EP0570422

**PCT**

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)


(51) International Patent Classification 5 :
C12N 15/82, 15/29, A01H 5/00
C12N 5/10

A1

(11) International Publication Number: WO 92/13957

(43) International Publication Date: 20 August 1992 (20.08.92)

(21) International Application Number: PCT/EP92/00275

(22) International Filing Date: 5 February 1992 (05.02.92)

(30) Priority data:
91400300.9 7 February 1991 (07.02.91) EP
*(34) Countries for which the regional or international application was filed:* AT et al.
91401787.6 28 June 1991 (28.06.91) EP
*(34) Countries for which the regional or international application was filed:* AT et al.



(71) Applicant *(for all designated States except US)*: PLANT GENETIC SYSTEMS, N.V. [BE/BE]; Kolonel Bourgstraat 106, B-1040 Brussels (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : DE BEUCKELEER, Marc [BE/BE]; HERDIES, Lydia [BE/BE]; GOSSELE, Véronique [BE/BE]; MARIANI, Celestina [IT/BE]; Jozef Plateaustraat 22, B-9000 Gent (BE).

(74) Agents: PLASSERAUD, Yves et al.; Ernest Gutmann-Yves Plasseraud S.A., 67, boulevard Haussmann, F-75008 Paris (FR).



(81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC (European patent), MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL (European patent), NO, PL, RO, RU, SD, SE (European patent), SN (OAPI patent), TD (OAPI patent), TG (OAPI patent), US.

**Published**
*With international search report.*


(54) Title: STAMEN-SPECIFIC PROMOTERS FROM CORN


(57) Abstract

Corn anther-specific promoters which are of particular utility in the production of transgenic male-sterile monocots and plants for restoring their fertility.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AT** | Austria | **FI** | Finland | **ML** | Mali |
| **AU** | Australia | **FR** | France | **MN** | Mongolia |
| **BB** | Barbados | **GA** | Gabon | **MR** | Mauritania |
| **BE** | Belgium | **GB** | United Kingdom | **MW** | Malawi |
| **BF** | Burkina Faso | **GN** | Guinea | **NL** | Netherlands |
| **BG** | Bulgaria | **GR** | Greece | **NO** | Norway |
| **BJ** | Benin | **HU** | Hungary | **PL** | Poland |
| **BR** | Brazil | **IE** | Ireland | **RO** | Romania |
| **CA** | Canada | **IT** | Italy | **RU** | Russian Federation |
| **CF** | Central African Republic | **JP** | Japan | **SD** | Sudan |
| **CG** | Congo | **KP** | Democratic People's Republic of Korea | **SE** | Sweden |
| **CH** | Switzerland | | | **SN** | Senegal |
| **CI** | Côte d'Ivoire | **KR** | Republic of Korea | **SU** | Soviet Union |
| **CM** | Cameroon | **LI** | Liechtenstein | **TD** | Chad |
| **CS** | Czechoslovakia | **LK** | Sri Lanka | **TG** | Togo |
| **DE** | Germany | **LU** | Luxembourg | **US** | United States of America |
| **DK** | Denmark | **MC** | Monaco | | |
| **ES** | Spain | **MG** | Madagascar | | |

# STAMEN-SPECIFIC PROMOTERS FROM CORN

This invention relates to promoters isolated from corn which can provide gene expression predominantly or specifically in stamen cells of a plant, particularly a monocotyledonous plant, and thereby provide little or no gene expression in other parts of the plant that are not involved in the production of fertile pollen. The promoters are useful in the production of transformed plants, in which a gene is to be expressed at least predominantly, and preferably specifically, in the stamen cells, preferably in the anther cells. The promoters are especially useful in the production of male-sterile plants and male fertility-restorer plants as described in European patent applications ("EPA") 89401194.9 and 90402281.1, respectively (which are incorporated herein by reference), particularly in the production of hybrids of monocotyledonous plants, such as corn, rice or wheat.

## Summary of the Invention

In accordance with this invention are provided: male flower-specific cDNA sequences isolated from corn comprising the sequences, SEQ ID no. 1 and SEQ ID no. 2, shown in the sequence listing. Also in accordance with this invention are provided stamen-specific, preferably anther-specific, promoters of the corn genes corresponding to such cDNA sequences, particularly the promoter which controls the expression of the genomic coding sequence corresponding to the cDNA of SEQ ID no. 2 and which is contained within the sequence of nucleotides 1 to 1179 of SEQ ID no. 3 (the "CA55 promoter" or "PCA55"). Each of such promoters can be used in a foreign DNA sequence, preferably a foreign chimaeric DNA sequence, which contains a structural gene, preferably a male-sterility DNA or a male fertility-restorer DNA, under the transcriptional control of the

promoter and which can be used to transform the nuclear genome of a cell of a plant, particularly a monocotyledonous plant. Further in accordance with this invention are provided: the male-sterile plant or male fertility-restorer plant which can be regenerated from such a cell transformed with the foreign DNA sequence of this invention; the transformed cell, itself; a culture of such a transformed cell; seeds of such a regenerated plant and its progeny; and a fertility-restored plant and its seeds resulting from crossing such male-sterile and male fertility-restorer plants.

Detailed Description of the Invention

In accordance with this invention, a male-sterile plant or a male fertility-restorer plant can be produced from a single cell of a plant by transforming the plant cell in a known manner to stably insert, into its nuclear genome, the foreign DNA sequence of this invention. The foreign DNA sequence comprises at least one male-sterility DNA or male fertility-restorer DNA that is: under the control of, and fused in frame at its upstream (i.e., 5') end to, one of the stamen-specific, preferably anther-specific, particularly tapetum-specific, promoters of this invention, such as the promoter and optionally the leader sequence of SEQ ID no. 3; and fused at its downstream (i.e., 3') end to suitable transcription termination (or regulation) signals, including a polyadenylation signal. Thereby, the RNA and/or protein or polypeptide, encoded by the male-sterility or male fertility-restorer DNA, is produced or overproduced at least predominantly, preferably exclusively, in stamen cells of the plant. The foreign DNA sequence can also comprise at least one marker DNA that: encodes a RNA and/or protein or polypeptide which, when present at least in a specific tissue or specific cells of the plant, renders the plant easily separable or

distinguishable from other plants which do not contain such RNA and/or protein or polypeptide at least in the specific tissue or specific cells; is under the control of, and is fused at its 5' end to, a second promoter which is capable of directing expression of the marker DNA at least in the specific tissue or specific cells; and is fused at its 3' end to suitable transcription termination signals, including a polyadenylation signal. The marker DNA is preferably in the same genetic locus as the male-sterility or male fertility-restorer DNA. This linkage between the male-sterility or male fertility-restorer DNA and the marker DNA guarantees, with a high degree of certainty, the joint segregation of both the male-sterility or male fertility-restorer DNA and the marker DNA into offspring of the plant regenerated from the transformed plant cell. However in some cases, such joint segregation is not desirable, and in such cases, the marker DNA should be in a different genetic locus from the male-sterility or male fertility-restorer DNA.

The male-sterility DNA of this invention can be any gene or gene fragment, whose expression product (RNA and/or protein or polypeptide) disturbs significantly the metabolism, functioning and/or development of stamen cells, preferably anther cells, and thus prevents the production of fertile pollen. Preferred male-sterility DNAs are described in EPA 89401194.9, for example those DNAs encoding: RNases such as RNase T1 or barnase; DNases such as endonucleases (e.g., EcoRI); proteases such as papain; enzymes which catalyse the synthesis of phytohormones (e.g., isopentenyl transferase or the gene products of gene 1 and gene 2 of the T-DNA of Agrobacterium; glucanases; lipases; lipid peroxidases; plant cell wall inhibitors; or toxins (e.g., the A-fragment of diphteria toxin or botulin). Other preferred examples of male-sterility DNAs

are antisense DNAs encoding RNAs complementary to genes, the products of which are essential for the normal development of fertile pollen. Further preferred examples of male-sterility DNAs encode ribozymes capable of cleaving specifically given target sequences of genes encoding products which are essential for the production of fertile pollen. Still other examples of male-sterility DNAs encode products which can render stamen cells, particularly anther cells - and not other parts of the plant - susceptible to specific diseases (e.g. fungi or virus infection) or stress conditions (e.g. herbicides).

The construction of a vector comprising a male-sterility DNA, such as a barnase-encoding DNA, under the control of a corn anther-specific promoter of this invention, is most conveniently effected in a bacterial host organism such as E. coli. However, depending on the nature of the male-sterility DNA and the specific configuration of the vector, problems can be encountered due to the expression of the male-sterility DNA in, and the concurrent decrease of viability of, the host organism. Such problems can be solved in a number of ways. For instance, the host organism can be provided, on the same or a different plasmid from that containing the male-sterility DNA or even on its chromosomal DNA, with another DNA sequence that prevents or inhibits significantly the effect of the expression of the male-sterility DNA in the host organism. Such an other DNA sequence can encode, for example: an antisense RNA so that the accumulation and translation of the male-sterility RNA is prevented; or a protein (e.g., barstar) which specifically inhibits the gene product of the male-sterility DNA (e.g., barnase; Hartley (1988) J. Mol. Biol. 202, 913). Alternatively, the male-sterility DNA can contain elements, such as a plant intron, which will only result in an active gene product in

a plant cell environment. Examples of introns that can be used for this purpose are introns of the transcriptional units of: the adh-1 gene of maize (Luehrsen and Walbot (1991) Mol. Gen. Genet. 225, 81; Mascarenhas et al (1990) Plant Mol. Biol. 15, 913), the shrunken-1 gene of maize (Vasil et al (1989) Plant Physiol. 91, 1575), the cat-1 gene of castor bean (Tanaka et al (1990) Nucleic Acids Research ("NAR") 18, 6767), the act-1 gene of rice (McElroy et al (1990) The Plant Cell 2, 163; PCT publication WO 91/09948) and the TA36 gene (intron shown in SEQ ID no. 4).

The male fertility-restorer DNA of this invention can be any gene or gene fragment, whose expression product (RNA and/or protein or polypeptide) inactivates, neutralizes, inhibits, blocks, offsets, overcomes or otherwise prevents the specific activity of the product of a male-sterility DNA in stamen cells, particularly in anther cells. Preferred male fertility-restorer DNAs are described in EPA 90402281.1, for example those DNAs encoding: barstar which is the inhibitor of barnase; EcoRI methylase which prevents the activity of EcoRI; or protease inhibitors (e.g. the inhibitors of papain). Other examples of male fertility-restorer DNAs are antisense DNAs encoding RNAs complementary to male-sterility DNAs. Further examples of male fertility-restorer DNAs encode ribozymes capable of cleaving specifically given target sequences of male-sterility DNAs.

The marker DNA of this invention can be any gene or gene fragment encoding an RNA and/or protein or polypeptide that allows plants, expressing the marker DNA, to be easily distinguished and separated from plants not expressing the marker DNA. Examples of the marker DNA are described in EPA 89401194.9, such as marker DNAs which encode proteins or polypeptides that: provide a distinguishable color to

plant cells, such as the A1 gene encoding dihydroquercetin-4-reductase (Meyer et al (1987) Nature 330, 677-678) and the glucuronidase gene (Jefferson et al (1988) Proc. Natl. Acad. Sci. USA ("PNAS") 83, 8447); provide a specific morphological characteristic to a plant such as dwarf growth or a different shape of the leaves; confer on a plant stress tolerance, such as is provided by the gene encoding superoxide dismutase as described in EPA 88402222.9; confer disease or pest resistance on a plant, such as is provided by a gene encoding a Bacillus thuringiensis endotoxin conferring insect resistance on a plant, as described in EPA 86300291.1; or confer on a plant a bacterial resistance, such as is provided by the bacterial peptide described in EPA 88401673.4. Preferred marker DNAs encode proteins or polypeptides inhibiting or neutralizing the activity of herbicides such as: the sfr gene and the sfrv gene encoding enzymes conferring resistance to glutamine synthetase inhibitors such as Bialaphos and phosphinotricine as described in EPA 87400544.0.

In order for the protein or polypeptide encoded by the marker DNA to function as intended, it is often preferred to have it produced in the plant cell as a precursor, in which the mature protein is linked at its N-terminal end to another polypeptide (a "targeting peptide") which will translocate the mature protein to a specific compartment such as the chloroplasts, the mitochondria, or the endoplasmic reticulum. Such targeting peptides and DNA sequences coding for them (the "targeting sequences") are well known. For example, if a marker DNA codes for a protein that confers tolerance or resistance to a herbicide or another selective agent that acts on chloroplast metabolism, such as the sfr (or bar) gene or the sfrv gene (European patent publication ("EP") 0,242,236), it may be

preferred that such gene also comprise a chloroplast targeting sequence such as that coding for the transit peptide of the small subunit of the enzyme 1,5-ribulose bisphosphate carboxylase (Krebbers et al (1988) Plant Mol. Biol. 11, 745; EPA 85402596.2), although other targeting sequences coding for other transit peptides, such as those listed by Von Heijne et al (1991) Plant Mol. Biol. Reporter 9, 104, can be used.

Each of the stamen-specific, preferably anther-specific, promoters of this invention, such as the CA55 promoter upstream from nucleotide 1180 in SEQ ID no. 3, which can be used to control the male-sterility DNA or the male fertility-restorer DNA, can be identified and isolated in a well known manner as described in EPA 89401194.9. In this regard, each of the SEQ ID no. 1 and no. 2 cDNAs of this invention can be used as a probe to identify (i.e., to hybridize to) the corresponding region of the corn genome (i.e., the region containing DNA coding for the stamen-specific mRNA, from which the cDNA was made). Then, the portion of the plant genome that is upstream (i.e., 5') from the DNA coding for such stamen-specific mRNA and that contains the promoter of this DNA can be identified. For instance, the cDNA of SEQ ID no. 2 can be used as a probe to identify and isolate a genomic clone from a genomic library of Zea mays, such as a lambda EMBL3 or EMBL4 Zea mays genomic library. In this way, a genomic DNA clone can be isolated and sequenced, such as the clone of SEQ ID no. 3. SEQ ID no. 3 contains a coding region which is homologous to the cDNA of SEQ ID no. 2, and upstream of this coding region is a promoter sequence, with a TATA box, which directs the anther-specific transcription of the coding region.

The second promoter, which controls the marker DNA, can also be selected and isolated in a well known manner, for example as described in EPA 89401194.9, so that the marker DNA is expressed either selectively in one or more specific tissues or cells or constitutively in the entire plant, as desired, depending on the nature of the RNA and/or protein or polypeptide encoded by the marker DNA.

In the foreign DNA sequence of this invention, 3' transcription termination signals or the "3' end" can be selected from among those which are capable of providing correct transcription termination and/or polyadenylation of mRNA in plant cells. The transcription termination signals can be the natural ones of the male-sterility or male fertility-restorer DNA, to be transcribed, or can be foreign or heterologous. Examples of heterologous 3' transcription termination signals are those of the octopine synthase gene (Gielen et al (1984) EMBO J. 3, 835-845) and of the T-DNA gene 7 (Velten and Schell (1985) NAR 13, 6981-6998). When the foreign DNA sequence of this invention comprises more than one structural gene (e.g., a male-sterility DNA or a fertility-restorer DNA and a marker DNA), it is preferred that the 3' ends of the structural genes be different.

In plants, especially in monocotyledonous plants, particularly cereals such as rice, corn and wheat, the expression in accordance with this invention of a marker DNA, as well as a male-sterility DNA or a fertility-restorer DNA, can be enhanced by the presence at one or more, preferably one, appropriate position(s) in the transcriptional unit of each foreign DNA sequence of this invention, of a suitable plant intron (Luehrsen and Walbot (1991) Mol. Gen. Genet. 225, 81; Mascarenhas et al (1990) Plant Mol. Biol. 15, 913; Vasil et al (1989) Plant Physiol.

91, 1575; Tanaka et al (1990) NAR 18, 6767; McElroy et al (1990) The Plant Cell 2, 163; PCT publication WO 91/09948). Preferably, each intron has a nucleotide sequence that: is recognizable by the cells of the plant species being transformed (for requirements of intron recognition by plants, see Goodall and Filipowicz (1989) Cell 58, 473; Hanley and Schuler (1988) NAR 16, 7159), is longer than about 70-73 bp (Goodall and Filipowicz (1990) Plant Mol. Biol. 14, 727), and is positioned close to the 5' end of the encoded mRNA, particularly in any untranslated leader sequence.

Cells of a plant can be transformed with the foreign DNA sequence of this invention in a conventional manner. Where the plant to be transformed is susceptible to Agrobacterium infection, it is preferred to use a vector, containing the foreign DNA sequence, which is a disarmed Ti-plasmid. The transformation can be carried out using procedures described, for example, in EP 0,116,718 and EP 0,270,822. Preferred Ti-plasmid vectors contain the foreign DNA sequence between the border sequences or at least located upstream of the right border sequence. Of course, other types of vectors can be used for transforming the plant cell, using procedures such as direct gene transfer (as described for example in EP 0,223,247), pollen mediated transformation (as described for example in EP 0,270,356, PCT publication WO/85/01856 and EP 0,275,069), in vitro protoplast transformation (as described for example in US patent 4,684,611), plant virus-mediated transformation (as described for example in EP 0,067,553 and US patent 4,407,956) and liposome-mediated transformation (as described for example in US patent 4,536,475).

Where the plant to be transformed is corn, recently developed transformation methods can be used such as the methods described for certain lines of corn by Fromm et al (1990) Bio/Technology 8, 833 and Gordon-Kamm et al (1990) The Plant Cell 2, 603.

Where the plant to be transformed is rice, recently developed transformation methods can be used such as the methods described for certain lines of rice by Shimamoto et al (1990) Nature 338, 274, Datta et al (1990) Bio/Technology 8, 736, Christou et al (1991) Bio/Technology 9, 957 and Lee et al (1991) PNAS 88, 6389.

Where the plant to be transformed is wheat, a method analogous to those described above for corn or rice can be used. Preferably for the transformation of a monocotyledonous plant, particularly a cereal such as rice, corn or wheat, a method of direct DNA transfer, such as a method of biolistic transformation or electroporation, is used. When using such a direct transfer method, it is preferred to minimize the DNA that is transferred so that essentially only the foreign DNA sequence of this invention, with its male-sterility DNA, fertility-restorer DNA and/or marker DNA, is integrated into the plant genome. In this regard, when a foreign DNA sequence of this invention is constructed and multiplied on a plasmid in a bacterial host organism, it is preferred that, prior to transformation of a plant with the foreign DNA sequence, plasmid sequences that are required for propagation in the bacterial host organism, such as an origin of replication, an antibiotic resistance gene for selection of the host organism, etc., be separated from the parts of the plasmid that contain the foreign DNA sequence.

The Examples, which follow, describe: the isolation and the characterization of the two corn cDNA sequences SEQ

ID no. 1 and no. 2 of this invention; their use for isolating the two stamen-specific promoters of this invention from the corn genome, such as the CA55 promoter upstream from nucleotide 1180 in SEQ ID no. 3; the construction of promoter cassettes for the fusion of the promoters with male-fertility and male fertility-restorer DNAs; the construction of plant transformation vectors from the promoter cassettes; as well as the transformation of corn, rice and tobacco with the resulting plant transformation vectors.

Unless stated otherwise in the Examples, all procedures for making and manipulating recombinant DNA were carried out by the standard procedures described in Maniatis et al, Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY (1982) and Sambrook et al, Molecular Cloning - A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, NY (1989). When making plasmid constructions, the orientation and integrity of cloned fragments were checked by means of restriction mapping and/or sequencing.

The sequence identification numbers referred to above and in the Examples are listed below.

Sequence Listing

SEQ ID no. 1: cDNA sequence of the CA444 gene.

SEQ ID no. 2: cDNA sequence of the CA455 gene.

SEQ ID no. 3: genomic DNA clone obtained from a Zea mays genomic library using the cDNA of SEQ ID no. 2 as a probe.

SEQ ID no. 4: intron from TA36 gene of Nicotiana tabacum linked to KpnI linkers

SEQ ID no. 5: sequence of plasmid pVE149.

Example 1

**Isolation and characterization of anther-specific cDNAs from corn**

For the cloning of cDNAs corresponding to genes which are expressed exclusively, or at least predominantly, in anthers of corn, a cDNA library was prepared from poly A$^+$ mRNA isolated from tassel spikelets from the publicly available corn line B73 bearing anthers in tetrad stage. By means of the Amersham cDNA Synthesis System Plus RPN 1256 Y/Z kit (Amersham International PLC, Buckinghamshire, England), cDNA was synthesized using reverse transcriptase and an oligo dT primer according to the directions set forth in the kit for its use.

The cDNAs were cloned in lambda gt10 vector, using the Amersham cDNA Cloning System - lambda gt10 - RPN1257 - kit, in accordance with the directions set forth in the kit for its use. From the cDNA library thus obtained (30,000 plaques), differential screening was performed with a labelled cDNA probe from corn B73 seedlings and with a labeled cDNA probe from corn B73 whole spikelets. 93 possible anther-specific cDNA clones were selected and again screened with labeled cDNA probes from corn B73 anthers, seedlings and ears. With the 66 remaining clones from these additional selections, a Southern analysis was performed with differential cDNA probes from anthers at tetrad stage and from opened tassels, silk and ears from the corn line B73. This led to the selection of 27 anther-specific clones which were subcloned in pGEM1 (Promega, Madison, Wisconsin, USA). Cross-hybridization between these subclones revealed the presence of at least 2 classes. Probes of some of these subclones were prepared and checked again for their specificity in Northern blots with 5 to 10

μg poly $A^+$ mRNA isolated from different corn B73 tissues (i.e., anthers, ears, silk, leaves, and spikelets at several stages). From this selection, two anther-specific clones, called "pCA444" and "pCA455", were identified. These clones were sequenced, and their sequences are shown in the sequence listing as SEQ ID no. 1 and SEQ ID no. 2, respectively. pCA455 was found to hybridize exclusively with mRNA from anthers in different stages of development. pCA444 was found to hybridize with mRNA from anthers and to hybridize very weakly with a mRNA of similar size from embryos.

The cDNA sequence of pCA444 reveals the presence of two open reading frames ("ORF") over a total of 323 and 376 nucleotides. The cDNA sequence of pCA455 reveals the presence of two ORFs over a total of 387 and 300 nucleotides.

**Example 2**

**Isolation of the anther-specific gene corresponding to the anther-specific cDNA clone, pCA444, of Example 1**

To isolate the genomic DNA clones carrying the regulatory sequences of the gene, CA444, corresponding to pCA444, two approaches are taken.

The first approach uses inverse polymerase chain reactions ("PCR") (Ochman et al (1989) in "PCR: Application & Protocols", Innis, M., Gelfand, D., Sninsky, J., and White, T., eds. Academic Press, New York) for the geometric amplification of the DNA sequences which flank, upstream and downstream, a chosen core region of the CA444 gene sequence corresponding to the sequence of pCA444. DNA digestions are carried out using conventional buffers and well known conditions. Fragments of a suitable size (less than 3 to 4 kb) for correct amplification and circularization are produced by using restriction enzymes

which do not cleave the chosen core region of the CA444 gene sequence and which are preliminarily identified by Southern hybridization. Circularization is performed with T4 DNA ligase in a dilute DNA concentration favoring monomeric circles (Collins and Weissman (1984) PNAS 81, 6812-6815). Three polymerase chain reactions are performed in parallel with three different oligonucleotide pairs under conventional conditions (Saiki et al (1985) Science 230, 1250-1354) using the Vent™ DNA polymerase (Catalog no. 254L - Biolabs New England, Beverly, MA 01915, U.S.A.) isolated from Thermococcus litoralis (Neuner et al (1990) Arch. Micobiol. 153, 205-207).

In one reaction, the flanking regions of the core region of CA444, from nucleotide 85 to nucleotide 358 of the corresponding cDNA sequence (SEQ ID no. 1), are amplified using the following pair of 22 and 20 oligonucleotides having the following respective sequences:

1) 5' CCG AGG ACC AGC AGG ACG AGG C 3' (nucleotide 64 to nucleotide 85 of pCA444 (SEQ ID no. 1)) and
2) 5' GGA TGG CAG GAG GGG AGA GG 3' (nucleotide 358 to nucleotide 377 of pCA444 (SEQ ID no. 1)).

In the second reaction, the flanking regions of the core region of CA444, from nucleotide 288 to nucleotide 392 of the corresponding cDNA sequence (SEQ ID no. 1), are amplified using the following pair of 20 and 23 oligonucleotides having the following respective sequences:

1) 5' GCA GGC TGT TGA TGA TGC CC 3' (nucleotide 269 to nucleotide 288 of pCA444 (SEQ ID no. 1)) and
2) 5' CCA TTT CAC AGT GAG AGC AGT CG 3' (nucleotide 392 to nucleotide 414 of pCA444 (SEQ ID no. 1)).

In the third reaction, the flanking regions of the core region of CA444, from nucleotide 43 to nucleotide 74 of the corresponding cDNA sequence (SEQ ID no. 1), are amplified using the following pair of 22 and 20 oligonucleotides having the following respective sequences:

1) 5' GGG GCG GTG GCT GCT TCT AGC G 3' (nucleotide 22 to nucleotide 43 of pCA444 (SEQ ID no. 1)) and
2) 5' GCT GGT CCT CGG CGG CGG CA 3' (nucleotide 74 to nucleotide 93 of pCA444 (SEQ ID no. 1)).

The second approach uses a lambda EMBL3 or EMBL4 Zea mays genomic library that is screened with the whole cDNA sequence of pCA444 as a probe. Corresponding genomic clones which hybridize to pCA444 are sequenced (Maxam and Gilbert (1977) PNAS 74, 560) and their orientation checked by Northern blot analysis with riboprobes of both senses. Comparison of the sequences of pCA444 with the genomic clone sequences leads to the identification of the homologous regions. At the 5' end of the region of each of these homologous genomic clones, the ATG codon and the consensus sequence TATA are determined. That the "TATA"-box is part of the promoter is confirmed by primer extension.

**Example 3**

**Isolation of the anther-specific gene corresponding to the anther-specific cDNA clone, pCA455, of Example 1**

To isolate the genomic DNA clones carrying the regulatory sequences of the gene, CA455, corresponding to pCA455, the two approaches of Example 2 are used.

In the first approach, inverse PCR (Ochman et al, 1989) is used for the geometric amplification of the DNA sequences which flank a chosen core region of the CA455 gene sequence corresponding to the sequence of pCA455. DNA digestion and circularization are carried out as in Example

2. Two polymerase chain reactions are performed in parallel with two different oligonucleotide pairs under conventional conditions (Saiki et al, 1985) using the Vent™ DNA polymerase isolated from T. litoralis (Neuner et al, 1990).

In one reaction, the flanking regions of the core region of CA455, from nucleotide 54 to nucleotide 87 of the corresponding cDNA sequence (SEQ ID no. 2), are amplified using the following pair of 21 and 23 oligonucleotides having the following respective sequences:

1) 5' GCT CGA TGT ATG CAG TGC AGC 3' (nucleotide 34 to nucleotide 54 of pCA455 (SEQ ID no. 2)) and
2) 5' CGT CGC CGT GTC GGT GCT TCT CG 3' (nucleotide 87 to nucleotide 109 of pCA455 (SEQ ID no. 2)).

In the second reaction, the flanking regions of the core region of CA455, from nucleotide 54 to nucleotide 557 of the corresponding cDNA sequence (SEQ ID no. 2), are amplified using the following pair of 21 and 24 oligonucleotides having the following respective sequences:

1) 5' GCT CGA TGT ATG CAG TGC AGC 3' (nucleotide 34 to nucleotide 54 of pCA455 (SEQ ID no. 2)) and
2) 5' CCG TTG CGT TGC GTT GCG TAG ACG 3' (nucleotide 557 to nucleotide 580 of pCA455 (SEQ ID no. 2)).

The second approach uses a lambda EMBL3 or EMBL4 Zea mays genomic library that is screened with the whole cDNA sequence of pCA455 as a probe. Corresponding genomic clones which hybridize to pCA455 are sequenced (Maxam and Gilbert, 1977), and their orientation is checked by Northern blot analysis with riboprobes of both senses. Comparison of pCA455 with the genomic clone sequences leads to the identification of the homologous regions. At the 5' end of the region of each of these homologous genomic clones, the

ATG codon and the consensus sequence TATA are determined. That the "TATA"-box is part of the promoter is confirmed by primer extension.

Using this second approach, an existing lambda EMBL4 Zea mays genomic library was screened with the whole cDNA sequence of pCA455, as a probe. The library was obtained from Dr. H. Saedler of the Max Planck Institute in Köln, Germany, with the designation "GH#1417". The library comprised Zea mays genomic DNA which was partially digested with MboI and the resulting restriction fragments of which were cloned between the BamHI sites of the bacteriophage lambda EMBL4 replacement vectors (Frischauff et al (1983) J. Mol. Biol. 170, 827; Pouwels et al (1988) Cloning vectors - a Laboratory Manual (supplementary update), Elsevier Science Publishers, Amsterdam). The restriction fragments of the library could be excised from the vectors as EcoRI fragments.

One EcoRI fragment of about 6 kb in length from the library was found to hybridize with pCA455 and was called "VG55". VG55 was found to contain an unique BamHI site, and one of the EcoRI-BamHI fragments of VG55 still hybridized with pCA455 while the other did not. The EcoRI-BamHI fragment that cross-hybridized with pCA455 was cloned between the EcoRI and BamHI sites of vector pGEM1 (Promega), yielding a plasmid called "pVG55.3". pVG55.3 was sequenced (Maxam and Gilbert, 1977), and its orientation was checked by Northern blot analysis with riboprobes of both senses. The sequence of pVG55.3, apart from some nucleotides at its 5' end (which includes its EcoRI site), is shown in SEQ ID no. 3 as having a high degree of homology with pCA455. The ATG codon of the presumed coding sequence of pVG55.3 is located at position 1180, the presumed coding sequence ends at position 1596, and the

"TATA"-box is located at position 1072. That the "TATA"-box is part of the promoter is confirmed by primer extension. The unique BamHI site, mentioned above, is located at position 2770 in SEQ ID no. 3.

The sequence upstream from position 1180 in SEQ ID no. 3 can be used as a promoter region for the anther-specific expression of a coding sequence of interest. This sequence is the CA55 promoter or PCA55. Preferably, the complete sequence from position 1 to position 1179 is used, but it appears that the minimum region which can serve as an anther-specific promoter extends about 300 to 500 bp upstream from position 1180 in SEQ ID no. 3. The use of the untranslated leader sequence in the PCA55 promoter region, between the transcription initiation site (which can be determined by means of primer extension) and the ATG start of translation, appears to be preferred but not essential for anther-specific expression of a heterologous structural gene under the control of the PCA55 promoter, and the leader sequence apparently can be replaced by the untranslated leader sequence of other genes, such as plant genes.

**Example 4**

**Construction of promoter cassettes derived from the anther-specific genes of Examples 2 and 3**

The 5' regulatory sequences, including the promoter, of each of the anther-specific genes of Examples 2 and 3 are subcloned into the polylinker of pMa5-8 and pMc5-8 (EPA 87402348.4). This produces vectors which can be used to isolate single stranded DNA for use in site-directed mutagenesis. Using site-directed mutagenesis (EPA 87402348.4), sequences surrounding the ATG translation initiation codon of the 5' regulatory sequences of each of the anther-specific genes are modified to create a unique

recognition site for a restriction enzyme for which there is a corresponding recognition site at the 5' end of each of the male-sterility and male fertility-restorer DNAs (that are to be fused to the 5' regulatory sequences in Example 5, below). The resulting plasmids each contain the newly created restriction site. The precise nucleotide sequence spanning each newly created restriction site is determined in order to confirm that it only differs from the 5' regulatory sequences of the corresponding corn anther-specific gene by the substitution, creating the new restriction site.

In using this procedure for constructing promoter cassettes, a NcoI site is introduced at the ATG translation initiation codon of pVG55.3 of Example 3 as follows. A 1280 bp EcoRI-AvaI fragment of pVG55.3 (the AvaI site is located at position 1276 of SEQ ID no. 3; the EcoRI site is derived from pGEM1 (Promega) and is located at the 5' end of SEQ ID no. 3 [not shown]) is cloned between the EcoRI and AvaI sites of the vectors pMa5-8 and pMc5-8 (Stanssens et al (1989) NAR 17, 4441; EPA 87402348.4), yielding plasmids called "pMa5-VG55.3" and "pMc5-VG55.3", respectively. These plasmids are used for site-directed mutagenesis by a gapped duplex DNA method using alternating selectable markers as described by Stanssens et al (1989) supra. The gapped duplex DNA is constructed from the single stranded pMc5-VG55.3 and the large EcoRI-AvaI fragment of pMa5-VG55.3. For mutagenesis, use is made of the oligonucleotide with the following sequence:

CAG GAG CGA GCC ATG GCT GCA G.

This mutagenesis introduces the NcoI site at the ATG codon of the coding sequence of SEQ ID no. 3. The resulting cassette comprises the promoter and leader sequence of SEQ ID no. 3 in a EcoRI-NcoI fragment that is to be fused to

the coding sequences of male-sterility and male fertility-restorer DNAs as described in Example 5, below.

Alternatively, the NcoI site is introduced at the ATG translation initiation codon of pVG55.3 during amplification by PCR of a DNA fragment, containing the CA55 promoter region, using the following two oligonucleotides as primers :

5'-GAT TCG AAT TCT GGT ATG CAT CAA TAG AGC CG-3'

5'-CAG GAG CGA GCC ATG GCT GCA G-3'

The amplified DNA fragment is used directly as a promoter cassette for constructing plant transformation vectors as described in Example 5.

Example 5

**Construction of plant transformation vectors from the promoter cassettes of Example 4**

Using the procedures described in EPA 89401194.9 and 90402281.1, the promoter cassettes of Example 4 are used to construct plant transformation vectors comprising foreign chimaeric DNA sequences of this invention, each of which contains the 5' regulatory sequences, including the anther-specific promoter, of one of the anther-specific genes isolated in Example 2 or 3. Each of these 5' regulatory sequences is upstream of, is in the same transcriptional unit as, and controls either a male-sterility DNA (from EPA 89401194.9) encoding barnase from Bacillus amyloliquefaciens (Hartley and Rogerson (1972) Preparative Biochemistry 2 (3), 243-250) or a male fertility-restorer DNA (from EPA 90402281.1) encoding barstar (Hartley and Rogerson (1972) supra; Hartley and Sweaton (1973) J. Biol. Chem. 248 (16), 5624-5626). Downstream of each male-sterility or male fertility-restorer DNA is the 3' end of the nopaline synthase gene

(An et al (1985) EMBO J. 4 (2), 277). Each chimaeric DNA sequence also comprises the 35 S'3 promoter (Hull and Howell (1987) Virology 86, 482-493) fused in frame with the neo gene encoding kanamycin resistance (EPA 84900782.8) and the 3' end of the octopine synthase gene (Dhaese et al (1983) EMBO J. 2, 419).

Alternatively, the plant transformation vectors pVE149, pVE139 and pVE136 are constructed as follows.

In a first step, the 1083 bp EcoRI-HindIII DNA fragment of pMT416, containing the barnase and barstar coding sequences (Hartley (1988) J.Mol.Biol. 202, 913), is ligated to the large EcoRI-HindIII fragment of plasmid pMa5-8, yielding plasmid pMa5tpbs1. By means of site-directed mutagenesis (PCT publication WO 89/03887), a NcoI site is then introduced at the ATG translation initiation codon of the barnase coding sequence. For this purpose, a gapped duplex DNA is constructed from the single stranded pMa5tpbs1, the large EcoRI-HindIII fragment of pMc5-8, and the following oligonucleotide:

5'-GAT AAC CGG TAC CAT GGT TGT CAC AGG GG-3'.

The resulting plasmid is designated as "pVE145A".

In a subsequent mutagenesis round, a NsiI site is introduced 14 bp downstream of the ATG translation initiation codon of the barnase coding sequence using a gapped duplex DNA consisting of single stranded DNA from pVE145A, the large EcoRI-HindIII fragment of pMa5-8, and the following oligonucleotide:

5'-CCC CGT CAA ATG CAT TGA TAA CCG G-3'.

The resulting plasmid is designated as "pVE145".

Plasmids pMa5-8 and pMc5-8 have been deposited on May 3, 1988 at the Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSM), Mascheroderweg 1B, D-330 Braunschweig,

Germany under accession numbers DSM 4567 and DSM 4566 respectively.

pVE145 is cut with NsiI, filled in with Klenow, and ligated to the 111 bp DNA fragment shown in SEQ ID no. 4 which contains the TA36 intron and which has been cleaved with KpnI and made blunt-ended with Klenow, yielding plasmid pVE146. The fragment of SEQ ID no. 4 is obtained by amplification, from genomic DNA of Nicotiana tabacum cv. Samsun, by means of PCR using the following two oligonucleotides as primers :

5'-CGA CGG TAC CAC GTA ATT AG-3'

5'-CAT AGG GTA CCT GTA TGT AAT AAA AAC-3'.

Plasmid pVE147 is constructed by ligation of the 2820 bp EcoRI-HindIII fragment of pGEM1 (Promega), the 979 bp HindIII-NcoI fragment of pVE146 (carrying the barnase gene with intron), and the 1184 bp PCR fragment of Example 4 carrying the CA55 anther-specific promoter from corn.

Finally, pVE149 is obtained by ligation of the following four DNA fragments:

- the 1715 bp EcoRI (filled-in with Klenow) - XbaI fragment of pVE147, carrying the barnase gene under control of the CA55 promoter,
- a 296 bp EcoRI-XbaI fragment of pTTM6 (deposited on March 7, 1988 at the DSM under DSM accession number 4468), carrying the 3' untranslated end of the nopaline synthase gene of Agrobacterium T-DNA,
- a 1728 bp BglII (filled-in with Klenow)-HindIII fragment, carrying the bar gene (EP 0,242,236) under the control of the 35S3 promoter (EP 0,359,617) and with a 3' untranslated end of the nopaline synthase gene of Agrobacterium T-DNA (this fragment corresponds to the sequence in SEQ ID no. 5 between positions 2409 and 4137), and

- the large EcoRI-HindIII fragment of pUC19 (New England Biolabs Inc. Beverly, MA, U.S.A.).

The complete sequence of pVE149 is shown in SEQ ID no. 5.

Plasmid pVE136 is identical to pVE149 except it lacks the TA36 intron in the barnase gene. pVE136 is constructed by replacing, in pVE149, the 534 bp NcoI-BamHI fragment, carrying the barnase gene, with the 449 bp NcoI-BamHI fragment of pVE145A.

pVE136 is constructed and maintained in E. coli WK6 containing the plasmid pMc5-BS. pMc5-BS contains the barstar gene under the control of the tac promoter (De Boer et al (1983) PNAS 80, 21) and is constructed by cloning the EcoRI-HindIII fragment of pMT416 (Hartley (1988) J.Mol.Biol. 202, 913) into pMc5-8. Then, the sequence, starting with the PhoA signal sequence and ending with the last nucleotide before the translation initiation codon of the barstar coding region, is deleted by looping-out mutagenesis according to the general procedures described by Sollazi et al (1985) Gene 37, 199. The availability of an ampicillin resistance gene on the pUC18-derived plasmids carrying the chimaeric barnase gene and the chloramphenicol resistance gene on pMc5-BS permits the strain to be kept stable on plates provided with two antibiotics or to select for any one plasmid. While normally repressed, gene expression from this promoter can be induced by addition of a commonly used inducer of the lac operon, IPTG (isopropyl-β-d- thiogalactopyranoside).

The 5843 bp NcoI-BamHI fragment of partially digested pVE149, carrying all of the plasmid except the barnase coding sequence, is filled in with Klenow and ligated to a DraI-HindIII fragment (filled- in with Klenow) of pVE151, carrying the barstar coding sequence. The resulting plasmid is designated as "pVE139". pVE151 is obtained by means of

site-directed mutagenesis of pMc5-BS, so that a DraI site is introduced at the ATG translation initiation codon of the barstar coding sequence. For this purpose, a gapped duplex DNA is constructed from the single stranded pMc5-BS, the large EcoRI-HindIII fragment of pMa5-8, and the following oligonucleotide:

5'-GCT TTT TTA AAT TTA TTT TCT CC-3'.

T-DNA vectors for Agrobacterium-mediated plant transformations are prepared by cloning the appropriate EcoRI (filled-in with Klenow)-HindIII fragments of pVE149 or pVE136 (containing the 35S3-bar and corn anther-specific promoter-barnase chimaeric genes) or pVE139 (containing the 35S3-bar and corn anther-specific promoter-barstar chimaeric genes) between the HindIII and XbaI (filled-in with Klenow) sites of the known T-DNA vectors pGSC1700 or pGSC1701A. pGSC1700 has been deposited on March 21, 1988 at the DSM under DSM accession number 4469, and pGSC1701A has been deposited on October 22, 1987 at the DSM under DSM accession number 4286. The T-DNA vectors containing pVE149, pVE139 and pVE136 are used for transformation of tobacco as described in Example 7.

**Example 6**

**Transformation of corn with the plant transformation vectors from Example 5.**

Using the procedures described by Fromm et al (1990) supra, embryogenic suspension cultures of a B73 X A188 corn line are transformed with the plant transformation vectors described in Example 5, including pVE149, pVE136 and pVE139 -- either directly or after suitable linearization (e.g., after digestion with EcoRI and/or HindIII). Transformed plants regenerated from the embryogenic suspension cultures, each containing an anther-specific promoter of Example 2 or 3 controlling either a male-sterility DNA or a

male fertility-restorer DNA, are normal except for their flowers. In this regard, each plant containing a male-sterility DNA under the control of one of the anther-specific promoters expresses such DNA at least predominantly in its anthers and produces no normal pollen, and each plant containing a male fertility-restorer DNA under the control of one of the anther-specific promoters expresses such DNA at least predominantly in its anthers but produces normal pollen.

Example 7

**Transformation of tobacco with the plant transformation vectors from Example 5**

Using the procedures described in EPA 89401194.9 and 90402281.1, tobacco plants are transformed by Agrobacterium-mediated transfer with the plant transformation vectors containing the foreign chimaeric DNA sequences from Example 5. The transformed tobacco plants, each containing an anther-specific promoter of Example 2 or 3 controlling either a male-sterility DNA or a male fertility-restorer DNA, are normal except for their flowers. In this regard, each plant containing a male-sterility DNA under the control of one of the anther-specific promoters expresses such DNA at least predominantly in its anthers and produces no normal pollen, and each plant containing a male fertility-restorer DNA under the control of one of the anther-specific promoters expresses such DNA at least predominantly in its anthers but produces normal pollen.

Example 8

**Transformation of rice with the plant transformation vectors from Example 5**

Using the procedures described by Datta et al (1990) supra, protoplasts of the rice line, Oryza sativa var.

Chinsurah Boro II, are transformed with the plant transformation vectors described in Example 5, including pVE149, pVE136 and pVE139 -- either directly or after suitable linearization (e.g., after digestion with EcoRI and/or HindIII). Transformed plants regenerated from the protoplasts, each containing an anther-specific promoter of Example 2 or 3 controlling either a male-sterility DNA or a male fertility-restorer DNA, are normal except for their flowers. In this regard, each plant containing a male-sterility DNA under the control of the anther-specific promoters expresses such DNA at least predominantly in its anthers and produces no normal pollen, and each plant containing a male fertility-restorer DNA under the control of the anther-specific promoter expresses such DNA at least predominantly in its anthers but produces normal pollen. Alternatively, immature embryos from rice varieties Gulfmont, Lemont, IR26, IR36, IR54, or IR72 are bombarded with gold particles, carrying appropriate plasmid DNA of Examples 5, and plants are regenerated according to the procedures described by Christou et al (1991) Bio/Technology 9, 957.

Needless to say, the use of the anther-specific corn promoters of this invention is not limited to the transformation of any specific plant(s). Such corn promoters can be useful in any crop where they are capable of controlling gene expression, and preferably where such expression occurs at least predominantly, preferably specifically, in stamen cells of the crop. Also, the use of such promoters is not limited to the control of male-sterility DNAs or male fertility-restorer DNAs but can be used to control the expression of any gene selectively in stamen cells.

Furthermore, this invention is not limited to the specific stamen-specific, preferably anther-specific, particularly tapetum-specific, promoters described in the foregoing Examples. Rather, this invention encompasses promoters equivalent to those of Examples 2 and 3 which can be used to control the expression of a structural gene, such as a male-sterility DNA or a male fertility-restorer DNA, selectively in stamen cells, preferably anther cells, particularly tapetum cells, of a plant. Indeed it is believed that the DNA sequences of the promoters of Examples 2 and 3 can be modified by replacing some of their nucleotides with other nucleotides, provided that such modifications do not alter substantially the ability of polymerase complexes, including transcription activators, of stamen cells, particularly anther cells, to recognize the promoters, as modified.

SEQUENCE LISTING

1. General Information

i) APPLICANT : PLANT GENETIC SYSTEMS N.V.

ii) TITLE OF INVENTION : stamen-specific promoters from corn

iii) NUMBER OF SEQUENCES : 5
SEQ. ID. NO 1 : cDNA CA444
SEQ. ID. NO 2 : cDNA CA455
SEQ. ID. NO 3 : genomic sequence from corn comprising the CA55 promoter
SEQ. ID. NO 4 : TA36 intron
SEQ. ID. NO 5 : plasmid pVE149

iv) CORRESPONDENCE ADDRESS :
A. ADDRESSEE : Plant Genetic Systems N.V.
B. STREET : Plateaustraat 22,
C. POSTAL CODE AND CITY : 9000 Ghent,
D. COUNTRY : Belgium

v) COMPUTER READABLE FORM :
A. MEDIUM TYPE 5.25 inch, double sided, high density 1.2 Mb floppy disk
B. COMPUTER : IBM PC/AT
C. OPERATING SYSTEM : DOS version 3.3
D. SOFTWARE : WordPerfect 5.1

vi) CURRENT APPLICATION DATA : Not Available

(vii) PRIOR APPLICATION DATA :
EPA 91400300.9, filed February 7, 1991
EPA 91401787.6, filed June 28, 1991

2. Sequence Description : SEQ ID NO. 1

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 533

STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA

ORIGINAL SOURCE
ORGANISM: Corn
ORGAN: anther

FEATURES: from nucleotide 2 TO nucleotide 376 : Open Reading Frame 1
from nucleotide 3 TO nucleotide 326 : Open Reading Frame 2

PROPERTIES: anther specific cDNA

MISCELLANEOUS : cDNA designated as CA444

```
GGATCCTCGA GGCAACAATG GCGCTAGAAG CAGCCACCGC CCCCCGCGCA      50
CTCCTCGCCG CGTGCCTCGT CCTGCTGGTC CTCGGCGGCG GCACCGGCCC     100
GTCGTCGGTG CTGCGCGGCG CCGGGGCGCA GGCCGGCGGG CAGTGCCTGC     150
CGCAGCTGAA CCGCCTCCTG GCGTGCCGCG CGTACCTGGT GCCCGGCGCG     200
CCGGACCCCA GCGCGGACTG CTGCAGCGCG CTGAGCGCCG TGTCGCACGA     250
GTGCGCCTGC AGCACCATGG GCATCATCAA CAGCCTGCCC GGCCGGTGCC     300
ACCTCGCCCA AGCCAACTGC TCCGCTTGAA GCAGGGACCT GGCACGCGTG     350
CTGCAATGGA TGGCAGGAGG GGAGAGGAAT AAGAAGTGTT TCCATTTCAC     400
AGTGAGAGCA GTCGAGCTCC AACGTTGTCG TCGTCGTCGT CTTCTTCTTT     450
TGATATTCAG ACTCTGTCTT GCGGTCTATA TCATCAGCAT AATAATAATA     500
AAATAAGTAA AACCAAAAAA AAAAAAAAAC CAT                       533
```

3. Sequence Description : SEQ ID NO. 2

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 796

STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULAR TYPE: cDNA to mRNA

ORIGINAL SOURCE
ORGANISM: corn
ORGAN: anther

FEATURES: from nucleotide 2 TO nucleotide 388 : Open Reading Frame 1
from nucleotide 70 TO nucleotide 369 : Open Reading Frame 2

PROPERTIES: anther specific cDNA

MISCELLANEOUS : cDNA designated as CA455

```
CCATGGTACC CGGATCCTCG CCAAAACGCA GAAGCTGCAC TGCATACATC      50
GAGCTAACTA TCTGCAGCGA TGTCTCGCTC CTGCTGCGTC GCCGTGTCGG     100
TGCTTCTCGC TGTCGCCGCG ACAGCCAGCG CCACCGCGCC GGCATGGCTG     150
CACGAGGAGG CCATGGCCAC GGGCCCGCTG GTCGCAGAGG GTGCAAGGGT     200
GGCGCCCTCC GCGTCCACCT GGGCTGCCGA CAAGGCGTCG CCGGCGAGGC     250
CGAGCGGCGG CATGGCCACG CAGGGCGACG ACCAGAGCTC GTCGGGCGGC     300
AGTGGCAGCA GCGGTGAGCA CGGCAAGGAG GAGGGCGAGA AGCAGGGCAA     350
GAGCTGCCTC ACCAAGGAGG AGTGCCACAA GAAGAAGATG ATCTGCGGCA     400
AGGGCTGCAC GCTCTCGGCG CACAGCAAGT GCGCCGCCAA GTGCACCAAG     450
TCCTGTGTCC CCACCTGCTA GGAGCCGAGG CCGGAGCTTG CCGGCGGCGA     500
GACCTCGATC GATCGAGTGC TTCACTTCAC TTCTTTGTTA TAGTTCTTGT     550
GTGTTGCCGT TGCGTTGCGT TGCGTAGACG AAGGGAATAA GGAAGGGTAA     600
TTGGATTACC TGTTCCAGAT CTCTGTGTAA GCGTGTTGTC GTGACAAGTC     650
ATTTGATCCA GAGCGAGCGA TGGATAGATC GCGCTCGCAG TTTTAATTGC     700
AATGCTAGTT CAATAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA     750
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA CCATGGTACC CGGATC        796
```

4. Sequence Description : SEQ ID NO. 3

SEQUENCE TYPE: nucleic acid
SEQUENCE LENGTH: 2784 basepairs

STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULAR TYPE: genomic DNA

ORIGINAL SOURCE: corn
ORGANISM: anther

FEATURES: - nucleotide 1 to nucleotide 1179 : region comprising the promoter and the leader sequence
- nucleotide 1072 : TATA Box
- nucleotide 1180 to nucleotide 1596 : presumed coding sequence
- nucleotide 1120 to nucleotide 11839 : region corresponding to cDNA of SEQ ID. 2

PROPERTIES: genomic DNA obtained by probing a Zea mays genomic library using the cDNA of SEQ ID 2 as a probe. The genomic DNA is designated as pVG55.3.

```
TGGTATGCAT CAATAGAGCC GGAAGATGGT CTGGAGTAAG GACCTGGCAG      50
TGTGATACGG GAACTTGACA TCTGAATAGA TATTCTCCCT TGTCCCTCTG     100
GTAAAAAAAA CTGTTGTCAC ATTTGCCTTC GCTGTGACTT GGATGTATCA     150
TGTATATCTT TGACCATTGA TATCTTGGTT AATCAGACGG TGCATTACAA     200
TCATGGCCTC ATTCATATAG GGTTTAGGGT TACCACGATT GGTTTGCATA     250
AGTAGTACCC CTCCGTTTCA AATTATGTCG TATTTTGATT TTTTAGATAC     300
ACTTTTTATA TAATTTTTTA TTTTAAATTA GGTGTTTTAT ATAATACGTA     350
TCTAAGTGTA TAATAAAATA TATGTATCTA AAAGCTGTAA TTTAGTATAA     400
ATTAGAATGG TGTATATCTT CAATGTATGA CAAATAATTT GAAATGGAGG     450
AGGGTATGAA AAGCCAAAAC CTCCTAGAAT ATGGAATGGA GGGAATACAT     500
ACAAATTCTT TGCTTCAGTT AAAAGAAACG AGAAAAGGAG GGGAATGGGG     550
AATCGTACTT CAGTTTTTAC GAGTTTTCAT CAAACATGTA TGCACGTCTT     600
CCCTTGGTTG ATGCATCTTT TTGGCAAATC TTCGTTTAAT TGCGGCTTCT     650
TTTTTATACC GTTCGAAGGT TTTCGTCGTC AATGCTGAAA CTCCACTTTC     700
ACCACCTTCG GTTGCATCTG CTTGCTTTCA ATTCACCTCT AATTAGTCCA     750
AGTGTTTCAT TGGACGAAGG TCCAAGTCCT TCAGATCATC TCAATTTTCT     800
TTGATCTGAA ACAACAATTT AAAACTGATT TTGTTACCTT GACCTGTCGA     850
AGACCTTCGA ACGAACGGTA CTGTAAAAAT ACTGTACCTC AGATTTGTGA     900
TTTCAATTCG ATTCGGGTCT CCTGGCTGGA TGAAACCAAT GCGAGAGAAG     950
AAGAAAAAAT GTTGCATTAC GCTCACTCGA TCGGTTACGA GCACGTAGTT    1000
GGCGCCTGTC ACCCAACCAA ACCAGTAGTT GAGGCACGCC CTGTTTGCTC    1050
ACGATCACGA ACGTACAGCA CTATAAAACA CGCAGGGACT GGAAAGCGAG    1100
ATTTCACAGC TCAAAGCAGC CAAAACGCAG AAGCTGCACT GCATATACAG    1150

AAGATACATC GAGCTAACTA GCTGCAGCG ATG TCT CGC TCC TGC       1194
                                MET Ser Arg Ser Cys
                                  1               5
```

```
TGC GTC GCC GTG TCG GTG CTT CTC GCT GTC GCC GCG ACA          1233
Cys Val Ala Val Ser Val Leu Leu Ala Val Ala Ala Thr
                 10                  15

GCC AGC GCC ACC GCG CCG GCA TGG CTG CAC GAG GAG CAG          1272
Ala Ser Ala Thr Ala Pro Ala Trp Leu His Glu Glu Gln
     20                  25                  30

CAC CTC GAG GAG GCC ATG GCC ACG GGC CCG CTG GTC GCA          1311
His Leu Glu Glu Ala MET Ala Thr Gly Pro Leu Val Ala
            35                  40

GAG GGT GCG AGG GTG GCG CCC TCC GCG TCC ACC TGG GCT          1350
Glu Gly Ala Arg Val Ala Pro Ser Ala Ser Thr Trp Ala
 45                  50                  55

GCC GAC AAG GCG TCG CCG GCG AGG CCG AGC GGC GGC ATG          1389
Ala Asp Lys Ala Ser Pro Ala Arg Pro Ser Gly Gly MET
          60                  65                  70

GCC ACG CAG GGC GAC GAC CAG AGC TCG TCG GGC GGC AGT          1428
Ala Thr Gln Gly Asp Asp Gln Ser Ser Ser Gly Gly Ser
                 75                  80

GGC AGC AGC GGT GAG CAC GGC AAG GCG GAG GGC GAG AAG          1467
Gly Ser Ser Gly Glu His Gly Lys Ala Glu Gly Glu Lys
     85                  90                  95

CAG GGC AAG AGC TGC CTC ACC AAG GAG GAG TGC CAC AAG          1506
Gln Gly Lys Ser Cys Leu Thr Lys Glu Glu Cys His Lys
            100                 105

AAG AAG ATG ATC TGT GGC AAG GGC TGC ACG CTC TCG GCG          1545
Lys Lys MET Ile Cys Gly Lys Gly Cys Thr Leu Ser Ala
110                 115                 120

CAC AGC AAG TGC GCC GCC AAG TGC ACC AAG TCC TGT GTC          1584
His Ser Lys Cys Ala Ala Lys Cys Thr Lys Ser Cys Val
        125                 130                 135

CCC ACC TGC TAGGAGCCGA GGCCGGAGCT TGCCGGCGGC GAGACCTCGA      1633
Pro Thr Cys

TCGATCGAGT GCTTCACTTC ACTTCTTTGT TATAGTTCTT GTGTGTTGCC      1683
GTTGCGTTGC GTTGCGTAGA CGAAGGGAAT AAGGAAGGGT AATTGGATTA      1733
CCTGTTCCAG ATCTCTGTGT AAGCGTGTTG TCGTGACAAG TCTTTTGATC      1783
CAGAGCGAGG GATGGATAGA TCGCGCTCGC AGTTTTAATT GCAATGCTAG      1833
TTCAATATGT GTGCATCATG TTGGCAACTA CATAGTCCAG ATTCAAACCG      1883
AGATCGCTGT TTAGCATGCC AGCACAATAA TAACGGTACA ATCATATTAT      1933
ATTTTATACA AATGCACAAT TTATCTCTAG AGATGTCAAT GGGAAATTCC      1983
TCATCGGGTT ATATCATCTC AGACTCATCC CCATCATATT TGATTCATCC      2033
TCATACTCAT CCTCATATCT ATCATGAGTG CAAAACTCAT TTCATACCCA      2083
```

```
TCTCTATTTT GGTTTAGGGT CTCCATCCCT AATTAAGGGA TAACTAGTAC      2133
TAACAACTAG CACAAACTAT CTAGATTTCA GATATCACCA CATTGACAAA      2183
CAATCATCCA TGAACTATGA TCCATTCATC CATCCATCAA AAAATAAATC      2233
GGTATTTCGA GAACGATAGA AGAAATGAAG TCGGCTCACC TTTCTTGGTC      2283
ACCATTTGAG TTTGTTGGTG CCTGAGAATC CATGGTCGTC ATCGTCGTCC      2333
TAGGGATCGG CGGTGCTCCT CGTTGTTGGT AAAGTCGCCA GTGTGTAGTG      2383
CTAGCGCAAC TGTCCAGGCG TGCAACGGTT GGCCGGCTGG AAAGGGCATA      2433
GCGTATGGCT GGTTATTTTT AGGGTTTTGT TTTTTTACTA ATCTGCTAGT      2483
TGCCTTGCCA TGTTGTCTTA TTGGGCTAGG ATCTAGGGCT TGTTACGCTG      2533
CTGTGTTGGG CTTGGTGTCC GGTTCAGCCT CAACTCATTC ATACAAATCA      2583
GATTCATACA AAACAGGTAT ACACGTATGA AATATCCATG GATAATCAGG      2633
TTCGAATTAT TGTCCCCTAA ACCCATACAC GTTTACCCAA TGGATGGATA      2683
TTTTGTCTCA TATCCATACA CATGAGACGA TTTTTGTCCC ATACCTGTGC      2733
TCTAATAGGA GAATTTCTCT CGGGATAGCG AGTATCGGAT CCTCTAGAGT      2783
C                                                           2784
```

5. Sequence Description : SEQ ID NO. 4

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 111 bp

STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULAR TYPE: genomic DNA amplified by means of PCR

ORIGINAL SOURCE
ORGANISM: Nicotiana tabacum cv. Samsun

FEATURES: from nt 1 TO nt 7 : spacer sequence
from nt 8 TO nt 13 : KpnI site
from nt 16 TO nt 100 : intron sequence
from nt 101 TO nt 106 : KpnI site
from nt 107 TO nt 111 : spacer sequence

PROPERTIES: intron of TA36 gene with KpnI linkers at 5' and 3' end

```
CTACGACGGT ACCACGTAAT TAGTTTATAC CTTTAAACTT TAATTTTCAA      50
CCGATTTTGT GTCGTCTGCT TAATCTAACT TATTGTTTTT ATTACATACA     100
GGTACCCTAT G                                               111
```

6. Sequence Description : SEQ ID NO. 5

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 6376 bp

STRANDEDNESS: double-stranded
TOPOLOGY: circular
MOLECULAR TYPE: plasmid DNA

FEATURES:
- nt 1 to nt 396 : pUC18 derived DNA
- nt 397 to nt 802 : 3' untranslated end derived from the nopaline synthase gene of Agrobacterium T-DNA
- nt 803 to nt 1223 ; barnase coding sequence with TA36 intron
- nt 1119 to nt 1203 : TA36 intron
- nt 1224 to nt 2408 : CA55 promoter from corn
- nt 2409 to nt 3272 : 35S3 promoter
- nt 3273 to nt 3824 : coding sequence of bar gene
- nt 3825 to nt 4137 : 3' untranslated end derived from the nopaline synthase gene of Agrobacterium T-DNA
- nt 4138 to nt 6376 : pUC18 derived DNA

PROPERTIES: plasmid DNA replicable in E. coli designated as pVE149

```
TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG      50
GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG     100
TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG     150
CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA     200
CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC ATTCGCCATT     250
CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT     300
TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA     350
ACGCCAGGGT TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT     400
CGAGCTCGGT ACCCGGGGAT CTTCCCGATC TAGTAACATA GATGACACCG     450
CGCGCGATAA TTTATCCTAG TTTGCGCGCT ATATTTTGTT TTCTATCGCG     500
TATTAAATGT ATAATTGCGG GACTCTAATC ATAAAAACCC ATCTCATAAA     550
TAACGTCATG CATTACATGT TAATTATTAC ATGCTTAACG TAATTCAACA     600
GAAATTATAT GATAATCATC GCAAGACCGG CAACAGGATT CAATCTTAAG     650
AAACTTTATT GCCAAATGTT TGAACGATCT GCTTCGGATC CTCTAGAGXX     700
XXCCGGAAAG TGAAATTGAC CGATCAGAGT TTGAAGAAAA ATTTATTACA     750
CACTTTATGT AAAGCTGAAA AAAACGGCCT CCGCAGGAAG CCGTTTTTTT     800
CGTTATCTGA TTTTTGTAAA GGTCTGATAA TGGTCCGTTG TTTTGTAAAT     850
CAGCCAGTCG CTTGAGTAAA GAATCCGGTC TGAATTTCTG AAGCCTGATG     900
TATAGTTAAT ATCCGCTTCA CGCCATGTTC GTCCGCTTTT GCCCGGGAGT     950
TTGCCTTCCC TGTTTGAGAA GATGTCTCCG CCGATGCTTT TCCCCGGAGC    1000
GACGTCTGCA AGGTTCCCTT TTGATGCCAC CCAGCCGAGG GCTTGTGCTT    1050
CTGATTTTGT AATGTAATTA TCAGGTAGCT TATGATATGT CTGAAGATAA    1100
TCCGCAACCC CGTCAAACTG TATGTAATAA AAACAATAAG TTAGATTAAG    1150
CAGACGACAC AAAATCGGTT GAAAATTAAA GTTTAAAGGT ATAAACTAAT    1200
TACGTGTTGA TAACCGGTAC CATGGCTGCA GCTAGTTAGC TCGATGTATC    1250
TTCTGTATAT GCAGTGCAGC TTCTGCGTTT TGGCTGCTTT GAGCTGTGAA    1300
```

```
ATCTCGCTTT CCAGTCCCTG CGTGTTTTAT AGTGCTGTAC GTTCGTGATC      1350
GTGAGCAAAC AGGGCGTGCC TCAACTACTG GTTTGGTTGG GTGACAGGCG      1400
CCAACTACGT GCTCGTAACC GATCGAGTGA GCGTAATGCA ACATTTTTTC      1450
TTCTTCTCTC GCATTGGTTT CATCCAGCCA GGAGACCCGA ATCGAATTGA      1500
AATCACAAAT CTGAGGTACA GTATTTTTAC AGTACCGTTC GTTCGAAGGT      1550
CTTCGACAGG TCAAGGTAAC AAAATCAGTT TTAAATTGTT GTTTCAGATC      1600
AAAGAAAATT GAGATGATCT GAAGGACTTG GACCTTCGTC CAATGAAACA      1650
CTTGGACTAA TTAGAGGTGA ATTGAAAGCA AGCAGATGCA ACCGAAGGTG      1700
GTGAAAGTGG AGTTTCAGCA TTGACGACGA AAACCTTCGA ACGGTATAAA      1750
AAAGAAGCCG CAATTAAACG AAGATTTGCC AAAAAGATGC ATCAACCAAG      1800
GGAAGACGTG CATACATGTT TGATGAAAAC TCGTAAAAAC TGAAGTACGA      1850
TTCCCCATTC CCCTCCTTTT CTCGTTTCTT TTAACTGAAG CAAAGAATTT      1900
GTATGTATTC CCTCCATTCC ATATTCTAGG AGGTTTTGGC TTTTCATACC      1950
CTCCTCCATT TCAAATTATT TGTCATACAT TGAAGATATA CACCATTCTA      2000
ATTTATACTA AATTACAGCT TTTAGATACA TATATTTTAT TATACACTTA      2050
GATACGTATT ATATAAAACA CCTAATTTAA AATAAAAAAT TATATAAAAA      2100
GTGTATCTAA AAAATCAAAA TACGACATAA TTTGAAACGG AGGGGTACTA      2150
CTTATGCAAA CCAATCGTGG TAACCCTAAA CCCTATATGA ATGAGGCCAT      2200
GATTGTAATG CACCGTCTGA TTAACCAAGA TATCAATGGT CAAAGATATA      2250
CATGATACAT CCAAGTCACA GCGAAGGCAA ATGTGACAAC AGTTTTTTTT      2300
ACCAGAGGGA CAAGGGAGAA TATCTATTCA GATGTCAAGT TCCCGTATCA      2350
CACTGCCAGG TCCTTACTCC AGACCATCTT CCGGCTCTAT TGATGCATAC      2400
CAGGAATTGA TCTAGAGTCG ACCTGCAGGC ATGCAAGCTC CTACGCAGCA      2450
GGTCTCATCA AGACGATCTA CCCGAGTAAC AATCTCCAGG AGATCAAATA      2500
CCTTCCCAAG AAGGTTAAAG ATGCAGTCAA AAGATTCAGG ACTAATTGCA      2550
TCAAGAACAC AGAGAAAGAC ATATTTCTCA AGATCAGAAG TACTATTCCA      2600
GTATGGACGA TTCAAGGCTT GCTTCATAAA CCAAGGCAAG TAATAGAGAT      2650
TGGAGTCTCT AAAAAGGTAG TTCCTACTGA ATCTAAGGCC ATGCATGGAG      2700
TCTAAGATTC AAATCGAGGA TCTAACAGAA CTCGCCGTGA AGACTGGCGA      2750
ACAGTTCATA CAGAGTCTTT TACGACTCAA TGACAAGAAG AAAATCTTCG      2800
TCAACATGGT GGAGCACGAC ACTCTGGTCT ACTCCAAAAA TGTCAAAGAT      2850
ACAGTCTCAG AAGACCAAAG GGCTATTGAG ACTTTTCAAC AAAGGATAAT      2900
TTCGGGAAAC CTCCTCGGAT TCCATTGCCC AGCTATCTGT CACTTCATCG      2950
AAAGGACAGT AGAAAAGGAA GGTGGCTCCT ACAAATGCCA TCATTGCGAT      3000
AAAGGAAAGG CTATCATTCA AGATGCCTCT GCCGACAGTG GTCCCAAAGA      3050
TGGACCCCCA CCCACGAGGA GCATCGTGGA AAAAGAAGAC GTTCCAACCA      3100
CGTCTTCAAA GCAAGTGGAT TGATGTGACA TCTCCACTGA CGTAAGGGAT      3150
GACGCACAAT CCCACTATCC TTCGCAAGAC CCTTCCTCTA TATAAGGAAG      3200
TTCATTTCAT TTGGAGAGGA CACGCTGAAA TCACCAGTCT CTCTCTATAA      3250
ATCTATCTCT CTCTCTATAA CCATGGACCC AGAACGACGC CCGGCCGACA      3300
TCCGCCGTGC CACCGAGGCG GACATGCCGG CGGTCTGCAC CATCGTCAAC      3350
CACTACATCG AGACAAGCAC GGTCAACTTC CGTACCGAGC CGCAGGAACC      3400
GCAGGAGTGG ACGGACGACC TCGTCCGTCT GCGGGAGCGC TATCCCTGGC      3450
TCGTCGCCGA GGTGGACGGC GAGGTCGCCG GCATCGCCTA CGCGGGCCCC      3500
TGGAAGGCAC GCAACGCCTA CGACTGGACG GCCGAGTCGA CCGTGTACGT      3550
CTCCCCCCGC CACCAGCGGA CGGGACTGGG CTCCACGCTC TACACCCACC      3600
TGCTGAAGTC CCTGGAGGCA CAGGGCTTCA AGAGCGTGGT CGCTGTCATC      3650
GGGCTGCCCA ACGACCCGAG CGTGCGCATG CACGAGGCGC TCGGATATGC      3700
CCCCCGCGGC ATGCTGCGGG CGGCCGGCTT CAAGCACGGG AACTGGCATG      3750
ACGTGGGTTT CTGGCAGCTG GACTTCAGCC TGCCGGTACC GCCCCGTCCG      3800
GTCCTGCCCG TCACCGAGAT CTGATCTCAC GCGTCTAGGA TCCGAAGCAG      3850
ATCGTTCAAA CATTTGGCAA TAAAGTTTCT TAAGATTGAA TCCTGTTGCC      3900
```

```
GGTCTTGCGA TGATTATCAT ATAATTTCTG TTGAATTACG TTAAGCATGT      3950
AATAATTAAC ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTTATGA      4000
TTAGAGTCCC GCAATTATAC ATTTAATACG CGATAGAAAA CAAAATATAG      4050
CGCGCAAACT AGGATAAATT ATCGCGCGCG GTGTCATCTA TGTTACTAGA      4100
TCGGGAAGAT CCTCTAGAGT CGACCTGCAG GCATGCAAGC TTGGCGTAAT      4150
CATGGTCATA GCTGTTTCCT GTGTGAAATT GTTATCCGCT CACAATTCCA      4200
CACAACATAC GAGCCGGAAG CATAAAGTGT AAAGCCTGGG GTGCCTAATG      4250
AGTGAGCTAA CTCACATTAA TTGCGTTGCG CTCACTGCCC GCTTTCCAGT      4300
CGGGAAACCT GTCGTGCCAG CTGCATTAAT GAATCGGCCA ACGCGCGGGG      4350
AGAGGCGGTT TGCGTATTGG GCGCTCTTCC GCTTCCTCGC TCACTGACTC      4400
GCTGCGCTCG GTCGTTCGGC TGCGGCGAGC GGTATCAGCT CACTCAAAGG      4450
CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG      4500
TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT      4550
GGCGTTTTTC CATAGGCTCC GCCCCCCTGA CGAGCATCAC AAAAATCGAC      4600
GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG      4650
TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT      4700
TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG GCGCTTTCTC      4750
AATGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG      4800
CTGGGCTGTG TGCACGAACC CCCCGTTCAG CCCGACCGCT GCGCCTTATC      4850
CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC      4900
TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT      4950
GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC      5000
AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG      5050
TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT      5100
TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA      5150
TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC      5200
GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC      5250
CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA      5300
AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG      5350
CGATCTGTCT ATTTCGTTCA TCCATAGTTG CCTGACTCCC CGTCGTGTAG      5400
ATAACTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT      5450
ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC      5500
CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCTCC      5550
ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT      5600
TAATAGTTTG CGCAACGTTG TTGCCATTGC TACAGGCATC GTGGTGTCAC      5650
GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG      5700
CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG      5750
TCCTCCGATC GTTGTCAGAA GTAAGTTGGC CGCAGTGTTA TCACTCATGG      5800
TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC      5850
TTTTCTGTGA CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT      5900
GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT AATACCGCGC      5950
CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAAAACG TTCTTCGGGG      6000
CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT CGATGTAACC      6050
CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT      6100
CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAATAAGG      6150
GCGACACGGA AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG      6200
AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA      6250
TTTAGAAAAA TAAACAAATA GGGGTTCCGC GCACATTTCC CCGAAAAGTG      6300
CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA CCTATAAAAA      6350
TAGGCGTATC ACGAGGCCCT TTCGTC                                6376
```

CLAIMS

1. A stamen-specific plant promoter which can be isolated from genomic corn DNA, upstream of a gene having a DNA sequence which corresponds to pCA444 of SEQ ID no. 1 or pCA455 of SEQ ID no. 2.

2. The stamen-specific promoter of claim 1 characterized by all or part of the DNA sequence between nucleotides 1 and 1179 of SEQ ID no. 3.

3. A foreign, preferably chimaeric, DNA sequence suitable for transforming a plant, which comprises the promoter of claim 1 or 2 and a structural gene under the control of the promoter.

4. The foreign DNA sequence of claim 3 wherein the structural gene is a male-sterility DNA or a male fertility-restorer DNA.

5. The foreign DNA sequence of claim 3 or 4 which is a foreign chimaeric DNA sequence.

6. A plant cell or plant cell culture transformed with the foreign DNA sequence of any one of claims 3-5.

7. A plant or its seeds consisting essentially of the plant cells of claim 6.

8. The genome of the plant of claim 7 containing the foreign DNA sequence of any one of claims 3-5.

9. A male-sterile plant of claim 7 in which the structural gene is a male-sterility DNA.

10. A male fertility-restorer plant of claim 7 in which the structural gene is a male fertility-restorer DNA.

11. A male fertility-restored plant which is a hybrid of the male-sterile plant of claim 9 and the male fertility-restorer plant of claim 10 or the seeds of the male fertility-restored plant.

12. The plant of any one of claims 7-11 which is a monocot.

13. The intron of SEQ ID no. 4.

14. A foreign, preferably chimaeric, DNA sequence for transforming a plant, such as the foreign DNA sequence of any one of claims 3-5, comprising: a promoter expressible in the plant; a structural gene under the control of the promoter; transcription termination signals, including a polyadenylation signal, for expressing the structural gene in the plant; and the intron of claim 13 within a portion of the foreign DNA sequence which is transcribed into mRNA in expressing the structural gene in the plant.

15. A plant cell, plant cell culture, plant or seed transformed with the foreign DNA sequence of claim 14.

# INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 92/00275

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/82; C12N15/29; A01H5/00; C12N5/10

## II. FIELDS SEARCHED

Minimum Documentation Searched[7]

| Classification System | Classification Symbols |
|---|---|
| Int.Cl. 5 | C12N ; A01H |

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

## III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

| Category * | Citation of Document, [11] with indication, where appropriate, of the relevant passages [12] | Relevant to Claim No.[13] |
|---|---|---|
| X<br>Y | WO,A,9 008 830 (ICI) 9 August 1990<br>see figure 18; examples 3,6 | 1,2<br>3-12 |
| Y | EP,A,0 344 029 (PLANT GENETIC SYSTEMS) 29 November 1989<br>see the whole document | 3-9,12 |
| Y | J. CELL. BIOCHEM. SUPPL.<br>vol. 15A, 1991, MEETING HELD JAN. 10-17<br>page 21;<br>MARIANI, C., ET AL.: 'Genetic destruction of tapetal cells results in the production of male sterile plants'<br>see abstract A039 | 10,11 |
| P,Y | WO,A,9 102 069 (PLANT GENETIC SYSTEMS) 21 February 1991<br>see the whole document | 10,11 |

-/--

* Special categories of cited documents :[10]

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

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

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

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

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

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

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

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

"&" document member of the same patent family

## IV. CERTIFICATION

| Date of the Actual Completion of the International Search | Date of Mailing of this International Search Report |
|---|---|
| 06 MAY 1992 | 20.05.92 |
| **International Searching Authority** | **Signature of Authorized Officer** |
| EUROPEAN PATENT OFFICE | MADDOX A.D. |

2