	
SEQ_9	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	272	288	(-)	0.777	0.827	ttgtcttTATTtcagat	
SEQ_9	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	273	289	(+)	1.000	0.984	tctgaaatAAAGacaaa	
SEQ_9	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	280	292	(-)	1.000	0.940	gggtTTGTcttta	
SEQ_9	P\$TELO/RPBX.01	Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression	0.84	285	299	(+)	1.000	0.864	acaaaCCCTgaactc	
SEQ_9	P\$AHBP/ATHB1.01	Arabidopsis thaliana homeo box protein 1	0.90	307	317	(-)	1.000	0.989	ctaATTAtttc	
SEQ_9	P\$AHBP/HAHB4.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0.87	307	317	(+)	1.000	0.943	gaaataATTAg	
SEQ_9	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	317	335	(+)	0.954	0.972	gccaaACACttgattccaa	
SEQ_9	P\$MADS/SQUA.01	MADS-box protein SQUAMOSA	0.90	334	354	(-)	1.000	0.906	ggtcgctATTtgttctgttt	
SEQ_9	P\$ERSE/ERSE_I.01	ERSE I (ER stress-response element I)-like motif	0.79	343	361	(+)	1.000	0.799	caaatagcgacctaaCACG	
SEQ_9	P\$IBOX/GATA.01	Class I GATA factors	0.93	356	372	(+)	1.000	0.958	aacacGATAatagatag	
SEQ_9	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	362	372	(-)	1.000	0.853	ctATCTattat	
SEQ_9	P\$AHBP/BLR.01	Transcriptional repressor BELLRINGER	0.90	383	393	(-)	1.000	0.928	tatATTAtttt	
SEQ_9	P\$IBOX/IBOX.01	I-Box in rbcS genes and other light regulated genes	0.81	385	401	(+)	0.750	0.817	aataaTATAaggatcag	
SEQ_9	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	398	412	(-)	1.000	0.864	aaccaagaAATCtga	

SEQ_9	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	0.87	411	423	(-)	1.000	0.909	ttTACTgtttaaa
SEQ_9	P\$GAPB/GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	0.88	431	445	(+)	1.000	0.984	acacATGAagacaag
SEQ_9	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	448	466	(+)	1.000	0.815	aaAAATtatagattttaca
SEQ_9	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	452	462	(-)	1.000	0.912	aaATCTataat
SEQ_9	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	455	469	(-)	1.000	0.889	ttttgtaaAATCtat
SEQ_9	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0.75	455	471	(-)	0.761	0.841	tgttttgtAAAAAtctat
SEQ_9	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGA-MOUS-like 2	0.82	496	516	(-)	0.968	0.856	agaagCCAAcattggcaacga
SEQ_9	P\$MADS/AG.01	Agamous, required for normal flower development, similarity to SRF (human) and MCM (yeast) proteins	0.80	497	517	(+)	0.902	0.806	cgtTGCCaatgttggtcttctt
SEQ_9	P\$LFYB/LFY.01	Plant specific floral meristem identity gene LEAFY (LFY)	0.93	500	512	(+)	0.885	0.938	tGCCAatgttggc
SEQ_9	P\$DOFF/PBOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0.75	510	526	(-)	1.000	0.801	tgtgtgtgAAAGaagcc
SEQ_9	P\$SALT/ALFIN1.01	Zinc-finger protein in alfalfa roots, regulates salt tolerance	0.93	516	530	(-)	1.000	0.986	tttgtGTGGtggaaa
SEQ_9	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	521	541	(-)	0.838	0.781	taACGGtcatatttgtgtgt
SEQ_9	P\$WBXF/ERE.01	Elicitor response element	0.89	527	543	(+)	1.000	0.897	caaataTGACcgttaag
SEQ_9	P\$MYBL/ATMYB77.01	R2R3-type myb-like transcription factor (l-type binding site)	0.87	530	546	(+)	0.857	0.916	atatgaCCGTaagact
SEQ_9	P\$MSAE/MSA.01	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0.80	531	545	(-)	1.000	0.889	gtcttAACGgtcata
SEQ_9	P\$AHBP/HAHB4.01	Sunflower homeodomain leucine-zipper protein	0.87	558	568	(+)	1.000	0.916	tttataATTAc

SEQ_9	P\$GTBX/GT1.01	Hahb-4 GT1-Box binding factors with a trihelix DNA-binding domain	0.85	558	574	(-)	0.968	0.859	catgtaGTAAttataaa
SEQ_9	P\$OPAQ/O2_GC4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	564	580	(+)	1.000	0.823	attactACATggaatga
SEQ_9	P\$HMGF/HMG_4Y.01	High mobility group I/Y-like proteins	0.89	577	591	(-)	1.000	0.912	tataTATTttataca
SEQ_9	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	580	596	(-)	0.777	0.793	cgatctaTATAttttat
SEQ_9	P\$TBPFF/TATA.02	Plant TATA box	0.90	581	595	(-)	1.000	0.909	gataTATAtatttta
SEQ_9	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	586	596	(-)	1.000	0.882	cgATCTatata
SEQ_9	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	617	633	(-)	1.000	0.891	gagaaaTAAAtggtcga

Table 4: cis- regulatory elements of SEQ ID NO: 14

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
SEQ_14	P\$OCSE/OC4.01	OCS-like elements	0.69	7	27	(-)	1.000	0.702	caaagtgtgactatACGTttt
SEQ_14	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	31	47	(-)	1.000	0.997	catagtcaAAAGcacao
SEQ_14	P\$WBXF/WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	34	50	(+)	1.000	0.961	tgcttTTGActatgtgt
SEQ_14	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	43	59	(+)	1.000	0.833	ctatgtgTATCtgttcc
SEQ_14	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expres-	0.85	59	69	(-)	1.000	0.892	cgATCTataag

protein Pur-alpha									
SEQ_14	P\$DOFF/DOF3.01	0.99	231	247	(-)	1.000	0.997	agccttcaAAAGcctaa	50
SEQ_14	P\$SUCB/SUCROSE.01	0.81	238	256	(-)	0.750	0.836	aaAAAAcatagcctcaaa	
SEQ_14	P\$LEGB/LEGB.01	0.59	268	294	(-)	0.750	0.593	accttcgACATgatctaagaacaaaga	
SEQ_14	P\$LREM/ATCTA.01	0.85	274	284	(-)	1.000	0.921	tgATCTaagaa	
SEQ_14	P\$CARM/CARICH.01	0.78	286	304	(-)	1.000	0.816	acatttcAACACccttcgac	
SEQ_14	P\$L1BX/PDF2.01	0.85	308	324	(+)	1.000	0.903	atgtaaTAAAtgttatt	
SEQ_14	P\$OCSE/OC SL.01	0.69	314	334	(-)	0.807	0.714	gcagctgagtaataACATtta	
SEQ_14	P\$OCSE/OC SL.01	0.69	322	342	(+)	1.000	0.717	attactcagctgctACGTtta	
SEQ_14	P\$CCAF/CCA1.01	0.85	368	382	(-)	1.000	0.852	tttacataAATCtca	
SEQ_14	P\$DOFF/PBOX.01	0.75	372	388	(+)	0.761	0.751	atttatgtAAAAAtccat	
SEQ_14	P\$IBOX/GATA.01	0.93	399	415	(-)	1.000	0.973	aaatgGATAagattgat	
SEQ_14	P\$MYBS/HVMCB1.01	0.93	402	418	(+)	1.000	0.963	aatcttATCCattttct	
SEQ_14	P\$AHBP/WUS.01	0.94	420	430	(+)	1.000	0.963	ctgatTAATct	
SEQ_14	P\$AHBP/WUS.01	0.94	421	431	(-)	1.000	0.963	cagatTAATca	
SEQ_14	P\$ABRE/ABRE.01	0.82	428	444	(-)	1.000	0.850	taattgcACGTgcaga	
SEQ_14	P\$MYBL/MYBPH3.02	0.76	437	453	(+)	1.000	0.782	tgcaatTAGTttgatca	
SEQ_14	P\$TEFB/TEF1.01	0.76	450	470	(-)	0.838	0.779	ccATGGctaataattgttgat	
SEQ_14	P\$GBOX/UPRE.01	0.86	493	513	(-)	0.767	0.862	cttcgtCCAAgtcaacataag	

SEQ_14	P\$RAV5/RAV1-5.01	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	0.96	493	503	(-)	1.000	0.960	gtcAACATAag
SEQ_14	P\$GBOX/BZIP911.01	bZIP transcription factor from Antirrhinum majus	0.77	494	514	(+)	1.000	0.833	ttatgtTGACttggacgaaga
SEQ_14	P\$OPAQ/O2.01	Opaque-2 regulatory protein	0.87	495	511	(-)	0.852	0.895	tcgtccaagTCAAcata
SEQ_14	P\$DOFF/DOF2.01	Dof2 - single zinc finger transcription factor	0.98	544	560	(-)	1.000	1.000	tatttattAAAGcaaac
SEQ_14	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	549	565	(+)	1.000	0.853	ctttaaTAAAtataagt
SEQ_14	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	559	577	(-)	0.750	0.816	caCAATcattctactata
SEQ_14	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	566	576	(+)	0.829	0.904	agaATGAttgt
SEQ_14	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	566	576	(-)	0.936	0.977	acaATCATtct
SEQ_14	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	575	587	(+)	1.000	0.926	gtgtTTGTcttct
SEQ_14	P\$GTBX/SBF1.01	SBF-1	0.87	581	597	(-)	1.000	0.885	tctgtgaTTAAgaagac
SEQ_14	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	583	593	(+)	1.000	0.963	cttctTAATca
SEQ_14	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	0.87	590	602	(-)	1.000	0.876	aaTACTcttgat
SEQ_14	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	610	620	(-)	1.000	0.963	aacctTAAATct
SEQ_14	P\$OCSE/OC SL.01	OCS-like elements	0.69	614	634	(+)	1.000	0.699	taaggttgaaatgaACGTcgt
SEQ_14	P\$EINL/TEIL.01	TEIL (tobacco EIN3-like)	0.92	624	632	(+)	0.964	0.932	aTGAAcgtc
SEQ_14	P\$GBOX/TGA1.01	Arabidopsis leucine zipper protein TGA1	0.90	633	653	(+)	1.000	0.903	gtaaaaTGACggttatgctcg
SEQ_14	P\$MYBL/ATMYB77.01	R2R3-type myb-like transcription factor (l-type binding site)	0.87	636	652	(+)	1.000	0.970	aatgaCGGTtatgctc
SEQ_14	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	639	659	(+)	0.838	0.778	tgACGGttatgctcgtagag

SEQ_14	P\$PREM/MGPROTORE.01	Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light- mediated induction	0.77	641	671	(-)	1.000	0.792	taagCGACgattctctcagcagcataaccgt
SEQ_14	P\$IBOX/GATA.01	Class I GATA factors	0.93	668	684	(+)	1.000	0.982	cttacGATAaggacgaa
SEQ_14	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	690	706	(+)	1.000	0.936	atttgattGTTAtcagg
SEQ_14	P\$MYBL/NTMYBAS1.01	Anther-specific myb gene from tobacco	0.96	701	717	(+)	1.000	0.967	atcagggtGTTAaaagt
SEQ_14	P\$GTBX/SBF1.01	SBF-1	0.87	703	719	(+)	1.000	0.921	cagggtgTTAAaagtgg
SEQ_14	P\$NCS1/NCS1.01	Nodulin consensus sequence 1	0.85	711	721	(+)	1.000	0.861	tAAAAGttgag
SEQ_14	P\$AREF/ARE.01	Auxin Response Element	0.93	715	727	(-)	1.000	0.932	gttTGTCtcaact
SEQ_14	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	717	729	(-)	1.000	0.927	tcgtTTGTctcaa
SEQ_14	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	747	765	(-)	0.954	0.972	acgtaACACctgttttagtc
SEQ_14	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	748	766	(+)	0.863	0.954	actaaACAGgtgttacggt
SEQ_14	P\$GBOX/HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.83	753	773	(-)	1.000	0.857	aatgtgaaACGTaacacctgt
SEQ_14	P\$GTBX/GT3A.01	Trihelix DNA-binding factor GT-3a	0.83	753	769	(+)	1.000	0.858	acagggtGTTAcgtttca
SEQ_14	P\$ABRE/ABRE.01	ABA response elements	0.82	755	771	(+)	1.000	0.820	agggtttACGTttcaca
SEQ_14	P\$NACF/TANAC69.01	Wheat NACdomain DNA binding factor Plant specific floral meristem identity gene LEAFY (LFY)	0.68	755	777	(-)	0.812	0.708	aaccaatgtgaAACGtaaacacct
SEQ_14	P\$LFYB/LFY.01	CCAAT-box in plant promoters	0.93	765	777	(-)	0.914	0.947	aACCAatgtgaaa
SEQ_14	P\$CAAT/CAAT.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.97	769	777	(-)	1.000	0.982	aaCCAAtgt
SEQ_14	P\$GTBX/GT1.01	Transcriptional repressor BELL- RINGER	0.85	790	806	(+)	1.000	0.906	cttgaaGTTActctatt
SEQ_14	P\$AHBP/BLR.01	Myb-like protein of Petunia hybrida	0.90	793	803	(+)	0.826	0.914	gaaGTTActct
SEQ_14	P\$MYBL/MYBPH3.01	Elicitor response element	0.80	802	818	(-)	1.000	0.828	ttggaccgGTTAaatag
SEQ_14	P\$WBXF/ERE.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.89	824	840	(-)	1.000	0.894	ttaaccTGACcggttgg
SEQ_14	P\$GTBX/GT1.01		0.85	830	846	(+)	1.000	0.854	ggtcagGTTAacaaaac

SEQ_14	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	833	849	(-)	1.000	0.930	agtgtttGTTAacctg
SEQ_14	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	845	861	(+)	1.000	0.910	acactgaaGTTAgccgc
SEQ_14	P\$GCCF/GCC-BOX.01	GCC-box, ethylene-responsive element (ERE)	0.86	853	865	(+)	1.000	1.000	gttAGCCgccaac
SEQ_14	P\$GBOX/HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.83	867	887	(+)	1.000	0.842	cgcttattACGTaaacggtag
SEQ_14	P\$NACF/TANAC69.01	Wheat NACdomain DNA binding factor	0.68	868	890	(-)	1.000	0.728	tggctaccgttTACGtaataagc
SEQ_14	P\$OCSE/OCSL.01	OCS-like elements	0.69	872	892	(-)	1.000	0.710	cgtggtaccggttACGTaat
SEQ_14	P\$MSAE/MSA.01	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0.80	875	889	(+)	1.000	0.827	acgtaAACGgtagcc
SEQ_14	P\$GBOX/GBF1.01	bZIP protein G-Box binding factor 1	0.94	881	901	(-)	1.000	0.943	tgctcgaaACGTggctaccgt
SEQ_14	P\$GBOX/HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	0.88	882	902	(+)	1.000	0.943	cggtagcCACGtttcgagcac
SEQ_14	P\$MYCL/MYCRS.01	Myc recognition sequences	0.93	882	900	(-)	1.000	0.936	gctcgaaACGTggctaccg
SEQ_14	P\$NACF/TANAC69.01	Wheat NACdomain DNA binding factor	0.68	883	905	(-)	0.812	0.717	gcagtgctcgaaACGtggtacc
SEQ_14	P\$ABRE/ABRE.01	ABA response elements	0.82	884	900	(-)	1.000	0.864	gctcgaaACGTggctac
SEQ_14	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	912	924	(+)	1.000	0.921	taatTTGTcttca
SEQ_14	P\$LEGB/LEGB.01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.59	975	1001	(-)	0.750	0.657	tcactagCCTTgcatgcgaatcagtag
SEQ_14	P\$LEGB/RV.01	RY and Sph motifs conserved in seed-specific promoters	0.87	978	1004	(+)	1.000	0.899	ctgattcgCATGcaaggctagtgacac

Table 5: cis- regulatory elements of SEQ ID NO: 16

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
-----------	---------------	---------------------	--------------	------------	----------	--------	-----------	-------------	----------

SEQ_16	P\$OCSE/OCSL.01	OCS-like elements	0.69	7	27	(-)	1.000	0.702	caaagtgactatACGTttt
SEQ_16	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	31	47	(-)	1.000	0.997	catagtcaAAAGcaca
SEQ_16	P\$WBXF/WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	34	50	(+)	1.000	0.961	tgcttTTGActatgtgt
SEQ_16	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	43	59	(+)	1.000	0.833	ctatgtgTATCtgttcc
SEQ_16	P\$LREM/ATCTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	59	69	(-)	1.000	0.892	cgATCTataag
SEQ_16	P\$MADS/AG.01	Agamous, required for normal flower development, similarity to SRF (human) and MCM (yeast) proteins	0.80	92	112	(-)	0.962	0.820	tagTTCccaaaccggttccaa
SEQ_16	P\$MIIG/MYBC1.01	Maize C1 myb-domain protein	0.92	104	118	(-)	1.000	0.935	tgttgGTAGttccca
SEQ_16	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein	0.82	108	128	(+)	0.968	0.828	aactaCCAacacaagcaatgc
SEQ_16	P\$GAGA/BPC.01	AGAMOUS-like 2	1.00	129	153	(+)	1.000	1.000	tcttagAGAGagaaaaagagtaataa
SEQ_16	P\$GAGA/BPC.01	Basic pentacycstine proteins	1.00	131	155	(+)	1.000	1.000	ttagagAGAGaaaaagagtaataaaa
SEQ_16	P\$DOFF/PBF.01	Basic pentacycstine proteins	0.97	134	150	(+)	1.000	0.988	gagagagaAAAGagtaa
SEQ_16	P\$SPF1/SP8BF.01	PBF (MPBF)	0.87	139	151	(-)	1.000	0.919	atTACTcttttct
SEQ_16	P\$NCS1/NCS1.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	0.85	140	150	(+)	1.000	0.855	gAAAAGagtaa
SEQ_16	P\$DOFF/PBF.01	Nodulin consensus sequence 1	0.97	145	161	(+)	1.000	0.990	gagtaataAAAGagagg
SEQ_16	P\$NCS1/NCS1.01	PBF (MPBF)	0.85	163	173	(+)	1.000	0.862	gAAAAGgttttg
SEQ_16	P\$WBXF/WRKY.01	Nodulin consensus sequence 1	0.92	166	182	(-)	1.000	0.942	agagtTTGAcaaaactt
SEQ_16	P\$IDDF/ID1.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	167	179	(+)	1.000	0.935	agttTTGTcaaac
SEQ_16	P\$OCSE/OCSL.01	Maize INDETERMINATE1 zinc finger protein	0.69	194	214	(-)	0.769	0.708	gaaagtagatcttACTTtct
SEQ_16	P\$OCSE/OCSL.01	OCS-like elements	0.69	195	215	(+)	0.769	0.733	gaaagtaagatctaACTTtct
SEQ_16	P\$AGP1/AGP1.01	OCS-like elements	0.91	199	209	(-)	1.000	0.915	ttaGATCttac
		AG-motif binding protein 1							

SEQ_16	P\$AGP1/AGP1.01	AG-motif binding protein 1 Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.91	200	210	(+)	1.000	0.983	taaGATCtaac
SEQ_16	P\$LREM/ATCTA.01	Anther-specific myb gene from tobacco	0.85	202	212	(+)	1.000	0.910	agATCTaactt
SEQ_16	P\$MYBL/NTMYBAS1.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.96	228	244	(+)	1.000	0.976	tttgggatGTTAggctt
SEQ_16	P\$TELO/ATPURA.01	Dof3 - single zinc finger transcription factor	0.85	233	247	(-)	0.750	0.868	caaaAGCCtaacatc
SEQ_16	P\$DOFF/DOF3.01	Sequence motif from the promoters of different sugar-responsive genes	0.99	237	253	(-)	1.000	0.997	agccttcaAAAGcctaa
SEQ_16	P\$SUCB/SUCROSE.01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.81	244	262	(-)	0.750	0.836	aaAAAAcatagcctcaaa
SEQ_16	P\$LEGB/LEGB.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.59	274	300	(-)	0.750	0.593	accttcgACATgatctaagaacaaaga
SEQ_16	P\$LREM/ATCTA.01	CA-rich element	0.85	280	290	(-)	1.000	0.921	tgATCTaagaa
SEQ_16	P\$CARM/CARICH.01	CA-rich element	0.78	292	310	(-)	1.000	0.787	acacttcAACAccttcgac
SEQ_16	P\$OCSE/OCSE.01	Protodermal factor 2	0.78	300	318	(-)	1.000	0.829	taataacatttattACGTcca
SEQ_16	P\$L1BX/PDF2.01	OCS-like elements	0.69	311	331	(-)	1.000	0.695	acgtaaTAAAtgttatt
SEQ_16	P\$OCSE/OCSE.01	OCS-like elements	0.85	314	330	(+)	1.000	0.903	gcagctgagtaataACATtta
SEQ_16	P\$OCSE/OCSE.01	Circadian clock associated 1	0.69	320	340	(-)	0.807	0.714	attactcagctgctACGTtta
SEQ_16	P\$CCAF/CCA1.01	Sequence motif from the promoters of different sugar-responsive genes	0.85	328	348	(+)	1.000	0.717	tttacataAATCtca
SEQ_16	P\$SUCB/SUCROSE.01	Class I GATA factors	0.81	374	388	(-)	1.000	0.852	
SEQ_16	P\$IBOX/GATA.01	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	0.93	386	404	(+)	1.000	0.818	aaAAATccgcaatcttatc
SEQ_16	P\$MYBS/HVMCB1.01	Homeodomain protein WUSCHEL	0.93	393	409	(-)	1.000	0.973	aaatgGATAagattgcg
SEQ_16	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.93	396	412	(+)	1.000	0.963	aatcttATCCattttct
SEQ_16	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	414	424	(+)	1.000	0.963	ctgatTAATct
SEQ_16	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	415	425	(-)	1.000	0.963	cagatTAATca

SEQ_16	P\$ABRE/ABRE.01	ABA response elements	0.82	422	438	(-)	1.000	0.850	taattgcACGTgcaga
SEQ_16	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	431	447	(+)	1.000	0.782	tgcaatTAGTttgatca
SEQ_16	P\$HMGF/HMG_IY.01	High mobility group I/Y-like proteins	0.89	441	455	(-)	1.000	0.907	atatTATTtgatcaa
SEQ_16	P\$AHBP/BLR.01	Transcriptional repressor BELLRINGER	0.90	446	456	(-)	1.000	0.928	aatATTAttg
SEQ_16	P\$RAV5/RAV1-5.01	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	0.96	487	497	(-)	1.000	0.960	gtcAACAtaag
SEQ_16	P\$OPAQ/O2.01	Opaque-2 regulatory protein	0.87	489	505	(-)	0.852	0.895	tcctccaagTCAAcata
SEQ_16	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	538	554	(-)	1.000	0.994	tattataAAAGGcaaac
SEQ_16	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	538	554	(-)	1.000	0.859	tatttaTAAAGgcaaac
SEQ_16	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	543	559	(+)	1.000	0.834	cttttaTAAAtataagt
SEQ_16	P\$TBPF/TATA.01	Plant TATA box	0.88	543	557	(+)	1.000	0.971	ctttTATAaatataa
SEQ_16	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	553	571	(-)	0.750	0.816	caCAATcattctactata
SEQ_16	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	560	570	(+)	0.829	0.904	agaATGAttgt
SEQ_16	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	560	570	(-)	0.936	0.977	acaATCAttct
SEQ_16	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	569	581	(+)	1.000	0.926	gtgtTTGTcttct
SEQ_16	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	0.87	586	598	(-)	1.000	0.876	aaTACTcttgat
SEQ_16	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	606	616	(-)	1.000	0.963	aacctTAAAtct
SEQ_16	P\$GBOX/TGA1.01	Arabidopsis leucine zipper protein TGA1	0.90	625	645	(+)	1.000	0.903	gtaaaaTGACgggttatgctcg
SEQ_16	P\$MYBL/ATMYB77.01	R2R3-type myb-like transcription factor (I-type binding site)	0.87	628	644	(+)	1.000	0.970	aaatgaCGGTtatgctc
SEQ_16	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	631	651	(+)	0.838	0.778	tgACGGttatgctcgtagag
SEQ_16	P\$IBOX/GATA.01	Class I GATA factors	0.93	660	676	(+)	1.000	0.950	cttgcGATAaggacgaa
SEQ_16	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	682	698	(+)	1.000	0.986	atttggttGTTAtcagg

SEQ_16	P\$MIIG/PALBOXL.01	Cis-acting element conserved in various PAL and 4CL promoters	0.80	693	707	(+)	0.750	0.818	atcaggttGTTGaaa
SEQ_16	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	693	709	(+)	0.817	0.777	atcaggTTGTtgaaga
SEQ_16	P\$NCS2/NCS2.01	Nodulin consensus sequence 2	0.79	705	719	(-)	0.750	0.801	gtttgtCTCAtcctt
SEQ_16	P\$OPAQ/GCN4.01	GCN4, conserved in cereal seed storage protein gene promoters, similar to yeast GCN4 and vertebrate AP-1	0.81	705	721	(+)	1.000	0.846	aaagaTGAGacaaaaga
SEQ_16	P\$AREF/ARE.01	Auxin Response Element	0.93	707	719	(-)	1.000	0.951	gttTGTCtcattct
SEQ_16	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.92	709	721	(-)	1.000	0.927	tcgtTTGTctcat
SEQ_16	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	739	757	(-)	0.954	0.972	acgtaACACcctgttagtc
SEQ_16	P\$MYCL/ICE.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.95	740	758	(+)	0.863	0.954	actaaACAGggtgttacgtt
SEQ_16	P\$GBOX/HBP1B.01	Trihelix DNA-binding factor GT-3a	0.83	745	765	(-)	1.000	0.857	aatgtaaACGTaacactgt
SEQ_16	P\$GTBX/GT3A.01	ABA response elements	0.83	745	761	(+)	1.000	0.858	acaggtGTTAcgtttca
SEQ_16	P\$ABRE/ABRE.01	Wheat NACdomain DNA binding factor	0.82	747	763	(+)	1.000	0.820	agggttACGTttcaca
SEQ_16	P\$NACF/TANAC69.01	Plant specific floral meristem identity gene LEAFY (LFY)	0.68	747	769	(-)	0.812	0.708	aaccaatgaaAACGtaaacact
SEQ_16	P\$LFYB/LFY.01	CCAAT-box in plant promoters	0.93	757	769	(-)	0.914	0.947	aACCAatgtgaaa
SEQ_16	P\$CAAT/CAAT.01	Myb-like protein of Petunia hybrida	0.97	761	769	(-)	1.000	0.982	aaCCAAtgt
SEQ_16	P\$MYBL/MYBPH3.01	Maize INDETERMINATE1 zinc finger protein Ethylene-responsive elements (ERE) and jasmonate- and elicitor-responsive elements (JERE)	0.80	794	810	(-)	1.000	0.818	cgggaccgGTTAaatag
SEQ_16	P\$IDDF/ID1.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.92	809	821	(-)	1.000	0.929	agttTTGTcaacc
SEQ_16	P\$GCCF/ERE_JERE.01	Wheat NACdomain DNA binding factor OCS-like elements	0.85	826	838	(+)	1.000	0.907	aaaaggCGCCaac
SEQ_16	P\$GBOX/HBP1B.01	M-phase-specific activators (NtmybA1, NtmybA2, NtmybA3)	0.83	840	860	(+)	1.000	0.858	cgttgttACGTaaacgtag
SEQ_16	P\$NACF/TANAC69.01		0.68	841	863	(-)	1.000	0.764	tggctaccgttTACGtaacaagc
SEQ_16	P\$OCSE/OCSE.01		0.69	845	865	(-)	1.000	0.710	cgtggctaccggttACGTaac
SEQ_16	P\$MSAE/MSA.01		0.80	848	862	(+)	1.000	0.827	acgtaAACGgtagcc

SEQ_16	P\$GBOX/GBF1.01	NtmybA2, NtmybB) bZIP protein G-Box binding factor 1 HBP-1a, suggested to be involved in the cell cycle-dependent expression	0.94	854	874	(-)	1.000	0.944	tgctcaaaACGTggctaccgt
SEQ_16	P\$GBOX/HBP1A.01	Wheat NACdomain DNA binding factor	0.88	855	875	(+)	1.000	0.914	cggtagcCACGttttgagcac
SEQ_16	P\$NACF/TANAC69.01	ABA response elements	0.68	856	878	(-)	0.812	0.712	gcagtgctcaaAACGtggtacc
SEQ_16	P\$ABRE/ABRE.01	Myc recognition sequences	0.82	857	873	(-)	1.000	0.860	gctcaaaACGTggctac
SEQ_16	P\$MYCL/MYCRS.01	Maize INDETERMINATE1 zinc finger protein	0.93	872	890	(+)	0.863	0.946	gcactgcATGTgctaattt
SEQ_16	P\$IDDF/ID1.01	TCP class I transcription factor (Arabidopsis)	0.92	885	897	(+)	1.000	0.921	taatTTGTcttca
SEQ_16	P\$TCPF/ATTCP20.01	Myb-like protein of Petunia hybrida	0.94	924	936	(-)	1.000	0.947	ttaaGCCCcaagtg
SEQ_16	P\$MYBL/MYBPH3.02	Homeodomain protein WUSCHEL	0.76	940	956	(-)	1.000	0.781	attaatTAGTtccacga
SEQ_16	P\$AHBP/WUS.01	HDZip class I protein ATHB5	0.94	948	958	(+)	1.000	1.000	ctaataATATga
SEQ_16	P\$AHBP/ATHB5.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0.89	952	962	(+)	0.829	0.902	ttaATGAttcg
SEQ_16	P\$AHBP/HAHB4.01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.87	952	962	(-)	1.000	0.979	cgaatcATTAA
SEQ_16	P\$LEGB/LEGB.01	RY and Sph motifs conserved in seed-specific promoters	0.59	952	978	(-)	0.750	0.657	tcactagCCTTgcatgcgaatcataa
SEQ_16	P\$LEGB/RY.01		0.87	955	981	(+)	1.000	0.899	atgattcgCATGcaaggctagtgacac

Table 6: cis- regulatory elements of SEQ ID NO: 22

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
SEQ_22	P\$AHBP/BLR.01	Transcriptional repressor BELL-RINGER	0.90	4	14	(+)	0.826	0.936	gaaGTTAttag
SEQ_22	P\$MYBL/CARE.01	CAACTC regulatory elements, GA-inducible	0.83	27	43	(-)	1.000	0.875	gttggctAGTTgtaagt

SEQ_22	P\$WBXF/WRKY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	49	65	(-)	1.000	0.975	catgtTTGAaccttaca
SEQ_22	P\$OPAQ/O2_GC4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	56	72	(-)	1.000	0.853	ttgtAACATgtttgac
SEQ_22	P\$GTBX/GT3A.01	Trihelix DNA-binding factor GT-3a	0.83	59	75	(+)	1.000	0.867	aaacatGTTAcaaaactc
SEQ_22	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	61	77	(-)	0.817	0.798	ttgagtTTGTaacatgt
SEQ_22	P\$HMGF/HMG_IY.02	High mobility group I/Y-like protein isolated from pea	1.00	72	86	(-)	1.000	1.000	ctttTATTtttgagt
SEQ_22	P\$PSRE/GAAA.01	GAAA motif involved in pollen specific transcriptional activation	0.83	81	97	(+)	1.000	0.879	taaaaGAAAcagtgag
SEQ_22	P\$LFYB/LFY.01	Plant specific floral meristem identity gene LEAFY (LFY)	0.93	85	97	(-)	1.000	0.969	cTCCActgtttct
SEQ_22	P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	0.85	100	116	(-)	1.000	0.904	gttcagGTTActcgatt
SEQ_22	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	109	127	(-)	0.954	0.972	atcaaaACACctgttcaggt
SEQ_22	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	110	128	(+)	0.863	0.954	cctgaACAGgtgtttgac
SEQ_22	P\$CARM/CARICH.01	CA-rich element	0.78	112	130	(-)	1.000	0.809	ttgatcaAACACctgttca
SEQ_22	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	140	156	(-)	1.000	0.781	aagagaTAGTgacacac
SEQ_22	P\$MYBS/ZMMRP1.01	Zea mays MYB-related protein 1 (transfer cell specific)	0.79	142	158	(+)	1.000	0.824	gtgtcacTATCtcttgg
SEQ_22	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	169	185	(+)	1.000	0.984	acacaaatAAAGaccct
SEQ_22	P\$MIIG/PALBOXL.01	Cis-acting element conserved in various PAL and 4CL promoters	0.80	188	202	(-)	0.785	0.801	agcagcttGGTTagg
SEQ_22	P\$CAAT/CAAT.01	CCAAT-box in plant promoters	0.97	204	212	(+)	1.000	0.985	atCCAAatcc
SEQ_22	P\$OPAQ/O2_GC4.01	Recognition site for BZIP transcription factors that belong to the group	0.81	206	222	(-)	0.829	0.819	tgtgtgACTTggattgg

of Opaque-2 like proteins									
SEQ_22	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	229	245	(+)	1.000	0.921	cagctaTAAAtgaaaca
SEQ_22	P\$TBPF/TATA.02	Plant TATA box	0.90	229	243	(+)	1.000	0.935	cagcTATAaatgaaa
SEQ_22	P\$MYBL/MYBPH3.01	Myb-like protein of Petunia hybrida	0.80	250	266	(-)	1.000	0.808	attcatctGTTAaagtt
SEQ_22	P\$GAPB/GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	0.88	257	271	(+)	1.000	0.964	acagATGAatactag
SEQ_22	P\$HEAT/HSE.01	Heat shock element	0.81	284	298	(-)	1.000	0.811	aggagacactAGAAc
SEQ_22	P\$AREF/ARE.01	Auxin Response Element	0.93	288	300	(+)	1.000	0.961	tagTGTCTctctca
SEQ_22	P\$NCS2/NCS2.01	Nodulin consensus sequence 2	0.79	288	302	(+)	0.750	0.819	tagtgtCTCCtcatt
SEQ_22	P\$ROOT/RHE.01	Root hair-specific element with a 2-nucleotid spacer between left part (LP) and right part (RP)	0.77	298	322	(-)	1.000	0.852	atogtagcttgaattCACGtaatga
SEQ_22	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	317	333	(-)	1.000	0.775	ttgagaTAGTgatcgta
SEQ_22	P\$MYBL/CARE.01	CAACTC regulatory elements, GA-inducible	0.83	326	342	(-)	1.000	0.838	atgtaggAGTTgagata
SEQ_22	P\$LEGB/LEGB.01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.59	357	383	(+)	0.750	0.595	tacaaaaCTATgcacaaaaaacaagc
SEQ_22	P\$EINL/TEIL.01	TEIL (tobacco EIN3-like) Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.92	380	388	(-)	1.000	0.934	aTGTAgtt
SEQ_22	P\$LREM/ATCTA.01	ABA insensitive protein 4 (ABI4)	0.85	385	395	(+)	1.000	0.931	acATCTaatac
SEQ_22	P\$CE1F/ABI4.01	Myb-like protein of Petunia hybrida	0.87	432	444	(+)	1.000	0.872	caatCACCGgtcga
SEQ_22	P\$MYBL/MYBPH3.01	Oryza sativa CaM-binding transcription factor	0.80	450	466	(+)	1.000	0.849	aggattcaGTTAattga
SEQ_22	P\$CGCG/OSCBT.01	Root hair-specific element with a 3-nucleotid spacer between left part (LP) and right part (RP)	0.78	475	491	(-)	0.817	0.796	cttCGAGtttgatcgga
SEQ_22	P\$ROOT/RHE.02		0.77	486	510	(+)	1.000	0.801	tcgaagactggtgagCACGaggacg

SEQ_22	P\$NCS2/NCS2.01	Nodulin consensus sequence 2	0.79	525	539	(-)	0.750	0.823	tggtgtATCTTcgag
SEQ_22	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	539	553	(+)	1.000	0.859	aagcaagaAATCtac
		Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes							
SEQ_22	P\$LREM/ATCTA.01	Arabidopsis thaliana meristem layer 1	0.85	546	556	(+)	1.000	0.869	aaATCTactga
SEQ_22	P\$L1BX/ATML1.02	Transcriptional repressor BELL-RINGER	0.76	563	579	(-)	0.890	0.767	cgcCAATaacttcagga
SEQ_22	P\$AHBP/BLR.01	Homeodomain glabrous 9	0.90	567	577	(+)	0.826	0.936	gaaGTTAttgg
SEQ_22	P\$L1BX/HDG9.01	Homeodomain glabrous 9	0.77	595	611	(-)	0.796	0.777	ccgaaaTTAAttcggat
SEQ_22	P\$L1BX/HDG9.01	Transcriptional repressor BELL-RINGER	0.77	597	613	(+)	0.750	0.796	ccgaatTAATttcgggg
SEQ_22	P\$AHBP/BLR.01	Class I GATA factors	0.90	599	609	(-)	1.000	1.000	gaaATTAattc
SEQ_22	P\$IBOX/GATA.01	Nodulin consensus sequence 1	0.93	614	630	(+)	1.000	0.961	aaaaaGATAaattagat
SEQ_22	P\$NCS1/NCS1.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	614	624	(+)	1.000	0.948	aaaaAAgataaa
		ABA (abscisic acid) inducible transcriptional activator							
SEQ_22	P\$LREM/ATCTA.01	TCP class I transcription factor (Arabidopsis)	0.85	622	632	(-)	1.000	0.921	gtATCTaattt
SEQ_22	P\$ABRE/ABF1.01	OCS-like elements	0.79	657	673	(+)	0.750	0.797	aagaaACAGgttggaat
SEQ_22	P\$TCPF/ATTCP20.01	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0.94	670	682	(-)	1.000	0.949	tccaGCCCaattg
SEQ_22	P\$OCSE/OCSL.01	Myb-like protein of Petunia hybrida MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	0.69	715	735	(+)	0.807	0.693	aaaaaaaaacggataACATatt
SEQ_22	P\$MSAE/MSA.01	Class I GATA factors	0.80	716	730	(+)	1.000	0.851	aaaaaAACGgataac
SEQ_22	P\$MYBL/MYBPH3.01		0.80	717	733	(+)	0.750	0.845	aaaaaacgGATAacata
SEQ_22	P\$MYBS/MYBST1.01		0.90	717	733	(-)	1.000	0.962	tatgtATCCgttttt
SEQ_22	P\$IBOX/GATA.01		0.93	720	736	(+)	1.000	0.935	aaacgGATAacatattt

SEQ_22	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	722	738	(-)	1.000	0.920	ataaatGTTAtccgt
SEQ_22	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	764	778	(-)	0.757	0.857	aaaaagaaAATAtct
SEQ_22	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	771	787	(+)	1.000	0.914	ttcttttGTTAggaaa
SEQ_22	P\$SEF4/SEF4.01	Soybean embryo factor 4	0.98	772	782	(+)	1.000	0.981	tcTTTTtgta
SEQ_22	P\$HEAT/HSE.01	Heat shock element	0.81	783	797	(+)	1.000	0.873	ggaaaaatttAGAAa
SEQ_22	P\$HMGF/HMG_IY.01	High mobility group I/Y-like proteins	0.89	790	804	(-)	1.000	0.895	ccatTATTttctaaa
		Sequence motif from the promoters							
SEQ_22	P\$SUCB/SUCROSE.01	of different sugar-responsive genes	0.81	794	812	(+)	1.000	0.851	gaAAATaatgaaataaa
SEQ_22	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	795	805	(-)	1.000	0.918	tccATTAtttt
		S1F, site 1 binding factor of spinach							
SEQ_22	P\$GTBX/S1F.01	rps1 promoter	0.79	797	813	(+)	1.000	0.841	aataATGGAaataaat
SEQ_22	P\$GTBX/SBF1.01	SBF-1	0.87	801	817	(+)	1.000	0.908	atgaaaTTAAatagcg
SEQ_22	P\$L1BX/HDG9.01	Homeodomain glabrous 9	0.77	803	819	(+)	1.000	0.822	ggaaatTAAAtagcgat
SEQ_22	P\$GTBX/GT3A.01	Trihelix DNA-binding factor GT-3a	0.83	817	833	(+)	1.000	0.843	gattatGTTAcaagata
SEQ_22	P\$OCSE/OC SL.01	OCS-like elements	0.69	819	839	(+)	0.807	0.707	ttatgttacaagatACGAtca
		Type-B response regulator (ARR10),							
		member of the GARP-family of plant							
SEQ_22	P\$GARP/ARR10.01	myb-related DNA binding motifs	0.97	829	837	(+)	1.000	0.985	AGATacgat
		Root hair-specific element with a 3-							
		nucleotid spacer between left part							
SEQ_22	P\$ROOT/RHE.02	(LP) and right part (RP)	0.77	838	862	(-)	0.750	0.777	tagcatttgcactgCCCGatgctg
SEQ_22	P\$NCS1/NCS1.01	Nodulin consensus sequence 1	0.85	877	887	(-)	0.804	0.899	aAAGgatcaa
SEQ_22	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	882	898	(+)	1.000	0.926	ccttttgGTTAtctcc
SEQ_22	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	903	919	(+)	1.000	0.785	gacaatTAGTttaggat
		Storekeeper (STK), plant specific							
		DNA binding protein important for							
		tuber-specific and sucrose-inducible							
SEQ_22	P\$STKM/STK.01	gene expression	0.85	904	918	(-)	1.000	0.865	tccTAAActaattgt
		Arabidopsis Telo-box interacting							
		protein related to the conserved							
SEQ_22	P\$TELO/ATPURA.01	animal protein Pur-alpha	0.85	909	923	(-)	0.750	0.867	caaaATCCtaaacta

SEQ_22	P\$AHBP/BLR.01	Transcriptional repressor BELL-RINGER	0.90	929	939	(+)	1.000	0.930	tatATTAatac
SEQ_22	P\$IBOX/GATA.01	Class I GATA factors	0.93	935	951	(-)	1.000	0.932	tgctgGATAatagatt
SEQ_22	P\$SPF1/SP8BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	0.87	935	947	(+)	1.000	0.995	aaTACTattatcc
SEQ_22	P\$MYBS/MYBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	0.90	938	954	(+)	1.000	0.943	actattATCCgacaaca
SEQ_22	P\$IDDF/ID1.01	Maize INDETERMINATE1 zinc finger protein	0.92	944	956	(-)	1.000	0.922	agtgTTGTcggat
SEQ_22	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	946	964	(-)	0.954	0.972	ctgaaACAAggtgtgtcgg
SEQ_22	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	947	965	(+)	0.954	0.984	cgacaACACtgtttcagc
SEQ_22	P\$GTBX/GT3A.01	Trihelix DNA-binding factor GT-3a	0.83	968	984	(-)	1.000	0.839	aaaaatGTTAaaaataag
SEQ_22	P\$MYBL/MYBPH3.01	Myb-like protein of Petunia hybrida	0.80	970	986	(-)	1.000	0.807	caaaaaatGTTAaaata
SEQ_22	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	975	989	(-)	0.766	0.872	aaacaaaaAATGtta
SEQ_22	P\$GTBX/S1F.01	S1F, site 1 binding factor of spinach rps1 promoter	0.79	997	1013	(-)	1.000	0.821	gctgATGGaagaagaa
SEQ_22	P\$GTBX/SBF1.01	SBF-1	0.87	1016	1032	(+)	1.000	0.875	tttctttTAAaaaatt
SEQ_22	P\$STKM/STK.01	Storekeeper (STK), plant specific							
SEQ_22	P\$NCS1/NCS1.01	DNA binding protein important for tuber-specific and sucrose-inducible gene expression	0.85	1021	1035	(+)	1.000	0.881	tttTAAaaaattgaa
SEQ_22	P\$GTBX/SBF1.01	Nodulin consensus sequence 1	0.85	1039	1049	(-)	1.000	0.852	aaaaAgttaaa
SEQ_22	P\$OPAQ/O2.01	SBF-1	0.87	1042	1058	(-)	1.000	0.874	gaaatttTTAAaaagt
SEQ_22	P\$AGP1/AGP1.01	Opaque-2 regulatory protein	0.87	1066	1082	(-)	1.000	0.898	tccataataTCATctga
SEQ_22	P\$AGP1/AGP1.01	AG-motif binding protein 1	0.91	1080	1090	(-)	1.000	0.915	tgaGATCttcc
SEQ_22	P\$AGP1/AGP1.01	AG-motif binding protein 1	0.91	1081	1091	(+)	1.000	0.911	gaaGATCtcaa
SEQ_22	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	1090	1106	(+)	1.000	0.897	aagagtTAAAtgtatcc

SEQ_22	P\$MYBS/OSMYBS.01	Rice MYB proteins with single DNA binding domains, binding to the amy-lase element (TATCCA)	0.82	1097	1113	(+)	1.000	0.897	aaatgTATCcatcttgg
SEQ_22	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	1111	1127	(-)	0.848	0.899	ccggttTTAAtgccccca
SEQ_22	P\$L1BX/ATML1.02	Arabidopsis thaliana meristem layer 1	0.76	1113	1129	(+)	1.000	0.854	gggCATTaaaaccggtg
SEQ_22	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	1115	1125	(-)	1.000	1.000	ggtttTAATgc
SEQ_22	P\$IBOX/GATA.01	Class I GATA factors	0.93	1133	1149	(+)	1.000	0.949	gggatGATAaatacaga
SEQ_22	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	1134	1150	(+)	1.000	0.855	ggatgaTAAAtacagac
SEQ_22	P\$ROOT/RHE.02	Root hair-specific element with a 3-nucleotid spacer between left part (LP) and right part (RP)	0.77	1172	1196	(+)	1.000	0.804	gtaattcatatttatCACGtttgcta
SEQ_22	P\$NACF/TANAC69.01	Wheat NACdomain DNA binding factor	0.68	1176	1198	(+)	0.895	0.684	ttcatatttatCACGtttgctaaa
SEQ_22	P\$GBOX/HBP1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.83	1179	1199	(-)	1.000	0.835	ttttagcaACGTgataaatat
SEQ_22	P\$MYCL/MYCRRS.01	Myc recognition sequences	0.93	1180	1198	(-)	1.000	0.967	tttagcaACGTgataaata
SEQ_22	P\$ABRE/ABRE.01	ABA response elements	0.82	1181	1197	(+)	1.000	0.826	atttatacACGTtgctaa
SEQ_22	P\$OPAQ/O2_GCIN4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	1182	1198	(+)	0.951	0.830	tttatacACGTtgctaaa
SEQ_22	P\$STKM/STK.01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	0.85	1192	1206	(+)	1.000	0.888	tgcTAAAaaaaattat
SEQ_22	P\$LEGB/LEGB.01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.59	1223	1249	(+)	0.750	0.609	acaaaatCAATtaaagagaaagaaga
SEQ_22	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	1227	1243	(+)	1.000	0.991	aatcaattAAAGagaaa

SEQ_22	P\$GAGA/GAGABP.01	(GA)n/(CT)n binding proteins (GBP, soybean; BBR, barley)	0.75	1231	1255	(+)	1.000	0.790	aattaaAGAGaaagaaagaaacgca
SEQ_22	P\$NCS2/NCS2.01	Nodulin consensus sequence 2	0.79	1232	1246	(-)	1.000	0.808	ttctttCTCTtaat
SEQ_22	P\$OCSE/OCSE.01	OCS-like elements	0.69	1283	1303	(+)	0.769	0.735	taaaataaaaaaattAACGCatg
SEQ_22	P\$SEF4/SEF4.01	Soybean embryo factor 4	0.98	1285	1295	(-)	1.000	0.983	aaTTTTtattt
SEQ_22	P\$OPAQ/O2_GCIN4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.81	1294	1310	(-)	1.000	0.825	cattcaACATgcgtaa
SEQ_22	P\$IDRE/IDE1.01	Iron-deficiency-responsive element 1	0.77	1297	1311	(+)	1.000	0.773	acGCATgttgaatgc
SEQ_22	P\$GBOX/BZIP910.02	bZIP transcription factor from Antirrhinum majus	0.84	1300	1320	(+)	1.000	0.864	catgttgaatgcTGACatgtc
SEQ_22	P\$GBOX/BZIP910.01	bZIP transcription factor from Antirrhinum majus	0.77	1306	1326	(+)	1.000	0.784	gaatgcTGACatgtcagtatg
SEQ_22	P\$ABRE/ABF1.03	ABA (abscisic acid) inducible transcriptional activator	0.82	1307	1323	(+)	0.750	0.829	aatgctgaCATGtcagt
SEQ_22	P\$OPAQ/O2.01	Opaque-2 regulatory protein	0.87	1307	1323	(-)	0.794	0.901	actgacatgTCAGGcatt
SEQ_22	P\$GTBX/S1F.01	S1F, site 1 binding factor of spinach rps1 promoter	0.79	1319	1335	(-)	1.000	0.818	attcATGGacatactga
SEQ_22	P\$ROOT/RHE.01	Root hair-specific element with a 2-nucleotid spacer between left part (LP) and right part (RP)	0.77	1322	1346	(+)	1.000	0.844	glatgtccatgaatcCACGtatcaa
SEQ_22	P\$GBOX/GBF1.01	bZIP protein G-Box binding factor 1	0.94	1329	1349	(-)	1.000	0.956	cgcttgatACGTggattcatg
SEQ_22	P\$GBOX/HBP1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	0.88	1330	1350	(+)	1.000	0.899	atgaatcCACGtatcaagcgc
SEQ_22	P\$ABRE/ABRE.01	ABA response elements	0.82	1332	1348	(-)	1.000	0.866	gcttgatACGTggattc
SEQ_22	P\$L1BX/ATML1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.82	1370	1386	(+)	1.000	0.889	tccttcTAAAtgaaaac
SEQ_22	P\$GAPB/GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	0.88	1375	1389	(+)	1.000	0.958	ctaaATGAaaacaac
SEQ_22	P\$URNA/USE.01	Upstream sequence elements in the promoters of U-snRNA genes of	0.75	1387	1403	(+)	0.750	0.781	aacttcACACatcaca

		higher plants							
SEQ_22	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida bZIP factors DPBF-1 and 2 (Dc3	0.76	1396	1412	(-)	0.817	0.771	tattgtTTGTtgtgatg
SEQ_22	P\$DPBF/DPBF.01	promoter binding factor-1 and 2)	0.89	1413	1423	(+)	1.000	0.898	cACACaagacc
SEQ_22	P\$GAGA/BPC.01	Basic pentacysteine proteins (GA)n/(CT)n binding proteins (GBP, soybean; BBR, barley)	1.00	1413	1437	(-)	1.000	1.000	aacgagAGAGaggggtcttgttg
SEQ_22	P\$GAGA/GAGABP.01	Basic pentacysteine proteins	0.75	1423	1447	(-)	0.750	0.805	gcagagAGACaacgagagagagggg
SEQ_22	P\$GAGA/BPC.01	Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light-mediated induction	1.00	1425	1449	(-)	1.000	1.000	tggcagAGAGGacaaacgagagagagg
SEQ_22	P\$PREM/MGPROTORE.01		0.77	1447	1477	(+)	1.000	0.794	ccagCGACccaaatcgaagcttgagaagaaca

Table 7: cis- regulatory elements of SEQ ID NO: 25

Seq. name	Family/matrix	Further Information	Opt. thresh.	Start pos.	End pos.	Strand	Core sim.	Matrix sim.	Sequence
SEQ_25	P\$IBOX/GATA.01	Class I GATA factors	0.93	27	43	(+)	1.000	0.946	aaaaaGATAaccacccc
SEQ_25	P\$NCS1/NCS1.01	Nodulin consensus sequence 1	0.85	27	37	(+)	1.000	0.948	aAAAAGataac
SEQ_25	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	29	45	(-)	1.000	0.980	agggggtgGTTAtcttt
SEQ_25	P\$SALT/ALFIN1.02	Zinc-finger protein in alfalfa roots, regulates salt tolerance	0.95	34	48	(-)	1.000	0.977	gctagggGGTGgtta
SEQ_25	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	57	75	(+)	1.000	0.852	ccAAATcataactatcaga
SEQ_25	P\$AHBP/ATHB9.01	HD-ZIP class III protein ATHB9	0.77	58	68	(-)	1.000	0.772	gttATGAtttg
SEQ_25	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	60	76	(-)	1.000	0.808	tctgaTAGTtatgatt
SEQ_25	P\$IBOX/GATA.01	Class I GATA factors	0.93	76	92	(+)	1.000	0.975	acaaaGATAaaaagccc
SEQ_25	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription	0.99	78	94	(+)	1.000	0.996	aaagataaAAAGcccca

SEQ_25	P\$SBPD/SBP.01	factor	0.88	99	115	(-)	1.000	0.882	ctatgGTACaactggt
SEQ_25	P\$NCS3/NCS3.01	SQUA promoter binding proteins Nodulin consensus sequence 3	0.89	119	129	(+)	1.000	0.893	caCACCCctcta
SEQ_25	P\$LEGB/LEGB.01	Legumin box, highly conserved se- quence element about 100 bp up- stream of the TSS in legumin genes	0.59	122	148	(-)	0.750	0.626	ttcataaCTATgtatgagatagagggt
SEQ_25	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	133	149	(+)	1.000	0.778	catacaTAGTtatgaat
SEQ_25	P\$MIIG/P_ACT.01	Maize activator P of flavonoid bio- synthetic genes	0.93	150	164	(-)	1.000	0.977	ttacGGTAaggtttca
SEQ_25	P\$MYBL/ATMYB77.01	R2R3-type myb-like transcription factor (l-type binding site)	0.87	172	188	(-)	1.000	0.909	ttgtggCGGTtcttgct
SEQ_25	P\$DREB/HVDRF1.01	H. vulgare dehydration-response factor 1	0.89	175	189	(+)	1.000	0.953	aggaACCGGccacaat
SEQ_25	P\$GCCF/ERE_JERE.01	Ethylene-responsive elements (ERE) and jasmonate- and elicitor- responsive elements (JERE)	0.85	175	187	(+)	1.000	0.865	aggaacCGCCaca
SEQ_25	P\$MYBS/TAMYB80.01	MYB protein from wheat	0.83	191	207	(-)	1.000	0.874	gcttATATtcccctcac
SEQ_25	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	192	212	(+)	1.000	0.796	tgAGGGaatataagccaaag
SEQ_25	P\$LEGB/LEGB.01	Legumin box, highly conserved se- quence element about 100 bp up- stream of the TSS in legumin genes	0.59	206	232	(+)	0.750	0.649	gccaaagCCCTgcaattttcagtgaga
SEQ_25	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	230	250	(+)	0.956	0.795	agAAGGgtaagattattaaag
SEQ_25	P\$AHBP/HAHB4.01	Sunflower homeodomain leucine- zipper protein Hahb-4	0.87	238	248	(+)	1.000	0.903	aagattATTaa
SEQ_25	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	239	255	(+)	1.000	0.980	agattattAAGgcagc
SEQ_25	P\$L1BX/ATML1.01	L1-specific homeodomain protein	0.82	240	256	(+)	1.000	0.835	gattatTAAAggcagcc

SEQ_25	P\$DREB/CRT_DRE.01	ATML1 (A. thaliana meristem layer 1)	0.89	262	276	(-)	1.000	0.968	ctttgCCGAcattgt
SEQ_25	P\$URNA/USE.01	C-repeat/dehydration response element	0.75	284	300	(-)	1.000	0.858	aatgtcCCACctcgaat
SEQ_25	P\$GTBX/SBF1.01	Upstream sequence elements in the promoters of U-snRNA genes of higher plants	0.87	292	308	(+)	1.000	0.908	tgggacaTTAAatttaa
SEQ_25	P\$L1BX/HDG9.01	SBF-1	0.77	292	308	(-)	0.796	0.792	ttaaattTAAAtgtccca
SEQ_25	P\$L1BX/PDF2.01	Homeodomain glabrous 9	0.85	294	310	(+)	1.000	0.851	ggacatTAAAtttaaaa
SEQ_25	P\$AHBP/WUS.01	Protodermal factor 2	0.94	296	306	(-)	1.000	1.000	aaattTAATgt
SEQ_25	P\$GTBX/SBF1.01	Homeodomain protein WUSCHEL	0.87	298	314	(+)	1.000	0.875	attaaattTAAaaaagaa
SEQ_25	P\$GTBX/SBF1.01	SBF-1	0.87	299	315	(-)	1.000	0.885	cttctttTAAatttaa
SEQ_25	P\$IDRE/IDE1.01	Iron-deficiency-responsive element 1	0.77	325	339	(-)	0.809	0.806	aaGCTTgctactttc
SEQ_25	P\$AGP1/AGP1.01	AG-motif binding protein 1	0.91	376	386	(+)	1.000	0.912	caaGATCtcc
SEQ_25	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	383	393	(+)	1.000	0.963	ttcctTAATcc
SEQ_25	P\$GTBX/S1F.01	S1F, site 1 binding factor of spinach rps1 promoter	0.79	383	399	(-)	1.000	0.810	tgttATGGattaaggaa
SEQ_25	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	0.82	387	407	(+)	1.000	0.865	ttaatCCA Taacaagaagtcc
SEQ_25	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	435	445	(+)	0.829	0.940	acaATGAttct
SEQ_25	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	435	445	(-)	0.936	0.941	agaATCAttgt
SEQ_25	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	0.81	444	462	(+)	1.000	0.875	ctAAATcatacatattacc
SEQ_25	P\$MADS/AGL15.01	AGL15, Arabidopsis MADS-domain protein AGAMOUS-like 15	0.79	487	507	(-)	0.850	0.804	tttTGCTacacciggtagtag
SEQ_25	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	487	505	(-)	0.954	0.966	ttgctACACcctggtagtag
SEQ_25	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	0.82	488	508	(+)	0.869	0.855	tactaCCAGgtgtagcaaaat
SEQ_25	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	510	526	(+)	1.000	0.880	cccgatTAAAttcataa

SEQ_25	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	512	522	(-)	1.000	0.963	gaattTAATcg
SEQ_25	P\$RAV5/RAV1-5.01	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	0.96	537	547	(-)	1.000	0.974	agcAACAAaat
SEQ_25	P\$CARM/CARICH.01	CA-rich element	0.78	548	566	(+)	1.000	0.837	accttcaAACAcagatgc
SEQ_25	P\$MYCL/ICE.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.95	554	572	(+)	0.863	0.954	aaacaACAGatgctgcaa
SEQ_25	P\$RAV5/RAV1-5.01	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	0.96	555	565	(+)	1.000	0.961	aacAACAgatg
SEQ_25	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	596	612	(+)	1.000	0.771	tctagcTAGTaacgacc
SEQ_25	P\$MYBL/NTMYBAS1.01	Anther-specific myb gene from tobacco	0.96	600	616	(-)	1.000	0.975	gtaggtcGTTActagc
SEQ_25	P\$MYBL/GAMYB.01	GA-regulated myb gene from barley	0.91	608	624	(-)	1.000	0.927	atagtgttGTTAggtcg
SEQ_25	P\$MIIG/PALBOXL.01	Cis-acting element conserved in various PAL and 4CL promoters	0.80	622	636	(-)	1.000	0.812	atctgtttGGTGata
SEQ_25	P\$CARM/CARICH.01	CA-rich element	0.78	623	641	(+)	1.000	0.788	atcacccaAACAgataaaca
SEQ_25	P\$MYBL/CARE.01	CAACTC regulatory elements, GA-inducible	0.83	646	662	(+)	1.000	0.834	tctagcgAGTTccagca
SEQ_25	P\$CE1F/ABI4.01	ABA insensitive protein 4 (ABI4)	0.87	680	692	(+)	1.000	0.900	tcgcCACCGcagca
SEQ_25	P\$DREB/CRT_DRE.01	C-repeat/dehydration response element	0.89	681	695	(+)	1.000	0.914	cgccaCCGAcgatta
SEQ_25	P\$PREM/MGPROTORE.01	Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light-mediated induction	0.77	683	713	(+)	1.000	0.774	ccacCGACgattatgattcactaaagctac
SEQ_25	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	698	714	(+)	1.000	1.000	gattcactAAAGctaca
SEQ_25	P\$CGCG/ATSR1.01	Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1	0.84	714	730	(-)	1.000	0.865	ccaCGCGtgtacttgtt
SEQ_25	P\$CE3S/CE3.01	(tobacco early ethylene-responsive gene)	0.77	715	733	(-)	1.000	0.787	tatccaCGCGtgtacttgt
		Coupling element 3 (CE3), non-							

ACGT ABRE

Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)
MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat
Octamer motif found in plant histone H3 and H4 genes
Zea mays MYB-related protein 1 (transfer cell specific)
Nodulin consensus sequence 2
M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)
Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light-mediated induction
Maize INDETERMINATE1 zinc finger protein
MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat
H. vulgare dehydration-response factor 1
Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs
Zinc-finger protein in alfalfa roots, regulates salt tolerance
Promoter elements involved in MgProto (Mg-protoporphyrin IX) and

SEQ_25 P\$CGCG/ATSR1.01
SEQ_25 P\$MYBS/MYBST1.01
SEQ_25 P\$HOCT/HOCT.01
SEQ_25 P\$MYBS/ZMMRP1.01
SEQ_25 P\$NCS2/NCS2.01
SEQ_25 P\$MSAE/MSA.01
SEQ_25 P\$PREM/MGPROTORE.01
SEQ_25 P\$IDDF/ID1.01
SEQ_25 P\$MYBS/MYBST1.01
SEQ_25 P\$DREB/HVDRF1.01
SEQ_25 P\$GARP/ARR10.01
SEQ_25 P\$SALT/ALFIN1.01
SEQ_25 P\$PREM/MGPROTORE.01

0.84 721 737 (+) 1.000 0.870 acaCGCGtgtagtagtg
0.90 722 738 (-) 1.000 0.936 tccactATCCacgcgtg
0.76 723 739 (-) 1.000 0.768 ttccactATCCacgcgt
0.79 732 748 (-) 0.777 0.852 atttctcTATTccacta
0.79 746 760 (+) 0.750 0.803 aattgcCTGTtcaac
0.80 753 767 (+) 1.000 0.875 tgttcAACGgggaga
0.77 786 816 (+) 1.000 0.801 atagCGACaaggaggagcgatattgcta
0.92 787 799 (-) 1.000 0.921 ctccTTGTcgcta
0.90 810 826 (+) 1.000 0.928 attgctATCCggaaagt
0.89 842 856 (+) 0.826 0.914 actcGCCGccatata
0.97 857 865 (-) 1.000 0.970 AGATccttg
0.93 872 886 (-) 1.000 0.930 ccttgGTGGcgccgt
0.77 888 918 (-) 1.000 0.774 actaCGACggcgatagggtgaccattcgag

SEQ_25	P\$NCS3/NCS3.01	Nodulin consensus sequence 3	0.89	896	906	(+)	1.000	0.947	gtCACCCctcat
SEQ_25	P\$AHBP/ATHB9.01	HD-ZIP class III protein ATHB9	0.77	920	930	(+)	0.750	0.773	gtaTTGAtctc
SEQ_25	P\$GARP/ARR10.01	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	0.97	941	949	(+)	1.000	0.973	AGATcctgg
SEQ_25	P\$ABRE/ABF1.03	ABA (abscisic acid) inducible transcriptional activator	0.82	952	968	(-)	1.000	0.833	ggcgaggcCGTGgctca
SEQ_25	P\$L1BX/ATML1.02	Arabidopsis thaliana meristem layer 1	0.76	987	1003	(-)	1.000	0.791	aggCATTcaaattctggc
SEQ_25	P\$L1BX/PDF2.01	Protodermal factor 2	0.85	989	1005	(+)	0.787	0.854	cagattGAAtgctcc
SEQ_25	P\$HEAT/HSE.01	Heat shock element	0.81	1015	1029	(-)	1.000	0.822	gcatactctccAGAAat
SEQ_25	P\$MYBL/MYBPH3.02	Myb-like protein of Petunia hybrida	0.76	1076	1092	(+)	0.817	0.775	gagagaTTGTtgctttc
SEQ_25	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	1077	1091	(-)	1.000	0.893	aaagcaacAATCtct
SEQ_25	P\$GTBX/GT3A.01	Trihelix DNA-binding factor GT-3a	0.83	1107	1123	(+)	1.000	0.847	atcagtGTTActtgat
SEQ_25	P\$CCAF/CCA1.01	Circadian clock associated 1	0.85	1126	1140	(-)	1.000	0.898	gcaaacagAATCtca
SEQ_25	P\$GAGA/BPC.01	Basic pentacysteine proteins	1.00	1128	1152	(-)	1.000	1.000	cgagagAGAGaggcaaacagaatct
SEQ_25	P\$GAGA/BPC.01	Basic pentacysteine proteins	1.00	1130	1154	(-)	1.000	1.000	aacgagAGAGagaggcaaacagaat
SEQ_25	P\$HEAT/HSE.01	Heat shock element	0.81	1130	1144	(-)	1.000	0.855	agaggcaaacAGAAat
SEQ_25	P\$GAGA/GAGABP.01	(GA)n/(CT)n binding proteins (GBP, soybean; BBR, barley)	0.75	1134	1158	(-)	0.750	0.760	ccagaaCGAGagagagaggcaaac
SEQ_25	P\$GAGA/GAGABP.01	(GA)n/(CT)n binding proteins (GBP, soybean; BBR, barley)	0.75	1138	1162	(-)	0.750	0.757	gaaaccAGAAcgagagagagaggca
SEQ_25	P\$PSRE/GAAA.01	GAAA motif involved in pollen specific transcriptional activation	0.83	1164	1180	(+)	1.000	0.843	ctgtaGAAAaacttttt
SEQ_25	P\$NCS1/NCS1.01	Nodulin consensus sequence 1	0.85	1170	1180	(-)	1.000	0.853	aAAAAgttttt
SEQ_25	P\$PSRE/GAAA.01	GAAA motif involved in pollen specific transcriptional activation	0.83	1185	1201	(-)	1.000	0.838	ctaaaGAAAatgttcgc
SEQ_25	P\$DOFF/DOF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0.98	1191	1207	(-)	1.000	0.984	ttgggcctAAAGaaaat

SEQ_25	P\$TCPF/ATTCP20.01	TCP class I transcription factor (Arabidopsis)	0.94	1198	1210	(+)	1.000	0.968	ttagGCCCaaagt
SEQ_25	P\$TCPF/ATTCP20.01	TCP class I transcription factor (Arabidopsis)	0.94	1208	1220	(-)	1.000	0.943	aaaaGCCCaaact
SEQ_25	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	1211	1227	(-)	1.000	0.994	cgttattaAAAGGcccaa
SEQ_25	P\$L1BX/HDG9.01	Homeodomain glabrous 9	0.77	1211	1227	(-)	1.000	0.818	cgttatTAAAgagcccaa
SEQ_25	P\$GTBX/SBF1.01	SBF-1	0.87	1213	1229	(-)	1.000	0.927	ggcgtaTTAAaagccc
SEQ_25	P\$DOFF/DOF3.01	Dof3 - single zinc finger transcription factor	0.99	1223	1239	(+)	1.000	0.995	taacgcctAAAGGcccaa
SEQ_25	P\$CAAT/CAAT.01	CCAAT-box in plant promoters	0.97	1250	1258	(-)	1.000	0.979	acCCAAtaa
SEQ_25	P\$TEFB/TEF1.01	TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences	0.76	1288	1308	(+)	0.956	0.767	atAAGGggaatctattttt
SEQ_25	P\$MADS/SQUA.01	MADS-box protein SQUAMOSA	0.90	1294	1314	(+)	1.000	0.917	ggaatctATTTatttaattgt
SEQ_25	P\$STKM/STK.01	Storekeeper (STK), plant specific	0.85	1298	1312	(-)	1.000	0.859	aatTAAAtaaataga
SEQ_25	P\$L1BX/HDG9.01	Homeodomain glabrous 9	0.77	1299	1315	(-)	1.000	0.791	aacaatTAAAtaaatag
SEQ_25	P\$HMGF/HMG_IY.01	High mobility group I/Y-like proteins	0.89	1300	1314	(+)	1.000	0.907	tattTATTtaattgt
SEQ_25	P\$GTBX/SBF1.01	SBF-1	0.87	1307	1323	(+)	1.000	0.872	ttaattgTTAAAtcattc
SEQ_25	P\$AHBP/WUS.01	Homeodomain protein WUSCHEL	0.94	1310	1320	(+)	1.000	0.963	attgtTAATca
SEQ_25	P\$OCSE/OCSE.01	OCS-like elements	0.69	1310	1330	(+)	1.000	0.712	attgtaatcattcACGTtga
SEQ_25	P\$LEGB/LEGB.01	Legumin box, highly conserved se- quence element about 100 bp up- stream of the TSS in legumin genes	0.59	1311	1337	(+)	0.750	0.625	ttgttaaTCATtcacgttgaccattga
SEQ_25	P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	0.89	1314	1324	(+)	0.936	0.939	ttaATCAttca
SEQ_25	P\$AHBP/HAHB4.01	Sunflower homeodomain leucine- zipper protein Hahb-4	0.87	1314	1324	(-)	1.000	0.945	tgaatgATTAA
SEQ_25	P\$GBOX/HBP1B.01	Wheat bZIP transcription factor	0.83	1315	1335	(-)	1.000	0.834	aatgtcaACGTgaatgatta

SEQ	Accession	Gene	Position	Score	Significance	Sequence
SEQ_25	P\$OPAQ/O2_GC	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	1317	0.81	(-)	tgttcaACGTgaatgat
SEQ_25	P\$WBXF/ERE.01		1322	0.89	(+)	tcacgtTGACcattgaa
SEQ_25	P\$HEAT/HSE.01		1328	0.81	(+)	tgaccattgaAGAAc
SEQ_25	P\$NCS1/NCS1.01		1373	0.85	(-)	cAAAGgatcaa
SEQ_25	P\$MADS/AGL2.01	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	1408	0.82	(+)	tgtctCCAGctctagtaatga
SEQ_25	P\$MYBL/MYBPH3.02		1427	0.76	(+)	gacaatTAGTttagttt
SEQ_25	P\$STKM/STK.01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	1428	0.85	(-)	aacTAAActaattgt
SEQ_25	P\$SPF1/SP8BF.01		1459	0.87	(-)	aaTACTattacac
SEQ_25	P\$GTBX/SBF1.01		1495	0.87	(-)	aaaattgTTAAataata
SEQ_25	P\$MYBL/MYBPH3.01		1497	0.80	(-)	agaaaattGTTAataa
SEQ_25	P\$E2FF/E2F.01	E2F class I sites	1527	0.82	(+)	ttttTTCagcaacg
SEQ_25	P\$GTBX/SBF1.01		1548	0.87	(+)	tgaattTTAAaacttg
SEQ_25	P\$SUCB/SUCROSE.01	Sequence motif from the promoters of different sugar-responsive genes	1552	0.81	(-)	aaATATcaagttttaaaaa
SEQ_25	P\$CCAF/CCA1.01		1568	0.85	(-)	ttaaaaaaAATCaaa
SEQ_25	P\$AHBP/BLR.01	Transcriptional repressor BELL-RINGER	1581	0.90	(-)	taaATTActtt
SEQ_25	P\$GTBX/SBF1.01		1583	0.87	(-)	tttgcattTAAAttact
SEQ_25	P\$L1BX/ATML1.01	L1-specific homeodomain protein	1584	0.82	(+)	gtaattTAAAtgcaaaa
SEQ_25	P\$ATML1	(A. thaliana meristem layer)	1600	0.918	(+)	

SEQ_25	P\$NCS1/NCS1.01	1)	Nodulin consensus sequence 1	0.85	1606	1616	(+)	0.878	0.862	cAAATgatatt
SEQ_25	P\$MYBL/CARE.01		CAACTC regulatory elements, GA-inducible	0.83	1625	1641	(+)	1.000	0.834	actcaagAGTTgtgtga
SEQ_25	P\$NCS2/NCS2.01		Nodulin consensus sequence 2	0.79	1653	1667	(-)	1.000	0.809	actgcCTCTtgccc
			TEF cis acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences							
SEQ_25	P\$TEFB/TEF1.01		Hordeum vulgare Myb-related CAB-promoter-binding protein 1	0.76	1665	1685	(-)	0.956	0.816	agAAGGatacaccagtgcaact
SEQ_25	P\$MYBS/HVMCB1.01		SBF-1	0.93	1672	1688	(+)	1.000	0.933	tggtgtATCCttctcgg
SEQ_25	P\$GTBX/SBF1.01		Nodulin consensus sequence 3	0.87	1686	1702	(+)	1.000	0.878	cggggcgTTAAaaccgt
SEQ_25	P\$NCS3/NCS3.01		Myb-like protein of Petunia hybrida	0.89	1702	1712	(-)	1.000	0.965	gtCACCTtcca
SEQ_25	P\$MYBL/MYBPH3.01		M-phase-specific activators	0.80	1704	1720	(-)	0.750	0.832	attaaacgGTCAccttc
SEQ_25	P\$MSAE/MSA.01		(NtmybA1, NtmybA2, NtmybB)	0.80	1707	1721	(-)	1.000	0.897	tattaAACGgtcacc
SEQ_25	P\$L1BX/HDG9.01		Homeodomain glabrous 9	0.77	1708	1724	(-)	1.000	0.791	atgtatTAAAcggtcac
			Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins							
SEQ_25	P\$OPAQ/O2_GC4.01		Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.81	1716	1732	(-)	1.000	0.834	agatagACATgtattaa
			Root hair-specific element with a 3-nucleotid spacer between left part (LP) and right part (RP)							
SEQ_25	P\$LREM/ATCTA.01		Wheat bZIP transcription factor	0.85	1727	1737	(+)	1.000	0.853	ctATCTattat
			HBP1B (histone gene binding protein 1b)							
SEQ_25	P\$ROOT/RHE.02		OCS-like elements	0.77	1747	1771	(+)	1.000	0.774	gtactcatagctatCACGttgctc
			Sequence motif from the promoters							
SEQ_25	P\$GBOX/HBP1B.01			0.83	1754	1774	(-)	1.000	0.843	gaggagcaACGTgatagctat
SEQ_25	P\$OCSE/OCSE.01			0.69	1760	1780	(-)	1.000	0.700	atttgtgaggagcaACGTgat
SEQ_25	P\$SUCB/SUCROSE.01			0.81	1799	1817	(+)	1.000	0.911	gcAAATcaattttataaag

SEQ_25	P\$TBPF/TATA.01	of different sugar-responsive genes	0.88	1807	1821	(+)	1.000	0.910	atttTATAaagacgc
SEQ_25	P\$GTBX/GT3A.01	Plant TATA box	0.83	1820	1836	(-)	1.000	0.865	tcagctGTTActcatgc
SEQ_25	P\$MYBL/MYBPH3.01	Trihelix DNA-binding factor GT-3a	0.80	1822	1838	(-)	1.000	0.823	attcagctGTTActcat
SEQ_25	P\$DOFF/PBF.01	Myb-like protein of Petunia hybrida	0.97	1834	1850	(+)	1.000	0.988	tgaataaaaAAGagagg
SEQ_25	P\$IDRE/IDE1.01	PBF (MPBF)	0.77	1856	1870	(+)	1.000	0.777	acGCA Tgttgattgc
SEQ_25	P\$GBOX/BZIP910.02	Iron-deficiency-responsive element 1	0.84	1859	1879	(+)	1.000	0.962	catgttgattgcTGACgtgtc
SEQ_25	P\$OCSE/OCSE.01	bZIP transcription factor from Antirrhinum majus	0.69	1859	1879	(+)	1.000	0.751	catgttgattgcTgACGTgtc
SEQ_25	P\$GBOX/HBP1B.01	OCS-like elements	0.83	1864	1884	(-)	1.000	0.973	ctacggacACGTcagcaatca
SEQ_25	P\$GBOX/GBF1.01	Wheat bZIP transcription factor	0.94	1865	1885	(+)	1.000	0.999	gattgctgACGTgtccgtagg
SEQ_25	P\$ABRE/ABF1.03	HBP1B (histone gene binding protein 1b)	0.82	1866	1882	(+)	1.000	0.960	attgcigaCGTGtccgt
SEQ_25	P\$OPAQ/RITA1.01	bZIP protein G-Box binding factor 1	0.95	1866	1882	(-)	1.000	0.956	acggacACGTcagcaat
SEQ_25	P\$OPAQ/RITA1.01	ABA (abscisic acid) inducible transcriptional activator	0.95	1867	1883	(+)	1.000	0.958	ttgctgACGTgtccgta
SEQ_25	P\$EINL/TEIL.01	Rice transcription activator-1 (RITA), basic leucin zipper protein, highly expressed during seed development	0.92	1882	1890	(-)	0.863	0.966	aTGGAaccta
SEQ_25	P\$GAGA/BPC.01	Rice transcription activator-1 (RITA), basic leucin zipper protein, highly expressed during seed development	1.00	1937	1961	(-)	1.000	1.000	ctcgagAGAGaggggacttatgatg
		TEIL (tobacco EIN3-like)							
		Basic pentacysteine proteins							

In the following table 9, cis-regulatory element families commonly found in at least 13 of the listed expression control sequences of the present invention are shown.

Table 8 describes the cis-regulatory element families.

5

Table 8: Description of cis-regulatory elements families commonly found in expression control sequences of the present invention.

TF-Family	Description
O\$VTBP	Vertebrate TATA binding protein vector
P\$AHBP	Arabidopsis homeobox protein
P\$DOFF	DNA binding with one finger (DOF)
P\$GTBX	GT-box elements
P\$IBOX	Plant I-Box sites
P\$L1BX	L1 box, motif for L1 layer-specific expression
P\$LREM	Light responsive element motif, not modulated by different light qualities
P\$MYBL	MYB-like proteins
P\$MYBS	MYB proteins with single DNA binding repeat
P\$NCS1	Nodulin consensus sequence 1
P\$CCAF	Circadian control factors
P\$GBOX	Plant G-box/C-box bZIP proteins
P\$IDDF	D domain factors, the ID domain includes a cluster of three different types of zinc fingers separated from a fourth C2H2 finger by a long spacer
P\$MADS	floral determination
P\$MYCL	Myc-like basic helix-loop-helix binding factors
P\$OPAQ	Opaque-2 like transcriptional activators
P\$SPF1	Sweet potato DNA-binding factor with two WRKY-domains
P\$SUCB	Sucrose box
P\$WBXF	W Box family
O\$INRE	Core promoter initiator elements
O\$PTBP	Plant TATA binding protein vector
P\$ABRE	ABA response elements
P\$LEGB	Legumin Box family
P\$NACF	Plant specific NAC [NAM (no apical meristem), ATAF172, CUC2 (cup-shaped cotyledons 2)] transcription factors
P\$OCSE	Enhancer element first identified in the promoter of the octopine synthase gene (OCS) of the Agrobacterium tumefaciens T-DNA

Table 9: Cis- regulatory elements families commonly found in expression control sequences in at least 13 of the listed expression control sequences of the present invention.

TF-Family	No of occurrences														
	p-Bn SC P (SEQ-ID 70)	p-BnGR PL (SEQ-ID 1)	p-BnCR U4 (SEQ-ID 77)	p-BnMY R (SEQ-ID 85)	p-BnSET L-var1 (SEQ-ID 22)	p-BnSET L-var2 (SEQ-ID 25)	p-BnSCT 2-var1 (SEQ-ID 14)	p-BnSCT 2-var2 (SEQ-ID 16)	p-BnMD P (SEQ-ID 95)	p-BnRT I-4 (SEQ-ID 103)	p-BnMT FL (SEQ-ID 111)	p-BnGS TF (SEQ-ID 119)	p-BnPE F-var1 (SEQ-ID 6)	p-BnPE F-var2 (SEQ-ID 9)	p-BnLS P (SEQ-ID 125)
O\$VTBP	7	18	17	39	6	10	5	8	19	9	10	13	28	4	12
P\$AHBP	16	8	22	25	6	12	7	9	15	22	9	9	18	5	14
P\$DOFF	14	8	11	9	2	7	5	5	4	8	6	4	10	8	5
P\$GTBX	12	13	10	16	10	12	5	1	7	8	10	5	22	5	10
P\$IBOX	4	4	4	2	4	2	2	2	7	3	1	1	8	2	3
P\$L1BX	9	7	11	18	10	10	2	3	10	10	2	1	10	3	9
P\$LREM	6	4	4	6	3	1	3	3	5	6	2	2	8	4	5
P\$MYBL	9	11	12	7	16	14	9	7	9	10	9	15	18	5	6
P\$MYBS	9	3	8	2	4	5	2	2	10	13	3	4	5	3	5
P\$NCS1	2	2	4	3	3	4	2	2	2	1	3	2	1	1	2
P\$CCAF	4	4	4	2	3	3	1	1	3	4	3	3	9	2	0
P\$GBOX	9	3	12	2	5	5	7	5	3	12	1	5	2	0	7
P\$IDDF	3	2	1	1	1	1	4	5	3	1	3	3	3	1	0
P\$MADS	6	9	9	2	0	5	2	2	7	12	6	3	8	5	6
P\$MYCL	1	0	6	3	5	2	3	3	3	3	1	6	4	1	1
P\$OPAQ	9	5	7	3	6	4	1	2	4	12	0	5	4	2	7
P\$SPF1	2	6	1	5	1	1	2	2	2	3	0	2	7	1	2
P\$SUCB	6	4	3	7	1	4	2	3	0	1	2	1	7	2	3
P\$WBXF	2	3	1	0	1	1	3	2	1	1	2	1	3	1	5

TF-Family	No of occurrences														
	p-Bn SC P (SEQ-ID 70)	p-BnGR PL (SEQ-ID 1)	p-BnCR U4 (SEQ-ID 77)	p-BnMY R (SEQ-ID 85)	p-BnSET L-var1 (SEQ-ID 22)	p-BnSET L-var2 (SEQ-ID 25)	p-BnSCT 2-var1 (SEQ-ID 14)	p-BnSCT 2-var2 (SEQ-ID 16)	p-BnMD P (SEQ-ID 95)	p-BnRT I-4 (SEQ-ID 103)	p-BnMT FL (SEQ-ID 111)	p-BnGS TF (SEQ-ID 119)	p-BnPE F-var1 (SEQ-ID 6)	p-BnPE F-var2 (SEQ-ID 9)	p-BnLS P (SEQ-ID 125)
O\$INRE	5	1	1	2	1	4	0	0	1	2	4	2	7	2	2
O\$PTBP	2	7	8	24	1	1	0	1	5	4	3	8	7	2	0
P\$ABRE	1	1	3	0	4	2	3	3	1	3	1	3	1	0	4
P\$LEGB	6	4	8	2	1	2	2	2	7	3	0	1	2	0	2
P\$NACF	2	2	2	1	1	0	3	3	2	3	1	0	1	1	1
P\$OCSE	3	4	3	5	3	3	7	7	3	7	0	4	4	0	3

Figure 1 shows a developmental expression analysis. Tissue types and developmental stages are given as listed in table 8. Samples 10, 11, 12 were pooled (assigned as 10), as well as samples 13, 14, 15 (assigned as 13).

Figure 2 shows the 18:3n-6 (GLA) content of seeds-oil of seeds harvested from transgenic plants harboring the T-DNA from vectors described in example 3. Shown are data from T1 seeds and T2 seeds as indicated on top of the figure. Measurements on T1 on seeds are on individual single seeds; measurements on T2 seeds are on seed batches. The black line indicates the minimal and the maximal observation, the box reaches from the 25% quartil to the 75 quartil; the median is indicated as black line within the box. The number of individual measurements is indicated as number above each box.

Figure 3 shows the 18:3n-6 (GLA) and 18:4n-3 (SDA) content of seeds-oil of seeds harvested from transgenic Arabidopsis plants harboring the T-DNA from vectors described in example 3.

Figure 4 illustrates for the different promoters the different ratios of the omega-3 fatty acid SDA to the omega-6 fatty acid GLA.

The invention will now be illustrated by the following Examples which are not intended, whatsoever, to limit the scope of this application.

Example 1: General Cloning Methods

General Cloning Methods including enzymatic digestion by restriction enzymes, agarous gel electrophoresys, purification of DNA fragments, transfer of nucleic acids to nitrocellulose on nylon membranes, maligation of DNA fragments, transformation of E. coli bacteria as well as culture of bacteria and sequence analysis of recombinant DNA have been carried out as described in Sambrook et al. (1989, Cold Spring Harbour Laboratory Press. ISBN 0-87969-309-6).

Example 2: Cloning of Promotor Elements from Brassica napus

For the analysis of seed specific expressed genes in Brassica napus, different tissues of various developmental stages (table 8) have been investigated.

5 Table 10: tissue types used for transcript analysis

Specific developmental stages are given using the Biologische Bundesanstalt, Bundes-sortenamt, and Chemical Industry (BBCH) code (Meier, 1997)

Sample Nr.	Tissue	Developmental stage	type sample
1	immature embryos	walking stick	purpose prio 1
2	immature embryos	fully developped (approx 20 days)	purpose prio 1
3	immature embryos	fully developped pre dessic. green	purpose prio 1
4	immature embryos	fully developped start dessic. Yellow	control/purpose prio 2
5	embryo sac complete	heart	control/purpose prio 2
6	immature seed coat and endosperm	weight mix sample 1 - 3	control/purpose prio 2
7	flower buds	BBCH 57	control/purpose prio 2
8	anthers and stigma	BBCH 68	control/purpose prio 2
9	flowers (exclude anthers and stigma)	BBCH 67	control
10	leaves	BBCH 32	control
11	leaves	BBCH 75	control
12	leaves	BBCH 80, senescing (yellow)	control
13	stem	BBCH 32	control
14	stem	BBCH 75	control
15	stem	BBCH 80, senescing (yellow)	control
16	seedling: hypocotyl and cotyledons	BBCH 12	control
17	seedling: roots (sand)	BBCH 12	control
18	empty siliques	weight mix sample 1 - 3	control

10

To this end, Brassica napus cv. Kumily plants were raised under standard conditions (Moloney et al. 1992, Plant Cell Reports 8: 238-242). The tissues where harvested at the indicated developmental stages (table 8) and used for the preparation of RNA (RNAeasy, Qiagen) according to the manufactures manual. To idendtify in further experiments seed specifically expressed mRNA transcripts, three pools were created by mixing the RNA: A seed specific pool P-S was created from sample 1, 2, 3 (see table 8). A control pool P-C1 was created from sample 9, 10, 11, 12, 13, 14, 15, 16, 17, 18. A

15

third pool P-C2 consisting of sample 4, 5, 6 7, 8 was used as a more stringent control: transcripts not expressed in control pool P-C2 are solely expressed in early seed development and in no other tissue/developmental stage. Using the amplified fragment length polymorphism (AFLP) method, known to a person skilled in the art, 384 primer combinations were used to identify 96 candidate transcript fragments being present solely in pool P-S and/or also weakly in pool P-C2 but absent on pool P-C1. A selection of 42 primer combinations that identified the 55 most promising transcript fragments were used to analyse in detail all samples listed in table 8, resulting in 26 fragments with confirmed expression solely in the developing embryo samples 1, 2, 3 and 4. Sequencing of those candidate fragments resulted in 20 unique sequences. Basic Local Alignment Search Tool (BLAST) was used to identify the putative full length transcripts corresponding to the identified transcript fragments. Surprisingly, besides genes known in the art for seed specific expression, e.g. napin and 3-ketoacyl-CoA synthase, a number of fragments listed in table 9 showed no homology to any known Brassica sequence, or to genes which were not known to be seed specifically expressed and/or to genes with unknown sequence upstream of the known mRNA sequence.

Table 11 Candidate fragments and homologue sequences identified using BLAST.

Candidate	fragment length	SEQ-ID	Brassica homologue	Arabidopsis homologue
BnSCP	157 bp	58		At5g36100 (unknown protein)
BnGRPL	174 bp	4		At3g10185 (similar to gibberelin responsive protein)
BnCRU4	127 bp	59	BnCRU4rn (Cruciferin)	
BnMYR	101 bp	60	BnMYRmc (Myrosinase)	
BnSETL	120 bp	26		At1g03270 (unknown protein)
BnSCT2	172 bp	19	BnSCT2 (Sinapoyl-cholin transferase 2)	
BnMDP	273 bp	61		At3g20370 (unknown, similar)
BnRTI-4	237 bp	62	BnRTI-4 (Trypsin Inhibitor family)	
BnMTFL	144 bp	63		At2g21650 (put. Myb factor transcr. Factor)
BnGSTF	186 bp	64		At3g62760 (AtGSTF13)

BnSRP	131 bp	65		
BnPEF	121 bp	12	Pectinesterase family protein	At5g47500 (pectinesterase family protein)
BnWSP	83 bp	66		
BnLSP	164 bp	67		At5g62200 (weak similarity to embryo-specific protein 3)

From leaf material of *Brassica napus* cv. Kumily, genomic DNA has been isolated using the DNAeasy kit (Qiagen) according to the manufacturer's manual. Culture conditions for the *Brassica napus* cv. Kumily were as discussed above. Using this genomic DNA, as template, multiple rounds of thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) – a method known to a person skilled in the art – was performed to isolate sequences 5' upstream and 3' downstream of the 20 identified expressed fragments. The amplified products were either sequences directly, or subcloned into the pGEM-T (Promega) vector prior sequencing. Subcloning was required in some cases, as *Brassica napus* is amphidiploid, that is, *Brassica napus* contains 10 chromosomes common with *Brassica rapa* (A genome) and 9 chromosomes common with *Brassica oleracea* (C genome). Subcloning of PCR products containing a mixture of two sequences amplified from both genomes allows to separate these two sequences. Sequencing was done by standard techniques (laser fluorescent DNA-sequencing, ABI according to the method of Sanger et al. 1977 Proc. Natl. Acad. Sci. USA 74, 5463-5467).

For candidate expressed sequences where no *Brassica* gene was known to the art, the open reading frame was identified with help of software prediction and using alignments with homologous genes from *Arabidopsis*. The following *Brassica napus* open reading frames have been identified (Table 10):

Table 12: Open reading frames (ORF) identified in *Brassica napus* cDNA sequences.

Brassica napus Sequence	ORF in bp	SEQ ID NO:
BnGRPL	360	5
BnPEF	1083	13
BnSCT1	1401	20
BnSCT2	1401	21
BnSETL1	1512	27

BnSETL2	1512	28
BnSCP	773	72
BnCRU4	1107	80
BnMYR	1644	88
BnMDP	1152	98
BnRTI-4	300	106
BnMTFL	303	114
BnGSTF	660	121
BnLSP	574	133

For the expressed sequence SEQ-ID 126, no open reading frame could be identified.

5 The following Brassica napus sequences upstream of the expressed sequence SEQ-ID 126 or the identified expressed open reading frames (ORF) have been obtained (Table 13).:

Table 13: Genomic 5' upstream sequences from the Brassica napus cDNA sequences.

10

Brassica napus Sequence	genomic 5' sequence in bp	SEQ ID NO:
p-BnGRPL	1790	1
p-BnPEF-var1	2027	6
p-BnPEF-var2	636	9
p-BnSCT2-var1	1019	14
p-BnSCT2-var2	996	16
p-BnSETL-var1	1490	22
p-BnSETL-var2	2010	25
p-BnSCP	2052	70
p-BnCRU4	1951	77
p-BnMYR	1360	85
p-BnMDP	1428	95
p-BnRTI-4	1820	103
p-BnMTFL	1335	111
p-BnGSTF	1565	119
p-BnSRP	2447	124

p-BnLSP	1593	121
---------	------	-----

The analysis of the 5' upstream sequences using Genomatix software Gems-Launcher showed that the sequences comprised promoter elements. This was confirmed by the presence of a TATA-Box which is required for transcription by RNA-polymerases. Also in the isolated fragments elements specific for seed-transcription factors (e.g. Prolamin-box, legumin box, RITA etc.) were found.

The following Brassica napus sequences downstream of the identified expressed open reading frames (ORF) have been obtained (Table 14).

Table 14: Genomic 3' downstream sequences from the Brassica napus cDNA sequences.

Brassica napus Sequence	genomic 3' sequence in bp	SEQ ID NO:
t-BnGRPL	581	2
t-BnPEF-var1	477	7
t-BnPEF-var2	538	10
t-BnSCT2-var1	573	15
t-BnSCT2-var2	576	17
t-BnSETL-var1	614	23
t-BnSCP	587	71
t-BnCRU4	514	78
t-BnMYR	652	86
t-BnMDP	483	96
t-BnRTI-4	572	104
t-BnMTFL	521	112
t-BnSRP	865	125

Example 3: Production of test constructs for demonstrating promoter activity

For the testing of the promoter elements in a first step promoter terminator cassettes were generated. To this end, fusion PCRs have been used wherein via two PCR steps promoter elements were linked with terminator elements. At the same time, a multiple cloning site was introduced in between the promoter and terminator elements. The primers used to generate cassettes using corresponding native Brassica terminators are shown in Table 15, Table 16 list the generated cassettes using the OCS terminator.

Table 15: Primer pairs used for the generation of promoter-multiple cloning site-terminated-cassettes via Fusion-PCR using native Brassica terminator sequences.

Brassica napus Promoter/Termin ator cassette	Primer pair 1. PCR Promoter	Primer pair 1. PCR Terminator	Primer pair 2. PCR
p-BnGRP_MCS_t- BnGRP	Forw: ataccggga- tacctgcaggt- taggccggccacaa- aacgaaatcatca- aatcgtg (SEQ ID NO: 29) Rev: caatcaatta- taggcctcgcagctt- taattaacgatcgagc- catggtttagagagccg- gaattattg (SEQ ID NO: 30)	Forw: ccatggctc- gatcggttaattaaag- catgcgaggcctataatt- gattgtttctcctcttagttg- taaataatctatc (SEQ ID NO: 31) Rev: taagcggccg- caatcggaccga- taccggtaggcgccgat- tatctacctaacctaaca- aacaaaag (SEQ ID NO: 32)	Forw:ataccgggatac ctgcaggttaggccggc- cacaaaacgaaatcat- caaatcgtg (SEQ ID NO:29) Rev: 'taagcggccg- caatcggaccga- taccggtaggcgccgat- tatctacctaacctaaca- aacaaaag (SEQ ID NO:32)
p-BnPEF- var1_MCS_t- BnPEF-var1	Forw: ataccgggatacctgca ggtaggccggccaaattt attaacccatctattgttca c (SEQ ID NO:33) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatgggacgagaaa gaaaatggtcggag (SEQ ID NO:34)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgtaagggaatca actcaaatgcttttc (SEQ ID NO:35) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgacgttaagcgagagt ggatag (SEQ ID NO:36)	Forw: ataccgggatacctgca ggtaggccggccaaattt attaacccatctattgttca c (SEQ ID NO:33) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgacgttaagcgagagt ggatag (SEQ ID NO:36)
p-BnPEF- var2_MCS_t- BnPEF-var2	Forw: ataccgggatacctgca ggtaggccggccagaat tacgtttgagttcaaattca	Forw: Ccatggctcgatcgtaat taaagcatgcgaggccta taattgattgtaggaatca	Forw: ataccgggatacctgca ggtaggccggccagaat tacgtttgagttcaaattca

	g (SEQ ID NO:37) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatgggacgagaaa taaattggtcgaag (SEQ ID NO:38)	actcaaatgctttt (SEQ ID NO:39) Rev: taagcggccgcaatcgg accgataccggtaggcg ccagaggtgaggaggag ttgcac (SEQ ID NO:40)	g (SEQ ID NO:37) Rev: taagcggccgcaatcgg accgataccggtaggcg ccagaggtgaggaggag ttgcac (SEQ ID NO:40)
p-BnSCT2- var1_MCS_t- BnSCT2-var1	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:41) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggtttctctgcttct tggtgtcac (SEQ ID NO:42)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgAgttcctcact cacttctctc (SEQ ID NO:43) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgccttatatggattttgtg ttactgacc (SEQ ID NO:44)	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:41) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgccttatatggattttgtg ttactgacc (SEQ ID NO:44)
p-BnSCT2- var2_MCS_t- BnSCT2-var2	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:45) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggtttctctgcttct tggtgtcac (SEQ ID NO:46)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgAgttcctcact cacttctctc (SEQ ID NO:47) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgccttatatggattttgtg ttactgacc (SEQ ID NO:48)	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:45) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgccttatatggattttgtg ttactgacc (SEQ ID NO:48)
p-BnSETL- var1_MCS_t- BnSETL-var1	Forw: ataccgggatacctgca ggtaggccggccagtag	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat	Forw: ataccgggatacctgca ggtaggccggccagtag

	aagttattagcaactgtac acac (SEQ ID NO:49) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggttgaccccttc ttgttcttc (SEQ ID NO:50)	aattgattgtacatactatat ttttgtttaccttg (SEQ ID NO:51) Rev: taagcggccgcaatcgg accgataccggtaggcg cccaaacacggctcaga gaagc (SEQ ID NO:52)	aagttattagcaactgtac acac (SEQ ID NO:49) Rev: taagcggccgcaatcgg accgataccggtaggcg cccaaacacggctcaga gaagc (SEQ ID NO:52)
p-BnSETL- var2_MCS_t- BnSETL-var1	Forw: ataccgggatacctgca ggtaggccggccatcgg ctacaaatccaactgg (SEQ ID NO:53) Rev:- caatcaattataggcctcg catgctttaattaacgatcg agccatggttgtcgtttcc tcagcttc (SEQ ID NO:54)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgtacatactatat ttttgtttaccttg (SEQ ID NO:51) Rev: taagcggccgcaatcgg accgataccggtaggcg cccaaacacggctcaga gaagc (SEQ ID NO:52)	Forw: ataccgggatacctgca ggtaggccggccatcgg ctacaaatccaactgg (SEQ ID NO:53) Rev: taagcggccgcaatcgg accgataccggtaggcg cccaaacacggctcaga gaagc (SEQ ID NO:52)
p-BnSCP_MCS_t- BnSCP	Forw: ataccgggatacctgca ggtaggccggccaatca taagttgtatcagttcatc (SEQ-ID No:73) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatgggtgttgactcat actggtggta (SEQ-ID No:74)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgggagaaaat atgggagaagatggaa (SEQ-ID No:75) Rev: taagcggccgcaatcgg accgataccggtaggcg cctatagacctgccaatc aaaccaac (SEQ-ID No:76)	Forw: ataccgggatacctgca ggtaggccggccaatca taagttgtatcagttcatc (SEQ-ID No:73) Rev: taagcggccgcaatcgg accgataccggtaggcg cctatagacctgccaatc aaaccaac (SEQ-ID No:76)
p- BnCRU4_MCS_t-	Forw: ataccgggatacc-	Forw: ccatggctcgatcgtaat-	Forw: ataccgggatacc-

BnCRU4	tgcaggtagg- ccggccaatgtacatg- gatgcgtatagatg (SEQ-ID No:81) Rev: caatcaattatagg- cctcgcatgctttaat- taacgatcgagccatg- gtagagcatg- tccgtgaacgacg (SEQ-ID No:82)	taaagcatgcgagg- cctataattgattgagcat- gagttagtgatgtaa- cagcg (SEQ-ID No:83) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccggaaga- gatggaagcttaca- gaatg (SEQ-ID No:84)	tgcaggtagg- ccggccaatgtacatg- gatgcgtatagatg (SEQ-ID No:81) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccggaaga- gatggaagcttaca- gaatg (SEQ-ID No:84)
p-BnMYR_MCS_t- BnMYR	Forw: ataccgggatacctgca ggtaggccggccacata aaaatatgttgagaaaat atc (SEQ-ID No:89) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggggttaatgtgt agattgtatatatg (SEQ-ID No:90)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgaacactttatc cacatcaagatcgc (SEQ-ID No:91) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgatgtgtgctcgtaattg cactttt (SEQ-ID No:92)	Forw: ataccgggatacctgca ggtaggccggccacata aaaatatgttgagaaaat atc (SEQ-ID No:89) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgatgtgtgctcgtaattg cactttt (SEQ-ID No:92)
p-BnMDP_MCS_t- BnMDP	Forw: ataccgggatacctgca ggtaggccggccatgcg agagagaagaaatgaa ataag (SEQ-ID No:99) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggcttttgagattg tatatatgaatg (SEQ-	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgatttctctaact aagaaactttgtag (SEQ-ID No:101) Rev: taagcggccgcaatcgg accgataccggtaggcg cccgataaagatgggtcta atgtccatc (SEQ-ID	Forw: ataccgggatacctgca ggtaggccggccatgcg agagagaagaaatgaa ataag (SEQ-ID No:99) Rev: taagcggccgcaatcgg accgataccggtaggcg cccgataaagatgggtcta atgtccatc (SEQ-ID

	ID No:100)	No:102)	No:102)
p-BnRTI-4_MCS_t-BnRTI-4	Forw: ataccgggatacctgca ggtagggccggccacttg cgccgaagatatatccga c (SEQ-ID No:107) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggcttctctata tatctcttac (SEQ-ID No:108)	Forw: ccatggctcgatcgtaatt aaagcatgagggcctat aattgattgatacttcacaa ctttgcataagcc (SEQ-ID No:109) Rev: taagcggccgcaatcgg accgataccggtaggcg ccaccaaaccgaaacc gaatccgaac (SEQ- ID No:110)	Forw: ataccgggatacctgca ggtagggccggccacttg cgccgaagatatatccga c (SEQ-ID No:107) Rev: taagcggccgcaatcgg accgataccggtaggcg ccaccaaaccgaaacc gaatccgaac (SEQ- ID No:110)
p-BnMTFL_MCS_t-BnMTFL	Forw: ataccgggatacctgca ggtagggccggccatattc actactatagagaacac (SEQ-ID No:115) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatgggtgtgaaaaag ttggcgtatttc (SEQ- ID No:116)	Forw: ccatggctcgatcgtaatt aaagcatgagggcctat aattgattgaacaagaag caacaaagcctaaactat (SEQ-ID No:117) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgatttgagtttcgtgaa gttcaga (SEQ-ID No:118)	Forw: ataccgggatacctgca ggtagggccggccatattc actactatagagaacac (SEQ-ID No:115) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgatttgagtttcgtgaa gttcaga (SEQ-ID No:118)
p-BnSRP_MCS_t-BnSRP	Forw: ataccgggatacctgca ggtagggccggccaaac ggcaattgattctcgccct g (SEQ-ID No:127) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggagagaggtta ttgaaaccacgt (SEQ-	Forw: ccatggctcgatcgtaatt aaagcatgagggcctat aattgattgttacatattgg cccaagaggcataat (SEQ-ID No:129) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgttaatgaaaacggta	Forw: ataccgggatacctgca ggtagggccggccaaac ggcaattgattctcgccct g (SEQ-ID No:127) Rev: taagcggccgcaatcgg accgataccggtaggcg ccgttaatgaaaacggta ccgaatttc (SEQ-ID

	ID No:128)	ccgaatttc (SEQ-ID No:130)	No:130)
--	------------	------------------------------	---------

Table 16: Primer pairs used for the generation of promoter-multiple cloning site-terminated-cassettes via Fusion-PCR using the OCS terminator sequence.

5

p-BnGRP_MCS_t-OCS	Forw: ataccggga-tacctgcaggtagg-ccggccacaaaac-gaaatcatcaaatcgtg (SEQ ID NO: 29) Rev: caatcaatta-taggcctcgcgcatg-ctttaattaacgacgagccatggttttaga-gagccggaattattg (SEQ ID NO: 30)	Forw: ccatg-gctcgatcgtaat-taaagcatgcgagg-cctataattgatt-gctgctttaatgagatatg-cgagacg (SEQ ID NO:55) Rev: taagc-ggccgcaatcggacc-gataccggtagg-cgccacaatcag-taaattgaacggagaa-tattattc (SEQ ID NO:56)	Forw: ataccggga-tacctgcaggtagg-ccggccacaaaac-gaaatcatcaaatcgtg (SEQ ID NO: 29) Rev: taagc-ggccgcaatcggacc-gataccggtagg-cgccacaatcag-taaattgaacggagaa-tattattc (SEQ ID NO:56)
p-BnPEF-var1_MCS_t-OCS	Forw: ataccggga-tacctgcaggtagg-ccggccaaatttat-taaccatctattgttcac (SEQ ID NO:33) Rev: caatcaatta-taggcctcgcgcatg-ctttaattaacgacgagccatgggacgagaaagaaatg-gtcggag (SEQ ID NO:34)	Forw: ccatg-gctcgatcgtaat-taaagcatgcgagg-cctataattgatt-gctgctttaatgagatatg-cgagacg (SEQ ID NO:55) Rev: taagc-ggccgcaatcggacc-gataccggtagg-cgccacaatcag-taaattgaacggagaa-tattattc (SEQ ID NO:56)	Forw: ataccggga-tacctgcaggtagg-ccggccaaatttat-taaccatctattgttcac (SEQ ID NO:33) Rev: taagc-ggccgcaatcggacc-gataccggtagg-cgccacaatcag-taaattgaacggagaa-tattattc (SEQ ID NO:56)

p-BnPEF- var2_MCS_t-OCS	Forw: ataccggga- tacctgcaggtagg- ccggccagaattacg- tttgagttcaaattcag (SEQ ID NO:37) Rev: caatcaatta- taggcctcgcg- ctttaattaacgatc- gagccatgggacga- gaaataaatggcgaag (SEQ ID NO:38)	Forw: ccatg- gctcgatcgtaat- taaagcatgcgagg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ ID NO:55) Rev: taagc- ggccgcaatcggacc- gataccggtagg- cgccacaatcag- taaattgaacggagaa- tattattc (SEQ ID NO:56)	Forw: ataccggga- tacctgcaggtagg- ccggccagaattacg- tttgagttcaaattcag (SEQ ID NO:37) Rev: taagc- ggccgcaatcggacc- gataccggtagg- cgccacaatcag- taaattgaacggagaa- tattattc (SEQ ID NO:56)
p-BnSCT2- var1_MCS_t-OCS	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:41) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggtttctctgcttct tggtgtcac (SEQ ID NO:42)	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgctgctttaatga gatatgcgagacg (SEQ ID NO:55) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:41) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)
p-BnSCT2- var2_MCS_t-OCS	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:45) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggtttctctgcttct	Forw: ccatggctcgatcgtaatt aaagcatgcgaggcctat aattgattgctgctttaatga gatatgcgagacg (SEQ ID NO:55) Rev: taagcggccgcaatcgg accgataccggtaggcg	Forw: ataccgggatacctgca ggtaggccggccagatg caaaaacgtatagtcaca c (SEQ ID NO:45) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga

	tggtgtcac (SEQ ID NO:46)	ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)	acggagaatattattc (SEQ ID NO:56)
p-BnSETL- var1_MCS_t-OCS	Forw: ataccgggatacctgca ggtaggccggccagtag aagttattagcaactgtac acac (SEQ ID NO:49) Rev: caatcaattataggcctcg catgctttaattaacgatcg agccatggttgaccccttc ttgttcttc (SEQ ID NO:50)	Forw: ccatggctcgatcgtaatt aaagcatgagggcctat aattgattgctgctttaatga gatatgagagacg (SEQ ID NO:55) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)	Forw: ataccgggatacctgca ggtaggccggccagtag aagttattagcaactgtac acac (SEQ ID NO:49) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)
p-BnSETL- var2_MCS_t-OCS	Forw: ataccgggatacctgca ggtaggccggccatcgg ctacaaatccaactgg (SEQ ID NO:53) Rev:- caatcaattataggcctcg catgctttaattaacgatcg agccatggttgctgctttcc tcagcttc (SEQ ID NO:54)	Forw: ccatggctcgatcgtaatt aaagcatgagggcctat aattgattgctgctttaatga gatatgagagacg (SEQ ID NO:55) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)	Forw: ataccgggatacctgca ggtaggccggccatcgg ctacaaatccaactgg (SEQ ID NO:53) Rev: taagcggccgcaatcgg accgataccggtaggcg ccacaatcagtaaattga acggagaatattattc (SEQ ID NO:56)
p-BnSCP_MCS_t- OCS	Forw: ataccgggatacctg- caggtaggccggc- caatcataagttgtat- cagttcatc (SEQ-ID No:73) Rev: caatcaatta-	Forw: ccatggctcgatcgtaat- taaagcatgagggc- tataattgattgctgctt- taatgagatatgcga- gacg (SEQ-ID No:55) Rev:	Forw: ataccgggatacctg- caggtaggccggc- caatcataagttgtat- cagttcatc (SEQ-ID No:73) Rev: taagcggccgcaatcg-

	taggcctcgcgatgctt- taattaacgatcgagc- catggtgtttgactca- tactggtggta (SEQ-ID No:74)	taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)	gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)
p- BnCRU4_MCS_t- OCS	Forw: atacccgggatacc- tgcaggtagg- ccggccaatgtacatg- gatgcgtatagatg (SEQ-ID No:81) Rev: caatcaattatagg- cctcgcgatgctttaat- taacgatcgagccatg- gtagagcatg- tccgtgaacgacg (SEQ-ID No:82)	Forw: ccatggctcgatcgtaat- taaagcatgcgagg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	Forw: atacccgggatacc- tgcaggtagg- ccggccaatgtacatg- gatgcgtatagatg (SEQ-ID No:81) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
p-BnMYR_MCS_t- OCS	Forw: atacccgggatacc- tgcaggtagg- ccggccacataaaaa- tatgttgagaaaatac (SEQ-ID No:89) Rev: caatcaattatagg- cctcgcgatgctttaat- taacgatcgagccatg- gggttaatgtgtagatt- tgtatatatg (SEQ-ID No:90)	Forw: ccatggctcgatcgtaat- taaagcatgcgagg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	Forw: atacccgggatacc- tgcaggtagg- ccggccacataaaaa- tatgttgagaaaatac (SEQ-ID No:89) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
p-BnMDP_MCS_t-	Forw:	Forw:	Forw:

OCS	ataccgggatacctg- caggtaggccggc- catgagagagagaaga- atgaaataag (SEQ- ID No:99) Rev: caatcaatta- taggcctcgcatgctt- taattaacgatcgagc- catggctttgagattgta- tatatgaatg (SEQ-ID No:100)	ccatggctcgatcgtaat- taaagcatgagggcc- tataattgattgctgctt- taatgagatatgcga- gacg (SEQ-ID No:55) Rev: taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)	ataccgggatacctg- caggtaggccggc- catgagagagagaaga- atgaaataag (SEQ- ID No:99) Rev: taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)
p-BnRTI- 4_MCS_t-OCS	Forw: ataccgggataacc- tgcaggtagg- ccggccact- tgcgccgaagatatatc- cgac (SEQ-ID No:107) Rev: caatcaattatagg- cctcgcatgctttaat- taacgatcgagccatg- gcttctcttatatc- tcttac (SEQ-ID No:108)	Forw: ccatggctcgatcgtaat- taaagcatgaggg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	Forw: ataccgggataacc- tgcaggtagg- ccggccact- tgcgccgaagatatatc- cgac (SEQ-ID No:107) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
p- BnMTFL_MCS_t- OCS	Forw: ataccgggataacc- tgcaggtagg- ccggccatattcactact- tatagagaacac (SEQ- ID No:115) Rev: caatcaattatagg-	Forw: ccatggctcgatcgtaat- taaagcatgaggg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev:	Forw: ataccgggataacc- tgcaggtagg- ccggccatattcactact- tatagagaacac (SEQ- ID No:115) Rev: taagcggccgcaatc-

	cctcgcatgctttaat- taacgatcgagccatg- gtgtgaaaaagt- ttggcgtatttc (SEQ-ID No:116)	taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
p- BnGSTF_MCS_t- OCS	Forw: atacccgggatacc- tgcaggtagg- ccggccagagcattatg- cttccaagcggac (SEQ-ID No:122) Rev: caatcaattatagg- cctcgcatgctttaat- taacgatcgagccatg- gttcaacttaact- tcgtaactcg (SEQ-ID No:123)	Forw: ccatggctcgatcgtaat- taaagcatgcgagg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	Forw: atacccgggatacc- tgcaggtagg- ccggccagagcattatg- cttccaagcggac (SEQ-ID No:122) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
p-BnSRP_MCS_t- OCS	Forw: atacccgggatacctg- caggtaggccggcca- aacggcaatt- gattctcgccctg (SEQ- ID No:127) Rev: caatcaatta- taggcctcgcatgctt- taattaacgatcgagc- catggagagaggttatt- gaaaccacgt (SEQ- ID No:128)	Forw: ccatggctcgatcgtaat- taaagcatgcgaggcc- tataattgattgctgctt- taatgagatatgcga- gacg (SEQ-ID No:55) Rev: taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)	Forw: atacccgggatacctg- caggtaggccggcca- aacggcaatt- gattctcgccctg (SEQ- ID No:127) Rev: taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)
p-BnLSP_MCS_t-	Forw:	Forw:	Forw:

OCS	ataccgggatacctg- caggtaggccggcca- aatagtcaagtttatgaat- cacag (SEQ-ID No:134) Rev: caatcaatta- taggcctcgcgctgctt- taattaacgatcgagc- catggtcttgaacttctt- gacattact (SEQ-ID No:135)	ccatggctcgatcgttaat- taaagcatgcgaggcc- tataattgattgctgctt- taatgagatatgcga- gacg (SEQ-ID No:55) Rev: taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)	ataccgggatacctg- caggtaggccggcca- aatagtcaagtttatgaat- cacag (SEQ-ID No:134) Rev: taagcggccgcaatcg- gaccgataccgg- taggcgccacaatcag- taaattgaacggagaa- tattattc (SEQ-ID No:56)
p-Napin_MCS_t- OCS	Forw: ataccgggataacc- tgcaggtagg- ccggccataaggat- gacctaccattcttga (SEQ-ID No:138) Rev: caatcaattatagg- cctcgcatgctttaat- taacgatcgagccatg- gtgttttaatacttgtttgtatt (SEQ-ID No:139)	Forw: ccatggctcgatcgttaat- taaagcatgcgagg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	Forw: ataccgggataacc- tgcaggtagg- ccggccataaggat- gacctaccattcttga (SEQ-ID No:138) Rev: taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
p-LuPXR_MCS_t- OCS	Forw: ataccgggataacc- tgcaggtagg- ccggccacacgggcag- gacatagggactact (SEQ-ID No:142) Rev: caatcaattatagg-	Forw: ccatggctcgatcgttaat- taaagcatgcgagg- cctataattgatt- gctgctttaatgagatatg- cgagacg (SEQ-ID No:55) Rev:	Forw: ataccgggataacc- tgcaggtagg- ccggccacacgggcag- gacatagggactact (SEQ-ID No:142) Rev: taagcggccgcaatc-

	cctcgcatgctttaat- taacgatcgagccatg- ggatttatgataaaaatg- tcggt (SEQ-ID No:143)	taagcggccgcaatc- ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)	ggaccgatacc- ggtaggcgccacaat- cagtaaattgaacgga- gaatattattc (SEQ-ID No:56)
--	---	--	--

The promoter-terminator cassettes were cloned into the pCR2.1 (Invitrogen) vector according to the manufacturer's manual and subsequently sequenced. In a further step, the delta 6 Desaturase Gene (SEQ ID NO: 68) was introduced via the *NcoI*, *PacI* restrictions site between the promoter and terminator sequence.

Using the Multisite Gateway System (Invitrogen), a multiple cloning site (SEQ ID 57) was introduced into each of the three pENTR vectors pENTR/A pENTR/B and pENTR/C via *HindIII* and *KpnI* restrictions sites. Into the first position of this MCS, the promoter-delta 6 Desaturase-terminator cassette was cloned via *FseI* and *KasI*. Similarly, the DsRed gene was introduced into pENTR/C between the Napin promoter and the OCS terminator. By performing a site specific recombination (LR-reaction), the created pENTR/B, pENTR/C and an empty pENTR/A vector were combined with the pSUN destination vector according to the manucaturers (Invitrogen) Multisite Gateway manual to generate the final binary vectors SEQ ID: 3 pSUN-p-GRPL_d6Des(Pir)_t-OCS, SEQ ID: 8 pSUN-p-PEF-var1_d6Des(Pir)_t-OCS, SEQ ID: 11 pSUN-p-PEF-var2_d6Des(Pir)_t-OCS, SEQ ID: 18 pSUN-p-SCT2-var2_d6Des(Pir)_t-OCS, SEQ ID: 24 pSUN-p-SETL-var1_t-OCS, SEQ ID: 79 pSUN-pBnCRU4_d6Des(Pir)_t-OCS, SEQ ID: 87 pSUN-pBnMYR_d6Des(Pir)_t-OCS, SEQ ID: 93 pSUN-pBnSETL-var2_d6Des(Pir)_t-OCS, SEQ ID: 94 pSUN-pBnSCT2-var1_d6Des(Pir)_t-OCS, SEQ ID: 97 pSUN-pBnMDP_d6Des(Pir)_t-OCS, SEQ ID: 105 pSUN-pBnRTI-4_d6Des(Pir)_t-OCS, SEQ ID: 113 pSUN-pBnMTFL_d6Des(Pir)_t-OCS, SEQ ID: 120 pSUN-pBnGSTF_d6Des(Pir)_t-OCS, SEQ ID: 132 pSUN-pBnLSP_d6Des(Pir)_t-OCS. Similarly, the two binary vectors SEQ ID: 137 pSUN-pNapin_d6Des(Pir)_t-OCS and SEQ ID: 141 pSUN-pLuPXR_d6Des(Pir)_t-OCS were cloned as a positive control, that is, these two vectors are known to be capable to drive seed specific expression of PUFA genes, e.g. the delta-6-desaturase SEQ ID NO: 68.

The resulting vectors were subsequently used for the production of transgenic plants. The promoter activity in the transgenic plant seeds was measured based on the expression of delta 6 Desaturase and an observed modification in the lipid pattern of the seeds as described in example 5.

Example 4: Production of transgenic plants

a) Generation of transgenic rape seed plants (amended protocol according to Moloney et al. 1992, Plant Cell Reports, 8:238-242)

5 For the generation of transgenic rapeseed plants, the binary vectors were transformed into *Agrobacterium tumefaciens* C58C1:pGV2260 (Deblaere et al. 1984, Nucl. Acids. Res. 13: 4777-4788). For the transformation of rapeseed plants (cv. Kumily,) a 1:50 dilution of an overnight culture of positive transformed acrobacteria colonies grown in Murashige-Skoog Medium (Murashige and Skoog 1962 Physiol. Plant. 15, 473) supplemented by 3% saccharose (3MS-Medium) was used. Petiols or Hypocotyledones of
10 sterial rapeseed plants were incubated in a petri dish with a 1:50 acrobacterial dilution for 5-10 minutes. This was followed by a three day co-incubation in darkness at 25°C on 3MS-Medium with 0.8% bacto-Agar. After three days the culture was put on to 16 hours light/8 hours darkness weekly on MS-medium containing 500mg/l Claforan (Cefotaxime-Natrium), 100 nM Imazetapyr, 20 mikrom Benzylaminopurin (BAP) and 1,6g/l
15 Glucose. Growing sprouts were transferred to MS-Medium containing 2% saccharose, 250mg/l Claforan and 0.8%Bacto-Agar. Even after three weeks no root formation was observed, a growth hormone 2-Indolbutyl acid was added to the medium for enhancing root formation.

20

Regenerated sprouts have been obtained on 2MS-Medium with Imazetapyr and Claforan and were transferred to the green house for sprouting. After flowering, the mature seeds were harvested and analysed for expression of the Desaturase gene via lipid analysis as described in Qui et al. 2001, J. Biol. Chem. 276, 31561-31566.

25

b) Production of transgenic flax plants

The production of transgenic flax plants can be carried out according to the method of Bell et al., 1999, In Vitro Cell. Dev. Biol. Plant 35(6):456-465 using particle bombardment. Acrobacterial transformation could be carried out according to Mlynarova et al.
30 (1994), Plant Cell Report 13: 282-285.

Example 5: Lipid Extraction

35

Lipids can be extracted as described in the standard literature including Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et al. (1993) Biotechnology, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH:
40

Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988) Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications.

Alternatively, extraction will be carried out as described in Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145. Quantitative and qualitative analysis of lipids or fatty acids are described in Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

Based on the analysed lipids, the expression of the Desaturase can be determined since the lipid pattern of successfully transformed plant seeds will differ from the pattern of control plant seeds. Seed specific expression of a delta-6-desaturase would result in formation of 18:3n-6 (GLA) and/or 18:4n-3 (SDA), depending on whether the delta 6 desaturase uses 18:2n-6 (LA) and/or 18:3n-3 (ALA) as substrate. Surprisingly, not only the two control promoters Napin and LuPXR harbored by the vectors SEQ ID: 137 pSUN-pNapin_d6Des(Pir)_t-OCS and SEQ ID: 141 pSUN-pLuPXR_d6Des(Pir)_t-OCS were capable to drive seed specific expression of the delta-6-desaturase as indicated by the formation of GLA, but also the promoters of the present invention (figure 2). Interestingly, the promoters influenced the ratio of the omega-3 fatty acid SDA to the omega-6 fatty acid GLA.

100

Claims

1. An expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or SEQ ID NO: 25; wherein said expression control sequence governs seed-specific expression of said operatively linked nucleic acid of interest.
2. The expression cassette of claim 1, wherein said expression control sequence comprises the nucleic acid sequence of SEQ ID NO: 22 or SEQ ID NO: 25.
3. A vector comprising the expression cassette of claim 1 or 2.
4. The vector of claim 3, wherein said vector is an expression vector.
5. A plant cell comprising the expression cassette of claim 1 or 2, or the vector of claim 3 or 4.
6. A seed cell comprising the expression cassette of claim 1 or 2, or the vector of claim 3 or 4.
7. A method of expressing a nucleic acid of interest in a plant cell or seed cell comprising:
 - (a) introducing an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or 25 or a vector comprising said expression control sequence, into said plant cell or seed cell; and
 - (b) expressing said nucleic acid of interest in said plant cell or seed cell.
8. The method of claim 7, wherein said host cell is a plant cell.
9. A method of expressing a nucleic acid of interest in a plant cell or seed cell comprising:

101

- 5 (a) introducing an expression cassette comprising an expression control sequence operatively linked to a nucleic acid of interest, wherein said expression control sequence is heterologous to and governs seed-specific expression of said nucleic acid of interest, and wherein said expression control sequence is at least 95% identical to the nucleic acid sequence of SEQ ID NO: 22 or 25 or a vector comprising said expression control sequence, into said plant cell or seed cell; and
- (b) expressing said nucleic acid of interest in said plant cell or seed cell.

- 10 10. The method of claim 9, wherein said nucleic acid of interest is expressed seed-specifically.
11. The method of claim 7, wherein said nucleic acid of interest is expressed seed-specifically.
- 15 12. The expression cassette of claim 1, wherein said expression control sequence is at least 98% identical to the nucleic acid sequence of SEQ ID NO: 22 or SEQ ID NO: 25.
- 20 13. The expression cassette of claim 1, wherein said expression control sequence is at least 99% identical to the nucleic acid sequence of SEQ ID NO: 22 or SEQ ID NO: 25.

Fig. 1

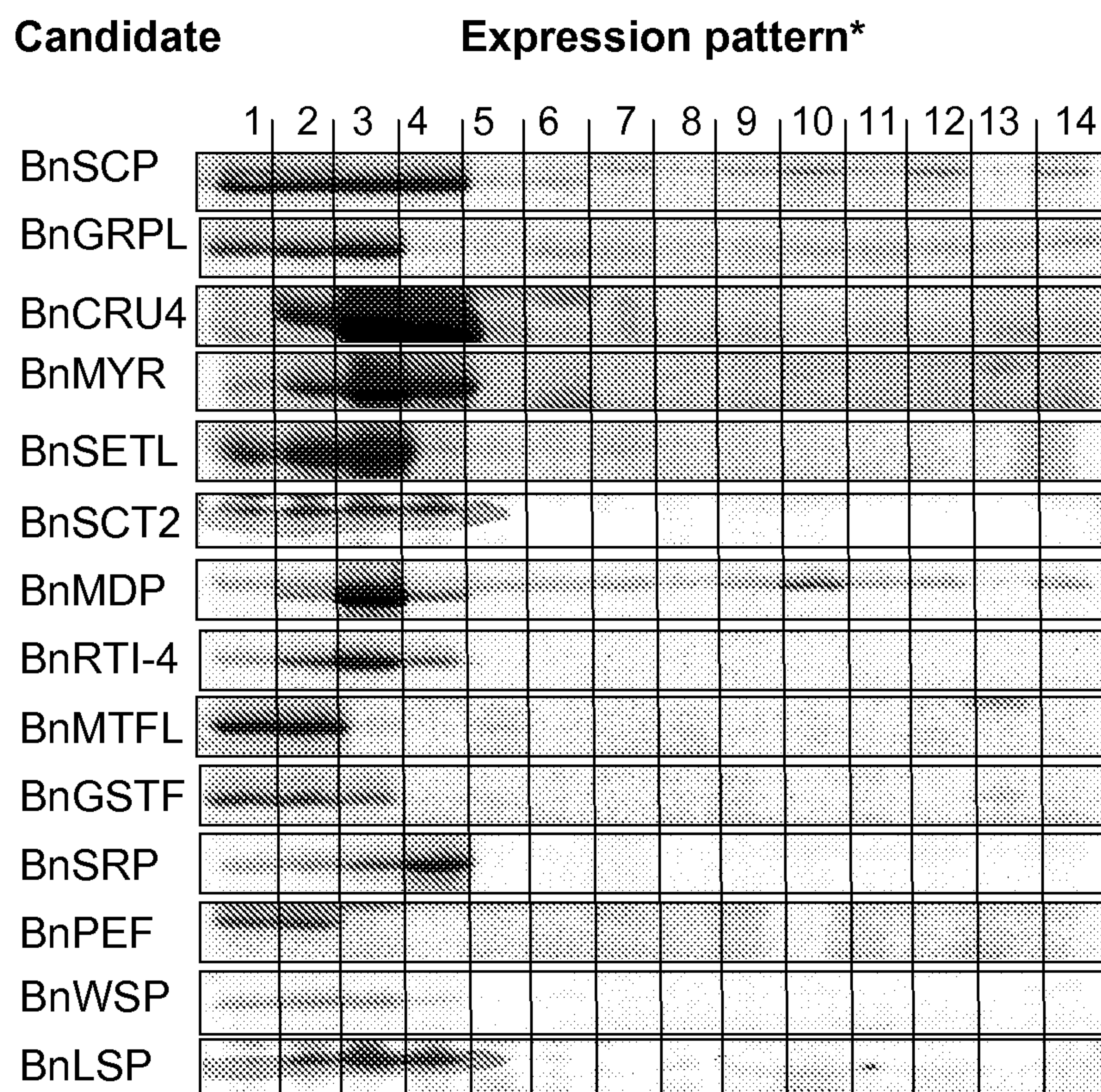


Fig. 2

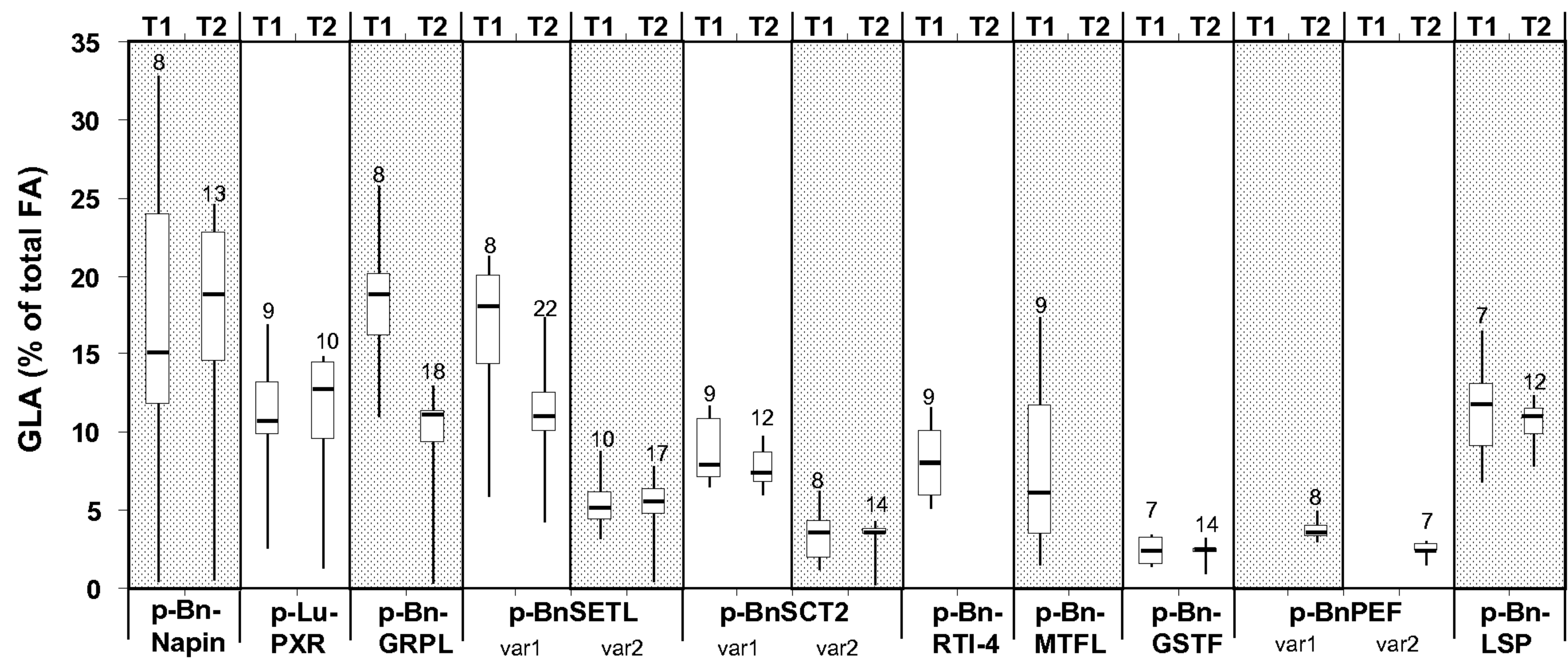


Fig. 3

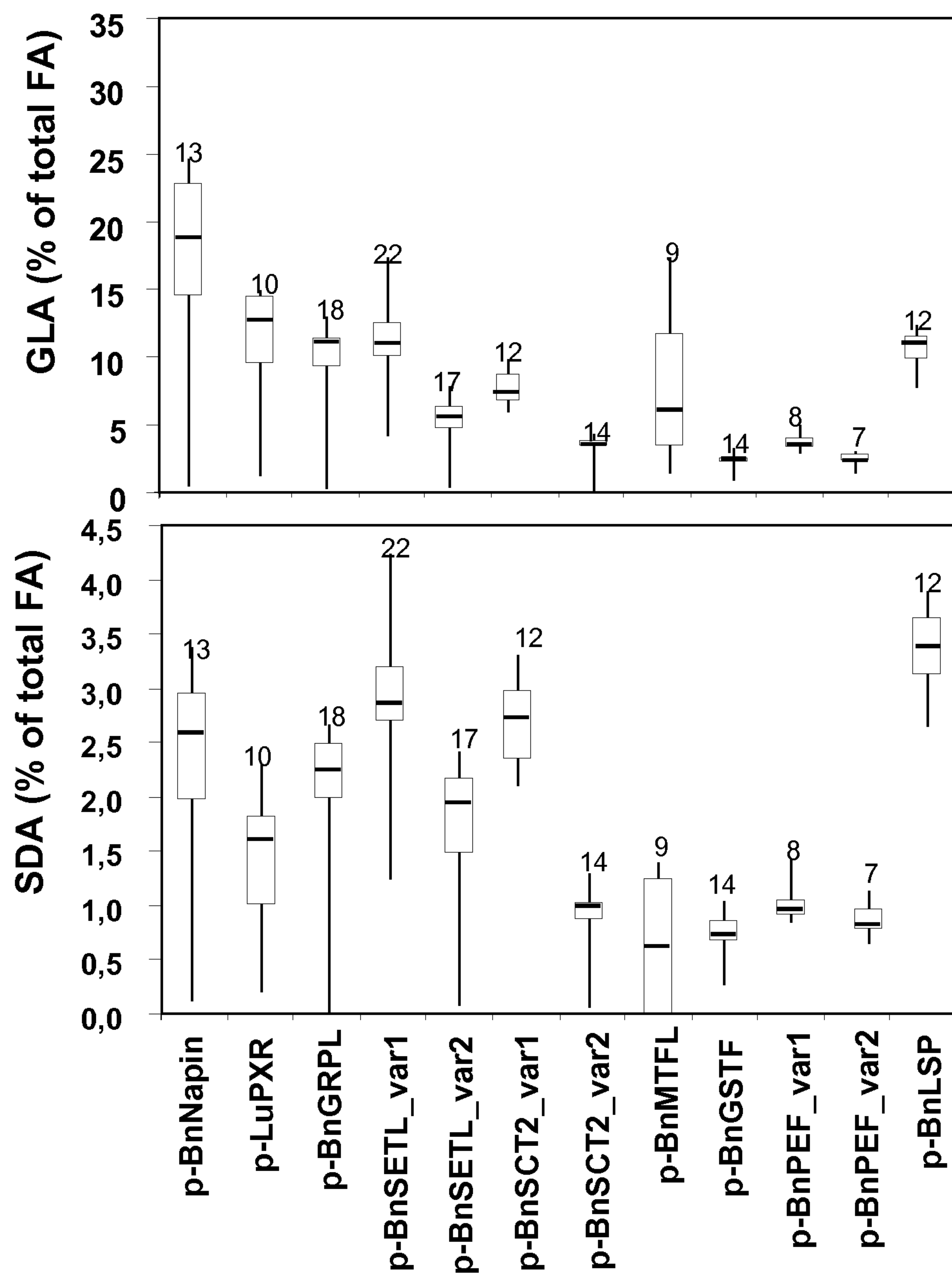


Fig. 4

