

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
25 June 2009 (25.06.2009)

PCT

(10) International Publication Number
WO 2009/077478 A2

(51) International Patent Classification:
C12N 15/82 (2006.01)

(21) International Application Number:
PCT/EP2008/067492

(22) International Filing Date:
15 December 2008 (15.12.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
07150029.2 14 December 2007 (14.12.2007) EP

(71) Applicants (for all designated States except US): **BASF PLANT SCIENCE GMBH** [DE/DE]; 67056 Ludwigshafen (DE). **Bioriginal Food & Science Corp.** [CA/CA]; 102 Melville Street, Saskatoon, Saskatchewan S7J 0R1 (CA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BAUER, Jörg** [DE/DE]; Maxburgstrasse 6, 67117 Limburgerhof (DE).

WETJEN, Tom [DE/DE]; Meerwiesenstrasse 21, 68163 Mannheim (DE). **QIU, Yiao** [CA/CA]; 403 Kendarine Road, Saskatoon, Saskatchewan S7N 3S5 (CA). **WU, Guohai** [CA/CA]; 2103 Kendarine Road, Saskatoon, Saskatchewan S7N 4A9 (CA).

(74) Agent: **DICK, Alexander**; Isenbruck Bösl Hörschler Wichmann Huhn LLP, Theodor-Heuss-Anlage 12, 68165 Mannheim (DE).

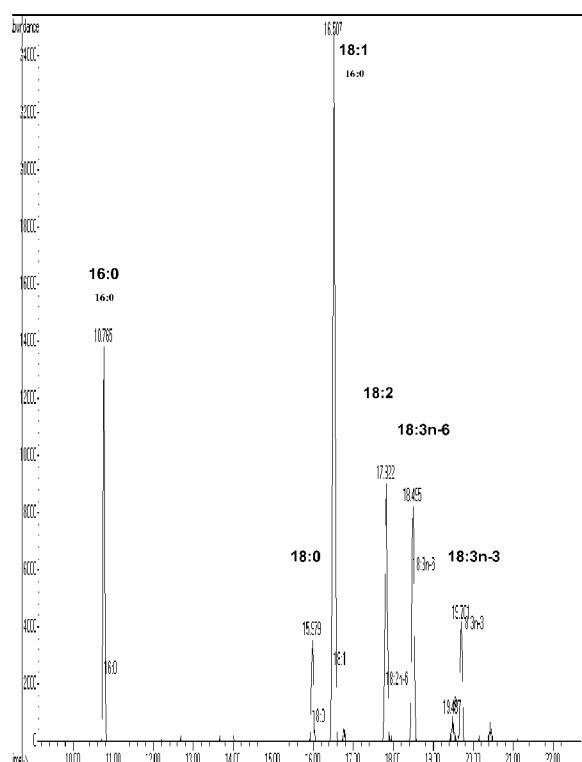
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, 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,

[Continued on next page]

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

Fig.1



(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 2009/077478 A2



GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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:

- *without international search report and to be republished upon receipt of that report*
- *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

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.

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.

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.

However, there is a clear need for further promoters which allow for a reliable and efficient 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: 7 to 12;
 - (b) 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: 7 to 12;
 - 20 (c) 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: 1 to 6;
 - (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 being at least 80% identical to an open reading frame sequence as shown in any one of SEQ ID NOs: 1 to 6;
 - 25 (e) an expression control sequence obtainable by 5' genome walking from an open reading frame sequence as shown in any one of SEQ ID NOs: 1 to 6; and
 - 30 (f) an expression control sequence obtainable by 5' genome walking from an open reading frame sequence being at least 80% identical to an open reading frame as shown in any one of SEQ ID NOs: 1 to 6.

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 for 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 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: 7 to 12.

- 5 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: 1 to 6, 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: 1 to 6. 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. 1 to 6, i.e. the expression control sequences shown in any one of SEQ ID NOs.: 7 to 12. 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 $0.1 \times$ 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, $0.1 \times$ SSC and 30°C to 55°C , preferably between 45°C and 55°C . 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 least 99% identical to the expression control sequences as shown in any one of SEQ ID NOs.: 7 to 12. 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 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 et al.,

CABIOS, 5 1989: 151-153) or the programs Gap and BestFit [Needleman and Wunsch (J. Mol. Biol. 48; 443-453 (1970)) and Smith and Waterman (Adv. Appl. Math. 2; 482-489 (1981))], which are part of the GCG software packet [Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711 (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.

10

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. 1 to 6. Rather, expression control sequences which allow for seed specific expression can also be found upstream of orthologous, paralogous or homologous genes (i.e. open reading frames). Thus, also preferably, a 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 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: 1 to 6. 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 one of SEQ ID NOs.: 1 to 6. In this context it should be mentioned that the open reading frame shown in SEQ ID NO: 1 encodes a polypeptide having pectinesterase activity, the open reading frames shown in SEQ ID NO: 2 and 5 encode "late embryogenesis abundant" (LEA) polypeptides, the open reading frame shown in SEQ ID NO: 3 encodes a polypeptide having anthocyanidin reductase activity, the open reading frame shown in SEQ ID NO: 4 encodes a polypeptide having proteinase inhibitor activity, and the open reading frame shown in SEQ ID NO: 6 encodes a polypeptide having lipid transfer activity. These biological activities can be determined by those skilled in the art without further ado.

Also preferably, a variant expression control sequence comprised by a polynucleotide of the present invention is (i) obtainable by 5' genome walking from an open reading frame sequence as shown in any one of SEQ ID NOs: 1 to 6 or (ii) obtainable by 5' genome walking 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: 1 to 6. Variant expression control sequences are obtainable without further by the genome walking technology

35

which can be carried out as described in the accompanying Examples by using, e.g., commercially available kits.

Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 7, preferably, comprise at least 80, at least 90, at least 100, at least 110, at least 120 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: 8, 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 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 80, at least 90, at least 100, at least 110 or all of the sequence motifs recited in Table 3. Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 10, preferably, comprise at least 40, at least 50, at least 60, at least 70 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: 11, preferably, comprise at least 80, at least 150, at least 200, at least 210, at least 220, at least 230, at least 240 or all of the sequence motifs recited in Table 5. Variant expression control sequences referred to in this specification for the expression control sequence shown in SEQ ID NO: 12, preferably, comprise at least 80, at least 90, at least 100, at least 110, at least 120 or all of the sequence motifs recited in Table 6. Specifically, the following elements are preferably comprised by all variant expression control sequences referred to in accordance with the present invention: CA-rich element, CCAAT box, G-box binding factor 1, RY repeat element, Prolamin box legumin box, Dof box and RITA motif. The specific sequences for the elements are shown in the Tables, below (marked in bold). These elements are characteristic for seed-specific promoters (Kim 2006, *Mol Genet Genomics* 276(4):351-368).

The term "seed specific" as used herein means that a nucleic acid of interest being 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-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.

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, most preferably, at 40 days after flowering.

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 person 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 the 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.

As set forth above, the polynucleotide of the present invention, in a preferred embodiment, comprises also a termination sequence for transcription downstream of the nucleic acid of interest. A termination sequence for transcription relates to a nucleic acid sequence which terminates the process of RNA transcription. Suitable termination sequences are well known in the art and comprise, preferably, the SV40-poly-A site, the tk-poly-A site, the nos or ocs terminator from *Agrobacterium tumefaciens* or the 35S terminator from Cauliflower mosaic virus.

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 eicosatetraenoic 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.

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

The term "vector", preferably, encompasses phage, plasmid, viral or retroviral vectors as well as artificial chromosomes, such as bacterial or yeast artificial chromosomes.

10 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

15 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

20 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 co-

25 precipitation, 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

30 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

35 host/cells.

Preferably, the vector referred to herein is suitable as a cloning vector, i.e. replicable in

40 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 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. 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.

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).

Suitable expression vector backbones are, preferably, derived from expression vectors known in the art such as Okayama-Berg cDNA expression vector pcDV1 (Pharmacia),
5 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
10 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
15 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 pYeDesaturasec1 (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
20 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
25 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
30 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).

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,
35 *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 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",
40 *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.

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 is 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*, Saccharo-

mycetaceae 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 satumus*, *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 acetii*, *Saccharomyces baillii*, *Saccharomyces bayanus*, *Saccharomyces bisporus*, *Saccharomyces capensis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces cerevisiae* var. *ellipsoideus*, *Saccharomyces chevalieri*, *Saccharomyces delbrueckii*, *Saccharomyces diastaticus*, *Saccharomyces drosophilae*, *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*, *Thraustochytrium aureum*, *Thraustochytrium benthiicola*, *Thraustochytrium globosum*, *Thraustochytrium indicum*, *Thraustochytrium kerguelense*, *Thraustochytrium kinnei*, *Thraustochytrium motivum*, *Thraustochytrium multirudimentale*, *Thraustochytrium pachydermum*, *Thraustochytrium proliferum*, *Thraustochytrium roseum*, *Thraustochytrium rossii*,

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

5 *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 poly-*

10 *myxa*, *Bacillus psychrosaccharolyticus*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus subtilis* subsp. *spizizenii*, *Bacillus subtilis* subsp. *subtilis* or *Bacillus thuringiensis*; Enterobacteriaceae such as the genera *Citrobacter*, *Edwardsiella*, *Enterobacter*, *Erwinia*, *Escherichia*, *Klebsiella*, *Salmonella* or *Serratia* for example the genera and species

15 *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*

20 subsp. *odorifera*, *Erwinia carotovora* subsp. *wasabiae*, *Erwinia chrysanthemi*, *Erwinia cypripedii*, *Erwinia dissolvens*, *Erwinia herbicola*, *Erwinia mallotivora*, *Erwinia milletiae*, *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*,

25 *Escherichia anindolica*, *Escherichia aurescens*, *Escherichia blattae*, *Escherichia coli*, *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.

35 *salamae*, *Salmonella enteritidis*, *Salmonella gallinarum*, *Salmonella heidelberg*, *Salmonella 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 plymouthisensis*, *Serratia plymuthica*, *Serratia proteamaculans*, *Serra-*

40

tia 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 meteor*, *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 terangae*, *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*.

20

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

25 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.

30 The term "non-human transgenic organism", preferably, relates to a plant, a plant seed, an 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, 35 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 40 mosses. Advantageous plants are selected from the group of the plant families

Adelotheciaceae, Anacardiaceae, Asteraceae, Apiaceae, Betulaceae, Boraginaceae, Brassicaceae, Bromeliaceae, Caricaceae, Cannabaceae, Convolvulaceae, Chenopodiaceae, Crypthecodiniaceae, Cucurbitaceae, Ditrachaceae, Elaeagnaceae, Ericaceae, Euphorbiaceae, Fabaceae, Geraniaceae, Gramineae, Juglandaceae, Lauraceae, Leguminosae, Linaceae, Prasinophyceae or vegetable plants or ornamentals such as Tagetes. Examples which may be mentioned are the following plants selected from the group consisting of: Adelotheciaceae 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 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. *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 *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 *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 the genera and species *Beta vulgaris*, *Beta vulgaris* var. *altissima*, *Beta vulgaris* var. *Vulgaris*, *Beta maritima*, *Beta vulgaris* var. *perennis*, *Beta vulgaris* var. *conditiva* or *Beta vulgaris* var. *esculenta* [sugarbeet], Crypthecodiniaceae, such as the genus *Cryptecodinium*, 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*,
5 *Distichium*, *Eccremidium*, *Lophidion*, *Philibertiella*, *Pleuridium*, *Saelania*, *Trichodon*, *Skottsbergia*, for example the genera and species *Ceratodon antarcticus*, *Ceratodon columbiae*, *Ceratodon heterophyllum*, *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 crispatis-*
10 *simum*, *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*,
15 *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*,
20 *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
25 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*,
30 *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*,
35 *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
40 genera *Aphanorrhagma*, *Entosthodon*, *Funaria*, *Physcomitrella*, *Physcomitrium*, for example the genera and species *Aphanorrhagma serratum*, *Entosthodon attenuatus*,

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*,
 5 *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 eurytostomum*, *Physcomitrium flexifolium*, *Physcomitrium hookeri*, *Physcomitrium hookeri* var. *serratum*, *Physcomitrium immersum*, *Physcomitrium kellermanii*, *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 *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
 10 genera and species *Linum usitatissimum*, *Linum humile*, *Linum austriacum*, *Linum bienne*, *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],
 15 Malvaceae, such as the genus *Gossypium*, for example the genera and species *Gossypium hirsutum*, *Gossypium arboreum*, *Gossypium barbadense*, *Gossypium herbaceum* 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
 20 and species *Musa nana*, *Musa acuminata*, *Musa paradisiaca*, *Musa* spp. [banana],
 25
 30
 35
 40

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 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 elongata* [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* [barley], *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 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 hybernum*, *Triticum macha*, *Triticum sativum* or *Triticum vulgare* [wheat], Porphyridiaceae, such as the genera *Chrootheca*, *Flintiella*, *Petrovanella*, *Porphyridium*, *Rhodela*, *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 example the genera and species *Nephroselmis olivacea*, *Prasinococcus capsulatus*, *Scherffelia dubia*, *Tetraselmis chui*, *Tetraselmis suecica*, *Mantoniella squamata*, *Ostreococcus tauri*, Rubiaceae such as the genus *Coffea*, for example the genera and species *Coffea* spp., *Coffea arabica*, *Coffea canephora* or *Coffea liberica* [coffee], Scrophulariaceae such as the genus *Verbascum*, for example the genera and species *Verbascum blattaria*, *Verbascum chaixii*, *Verbascum densiflorum*, *Verbascum lagurus*, *Verbascum longifolium*, *Verbascum lychnitis*, *Verbascum nigrum*, *Verbascum olympicum*, *Verbascum phlomoides*, *Verbascum phoenicum*, *Verbascum pulverulentum* or *Verbascum thapsus* [mullein], Solanaceae such as the genera *Capsicum*, *Nicotiana*, *Solanum*, *Lycopersicon*, for example the genera and species *Capsicum annuum*, *Capsicum an-*

num 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-*
5 *gena* [eggplant], *Lycopersicon esculentum*, *Lycopersicon lycopersicum*, *Lycopersicon*
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
10 accordance with the present invention are oil fruit crops which comprise large amounts
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
15 crops such as maize, wheat, rye, oats, triticale, rice, barley, cotton, cassava, pepper,
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,
20 hemp, castor-oil plant, olive, Calendula, Punica, evening primrose, pumpkin/squash,
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
25 safflower, sunflower, poppy, evening primrose, walnut, linseed, or hemp.

Preferred mosses are Physcomitrella or Ceratodon. Preferred algae are Isochrysis,
Mantoniella, Ostreococcus or Cryptocodinium, and algae/diatoms such as Phaeodac-
tylum or Thraustochytrium. More preferably, said algae or mosses are selected from
30 the group consisting of: Shewanella, Physcomitrella, Thraustochytrium, Fusarium, Phy-
tophthora, Ceratodon, Isochrysis, Aleurita, Muscarioides, Mortierella, Phaeodactylum,
Cryptocodinium, specifically from the genera and species Thallasiosira pseudonona,
Euglena gracilis, Physcomitrella patens, Phytophthora infestans, Fusarium gramineum,
Cryptocodinium cohnii, Ceratodon purpureus, Isochrysis galbana, Aleurita farinosa,
35 Thraustochytrium sp., Muscarioides viallii, Mortierella alpina, Phaeodactylum tricor-
num or Caenorhabditis elegans or especially advantageously Phytophthora infestans,
Thallasiosira pseudonona and Cryptocodinium cohnii.

Transgenic plants may be obtained by transformation techniques as published, and
40 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: *Crypthecodinium 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

- (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.

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 nucleic acid of interest will be expressed, preferably, seed specific in the said non-human transgenic organism.

In the following tables 1 to 6, the cis-regulatory elements found in the expression control sequences of the present invention are shown.

Table 1: cis-regulatory elements of SEQ ID NO: 7

Seq. name	Fam-ily/matrix	Further Information	Opt. thres h.	Start pos.	End pos.	Str an d	Core sim.	Mat-rix sim.	Sequence
SEQ 7	P\$TBPF/TA	Plant TATA box	0.90	4	18	+	1.00	0.93	gtcaTATAatatga
SEQ 7	P\$TBPF/TA	Plant TATA box	0.90	5	19	-	1.00	0.93	gtcaTATAatatga
SEQ 7	P\$PSRE/G	GAAA motif involved in pollen specific transcriptional activation	0.83	20	36	-	1.00	0.86	caaaaGAAAc-tatggaa
SEQ 7	P\$GTBX/S	SBF-1	0.87	33	49	+	1.00	0.87	tttggagTTA-Aacgcat
SEQ 7	P\$NACF/T	Wheat NACdomain DNA binding factor	0.68	60	82	+	1.00	0.68	tatgatttagc-TACGtgacagaa
SEQ 7	P\$GBOX/O	Oryza sativa bZIP protein 8	0.84	64	84	+	1.00	0.89	atttagctACGT-gacagaaaa
SEQ 7	P\$ABRE/A	ABA response elements	0.82	65	81	+	1.00	0.83	tttagctACGT-gacaga
SEQ 7	P\$OPAQ/RI	Rice transcription activator-1 (RITA), basic leucin zipper protein, highly expressed during seed development	0.95	65	81	-	1.00	0.98	tctgtcACG-Tagctaaa
SEQ 7	P\$OPAQ/RI	Rice transcription activator-1 (RITA), basic leucin zipper protein, highly expressed during seed development	0.95	66	82	+	1.00	0.98	tttagctACGTga-cagaa
SEQ 7	P\$AREF/S	Silencing element binding factor - transcriptional repressor	0.96	70	82	-	1.00	0.96	ttcTGTcagctag
SEQ 7	EBF.01	Homeodomain protein of the Knotted class1	1.00	73	85	+	1.00	1.00	cgTGACagaa-aat
SEQ 7	P\$TALE/HV	Avian C-type LTR PolyA signal	0.71	88	108	+	0.75	0.73	cagatCA-AAgtgtcgttttt
SEQ 7	O\$RPOA/A	Avian C-type LTR PolyA signal	0.71	91	111	-	0.75	0.72	tataaAAAACga-cactttgat

SEQ _7	P\$NACF/T	Wheat NACdomain DNA binding factor	0.81	0.71	ggctataaaa-
SEQ _7	ANAC69.01		2	7	AACGacacttg
SEQ _7	P\$TBPF/TA	Plant TATA box	1.00	0.96	ggctATATAaaa-
SEQ _7	TA.02		0	8	aacg
SEQ _7	P\$AHBP/W	Homeodomain protein WUSCHEL	1.00	1.00	
SEQ _7	US.01		0	0	ttaatTAATgt
SEQ _7	P\$DOFF/D		1.00	0.98	cattaattAAAG-
SEQ _7	OF1.01	Dof1 / MNB1a - single zinc finger transcription factor	0	0	gataa
SEQ _7	P\$MYBS/M		1.00	0.98	tactttATCCtt-
SEQ _7	YBST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	0	3	taatt
SEQ _7	P\$IBOX/GA		1.00	0.96	taagGATA-
SEQ _7	TA.01	Class I GATA factors	0	7	aagtaaga
SEQ _7	P\$NCS1/N		0.80	0.88	
SEQ _7	CS1.01	Nodulin consensus sequence 1	4	4	tAAAGgataaa
SEQ _7	P\$MYCL/IC		0.95	0.96	aggaaACAAat-
SEQ _7	E.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	4	6	gatttcca
SEQ _7	P\$PSRE/G	GAAA motif involved in pollen specific transcriptional activation	1.00	0.84	ttgtGAAAt-
SEQ _7	AAA.01		0	2	catttgt
SEQ _7	P\$AHBP/A	Arabidopsis thaliana homeo box protein 1	0.78	0.90	
SEQ _7	THB1.01		9	0	caaATGAttc
SEQ _7	P\$AHBP/A	HDZip class I protein ATHB5	0.93	0.93	
SEQ _7	THB5.01		6	6	gaaATCAttg
SEQ _7	P\$NCS1/N	Nodulin consensus sequence 1	0.87	0.90	
SEQ _7	CS1.01	CAACTC regulatory elements, GA-inducible	8	4	cAAATgatttc
SEQ _7	P\$MYBL/C		1.00	0.83	attcgaaAGTT-
SEQ _7	ARE.01	Dof1 / MNB1a - single zinc finger transcription factor	0	6	gattca
SEQ _7	P\$DOFF/D		1.00	0.98	accaaattAAAG-
SEQ _7	OF1.01		0	7	tattc
SEQ _7	P\$WBXF/E	Elicitor response element	1.00	0.91	tgagctTGAC-
SEQ _7	RE.01		0	8	caaatta
SEQ _7	O\$RVUP/L	Upstream element of C-type Long Terminal Repeats	1.00	0.78	actcacat-
SEQ _7	TRUP.01	Legumin box, highly conserved sequence element about 100 bp up-	0	3	gagTTTCgtgtga
SEQ _7	P\$LEGB/LE	stream of the TSS in legumin genes	0.75	0.60	tacatatCCAAat-
SEQ _7	GB.01		0	1	cagagtagaaggc

SEQ _7	P\$MYBS/O	Rice MYB proteins with single DNA binding domains, binding to the amy-	0.82	274	290	-	1.00	0.85	gtacaTATCca-
SEQ _7	SMYBS.01	lase element (TATCCA)					0	0	aatcag
SEQ _7	P\$GTBX/S		0.87	284	300	+	1.00	0.89	tatgtacTTAA-
SEQ _7	BF1.01	SBF-1					0	7	aacact
SEQ _7	P\$EINL/TEI	TEIL (tobacco EIN3-like)	0.92	285	293	+	1.00	0.93	aTGTAActta
	L.01						0	5	acttaaaa-
SEQ _7	O\$RVUP/L	Upstream element of C-type Long Terminal Repeats	0.76	289	309	+	1.00	0.78	cacTTTCtga-
SEQ _7	TRUP.01	Arabidopsis Telo-box interacting protein related to the conserved animal					0	4	gaa
SEQ _7	P\$TELO/AT	protein Pur-alpha	0.85	308	322	+	1.00	0.89	aaacACCC-
SEQ _7	PURA.01						0	6	taacgct
SEQ _7	P\$MIIG/P_	Maize activator P of flavonoid biosynthetic genes	0.93	320	334	+	0.96	0.97	gcttGGTTggtgg
SEQ _7	ACT.01						6	3	at
SEQ _7	P\$MYBL/M	Myb-like protein of Petunia hybrida	0.80	339	355	-	1.00	0.80	ctaaaattGTTAt-
SEQ _7	YBPH3.01						0	7	gatg
SEQ _7	O\$RPOA/D	PolyA signal of D-type LTRs	0.78	348	368	-	0.75	0.83	aCCAAtaaaa-
SEQ _7	TYPEPA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the ex-					0	4	aatctaaaaatt
SEQ _7	P\$LREM/A	pression of photosynthesis-related genes	0.85	349	359	-	1.00	0.99	
SEQ _7	TCTA.01						0	0	aaATCTaaaat
SEQ _7	P\$CCAF/C	Circadian clock associated 1	0.85	352	366	-	1.00	0.97	caataaaa-
SEQ _7	CA1.01						0	1	AATCtaa
SEQ _7	P\$CAAT/C	CCAAT-box in plant promoters	0.97	361	369	-	1.00	0.98	
SEQ _7	AAT.01						0	1	caCCAAtaa
SEQ _7	P\$OPAQ/O	Opaque-2 regulatory protein	0.87	363	379	-	0.79	0.87	aaccacataT-
SEQ _7	2.01						4	1	CACcaat
SEQ _7	O\$RPAD/P	Mammalian C-type LTR Poly A downstream element	0.87	371	383	+	1.00	0.87	
SEQ _7	ADS.01						0	4	tatGTGGtttgt
SEQ _7	P\$MIIG/PA	Putative cis-acting element in various PAL and 4CL gene promoters	0.81	380	394	+	0.93	0.85	ttGTGGttgga-
SEQ _7	LBOXP.01						6	9	gaag
SEQ _7	O\$RPOA/L	Lentiviral Poly A signal	0.94	398	418	-	1.00	0.98	tgaAATAaagt-
SEQ _7	POLYA.01						0	2	tattagaact
SEQ _7	P\$HMGF/H	High mobility group I/Y-like proteins	0.89	408	422	+	1.00	0.89	
SEQ _7	MG_IY.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0.87	415	425	+	0	5	acttTATTtcatcat
SEQ _7	P\$AHBP/H						1.00	0.94	ttcatcATTAt

7	AHB4.01						0	0		tactgGTACaa-
SEQ	P\$SBPD/S						1.00	0.90		tacagt
7	BP.01						0	2		tatactGGTA-
SEQ	P\$GTBX/G						0.75	0.83		caataca
7	T3A.01						0	9		gaaTGCCta-
SEQ	P\$MADS/A						1.00	0.89		tactgtacaat
7	GL1.01						0	0		ttgTACCagta-
SEQ	P\$MADS/A						0.99	0.86		taggcattct
7	GL1.01						5	2		tcctcGATAata-
SEQ	P\$IBOX/GA						1.00	0.93		caaat
7	TA.01						0	7		aagttTTGActat-
SEQ	P\$WBXFW						1.00	0.97		tata
7	RKY.01						0	1		tatatGATAat-
SEQ	P\$IBOX/GA						1.00	0.94		tagcaa
7	TA.01						0	6		
SEQ	P\$AHBP/H						1.00	0.91		atgataATTAg
7	AHB4.01						0	0		
SEQ	P\$NCS1/N						0.87	0.86		cAAATgatgtg
7	CS1.01						8	8		taagtaTAAAag-
SEQ	P\$L1BX/AT						1.00	0.82		tattg
7	ML1.01						0	3		
SEQ	P\$SPF1/SP						1.00	0.92		aaTACTtttatac
7	8BF.01						0	3		accTAAAtaat-
SEQ	P\$STKM/S						1.00	0.85		caaa
7	TK.01						0	2		ctcatttTTAAta-
SEQ	P\$GTBX/S						1.00	0.87		taga
7	BF1.01						0	2		
SEQ	P\$SEF4/SE						1.00	0.98		caTTTTtaata
7	F4.01						0	3		gtttaGAAAttcta-
SEQ	P\$PSRE/G						1.00	0.87		tat
7	AAA.01						0	2		aaaaaacgGTT-
SEQ	P\$MYBL/M						0.75	0.84		Tagaaa
7	YBPH3.01						0	5		tcta-
SEQ	P\$MYBL/M						0.75	0.80		

SQUA promoter binding proteins

Trihelix DNA-binding factor GT-3a

AGL1, Arabidopsis MADS-domain protein AGAMOUS-like 1

AGL1, Arabidopsis MADS-domain protein AGAMOUS-like 1

Class I GATA factors

WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box

Class I GATA factors

Sunflower homeodomain leucine-zipper protein Hahb-4

Nodulin consensus sequence 1

L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1) DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression

SBF-1

Soybean embryo factor 4

GAAA motif involved in pollen specific transcriptional activation

Myb-like protein of Petunia hybrida

Myb-like protein of Petunia hybrida

7	YBPH3.01	0	3	aaccGTTTtttt
SEQ	P\$MSAE/M	1.00	0.86	aaaaAACGgtt-
7	SA.01	0	7	tag
SEQ	P\$DOFF/P	0.76	0.80	tgaaatgaAAA-
7	BOX.01	1	2	Aaaaa
SEQ	P\$GAPB/G	1.00	0.96	tgaaATGAaaa-
7	AP.01	0	0	aaaa
SEQ	P\$OPAQ/O	1.00	0.92	ttttcattTCAT-
7	2.01	0	4	catc
SEQ	P\$DOFF/P	1.00	0.75	tggttagaAAA-
7	BOX.01	0	8	Gatgat
SEQ	P\$NCS1/N	1.00	0.96	
7	CS1.01	0	3	gAAAAgatgat
SEQ	P\$DOFF/P	1.00	0.81	gagactgtAAA-
7	BOX.01	0	9	Gatgaa
SEQ	P\$AHBP/H	1.00	0.89	
7	AHB4.01	0	2	tatatgATTAg
SEQ	P\$HMGF/H	1.00	0.90	agttTATTt-
7	MG_IY.01	0	5	cattcg
SEQ	P\$TBPF/TA	1.00	0.91	
7	TA.02	0	3	tttcTATAtattaaa
SEQ	P\$AHBP/H	1.00	0.93	
7	AHB4.01	0	6	togattATTAt
SEQ	P\$MSAE/M	0.75	0.80	cttcaAATGgt-
7	SA.01	0	3	gata
SEQ	P\$MYCL/IC	1.00	0.95	ttgtgACATttgg-
7	E.01	0	3	cattac
SEQ	P\$OCSE/O	0.97	0.81	tgacattggcat-
7	CSTF.01	4	1	TACGggaa
SEQ	P\$MADS/M	1.00	0.76	aggtcCCATg-
7	ADS.01	0	5	tatctaaact
SEQ	P\$EINL/TEI	1.00	0.98	
7	L.01	0	0	aTGAtctc
SEQ	P\$MADS/A	0.90	0.82	gttTGCCa-
M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)				
Prolamin box, conserved in cereal seed storage protein gene promoters				
Cis-element in the GAPDH promoters conferring light inducibility				
Opaque-2 regulatory protein				
Prolamin box, conserved in cereal seed storage protein gene promoters				
Nodulin consensus sequence 1				
Prolamin box, conserved in cereal seed storage protein gene promoters				
Sunflower homeodomain leucine-zipper protein Hahb-4				
High mobility group I/Y-like proteins				
Plant TATA box				
Sunflower homeodomain leucine-zipper protein Hahb-4				
M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)				
ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)				
bZIP transcription factor binding to OCS-elements				
Binding sites for AP1, AP3-PI and AG dimers				
TEIL (tobacco EIN3-like)				
Agamous, required for normal flower development, similarity to SRF (hu-				

7	G.01	man) and MCM (yeast) proteins	2	5	catgtgtaagaag
SEQ	P\$ABRE/A		0.75	0.79	cttacA-
7	BF1.01	ABA (abscisic acid) inducible transcriptional activator	0	0	CATgtggcaaa
SEQ	P\$OPAQ/O	Recognition site for BZIP transcription factors that belong to the group of	1.00	0.84	ttggcACATgtg-
7	2_GCN4.01	Opaque-2 like proteins	0	1	taag
SEQ	P\$DOFF/P		0.76	0.77	gcactgtAAAAAt-
7	BOX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	1	6	caaa
SEQ	O\$RPOA/A		1.00	0.81	gtaaaTAAAc-
7	POLYA.01	Avian C-type LTR PolyA signal	0	6	taacttttga
SEQ	P\$MYBL/M		1.00	0.91	aaaagtTAGTt-
7	YBPH3.02	Myb-like protein of Petunia hybrida	0	9	tattta
SEQ	P\$MADS/M		1.00	0.75	acaaaCCATg-
7	ADS.01	Binding sites for AP1, AP3-PI and AG dimers	0	2	taataaaacta
SEQ	P\$TBPF/TA		0.78	0.88	accaTGTAaa-
7	TA.01	Plant TATA box	2	6	taaac
SEQ	P\$OCSE/O		0.97	0.75	taacacaaa-
7	CSTF.01	bZIP transcription factor binding to OCS-elements	4	3	caacTACGtagg
SEQ	P\$LFYB/LF		0.88	0.96	
7	Y.01	Plant specific floral meristem identity gene LEAFY (LFY)	5	2	gGCCAttggttg
SEQ	P\$MSAE/M		0.75	0.80	aaaccAATGgc-
7	SA.01	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0	0	cata
SEQ	P\$CAAT/C		1.00	0.98	
7	AAT.01	CCAAT-box in plant promoters	0	6	aaCCAAtgg
SEQ	P\$HEAT/H		1.00	0.85	agatatattgaA-
7	SE.01	Heat shock element	0	6	GAAt
SEQ	P\$MYCL/IC		1.00	0.97	ttcatACATatgt-
7	E.01	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0	0	taaca
SEQ	P\$MYBL/M		1.00	0.89	cgtggtTAGT-
7	YBPH3.02	Myb-like protein of Petunia hybrida	0	6	taataac
SEQ	P\$OCSE/O		1.00	0.70	gttattaactaac-
7	CSL.01	OCS-like elements	0	6	cACGTaaa
SEQ	P\$MIIG/PA		0.93	0.85	acGTGGttagt-
7	LBOXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	6	5	taat
SEQ	P\$GBOX/G	bZIP protein G-Box binding factor 1	1.00	0.97	acatttt-

SEQ _7	P\$MYBS/T	MYB protein from wheat	115	117	1.00	0.91	atgcATATaca-
SEQ _7	AMYB80.01	MYB protein from wheat	4	0	0	7	cacgta
SEQ _7	P\$MYBS/T	MYB protein from wheat	115	117	1.00	0.89	gtgtATATgcatt-
SEQ _7	AMYB80.01	MYB protein from wheat	9	5	0	3	tatg
SEQ _7	P\$L1BX/AT	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	116	117	1.00	0.93	ctctcaTAAAtg-
SEQ _7	ML1.01		3	9	0	8	catat
SEQ _7	P\$DOFF/D		117	119	1.00	0.98	tgagagctAAA-
SEQ _7	OF1.01	Dof1 / MNB1a - single zinc finger transcription factor	4	0	0	2	Gagtat
SEQ _7	P\$MYBS/O	Rice MYB proteins with single DNA binding domains, binding to the amy-	118	119	1.00	0.90	aagagTATC-
SEQ _7	SMYBS.01	lase element (TATCCA)	3	9	0	5	cattcatt
SEQ _7	P\$MYBS/T	MYB protein from wheat	119	121	1.00	0.89	aagtATATgca-
SEQ _7	AMYB80.01		4	0	0	6	aatgaa
SEQ _7	P\$L1BX/AT	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	119	121	0.75	0.82	tgaaagTATAtg-
SEQ _7	ML1.01		7	3	0	7	caaat
SEQ _7	P\$MYBS/T	MYB protein from wheat	119	121	1.00	0.90	tigcATATacttt-
SEQ _7	AMYB80.01		9	5	0	9	cata
SEQ _7	P\$OCSE/O	OCS-like elements	121	123	1.00	0.69	agagttatatata-
SEQ _7	CSL.01	Plant TATA box	2	2	0	2	aACGTatg
SEQ _7	P\$TBPF/TA	Plant TATA box	121	122	1.00	0.88	tataTATAaacg-
SEQ _7	TA.01	Plant TATA box	3	7	0	9	tat
SEQ _7	P\$TBPF/TA	Plant TATA box	121	122	1.00	0.91	gttaTATAta-
SEQ _7	TA.02	Plant TATA box	5	9	0	7	aacgt
SEQ _7	P\$TBPF/TA	Plant TATA box	121	123	1.00	0.93	cgttTATAta-
SEQ _7	TA.02	Plant TATA box	6	0	0	7	taact
SEQ _7	P\$TBPF/TA	Plant TATA box	121	123	1.00	0.93	gagtTATAata-
SEQ _7	TA.02	Plant TATA box	7	1	0	1	aac
SEQ _7	O\$RPAD/P	Mammalian C-type LTR Poly A downstream element	124	125	1.00	0.87	
SEQ _7	ADS.01	Prolamin box, conserved in cereal seed storage protein gene promoters	2	4	0	7	attGTGGttcat
SEQ _7	P\$DOFF/P		125	127	1.00	0.75	tgattagA-
SEQ _7	BOX.01	CCAAT-box in plant promoters	5	1	0	1	AAGtgtat
SEQ _7	P\$CAAT/C		126	127	1.00	0.99	atCCAAataa
SEQ _7	AAT.01	Arabidopsis thaliana homeo box protein 1	7	5	0	2	
SEQ _7	P\$AHBP/A		126	127	1.00	0.99	agaATTAttgg
SEQ _7	THB1.01		9	9	0	0	

[illegible]

SE	Q_8	P\$DREB/CRT _DRE.01	C-repeat/dehydration response element	0. 10 11 1.0 0.9 89 4 8 + 00 23	catttCCGA- cagggg
SE	Q_8	P\$MADS/AGL 1.01	AGL1, Arabidopsis MADS-domain protein AGAMOUS-like 1	0. 10 12 0.9 0.8 84 4 4 + 75 56	catTTCCga- caggggcagaac
SE	Q_8	P\$LEGB/R.Y.0 1	RY and Sph motifs conserved in seed-specific promoters	0. 15 18 1.0 0.8 87 4 0 - 00 73	gagacattCATGcac- cagggcggctgt
SE	Q_8	P\$NCS3/NCS 3.01	Nodulin consensus sequence 3	0. 20 21 1.0 0.9 89 3 3 + 00 60	gtCACCCctccc
SE	Q_8	P\$MADS/AGL 2.01	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	0. 23 25 0.7 0.8 82 9 9 + 63 21	caaatCCG- Taactcgtaaata
SE	Q_8	P\$OCSE/OCS TF.01	bZIP transcription factor binding to OCS-elements	0. 24 26 0.9 0.7 73 8 8 - 74 70	tggcggtagtatt- TACGagtt
SE	Q_8	P\$DREB/HVD RF1.01	H. vulgare dehydration-response factor 1	0. 25 27 1.0 0.8 89 8 2 + 00 97	tactACCGc- caccgg
SE	Q_8	P\$CE1F/ABI4. 01	ABA insensitive protein 4 (ABI4)	0. 26 27 1.0 0.8 87 3 5 + 00 93	ccgcCACCCggcca
SE	Q_8	O\$RPOA/AP OLYA.01	Avian C-type LTR PolyA signal	0. 26 28 0.7 0.7 71 4 4 - 50 54	agaaaaAA- AAtggccggtggcg
SE	Q_8	P\$MADS/SQ UA.01	MADS-box protein SQUAMOSA	0. 26 28 1.0 0.9 90 8 8 + 00 02	accggc- cATTTtttcttaga
SE	Q_8	P\$LREM/ATC TA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0. 28 29 1.0 0.9 85 1 1 - 00 36	aaATCTaagaa
SE	Q_8	P\$CCAF/CCA	Circadian clock associated 1	0. 28 29 - 1.0 0.9	caaaaaaaAATC-

Q_8	1.01			85	4	8	00	81	taa
SE									
Q_8	P\$NCS1/NCS	Nodulin consensus sequence 1		0.	29	30	1.0	0.9	
SE	1.01			85	8	-	00	49	aAAAAGatttc
Q_8	P\$CCAF/CCA	Circadian clock associated 1		0.	31	32	1.0	0.8	gaaaaaaaaAAT-
SE	1.01			85	2	6	-	00	Caga
Q_8	P\$CCAF/CCA	Circadian clock associated 1		0.	32	33	1.0	0.8	gaaaaaaaa-
SE	1.01			85	5	9	-	00	AATCoga
Q_8	P\$MSAE/MSA	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)		0.	33	35	1.0	0.8	acacaAACGgga-
SE	.01			80	7	1	-	00	gaa
Q_8	P\$CARM/CA	CA-rich element		0.	34	36	1.0	0.7	tcctcacAACAc-
SE	RICH.01			78	3	1	-	00	caaacg
Q_8	P\$MYBL/CAR	CAACTC regulatory elements, GA-inducible		0.	35	37	1.0	0.8	ggaaga-
SE	E.01			83	9	5	+	00	gAGTTgtgga
Q_8	P\$URNA/USE	Upstream sequence elements in the promoters of U-snrRNA genes of higher plants		0.	36	37	1.0	0.7	catttcCCA-
SE	.01			75	3	9	-	00	Caactctc
Q_8	P\$OCSE/OCS	OCS-like elements		0.	36	38	0.8	0.6	agttgtgggaaat-
SE	L.01			69	6	6	+	07	gACATata
Q_8	P\$OPAQ/O2_	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins		0.	37	39	1.0	0.8	gaaatgACA Tata-
SE	GCN4.01			81	4	0	+	00	tata
Q_8	P\$TBPF/TAT	Plant TATA box		0.	37	39	1.0	0.9	gacaTATAtata-
SE	A.02			90	9	3	+	00	gag
Q_8	P\$TBPF/TAT	Plant TATA box		0.	38	39	1.0	0.9	tctcTATAtatgt
SE	A.02			90	0	4	-	00	60

SE	Q_	P\$MADS/AG.	Agamous, required for normal flower development, similarity to SRF (human) and	0.	49	51	1.0	0.8	actTACC-
8	01		MCM (yeast) proteins	80	0	0	-	00	34 taatttcgtgtaa
SE	Q_	P\$GTBX/GT1.		0.	49	51	0.9	0.9	aattagGTAAgtta-
8	01		GT1-Box binding factors with a trihelix DNA-binding domain	85	9	5	+	68	07 aag
SE	Q_	P\$DOFF/DOF		0.	50	52	1.0	0.9	ggtaagttAAAGca-
8	2.01		Dof2 - single zinc finger transcription factor	98	4	0	+	00	93 cat
SE	Q_	P\$L1BX/ATM		0.	51	52	1.0	0.8	
8	L1.01		L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	82	3	9	-	00	38 gtgtatTAAAtgtgctt
SE	Q_	P\$GBOX/GBF		0.	51	53	1.0	0.9	aggtgaaaACGTg-
8	1.01		bZIP protein G-Box binding factor 1	94	9	9	-	00	56 tattaaat
SE	Q_	P\$OCSE/OCSE		0.	52	54	1.0	0.7	atatttagtgaa-
8	L.01		OCS-like elements	69	5	5	-	00	77 aACGTgta
SE	Q_	P\$GTBX/GT3		0.	53	55	1.0	0.8	gttaccGTTAtatt-
8	A.01		Trihelix DNA-binding factor GT-3a	83	8	4	-	00	39 tag
SE	Q_	P\$MYBL/MYB		0.	54	55	1.0	0.8	
8	PH3.01		Myb-like protein of Petunia hybrida	80	0	6	-	00	37 atgttaccGTTAtattt
SE	Q_	P\$MSAE/MSA		0.	54	55	1.0	0.9	aatatAACGgtaa-
8	.01		M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	80	1	5	+	00	42 ca
SE	Q_	P\$GTBX/GT3		0.	54	56	1.0	0.8	aagcatGTTAccgt-
8	A.01		Trihelix DNA-binding factor GT-3a	83	4	0	-	00	76 tat
SE	Q_	P\$NACF/TAN		0.	57	59	1.0	0.6	gatgtaatgatTAC-
8	AC69.01		Wheat NACdomain DNA binding factor	68	3	5	+	00	92 Gacgtattt
SE	Q_	P\$AHBP/ATH		0.	57	58	+	1.0	1.0 gtaATGAttac
			HD-ZIP class III protein ATHB9						

Q_8	B9.01		77	6	6	00	00	
SE								
Q_8	P\$AHBP/HAH	Sunflower homeodomain leucine-zipper protein Hahb-4	0.	57	58	1.0	0.9	
SE	B4.01		87	6	6	00	81	gtaatcATTAc
Q_8	P\$GBOX/ABZ	Anaerobic basic leucine zipper	0.	59	61	1.0	0.9	acgaatgtACGTg-
SE	1.01		91	5	5	00	10	gaatgaga
Q_8	P\$OPAQ/O2.	Opaque-2 regulatory protein	0.	59	61	1.0	0.9	cattCCACgta-
SE	02		87	8	4	00	01	cattcg
Q_8	P\$EINL/TEIL.	TEIL (tobacco EIN3-like)	0.	60	61	1.0	0.9	
SE	01		92	3	1	00	41	aTGTAcgtg
Q_8	P\$NACF/TAN	Wheat NACdomain DNA binding factor	0.	62	64	0.8	0.6	aaggagttttacGAC-
SE	AC69.01		68	7	9	00	81	Gaaaccagt
Q_8	P\$DOFF/PBO	Prolamin box, conserved in cereal seed storage protein gene promoters	0.	65	67	0.7	0.7	cgccatgtAAATt-
SE	X.01		75	4	0	00	68	tacg
Q_8	P\$GBOX/OSB	Oryza sativa bZIP protein 8	0.	65	67	0.7	0.8	gtaaatttA-
SE	Z8.01		84	5	5	00	43	CATggcggtttac
Q_8	P\$LREM/ATC	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of	0.	68	69	1.0	0.9	
SE	TA.01	photosynthesis-related genes	85	1	1	00	00	gaATCTaactt
Q_8	P\$DOFF/DOF	Dof3 - single zinc finger transcription factor	0.	70	71	1.0	0.9	ttcgttgtAAAGcc-
SE	3.01		99	0	6	00	95	cat
Q_8	P\$ERSE/ERS	ERSE I (ER stress-response element I)-like motif	0.	71	73	0.7	0.7	cccatgtaaatttac-
SE	E.I.01		79	2	0	00	91	GACG
Q_8	P\$MYBL/MYB	Myb-like protein of Petunia hybrida	0.	74	76	0.7	0.7	aaatgtTCGTgt-
SE	PH3.02		76	6	2	00	61	tatg

[illegible]

SE	Q_8	P\$MADS/AGL3.01	AGL3, MADS Box protein	0.82	84	0.9	0.8	ttatcCCAAatact- gatttag
SE	Q_8	P\$MYBS/MYB ST1.01	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	0.83	84	1.0	0.9	atcattATCCcaaa- tac
SE	Q_8	P\$IBOX/GAT A.01	Class I GATA factors	0.83	85	1.0	0.9	tttgGATAat- gatgct
SE	Q_8	P\$AHBP/HAH B4.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0.84	85	1.0	0.9	agcatcATTAT
SE	Q_8	P\$MADS/AG.01	Agamous, required for normal flower development, similarity to SRF (human) and MCM (yeast) proteins	0.84	86	1.0	0.8	ctaTACCtaataa- gagcatca
SE	Q_8	P\$MIIG/MYB C1.01	Maize C1 myb-domain protein	0.87	88	1.0	0.9	cgttgGTAGttgcat
SE	Q_8	P\$MYBL/CAR E.01	CAACTC regulatory elements, GA-inducible	0.87	88	1.0	0.8	cgttggtAGTTg- cattg
SE	Q_8	P\$NCS1/NCS 1.01	Nodulin consensus sequence 1	0.88	89	0.8	0.9	cAAATgatgat
SE	Q_8	P\$GBOX/HBP 1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	0.89	91	1.0	0.9	tgcatgcCACGtt- gaatcaaa
SE	Q_8	P\$GBOX/OSB Z8.01	Oryza sativa bZIP protein 8	0.89	91	1.0	0.9	tgattcaACGTgg- catgcag
SE	Q_8	P\$ABRE/ABR E.01	ABA response elements	0.89	91	1.0	0.9	tgattcaACGTgg- catg
SE	Q_8	P\$LEGB/RV.0	RY and Sph motifs conserved in seed-specific promoters	0.90	92	1.0	0.8	caacgtggCATG-

Q_8_1	SE	1	87	3	9	00	84	cagttcattggctaa
Q_8_2.01	SE	P\$PEFF/ZPT2	0.	90	92	1.0	0.7	aacgtggcatg-
Q_8_2.01	SE	Member of the EPF family of zinc finger transcription factors	75	4	6	+	00	51 CAGTtcattggc
Q_8_T.01	SE	P\$CAAT/CAA	0.	91	92	1.0	0.9	agCCAAAtga
Q_8_T.01	SE	CCAAT-box in plant promoters	97	9	7	-	00	78
Q_8_A.01	SE	P\$GTBX/GT3	0.	92	94	1.0	0.8	ctaattGTTAca-
Q_8_A.01	SE	Trihelix DNA-binding factor GT-3a	83	6	2	+	00	84 catct
Q_8_B4.01	SE	P\$AHBP/HAH	0.	95	96	1.0	0.9	cgtataATTAg
Q_8_B4.01	SE	Sunflower homeodomain leucine-zipper protein Hahb-4	87	3	3	-	00	13
Q_8_1.01	SE	P\$GBOX/GBF	0.	95	97	1.0	0.9	ctaatta-
Q_8_1.01	SE	bZIP protein G-Box binding factor 1	94	3	3	+	00	56 tACGTggtggtgag
Q_8_E.01	SE	P\$ABRE/ABR	0.	95	97	1.0	0.8	taatta-
Q_8_E.01	SE	ABA response elements	82	4	0	+	00	74 tACGTggtggt
Q_8_N1.02	SE	P\$SALT/ALFI	0.	95	97	1.0	0.9	tatacgtGGTGgtga
Q_8_N1.02	SE	Zinc-finger protein in alfalfa roots, regulates salt tolerance	95	8	2	+	00	55
Q_8_AC69.01	SE	P\$NACF/TAN	0.	96	98	1.0	0.7	gtggtggtgag-
Q_8_AC69.01	SE	Wheat NACdomain DNA binding factor	68	3	5	+	00	47 TACGtagttcaa
Q_8_MYBS.01	SE	P\$MYBS/OS	0.	10	10	1.0	0.8	aaaagTATCcac-
Q_8_MYBS.01	SE	Rice MYB proteins with single DNA binding domains, binding to the amylase element (TATCCA)	82	09	25	-	00	42 caaaa
Q_8_O\$RVUP/LTR	SE	UP.01	0.	10	10	0.7	0.8	tgcataat-
Q_8_O\$RVUP/LTR	SE	Upstream element of C-type Long Terminal Repeats	76	25	45	+	61	05 cagTTTTTgattgt
Q_8_P\$LREM/ATC	SE	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.	10	10	1.0	0.9	atATCTaaatc
Q_8_TA.01	SE	TA.01	85	49	59	+	00	90

8	SE	P\$LREM/ATC	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of	0.	10	10	1.0	0.8	
8	Q_	TA.01	photosynthesis-related genes	85	55	65	+	00	74 aaATCTacgtg
8	SE	P\$GBOX/TGA		0.	10	10	0.8	0.9	tctacgTGATg-
8	Q_	1.01	Arabidopsis leucine zipper protein TGA1	90	58	78	+	18	01 taacgaccga
8	SE	P\$GTBX/GT3		0.	10	10	0.7	0.8	acgaccGATAcaa-
8	Q_	A.01	Trihelix DNA-binding factor GT-3a	83	71	87	+	50	52 taaa
8	SE	O\$RPOA/LPO		0.	10	10	1.0	0.9	tacAATAaa-
8	Q_	LYA.01	Lentiviral Poly A signal	94	79	99	+	00	41 caattctgattg
8	SE	P\$STKM/STK.	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific	0.	10	10	1.0	0.9	
8	Q_	01	and sucrose-inducible gene expression	85	81	95	+	00	02 caaTAAAcattctg
8	SE	O\$RPOA/AP		0.	10	11	0.7	0.7	ttgaaTATAtatcca-
8	Q_	OLYA.01	Avian C-type LTR PolyA signal	71	97	17	+	50	42 tatgat
8	SE	P\$GTBX/S1F.		0.	11	11	1.0	0.8	tcatATGGatata-
8	Q_	01	S1F, site 1 binding factor of spinach rps1 promoter	79	00	16	-	00	11 tatt
8	SE	P\$MYBS/OS	Rice MYB proteins with single DNA binding domains, binding to the amylase element	0.	11	11	1.0	0.8	atataTATCcatat-
8	Q_	MYBS.01	(TATCCA)	82	01	17	+	00	82 gat
8	SE	P\$MADS/AGL		0.	11	11	1.0	0.7	tcaTACTatca-
8	Q_	15.01	AGL15, Arabidopsis MADS-domain protein AGAMOUS-like	79	04	24	-	00	93 tatgatata
8	SE	P\$ERSE/ERS		0.	11	11	1.0	0.8	caactatggtgttc-
8	Q_	E.I.01	ERSE I (ER stress-response element I)-like motif	79	38	56	+	00	42 CACG
8	SE	P\$NACF/TAN		0.	11	11	0.8	0.6	tatggtgttcCACG-
8	Q_	AC69.01	Wheat NACdomain DNA binding factor	68	42	64	+	95	87 taacgacc

SE	P\$GBOX/GBF 1.01	bZIP protein G-Box binding factor 1		0.	11	11	1.0	0.9	tgtctggtACGTgg-
Q_8				94	45	65	-	00	73 caacacc
SE	P\$GBOX/HBP 1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression		0.	11	11	1.0	0.9	gtgttgcCACG-
Q_8				88	46	66	+	00	52 taacgacct
SE	P\$OPAQ/RIT A1.01	Rice transcription activator-1 (RITA), basic leucine zipper protein, highly expressed during seed development		0.	11	11	1.0	0.9	gtcgttACGTgg-
Q_8				95	47	63	-	00	81 caaca
SE	P\$ABRE/ABF 1.03	ABA (abscisic acid) inducible transcriptional activator		0.	11	11	1.0	0.8	ggtcgttaCGTGg-
Q_8				82	48	64	-	00	97 caac
SE	P\$OPAQ/RIT A1.01	Rice transcription activator-1 (RITA), basic leucine zipper protein, highly ex-pressed during seed development		0.	11	11	1.0	0.9	gttgccACGTaac-
Q_8				95	48	64	+	00	85 gacc
SE	P\$MYBL/NTM YBAS1.01	Anther-specific myb gene from tobacco		0.	11	11	1.0	0.9	agatggctcGT-
Q_8				96	52	68	-	00	74 TAogtgg
SE	O\$RPOA/LPO LYA.01	Lentiviral Poly A signal		0.	11	11	1.0	0.9	gttAATAaa-
Q_8				94	56	76	-	00	89 gatggtcgttac
SE	P\$MADS/MA DS.01	Binding sites for AP1, AP3-PI and AG dimers		0.	11	11	1.0	0.7	aacgaCCATcttat-
Q_8				75	58	78	+	00	68 taacca
SE	P\$DOFF/DOF 1.01	Dof1 / MNB1a - single zinc finger transcription factor		0.	11	11	1.0	0.9	tggtaatAAA-
Q_8				98	62	78	-	00	84 Gatggt
SE	P\$GTBX/SBF 1.01	SBF-1		0.	11	11	1.0	0.8	gtcatggTTAAata-
Q_8				87	66	82	-	00	94 aaga
SE	P\$GBOX/HBP 1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)		0.	11	11	1.0	0.9	gttgcggcACGT-
Q_8				83	72	92	-	00	43 catggttaa
SE	P\$GBOX/TGA	Arabidopsis leucine zipper protein TGA1		0.	11	11	+ 1.0	0.9	taaccaT-

Q_8	1.01		90	73	93	00	71	GACgtgccgcaacc
SE								
Q_8	P\$ABRE/ABF	ABA (abscisic acid) inducible transcriptional activator	0.	11	11	1.0	0.8	aaccat-
SE	1.03		82	74	90	+	00	63 gaCGTGccgca
Q_8	P\$ERSE/ERS		0.	11	12	1.0	0.8	cctctgtgttgccgC
SE	E_1.01	ERSE I (ER stress-response element I)-like motif	79	82	00	-	00	04 ACG
Q_8	P\$IDDF/ID1.0		0.	11	12	1.0	0.9	
SE	1	Maize INDETERMINATE1 zinc finger protein	92	96	08	-	00	27 gttgTTGTcctct
Q_8	P\$MYBL/GAM		0.	11	12	0.8	0.9	
SE	YB.01	GA-regulated myb gene from barley	91	97	13	-	84	11 tatgtgttGTTGtcctc
Q_8	O\$L TUP/TAA		0.	12	12	1.0	0.7	acagcatcgcaga-
SE	CC.01	Lentiviral TATA upstream element	71	13	35	+	00	25 AACCGcatact
Q_8	P\$MADS/AGL		0.	12	12	1.0	0.8	gcaTACTaa-
SE	15.01	AGL15, Arabidopsis MADS-domain protein AGAMOUS-like 15	79	29	49	+	00	08 cactcgcaagt
Q_8	P\$CARM/CA		0.	12	12	1.0	0.8	aaagtgcAACAcc-
SE	RICH.01	CA-rich element	78	45	63	+	00	12 caaaac
Q_8	O\$MINI/MUS		0.	12	13	1.0	0.8	gcaacaCCACg-
SE	CLE_INI.02	Muscle Initiator Sequence	86	90	08	+	00	60 cagctata
Q_8	P\$OCSE/OC		0.	12	13	1.0	0.6	ccacgcagctata-
SE	L.01	OCS-like elements	69	96	16	+	00	90 cACGTatc
Q_8	P\$GBOX/GBF		0.	13	13	1.0	0.9	tagaagatACGTg-
SE	1.01	bZIP protein G-Box binding factor 1	94	01	21	-	00	67 tatagctg
Q_8	P\$OCSE/OC		0.	13	13	1.0	0.7	ttagcatagaaga-
SE	L.01	OCS-like elements	69	07	27	-	00	37 tACGTgta

8	SE	P\$GARP/ARR	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	0.	13	13	1.0	0.9	
Q_	8	10.01		97	09	17	-	00	85 AGATacgtg
SE	Q_	P\$GBOX/GBF	bZIP protein G-Box binding factor 1	0.	13	13	1.0	0.9	
8	8	1.01		94	20	40	-	00	76 gacatgacACGTgt-tagcata
SE	Q_	P\$GBOX/GBF	bZIP protein G-Box binding factor 1	0.	13	13	1.0	0.9	
8	8	1.01		94	21	41	+	00	77 atgctaacACGTgt-catgtct
SE	Q_	P\$MYCL/ICE.	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.	13	13	0.9	0.9	
8	8	01		95	21	39	-	54	66 acatgACACgtgt-tagcat
SE	Q_	P\$ABRE/ABR	ABA response elements	0.	13	13	1.0	0.9	
8	8	E.01		82	22	38	+	00	51 tgctaacACGTgt-catg
SE	Q_	P\$CE3S/CE3.	Coupling element 3 (CE3), non-ACGT ABRE	0.	13	13	0.7	0.8	
8	8	01		77	22	40	+	50	06 tgctaaCACGtgt-catgtc
SE	Q_	P\$MYCL/OSB	Rice bHLH protein	0.	13	13	1.0	0.9	
8	8	HLH66.01		85	22	40	+	00	01 tgctaaCACGtgt-catgtc
SE	Q_	P\$ABRE/ABR	ABA response elements	0.	13	13	1.0	0.8	
8	8	E.01		82	23	39	-	00	97 acatgacACGTgt-tagc
SE	Q_	P\$DPBF/DPB	bZIP factors DPBF-1 and 2 (Dc3 promoter binding factor-1 and 2)	0.	13	13	1.0	0.9	
8	8	F.01		89	26	36	+	00	20 aACACgtgtca
SE	Q_	P\$DREB/CRT	C-repeat/dehydration response element	0.	13	13	1.0	0.9	
8	8	_DRE.01		89	41	55	+	00	04 ttgaaCCGACcaaga
SE	Q_	P\$MYCL/ICE.	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	0.	13	13	1.0	0.9	
8	8	01		95	56	74	-	00	65 tcgatACATttg-tagtgtg

SE	Q_8	O\$LDPS/LDS POLYA.01	Lentiviral Poly A downstream element	0. 13 13 1.0 0.9 89 84 98 - 00 00	gtGTGTatggtcttc
SE	Q_8	P\$MADS/AGL 15.01	AGL15, Arabidopsis MADS-domain protein AGAMOUS-like 15	0. 13 14 0.8 0.8 79 98 18 - 50 23	tgtTGCTgtgtaaa- gaaagt
SE	Q_8	P\$DOFF/PBO X.01	Prolamin box, conserved in cereal seed storage protein gene promoters	0. 13 14 1.0 0.8 75 99 15 - 00 82	tgctgtgTAAAGa- aagt
SE	Q_8	P\$MADS/AGL 15.01	AGL15, Arabidopsis MADS-domain protein AGAMOUS-like 15	0. 13 14 0.9 0.8 79 99 19 + 25 82	actTTCTttaca- cagcaacat
SE	Q_8	P\$RAV5/RAV 1-5.01	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	0. 14 14 1.0 0.9 96 12 22 + 00 67	agcAACAtaca
SE	Q_8	P\$DOFF/DOF 2.01	Dof2 - single zinc finger transcription factor	0. 14 14 1.0 0.9 98 31 47 - 00 83	aattatatAAA Gctttt
SE	Q_8	P\$TBPF/TAT A.01	Plant TATA box	0. 14 14 1.0 0.8 88 32 46 - 00 92	attaTATAaagcttt
SE	Q_8	P\$TBPF/TAT A.02	Plant TATA box	0. 14 14 1.0 0.9 90 34 48 - 00 21	taatTATAtaaagct
SE	Q_8	P\$AHBP/HAH B4.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0. 14 14 1.0 0.9 87 39 49 + 00 02	tatataATTAt
SE	Q_8	P\$IBOX/GAT A.01	Class I GATA factors	0. 14 14 1.0 0.9 93 39 55 - 00 46	gaaatGATAatta- tata
SE	Q_8	P\$AHBP/HAH B4.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0. 14 14 1.0 0.9 87 42 52 - 00 10	atgataATTAt
SE	Q_8	P\$AHBP/ATH	HD-ZIP class III protein ATHB9	0. 14 14 - 1.0 0.7	gaaATGAtaat

SE
Q_ P\$NCS1/NCS

Nodulin consensus sequence 1

77 45 55 00 75

0. 14 14 0.8 0.9

85 45 55 - 78 15 gAATgataat

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

SE	P\$CGCG/ATS	Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)	0.	84	Q_9 R1.01
Q_9	R1.01		0.84	2	18 - 00 59
SE	P\$CE3S/CE3.	Coupling element 3 (CE3), non-ACGT ABRE	0.		cga-
Q_9	01		77	3	21 - 00 05
SE	P\$CGCG/ATS	Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)	0.		gcaCGCGtggtc-
Q_9	R1.01		84	9	25 + 00 56
SE	P\$MIG/P_AC		0.		gacgg
Q_9	T.01	Maize activator P of flavonoid biosynthetic genes	93	30	44 + 66 83
SE	P\$GBOX/RO		0.		aaa
Q_9	M.01	Regulator of MAT (ROM1, ROM2)	85	39	59 + 00 91
SE	P\$GBOX/BZI		0.		cagcctcog
Q_9	P910.02	bZIP transcription factor from Antirrhinum majus	84	44	64 - 50 55
SE	O\$RVUP/LTR		0.		GAGgtggt
Q_9	UP.01	Upstream element of C-type Long Terminal Repeats	76	55	75 + 00 04
SE	P\$GBOX/BZI		0.	10	12 0.7 0.7
Q_9	P911.02	bZIP transcription factor from Antirrhinum majus	76	7	7 - 50 65
SE	P\$DREB/HVD		0.	11	13 0.7 0.9
Q_9	RF1.01	H. vulgare dehydration-response factor 1	89	7	1 + 82 00
SE	P\$MADS/AGL		0.	13	15 0.7 0.8
Q_9	1.01	AGL1, Arabidopsis MADS-domain protein AGAMOUS-like 1	84	0	0 + 61 60
SE	P\$E2FF/E2F.	E2F class I sites	0.	13	15 1.0 0.8
Q_9	01		82	6	0 - 00 25

SE Q_9 01	P\$GTBX/GT1.	GT1-Box binding factors with a trihelix DNA-binding domain	0. 47 48 0.9 0.8	gattgtGTAAgtc-
SE Q_9 A.01	P\$MSAE/MS	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	85 1 7 + 68 52	tata
SE Q_9 2.01	P\$NCS2/NCS	Nodulin consensus sequence 2	0. 51 52 1.0 0.9	agtcAACGG-
SE Q_9 01	P\$LFYB/LFY.	Plant specific floral meristem identity gene LEAFY (LFY)	80 3 7 + 00 68	caaga
SE Q_9 A.01	P\$MSAE/MS	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0. 53 55 1.0 0.7	ggttgtCTCTgtga
SE Q_9 PH3.01	P\$MYBL/MYB	Myb-like protein of Petunia hybrida	79 8 2 + 00 95	a
SE Q_9 P911.02	P\$GBOX/BZI	bZIP transcription factor from Antirrhinum majus	0. 58 59 0.9 0.9	tCCCAatgtaaa
SE Q_9 01	P\$OPAQ/O2.	Opaque-2 regulatory protein	93 0 2 + 14 48	ggtaaAACGgtt-
SE Q_9 GCN4.01	P\$OPAQ/O2_	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0. 58 60 1.0 0.8	gat
SE Q_9 T.01	P\$CAAT/CAA	CCAAT-box in plant promoters	80 7 1 + 00 90	gataaacgGTT-
SE Q_9 T.01	P\$MIIG/P_AC	Maize activator P of flavonoid biosynthetic genes	0. 58 60 0.7 0.8	gtaaacgGTT-
SE Q_9 T.01	P\$IBOX/IBOX	I-Box in rbcS genes and other light regulated genes	80 8 4 + 50 31	Gatgat
SE Q_9 RICH.01	P\$CARM/CA	CA-rich element	0. 60 62 1.0 0.8	tggtggTGACatg-
SE Q_9 RAV5/RAV	P\$RAV5/RAV	5'-part of bipartite RAV1 binding site, interacting with AP2 domain	76 4 4 + 00 04	tatgattg
SE Q_9 1-5.01	P\$LEGB/RV.0	RY and Sph motifs conserved in seed-specific promoters	0. 60 62 0.7 0.9	tcatacatgTCAC-
SE Q_9 1	P\$LEGB/RV.0	Circadian clock associated 1	87 5 1 - 94 01	cacc
SE Q_9 1.01	P\$CCAF/CCA	OCS-like elements	0. 60 62 1.0 0.9	gtgggtACATg-
SE Q_9 P\$OCSE/OC	P\$OCSE/OC		81 6 2 + 00 41	tatgat
Q_9 SL.01	SL.01		0. 61 62 1.0 0.9	aaCCAAtca
			97 9 7 - 00 84	tggaGGTTgtccc
			0. 69 70 0.9 0.9	g
			93 5 9 + 66 69	g
			0. 72 74 0.7 0.8	ttgaaGAGAaggt-
			81 9 5 + 50 11	taag
			0. 73 75 1.0 0.8	ggcctgcAACAtct-
			78 9 7 - 00 03	taacc
			0. 77 78 1.0 0.9	g
			96 0 0 - 00 79	tgcaAACAcaaa
			0. 78 80 1.0 0.8	ggtagcttCATGca-
			87 2 8 - 00 95	aaatctcagctt
			0. 78 80 1.0 0.9	tcatgcaaAATCt-
			85 7 1 - 00 45	ca
			0. 79 81 0.7 0.7	caatcaaagaggt-
			69 8 8 - 69 06	gACTTcat

SE P\$GBOX/CP	Common plant regulatory factor (CPRF) from parsley	0.	13	14	1.0	0.9	gttagt-
Q_9 RF.01		95	80	00	+	00	cACGTgcttacctt
SE P\$MYCL/OSB		0.	13	13	1.0	0.9	ggtaagCACGT-
Q_9 HLH66.01	Rice bHLH protein	85	80	98	-	00	gactcaac
SE P\$MYCL/MY		0.	13	13	1.0	0.9	ttgagtcACGTgct-
Q_9 CRS.01	Myc recognition sequences	93	81	99	+	00	lacct
SE P\$OPAQ/RIT	Rice transcription activator-1 (RITA), basic leucine zipper protein, highly expressed	0.	13	13	1.0	0.9	gtaagcACGT-
Q_9 A1.01	during seed development	95	81	97	-	00	gactcaa
SE P\$ABRE/ABR		0.	13	13	1.0	0.8	ggtaagcACGT-
Q_9 E.01	ABA response elements	82	82	98	-	00	gactca
SE O\$RPOA/AP		0.	13	14	1.0	0.7	ccctcTAAA-
Q_9 OLYA.01	Avian C-type LTR PolyA signal	71	97	17	+	00	aagccttttga
SE P\$DOFF/DOF		0.	13	14	1.0	0.9	ccctctaaA-
Q_9 3.01	Dof3 - single zinc finger transcription factor	99	97	13	+	00	AAGccttt
SE O\$RPOA/AP		0.	14	14	0.7	0.7	ttgatCAAA-
Q_9 OLYA.01	Avian C-type LTR PolyA signal	71	01	21	-	50	aaggcttttag
SE P\$LFYB/LFY.		0.	14	14	0.9	0.9	
Q_9 01	Plant specific floral meristem identity gene LEAFY (LFY)	93	14	26	-	14	aACCAttgatcaa
SE P\$SBPD/SBP		0.	14	14	1.0	0.8	atttgGTACAac-
Q_9 .01	SQUA promoter binding proteins	88	19	35	-	00	cattg
SE P\$SBPD/SBP		0.	14	14	1.0	0.9	tggttGTACca-
Q_9 .01	SQUA promoter binding proteins	88	22	38	+	00	aatgag
SE P\$MADS/AG.	Agamous, required for normal flower development, similarity to SRF (human) and	0.	14	14	1.0	0.8	ttgTACCaatga-
Q_9 01	MCM (yeast) proteins	80	25	45	+	00	gaagagag
SE P\$GAGA/BPC		1.	14	14	1.0	1.0	gagaagAGAGac-
Q_9 .01	Basic pentacycstine proteins	00	36	60	+	00	tataaagcgttgca
SE P\$SPF1/SP8	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b	0.	14	14	1.0	0.8	
Q_9 BF.01	(TACTATT) sequences of sporamin and beta-amylase genes	87	64	76	+	00	ttTACTtttactt
SE P\$SPF1/SP8	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b	0.	14	14	1.0	0.8	
Q_9 BF.01	(TACTATT) sequences of sporamin and beta-amylase genes	87	70	82	+	00	ttTACTtttactt
SE P\$TBPF/TAT		0.	14	14	1.0	0.9	ttccTATAtaaaa-
Q_9 A.02	Plant TATA box	90	81	95	+	00	aag
SE P\$TBPF/TAT		0.	14	14	1.0	0.9	cctaTATAAaa-
Q_9 A.01	Plant TATA box	88	83	97	+	00	aagtc

Table 4: cis-regulatory elements of SEQ ID NO: 10

SEQ	P\$PSRE/GAAA.0	GAAA motif involved in pollen specific transcriptional activation	0.	83	1	1.0	0.8	tgatGAAAtcgtatcg
SEQ _10	1				3	9	+ 00 81	
SEQ	P\$MIIG/P_ACT.0	Maize activator P of flavonoid biosynthetic genes	0.	93	9	1	0.9 0.9	agctGTTggtacga
SEQ _10	1				9	3	+ 66 83	
SEQ					1	1		
SEQ _10	P\$SBPD/SBP.01	SQUA promoter binding proteins	0.	88	0	1	1.0 0.9	ttatcGTACcaaccagc
SEQ _10	1				6	- 00 14		
SEQ	P\$IBOX/GATA.0	Class I GATA factors	0.	93	0	2	1.0 0.9	ggtacGATAataatgta
SEQ _10	1				7	3	+ 00 58	
SEQ	P\$AHBP/HAHB4.	Sunflower homeodomain leucine-zipper protein Hahb-4	0.		1	2	1.0 0.9	tacattATTAt
SEQ _10	01				3	3	- 00 09	
SEQ	P\$OPAQ/O2_GC	Recognition site for BZIP transcription factors that belong to the group of	0.	81	2	4	0.8 0.8	tgcacACCTgcttaaa
SEQ _10	N4.01	Opaque-2 like proteins			7	3	+ 29 18	
SEQ		Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	0.	85	9	3	- 33 63	accCAAAActatttaa
SEQ _10	P\$STKM/STK.01				1	1		

SEQ _10	P\$CARM/CARIC H.01	CA-rich element	0.	78	1 4 5 1	1 6 3 1	1.0 0.0 - 0.0	0.8 32 0.9	tgaacaAACAccca- aact
SEQ _10	P\$AREF/ARE.01	Auxin Response Element	0.	93	8 2 1	9 4 2	1.0 0.0 + 0.0	0.9 51	ttaTGTCtcagtt
SEQ _10	P\$MYBL/GAMYB .01	GA-regulated myb gene from barley	0.	91	8 4 2	0 0 2	1.0 0.0 + 0.0	0.9 15	atgtctcaGTTAtatct
SEQ _10	P\$OPAQ/O2_GC N4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.	81	0 1 2	1 7 2	0.8 - 0.0	0.8 21	gaaattACTTggtttac
SEQ _10	P\$CARM/CARIC H.01	CA-rich element	0.	78	1 2 2	3 0 2	1.0 0.0 - 0.0	0.7 90	gtgaacgAACACcga- aatt
SEQ _10	P\$GBOX/HBP1B .01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.	83	3 4 2	5 4 2	1.0 0.0 - 0.0	0.8 34	actttgggACGTaag- taat
SEQ _10	P\$DOFF/DOF2.0 1	Dof2 - single zinc finger transcription factor	0.	98	5 6 2	7 2 2	1.0 0.0 + 0.0	0.9 95	atatctatAAAGcaagt
SEQ _10	P\$LREM/ATCTA. 01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.	85	5 6 2	6 6 2	1.0 0.0 + 0.0	0.9 22	atATCTataaa
SEQ _10	P\$TBPF/TATA.0 1	Plant TATA box	0.	88	5 7 2	7 1 2	1.0 0.0 + 0.0	0.8 82	tatcTATAaagcaag
SEQ _10	P\$OPAQ/O2_GC N4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0.	81	6 2 2	7 8 2	0.8 - 0.0	0.8 46	cagatcACTTgctttat
SEQ _10	O\$RPAD/PADS. 01	Mammalian C-type LTR Poly A downstream element	0.	87	6 8 0	8 0 +	0.8 83	0.8 76	caaGTGAtctgat
SEQ _10	P\$MYBS/HVMC	Hordeum vulgare Myb-related CAB-promoter-binding protein 1	0.	0.	2 2 2	2 2 +	1.0 0.0 + 0.0	0.9	tcgatATCCtatgaaa

	SEQ	P\$GAPB/GAP.01	Cis-element in the GAPDH promoters conferring light inducibility	0.	8	9	1.0	0.9	
	_10			88	2	6	+	00	17
					3	3			
SEQ				0.	3	5		1.0	0.8
_10		P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	85	5	1	+	00	55
					3	3			
SEQ				0.	4	6		0.7	0.8
_10		P\$TELO/ATPUR A.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	85	7	1	+	50	68
					4	4			
SEQ				0.	3	4		1.0	0.9
_10		P\$TBPF/TATA.0	Plant TATA box	88	2	6	-	00	85
					4	4			
SEQ				0.	4	5		0.8	0.9
_10		P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	89	3	3	+	29	04
					4	4			
SEQ				0.	4	5		0.9	0.9
_10		P\$AHBP/ATHB5.01	HDZip class I protein ATHB5	89	3	3	-	36	77
					4	4			
SEQ				0.	5	6		1.0	0.9
_10		P\$AHBP/WUS.0	Homeodomain protein WUSCHEL	94	1	1	-	00	63
					4	4			
SEQ				0.	5	7		1.0	0.8
_10		P\$PSRE/GAAA.0	GAAA motif involved in pollen specific transcriptional activation	83	5	1	+	00	76
					4	4			
SEQ				0.	7	8		0.7	0.8
_10		P\$MYBL/MYBPH 3.01	Myb-like protein of Petunia hybrida	80	2	8	-	50	33
					4	5			
SEQ				0.	8	0		1.0	0.8
_10		P\$PSRE/GAAA.0	GAAA motif involved in pollen specific transcriptional activation	83	8	4	+	00	37
SEQ				0.	4	5		0.9	0.8
_10		P\$GTBX/GT1.01	GT1-Box binding factors with a trihelix DNA-binding domain	85	9	0	-	68	91
								aattttGTAAatttct	

0.	1	3	1.0	0.8	
81	5	1	-	00	55

ctgtgtACATggctaaa

ANT (Arabidopsis protein INTEGUMENTA), member of the plant-specific family of AP2/EREBP-transcription factors

0.	4	5	1.0	0.8
81	1	7	-	00
				20

agaatatTCCaatgct

SEQ P\$MYBS/TAMYB
_10 80.01 MYB protein from wheat

0. 4 5 1.0 0.9
83 2 8 - 00 87 gagaATATtcccaatgc

SEQ P\$MYBS/TAMYB
_10 80.01 MYB protein from wheat

0. 4 6 1.0 0.9
83 7 3 + 00 41 gggaATAttcttccac

SEQ	Heat shock element
10	P\$HEAT/HSE.01

0.	6	8	1.0	0.8
81	7	1	-	00
				35
				tgaagaaaaAGAAt

SEQ P\$IBOX/GATA.0
_10 1

0.79	1.0	0.9	
93	84	-	gcataGATAaigtgaa

SEQ P\$DOFF/DOF3.0
10 1

0.80	1.00	0.97
aaatttcaAAAGcatag		

SEQ 10	P\$PREM/MGPR OTORE.01	Promoter elements involved in MgProto (Mg-protoporphyrin IX) and light-mediated induction
-----------	--------------------------	---

0.	9	2	1.0	0.7	catCGACaacaagtta
77	2	2	-	00 78	aatttcaaaaqca

SEQ P\$PSRE/GAAA.0
10 1
GAAA motif involved in pollen specific transcriptional activation

0. 9 1 1.0 0.8
83 4 0 + 00 37 ctttGAAAttttaact

SEQ
10 P\$IDDF/ID1.01 Maize INDETERMINATE1 zinc finger protein

0. 0 2 1.0 0.9
92 9 1 + 00 39 cttgTTGTcgtg

SEQ P\$MYBS/HVMC
10 B1.01
Hordeum vulgare Myb-related CAB-promoter-binding protein 1

	0	1	3	0	0.9
0.	93	4	0	-	00
93					45

aaqcatATCCatqaca

SEQ _10	P\$MYBS/TAMBYB 80.01	MYB protein from wheat	0. 6 83 1 3 9 5 + 6 6 1.0 0.9 atggATATgcttagaga
SEQ _10	P\$PSRE/GAAA.0 1	GAAA motif involved in pollen specific transcriptional activation	0. 6 83 2 4 9 5 + 6 6 1.0 0.9 ttagaGAAAttttcaaa
SEQ _10	P\$AHBP/WUS.0 1	Homeodomain protein WUSCHEL	0. 6 94 3 3 6 6 1.0 0.9 aacatTAATca
SEQ _10	P\$AHBP/WUS.0 1	Homeodomain protein WUSCHEL	0. 6 94 4 4 6 6 1.0 1.0 - 00 00 ttgatTAATgt
SEQ _10	P\$IBOX/GATA.0 1	Class I GATA factors	0. 7 93 0 6 6 7 1.0 0.9 + 00 71 atcaaGATAagctagca
SEQ _10	P\$PALA/PALBO XA.01	Putative cis-acting element on various PAL and 4CL gene promoters	0. 9 84 8 6 7 7 0.8 0.8 + 25 58 tgactgtCCATccaatata
SEQ _10	P\$CAAT/CAAT.0 1	CCAAT-box in plant promoters	0. 0 97 7 5 7 7 1.0 0.9 + 00 83 atCCAAat
SEQ _10	O\$RPOA/APOLY A.01	Avian C-type LTR PolyA signal	0. 4 71 0 6 7 7 0.7 0.7 + 50 65 tataaTATAtcgacaattttt
SEQ _10	P\$GTBX/SBF1.0 1	SBF-1	0. 5 87 3 9 7 7 1.0 0.8 - 00 85 cgttttcTTAAaaattg
SEQ _10	P\$MSAE/MSA.0 1	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0. 6 80 1 5 7 7 1.0 0.8 + 00 71 aagaaAACGgtataa
SEQ _10	O\$RVUP/LTRUP .01	Upstream element of C-type Long Terminal Repeats	0. 6 76 6 6 7 8 + 1.0 0.8 aacggtataacTTTCa- aagct
SEQ	P\$L1BX/ATML1.	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0. 7 8 + 1.0 0.8 ggtttaTAAAGtgcaaa

SEQ _10 2	P\$TBPF/TATA.0	Plant TATA box	0. 90	9 1	0 5	1.0 + 00	0.9 14	ggttTATAaatgta
SEQ _10	P\$OPAQ/O2_GC N4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0. 81	9 3	0 9	1.0 - 00	0.8 58	ccttgACAItataaa
SEQ _10 01	P\$WBXF/WRKY.	WRKY plant specific zinc-finger-type factor associated with pathogen de- fence, W box	0. 92	9 5	1 - 00	1.0 00	0.9 36	gtcctTTGAcattata
SEQ _10 3.01	P\$MYBL/MYBPH	Myb-like protein of Petunia hybrida	0. 80	1 8	3 4	1.0 + 00	0.9 08	tcaaacccGTTAgtcaa
SEQ _10 1	P\$MSAE/MSA.0	M-phase-specific activators (NtmybA1, NtmybA2, NtmybB)	0. 80	1 9	3 - 00	1.0 00	0.8 54	tgactAACGggtttg
SEQ _10 3.02	P\$MYBL/MYBPH	Myb-like protein of Petunia hybrida	0. 76	2 2	3 8	1.0 + 00	0.7 64	accgtTAGTcaaggtt
SEQ _10 01	P\$OCSE/OCSL.	OCS-like elements	0. 69	5 0	7 0	0.7 + 69	0.7 05	atacacaagcactcACC- Tact
SEQ _10 L.01	P\$MIIG/PALBOX	Cis-acting element conserved in various PAL and 4CL promoters	0. 80	6 0	7 4	1.0 - 00	0.8 51	gtgtagtaGGTGagt
SEQ _10 RE.01	P\$DREB/CRT_D RE.01	C-repeat/dehydration response element	0. 89	6 9	8 3	1.0 + 00	0.9 23	ctacaCCGAcactga
SEQ _10 .01	P\$OCSE/OCSTF	bZIP transcription factor binding to OCS-elements	0. 73	7 5	9 + 00	1.0 00	0.7 73	cgacactgacatt- GACGtctc
SEQ _10 .01	P\$GBOX/HBP1B	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0. 83	8 8	9 0	1.0 - 00	0.8 91	aatcagagACGT- caatgtcag

Table 5: cis- regulatory elements of SEQ ID NO: 11

SEQ P\$AHBPWUS. _11 01	Homeodomain protein WUSCHEL	0. 9 1.0 1.0 4 9 19 + 00 00 tcctTAATgg
SEQ P\$GTBX/SBF1 _11 .01	SBF-1	0. 8 1.0 0.8 7 16 32 + 00 86 atggagcTTAAactct
SEQ P\$L1BX/ATML _11 1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0. 8 1.0 0.8 2 31 47 - 00 35 ttgtatTAAAttcagaa
SEQ P\$MYBL/MYB _11 PH3.02	Myb-like protein of Petunia hybrida	0. 7 0.7 0.8 6 44 60 + 78 32 ca
SEQ P\$WBXF/ERE. _11 01	Elicitor response element	0. 8 1.0 0.9 9 91 7 - 00 03 ttattcTGACcaattgta
SEQ P\$MYBL/GAM _11 YB.01	GA-regulated myb gene from barley	0. 9 1.0 0.9 1 0 6 - 00 19 ttttcttGTTAttctg
SEQ P\$GTBX/GT1. _11 01	GT1-Box binding factors with a trihelix DNA-binding domain	0. 8 1.0 0.9 5 0 6 - 68 58 attgtaGTAAtttctt
SEQ P\$NCS1/NCS1 _11 .01	Nodulin consensus sequence 1	0. 8 0.8 0.8 5 5 5 + 04 82 aAAACgatttg
SEQ P\$GBOX/UPR _11 E.01	UPRE (unfolded protein response element) like motif	0. 8 1.0 0.9 6 7 7 - 00 50 aatcggt
SEQ P\$GBOX/GBF _11 1.01	bZIP protein G-Box binding factor 1	0. 9 1.0 0.9 4 8 8 + 00 87 taagtat

SEQ _11	P\$WBXF/WRK Y.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.9	13	15	1.0	0.9	acgatTTGAcgtgg- taa
SEQ _11	P\$ABRE/ABRE .01	ABA response elements	0.8	13	15	1.0	0.8	cgatttgACGTgg- taag
SEQ _11	P\$OPAQ/O2.0 2	Opaque-2 regulatory protein	0.8	13	15	1.0	0.9	cttaCCACgtca- aatcg
SEQ _11	P\$GAPB/GAP. 01	Cis-element in the GAPDH promoters conferring light inducibility	0.8	15	16	1.0	0.8	
SEQ _11	P\$CAAT/CAAT .01	CCAAT-box in plant promoters	0.9	16	16	1.0	0.9	agCCAAAtga
SEQ _11	P\$GBOX/GBF 1.01	bZIP protein G-Box binding factor 1	0.9	17	19	1.0	0.9	ctgttatACGTgtga- gaact
SEQ _11	P\$ABRE/ABRE .01	ABA response elements	0.8	17	18	1.0	0.8	ttgttatACGTgtgaga
SEQ _11	P\$AHBP/HAH B4.01	Sunflower homeodomain leucine-zipper protein Hahb-4	0.8	19	20	1.0	0.9	catataATTAg
SEQ _11	O\$LTUP/TAAC C.01	Lentiviral TATA upstream element	0.7	26	28	1.0	0.7	tacttaagtcAACC- caaacag
SEQ _11	P\$CGCG/ATS R1.01	Arabidopsis thaliana signal-responsive gene1, Ca2+/ calmodulin binding protein homolog to NtER1 (tobacco early ethylene-responsive gene)	0.8	29	31	1.0	0.9	cccCGCGtaatttcc- ga
SEQ _11	P\$AHBPWUS. 01	Homeodomain protein WUSCHEL	0.9	31	32	1.0	0.9	
SEQ _11	P\$CAAT/CAAT	CCAAT-box in plant promoters	0.4	1	1	-	00	63 ccaatTAATcc
			0.31	32	-	1.0	0.9	ccCCAAAtta

[illegible]

SEQ P\$CAAT/CAAT
11 .01
CCAAT-box in plant promoters

7	7	3	00	71	
5					
0.					
9	89	90	1.0	1.0	
4	1	1	+	00	00
0.					tatttTAATga
8	92	93	0.8	0.8	
5	0	6	+	43	81
0.					tctgtgGTGAatgattga
8	92	93	0.9	0.9	
9	7	7	-	36	39
0.					ttaATCAttca
8	92	93	1.0	0.9	
7	7	7	+	00	45
0.					tgaatgATTAA
8	92	94	1.0	0.9	
7	7	3	+	00	13
0.					tgaatgaTTAAAttcac
8	92	94	1.0	0.8	
2	9	5	+	00	40
0.					aatgatTAAAttcacca
9	93	94	1.0	0.9	
4	1	1	-	00	63
0.					gaattTAATca
8	93	94	0.8	0.9	
5	3	9	-	43	19
0.					agtggtGTGAatttaaat
9	94	95	1.0	0.9	
7	3	9	-	00	84
0.					ggttatgaAAAGtggtgg
9	95	96	1.0	0.9	
1	0	6	-	00	19
0.	95	96	1.0	0.9	
9	6	4	+	00	84
					aaCCCAAttca

SEQ _11 .01	P\$GTBX/SBF1	SBF-1	7	0.	97	99	1.0	0.9	ggagtagTTAAaa-
			7	7	3	+	00	49	aaga
SEQ _11 .01	P\$NCS2/NCS2	Nodulin consensus sequence 2	7	0.	98	99	0.7	0.8	
			9	4	8	-	50	46	tattgtCTTTttaa
SEQ _11 _IY.02	P\$HMGF/HMG	High mobility group I/Y-like protein isolated from pea	0	99	10	1.0	1.0		
			0	4	08	-	00	00	atttTATTtttattg
SEQ _11 _IY.01	P\$HMGF/HMG	High mobility group I/Y-like proteins	8	99	10	1.0	0.9		
			9	9	13	-	00	24	ttttTATTttattt
SEQ _11 A.01	P\$MADS/SQU	MADS-box protein SQUAMOSA	0.	10	10	1.0	0.9		ttcagctATTTttattt-
			0	01	21	-	00	50	tatt
SEQ _11 01	P\$SEF4/SEF4.	Soybean embryo factor 4	0.	9	10	1.0	0.9		
			8	05	15	-	00	85	taTTTTtattt
SEQ _11 01	P\$CE1F/ABI4.	ABA insensitive protein 4 (ABI4)	0.	8	10	1.0	0.8		
			7	26	38	+	00	91	acaaCACCGacat
SEQ _11 DRE.01	P\$DREB/CRT_	C-repeat/dehydration response element	0.	8	10	1.0	0.9		
			9	27	41	+	00	62	caacaCCGAcatga
SEQ _11 01	P\$AHBPWUS.	Homeodomain protein WUSCHEL	0.	9	10	1.0	0.9		
			4	39	49	-	00	63	gacgtTAATca
SEQ _11 RA.01	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8	10	10	0.7	0.8		
			5	53	67	+	50	63	taaaAACCTagacta
SEQ _11 02	P\$TBP/ATA.	Plant TATA box	0.	9	10	1.0	0.9		
			0	62	76	+	00	24	agacTATAaaacat

SEQ	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.
_11	RA.01		8
			5
			68
			82
			+ 50
			68
			taaaACCAataaatcc
			0.
SEQ	O\$RPOA/POL		7
_11	YA.01		6
		Mammalian C-type LTR Poly A signal	71
			91
			+ 00
			04
			aatctga
			0.
SEQ	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8
_11	RA.01		5
			75
			89
			+ 50
			60
			ataaATCCtaaatct
			0.
SEQ	P\$CCAF/CCA1		8
_11	.01		5
		Circadian clock associated 1	77
			91
			+ 00
			86
			aaatcctaAATCtga
			0.
SEQ	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8
_11	RA.01		5
			89
			03
			+ 50
			58
			tgaaACACtaaacca
			0.
SEQ	O\$RPOA/POL		7
_11	YA.01		6
			92
			12
			+ 00
			07
			aatctca
			0.
SEQ	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8
_11	RA.01		5
			96
			10
			+ 50
			62
			ctaaACCAataaatct
			0.
SEQ	P\$CCAF/CCA1		8
_11	.01		5
		Circadian clock associated 1	98
			12
			+ 00
			05
			aaaccataAATCtca
			0.
SEQ	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8
_11	RA.01		5
			11
			11
			0.7
			0.8
			16
			30
			+ 50
			57
			ctaaACTCtaaaccc
			0.
SEQ	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8
_11	RA.01		5
			11
			11
			1.0
			0.9
			23
			37
			+ 00
			85
			ctaaACCCtaaatct
			0.
SEQ	P\$CCAF/CCA1		8
_11	.01		5
		Circadian clock associated 1	11
			11
			1.0
			0.8
			25
			39
			+ 00
			92
			aaaccctaAATCtta
			0.
SEQ	P\$DOFF/PBO	Prolamin box, conserved in cereal seed storage protein gene promoters	0.
			11
			11
			- 1.0
			0.7
			ttgggtttAAAAGtttaa

	SEQ	P\$GTBX/SBF1 _11 .01	SBF-1		8	11	11	1.0	0.8	
					7	38	54	-	00	tttgggtTTAAagttta
					0.					
	SEQ	P\$SEF3/SEF3. _11 .01	SEF3, Soybean embryo factor 3		8	11	11	1.0	0.8	
					7	43	57	+	00	tttaaACCCaaactc
					0.					
	SEQ	P\$TELO/ATPU _11 RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha		8	11	11	1.0	0.8	
					5	44	58	+	00	ttaaACCCaaactct
					0.					
	SEQ	P\$TELO/ATPU _11 RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha		8	11	11	0.7	0.8	
					5	50	64	+	50	ccaaACTCtaaacaaa
					0.					
	SEQ	P\$STKM/STK. _11 .01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression		8	11	11	1.0	0.8	
					5	55	69	+	00	ctcTAAAacaataaac
					0.					
	SEQ	O\$RPOA/LPO _11 LYA.01	Lentiviral Poly A signal		9	11	11	1.0	0.9	aacAATAaacctta-
					4	60	80	+	00	aatccta
					0.					
	SEQ	P\$TELO/ATPU _11 RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha		8	11	11	0.7	0.8	
					5	64	78	+	50	ataaACCCTtaaattcc
					0.					
	SEQ	P\$TELO/ATPU _11 RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha		8	11	11	0.7	0.8	
					5	71	85	+	50	ttaaATCCtaaaaatc
					0.					
	SEQ	P\$CCAF/CCA1 _11 .01	Circadian clock associated 1		8	11	11	1.0	0.9	
					5	74	88	+	00	aatcctaaAATCtaa
					0.					
	SEQ	P\$LREM/ATCT _11 A.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes		8	11	11	1.0	0.9	
					5	81	91	+	00	aaATCTaaacc
	SEQ	P\$TELO/ATPU _11 RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha		0.	11	11	1.0	0.9	
					8	85	99	+	00	ctaaACCCtaaacc

SEQ _11 .01	P\$TELO/RPBX	Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression	5 0.	11 92	12 06	0.7 + 55	0.8 47	ctaaaCCCAaagcta
SEQ _11 .01	P\$LEGB/LEGB	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	5 9	11 97	12 23	0.7 + 50	0.6 07	ccaaagCTATaaac- cataaaccataa
SEQ _11 RA.01	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8 5	12 06	12 20	0.7 + 50	0.8 55	ataaACCAtaaacca
SEQ _11 RA.01	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8 5	12 13	12 27	0.7 + 50	0.8 52	ataaACCAtaaaatc
SEQ _11 .01	P\$CCAF/CCA1	Circadian clock associated 1	8 5	12 16	12 30	1.0 + 00	0.9 50	aaccataaAATCtaa
SEQ _11 PEPA.01	O\$RPOA/DTY	PolyA signal of D-type LTRs	7 8	12 17	12 37	1.0 + 00	0.8 52	aCCATaaatcta- aaccctaa
SEQ _11 C.01	O\$LTUP/TAAC	Lentiviral TATA upstream element	7 1	12 18	12 40	1.0 + 00	0.7 13	ccataaaaatcta- AACCCtaaatc
SEQ _11 A.01	P\$LREM/ATCT	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	8 5	12 23	12 33	1.0 + 00	0.9 70	aaATCTaaacc
SEQ _11 RA.01	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8 5	12 27	12 41	1.0 + 00	0.9 85	ctaaACCCtaaatcc
SEQ _11 RA.01	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8 5	12 34	12 48	0.7 + 50	0.8 62	ctaaATCCtaaatcc
SEQ _11 RA.01	P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	8 5	12 41	12 55	0.7 + 50	0.8 57	ctaaATCCtaaaacc

SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.8	12	12	1.0	0.9	
5			48	62	+	00	80	ctaaACCCtaaactt
0.								
8			12	12	1.0	0.8		
5	SEQ _11	P\$NCS1/NCS1 .01	55	65	-	00	95	aAAAAgttag
0.		Nodulin consensus sequence 1						
8			12	12	1.0	0.8		
7	SEQ _11	P\$GTBX/SBF1 .01	62	78	-	00	89	aa
0.		SBF-1						
8			12	12	1.0	0.8		tttagggTTAAaaaa-
0.								
8			12	12	1.0	0.8		
4	SEQ _11	P\$TELO/RPBX .01	67	81	+	00	98	tttaaCCCTaaactc
0.		Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression						
8			12	13	0.7	0.8		
5	SEQ _11	P\$TELO/ATPU RA.01	93	07	+	50	55	tcaaATCCtaaactc
0.		Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha						
8			13	13	0.7	0.8		
5	SEQ _11	P\$TELO/ATPU RA.01	00	14	+	50	57	ctaaACTCtaaaccc
0.		Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha						
8			13	13	1.0	0.8		
7	SEQ _11	P\$SEF3/SEF3. .01	06	20	+	00	91	tctaaACCCaaaact
0.		SEF3, Soybean embryo factor 3						
8			13	13	0.7	0.8		
4	SEQ _11	P\$TELO/RPBX .01	07	21	+	55	58	ctaaaCCCCAaaactt
0.		Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression						
8			13	13	0.7	0.8		
5	SEQ _11	P\$TELO/ATPU RA.01	21	35	+	50	55	tcaaATCCtaaaccc
0.		Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha						
8			13	13	1.0	0.8		
5	SEQ _11	P\$TELO/ATPU RA.01	28	42	+	00	61	ctaaACCCaaacctc
0.		Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha						
8			13	13	1.0	0.9		
5	SEQ _11	P\$TELO/ATPU RA.01	42	56	+	00	80	ctaaACCCtaaactt
0.	SEQ	P\$MYBL/MYB	13	13	+	1.0	0.8	atctgcTAGTtaa-
		Myb-like protein of Petunia hybrida						

_11 PH3.02			7	57	73	00	08	taag
			6					
SEQ P\$OCSE/OCS			0.					
_11 L.01	OCS-like elements		6	13	13	0.7	0.7	taagattaaggtt-
			9	70	90	+	69	06 tACGGttt
			0.					
SEQ P\$AHBPWUS.			9	13	13	1.0	0.9	
_11 01	Homeodomain protein WUSCHEL		4	72	82	-	00	63 aacctTAAATct
			0.					
SEQ P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein		8	13	13	0.7	0.8	
_11 RA.01	Pur-alpha		5	78	92	-	50	57 ctaaACCGtaaacct
			0.					
SEQ P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein		8	13	14	1.0	0.8	
_11 RA.01	Pur-alpha		5	92	06	-	00	59 ataaACCCaaacdc
			0.					
SEQ P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein		8	13	14	0.7	0.8	
_11 RA.01	Pur-alpha		5	99	13	-	50	55 ataaACCAtaaaccc
			0.					
SEQ P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein		8	14	14	1.0	0.9	
_11 RA.01	Pur-alpha		5	13	27	-	00	88 ccaaACCCtaaatca
			0.					
SEQ P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein		8	14	14	1.0	0.8	
_11 RA.01	Pur-alpha		5	19	33	-	00	57 ttaaACCCaaaccc
			0.					
SEQ P\$CCAF/CCA1			8	14	14	1.0	0.8	
_11 .01	Circadian clock associated 1		5	31	45	-	00	72 aaactttaAATCtta
			0.					
SEQ P\$TELO/RPBX	Ribosomal protein box, appears unique to plant RP genes and genes associated		8	14	14	0.7	0.8	
_11 .01	with gene expression		4	40	54	-	55	58 ctaaaCCCAaaact
			0.					
SEQ P\$SEF3/SEF3.			8	14	14	1.0	0.8	
_11 01	SEF3, Soybean embryo factor 3		7	41	55	-	00	72 cctaaACCCaaaact
SEQ P\$TELO/ATPU	Arabidopsis Telo-box interacting protein related to the conserved animal protein		0.	14	14	1.0	0.9	
_11 RA.01	Pur-alpha		8	47	61	-	00	80 ctaaACCCtaaaccc

SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.	8	15	15	0.7	0.8	
			5	23	37	-	50	62	ctaaACCAtaaactct
SEQ _11	P\$LREM/ATCT A.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.	8	15	15	1.0	0.9	
			5	31	41	-	00	70	aaATCTaaacc
SEQ _11	P\$CCAF/CCA1 .01	Circadian clock associated 1	0.	8	15	15	1.0	0.9	
			5	34	48	-	00	38	aatcctaaAATCtaa
SEQ _11	O\$RPOA/POL YA.01	Mammalian C-type LTR Poly A signal	0.	7	15	15	1.0	0.7	aaccaTAAAtcc-
			6	41	61	-	00	84	caatcctaa
SEQ _11	O\$RPOA/POL YA.01	Mammalian C-type LTR Poly A signal	0.	7	15	15	1.0	0.7	aacacTAAAccata-
			6	48	68	-	00	68	aatccca
SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.	8	15	15	0.7	0.8	
			5	50	64	-	50	62	ctaaACCAtaaattcc
SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.	8	15	15	0.7	0.8	
			5	57	71	-	50	67	caaaACACtaaacca
SEQ _11	P\$TELO/RPBX .01	Ribosomal protein box, appears unique to plant RP genes and genes associated with gene expression	0.	8	15	15	1.0	0.9	
			4	71	85	-	00	89	ataaaCCCTaaattcc
SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.	8	15	15	0.7	0.8	
			5	78	92	-	50	63	taaaACCAtaaacc
SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.	8	15	16	0.7	0.8	
			5	86	00	-	50	54	ctaaACTCtaaaacc
SEQ _11	P\$TELO/ATPU RA.01	Arabidopsis Telo-box interacting protein related to the conserved animal protein Pur-alpha	0.	8	15	16	0.7	0.8	
			5	93	07	-	50	63	taaaAACCTaaactc
SEQ	P\$HOCT/HOC	Octamer motif found in plant histone H3 and H4 genes	0.	16	16	-	1.0	0.8	tagcgcgATCCgca-

_11	T.01		7	18	34	00	38	aaa
			6					
			0.					
SEQ P\$AHBPWUS.		Homeodomain protein WUSCHEL	9	16	16	1.0	1.0	
_11	01		4	32	42	+	00	00 ctaatTAATgt
			0.					
SEQ P\$DREB/CRT_		C-repeat/dehydration response element	8	16	16	1.0	0.9	
_11	DRE.01		9	36	50	-	00	65 caacaCCGAacattaa
			0.					
SEQ P\$CE1F/ABI4.		ABA insensitive protein 4 (ABI4)	8	16	16	1.0	0.8	
_11	01		7	39	51	-	00	91 acaaCACCCgacat
			1.					
SEQ P\$HMGF/HMG		High mobility group I/Y-like protein isolated from pea	0	16	16	1.0	1.0	
_11	_IY.02		0	49	63	+	00	00 tgttTATTtttttag
			0.					
SEQ P\$STKM/STK.		Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	8	16	16	1.0	0.8	
_11	01		5	51	65	-	00	77 agcTAAAAaaaaataaa
			0.					
SEQ P\$OCSE/OCS		OCS-like elements	6	16	16	0.7	0.7	
_11	L.01		9	64	84	+	69	26 ctatttaatttttACTTtta
			0.					
SEQ P\$STKM/STK.		Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	8	16	16	1.0	0.8	
_11	01		5	67	81	-	00	86 aagTAAAAaaaataaa
			0.					
SEQ P\$L1BX/ATML		L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	8	16	16	1.0	0.8	
_11	1.01		2	74	90	-	00	46 aaa
			0.					
SEQ P\$SPF1/SP8B		DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	8	16	16	1.0	0.8	
_11	F.01		7	75	87	+	00	77 ttTACTttttatg
			0.					
SEQ P\$STKM/STK.		Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	8	16	17	1.0	0.8	
_11	01		5	91	05	+	00	72 ttTAAAcattttat
SEQ P\$MADS/SQU		MADS-box protein SQUAMOSA	0.	16	17	1.0	0.9	
_11	A.01		9	93	13	+	00	42 ttaaacATTtTatat-gaca

SEQ P\$TBPF/TATA. _11 02	Plant TATA box	0	16 17	1.0 0.9	
		0.	96 10 - 00 56	cataTATAaatagtt	
SEQ P\$TBPF/TATA. _11 02	Plant TATA box	0.	16 17	1.0 0.9	
		9	98 12 - 00 13	gtcaTATAtaaatag	
SEQ P\$LEGB/LEGB _11 .01	Legumin box, highly conserved sequence element about 100 bp upstream of the TSS in legumin genes	0.	17 17	0.7 0.6	ttcataaCCAAcccaa-
		5	04 30 - 50 07	aatgcatatat	
SEQ P\$MIIG/P_ACT _11 .01	Maize activator P of flavonoid biosynthetic genes	0.	17 17	0.9 0.9	
		9	14 28 + 66 73	ttttGGTTggttatg	
SEQ P\$MYBL/GAM _11 YB.01	GA-regulated myb gene from barley	0.	17 17	1.0 0.9	
		9	15 31 + 00 43	tttggttgGTTAtgaaa	
SEQ P\$GAPB/GAP. _11 01	Cis-element in the GAPDH promoters conferring light inducibility	0.	17 17	1.0 0.9	
		8	22 36 + 00 33	ggttATGAaaagtac	
SEQ P\$GTBX/GT1. _11 01	GT1-Box binding factors with a trihelix DNA-binding domain	0.	17 17	0.8 0.8	agtacgGTGAatt-
		8	32 48 + 43 89	taac	
SEQ P\$L1BX/ATML _11 1.01	L1-specific homeodomain protein ATML1 (A. thaliana meristem layer 1)	0.	17 17	1.0 0.8	
		8	36 52 - 00 23	gaatgtTAAAttcaccg	
SEQ P\$GTBX/SBF1 _11 .01	SBF-1	0.	17 17	1.0 0.8	
		8	38 54 - 00 94	gtgaatgTTAAAttcac	
SEQ P\$GTBX/GT1. _11 01	GT1-Box binding factors with a trihelix DNA-binding domain	0.	17 17	0.8 0.8	cctacgGTGAatgt-
		8	44 60 - 43 89	taa	
SEQ O\$RPOA/DTY _11 PEPA.01	PolyA signal of D-type LTRs	0.	17 17	0.7 0.8	aCAATtaaaatataa-
		7	71 91 - 50 34	caatac	

SEQ O\$RPOA/APO _11 LYA.01	Avian C-type LTR PolyA signal	0. 7 17 18 0.7 0.7 0.7 0.7 1 98 18 - 50 17 0. 17 cacttgga
SEQ P\$CARM/CARI _11 CH.01	CA-rich element	0. 7 18 18 1.0 0.7 0.7 0.7 8 07 25 - 00 85 0. 17 aaca
SEQ P\$WBXF/WRK _11 Y.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0. 9 18 18 1.0 0.9 0.9 0.9 2 13 29 + 00 64 0. 17 ca
SEQ P\$LEGB/R.Y.01	RY and Sph motifs conserved in seed-specific promoters	0. 8 18 18 1.0 0.9 0.9 0.9 7 25 51 - 00 39 0. 17 catgtttccgtgacc
SEQ P\$IDRE/IDE1.0 _11 1	Iron-deficiency-responsive element 1	0. 7 18 18 1.0 0.7 0.7 0.7 7 28 42 - 00 86 0. 17 atGCATgtttccgtg
SEQ P\$LEGB/R.Y.01	RY and Sph motifs conserved in seed-specific promoters	0. 8 18 18 1.0 0.9 0.9 0.9 7 28 54 + 00 39 0. 17 cactgttaacacgat
SEQ P\$OPAQ/O2_ _11 GCN4.01	Recognition site for BZIP transcription factors that belong to the group of Opaque-2 like proteins	0. 8 18 18 1.0 0.8 0.8 0.8 1 34 50 - 00 38 0. 17 gttgtaACATgcatgtt
SEQ P\$GTBX/GT3A _11 .01	Trihelix DNA-binding factor GT-3a	0. 8 18 18 1.0 0.8 0.8 0.8 3 37 53 + 00 43 0. 17 ga
SEQ P\$GTBX/GT3A _11 .01	Trihelix DNA-binding factor GT-3a	0. 8 18 18 0.7 0.8 0.8 0.8 3 46 62 + 50 52 0. 17 gat
SEQ P\$GBOX/GBF _11 1.01	bZIP protein G-Box binding factor 1	0. 9 19 19 1.0 0.9 0.9 0.9 4 14 34 - 00 61 0. 17 caacacc
SEQ P\$GBOX/HBP _11 1A.01	HBP-1a, suggested to be involved in the cell cycle-dependent expression	0. 8 19 19 1.0 0.8 0.8 0.8 8 15 35 + 00 99 0. 17 caagcat
SEQ P\$ABRE/ABRE	ABA response elements	0. 0. 19 19 + 1.0 0.8 0.8 0.8 17 tgttgacACGTag-

[illegible]

SEQ P\$MYBS/TAM _11 YB80.01	MYB protein from wheat	0. 8 19 20 1.0 0.9 ccgcATATactaaa- 3 95 11 + 00 11 cac
SEQ P\$LFYB/LFY.0 _11 1	Plant specific floral meristem identity gene LEAFY (LFY)	0. 9 20 20 0.8 0.9 3 04 16 - 85 50 tGCCAgtgttttag
SEQ P\$DOFF/DOF2 _11 .01	Dof2 - single zinc finger transcription factor	0. 9 20 20 1.0 0.9 ggagattaAAAGGc- 8 45 61 - 00 81 taat
SEQ P\$GTBX/SBF1 _11 .01	SBF-1	0. 8 20 20 1.0 0.8 gtggagaTTAA- 7 47 63 - 00 97 aagcta
SEQ P\$AHBPWUS. _11 01	Homeodomain protein WUSCHEL	0. 9 20 20 1.0 0.9 4 49 59 + 00 63 gctttTAATct
SEQ P\$MIIG/PALB _11 OXP.01	Putative cis-acting element in various PAL and 4CL gene promoters	0. 8 20 20 0.9 0.8 1 58 72 - 36 65 gcGTGGtggtgtggag
SEQ O\$MINI/MUSC _11 LE_INI.02	Muscle Initiator Sequence	0. 8 20 20 1.0 0.8 cacacaCCACg- 6 61 79 + 00 79 cagctata
SEQ P\$GBOX/GBF _11 1.01	bZIP protein G-Box binding factor 1	0. 9 20 20 1.0 0.9 tagaagacACGTgta- 4 72 92 - 00 76 tagctg
SEQ P\$GBOX/CPR _11 F.01	Common plant regulatory factor (CPRF) from parsley	0. 9 20 20 1.0 0.9 agctata- 5 73 93 + 00 85 cACGTgtcttctat
SEQ P\$MYCL/MYC _11 RS.01	Myc recognition sequences	0. 9 20 20 1.0 0.9 agaagacACGTgta- 3 73 91 - 00 65 tagct
SEQ P\$ABRE/ABRE _11 .01	ABA response elements	0. 8 20 20 1.0 0.8 2 74 90 + 00 84 gctatacACGTgtcttc
SEQ P\$CE3S/CE3.0	Coupling element 3 (CE3), non-ACGT ABRE	0. 0. 20 20 + 0.7 0.7 gctataCACGgtgtcttc-

SEQ_P\$MYCL/ICE.0_11_1	ICE (inducer of CBF expression 1), AtMYC2 (rd22BP1)	9 20 0.9 0.9 5 74 92 + 54 55 0.	gctatACACGtgtgtcttc ta
SEQ_P\$ABRE/ABF1_11_.01	ABA (abscisic acid) inducible transcriptional activator	7 20 20 1.0 0.8 9 75 91 - 00 24 0.	agaagACACgtgtgta- tag
SEQ_P\$DPBF/DPBF_11_.01	bZIP factors DPBF-1 and 2 (Dc3 promoter binding factor-1 and 2)	8 20 20 1.0 0.9 9 78 88 + 00 08 0.	tACACgtgtct
SEQ_P\$CGCG/OSC_11_BT.01	Oryza sativa CaM-binding transcription factor	7 20 20 0.9 0.8 8 79 95 + 06 04 0.	acaCGTgttcttatgc
SEQ_P\$GBOX/CPR_11_F.01	Common plant regulatory factor (CPRF) from parsley	9 20 21 1.0 0.9 5 91 11 - 00 68 0.	aacaagg- cACGTgttggcata
SEQ_P\$GBOX/CPR_11_F.01	Common plant regulatory factor (CPRF) from parsley	9 20 21 1.0 0.9 5 92 12 + 00 68 0.	atgccaa- cACGTgcctgttc
SEQ_P\$MYCL/OSB_11_HLH66.01	Rice bHLH protein	8 20 21 1.0 0.9 5 92 10 - 00 37 0.	acaagg- CACGtgttggcat
SEQ_P\$ABRE/ABF1_11_.01	ABA (abscisic acid) inducible transcriptional activator	7 20 21 1.0 0.8 9 93 09 + 00 52 0.	tgccaa- CACgtgccttg
SEQ_P\$CE3S/CE3.0_11_1	Coupling element 3 (CE3), non-ACGT ABRE	7 20 21 0.7 0.7 7 93 11 + 50 76 0.	tgccaa- CACGtgcctgtt
SEQ_P\$MYCL/OSB_11_HLH66.01	Rice bHLH protein	8 20 21 1.0 0.9 5 93 11 + 00 50 0.	tgccaa- CACGtgcctgtt
SEQ_P\$MIIG/PALB_11_OXL.01	Cis-acting element conserved in various PAL and 4CL promoters	8 13 27 - 85 06 0.	ccttggtcGGTTga

38

[illegible]

Q_1 2	X.01		3	9	50	29	TAaggtttta
SE							
Q_1 2	P\$CARM/CA RICH.01	CA-rich element	22	24	1.0	0.7	gtcctgaAACAA-
SE			0.78	2	0	+ 00	85 aataaac
Q_1 2	P\$TBPF/TAT A.02	Plant TATA box	23	25	1.0	0.9	aacaTATAta-
SE			0.90	8	2	+ 00	15 aactg
Q_1 2	P\$TBPF/TAT A.01	Plant TATA box	24	25	1.0	0.8	cataTATA-
SE			0.88	0	4	+ 00	92 aactgat
Q_1 2	P\$MYBL/GA MYB.01	GA-regulated myb gene from barley	27	28	1.0	0.9	attggtttGTTA-
SE			0.91	2	8	- 00	36 taagg
Q_1 2	P\$CAAT/CA AT.01	CCAAT-box in plant promoters	28	29	1.0	0.9	
SE			0.97	2	0	+ 00	75 aaCCAAtct
Q_1 2	P\$AHBP/AT HB9.01	HD-ZIP class III protein ATHB9	28	29	0.7	0.7	
SE			0.77	4	4	- 50	80 gtaAAGAttgg
Q_1 2	P\$DOFF/PB OX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	28	30	1.0	0.7	ataactgtAAA-
SE			0.75	4	0	- 00	77 Gattgg
Q_1 2	P\$MYBL/AT MYB77.01	R2R3-type myb-like transcription factor (l-type binding site)	28	30	1.0	0.8	tcittaCAGTta-
SE			0.87	8	4	+ 00	89 tagtg
Q_1 2	P\$NACF/TA NAC69.01	Wheat NACdomain DNA binding factor	30	32	0.8	0.7	gaggtataaca-
SE			0.68	6	8	- 12	49 GACGatagtata
Q_1 2	P\$MYBL/GA MYB.01	GA-regulated myb gene from barley	31	32	1.0	0.9	tatcgtctGTTA-
SE			0.91	1	7	+ 00	36 tacct
Q_1 01	P\$IDDF/ID1. 01	Maize INDETERMINATE1 zinc finger protein	32	33	1.0	0.9	ctctTTGTcggg
			0.92	6	8	+ 00	53 g

[illegible]

Q_1 2	AA.01			2	8	50	45	tatct
SE								
Q_1 2	P\$MADS/SQ UA.01	MADS-box protein SQUAMOSA		66 2	68 2	1.0 + 00	0.9 25	aatatctATTTttt- tacataa
SE								
Q_1 2	P\$LREM/AT CTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes		66 3	67 3	1.0 + 00	0.8 58	atATCTatttt
SE								
Q_1 2	P\$AHBP/AT HB5.01	HDZip class I protein ATHB5		67 9	68 9	0.9 + 36	0.9 39	ataATCAttgt
SE								
Q_1 2	P\$AHBP/HA HB4.01	Sunflower homeodomain leucine-zipper protein Hahb-4		67 9	68 9	1.0 - 00	0.9 66	acaatgAT- TAt
SE								
Q_1 2	P\$NCS2/NC S2.01	Nodulin consensus sequence 2		68 7	70 1	1.0 - 00	0.7 96	aatagaCTCT- taaca
SE								
Q_1 2	P\$GAPB/GA P.01	Cis-element in the GAPDH promoters conferring light inducibility		69 4	70 8	1.0 - 00	0.9 02	aaatATGAata- gact
SE								
Q_1 2	P\$IDDF/ID1. 01	Maize INDETERMINATE1 zinc finger protein		71 6	72 8	1.0 - 00	0.9 35	gaatTTGTcca- ca
SE								
Q_1 2	P\$AHBP/HA HB4.01	Sunflower homeodomain leucine-zipper protein Hahb-4		72 8	73 8	1.0 + 00	0.9 21	cgtattATTAc
SE								
Q_1 2	P\$SPF1/SP8 BF.01	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes		73 4	74 6	1.0 + 00	0.8 94	atTACTtttcttg
SE								
Q_1 2	P\$SEF4/SEF 4.01	Soybean embryo factor 4		74 5	75 5	1.0 + 00	0.9 82	tgTTTTTgtttt
SE								
Q_1 2	P\$NCS1/NC S1.01	Nodulin consensus sequence 1		76 4	77 4	0.8 - 78	0.8 65	aaaaTgatagt

SE	Q_1	P\$MYBL/GA	GA-regulated myb gene from barley	88	90	1.0	0.9	tttgttttGTTAact
SE	Q_1	P\$MYBL/GA	GA-regulated myb gene from barley	88	90	1.0	0.9	tttgttttGTTAact
SE	Q_1	P\$MYBS/MY	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	89	91	1.0	0.9	atagttATCCa-
SE	Q_1	P\$IBOX/GAT	Class I GATA factors	90	91	1.0	0.9	ttctgGATAAc-
SE	Q_1	P\$TBPFF/TAT	Plant TATA box	90	92	1.0	0.9	taacTATAAat-
SE	Q_1	P\$AHBP/AT	Arabidopsis thaliana homeo box protein 1	91	92	1.0	0.9	taaATTAttg
SE	Q_1	P\$AHBP/AT	Arabidopsis thaliana homeo box protein 1	91	92	0.7	0.9	caaATAAttta
SE	Q_1	P\$SPF1/SP8	DNA-binding protein of sweet potato that binds to the SP8a (ACTGTGTA) and SP8b (TACTATT) sequences of sporamin and beta-amylase genes	94	95	0.7	0.8	atAACTattgtga
SE	Q_1	P\$LREM/AT	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of pho-	95	96	1.0	0.9	atATCTa-
SE	Q_1	P\$PSRE/GA	GAAA motif involved in pollen specific transcriptional activation	95	96	0.7	0.8	ttataGATA-
SE	Q_1	P\$MYBL/GA	GA-regulated myb gene from barley	97	99	1.0	0.9	tttgttttGTTAact
SE	Q_1	P\$MYBS/MY	MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat	98	10	1.0	0.9	gtagttATCCa-
SE	Q_1	P\$IBOX/GAT	Class I GATA factors	98	10	1.0	0.9	ttctgGATAAc-

Q_1 2	A.01		9	05	00	35	tacaaa
SE							
Q_1 2	P\$MYBL/GA MYB.01	GA-regulated myb gene from barley	99	10	1.0	0.9	gatttgtaGT- TAtccag
SE			0.91	1	07	- 00	57
Q_1 2	P\$AHBP/AT HB1.01	Arabidopsis thaliana homeo box protein 1	10	10	0.7	0.9	caaATGAttg
SE			0.90	02	12	- 89	00
Q_1 2	P\$AHBP/AT HB5.01	HDZip class I protein ATHB5	10	10	0.9	0.9	caaATCAttg
SE			0.89	02	12	+ 36	36
Q_1 2	P\$NCS1/NC S1.01	Nodulin consensus sequence 1	10	10	0.8	0.9	cAAATgattg
SE			0.85	02	12	- 78	04
Q_1 2	P\$DOFF/PB OX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	10	10	0.7	0.7	tgatatgcAAAT- gattt
SE			0.75	03	19	- 76	57
Q_1 2	P\$STKM/ST K.01	Storekeeper (STK), plant specific DNA binding protein important for tuber-specific and sucrose-inducible gene expression	10	10	1.0	0.9	cccTAAAAa- aattgat
SE			0.85	16	30	- 00	17
Q_1 2	O\$RVUP/LT RUP.01	Upstream element of C-type Long Terminal Repeats	10	10	1.0	0.7	tacagcac- taaTTTCccta- aa
SE			0.76	24	44	- 00	79
Q_1 2	P\$DOFF/PB OX.01	Prolamin box, conserved in cereal seed storage protein gene promoters	10	10	0.7	0.7	agtgtgtA- AATTttca
SE			0.75	36	52	+ 76	81
Q_1 2	P\$GTBX/GT 1.01	GT1-Box binding factors with a trihelix DNA-binding domain	10	10	0.9	0.8	agtgtGT A- Aattttca
SE			0.85	36	52	+ 68	81
Q_1 2	P\$CCAF/CC A1.01	Circadian clock associated 1	10	10	1.0	0.9	caaaaaaa- AATCtat
SE			0.85	51	65	+ 00	73
Q_1 2	P\$LREM/AT CTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of pho- tosynthesis-related genes	10	10	1.0	0.8	aaATCTataga
SE			0.85	58	68	+ 00	97

[illegible]

SE	Q_1	P\$MADS/AG	AGL2, Arabidopsis MADS-domain protein AGAMOUS-like 2	12	12	0.8	0.8	tacatCCA-		
	2	L2.01		0.82	19	39	-	69	22	Gattttgatacat
SE	Q_1	P\$CCAF/CC	Circadian clock associated 1	12	12	1.0	0.8	tgtatcaa-		
	2	A1.01		0.85	20	34	+	00	99	AATCtgg
SE	Q_1	P\$OCSE/OC	OCS-like elements	12	12	0.8	0.7	ggatgtatggata-		
	2	SL.01		0.69	33	53	+	07	19	tACATatc
SE	Q_1	P\$MYBS/TA	MYB protein from wheat	12	12	1.0	0.8	atgtATATcca-		
	2	MYB80.01		0.83	34	50	-	00	96	tacatc
SE	Q_1	P\$MYBS/TA	MYB protein from wheat	12	12	1.0	0.9	atggATATaca-		
	2	MYB80.01		0.83	39	55	+	00	05	tatctt
SE	Q_1	P\$AHBP/AT	Arabidopsis thaliana homeo box protein 1	12	12	1.0	0.9			
	2	HB1.01		0.90	56	66	-	00	89	gtaATTAttat
SE	Q_1	P\$AHBP/HA	Sunflower homeodomain leucine-zipper protein Hahb-4	12	12	1.0	0.9	ataataAT-		
	2	HB4.01		0.87	56	66	+	00	57	TAc
SE	Q_1	P\$GTBX/GT	GT1-Box binding factors with a trihelix DNA-binding domain	12	12	0.9	0.9	actttgGTAAt-		
	2	1.01		0.85	56	72	-	68	23	tattat
SE	Q_1	P\$GTBX/GT	Trihelix DNA-binding factor GT-3a	12	12	1.0	0.8	aggtatGT-		
	2	3A.01		0.83	65	81	-	00	89	TActttggt
SE	Q_1	P\$MYBS/OS	Rice MYB proteins with single DNA binding domains, binding to the amylase element (TATCCA)	12	12	1.0	0.8	tttcgTATC-		
	2	MYBS.01		0.82	80	96	-	00	25	tatgtgag
SE	Q_1	P\$GARP/AR	Type-B response regulator (ARR10), member of the GARP-family of plant myb-related DNA binding motifs	12	12	1.0	0.9			
	2	R10.01		0.97	87	95	+	00	85	AGATacgaa
SE	Q_1	P\$NCS1/NC	Nodulin consensus sequence 1	0.85	12	13	+	1.0	0.8	gAAAAgcttac

Q_1 2	S1.01			93	03	00	93	
SE								
Q_1 2	P\$WBXF/MR KY.01	WRKY plant specific zinc-finger-type factor associated with pathogen defence, W box	0.92	12	13	1.0	0.9	aatttTTGActg-
SE				98	14	-	00	75 taagc
Q_1 2	P\$NCS2/NC S2.01	Nodulin consensus sequence 2	0.79	13	13	0.7	0.8	tagtgtATCTt-
SE				12	26	-	50	12 gaat
Q_1 2	P\$GBOX/TG A1.01	Arabidopsis leucine zipper protein TGA1	0.90	13	13	1.0	0.9	gagggcT-
SE				25	45	-	00	82 GACgttttaggta
Q_1 2	P\$GBOX/HB P1B.01	Wheat bZIP transcription factor HBP1B (histone gene binding protein 1b)	0.83	13	13	1.0	0.9	aACGT-
SE				26	46	+	00	38 cagccctct
Q_1 2	P\$GBOX/BZI P910.02	bZIP transcription factor from Antirrhinum majus	0.84	13	13	1.0	0.9	ttgtaagagggcT-
SE				31	51	-	00	52 GACgtttt
Q_1 2	P\$OCSE/OC SL.01	OCS-like elements	0.69	13	13	1.0	0.7	ttgtaagagggc-
SE				31	51	-	00	27 gACGTttt
Q_1 2	P\$LREM/AT CTA.01	Motif involved in carotenoid and tocopherol biosynthesis and in the expression of photosynthesis-related genes	0.85	13	13	1.0	0.8	ccATCTata-
SE				67	77	+	00	82 ta
Q_1 2	O\$LTUP/TA ACC.01	Lentiviral TATA upstream element	0.71	13	14	1.0	0.7	cttgactatca-
SE				82	04	+	00	69 gAACCCtcaaat
Q_1 2	P\$OCSE/OC SL.01	OCS-like elements	0.69	13	14	0.7	0.7	gaacctcaaat-
SE				93	13	+	69	01 taACTTctc
Q_1 2	P\$GTBX/SB F1.01	SBF-1	0.87	13	14	1.0	0.8	tgagaagT-
SE				98	14	-	00	83 TAAtttga
Q_1 2	P\$PSRE/GA AA.01	GAAA motif involved in pollen specific transcriptional activation	0.83	14	14	1.0	0.8	tcigaGA-
SE				11	27	-	00	96 AAtgtttgag

2	SE	P\$GCCF/ER	Ethylene-responsive elements (ERE) and jasmonate- and elicitor-responsive elements	14	14	1.0	0.9	aagagtCGC-
2	Q_1	E_JERE.01	(JERE)	0.85	50	62	+ 00	24 Caag

All references cited in this specification are herewith incorporated by reference with respect to their entire disclosure content and the disclosure content specifically mentioned in this specification.

5

The Figures show:

Fig. 1: Gas chromatogram of a transgenic line transformed with binary vector pSUN-BN3.

10

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

15

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).

20

Example 2: Cloning of Promotor Elements from Brassica napus

25

For the analysis of potentially seeds specific expressed genes in Brassica napus, three different seed stages have been investigated. To this end, Brassica napus cv. Westar plants were raised under standard conditions (Moloney et al. 1992, Plant Cell Reports 8: 238-242). The seeds have been harvested 20 days, 25 days and 40 days after flowering. The seeds were used for the preparation of RNA (RNAeasy, Qiagen) according to the manufactures manual. In parallel, plant material from roots, leaves and stipes has been used for preparation of RNA. The said RNA was mixed and used as a control for the further experiments. RNA from the seed stages as well as control RNA were treated by the one-color gene expression kit (Agilent) for microarray-analysis. The Arapidopsis whole genome chip (Agilent) was hybridized with the treated RNA. Based on different labelled RNAs, the genes from Brassica napus could be identified which are expressed in the seeds solely but not in other organs or tissues. Six genes from Arabidopsis thaliana have been identified which hybridized with the probes from Brassica napus (Table 7).

30

35

40

Table 7: Arabidopsis genes which were capable of hybridizing the seeds specific probes from *Brassica napus*:

Arabidopsis sequence	Protein function	Expression pattern
At1g23200	pectinesterase	seed
At1g52690	LEA gene	seed
At1g61720	anthocyanidin reductase	seed
At2g38900	Proteinase inhibitor	seed
At3g15670	LEA gene	seed
At5g38170	Lipid transfer protein	seed

- 5 Based on the gene sequences from *Arabidopsis thaliana*, homologs have been identified in *Brassica napus* cDNA libraries. For all six *Arabidopsis* genes, homologous coding sequences in *Brassica napus* could be identified (Table 8).

- 10 Table 8: Homology sequence from *Brassica napus* corresponding to the *Arabidopsis* sequence shown in Table 7.

Brassica napus sequence	SEQ ID NO:	Arabidopsis homolog	Protein function
BN1	1	At1g23200	pectinesterase
BN2	2	At1g52690	LEA gene
BN3	3	At1g61720	anthocyanidin reductase
BN4	4	At2g38900	Proteinase inhibitor
BN6	5	At3g15670	LEA gene
BN8	6	At5g38170	Lipid transfer protein

- 15 From leaf material of *Brassica napus* cv. Westar, genomic DNA has been isolated using the DNAeasy kit (Qiagen) according to the manufacturer's manual. Culture conditions for the *Brassica napus* cv. Westar were as discussed above. Based on the genomic DNA, a genomic DNA library was established using the Genome Walker kit (Clontech). The following primer sequences were derived from *Brassica napus* cDNA

sequences in order to isolate upstream sequences of the Brassica napus genomic sequences (Table 9).

Table 9: Primer sequences for the implication of 5 prime upstream sequences in combination with AP1/ AP2 (Clontech).

Brassica napus sequence	Primer sequence 5'-3'	SEQ ID NO:
BN1	ATTGGTATAATATATTTGG	14
BN2	GTTTCTGTGTAGAGAACTG	15
BN3	CTGATTAAATTCTTAAGACCAG	16
BN4	CCAAAATTACCAGCACATTC	17
BN6	GTTGCTGTGTATAAACTGTG	18
BN8	TCTGAGAAATGTTTGAGAAG	19

The indicated primers were used in combination with the AP1 and AP2 primers according to the manufacturer's manual of the Genome Walker kit in a PCR. PCR conditions were as follows:

Primary PCR:

7 cycles 94 C°, 25 seconds, 72 C°, 4 minutes,
32 cycles 94 C°, 25 seconds 67 C° 4 minutes,
final cycle 67 C°, 4 minutes.

Secondary PCR

5 cycles 94 C°, 25 seconds, 72 C° 4 minutes,
22 cycles 94 C°, 25 seconds, 67 C°, 4 minutes,
final cycle 67C°, 4 minutes.

Using the former specific primers, specific fragments for the primer combinations were obtained. These fragments were cloned into the pGEM-T (Pomega) vector using the manufacturer's manual and sequenced 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). The following Brassica napus sequences have been obtained (Table 10).

Table 10: Genomic five prime upstream sequences from the Brassica napus cDNA sequences.

Brassica napus Sequence	genomic 5' sequence in bp	SEQ ID NO:
BN1	1336	7
BN2	1532	8
BN3	1612	9
BN4	1767	10
BN6	2281	11
BN8	1490	12

5

The analysis of the 5' upstream sequences using Genomatix software GemsLauncher 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.

10

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 a CaMV35S terminator was linked with promoter elements. In a further step, a multiple cloning site was introduced in between the promoter and terminator elements. The primers used are shown in Table 11.

20

Table 11: Primer pairs used for the generation of promoter-terminated-cassettes via Fusion-PCR.

Brassica napus Promoter/Terminator cassette	Primer pair 1. PCR Promoter	Primer pair 1. PCR Terminator	Primer pair 2. PCR
p-BN1_t-35S	5'- acctgcaggttaggccggc- cactgtcatatatatatgac (SEQ	5'- ccatggacttaggccttagcttaat- taactaagtcgacaagctc-	5'- acctgcaggttaggccggc- cactgtcatatatatatgac (SEQ

	ID NO: 20) 3'- tcacaaacctcccgatgtttataca- cacatggacttaggccttagct- taattaactaagtcgacaagctc- gag (SEQ ID NO: 21)	gagtttcccataataatg (SEQ ID NO: 22) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO: 23)	ID NO:24) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:25)
p-BN2_t-35S	5'- acctgcaggtaggcccggcca- taaccctctccatgttgatac (SEQ ID NO:26) 3'- ccttgaagaaaagaaacatg- gacttaggccttagcttaataac- taagtcgacaagctcgag (SEQ ID NO:27)	5'- ccatggacttaggccttagcttaat- taactaagtcgacaagctc- gagtttcccataataatg (SEQ ID NO:28) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:29)	5'- acctgcaggtaggcccggcca- taaccctctccatgttgatac (SEQ ID NO:30) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:31)
p-BN3_t-35S	5'- acctgcaggtaggcccggccacta- tagggcacgcgtggtcg (SEQ ID NO:32) 3'- gcgttaagaattataatatacagc- catggacttaggccttagcttaat- taactaagtcgacaagctcgag (SEQ ID NO:33)	5'- ccatggacttaggccttagcttaat- taactaagtcgacaagctc- gagtttcccataataatg (SEQ ID NO:34) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:35)	5'- acctgcaggtaggcccggccacta- tagggcacgcgtggtcg (SEQ ID NO:36) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:37)
p-BN4_t-35S	5'- acctgcaggtaggcccggccgtt- gatggaaatcgatcgctg (SEQ ID NO:38) 3'- ctgcaaagataaaaaaaagggg- tagcaacccatggacttaggcct- tagcttaattaactaagtcga- caagctcgag (SEQ ID NO:39)	5'- ccatggacttaggccttagcttaat- taactaagtcgacaagctc- gagtttcccataataatg (SEQ ID NO:40) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:41)	5'- acctgcaggtaggcccggccgtt- gatggaaatcgatcgctg (SEQ ID NO:42) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:43)
p-BN6_t-35S	5'- acctgcaggtaggcccggccttg- tactctccctaatggag (SEQ ID NO:44) 3'- cctatttgcgcatttgaagaaagaa- aaccatggacttaggccttagct- taattaactaagtcgacaagctc- gag (SEQ ID NO:45)	5'- ccatggacttaggccttagcttaat- taactaagtcgacaagctc- gagtttcccataataatg (SEQ ID NO:46) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:47)	5'- acctgcaggtaggcccggccttg- tactctccctaatggag (SEQ ID NO:48) 3'gaattaattcggcggttaattcaggg cgcc (SEQ ID NO:49)
p-BN8_t-35S	5'- acctgcaggtaggcccggccgt- gaatcacaagcaaagag (SEQ	5'- ccatggacttaggccttagcttaat- taactaagtcgacaagctc-	5'- acctgcaggtaggcccggccgt-

	ID NO:50) 3'- gagtcgccaagcttacaacc- catggacttaggccttagctaat- taactaagtcgacaagctcgag (SEQ ID NO:51)	gagtttccataataatg (SEQ ID NO:52) 3'gaattaattcggcgtaattcaggg cgcc (SEQ ID NO: 53)	ID NO:54) 3'gaattaattcggcgtaattcaggg cgcc (SEQ ID NO:55)
--	---	---	--

The promoter-terminator cassettes were cloned into the pGEMT (Promega vector) according to the manufacturer's manual and subsequently sequenced. Via the restriction site of *Sbf1-EcoRV* (New England Biolabs), cassettes were transferred into the vector pENTRB (Invitrogen) according to standard techniques. In a further step, the delta 6 Desaturase Gene (SEQ ID NO: 13) was introduced via the *Nco1-Pac1* restriction sites into the generated pENTRB vectors pENTRB-p-BN1_t-35S, pENTRB-p-BN2_t-35S, pENTRB-p-BN3_t-35S, pENTRB-p-BN4_t-35S, pENTRB-p-BN6_t-35S, pENTRB-p-BN8_t-35S.

The resulting vectors were subsequently used for Gateway (Invitrogen) reactions together with the binary plasmid pSUN to generate binary vectors 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.

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)

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 (Var. Drakkar, NPZ Norddeutsche Pflanzenzucht, Hohenlieth, Deutschland) 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. Petioles or Hypocotyledones of 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), 50mg/l Kanamycine, 20 mikrom Benzylaminopurin (BAP) and 1,6g/l 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.

5 Regenerated sprouts have been obtained on 2MS-Medium with Kanamycine 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.

10 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. (1994), Plant Cell Report 13: 282-285.

15

c) Production of transgenic Arabidopsis plants

20 Transgenic Arabidopsis plants were generated according to the protocol of Bechthold et al. 1993 (Bechthold, N., Ellis, J., Pelletier, G. (1993) *In planta Agrobacterium*-mediated gene transfer by infiltration of *Arabidopsis thaliana* plants. *C.R. Acad. Sci. Ser. III Sci. Vie.*, **316**, 1194-1199). Arabidopsis plants of the ecotype Col0fae1 were grown on soil after a vernalisation of the seeds for 3 days at 4°C. After plants started to flower, they were dipped into an *Agrobacterium tumefaciens* solution containing *Agrobacterium* strain pMP90 transformed with the binary plasmids as described in Example 3 and following other components: ½ MS pH 5,7, 5 % (w/v) Sacharose, 4,4 µM Benzylaminopurin, 0,03 % Silwet L-77 (Lehle Seeds, Round Rock, TX, USA). *Agrobacterium* solution was diluted to a final concentration of OD₅₄₆ 0,8. Plants were dipped two times into above described solution and kept 4-6 weeks for normal growth and seed formation. Dried seeds were harvested and subjected to selective growth based on the tolerance against the herbicide Pursuit (BASF). Seeds of this generation of selected plants were then subjected to lipid analysis.

30

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: 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,

40

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.

- 5 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.
- 10

- Based on the analysed lipids, the expression of the Desaturase were determined since the lipid pattern of successfully transformed plant seeds are differing from the pattern of control plant seeds.
- 15

Analysis of promoter BN3:

- 20 Seeds from different *Arabidopsis* plants containing the T-DNA of binary vector pSUN-BN3 (see Example 3) were subjected to lipid analysis as described above (Tab. 12 and Fig. 1). Compared to the non-transgenic control plants (WT), plants containing pSUN-BN3 produced in addition to the fatty acids found in the control plants a novel fatty acid, γ -linolenic acid (18:3 Δ 6,9,12). The synthesis of this novel fatty acid is subject to the enzyme Δ 6-desaturase, which gene is behind the BN3 promoter. The fact that the novel fatty acid can be detected in significant amounts in the seeds of *Arabidopsis* plants containing the T-DNA of binary vector pSUN-BN3 is explained by the functional expression of the gene Δ 6-desaturase from *Pythium irregulare*. According to the identification of the novel generated fatty acid γ -linolenic acid, the promoter BN3 is enabling the functional expression of the respective gene.
- 25
- 30

Therefore the promoter BN3 is a functional promoter, driving expression of genes in seeds.

- Other parts of the transgenic *Arabidopsis* plants were also subjected to gas chromatographic analysis, but no other fatty acids than in the non-transgenic *Arabidopsis* plants were observed.
- 35

Therefore, the promoter BN3 is driving functional expression in a seed-specific manner, thereby only allowing the transcription of attached genes in seeds.

Table 12: Gas chromatographic analysis of seeds from different *Arabidopsis* plants either being non-transgenic controls (WT) or transgenic lines containing the T-DNA of plasmid pSUN-BN3 derived from chromatograms as shown in Fig. 1. The respective fatty acids are given in chemical nomenclature (16:0 palmitic acid, 18:0 stearic acid, 18:1-n-9 oleic acid, 18:2n-6 linoleic acid, 18:3n-6 γ -linolenic acid, 18:3n-3 α -linolenic acid). The fatty acid 18:3n-6 is a product of the enzymatic reaction of the $\Delta 6$ -desaturase from *Pythium irregulare*, which is not observed in the non-transgenic control lines.

sample name	16:0	18:0	18:1n-9	18:2n-6	18:3n-6	18:3n-3
WT						
WT	13,03	3,68	27,65	35,37	0,00	20,27
WT	14,04	3,51	25,16	42,81	0,00	14,49
WT	9,63	2,66	36,85	35,07	0,00	15,80
WT	10,16	2,80	37,84	35,35	0,00	13,86
WT	8,94	3,02	31,66	36,93	0,00	19,44
WT	9,52	2,93	29,74	37,15	0,00	20,66
pSUN-BN3_1	13,19	2,82	35,20	25,31	12,20	11,28
pSUN-BN3_2	14,23	5,25	44,38	14,58	14,57	7,00
pSUN-BN3_3	13,27	3,45	46,13	18,62	9,86	8,67
pSUN-BN3_4	11,76	2,08	43,63	29,17	3,97	9,39
pSUN-BN3_5	12,26	1,76	42,01	25,92	8,12	9,94
pSUN-BN3_6	10,25	3,22	35,82	26,64	10,43	13,65
pSUN-BN3_7	10,65	3,78	34,04	23,74	14,62	13,17

Claims

1. 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:
 - (a) an expression control sequence having a nucleic acid sequence as shown in any one of SEQ ID NOs: 7 to 12;
 - (b) an expression control sequence having a nucleic acid sequence which hybridizes under stringent conditions to a nucleic acid sequence as shown in any one of SEQ ID NOs: 7 to 12
 - (c) 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: 1 to 6;
 - (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 being at least 80% identical to an open reading frame sequence as shown in any one of SEQ ID NOs: 1 to 6;
 - (e) an expression control sequence obtainable by 5' genome walking from an open reading frame sequence as shown in any one of SEQ ID NOs: 1 to 6; and
 - (f) an expression control sequence obtainable by 5' genome walking from an open reading frame sequence being at least 80% identical to an open reading frame as shown in any one of SEQ ID NOs: 1 to 6.
2. The polynucleotide of claim 1, wherein said expression control sequence comprises at least 1,000 nucleotides.
3. The polynucleotide of claim 1 or 2, wherein said polynucleotide further comprises a nucleic acid of interest being operatively linked to the expression control sequence.
4. The polynucleotide of any one of claims 1 to 3, wherein said polynucleotide further comprises a termination sequence which allows for termination of the transcription of a nucleic acid of interest.
5. A vector comprising the polynucleotide of any one of claims 1 to 4.
6. The vector of claim 5, wherein said vector is an expression vector.
7. A host cell comprising the polynucleotide of any one of claims 1 to 4 or the vector of claim 5 or 6.

8. The host cell of claim 7, wherein said host cell is a plant cell.
9. A non-human transgenic organism comprising the polynucleotide of any one of
5 claims 1 to 4 or the vector of claim 5 or 6.
10. The non-human transgenic organism of claim 9, wherein said organism is a plant
or a seed thereof.
- 10 11. A method for expressing a nucleic acid of interest in a host cell comprising
 - (a) introducing the polynucleotide of any one of claims 1 to 4 or the vector of
claim 5 or 6 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.
15
12. The method of claim 11, wherein said host cell is a plant cell.
13. A method for expressing a nucleic acid of interest in a non-human organism
comprising
20
 - (a) introducing the polynucleotide of any one of claims 1 to 4 or the vector of
claim 5 or 6 into the non-human organism, whereby the nucleic acid se-
quence of interest will be operatively linked to the expression control se-
quence ; and
 - (b) expressing the said nucleic acid sequence in said non-human transgenic
25 organism.
14. The method of claim 13, wherein said non-human transgenic organism is a plant
or seed thereof.
- 30 15. The method of claim 13 or 14, wherein said nucleic acid of interest is expressed
seed-specific.
16. Use of the polynucleotide of any one of claims 1 to 4, the vector of claim 5 or 6,
the host cell of claim 7 or 8 or the non-human transgenic organism of claim 9 or
35 10 for the expression of a nucleic acid of interest.
17. The use of claim 16, wherein said expression is seed specific.
- 40 18. The polynucleotide of any one of claims 1 to 4, the vector of claim 5 or 6, the host
cell of claim 7 or 8, the non-human transgenic organism of claim 9 or 10, the
method of any one of claims 11 to 15 or the use of claim 16 or 17, wherein said

nucleic acid of interest encodes a seed storage protein or is involved in the modulation of seed storage compounds.

Fig.1

