(72) Inventors; and
(75) Inventors/Applicants (for US only): BADE, Jacob, Bernardus [NL/NL]; Buys Ballotstraat 60, NL-2313 TZ Leiden (NL). CUSTERS, Jerôme, Hubertina, Henricus, Victor [NL/NL]; Zaalbergstraat 24, NL-2405 XB Alphen a/d Rijn (NL).

(74) Agent: PRINS, A.W.; c/o Vereenigde, Nieuwe Parklaan 97, NL-2587 BN The Hague (NL).

(54) Title: NEW CONSTITUTIVE PLANT PROMOTER

(57) Abstract: The present invention provides a constitutive promoter obtainable from *Brassica napus* plants. According to the present invention there is provided aDNA fragment harbouring a constitutive promoter, said DNA fragment being present in clone pJB1178-29, deposited with the Centraal Bureau of Schimmelcultures (Baarn, the Nederlands) on 6 February 2001 under no. CBS 109272. The DNA fragment according to the present invention is further characterised in that it comprises whole or parts of the sequence represented by SEQ ID NO: 7. Further, the invention comprises the homologue sequence *in Arabidopsis thaliana* as depicted in SEQ ID NO:8.

NEW CONSTITUTIVE PLANT PROMOTER

Field of the Invention

5   The present invention relates to new plant promoters and more specifically to promoters that act constitutively.

Background

In the context of this disclosure, the term 'promoter' or 'promoter region' refers

10   to a sequence of DNA, usually upstream (5') to the coding sequence of a structural gene, which controls the expression of the coding region by providing the recognition for RNA polymerase and/or other factors required for transcription to start at the correct site.

There are generally two types of promoters, inducible and constitutive

15   promoters. An inducible promoter is a promoter that is capable of directly or indirectly activating transcription of one or more DNA sequences or genes in response to an inducer. In the absence of an inducer the DNA sequences or genes will not be transcribed. Typically, the protein factor, which binds specifically to an inducible promoter to activate transcription, is present in an inactivated form, which is then

20   directly or indirectly converted to the active form by the inducer. The inducer can be a chemical agent; a physiological stress caused by environmental conditions, or can be an endogenously generated compound in response to changes in the development of the plant.

Constitutive promoters direct the expression of the DNA sequence (gene),

25   which they control, throughout the various parts of the plants and continuously throughout plant development. However, the term 'constitutive' as used herein does not necessarily indicate that a gene is expressed at the same level in all cell types, but that the gene is expressed in a wide range of cell types, although some variation in abundance is often observed.

30   One of the earliest and most important inventions in the field of plant protein expression is the use of (plant) viral and *Agrobacterium*-derived promoters that

provide a powerful and constitutive expression of heterologous genes in transgenic plants. Several of these promoters have been used very intensively in plant genetic research and are still the promoter of choice for rapid, simple and low-risk expression studies. The most famous are the 35S and 19S promoter from Cauliflower Mosaic Virus (CaMV), which was already found to be practical useful in 1984 (EP 0 131 623), and the promoters which can be found in the T-DNA of *Agrobacterium*, like the nopaline synthase (nos), mannopine synthase (mas) and octopine synthase (ocs) promoters (EP 0 122 791, EP 0 126 546, EP 0 145 338). A plant-derived promoter with similar characteristics is the ubiquitin promoter (EP 0 342 926).

However, all promoters described above, although they are commonly known as constitutive promoters, still show patterns of organ- or developmental-specific expression, and frequently the pattern of expression found with these promoters is not ideal for some applications. Also, it has been found that duplication of promoters to drive expression of two different genes can cause problems because of DNA-dependent silencing. This risk especially appears in gene-stacking approaches in which several genes need to be expressed simultaneously. Further, the virus or *Agrobacterium* derived promoters are less attractive from a regulatory point of view.

Therefore, still the need exist for new constitutively active promoters from plant origin.

It is therefore an aim of the present invention to provide a new plant-derived constitutive promoter.

It is a further aim of the present invention to provide fragments of DNA comprising the promoter according to the present invention.

It is a further aim of the present invention to provide transgenic plants (or parts and/or seeds thereof) comprising the promoter according to the present invention, including plants (or parts of plants) and seeds derived from said transgenic plants.

Summary of the Invention

The present invention provides constitutive promoters obtainable from *Brassica napus* plants.

5   According to the present invention there is provided a DNA fragment harbouring a constitutive promoter, said DNA fragment being present in the EcoRI fragment present in clone pJB1178-29 deposited with the Centraal Bureau of Schimmelcultures (Baarn, the Netherlands) on 6 February 2001 under no CBS 109272.

The DNA fragment according to the present invention is further characterised in

10   that it comprises the nucleotide sequence represented by nucleotide 1 to 641 of SEQ ID NO: 1. Further, the DNA fragment is characterised in that it comprises the nucleotide sequence represented by SEQ ID NO:7

Also part of the invention is a DNA fragment, characterised in that it comprises the nucleotide sequence represented by SEQ ID NO:8 or parts thereof, wherein said

15   fragment or parts thereof are capable of promoting constitutive expression of a DNA sequence on reintroduction into a plant.

In more detail the invention provides for a DNA fragment capable of promoting constitutive expression of a DNA sequence on reintroduction into a plant, characterised in that it comprises a nucleotide sequence represented in SEQ ID NO:7

20   starting at the nucleotide selected from the group consisting of: nucleotide 4, nucleotide 34, nucleotide 341, nucleotide 588, nucleotide 650, nucleotide 782, nucleotide 825, nucleotide 937, nucleotide 1246, nucleotide 1556, nucleotide 1780, nucleotide 1849, nucleotide 1912, nucleotide 1960, nucleotide 2044, nucleotide 2243, nucleotide 2408 and nucleotide 2638; and ending at the nucleotide selected from the

25   group consisting of: nucleotide 341, nucleotide 588, nucleotide 650, nucleotide 782, nucleotide 825, nucleotide 937, nucleotide 1246, nucleotide 1556, nucleotide 1780, nucleotide 1849, nucleotide 1912, nucleotide 1960, nucleotide 2044, nucleotide 2243, nucleotide 2408, nucleotide 2638 and nucleotide 2654.

30   The present invention further includes a chimeric DNA sequence comprising, in the direction of transcription, at least one DNA fragment as hereinbefore described and at least one DNA sequence to be expressed under the transcriptional control of

said DNA fragment, wherein the DNA sequence to be expressed is not naturally under the transcriptional control of the DNA fragment.

The present invention further provides replicons comprising the abovementioned chimeric DNA sequences.

Also included in the present invention are microorganisms containing such a replicon, specifically pJB1178-29, plant cells having incorporated into their genome, a chimeric DNA sequence as described above and plants essentially consisting of said cells. Such a plant may be a dicotyledonous plant or a monocotyledonous plant. Also parts of said plants selected from seeds, flowers, tubers, roots, leaves, fruits, pollen and wood, form part of the invention.

According to a further aspect of the present invention, there is provided use of a chimeric DNA sequence in the transformation of plants and use of a portion or variant of the DNA fragments according to the invention for making hybrid regulatory DNA sequences.

Description of the Figures

The present invention will now be described with reference to the following Figures, which are by way of example.

Figure 1. Schematic overview of T-DNA structures in the constructs used for promoter tagging in *Brassica napus*. All binary vectors contain a 35S-hpt-nos cassette as in pMOG22 (Goddijn *et al.*, 1993). Construct pMOG448 was used as selection control (+hygromycin, -kanamycin). Construct pMOG964 contains a gus::nptII fusion gene (Datla *et al.*, 1991) combined with a double enhanced 35S promoter, whereas tagging construct pMOG1178 has a promoterless version of the same coding region. The gus::nptII gene contains an intron in the gus part as described by Vancanneyt *et al.* (1990). Spectinomycin resistance and ColE1 origin of replication are included as plasmid rescue features. The ampicillin gene is disrupted (Δ) to avoid resistance of *Agrobacterium* to carbenicillin. Restriction sites used for the Southern blot analysis

and plasmid rescue experiments are mapped (HindIII, EcoRI and BamHI). The waved line represents genomic plant DNA adjacent to the right border. LB left border, RB right border, p promoter, t transcriptional terminator sequence, hpt hygromycin phosphotransferase, als acetolactate synthase gene, gus-I GUS reporter gene plus intron, nptII neomycin phosphotransferase gene spec 3 kb fragment including spectinomycin resistance gene, Δamp ampicillin resistance gene (non-functional), ColE1 ColE1 origin of replication.

Figure 2. GUS activity in leaf and root parts of transgenic *Brassica napus* line 1178-29

Figure 3. Restriction analysis of rescued plasmids. Eleven different fragments of genomic sequence upstream of the gus::nptII tagging region were isolated via plasmid rescue (see Fig 1). DNA was isolated from the bacterial cultures, digested with EcoRI or EcoRI+BamHI and separated over an agarose gel (sets 1-11). Positions of 1 kb markers (Gibco-BRL) are indicated. The genomic fragments isolated from the three single copy lines (1178-21, 1178-29 and 1178-43) were used for sequence analysis, construction of binary vectors and re-transformation to wildtype *Brassica napus*.

Figure 4. Comparison of nucleotide sequences upstream of the gus::nptII coding region in tagging construct pMOG1178 and three transgenic lines (1178-21, 1178-29 and 1178-43). T-DNA right border, restriction sites (HindIII and BamHI) and start codon (ATG) of gus::nptII are underlined. Approximately 600 base pairs of sequence was determined for each line (single strand) and analysed via BLASTN searching. Homology was found with three BAC clones of *Arabidopsis* and cDNA clones of *Arabidopsis* and a *B. napus*. Start of homologous sequence is indicated (dashed line).

Figure 5. GUS activity in young callus driven by new genomic sequences with promoter activity. Fragments upstream of the gus::nptII tagging gene (pMOG1178, Fig

1) integrated in the genome of *Brassica napus* were isolated via plasmid rescue, cloned in binary vector pMOG22 (35S-hpt-nos, Fig 1) and re-transformed to *Brassica napus* hypocotyl explants (Table 3). Histochemical XGluc staining (24 hours) was performed after 3 weeks of culture on hygromycin containing medium. A: very high expression of 35S-gus::nptII control construct pMOG964 (Fig 1); B: good expression (pJBBIN1178-21); C: good expression (pJBBIN1178-29); D: moderate expression (pJBBIN1178-43).

Figure 6. Outline of the promoter and the sequence boxes containing promoter-elements. For explanation of the boxes, see example 7.

Detailed Description of the Invention

The present invention primarily concerns promoters or regulatory sequences naturally occurring in *Brassica napus* (oil seed rape). It has been found that genes under the regulatory control of these promoter or regulatory sequences are expressed in many tissues of the plant at different developmental stages, indicating constitutive expression. Specifically the promoter of the invention is the promoter driving the gus::nptII gene in the construct pJB1178-29, deposited with the Centraal Bureau of Schimmelcultures (Baarn, the Netherlands) on 6 February 2001 under no. CBS 109272.

At first, approximately 600 base pairs of sequence was determined (single strand) (nucleotides 1-641 of SEQ ID NO:1) for pJB1178-29 and analysed via BLASTN searching. This revealed significant homology for with the *Arabidopsis* clone BACF9D24 (3e–39) and the *Arabidopsis* cDNA clone 701549985 (3e-28). This cDNA is indicated as phenyl alanine tRNA synthetase protein like. Further, the complete sequence of the insert was determined (SEQ ID NO:7).

It is emphasized that the nucleotide sequence of the promoter of the invention may be subject to variations without significantly affecting the functionality, i.e. the specificity of the promoter. One of the possibilities to change the promoter is to delete certain fragments of the promoter while maintaining the elements that are necessary

for the specificity. This can be accomplished by making several deletion mutants of the promoter, linking them up in a construct with a reporter gene (e.g. the gus gene or a gene coding for a fluorescent protein, like the GFP gene of Aequoria) and subsequently performing expression studies on plants transformed with said construct.

Accordingly, also fragments of the promoter sequences of the construct pJB1178-29 driving predominantly constitutive expression, are part of this invention.

Furthermore, also promoter sequences formed by small changes in the nucleotide sequence by substitution or addition of nucleotides of the promoter sequence of the construct pJB1178-29 are included in this invention. It is envisaged that also in other species of plants homologous sequences can be found which have the same functionality as the sequence of the invention.

When comparing nucleic acid sequences for the purposes of determining the degree of homology or identity one can use programs such as BESTFIT and GAP (both from the Wisconsin Genetics Computer Group (GCG) software package) BESTFIT, for example, compares two sequences and produces an optimal alignment of the most similar segments. GAP enables sequences to be aligned along their whole length and finds the optimal alignment by inserting spaces in either sequence as appropriate. Suitably, in the context of the present invention when discussing homology of nucleic acid sequences, the comparison is made by alignment of the sequences along their whole length.

Preferably, sequences which have substantial homology have at least 50% sequence homology, desirably at least 70% sequence homology and more desirably at least 80%, 90% or at least 95% sequence homology, in increasing order of preference, with said sequences. In some cases the sequence homology may be 99% or above.

Desirably, the term "substantial identity" indicates that said sequence has a greater degree of identity with any of the sequences described herein than with prior art nucleic acid sequences.

The terms "regulatory sequence" or "regulatory region" and "promoter" are used interchangeably herein.

The present invention further provides chimeric DNA sequences comprising the DNA fragments of the present invention. The expression chimeric DNA sequence, as used herein, shall encompass any DNA sequence comprising DNA sequences not naturally found. For instance, chimeric DNA, as used herein, shall encompass DNA

5   comprising the regulatory region which is inducible in a non-natural location of the plant genome, notwithstanding the fact that said plant genome normally contains a copy of said regulatory region in its natural chromosomal location. Similarly, said regulatory region may be incorporated into a part of the plant genome where it is not naturally found, or in a replicon or vector where it is not naturally found, such as a

10   bacterial plasmid or a viral vector. The term "chimeric DNA", as used herein, shall not be limited to DNA molecules which are replicable in a host, but shall also encompass DNA capable of being ligated into a replicon, for instance by virtue of specific adaptor sequences, physically linked to the regulatory region according to the invention. The regulatory region may or may not be linked to its natural downstream

15   open reading frame.

The open reading frame of the gene whose expression is driven by the regulatory regions of the invention may be derived from a genomic library. In this situation, it may contain one or more introns separating the exons making up the open reading frame that encodes a protein according to the invention. The open reading

20   frame may also be encoded by one uninterrupted exon, or by a cDNA to the mRNA encoding a protein according to the invention. Chimeric DNA sequences according to the invention also comprise those in which one or more introns have been artificially removed or added. Each of these variants is embraced by the present invention.

In order to be capable of being expressed in a host cell, a regulatory region

25   according to the invention will usually be provided with a transcriptional initiation region, which may be suitably derived from any gene capable of being expressed in the host cell of choice, as well as a translational initiation region for ribosome recognition and attachment. In eukaryotic cells, an expression cassette usually also comprises a transcriptional termination region located downstream of said open reading frame,

30   allowing transcription to terminate and polyadenylation of the primary transcript to occur. Also, it is often the case that a signal sequence may be encoded, which is responsible for the targeting of the gene expression product to subcellular

compartments. The principles governing the expression of a chimeric DNA construct, in a chosen host cell, are commonly understood by those of ordinary skill in the art. Furthermore, the construction of expressible chimeric DNA constructs is now routine for any sort of host cell, be it prokaryotic or eukaryotic.

5   In order for the chimeric DNA sequence to be maintained in a host cell, it will usually be provided in the form of a replicon comprising said chimeric DNA sequence (according to the invention) linked to DNA, which is recognised and replicated by the chosen host cell. Accordingly, the selection of the replicon is determined largely by the host cell of choice. Such principles as govern the selection of suitable replicons

10   for a particular chosen host are well within the realm of the ordinary person skilled in the art.

       A special type of replicon is one capable of transferring itself, or a part thereof, to another host cell, such as a plant cell, thereby co-transferring the open reading frame to the plant cell. Replicons with such capability are herein referred to as

15   vectors. An example of such vector is a Ti-plasmid vector which, when present in a suitable host, such as *Agrobacterium tumefaciens*, is capable of transferring part of itself, the so-called T-region, to a plant cell. Different types of Ti-plasmid vectors (*vide:* EP 0 116 718 B1) are now routinely being used to transfer chimeric DNA sequences into plant cells, or protoplasts, from which new plants may be generated

20   which stably incorporate said chimeric DNA in their genomes. Particularly preferred forms of Ti-plasmid vectors are the so-called binary vectors as claimed in (EP 0 120 516 B1 and US 4,940,838). Other suitable vectors, which may be used to introduce DNA according to the invention into a plant host, may be selected from the viral vectors, for example, non-integrative plant viral vectors, such as derivable from the

25   double stranded plant viruses (for example, CaMV) and single stranded viruses, gemini viruses and the like. The use of such vectors may be advantageous, particularly when it is difficult to stably transform the plant host. Such may be the case with woody species, especially trees and vines.

       The expression "host cells incorporating a chimeric DNA sequence according to

30   the invention in their genome" shall encompass cells and multicellular organisms comprising or essentially consisting of such cells which stably incorporate said chimeric DNA into their genome thereby maintaining the chimeric DNA, and

preferably transmitting a copy of such chimeric DNA to progeny cells, be it through mitosis or meiosis. According to a preferred embodiment of the invention, plants are provided which essentially consist of cells which incorporate one or more copies of said chimeric DNA into their genome, and which are capable of transmitting a copy or

5   copies to their progeny, preferably in a Mendelian fashion. By virtue of the transcription and translation of the chimeric DNA of the invention in some or all of the plant's cells, those cells that comprise said regulatory region will respond to wounding and thus produce the protein encoded by the open reading frame which is under control of the regulatory region. In specific embodiments of the invention, this

10 protein will be an antipathogenic protein capable of conferring resistance to pathogen infections.

As is well known to those skilled in the art, regulatory regions of plant genes consist of disctinct subregions with interesting properties in terms of gene expression. Examples of such subregions include enhancers and silencers of transcription. These

15 elements may work in a general (constitutive) way, or in a tissue-specific manner. Deletions may be made in the regulatory DNA sequences according to the invention, and the subfragments may be tested for expression patterns of the associated DNA. Various subfragments so obtained, or even combinations thereof, may be useful in methods or applications involving the expression of heterologous DNA in plants. The

20 use of DNA sequences according to the invention to identify functional subregions, and the subsequent use thereof to promote or suppress gene expression in plants is also encompassed by the present invention.

Furthermore, it is generally believed that use of a transcriptional terminator region enhances the reliability as well as the efficiency of transcription in plant cells.

25 Use of such a region is therefore preferred in the context of the present invention.

Although the application only contains examples in *Brassica* and potato, the application of the present invention is advantageously not limited to certain plant species. Any plant species may be transformed with chimeric DNA sequences

30 according to the invention.

Although some of the embodiments of the invention may not be practicable at present, for example, because some plant species are as yet recalcitrant to genetic

transformation, the practising of the invention in such plant species is merely a matter of time and not a matter of principle, because the amenability to genetic transformation as such is of no relevance to the underlying embodiment of the invention.

5   Transformation of plant species is now routine for an impressive number of plant species, including both the *Dicotyledoneae* as well as the *Monocotyledoneae*. In principle, any transformation method may be used to introduce chimeric DNA according to the invention into a suitable ancestor cell, as long as the cells are capable of being regenerated into whole plants. Methods may suitably be selected from the

10   calcium/polyethylene glycol method for protoplasts (Krens, F.A. *et al.*, Nature 296, 72-74, 1982; Negrutiu I. *et al,*, Plant Mol. Biol. 8, 363-373, 1987), electroporation of protoplasts (Shillito R.D. *et al.*, Bio/Technol. 3, 1099-1102, 1985), microinjection into plant material (Crossway A. *et al.*, Mol. Gen. Genet. 202, 179-185, 1986), DNA (or RNA-coated) particle bombardment of various plant material (Klein T.M. *et al.*,

15   Nature 327, 70, 1987), infection with (non-integrative) viruses and the like. A preferred method according to the invention comprises *Agrobacterium*-mediated DNA transfer. Especially preferred is the use of the so-called binary vector technology as disclosed in EP A 120 516 and U.S. Patent 4,940,838. A further preferred method for transformation is the floral dip method essentially as described by Clough and Bent

20   (1998) Plant J. 16: 735-743.

Tomato transformation is preferably essentially as described by Van Roekel *et al.* (Plant Cell Rep. 12, 644-647, 1993). Potato transformation is preferably essentially as described by Hoekema *et al.* (Hoekema, A. *et al.,* Bio/Technology 7, 273-278, 1989).

25   Generally, after transformation, plant cells or cell groupings are selected for the presence of one or more markers which are encoded by plant expressible genes co-transferred with the nucleic acid sequence encoding the protein according to the invention, after which the transformed material is regenerated into a whole plant.

Although considered somewhat more recalcitrant towards genetic

30   transformation, monocotyledonous plants are amenable to transformation and fertile transgenic plants can be regenerated from transformed cells or embryos, or other plant material. Presently, preferred methods for transformation of monocots are

microprojectile bombardment of embryos, explants or suspension cells, and direct DNA uptake or electroporation (Shimamoto *et al.,* Nature <u>338</u>, 274-276, 1989). Transgenic maize plants have been obtained by introducing the *Streptomyces hygroscopicus bar*-gene, which encodes phosphinothricin acetyltransferase (an enzyme which inactivates the herbicide phosphinothricin), into embryogenic cells of a maize suspension culture by microprojectile bombardment (Gordon-Kamm, Plant Cell, <u>2</u>, 603-618, 1990). The introduction of genetic material into aleurone protoplasts of other monocot crops such as wheat and barley has been reported (Lee, Plant Mol. Biol. <u>13</u>, 21-30, 1989). Wheat plants have been regenerated from embryogenic suspension culture by selecting only the aged compact and nodular embryogenic callus tissues for the establishment of the embryogenic suspension cultures (Vasil, Bio/Technol. <u>8</u>, 429-434, 1990). The combination with transformation systems for these crops enables the application of the present invention to monocots.

Monocotyledonous plants, including commercially important crops, such as rice and corn are also amenable to DNA transfer by *Agrobacterium* strains (*vide* WO 94/00977; EP 0 159 418 B1; Gould J, *et al.,* Plant. Physiol. <u>95</u>, 426-434, 1991).

Following DNA transfer and regeneration, putatively transformed plants may be evaluated, for instance using Southern analysis to monitor the presence of the chimeric DNA according to the invention, copy number and/or genomic organization. Additionally or alternatively, expression levels of the newly introduced DNA may be undertaken, using Northern and/or Western analysis, techniques well known to persons having ordinary skill in the art.

Following such evaluations, the transformed plants may be grown directly, but usually they may be used as parental lines in the breeding of new varieties or in the creation of hybrids and the like.

To obtain transgenic plants capable of constitutively expressing more than one chimeric gene, a number of alternatives are available including the following:

<u>A</u>.   The use of DNA, for example, a T-DNA on a binary plasmid, with a number of modified genes physically coupled to a selectable marker gene. The advantage of this method is that the chimeric genes are physically coupled and therefore migrate as a single Mendelian locus.

<u>B</u>.   Cross-pollination of transgenic plants each already capable of expressing one or more chimeric genes, preferably coupled to a selectable marker gene, with pollen from a transgenic plant which contains one or more chimeric genes coupled to another selectable marker. The seed, obtained by this crossing, maybe selected on the basis of the presence of the two selectable markers, or on the basis of the presence of the chimeric genes themselves. The plants obtained from the selected seeds can then be used for further crossing. In principle, the chimeric genes are not on a single locus and the genes may therefore segregate as independent loci.

<u>C</u>.   The use of a number of a plurality of chimeric DNA molecules, for example, plasmids, each having one or more chimeric genes and a selectable marker. If the frequency of co-transformation is high, then selection on the basis of only one marker is sufficient. In other cases, the selection on the basis of more than one marker is preferred.

<u>D</u>.   Consecutive transformation of transgenic plants already containing a first, second etc., chimeric gene with new chimeric DNA, optionally comprising a selectable marker gene. As in method B, the chimeric genes are in principle not on a single locus and the chimeric genes may therefore segregate as independent loci.

<u>E</u>.   Combinations of the above mentioned strategies.

The actual strategy may depend on several easily determined considerations, such as the purpose of the parental lines (direct growing, use in a breeding programme, use to produce hybrids). The actual strategy is not critical with respect to the described invention.

14

Examples

General

Plant material

5   All *Brassica* transformation experiments described were performed with hypocotyl
segments of *Brassica napus* variety 'Westar'. Tissue culture conditions were
essentially as described by Bade and Damm (1995, In: Gene Transfer to Plants;
Potrykus, I.; Spangenberg, G. Eds. Springer Verlag: Berlin; pp 32-38). For seed
production or chromosomal DNA isolation transgenic plants were grown in pots
10   (diameter 15 cm) in a greenhouse with the following conditions: 21-24 °C, 60-80 %
humidity and a 16 hour light cycle.

The potato material used for transformation experiments were *in vitro* stem explants
from *Solanum tuberosum* variety 'Desiree'.

Bacterial strains

15   *Escherichia coli* strain DH5α (Clonetech) and DH10B (Clonetech) were used for
bacterial cloning. Strains were grown at 37°C in LB medium supplemented with
carbenicillin (100 mg/L), kanamycin (50 mg/L) or spectinomycin (50 mg/L)
depending on the type of plasmid. *Agrobacterium tumefaciens* strain MOG301 (Hood
et al., 1993, Transgenic Research 2:208-218), harbouring a non-oncogenic nopaline
20   Ti-helper plasmid in a C58 chromosomal background, was grown at 29 °C in LB
medium supplemented with kanamycin (100 mg/L) and rifampicin (20 mg/L).

Plasmid constructions

Construct pMOG22 was described by Goddijn *et al.* (1993, Plant Journal 4(5):863-
873). Vector pMOG448 was made in two steps. In the first step the *Hind*III 35S-
25   *gus*intron fragment of p35SGUS.INT (Vancanneyt et al., 1990, Molecular and
General Genetics 220:245-250) was cloned into pMOG22. Then the 5.8 kb *Xba*I
fragment of pGH1 (Haughn et al., 1988, Molecular and General Genetics 211:266-
271) was cloned in between the *hpt* and *gus*-intron parts. This particular fragment
contains a mutant Arabidopsis acetolactate synthase gene (*csr*-1), which confers

resistance to the herbicide chlorsulfuron. The coding region of the mutant *als* gene is still accompanied by its own 5' (2.5 kb) and 3' (1.3 kb) regulatory sequences.

Tagging constructs pMOG1178 and control pMOG964 contain plasmid rescue features (Koncz *et al.*, 1989, Proceedings of the National Academy of Sciences of the

5   United States of America   86:8467-8471.) but some modifications were made specifically for application in the *Brassica napus* transformation protocol. Due to the routine use of carbenicillin as antibiotic to control *Agrobacterium* it was decided to destroy the functionality of the *amp* gene and add a spectinomycin resistance gene instead. The constructs were made as follows.

10   The *EcoR* I site in pUC9 (Vieira and Messing, 1982, Gene 19:259-268) was modified by inserting an adapter made from oligo LS216 (5'AATTAGATCT 3')(SEQ ID NO: 2). The *Bgl*II site was then used for insertion of a 3 kb *Bam*HI fragment containing a bacterial spectinomycin resistance gene isolated from plasmid Cel369 (unpublished, Leiden University). The *amp* resistance gene was disrupted by partial digestion with

15   *Ava*II. Positive clones were selected for resistance to spectinomycin and susceptibility for carbenicillin.

A p35S-*gus::npt*II-tnos fusion gene (Datla *et al.*, 1991, Gene 101:239-246) was isolated as *Hin*dIII-*Bgl*II fragment from pBI426 (Charest *et al.*, 1993, Plant Cell Reports 12:189-193) and introduced in our spectinomycin vector, which was digested

20   with *Hin*dIII and *Bam*HI. This intermediate vector was linearised using *Hin*dIII, cloned in binary vector pMOG22 and named pMOG964.

The *Hin*dIII site of the above mentioned intermediate vector was changed into *Eco*RI using an adapter made out of primers SV5 (5'-AGCTCACGAATTCTCAGG-3') (SEQ ID NO: 3) and SV6 (5'-AGCTCCTGAGAATTCGTG-3')(SEQ ID NO: 4). The

25   resulting vector was digested with *Bst*BI and *Eco*RI and ligated into the likewise digested tagging vector pMOG553 (Goddijn *et al.*, 1993; EMBL database accession number X84105). In this way a stretch of 1038 base pairs in pMOG553 was replaced by about 8 kb of new sequence without altering the right border configuration (*gus*-intron and octopine border). The new vector pMOG1178 appeared unstable in *E. coli*

30   in the desired orientation. Hence the final cloning step was performed in *Agrobacterium*.

Constructs pJB1178-21, pJB1178-29 and pJB1178-43 were obtained via plasmid rescue (see below) from transgenic lines 1178-21, 29 and 43 respectively. These multicopy plasmids were linearised (*Eco*RI) and cloned as fragments in pMOG22 resulting in binary vectors pJBbin1178-21, pJBbin1178-29 and pJBbin1178-43

5   respectively.

Binary vectors were introduced in *Agrobacterium* strain MOG301 using electroporation (protocol Gibco BRL).

Plant transformation

10   Hypocotyl segments were transformed according to the procedure described by Bade and Damm (1995, In: Gene Transfer to Plants; Potrykus, I.; Spangenberg, G. Eds. Springer Verlag: Berlin; pp 32-38;). Minor modifications were included as follows. Kinetin was omitted from the callus induction medium (CIM) and NAA (0.1 mg/L) was added to the regeneration medium (SIM). The concentration of kanamycin as

15   selective agent was 15 mg/L. The sucrose level of the regeneration medium was lowered to 10 g/L to increase visual contrast between wildtype and transgenic callus. Shoot elongation was performed on non-selective medium (SEM). Transgenic nature of the plants produced was confirmed by rooting on hygromycin (5 mg/L) containing medium (SEM).

20   Potato *in vitro* stem explants were isolated one day prior to *Agrobacterium* inoculation. They were cultured in liquid callus induction medium (MS salts, B5 vitamines, sucrose 30 g/l, zeatin riboside 0.5 mg/l and 2,4-D 1.0 mg/l). After Agrobacterium inoculation (OD600 0.2, 20 minutes) the explants were cocultivated for 2 days on solidified callus induction medium (agar 8 g/l) and subsequently

25   transferred to regeneration medium (MS salts, B5 vitamines, sucrose 30 g/l, cefotaxim 200 mg/l, vancomycin 100 mg/l and zeatin riboside 3.0 mg/l). About one week after transformation explants were transferred to fresh regeneration medium, which was supplemented with hygromycin (10 mg/l) or kanamycin (100 mg/l). This medium was refreshed biweekly. Shoots were harvested 8 weeks later and placed on selective

30   rooting medium (1/2 concentrated MS salts, 1/2 concentrated B5 vitamines, sucrose 10g/l, IBA 0.1 mg/l and hygromycin 5 mg/l).

GUS histochemical assay

GUS activity in different plant parts of transgenic lines was investigated using histochemical analysis as described by Jefferson *et al.* (1987, Plant Molecular Biology Reporter 5:387-405). Samples of in vitro and in vivo plants were vacuum infiltrated

5   for 5 minutes in a solution containing 5-Bromo-4-Chloro-3-Indolyl-ß-D-glucuronicacid-cyclohexyl ammoniumsalt (0.5 mg/l); Na-P-Buffer (50 mM pH7); Na2-EDTA (5 mM pH 8.0); Triton X-100 (0.05 % v/v); Potasium Ferrocyanyde (0.5 mM); Potasium Ferricyanyde (0.5 mM). Samples were incubated overnight at 37 °C and subsequently cleared from chlorophyll by washing with ethanol (70 %). A

10   classification of GUS activity (0-5 = zero - very high) was made based on intensity of blue staining.

Callus induction assay

Small *Brassica napus* leaf disks (5*5 mm) of in vitro plants were placed adaxial side

15   up on regeneration medium (SIM) supplemented with 2,4-D (1 mg/L). Sucrose level in the medium was kept at 10 g/L. After 3 weeks of culture new green callus was formed at the cutting edges and complete explants were histochemically stained for GUS activity. Leaf samples of transgenic potato lines were similarly placed on potato regeneration medium supplemented with 2,4-D (1.0 mg/l). These explants were

20   assayed for GUS activity after 2 weeks.

Auxin induction assay

Node segments of in vitro plants were subculture on hormone free medium (SEM) supplemented with or without NAA (0.1 mg/L). After 3 weeks new leaves and roots

25   were formed. At that time complete plants were histochemically stained for GUS activity.

PCR analysis

Transgenic plants were analysed by PCR using the DNA sample preparation method

30   as described by Thompson and Henry (1995). Small leaf pieces (±2 mm2) were taken

from in vitro grown plantlets, sealed in micro-centrifuge tubes (1.5 mL) and frozen in liquid nitrogen. Twenty microliter extraction buffer (100 mM TrisHCL pH9.5; 1 M KCL; 10 mM EDTA) was added and samples were heated for 10 minutes at 95 °C. After cooling down on ice samples were used directly or stored at 4 °C until use. PCR primers were: 5'-GTGACATCTCCACTGACGTAAG-3' (35S-P4) (SEQ ID NO: 5) and 5'-CGAACTGATCGTTAAAACTGCC-3' (SQ-GUS-192)(SEQ ID NO: 6). The primer annealing sites are indicated in Figure 1. One PCR cycle of 5' 95 °C, 5' °C, 5' 72 °C was followed by 30 cycles of 1' 95 °C, 1' 55°C, 1' 72 °C. A last cycle was carried out for 1' 95 °C, 1' 55°C, 10' 72 °C. The reaction volume was 50 µl, containing 1 µl of DNA sample, Taq buffer, 1.5 mM MgCl2, 2*25 pmol primer, 200 µM dNTPs and 2.5 units Platinum Taq polymerase. PCR samples were analysed using electrophoresis in agarose gels.

Plasmid rescue

Approximately 5 µg of EcoRI digested genomic DNA of individual transgenic lines was dissolved in 25 µl of H2O. Five µl T4 ligase (Gibco BRL), 60 µl T4 ligase buffer and 210 µl H2O were added and the mixture was incubated for 20 hours at 14 °C. Ligated DNA was cleaned once using phenol-chloroform extraction (Sambrook et al. 1989, Molecular cloning: A laboratory manual, 2nd edition; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY) and subsequently dissolved in 10 µl TE. One µl of the solution was used for electroporation of one sample DH10B electromax (Gibco BRL) competent cells using a Cell porator system (Gibco BRL). Settings were used as suggested by the manufacturer. After one hour recovery in SOC medium, cells were spinned down, dissolved in 100 µl LB and plated on LB plates containing spectinomycin (50 mg/L). Colonies became visible after 24-48 hours incubation at 37°C. Subculture of individual colonies on plates and in liquid LB (Spec 50 mg/L) was used to confirm true resistant nature of the clones rescued.

Sequencing

Transition zones from the gus::nptII gene into the plant genome were sequenced using the ABI sequencing kit (Prism BigDye Terminator Cycle) and 5'-

CGAACTGATCGTTAAAACTGCC-3' (SQ-GUS-192) (SEQ ID NO: 6) as single primer. Rescued plasmids or binary vectors were used as template DNA. Conditions were applied as suggested by the manufacturer. Approximately 500-600 bp of sequence was determined. Sequence data were analysed using BLASTN computer searching.

Example 1 Transformation of *Brassica* with tagging construct

Hypocotyl explants of *Brassica napus* were transformed with the tagging construct pMOG1178 (Fig 1) and placed on medium containing 15 mg/L kanamycin which is the lowest concentration discriminating between resistant cell clusters and non-transgenic tissue. In the transformation experiments part of the explants was placed on hygromycin containing medium to select for expression of the 35S-hpt cassette. The frequency at which hygromycin resistant calli were obtained was used as a measure for the efficacy of T-DNA integration in a particular experiment. Construct pMOG448 (Fig 1) was used as a negative control for kanamycin selection. The same construct and construct pMOG964 (Fig 1) were used as positive controls for hygromycin selection. The latter construct was also used as the positive control for kanamycin selection.

From a series of 15 transformation experiments the results of three typical tagging experiments are presented in Table 1. The frequency at which hygromycin resistant calli were formed after transformation with pMOG1178 (number of resistant calli / explant * 100%) ranged from 55 to 99 percent. The callus frequency of the positive controls pMOG448 and pMOG964 ranged from 37 to 66 percent. After kanamycin selection, the callus frequency of the negative control pMOG448 was zero. The frequency obtained with the positive control pMOG964 was 81 - 119 %. A low but significant number of kanamycin resistant calli (1.4 - 3.5%) were produced on explants transformed with the tagging construct pMOG1178.

The relative tagging frequency is the ratio between kanamycin and hygromycin resistant callus formation within a certain tagging experiment. This number represents the fraction of T-DNA inserts integrated behind a genomic promoter sequence that is

active in callus tissue. The relative tagging frequency ranged from 2.6 to 3.8 percent

between the different experiments.

Table 1. Transgenic callus formation in *Brassica napus* experiments with a promoterless *gus::npt*II tagging construct. Hygromycin and kanamycin callus frequencies are calculated as number of resistant calli/explant*100%; relative tagging frequencies as kanamycin freq/hygromycin freq*100%. Datapoints are based on transformation of at least 150 hypocotyl explnats. Kanamycin selection experiments after transformation with pMOG1178 were performed with at least 500 explants.

| Construct | Marker / reporter | Exp. I | | | Exp II | | | Exp III | | |
		Hyg	Kana	kan/hyg	Hyg	Kana	kan/hyg	Hyg	Kana	kan/hyg
pMOG448	35S-*hpt* / 35S-*gus*	37	0	0	66	0	0	-	-	-
pMOG964	35S-*hpt* / 35S-*gus::npt*II	40	81	203	59	119	202	-	-	-
pMOG1178	35S-*hpt* / -*gus::npt*II	92	3.5	3.8	99	3.2	3.2	55	1.44	2.6

5   Eighty-seven kanamycin resistant calli were obtained from all tagging experiments. In

total 36 calli were successfully regenerated. This callus regeneration frequency (41 %)

is within the range obtained normally. Shoot primordia were isolated and subcultured

on medium without kanamycin. This non-selective step was used to allow the

development of tag lines with limited or no *npt*II expression after differentiation.

10   Twenty out of the 36 regenerated plants showed normal root formation on hygromycin

containing medium (5 mg/l), which indicated expression of the 35S-*hpt* cassette in

roots. This observation confirmed the successful introduction of the promoterless

*gus::npt*II tagging construct via kanamycin selection. Hygromycin-sensitive lines were

excluded from further analysis.

15

Example 2 GUS activity in differentiated plant tissue

Kanamycin resistant control plants containing the 35S-*gus::npt*II construct (pMOG964) showed high levels of constitutive GUS activity (data not shown). Thus,

5   the transgenic *gus::npt*II tag lines were expected to show GUS staining when the tagged genomic promoter was still active in certain plant tissues. Leaf, stem and root tissue of in vitro and greenhouse grown plants were histochemically assayed (Table 2). Fifteen of the 20 lines showed a detectable level of expression in one or more parts of the plant in either greenhouse or in vitro. The blue staining was usually very weak

10 and was often restricted to the vascular tissue of leaves and stem. In general, expression under greenhouse conditions was lower compared to in vitro conditions.

Four lines (1178-1, 26, 29 and 45) showed moderate to high levels of GUS activity in leaf, stem and root. Only one line (1178-26) showed a high constitutive expression pattern also after transfer to soil. Enhanced expression was observed in some shoot

15 (1178-2 and 30) or root meristems (1178-21, 29, 33, 43 and 45).

7.1.1 GUS activity in re-induced callus

As a control a set of analyses was carried out to investigate the GUS enzyme levels in re-induced callus. It was expected that most of the lines showed some level of GUS activity in this phase, because the original promoterless *gus::npt*II insertion resulted in

20 kanamycin resistant callus. Leaf disks of all lines were placed on shoot induction medium supplemented with 2,4-D (1 mg/l), which resulted in formation of green non-regenerating callus at the edges of the explants. Eighteen of the 20 lines showed a detectable level of GUS activity after 14 days of culture on this 2,4-D containing medium. Expression was mainly localised in callus tissue newly formed on the edges

25 of the explants (Fig 2a+b). Relative upregulated expression in callus compared to the rest of the explant, was found for 12 lines (1178-2, 5, 10, 11, 18, 21, 22, 30, 37, 40, 43 and 45). No detectable enzyme levels were found in other plant tissues investigated (see below).

Upregulated expression in callus was also observed when T1 hypocotyl segments of

30 tag lines were placed on regeneration medium (Fig 2c+d).

Table 2. Semi-quantitative analysis of GUS activity in *Brassica napus* tagged lines. Lines with relative upregulation are marked *

Promoter tag line	In vitro						In vivo					Comments
	Leaf	Stem	root	shoot meristem	2,4-D callus	NAA leaf	leaf	stem	root	shoot meristem	root meristem	
1178-1	3	3	2	3	3	3	1	1	0	1	0	35S
1178-2	0	1	1	1	1 *	0	0	0	0	1	* 0	
1178-5	1	1	1	2	3 *	2 *	0	1	1	1	0	
1178-10	0	0	0	0	4 *	0	0	0	1	0	0	
1178-11	0	0	0	0	1 *	0	0	0	0	0	0	
1178-18	0	0	0	0	1 *	1 *	0	0	0	0	0	
1178-21	1	1	1	2	4 *	2 *	0	1	0	0	1	* Single copy
1178-22	0	0	0	0	5 *	0	0	0	1	0	0	
1178-26	5	3	2	0	5	5	5	3	1	2	0	35S
1178-29	4	3	4	3	4	4	2	2	1	1	3	* Single copy
1178-30	0	0	0	0	2 *	0	0	0	0	1	* 0	
1178-32	1	1	0	1	1	0	0	1	0	1	0	
1178-33	1	1	0	1	2 *	2 *	1	1	0	0	2	*
1178-34	0	0	0	0	0	0	0	0	0	0	0	
1178-37	1	1	0	2	2 *	1	0	1	0	1	0	
1178-38	2	0	0	1	2	2	1	0	0	0	0	
1178-40	0	0	0	0	4 *	0	0	0	0	0	0	
1178-42	0	0	0	0	0	0	0	0	0	0	0	
1178-43	1	3	2	0	4 *	2 *	0	1	1	0	2	* Single copy
1178-45	2	2	1	1	5 *	4 *	1	1	1	1	2	*
# tg: 20	11	11	8	10	18	(13) 11	(6) 6	10	7	9	(2) 5	(5)

7.1.2 GUS activity upregulated by auxin treatment

5 Since T-DNA integration is thought to occur in actively transcribed regions of the genome (Koncz *et al.*, 1989, Proceedings of the National Academy of Sciences of the United States of America 86:8467-8471) and in this occasion integration of the promoterless *gus::npt*II tagging construct took place during culture on auxin containing medium (cocultivation with 2,4-D and selection with NAA), we wished to

10 check upregulation of the tagged promoters by auxin. In vitro plant material of all tag

lines was cloned. At least one node segment of each line was propagated on NAA (0.1 mg/l) containing medium. Plantlets grown on this medium developed significantly more and thicker roots, but were otherwise comparable to the control clones grown on hormone free medium. Histochemical GUS assays were performed 4 weeks after

5   subculture. Six lines (1178-5, 18, 21, 33, 43 and 45) showed NAA induced GUS activity in the leaves.

An overview of the GUS expression data is presented in Table 2. In two transgenic lines (1178-34 and 1178-42) no GUS staining was observed.

10   **7.1.3 Example 3 Isolation of genomic 'promoter' sequences**

Before the actual isolation of genomic sequences upstream of the tagging construct, all lines were screened by PCR to detect possible scrambled T-DNA insertions. Using a 35S-*gus* primer-set it was found that two lines (1178-1 and 26) contained at least part of the 35S promoter, presumably originating from the 35S-*hpt* cassette, upstream of

15   the promoterless *gus::npt*II gene. This was confirmed by sequencing the amplified fragments (data not shown). These lines were excluded from further analysis.

Genomic DNA of the remaining lines was digested with *Eco*RI (Fig 1) and used for Southern blotting analysis. Single T-DNA insertions were detected in lines 1178-21, 29 and 43. The T-DNA right border fragments were approximately 12 kb in size (data

20   not shown), which indicates ±3kb genomic sequence between T-DNA and *Eco*RI restriction site. Results of the other lines were difficult to interpret due to the very large sizes of the *Eco*RI fragments.

Digested DNA was also used for plasmid rescue experiments. Despite the large

25   fragment sizes observed on the Southern blot, spectinomycin resistant colonies were readily obtained. Rescued plasmids were checked by restriction enzyme analyses. These analyses were done using *Eco*RI, expected to linearize the plasmids. A double digest with *Eco*RI plus *Bam*HI was used to separate the original T-DNA (vector) of 9 kb from the newly isolated genomic sequences (Fig 1). Plasmid rescue from the single

30   copy T-DNA lines (1178-21, 29 and 43) resulted in identical clones within lines, whereas other lines showed 2 or more different restriction patterns (data not shown). It

is likely that these different restriction patterns represented rescued fragments of different T-DNA insertions. Some examples of rescued plasmids digested with the enzyme combinations are shown in Figure 3. The plasmids rescued from the single copy lines (1178-21, 1178-29 and 1178-43) are indicated and named according to the originating tag line (pJB1178-21(=pMOG2001), pJB1178-29(=pMOG2002) and pJB1178-43(=pMOG2003)).

The linear fragments (*Eco*RI) range in size from ±11 kb (clone 1) to ±40 kb (clones 4 and 5). The fragment sizes (±12kb) of the plasmids from single copy T-DNA lines (pJB1178-21, pJB1178-29 and pJB1178-43) match with the results obtained by Southern blotting (see above). In some cases (clones 7, 9 and 11) it appeared impossible to linearize the rescued plasmids by *Eco*RI digestion. Apparently the *Eco*RI site was destroyed. From the lanes with the double digestions, it can be seen that most of the clones show the expected 9 kb vector band (Fig 3). Exceptions are those clones without the *Eco*RI site as mentioned above and pJB1178-43. All other clones contain, besides the 9 kb fragment, 1-3 other bands originating from the isolated plant DNA.

DNA sequences upstream of the *gus::npt*II gene were determined for each of the 3 single copy lines (1178-21, 1178-29 and 1178-43). The original right border and *Hind*III site of pMOG1178 (Fig 1) were absent in all three lines (Fig 4). In line 1178-43 the *Bam*HI site (Fig1) was also not present anymore, which explains the absence of the expected 9 kb fragment after *Eco*RI**Bam*HI digestion (see above).

Analysis of tagged *Brassica napus* promoter sequence

The rescued *Brassica napus* promoter sequence of line 1178-29 (SEQIDNO: 7) was used in a BLAST (Altschul *et al.*, Nucleic Acids Res. 1997; 25:3389-3402) search against the *Arabidopsis* genome sequence (TIGR: www.tigr.org/tdb/e2k1/ath1/). Extensive homology was found to a certain portion of the *Arabidopsis* genome. The 2624 bp *Brassica* sequence displays high homology with a region on *Arabidopsis* chromosome 3 on BAC F9D24. The homology covers almost the entire region on *Arabidopsis* BAC F9D24 that represents predicted ORF F9D24.50 (phenylalanine-tRNA synthase-like protein) with 10 exons. The homology with the *Brassica* sequence is the strongest in regions were predicted exons are located but the homology is also present, although more limited, in (predicted) intron regions. The *Arabidopsis* sequence homologous to the *Brassica* 1178-29 sequence is listed as SEQIDNO:8.

The extensive homology found between the chromosomal regions of *Arabidopsis thaliana* ecotype Columbia and *Brassica napus* c.v. Westar shows the high level of genomic co-linearity between these closely related plant species.

5   The promoter sequence was analysed for the presence of promoter motifs known to play a regulatory role in auxin induced, pathogen induced (plant defence hormone responsive) and constitutive gene expression. Both promoters are very active in callus tissue and respond to auxin treatment. Next to this proven promoter activity there might be an involvement of pathogen and wound responsive elements in the regulation of gene expression driven by this promoter sequence as it was identified in promoter

10   trapping experiments during exposure to wounding and *A. tumefaciens* infection. Promoter elements identified in the promoter are indicated in Figure 6. Elements were found containing the core sequence (TGTCTC) of the auxin responsive elements (AuxREs) required for auxin responsiveness of a soybean *GH3* promoter (Ulmasov *et al.*, Plant Cell 1995 Oct;7(10):1611-1623). Next to the presence of these auxin responsive elements a sequence identical to a tobacco Dof protein NtBBF1 binding site found in

15   the *RolB* oncogene promoter is found. The NtBBF1 is probably the protein involved in mediating tissue specific and auxin inducible expression of *RolB* in plants (Baumann *et al.*, Plant Cell 1999 March;11(3):323-334). Dof zinc finger proteins are also thought to be auxin inducible (Kang and Singh, Plant J. 2000; 21:329-339). Elements that occur in a high frequency in genes that are pathogen and/or stress inducible were also identified in both promoter sequences. The W-box motifs that are able to bind

20   members of the plant WRKY family of transcription factors are present in both promoter sequences (8 and 4 copies respectively). The presence of a high frequency of W-box sequences (TTGACn) is associated with pathogen, elicitor and salicylic acid responsiveness (Eulgem *et al.*, Trends Plant Sci. 2000;5(5):199-206). Sequences very similar to the H-box consensus (CCTAnC) as described by Lois *et al.* (EMBO J. 1989;8(6):1641-1648) and Fischer (Ph.D. thesis, University of Hohenheim, 1994) were

25   found and these boxes are known to confer fungal elicitor and wound induced expression when fused as multimers to a plant minimal promoter (Takeda *et al.*, Plant J. 1999;18(4):383-393). Also identified were boxes similar to the so-called G-box regulatory motif (CAmGTG, Loake *et al.*, Proc. Natl. Acad. Sci. USA 1992; 89:9230-9234) and to the GCC-box (AGCCGCC) which is mainly found in the 5' upstream region of genes upregulated by the plant hormone ethylene (Ohme-Takagi and Shinshi, Plant

30   Cell 1995 Feb;77(2):173-182). A tetramer of an extended G-box motif confers high level constitutive expression in transgenic plants when coupled to a minimal promoter (Ishige *et al.*, Plant J. 1999;18:443-448). The S-box is a very strong elicitor responsive element, which can confer very strong inducibility (WO 00/29592). The transcription start site in the 1178-29 promoter fusion was not mapped and therefore it remains difficult to predict the location of a presumed TATA box. Nevertheless

35   there are sequences present that very well might function as a RNA polymerase II binding site.

7.1.4 Example 4 Evaluation of isolated 'promoter'-*gus::npt*II plasmids

The three plasmids rescued from the single copy T-DNA lines (1178-21, 29 and 43) were selected for further analysis of promoter activity. Binary vectors were constructed by using a double selection strategy. Linear fragments (*Eco*RI) of the rescued plasmids were ligated with a linear fragment (*Eco*RI) of the binary vector pMOG22 (Fig 1) and transformed to *E. coli*. Colonies were selected for kanamycin and spectinomycin resistance, indicating successful ligation. However, most of the binary clones appeared to contain a certain deletion, as evidenced by a reduction in size of the original 9kb *Eco*RI/*Bam*HI vector fragment (not shown). Sequence analysis of the promoter-*gus* fusions of the new binary vectors indicated unaltered presence of the genomic sequences directly upstream of the *gus::npt*II gene.

Three to five binary vectors were selected per rescued plasmid. This selection was based on best resemblance with the expected restriction patterns using *Eco*RI and *Bam*HI (data not shown). Twelve clones (4*pJBBIN1178-21, 5*pJBBIN1178-29 and 3*pJBBIN1178-43) were transferred to *Agrobacterium* strain MOG301 and subsequently transformed to *Brassica napus*. All transformations were carried out in duplo using ± 100 hypocotyl explants each. Transformations with pMOG964 and pMOG1178 (Fig1) served as controls in these experiments. Five days after transformation about twenty explants per construct were evaluated for transient GUS activity. Five out of twelve clones showed GUS activity in hygromycin resistant callus. Three weeks after transformation hygromycin resistant calli were produced in all transformations, except for one 1178-43 binary clone. At this time point eight out of twelve clones showed GUS activity in hygromycin resistant callus (Fig 5). Specifically clones from tag lines 1178-21 and 29 showed dark blue staining in the histochemical GUS assay.

Six binary vectors (2*pJBBIN1178-21, 2*pJBBIN1178-29 and 2*pJBBIN1178-43) were also used for potato transformation. Four of these showed transient GUS activity early after transformation. Five out of six vectors revealed stable GUS activity in developing hygromycin resistant calli. Approximately 50 potato explants per construct were used for a transformation experiment using kanamycin as selective agent. Putative transgenic shoots were harvested and tested for their true transgenic nature by

placing them on hygromycin containing rooting medium. Only those shoots having an active35S-hpt cassette integrated in the genome would produce a normal root system. Five out of 6 constructs transformed produced one or more transgenic potato lines. The transformation frequencies (number of transgenic plants per explant * 100%)

5   ranged form 2 to 11 percent, which is comparable to the level obtained with the positive control construct pMOG964 (5%).

Preliminary results indicated that GUS activity in leaf samples of the transgenic potato lines varies from zero to relatively high. Some of the transgenic lines show only low GUS activity in leafs, but this level could be up-regulated when leaf explants were

10   placed on callus induction medium (Fig 5). An overview of the Brassica and potato transformation results is shown in Table 3.

Table 3. Transformation of B. napus and S. tuberosum with tagged promoter sequences

Binary vectors	B. napus (hyg)		S. tuberosum (hyg)		S. tuberosum (kana)							
	transient GUS	stable GUS	transient GUS	stable GUS	stable GUS	n	nr of reg expl	nr of shoots	nr hyg rooting	nr leaf GUS+	nr callus GUS+	transf freq
pJBBIN1178-21-1	+++	+++	+++	+++	+++	44	20	11	5	5	2	11
pJBBIN1178-21-3	-	-										
pJBBIN1178-21-7	-	-		+								
pJBBIN1178-21-11	+	++	++	+	++	48	15	13	2	2	1	4
pJB1178-29-1	-	+	+	+	+	46	3	7	5	4	2	11
pJB1178-29-3	+++	+++										
pJB1178-29-4	-	-										
pJB1178-29-5	-	+										
pJB1178-29-11	++	+++	++++	++	++++	54	7	7	1	1	0	2
pJB1178-43-1	-	+	-	-	+	64	1	4	0	-	-	0
pJB1178-43-2	-	-										
pJB1178-43-3	+	++	-	+	++	65	5	10	2	2	1	3
pMOG964	+++++	+++++	+++++	+++++	+++++	38	18	18	2	2	0	5
pMOG1178	-	-	-	-	-	43	1	0	0	-	-	0

Table 3. Retransformation of *Brassica napus* and *Solanum tuberosum* with tagged promoter sequences. Transformation frequencies are calculated as number of transgenic lines per explant * 100%.

BUDAPEST TREATY ON THE INTERNATIONAL
RECOGNITION OF THE DEPOSIT OF MICROORGANISMS
FOR THE PURPOSES OF PATENT PROCEDURE

INTERNATIONAL   FORM

Syngenta Mogen BV Postbus 628 2300 AP LEIDEN Nederland *name and address of depositor*	RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT issued pursuant to Rule 7.1 by the INTERNATIONAL DEPOSITARY AUTHORITY identified at the bottom of this page

I.   IDENTIFICATION OF THE MICROORGANISM

Identification reference given by the DEPOSITOR: E. coli DH10B (Gibco BRL) pMOG2002	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: CBS 109272

II.   SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION

The microorganism identified under I above was accompanied by:

[X] a scientific description

[ ] a proposed taxonomic designation

(mark with a cross where applicable)

III. RECEIPT AND ACCEPTANCE

This International Depositary accepts the microorganism identified under I above, which
received by it on   **06-02-2001**   *(date dd-mm-yy of the original deposit)* 1

IV.   RECEIPT OF REQUEST FOR CONVERSION

The microorganism identified under I above was received by this International Depositary
Authority on   **not applicable**   *(date dd-mm-yy of the original deposit)*   and a
request to convert the original deposit to a deposit under the Budapest Treaty was received
by it on   **not applicable**   *(date dd-mm-yy of receipt of request for conversion)*

V.   INTERNATIONAL DEPOSITARY AUTHORITY

Name: **Centraalbureau voor Schimmelcultures** Address **Uppsalalaan 8 P.O. Box 85167 3508 AD UTRECHT The Netherlands**	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): Mrs F.B. Snippe-Claus Dr J.A. Stalpers Date (dd-mm-yy): 19-02-2001

1   Where Rule 6.4(d) applies, such date is the date on which the status of international
    depositary authority was acquired.

Form BP/4   (sole page)   CBS/910
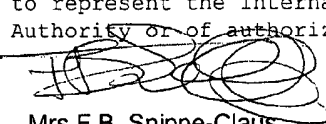
BUDAPEST TREATY ON THE INTERNATIONAL
RECOGNITION OF THE DEPOSIT OF MICROORGANISMS
FOR THE PURPOSES OF PATENT PROCEDURE

INTERNATIONAL FORM

Syngenta Mogen BV
Postbus 628
2300 AP LEIDEN
Nederland

*name and address of the party to whom the
viability statement is issued*

VIABILITY STATEMENT
issued pursuant to Rule 10.2 by the
INTERNATIONAL DEPOSITARY AUTHORITY
identified on the following page

I. DEPOSITOR	II. IDENTIFICATION OF THE MICROORGANISM
Name: Syngenta Mogen BV	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: **CBS 109272**
Address Postbus 628 2300 AP LEIDEN Nederland	Date (dd-mm-yy) of the deposit or of the transfer: 1 06-02-2001

III. VIABILITY STATEMENT

The viability of the microorganism identified under II above was tested
on 08-02-2001   2 .   On that date (dd-mm-yy), the said microorganism was

$\boxed{X}$³ viable

$\boxed{}$³ no longer viable

¹ Indicate the date of the original deposit or, where a new deposit or a transfer has been
made, the most recent relevant date (date of the new deposit or date of the transfer).

² In the cases referred to in Rule 10.2(a)(ii) and (iii), refer to the most recent
viability test.

³ Mark with a cross the applicable box.
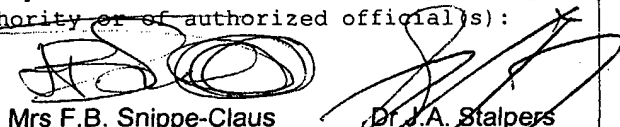
IV. CONDITIONS UNDER WHICH THE VIABILITY HAS BEEN PERFORMED

V. INTERNATIONAL DEPOSITARY AUTHORITY

Name:	Centraalbureau voor Schimmelcultures	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s):
Address	Uppsalalaan 8 P.O. Box 85167 3508 AD UTRECHT The Netherlands	Mrs F.B. Snippe-Claus Dr J.A. Stalpers Date (dd-mm-yy): 19-02-2001

4 Fill in if the information has been requested and if the results of the test were negative.

CLAIMS

1. A DNA fragment as present within the EcoRI fragment of clone pJB1178-29, which is deposited with the Centraal Bureau of Schimmelcultures (Baarn, the Netherlands) under number CBS 109272 on 6 February 2001 wherein said DNA fragment is capable of promoting constitutive expression of an associated DNA sequence on reintroduction into a plant.

2. A DNA fragment according to claim 1, characterised in that said fragment comprises the nucleotide sequence represented by SEQ ID NO: 1.

3. A DNA fragment according to claim 1, characterised in that it comprises the nucleotide sequence represented by SEQ ID NO:7

4. A DNA fragment, characterised in that it comprises the nucleotide sequence represented by SEQ ID NO:8 or parts thereof, wherein said fragment or parts thereof are capable of promoting constitutive expression of a DNA sequence on reintroduction into a plant.

5. A DNA fragment capable of promoting constitutive expression of a DNA sequence on reintroduction into a plant, characterised in that it comprises a nucleotide sequence represented in SEQ ID NO:7 starting at the nucleotide selected from the group consisting of: nucleotide 4, nucleotide 34, nucleotide 341, nucleotide 588, nucleotide 650, nucleotide 782, nucleotide 825, nucleotide 937, nucleotide 1246, nucleotide 1556, nucleotide 1780, nucleotide 1849, nucleotide 1912, nucleotide 1960, nucleotide 2044, nucleotide 2243, nucleotide 2408 and nucleotide 2638; and ending at the nucleotide selected from the group consisting of: nucleotide 341, nucleotide 588, nucleotide 650, nucleotide 782, nucleotide 825, nucleotide 937, nucleotide 1246, nucleotide 1556, nucleotide 1780, nucleotide 1849, nucleotide 1912, nucleotide 1960, nucleotide 2044, nucleotide 2243, nucleotide 2408, nucleotide 2638 and nucleotide 2654.

6. A chimeric DNA sequence comprising in the direction of transcription:

    a) a DNA fragment according to any of claims 1 to 5; and

    b) a DNA sequence to be expressed under transcriptional control of said DNA fragment, which DNA sequence is not naturally under transcriptional control of said DNA fragment.

7. A replicon comprising a chimeric DNA sequence according to claim 6.

8. A microorganism containing a replicon according to claim 7.

9. A plant cell comprising a chimeric DNA sequence according to claim 6.

10. A plant comprising cells according to claim 9.

11. A part of a plant selected from seeds, flowers, tubers, roots, leaves, fruits, pollen and wood, obtained from a plant according to claim 10.

12. Use of a DNA fragment according to any of claims 1 to 5 in identifying homologues capable of promoting expression in a plant.

13. Method of transforming plants using a chimeric DNA sequence according to claim 6.

14. Method of transforming plants characterised in that a DNA fragment according to any of claims 1 to 5 is driving expression of a selectable marker gene.

15. Method of producing hybrid regulatory DNA sequences using a fragment or homologue of said DNA fragment according to any of claims 1 to 5.

Fig. 1

Fig. 2

Fig. 3

pMOG1178

HindIII

T-DNA right border

BamHI

gus::nptII

GCGACTTAATGATTACAACGGTATATATCCTGCCAAAGCTTGGATCCCGGGTAGGTCAGTCCCTTATGTTACGTCCTGTAG

pJB1178-21

plant

BamHI

gus::nptII

GCTAGCTAAGGTATTATATAAAACTGTGGGTGATTCGAGCTTGGATCCCGGGTAGGTCAGTCCCTTATGTTACGTCCTGTAG

A. thal.: BACT22C5 (2e-24)
B. napus: cDNA pMB68-4F; L. maculans defence respons (3e-19)

pJB1178-29

plant

BamHI

gus::nptII

TTTTTCAAATGCAGAGTAAGGTGCGTGATGAGGTGCAGAGCTTGGATCCCGGGTAGGTCAGTCCCTTATGTTACGTCCTGTAG

A. thal.: BACF9D24 (3e-39)
A. thal. cDNA 701549985; phenyl alanine-tRNA synthetase like protein (3e-28)

pJB1178-43

plant

gus::nptII

TTTCACTTTAATTTCTACTCATATCCATCATTCCCTGACCATTTCACCCGGGTAGGTCAGTCCCTTATGTTACGTCCTGTAG

A. thal.: BACF15L12 (2e-12)
(promoter region of a DOF6 zinc finger like protein)

Fig. 4

5/6

Fig. 5

Fig. 6

SEQUENCE LISTING

```
<110> Syngenta Mogen B.V.

<120> NEW PLANT CONSTITUTIVE PROMOTER

<130> brassica constitutive promoter

<140>
<141>

<150> 01202056.6
<151> 2001-05-31

<160> 8

<170> PatentIn Ver. 2.1

<210> 1
<211> 782
<212> DNA
<213> Brassica napus

<220>
<221> misc_feature
<222> (642)..(782)
<223> gus::nptII-sequence

<220>
<221> misc_feature
<222> (617)..(622)
<223> BamHIrestriction site

<400> 1
tcagtgaaag atccagtagt cctgttagtt ttgacgggtg taaacaaaac cattgtttta 60
cagtatccct ccttgtataa gacaccagtt tctggatcag tgaatcattc acagagaata 120
acttttgtga agttgtgaga ggaatcgctg gggatcttgt tgaagaggta cataacttat 180
cctttgattg gtatttggtt gagaaagaat gctaactctc tatatctcaa ctttacttgt 240
atcataatca tgttcttggg agtgattgtt tatagccttt tacaaattga ttcacaggtg 300
aagttgatag acagtttcac caataagaaa gggatgacga gtcactgtta cagaattgtg 360
ttccgttcca tggagcgctc tcttacagac gaggaggtca atgatctgca ggtaatcact 420
gttgcttgtt ttgtcattaa tccagaaacg acatttactt gtttataatt caaaaccttt 480
tgtagctaaa ttacactctc catataacca accataagaa gataggaagg ttgcatttgg 540
ctaattgctt gttagtgtta aaaattggcg tgttttttca aatgcagagt aaggtgcgtg 600
atgaggtgca gagcttggat ccccgggtag gtcagtccct tatgttacgt cctgtagaaa 660
ccccaacccg tgaaatcaaa aaactcgacg gcctgtgggc attcagtctg gatcgcgaaa 720
actgtggaat tggtcagcgt tggggggaaag cgcgttacaa gaaagccggg caattgcggg 780
ca                                                               782

<210> 2
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 2
aattagatct                                                        10

<210> 3
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer
```
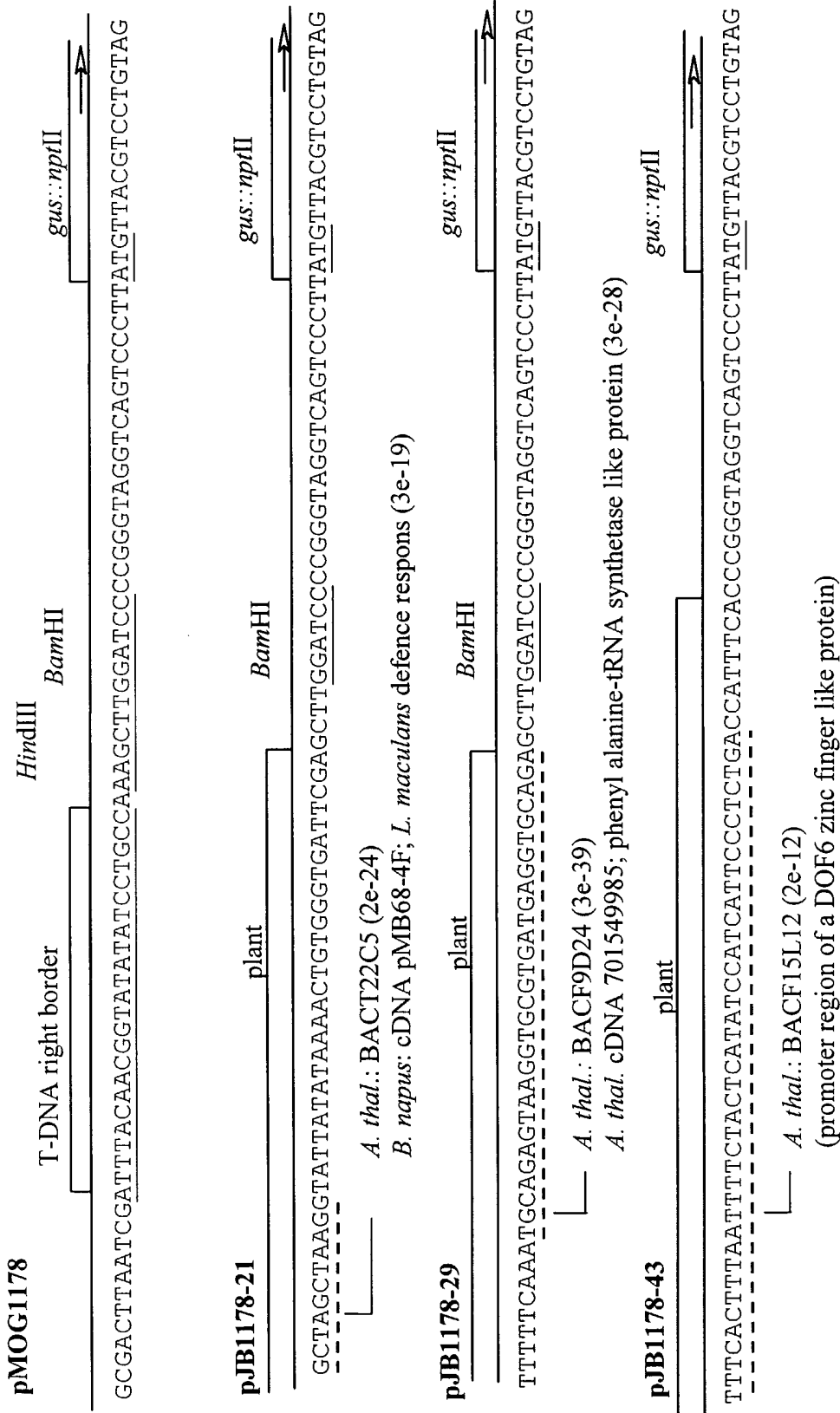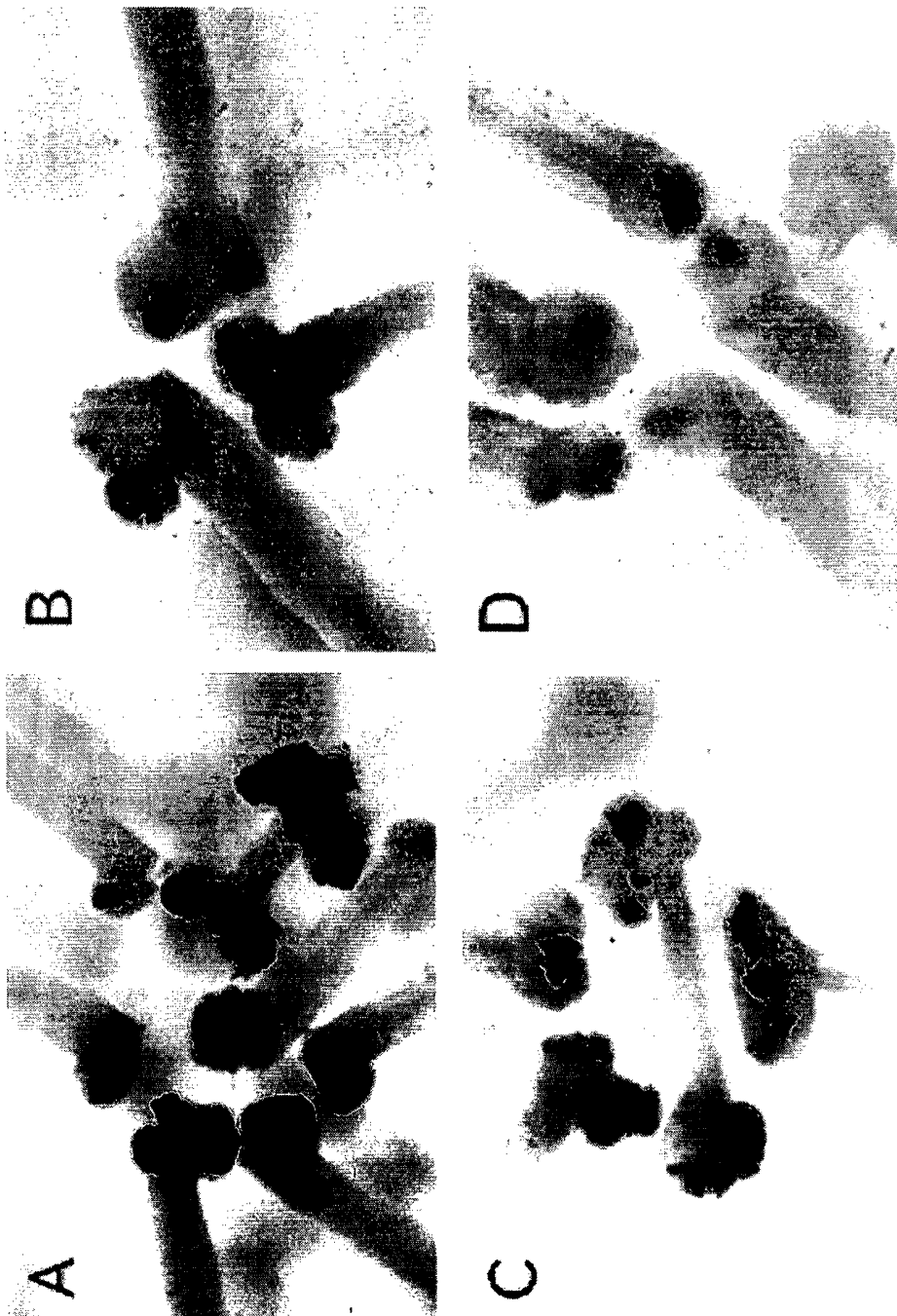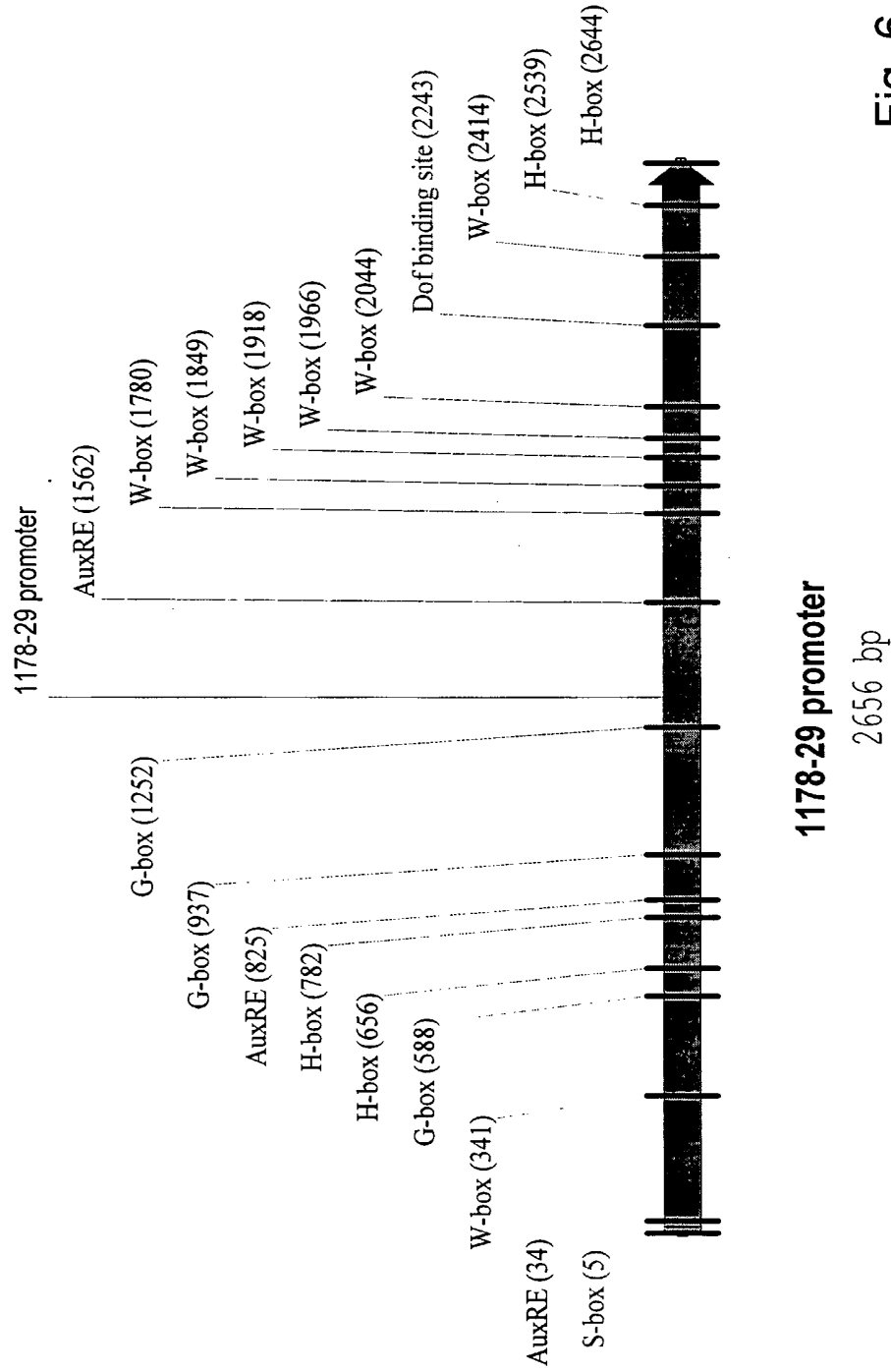
```
        <400> 3
        agctcacgaa ttctcagg                                          18


  5     <210> 4
        <211> 18
        <212> DNA
        <213> Artificial Sequence

 10     <220>
        <223> Description of Artificial Sequence:primer

        <400> 4
        agctcctgag aattcgtg                                          18
 15

        <210> 5
        <211> 22
        <212> DNA
 20     <213> Artificial Sequence

        <220>
        <223> Description of Artificial Sequence:primer

 25     <400> 5
        gtgacatctc cactgacgta ag                                     22


        <210> 6
 30     <211> 22
        <212> DNA
        <213> Artificial Sequence

        <220>
 35     <223> Description of Artificial Sequence:primer

        <400> 6
        cgaactgatc gttaaaactg cc                                     22

 40
        <210> 7
        <211> 2656
        <212> DNA
        <213> Brassica napus
 45
        <220>
        <221> misc_feature
        <222> (4)..(10)
        <223> S-box
 50
        <220>
        <221> misc_feature
        <222> (34)..(39)
        <223> Auxin-responsive element
 55
        <220>
        <221> misc_feature
        <222> (341)..(346)
        <223> W-box
 60
        <220>
        <221> misc_feature
        <222> (588)..(593)
        <223> G-box
 65
        <220>
        <221> misc_feature
        <222> Complement((650)..(655))
        <223> H-box
 70
```

```
       <220>
       <221> misc_feature
       <222> (782)..(787)
       <223> H-box
  5
       <220>
       <221> misc_feature
       <222> (825)..(830)
       <223> auxin-responsive element
 10
       <220>
       <221> misc_feature
       <222> (937)..(942)
       <223> G-box
 15
       <220>
       <221> misc_feature
       <222> Complement((1246)..(1251))
       <223> G-box
 20
       <220>
       <221> misc_feature
       <222> Complement((1556)..(1561))
       <223> auxin-responsive element
 25
       <220>
       <221> misc_feature
       <222> (1780)..(1785)
       <223> W-box
 30
       <220>
       <221> misc_feature
       <222> (1849)..(1854)
       <223> W-box
 35
       <220>
       <221> misc_feature
       <222> Complement((1912)..(1917))
       <223> W-box
 40
       <220>
       <221> misc_feature
       <222> Complement((1960)..(1965))
       <223> W-box
 45
       <220>
       <221> misc_feature
       <222> (2044)..(2049)
       <223> W-box
 50
       <220>
       <221> protein_bind
       <222> (2243)..(2248)
       <223> Dof binding site
 55
       <220>
       <221> misc_feature
       <222> Complement((2408)..(2413))
       <223> W-box
 60
       <220>
       <221> misc_feature
       <222> Complement((2533)..(2538))
       <223> H-box
 65
       <220>
       <221> misc_feature
       <222> Complement((2638)..(2643))
       <223> H-box
 70
```

```
<220>
<221> misc_feature
<222> (2654)..(2656)
<223> translation start site

<400> 7
actcgccacc gcgattctcg tcgtcggaga ctttgtctcc cccccccctc ttcatcaacg 60
gtgttcctcg ataaacttcg tgtttactca tctccgacct cgaatacgcc atgaccatat 120
tttcagtcca gtccactatc ttcacccgag cttccgtcgc tcttctctcc agcaacggac 180
tcaaacgctt ttctctcgct tcttcgtttt cctccaacgc tctatactct ccacctctcc 240
ccaaaacgaa gaagcgccgc ttccccatcg tctctgccgt tgatatcggc ggtgtcacag 300
tcgctagaaa cggttcgtgc ctctgattta cagattgagt ttgacttagt ggaagctcgt 360
tagcttgaaa tgttcaactt cttttttttt ttgcagatgt ggtgagagat gatgatccta 420
caaacaatgt acccgactcc atcttctcta aacttggaat gcagctacac agaagggaca 480
agcatcctat tgggatcata aagaatgcta tctacgacta cttcgagtcc aattacgcta 540
aaaagtttga gactttcgaa gatctttcac caattgttac caccaagcaa gtgagtgtac 600
ttcccccctta ctcaaagctt gcatctttaa ctggaaccat catcatgagg ataggagact 660
ctctgtttca catagagtgt tttctgttag agactgagag ttgttagagt aacatgctga 720
atattgtgtg agactctctg atgagatctt agcttgagtt ttcgtttgta tttgctctag 780
tcctatcaat aaaaagagtt acatagtcgc aatcataata acaatgtctc ctttggtgtc 840
agaactttga tgatgtgcta gtccctgctg atcagtaag cagaagcctt aacgacacct 900
actacgtaga ttcccaaact gttttgagat gccatacaag tgctcaccaa gctgagctgt 960
tgcgggatgg tcatagacgt ttccttgtca ctggagatgt ttaccgcaga gattctatcg 1020
actctactca ttatccggtt ttccatcagg tgttctttc tttcactttg gctgtttgg 1080
tcgacatcat gtgtttctta tactagttgt ttatgttttc tcttgaatct aatacagatg 1140
gaaggctttt gtgtcttttc tcctgaggac tggaacgagt ctggcaagga ttccacgttg 1200
tatgctgctg aggatttgaa gaaatgtcta gaggggttgg cacgtcactt gtttggtaag 1260
ctaagatcta ttcaacagaa gtaatcttca gatagttagc agtacttctt tgggtgtttt 1320
ataggatgat tatagcatcg actataaaat gaagatgcat atatttagat gaaacgattt 1380
aacatagaac aaacttgtga atctgttact ctttgattta agctttcttt ggctgccagg 1440
tgctgtggaa atgagatggg ttgatacata ctttccattt actgagcctt ctttcgagct 1500
tgagatttat tttaaggtag tcttttttcca tcttaaatac ttgctttgct ttaaagagac 1560
atacttttgt ttttgtgcag acatacatat actggagttg tgttttggct accaatctta 1620
gcaagctaac aagctttatc tggttgttga ttcaggaaga ctggttagag gttttgggct 1680
gtggggtgac ggagcaaaga attttgaagc agagtggatt agaaaacaat gttgcttggg 1740
cctttggact aggattggaa cggcttgcta tggttttgtt tgacatccct gatataagac 1800
ttttttttggt cagacgatga acggtttact tcccaggtga ttacttgatt gacaacttac 1860
aaaattggat gatgaagttt attctctcaa acgtttgag ttcttttctc acgtcaatgt 1920
atcttgatgt tatgtgcagt ttggaaaagg agagcttgga gtcaagttca agccattctc 1980
aaaggttata caattttttca ctcttcatgt ctatcagtga aagattcagt agtcctgtta 2040
gttttgacgg tgctaaacaa aacattgttt tacagtatcc tccttgttat aaggacatca 2100
gtttctggat cagtgaatca ttcacagaga ataacttttg tgaagttgtg agaggaatcg 2160
ctggggatct tgttgaagag gtacataact tatcctttga ttggtatttg gttgagaaag 2220
aatgctaact ctctatatct caactttact tgtatcataa tcatgttctt gggagtgatt 2280
gtttatagcc ttttacaaat tgattcacag gtgaagttga tagacagttt caccaataag 2340
aaagggatgc gagtcactgt tacagaattg tgttccgttc catggagcgc tctcttacag 2400
acgaggaggt caatgatctg caggtaatca ctgttgcttg ttttgtcatt aatccagaaa 2460
cgacatttac ttgtttataa ttcaaaacct tttgtagcta aattacactc tccatataac 2520
caaccataag aagataggaa ggttgcattt ggctaattgc ttgttagtgt taaaaattgg 2580
cgtgtttttt caaatgcaga gtaaggtgcg tgatgaggtg cagagcttgg atccccgggt 2640
aggtcagtcc cttatg                                              2656


<210> 8
<211> 2485
<212> DNA
<213> Arabidopsis thaliana

<400> 8
aatcgccgcc gcaatcttct tcatcggcct ccgttctaca tcgacggtgt ttgccgtaac 60
ttctgtcaaa ctctcagaat ttgcttaagt ataccaccta actcgagacg ctatgaccgt 120
tttctcagtt cagtccacta tcttcagtcg agcctccgta gctcttctct cgagcaatgg 180
cttcaaacga ttttcattcg tttcttcgtt ttcttcctcc gccgcttact ctccacctaa 240
aatgaggaag cgtcgctacc caatcgtctc tgctgttgat attggtggcg tcgcaatcgc 300
tagaaatggt tcgttcttag attcgattct taaaagtgaa gttcataaaa catcgcactt 360
gctccaaaag aagttatatt tgacattttt tagtgtacac ttattgaatt ttcagatgtg 420
gtgagagagg atgatccaac aaataatgta ccagattcga ttttctctaa actaggaatg 480
cagctacaca gaagagataa gcatccgatt ggtatcttaa aaaacgctat ctacgattac 540
tttgattcca attactcaaa caagtttgag aagttcgaag acctttcccc aattgttacc 600
acaaagcaag tacgttttca gtactcaagt ttgcatcttt ctagaagtat cacttggttt 660
```

```
      tcaatgtgat cattattggt ttttggtacc agaactttga tgatgtgcta gtccctgctg 720
      atcatgtaag cagaagtctt aatgacacgt actatgtaga ctcacaaact gttttgagat 780
      gtcatacgag tgctcaccaa gctgagctgt tgaggaaagg tcatagtcgt ttccttgtaa 840
      ccggggatgt ttaccgaaga gattctattg actctactca ttatccggtt ttccatcagg 900
  5   tgttctattc ttgaggtccc tgtgtttttc ttttactttg gctgttttgc tcgacaggtg 960
      tattatgttt tttatctatt acagatggaa ggttttgtg ttttctctcc tgaggactgg 1020
      aacgggtctg gcaaggattc cactttgtat gctgctgagg atttgaagaa atgtcttgag 1080
      ggattggcac gccacttatt tggtacatta agatccaata aacaatctat attcttcagc 1140
      aagtgtaaat aacttcaaag atggtttatt aagagttgtt taggatgatt atttcattaa 1200
 10   tttaagaaga agttggggaa atatacatga aataatttga tctgagcttc ttttttggc 1260
      tgccaggttc ggtggagatg agatgggttg atacatattt cccatttacc aatccatctt 1320
      ttgagcttga gatatatttt aaggtagtct atgagtcttt cgtttttcata tctttgcttt 1380
      aaagagacat ataatacttc tattttgtg tggtctcctt ttcccaaata catattggtg 1440
      ttattggata gaaattatag catctaacac aaacttcagt ttctcatcta acacaagctt 1500
 15   tatctggtta tcggttcagg aagactggtt ggaggttttg ggctgtgggg tgaccgagca 1560
      agtaattctg aaacaaagtg gattagaaaa taatgttgct tgggccttcg gacttggact 1620
      tgagagactt gctatggttt tgtttgacat acctgatata cgattttttct ggtcatccga 1680
      tgaacgattc acgtcccagg tgattacttg ggtgacaact tcaaatttta ggttatgaga 1740
      ctaatcgtct aaaaatatga attattttct cacattaatg tatttgatgt tacatgcagt 1800
 20   ttggaaaagg agaacttgga gtgaaattca agccatattc aaaggtaaaa cacttaatgt 1860
      ccatgtctcg tagactaatg aactctagtt gaagacttat ctgtattgta tgttaacga 1920
      tggcaaacaa aattttgttc tgcagtatcc tccttgttac aaggacatca gtttctggat 1980
      aagtgatttg ttcacagaga ataatttttg tgaagttgtt agaggaattg ctggggatct 2040
      tgttgaagag gtatcttata tcttgattgt ttgggagaga atattagctt ttacaggaat 2100
 25   caaatttact tttcagctat cctttaattg tatcatcata ttcttcagtg ttcctttgtt 2160
      tatagccttt attgcattca caggtgaagt taattgacca attcaccaat aagaagaaag 2220
      ggctgacgag tcattgttac agaatcgtgt tccgttccat ggagcggtct cttacggacg 2280
      aggaggtcaa tgatctgcag gtaatcactt ttgcttctcc tttcatcatt aacatgtaag 2340
      atattcaaaa ccgtttcata gacaaaatga aattttccaa atcgtgatag caagatagaa 2400
 30   ttggttgtgt atatgttgtt agtgttaaat gtgttgaaat ggtgacttat cgaaatgcag 2460
      agtaaagtgc gtgatgaggt gcaga                                       2485
```

35